The Effect of Negative Supercoiling on the Formation and Positioning of Nucleosome Cores *In Vitro*.

Hugh-George Patterton

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Biochemistry

University of Cape Town

September 1991

ABSTRACT

The Effect of Negative Supercoiling on the Formation and Positioning of Nucleosome Cores *In Vitro*.

Hugh-George Patterton

Department of Biochemistry
University of Cape Town
September 1991

The effect of the negative supercoiling of DNA on the formation and positioning of nucleosome cores was investigated in a 1915bp plasmid (pHP2) containing a section of the early H1-H4 histone gene spacer of *Psammechinus miliaris*, previously shown to position the histone octamer (Retief *et al.*, 1987, *Biochemistry*, 26, 4449-4453).

It is shown for the first time, by determinations of the linking difference of reconstituted supercoiled plasmid following topoisomerase I relaxation and the yield and fragment size distribution of micrococcal nuclease digests of supercoiled and linearized plasmid, that nucleosome core reconstitution by urea/salt and salt dialysis proceeds cooperatively on both linearized and supercoiled plasmids. Evidence is further presented which indicates that the nucleosome core reconstitutes more efficiently on negatively supercoiled plasmids compared to linearized plasmids. The free energy of supercoiling is shown to be sufficient to account for this difference, and may contribute to the observed preferential migration of the octamer to negatively supercoiled plasmid compared to linear fragments. This migration is facilitated by high ionic strength, but not by high concentrations of poly[L-glutamate] or 146bp core DNA.

It is further shown for the first time, by DNase I digestion and primer extension, that identical translational and rotational positions are adopted by nucleosome cores on linearized plasmids and circular plasmids in the absence and presence of negative superhelical stress. This conservation of the positioning frames is shown to persist, irrespective of the precision of the core placement, or alterations of the relative angular orientation of the positioning frames of adjacent cores. This finding suggests that the topological and geometric constraints of chromatin loops may be negligible in the determination of nucleosome positioning *in vivo*.

The positioning of the core incorporating a $d(A\text{-}G)_{16}.d(C\text{-}T)_{16}$ stretch, shown not to adopt a H-DNA conformation in the reconstituted supercoiled plasmid, is analyzed in terms of known rotational determinants, and possible translational determinants proposed. The biological significance of the determined position of the core is discussed, and lastly, the conclusions related to other studies.

UNIVERSITY OF CAPE TOWN

PRIVATE BAG, RONDEBOSCH, 7700. SOUTH AFRICA

DOCTORAL DEGREES BOARD
Room 123, Bremner Building
Telephone: 650-2110

Telex: 5-22208
Fax: (021) 650 2138

Please complete and return to the Doctoral Degrees Board, University of Cape Town, when submitting your thesis for examination

PhD THESIS TITLE: The effect of negative supercoiling on the formation and positioning of nucleosome cores in vitro.

I hereby

(a) grant the University of Cape Town free licence to reproduce the above thesis in whole or in part, for the purpose of research;

(b) declare that:

(i) the above thesis is my own unaided work, both in concept and execution, and that apart from the normal guidance from my supervisor, I have received no assistance except as stated below:

(ii) except as stated below, neither the substance nor any part of the above thesis has been submitted in the past, or is being, or is to be submitted for a degree in the University or any other university.

The thesis has been presented by me in fulfilment of the requirements for the degree of PhD.

SIGNED: Signed by candidate

DATE: 12/9/91.

The Effect of Negative Supercoiling on the Formation and Positioning of Nucleosome Cores *In Vitro*.

Hugh-George Patterton

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Biochemistry

University of Cape Town

September 1991

CERTIFICATE OF SUPERVISOR

In terms of paragraph eight of "General regulations for the degree of Ph.D." I, as supervisor of the candidate, H. G. Patterton, certify that I approve of the incorporation into this thesis of material that has already been published or submitted for publication.

Signed by candidate

Prof Claus von Holt

Director, UCT-CSIR Research Center for Molecular Biology
in the Department of Biochemistry
University of Cape Town

CONTENTS

CHAPTER 4

CHAPTER 5

ABSTRACT

The Effect of Negative Supercoiling on the Formation and Positioning of Nucleosome Cores *In Vitro*.

Hugh-George Patterton

Department of Biochemistry
University of Cape Town
September 1991

The effect of the negative supercoiling of DNA on the formation and positioning of nucleosome cores was investigated in a 1915bp plasmid (pHP2) containing a section of the early H1-H4 histone gene spacer of *Psammechinus miliaris*, previously shown to position the histone octamer (Retief *et al.*, 1987, *Biochemistry*, 26, 4449-4453).

It is shown for the first time, by determinations of the linking difference of reconstituted supercoiled plasmid following topoisomerase I relaxation and the yield and fragment size distribution of micrococcal nuclease digests of supercoiled and linearized plasmid, that nucleosome core reconstitution by urea/salt and salt dialysis proceeds cooperatively on both linearized and supercoiled plasmids. Evidence is further presented which indicates that the nucleosome core reconstitutes more efficiently on negatively supercoiled plasmids compared to linearized plasmids. The free energy of supercoiling is shown to be sufficient to account for this difference, and may contribute to the observed preferential migration of the octamer to negatively supercoiled plasmid compared to linear fragments. This migration is facilitated by high ionic strength, but not by high concentrations of poly[L-glutamate] or 146bp core DNA.

It is further shown for the first time, by DNase I digestion and primer extension, that identical translational and rotational positions are adopted by nucleosome cores on linearized plasmids and circular plasmids in the absence and presence of negative superhelical stress. This conservation of the positioning frames is shown to persist, irrespective of the precision of the core placement, or alterations of the relative angular orientation of the positioning frames of adjacent cores. This finding suggests that the topological and geometric constraints of chromatin loops may be negligible in the determination of nucleosome positioning *in vivo*.

The positioning of the core incorporating a $d(A\text{-}G)_{16}.d(C\text{-}T)_{16}$ stretch, shown not to adopt a H-DNA conformation in the reconstituted supercoiled plasmid, is analyzed in terms of known rotational determinants, and possible translational determinants proposed. The biological significance of the determined position of the core is discussed, and lastly, the conclusions related to other studies.

ACKNOWLEDGEMENTS

Several people have contributed to the research project presented in this manuscript and are gratefully acknowledged:

Prof Claus von Holt, my supervisor, for his advice, support, and for having the original idea to investigate the effect of negative supercoiling on nucleosome core positioning.

Dr Aaron Klug for discussions and his many helpful suggestions.

Prof Max Birnstiel for supplying a plasmid containing the entire early histone gene battery of *Psammechinus miliaris*, and also for supplying the unpublished nucleotide sequence of the battery.

Peter Becker for discussion and for writing and assisting with the implementation of the computer program to remove hidden lines used in the generation of figures 4.6, 4.8 and 4.14.

Dr Ruth Smart for discussion on mathematical methods of surface fitting.

Drs David Clark and Gary Felsenfeld for advice on reconstitution methodology and accurate determination of the histone octamer concentration.

Dr Jerry Rodrigues for performing amino acid analyses.

My colleagues at the Department of Biochemistry at the University of Cape Town for discussion.

The Foundation of Research Development (CSIR) for financial assistance.

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS

A_{230}	Absorbance at 230nm
A_{260}	Absorbance at 260nm
bp	base-pair
BSA	Bovine serum albumin
DMSO	dimethylsulphoxide
DNase I	deoxyribonuclease I
DTT	dithiothreitol
EDTA	ethylenediaminetetra-acetic acid
EGTA	ethyleneglycol-bis-(2-amino-ethyl ether) N,N'-tetra-acetic acid
LB	Luria-Bertani bacterial growth medium
MNase	micrococcal nuclease
PAGE	polyacrylamide gel electrophoresis
PMSF	phenylmethylsulfonyl fluoride
SDS	sodiumdodecyl sulphate
TRIS	Tris(hydroxymethyl)-aminomethane

CHAPTER 1

1 Literature review.

1.1 Nucleosome structure:

1.1.1 Introduction:

Mild nuclease digestion of eukaryotic chromatin releases the fundamental genetic packaging unit, the nucleosome, which consists of 166bp of duplex DNA wound in two negative superhelical turns around an octameric protein complex composed of two each of the core histones H2A, H2B, H3 and H4 (reviewed in 34, 66, 126, 155, 254 and 271). The superhelical DNA continues into a variable amount of linker DNA not directly bound to the protein surface. A single copy of H1 (4) or one of its variants is associated on the outside of the structure at the point where the DNA enters and leaves the nucleosome (164, 178, 236). Further nuclease digestion results in the dissociation of histone H1 (178) concomitant with trimming the DNA down to 146bp to produce the core particle, a discoid structure 57Å thick and 110Å in diameter with a two-fold axis of symmetry (121). The structure of the core particle was independently solved by X-ray crystallography to 7Å resolution by Richmond *et al.* (208), and to 8Å resolution by Uberbacher and Bunick (251).

These two studies are structurally in close agreement and placed the B-form DNA double helix as 1,8 turns of a slightly irregular left-handed superhelix at a mean pitch of 28-30Å around the outside of the histone octamer (208, 251). The superhelix contained 7,6 helical periods per turn which gave the duplex DNA a (not necessarily uniform) helical period of approximately 10,2bp with respect to the octamer surface (251).

The histones constituting the protein core were visible as discrete electron dense domains that contacted the inside of the superhelix (208, 251). At several of these locations rod-like features were evident which were proposed to represent α-helixes (208). No protein density was seen to extend around the DNA duplex (208, 251). Individual histones were assigned by their spatial proximities deduced from histone-histone crosslinking data (reviewed in 155), and from DNA-histone contacts relative to the 5'-side of each strand (15, 163).

1.1.2 Structure of the core particle:

The DNA duplex was proposed to contact the histones in the order $H2A^1$-$H2B^1$-$H4^1$-$H3^1$-$H3^2$-$H4^2$-$H2B^2$-$H2A^2$, where the superscript denotes the histone copy (121, 208, 251). In particular, the two H4 molecules were placed furthest from the point of DNA entry and exit, making extensive contacts with the DNA on either side of the dyad axis in the region between sites ±3 and ±4* (208, 251). Additional contacts may take place in the region between sites ±4 and ±5 (208).

The two H3 molecules contacted the central turn of the superhelix in the vicinity of the dyad in the region from -2,5 to +2,5 (251). In the Richmond *et al.* (208) electron density map, postulated α-helixes made extensive contact with the minor groove on either side of the dyad axis at sites ±0,5.

---

* This notation represents contact in the region from site +3 to +4 by one copy of H4, and from site -3 to -4 by the other H4 molecule, symmetrically related through the dyad axis. The sites refer to the approximate positions of maximal DNase I cleavage in the core DNA where the minor groove is orientated away from the protein surface (142). These sites are numbered from -7 to +7 with site 0 lying on the dyad axis on the side of DNA entry and exit (119).

H2B was in close proximity to the DNA helix from sites ±4 to ±5,5 and H2A from ±5,5 to the terminus at site ±7 (251). The assignments of the dimer contact points form a departure between the two crystallographic studies, since Richmond *et al.* (208) found a shift of the one dimer relative to its position of symmetry and an obstruction of the smooth exit angle of the DNA superhelix by both H2A molecules. Neither of these features were observed by Uberbacher and Bunick (251), who ascribed them to packing forces, as suggested by Richmond *et al.* (208).

1.1.3 DNA-octamer association:

The complex between the histone octamer and the DNA double helix is stabilized by electrostatic interaction, possibly along a left-handed ridge of about 25Å pitch. This ramp was observed after reconstruction of the three-dimensional structure of the histone octamer from electron microscopic images of ordered tubular aggregates (121). The crystallographic studies of Richmond *et al.* (208) and Uberbacher and Bunick (251) suggested that the DNA duplex contacts the octamer surface once every helical period, consistent with the approximately ten base pair periodicity found in DNA-histone crosslinking sites (15, 163). Determination of the net ion release upon thermal denaturation allowed McGhee and Felsenfeld (153) to estimate that about 15% of the phosphate groups in the terminal core DNA are neutralized. If this value is representative of all the DNA in the core, it would be consistent with about three strong interactions per helical period (153). Ichimura *et al.* (109) found that upon release of the histone octamer from the core DNA, 14 additional arginine residues became available for chemical modification and proposed that these were involved in strong interactions via ringed-structure salt-bridges with the tetrahydral phosphates of the DNA backbone. No increase in the fraction of reactive lysine

residues was found upon octamer dissociation (109), which is in disagreement with the study of Lambert and Thomas (128) who found that several lysine side-chains became more accessible to reductive methylation in each of the four core histones after removal of the core DNA. It is currently accepted that both arginine and lysine residues are involved in the electrostatic binding of the core DNA to the octamer surface (128).

1.1.4 Location of histone H1/H5 in the nucleosome:

The position of the H1/H5 molecule in the nucleosome, not being part of the crystallized core particle (208, 251), has been inferred indirectly from several lines of evidence. Micrographs of chromatin showed the linker DNA to enter and exit the nucleosome at juxtaposed sites in native chromatin, whereas, in H1-stripped chromatin, these points appeared on opposite sides of the nucleosome (see 1.2.2). This led Thoma *et al.* (236) to propose that H1 is bound to the linker DNA at the point of entry and exit, thus stabilizing two full turns of the nucleosomal DNA. This is supported by Noll and Kornberg (178) who demonstrated the release of H1 from chromatin during micrococcal nuclease digestion past the 166bp pause down to the 146bp particle. The stabilization of the additional 20bp of linker DNA was ascribed to the central, globular domain of H1/H5 by Allan *et al.* (5) who demonstrated the presence of a transient pause at 166bp during micrococcal nuclease digestion of stripped chromatin reconstituted with only the globular peptide fragment of either H1 or H5. Boulikas *et al.* (27) reported the crosslinking of a section from the globular domain of H1 to H2A using a "zero-length" crosslinker. The position assigned to H2A in the core particle (208, 251) and the inferred proximity of H1 to H2A (27) support the suggested placement of the central domain of H1/H5 at the core DNA termini. Mirzabekov *et al.* (164) reported the

chemical crosslinking of histidine 25 and 62, both within the globular domain of H5, to proximal and distal linker DNA respectively. Drew and McCall (56), studying a 440bp fragment containing three histone octamers, reported decreased nuclease susceptibility at the ends of the DNA molecule as well as in the linker areas and dyad sections of two of the three cores after reconstitution with H5. From all of the available evidence it can reasonably be concluded that the globular domain of H1/H5 binds to the terminal region of the core DNA (164).

Little is known about the structural features governing the placement of H1/H5 on the nucleosome (234). Turnell *et al.* (250) proposed the involvement of a decapeptide motif identified in a range of core particle binding proteins. This decapeptide motif was suggested to fold as an amphipathic α-helix with suitably orientated basic residues making electrostatic contact with phosphate groups on either side of a narrow minor groove at the core dyad (250). The required proximity of the decapeptide motif to the DNA is supported by crosslinking (27, 164) and chemical modification (241) studies.

The binding of a single H1/H5 molecule confers a polarity to the core particle having a two-fold symmetry axis (121, 208, 251). This property was discussed fully by Widom (271) and van Holde (254).

1.2 Chromatin structure:

1.2.1 Introduction:

The winding of 166bp of duplex DNA in two superhelical turns onto the discoid histone octamer with a diameter of approximately 110Å (121, 208, 251), results in a lengthwise compaction of the DNA molecule of less than one order of

magnitude. Since the single DNA molecule from a typical eukaryotic chromosome is several centimeters in length (254), it seems obvious that the eukaryotic genome must be arranged at levels of organization above a linear array of nucleosomes to fit into the cell nucleus of 10μm diameter (34). These higher levels of folding must not only achieve the required 10 000 fold compaction of the metaphase chromosome (271), but must also be sufficiently reversible during the replication cycle of the cell to allow processes such as DNA transcription, replication and recombination to occur during interphase (66, 271).

The relevance of the various levels of structural organization of chromatin to the basic genetic and metabolic requirements of the cell (236, 254, 271) has ensured intense interest in this field, but, unlike the structure of the core particle (15, 163, 208, 251), no definitive model exists for the higher order structures of chromatin (254, 271). The reason for this is twofold: firstly, it is an extremely complex problem that is, by necessity, approached using techniques suited to study spatially ordered structures (274); the X-ray diffraction pattern of the chromatin fibre contains broad bands, indicating that the fibre is not completely regular (273). Secondly, the sensitivity of the chromatin fibre to preparative methodology and ionic environment can easily result in artifactual structures (34, 236, 273). A large body of information on the general morphological behavior of the chromatin fibre isolated from different sources under various ionic conditions has consequently been generated by electron microscopic and hydrodynamic studies (13, 33, 68, 236, 240, 270).

1.2.2 The role of ionic strength and histone H1/H5:

In a detailed electron microscopic study by Thoma *et al.* (236) it was reported that at very low ionic strengths the chromatin assumed an open, zigzag shape, with the linker DNA seen to enter and exit the clearly identifiable nucleosomes at almost the same position. Increasing the ionic strength lead to a gradual compaction of the chromatin with the nucleosomes packed more closely and the linker DNA generally not visible. At 40mM NaCl a compact fibre with 3 nucleosomes discernable across its width of 20 to 25nm was seen. Finch and Klug (68) had in an earlier electron microscopic study observed similar structures at low Mg^{2+} concentrations and found them, in conjunction with X-ray diffraction data, compatible with a helical arrangement of nucleosomes into a solenoid (single start helix) with between 6 and 7 nucleosomes per turn, a finding supported by the evidence of Thoma *et al.* (236). At 60mM NaCl individual nucleosomes were generally not visible, but cross-striation, indicative of a helical structure, was evident in fibres with a width of approximately 30nm (hereafter referred to as the 30nm fibre).

The results of this electron microscopic investigation are entirely consistent with the hydrodynamic study of Butler and Thomas (33) on nucleosome oligomers at different ionic strengths. These authors found a linear relationship in a double logarithmic plot of sedimentation coefficient *versus* ionic strength for nucleosome trimers to pentamers between 5mM and 25mM NaCl. Above this salt concentration, the sedimentation coefficient (s) was independent of ionic strength (I). However, the linear relationship, obeying the simple power law $s \propto I^n$, was found to extend over a wider range between 5mM and 125mM NaCl for hexameric and larger oligomers up to 30mers. The oligomer size required for this transition was proposed to correspond to one helical turn in

a solenoidal structure that condenses continuously with increasing ionic strength. Although a transition in the sedimentation behavior of similar magnitude was not seen for oligomers at higher multiples of six, the value of the exponent "n" in the power law changed fairly sharply between the 11mer and 13mer, possibly indicating the formation of two complete helical turns in the solenoid. Oligomers of size 50mer and above (33, 240) displayed an abrupt decrease in the sedimentation coefficient (dependent on the presence of histone H1; see below) at ionic strengths below approximately 50mM NaCl, with linear relations of different slope on the two sides of the jump. This discontinuity was shown not to be due to aggregation and was proposed to represent the disruption of the large solenoidal structure below about 50mM NaCl where insufficient axial stabilization renders it susceptible to hydrodynamic shear. This proposed change to a more stable structure above 50mM NaCl was supported by the observation that crosslinking the sample above (65mM) or just below (40mM) the ionic strength where the jump occurred, abolished this transition. The authors suggested that the transiently present structures were sufficiently compacted at 40mM NaCl, just under the point where the jump occurred, allowing the crosslinking reagent to accumulatively capture the structure of the solenoid (33) which was not axially stabilized enough to withstand the hydrodynamic shearing forces in the uncross-linked state at low salt (240). The jump could be regenerated by cleaving the crosslinking reagent, and was also present when crosslinking was performed at low ionic strengths (15mM NaCl) where the structure was insufficiently compacted (33). Bates *et al.* (13) found that chicken erythrocyte chromatin exhibited hydrodynamic properties similar to rat liver chromatin discussed above, but that the former seemed more stable since at identical ionic strengths a 60mer withstood the same frictional shear as a 50mer from rat liver chromatin. It was proposed that the presence of substantial

amounts of histone H5, an extreme variant of H1, stabilizes the virtually transcriptionally inactive chromatin of chicken erythrocytes more than H1 does the transcriptionally active rat liver chromatin (13, 239).

The gradual folding of chromatin with increasing ionic strength could be reproduced by substituting NaCl with equivalent concentrations of KCl or much lower (less than 2mM) concentrations of $MgCl_2$, as judged by the appearance of the chromatin fragments under the electron microscope (236, 270), and by X-ray diffraction studies (270). The response of chromatin to ionic strength and the valency of the cation was quantitatively described (44) by the counterion condensation theory of polynucleotides (147). In this treatment, the partial neutralization of the negatively charged phosphate groups by cations localized within a 7Å distance of the DNA molecule, and increased screening of the residual charge by delocalized cations (147), diminished coloumbic repulsion between the linker areas (44) and closely juxtaposed DNA supercoils on the nucleosomes of adjacent turns of the solenoid (270, 271), allowing the structure to become very compact. This intramolecular folding into a fibre structure is not only dependent on charge neutralization and shielding, but also requires the histone H1 (or H5), as referred to above.

Thoma *et al.* (236) showed that chromatin stripped of histone H1 (and some nonhistone proteins) by treatment with an ion-exchange resin, responded differently to ionic strength variations compared to native chromatin. At very low ionic strengths the H1-depleted chromatin appeared as extended filaments of irregular thickness with the nucleosomes opened up. At higher ionic strengths (40mM NaCl) the nucleosomes were stabilized by the reduction in charge repulsion, but still differed from the nucleosomes observed at similar ionic strengths in H1-containing chromatin. The most

striking difference was that the DNA could be seen to enter and leave the nucleosomes at opposite sites giving rise to a "beads-on-a-string" appearance as opposed to the zigzag filament of native chromatin. Above 40mM NaCl the H1-depleted chromatin condensed into irregular clumps of varying diameter connected by beads-on-a-string filaments.

Butler and Thomas (33) also found a gradual compaction of H1-depleted chromatin with increasing ionic strength as demonstrated by the linear relation between the sedimentation coefficient and the ionic strength in double logarithmic plots of these two parameters, but demonstrated that this relation was shifted to lower sedimentation values compared to native chromatin, which can be explained by the more extended structures observed by Thoma *et al.* (236). This salt-dependent condensation was also demonstrated by sedimentation studies on reconstitutes of twelve tandem repeats of the sea urchin 5S rRNA gene (99). The structural homogeneity of the sample allowed Hansen *et al.* (99) to demonstrate by comparison of experimentally determined and model-dependent predicted sedimentation values that H1-depleted "chromatin" existed as an extended structure within a limited range of conformations at low ionic strength, and as a more compacted, non-solenoidal structure at moderate ionic strength (40mM and 100mM NaCl). This difference in structure is also underscored by the absence of dramatic shifts in sedimentation behavior of large H1-depleted chromatin fragments at lower ionic strengths (33). When H1 is reconstituted with H1-depleted chromatin, the resulting preparation behaves morphologically similar to native chromatin (33).

Clark (44) proposed that the difference in sedimentation coefficient (and hence degree of compaction) of native and H1-depleted chromatin over a given range of ionic strength was explicable in terms of the contribution of H1 to the

partial charge neutralization of the DNA phosphates. It was further shown (44) that the same linear relation existed between the frictional coefficients of both native and H1-stripped chromatin and the electrostatic free energy of the linker DNA (calculated from Manning's counterion condensation theory of polynucleotides (147)) when the contribution of H1 binding to fractional charge reduction was considered. This suggests that an important role of H1 is the partial neutralization of charges on the linker DNA, allowing these sections to approach much closer in the fibre structure without raising the electrostatic free energy to unacceptable levels. It was also pointed out that any factor such as histone variants, acetylation, phosphorylation or ADP-ribosylation which influenced the residual charge or affected the degree of charge neutralization of particularly H1, would modulate the extent of folding of the 30nm fibre (44). The involvement of H1 in non-electrostatic interactions is not, however, excluded, and in a broader sense H1 may be seen to thermodynamically stabilize the higher order structures of chromatin (68, 236).

Apart from the electron microscopic and hydrodynamic studies on chromatin which gave information on its general shape, various other techniques, aimed at elucidating the internal architecture of the 30nm fibre, have also been utilized. No attempt was made to exhaust the literature in this area which is fully covered by Butler (34), van Holde (254) and Widom (271), and reference to only the most significant studies in the context of this manuscript is made in the discussion below.

1.2.3 Fine structure of the 30nm chromatin fibre:

Any reasonable model of the 30nm fibre must incorporate at least two key aspects of fibre geometry: the orientation of

individual nucleosomes with respect to the fibre axis, and the arrangement of the linker DNA between nucleosomes (connectivity).

There are only three possible orthogonal orientations of the nucleosomes with respect to the fibre axis: the nucleosomes can be arranged radially, with their flat faces parallel to the fibre axis and their inner edges touching, or tangentially, with their faces still parallel to, but now with one face of each orientated towards the fibre axis, or perpendicular, with the faces perpendicular to the fibre axis, as the name suggests.

Suau *et al.* (230) performed neutron scattering experiments on Mg^{2+} condensed chicken erythrocyte chromatin and showed that the obtained radius of gyration was incompatible with that expected for a fibre with tangentially arranged nucleosomes. It was also noted (230) that the radial distribution function lacked the sharp inflection suggested to signify a tangential packing with a large hole in the center of the fibre. McGhee *et al.* (157) were able to discriminate between the two remaining possibilities by measuring the dichroism of chicken erythrocyte chromatin fibres at 260nm. This measurement was repeated at different electric field strengths and extrapolated to infinite field strength where all the fibres would be orientated with their axes parallel to the applied field. Since the measured dichroism is composed of the dichroism signals of the linker DNA and of the chromatosomal DNA, and since the magnitude of these signals depend on the orientation of the linker or chromatosomal DNA relative to the aligned fibre axis, it is possible to deduce a range of conformations from the measured net dichroism. The overall dichroism was found to be negative by McGhee *et al.* (157) and therefore excluded the perpendicular arrangement of the chromatosomes (maximum positive contribution) even if the linker DNA is assumed to

be aligned parallel to the fibre axis (maximum negative contribution). In fact, if the contribution by the linker DNA is maximally negative, the largest permitted tilt angle of the chromatosome relative to the fibre axis was found to be approximately 40°. If the linker DNA is assumed to trace a supercoiled path between adjacent nucleosomes, the average tilt angle is estimated at between 20° and 25° (157). In a later study by the same group (156) the field-free relaxation times and dichroism were measured for a variety of chromatin fibres with linker lengths ranging from 11bp (CHO cells) to 74bp* (sea urchin sperm). The measured relaxation time was compared to the relaxation times predicted by hydrodynamic theory for three different solenoid geometries and found to be compatible with a one start contact helix with 6 nucleosomes per turn, 11nm pitch and 30nm diameter, irrespective of linker length. The investigators did note, however, that the calculated dimensions could accommodate a slight variation in fibre diameter. The relation between the dichroism and the two dependent variables, nucleosome tilt and linker pitch angle (defining the supercoiled path of the linker DNA), allowed McGhee *et al.* (156) to estimate the tilt angle by fixing the pitch angle at a value compatible with the deduced solenoid dimensions. The result suggested that the nucleosome faces are tilted between 20° and 33°, in inverse proportion to the linker length.

The dependence of the tilt angle on the assumed orientation of the linker DNA was avoided by Mitra *et al.* (161) who, using the technique of photochemical dichroism, could discriminate between the contributions of the linker and chromatosomal DNA to the net dichroism and obtained a result

*The linker lengths presented represent generally accepted values and are not necessarily those used by the authors of the cited publication. This is done to maintain consistency.

indicating a tilt angle of 33°±3° for the chromatosomes relative to the fibre axis. It is generally excepted that the nucleosomes are tilted by 20°-40° in the 30nm fibre (254).

The X-ray diffraction study on chicken erythrocyte chromatin fibres by Widom and Klug (273) gave an indication of the arrangement of nucleosomes in the 30nm fibre with respect to each other. In that study, fibers were partially oriented in a X-ray capillary, and a series of diffraction bands which could be correlated with the direction of the 30nm fibre axis were recorded at 34nm, 11nm, 5,7nm, 3,7nm and 2,7nm. The 34nm band, for instance, was found to be equatorial, and ascribed to interference from laterally packed fibres oriented parallel to the capillary direction. The 11nm feature was found to be meridional and proposed to originate from the vertical, edge-to-edge stacking of nucleosomes along the fibre axis on progressive turns of the solenoid, in agreement with the cross-striations of about 11nm pitch observed by Thoma *et al.* in earlier EM work (236) (see 1.2.2). The 5,7nm equatorial reflection was ascribed to the radial packing of nucleosomes around the fibre axis, and both the 3,7nm and 2,7nm equatorial reflections shown to originate from intranucleosomal structural features. To confirm the proposed orientation of the nucleosomes in the 30nm fibre, Widom and Klug (273) calculated the fibre diffraction pattern from the electron density map of a single core particle (208) and found that only an orientation similar to that proposed, gave a diffraction pattern in the 2,7nm-3,7nm region similar to that obtained for the whole fibre in the equivalent higher angle region.

The second aspect of fibre geometry, the trajectory of the linker DNA, has proven to be rather elusive to rigorous investigation. It was not visible in electron micrographs of the 30nm fibre (236), could not be unambiguously assigned

from electric dichroism data (156, 157, 161) and contributed only weakly to X-ray diffraction patterns (273). The effect of linker length on fibre diameter was also contested. Butler (35) stated that the fibre diameter was independent of linker length on the grounds of electron microscopic and sedimentation studies on different repeat length chromatins, a view supported by the EM study on chicken erythrocyte and sea urchin sperm chromatin by Widom *et al.* (272). This conclusion was, however, disputed by Williams *et al.* (274) who measured fibre diameters by electron microscopy and found an average of 30nm for *Necturus* chromatin (28bp linker) in comparison with 40nm for *Thyone* chromatin (67bp linker).

A quantity that might reflect on this contention is the mass per unit length of different repeat length fibres. Widom and colleagues (272, 273) have shown that partially oriented fibres from sea urchin sperm (74bp linker) and chicken erythrocyte (46bp linker) gave identical X-ray diffraction patterns. Of special significance was the fact that both gave an 11nm meridional and 5,7nm equatorial reflection, from which it could reasonably be concluded that the radial disposition and edge-to-edge packing of nucleosomes along the fibre axis were insensitive to linker length (272). The corollary was that if the fibre diameter was independent of linker length, the mass per unit length (or number of nucleosomes per helical turn) must likewise have been independent, provided the fibre structure was otherwise conserved. Here, too, was disagreement. In neutron scattering experiments on chicken erythrocyte chromatin (46bp linker), Suau *et al.* (230) found 6-7 nucleosomes per 11nm helical turn. An identical value for chicken erythrocyte chromatin was also found by both scanning transmission electron microscopy (STEM) and small-angle neutron scattering studies by Gerchman and Ramakrishnan (80). A significantly different result was obtained by

Woodcock *et al.* (276) who found close to 12 nucleosomes per 11nm by STEM studies of chicken erythrocyte chromatin, although this result has been ascribed to aggregation (80). The mass per unit length of *Necturus* fibres with a 28bp linker length was determined by STEM by Williams *et al.* (274) and found to be 8,1 nucleosomes/11nm, slightly higher than the result obtained for chicken erythrocyte fibers with a 46bp linker length (230, 80). *Thyone* (67bp linker) fibres were found to have close to 13 nucleosomes/11nm (274). This apparent relationship between linker length and mass per unit length (and hence fibre diameter), was further pursued by Williams *et al.* (274). This group estimated fibre diameters from the intensity and position of a "20nm reflection", proposed to correspond to fiber diameter, found in X-ray diffraction studies of isolated nuclei from species with different repeat lengths. A linear relationship between nucleosome repeat and fibre diameter was found with values ranging from 29nm for mouse myeloma nuclei (25bp linker) to approximately 41nm for sea urchin sperm (74bp linker). This result was dependent on the interpretation of the "20nm reflection", and although the proposed origin seems likely, it has not yet been confirmed (271).

The various models proposed for the 30nm fibre are reasonably consistent in terms of nucleosome arrangement, but diverge significantly in placement of the linker DNA and hence the connectivity, which is designed to reflect different experimental data sets.

The solenoid model proposes the 30nm fibre as a single start contact helix with 6-7 nucleosomes (not necessarily an integral number) per helical turn arranged at a constant pitch of 11nm (68, 236, 273). The nucleosomes were suggested to be slightly tilted relative to the fibre axis with lateral connectivity (273). Slight variations to the solenoid model in terms of linker disposition are those of

McGhee *et al.* (156) placing the linker DNA as supercoils between adjacent nucleosomes, and Butler (35) who suggested that the linker DNA bulges into the central hole of the solenoid.

The twisted-ribbon model (276) is a single start helix in which the zigzag form is compressed into a parallel stack of nucleosomes wound onto the surface of a cylinder with the linker DNA repeatedly running up and down the fibre axis, connecting the two parallel rows of nucleosomes. Every other nucleosome would thus be juxtaposed. The lengthwise arrangement of the linker DNA of this model is not supported by dichroism data (156, 161), nor has a linker length dependent meridional X-ray reflection been observed (272, 273).

The crossed-linker model (274) essentially proposes a left-handed, two start helix with two parallel rows arranged in a helical manner on opposing sides of a cylinder with the linker DNA connecting nucleosomes transversely through the center of the fibre. Such an arrangement is intended to explain the proposed dependence of the fibre diameter on the linker length and the high mass per unit length and pitch angles obtained (274). Some of this experimental data has been questioned by Thoma (234) and Widom (271).

It would be premature to dismiss any of the models in their broader terms on the grounds of current information (254), and attempts to arrive at a definitive model must address the following questions conclusively: is the mass per unit length and fibre diameter independent of variations in linker length, is the structure a one-start or a two-start helix, and how is the linker DNA arranged?

1.2.4 *In vivo* relevance of *in vitro* structures:

A large proportion of the results discussed above have been obtained by studying "refolded" chromatin fibres. A natural question would be: what is the structural integrity of these fibres with regard to the *in vivo* conformation of chromatin? It has been demonstrated (33, 236) that the EM appearance and hydrodynamic behavior of fibres that have been isolated at low ionic strength and subsequently refolded were indistinguishable from chromatin isolated at ionic strengths where it existed as a compacted structure.

The faithfulness of fibres isolated under conditions where it exists as a condensed structure was demonstrated by the study of Langmore and Schutt (129), where the low-angle X-ray scatter of whole living cells and isolated nuclei were recorded. A strong band at approximately 40nm was found which could be eliminated by nuclease treatment and was proposed to correspond to the fibre diameter. Some of the cells and all of the nuclei studied also gave scattering bands at 20nm, 11nm, 6nm, 3,8nm and 3nm suggested to reflect internal structural features. This same series of reflections was also obtained by X-ray diffraction of isolated nuclei by Williams *et al.* (274), and corresponds closely to the diffraction patterns obtained *in vitro* on partially oriented fibres by Widom and Klug (273) and Widom *et al.* (272). It therefore appears that *in vivo* structures are faithfully reflected by *in vitro* structures in terms of diameter and internal nucleosomal arrangement (129), but subtle differences are not excluded (271).

1.3 Chromatin domains:

There is some evidence that in the metaphase chromosome the 30nm fibre coils round itself to form supercoiled loops of approximately 50nm diameter which are attached at their

bases to the proteinaceous chromosome scaffold enriched in topoisomerase II and other non-histone proteins (reviewed in 34, 77 and 254). Some form of higher order organization of the 30nm fibre is conserved during the cell cycle, and during interphase topologically sequestered chromatin loops still exist. Benyajati and Worcel (16) reported a gradual increase in the sedimentation coefficient of interphase *Drosophila* genome up to a plateau during mild DNase I digestion. If the genome had been topologically continuous, a single nick would have produced complete relaxation and the existence of only two conformational states would have been evident. The requirement for several nicks in the DNA before it reaches a limit conformation was therefore consistent with the existence of topologically isolated loops.

The average loop size from the Benyajati and Worcel study (16) was estimated at 85Kbp. A wide variety of average values ranging from 35 to 250Kbp have been reported in the literature (see 254 and references cited therein). The reason for this broad range was addressed by Jackson *et al.* (111) who showed in a restriction enzyme study on agarose encapsulated HeLa cells that the average loop size was strongly dependent on the isolation procedure, 86Kbp being the average value obtained under the most "physiological" conditions. The average loop size was further shown to be insensitive to the replication state of the cell (111).

In the interphase nucleus the chromatin loops appear attached to a scaffold comprised of a complex set of non-histone proteins from the peripheral lamina and to the equivalent of the metaphase scaffold, the internal matrix (76, 110). Scaffold attachment regions (SARs) have been identified in the non-transcribed regions of the histone gene repeat, heat-shock hsp70, alcohol dehydrogenase (*Adh*), fushi tarazu (*ftz*) and glue protein (*Sgs-4*) genes from

Drosophila (75, 162), as well as the chicken lysozyme (192) and human β-globin (113) genes. In the case of the heat-shock, *Adh*, *ftz* and *Sgs-4* genes the SARs were found to be in close proximity to regulatory/enhancer elements (75, 162). In agreement with the cell cycle invariance of the loop size (111), no rearrangements of SARs were detected upon induction of transcription on the heat-shock gene (162), nor was any redistribution of SARs of histone, heat-shock or *ftz* genes evident in a comparison of scaffolds from interphase and metaphase nuclei (161).

The base sequence or DNA structure recognized by the scaffold in the SARs seems highly conserved, since the cloned SAR fragments from *Drosophila* bind equally well to rat liver and HeLa scaffolds (110, 161). The areas of protein binding in the scaffold attachment region of the H1-H3 histone gene spacer from *Drosophila* have been identified as two 200bp exonuclease III resistant regions separated by a 100bp nuclease accessible area (76). Sequence analysis of SARs revealed the presence of several copies of the topoisomerase II consensus sequence as well as 10bp A-rich and T-rich elements (note that these are not cross-strand complementary sequences) (75, 76). Postulated roles for the latter sequence motifs include sequence-specific protein binding, curvature of the DNA, and maintenance of nucleosome free stretches (75).

The role of the structurally demarcated chromatin domains has been postulated to include independent transcription units topologically regulated by topcisomerase II at the bases of the loops (75) and allowing the close approach of regulatory elements and their promotors within a subcompartment enriched in recognition sequences for transcription factors and RNA polymerase (162).

1.4 Nucleosome positioning:

1.4.1 Introduction:

The placement of a nucleosome on DNA can be defined in terms of two structural parameters of the DNA helix: a translational and a rotational component (210, 243, 246). The translational component denotes the frame extending linearly along the DNA molecule that is occupied by the nucleosome (59, 246). This frame is experimentally determined by measuring the difference in accessibility of the DNA, nucleosomally wrapped and nucleosome-free in solution, to DNase I (55, 200), micrococcal nuclease (18, 21, 55, 235), exonuclease III (135, 199), or various restriction endonucleases (54, 74). Trimethylpsoralen crosslinking studies have also demarcated translational frames (36).

The rotational component refers to the angular orientation of the DNA helix relative to the histone octamer surface (59, 246). This frame is experimentally obtained by determining the phase of the periodic difference in the accessibility along the DNA helix wrapped into nucleosomes and in an appropriate nucleosome-free control to enzymatic or chemical probes (246). These include DNase I (55, 59, 205), DNase II (200), micrococcal nuclease (55) or hydroxyl radicals (194, 249). The oscillating reactivity results from the structural feature or chemical group recognized by the probe being exposed only periodically, due to the proximity of the helical DNA molecule to the octamer surface (142).

The independence by which these two positional parameters can be altered is constrained by the structural requirements of the nucleosome core. Movement of the nucleosome along the length of the DNA molecule introduces new sequences into

the one side of the nucleosomal DNA while removing sequences from the other side, thereby altering the translational frame of the nucleosome. The size of the increment by which this translational shift can occur depends on the variance allowed in the rotational frame. The X-ray crystallographic structure of the core particle (208, 251) showed that the DNA helix is associated with the surface of the histone octamer in a defined fashion. For instance, on the dyad axis at position 0, the minor groove of the DNA points away from the protein surface (208). This orientation may be due to the interaction between putative α-helixes of the H3 histones and the minor groove on either side of the dyad axis at positions +0,5 and -0,5 (208). This structural arrangement is maintained when the core DNA is shifted over the octamer surface in steps of an integral helical turn, which results in the octamer remaining associated with the same side of the DNA helix. It is thus clear that nucleosomes in different translational frames may share the same rotational frame. If, however, the translational frame is shifted by a fraction of a helical turn, it must be accompanied by a simultaneous rotation of the DNA. In such a case both the translational and rotational setting of the nucleosome will be altered.

Since the core DNA does not on average contain an integral number of base pairs per helical period (55), it seems unlikely that the rotational setting is graduated by the repeated interaction of phosphate groups, aligned on one side of the DNA helix, with the octamer surface over the whole length of the core DNA (246). This assumption is supported by the finding that the helical period of the core DNA is not necessarily uniform throughout the whole core (194, 246, 251) and that the average period can differ from 10,2bp to 10,5bp amongst cores on an 860bp fragment (55)*.

* The equivalence in nuclease periodicity and helical period requires the angle of attack of the nuclease to remain

This heterogeneity does not, however, imply that the rotational setting can be altered in a continuous, smooth fashion: the increments by which the rotational frame can be altered may be governed by the combination of a requirement for an unique rotational orientation of the DNA molecule at positions ±0,5 and ±3,5 to allow the appropriate atomic arrangements for the interaction of the putative α-helixes of the H3 and H4 histones (208) with the DNA molecule in the minor groove at these positions, and the proper orientation of phosphate groups relative to the octamer surface in distal regions of the core DNA, where the helical period may approach 10,0bp per turn (246). In these distal regions an integral helical period would also allow repeated interactions, stabilized by divalent cations, between phased phosphate groups on adjacent superhelical turns, as noted by Klug and Lutter (119). It seems reasonable to conclude that the gradation of rotational settings may be defined by the combination of these interactions stabilizing the DNA helix on the octamer surface and the helical period of the core DNA at these points of interaction.

It follows from the discussion above that the experimental determination of only the rotational frame of a nucleosome does not unambiguously define the translational frame, and if the translational frame is known, it is not possible to deduce the rotational frame without knowledge of the twist

---

constant relative to the octamer surface along the whole length of the core DNA. Klug and Lutter (119) argued for a change in the nuclease attack angle due to the steric influence of the two superhelical turns on each other, and the difference in the structural environment sensed by the nuclease in the central region of the core DNA in the vicinity of the dyad. If it is assumed, however, that the *change* in the angle of attack in the cores representing the two extremes of nuclease periodicities mentioned in the text, above, are identical, the difference in cutting frequency must represent a difference in the helical period. Moreover, Drew and Calladine (55) proposed that the change in nuclease attack angle may be less variable for several cores linked in series than for isolated core particles.

in the core DNA. Furthermore, this will only be possible if the borders of the translational frame are reported in terms of rotationally unique features, such as the sequences underlying the minor groove oriented towards the octamer surface of the proximal helical turns of the nucleosomal DNA.

Nucleosome positioning (reviewed in references 217, 233 243 and 246) refers to the phenomenon where a nucleosome is placed in a unique translational and rotational setting relative to the underlying DNA sequence. Several examples are listed in Table 1.

The experimentally determined location of a nucleosome represents the average position of nucleosomes positioned in several closely spaced translational and rotational settings over the sequence of interest in a large population of DNA molecules (233). These closely spaced settings may be assumed to differ in the free energy change associated with nucleosome accommodation (200, 243). A Boltzmann distribution (56) of nucleosomes in the various possible frames is therefore likely to exist in the total population of DNA molecules where the major sub-population will be constituted by the nucleosome in the most favorable translational and rotational frame (199). Such a nucleosome is operationally defined as positioned (233, 246). A group of sub-populations representing different positioning frames not differing much in the free energy change of nucleosome accommodation will result in the positioning of the nucleosome in overlapping frames (135). If there is no clear energetically most favorable frame, the positioning of a nucleosome over the region of interest is likely not to be experimentally evident (69), in which case the position is defined as random (233).

Table 1: Naturally occurring DNA sequences on which positioned nucleosomes or histone cores have been identified.

DNA source*	DNA sequence	Ref.
Eukaryotic	5S rRNA gene *Lytechinus variegatus*	220,221,69
Eukaryotic	β^A-globin gene promoter region chicken	117
Eukaryotic	histone gene spacers H2A-H1 and H1-H4 *Psammechinus miliaris*	203
Eukaryotic	234bp satellite repeat mouse	135
Eukaryotic	Section of α-satellite DNA African green monkey	175
Eukaryotic	250bp α-satellite fragment mouse	187
Eukaryotic	5S rRNA gene *Xenopus*	205,55
Prokaryotic	*tyrT* promoter fragment	59
Prokaryotic	145bp sequence (random selection)	200,199
Prokaryotic	234bp pBR322 fragment	135
Prokaryotic	203bp and 144bp fragments *lac* control region	42
Prokaryotic	pBR322 inserted into yeast	64
Prokaryotic	Section of tetracycline resistance gene	56
Viral	SV40	195,8
Viral	MMTV Long Terminal Repeat	194,193, 189
Yeast	*trp1,ars, URA3* in TRP1ARS1	235,232, 238
Yeast	centromere region of chromosome III and XI	21
Yeast	*pet56-his3-ded1* fragments	141
Yeast	*PHO5* acid phosphatase gene	18
Protozoan	5S rRNA genes macronuclei of *Tetrahymena*	187

* The association of nucleosomes with DNA has structural and functional significance only in eukaryotic systems. Studies involving the reconstitution of the histone octamer with DNA from sources other than eukaryotic have, however, contributed significantly to an understanding of the rules governing nucleosome positioning, and have been included.

1.4.2 Biological significance of nucleosome positioning:

The question of nucleosome positioning is central to an understanding of the relationship between the structural organization of eukaryotic chromatin and its function. The close association of the DNA duplex with the histone octamer surface may render recognition sequences inaccessible to direct approach from the solvent (194). Similarly, the strong electrostatic interactions between arginine and lysine side chains of the histones (109, 128) and the DNA duplex every helical turn (208, 251) may displace the equilibrium of conformational features of the duplex which may act as recognition elements (180). Conversely, the processive mechanisms of transcription and replication may cause the transient dissociation of nucleosomes from the DNA (23, 112, 138, 139, 140, 167).

This reciprocal relation between nucleosome placement and chromatin function has been demonstrated for several genetic processes. The initiation of transcription by bacterial SP6 RNA polymerase (138, 167) and T7 RNA polymerase (275) as well as mammalian RNA polymerase II (138) and *Xenopus* RNA polymerase III (167) is inhibited when the corresponding promoters form part of DNA complexed with a histone octamer, probably due to the inability of the polymerase to form an open complex with the DNA (138, 167). This interference with the normal function of DNA elements when covered by a nucleosome core was also demonstrated by Simpson (218) who reported a decrease in the copy number of deletion constructs of a yeast plasmid when the autonomously replicating sequence (ARS) was situated within a few helical turns of a core dyad axis. It was proposed (218) that the precise positioning of a nucleosome next to the ARS in the wild-type plasmid enabled unimpeded interaction of proteins with the ARS resulting in a normal high level of autonomous

replication. This absence of nucleosomal structures was also found on the centromeric regions of yeast chromosomes III and XI which contain elements important for protein binding during chromosome segregation (21)

The involvement of nucleosome placement as a regulatory mechanism in gene expression has been demonstrated *in vivo* for the yeast acid phosphatase gene (*PHO5*), where the removal of four positioned nucleosomes on the upstream activation site (UAS) is a prerequisite to activation of transcription (6). Apart from repressing the gene, the precise placement of these nucleosomes enabled their removal by two *trans*-acting regulatory proteins, PHO2 and PHO4 (65). The mechanism of nucleosome removal has been suggested to involve the binding of PHO4 to its *cis*-element which is exposed in the linker area between two positioned nucleosomes. This causes, in conjunction with PHO2, the removal of the downstream nucleosome and allows binding of another PHO4 molecule to an exposed site, eventually opening approximately 600bp of the promoter area of *PHO5* (6, 65). The down-regulatory effect of nucleosomally wrapped promoter on *PHO5* transcription was also directly shown by Han *et al.* (97), who reported the active transcription of *PHO5*, even under normally repressive conditions, in mutant yeast cells depleted of nucleosomes. Another level of gene regulation involving the presence of nucleosomes is the competition between transcription factor binding and chromatin assembly on nascent DNA after passage of the DNA replication complex (7), and between histone H1 and TFIIIA binding (205).

The positioned nucleosome on the mouse mammary tumor virus (MMTV) promoter represents a structure where the precise rotational orientation of the DNA in the core may be essential for transcriptional activity of the gene (194). The major grooves of two hormone responsive elements are exposed to the solvent, allowing binding of both

glucocorticoid and progesterone receptors to the core DNA, in contrast to the cognate sequence of nuclear factor I (NFI) where the major groove is orientated towards the octamer surface, impeding binding of NFI required for full transcriptional activity of MMTV (194). *In vitro* studies demonstrated an increased accessibility of the NFI recognition area to exonuclease III following steroid receptor binding, suggesting part of a transcriptional activation mechanism that might operate on this promoter (194).

It is evident, then, that the precise positioning or absence of nucleosomes in some regulatory areas is essential to normal genetic function, and that an understanding of the rules governing nucleosome positioning constitutes a fundamental section in the systematic study of chromatin biochemistry.

1.4.3 Sequence dependence of DNA-octamer affinity:

Not all DNA sequences can be accommodated in cores equally well (195). This differential affinity was first demonstrated by *in vitro* reconstitution (reviewed in 130) of the histone octamer with different synthetic polynucleotides. The homopolymers poly(dG).poly(dC) and poly(dA).poly(dT) could not be reconstituted into polycores (204, 219; but see 114 for an opposing view), whereas the heteropolymers poly(dG-dC).poly(dG-dC) (219), poly(dA-dT).poly(dA-dT) (204, 219) and poly(dA-dG).poly(dC-dT) (115) could. The inability of homopolymers to be reconstituted into polycores may be traced to the X-ray structures of single crystals of d(GGGGCCCC).d(GGGGCCCC) (151) and d(CGCAAAAAAGCG).d(CGCTTTTTTCGC) (173). McCall *et al.* (151) suggested, on the grounds of their crystal structure, that the base pairs in poly(dG).poly(dC) slide outwards to allow

overlapping of the five- and six-membered rings of adjacent heterocyclic guanines. In the crystal structure of the oligo(dA) tract containing oligomer, Nelson *et al.* (173) found that the (dA).(dT) base pair adopted a highly propeller twisted conformation which allowed the parallel stacking of the adenine residues on the one strand, and, similarly, the thymine residues on the other strand. The high propeller twist also allowed a set of bifurcated hydrogen bonds between the N_6 of adenine and the O_4 of two thymine residues: the one thymine in the same base pair and the other in the adjacent base pair (173). These bifurcated hydrogen bonds would confer a relative axial inflexibility to poly(dA).poly(dT) (173, 61). Thus, although both homopolymers appear axially rigid, poly(dA).poly(dT) may be regarded as less disposed to the bending required for the wrapping of the DNA around the histone octamer (61). It is interesting to note in this regard that Jayasena and Behe (114) reported the successful reconstitution of poly(dG).poly(dC) with the histone octamer and that $(dG)_n.(dC)_n$ with $n \leq 9$ were not statistically underrepresented in 177 core sequences (210). The inability of earlier studies (219, 204) to obtain assembled cores may lie in the different reconstitution methodologies employed (high salt exchange (114) *versus* salt dialysis (204 and 219)). The high initial salt concentrations employed in salt dialysis assembly may have caused the formation of tetrameric structures consisting of four poly(dG) strands (182) impeding proper core formation (204).

The histone octamer was also reconstituted with bacteriophage T4 DNA in which all the cytosine residues are replaced by hydroxymethylcytosine (154), as well as with poly(dG-m^5dC).poly(dG-m^5dC) (177) and with poly(rG-dC).poly(rG-dC) (115). Apart from assembling onto a wide range of polynucleotides, the histone octamer also associates with an immense variety of naturally occurring

DNA sequences, which are listed in Table 1. The affinity of the histone octamer for these naturally occurring sequences also differs, as demonstrated by Poljak and Gralla (195), who found an order of magnitude difference in the affinity of the histone octamer for different restriction fragments from simian virus 40 DNA.

Although this gross difference in ability of particular fragments to accommodate the histone octamer can, at a low level, contribute to nucleosome placement, the precise positioning of nucleosomes relative to the underlying base sequence would seem to require a more subtle mechanism.

1.4.4 Mechanism of nucleosome positioning:

1.4.4.1 Consensus sequence:

The precise interaction of proteins with DNA at a particular base pair sequence often involves the recognition of a consensus sequence, i.e., a series of nucleotides of limited extent and variability. The mechanism of recognition may involve direct hydrogen bond mediated interaction between a series of amino acid side-chains and a complementary series of electron acceptor and donor groups on the major groove side of base pairs (17) as in the binding of *434* repressor (2) or λ repressor (181) to their respective operators, or the recognition by the restriction endonuclease *Eco* RI of its restriction site (152). An alternative mechanism involves the recognition of a sequence dependent conformation (137) of the sugar-phosphate backbone of the DNA molecule, as exemplified by the binding of the *trp* repressor to its operator (180).

Initial attempts to elucidate the mechanism of nucleosome positioning, in the absence of a large experimental data base of positioning sequences, were directed towards the

identification of base sequences occurring at similar positions in different core particles (69) or displaying a symmetry across the dyad axis (220). An analysis by FitzGerald and Simpson (69) of sequence homologies in three positioned cores, of which two were 7bp and 17bp deletion constructs of the parental sequence, revealed no sequence or structural region occurring in the same area in all three cores. Satchwell *et al.* (210) came to the same conclusion after analysis of 177 core sequences. Ramsay *et al.* (200) noted that the absence of extensive contacts between the histones and the DNA bases, particularly in the major groove, argued against a direct sequence specific interaction.

In a deletion analysis of a positioning 145bp sequence, Ramsay (199) reported the continued unique positioning of the octamer until deletion of the original sequence, from either side independently, past the border of the central 40 base pairs, which offered support to the suggestion that nucleosome positioning is governed by a number of weak interactions along a length of DNA in the core (69, 135, 199 and 200). This proposal was also supported by Neubauer *et al.* (175) who reported that insertion of 5bp into the center of an α-satellite DNA fragment, which contained several overlapping positioning frames (135), resulted in cores positioning in frames that maximized contact with the parental frames (175).

These experimental findings, together with the large variety of sequences with which the histone octamer can associate, make the existence of a consensus sequence as the single determinant of nucleosome positioning improbable (135, 210, 243, 246) but argue, rather, for the involvement of the cumulative effect of several weak interactions between the octamer and the entire length of core DNA (175, 199). It was suggested that these interactions are permitted by the

distribution of short sequences in the core DNA at positions conducive to the bending of the DNA around the octamer (59, 175).

1.4.4.2 Rotational determinants:

A single superhelical turn in the core particle is bent through 360° over only 78bp (251) resulting in a 43Å radius of curvature (208, 251). This causes a structural deformation of the DNA clearly visible in the crystal structures of the core particle as a compression of the DNA grooves where they face in towards the octamer surface and an expansion where they face out. The groove widths between the inside and outside change from 8Å to 13Å for the minor groove, and from 12Å to 20Å for the major groove (208, 251). The determination of the structure of several oligonucleotides by X-ray crystallography (reviewed by Drew *et al.* (61)) has revealed a dependence of the local helix conformation on the base pair sequence. In particular, in runs of $(dG)_n.(dC)_n$ the minor groove is wide or can be expanded (151), whereas, in runs of $(dA)_n.(dT)_n$, the minor groove is narrow or can be compressed (173). Drew and Travers (59) had earlier proposed that the tight winding of the DNA around the histone octamer may determine the rotational orientation of the DNA by placing A+T rich segments where the minor groove faced in towards the octamer surface and must be compressed, while G+C rich segments are placed where the minor groove faced out, and must be expanded, insofar as this propensity complied with the bending constraint of the molecule as a whole.

This was tested by comparing the rotational orientation, determined by DNase I digestion, of a 169bp fragment from the *tyrT* promoter of *Escherichia coli* ligated into a circle or reconstituted into a core (59). The authors found that the circularly constrained molecule adopted a rotational

orientation placing three A+T segments where the minor groove faced in towards the center of the circle and two G+C segments were the minor groove faced out, and that an identical rotational orientation was assumed by the DNA molecule in the reconstituted core (59). This result was a strong indication that the rotational setting of a nucleosome is governed by the sequence dependent bending preference of the DNA molecule (59). The distribution of A+T and G+C segments in core DNA was further investigated by performing DNase I protection studies on purified chicken erythrocyte core DNA complexed with either distamycin, which selectively binds to runs of four or more (dA).(dT) base pairs, or chromomycin, which binds specifically to runs of four or more (dG).(dC) base pairs, both antibiotics protecting the DNA from nucleolytic cleavage ("statistical sequencing") (59). A strong modulation with a periodicity of approximately 10,2bp was found for the distribution of both antibiotics, with distamycin binding preferentially at positions where the minor groove would face in, and chromomycin where the minor groove would face out, confirming the angular disposition of A+T and G+C segments and supporting the proposal that the rotational frame occupied by a nucleosome is a function of the sequence dependent anisotropic flexibility of the DNA molecule (59).

The periodic distribution of di- and trinucleotides in core DNA was further investigated by Satchwell *et al.* (210). The authors performed Fourier analyses on the distribution of 10 unique dinucleotides and 32 unique trinucleotides in 177 cloned core DNA sequences aligned at their midpoints. A strong modulation at a period of approximately 10,2bp was found for several di- and trinucleotides, with the largest fractional variation in occurrence in the core DNA exhibited by ApA/TpT, ApApA/TpTpT and ApApT/ApTpT, present more often at positions where the minor groove faced towards the octamer surface, and GpC, GpG, TpG, GpGpC/GpCpC and

ApGpC/GpCpT, distributed with maxima at positions where the minor groove faced out (210). The effect of the sequence context of dinucleotide steps on their angular preference is exemplified by the lack of positional choice in core DNA exhibited by the isolated dinucleotide ApA, which showed a strong periodic modulation mainly because it constituted part of ApApA/TpTpT and ApApT/ApTpT. In fact, the weak modulation shown by ApApG/CpTpT placed this trinucleotide at positions where the minor groove would face out. Another significant feature of the periodic distribution of di- and trinucleotides was the phase inversion found in the central two helical periods where the DNA connects the two superhelical turns of the core particle. It is unclear how this feature fits in with the structural requirements present at the dyad (37, 210), but may reflect the requirements of H1/H5 binding (250). It is further important to note that the various short sequences do not exclusively occur at positions where the minor groove faces either in or out. Some, such as CpGpC/GpCpG or TpApT/ApTpA, showed a distribution that would place them at intermediate angular orientations (210).

All the experimental findings discussed above would strongly suggest that the periodic distribution of short sequences in phase with the helical period of nucleosomal DNA would define a preferred direction of flexibility for the DNA molecule and form an essential determinant for the rotational orientation of the DNA molecule relative to the octamer surface in nucleosomes (59, 210, 244, 245, 246).

This feature of nucleosomal DNA was used by Calladine and Drew (37) in the formulation of an algorithm designed to predict the position of nucleosomes given only the base pair sequence. The algorithm, in summary, consists of calculating the accumulated difference over a 124 base step "window" between a series of base pair roll values required

for uniform planer bending of a magnitude consistent with that of DNA in the core, and a series of roll values representing the particular dinucleotides present over the 124 base steps. These latter roll values were deduced from X-ray crystallographic data for each of the ten unique dinucleotide steps and adjusted to reflect the fractional variation in occurrence and sequence context found by Satchwell *et al.* (210). The calculation of this difference function over a 124 base step window, serially displaced along the sequence of interest, gave good fits when the center of the window approached the position of the dyad axis published for several positioning sequences (37, 55, 246). A second algorithm by Drew and Calladine (55), not making reference to any structural model but simply determining the accumulated likelihood of finding specific di- and trinucleotides at angular orientations governed by the serial application of a 124 base step window to the test sequence, gave similar good agreement between the predicted and experimentally determined rotational orientation of the histone octamer (55, 56, 117).

Abundant evidence is therefore available (59, 210) that the distribution of short anisotropically flexible sequences in phase with the helical period of nucleosomal DNA imparts a directional propensity to DNA bending and places a constraint on the rotational frames energetically available for nucleosome occupation (56).

1.4.4.3 Translational determinants:

As discussed in section 1.4.1, movement of the core DNA by one or more integral helical turns over the octamer surface alters the translational frame of the core, but not the rotational frame. It is conceivable that an extended length of DNA with an appropriate sequence periodicity matching the local helical periodicity of core DNA can accommodate the

core in several translational frames sharing the same rotational frame. Additional determinants must therefore exist to explain the precise translational placement of the histone octamer observed in several studies (55, 117, 194, 200, 220).

Several conformational features of the DNA duplex that may influence the translational placement of the octamer have been identified (245). The study by Satchwell *et al* (210) demonstrated an inversion in the phase of the distribution of several strongly modulated di- and trinucleotides over the two central helical turns of the core DNA. Although the structural element requiring this phase inversion is not clear (37, 210), it may reasonably be assumed that on a length of anisotropically flexible DNA containing a short section with a bending preference opposite to that of the rest of the molecule, the core will position with this area at the dyad to comply as closely as possible with the bending character of the DNA molecule as a whole (245, 246). Another feature of the sequence content of core DNA that highlights a possible translational determinant is the preferential placement of runs of (dA).(dT) longer than 5bp towards the extremities of the core (210) which may be understood in terms of the axial inflexibility of such sequences (173). This feature was not, however, as prominent in the core sequences of nucleosome dimers (211). The trinucleotides TpGpG/CpCpA and CpGpG/CpCpG similarly seem to avoid the central region of the core (55), whereas TpTpG/CpApA displayed a preference to be placed at the dyad (250). Sequences displaying a torsional flexibility have also been shown to be enriched in the linker areas and a specific region in the core (211).

An algorithm to predict the translational placement of the histone octamer was proposed by Drew and Calladine (55) based on the statistical distribution of particular

trinucleotides which were depleted in the central sections of cores reconstituted on a DNA molecule containing the 5S rRNA gene of Xenopus. This algorithm entails calculation of the accumulated likelihoods of finding these trinucleotides at particular positions over a 124 base step window serially displaced along the test sequence, but is not as successful as the rotational placement algorithms (see section 7.2).

It is evident, then, that *in vitro* the positioning of the histone octamer on relatively short pieces of DNA can be ascribed largely to the sequence dependent structural properties of the DNA duplex itself. *In vivo*, however, the DNA is not only packaged into nucleosomal arrays much longer than the oligocores typically studied *in vitro*, but these nucleosome fibres are additionally arranged in three dimensions into higher order structures (see section 1.2) and the DNA is involved in processes requiring the binding of regulatory proteins such as activators and repressors of gene expression. This full structural and functional participation of DNA *in vivo* may affect the positioning frames defined by the sequence dependent determinants, as is discussed below.

1.4.5 Determinants of nucleosome positioning *in vivo*:

1.4.5.1 Nucleosome-nucleosome interactions:

Drew and Calladine (55) demonstrated that an octamer reconstituted on a 150bp fragment from the *Xenopus* 5S rRNA gene *in vitro*, positioned in two rotational frames differing by approximately half a helical turn, whereas, when forming part of an 860bp fragment, the octamer positioned over this sequence in a single rotational frame different from both of those found on the short fragment. A similar observation was made on the placement of the histone octamer on a 154bp fragment and the core positioned over this region in a 440bp

gene in the opposite orientation, however, resulted in the four nucleosomes present on *TRP1* adopting precise positions (232). It was proposed that the hypersensitive site present *in vivo* at the 5' region of the *URA3* gene, that contains transcription regulatory and initiation sites, acted as a boundary to redirect nucleosome placement on the *TRP1* gene (232).

These boundary structures have also been suggested as a mechanism in the positioning of long nucleosomal arrays by a purely statistical mechanism in the absence of positioning signals inherent to the DNA sequence (124, 125). Experimental support for this proposal is found in the increased precision of nucleosome placement in long nucleosomal arrays closer to the chromosomal telomeres of *Oxytrichia* (86), in accordance with the mathematical treatment of the effect of boundary structures on nucleosome placement (125). DNA structures that may similarly act as border conditions include hairpin loops (176), Z-DNA (74, 177), and extended runs of (dA).(dT) (210) by virtue of their preferential location between nucleosome cores.

1.4.5.3 Chromatin folding:

Another mechanism that has been implicated in nucleosome positioning *in vivo* is the folding of chromatin into a three dimensional structure (238). This long range effect was demonstrated by the change caused by the insertion of the *URA3* gene into the TRP1 region on the nucleosome positions on a tandem insert of a section of the *URA3* gene in the *UNF* region of TRP1ARS1 (238). The tandem *URA3* insert is flanked on one side by nucleosomes that did not change their positions due to modification of the *TRP1* gene, and on the other side by a nuclease sensitive site that persists after TRP1 modification and presumably blocked linear transmission

of nucleosome-nucleosome interactions from the *TRP1* region to the tandem *URA3* insert (238).

1.4.5.4 Non-histone proteins:

Apart from the effect of hypersensitive sites, a modulating effect of non-histone protein binding on nucleosome placement was also directly demonstrated by the shift of a nucleosome upon binding of *α2* repressor to its operator in the TRP1ARS1 yeast plasmid *in vivo* (218). It is not clear by which mechanism *α2* repressor affects the placement of the nucleosome, which moves towards the *α2* operator upon repressor binding (218).

1.4.5.5 Sequence-dependent structural determinants:

Although several mechanisms have clearly been shown to influence the precise placement of the histone octamer *in vivo*, the favorable energy change accompanying the formation of nucleosomes on strongly anisotropically flexible sequences does seem to remain a dominant determinant of nucleosome positioning. This was shown by Linxweiler and Hörz (135) in a comparative study on *in vitro* and *in vivo* positioning on mouse satellite DNA, where all but two of the multiple overlapping frames present on a 234bp satellite monomer fragment *in vitro*, were also present *in vivo*. Thoma and Simpson (237) similarly reported that insertion of a 300bp tandem repeat of the 5S rRNA gene between two of the three precisely positioned nucleosomes on the *UNF* region of the TRP1ARS1 yeast plasmid, resulted in the formation of a nucleosome on one 5S DNA monomer at a position identical to that occupied *in vitro*.

All the above experimental evidence would support the conclusion that *in vivo* nucleosome positions are defined by several determinants acting simultaneously (238). In

regulatory areas the precise placement of nucleosomes may be biologically essential and governed by strong local sequence dependent determinants of the DNA duplex such as anisotropically flexible regions flanked by sequences favouring placement in the linker areas. Candidates for nucleosomes thus positioned would include those on the somatic 5S rRNA internal control region (205), the mouse mammary tumor virus promoter (194), the *PHO5* upstream activation site (18) and the nucleosome downstream from the autonomously replicating sequence of the yeast plasmid TRP1ARS1 (218). On structural genes precise positioning may be less important and nucleosomes may simply perform a packaging role. Here the observed positioning (21, 86) may be due to statistical placement between border structures such as nuclease hypersensitive sites, sequence specific non-histone proteins, or inflexible DNA structures modulating weak determinants present in fortuitous sequence periodicities.

1.5 Topology of DNA:

The helical repeat (h) of a linear, mixed-sequence DNA molecule in solution is equal to 10,5bp/turn under standard conditions (184, 206, 207, 260, 261). If the helical axis of such a DNA molecule is constrained to lie co-planer with a flat surface, the one strand will complete exactly two turns around the helical axis every 21bp. In the general case, a planerly constrained molecule will have N/h_0 such turns, where N equals the number of base-pairs in the molecule and h_0 the helical repeat defined relative to the surface normal under standard conditions when the DNA ends are free to rotate. If the free rotation of the ends of the DNA molecule is impeded by some anchorage structure or by circularization of the molecule, the number of turns can be changed by nicking and rotation of one strand around the helix axis followed by religation. Depending on the

magnitude and direction of rotation the one strand will revolve around the helix axis an integer number of times more or less, and this molecule will differ from the parental molecule in a topological quantity, the linking number (Lk), which is constant under continuous deformation and can only be altered by a strand breakage, rotation and closure event (49, 73, 266). If the resulting Lk is less than the linking number of the identical molecule where $h=h_0$ (Lk_0), the former molecule is negatively supercoiled or underwound, and if the Lk is more, positively supercoiled or overwound (14). The difference between Lk and Lk_0 is known as the linking difference (ΔLk), and the size-independent specific linking difference or superhelical density (σ) is the ratio of ΔLk to Lk_0 (48).

In solution DNA molecules are generally not constrained to flat surfaces and an insight into the linkage of such molecules requires a more formal approach. In mathematical treatments of the topology of DNA, Fuller (73) and White (266) have shown that the linking number is equal to the summation of two geometric properties of the molecule, the twist (Tw) and the writhe (Wr). This equality has been refined by White *et al.* (267, 269) in the equation $Lk=SLk+\Phi$ which expresses the linking number in parameters more suited to experimental manipulation. SLk is the surface linking number which is the sum of the surface twist (STw) and the writhe, and Φ is the winding number. The mathematical detail of this model of DNA topology (surface linking theory) has been discussed by White *et al.* (269) and reviewed by Cozzarelli *et al.* (48) and will not be elaborated on further here.

1.5.1 <u>The conformation of supercoiled DNA in solution</u>:

In solution supercoiled DNA, isolated as covalently closed circular molecules or prepared by enzymatic means, exists as

a population of topoisomers due to thermally induced torsional fluctuations of the DNA prior to ring closure (197). The mass proportioning through such a topoisomer population follows a Boltzmann distribution which indicates a quadratic dependence of the free energy change of supercoiling on the linking difference (197). In particular, $\Delta G=K(\Delta Lk)^2$ (50, 197) where K is a constant dependent on the number of base-pairs in the DNA molecule (103, 197).

Although the linking number is a topological property of the molecule and invariant under conditions of continuous deformation, it is proportioned between twist and writhe at a ratio determined by the ionic (9) and thermal (9, 50) environment of the molecule, as well as physical distortions of the DNA helix caused by binding of intercalative drugs (215) or of proteins such as the histone octamer (81) or DNA gyrase (201). This property of the topology of DNA has been exploited in the electrophoretic separation of topoisomers (215) and the preparation of both positively (43) and negatively (24) supercoiled DNA.

In solution a circular DNA molecule can assume either of two topologically equivalent forms: a left-handed toroidal shape, or a right-handed plectonemic (interwound) shape (14). Brady *et al.* (28, 29) proposed on the ground of X-ray scatter data that the toroidal form dominates, whereas electronmicrographs would indicate a predisposition towards a plectonemic shape (22). The criticism that sample preparation methodology in electronmicroscopic studies may force a structural preference was avoided by Adrian *et al.* (1) who performed cryo-electron microscopy on DNA molecules suspended in a vitrified layer. In that study it was shown that the DNA assumes a right-handed plectonemic shape (1). Further evidence, independent of any structural transition that may occur during sample preparation, comes from

electron microscopic studies of recombination products, where the inversion of substrate by λint, which conserves the handedness of the substrate, produced knotted structures which were exclusively right-handed (positive nodes)* (225). Apart from the sign of the nodes, the number of nodes observed in knotted torii lends additional support to the existence of the interwound form in solution (225). It has been proposed that the plectonemic shape is preferred by free DNA due to the larger pitch angle (less DNA flexure) and the smaller twist (14, 48) compared to the toroidal form. This is seen from the dependence of writhe on the superhelix pitch angle (γ) related by $-n(\sin\gamma)$ for plectonemic DNA and $-n(1-\sin\gamma)$ for toroidal DNA (48). Thus, as γ increases, the required compensatory contribution of Tw will decrease for plectonemic DNA, but increase for toroidal DNA. Therefore, qualitatively, plectonemic supercoils allow the DNA to reduce flexural tension without imparting excessive torsion to the DNA duplex. Several experimental studies have found that $\gamma=54°$ (1, 22) independent of the ΔLk of the circular molecule (22). This may represent a value where the energies of bending and twisting are at a minimum (48).

1.5.2 Nucleosomes and linking numbers:

The change in linking number accompanying nucleosome core formation has been determined at approximately -1,1 linking units per core by a combination of electrophoretic and electron microscopic studies (81, 221, 283). This was thought to be incompatible with the 1,8 turns of the DNA duplex round the histone octamer reported in the X-ray crystallographic study of Richmond *et al.* (208), and came to

* The equivalence of positive nodes and a right-handed helical structure is valid only for knotted recombination products. Resolution of a right-handed interwound helix to form catenated products retains the node sign as negative. See Spengler *et al.* (225).

be known as the "linking number paradox". The resolution of the paradox, as first proposed by Finch and Klug (68), lies in the decrease of the helical period from 10,5bp/turn (184, 206, 207, 260, 261) to approximately 10,2bp/turn following core formation (55, 56, 59, 210, 251), as demonstrated by White and colleagues (48, 267, 268) and Travers and colleagues (120, 245, 246). Following the argument of these two groups and using the notation of White and Bauer (267), the change in linking number, ΔLk, equals the summation of the surface linking number, SLk, and the change in the winding number, $\Delta\Phi$ (SLk of relaxed DNA is zero). In toroidally supercoiled DNA, SLk simply equals the signed number of superhelical turns (n), which is -1,8 for the 146bp core particle (208, 251). The change in winding number of the DNA free in solution to wrapped around the histone octamer is (146/10,2)-(146/10,5) which equals +0,41. Therefore, ΔLk of core formation equals -1,8+0,41 (=-1,39). If the helical repeat of core DNA is closer to 10,0bp/turn (Lutter), ΔLk=-1,1. These calculations do not incorporate a contribution from linker topology currently unknown (48, 246). A value close to -1,6 found for the change in linking number of H5 containing particles (282) may be obtained as calculated above with SLk=-2 and $\Delta\Phi$=+0,46.

1.5.3 Average superhelical stress *in vivo*:

In prokaryotes, episomal DNA segments are usually circularized and the chromosome arranged into topologically isolated domains (223). The eukaryotic genome is organized into similar topologically sequestered loops both in its condensed state during metaphase (183) and decondensed state during interphase (16) (see section 1.3). In the absence of a swivel mechanism allowing the ordered occurrence of a strand breaking, rotation and religation event, the DNA in such a circular molecule or loop would be unable to generate or dissipate any torsional stress by a linkage change. This

swivel mechanism is supplied by the topoisomerase enzymes (reviewed in 78, 262 and 263). In prokaryotes two types of topoisomerases cooperate to maintain the DNA within a narrow topological range: topoisomerase I, which relaxes negatively supercoiled DNA (but can relax positively supercoiled DNA in the presence of a single-stranded region), and DNA gyrase (a type II topoisomerase) which removes positive supercoils and introduces negative supercoils linked to the hydrolysis of ATP (78). In eukaryotes two topoisomerases have been identified: topoisomerase I and II, which will relax both negatively and positively supercoiled DNA by 1 or 2 linking units respectively (78). The latter topoisomerase is a major constituent of the nuclear scaffold (77).

The concerted diametric actions of topoisomerase I and DNA gyrase in prokaryotes maintain a homeostatic superhelicity in bacterial plasmids with $\sigma \approx -0,06$ in purified plasmids as evidenced by the inhibition of either topoisomerase *in vivo* (63, 83). This does not necessarily reflect the torsional and flexural state of the DNA *in vivo* where it is complexed with two dimers of the HU protein per 60bp (31, reviewed in 191). Pettijohn and Pfenninger (190) demonstrated that plasmids isolated from *E. coli* cells where religation of γ-irradiation nicked plasmids was allowed in the absence of gyrase activity, retained σ at -0,036, which indicated that about 60% of the superhelicity in the plasmids was restrained, probably by bound proteins including the HU protein (190, 223). This finding is supported by studies on DNA structural transitions *in vivo*. Negative superhelicity stabilizes several non-B-DNA secondary structures, including hairpin formation, cruciform extrusion, and Z- and H-DNA flips (reviewed in 72). The absence of cruciform (222) and Z-DNA (224) structures *in vivo*, which occur readily in the identical purified plasmids *in vitro*, would suggest an uncurbed superhelicity less than required to facilitate these transitions *in vivo*. Although no direct measurement

of the free superhelical stress in the bacterial chromosome has been reported, a comparison of the ratio of promoter strengths *in vivo* to that found *in vitro* under different degrees of supercoiling (see below), suggest a value of -0,048 (25), which is similar to that reported for bacterial plasmids (190).

In eukaryotic organisms most of the negative superhelicity appears to be restrained due to the binding of nucleosomes. Giaever *et al.* (83) reported that the linking difference of the 2μm plasmid from yeast changed little in response to manipulations of the intracelluler level of topoisomerases, and Sinden *et al.* (223) reported that the rate of trimethylpsoralen photobinding to DNA, which is proportional to superhelical density *in vitro*, was independent of γ-irradiation induced nicking, indicating that most of the superhelical stress is restrained in eukaryotes. It must be emphasized that these results reflect the average superhelical stress, and do not indicate the absence of transient and localized areas of superhelically strained DNA, as discussed below.

1.5.4 Biological significance of supercoiling:

1.5.4.1 Prokaryotes:

The efficiency of several genetic processes in prokaryotes are influenced by negative superhelicity. Borowiec and Gralla (24) demonstrated that the rate of open complex formation on the *lac* p^S promoter increased to 36 fold at σ = -0,065 compared to the relaxed template, followed by a decline at higher negative superhelicities. Although this can in part be ascribed to torsional stress promoted melting of the TATA box region at -10, which became S1 nuclease sensitive in the *tyrT* promoter under conditions of negative supercoiling (62), all three elements of bacterial promoters

appear to be involved in mediating the effect (25). The additional contribution may come from the writhe induced bending (51) of the promoter area (25, 244). Indeed, Collis *et al.* (45) have shown that synthetic curved sequences with little sequence homology to natural promoters except in the Pribnow box area acted as promoters for *E. coli* RNA polymerase, and that sequence alterations reducing this curvature between sites -35 and -10 caused a significant decrease in promoter activity. It has further been shown that different promoters reach optimum activities at different superhelical densities, suggesting that superhelical stress may modulate gene expression (25). This was in fact demonstrated for transcription of the topoisomerase I and DNA gyrase genes. As expected from the opposing roles of these two enzymes in maintaining an average superhelical density in prokaryotes, transcription of *topA* is activated by high negative superhelicities, and the *gyrA* and *gyrB* genes by a low negative superhelicity (159, reviewed in 196).

Apart from the modulating effect of superhelicity on transcription, it has also been shown to influence site-specific recombination catalyzed by *Tn*21 resolvase (41) and λint (165), as well as the formation of looped structures, which are especially favoured by the plectonemic form of DNA in solution (22), allowing simultaneous binding of the *lac* repressor to its operator and pseudo-operator separated by 93bp (26). The optimum distance between two *lac* operators for simultaneous binding by the *lac* repressor was further shown by Krämer *et al.* (127) to depend on the degree of superhelicity, suggesting that torsionally induced twist changes alters the relative rotational orientation of the operator sequences, modulating the stability of loops bridged by *lac* repressor at their bases. Supercoiling has also been implicated in the initiation step of DNA replication (79).

The genetic processes discussed above are not only responsive to the topological state of the DNA but some also influence the superhelicity of the DNA molecule on which they occur. Liu and Wang (136) proposed that the viscous drag exerted by solvent molecules on the nascent RNA chain and associated proteins may impede rotation of the RNA polymerase complex round the DNA, leading, due to the handedness of the DNA duplex, to the generation of positive supercoils ahead of the transcription complex, and negative supercoils behind it. This supercoiling action was suggested to continue to a point where the superhelical torsion equals that of rotational friction at which point the transcription complex rotates round the DNA molecule (136). This proposal was confirmed by both *in vitro* and *in vivo* studies and a dependency on the length of the transcript and the relative orientation of complexes simultaneously transcribing from different promoters was demonstrated (179, 247). The direction of transcribing complexes becomes important in circular molecules due to one-dimensional diffusion of the superhelical stress through the DNA molecule and destruction of this stress where positively and negatively supercoiled domains meet (136). Note that the generation of this localized superhelicity does not involve a linkage change and will dissipate after termination of transcription if it is not captured by some other means such as a selective linkage change in either of the two domains (82, 247).

The generation of these supercoiled domains may further influence genetic regulation by inducing structural transitions in the DNA helix (136). Such a transcription dependent transition *in vivo* was reported by Rahmouni and Wells (198) who demonstrated a B-to-Z transition in a $d(C\text{-}G)_6.d(C\text{-}G)_6$ stretch when present upstream (but not downstream) of the *tet* gene in a bacterial plasmid.

1.5.4.2 Eukaryotes:

Benyajati and Worcel (16) demonstrated that in eukaryotic chromatin superhelical stress can be localized to a loop (see section 1.3). This suggests that the genetic activity of a single loop may be modulated by confined superhelical stress (75). Although several nucleosome-free eukaryotic genes (212) and SV40 (108) responded to torsional stress with a structural transition in their 5' areas *in vitro*, and a correlation between superhelical stress and transcription of 5S rRNA genes injected into frog oocytes (209) was directly demonstrated, it is not known whether such a mechanism is utilized by the eukaryotic cell. It is clear that superhelical stress generating processes do occur in these looped domains, as underscored by the association of topoisomerase II with the loop bases. One such process is transcription, which was shown to supercoil yeast plasmids *in vivo* (30, 82) by a mechanism suggested to be identical to that of prokaryotes (82, 136; see above). The extent to which these localized areas of supercoiling may influence additional genetic processes on the same chromatin loop depends on the rate of dissipation of the superhelical stress, either by diffusion through the chromatin loop (82, 136), or by relaxation by topoisomerases (30, 82). In eukaryotes the utilization of localized superhelical stress to stabilize structural transitions in the DNA duplex would further depend on the extent to which this stress can be propagated through nucleosomal DNA, should the sequences facilitating such structural transitions be nucleosomally wrapped. The degree to which torsional stress can be accommodated by nucleosomal DNA seems histone specific: the amount of thermal untwisting allowed by yeast cores suggests that the DNA equivalent of a full superhelical turn can respond to temperature changes with an alteration in twist similar to that of free DNA (168). This is in contrast to

the resistance to thermal unwinding displayed by mammalian (143) and chicken (168) core DNA. Apart from thermally induced twist changes, little is known about the response of nucleosomal DNA to superhelical stress. A comparison of the designated sites of maximal DNase I cleavage in a core reconstituted onto a 208bp linear fragment (194) or onto a 212bp circularized fragment with $\Delta Lk=-2$ containing the same positioning sequence (193), revealed a decreased helical winding of 2bp to 3bp over the last three helical periods. This would indicate that mammalian cores respond to excess negative superhelical stress with a twist reduction in the terminal core regions. The formation of a single core on a circular molecule of 212bp with a linking difference of -2 would cause a specific linking difference of -0,157 in the remaining free DNA if 146bp is accommodated by the core. It is not known whether such a degree of superhelical stress is ever present in chromatin loops. Rahmouni and Wells (198) have demonstrated that the local superhelical density can reach a value of -0,038 behind the transcription complex in prokaryotes. The related value in eukaryotes has not been determined. Currently available evidence would therefore suggest that cores are stable at the levels of excess negative supercoiling present *in vivo* (87, 172, 193).

The same is not true for positively supercoiled DNA however. Goulet *et al.* (87) and Negri *et al.* (172) reported that octamer reconstitution onto small DNA rings proceeded negligibly if this resulted in the introduction of nett positive supercoils into the remainder of the molecule. Clark and Felsenfeld (43), however, reported successful reconstitution onto positively supercoiled molecules of 3293bp where the number of associated cores would result in $\sigma=+0,57$ in the linker areas. The size of this plasmid might have allowed the accommodation of this level of positive superhelical stress by writhe changes not possible in the very small DNA molecules mentioned above.

The cores formed on large positively supercoiled plasmids seemed structurally identical to those formed on negatively supercoiled DNA as judged by CD spectra and the results of crosslinking studies (43). An important difference, however, was the migration of the histone octamer from assembled positively supercoiled plasmids to free negatively supercoiled molecules under conditions of elevated salt (43). Pederson and Morse (186) suggested that the twin supercoiled domain generated by a transcribing RNA polymerase (82, 136) may facilitate transcription by the displacement of an octamer from the positively supercoiled domain in front of the complex, where dissolution of the octamer would introduce one negative supercoil and decrease the superhelical stress, and reassociation of an octamer with the negatively supercoiled domain behind the polymerase molecule, where assembly of the octamer would introduce one positive supercoil and once again decrease the superhelical stress. A similar model was subsequently proposed by Clark and Felsenfeld (43). An investigation into topological changes accompanying transcription on yeast plasmids failed to reveal any linkage change which might result from octamer displacement, but this result may be ascribed to a rapid migration compared to the reaction rate of the topoisomerases (186).

1.6 <u>The purpose of this investigation</u>:

It is clear from the material reviewed above that the structure and properties of the fundamental genetic packaging unit of eukaryotes is understood in considerable detail. The crystal structure of the core particle has been solved (208, 251) and the relative spatial arrangement of the constituent histones (121, 155) and the path of the core DNA over the octamer surface largely defined (15, 121, 163). The folding of the nucleosomally wrapped DNA into higher

order structures has been investigated in detail (33, 68, 157, 236, 273) and structural models proposed (68, 156, 236, 273, 274, 276). The rules governing the precise placement of the histone octamer relative to the underlying DNA sequence have been formulated (37, 55, 59, 175, 200, 210, 220) and the biological significance of positioned nucleosomes demonstrated (65, 194, 218). However, the effect of DNA topology on nucleosomal core formation and positioning has remained largely unexplored. This is an eminently relevant area due to the established effect of negative supercoiling on various genetic processes (108, 209, 212), the demonstrated topological sequestration of chromatin loops in eukaryotes (16, 183) and the demonstrated generation of superhelical stress by transcription *in vivo* (82).

This study has therefore addressed the effect of negative superhelical stress on the efficiency of core formation and the positioning of the histone octamer. Specific questions asked were: i) how does the efficiency of reconstitution onto negatively supercoiled DNA compare to that of linearized DNA, ii) what is the influence of reconstitution method on this efficiency, iii) what is the effect of the topological state of the DNA molecule on positioning of the histone octamer, and iv) does the rotational orientation of a positioning fragment relative to the rest of a circular molecule influence the positioning of the octamer?

CHAPTER 2

2 Materials and Methods.

All chemicals were analytical grade. Water was deionized double glass-distilled.

2.1 Plasmid construction, propagation, and isolation.

The 55bp *Eco* RI-*Hind* III fragment containing the polylinker region of pUC18 was isolated from an 8% acrylamide gel, purified, made blunt-ended with the Klenow fragment of DNA polymerase and the appropriate deoxyribonucleotides, and ligated into the *Sma* I site of the plasmid pGV403 (258). The resulting plasmid was denoted pHP5. This plasmid was propagated in *Esherichia coli* strain JM109 cells. It was found that propagation of pHP5 in *E. coli* strain HB101 resulted in complete deletion of the polylinker region.

The plasmid pHP2 was derived from pHP5 by ligation of a 199bp *Afl* III-*Hind* III fragment from the H1-H4 spacer region of h22 (see figure 2.1), containing the entire early histone gene battery of *Psammechinus miliaris* (kindly provided by Prof M Birnstiel), into the *Sma* I site of pHP5. The 199bp fragment was isolated from an 8% acrylamide gel, purified, and filled-in with the Klenow fragment of DNA polymerase and the appropriate deoxyribonucleotides prior to ligation into pHP5. The partial nucleotide sequence of pHP2 is given in Figure 2.1.

The plasmid pHP2R was obtained from pHP2 by cleavage at the *Xba* I site, filling-in with the Klenow fragment of DNA polymerase and the appropriate deoxyribonucleotides, and recirculization. This manipulation should have introduced an additional four base-pairs into the pHP2 plasmid, but

```
        10          20          30          40          50
ACGCCAGCAA CGCGGCCCGA AGCTTATCGG TGACCCTGAC TAAGTCGAGC
                    primer 2
        60          70          80          90          100
CCAGCTTGCA TGCCTGCAGG TCGACTCTAG AGGATCCCCA GCTTATAATC

        110         120         130         140         150
ATCCTTATAC ACGCGCAGTC GATGAGATGA AAAGTTCATT AACGCTACAT

        160         170         180         190         200
TTACAGTGTT TTGGGCAATT CTCCCTCCCC CCCCCCCTCT CTCTCTCTCT

        210         220         230         240         250
CTCTCTCTCT CTCTCTCTCC CTTCCTCTAA ATATGTTGAA TTGAGTGTGT

        260         270         280         290         300
TTCATGTCAG GAGTGGTATA ACTCTCTTTT CCTCTACGTG GGGTACCGAG

        310         320         330         340         350
CTCGAATTGG GCTCGACTCA GTCAGGGCGA TGATAAGCTG TCAAACATGA
                                primer 1
```

Figure 2.1: Partial nucleotide sequence of the Watson strand of plasmid pHP2. The numbering origin of plasmid pGV403 is retained. The positions at which primer 1 and primer 2 anneal are underlined. Primer 1 anneals to the Watson strand and primer 2 anneals to the complementary Crick strand. The sequence of the *Afl* III-*Hind* III fragment obtained from the H1-H4 spacer of h22 (see text), is enclosed in square brackets.

sequencing the plasmid revealed the presence of only three additional bases. This may have resulted from the exonucleolytic removal of the terminal d(G.C) base-pair by a contaminating nuclease present in one of the modification steps. The partial nucleotide sequence of plasmid pHP2R is given in figure 2.2.

The plasmid pHP2RR was derived from pHP2R by filling-in the restricted *Asp* 718 site with the Klenow fragment of DNA polymerase and the appropriate deoxyribonucleotide triphosphates, followed by religation. The partial nucleotide sequence of pHP2RR is given in figure 2.3.

The modifications were initially confirmed by restriction enzyme analysis, and subsequently nucleotide sequencing.

All cloning techniques were according to Maniatis *et al.* (146).

The three plasmids from the pHP2-series were all propagated in *E. coli* HB101. The plasmids were isolated by the Triton-lysis method (12) from un-amplified 16h cultures in LB containing the antibiotic chloramphenicol at 25μg/ml. Supercoiled plasmid was isolated by banding the plasmid preparation twice in cesium chloride/ethidium bromide gradients. The purified, supercoiled plasmid was resuspended in TE (pH8,0) and stored in small aliquots at -20°C. This final preparation contained <5% concatenated dimers and nicked molecules.

2.2 Synthesis, purification and end-labeling of the 20mer oligonucleotide primers.

Two 20mer oligonucleotide primers were synthesized on a Beckman system 1+ DNA synthesizer, purified by established

```
        10          20          30          40          50
ACGCCAGCAA CGCGGCCCGA AGCTTATCGG TGACCCTGAC TAAGTCGAGC
                        primer 2
        60          70          80          90          100
CCAGCTTGCA TGCCTGCAGG TCGACTCTAC TAGAGGATCC CC[AGCTTATA
                          ***
        110         120         130         140         150
ATCATCCTTA TACACGCGCA GTCGATGAGA TGAAAAGTTC ATTAACGCTA

        160         170         180         190         200
CATTTACAGT GTTTTGGGCA ATTCTCCCTC CCCCCCCCCC TCTCTCTCTC

        210         220         230         240         250
TCTCTCTCTC TCTCTCTCTC TCCCTTCCTC TAAATATGTT GAATTGAGTG

        260         270         280         290         300
TGTTTCATGT CAGGAGTGGT ATAACTCTCT TTTCCTCTAC GT]GGGGTACC

        310         320         330         340         350
GAGCTCGAAT TGGGCTCGAC TCAGTCAGGG CGATGATAAG CTGTCAAACA
                                  primer 1
```

Figure 2.2: Partial nucleotide sequence of the Watson strand of plasmid pHP2R. The numbering origin of pGV403 is retained. The positions at which primer 1 and primer 2 anneal are underlined. Primer 1 anneals to the Watson strand, whereas primer 2 anneals to the complementary Crick strand. The three bases introduced at position 77-79 by filling-in of the *Xba* I site of pHP2 are marked by asterisks. The *Afl* III-*Hind* III fragment obtained from the H1-H4 spacer of h22 (see text), is enclosed in square brackets.

```
        10          20          30          40          50
ACGCCAGCAA CGCGGCCCGA AGCTTATCGG TGACCCTGAC TAAGTCGAGC
                    primer 2
        60          70          80          90          100
CCAGCTTGCA TGCCTGCAGG TCGACTCTAC TAGAGGATCC CCAGCTTATA

        110         120         130         140         150
ATCATCCTTA TACACGCGCA GTCGATGAGA TGAAAAGTTC ATTAACGCTA

        160         170         180         190         200
CATTTACAGT GTTTTGGGCA ATTCTCCCTC CCCCCCCCCC TCTCTCTCTC

        210         220         230         240         250
TCTCTCTCTC TCTCTCTCTC TCCCTTCCTC TAAATATGTT GAATTGAGTG

        260         270         280         290         300
TGTTTCATGT CAGGAGTGGT ATAACTCTCT TTTCCTCTAC GTGGGGTACG
                                                    ****
        310         320         330         340         350
TACCGAGCTC GAATTGGGCT CGACTCAGTC AGGGCGATGA TAAGCTGTCA
                                  primer 1
```

Figure 2.3: Partial nucleotide sequence of the Watson strand of plasmid pHP2RR. The numbering origin of plasmid pGV403 is retained. The positions at which primer 1 and primer 2 anneal are underlined. Primer 1 anneals to the Watson strand and primer 2 to the complementary Crick strand. The four bases introduced at positions 296-299 by filling-in of the *Asp* 718 site of pHP2 are indicated by asterisks. The *Afl* III-*Hind* III fragment obtained from the H1-H4 spacer of h22 (see text), is enclosed in square brackets.

procedures (12), and stored in small aliquots at -20°C in TE (pH8,0). The nucleotide sequence of primer 1 is d(ACAGCTTATCATCGCCCTGA) and of primer 2, d(CGAAGCTTATCGGTGACCCT). The positions of the complementary sequences in pHP2, to which these primers would anneal, are indicated in Figures 2.1-2.3.

[γ-^{32}P]ATP was synthesized from ADP (Sigma) and 32-phosphorus (Amersham) by the method of Johnson and Walseth (116) to yield a product with a specific activity >5000Ci/mmole.

The primers were end-labeled at their 5'-ends by incubating 300pmoles of 5'-ends (2μg primer) in 50μl of 50 mM TRIS.Cl (pH7,6), 10mM $MgCl_2$, 5mM dithiothreitol, 0,1mM spermidine, 150μCi [γ-^{32}P]ATP and 20 units of T4 polynucleotide kinase for 30min at 37°C. After this period, the sample was extracted once with each of equal volumes phenol and chloroform, passed through a Sephadex G-50 spincolumn equilibrated with TE (pH8,0) to remove excess ATP, and precipitated with ethanol. The end-labeled primer, resuspended in 25μl TE (pH8,0), was routinely recovered at a specific activity of 3X10^6 DPM/μl (liquid scintillation).

2.3 Isolation of the chicken erythrocyte core particle and histone octamer.

2.3.1 Isolation of the histone octamer:

The histone octamer was isolated from 250ml fresh chicken blood by the method of Murray *et al.* (169) modified by Greyling *et al.* (90). Briefly, packed chicken erythrocytes, from which the Buffy coat had been removed, were carefully resuspended in 170ml SSC (140mM NaCl and 10mM trisodiumcitrate) and centrifuged in a benchtop centrifuge to gently pellet the erythrocytes. This step was

repeated (usually 5 times) until the supernatant was no longer pink, at which point the packed erythrocytes were resuspended in 170ml buffer A (15mM TRIS.Cl (pH7,4), 65mM KCl, 65mM NaCl, 0,15mM spermine, 0,5mM spermidine, 5mM 2-mercaptoethanol, 0,2mM EDTA and 0,5mM EGTA), transferred to a 500ml conical flask, and adjusted to 1% (v/v) Nonidet P40 and 0,1mM PMSF. The nuclei were then pelleted by a 5min centrifugation step at top speed in a benchtop centrifuge, the supernatant removed, and the nuclei resuspended in 170ml buffer A containing 0,1mM PMSF. The nuclei were washed four times in the above buffer, the optical density of a 1:100 dilution determined in 4M NaCl (OD_{260} $20cm^{-1}$ = 1mg/ml; 12), and resuspended in a volume of 50mM TRIS.Cl (pH7,4), 25mM KCl, 4mM $MgCl_2$, 1mM $CaCl_2$ and 0,2mM PMSF to a final concentration of 5mg/ml. Micrococcal nuclease was added (1000 units [absorbance units] per 10ml of nuclei) and the digestion allowed to proceed at 37°C for 45min and then quenched by addition of EDTA to a final concentration of 10mM and placed on ice. The nuclei were then centrifuged for 5min in a benchtop, and the pellet, after addition of 15ml 0,25mM EDTA (pH8,0), homogenized with four strokes of a tight dounce. The preparation was then dialysed overnight against 3l of 0,25mM EDTA (pH8,0). The recovered dialysate was centrifuged for 10min in a benchtop, and the supernatant (approximately 12ml at 10mg DNA per ml) applied to a 4,5X5cm hydroxylapatite column equilibrated in 10mM sodium phosphate and 0,1mM PMSF. The hydroxylapatite was prepared by the method of Bernardi (19). The column was eluted with 150ml sodium phosphate (pH7,0), 0,1mM PMSF followed by 500ml of 3M NaCl, 10mM sodium phosphate (pH7,0) and 0,1mM PMSF. The optical density of 2,5ml fractions were determined at 230nm, the slow eluting peak, containing the histone octamer and some dissociated histones, pooled, concentrated to 13mg/ml (OD_{230} $3cm^{-1}$ = 1mg/ml) on an Amicon with a P10 membrane, and applied to a 90X2,5cm Sephadex G-100 column equilibrated in 2M NaCl, 10mM sodium

phosphate (pH7,4) and 0,1mM PMSF. The column was eluted in the same buffer and the optical density of 2,5ml fractions determined at 230nm. The faster eluting peak, containing the histone octamer, was pooled and concentrated 50-fold in an Amicon device with a P10 membrane. The sample was made to 50% (v/v) glycerol and stored at -20°C. The concentration of the histone octamer was determined by performing amino acid analysis in triplicate with norleucine as internal standard, and the integrity of the octamer checked by DMS crosslinking and possible histone degradation monitored by electrophoresis on a 15% SDS-PAGE gel (data not shown).

2.3.2 Isolation of chicken cores:

The 146bp core particle was isolated from fresh chicken erythrocytes exactly as described by Drew and Calladine (55). The resulting preparation was stored in 50% (v/v) glycerol at -20°C, possible histone degradation monitored as above, and the concentration of the core particle determined at 260nm in 100mM NaOH (A_{260} of 10 = 1mg/ml (55)).

2.4 Reconstitution of nucleosome cores on plasmid DNA.

2.4.1 Poly[L-glutamate] reconstitution.

Poly[L-glutamate] facilitated assembly was performed as described by Retief *et al.* (203). An appropriate amount of histone octamer, to give the required molar octamer:DNA ratio, was mixed with a two-fold weight excess of poly[L-glutamate] (M_r 65 000; Sigma) in TE (pH8,0) at room temperature. A volume from a stock solution of NaCl that would give a concentration of 100mM NaCl after dilution of the sample to the final volume, was added, followed by dilution of the sample with TE (pH8,0) and addition of the plasmid DNA. The final DNA concentration was 100µg/ml. The

sample was then incubated at 37°C for 3h to allow proper assembly.

2.4.2 High salt exchange reconstitution.

High salt exchange assembly was performed as described by Drew and Calladine (55). Plasmid DNA (10μg) was mixed with 400μg core particles in 100μl of 700mM NaCl, 20mM TRIS.Cl (pH8,0), 0,2mM EDTA and 0,2mM PMSF. The exchange of the histone octamer from the core particles to the plasmid DNA was allowed at 37°C for the times specified in the text (nominally 20min), and the sample then equilibrated at room temperature for 10min The intermolecular migration of the octamer was then impeded by a gradual reduction of the NaCl concentration in the sample by the addition of 25μl volumes of 5mM TRIS.Cl (pH8,0), 0,1% (v/v) Nonidet P40 and 0,2mM EDTA every 10min at room temperature to a final volume of 1ml.

2.4.3 Urea/salt reconstitution.

Urea/salt dialysis reconstitution of nucleosome cores was performed as described by Camerini-Otero and Felsenfeld (40). Plasmid DNA (10μg) was mixed with an appropriate amount of histone octamer to give the required molar octamer:DNA ratio in 200μl of 5mM TRIS.Cl (pH8,0), 10mM 2-mercaptoethanol, 0,5mM EDTA and 0,2mM PMSF (dialysis buffer I) containing 5M urea and 2M NaCl. Step gradient dialysis was then performed for 80min periods against the dialysis buffer I containing 5M urea and decreasing NaCl concentrations of 1,2M, 1M, 0,8M and 0,6M. The urea was then dialysed out against dialysis buffer I containing 0,6M NaCl for 80min, and the preparation finally dialysed against dialysis buffer I overnight. All dialysis steps were performed at 4°C with continual agitation.

2.4.4 Salt dialysis reconstitution.

Salt dialysis reconstitution of nucleosome cores was performed as described by Hansen *et al.* (99). Plasmid DNA (10μg) was mixed with an appropriate amount of histone octamer (in 2M NaCl, 10mM TRIS.Cl (pH8,0), 1mM EDTA and 0,2mM PMSF) to give the required molar octamer:DNA ratio in 200μl of 10mM TRIS.Cl (pH8,0), 1mM EDTA and 0,2mM PMSF (dialysis buffer II) containing 2M NaCl. Step gradient dialysis was then performed at 4°C against dialysis buffer II containing the following decreasing amounts of NaCl and for the following times: 1,5M (4h), 1,0M (4h), 0,75M (3h) and 0,5M (3h). The preparation was finally dialysed to low salt against dialysis buffer II for a 12h period. All steps were at 4°C with continual agitation.

2.4.5 Native agarose gels.

The completeness of reconstitutions and possible aggregation were routinely monitored by the application of small aliquots (approximately 0,5μg DNA) to 0,7% (w/v) agarose gels in 45mM TRIS.borate (pH8,3), 45mM borate and 1mM EDTA (0,5X TBE) as described by Rhodes (205). The gels, with appropriate unreconstituted DNA controls, were electrophoresed in the same buffer at room temperature at 10mA for approximately 2h, and the DNA visualized on an ultra-violet transilluminator after staining in 0,5μg ethidium bromide/ml 0,5X TBE.

2.4.6 Sucrose gradient ultracentrifugation.

Reconstituted and unreconstituted plasmids were analysed on 5-20% (w/v) sucrose gradients in 10mM TRIS.Cl (pH8,0), 60mM NaCl, 1mM EDTA and 0,2mM PMSF spun in a SW40Ti (Beckman) rotor at 36 000rpm at 4°C for 5h. The

sedimentation profiles were determined by measurements of the optical density of the fractionated gradients at 260nm.

2.5 Topoisomerase I relaxation.

A 50µl volume containing 2µg of reconstituted DNA in 72mM KCl, 30mM TRIS.Cl (pH8,0), 5mM DTT, 5mM $MgCl_2$ and BSA (100µg/ml) was incubated in the presence of 8U topoisomerase I (Amersham) at 37°C for 60min. The sample was then adjusted to 0,5% (w/v) SDS and 50µg/ml proteinase K and incubated for 20min at 37°C. The sample was then deproteinised by extracting once with equal volumes phenol and chloroform, and the DNA precipitated from the sample by addition of 3M $NaCH_3COO$ to a final concentration of 300mM, followed by 2,5 volumes of cold (-20°C) ethanol (96%). The DNA was pelleted by centrifugation for 15min in a microfuge (Eppendorf) or a SM24 rotor (Sorvall) at 15 000rpm at 4°C. The DNA was further desalted by pelleting the sample resuspended in ethanol (70%) by centrifugation for 5min as above, dried in a rotatory evacuator, and finally resuspended in 20µl TE (pH8,0).

2.5.1 Preparation of topoisomer populations of different average linking difference.

A series of DNA topoisomer populations, differing in average specific linking difference, was prepared by the relaxation of supercoiled DNA with topoisomerase I in the presence of different concentrations ethidium bromide as described by Borowiec and Gralla (24). The 20µl samples containing 5µg of supercoiled DNA in 72mM KCl, 30mM TRIS.Cl (pH8,0), 5mM DTT, 5mM $MgCl_2$, BSA (100µg/ml) and ethidium bromide at 0µM, 12µM, 25µM, 37µM or 50µM were incubated with 16U topoisomerase I at 37°C for 3h. The samples were then extracted once with each of equal volumes phenol and

chloroform, ethanol precipitated, desalted and dried as described above, and finally resuspended in 12μl TE (pH8,0).

2.5.2 Electrophoretic resolution of DNA topoisomers.

DNA topoisomers were resolved on 1,5% (w/v) horizontal agarose gels electrophoresed at 3-4V/cm for 24-48h at 4°C in TBE containing 0, 1 or 3μg/ml chloroquine diphosphate (Sigma). Where topoisomers were resolved in two dimensions, the agarose gels, with a $5mm^3$ circular well placed 2cm from either edge in one corner, were run as above in the first dimension in TBE, soaked in TBE containing 0,5μg/ml chloroquine for 30min, rotated through 90°, and run in the second dimension in the chloroquine containing buffer, as described by Peck and Wang (185). The DNA was visualized by staining the gels in TBE containing 1μg/ml ethidium bromide for 30min, destained in 1mM $MgCl_2$ for 30min, and photographed on an ultra-violet transilluminator.

2.6 Nuclease digestions.

Prior to performing any nuclease digestion, the optical density of supercoiled and linearised plasmids assembled at different molar ratios of octamer:DNA were measured at 260nm in 100mM NaOH to ensure that equal concentrations of DNA were used in the experiments described below. An optical density of $20cm^{-1}$ was taken to equal 1mg/ml of DNA (12). When found necessary, the concentrations were adjusted by the addition of TE (pH8,0).

2.6.1 Micrococcal nuclease.

Micrococcal nuclease (MNase) digestions were performed on 4μg of DNA in 100μl volumes of 10mM TRIS.Cl (pH8,0), 1mM EDTA, 0,2mM PMSF and 2mM $CaCl_2$. The samples were equilibrated at 37°C for 15min before addition of 0,13U

micrococcal nuclease (absorbance units) and incubated for exactly 10min at 37°C. The digestions were quenched by addition of 2μl of 500mM EDTA, adjusted to 0,5% (w/v) SDS and 100μg/ml proteinase K, and incubated for an additional 20min at 37°C. The samples were deproteinised by extracting twice with equal volumes phenol and once with an equal volume of chloroform and the DNA purified by ethanol precipitation as described above. The DNA pellets were resuspended in 10μl TE (pH8,0) and the whole sample applied to a 10% (w/v) acrylamide gel electrophoresed at 8V/cm for 16h in TBE at 22°C. The gels were stained in TBE containing 1μg/ml ethidium bromide for 1h and photographed on an ultra-violet transilluminator. The products of MNase digestion of supercoiled and linearised plasmids of a particular experiment were always run on the same acrylamide gel to minimize errors in the comparison due to the electrophoresis and subsequent handling and photography of the gel.

2.6.2 DNase I.

DNase I was used as a probe for the location of the histone octamer on the DNA (55, 59, 119, 205). Because the effect of supercoiling on octamer positioning was investigated in some of the experiments, and DNase I, by nicking the DNA, removes this supercoiling, it is important that DNA molecules are mostly singly cut. Any subsequent cuts on the same molecule might reflect a shifted position of the octamer due to the removal of supercoiling. It is possible to obtain molecules that, when cut, are cut mostly once, by digesting the DNA very lightly with DNase I. This question is addressed in detail in section 5.1. Trial runs on all reconstituted samples were therefore always performed to obtain the DNase I concentration and digestion time required to nick less than 50% of the plasmids in the sample.

The trial digest was performed by dispensing a 50μl volume containing 2μg of supercoiled reconstituted or free plasmid DNA in 10mM TRIS.Cl (pH8,0), 1mM EDTA, 0,2mM PMSF and 2mM $MgCl_2$ into an eppendorf, and removing a 10μl aliquot for zero time. To the remainder, 1μl of DNase I (0,03U for reconstituted plasmids and 0,003U for free plasmids) was added and 10μl aliquots withdrawn at 10, 30, 60 and 120sec. The aliquots withdrawn were immediately mixed with 1μl volumes of 100mM EDTA to quench the digestions and then adjusted to 1%(v/v) sarkosyl, 5%(v/v) glycerol and 0,025%(w/v) bromophenol blue and electrophoresed in 1%(w/v) agarose gels in TBE at 100V for 30min at 22°C. The DNA was visualized under ultra-violet illumination after staining the gel in TBE containing 0,5μg/ml ethidium bromide. The time required for 40-50% nicking was determined by a visual comparison of the intensities of the bands corresponding to supercoiled and nicked circular DNA.

In the proper digest, a volume of 160μl of 10mM TRIS.Cl (pH8,0), 1mM EDTA, 0,2mM PMSF and 2mM $MgCl_2$ containing the reconstituted or free plasmid at a concentration of 40μg/ml was incubated at 22°C with DNase I (0,12U for plasmid reconstituted with the octamer and 0,012U for free plasmid) for the period determined in the trial digest (usually 30sec to 1min). After completion of the determined incubation period, the digestion was quenched by addition of EDTA to a final concentration of 10mM and adjusted to 0,5%(w/v) SDS and 100μg/ml proteinase K. The sample was incubated at 37°C for 20min and deproteinised by extraction once with equal volumes of phenol and chloroform. The DNA was purified by ethanol precipitation as described above, and resuspended in 100μl volumes of water. The extent of DNase I digestion was monitored by electrophoresis of an aliquot on an agarose gel as described above. The DNase I digestion of reconstituted or free linearised plasmid was performed under conditions identical to the

corresponding supercoiled plasmid sample. The linearization of the plasmid molecule following *Pvu* II restriction enzyme digestion was confirmed by agarose gel electrophoresis prior to assembly and DNase I digestion.

2.6.3 P1-nuclease digestions.

Structural transitions stabilised by negative supercoiling or low pH were investigated with the single-strand specific P_1-nuclease (20). Samples containing 1,5μg of plasmid in 50μl volumes of 50mM NaCl and 10mM $MgCl_2$ buffered with either 10mM TRIS.Cl (pH7,5) or 10mM $NaCH_3COO$ (pH4,5) depending on the required pH, were briefly incubated at 37°C followed by the addition of an appropriate number of units of P_1-nuclease indicated in the text. The reactions were allowed to continue at 37°C for exactly 10min and quenched by addition of EDTA to a final concentration of 40mM. Samples of DNA reconstituted with the histone octamer were extracted with equal volumes of phenol and chloroform and all samples dialysed against 20ml volumes of TE (pH8,0) on 25nm dialysis type VS filters (Millipore) in petri-dishes for 2h at 22°C. Recovered samples of supercoiled DNA were divided into two equal volumes, the one of which was digested with *Nco* I and the other with *Pvu* II under the conditions recommended by the supplier. Linearised plasmid was always obtained by restriction with one of these two enzymes prior to P_1-nuclease digestion. All samples were then extracted with equal volumes of phenol and chloroform, the DNA purified by ethanol precipitation as described above, and the resuspended samples applied in equal amounts to 2%(w/v) agarose gels (20X25cm) run in TBE containing 0,5μg/ml ethidium bromide at 150V for 2h at 22°C. The gels were photographed on an ultra-violet transilluminator.

2.7 Primer extension of DNase I digests.

Primer extensions were performed essentially as described by Gralla (88) and Borowiec *et al.* (26). Briefly, the concentration of DNA in DNase I digested samples in water was determined spectroscopically at 260nm in 100mM NaOH with an optical density of $20cm^{-1}$ taken to equal 1mg/ml of DNA (12). Samples (35μl) containing 1,5μg of DNase I digested plasmid were dispensed into 1,5ml eppendorf tubes and 1μl of 5'-end labeled primer 1 or primer 2 (approximately 30-fold molar excess over plasmid DNA) added followed by 4μl of a freshly prepared 10mM NaOH solution. The samples were heated at 80°C for exactly 2min and returned to ice for at least 5min. During this time the samples were adjusted to 50mM TRIS.Cl (pH7,2), 10mM $MgSO_4$ and 0,2mM DTT by a single addition from a freshly prepared 10X stock solution, heated at 45°C for 3min and returned to ice for at least 5min During this period 5μl of a solution containing each of the four deoxyribonucleotide triphosphates at 5mM was added, whereafter the samples were equilibrated at 52°C for 5min The Klenow fragment of DNA polymerase (1U) was added to each sample at 30 second staggered intervals and the polymerization of the nucleotides to the primer allowed at 52°C for exactly 10min After this period the extension reactions were quenched by addition, once again at staggered 30sec intervals, of 17μl of 4M NH_4CH_3COO and 40mM EDTA followed by 125μl of cold (-20°C) absolute ethanol. The DNA was allowed to precipitate for 15min at -70°C and pelleted in a microfuge (Eppendorf) for 10min at 4°C. The supernatants were carefully discarded and the pelleted DNA desalted by vortexing for 10sec with 200μl of 70%(v/v) ethanol. The samples were then returned to ice for 5min, followed by centrifugation in a microfuge for 2min at 4°C. The supernatants were once again carefully discarded, and the DNA pellets dried in a rotatory evacuator. The dried samples were resuspended in 3μl of 98%(v/v) formamide, 1mM

EDTA and 1mM NaOH and $2X10^5$CPM (Cerenkov) applied to individual lanes of a 6% acrylamide/50% urea gel (45X20X0,05cm) run at 1500V for approximately 1,5h with the temperature regulated at approximately 45°C by recirculating water from a waterbath through the cooling jackets. The gels were fixed in a mixture of 10%(v/v) methanol and 10%(v/v) acetic acid for 10min, washed in water for 10min, and dried onto Whatmann 3MM paper. The dried gels were exposed to Cronex X-ray film (pre-flashed twice on either side) backed with an intensifying screen (Kodak) at -70°C for 1 to 3 days as appropriate. Care was taken to obtain exposures within the linear range of the film.

Although the total number of Cerenkov counts applied per lane was identical, the large molar excess of 5' end-labeled primer (required to obtain a reasonable amount of extended primer) often complicated attempts to obtain even amounts of extended primer per lane. It was also found that primer extension consistently gave a lower resolution than directly end-labeled DNase I digested fragments (data not shown). This resolution could be improved by labeling primers with [γ-^{32}P]ATP within 24h after synthesis of the radionucleotide, and performing the extension of the 5' end-labeled primers, stored at -70°C, within 24h after end-labeling.

2.8 Data analysis and calculations.

2.8.1 Determination of the linking difference of a topoisomer population.

The lanes of agarose gels on which DNA topoisomers had been resolved were densitometrically scanned on the film negative. The peak heights corresponding to individual topoisomers were determined on a plot of the densitometric scans, and the offset of the Gaussian center relative to the

most abundant topoisomer of the distribution calculated with the algorithm of Kolb and Buc (122). The linking difference of a given topoisomer population relative to that of a completely relaxed population was determined by counting the number of topoisomer bands electrophoretically resolved between the most abundant topoisomer of the one population and (up to and including) that of the second population, and adding to this the offsets of the two populations as calculated above.

2.8.2 Calculation of the number of octamers assembled onto linearised DNA at different molar ratios of octamer:DNA.

The lanes of acrylamide gels of micrococcal nuclease digested supercoiled and linearised plasmids reconstituted with the histone octamer at different molar octamer:DNA ratios were densitometrically scanned on the film negative, and the scans stored and further manipulated in computer memory. A series of scans of either supercoiled or linearised molecules representing the experimental range of molar octamer:DNA assembly ratios were aligned at the position of the well on the acrylamide gel, and clipped to produce shortened scans, each containing 991 datapoints, corresponding to the range of DNA fragments from 90bp to 700bp.

Densitometric patterns corresponding to the MNase digests of molecules assembled at molar octamer:DNA ratio intermediate to those determined by experiment were calculated by cubic spline interpolation between corresponding data points of the experimental scans flanking the area of interest. For instance, if the series of datapoints of a pair of experimental scans representing the MNase digestion profiles on a acrylamide gel of supercoiled plasmids reconstituted at consecutive molar octamer:DNA ratios, A and B, are represented by A_1, A_2, ... ,A_{n-1}, A_n and B_1, B_2, ... ,B_{n-1},

B_n, the value of the first datapoint X_1 of the scan representing the MNase digestion profile at a molar octamer:DNA ratio of X, intermediate to A and B, would be calculated by cubic spline interpolation between A_1 and B_1. The next datapoint, X_2, would be calculated by interpolation between A_2 and B_2, and so forth, until a series X_1, X_2, ... , X_{n-1}, X_n , representing the MNase digestion profile at a molar octamer:DNA ratio intermediate to A and B was generated. This procedure was repeated between all experimental scans to obtain a continuous surface representing the densitometric scans from molar octamer:DNA ratios of 0 to 11. A surface generated by this procedure is shown in figure 4.5. The one set of parallel edges of such a surface would represent the axes on which the distance of electrophoresis or DNA fragment size can be measured (scan axis), the other set of edges would represent the molar octamer:DNA assembly ratio (input axis), and the third axis, perpendicular to these two would correspond to the optical density measured with the densitometer (optical density axis). Such a surface allows one to obtain the densitometric profile of a MNase digest of plasmid assembled at any molar octamer:DNA ratio between the terminal ratios of 0 and 11.

A few comments on the method of interpolation might be in order here. A number of mathematical methods can be used to fit a surface to suitably spaced "slices" (3, 166). The degree of accuracy of some of these methods depends on the features of the surface. A n*th*-order polynomial in two variables can accurately describe the surface of a perfect parabolic dish, but is only approximate in the case of more convoluted surfaces. A linear interpolation between slices can accurately describe a segmented surface, but is clearly unsatisfactory for smoothly curved surfaces. At the other extreme, a two dimensional Fourier analysis can describe surfaces to very high accuracy if datapoints can be sampled

at a high frequency, but becomes progressively approximate as the sampling frequency drops. I suggest that a cubic spline interpolation is the method of choice, since the calculated function is not fitted to the datapoints, but includes them, and between experimental data points the function changes smoothly as a third-order polynomial at a slope described by a set of three data points. The success of this procedure in calculating densitometric profiles at intermediate molar octamer:DNA ratios is demonstrated at the end of section 4.2. The cubic spline interpolation procedure was performed with the program *CUBESPLN*, of which a listing is given in appendix B.

The number of octamers assembled onto linearised plasmid at a given input ratio of octamer:DNA (w/w) was determined by finding the position on the surface representing the densitometric scans of supercoiled plasmids that matched the densitometric scan of the linearised plasmid, assembled at the given molar octamer:DNA ratio, best. This was performed by the least-squares method, where the summed squared differences between a pair of densitometric profiles, the one representing the linearized plasmid at the assembly ratio of interest and the other representing the supercoiled plasmid chosen at intervals of 0,1 from the supercoiled surface, was calculated. The assembly ratio of octamer:DNA where the densitometric profile from the supercoiled surface matched that of the linearized plasmid best, was indicated by the smallest summed squared differences value. Since the number of nucleosome cores formed on supercoiled DNA can be determined from the linking difference following topoisomerase I relaxation (see section 4.2), the number of cores formed at a particular molar octamer:DNA ratio on the linearized assembled plasmid could be directly obtained. The position on the surface representing the assembled supercoiled plasmid that best matched the relevant densitometric profile of the assembled linearized plasmid

was determined with the program *FINDFIT*, of which a listing is given in appendix C.

2.8.3 Probability difference plots.

The appropriate lanes on autoradiographs of electrophoretically separated primer extended DNase I digestion products were densitometrically scanned, and the traces plotted onto A3 size paper. The areas of peaks (A) in the region of interest were calculated from measurements of the peak heights and base widths, and the extent of DNase I cleavage (R_i) at every location i along the DNA molecule calculated, after Lutter (142), by the equation $R_i=A_i/(\sum_{j=1}^{N} A_j)$. Thus, the extent of cleavage at peak i is calculated from the ratio of the peak area of peak i to that of all peaks above and including i. An indication of the degree of protection at peak i due to association of the histone octamer can now be obtained by comparing the extents of DNase I cleavage at point i in assembled and free DNA. This is formalized in the probability difference equation $P_i=\ln(R_{ai}/R_{fi})$, where R_{ai} represents the extent of DNase I cleavage at position i in the DNA molecule assembled with the histone octamer, and R_{fi} the extent of cleavage at the identical position i in the free DNA molecule (57). An implicit assumption of calculating the extents of DNase I digestion at particular bonds from digestion data obtained at a single time point (57), is that the extent of cleavage at each particular bond is within the linear range.

2.8.4 Correlation analysis.

In order to determine whether nucleosome cores assumed identical rotational and translational (*modulus* average helical period) frames on DNA molecules under different experimental conditions, a correlation analysis was performed between the difference probability data of the

relevant experiments. This was performed by calculating the summed difference between corresponding datapoints at different offsets of the two difference probability datasets. In particular, the correlation coefficient (C_j) was calculated from:

$$(1/N)\sum_{i=0}^{N}|S_i-L_{i+j}|$$

where S_i represents the difference probability value of point i of the one dataset, and L_{i+j} represents the difference probability value at point i+j (offset by j relative to point S_i). The summed value was calculated at integer offsets of -4 to 5 and normalized to N, the number of datapoints compared. In all the correlation plots shown in the text, N=120. The minimum value in a plot of j *versus* C_j indicates the offset where the two difference probability datasets are optimally aligned.

CHAPTER 3

3 Non-B-DNA structural transitions in the plasmid pHP2.

3.1 Introduction.

In *P. miliaris* (sea urchin) the early histone genes are arranged in batteries of tandem repeats, where each repeat contains a single copy of each of the five histone genes (150). The structural organisation of one such repeat is shown schematically in figure 3.1. In the histone gene repeat of *Drosophila*, where the exact order of the histone genes differ from that of *P. miliaris*, a single scaffold attachment region was identified in the A+T-rich area of the H1-H3 histone spacer (76, 162, see section 1.3). A similar investigation on the histone gene repeat of *P. miliaris* has not been performed, but a search of the entire repeat for topoisomerase II consensus sequences and A-rich and T-rich boxes, which were shown to be associated with scaffold attachment regions (75 and 76), revealed an enrichment of these sequences in the H3-H2A and H2A-H1 spacer regions (see figure 3.1). The particularly high frequency with which these sequences occur in the H2A-H1 spacer suggests that the region bordered by the two extreme topoisomerase II recognition sites (4406bp-4933bp) may constitute part of a scaffold attachment region. This suggested arrangement of one attachment site per histone gene repeat implies that the structural organisation of the early histone gene battery from *P. miliaris* into loops, may be very similar to that of *Drosophila* (162).

An investigation of the *in vitro* positioning of nucleosome cores at a low resolution on the entire gene quintet of *P. miliaris* by Retief (202) did not reveal the presence of any precisely positioned cores. This is in contrast to the situation *in vivo* on the histone gene repeat of *Drosophila*

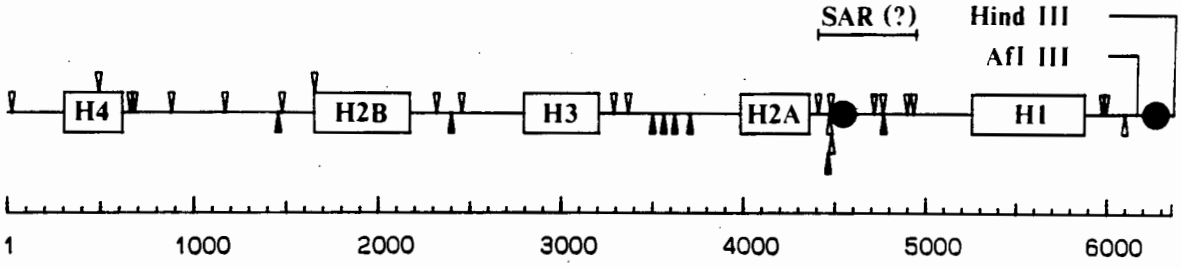

Figure 3.1: Organization of the early histone gene battery of *Psammechinus miliaris*. The histone structural genes are represented by rectangles. The locations of topoisomerase II recognition sequences (open triangles above the line; 11/15 overall agreement with the consensus sequence GTNRAQATTNATNNP, and 100% homologous to the six underlined nucleotides), A-boxes (open triangles below the line; 90% homologous to the consensus sequence AATAAAQAAA), and T-boxes (solid triangles below the line; 100% homologous to the consensus sequence TTRTRTTRTT) are indicated. The consensus sequences were taken from the study by Gasser and Laemmli (75), where R = A or T, Q = C or T, P = A or G, and N represents any of the four deoxyribonucleotides. The location of two positioned cores (R3) are indicated by solid circles. The scale bar at the bottom is drawn in units of base-pair.

where ordered nucleosome core arrays were present on the spacer regions (277) whereas the regions corresponding to the structural genes showed a more random arrangement. Subsequent studies performed at single nucleotide level on three selected regions from the *P. miliaris* histone gene repeat did, however, reveal the presence of a positioned nucleosome core at position 4475bp-4613bp on the H2A-H1 spacer region, and at position 6189bp-6329bp on the H1-H4 histone gene spacer (203). The location of these cores are shown in figure 3.1. The core formed on the H1-H4 spacer occupied a single frame with a well defined rotational placement and was for this reason chosen for the present study. A conspicuous feature in the sequence of this frame, isolated as a 199bp *Afl* III-*Hind* III fragment, is the presence of a 32bp d(A-G).d(C-T) run. A similar d(A-G).d(C-T) run of 46bp was present in the frame of the octamer positioned on the H2A-H1 spacer (203). It is known that the heteropolymer poly(dA-dG).poly(dC-dT) can be reconstituted into nucleosome cores (115), and one of the cloned core sequences in the study by Satchwell *et al.* (210) consisted almost entirely of this motif. Extended poly(purine).poly(pyrimidine) stretches of this form are often present in the eukaryotic genome (67, 106, 145, 148 and 216) and can assume a triple helical H-DNA conformation *in vitro* under appropriate conditions of negative superhelicity and/or acid pH (reviewed in 72 and 265). Since non-B-DNA structures were shown to influence core formation (177) and positioning (74 and 176), the presence of such structures was investigated in the plasmid pHP2, which contains the 199bp *Afl* III-*Hind* III fragment, to ensure that differences that may be observed between the positioning of nucleosome cores on supercoiled and linearised plasmids are not due to supercoil-induced transitions in the structure of the DNA duplex.

3.2 The presence of structural transitions to a non-B-DNA structure in the plasmid pHP2.

In a closed circular DNA molecule any structural transitions of the DNA duplex which reduces the twist of the molecule, such as a B-to-Z (185) or B-to-H (144) flip or cruciform formation (89, 98), is accompanied by a compensatory increase in the writhe of the molecule due to the topological invariance of the linking number (73, 266, reviewed by Cozzarelli *et al.* (48), see section 1.5). This is immediately apparent from the mathematical relation between these parameters,

$$\Delta Lk=\Delta Wr+\Delta Tw \qquad (3.1)$$

where ΔLk is the linking difference, ΔWr the change in writhe, and ΔTw the change in twist (73, 266). ΔLk is a topological parameter and is invariant under conditions of smooth deformation, including structural transitions in the DNA duplex in the absence of any strand-breaking event (see section 1.5). Thus, under such conditions, any change in twist will be accompanied by a change in writhe of similar magnitude but opposite in sign, since $\Delta Lk=0$. The writhe change caused by a supercoil-stabilised structural transition is readily detected in a two dimensional electrophoretic separation of topoisomers in agarose gels as a discontinuity in the gradual increase in electrophoretic mobility (which is a function of the writhe of the molecule) with increasing linking difference (105, 144, 185 and 264).

Figure 3.2 A shows the electrophoretic separation of a series of topoisomers of pHP2 in an agarose gel in the absence of chloroquine in the first dimension, and in the presence of chloroquine in the second dimension, where intercalation of the drug between the DNA base-pairs titrates out negative supercoils and leads to the loss of

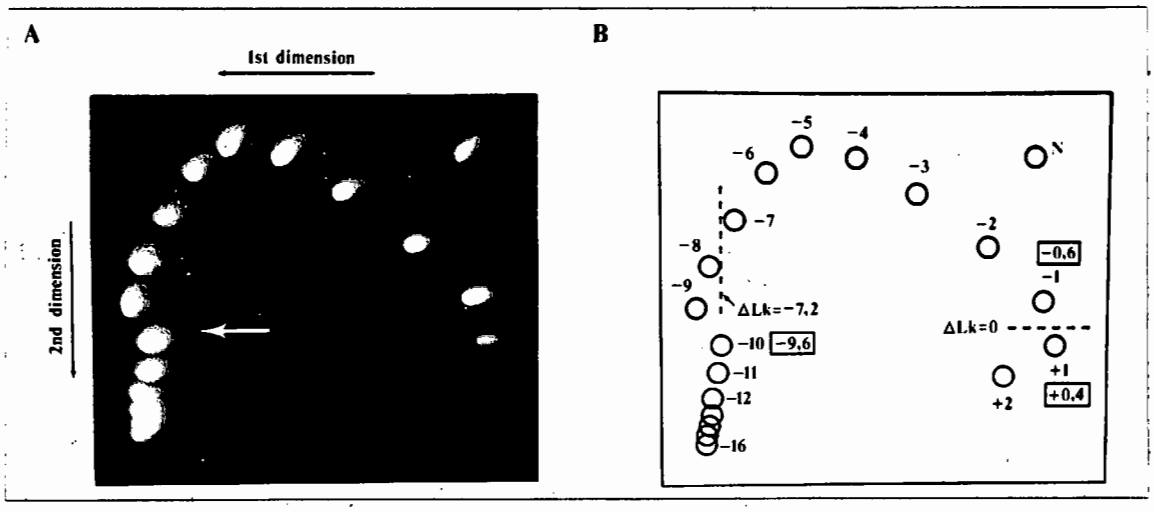

Figure 3.2 : Two-dimensional electrophoresis of topoisomers of plasmid pHP2. A mixture of DNA topoisomers of pHP2 was electrophoresed in 1,5% (w/v) agarose in TBE in the first dimension, and in TBE containing 0,5μg/ml chloroquine in the second dimension. A discontinuity due to a structural transition of the DNA duplex is indicated by an arrow in (A). Schematic representation (B) of the gel shown in (A). The individual topoisomers are identified and the ΔLk of selected topoisomers indicated in rectangles. The positions in the first dimension that correspond to ΔLk of 0 and -7,2 are indicated by stippled lines.

structural transitions in topoisomers where the resulting unrestrained negative superhelical stress is insufficient to stabilise such transitions (185 and 264). Two important points related to the structure of the DNA duplex in the plasmid under the investigated conditions of negative supercoiling can be deduced from figure 3.2 A. Firstly, a single deviation in the trend of increasing electrophoretic mobilities with increasing negative superhelicities is visible (indicated by an arrow in figure 3.2 A) for topoisomers $\leq$-10, indexed relative to the slowest migrating negatively supercoiled topoisomer in the first dimension of electrophoresis (see Figure 2 B). This suggests that only one type of secondary structure is stably present at a specific linking difference $\leq$-0,053. The presence of more than one type of non-B-DNA structure could reasonably be expected to result in the appearance of more than one electrophoretic member for a given topoisomer species in the first direction of electrophoresis, as reported for supercoiled molecules containing either Z-DNA or cruciform structures (20) or multiple conformers of H-DNA (105). The formation of one type of structure may, however, result in a decreased negative superhelical stress insufficient to stabilise a second structural transition. A second transition may then only be evident at superhelicities sufficiently high to stabilise more than one secondary structure, as demonstrated by Blaho *et al.* (20) and Htun and Dahlberg (105). Such a second transition is not evident in Figure 3.2 A up to σ=-0,086, corresponding to the most negative topoisomer resolved (ΔLk=-15,6). One may therefore conclude that a single structural transition is present in the plasmid pHP2 in the range -0,086$\leq \sigma \leq$-0,053. Secondly, due to the sequence specificity of such structural transitions (reviewed in 72), one may conclude that it is localised to a single region of the plasmid pHP2.

A careful comparison of the positions at which the individual topoisomers resolve indicate that topoisomer -10 (ΔLk=-9,6) migrates similarly to a topoisomer with a linking difference of -7,2 (refer to Figure 3.2 B). Since identical electrophoretic mobilities imply identical values for writhe (264), it is clear that the secondary structure present in topoisomer -10 caused a twist reduction of -2,4.

The native plasmid, isolated from a bacterial cell-culture, has a specific linking difference of -0,06 (see section 4.2). This level of supercoiling corresponds to a point beyond that at which the structural transition occurs (see figure 3.2 A), which would indicate that the structure is present in the native supercoiled plasmid. This implies that the structure of the DNA duplex in a section of the plasmid presented to the histone octamer at the onset of assembly (see below) differs between the supercoiled and the linearised molecule. Since the intended variable parameter in this investigation is the superhelicity of the DNA molecule, and not supercoil stabilised secondary structures, the identity and position of the structural transition was further investigated with nuclease P_1 in both free and assembled supercoiled plasmids.

3.3 Identification of the supercoil-stabilised non-B-DNA structure in free pHP2.

The exact structural features recognised by nuclease P_1 in the DNA molecule is not currently known, unlike that of micrococcal nuclease (70) and DNase I (57, 137, 231) which are well characterised in this regard. It was proposed that both the related nucleases P_1 and S_1 recognise a structural distortion of the sugar-phosphate backbone (265) where at least two unpaired bases are required by nuclease S_1 to produce a single-strand scission (53). Nuclease P_1 has the advantage over S_1 by virtue of a wider pH optimum and more

physiological reaction conditions and has been used as a structural probe for B-Z junctions (20), cruciform structures (20) and H-DNA (101 and 102).

Supercoiled and linearised pHP2 were digested with nuclease P_1 at pH4,5 and pH7,5 followed by cleavage with either of restriction enzymes *Pvu* II or *Nco* I in the case of supercoiled plasmids. The linearised plasmid digested with nuclease P_1 was obtained from supercoiled plasmid by cutting with either of these two restriction enzymes prior to treatment with nuclease P_1. The results of the nuclease P_1 digestions are shown in figure 3.3. It is clear from the appearance of only two major bands below that corresponding to the intact linearised plasmid (see lanes 1-4 of figure 3.3) that nuclease P_1 cleaves predominantly at only one position in the supercoiled plasmid at both pH-values. It is further evident from the similarity in the size of these lower two bands following *Pvu* II (lanes 1 and 3) or *Nco* I (lanes 2 and 4) digestion, that this position must be very similar, if not identical, at both neutral and acidic pH. Looking next at the linearised plasmid (lanes 5-8 in figure 3.33), bands migrating similar to those of the supercoiled plasmid are visible at pH4,5 (lanes 7 and 8), whereas no digestion by nuclease P_1 is visible for the linearised molecule at pH7,5 (lanes 5 and 6). This result would indicate that nuclease P_1 recognises and cleaves at a secondary structure which is stabilised by negative supercoiling and is also present on a linearized molecule under acidic conditions.

The position of nuclease P_1 cleavage in the plasmid was subsequently determined relative to the restriction sites. The fragment size corresponding to the lower two resolved bands in lanes 1 to 4 and lanes 7 and 8 of figure 3.3 were calculated with a third-order polynomial fitted by the method of least-squares to the standards in lanes M1 and M2.

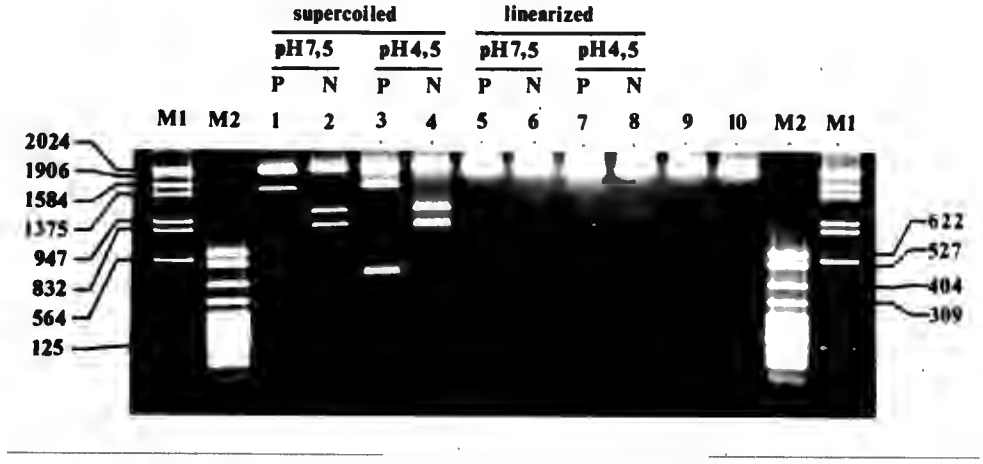

Figure 3.3: Nuclease P_1 digestion of free pHP2. Supercoiled (lanes 1-4) and linearized (lanes 5-8) plasmid pHP2 were digested with 2U nuclease P1 followed by restriction with *Pvu* II (P) or *Nco* I (N) in the case of the supercoiled plasmid. The linearized plasmid was cut with these two restriction endonucleases prior to digestion with nuclease P1. The digestions were carried out at pH7,5 (lanes 1,2,5 and 6) or pH4,5 (lanes 3,4,7 and 8). A *Eco* RI/*Hind* III digest of λ-DNA (M1) and *Hpa* II digest of pBR322 (M2) were used as size standards with fragment sizes in base-pairs indicated in the left- and right-hand margins respectively.

The top fragment of the lower two in lanes 1,3 and 7 is approximately 1500bp and the bottom fragment in the same lanes about 500bp (the approximate sizes are given for illustrative purposes). Fragments of approximately this size can be obtained relative to the *Pvu* II site (670bp) by cleavage in the region of either positions 200 or 1100, which are symmetrically related in the plasmid relative to the *Pvu* II site. The top and bottom fragments present in lanes 2,4 and 8 correspond to approximately 1100bp and 900bp respectively, and can be obtained relative to the *Nco* I site (1073bp) by cleavage in either the region of position 2000 or 200. Thus, by elimination, nuclease P_1 cleaves the plasmid in the region of position 200. The exact value, which was obtained from three independent experiments, is 210±11bp. By referring to the sequence of the plasmid in this region (figure 2.1) it is clear that the nuclease recognises a structural transition in the d(A-G).d(C-T) run. The appearance of a P_1 nuclease sensitive structure in the d(A-G).d(C-T) run in the supercoiled plasmid at both pH7,5 and pH4,5 but only at pH4,5 in linearised plasmids is indicative of a structural transition to a triple helical H-DNA conformation (105, 144 and 160). Further support for such a H-DNA conformation comes from the magnitude of the deduced twist change ($\Delta Tw=-2,4$) which is in good agreement with an expected change of -2,9 if the entire 32bp d(A-G).d(C-T) run assumes the H-DNA conformation (105). The existence of the d(A-G).d(C-T) run in the H1-H4 histone spacer region in a H-DNA conformation under appropriate conditions is in agreement with the results of other studies (144, 256, 257). The only difference between the results reported above and that of other workers is the detection by two-dimensional electrophoresis of a structural transition at $\sigma \leq -0,053$ at pH7,5. In the study by Lyamichev *et al.* (144) it was reported that a structural transition at pH7,0 and above was only evident in two-dimensional electrophoresis at $\sigma \leq -0,08$. This may be ascribed to the

increased stability of the hydrogen-bond stabilised structure under the electrophoretic conditions of this study (4°C) compared to that of Lyamichev *et al.* (144) (room-temperature). The presence of an unusual structure in the d(A-G).d(C-T) run at slightly alkaline pH (the two-dimensional electrophoresis was performed in TBE buffer at pH8,3) is clearly indicated by the nuclease digestion studies reported above (see also 101), where a positive result will be obtained even if the non-B-DNA structure is present only transiently at the higher temperature (37°C).

3.4 Presence of the d(A-G).d(C-T) run in a H-DNA conformation in the plasmid pHP2 reconstituted into nucleosome cores.

The presence of H-DNA in a supercoiled plasmid reconstituted into nucleosome cores was investigated by the nuclease P_1 digestion of free supercoiled pHP2 and supercoiled pHP2 assembled with the histone octamer at a ratio of 1:1 (octamer:DNA (w/w)). Under these conditions the plasmid is fully occupied with nucleosome cores (see section 4.2). The results are shown in figure 3.4. It is clear that the number of units of nuclease P_1 sufficient to significantly cleave at the H-DNA structure present in the d(A-G).d(C-T) run in the free supercoiled plasmid (lane 2 and 3), is insufficient to cleave at the corresponding position in the supercoiled plasmid reconstituted with nucleosome cores (lane 4 and 5). Increasing the nuclease P_1 concentration 15 fold is still insufficient to cause significant cleavage in the assembled plasmid (lane 6 and 7). To demonstrate that the nuclease is active and can recognise the H-DNA conformation under the conditions of the reconstituted sample, free supercoiled plasmid was mixed with assembled plasmid at an equal DNA ratio (w/w) prior to addition of the nuclease. Subsequent digestion with P_1 nuclease resulted in the appearance of fragments (lanes 8 and 9) corresponding in

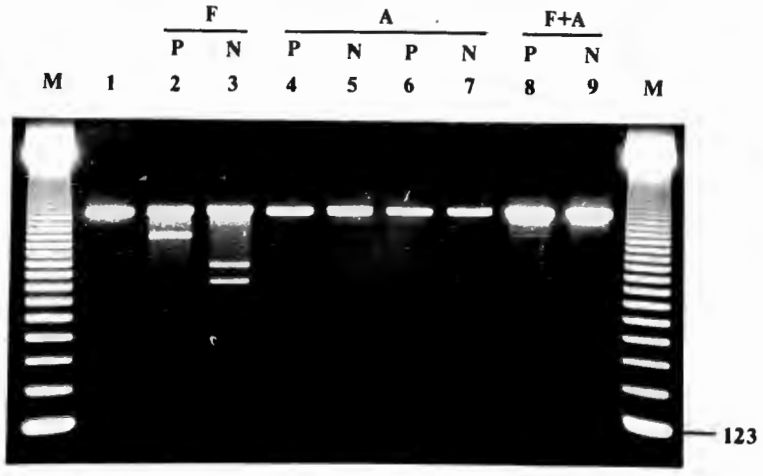

Figure 3.4: Nuclease P1 digestion of plasmid pHP2 assembled with the histone octamer. Free (lanes 2 and 3) and assembled (lanes 4-7) plasmid pHP2 were digested with 2U (lanes 2-5), and 30U (lanes 6 and 7) nuclease P1, followed by cleavage with either restriction endonuclease *Pvu* II (P; lanes 2, 4 and 6) or *Nco* I (N; lanes 3, 5 and 7). A 1:1 (w/w) mixture of free and assembled pHP2 was digested with 2U nuclease P1 (lanes 8 and 9) and similarly cut with *Pvu* II (P; lane 8) or *Nco* I (N; lane 9). A 123bp ladder was included as a size standard (M) with the position of the monomer indicated in the right-hand margin.

size to those present after digestion of the free supercoiled plasmid (lanes 2 and 3), which indicates that nuclease P_1 recognised the H-DNA conformation under the conditions of the assembled plasmid. The absence of any detectable cleavage in lanes 4 and 5 as well as in lanes 6 and 7 suggests that the d(A-G).d(C-T) run did not assume a H-DNA conformation in the plasmid reconstituted at a 1:1 (w/w) ratio with the histone octamer. This result is not unexpected. The association of one octamer with DNA results in the effective loss of negative superhelical stress equal to -1,1 linking units from the plasmid (81, 221, 283). The native plasmid has an average ΔLk of -10,9 (see section 4.2). Thus, the association of two octamers with the plasmid will result in an unconstrained superhelical stress equal to -8,7 which corresponds to a specific linking difference of -0,048. By referring to the results of the two-dimensional electrophoretic investigation above (see figure 3.2), such a specific linking difference is clearly insufficient to stabilise a structural transition in the d(A-G).d(C-T) run to a H-DNA conformation. It follows directly that the H-DNA conformation will be very unstable in the plasmid fully occupied with nucleosome cores.

3.5 Discussion.

In a study on the effect of negative superhelicity on nucleosome core formation and positioning, particularly where the variable parameters of interest include the torsional stress of the duplex and the circular constraint of the molecule, the occurence of supercoil-stabilised secondary structures, which can influence these parameters, must be investigated to ensure interpretable results. The results presented above indicate the presence of a supercoil-dependent structural transition, most likely a triple-helical H-DNA conformation, in the d(A-G).d(C-T) run of the plasmid pHP2. This structure was further shown to be

absent in supercoiled plasmids fully occupied with nucleosome cores, most likely due to an unconstrained negative superhelical stress insufficient to facilitate a structural transition.

Reports on the ionic strength dependency of molecular triplex stability, based on thermal denaturation (214) and nuclease digestion studies (101), indicate an increasing stability with increasing sodium or potassium cation concentration in the range tested (10mM to 1M) (214). It appears reasonable to assume that under the ionic conditions present during the salt-dialysis reconstitution of nucleosome cores (initially 2M NaCl; see section 2.4.4 of Material and Methods), the d(A-G).d(C-T) run will form a triple helix in supercoiled plasmids. It is therefore possible that the H-DNA conformation will be present during the early stages of assembly, until a sufficient number of octamers have assembled to reduce the unrestrained negative superhelical stress to a level unable to support the triple helical structure. It may be argued, in such a scenario, that the initial presence of the H-DNA conformation will influence the placement of at least the first two octamers (or $(H3\text{-}H4)_2$ tetramer which associates with the DNA before two (H2A-H2B) dimers (100)) assembled onto the DNA, which in turn will govern the distribution of available free sites, thereby indirectly influencing the positions assumed by the nucleosome cores. This seems unlikely. Apart from possible intermolecular migration of the octamer at intermediate ionic strengths (200), the octamer also has a considerable intramolecular translational freedom ("slipping") at lower ionic strengths (155). It is therefore plausible that at and below the point where the H-DNA conformation is no longer present, the core will still have a dynamic distribution, allowing settlement into final positions not influenced by the initially present triple helical conformation of the d(A-G).d(C-T) run. It also follows

directly from this argument that the stable incorporation of the H-DNA conformation into cores is unlikely, since any transient exposure, due to octamer movement, will lead to the collapse of the triple helical structure.

Another line of evidence against the existence of the d(A-G).d(C-T) run in a H-conformation in the core DNA comes from relaxation studies of assembled supercoiled plasmids. Here, the linking difference of supercoiled plasmids assembled at various ratios of octamer:DNA (w/w) were determined after relaxation with topoisomerase I (see section 4.2). If the H-DNA can be incorporated into core DNA, the likelihood of such an occurrence will increase at higher weight ratios of octamer:DNA. Relaxation of the assembled plasmid with topoisomerase I will lead to the removal of excess superhelical stress sensed by the enzyme in the linker areas which will exclude that constrained by the nucleosome cores and in the H-DNA structure. It appears unlikely that the topoisomerase will bind to and relax the stress in the triple helix conformation itself since the required exposure of the structure should have enabled cleavage by nuclease P_1. Thus, at most, a displacement of the curve relating the input ratio of octamer:DNA (w/w) to ΔLk towards higher values should be apparent, or, at least, a similar deviation of this curve at high input ratios of octamer:DNA (w/w). Neither of these features are apparent (see figure 4.3). It can therefore be concluded that the d(A-G).d(C-T) run is not present as a triple helix in supercoiled plasmids assembled at the ratios of octamer:DNA (w/w) employed in this study, and that the positioning of the histone octamer on supercoiled plasmids is not influenced by secondary structures that are stabilised by negative superhelical stress in the free plasmid.

In contrast, core formation may be influenced by the initial presence of H-DNA. Nickol *et al.* (177) demonstrated the inability of poly(dG-m^5dC).poly(dG-m^5dC) in the Z-form to be reconstituted into nucleosome cores. This inability has been suggested to lie in the inflexible nature of the Z-DNA duplex (74). No study on the flexibility of H-form $d(A\text{-}G)_n.d(C\text{-}T)_n$ has been published, but it would seem that the presence of a third strand bound by Hoogsteen base-pairing in the major groove would confer some axial rigidity to the structure. Confined areas of the DNA duplex that contain axially inflexible sequences such as oligo(dA).oligo(dT) (173) or inflexible structures such as Z-DNA (242) have been shown to be excluded from the central part of the core DNA, but can occupy the terminal regions (74 , 210) where the requirement for bending the DNA is less stringent (135, 243). It therefore appears likely that a rigid triple helical H-DNA conformation of between 1 and 2 helical turns in length may be similarly excluded from the central core DNA. Thus, the effect of H-DNA on core formation in supercoiled plasmids is two-fold: firstly, the H-DNA may exhibit a decreased tendency to be reconstituted into cores, and, secondly, the presence of the H-DNA will decrease the unconstrained negative superhelicilty of the DNA. It follows that in studies where the rate of initial octamer assembly is investigated as a function of negative superhelicily, the measurements will include an error due to both the decreased negative superhelical stress, and a chance that a random selection of 145bp stretches on the DNA will include the H-conformation which may not assemble with the same efficiency.

In contrast, where measurements are made of the efficiency of reconstitution as a function of negative superhelicily after the assembly system has reached equilibrium, the effect of the initial presence of the H-DNA structure is likely to be negligible. This follows from the dynamic

intermolecular exchange of the octamer at intermediate ionic strength, which is a common condition of most assembly systems (poly[L-glutamate] facilitated assembly is a notable exception (43)). Thus, under these conditions, the system can proceed to a final state independent of initially present secondary structures.

The absence of a triple-helical structure in the d(A-G).d(C-T) run in assembled plasmids immediately raises the question of the possible function of these ubiquitous sequences. As discussed in section 1.6.3, most of the negative superhelical stress in eukaryotes is constrained by bound nucleosomes. Furthermore, heteropolymeric sequences of the $d(A\text{-}G)_n.d(C\text{-}T)_n$ form are often found associated with cores (203, 210), which have been shown to stabilise the DNA duplex in the B-form (177). It is currently unclear whether the demonstrated structural polymorphism of these sequences have biological significance. Kinniburgh (118) suggested that a heteropolymeric sequence present in the promoter of the human *c-myc* gene, which can adopt a H-DNA conformation *in vitro*, may be induced into a triple-helical structure *in vivo* by a *trans*-acting ribonucleoprotein. In an *in vivo* study by Glaser *et al.* (84), the numerous repeats of $d(A\text{-}G)_n.d(C\text{-}T)_n$, required in the promoter area for optimal expression of the *Drosophila hsp26* gene, were shown not to adopt a H-DNA conformation in isolated nuclei. Cooney *et al.* (46) demonstrated a repression of transcription upon binding of a 27bp oligonucleotide to the promoter area of the *c-myc* gene to form a triple-helical structure. This clearly does not involve H-DNA formation, but suggests that the ability of poly(purine).poly(pyrimidine) sequences to form triple-helixes may be exploited in other ways, such as the binding of RNA or ribonucleoproteins to such regions. It should also be noted that the negatively supercoiled domain present behind the transcription complex in eukaryotes (82) may be sufficient to induce a B-to-H

transition *in vivo* in nucleosome-free regions, and in this regard the possible formation and influence of H-DNA *in vivo* clearly warrants further investigation (72).

CHAPTER 4

4 The effect of negative supercoiling on nucleosome core reconstitution *in vitro*.

4.1 Introduction:

In vitro studies on reconstituted nucleosome cores have contributed significantly to an understanding of nucleosome (59, 200, 205, 220) and chromatin (99) structure, the rules governing the positioning of cores (37, 55, 59, 175, 200, 220) and the regulatory effect of nucleosomes on chromatin function (65, 194, 218). These studies employed a range of reconstitution methodologies (reviewed by Laskey *et al.* (130)) which included, depending on the specific application, the stepwise dialysis of mixtures of histone octamer and DNA from high to low concentrations of NaCl with (40) or without (99) urea, the exchange at high salt of the histone octamer from H1-stripped chromatin or core particles to the DNA molecule of interest (55, 200), and the assembly of DNA with the histone octamer facilitated by polyanions such as poly[L-glutamate] (228) or RNA (174) at intermediate ionic strength.

Most reconstitution studies were performed with linearized DNA as often dictated by the experimental approach, such as the employment of end-labeled DNA molecules. The exceptions include studies on the ability of various combinations of the core histones to induce superhelical turns into DNA (40), measurements of the topological properties of nucleosome cores (81, 221), the ability of the histone octamer to be reconstituted with very small circular DNA molecules (172, 283), and the ability of cores to accommodate non-B-DNA structures (36, 74). No comparative study on the relative efficiencies of reconstitution on supercoiled and linearized DNA has, however, been published.

Such a study is clearly relevant since the chromatin in eukaryotes is organized into topologically isolated loops (16) where the dissociation of the octamer will liberate one negative superhelical turn and association of the octamer will induce one positive superhelical turn (81, 221, 283). Although most of the superhelical stress in eukaryotes is restrained by bound nucleosomes (83, 223), the association of topoisomerase II with the scaffold attachment regions (75) suggest that free superhelical stress must be present at times, such as after decondensation of the 30nm fiber, should this occur over a significant section of a single chromatin loop. It was also shown that transcription (82, 136) leads to the supercoiling of DNA. There is therefore ample evidence that the nucleosome core, in its native environment, constitutes part of a system that is not only topologically constrained, but where the topological properties are dynamic. In this regard a comparative study on the efficiencies of nucleosome core reconstitution on linearized and supercoiled DNA would be informative in terms of the possible contribution of supercoiling to nucleosome formation and dissolution *in vivo*.

Apart from the insight that can be gained on core formation in a system that resembles the *in vivo* situation more closely, a quantitation of assembly efficiencies on linearized and supercoiled DNA has direct relevance to the question in hand, namely the effect of negative supercoiling on nucleosome core positioning. Determinations of nucleosome core placement often employ endonucleolytic enzymes such as DNase I (55, 59, 200, 205, 220; see Chapter 5). Here the translational and rotational frame of the core is deduced from the extent and phase, respectively, of the uninterrupted modulation in DNase I cleavage. It is also possible to obtain an indication of the precision of the rotational placement from the amplitude of the DNase I cutting periodicity. This can be seen intuitively by

considering that the observed DNase I cutting pattern is the average obtained from the superimposition of the individual cutting patterns from a large population of molecules. In a rotationally weakly positioned core, a wide distribution of rotational settings will be occupied, and the individual cleavage periodicities will be mostly out of phase. In contrast, a core that is rotationally precisely positioned will yield periodicities that are mostly in phase. The amplitude in the modulation of the DNase I cleavage pattern will thus be larger for the latter core. It further follows that erroneous conclusions on the rotational precision of core positioning may be drawn from comparative studies if the two samples are not occupied to the same extent by nucleosome cores. This can be demonstrated by referring to the theoretical treatment by Drew and McCall (56) of the amplitude of DNase I cutting periodicity as a function of core rotational distribution. Drew and McCall (56) proposed that the rate of DNase I cleavage (ϕ_x) at a particular phosphodiester bond of core DNA is proportional to:

$$\sum_{i=-4}^{5} N_i \cos^4[\pi(x-i)/10], \qquad (4.1)$$

where x is the phosphodiester bond index from 0 (faces away from core surface; maximally cleaved) to 5 (faces towards core surface; minimally cleaved) in the preferred rotational setting (i=0), and i is one of ten quantitized rotational settings of the DNA relative to the core surface, indexed from -4 to 5. N_i is the number of cores in a particular setting i and is proportional to:

$$\exp(-i^2\Delta G/25RT). \qquad (4.2)$$

ΔG is the free energy change accompanying a 180° rotation of the DNA helix on the core surface relative to the preferred orientation. It was shown (56) that the natural logarithm of the ratio of the rate of cleavage at bond 0 to bond 5

$[\ln(\phi_0/\phi_5)]$ was equal to 3.08 for an experimentally determined core. This identical ratio is predicted by equation 4.1 if ΔG=5,6kcal/mole in equation 4.2 (56).

Although the model was formulated to deduce free energy changes for rotationally positioned cores, it can be extended to describe the amplitude of DNase I cleavage in a large population of DNA molecules where not all the molecules are assembled. In this case the average rate of cleavage at bond x is normalized per molecule to allow comparison at different degrees of saturation, to yield Φ_x, which is proportional to:

$$(1-F)+(1/N_d)\phi_x, \qquad (4.3)$$

where ϕ_x is defined by equation 4.1 and N_d is the total number of DNA molecules, which is proportional to:

$$(1/F)\sum_{i=-4}^{5} N_i. \qquad (4.4)$$

F is the fraction of assembled molecules. The summation factor, 1-F, is the contribution made to the cleavage at bond x by the free DNA molecules. By referring to the original expression of ϕ_x (equation 4.1), it is clear that 1-F is a special case of this expression where all bonds can be maximally cleaved by DNase I (assuming no sequence specificity, after Drew and McCall (56)) and N_i equals the number of free DNA molecules normalized relative to the total number of DNA molecules.

If the natural logarithm of the ratio of cleavage at bond 0 to bond 5 is calculated with F=1, equation 4.3 simply reduces to $\ln(\phi_0/\phi_5)$, as given by equation 4.1. However, if not all the DNA molecules are occupied, the ratio of cleavage deviates from that proposed by equation 4.1. This deviation is mostly due to the contribution made by free DNA

molecules to cleavage at sites that are protected in the core DNA. For instance, if ΔG is set to 5,6kcal/mole, as above, and only 80% of the molecules are assembled, the ratio $\ln(\Phi_0/\Phi_5)$ is equal to 1,44. This is an important result since it means that if the rotational distribution of a nucleosome core on two unequally assembled DNA molecules is compared, one may erroneously deduce that the rotational distribution of one is more restrained. Conversely, if the inherent rotational distribution of a nucleosome core on two DNA molecules differ, the amplitude in the periodic cutting by DNase I can suggest similar rotational constraints if the core of higher rotational preference is underrepresented. This can be seen by substituting ΔG with 1,9kcal/mole in equation 4.2, and calculating the ratio of cleavage with equation 4.3 with F=1. The result is identical to the ratio calculated with equation 4.3 with ΔG=5,6kcal/mole and only 80% of the molecules assembled (F=0,8).

To interpret possible differences in core positions between linearized and supercoiled molecules, where the circular constraint of supercoiled molecules may, for instance, impose a narrowed rotational distribution, it is essential to undertake a detailed quantitation of the efficiencies of assembly on linearized and supercoiled DNA.

4.2 The quantitation method:

Several different approaches were attempted to quantitate the number of histone octamers associated as nucleosome cores with a known amount of DNA at different molar input ratios of octamer:DNA. These included spectroscopic measurements of the absorbance of reconstituted samples at 230nm and 260nm, ultracentrifugation of DNA samples reconstituted with [^{125}I]-labeled octamer on poly[L-glutamate] containing sucrose gradients to resolve and quantitate unassembled octamer, electrophoretic

separation on "nucleosome gels" of MNase digests of known amounts of DNA assembled with [^{125}I]-labeled octamer to quantitate the amount of DNA-associated octamer, and, similarly, the quantitation of the amount of [^{125}I]-labeled octamer associated with DNA by chromatography on small hydroxylapatite columns. These approaches gave either inconsistent results, or, in the case of [^{125}I]-labeled octamer, were subject to a large error due to non-specific binding (to vessel surfaces) of the octamer.

A successful method of quantitation was developed by measurements of the yield and fragment size distribution of MNase digests of supercoiled plasmid DNA containing a known amount of nucleosome cores. Supercoiled DNA was used since it is possible to determine the number of assembled cores by measurements of the linking difference following topoisomerase I relaxation.

The topoisomer distribution of supercoiled plasmid relaxed with topoisomerase I after urea/salt reconstitution with different molar input ratios of octamer:DNA is shown in figure 4.1. The systematic displacement of the center of the topoisomer distribution towards higher negative superhelicities with increasing octamer:DNA assembly ratio (lanes 1 - 6 of figure 4.1 A; the plasmid in lane 1 migrates as positively supercoiled DNA under the conditions of electrophoresis) indicates that more nucleosome cores form on the supercoiled plasmid at higher assembly ratios, since each core is associated with a linking difference of approximately -1,1 (81, 221, 283; see also section 1.6.2). The number of cores formed at different assembly ratios on the supercoiled DNA can be determined from a measurement of the number of linking units between the Gaussian centers of the topoisomer distributions corresponding to unassembled plasmid and the plasmid assembled at the ratio of interest. The highly negatively supercoiled samples on the gel are not

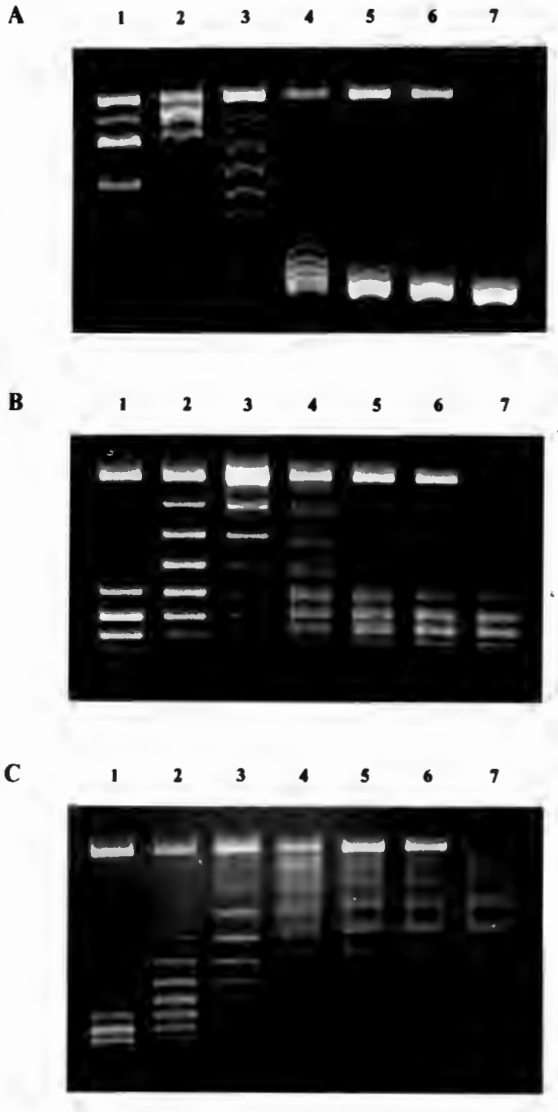

Figure 4.1: DNA topoisomer distribution of supercoiled pHP2 assembled at different octamer:DNA ratios by the method of urea/salt dialysis. Supercoiled pHP2 was assembled at octamer:DNA (mole/mole) ratios of 0, 3, 5, 7, 9 and 11 (lanes 1 - 6) and relaxed with 8U of eukaryotic topoisomerase I. Aliquots of the sample at each assembly ratio were electrophoresed on 1,5% (w/v) agarose gels run in TBE containing chloroquine at 0μg/ml (A), 1μg/ml (B) or 3μg/ml (C) in order to resolve individual topoisomers and the gel stained with ethidium bromide. Native supercoiled pHP2 (no topoisomerase I treatment) was included in lane 7 of each gel.

amenable to such an analysis, since the individual topoisomers are not well resolved. This may be due to the hyperbolic dependence of the radius of plectonemically supercoiled DNA on the specific linking difference of the molecule (48). Adjacent topoisomers in a highly supercoiled population do not differ drastically in structure. In order to determine the Gaussian centers of the highly negatively supercoiled populations that are not resolved into individual topoisomers in figure 4.1 A (lanes 5 - 7), it is necessary to perform the electrophoresis in the presence of the intercalative dye, chloroquine, to decrease the number of negative superhelical turns to a level where the individual topoisomers can be resolved on the gel (215). The results are shown in figure 4.1 B and C. At every assembly ratio an electrophoretic separation that clearly resolved individual topoisomers was selected and scanned on the film negative. The scans are shown in figure 4.2. The linking difference of the topoisomer population at each assembly ratio was determined from these scans as described in section 2.8.1 of materials and methods. The average of three independent experimental determinations of ΔLk as a function of the molar input ratio of octamer:DNA is shown in figure 4.3.

The distribution of datapoints in figure 4.3 clearly suggests a linear dependence of the number of cores assembled on the molar input ratio of octamer:DNA. The slope of the least-squares line further indicates a 1:1 correspondence, which shows that the number of cores assembled per molecule is equal to the molar input ratio of octamer:DNA (if $\Delta Lk=-1,1$ per core, as reported in 81 and 283). A slight deviation from this linearity appears at high assembly ratios. The datumpoint corresponding to a molar input ratio of 11 in figure 4.3 is below the point to which the least-squares line will be extrapolated. This is perhaps surprising since a 1915bp plasmid should easily

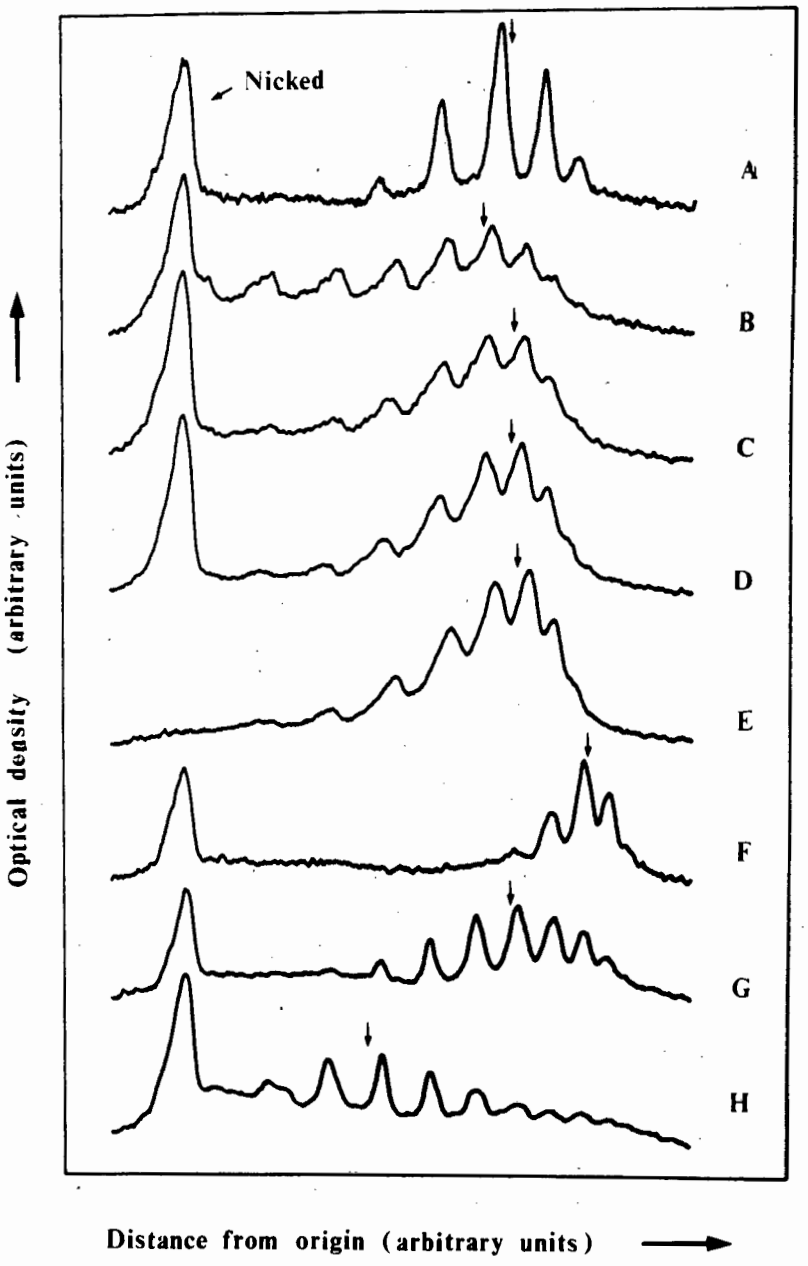

Figure 4.2: Densitometric scans of resolved topoisomers. Selected lanes of the gels shown in figure 4.1 were densitometrically scanned on the film negative and the scans aligned relative to the electrophoretic origin. The lanes where chosen to clearly show the Gaussian center of the topoisomer distribution (indicated by an arrow above each scan; see section 2.8.1 of materials and methods) at every experimental assembly ratio of octamer:DNA and represent distributions obtained in the presence of chloroquine at a concentration of 1μg/ml (A - E) or 3μg/ml (F - H). Individual scans (A - H) correspond to assembly ratios of octamer:DNA (mole/mole) of 0, 7, 9, 11, native supercoiled (no topoisomerase I treatment), 0, 3 and 5. The position of the nicked, circular plasmid is indicated in the top, left-hand side of the figure.

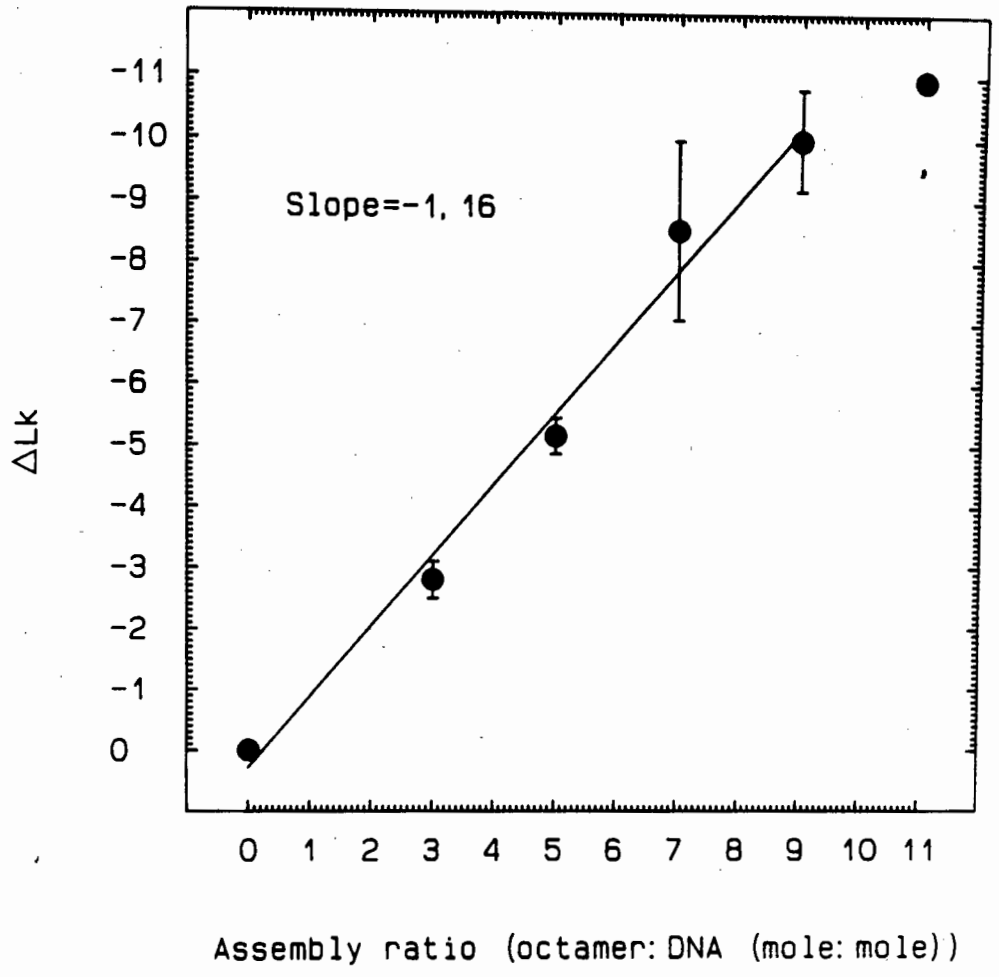

Figure 4.3: The efficiency of reconstitution of supercoiled DNA by urea/salt dialysis. Plot of the assembly ratio of octamer:DNA (mole/mole) *versus* the linking difference following topoisomerase I treatment. The average of three independent determinations is shown at each assembly ratio with the standard deviation indicated by brackets where this exceeds the diameter of the plotted datumpoint (±0,16). The line was fitted to the datapoints by the method of least squares. The datumpoint corresponding to an assembly ratio of 11moles/mole was omitted from the fitting procedure (see text).

accommodate 11 nucleosome cores (174bp/core). A possible reason for this deviation may be seen in measurements of the linking difference of the free plasmid, where $\Delta Lk=-10,9\pm0,09$ (average of three independent determinations $\pm$ standard deviation). Thus, with +1,1 positive supercoils introduced into the plasmid per nucleosome core formed, the unrestrained superhelical stress of the plasmid occupied by 11 cores would equal 12,1-10,9=+1,2. If adjacent cores are joined by 28bp of linker DNA [(1915/11)-146], the specific linking difference of the linker DNA will equal +0,04 if all the superhelical stress is concentrated in the linker areas. It was shown that the assembly of the octamer onto small circular DNA molecules of 185bp or 359bp proceeded negligibly if associated with the introduction of net positive superhelical stress (172, 87). In contrast, the octamer could be assembled on a large (3293bp) *positively* supercoiled plasmid where the superhelical stress was proposed to be accommodated by writhe and/or twist changes in the large molecule (43). Measurements that were made of the linking difference after assembly at an octamer:DNA ratio that should have fully occupied the positively supercoiled plasmid with nucleosome cores suggest that the average number of cores assembled per molecule was less than the input ratio (43). It therefore appears that at assembly conditions close to saturation, the supercoiled molecule is less able to accommodate superhelical stress, which may be related to a decrease in the flexibility of a long run of closely juxtaposed cores. It is therefore likely that the supercoiled molecule in this study, similarly, cannot efficiently accommodate positive superhelical stress at conditions close to saturation, leading to the observed decreased assembly efficiency (figure 4.3). This point is raised again and discussed further at the end of this chapter.

Thus, having obtained a measure of the number of nucleosome cores formed on supercoiled DNA at different molar input ratios of octamer:DNA, one may now proceed to investigate MNase digestion of assembled plasmids as a quantitation method. The result of an electrophoretic separation of MNase digests of supercoiled plasmid pHP2 assembled at different molar input ratios of octamer:DNA is shown in figure 4.4. The digestions were carried out as explained in section 2.6.1 of materials and methods. Prior to addition of the nuclease, the samples were pre-equilibrated at 37°C for 15min. to ensure equal initial enzyme activities in different samples.

It is clear when referring to figure 4.4 that different degrees of octamer occupation of DNA lead to significant differences in the yield and size distribution of recovered DNA fragments. In lane 2 of figure 4.4, two major bands, both originating from the core monomer, are visible. The bottom band of approximately 146bp corresponds to the 1,8 superhelical turns partially protected from MNase cleavage in the core particle (126), and the top band of approximately 166bp is due to the protection at low ionic strength of an additional double helical turn at each of the two termini of the core particle (134). At higher molar input ratios of octamer:DNA, fragments corresponding to core dimers (lanes 3 - 6) and trimers (lanes 5 and 6) are visible.

This dependence of the electrophoretic profile of MNase digestion on the average number of cores per plasmid molecule is clearly seen in the densitometric scan of the gel, shown in figure 4.5. At a molar assembly ratio of 3 (scan F), a significant peak at 146bp and a smaller peak corresponding to the 166bp fragment is visible. Very little material originating from the core dimer is visible. At a molar ratio of 5 (scan E), the amount of DNA fragment of

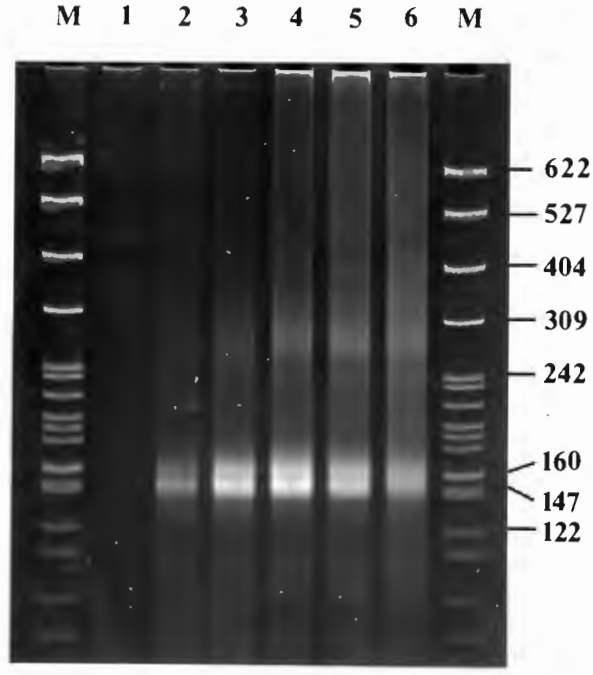

Figure 4.4: Micrococcal nuclease digestion of supercoiled plasmid pHP2 reconstituted into nucleosome cores. Supercoiled pHP2 was assembled by urea/salt dialysis at different molar ratios of octamer:DNA followed by digestion with 1U micrococcal nuclease at 37°C for 10min. and the digestion products resolved by electrophoresis on a 10% (w/v) acrylamide gel. The gel was stained in ethidium bromide. The assembly was performed at octamer:DNA (mole/mole) ratios of 0, 3, 5, 7, 9 and 11 (lanes 1 - 6). A *Hpa* II digest of plasmid pBR322 was included as size standard (lanes labeled M) with fragment sizes in base-pairs indicated in the right-hand margin.

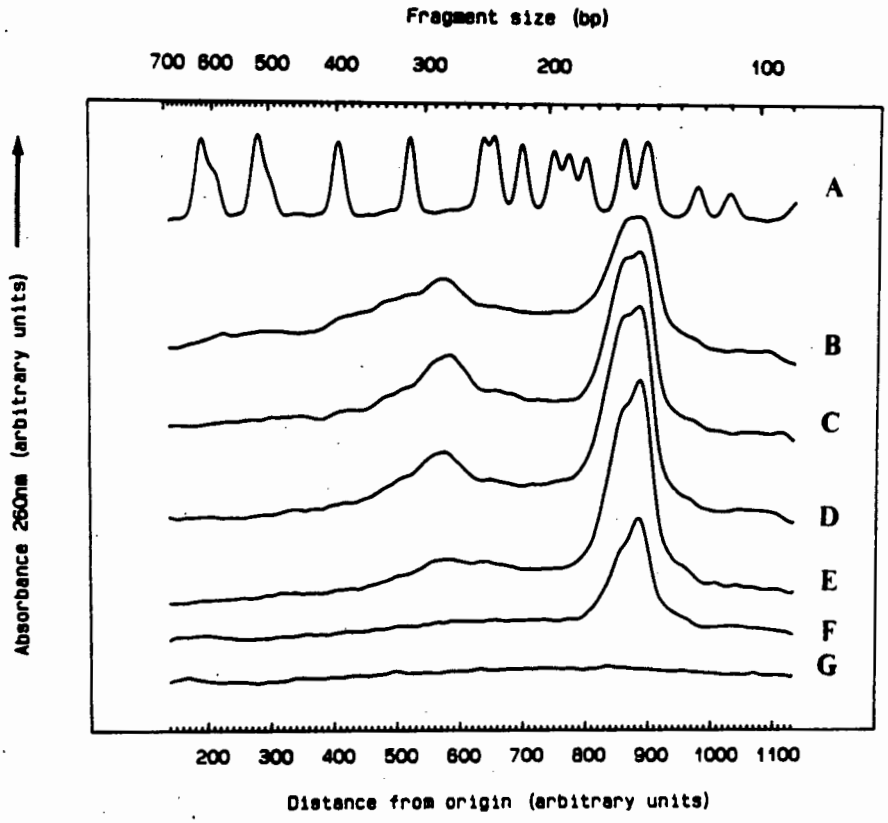

Figure 4.5: Assembly ratio dependent differences in the electrophoretic profiles of micrococcal nuclease digested assembled pHP2. The individual lanes of the gel shown in figure 4.4 were densitometrically scanned on the film negative and the scans clipped and aligned relative to the electrophoretic origin to produce shortened scans corresponding to fragment sizes from approximately 700bp - 90bp (see section 2.8.2 of materials and methods). The clipped densitometric scans were plotted to the same scale in the direction of increasing optical density and are directly comparable. Scan A is of a *Hpa* II digest of plasmid pBR322 (right-hand lane M of Figure 4.1). Scans B - G are of the micrococcal nuclease digested assembled plasmid pHP2 shown in lanes 6 - 1 of figure 4.4 corresponding to assembly ratios of octamer:DNA (mole/mole) of 11, 9, 7, 5, 3 and 0.

core particle length has increased, and a clear peak corresponding to the core dimer is now visible. Note that the ratio of the 166bp to the 146bp fragment has increased. At a ratio of 7 (scan D), this increase is continued although the total amount of monomeric fragment recovered appears similar to that of the lower assembly ratio (compare with scan E). The size of the core dimer peak has, in contrast, continued to increase. At an assembly ratio of 9 (scan C), the ratio of recovered 166bp to 146bp fragment appears similar to that of the lower assembly ratio, although the total amount of core length fragment has decreased (compare with scan D). The peak corresponding to the core dimer has increased further. At an assembly ratio of 11 (scan B), the monomeric peak is further reduced, and more material corresponding to larger fragments appear on the left-hand side of the scan. It is thus clear that under carefully controlled conditions of MNase digestion the fragment yield and distribution changes in a smooth fashion that depends on the average number of cores per plasmid molecule. The gradual shift of the material towards larger fragments with increasing assembly ratio is most likely due to a reduction in the rate of MNase cleavage per DNA molecule as the average length of free DNA, where the MNase approach to the DNA is less obstructed than in the core, decreases (126).

In order to present this dependence graphically, one may place the densitometric scans shown in figure 4.5 on a surface, where each scan is aligned with a point on one edge of the surface corresponding to the molar octamer:DNA assembly ratio of the particular scan. Scans corresponding to assembly ratios intermediate to those determined by experiment, may be obtained by interpolation between the experimental scans. For instance, referring to scans E and F of figure 4.5, it is clear that increasing the molar input ratio of octamer:DNA from 3 to 5 results in an increase in

the ratio of 166bp to 146bp fragment, an increase in the total yield of monomeric DNA, as well as the appearance of core dimer DNA. This change in the yield of DNA at every fragment size is unlikely to occur as a single jump somewhere between the two assembly ratios and more likely occurs in a smooth fashion proceeding from a molar ratio of 3 to a molar ratio of 5. It therefore seems highly probable that at a molar input ratio of octamer:DNA of 4, the densitometric scan of the MNase digest will have a profile that is some average of the two bordering experimental profiles. The question is, what average? It was decided that a cubic spline interpolation between corresponding datapoints on the two experimental scans bordering the interval of interest would be best suited to this application, since the calculated function is a smooth third-order polynomial that is forced through the experimental datapoints bordering the interval of interest at a slope that is governed by an experimental datapoint on either side. A cubic spline interpolation was performed for every interval between the scans shown in figure 4.5 (see section 2.8.2 of materials and methods), and the resulting scans placed on the surface at the appropriate molar assembly ratio of octamer:DNA between the experimental scans. The surface so constructed is shown in figure 4.6.

Note that the densitometric profiles on the surface corresponding to molar assembly ratios of octamer:DNA of 0, 3, 5, 7, 9 and 11 are identical to the experimentally determined scans shown in figure 4.5. This is because the surface is not fitted to the experimental scans, but obtained by interpolation between them. If attention is next drawn to the modulation of the surface itself, several features of the dependence of the MNase digestion profile on the assembly ratio of octamer:DNA discussed above, is evident. The yield of core length DNA fragment (peak on the right-hand side of the surface) increases with increasing

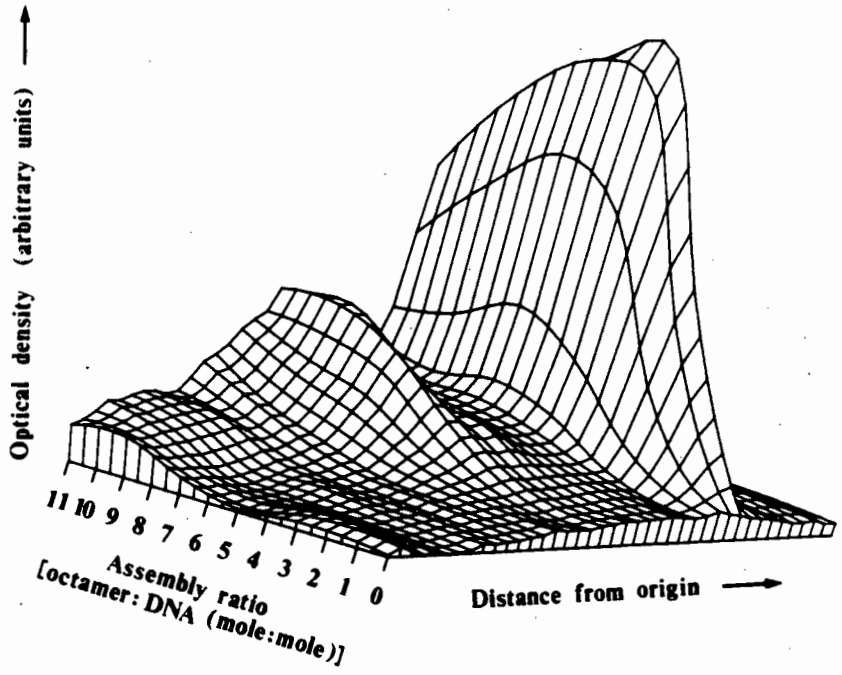

Figure 4.6: Densitometric surface of micrococcal nuclease digested assembled pHP2. The surface was composed by the alignment of densitometric scans of electrophoretically separated micrococcal nuclease digests of supercoiled pHP2 reconstituted by urea/salt dialysis at octamer:DNA (mole/mole) ratios of 0, 3, 5, 7, 9 and 11. Densitometric profiles corresponding to intermediate assembly ratios were obtained by cubic spline interpolation of these experimental scans (see section 2.8.2 of materials and methods). The assembly ratio corresponding to individual scans is shown on the left-hand side of the surface. The edge of the surface approximately in the plane of the page indicates the distance from the electrophoretic origin and incorporates fragment sizes from about 700bp - 90bp, with the direction of electrophoresis from left to right. The large peak on the far right of the surface corresponds to the nucleosome core monomer (both 166bp and 146bp) and the peak towards the left of the surface to the nucleosome core dimer. Every fifth datumpoint is represented on the surface which was plotted in perspective with hidden lines removed for clarity. See text for detail.

assembly ratio to reach a maximum at an assembly ratio of approximately 5, and decreases at higher assembly ratios. DNA fragments corresponding to the core dimer (left-hand side peak on the surface) start to appear at an assembly ratio of approximately 3, and the yield of the fragment constituting this peak increases up to an assembly ratio of approximately 10, where it appears to level off. At very high assembly ratios (8 - 11 moles octamer per mole plasmid), the appearance of DNA fragments longer that those corresponding to the core dimer is visible in the left-hand, back corner of the surface.

All this is very sensible in terms of what has been learned from the densitometric profiles shown in figure 4.5. However, one now has a continuous surface corresponding to the MNase digestion profile at molar assembly ratios of octamer:DNA from 0 to 11, and, by reference to figure 4.3, the average number of octamers associated with the supercoiled DNA molecule at each of these assembly ratios. Thus, if a carefully controlled MNase digestion of assembled linearized plasmid is performed in parallel to the assembled supercoiled molecule, densitometric scans will be obtained from which the average number of nucleosome cores per linearized molecule can be deduced by finding the scan on the densitometric surface of the supercoiled molecule that fits the scan of the linearized molecule most closely.

This approach makes the assumption that the rate of cleavage by MNase of supercoiled and linearized assembled molecules is identical. This seems fair. It has been proposed that MNase binds to the sugar-phosphate backbone of duplex DNA (70, 248) and was shown to catalyze hydrolysis of the P-O5' phosphodiester bond preferentially in regions of low helix stability such as at d(T-A) dinucleotides where the nuclease may recognize a structural feature of single stranded DNA (70). It has also been shown that negative superhelicity

facilitates melting of the d(T-A) containing TATA-box region of the *tyrT* promoter (62) and in this regard one may propose that supercoiling would stabilize single-stranded transitions of the DNA duplex recognized by MNase. The initial rate of cleavage of supercoiled DNA would therefore be higher than of linearized DNA. Flick *et al* (70), however, proposed that the initial MNase-DNA interaction involves duplex DNA, which is supported by the cleavage stagger observed on the two stands of the duplex, and Tucker *et al* (248) reported that the activation energy of cleavage, and therefore the rate-limiting step, is the same for double- and single-stranded DNA. It is therefore unclear what, if any, the difference between the initial rate of cleavage by MNase of supercoiled and linearized DNA would be. However, the extent of digestion of the assembled supercoiled and linearized DNA required by the quantitation method is such to allow the appearance of mostly core length and core dimer length fragments, with some fragments shorter that core length visible (refer to figure 4.4). The appearance of fragments trimmed to this length would have involved several cuts by MNase to the same molecule. In the case of a supercoiled molecule, any distinction that MNase may make between a supercoiled and linearized molecule disappears after the first cut. Several additional cuts must then still be made to the same molecule, of which the origin is now irrelevant, to appear as a fragment on the gel at a position where the majority of the digested DNA migrates (700bp - 90bp). The effect of differences between the rate of the initial cut to supercoiled and linearized DNA, if any, is therefore likely to be negligible in the electrophoretic profile of the MNase digestion ultimately obtained.

Once the densitometric surfaces representing the MNase digestion of supercoiled and linearized DNA assembled at different input ratios of octamer:DNA have been obtained, a

method must be decided on to find the position on the supercoiled surface where the densitometric profile best fits that of the relevant profile from the linearized surface. This entails comparing each profile from the linearized surface, in turn, with all the profiles on the supercoiled surface. The method of least-squares was used for this purpose. The sum of the squares of the differences between corresponding datapoints was calculated for pairs of profiles chosen from the linearized and supercoiled surfaces at intervals of 1 and 0,1 respectively. The number of cores associated with a linearized DNA molecule at a given molar input ratio of octamer:DNA was indicated by the profile on the supercoiled surface that gave the smallest sum of the squares of the differences between the two profiles. The number of cores represented by the identified profile on the supercoiled surface was known from the topoisomerase I relaxation results. See section 2.8.1 of materials and methods for more detail.

Before proceeding further, one may now verify the ability of the cubic spline interpolation procedure to calculate densitometric scans at assembly ratios intermediate to those determined by experiment. This was performed by recalculating the surface shown in figure 4.6 with one of the experimental scans omitted. The densitometric profile at the relevant ratio was thus replaced with a scan obtained by interpolation. The closeness with which this interpolated scan matched the omitted experimental scan was then determined by fitting the experimental scan to the surface as explained above. This was performed, in turn, for assembly ratios of 3, 5, 7 and 9. The average fit was within 0,7 cores of the real experimental value. This result indicates that the intermediate densitometric scans, where the interpolation interval is even smaller than the test cases above, will match the true experimental scans with high accuracy. With the integrity of the interpolation

method firmly established, one may now proceed to investigate the assembly efficiency of linearized DNA.

4.3 Efficiency of core reconstitution by urea/salt dialysis:

Supercoiled and linearized pHP2 were assembled by urea/salt dialysis over a range of molar octamer:DNA ratios, an aliquot of each sample digested under carefully controlled conditions with MNase, and the digestion products resolved by electrophoresis (see section 2.6.1 of materials and methods). The result is presented in figure 4.7. An aliquot of each supercoiled sample was also used for a linking difference determination after relaxation with topoisomerase I. These result were discussed above (section 4.2).

If attention is turned first to the MNase digested supercoiled plasmid (lanes 1 - 6 of figure 4.7), it is seen that a gradual change in the yield and distribution of digestion fragments, similar to that encountered previously (figure 4.4), was obtained. This is as expected and simply reflects the decrease in the rate of cleavage by MNase of DNA molecules containing, on average, shorter stretches unoccupied by the histone octamer at higher assembly ratios. If a closer comparison of the current and the previous result is made (figure 4.7 and figure 4.4), it is clear that the total amount of digestion obtained differs between the two experiments. In lane 6 of figure 4.7, for instance, a significant amount of material migrates above the position of the fragment corresponding to the core dimer (280bp - 290bp), whereas, in lane 6 of figure 4.4, corresponding to an identical assembly ratio, the majority of the digested material migrate as core monomer and dimer fragment. This is probably due to small differences in the enzyme concentration of the two experiments, where the enzyme was taken from different diluted stock solutions. To

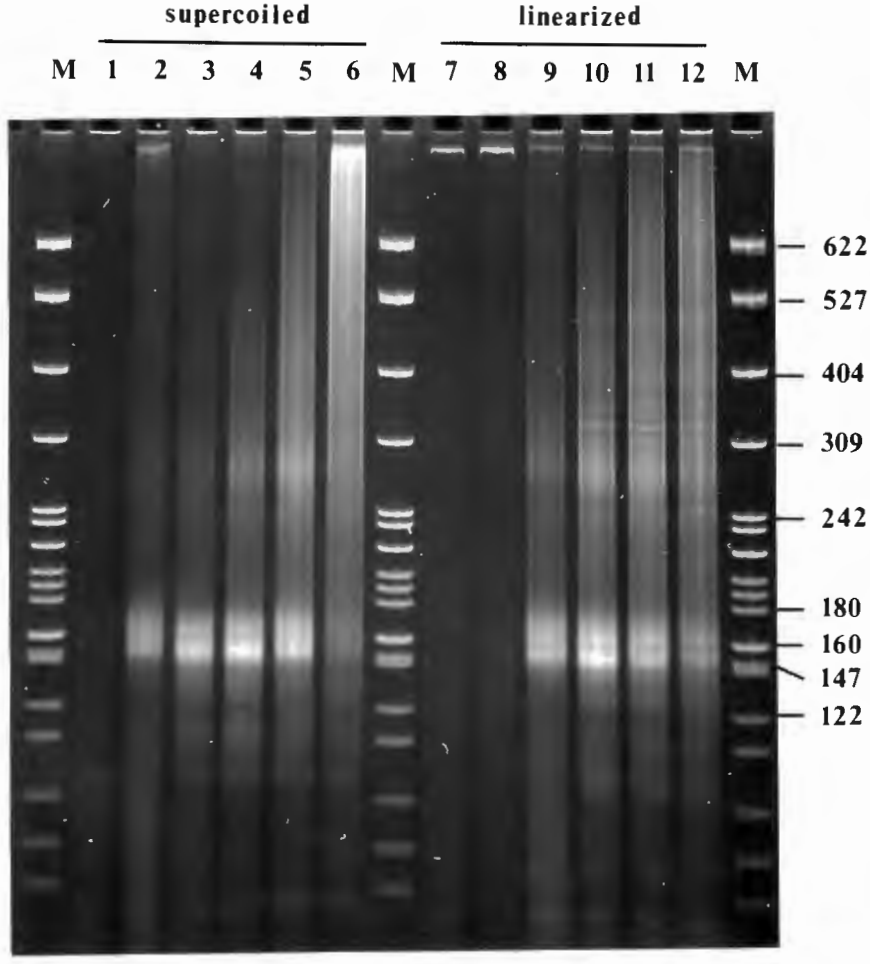

Figure 4.7: Micrococcal nuclease digestion of plasmid pHP2 reconstituted into nucleosome cores. Supercoiled (lanes 1 - 6) and linearized (lanes 7 - 12) pHP2 were assembled by urea/salt dialysis at different molar ratios of octamer:DNA followed by digestion with 1U micrococcal nuclease at 37°C for exactly 10min and the digestion products resolved by electrophoresis on a 10% (w/v) acrylamide gel. The gel was stained with ethidium bromide. The assembly was performed at octamer:DNA (mole/mole) ratios of 0 (lanes 1 and 7), 3 (lanes 2 and 8), 5 (lanes 3 and 9), 7 (lanes 4 and 10), 9 (lanes 5 and 11) and 11 (lanes 6 and 12). A *Hpa* II digest of plasmid pBR322 was included as size standard (lanes labeled M) with fragment sizes in base-pairs indicated in the right-hand margin.

minimize the contribution of this factor to differences between supercoiled and linearized plasmid, the comparative experiments were carried out in parallel with enzyme dispensed from the same freshly prepared solution.

If the MNase digestion result of the linearized plasmid is considered next (lanes 7 - 12 of figure 4.7), it is immediately obvious that a significant difference exists between the assembly efficiencies of supercoiled and linearized DNA. At a molar assembly ratio of octamer:DNA of 3, the yield of MNase digestion of the linearized plasmid is indistinguishable from that of free linear plasmid (lanes 7 and 8 of figure 4.7). This is not due to the complete loss of the DNA sample during the purifications steps preceding electrophoresis, as indicated by the presence of residual full length linear fragment at similar yields, and the appearance of background fluorescence of similar intensities, in lane 7 and 8. Moreover, the same result was obtained from three independent repetitions of the experiment (data not shown). The result therefore indicates that, in contrast to supercoiled DNA, very little octamer associates with linearized DNA as nucleosome cores at low assembly ratios. At intermediate molar assembly ratios of 5 - 9, the electrophoretic profiles of the linearized plasmid (lanes 9 - 11) appear similar to that of supercoiled plasmid (lanes 3 - 5). At a molar assembly ratio of 11, the electrophoretic profile of linearized DNA (lane 12) appears to correspond to that of supercoiled DNA between the assembly ratios of 9 and 11 (lanes 5 and 6).

A visual inspection of the gel therefore indicates that the efficiency of nucleosome core reconstitution on linearized DNA is less than on supercoiled DNA at very low and at very high ratios of octamer:DNA. However, to obtain a quantitative measure over the full assembly range, it is necessary to determine the average number of nucleosome

cores per DNA molecule at every assembly ratio. This was achieved by densitometric scanning of lanes 1 - 6 and 7 - 12 of figure 4.7 on the film negative, the clipping and alignment of the experimental scans, and the calculation of the densitometric surface for supercoiled and linearized DNA as described above (see also section2.8.2 of materials and methods). The result is shown in figure 4.8.

If the densitometric surface corresponding to the MNase digestion of the assembled supercoiled molecules (figure 4.8 A) is considered first, features previously encountered (figure 4.6) are visible. The yield of monomer fragment (right-hand side peak on surface) is seen to increase steadily with increasing assembly ratio to reach a maximum at intermediate ratios, and then decline. The dimer peak (left-hand side of surface) appears at intermediate ratios and increases to reach a maximum at high assembly ratios where there is a pronounced presence of fragment above dimer size (left-hand, back corner of surface). On the densitometric surface corresponding to the MNase digests of the assembled linearized plasmid (figure 4.8 B), the striking feature is the absence of core length fragment at low ratios of assembly. Apart from this feature, the general trend in the change of the densitometric profiles appears very similar between supercoiled and linearized DNA.

The number of octamers associated with the linearized DNA at molar assembly ratios of 0 - 11 was calculated at intervals of 1 by finding the densitometric profile on the supercoiled surface that fits the relevant profile from the linearized surface best (see section 4.2 above and section 2.8.2 of materials and methods). Once the corresponding profile on the supercoiled surface had been identified, the number of nucleosome cores on the linearized molecule was obtained from figure 4.3. This was performed in three independent experiments. The result is shown in figure 4.9.

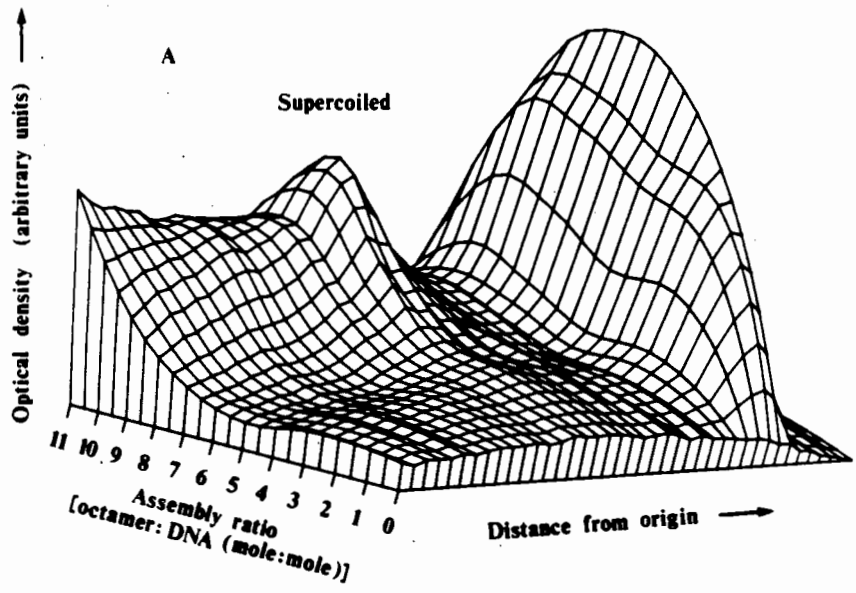

Figure 4.8 A: Densitometric surface of micrococcal nuclease digested assembled pHP2. The surface was obtained by the alignment of densitometric scans of electrophoretically separated micrococcal nuclease digests of supercoiled pHP2 reconstituted by urea/salt dialysis at octamer:DNA (mole/mole) ratios of 0, 3, 5, 7, 9 and 11. Densitometric profiles corresponding to intermediate assembly ratios were obtained by cubic spline interpolation of the experimental scans (see section 2.8.2 of materials and methods). The assembly ratio corresponding to individual scans is shown on the left-hand side of the surface. The edge of the surface approximately in the plane of the page indicates the distance from the electrophoretic origin and incorporates fragment sizes from about 700bp - 90bp, with the direction of electrophoresis from left to right. The nucleosome core monomer (both 166bp and 146bp) and dimer peaks are on the right- and left-hand side of the surface respectively. Every fifth datumpoint is represented on the surface which was plotted in perspective with hidden lines removed for clarity. The surfaces in figure 4.8 A and B were plotted to the same scale and individual densitometric profiles from the two surfaces are directly comparable.

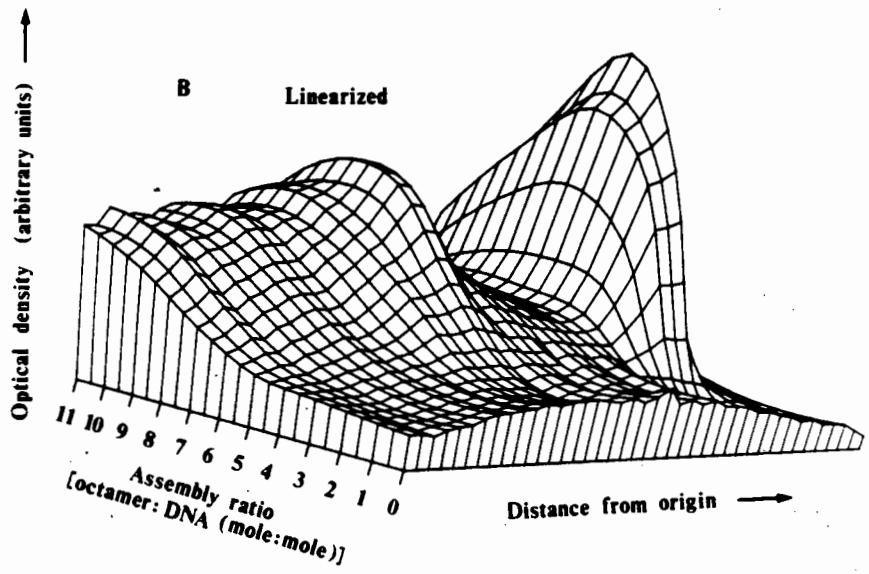

Figure 4.8 B: Densitometric surface of micrococcal nuclease digested assembled pHP2. The surface was obtained by the alignment of densitometric scans of electrophoretically separated micrococcal nuclease digests of linearized pHP2 reconstituted by urea/salt dialysis at octamer:DNA (mole/mole) ratios of 0, 3, 5, 7, 9 and 11. Densitometric profiles corresponding to intermediate assembly ratios were obtained by cubic spline interpolation of the experimental scans (see section 2.8.2 of materials and methods). The assembly ratio corresponding to individual scans is shown on the left-hand side of the surface. The edge of the surface approximately in the plane of the page indicates the distance from the electrophoretic origin and incorporates fragment sizes from about 700bp - 90bp, with the direction of electrophoresis from left to right. The nucleosome core monomer (both 166bp and 146bp) and dimer peaks are on the right- and left-hand side of the surface respectively. Every fifth datumpoint is represented on the surface which was plotted in perspective with hidden lines removed for clarity. The surfaces in figure 4.8 A and B were plotted to the same scale and individual densitometric profiles from the two surfaces are directly comparable.

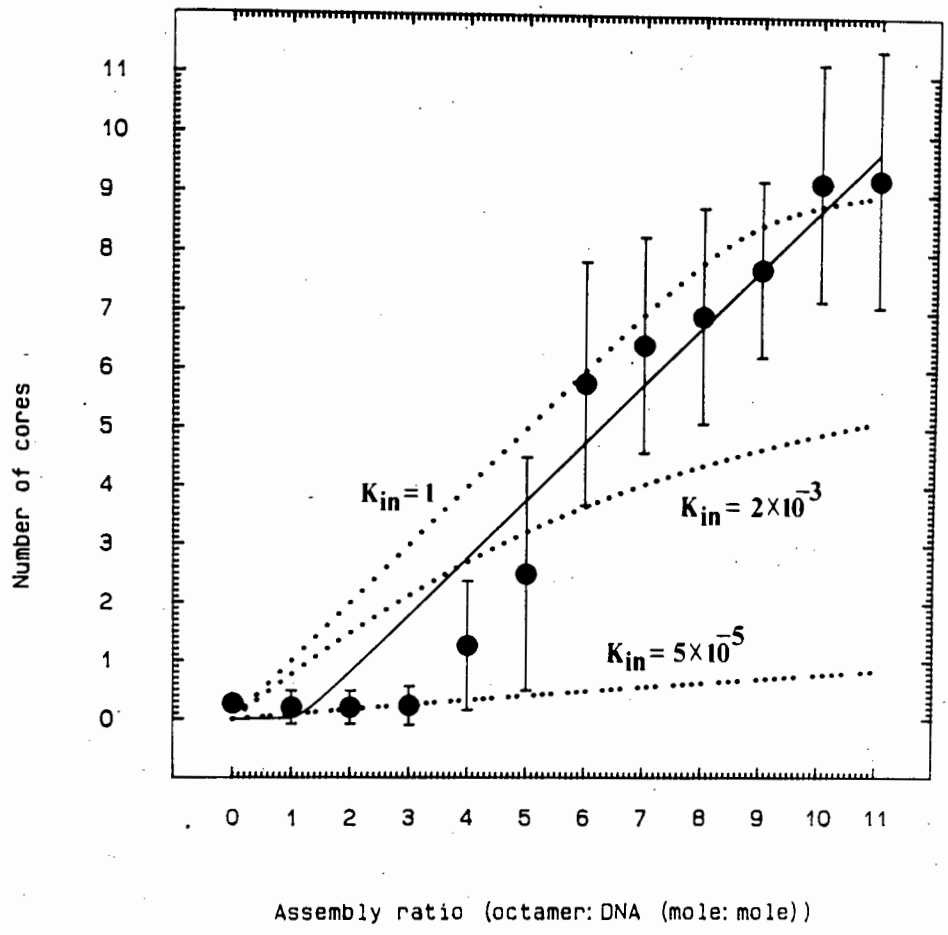

Figure 4.9: Assembly efficiency of linearized DNA. Plot of the number of nucleosome cores assembled by urea/salt dialysis onto linearized pHP2 as a function of the input ratio of octamer:DNA. Each datumpoint is the average of three independent determinations, performed as explained in the text, with the standard deviation indicated by brackets where this exceeds the width of the plotted datumpoint (±0,16). The solid line (——) represents the binding curve predicted by the homogeneous lattice theory of multiple equilibria which was minimized to the datapoints by the method of least squares in terms of the intrinsic association constant (K_{in}=1X10^{-7}) and the cooperativity constant (ω=8X10^{6}). The dotted lines (··) represent the binding curve in the absence of any cooperative effects (ω=1) with K_{in} = 1, 2X10^{-3} and 5X10^{-5}, as indicated in the figure. See text for detail.

The distribution of the datapoints in figure 4.9 indicates that the assembly efficiency of linearized DNA is less than that of supercoiled DNA over the entire assembly range. At low molar ratios of octamer:DNA, very little core reconstitution takes place. In order to identify the factors responsible for this assembly behavior, one may analyze the binding data in terms of ligand interactions.

The binding of the histone octamer to DNA cannot be viewed in terms of simple independent binding-site interactions, since the association of one octamer with DNA does not involve the binding to a single, well-defined site, but an extended length of DNA (nucleosome positioning may be neglected). Furthermore, if the length of free DNA between two closely placed cores is less than core length, binding by a third octamer to this stretch is impeded. It is thus clear that the type of model required to describe the interaction of the histone octamer with DNA must incorporate this binding-site overlap component.

A suitable model is the homogeneous lattice theory of multiple equilibria of McGhee and von Hippel (158). It was proposed (158) that the ratio of the average number of ligand bound per lattice residue (v) to the free ligand concentration (L) is equal to:

$$K(1-nv)\left[\frac{(2\omega+1)(1-nv)+v-R}{2(\omega-1)(1-nv)}\right]^{n-1}\left[\frac{1-(n+1)v+R}{2(1-nv)}\right]^{2} \qquad (4.6)$$

where K is the intrinsic association constant, n is the number of residues per ligand binding site, v is as defined above, ω is the cooperativity constant, and R is equal to:

$$\sqrt{\{[1-(n+1)v]^2+4\omega v(1-nv)\}} \qquad (4.7)$$

In the present context the ligand is the octamer, the lattice is the DNA molecule, and the lattice residue is the nucleotide. Thus, a specific value of v can be substituted into equation 4.6 and, with appropriate choices for K and ω, the ratio v/L calculated. Since v is known, it possible to calculate L, the molar concentration of free octamer. v can be translated to the number of moles of octamers bound per mole of DNA molecule by multiplying by the total number of nucleotides per DNA molecule (1915bp). Thus, if K and ω is known, it is possible to calculate the molar assembly ratio of octamer:DNA that corresponds to a specific value for the average number of assembled cores. The question is, what are appropriate values for K (or K_{in}) and ω? Before undertaking a search for numeric values, one may obtain a qualitative indication by fitting selected binding curves to the datapoints shown in figure 4.9.

In figure 4.9 a dotted line which corresponds to the binding curve where $K_{in}=2X10^{-3}$ and ω=1 (no cooperativity) is shown. The shape of this binding curve epitomizes a quality of overlapping-site ligand binding. McGhee and von Hippel (158) showed that the initial slope of the binding curve is a function of the intrinsic association constant and the binding site size. However, at higher degrees of saturation, there is an accumulation of octamer free gaps smaller than core size, and the DNA molecule now entropically resists being saturated (158). This is mostly a function of the binding site size. In order to obtain a binding curve for non-cooperative binding that reflects the experimental degree of binding at high input ratios, the association constant must be increased. Note, however, that the initial slope of the dotted line in figure 4.9 does not represent the poor binding of the octamer at low assembly ratios well. If a larger association constant is used, this feature is even less well represented (dotted line with $K_{in}=1$). It can only be reproduced by choosing a very small

association constant (dotted line with $K_{in}=5X10^{-5}$), although the binding of the octamer at higher assembly ratios is then not correctly predicted. It is thus clear that the experimental binding response must reflect a small association constant combined with positive cooperativity.

Numeric values for K and ω were obtained by performing a semi-iterative procedure whereby the binding curve relating the number of assembled cores to the assembly ratio was calculated as outlined above with n=174, and fitted by least-squares to the experimental datapoints shown in figure 4.9. This was performed over a range of values for K and ω to obtain those that gave a binding curve that fitted the experimental data best. The result obtained from this procedure set $K=1X10^{-7}$ and $\omega=8X10^{6}$ (solid line in figure 4.9). These are very small and very large values respectively. An indication of the soundness of this combined choice may be obtained from considerations of the standard free energy change that would accompany ligand-lattice interactions with these binding parameters. If the octamer binds to an available site closely bordered on either side by bound octamer, the apparent association constant will be equal to $K\omega^2$. Thus, if measurements of the free energy change of core formation is made by reconstituting polycores, the measured quantity will be equal to $-RTlnK\omega^2$. If the values of K and ω obtained above are substituted into this equation, ΔG=-10kcal/mole. This is close to the value of -15.6kcal/mole obtained by Stein (227) for core formation on 140bp DNA molecules in 0,6M NaCl at 26°C. It would therefore seem that the combined choice of values for K and ω is not unrealistic. However, the quantities measured by experiment is not entirely suitable for an accurate determination of binding constants. The classical approach in determinations of the association constant involves several measurements of the *free* ligand concentrations at different total ligand concentrations

(158). This was not the quantity of interest in the design of the experiment. In this regard it may be mentioned that combined values of K and ω ranging from $5X10^{-5}$ and $1X10^{4}$ to $1X10^{-10}$ and $8X10^{9}$, gave subjectively acceptable fits to the experimental datapoints of figure 4.9. The values of K and ω were, however, consistently very small and very large, respectively, as suggested by inspection above. Thus, one may accept the values in a qualitative fashion, and see what can be learned from them. This is done in the discussion at the end of this chapter.

4.4 Efficiency of core reconstitution by salt dialysis:

The assembly efficiency of core reconstitution by salt dialysis was investigated by an identical approach to that described for urea/salt dialysis above. In figure 4.10 the electrophoretic separation of topoisomers of supercoiled pHP2 assembled over a range of molar octamer:DNA ratios is shown after topoisomerase I relaxation. The electrophoretic distribution of the topoisomer populations exhibit a shift towards more negative linking differences with increasing assembly ratios (figure 4.10 A), indicating the association of more cores with the DNA as the input ratio of octamer:DNA is increased. The lanes in figure 4.10 where individual topoisomers were clearly resolved were densitometrically scanned on the film negative. The scans are shown in figure 4.11. The linking difference of the topoisomer population at every assembly ratio following topoisomerase I relaxation was calculated (see section 2.8.1 of materials and methods), and is shown in figure 4.12 versus the molar assembly ratio.

The distribution of the datapoints in figure 4.12 suggests that the number of assembled cores is a linear function of the assembly ratio, and for this reason a line was fitted to the datapoints by the method of least squares. The obtained

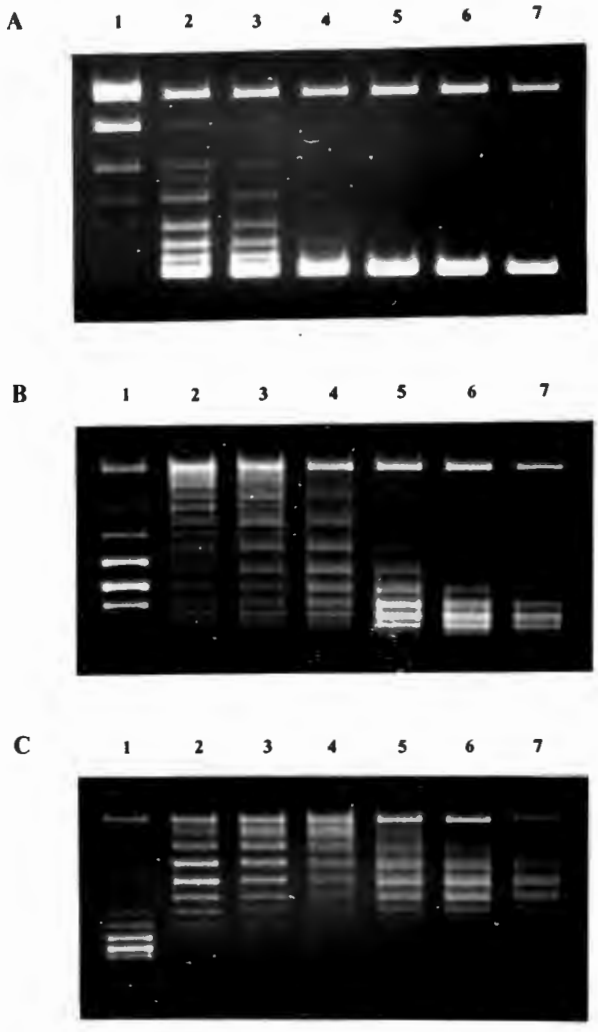

Figure 4.10: DNA topoisomer distribution of supercoiled pHP2 assembled at different octamer:DNA ratios by the method of salt dialysis. Supercoiled pHP2 was assembled by salt dialysis at octamer:DNA (mole/mole) ratios of 0, 3, 5, 7, 9 and 11 (lanes 1 - 6) and relaxed with 8U of eukaryotic topoisomerase I. Aliquots of the sample at each assembly ratio were electrophoresed on 1,5% (w/v) agarose gels run in TBE containing chloroquine at 0μg/ml (A), 1μg/ml (B) or 3μg/ml (C) in order to resolve individual topoisomers and the gels stained with ethidium bromide. Native supercoiled pHP2 (no topoisomerase I treatment) was included in lane 7 of each gel.

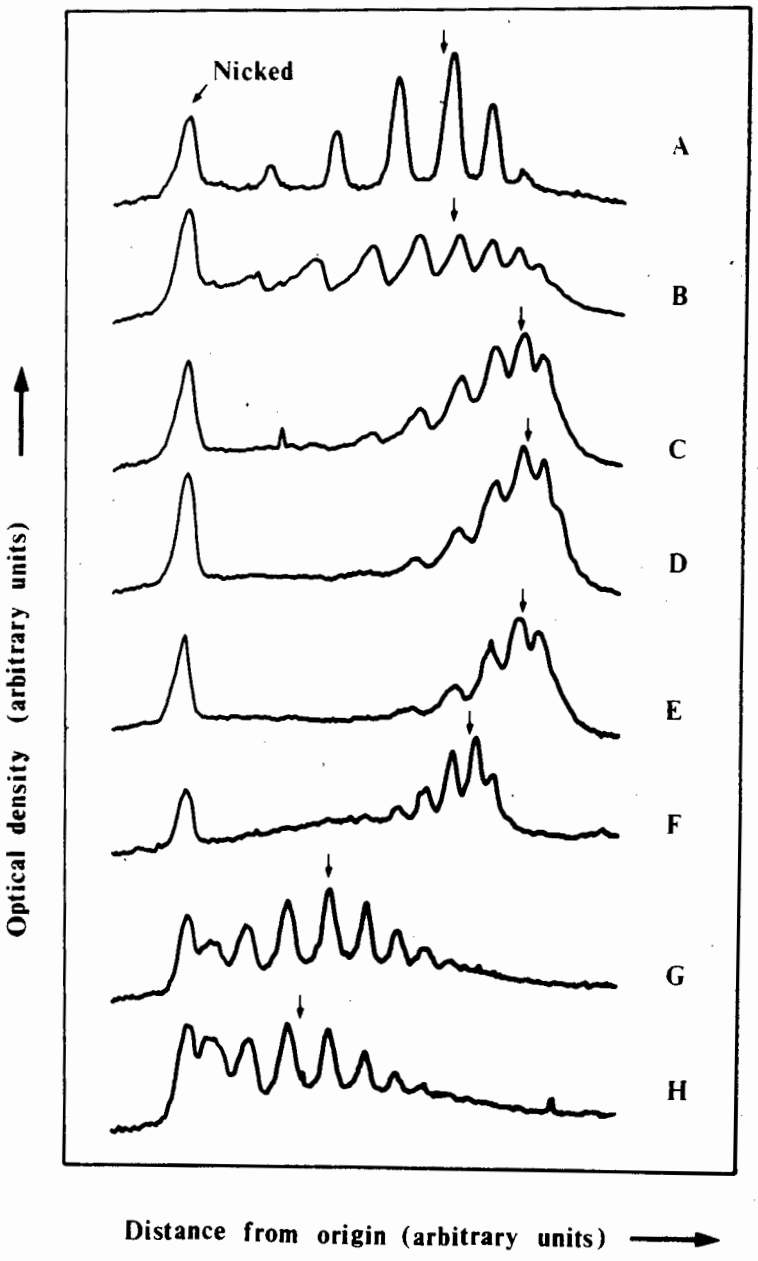

Figure 4.11: Densitometric scans of resolved topoisomers. Selected lanes of the gel shown in figure 4.10 were densitometrically scanned on the film negative and the scans aligned relative to the electrophoretic origin. The lanes where chosen to clearly show the Gaussian center of the topoisomer distribution (indicated by an arrow above each scan; see section 2.8.1 of materials and methods) at every experimental assembly ratio of octamer:DNA and represent distributions obtained in the presence of chloroquine at a concentration of 1μg/ml (A - E) or 3μg/ml (F - H). Individual scans (A - H) correspond to assembly ratios of octamer:DNA (mole/mole) of 0, 7, 9, 11, native supercoiled (no topoisomerase I treatment), 0, 3 and 5. The position of the nicked, circular plasmid is indicated in the top, left-hand side of the figure.

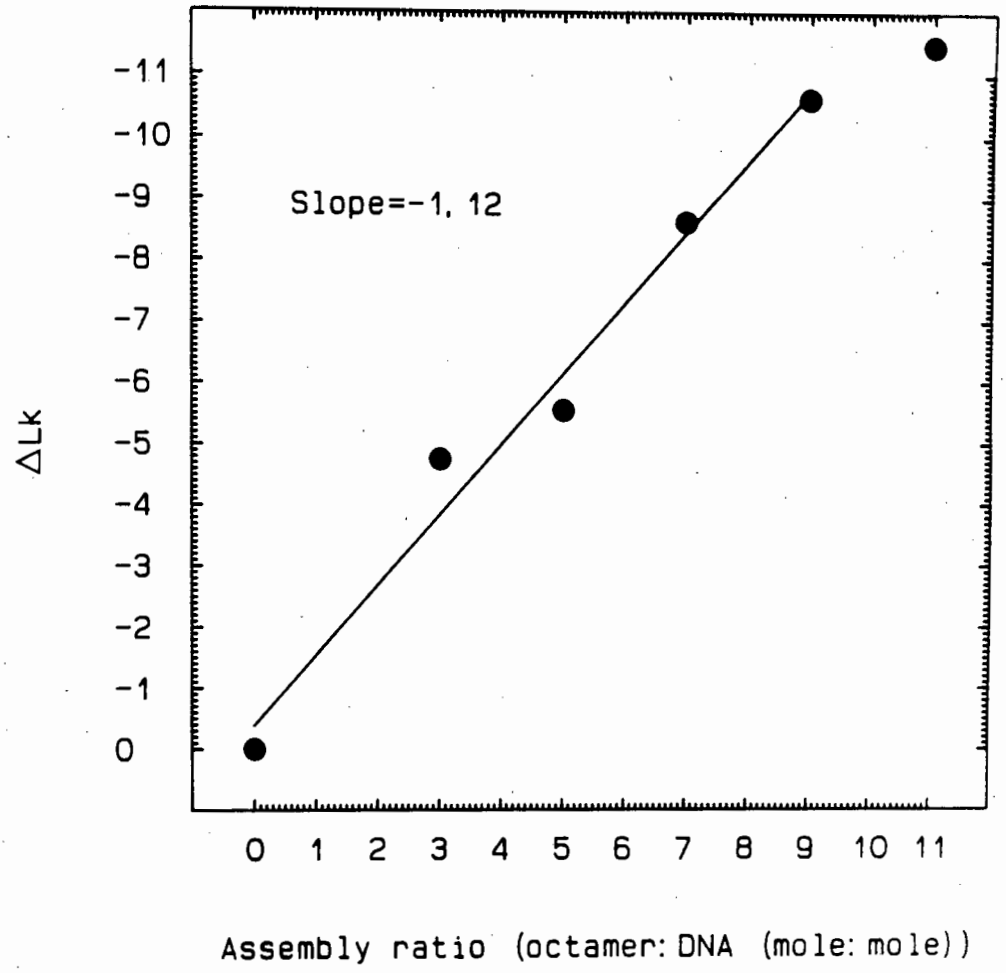

Figure 4.12: The efficiency of reconstitution of supercoiled DNA by salt dialysis. Plot of the assembly ratio of octamer:DNA (mole/mole) *versus* the linking difference following topoisomerase I treatment. The line was fitted to the datapoints by the method of least squares. The datumpoint corresponding to an assembly ratio of 11moles/mole was omitted from the fitting procedure (see text).

slope (-1,1) indicates that the average number of cores assembled per molecule at a given assembly ratio is equal to the molar input ratio of octamer:DNA, since it is well established that the number of supercoils constrained by the core corresponds to a linking difference of -1,0 (221) to -1,1 (81, 283). The efficiency of assembly of supercoiled plasmid by salt dialysis therefore appears identical to that of urea/salt dialysis. The linking difference of the relaxed, assembled plasmid at a molar ratio of octamer:DNA of 11 is again well below the extrapolation point of the line fitted to the rest of the experimental data set. This apparent decrease in the efficiency of assembly of supercoiled DNA at octamer:DNA ratios close to saturation was proposed to reflect the introduction of net positive superhelical stress into the plasmid, which may be assumed to be relatively inflexible when almost fully occupied with the histone octamer. This phenomenon is discussed in detail at the end of the chapter.

The assembly efficiency of salt dialysis on linearized plasmid was investigated as described above. Figure 4.13 shows the result of the electrophoretic separation of MNase digested supercoiled and linearized plasmid assembled over a range of molar octamer:DNA ratios. Visual inspection of the gel shown in figure 4.13 reveals a trend in the shift of fragment yield and distribution similar to that of urea/salt dialysis assembly (see section 4.2 and 4.3). If the individual lanes corresponding to identical assembly ratios of supercoiled and linearized DNA in figure 4.13 are compared, it is clear that the yield of core length fragment is higher at low assembly ratios for supercoiled DNA (lane 2) than for linearized DNA (lane 8). A similar decreased yield at higher assembly ratios is also apparent (compare lanes 12 and 6). Inspection of the gel therefore suggests that the assembly efficiency of supercoiled *versus*

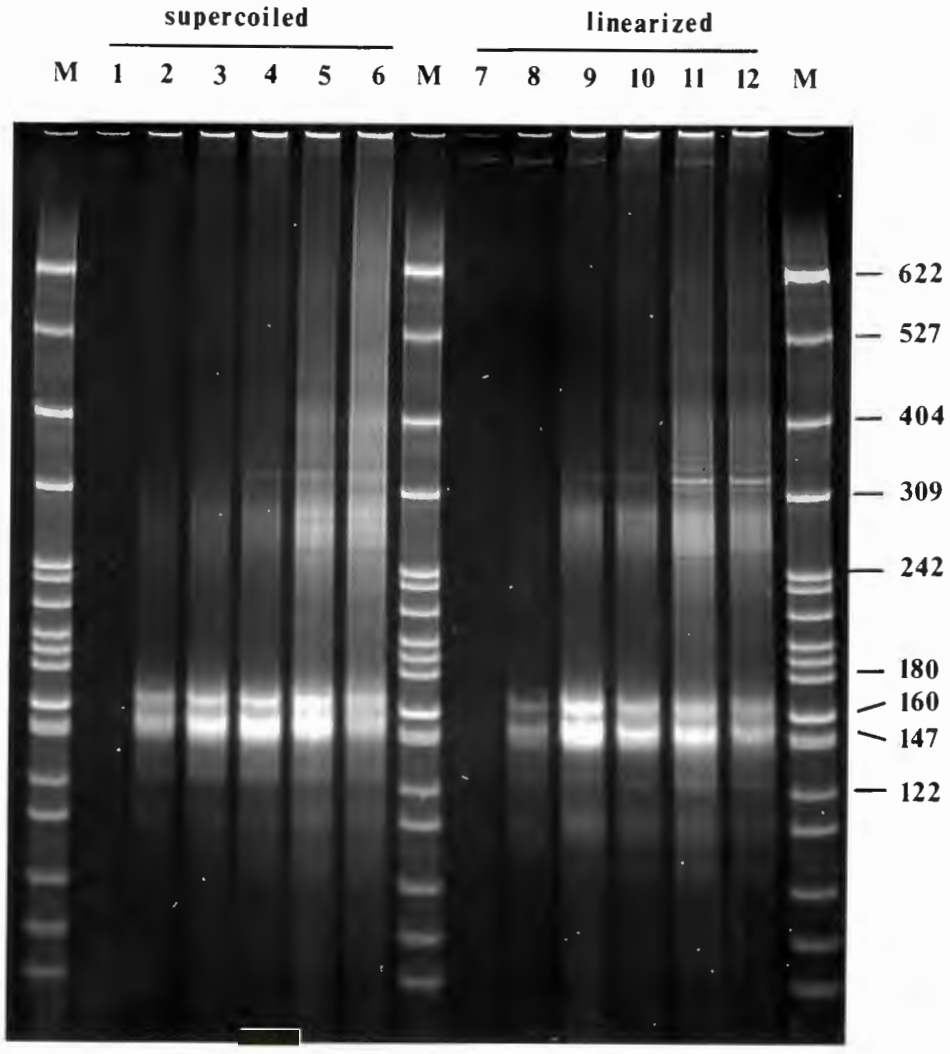

Figure 4.13: Micrococcal nuclease digestion of plasmid pHP2 reconstituted into nucleosome cores. Supercoiled (lanes 1 - 6) and linearized (lanes 7 - 12) pHP2 were assembled by salt dialysis at different molar ratios of octamer:DNA followed by digestion with 1U micrococcal nuclease at 37°C for exactly 10min and the digestion products resolved by electrophoresis on a 10% (w/v) acrylamide gel. The gel was stained with ethidium bromide. The assembly was performed at octamer:DNA (mole/mole) ratios of 0 (lanes 1 and 7), 3 (lanes 2 and 8), 5 (lanes 3 and 9), 7 (lanes 4 and 10), 9 (lanes 5 and 11) and 11 (lanes 6 and 12). A *Hpa* II digest of plasmid pBR322 was included as size standard (lanes labeled M) with fragment sizes in base-pairs indicated in the right-hand margin.

linearized DNA of salt dialysis may be very similar to that of urea/salt dialysis.

A quantitative measure of this difference was obtained by densitometric scanning of the film negative and the construction and systematic comparison of densitometric surfaces corresponding to supercoiled and linearized DNA as described above (section 4.2 and section 2.8.2 of materials and methods). The densitometric surfaces obtained by interpolation of the experimental scans are shown in figure 4.14, and the result of the stepwise comparison in figure 4.15.

The result shown in figure 4.15 clearly indicates that, similar to urea/salt dialysis, the assembly efficiency of linearized DNA assembled by salt dialysis is less than that of supercoiled DNA over the whole assembly range tested, and that linear DNA displays an especially low affinity for the histone octamer at low molar input ratios of octamer:DNA.

An inspection of the distribution of the datapoints in figure 4.15 similarly suggests that the octamer binds in a positively cooperative way to the DNA, whereas binding of the octamer to the DNA at low assembly ratios, when the cooperative effect will be mostly absent (158) indicates a low intrinsic association constant. Values for these two parameters were calculated as described above (see section 4.3) and yielded $K=1X10^{-8}$ and $\omega=1X10^{8}$. As before, the values are respectively very small and very large. If the magnitude of these values are excepted in a qualitative fashion, one may compare theoretical predictions of binding behavior with that determined by experiment. For instance, the magnitude of the association constant suggests that core particles, where there can be no cooperative contribution to binding, should be unstable in solution, and that reconstitution of cores on core length DNA, where

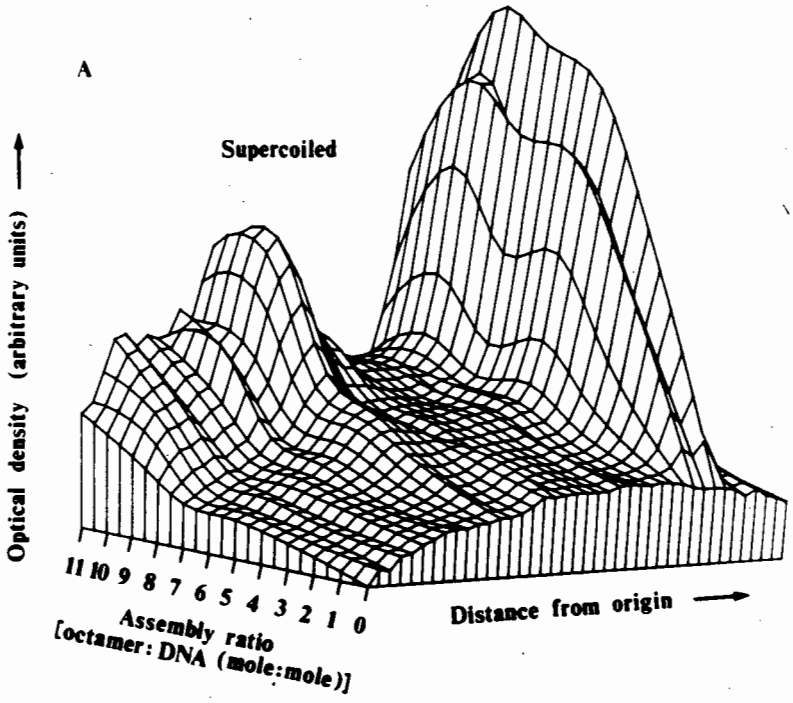

Figure 4.14 A: Densitometric surface of micrococcal nuclease digested assembled pHP2. The surface was obtained by the alignment of densitometric scans of electrophoretically separated micrococcal nuclease digests of supercoiled pHP2 reconstituted by salt dialysis at octamer:DNA (mole/mole) ratios of 0, 3, 5, 7, 9 and 11. Densitometric profiles corresponding to intermediate assembly ratios were obtained by cubic spline interpolation of the experimental scans (see section 2.8.2 of materials and methods). The assembly ratio corresponding to individual scans is shown on the left-hand side of the surface. The edge of the surface approximately in the plane of the page indicates the distance from the electrophoretic origin and incorporates fragment sizes from about 700bp - 90bp, with the direction of electrophoresis from left to right. The nucleosome core monomer (both 166bp and 146bp) and dimer peaks are on the right- and left-hand side of the surface respectively. Every fifth datumpoint is represented on the surface which was plotted in perspective with hidden lines removed for clarity. The surfaces in figure 4.14 A and B were plotted to the same scale and individual densitometric profiles from the two surfaces are directly comparable.

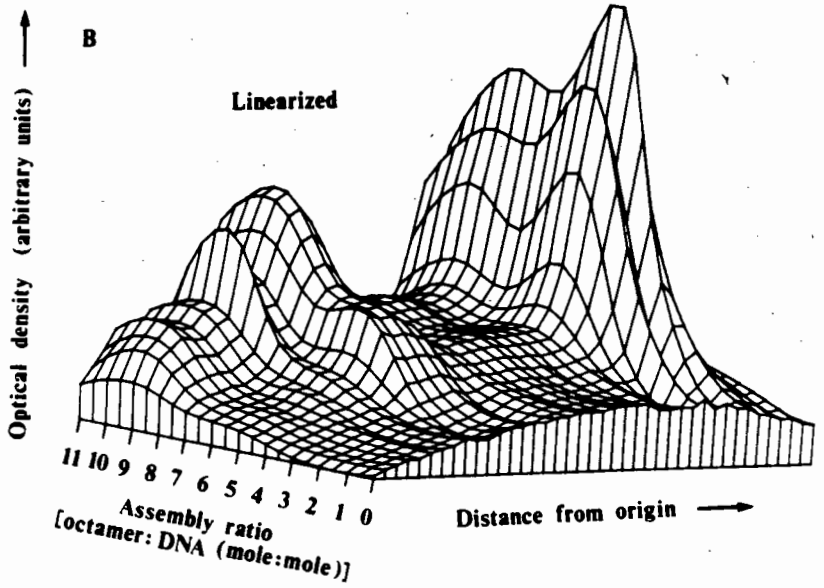

Figure 4.14 B: Densitometric surface of micrococcal nuclease digested assembled pHP2. The surface was obtained by the alignment of densitometric scans of electrophoretically separated micrococcal nuclease digests of linearized pHP2 reconstituted by salt dialysis at octamer:DNA (mole/mole) ratios of 0, 3, 5, 7, 9 and 11. Densitometric profiles corresponding to intermediate assembly ratios were obtained by cubic spline interpolation of the experimental scans (see section 2.8.2 of materials and methods). The assembly ratio corresponding to individual scans is shown on the left-hand side of the surface. The edge of the surface approximately in the plane of the page indicates the distance from the electrophoretic origin and incorporates fragment sizes from about 700bp - 90bp, with the direction of electrophoresis from left to right. The nucleosome core monomer (both 166bp and 146bp) and dimer peaks are on the right- and left-hand side of the surface respectively. Every fifth datumpoint is represented on the surface which was plotted in perspective with hidden lines removed for clarity. The surfaces in figure 4.14 A and B were plotted to the same scale and individual densitometric profiles from the two surfaces are directly comparable.

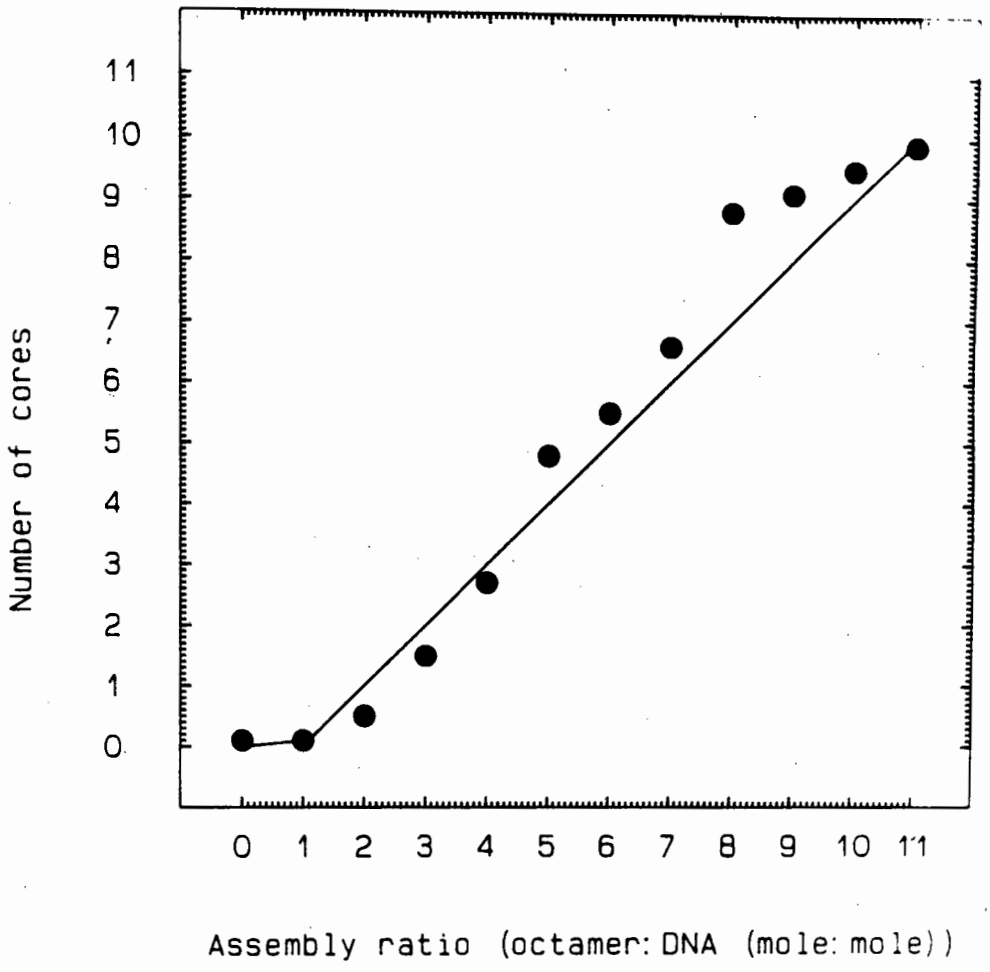

Figure 4.15: Assembly efficiency of linearized DNA. Plot of the number of nucleosome cores assembled by salt dialysis onto linearized pHP2 as a function of the input ratio of octamer:DNA. Each datumpoint was calculated as explained in the text. The solid line (—) represents the binding curve predicted by the homogeneous lattice theory of multiple equilibria, which was minimized to the datapoints by the method of least squares in terms of the intrinsic association constant ($K_{in}=1X10^{-8}$) and the cooperativity constant ($\omega=1X10^{8}$). See text for detail.

cooperative effects will similarly be absent, must proceed with very low efficiency. There is experimental support for both these predictions. The first comes from studies by Yager and von Holde (279, 280) and Ausio *et al.* (10, 11) on the dissociation of DNA from core particles at different ionic strengths and at different concentrations of core particle. It was found that a significant proportion of DNA dissociated from the core even at very low ionic strengths (the final dialysis in the salt dialysis procedure is against 10mM TRIS.Cl, 1mM EDTA and 0,2mM PMSF), and, as expected from the law of mass-action, this phenomenon is dependent on the concentration of core particles. The concentration of the histone octamer during assembly at molar ratios of octamer:DNA of 3 is 120nM. The concentration of the DNA at each assembly ratio is 40nM. Herein may lie the explanation for the very low association constant. The measured inherent affinity of the octamer for the DNA may have been made at a level of dilution where mass-action favours dissociation. If the same series of measurements is repeated at a higher concentration of DNA (and therefore octamer), the deduced intrinsic association constant may be different. This is unfortunately not possible for either urea/salt or salt dialysis assembly, since, at a final DNA concentration much in excess of 40nM (50μg/ml for plasmid pHP2), aggregation occurs (data not shown;40).

Apart from the dissociation data of core particles at low ionic strength and high dilution, the magnitude of the association constant calculated above can also be tested by considering the results of salt dialysis assembly onto core length DNA. In the study by Lorch *et al* (138) the efficiency of reconstitution of core length DNA by salt dialysis was monitored by the electrophoresis of the reconstitution products on native agarose gels. It was found (138) that even at assembly ratios of octamer:DNA

(w/w) of 1, a significant amount of core DNA remained unassembled. An increase in the assembly ratio resulted in aggregation. This is unlike the result of the assembly onto large DNA molecules, as is shown by figure 4.16. Supercoiled and linearized pHP2 were reconstituted by salt dialysis at different molar octamer:DNA ratios and the reconstituted material resolved on native agarose gels (see sections 2.4.4 and 2.4.5 of materials and methods). Referring to the result presented in figure 4.16, well defined bands are seen that migrate at progressively slower rates at increasing assembly ratios. This is due to both an increase in the size of the molecular complex, and a reduction in the net charge of the assembled DNA molecule (43, 55, 205). It is further evident that at each assembly ratio a well-defined homogeneous assembly product is formed, as suggested by the topoisomer distributions following relaxation of reconstituted samples with topoisomerase I. Note that the shift of decreased magnitude observed for core formation on linearized DNA compared to that of supercoiled DNA is not directly due to the decreased assembly efficiency, but rather the lower electrophoretic mobility of free linearized DNA compared to free supercoiled DNA under the electrophoretic conditions. This is evident when both samples are run on the same gel (data not shown). Although it is not possible to discriminate free from assembled DNA at low ratios of octamer:DNA for the linearized sample, at intermediate ratios, where a clear shift is evident (lanes 4 - 7 of figure 4.16), no free DNA is visible. This is in contrast to the result of Lorch *et al* (138) where a significant proportion (approximately 50%) of the core DNA remains unassembled at corresponding intermediate molar assembly ratios (see lanes 3 and 4 of figure 2 of reference 138). This difference in itself suggests the presence of a degree of cooperativity on longer DNA molecules, and supports the suggestion that the intrinsic association

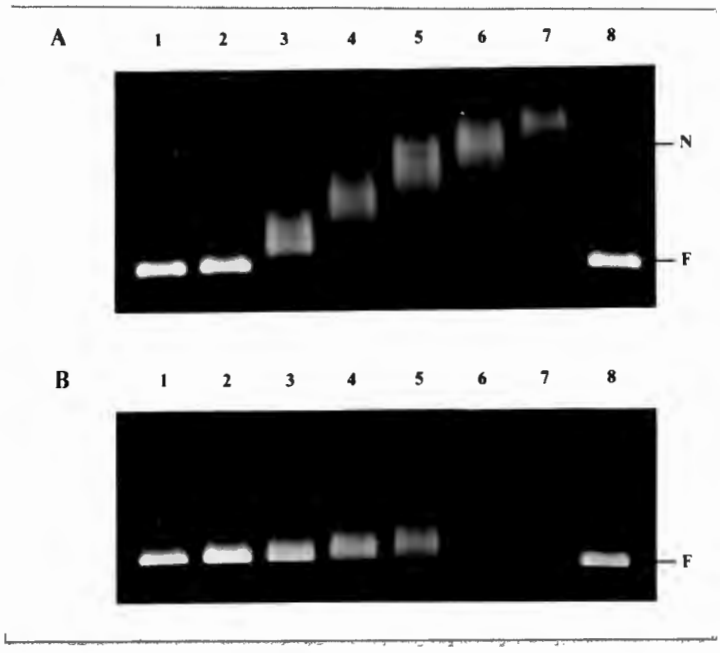

Figure 4.16: Native agarose gel electrophoresis of assembled pHP2. Supercoiled (A) and linearized (B) pHP2 were assembled at different molar ratios of octamer:DNA and electrophoresed in 0,7% (w/v) agarose gels in 45mM TRIS.borate at 10mA. The gel was stained with ethidium bromide. The assembly ratios of the corresponding lanes in A and B were identical. Salt dialysis assembly was performed at octamer:DNA (mole/mole) ratios of 0, 3, 5, 7, 9 and 11 (lanes 2 - 7). Lanes 1 and 8 correspond to free supercoiled (A) or linearized (B) plasmid DNA not exposed to the assembly conditions. The position of the free plasmid (F) is indicated in both A and B, and that of nicked, circular plasmid (N) in A.

constant of octamer-DNA interaction at dilute concentrations is small.

4.5 Efficiency of reconstitution by poly[L-glutamate] and high salt exchange:

Results from studies on the reconstitution of nucleosome cores onto supercoiled DNA by poly[L-glutamate] and high salt exchange, revealed fundamental differences between the assembly methods employing intermediate ionic strengths (<1M NaCl) and those employing high ionic strengths (2M NaCl) discussed above. The electrophoretic separation of topoisomers following topoisomerase I relaxation of supercoiled pHP2 assembled at low molar ratios of octamer:DNA by poly[L-glutamate] facilitated assembly (see section 2.4.1 of materials and methods) is presented in figure 4.17. The topoisomer distributions obtained (lanes 2 and 3 of figure 4.17) differ in an important way from those obtained at similar assembly ratios by urea/salt and salt dialysis assembly (see lanes 2 and 3 of figure 4.1 A and lanes 2 and 3 of figure 4.10 A). Unlike the homogeneous distribution previously obtained, the distribution now has a bimodial character, suggesting that a small fraction of the supercoiled plasmid is saturated with octamer at a low molar input ratio of octamer:DNA and the remaining fraction unassembled. This interpretation is supported by the result of the electrophoretic separation of poly[L-glutamate] assembly product on a native agarose gel. A fraction was found to migrate at the position of the fully assembled plasmid, and a fraction at a position close to that of free DNA (data not shown). Thus, the core distribution of poly[L-glutamate] facilitated assembly is, unlike the statistical analogy of diners taking their places in a dining hall, decidedly non-statistical. This feature was further investigated by ultracentrifugation of reconstitution products on sucrose gradients.

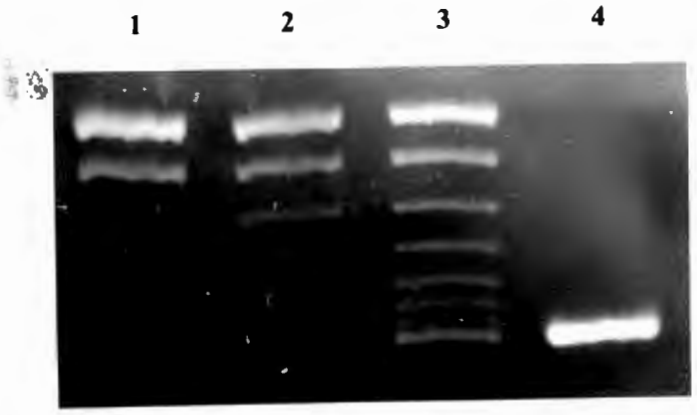

Figure 4.17: DNA topoisomer distribution of supercoiled pHP2 assembled at different octamer:DNA ratios by poly[L-glutamate] facilitated assembly. Supercoiled pHP2 was assembled at octamer:DNA (mole/mole) ratios of 0, 3 and 5 (lanes 1 - 3), relaxed with 8U of topoisomerase I, and electrophoresed in a 1,5% (w/v) agarose gel in the absence of chloroquine. Native supercoiled pHP2 (no topoisomerase I treatment) was included in lane 4. The direction of electrophoresis is from top to bottom.

Supercoiled pHP2 was assembled at a molar input ratio of octamer:DNA of 6 by urea/salt dialysis, salt dialysis, poly[L-glutamate] facilitated assembly, and by high salt exchange, and the reconstitution products analyzed by ultracentrifugation. The result is presented in figure 4.18. As expected from the topoisomer distribution results (see figure 4.1 and 4.10) the population of DNA molecules containing nucleosome cores is homogeneous in both urea/salt and salt dialysis assembly, and resolve as a single peak (trace B and C of figure 4.18). The assembly of the supercoiled DNA by poly[L-glutamate] facilitated assembly and by high salt exchange, in contrast, produce two peaks, one close to the position of the fully assembled plasmid, and the other close the position of the free plasmid (trace D and E of figure 4.18; the presence of the large peak on the very left-most side of trace E is due to the presence of a large molar excess of core particles). The position of the free and fully assembled plasmid are shown for comparative purposes (trace A and F).

At face value it would appear that the assembly efficiency of supercoiled DNA by high salt exchange or facilitated by poly[L-glutamate] represents an all-or-nothing mechanism characteristic of highly cooperative processes (158). It was found that this first stage of assembly, involving the exchange of the histone octamer from core particles to the supercoiled molecules in the high salt exchange method, is followed by a second phase of exchange of slower kinetics. This was demonstrated by allowing the initial exchange of the octamer at 700mM NaCl at 37°C for 20min or 90min before commencing the stepwise dilution to 70mM NaCl. The result is presented in figure 4.19. It is clear that at longer incubation times at 700mM NaCl, the octamer present on saturated supercoiled molecules is statistically proportioned between all supercoiled molecules. This

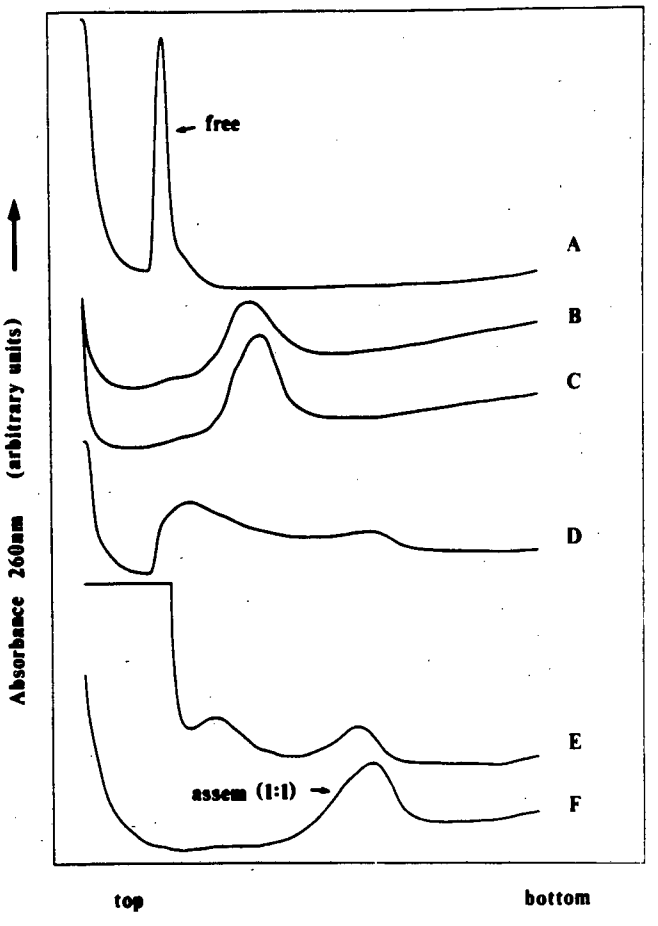

Figure 4.18: Non-statistical distribution of nucleosome cores on supercoiled DNA. The product of reconstitution following urea/salt dialysis (B), salt dialysis (C), poly[L-glutamate] facilitated (D) and high salt exchange (E) assembly of supercoiled pHP2 at an octamer:DNA weight ratio of 0,5:1 (octamer:DNA molar ratio of 6) was analyzed by ultracentrifugation on 5-20% (w/v) sucrose gradients. The absorbance at 260nm along each gradient was measured and is plotted to the same scale and aligned relative to the top of the gradient on the left-hand side of the figure.. Also shown is the position on the gradient of free supercoiled pHP2 (A) and supercoiled pHP2 assembled by urea/salt dialysis at an octamer:DNA (w/w) ratio of 1:1 (octamer:DNA molar ratio of 11) (F). All gradients were spun on the same rotor.

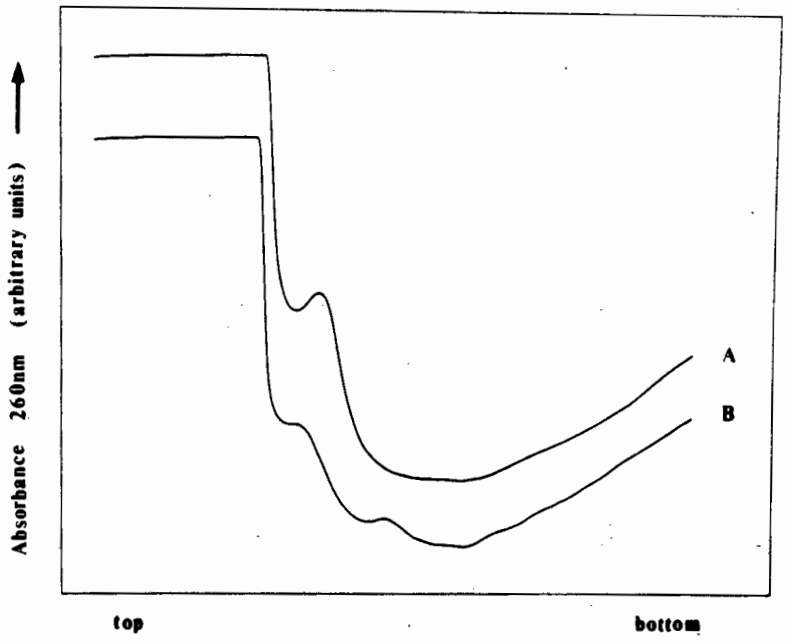

Figure 4.19: Statistical distribution of nucleosome cores on supercoiled DNA following prolonged incubation. The exchange of the histone octamer from core particles to supercoiled pHP2 was allowed at 700mM NaCl at 37°C for 90min (A) or 20min (B) before the stepwise dilution to 70mM NaCl (see section 2.4.2 of materials and methods). The product of reconstitution was analyzed by ultracentrifugation on 5-20% (w/v) sucrose gradients. The absorbance at 260nm was measured along each of the gradients and plotted to the same scale aligned relative to the top of the gradient on he left-hand side of the figure. The two gradients were spun on the same rotor.

mechanism is clearly not at work in poly[L-glutamate] facilitated assembly, where, even after a 180min incubation period at 37°C, two peaks corresponding closely to fully saturated and free plasmid remain (trace D of figure 4.18). To investigate whether this second exchange phase is a function of the presence of core length DNA or of ionic strength, supercoiled plasmid saturated with nucleosome cores was incubated under different experimental conditions with free supercoiled plasmid at an equal weight ratio (DNA:DNA). The result is presented in figure 4.20.

At 100mM NaCl, in the absence of both poly[L-glutamate] and core length DNA, no exchange is visible (trace B). The fully assembled and free plasmid still resolve at the corresponding positions on the gradient (compare with scan A and G). It is also clear that, irrespective of the concentration, neither core length DNA (trace C and D) nor poly[L-glutamate] (trace E and F) can facilitate this second phase exchange at 100mM NaCl. The experiment was therefore repeated at a higher ionic strength (700mM NaCl). The result is presented in figure 4.21.

It is clear that at the higher ionic strength, as expected from the results of other workers (55, 200), the octamer can efficiently migrate from the saturated supercoiled plasmid to the free supercoiled plasmid. This occurs in the absence of both poly[L-glutamate] and core length DNA (trace B of figure 4.21) where the peak corresponding to the fully assembled plasmid is replaced by a slower sedimenting peak corresponding to the supercoiled plasmid occupied by an intermediate number of cores. A small amount of free DNA is still visible. The result in the presence of 1,2μM poly[L-glutamate] is very similar (trace D of figure 4.21). In the presence of 1,2μM core length DNA, however, two peaks are still visible (trace C of figure 4.21). Comparison of the sedimentation position of these two peaks with that of

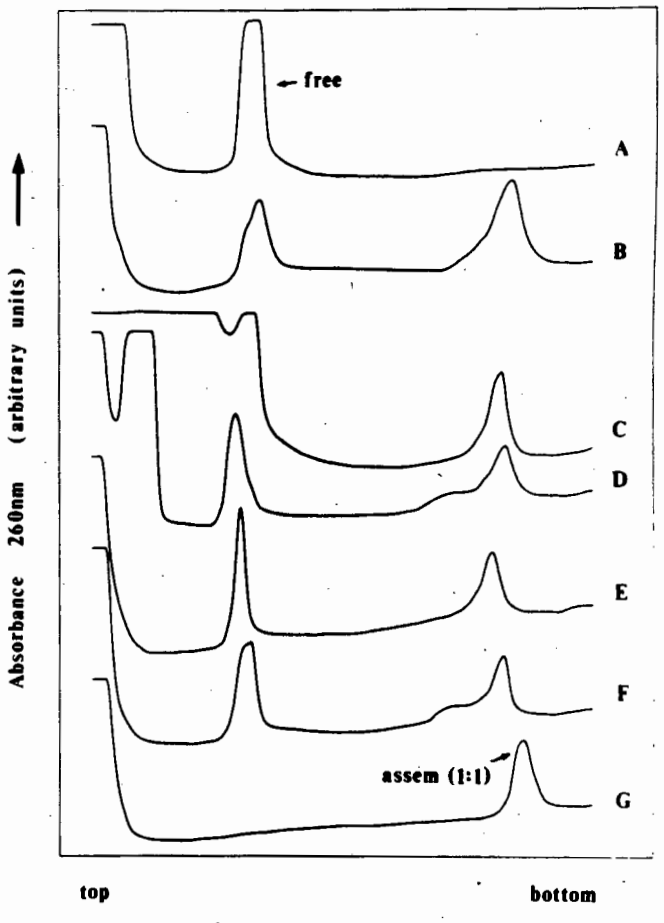

Figure 4.20: Migration of the histone octamer from assembled supercoiled pHP2 to free supercoiled pHP2 at 100mM NaCl. 10μg of supercoiled pHP2 assembled at an octamer:DNA (w/w) ratio of 1:1 (octamer:DNA molar ratio of 11) was incubated with 10μg of free supercoiled pHP2 for 90 min at 37°C in 10mM TRIS.Cl (pH8,0), 100mM NaCl, 1mM EDTA and 0,2mM PMSF (B), after adjustment to 24μM core DNA (C), 1,2μM core DNA (D), 24μM poly[L-glutamate] (E) or 1,2μM poly[L-glutamate] (F) and the distribution of the histone octamer analyzed by ultracentrifugation on 5-20% (w/v) sucrose gradients. The absorbance at 260nm was measured along each of the gradients and plotted to the same scale aligned relative to the top of the gradient on the left-hand side of the figure. The positions of free supercoiled pHP2 (A) and supercoiled pHP2 assembled at an octamer:DNA (w/w) ratio of 1:1 (octamer:DNA molar ratio of 11) (G) are indicated. Traces A - F are from gradients spun on the same rotor.

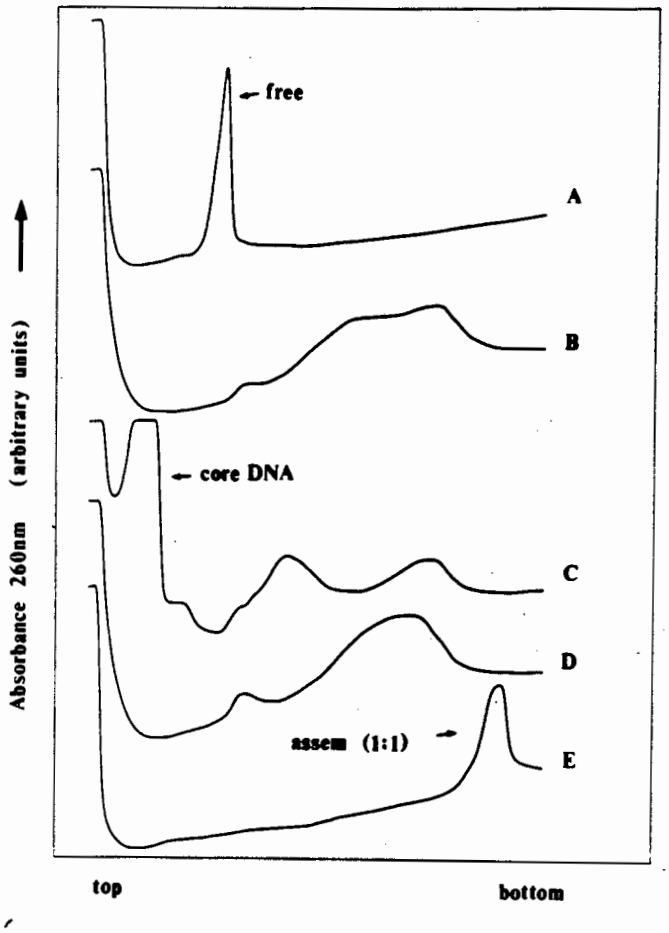

Figure 4.21: Migration of the histone octamer from assembled supercoiled pHP2 to free supercoiled pHP2 at 700mM NaCl. 10μg of supercoiled pHP2 assembled at an octamer:DNA (w/w) ratio of 1:1 (octamer:DNA molar ratio of 11) was incubated with 10μg of free supercoiled pHP2 for 90min at 37°C in 10mM TRIS.Cl (pH8,0), 700mM NaCl, 1mM EDTA and 0,2mM PMSF (B) after adjustment to 1,2μM core DNA (C) or 1,2μM poly[L-glutamate] (D) and the distribution of the histone octamer analyzed by ultracentrifugation on 5-20% (w/v) sucrose gradients. The absorbance at 260nm was measured along each of the gradients and plotted to the same scale aligned relative to the top of the gradient on the left-hand side of the figure. The positions of free supercoiled pHP2 (A) and supercoiled pHP2 assembled at an octamer:DNA (w/w) ratio of 1:1 (octamer:DNA molar ratio of 11) (E) are indicated. All gradients were spun on the same rotor.

free and saturated plasmid (traces A and E) indicate that they correspond to neither. The faster sedimenting peak appears less assembled than the fully saturated peak, and the slower sedimenting peak more assembled than the free DNA. The reason for this difference in the behavior of the exchange compared to that of poly[L-glutamate] and in the absence of both poly[L-glutamate] and core length DNA is seen when the peak corresponding to core length DNA (trace C in figure 4.21) is examined. On the right-hand side of this peak a shoulder is visible. This partly obscured peak corresponds to the core particle. This indicates that in the presence of free core length DNA, the octamer not only migrated from the donor supercoiled molecule to acceptor supercoiled molecule, but also to the smaller (linear) core length fragments. What is the relative efficiency of each of these different migratory paths? Since the peak corresponding to core length DNA exceeds the absorbance scale, an exact calculation is not possible. However, from the peak shape one may estimate that approximately 5-10% of the core DNA was incorporated into core particles. In contrast, essentially 100% of the free supercoiled DNA was assembled. If the distribution of the octamer among core length equivalent fragments proceeded with equal probability, it is expected that approximately 38% of the core DNA should have become assembled. This was clearly not the case and indicates that the assembly onto supercoiled molecules proceeds significantly more efficiently than onto core DNA, a result that provides supports to the finding of cooperativity above (see section 4.3 and 4.4). This result also emphasizes the need to take care in the selection of a competitor DNA in displacement experiments if the degree of assembly of the octamer "trap" is to be a true reflection of the displacement of the octamer from the donor molecule under the experimental conditions. Although insufficient experimental evidence is available to provide firm support to the suggestion, the observed tendency of high cooperative

binding with high salt exchange and poly[L-glutamate] facilitated assembly suggests that the degree of cooperativity of octamer:DNA association may increase at lower ionic strengths. It is interesting to note in this regard that in the study of Caffarelli *et al.* (36), the topoisomer distributions of supercoiled ColE1 DNA assembled by salt dialysis from a starting concentration of 1M NaCl (as opposed to 2M NaCl in this study), display a clear bimodial character similar to that of poly[L-glutamate] facilitated assembly reported above (see figure 3 of reference 36).

4.6 Discussion:

A clear difference was shown in the efficiency of assembly of nucleosome cores on supercoiled and linearized DNA (section 4.3 and 4.4). This was especially conspicuous at low molar assembly ratios of octamer:DNA. It is unlikely that this difference reflects a kinetic effect, since, in the case of salt dialysis assembly, an equilibrium condition is reached (100). It is therefore also unlikely that, due to kinetic differences, the octamer dissociates into constituent histones at high dilution in the less efficient (linearized) sample which is then not efficiently assembled. Moreover, the salt dialysis assembly has been shown to proceed by association of the H3-H4 tetramer, followed by two H2A-H2B dimers (100). The difference in assembly efficiency therefore most likely reflects the topology of the DNA molecule.

The association of the histone octamer with DNA can be described by several independent mechanistic steps, each with a characteristic free energy change associated. These include the association of the individual histones to form an octamer, the electrostatic interaction of the DNA-backbone phosphate groups with the lysine and arginine

side-chains of the octamer, and the deformation of the DNA double helix into 1,8 negative supercoils (40). If all these energy changes are incorporated into a single free energy change of core formation, ΔG_{core}, it would reflect the intrinsic association constant, K_i, of octamer-DNA interaction. K_i is independent of any cooperative interaction that may occur between adjacent cores and is equal to the association constant, K, of equation 4.6.

In a supercoiled molecule, however, there may be an additional contribution to ΔG_{core}. This contribution comes from the free energy of supercoiling (ΔG_{sup}), and is equal to:

$$-A(\Delta Lk)^2, \qquad (4.8)$$

where A is a constant depending on the size of the supercoiled molecule (50, 197). In a theoretical treatment of the possible contribution of supercoiling energy to core formation, Camerini-Otero and Felsenfeld (40) showed that the supercoiling energy was insufficient to balance the energy requirement of deforming the DNA into the required shape. Nevertheless, supercoiling may contribute to core formation, even if this energy is only a fractional contribution to ΔG_{core}. If it is assumed that all the energy of supercoiling is available for core formation, it follows that the association of one octamer with a supercoiled DNA molecule has an additional energy contribution (ΔG_{sup}) equal to $-A[(\Delta Lk)^2-(\Delta Lk+1)^2]$, where the subtracted term represents the supercoiling energy of the molecule with one less unconstrained negative superhelical turn. In the general case (assuming the induction of one positive supercoil per core formed) the amount of energy available for assembly of the a'th octamer is given by:

$$\Delta G_{sup}=-A[(\Delta Lk+a-1)^2-(\Delta Lk+a)^2], \qquad (4.9)$$

which, by resolving the quadratic parameters and adding like terms, reduces to:

$$\Delta G_{sup}=-A(-2\Delta Lk-2a+1). \qquad (4.10)$$

The contribution of this free energy to binding of an octamer to DNA augments the intrinsic association constant in a manner dependent on the number of cores on the DNA. Thus the apparent intrinsic association constant of an octamer binding to supercoiled DNA is given by:

$$K_{i,app}=K_i e^{A(-2\Delta Lk-2a+1)} \qquad (4.11)$$

In order to relate this result to experimental data, the variable a may be substituted with vN, where v is as previously defined, and N is the size of the plasmid in base-pairs. This gives the final expression of the apparent intrinsic association constant as:

$$K_{i,app}=K_i e^{A(-2\Delta Lk-2vN+1)} \qquad (4.12)$$

Thus, the apparent intrinsic association constant can be calculated at various degrees of reconstitution with the term vN equal to integer values from 0 to 11. The obtained value of $K_{i,app}$ can be substituted into equation 4.6 to obtain the ratio v/L at the appropriate fraction of octamer binding. As explained in section 4.3, the molar ratio of octamer:DNA corresponding to the relevant number of assembled octamers can be obtained from this relation. One may therefore calculate the binding curve of reconstitution onto supercoiled DNA from the binding data of linearized DNA by setting $K_i=1X10^{-7}$ in equation 4.12, calculating $K_{i,app}$ at the relevant degree of binding, and substituting K in equation 4.6 with this value of $K_{i,app}$ and with $\omega=8X10^6$. This was performed for integer values of vN from 0 to 11 and

the molar ratio of octamer:DNA corresponding to specific values of v calculated. The result is shown in figure 4.22.

It is clear that if all the free energy of supercoiling is available to core formation, the experimental binding curve relating the number of octamers assembled onto supercoiled DNA at various molar input ratios of octamer:DNA (solid line in figure 4.22) is accurately predicted (dashed line) from the experimental binding data of assembly onto linearized DNA (dotted line). This result indicates that the energy of supercoiling is sufficient to account for the experimentally obtained differences in the efficiencies of assembly onto supercoiled and linearized DNA. Another feature of the predicted binding curve (dashed line) that fits the experimental binding curve of supercoiled DNA (solid line), is the decrease in binding efficiency at high molar input ratios of octamer:DNA. This is due to a decrease in the contribution of supercoiling to octamer binding when most of the negative supercoils are restrained by bound cores. However, in this limit the application of equation 4.12 may be spurious. The constant A in equation 4.8 was derived from studies on free supercoiled molecules of different sizes (197, 50). It would seem unlikely that the energy constraints of accommodating one negative supercoil on a free circular plasmid would be the same as that of a circular plasmid nearly saturated with cores, where the flexibility characteristics are expected to be quite different. It is more likely that, as suggested, the decrease in binding efficiency experimentally found reflects the introduction of positive supercoils into the plasmid, although, in the limit, $K_{i,app}$ will approach K_i, and the binding efficiency to supercoiled plasmids will be similar to linearized plasmid.

Thus, with the efficiency of core formation on supercoiled and linearized plasmids characterized in detail, one may now

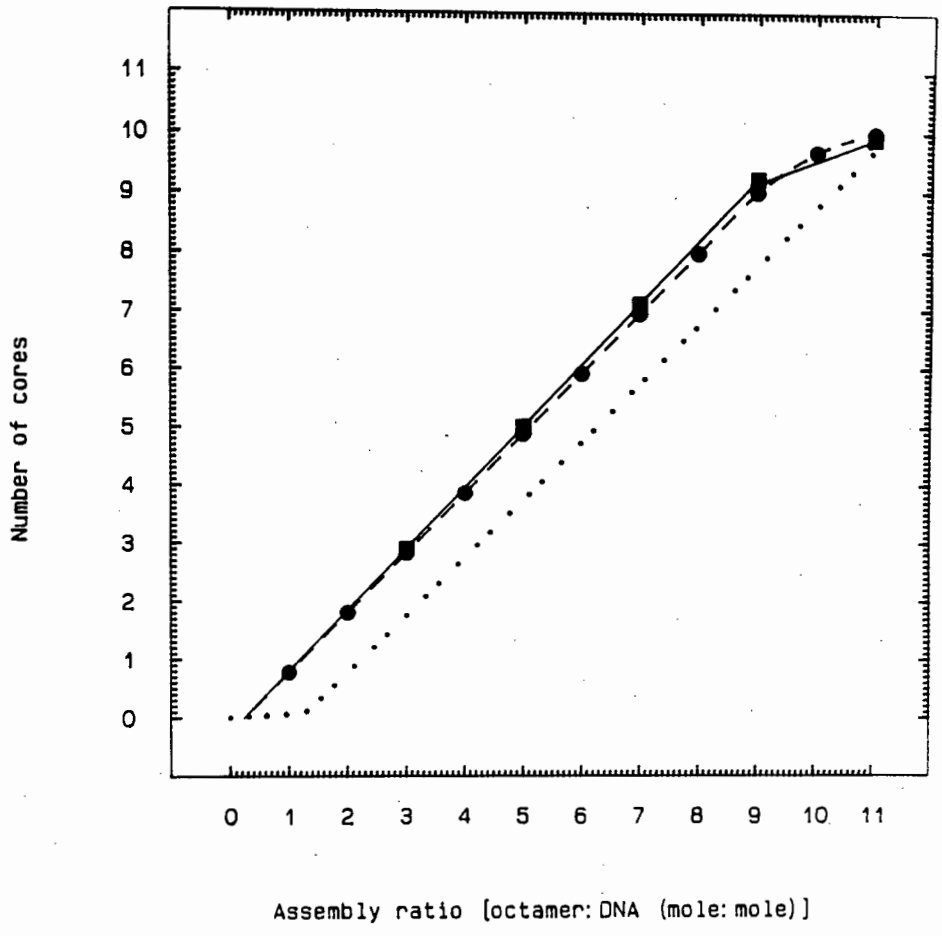

Figure 4.22. Contribution of negative supercoiling to the formation of nucleosome cores. The dotted line (··) represents the binding curve predicted by the homogeneous lattice theory of multiple equilibria (158) fitted to the experimental datapoints of octamer binding by linearized plasmid (see section 4.3 and figure 4.9). The solid line (——) connects the datapoints (squares) representing the number of nucleosome cores formed at different molar assembly ratios of octamer:DNA on supercoiled plasmid (see section 4.2 and figure 4.3; ΔLk of figure 4.3 was converted to the number of cores by dividing by 1,1 (81, 283)). The dashed line (--) connects the predicted datapoints (circles) if all the energy of supercoiling was available to core formation on the negatively supercoiled plasmid. See text for detail.

proceed to investigate the effect of negative superhelicity and the circular constraint of supercoiled molecules on the positioning of the nucleosome core. This is addressed in the next chapter.

CHAPTER 5

5 The effect of negative supercoiling of DNA molecules on the positioning of nucleosome cores *in vitro*.

5.1 Introduction:

The chromatin of an eukaryotic chromosome is demarcated into topologically sequestered domains (16) by attachment to the chromosome scaffold during metaphase (161) and internal matrix during interphase (76, 110; see section 1.3). Although most of the negative superhelical stress in these topologically isolated domains are restrained by bound nucleosomes (83, 223), it was shown that localized areas of torsionally stressed DNA are generated by the processes of transcription (82, 136), and may also be generated by the condensation and decondensation of chromatin during the cell-cycle, as suggested by the association of topoisomerase II with the bases of the chromatin loops (75). Although precise measurements of the magnitude of this superhelical stress in eukaryotes have not yet been performed, it appears that individual nucleosomes and nucleosomal arrays may transiently be exposed to sizeable superhelical tensions. In prokaryotes, for instance, the passage of the transcribing polymerase *in vivo* was shown to generate an area upstream of the transcription complex where the specific linking difference exceeded -0,038 (198).

Nucleosome positioning has been studied in detail on linearized DNA molecules. The positioning of cores may, similar to their formation reported in the previous chapter, also be influenced by negative supercoiling. This effect may be a combination of the general shape of the supercoiled molecule as determined by the apportionment of writhe in the naked supercoiled molecule in solution, and steric effects

of accommodating nucleosome cores on a molecule where the geometric shape is dictated by the topology of the DNA.

The H1-H4 histone gene spacer of *Psammechinus miliaris* was previously shown to be occupied by at least one positioned nucleosome core *in vitro* (203). The effect of the circular constraint and superhelical stress of supercoiled molecules on the positioning of this core was investigated in the 1915bp plasmid pHP2 which contains the positioning sequence from the H1-H4 gene spacer as a 199bp *Afl* III-*Hind* III insert (see section 2.1 of materials and methods). The partial nucleotide sequence of pHP2 is given in figure 1.1. The positioning of the core on the insert was investigated on both linearized and supercoiled (σ=-0,06) pHP2. The linearized pHP2 was obtained by digestion of supercoiled pHP2 with *Pvu* II prior to assembly. This restriction enzyme introduces a blunt-ended double-strand break between nucleotides 672bp and 673bp, a site well removed from the core of interest (203), thus minimizing possible contributions of end-effects (135) to the positioning of the core on linearized molecules.

The position of the octamer in the region of 100bp-250bp (see figure 2.1) was investigated by DNase I digestion (see Klug and Lutter (119), Drew and Calladine (55) and Drew and Travers (57) for a discussion of DNase I as a probe for nucleosome core placement). Since, in some of the experiments described below, the variable parameter of interest is the supercoiling of the DNA, and DNase I, by virtue of introducing a single-stranded nick, removes this constraint, it is essential that individual DNA molecules are nicked only once on average. The apportionment of single-stranded nicks produced by mild DNase I digestion of a large population of molecules is expected to follow a Poisson distribution. In figure 5.1 the fraction of molecules containing a single nick, predicted by such a

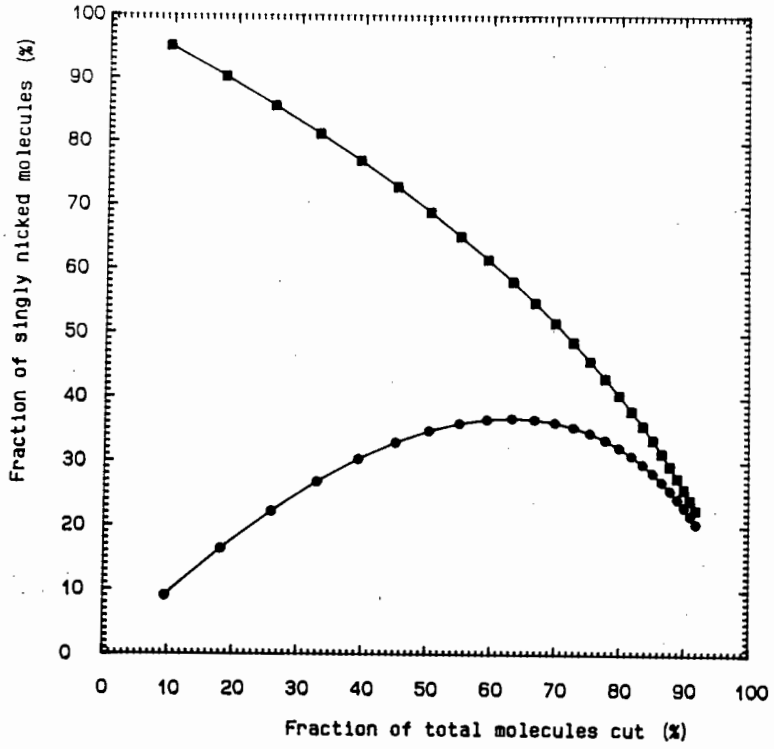

Figure 5.1. Distribution of nicks in a population of DNA molecules. The fraction of singly nicked DNA molecules, expressed as a ratio of the total number of DNA molecules (circles) or as a ratio of the total number of nicked molecules (squares) at different extents of cleavage, as predicted by a Poisson distribution.

Poisson distribution, is shown as a function of the total fraction of DNA molecules cleaved by DNase I. It is clear from the figure that if approximately 10% of the molecules are cleaved by DNase I, essentially 100% of these nicks appear as single nicks on DNA molecules. As the total fraction of DNA molecules cleaved by DNase I increases, the fraction of singly nicked molecules similarly increases to reach a maximum when approximately 60% of the input DNA has been digested by DNase I. However, by referring to the curve relating the ratio of the fraction of singly nicked molecules to the total fraction of nicked DNA molecules, it is evident that the percentage of singly nicked molecules decreases with increasing digestion ratios. For instance, when 40% of the DNA molecules have been nicked by DNase I, approximately 30% of the total number of DNA molecules contain single nicks and 9% two or more nicks. When 50% of the total number of DNA molecules have been nicked, 35% of the total number of molecules contain single nicks and approximately 16% two or more nicks. The fraction of nicked molecules that contain only a single nick therefore decreases from [30/(30+9)]100=77% to [35/(35+16)]100=69% when proceeding from 40% to 50% of the total number of DNA molecules cleaved. Thus, even though the absolute number of molecules containing single nicks increases up to where approximately 60% of all the DNA molecules have been cleaved by DNase I, a significant proportion of molecules containing two or more nicks is also produced. This is clearly undesirable, since, if the position of the nucleosome core is influenced by supercoiling, this effect may be removed after introduction of the first nick. The second nick in the core DNA would then reflect the position of the core in the absence of supercoiling, and the eventual DNase I footprint would represent a superimposition of core positions in the presence and absence of supercoiling. Unfortunately, the extent of DNase I cleavage cannot be chosen to cause single nicks only, since the total amount of

digested material would be insufficient to obtain a clear and accurate indication of the core position. A compromise is therefore required. In the experiments described below, the fraction of the total number of DNA molecules cleaved by DNase I was approximately 40% and never more than 50%. This resulted in approximately 70%-80% of the total number of nicked molecules containing only single nicks, and gave a sufficient signal-to-noise ratio to obtain a clear answer on the effect on supercoiling on nucleosome core positioning. The DNase I digestion conditions required to give this degree of cleavage was determined for each single experiment by a trial digestion as explained in section 2.6.2.

5.2 Nucleosome core positioning in the absence and presence of negative superhelical stress:

5.2.1 Nucleosome core positioning on circular molecules in the absence of superhelical stress:

The plasmid pHP2, in a linearized and supercoiled form (σ=-0,06), was reconstituted with the histone octamer at a molar octamer:DNA ratio of 11 by the method of salt dialysis (see section 2.4.4 of materials and methods). It was previously shown (see section 4.2) that reconstitution at this molar ratio resulted in an average of 11 nucleosome cores forming per supercoiled DNA molecule. Since the DNA sample used in this reconstitution procedure contained a mixture of topoisomers with a Gaussian distribution of linking differences (average ΔLk$\pm$standard deviation= -10,9$\pm$0,1), it is expected that, with the introduction of approximately +1,1 superhelical turns per core formed, very little torsional stress will be present on average in the whole population following reconstitution. This was demonstrated by the similar linking difference of the reconstituted molecule after topoisomerase I treatment and naked supercoiled pHP2 (see section 4.2). The average

structure of the reconstituted supercoiled molecule is therefore expected to be very similar to a linearized molecule that is circularized after reconstitution with 11 nucleosome cores. In this regard, an investigation of the positioning of nucleosome cores on the supercoiled molecule is expected to yield information on the effect of the circular constraint of the molecule rather than of superhelical stress.

After assembly, an aliquot of the reconstituted material was analysed on native agarose gels to ensure that aggregation had not taken place (result not shown), followed by digestion of the remainder with DNase I (see section 2.6.2 of materials and methods). The frequency of cleavage by DNase I in reconstituted and free DNA samples was determined by primer extension (26, 88) of 20-mer single-stranded oligonucleotide primers labeled at the 5'-end with [γ-^{32}P]ATP (see section 2.7 of materials and methods). Two primers (see section 2.2 of materials and methods) were employed to obtain information from both strands. Primer 1 is complementary to region 341bp-322bp of the Watson strand and primer 2 to region 18bp-37bp of the Crick strand (see figure 2.1 of section 2.1). Thus, the extension of primer 1 by the Klenow fragment of DNA polymerase proceeds in the direction of the numbering origin of the plasmid, and will yield information on the cleavage by DNase I of the Watson strand, whereas primer 2 is extended in the opposite direction and will yield information on the DNase I cleavage of the Crick strand. After extension, the DNA was purified, electrophoresed on a 6% acrylamide gel containing 7M urea, and the gel autoradiographed as explained in section 2.7 of materials and methods. The result is presented in figure 5.2.

If attention is first turned to the lanes representing the DNase I digestion of free DNA in figure 5.2 A (lanes 4, 6,

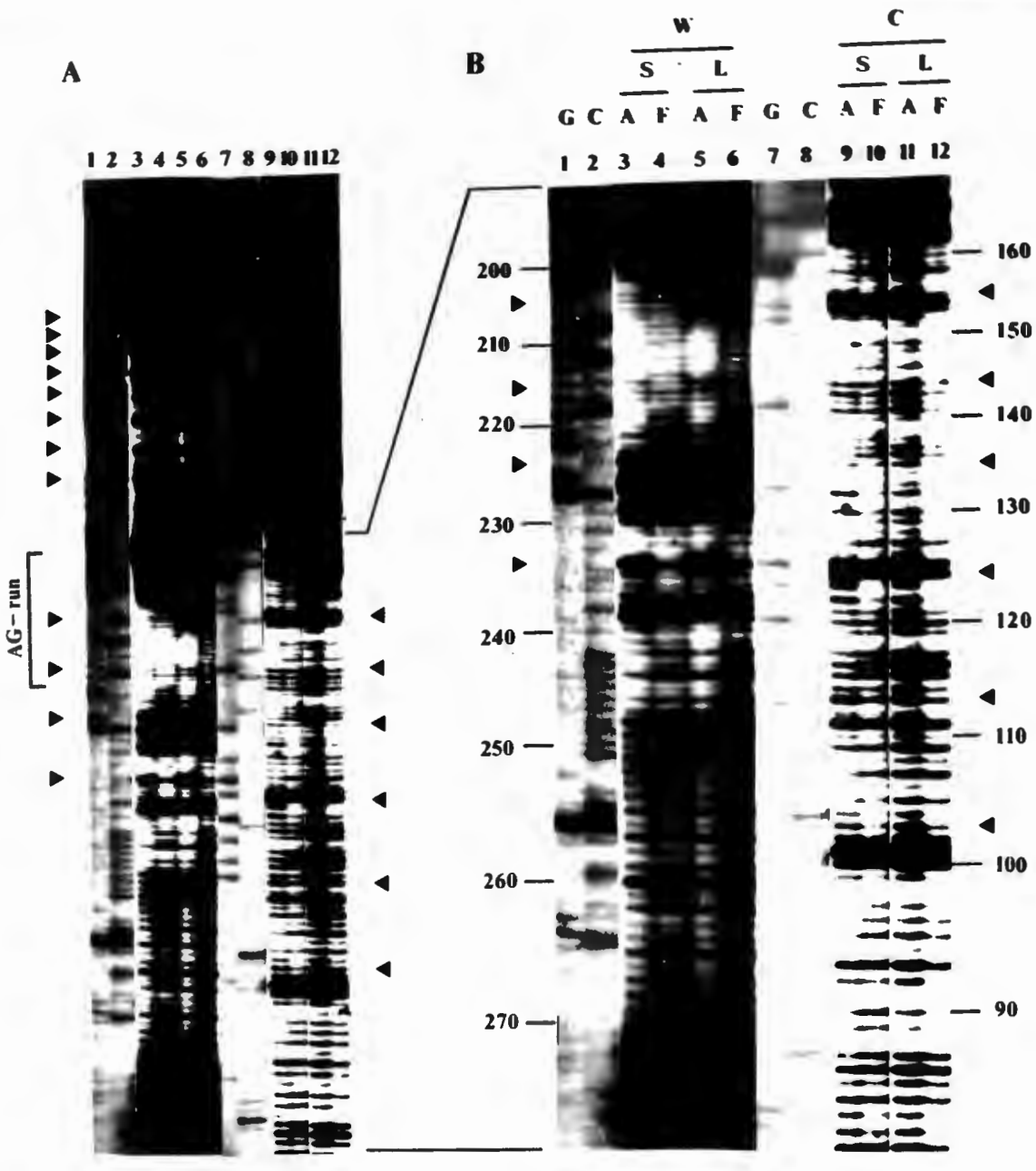

Figure 5.2. DNase I digestion of pHP2 reconstituted with the histone octamer at a molar octamer:DNA ratio of 11. The location of nucleosome cores on supercoiled (S; lanes 3, 4, 9 and 10) and linearized (L; lanes 5, 6, 11 and 12) pHP2 were determined by DNase I digestion of assembled (A; lanes 3, 5, 9 and 11) and free (F; lanes 4, 6, 10 and 12) plasmid followed by extension of 20mer primers annealed to either the Watson (W; lanes 3-6) or Crick (C; lanes 9-12) strand. The location of DNase I cleavage sites was identified relative to guanosine (G) or cytosine (C) chain terminating sequencing reactions of 20mer primers annealed to the Watson (lanes 1 and 2) or Crick (lanes 7 and 8) strands. The sequence positions of individual bands are shown in the margins of (B). The maxima in the modulation of DNase I cleavage with a period of approximately 10bp in reconstituted samples are indicated by triangles. Panel B is an enlarged section of panel A. Pairs of samples representing reconstituted and free plasmid were run on the same gel. An autoradiograph of the gel is shown. See text for detail.

10 and 12), a distribution of bands of different intensities is seen. These variances in intensity reflect differences in the frequency of nicking along the DNA molecule, and is due to the sequence specificity of the DNase I enzyme (57, 137). Considering the lanes representing the DNase I digestion of the DNA complexed with the histone octamer next (lanes 3, 5, 9 and 11 of figure 5.2 A), an additional periodic modulation in the cleavage frequency of DNase I is visible. The positions where this modulation reaches a maximum compared to the free DNA are indicated by solid triangles in the margins of figure 5.2 A. If the number of nucleotides between successive maxima are counted, it is clear that the modulation follows a period of approximately 10bp. This is indicative of a rotationally positioned nucleosome core (55, 56, 59, 193, 194, 200, 220) and results from the higher frequency of DNase I cleavage at positions where the DNA backbone points away from the octamer surface and the enzyme molecule is sterically less obstructed (119, 142).

Another feature in the DNase I cleavage pattern in figure 5.2 A warrants comment: this is the unusually high apparent rate of DNase I cleavage in a section of the d(A-G).d(C-T) motif (indicated by brackets in the margin of figure 5.2 A). This is in fact due to the low efficiency of extension by the Klenow fragment through the poly(purine).poly(pyrimidine) stretch, as demonstrated by the presence of fragments of similar length in the primer extension of undigested pHP2 (data not shown) and the absence of the corresponding "hypersensitive site" in DNase I digested directly end-labeled fragment (203). In order to alleviate this decreased efficiency of Klenow extension, several approaches were attempted, including the substitution of dGTP with 7'-deaza-dGTP, performing the extension reaction in the presence of 10% (v/v) DMSO, the substitution of the Klenow fragment of DNA polymerase with

the more processive "Sequenase" (United States Biochemical Corporation), and performing the extension reaction at elevated temperatures. It was found that the only modification that resulted in a more efficient extension over the d(A-G).d(C-T) motif was the pre-incubation of the DNase I digested samples (to which the primers had been annealed) at 52°C for at least 5 minutes prior to addition of Klenow, and proceeding with the extension reaction at 52°C. The amount of Klenow fragment used made little difference to the result. The gel shown in figure 5.2 and all subsequent DNase I digestion gels are of extension reactions performed in this fashion, as explained in section 2.7.

If attention is next turned to figure 5.2 B, which shows an enlarged section of the gel shown in A, the 10bp periodicity discussed above is clearly visible. In lane 5, which represents the Watson strand of linearized DNA reconstituted with the histone octamer, sites of maximal DNase I cleavage are visible at positions 205, 215, 224 and 234 (indicated by solid triangles). In the equivalent unreconstituted sample (lane 6), the corresponding positions do not represent cleavage maxima. This is most readily seen by studying the regions intermediate to the cleavage maxima in the reconstituted sample (lane 5). Between the maxima at positions 205, 215 and 224, a reduction in the rate of DNase I cleavage is visible, the centers of which occur at positions 210 and 219. At the equivalent positions in the free DNA (lane 6), significant cleavage by DNase I is still visible. Indeed, at position 219, the intensity of the band in the free sample (lane 6) is similar to that of the bands at the positions of maximal cleavage (205 and 215) in the reconstituted sample (lane 5).

If a careful comparison of the reconstituted (lane 5) and free DNA sample (lane 6) in the lower part of the gel below

position 234 is made, only subtle differences are discernable. The area between positions 239 and 248, for instance, appears less digested in the reconstituted than in the free sample (compare lanes 5 and 6). The bands at positions 242 and 244 appear more intense in the free sample than the equivalent bands in the reconstituted sample. If the 10bp modulation mentioned above had continued into the lower part of the gel, position 244 should have corresponded to a maximum in the reconstituted sample. The absence of this maximum suggests that position 234 represents the terminal helical turn of a nucleosome core that is positioned with its one border at position 234, and its other border towards the top of the gel, thus incorporating the 32bp d(A-G).d(C-T) motif in the core. This second border can more easily be investigated on the Crick strand (lanes 7-12 of figure 5.2 B) where the primer extension products of the DNase I digested DNA molecule are sufficiently resolved at single nucleotide level.

It is clear that significant differences exist between the rates of DNase I digestion on the Crick strand of linearized pHP2 reconstituted with the histone octamer (lane 11) and free in solution (lane 12). In the reconstituted sample (lane 11), maxima are visible at positions 82, 92, 103, 113, 124, 135, 144 and 155 (indicated by solid triangles). The intensity of the bands at the corresponding positions in the free DNA is much reduced and constitutes part of the DNase I digestion pattern resulting from the sequence specificity of the enzyme (57, 137).

The total extent of the modulation in DNase I cleavage identified above (234-82=152bp), exceeds the 146bp of the 1,8 superhelical turn core-particle. In contrast to the clear interruption in the modulation in DNase I cleavage at position 234, no similar discontinuity is evident on the Crick strand towards the lower extreme of the gel. This may

be due to two closely adjoining cores linked by a length of DNA close to an integral helical turn, thus causing little interruption in the periodicity of DNase I cleavage proceeding from the one core to the next, or, the association of more than 1,8 superhelical turns with the histone octamer, thereby causing a discontinuity at a position not visible in figure 5.2. This point is deferred for later discussion.

If attention is next turned to the supercoiled molecule (lanes 3, 4, 9 and 10 of figure 5.2 B), it is clear that reconstitution with the histone octamer produced maxima in the rates of DNase I cleavage similar to that of the linearized molecule. For instance, on the Watson strand (lane 3) maxima appear at positions 205, 215, 224 and 234. On the Crick strand (lane 9) maxima are visible at positions 82, 92, 103, 113, 124, 135, 144 and 155. The modulation in the DNase I cleavage displays a similar interruption at position 234 on the Watson strand, whereas no pause in the periodicity is visible on the Crick strand. A visual inspection of the gel therefore indicates little difference in both the apparent translational and rotational positioning of the nucleosome core present over the d(A-G).d(C-T) motif in the linearized and circular molecule. However, since it is difficult to obtain a reliable indication from the gel alone, it is necessary to perform a more objective and precise analysis of the digestion data in the form of a difference probability analysis.

5.2.2 Translational placement of the nucleosome core on the linearized molecule:

The probability analysis was performed by calculating the rates of cleavage by DNase I at individual phosphodiester bonds, and subtracting the natural logarithm of the rate of cleavage at a particular bond in the unreconstituted sample

from the natural logarithm of the rate of cleavage of the corresponding bond in the reconstituted sample, thus obtaining the difference in the probability of cleavage by DNase I at a particular bond in the reconstituted and free DNA molecule (see section 2.8.3 of materials and methods). This was performed for all the positions that were represented by clearly resolved fragments in the gel shown in figure 5.2. The difference probability plot calculated for the linearized sample is presented in figure 5.3, where the region between positions 172 and 197 is not represented. This region contains the d(A-G).d(C-T) motif, where, due to aborted primer extension products, no clear measurement on the rate of DNase I digestion was possible.

Looking first at the modulation in cleavage on the Crick strand in the linearized sample (figure 5.3), maxima are visible at positions 81, 93, 104, 113, 122, 133, 143, 153 and 164. On the Watson strand the positions of maximum cleavage are 206, 216, 226 and 235. These maxima simply reflect an increased probability of cleavage by DNase I at these positions compared to the equivalent positions in the unreconstituted DNA molecule, and represent positions on the backbone of the Watson and Crick strands offset by approximately 2bp in the 3′ direction relative to the position where the minor groove faces away from the surface of the histone octamer (55, 59, 205). This offset in the cleavage on the two strands relative to the point of maximum exposure of the minor groove is brought about by the way in which the DNase I molecule binds across the minor groove (57, 231) and hydrolyzes either of the most exposed phosphodiester bonds that are the nearest neighbors in space across the minor groove (59). Because of the structure of the DNA duplex (reviewed in 61), these phosphodiester bonds are not part of the same nucleotide base-pair, but staggered in the 3′ direction by approximately 2bp (55).

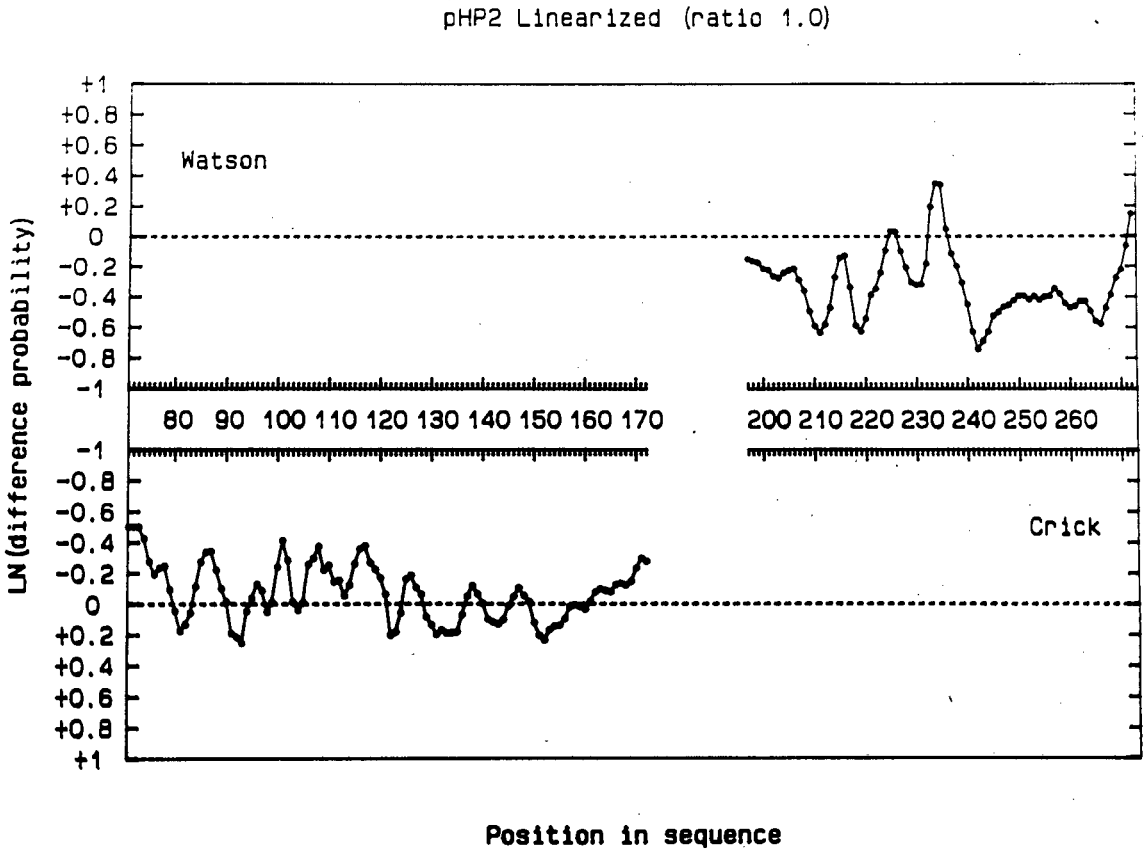

Figure 5.3. Probability difference of DNase I cleavage in linearized pHP2 reconstituted with the histone octamer at a molar octamer:DNA ratio of 11. The probability of cleavage, expressed as the natural logarithm of the difference between the rate of cleavage in reconstituted and free plasmid (see section 2.8.3 of materials and methods) is shown as a three point running average at different positions in the sequence of pHP2. Cleavage data on the Watson and Crick strands are shown in the top and bottom sections of the figure respectively. See text for detail.

The detailed picture on the differential rates of cleavage by DNase I in the reconstituted DNA molecule now allows an analysis of the translational and rotational frames of the nucleosome core(s) present between position 71 and 273 of pHP2.

Referring first to the Watson strand of the linearized pHP2, it is clear that the periodicity in DNase I cleavage displays an abrupt pause at position 235. If the modulation in cleavage visible between positions 197 and 235 had continued with the same period, maxima in the probability of DNase I cleavage were expected at positions 244/245 and 254/255. Instead, very weak maxima with a period of approximately 5bp are visible at positions 246, 251, 257 and 262. This is clearly not a continuation of the periodicity of approximately 10bp of the core DNA, and may represent a second core that positions in a mixture of two rotational frames differing by half a helical turn. Thus, as suggested by the visual inspection of the DNase I gel above, position 235 represents the right-hand side border, relative to the numbering origin of the plasmid, of a nucleosome core.

To be absolutely sure of this assignment, one may investigate the probability of cleavages at each of the maxima identified on the Watson strand, and compare these to the maxima found by other investigators. Lutter (142) demonstrated that the rates of cleavage by DNase I at positions maximally exposed to DNase I were not equivalent through the length of the core, but displayed a defined series of values suggested to represent differences in the molecular environment sensed by DNase I as it approached the core DNA at different positions in the core (119, 142). These rates of DNase I cleavage, plotted as a natural logarithm of the values reported by Lutter (142), are represented by the unfilled rectangles in figure 5.4. By aligning a similar plot of the probabilities of DNase I

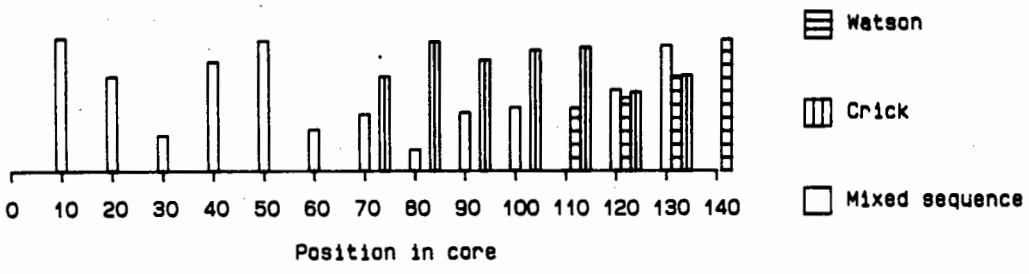

Figure 5.4. Probabilities of cleavage at identified maxima. The natural logarithms of the difference in the rates of cleavage by DNase I in reconstituted and free linearized pHP2 at maxima identified in the text on the Watson (horizontally hatched rectangles) and Crick (vertically hatched rectangles) strands are compared with the natural logarithm of the rates reported by Lutter (142) at individual sites in the core (unfilled rectangles). See text for detail.

cleavage at the maxima identified above with those of the mixed sequence cores, an indication of the border of a core may be obtained by finding the best correlation between the two sets of data (55, 56, 200). This was performed for the maxima on the Watson strand, and the probabilities of cleavage are indicated by the horizontally hatched rectangles shown in figure 5.4. It is clear that the alignment shown in figure 5.4 represents a good fit insofar as a decrease in the probability of DNase I cleavage proceeding from site 130-110 is concerned (the rate of cleavage at position 140 is not available from the Lutter study (142) where 140bp cores were employed), corroborating the suggested border of a core at position 235.

One may now proceed in a similar fashion to identify the other border on the Crick strand. In a nucleosome core containing 146bp (126, 254), this border is expected to be in the vicinity of position 235-146=89. A maximum in the probability of cleavage by DNase I is present at position 93 on the Crick strand (see figure 5.3). However, this probably does not represent the left-hand side border of the core, judged from the continuity of the cleavage periodicity in this region. If position 93 represented the position of cleavage by DNase I on the terminal helical turn, successive maxima were expected at approximately positions 103, 113 and so forth. However, if the modulation in the probability of cleavage by DNase I is closely studied in figure 5.3, apart from the maxima at the predicted positions, a clear maximum is also present at position 98. This feature is not characteristic of the modulation in the cutting frequency in the core DNA of a rotationally positioned octamer, but more indicative of the cutting of linker DNA not protected by histones, as seen in the difference probability plots of long DNA molecules occupied by several cores (55, 56). Thus, the periodicity of cleavage in the DNA molecule suggests that position 104 represents the left-hand terminal

helical turn of the core that has a right-hand border at position 235. Is this assignment borne out by the probability of cleavage at identified sites compared to the rates reported by Lutter (142)? The probabilities of cleavage at sites 104, 113, 122, 133, 143, 153 and 164 were aligned with the rates of cleavage at sites 130-70 in mixed sequence core DNA, and are shown as vertically hatched rectangles in figure 5.4. The fit is rather poor. Neither of the very low rates of cleavage at sites 80 or 110 is exhibited by the core on pHP2. However, similar poor fits were also obtained by Drew and Calladine (55) for cores reconstituted onto a 860bp molecule, and were suggested to result from a mixture of 10bp spaced translational positions assumed by the cores. This cannot be the explanation for the absence in correlation between the probabilities of cleavage on the Crick strand in region 104-164 and mixed sequence cores, since a clear border to the translational placement of the core is suggested by the pause in the periodicity of DNase I cleavage at position 235 (see figure 5.3). The possibility that this represents an "ultimate" right-hand side border to the core which assumes a mixture of translational placements to the left-hand side of this border, is argued against by the good fit that is obtained between the probability of cutting on the 3′ side of the Watson strand and the mixed sequence cores (see figure 5.4). The reason for the poor reproduction of the cutting rates on the 3′ side of the Crick strand must therefore lie elsewhere. One possibility is the sequence of the core DNA in pHP2. Lutter (142) employed a core preparation from rat liver that contained a wide range of DNA sequences, in contrast to the core on pHP2 where a single defined sequence was occupied. DNase I was shown to display a sequence specificity (137) suggested to reflect the sensitivity of the enzyme to the width of the minor groove (57). Both the width of the minor groove and the variability of this width was further shown to be sequence

dependent (57, 71), and for this reason the probability of cleavage by DNase I at a particular phosphodiester bond in pHP2 on and off the nucleosome core may be different than the average rate of cleavage at the same site in mixed sequence cores. It was furthermore also shown that the rates of cleavage change at the individual sites when the DNA is reconstituted as a fragment of approximately core length, compared to when it constitutes part of a longer molecule occupied by several cores, a phenomenon suggested to reflect the folding of the larger molecule where adjacent cores impose a steric constraint on DNase I cleavage (55, 56). Although the spatial arrangement of the reconstituted molecule is not known, the proximity of the rotationally positioned adjacent core of which the two terminal cleavage sites are visible at positions 81 and 93 renders this suggestion likely.

The discontinuity in the periodicity of DNase I cleavage, and the frequency of attack at individual sites on the 3' side of the Watson strand therefore indicate that a nucleosome core is positioned between positions 104 and 235 on linearized pHP2 (further evidence supporting the assignment of position 104 as the left-hand side border of the core is given in chapter 6). This core is denoted "core 2". The extent of this core (235-104=131bp) is slightly less than that of the canonical nucleosome core, but values other than 146bp are often found in DNase I digestion studies of nucleosome cores reconstituted *in vitro* (55, 56, 193, 220). It therefore appears that the DNA leaves the surface of the octamer at site 130 (position 104 on the Crick strand) and proceeds, through an 11bp linker, to the surface of the adjacent core where position 93 on the Crick strand represents the right-hand side border of the adjacent core, which is denoted "core 1". The core present in a mixture of rotational placements on the other side of core 2 (beyond position 235 in pHP2) is denoted "core 3".

5.2.3 Rotational placement of core 2 on the linearized molecule:

The assignment of the translational borders of core 2 and the identification of the individual sites of maximum DNase I cleavage now allows a determination of the position of the dyad axis for core 2. For instance, position 235 on the Watson strand (see figure 5.3) was found to correspond to site 140 on the core, and position 104 on the Crick strand, to site 130. It is thus expected that the core dyad will be removed by approximately 7 helical turns from position 235 (Watson) and approximately 6 helical turns from position 104 (Crick). Referring to figure 5.3, it is clear that 6 helical turns on the Crick strand coincide with the maximum at position 164. Cutting data at the related position on the Watson strand is unavailable. Thus, taking the 1bp stagger displayed by DNase I in the 3′ direction relative to the most exposed point of the minor groove into consideration (55, 56, 119, 142), the core dyad can be assigned to position 165. For a justification of this approach to assign the dyad, see Rhodes (205).

5.2.4 Translational and rotational placement of core 2 on a circular molecule in the absence of superhelical stress:

A plot relating the differential probability of cleavage by DNase I in reconstituted supercoiled pHP2 *versus* the sequence position was calculated from the gel shown in figure 5.2 as explained above (see also section 2.8.3 of materials and methods) and is presented in figure 5.5.

Sites of maximum cleavage on the Watson strand are visible at positions 206, 216, 226 and 235 (see figure 5.5). The cleavage site at position 235, similar to that on the linearized molecule, represents the terminal site of the preceding periodicity in the cleavage pattern. Beyond this

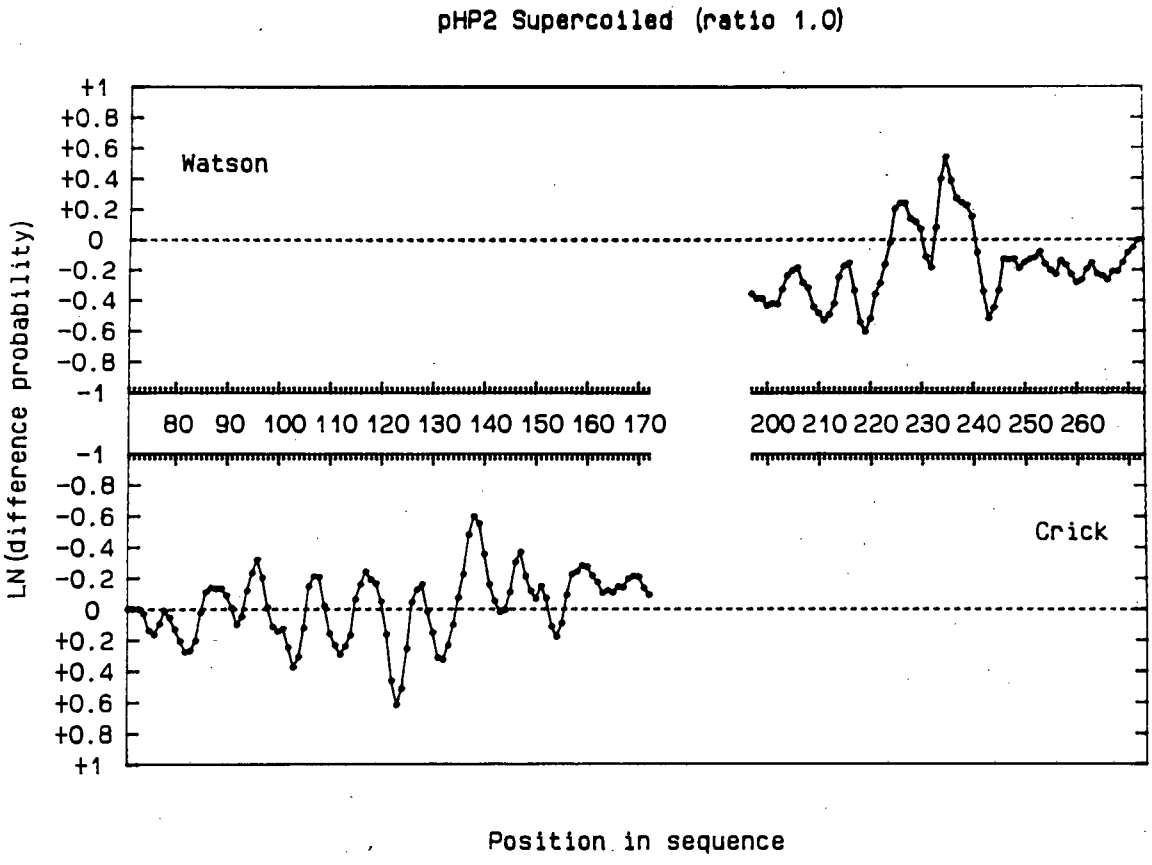

Figure 5.5. Probability difference of DNase I cleavage in supercoiled pHP2 reconstituted with the histone octamer at a molar octamer:DNA ratio of 11. The probability of cleavage, expressed as the natural logarithm of the difference between the rate of cleavage in reconstituted and free plasmid (see section 2.8.3 of materials and methods) is shown as a three point running average at different positions in the sequence of pHP2. Cleavage data on the Watson and Crick strands are shown in the top and bottom sections of the figure respectively. See text for detail.

site, weak maxima, which follow a periodicity of approximately 5bp, are visible at positions 247, 253, 257 and 263. On the Crick strand, maxima in the probability of DNase I cleavage are visible at positions 82, 94, 103, 112, 123, 132, 144, 154 and 164. A shoulder is visible at position 100 on the peak that displays a maximum at position 103 (see figure 5.5). This discontinuity can again be assigned to the linker DNA joining two adjacent cores, and thus, by analogy to the core on the linearized molecule discussed above, suggests that the borders of core 2 are situated at positions 103 (Crick) and 235 (Watson). The extent of the core (235-103=132bp) and the magnitude of the probabilities of cleavage on the Watson strand of core 2 again suggest that the maxima at positions 103 and 235 correspond to sites 130 and 140 of the mixed sequence core (142), thus, employing the same argument as above, placing the dyad axis of core 2 on the supercoiled molecule at position 165.

This result indicates that the positioning of core 2 is not changed in proceeding from a linearized to a supercoiled molecule. The individual sites of maximal DNase I cleavage in the linearized and supercoiled molecule also align well, with a discrepancy of, at most, plus or minus one base-pair. The non-uniformity of this offset is consistent with the absence of a twist change in the terminal helical turns of the supercoiled molecule, as expected if the structure of the supercoiled molecule reconstituted at a molar octamer:DNA ratio of 11 approximated that of a reconstituted circular molecule with little unrestrained torsional stress.

A further test on the alignment of the probabilities of cleavage in the core on the linearized and supercoiled molecules may be undertaken to ensure the correctness of the visual deduction. To test whether the positions of maximum cleavage by DNase I on the Watson and Crick strands of the

linearized and supercoiled molecules do indeed align properly, a difference correlation between the two data sets can be done.

The best alignment between the probabilities of cleavage in the supercoiled and linearized molecule for core 2 was calculated by a correlation function in the form of

$$C_j=(1/N)\sum_{i=0}^{N}|S_i-L_{i+j}| \qquad (5.1)$$

where j represent the offset between the two datasets and was set equal to integer values from -4 to +5, and i represents the index of comparison, and was successively set to each position of DNase I cleavage within the limit of core 2 on the supercoiled molecule. S_i is the probability of cleavage in the reconstituted supercoiled molecule at position i, and L_{i+j} the probability of cleavage at position i+j on the reconstituted linearized molecule. The summation of the differences between the probabilities of cleavage on the supercoiled and linearized molecules at a particular offset j, was normalized to N (=120), the total number of positions compared (see also section 2.8.4).

It is clear from equation 5.1 that if two datasets are poorly aligned, that is, the peaks in the one dataset is overlaid onto the troughs of the second dataset, the value of the correlation coefficient, C_j, will be large. However, if the two datasets are aligned well, that is, peaks aligned with peaks, C_j will be small. The correlation between the difference probability plots for core 2 on supercoiled and linearized DNA is shown in figure 5.6.

In figure 5.6 it is seen that the correlation coefficient varies from a maximum value at a base-pair offset of -4, to a minimum value at an offset of 0, and back to a maximum value at an offset of +5. This is exactly what is expected

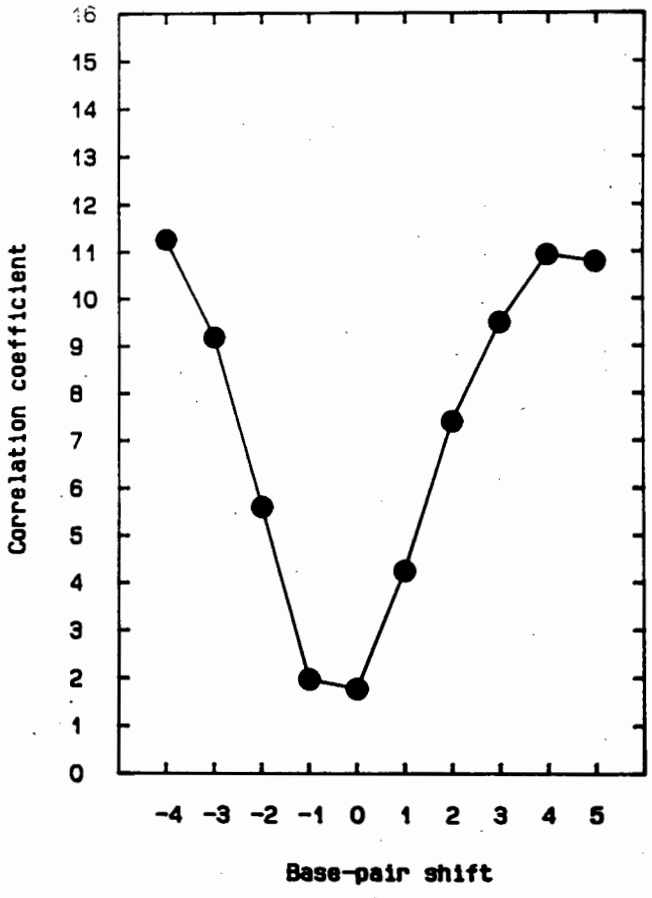

Figure 5.6. Correlation between the sites of maximum DNase I cleavage in core 2 on supercoiled and linearized pHP2 reconstituted at a molar octamer:DNA ratio of 11. The correlation coefficient, shown on the ordinate, was obtained by summation of the differences between the probability difference datasets of core 2 at different offsets, shown on the abscissa, of the dataset of the linearized molecule relative to that of the supercoiled molecule. See text for detail.

if the peaks in the difference probability plots of reconstituted supercoiled and linearized DNA is shifted from -144° out of phase, to 0° (in phase), to 180° out of phase. The minimum found at a base pair offset of 0 therefore indicates that identical positions in the sequence of reconstituted linearized and circular pHP2 are exposed to cleavage by DNase I, and confirms the identical rotational placements of core 2 on the linearized and supercoiled plasmid. It is formally possible that the translational position of core 2 on the linearized and circular molecule differ by an integer helical turn or a multiple thereof, a feature to which this type of correlation analysis is not sensitive. The discontinuity in the periodicity of DNase I cleavage at positions 104 (Crick) and 235 (Watson) on both the linearized and circular molecule suggest identical translational settings for core 2. The result of this comparison therefore indicates that core 2 positions in identical translational and rotational settings on linearized and supercoiled plasmids assembled at a molar octamer:DNA ratio of 11.

5.3 The effect of negative supercoiling on nucleosome core positioning:

In order to investigate the effect of negative superhelical stress on the positioning of the histone octamer, linearized and supercoiled (σ=-0,06) pHP2 were reconstituted with the histone octamer at a molar octamer:DNA ratio of 8, and DNase I digestions and primer extensions of the reconstituted and free DNA samples performed as described above (see also sections 2.4.4, 2.6.2 and 2.7 of materials and methods). It was found that this molar assembly ratio represented the lower limit that still gave a clear modulation in the periodicity of cleavage by DNase I. At lower assembly ratios, this modulation was difficult to detect.

At this ratio of assembly, where the difference between the average number of cores per linearized and supercoiled molecule was shown to be small (see section 4.4), the unrestrained negative superhelical stress in the supercoiled plasmid should equal -10,9+8X1,1=-2,1 which is equal to a specific linking difference of -0,03 in the linker DNA joining adjacent cores, assuming that 146bp is bound per nucleosome core. However, the position assumed by core 2 on this supercoiled molecule may not only reflect the influence of superhelical tension on the core position, but also the effect of core-core interactions, since the supercoiled molecule contains, on average, less nucleosome cores than the molecule investigated above at a molar assembly ratio of 11. Any such effect should, however, be apparent in a comparison of the positions of cores 1-3 on the linearized molecules assembled at molar ratios of 8 and 11 respectively. The result of the DNase I digestion of supercoiled and linearized DNA assembled at a molar octamer:DNA ratio of 8 is shown in figure 5.7.

In figure 5.7 A the maxima in the modulation of DNase I cleavage on the Watson strand (lanes 3-6) and Crick strand (lanes 9-12) of assembled pHP2 (lanes 3, 5, 9 and 11) are indicated by the solid triangles in the margins. Once again this modulation has a period of approximately 10bp indicating the presence of a rotationally positioned nucleosome core on both the linearized (lanes 3 and 9) and supercoiled (lanes 5 and 11) molecules. In figure 5.7 B an enlarged section of A is shown in which this modulation is even more clear. In order to obtain an accurate indication of the location of the core(s) on the linearized and supercoiled molecules, difference probability plots were calculated from the gel shown in figure 5.7. The results are presented in figure 5.8 and 5.10

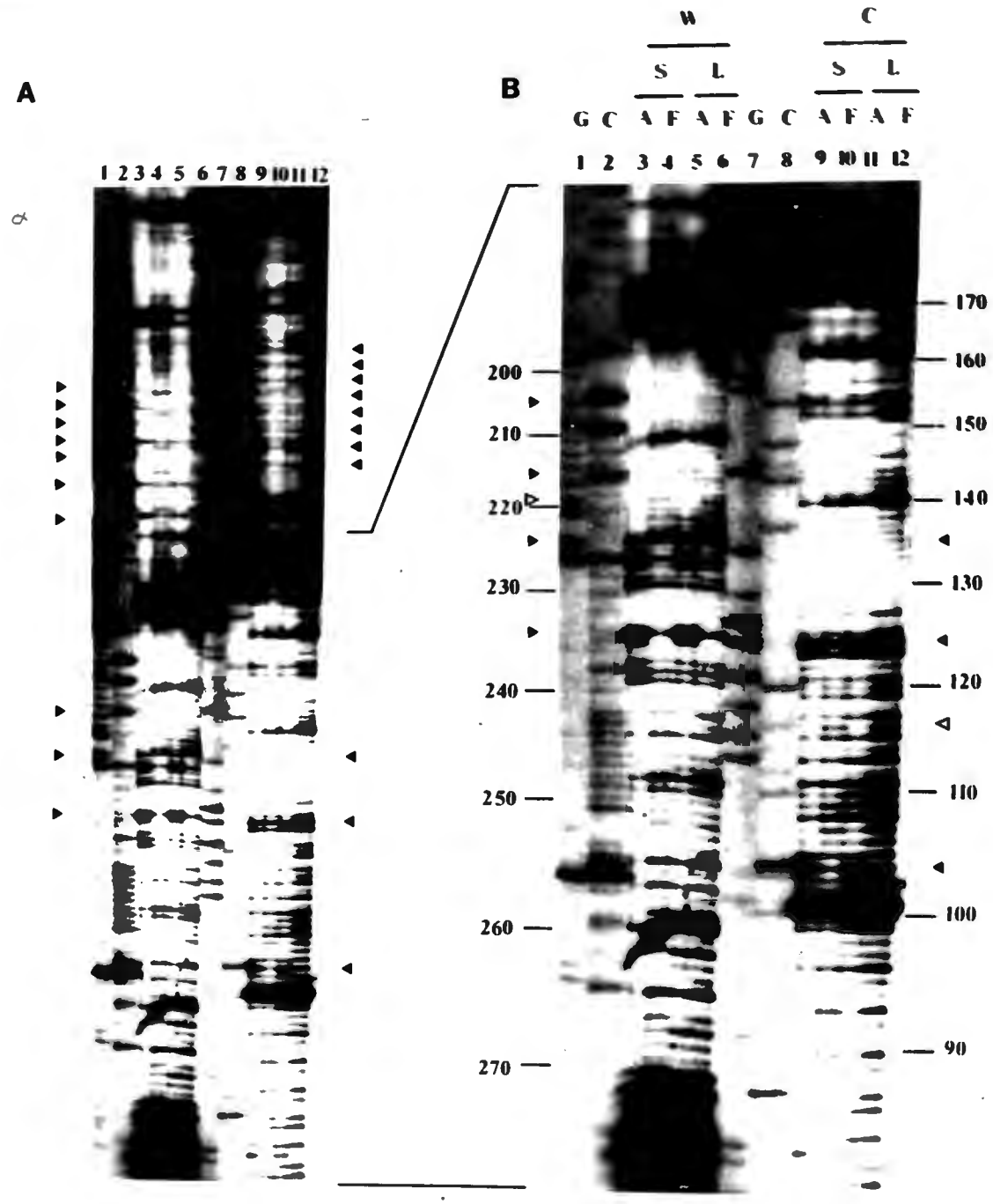

Figure 5.7. DNase I digestion of pHP2 reconstituted with the histone octamer at a molar octamer:DNA ratio of 8. The location of nucleosome cores on supercoiled (S; lanes 3, 4, 9 and 10) and linearized (L; lanes 5, 6, 11 and 12) pHP2 were determined by DNase I digestion of assembled (A; lanes 3, 5, 9 and 11) and free (F; lanes 4, 6, 10 and 12) plasmid followed by extension of 20mer primers annealed to either the Watson (W; lanes 3-6) or Crick (C; lanes 9-12) strand. The location of DNase I cleavage sites was identified relative to guanosine (G) or cytosine (C) chain terminating sequencing reactions of 20mer primers annealed to the Watson (lanes 1 and 2) or Crick (lanes 7 and 8) strands. The sequence positions of individual bands are shown in the margins of (B). The maxima in the modulation of DNase I cleavage with a period of approximately 10bp in reconstituted samples are indicated by solid triangles. Positions where DNase I cleaved at a lower frequency in the reconstituted samples relative to the free DNA are indicated by open triangles. Panel B is an enlarged section of panel A. An autoradiograph of the gel is shown. See text for detail.

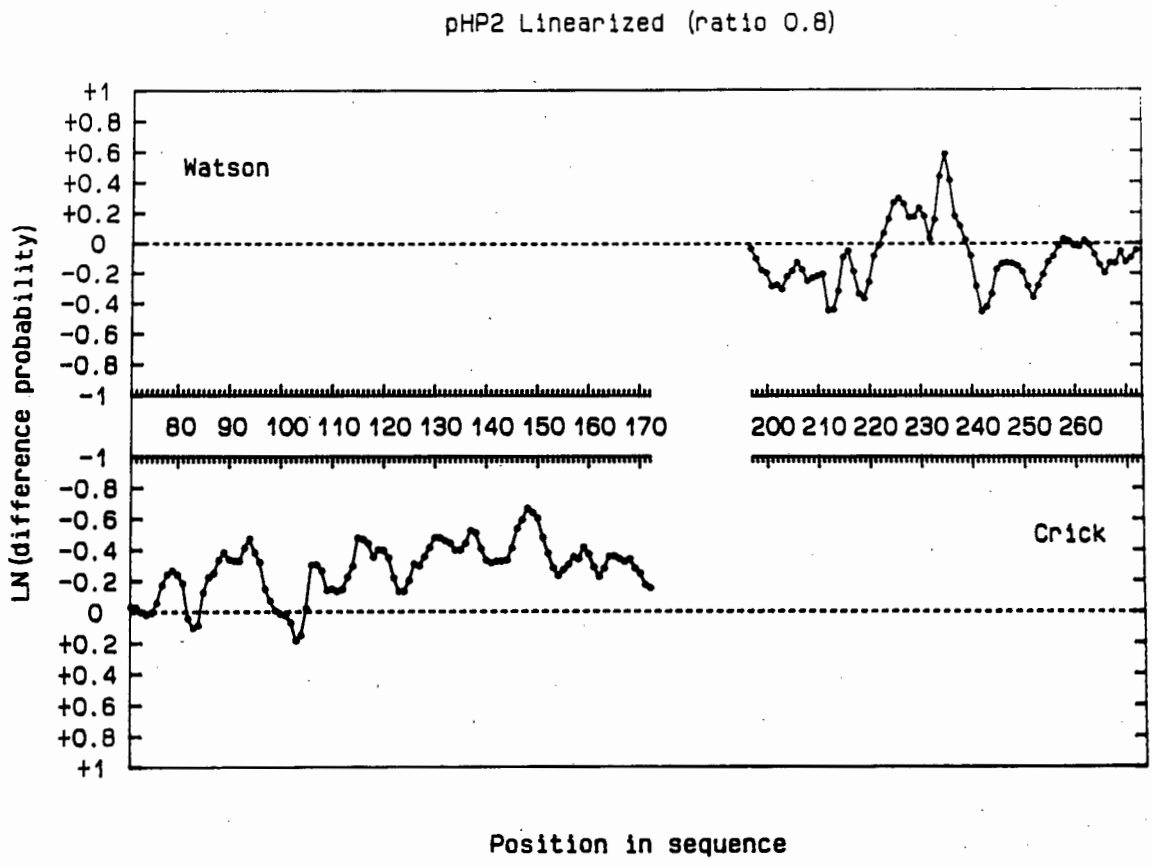

Figure 5.8. Probability difference of DNase I cleavage in linearized pHP2 reconstituted with the histone octamer at a molar octamer:DNA ratio of 8. The probability of cleavage, expressed as the natural logarithm of the difference between the rate of cleavage in reconstituted and free plasmid (see section 2.8.3 of materials and methods) is shown as a three point running average at different positions in the sequence of pHP2. Cleavage data on the Watson and Crick strands are shown in the top and bottom sections of the figure respectively. See text for detail.

In the linearized molecule (figure 5.8), maxima appear at positions 83, 92(weak), 103, 111, 124, 134, 142, 154 and 162(weak) on the Crick strand, and at positions 206, 216, 226 and 235 on the Watson strand. In order to determine whether the sites of maximal cleavage in core 2 have shown a consistent shift relative to the cleavage sites on the linearized molecule assembled at a molar octamer:DNA ratio of 11, the correlation between the cleavage patterns in figure 5.8 and 5.3 was calculated as explained above (see section 2.8.4 of materials and methods). The result is shown in figure 5.9. The minimum obtained in the correlation coefficient at a base-pair offset of 0 indicates that neither the rotational nor translational setting of core 2 changed due to a decrease in the molar ratio of assembly from 11 to 8. The possibility that the translational setting shifted by an integer number of times the local twist of the core DNA is countered by the discontinuity in DNase I cleavage and the magnitude of the probability of cleavage at the identified maxima on the 3'-side of the Watson strand, as discussed above. This result therefore suggests that the positioning of core 2 on supercoiled DNA assembled at a molar octamer:DNA ratio of 8, should reflect the effect of the unrestrained negative torsional tension in the molecule, rather than a possible decrease in core-core interactions at the lower assembly ratio.

Turning next to the difference probability plot of the supercoiled pHP2 assembled at a molar ratio of octamer:DNA of 8 (figure 5.10), maxima in the probability of cleavage by DNase I in the reconstituted molecule are seen at positions 83, 92, 103, 112, 123, 133, 144, 155 and 164(weak) on the Crick strand, and at positions 206, 216, 226 and 235 on the Watson strand. This series of cleavage maxima is in close agreement, not only with the cleavage series found on the linearized molecule, but also with that found on the

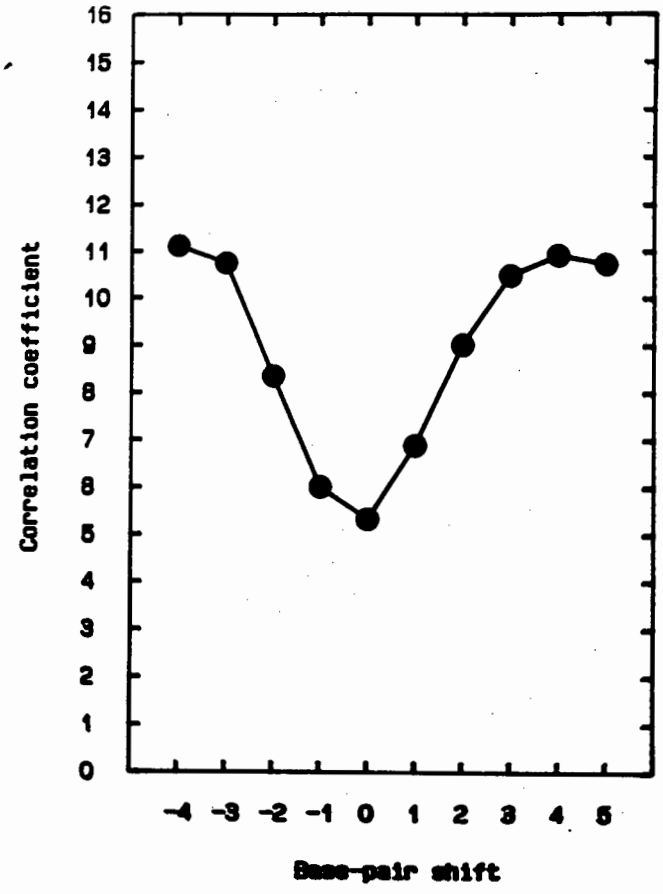

Figure 5.9. Correlation between the sites of maximum DNase I cleavage in core 2 on linearized pHP2 reconstituted at molar octamer:DNA ratios of 8 and 11. The correlation coefficient, shown on the ordinate, was obtained by summation of the differences between the probability difference datasets of core 2 at different offsets, shown on the abscissa, of the dataset of the linearized molecule assembled at a molar octamer:DNA ratio of 8 relative to that of the linearized molecule assembled at a molar octamer:DNA ratio of 11. See text for detail.

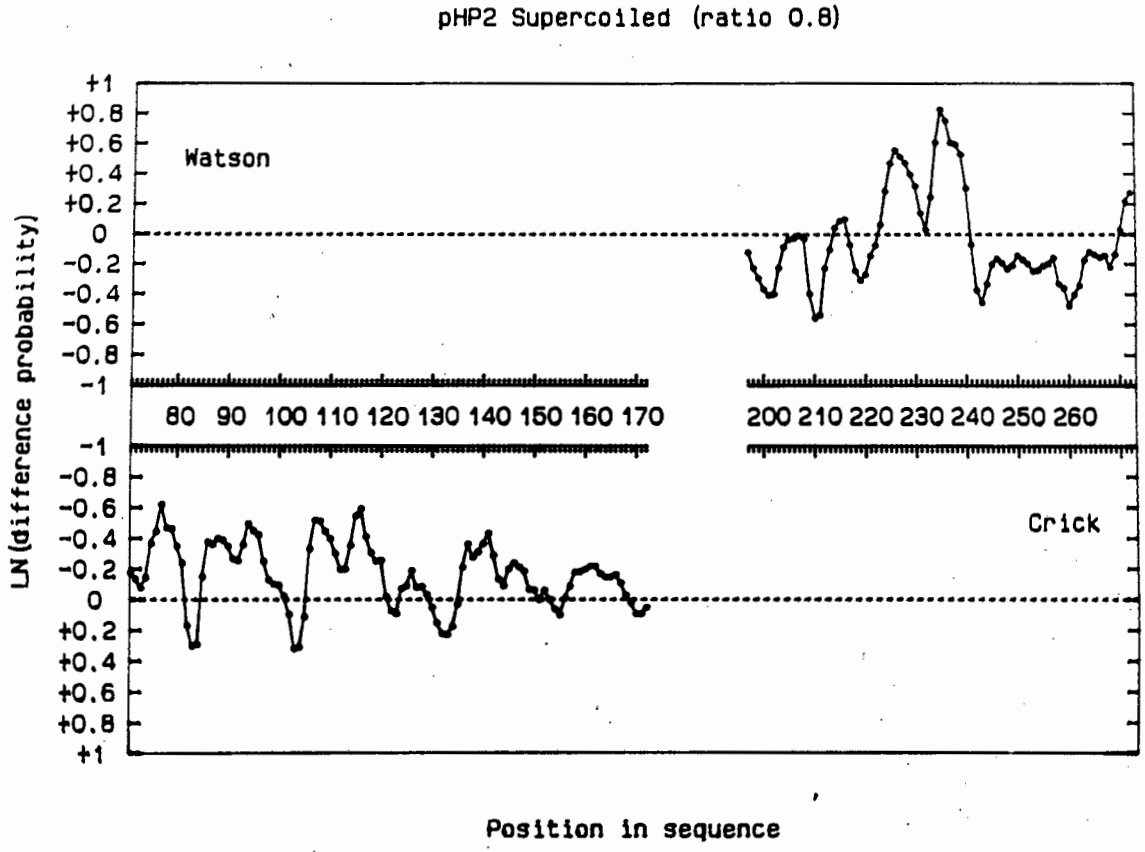

Figure 5.10. Probability difference of DNase I cleavage in supercoiled pHP2 reconstituted with the histone octamer at a molar octamer:DNA ratio of 8. The probability of cleavage, expressed as the natural logarithm of the difference between the rate of cleavage in reconstituted and free plasmid (see section 2.8.3 of materials and methods) is shown as a three point running average at different positions in the sequence of pHP2. Cleavage data on the Watson and Crick strands are shown in the top and bottom sections of the figure respectively. See text for detail.

supercoiled molecule assembled at a molar ratio of 11. The preliminary indication is therefore that the level of superhelical stress present in the supercoiled molecule (σ=-0,03 in the linker areas) has little effect on the positioning of core 2. In order to ensure that no subtle differences are overlooked, a correlation analysis was performed between the cleavage maxima found in the supercoiled molecule assembled at a molar ratio of 8, the corresponding linearized molecule, and also the supercoiled molecule assembled at a molar ratio of 11. The results of this comparison are presented in figure 5.11 and 5.12.

Both the curves relating the coefficient of correlation to the offset of the cleavage probabilities (figures 5.11 and 5.12) show well defined minima at an offset of 0. This result therefore indicates that in the presence of superhelical stress equivalent to a specific linking difference of -0,03 in the linker areas, core 2 occupies a position that is identical to the corresponding core on the circular molecule in the absence of superhelical stress, and also to the core on a linearized molecule.

5.4 Curvature in the positioning sequences of cores 1-3:

DNA exhibits a sequence-dependent conformation (137), as is evident in the structures of several double-stranded oligonucleotides solved by X-ray crystallography (reviewed by Drew *et al.* (61)). The distribution of appropriate short sequences in phase with the helical period of DNA can lead to the accumulated deflection of the local helical axis to describe an overall curve in a plane (94, 278). Experimental evidence for the existence of a helical axis that is, on average, curved, comes from several different approaches. In the electrophoretic separation of DNA fragments in acrylamide gels, curved fragments migrate anomalously slow compared to uncurved fragments of the same

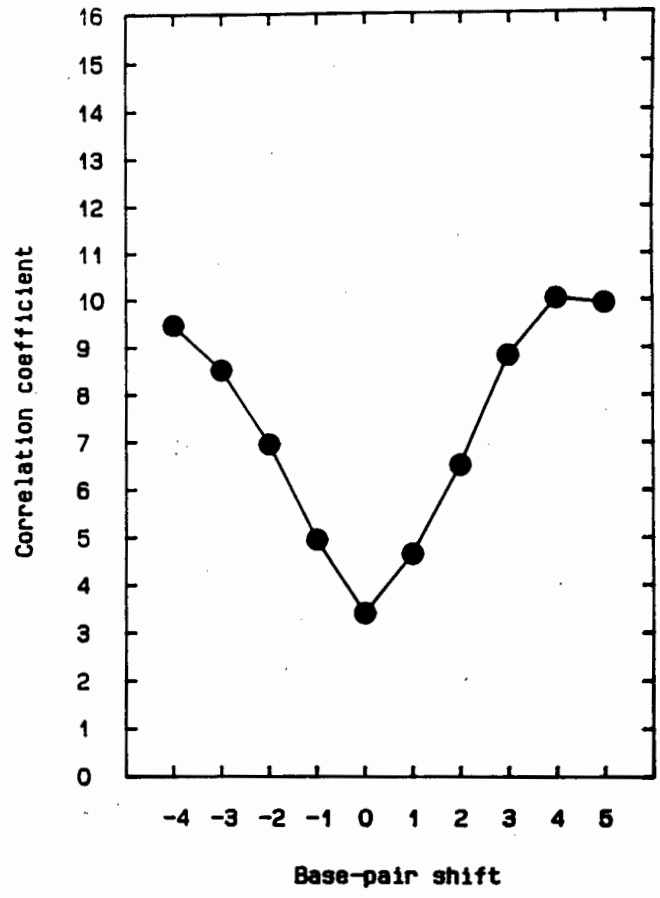

Figure 5.11. Correlation between the sites of maximum DNase I cleavage in core 2 on supercoiled and linearized pHP2 reconstituted at a molar octamer:DNA ratio of 8. The correlation coefficient, shown on the ordinate, was obtained by summation of the differences between the probability difference datasets of core 2 at different offsets, shown on the abscissa, of the dataset of the linearized molecule relative to that of the supercoiled molecule. See text for detail.

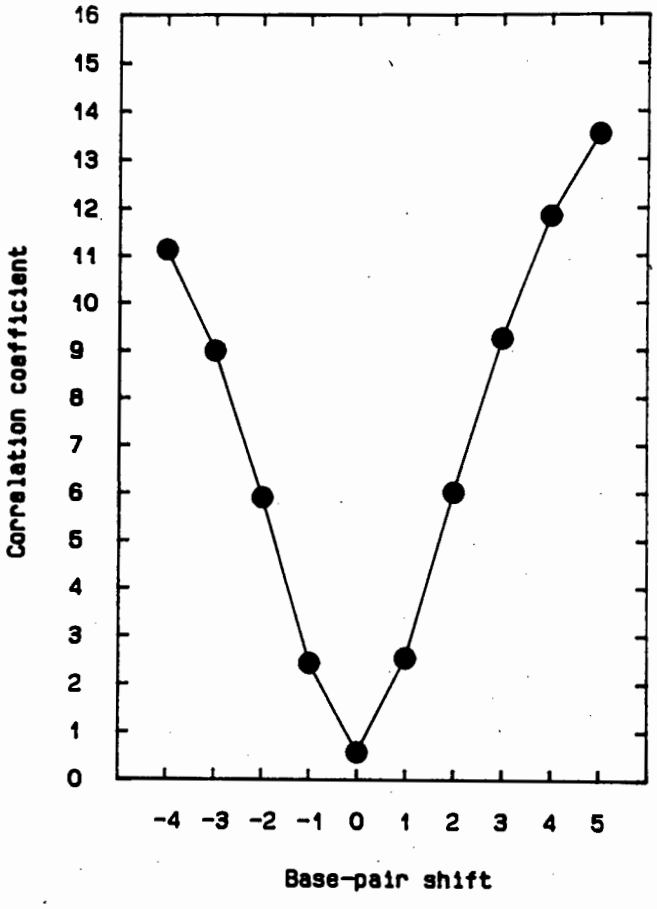

Figure 5.12. Correlation between the sites of maximum DNase I cleavage in core 2 on supercoiled pHP2 reconstituted at molar octamer:DNA ratios of 8 and 11. The correlation coefficient, shown on the ordinate, was obtained by summation of the differences between the probability difference datasets of core 2 at different offsets, shown on the abscissa, of the dataset of the supercoiled molecule assembled at a molar octamer:DNA ratio of 8 relative to that of the supercoiled molecule assembled at a molar octamer:DNA ratio of 11. See text for detail.

size (149, 229, 278). Levene and Zimm (133) proposed that this phenomenon reflects a smaller range of possible migration paths through the gel matrix for curved molecules, where the molecule chooses a path to deviate as little as possible from the elastic free energy of its preferred conformation (see also Calladine *et al.* (39)). In contrast, uncurved molecules can be aligned parallel to the applied field, and the probability is less that these molecules assume a conformation that represents a lower elastic free energy than the same molecule migrated one step further along in the direction of the applied field (133; see also reference 39). Further evidence for curvature comes from electric birefringence (93), electron microscopy (104, 132, 170), nuclear magnetic resonance (92, 171), hydroxyl radical cleavage (32) and from religation studies (252, 281). Several sets of theories to explain the phenomenon of curvature have been published (see, for example, 39, 123, 226 and 253) which were comprehensively reviewed by Diekmann (52), Hagerman (96), Travers and Klug (246) and Travers (244).

It was shown that curvature of DNA is often associated with a positioned core (47, 104, 188), a phenomenon that may be understood in the decreased energy required in wrapping the DNA onto a superhelical axis perpendicular to the azimuth of the curve as opposed to any other direction. (See also the discussion on anisotropic flexibility as a determinant of the rotational positioning of nucleosome cores in section 1.4.4.2).

In order to test the involvement of curvature in the positioning of cores 1-3 identified above, the sequences on which these cores position were isolated as *Hga* I-*Fok* I (1878bp-113bp), *Fok* I-*Asp* 718 (113bp-292bp) and *Asp* 718-*Acy* I (292bp-427bp) fragments of 150bp, 179bp and

136bp respectively, and electrophoretically resolved on a 10% acrylamide gel. The result is presented in figure 5.13.

The apparent size of each fragment resolving as the bottom band in lanes 1-3 of figure 5.13 was calculated with a sixth-order polynomial fitted to the average distance migrated by the standard fragments (lanes labelled M). By this method the apparent size of the fragment on which core 1 formed was calculated at 140bp, the fragment of core 2 at 212bp, and that of core 3 at 138bp. The ratio of the apparent size in base-pairs to the real size ("k-factor") of these three fragments is therefore equal to 0,93, 1,18 and 1,01. These values indicate that the positioning sequences of cores 1 and 3 are uncurved, whereas that of core 2 is curved in solution.

5.5 Discussion:

5.5.1 Core-core and folding interactions:

An indication of the contribution of core-core interactions to the positioning of the histone octamer can be obtained by comparing the position of core 2 on the linearized 1915bp pHP2, with that of the nucleosome core forming over the identical sequence on a 337bp fragment, investigated by Retief *et al.*(203). This fragment contained the sequence from position 6042 to 6356 of h22 (see 203), which included the whole of the positioning sequence of core 2, and a section of the positioning sequence of core 3. The positioning of the histone octamer on this 337bp fragment was investigated after poly[L-glutamate] facilitated assembly at an octamer:DNA (w/w) ratio of 0,6, an assembly ratio at which an average of 1 nucleosome core was expected to form per DNA molecule (203). Unfortunately no difference probability analysis of the DNase I digestion data was performed by Retief *et al.* (203), but by comparing the

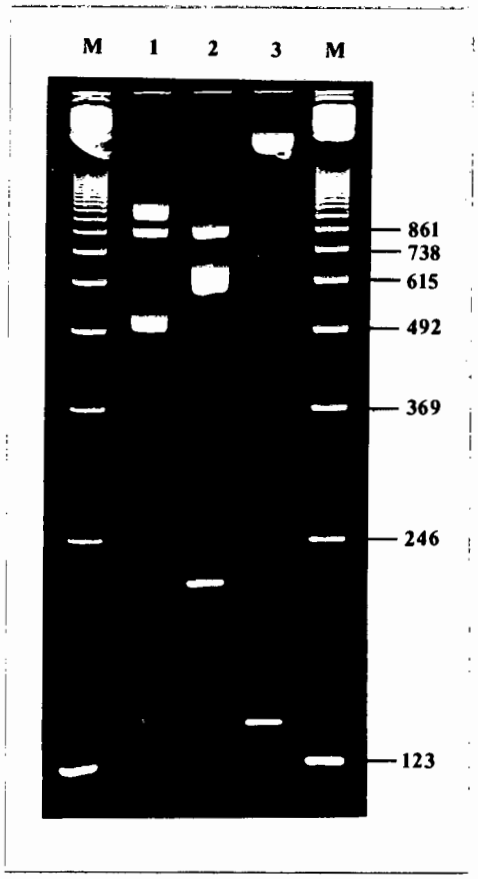

Figure 5.13. Curvature in the positioning sequences of cores 1-3. pHP2 was double digested with either of the restriction enzyme pairs *Hga* I-*Fok* I (lane 1), *Fok* I-*Asp* 718 (lane 2), or *Asp* 718-*Acy* I (lane 3) to yield the areas of pHP2 on which cores 1-3 position as 150bp, 179bp and 136bp fragments respectively. The digested plasmid was electrophoresed on a 10% (w/v) acrylamide gel run in TBE at 4°C where the core fragment resolved as the bottom band in each of the lanes. A 123bp ladder was included as a size standard (M) with fragment sizes in base-pair indicated in the right-hand margin. The gel was stained in ethidium bromide.

identified sites of maximal DNase I cleavage on the gels (202), an estimate of the rotational placements of the core on the long and short molecule can be obtained. On the short fragment the sites, converted to the numbering scheme shown in figure 1.1, occur at positions 125, 135, 145, 155 and 165. On the complementary Watson strand no sites of maximal DNase I cleavage in the region of core 2 were pertinently identified by Retief (202). A comparison of the sites on the Crick strand of the 337bp fragment with those of pHP2 reported above (see section 5.2) reveals a close match. This agreement would suggest that core 2 assumed identical positions on a 1915bp and a 337bp fragment. This result should however be treated with caution, since in the study by Retief (202) the sites of cleavage were identified relative to an end-labeled *Hpa* II standard of pBR322 and not a sequencing standard electrophoresed on the same gel, a procedure that may introduce a considerable error in assigning DNase I cleavage positions to positions in the sequence. However, a careful comparison of the local DNase I cleavage patterns in the corresponding regions of pHP2 and the 337bp fragment suggests that the rotational placement of core 2 is identical on the long and short fragment.

At face value this is an interesting finding, since on the linearized pHP2 reconstituted with the histone octamer at a molar octamer:DNA ratio of 11, core 2 should be flanked on the one side by 2-3 cores, and on the other side by 7-8 cores. In the study by Drew and McCall (56) it was found that the rotational placement of the central core on a 440bp fragment occupied by 3 cores shifted by approximately 2bp relative to the position of this core on the equivalent section of core-length DNA. In the study by Drew and Calladine (55), a difference in positioning of the two central cores on a 860bp fragment containing five octamer-DNA complexes relative to that on the corresponding

short core-length fragments was also reported. This change in the position of the core(s) was ascribed to the long-range folding of the larger molecule which forced the central cores to occupy positions other than that indicated by the local sequence-directed bending properties of the DNA molecule (56).

In the present study it appears that core 2 assumed identical positions on a 1915bp fragment that contained 11 cores, and on a 337bp fragment that contained 1 core. It therefore appears that the bending preferences of the DNA sequence occupied by core 2 is the dominant positioning determinant of this core. In fact, as will be shown in chapter 7, the predicted anisotropic flexibility of the core DNA in core 2 suggests that this length of DNA is an even stronger positioner than the 5S rRNA gene, a sequence previously shown to direct the histone octamer into a precise setting (55, 205). This result therefore indicates that the position of the core can be determined by the local bending properties of the DNA alone, even if abutted by several cores on either side, and need not be modulated by core-core interactions and the folding of chromatin *in vivo*.

The very favorable positioning frame dictated by the base-pair sequence of core 2 may be the reason for the conservation of this frame in a circular molecule in the presence and absence of superhelical tension. It may be argued that if a less well positioned core is investigated, the topological properties of the DNA molecule may influence the core position. However, the results presented above argue against this suggestion as a general rule. It was shown that core 3 fortuitously assumed a mixture of rotational settings differing by 5bp or approximately half a helical turn (see section 5.2.2). This mixture of rotational settings was maintained in a circular molecule in

both the presence and absence of unrestrained negative superhelical tension (see figures 5.10 and 5.5).

The 5bp offset in the rotational positions of core 3 is also interesting for another reason. Drew and Calladine (55) showed that a core that positioned in a mixture of rotational settings on a core-length DNA fragment, assumed a single rotational setting in a 860bp molecule occupied by 5 nucleosome cores. This single rotational setting differed from both those present on the short fragment (55). Core 3, on the other hand, displayed a mixture of rotational settings on a 1915bp fragment occupied by 11 nucleosome cores, and this demonstrates that, though core-core interactions/folding interactions can induce a core to assume a single rotational setting, this does not always happen.

5.5.2 Nucleosome core positioning on circular molecules in the absence and presence of superhelical stress:

The results of the positioning of the nucleosome core on circular molecules in the absence and presence of negative superhelical stress can be analyzed from two different directions. The first approach involves the possible contribution of the shape of the supercoiled molecule presented to the octamer on the positioning of the core, and requires an analysis of the geometry of the naked supercoiled molecule in solution. The second approach involves the steric constraints imposed on the nucleosome cores in a small circular molecule, and must include a proposal on the expected conformation of a reconstituted circular molecule.

5.5.2.1 Geometry of the free supercoiled molecule:

It was shown that a negatively supercoiled molecule assumes a plectonemic shape in solution (1, 225). This form of supercoiled DNA can be approximated as the right-handed winding of the DNA helix up and down the length of a hemi-spherically capped cylinder. Although it was shown that the binding of DNA-bending proteins was enhanced by negative supercoiling (reviewed in 78), suggested to reflect the higher affinity of the protein for the site that is already bent due to the geometry of the supercoiled molecule, this quality is unlikely to affect the positioning of the histone octamer. This is due to fundamental geometric differences between plectonemically and toroidally supercoiled DNA, as can be seen by calculating specific topological and geometric quantities of the two different forms of supercoiled DNA.

It was shown that the pitch angle (γ) of plectonemically supercoiled DNA equalled 54° over a range of specific linking differences (1, 22). It was further shown by Cozzarelli and colleagues (Boles *et al.*, 22) that the number of superhelical turns (n) in plectonemically supercoiled DNA was related to the linking difference (ΔLk) by $n=-0,89\Delta Lk$. Thus, using as example a 2000bp plasmid at $\sigma=-0,06$ (a close approximation of supercoiled pHP2), n=10,17, and the writhe (Wr) will equal $-n\sin\gamma=-8,23$. Since the surface linking number (SLk) of plectonemically supercoiled DNA is zero, the surface twist (STw) will equal 8,23 (from the relation SLk=STw+Wr). The winding number (ϕ) can now be calculated from the relation $Lk=STw+Wr+\phi$, and equals 179. The helical repeat (h; $=N/\phi$) of the plectonemically supercoiled DNA, expressed relative to the surface normal of the virtual cylinder, equals 11,2bp. The helical repeat of the supercoiled DNA is therefore not only larger than that of relaxed DNA (206, 207, 260, 261) but also larger than the

helical repeat of nucleosomally wrapped DNA, where h=10,0bp-10,2bp (55, 59, 142).

This is an important difference, since the sequence periodicity of di- and trinucleotides with similar bending preferences is close to 10,2bp in core DNA (59, 210). The match between the sequence periodicity and helical repeat of core DNA was proposed to represent the requirement for the placement of sequences with similar anisotropic flexibilities at equivalent angular orientations in the core DNA relative to the octamer surface (59, 210), and forms the basis for the dependence of the rotational placement of the histone octamer on the anisotropic flexibility of the core DNA (55, 56, 59) The bending imposed on plectonemically supercoiled DNA, in contrast, will be governed by the sequence periodicity at 11,2bp, a repeat at which the anisotropically flexible sequences of core DNA will not be in phase with the surface normal. The geometric shape of a supercoiled molecule described by the distribution of writhe is therefore expected to reflect the local bending properties of the DNA at a period of 11,2bp, and, as such, is not expected to influence the rotational placement of a nucleosome core. A similar point was made by Camerini-Otero and Felsenfeld, arguing against the contribution of the bending in right-handed plectonemically supercoiled DNA to the bending of the DNA into a left-handed toroidal supercoil around the histone octamer (40). A notable exception to these non-equivalent geometries is the DNA sections located at the end-loops of the plectonemically supercoiled molecule. Laundon and Griffith (132) reported that curved sections displayed a propensity to be placed at the end-loops, a phenomenon suggested to reflect the bending requirements in these areas. Curved DNA has often been shown to be associated with rotationally positioned cores (42, 47, 188, 193; this study), and in this regard the sequences located at the end loops of supercoiled molecules

may position a nucleosome core in a well-defined rotational setting. This will however be due to the inherent bending characteristics of the DNA, rather than an effect of the geometric shape of the supercoiled molecule on nucleosome core placement.

The argument presented above assumed the octamer as a single hard sphere, whereas the reconstitution results reported in this study was obtained by the step-wise dialysis of an octamer and DNA mixture from high to low salt, where first the $(H3\text{-}H4)_2$ tetramer associates with the DNA at 1,0M NaCl, followed by the two H2A-H2B dimers at 0,8-0,6M NaCl (100). The sequential manner by which the reconstitution proceeds will tend to further negate any influence of the initial shape of the supercoiled molecule on core placement. Furthermore, during the final stages of dialysis, the octamer will have significant freedom to assume the energetically most favorable local position on the DNA molecule. The initial geometric shape of the supercoiled DNA molecule presented to the octamer is therefore not expected to have any influence on the positioning of nucleosome cores.

5.5.2.2 Proposed conformation of a circular reconstituted molecule:

The discoid core particle has a diameter of 110Å, a width of 57Å, and the DNA supercoil an average pitch of 27Å (208, 251). Thus, if two cores are placed side by side with their short axes coinciding, the length of DNA required to connect the point of exit of the one core with the point of entry of the other core is equal to $\sqrt{(0,4\pi p)^2+(0,4\pi r)^2}$=70Å, where r is the radius of the core (55Å) and p the pitch of the linker supercoil (42Å/2π), assuming that 1,8 superhelical turns are bound to the surface of each core. This is equal to 21bp of B-DNA, where the base-pair to base-pair distance

along the local helix axis is equal to 3,4Å. In the 1915bp pHP2 reconstituted with an average of 11 nucleosome cores, each core occupies 174bp of DNA, which leaves 28bp of linker DNA per core. If the reconstituted supercoiled molecule is approximated as a torus with the edges of individual, radially arranged cores touching towards the center of the circle, the length of linker is sufficient to connect adjacent cores if the points of DNA entry and exit are also oriented towards the center of the circle, as schematically shown in figure 5.14 A. The edge-to-edge distance on the far side of the cores is approximately 62Å. If two cores are placed side by side at a distance of 62Å with their short axes coinciding, the length of the linker DNA required to connect adjacent exit and entry points is equal to 99Å, which is equal to 29bp of B-DNA. It therefore appears that the reconstituted molecule arranged with the points of DNA entry and exit orientated away from the center of the circle is disallowed by the length of available linker. Furthermore, the pitch angle of the superhelical path traced by the linker DNA in the latter model is equal to 46°, whereas the pitch angle of the linker DNA with the core entry and exit points oriented towards the center of the circle equals 7°. The bending requirements of the linker DNA as it leaves the surface of the one octamer or approaches the surface of the adjacent octamer is thus far less in the latter case. It is expected that the linker, in order to minimize the angle of the required bend, will act like a spring, and tend to orientate the cores with the points of DNA entry and exit towards the center of the circle. The arrangement of the cores shown in figure 5.14 A thus seems to represent the favorable conformation of the reconstituted molecule in the absence of superhelical stress.

It is possible that individual cores may rotate by 180° to occupy the central hole of the complex, which has a diameter

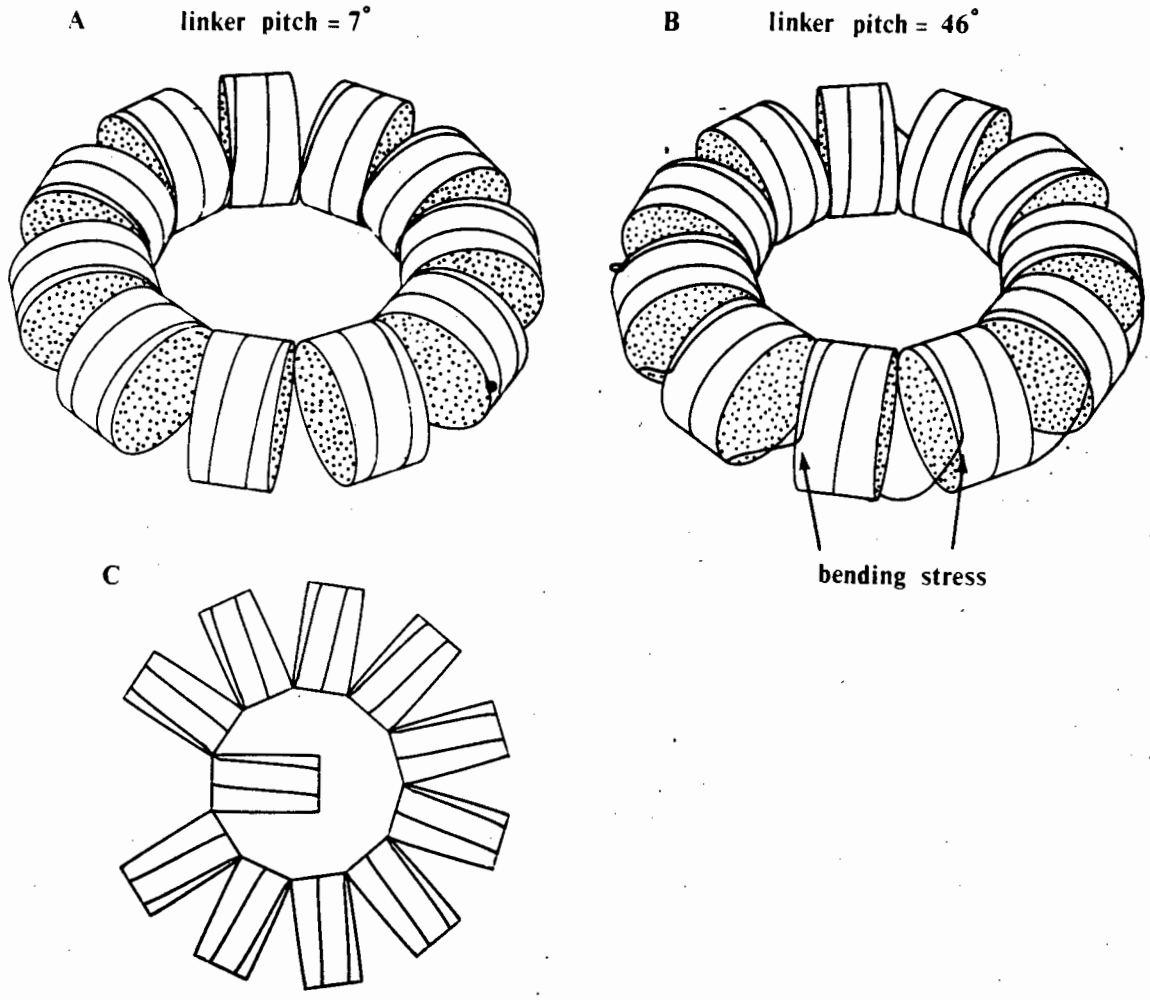

Figure 5.14. Schematic representation of a 1915bp circular plasmid occupied by 11 nucleosome cores. In the top part of the figure the circular molecule is shown with the points of DNA entry and exit of individual nucleosome cores located towards the inside (A) or outside (B) of the circle. The top view (C) of the circular molecule shows the size of the central hole. The dimensions of the core was taken from Richmond *et al.* (208) and Uberbacher and Bunick (251) and is in scale relative to the length of the 1915bp DNA molecule, but not the diameter of the DNA helix.

of 194Å, as schematically shown in figure 5.14 C. The size of this central hole would, however, suggest that no more than one core can occupy it at a time, or two at most if the reconstituted circle is slightly flattened. It is expected that more than two cores would be excluded on steric grounds. The alternating arrangement of the cores with the entry and exit points sequentially placed towards and away from the center of the circle cannot be supported by the length of available linker, although the rotation of consecutive cores to positions above or below the plane of the circle appears possible.

It is thus clear from such a purely mechanistic approach that the supercoiled molecule occupied by 11 nucleosome cores would have some conformational latitude. It is also clear from this model that conservative levels of supercoiling should be accommodated by apportionment between the writhe of the toroidal axis, and the twist in the linker areas, once again allowing individual octamers to comply with the sequence periodicity of the core DNA. The conformational flexibility of the circular molecule is also expected to be wider at an octamer:DNA molar ratio of 8, where the average linker length is increased to 93bp. Note, however, that the Wr of plectonemically supercoiled DNA is related to the supercoil radius (r) by $-np/\sqrt{r^2+p^2}$. Thus, if the circular molecule assumes a right-handed plectonemic shape in the presence of excess negative superhelical stress, the supercoil radius will decrease as a function of negative supercoiling if the superhelix pitch is maintained at a constant value, and cores located diametrically across the supercoil axis will be brought into closer proximity, increasing the chances of steric clashes. At very high levels of negative superhelical stress, the circular molecule may no longer be able to undergo a further writhe reduction due to the compact nature of the complex, and at this point the introduction of more negative helical tension

may cause a twist reduction in the core DNA from the termini. No decrease in the helical period of the terminal regions of the core DNA was visible for core 2 at σ=-0,03 in the linker areas. Such a twist reduction may only become apparent in more negatively supercoiled molecules.

It was suggested above that cores would tend to orient the side of DNA entry and exit towards the center of a circular molecule in order to minimize the bending requirements of the linker DNA. It thus appears that the average rotational orientation in a circular molecule saturated with nucleosome cores may be determined by a few rotationally strongly positioned cores. In a circular molecule occupied by several rotationally strongly positioned cores, the rotational orientation of the whole circle may be determined by the arrangement where the preferred orientation relative to the center of the circle of the individual cores is complied with as closely as possible in combination with avoiding unfavorable steric clashes across the central hole. This matter is experimentally addressed in the following chapter.

CHAPTER 6

6 Positioning of a nucleosome core on a rotationally permuted frame.

6.1 Introduction:

The effect of altering the angular orientation of the rotational frame of a core relative to those of adjacent cores on nucleosome core positioning, was investigated in two insertion constructs of pHP2.

It was shown in the previous chapter that position 93-104 (Crick) and 235-246 (Watson) constituted the linker DNA between cores 1 and 2 and between cores 2 and 3, respectively (see section 5.2.2). If the DNA in either of these regions is restricted with an enzyme producing 5'-protruding ends, the 3'-recessed ends extended with complementary bases to produce blunt ends, and the DNA molecule religated, the relative angular orientation of the two sides of the DNA molecule will be altered. The relative angular orientation of cores rotationally positioned on either side of the filled-in restriction site is expected to be similarly changed. Although one would ideally choose to rotate the DNA in the assigned linker areas to minimize the possibility of modifying the sequence-directed rotational placement of the cores (59, 210), the distribution of unique restriction sites did not allow this, and the most suitable sites were *Xba* I (76bp) and *Asp* 718 (292bp), both of which cleave the two DNA strands at positions staggered by 4bp in the 5'-direction. The conversion of these staggered ends to blunt ends with the Klenow fragment of DNA polymerase is therefore expected to produce 4bp insertions in the core DNA of cores 1 and 3 at positions removed by 1,7 and 4,5 helical turns from the proximal terminus of each core respectively. If the rotational orientations of cores 1-3 are conserved

relative to the bulk of the underlying DNA sequence, the filling-in of the *Xba* I site should rotate core 2 (and cores present to the right-hand side of core 2 on a linearized molecule) by 137° anti-clockwise relative to core 1 , and filling-in of the *Asp* 718 site should rotate core 3 (and other cores present on the right-hand side of core 3 on a linearized molecule) by 137° anti-clockwise relative to cores 1 and 2 (viewed down the length of the DNA molecule with core 3 proximal to the viewer). The investigation of nucleosome core positioning on linearized and supercoiled plasmids thus modified should give information on the influence of core-core interactions on nucleosome core positioning.

Two plasmids were constructed in this way. The first, denoted pHP2R, was obtained by religation of pHP2 with the *Xba* I site converted to blunt ends. This procedure should have introduced a 4bp insert, but nucleotide sequencing revealed the presence of only an additional 3bp, probably due to the exonucleolytic removal of the terminal (dG).(dC) base-pair by a contaminating nuclease present during one of the modifications steps. The positioning frame of core 2 (and all cores present on the right-hand side of core 2 on a linearized molecule) should therefore be rotated by 103° anti-clockwise relative to core 1. The partial nucleotide sequence of pHP2R is shown in figure 1.2.

The plasmid pHP2RR was obtained by inserting 4bp into the *Asp* 718 site of pHP2R. Thus, in pHP2RR, the rotational frame of core 2 is rotated by 103° anti-clockwise relative to core 1, and that of core 3 by 137° anti-clockwise relative to core 2. The partial nucleotide sequence of pHP2RR is shown in figure 1.3.

6.2 Positioning of nucleosome cores on linearized and supercoiled pHP2R:

Supercoiled (σ=-0,06) and linearized (*Pvu* II restricted) pHP2R were reconstituted into nucleosome cores at a molar octamer:DNA ratio of 11, and the placement of cores determined by DNase I digestion and primer extension as previously described (see sections 2.4.4, 2.6.2 and 2.7 of materials and methods) The autoradiograph of the gel is shown in figure 6.1.

A modulation with a period of approximately 10bp is visible in the DNase I cleavage pattern of both the linearized (lanes 5 and 11) and supercoiled (lanes 3 and 9) reconstituted molecules (indicated by solid triangles in figure 6.1), reflecting the presence of a rotationally positioned nucleosome core (55, 59, 200, 220). The enhanced cleavage at positions 101-104 in the free and reconstituted plasmid, both in the linearized and supercoiled DNA molecules, is absent at the corresponding nucleotides in pHP2 (see figure 5.2). It is not clear why insertion of 3 nucleotides after position 76 induced this hypersensitivity, but the phenomenon may be related to the presence of the TATA sequences (62, 85) in this region of the plasmid (see figure 2.2), or represents a decreased efficiency of primer extension across this region. Insufficient experimental evidence is, however, available to give a firm explanation. In order to accurately determine the placement of the nucleosome cores on pHP2R, a differential probability analysis was performed (see section 2.8.3). The results are presented in figures 6.2 and 6.5.

If the difference probability plot of linearized pHP2R is examined first (figure 6.2), maxima in the probability of DNase I cleavage are visible at positions 79, 91, 105, 116, 125(weak), 138, 146, 157 and 167 on the Crick strand, and at

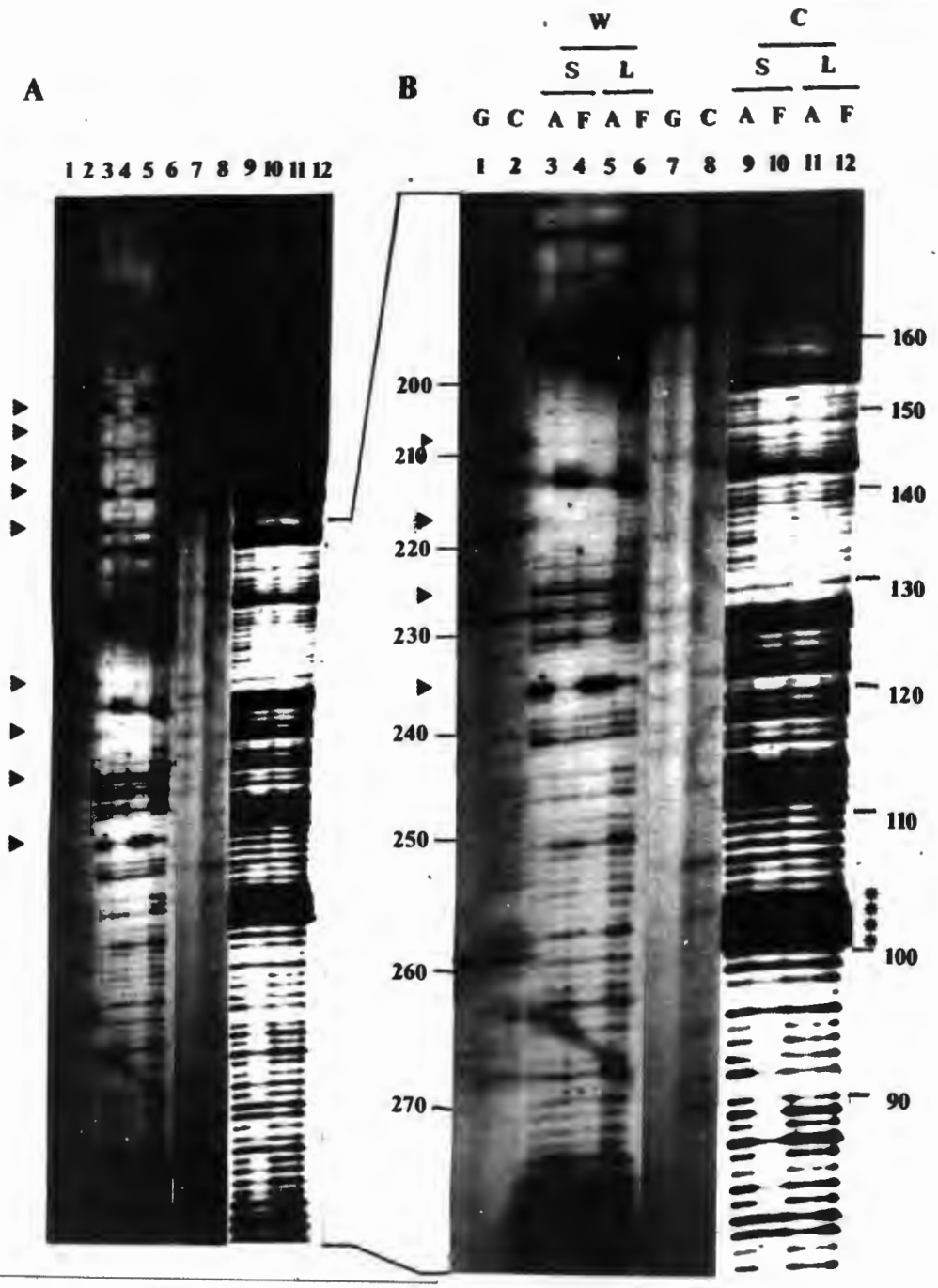

Figure 6.1. DNase I digestion of pHP2R reconstituted with the histone octamer at a molar octamer:DNA ratio of 11. The location of nucleosome cores on supercoiled (S; lanes 3, 4, 9 and 10) and linearized (L; lanes 5, 6, 11 and 12) pHP2R were determined by DNase I digestion of assembled (A; lanes 3, 5, 9 and 11) and free (F; lanes 4, 6, 10 and 12) plasmid followed by extension of 20mer primers annealed to either the Watson (W; lanes 3-6) or Crick (C; lanes 9-12) strand. The location of DNase I cleavage sites was identified relative to guanosine (G) or cytosine (C) chain terminating sequencing reactions of 20mer primers annealed to the Watson (lanes 1 and 2) or Crick (lanes 7 and 8) strands. The sequence positions of individual bands are shown in the margins of (B). The maxima in the modulation of DNase I cleavage with a period of approximately 10bp in reconstituted samples are indicated by triangles. The hypersensitive sites that appeared at positions 101-104 after insertion of 3bp into pHP2 are indicated by asterisks. Panel B is an enlarged section of panel A. An autoradiograph of the gel is shown. See text for detail.

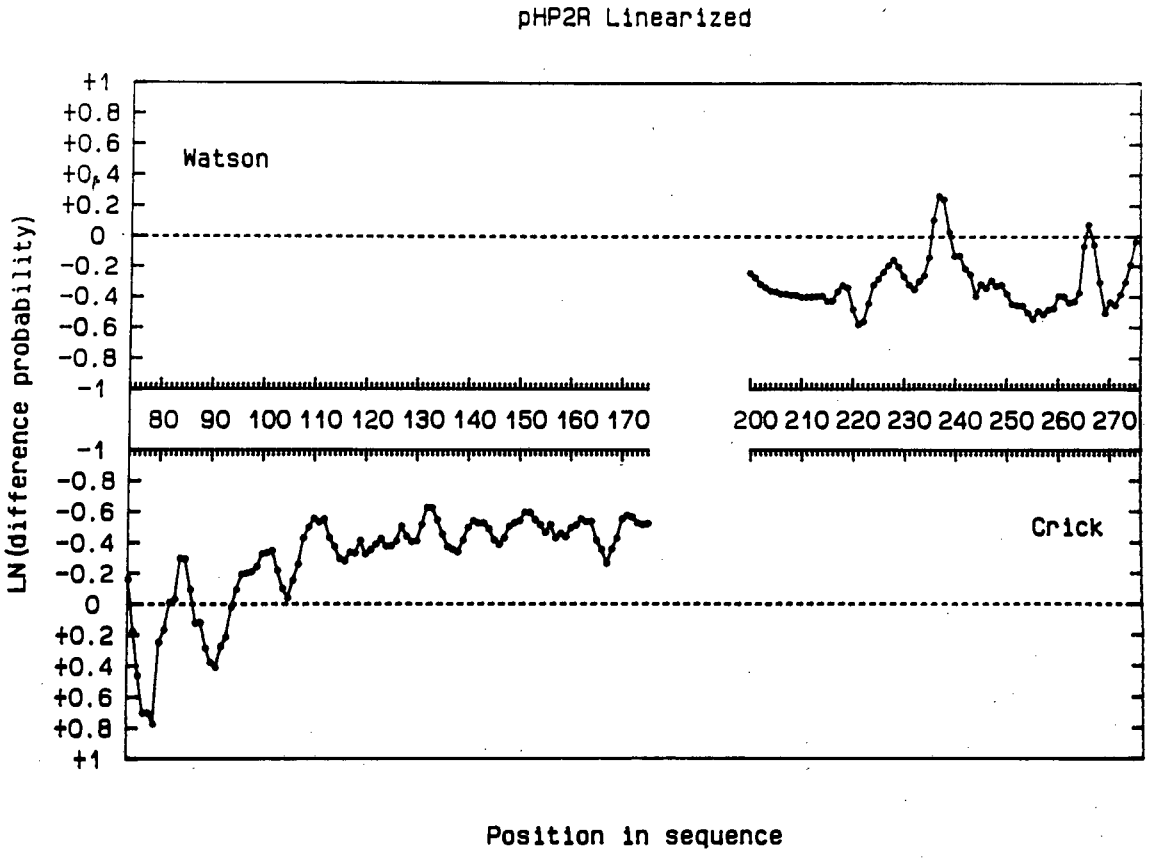

Figure 6.2. Probability difference of DNase I cleavage in linearized pHP2R reconstituted with the histone octamer at a molar octamer:DNA ratio of 11. The probability of cleavage, expressed as the natural logarithm of the difference between the rate of cleavage in reconstituted and free plasmid (see section 2.8.3 of materials and methods) is shown as a three point running average at different positions in the sequence of pHP2R. Cleavage data on the Watson and Crick strands are shown in the top and bottom sections of the figure respectively. See text for detail.

positions 218, 228 and 238 on the Watson strand. Note, however, due to numbering differences, these positions are not equivalent to those of pHP2. Nucleotides present on the right-hand side of position 79 in pHP2R are shifted by +3bp relative to the position of the corresponding nucleotide in pHP2. The nucleotides present at positions 77,78 and 79 in pHP2R are absent in pHP2. Since the only constant reference of rotational placement is the DNA sequence, it is necessary to convert the positions of maximal cleavage in pHP2R to the corresponding sequence positions in pHP2 by subtracting 3 from the sequence positions in pHP2R in order to compare the core placements in the two plasmids. The sites of maximal DNase I cleavage are then seen to correspond to positions 76, 88, 102, 113, 122, 135, 143, 154 and 164 on the Crick strand, and to positions 215, 225 and 235 on the Watson strand in the sequence of pHP2. In order to avoid confusion between the two numbering schemes, the sequence positions in pHP2R will be indicated by a subscript "r". The locations and probabilities of cleavage at each of these sites agree closely with those of the corresponding sites on pHP2 (see section 5.2.2), and thus, by analogy, positions 105_r-238_r can be assigned as the translational borders of a core incorporating the d(A-G).d(C-T) motif. In order to discriminate between nucleosome cores present on pHP2 and pHP2R, the latter cores will similarly be denoted with a subscript "r". The translational borders of core 2_r at positions 105_r (Crick) and 238_r (Watson) corresponds closely to the translational borders of core 2 (104 (Crick), 235 (Watson)). If the structural organisation of core 2 and 2_r is equivalent, as suggested by the similarities in the relative probabilities of cleavage at corresponding helical turns in the two cores, particularly on the right-hand side of the two cores (compare figures 6.2 and 5.3), the maximum in the cleavage periodicity at position 105_r of core 2_r is expected to correspond to site 130 of the mixed sequence cores (142). The dyad axis of core 2_r is therefore removed

by approximately 6 helical turns from this point, which corresponds to the maximum at position 167_r. This places the dyad axis of core 2_r, taking the 1bp stagger of DNase I cleavage across the minor groove into consideration (55, 57, 119, 142), at position 168_r. This position is exactly removed by 3bp (in terms of sequence positions) relative to the dyad axis of core 2, and it thus appears that the translational and rotational frames of core 2 and 2_r are equivalent relative to the base-pair sequence.

To ensure that no subtle differences in the positioning of cores 2 and 2_r are overlooked, a correlation analysis between the difference probability data of the two cores on linearized pHP2 and pHP2R was performed as previously explained (see section 2.8.3). The result is presented in figure 6.3.

Referring to this figure, it is seen that the difference in the probability of DNase I cleavage in the two cores is at a minimum when the corresponding nucleotides of the two cores are exactly aligned. This result indicates that the rotational and translational (*modulus* average helical period) position of cores 2 and 2_r is identical relative to the core sequences, demonstrating that the rotation of the positioning frame of core 2 by 103° anti-clockwise relative to that of core 1 did not cause core 2 to assume a different position. It therefore appears that the positioning of core 2 is determined by the sequence periodicity of the core DNA, and is not detectably influenced by adjacent cores. This is expected from the identical positions assumed by the core on a 1915bp linearized DNA molecule and a 337bp fragment (see section 5.5.1).

If the modulation in the periodicity of DNase I cleavage towards the left of core 2_r is considered next, positions of maximum cleavage are seen at positions 79_r and 91_r (see

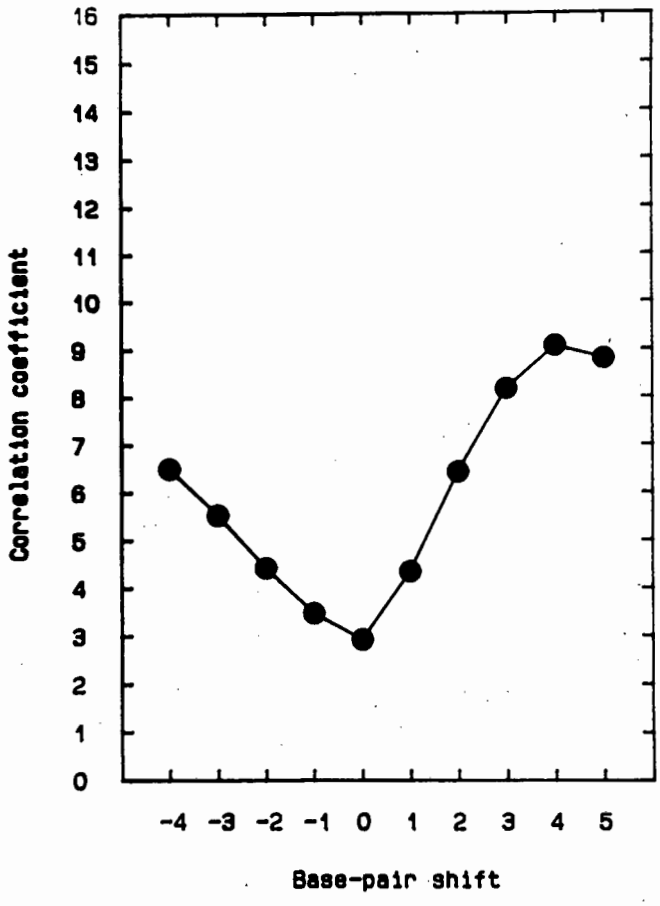

Figure 6.3. Correlation between the sites of maximum DNase I cleavage in core 2 on linearized pHP2 and pHP2R reconstituted at a molar octamer:DNA ratio of 11. The correlation coefficient, shown on the ordinate, was obtained by summation of the differences between the probability difference datasets of core 2 at different offsets, shown on the abscissa, of the dataset of linearized pHP2R relative to linearized pHP2. See text for detail.

figure 6.2). In order to examine the positions of cleavage by DNase I in this region, it is necessary to clearly identify the relevant individual nucleotides present at every position in this region of the sequence in pHP2 and pHP2R, which is presented in figure 6.4. The determined positions of maximal DNase I cleavage are indicated by asterisks above and below the relevant sequence positions in pHP2 and pHP2R in figure 6.4.

```
                          pHP2 sequence position
3' Crick                  *                                   *
         77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93
G  A  T  G  A  T  C  T  C  C  T  A  G  G  G  G  T  C  G  A
77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96
      *                                          *        5' Crick
                          pHP2R sequence position
```

Figure 6.4: Numbering of the sequences of pHP2 and pHP2R. Only the Crick strand is shown. The nucleotides present at positions 77-79 of pHP2R was introduced by filling-in of the *Xba* I site of pHP2, and are absent in pHP2. The sites of maximal DNase I cleavage in reconstituted pHP2 and pHP2R are indicated by the asterisks above and below the relevant sequence positions. See text for detail.

By referring to figure 6.4, it is clear that the sites of maximum cleavage at positions 81 and 93 in pHP2 did not shift by 3 positions to the right to appear at the identical nucleotides at positions 84_r and 96_r in pHP2R, but by 5 nucleotides to the left to appear at positions 79_r and 91_r. The direction of this shift is as expected if it represents DNase I cleavage in the exposed regions of the terminal two helical turns of a core that retained its position relative to the bulk of the DNA sequence present to the left of position 76. It was suggested that the bending preferences of the terminal core sequences exerted a lesser influence on the rotational placement of the nucleosome core (243), as

demonstrated by Neubauer *et al.* (175) and Piña *et al.* (193) where the rotational orientation of a core was retained relative to the bulk of the DNA sequence in angularly permuted positioning frames. The introduction of a short insert close to the terminus of the core DNA (175, 193) is therefore expected to rotate the formerly protected phosphodiester bonds in the terminal helical turns to a position where they face away from the octamer surface and become susceptible to DNase I cleavage. However, the shift in the positions of maximal DNase I cleavage in the terminal region of core 1 is offset by -2bp relative to that expected from the 3bp insert. This might reflect a change in the rotational placement of core 1_r relative to that of core 1 due to core-core interactions, but the paucity of the evidence precludes such a deduction to be made with confidence. The sites of maximal DNase I cleavage over the whole length of core 1 and core 1_r must be compared to exclude the possibility that the shift of the sites of maximal DNase I cleavage in core 1_r is due to a change in the helical period of the terminal core DNA (which may be caused by core-core interactions). The shift of these latter sites relative to those present on the 5' side of the Crick strand in pHP2R, however, confirms the assignment of position 91_r (93 in pHP2) and 105_r (102 in pHP2) to the right- and left-hand borders of two adjacent cores, as suggested in the previous chapter.

In the region of the sequence where core 3 resolved in a mixture of rotational frames on pHP2 (see section 5.2.2), no consistent periodicity in the probability of DNase I cleavage is visible in linearized pHP2R (see figure 6.2). This suggests that core 3_r continued to be rotationally weakly positioned after rotation of the sequence by 103° anti-clockwise relative to the positioning frame of core 1.

If the modulation in the probability of DNase I cleavage in the reconstituted supercoiled pHP2R is examined next (figure 6.5), maxima are observed at positions 78_r, 91_r, 96_r, 108_r, 119_r, 127_r (weak), 138_r, 148_r (weak) and 156_r on the Crick strand, and at positions 219_r, 229_r and 237_r on the Watson strand. In order to aid a comparison of the sites of maximal cleavage in core 2_r in linearized and supercoiled pHP2R, the sites are listed as linearized-supercoiled: 79_r-78_r, 91_r-91_r, 105_r-108_r, 116_r-119_r, 125_r-127_r, 138_r-138_r, 146_r-148_r and 157_r-156_r, 218_r-219_r, 228_r-229_r, 238_r-237_r. An examination of this series reveals that, whereas the sites of maximal cleavage in core 1_r on the supercoiled molecule were conserved relative to the sequence, a decided shift, which is on average to the right, occurred in the sites of core 2_r on the supercoiled molecule. Since subtle differences in the location and spread of peaks may be overlooked by such a one-on-one comparison of the peak maxima, a correlation analysis was performed on the difference probability data of core 2_r on the supercoiled and linearized pHP2R, the result of which is presented in figure 6.6.

Referring to the average correlation cf the two strands in figure 6.6 (datapoints indicated by circles), it is seen that the curve displays two local minima. This is due to two groups of peaks which are aligned at different base-pair offsets in the two difference probability datasets. To analyse this feature further, the correlation was performed on the peaks present on the Crick strand (triangles in figure 6.6) and on the Watson strand (squares in figure 6.6) individually. Referring to the relevant curves in figure 6.6, it is seen that the positions of maximal cleavage on the Crick strand in supercoiled pHP2R are shifted by 3 nucleotides to the right relative to those on the linearized pHP2R, and that the peaks on the Watson strand are shifted by one nucleotide to the left on

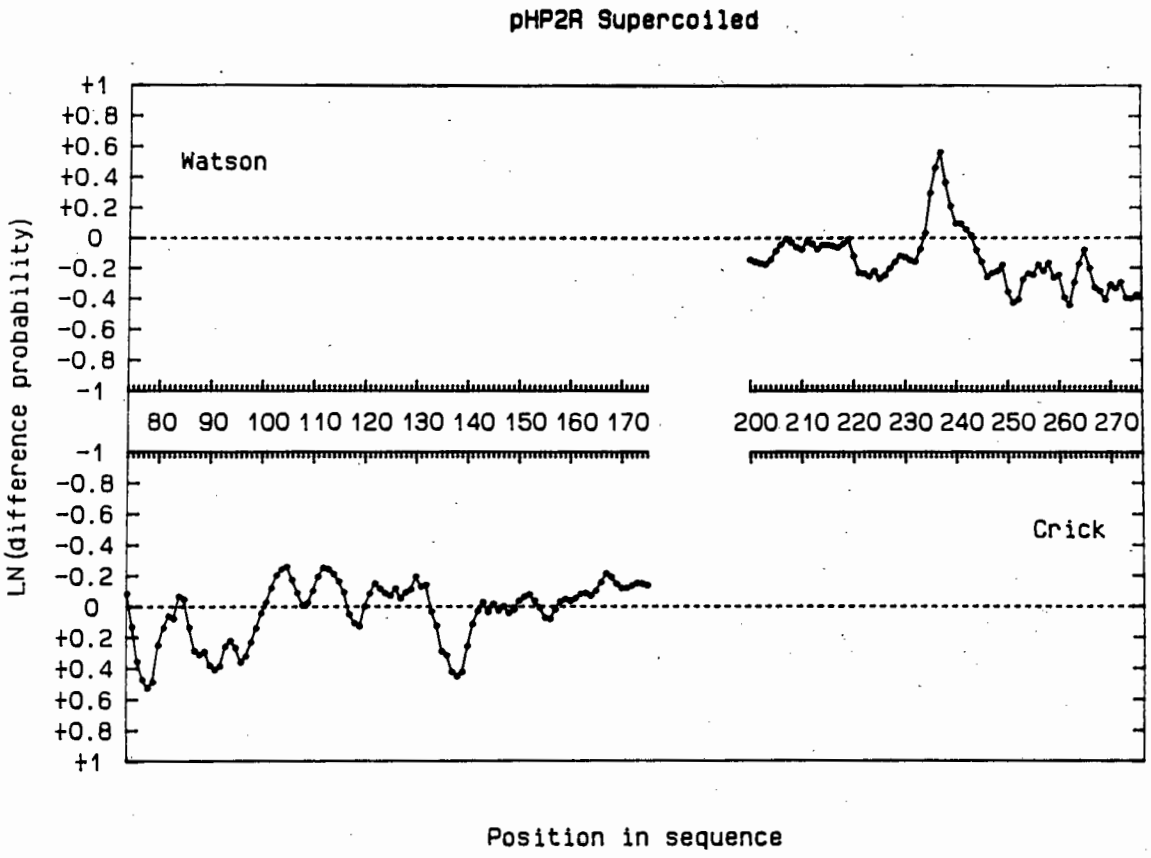

Figure 6.5. Probability difference of DNase I cleavage in supercoiled pHP2R reconstituted with the histone octamer at a molar octamer:DNA ratio of 11. The probability of cleavage, expressed as the natural logarithm of the difference between the rate of cleavage in reconstituted and free plasmid (see section 2.8.3 of materials and methods) is shown as a three point running average at different positions in the sequence of pHP2R. Cleavage data on the Watson and Crick strands are shown in the top and bottom sections of the figure respectively. See text for detail.

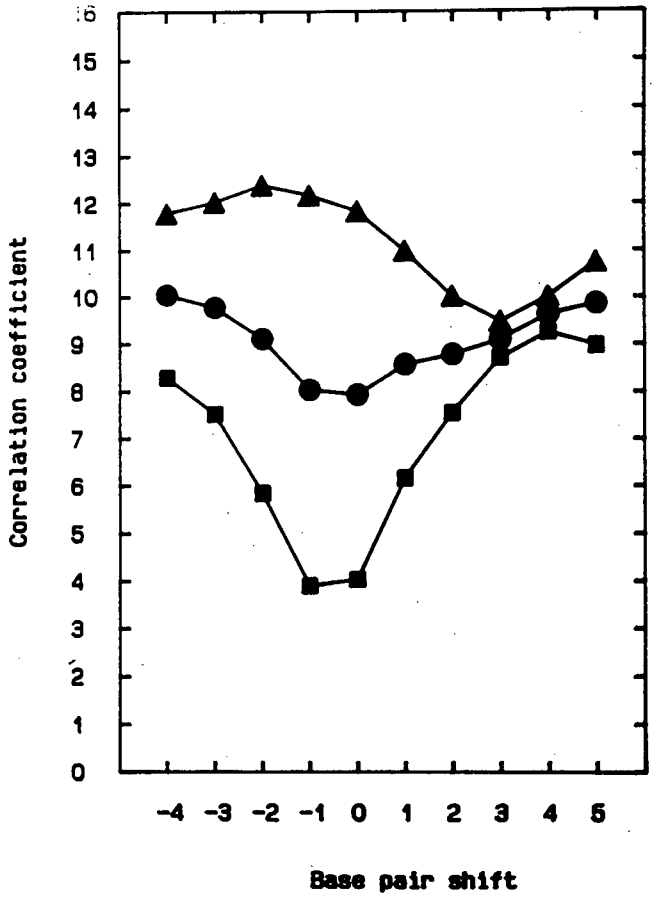

Figure 6.6. Correlation between the sites of maximum DNase I cleavage in core 2 on supercoiled and linearized pHP2R reconstituted at a molar octamer:DNA ratio of 11. The correlation coefficient, shown on the ordinate, was obtained by summation of the differences between the probability difference datasets of core 2 at different offsets, shown on the abscissa, of the dataset of supercoiled relative to linearized pHP2R (circles). The correlation between the sites of maximum cleavage on the 3′ side of the Watson strand (squares) or 5′ side of the Crick strand (triangles) at similar offsets are also shown. See text for detail.

supercoiled pHP2R relative to linearized pHP2R. Care should be taken in interpreting this shift, since, in the correlation analysis, the small offset of two peaks of large amplitude will contribute more significantly to a negative signal than a large offset between two peaks of small amplitude. However, the assigned location of the large peak present at position 237_r on the Watson strand is expected to be less prone to background noise compared to the weakly defined peaks at positions 219_r and 229_r (see figure 6.5). The shift of -1bp suggested by the correlation analysis for the peaks on the Watson strand would therefore appear justified.

The result of the correlation analysis therefore suggests that the position of core 2_r did not shift on supercoiled relative to linearized pHP2R. The difference in the direction of the shift on the two sides of core 2_r is rather an indication of an increase in the twist of the terminal core DNA. This suggestion is supported by the decrease in the offset of the sites of maximal DNase I cleavage in corresponding helical turns which is visible, on average, closer to the dyad of the two cores. For instance, the difference in the location of the terminal helical turn on the left-hand side of the core in supercoiled and linearized pHP2R is 3bp, the difference in the second helical turn is 2bp, the third is also 2bp, and the fourth set of peaks align perfectly.

Apart from this change in helical repeat visible in the terminal core DNA, a site of maximal DNase I cleavage is also visible at position 96_r in the supercoiled molecule (see figure 6.5), a site which is absent at the corresponding position in linearized pHP2R (see figure 6.2). If the sequence in this region of pHP2R is identified (refer to figure 1.2), it is seen than the maximum at position 96_r constituted cleavage by DNase I on the 3'-side of the third

thymidine in the sequence TTATAA. A similar increase in the probability of cleavage in this sequence was not observed in the reconstituted supercoiled molecule where the relative rotational orientation of the positioning frames of cores 1 and 2 was not modified (compare figures 5.3 and 5.5).

It is well established that the TpA base-step is more responsive to torsional stress in DNA than other base-steps (62, 85). It was further proposed that the rate of DNase I cleavage at a particular phosphodiester bond is sensitive to the width of the minor groove (57, 58, 71), alteration of which can modify the rate of cleavage at a specific nucleotide in the sequence (71). However, the width of the minor groove in AT-rich sequences is usually narrower than that of a mixed sequence helix (59, reviewed in 61), and in this regard the groove width in the region of the TATA sequence is expected to become even narrower under the influence of the positive torsional stress visible in the terminal helical turns of core 2_r. However, the minor groove may be wide in the negatively supercoiled free plasmid, and become of an intermediate width preferred by DNase I after reconstitution of the molecule, thus producing the site on enhanced cleavage in the linker between cores 1_r and 2_r. Whatever the precise mechanism of this enhancement, its appearance in the TTATAA sequence in the reconstituted supercoiled molecule and its absence in the reconstituted linearized molecule would suggest that the linker is under torsional stress, in agreement with the shift observed in the sites maximally exposed to DNase I cleavage on the surface of the octamer on supercoiled pHP2R.

No significant twist change is visible in the terminal region of core 1_r on the supercoiled molecule, where the sites of maximal cleavage at positions 79_r and 91_r agree well with those on the linearized molecule. Core 3_r

similarly seems to persist in its positioning in a mixture of rotational frames on supercoiled pHP2R.

6.3 Positioning of nucleosome cores on linearized and supercoiled pHP2RR:

Supercoiled (σ=-0,06) and linearized (*Pvu* II cut) pHP2RR were reconstituted into nucleosome cores at a molar octamer:DNA ratio of 11 by the method of salt dialysis. The reconstituted and free DNA samples were digested with DNase I and primer extended as described before (see sections 2.4.4, 2.6.2 and 2.7 of materials and methods). The result is shown in figure 6.7.

The modulation with a period of approximately 10bp in the cleavage pattern of DNase I in the reconstituted molecules (lanes 3, 5, 9 and 11), indicative of rotationally positioned nucleosome cores (55, 59, 200, 220), is indicated by the solid triangles in the margins of figure 6.7. This periodicity in cleavage is particularly pronounced over a section of the d(A-G).d(C-T) motif, were clear maxima are visible at positions 221 and 212 (see figure 6.7 B). In order to obtain an accurate indication of the placement of the nucleosome cores on pHP2RR, a differential probability analysis was performed as before (see section 2.8.3 of materials and methods). The result is shown in figure 6.8 and 6.10. Keeping with the convention of the previous section, nucleosome cores present on, and sequence positions in pHP2RR (which are identical to those of pHP2R in the region of interest) are indicated by a subscript "rr"

On linearized pHP2RR (figure 6.8) maxima in the probability of DNase I cleavage are visible at positions 79_{rr}, 93_{rr}, 107_{rr}, 114_{rr}, 125_{rr}, 135_{rr}, 144_{rr}, 154_{rr} and 163_{rr} on the Crick strand, and at positions 209_{rr}, 219_{rr}, 228_{rr}, 233_{rr}, 238_{rr}, 243_{rr}, 248_{rr}, 257_{rr} and 265_{rr} on the Watson strand.

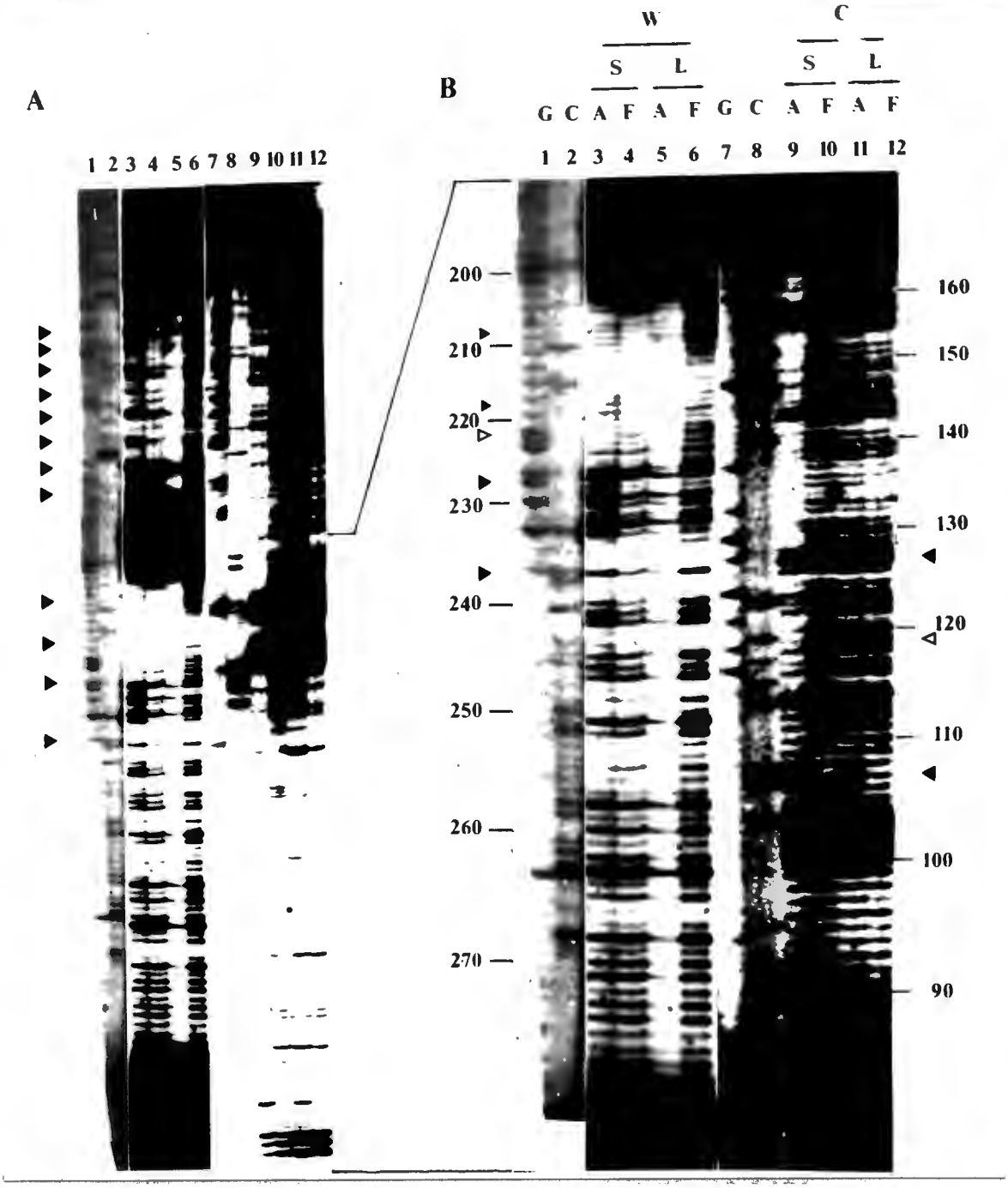

Figure 6.7. DNase I digestion of pHP2RR reconstituted with the histone octamer at a molar octamer:DNA ratio of 11. The location of nucleosome cores on supercoiled (S; lanes 3, 4, 9 and 10) and linearized (L; lanes 5, 6, 11 and 12) pHP2RR were determined by DNase I digestion of assembled (A; lanes 3, 5, 9 and 11) and free (F; lanes 4, 6, 10 and 12) plasmid followed by extension of 20mer primers annealed to either the Watson (W; lanes 3-6) or Crick (C; lanes 9-12) strand. The location of DNase I cleavage sites was identified relative to guanosine (G) or cytosine (C) chain terminating sequencing reactions of 20mer primers annealed to the Watson (lanes 1 and 2) or Crick (lanes 7 and 8) strands. The sequence positions of individual bands are shown in the margins of (B). The maxima in the modulation of DNase I cleavage with a period of approximately 10bp in reconstituted samples are indicated by solid triangles. Positions where DNase I cleaved at a lower frequency in the reconstituted samples relative to the free DNA are indicated by open triangles. Panel B is an enlarged section of panel A. An autoradiograph of the gel is shown. See text for detail.

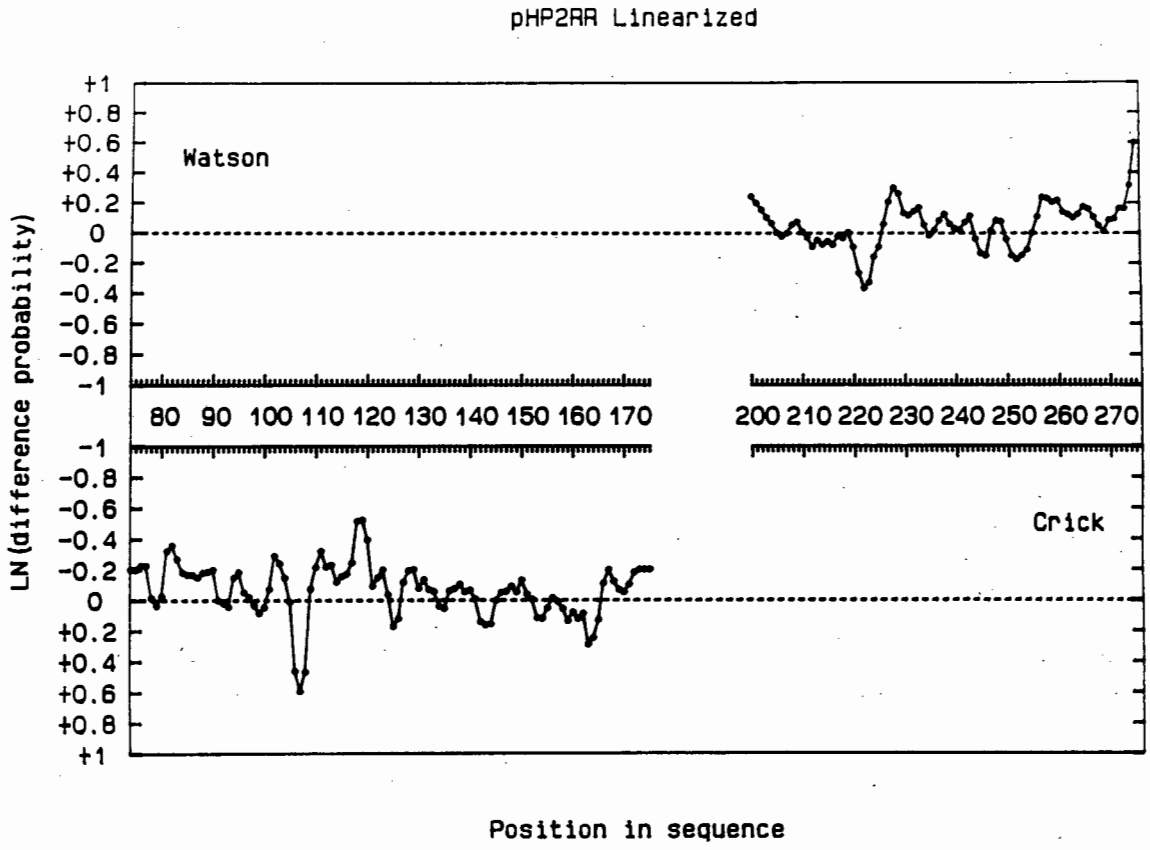

Figure 6.8. Probability difference of DNase I cleavage in linearized pHP2RR reconstituted with the histone octamer at a molar octamer:DNA ratio of 11. The probability of cleavage, expressed as the natural logarithm of the difference between the rate of cleavage in reconstituted and free plasmid (see section 2.8.3 of materials and methods) is shown as a three point running average at different positions in the sequence of pHP2RR. Cleavage data on the Watson and Crick strands are shown in the top and bottom sections of the figure respectively. See text for detail.

A conspicuous feature in the probability difference plot of the linearized pHP2RR is the appearance of split peaks, particularly pronounced on the Watson strand, where the maxima between positions 228_{rr} and 248_{rr} are separated by 5bp on average. If a nucleosome core is present in a mixture of two rotational frames differing by half a helical turn, two series of peaks, the one offset by 5bp relative to the other, should be present over the whole length of DNA protected by the histone octamer. The magnitude of the split peaks at positions 228_{rr}, 233_{rr}, 238_{rr} and 243_{rr} are similar, whereas weak local maxima present between the identified peaks on the Crick strand (see figure 6.8) are significantly less pronounced than the adjacent peaks, suggesting that these represent background noise rather than true peaks. Since the split peaks are present on only a section of core 2_{rr} DNA, it appears unlikely that core 2_{rr} positions in two rotational frames differing by approximately half a helical turn. It also appears unlikely than the right-hand terminal section of the core DNA is twisted by approximately 180° relative to the remainder of the core DNA in half of the core population. The appearance of the split peaks most likely reflect the invasion of core 2_{rr} DNA by core 3_{rr}, which was previously shown to resolve in a mixture of rotational settings, in a fraction of the linearized pHP2RR molecules. This change in the translational frame of core 3_{rr} may have been caused by the insertion of 4bp after position 296_{rr} in pHP2RR (see figure 1.3). This position is well within the core DNA of core 3_{rr}, and the rotation of the terminal 4,5 helical turns in core 3_{rr} may have aligned a sequence periodicity present in the terminal helical turns of core 3_{rr} with a weak periodicity present in the remainder of the core DNA, thus allowing a wider range of translational settings to core 3_{rr} where the bending preference of the DNA can be complied with along the whole length of the core. The peaks present at positions 228_{rr} and 238_{rr} can therefore be assigned to the

continuation of the 10bp periodicity defined by the peaks present at positions 209_{rr} and 219_{rr} and were thus produced by core 2_{rr}. Core 3_{rr} may have contributed these maxima and probably gave rise to the peaks at positions 233_{rr}, 243_{rr} and 248_{rr}. The proposed shift in the average location of core 3_{rr} suggests that the translational setting of this core was not simply induced by the presence of the strongly positioned core 2_{rr}.

In order to establish whether the rotation of the positioning frames of both core 1 and 3 relative to that of core 2 caused a shift in the location of the latter core, a correlation analysis between the difference probability data of core 2 on linearized pHP2 and core 2_{rr} on linearized pHP2RR was performed. The result is presented in figure 6.9.

Referring to this figure, a clear minimum at a base-pair offset of 0 is seen. This indicates that the rotational and translational (*modulus* average helical period) position of core 2_{rr} is identical to that of core 2 on linearized molecules, and, similar to the result obtained for core 2_r, suggests that rotation of a section of the positioning frame of core 3_r relative to that of core 2_r, did not induce a shift in the positioning of the latter core.

If the probability of DNase I cleavage in reconstituted supercoiled pHP2RR is considered next (figure 6.10), maxima are visible at positions 79_{rr}, 91_{rr}, 107_{rr}, 114_{rr}, 126_{rr}, 136_{rr}, 143_{rr}, 154_{rr} and 164_{rr} on the Crick strand, and at positions 209_{rr}, 218_{rr}, 228_{rr}, 232_{rr}, 238_{rr}, 242_{rr}, 247_{rr} and 258_{rr} on the Watson strand.

The split peaks discussed above are again visible at positions 228_{rr}, 232_{rr}, 238_{rr} and 242_{rr} on the Watson strand, but can now clearly be seen to represent a feature

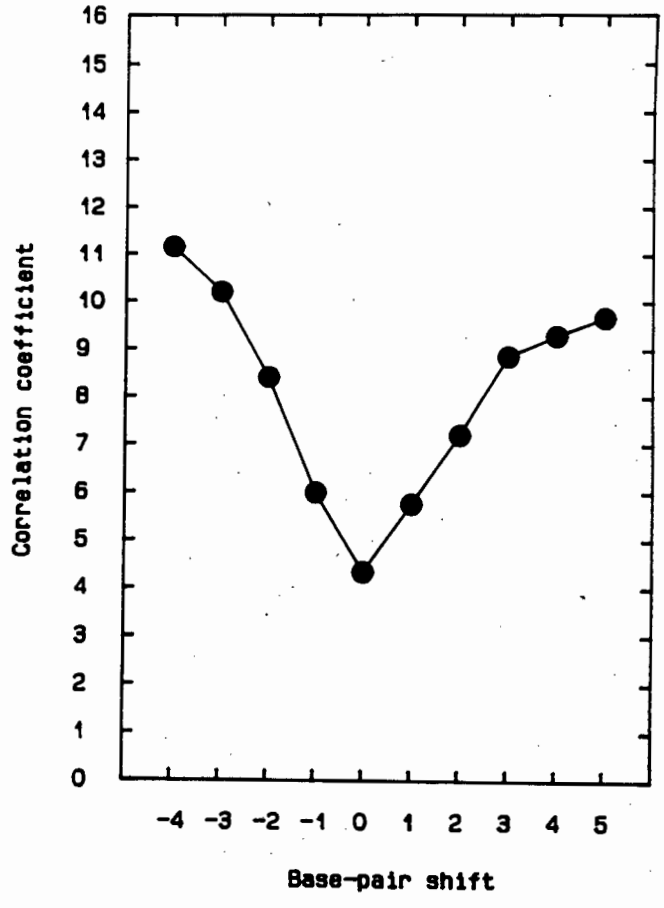

Figure 6.9. Correlation between the sites of maximum DNase I cleavage in core 2 on linearized pHP2 and pHP2RR reconstituted at a molar octamer:DNA ratio of 11. The correlation coefficient, shown on the ordinate, was obtained by summation of the differences between the probability difference datasets of core 2 at different offsets, shown on the abscissa, of the dataset of linearized pHP2RR relative to that of linearized pHP2. See text for detail.

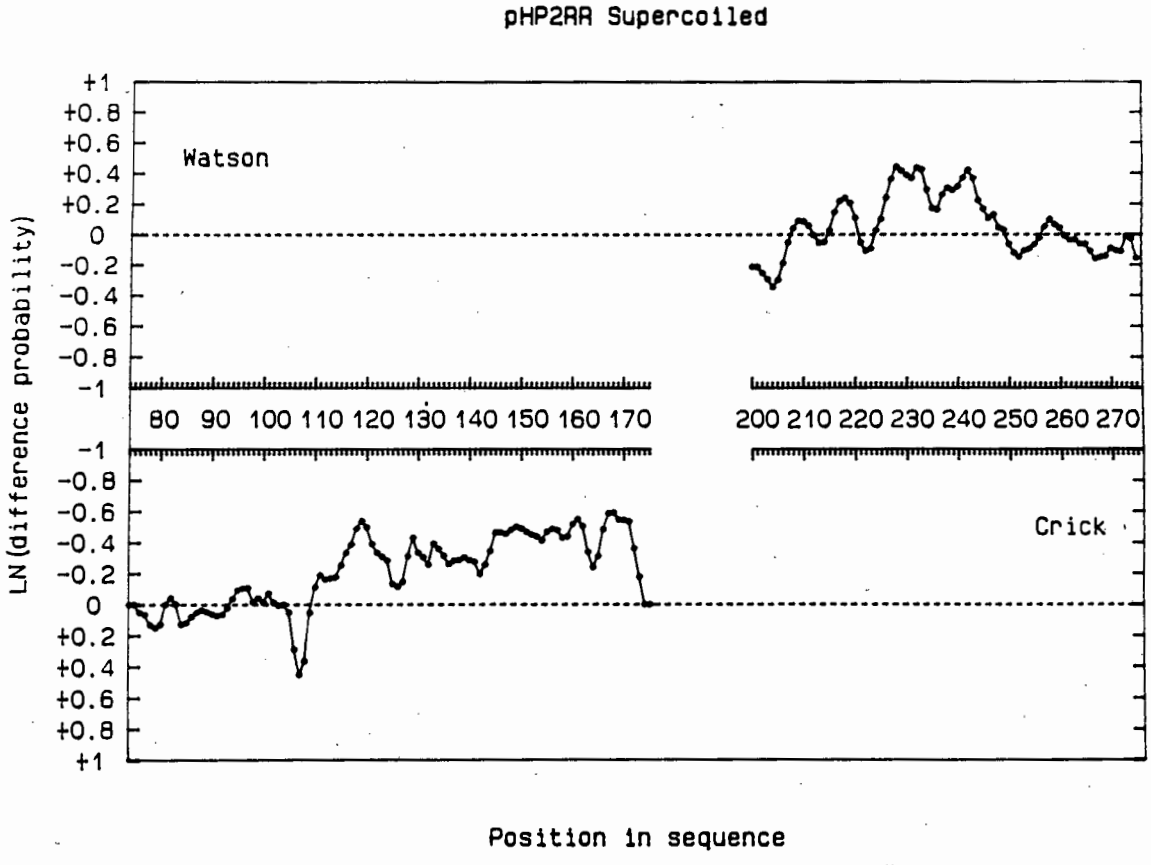

Figure 6.10. Probability difference of DNase I cleavage in supercoiled pHP2RR reconstituted with the histone octamer at a molar octamer:DNA ratio of 11. The probability of cleavage, expressed as the natural logarithm of the difference between the rate of cleavage in reconstituted and free plasmid (see section 2.8.3 of materials and methods) is shown as a three point running average at different positions in the sequence of pHP2RR. Cleavage data on the Watson and Crick strands are shown in the top and bottom sections of the figure respectively. See text for detail.

isolated to the right-hand side of core 2_{rr}. At positions 209_{rr} and 218_{rr} smooth, single peaks are visible. It therefore appears that the translational placement of core 3_{rr} on supercoiled pHP2RR still extends to within the positioning sequence of core 2_{rr}. In order to determine whether the positioning of core 2_{rr} has changed on the supercoiled pHP2RR relative to that of the corresponding core on the linearized plasmid, a correlation analysis was performed between the two difference probability data sets. The result is shown in figure 6.11.

In figure 6.11 the correlation between the probability difference data of the Watson and Crick stands of the two datasets are shown for comparison with the corresponding correlation curves of pHP2R (see figure 6.6). Examining the curve relating the effect of alternate alignments of the peaks present on the Crick strand first (triangles in figure 6.11), a rather wide through is visible. A minimum value is present at a base-pair offset of -1, but due to the very small differences in the values at offsets of -1 and 0, is probably not significant. The pronounced minimum present at an offset of +3 in the correlation of the Crick strand of supercoiled and linearized pHP2 (see section 6.2) is certainly not evident. This indicates that the change in helical period present in the left-hand side of core 2_r on supercoiled pHP2R, is absent after rotation of a section of the positioning frame of core 3_r. In this regard the enhanced cleavage present at position 96_r is also seen to be absent at the corresponding position in supercoiled pHP2RR. The difference probability data of the Watson stand (squares in figure 6.11) on the supercoiled and linearized pHP2RR is seen to be well aligned at a base-pair offset of 0.

It is clear from the minimum observed at a base-pair offset of 0 in the average correlation of the two datasets (circles in figure 6.11) that core 2_{rr} positions in identical

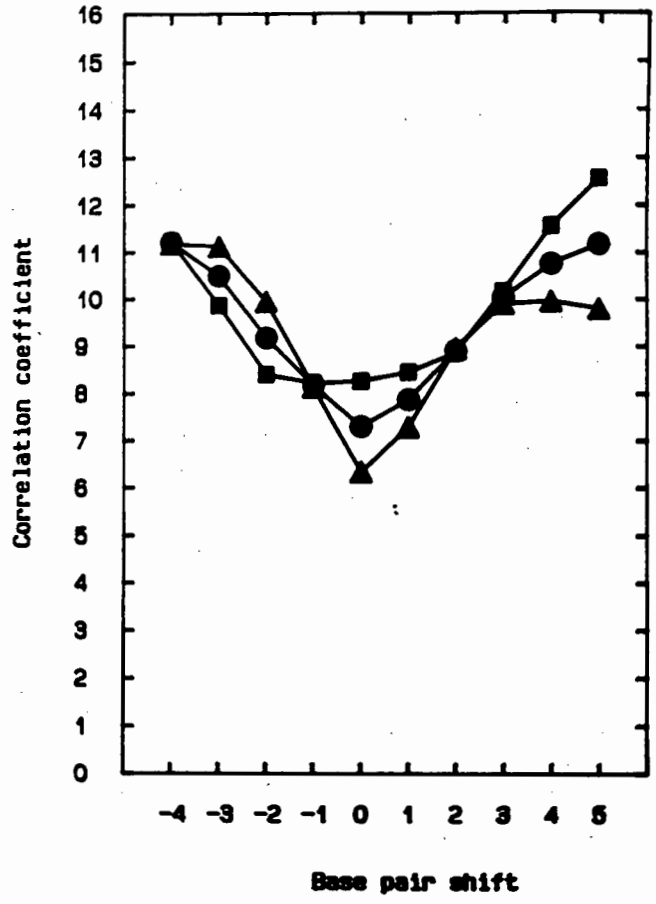

Figure 6.11. Correlation between the sites of maximum DNase I cleavage in core 2 on supercoiled and linearized pHP2RR reconstituted at a molar octamer:DNA ratio of 11. The correlation coefficient, shown on the ordinate, was obtained by summation of the differences between the probability difference datasets of core 2 at different offsets, shown on the abscissa, of the dataset of supercoiled relative to linearized pHP2RR (circles). The correlation between the sites of maximum cleavage on the 3′ side of the Watson strand (squares) or 5′ side of the Crick strand (triangles) at similar offsets are also shown. See text for detail.

rotational and translational (*modulus* average helical period) frames on supercoiled and linearized pHP2RR, and that the rotation of a section of the positioning frame of core 3_{rr} did not cause a shift in these positions.

6.4 Discussion:

It was shown above that on linearized pHP2R, where the rotational frame of core 2 was rotated by 103° anti-clockwise relative to that of 1 (see figure 6.12), core 2_r positioned in a location identical to that of the corresponding core on the unmodified plasmid. A shift in the sites of maximal DNase I cleavage in the terminal two helical turns of core 1_r was visible suggesting a shift in the rotational frame of core 1_r relative to that of core 2_r beyond that expected from the rotation between the sequences of these two frames. However, available experimental evidence was insufficient to provide firm grounds for such a claim. It was clear, however, that core 2_r retained its rotational and translational (*modulus* average helical period) frame, irrespective of the rotation between the positioning frames of cores 1 and 2. On the supercoiled molecule, the conservation of the positioning frame of core 2 persisted, although an increase in the twist in the terminal helical turns on either side of core 2_r was visible. This twist change must represent a steric clash between rotated core 2 and adjacent cores or cores diametrically placed opposite the DNA circle, and was not due to the supercoiled nature of the DNA molecule since similar twist changes were absent in assembled supercoiled pHP2. The finding that rotation of a section of the positioning frame of core 3 alleviated this overwinding of the DNA in the terminal regions of core 2_r suggests that core 3 may have contributed to the steric clash which caused a stable increase in the angle of rotation between cores 1 and 2, and cores 2 and 3. However,

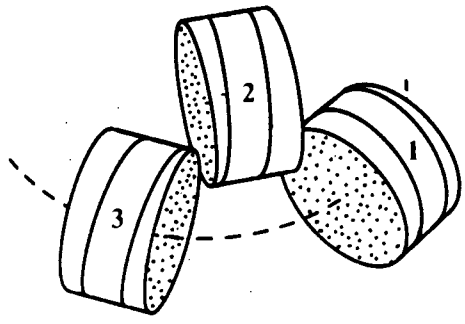

Figure 6.12. Schematic representation of the rotation of the nucleosome cores in plasmid pHP2RR. Core 2 is rotated by 103° anti-clockwise relative to core 1, and core 3 by 137° anti-clockwise relative to core 2. The relative spatial arrangement of the three cores shown in the figure is for illustrative purposes only.

in the absence of further experimental evidence on the spatial structure of the relevant supercoiled molecules, no consistent explanation for this phenomenon is possible.

The absence of any adjustment in the rotational and translational positioning of core 2 indicates that nucleosome cores can maintain their position relative to the underlying DNA sequence, irrespective of the orientation of this sequence relative to the positioning frames of adjacent cores. The results further also suggest that the orientation of the positioning frame of a core relative to the center of a circular molecule may make little difference to the positioning of the relevant core which may influence the rotational orientation of the circle itself. In this regard it was shown that core 2 is extraordinarily stably positioned relative to the sequence-directed position, which may indicate that the precise positioning of core 2 is required for biological function. This may include the correct exposure of specific sequences to the sequence recognition or structure recognition domains of DNA-binding proteins, as shown for the nucleosomes present on the *PHO5* (6, 65) and *MMTV LTR* (193, 194) promoter regions. This point is again raised in the next chapter.

CHAPTER 7

7 Discussion.

7.1 Introduction:

In the previous chapters results on the influence of negative supercoiling on the formation and positioning of nucleosome cores were presented. Here, these results are discussed in terms of the structural organisation of chromatin into topologically constrained loops in the interphase nucleus (see section 1.3), and genetic processes identified where the demonstrated response or absence thereof of nucleosome cores to negative supercoiling may be biologically significant. Before proceeding with such a discussion, it would be informative to analyze the sequences of the positioning frames of cores 1-3 in terms of suggested positioning determinants (37, 55, 56, 59, 210, 211; see also section 1.4.4), and determine how closely these account for the positioning of the nucleosome cores in this study.

7.2 Sequence-directed positioning of cores 1-3:

A nucleosome core was shown to adopt a precisely defined rotational and translational position on a section of the H1-H4 histone gene spacer of *Psammechinus miliaris*, corresponding to nucleotides 104-235 of pHP2 (see section 5.2). On either side of this core, the right- and left-hand borders of adjacent cores were visible at positions 93 and 248 respectively. What determined the positioning of these nucleosome cores?

It was suggested that the nucleosome core positions on a side of the DNA helix where the DNA bending required in wrapping the DNA around the octamer places base-steps with similar bending properties, repeated in phase with the

helical period of core DNA, at the preferred angular orientation of these steps, as far as this complies with the overall bending preference of the molecule (59, 210). It was further shown that the anisotropic flexibility of DNA alone, defined by the periodic distribution of particular di- and trinucleotides in phase with the helical period, can account for the rotational placement of several nucleosome cores on short DNA fragments *in vitro* (37, 55, 56, 61, 117). In order to determine whether the rotational placement of core 2 is similarly determined by the anisotropic flexibility of the DNA, the closeness with which the sequence periodicity of successive 125bp stretches matched the structural requirements of the core (59, 210) was calculated. At least two algorithms exist with which to perform this calculation (37, 55). It was decided to use the algorithm of Drew and Calladine (55), rather than that of Calladine and Drew (37), since the former is based solely on the statistical distribution of di- and trinucleotides in chicken core sequences (210), and makes no assumptions on the structural detail of individual base-steps (55, 56). The rotational algorithm was applied to a section of the sequence of pHP2 extending from 1733bp, through the numbering origin at position 1915/0, to position 542, as described by Drew and Calladine (55). This length of DNA corresponds to locations of the dyad axis (nucleotide 63 in the 125bp testing "window") from position 1795 to 480. The result is presented in figure 7.1.

It is clear from figure 7.1 that local areas of significantly differing predicted anisotropic flexibility occur in the length of DNA shown. For instance, in the regions of 100-140, 175-225 and 405-435 a large amplitude in the signal is observed. The particularly pronounced central area displays a range of approximately 7 units of likelihood (see reference 55), which is more than observed in any other published rotational placement plot (55, 56, 117). These

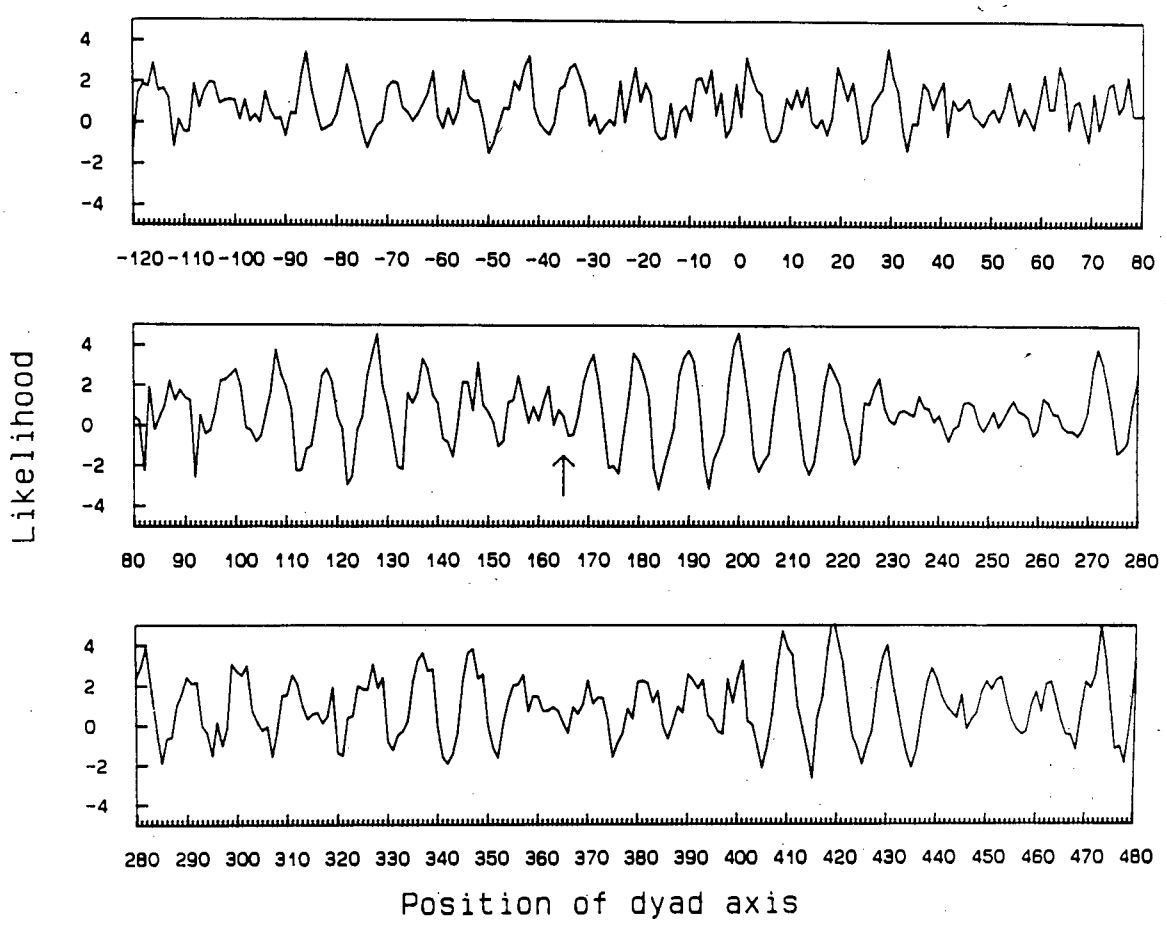

Figure 7.1. Predicted rotational positions of nucleosome cores on pHP2. The likelihood of the location of the core dyad at different positions in the sequence of pHP2 is indicated as the likelihood times RT (0,616 at room temperature) and varies from -4 for a favorable rotational setting to +4 for a bad one (after Drew and Calladine (55)) The location of the dyad of core 2 is indicated by the arrow at position 165. To maintain numbering consistency, the sequence positions beyond the numbering origin are indicated by negative numbers. To convert these to the sequence position in the plasmid, add 1915. See text for detail.

regions correspond to areas in the sequence where di- and trinucleotides of opposite bending preference are distributed approximately out of phase at a period equal to the helical repeat of core DNA (defined relative to the surface normal). In other areas of the plot the signal appears more irregular and correspond to 125bp sequences with little bending preference.

Core 2 was shown to be located between positions 104 (Crick) and 235 (Watson), with the dyad centered at position 165, indicated by an arrow in figure 7.1. It is clear that the experimentally determined location of the dyad axis of core 2 corresponds very closely with a minimum in the plot of rotational preferences. This would indicate that the rotational orientation of the octamer in core 2 relative to the DNA helix places di- and trinucleotides that prefer a wide minor groove at positions in the core where the minor groove points away from the octamer surface, and di- and trinucleotide preferring a narrow minor groove at locations where the minor groove is orientated towards the octamer surface. In order to identify these base-steps, the DNA sequence of core 2 is presented in figure 7.2 with the positions where the minor groove was found to point away from the octamer surface (see sections 5.2.2 and 5.2.3) identified with solid circles above the sequence.

It is clear from the location of the positions in the sequence where the minor groove points away from the octamer surface that these correspond to regions that are, on average G+C-rich, whereas positions where the minor groove points towards the octamer surface are, on average, A+T-rich. A structural basis for this phenomenon was proposed by Drew and Travers (59) and discussed in section 1.4.4.2. The preferred rotational orientation of the DNA in core 2 relative to the sequence periodicity is therefore in agreement with the principle that the

anisotropic flexibility of the DNA constitutes the rotational determinant of core placement. However, referring to figure 7.1 it would seem that areas on either side of the location of the dyad of core 2 have even lower minima, suggesting that the sequence periodicities in these areas comply even closer with the structural requirements of the core. It would be justified to ask why core 2 did not occupy either of these flanking regions? As stressed by Drew and McCall (56), the translational placement of a core cannot be deduced from a rotational preference plot such as

```
          100     •     110     •     120     •     130     •     140
GCTTATAATC  ATCCTTATAC  ACGCGCAGTC  GATGAGATGA  AAAGTTCATT

    •     150     •     160     •     170           180           190
AACGCTACAT  TTACAGTGTT  TTGGGCAATT  CTCCCTCCCC  CCCCCCCTCT
                              dy|ad

          200     •     210     •     220     •     230     •     240
CTCTCTCTCT  CTCTCTCTCT  CTCTCTCTCC  CTTCCTCTAA  ATATGTTGAA
```

Figure 7.2: Structural organisation of the DNA in core 2 on linearised pHP2. Only the nucleotide sequence of the Watson strand is shown. The translational frame of core 2, derived from the extent of the DNase I periodicity, is indicated by square brackets. The nucleotides where the minor groove is orientated away from the octamer surface in the reconstituted molecule are indicated by solid circles (•). These are offset by 1bp in the 3′ direction relative to the point of maximum cleavage by DNase I on the Crick strand in the core. The offset relative to the point of maximum cleavage on the Watson strand is 1bp in the 5′ direction. Regions that are A+T-rich are underlined. The dyad axis is indicated at position 165. This structural organisation of the DNA in core 2 was derived from the experimental results presented in section 5.2

presented in figure 7.1. In an extended length of anisotropically flexible DNA, the octamer can theoretically assume different translational placements differing by integral helical turns, thus maintaining the placement of the octamer on the same side of the DNA helix. However, since the free energy change of core formation is expected to be more favorable in areas where the sequence periodicity is more consistent with the bending requirements of core DNA, it appears that if the anisotropic flexibility of the DNA was the sole contributor to core positioning, the dyad axis of core 2 should have been located in one of the flanking regions. The appearance of the dyad axis at the indicated position therefore suggests that an additional translational signal must be present in the region of the sequence occupied by core 2. (The possibility that either core 1 or 3 is translationally strongly positioned, thus "forcing" core 2 onto the determined translational frame, is temporarily neglected).

Satchwell *et al.* (210) reported that runs of $d(A)_n.d(T)_n$ with n>5 appeared to be excluded from the central core area, particularly at positions between 40bp and 50bp from the core terminus, and located more often in the terminal regions of the core DNA. It was proposed that due to the axial rigidity (61) of $d(A)_n.d(T)_n$, conferred by the excellent same-strand stacking and bifurcated hydrogen-bonds (173), these sequences could not be accommodated in the central areas of the cores, where sharp bending of the DNA is required. It is not clear why the region specifically between positions 40bp and 50bp from the core terminus, which is situated between the kinks in the DNA superhelix at positions ± 1 and ± 4 identified in the core crystal structure (208), is avoided by extended runs of $d(A)_n.d(T)_n$. Interestingly, Garner and Felsenfeld (74) demonstrated that runs of $d(C\text{-}G)_n.d(C\text{-}G)_n$ in the Z-form, which has a

persistence length of 2000Å compared to 840Å of B-DNA (242) and are therefore expected to be axially more rigid, were also excluded from the central areas of the core DNA. It was therefore suggested that axially inflexible sequences might constitute translational signals (56, 60, 61, 246).

Drew and Calladine (55) reported that the trinucleotides TGG/CCA and CGG/CCG, similar to the trinucleotide AAA/TTT, were incorporated into cores with a decreased probability, and demonstrated that the distribution of these three trinucleotides could account for the translational placement of cores on an 860bp fragment containing the *Xenopus* 5S RNA gene. In order to test whether the translational placement of cores 2 on pHP2 was similarly determined by the distribution of these trinucleotides, an algorithm of translational positioning, proposed by Drew and Calladine (55), was applied to the region from 1736-560 in pHP2. The result is presented in figure 7.3.

It is clear that the predicted translational frames on pHP2 differ significantly from that assumed by the cores *in vitro*. For instance, the position of the dyad axis of core 2, indicated by an arrow in figure 7.3, is located on a maximum, and not a minimum suggested to represent a favorable translational setting (55). In figure 7.4 the occurrence of all 32 unique trinucleotides in the sequence of core 2 are graphically presented (complimentary trinucleotides are assumed to be equivalent). Referring to this figure, it can be seen that some of the trinucleotides that were suggested to be statistically excluded from the central area of the core, do appear in the DNA of core 2. Notable is the distribution of the trinucleotide AAA/TTT (row 14 of figure 7.4), which occurs 6 times, including as two runs of $d(A)_n.d(T)_n$ with n=4 at positions 130 and 162, the latter position within one helical turn from the dyad axis. The inclusion of oligo(dA).oligo(dT) runs of this

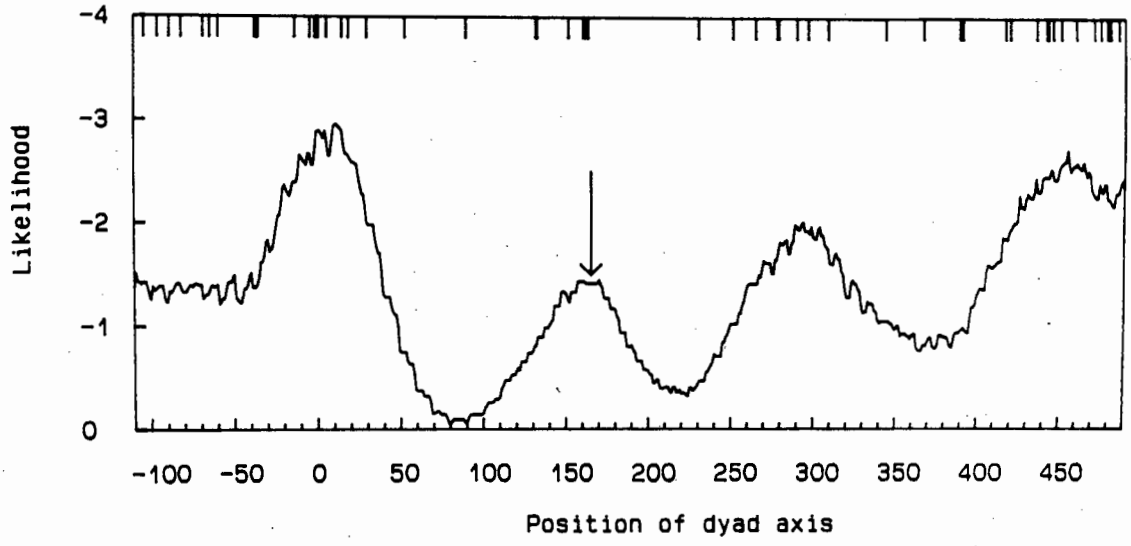

Figure 7.3. Predicted translational positions of nucleosome cores on pHP2. The likelihood of a translational setting at different locations in the sequence of pHP2, calculated after Drew and Calladine (55), is indicated on the ordinate as the likelihood times RT (0,616 at roomtemperature) and varies from 0 for a favorable position to -4 for an unfavorable position. The location of the dyad of core 2 is indicated by the arrow at position 165. The vertical bars at the top of the figure indicates the location of the trinucleotides AAA/TTT, TGG/CCA and CGG/CCG in the sequence of pHP2. See text for detail.

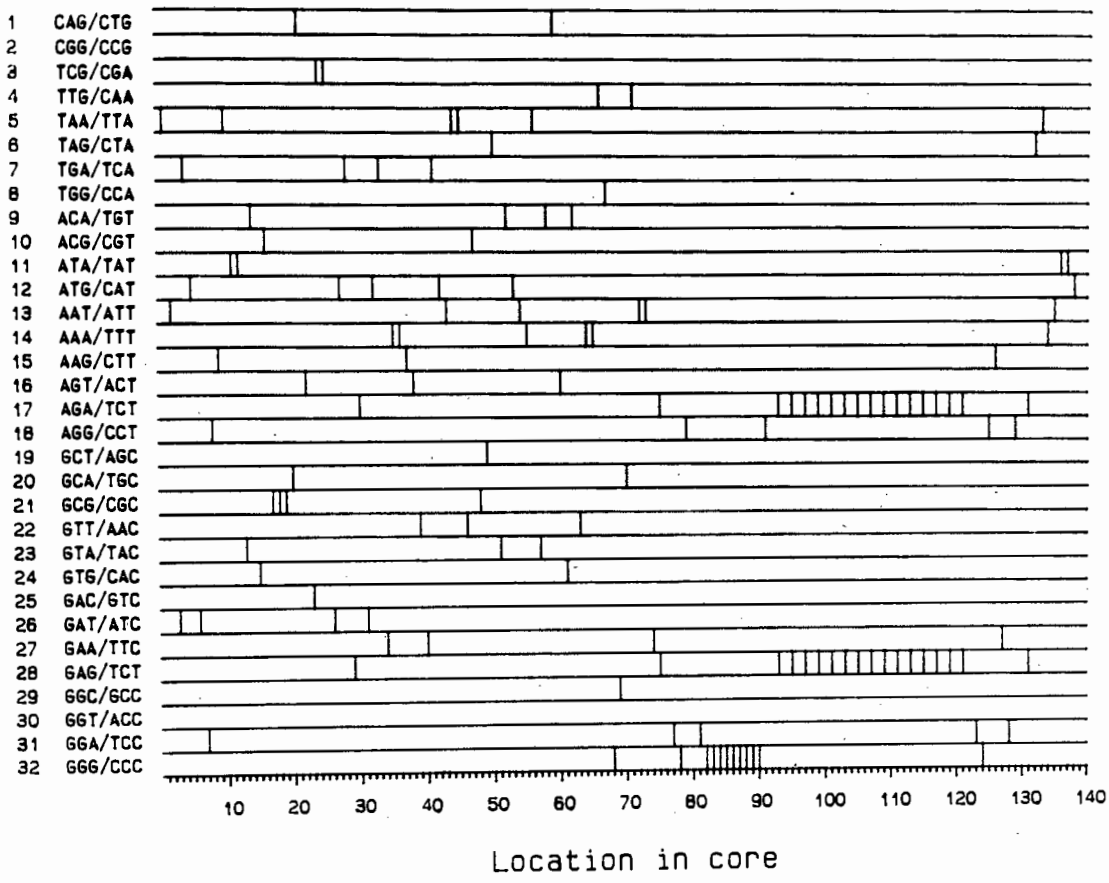

Figure 7.4. The distribution of the 32 possible unique trinucleotides in the sequence of core 2. The vertical lines indicate the locations of the relevant trinucleotides, listed on the left-hand side of the figure. Complimentary trinucleotides were treated as equivalent.

length is consistent with the results of Satchwell *et al.* (210). It can also be seen that the trinucleotide GTT/AAC occurs at position 158 (row 22 of figure 7.4), within the same helical turn. These sequences do not occur with a higher frequency in the flanking linker areas of core 2.

One reason for the poor agreement between the sequence dependency of core placement of the 5S RNA fragment and the H1-H4 fragment may lie in the frequency of occurrence of the relevant trinucleotides. In the 5S RNA fragment these trinucleotides occur on average 18 times per 140bp stretch. In the H1-H4 fragment in pHP2 the average occurrence is only 7 times per 140bp region (individual locations are indicated by the vertical bars at the top of figure 7.3). The decreased frequency of occurrence of these three nucleotides is emphasized by the small minimum values obtained in figure 7.3 compared to the minima obtained in the plot of the 5S RNA fragment (see figure 15 of reference 55). It is therefore doubtful whether the distribution of the trinucleotides in the H1-H4 fragment would give a statistically reliable indication of the placement of the core. Since core 2 was found to be translationally strictly positioned, some other feature in the sequence in the region of core 2 must determine this placement, possibly negating the proposed influence (55) of the trinucleotides mentioned above.

It was shown by Turnell *et al.* (250) that the sequences GGT/ACC, CAA/TTG, XA_4X/YT_4Y ($X \neq A$ and $Y \neq T$) and GGG/CCC are located more often at the dyad axis than at other positions in the core. In the helical turn centered at the dyad of core 2 the sequences GGG/CCC, XA_4X/YT_4Y occur once each, and the sequence CAA/TTG twice (see figure 7.2). It was proposed (211) that these sequences are preferentially placed at the dyad since they comply with the structural requirement of a narrow minor groove at this position (210),

and as such may define a translational setting for a core (246).

Turning next to sequences that are preferentially placed towards the termini and linker areas of cores, it is seen that no continuous run of $d(A)_n.d(T)_n$ with $n \geq 5$ is present in the flanking sequences of core 2 (refer to figure 7.2). However, the tetranucleotide TATA occurs at positions 94 and 106 in the linker area between cores 1 and 2 and in the terminal helical turn of core 2, respectively. On the other side of core 2, at position 228, the sequence TAAATAT appears. The number of core sequences available is insufficient to obtain a statistically reliable estimate of the distribution of tetranucleotides (210), however, by referring the sequences of positioned cores investigated by other workers, it was found that the tetranucleotide TATA occurred in the linker area of the core positioned on the β-globin promoter, and never in the core sequence (117). The sequence TAATATAAAA appeared in the linker area between two cores positioned on a 492bp fragment from the MMTV long terminal repeat (193), and repeated d(T-A).d(T-A) dinucleotides were absent from the core DNA. In the core present on a 260bp fragment containing the 5S rRNA gene from *Lytechinus variegatus*, the sequence TATA occurred in both the assigned terminal helical turn as well as in the core DNA (220). However, in this case, shifting of the core on the 260bp fragment to a position excluding both TATA sequences would have resulted in the incorporation of $d(A)_8.d(T)_8$, a sequence which was shown to be preferentially excluded (210), as well as the sequence TAGTA. In the study by Neubauer *et al.* (175), the single TATA sequence present in a fragment of the α-satellite DNA from African Green Monkey, on which nucleosome cores were shown to occupy multiple overlapping frames both *in vitro* and *in vivo*, was incorporated into the terminal helical turn of the nucleosome core positioned in the dominant translational

frame on the fragment. Thus, although these few examples clearly does not represent a database of sufficient size to claim statistical significance, it appears possible that the tandem d(T-A).d(T-A) base-step may be present more often in the core termini and linker areas than in the central core DNA. Although it was shown that the d(T-A).d(T-A) containing poly(dA-dT).poly(dA-dT) can be reconstituted into nucleosome cores (204, 219), it must be borne in mind that in such a reconstitution system, the association of the octamer with the DNA and concomitant release of counter-ions from the DNA surface constitutes a favorable entropic change (44, 227), and the reconstitution ability of the molecule, other than indicating gross structural features of the molecule such as the presence of pronounced axial rigidity in poly(dA).poly(dT) (204, 219), does not give an indication of possible subtle differences (at ionic conditions matching those *in vivo*) in the affinity of the octamer for molecules of different DNA sequence. In this regard it was shown that the TA step in a crystal of the tetranucleotide d(TATA).d(TATA) was unstacked (255) and assumed a high roll and high slide conformation in three other crystal structures (107, 213, 259), suggested to result from the cross-strand purine-purine clash (38) of highly propeller-twisted adjacent T.A and A.T base-pairs. The repetition of this conformation over three base-steps may render the sequence unsuitable for incorporation into the core DNA of defined anisotropic flexibility. There is indeed a decrease in the occurrence of the TA dinucleotide in core sequences, although this may have been induced by the preferential cleavage by MNase, used in the preparation of the cores, at the TA step (70), as suggested by the authors (210). The TA step was also shown to be thermally least stable of the dinucleotides (85) and easily deformed by torsional stress (62). Satchwell and Travers (211) reported an enrichment of torsionally flexible sequences in the linker areas of sequenced dinucleosomes, and although

the individual sequences were not stated, the presence of TATA sequences may have contributed significantly to this average physical property. It is possible to test for the preponderance of TATA sequences in the linker areas by determining the relative accessibility of the sequence GTATAC to the restriction enzyme *Sna* I and comparing this to a range of other restriction enzymes with recognition sequences expected not to be excluded from core DNA in long chromatin.

Another feature in the DNA sequence of core 2 that may have caused the strict translational placement is the curvature of the molecule (see section 5.4). It was shown that the histone octamer could be reconstituted with higher efficiency onto curved fragments (47, 188, 195), as opposed to uncurved fragments. This may represent a more favorable energy change in wrapping a length of DNA that is already curved round the octamer compared to a length that is, on average, straight. Note, however, that the phenomenon of curvature reflects structural features of the DNA distinct from those determining the anisotropic flexibility (61). Thus, curvature per se is not an indication of a favorable positioning frame. This may be demonstrated in an extreme case where an extended length of poly(dA).poly(dT) is interrupted, at its center, with the sequence $d(T)_5.d(A)_5$, expected to introduce a sharp deflection of the average helical axis (95), and conferring electrophoretic properties to the molecule indicative of curvature (278). Referring to the results of Simpson and Künzler (219) and Rhodes (204), this molecule is expected not to be wrapped into a nucleosome core. Along similar lines, if the azimuth of curvature does not coincide with the average direction of anisotropic flexibility of the molecule, the curvature is expected not to contribute to favorable positioning. However, if a molecule is simultaneously curved and anisotropically flexible in the same direction, the

curvature should not only enhance the rotational placement of the core, but also constitute a translational signal, since the energy change of core formation in the curved region should be more favorable than in neighboring uncurved regions with similar anisotropic flexibilities. Hsieh and Griffith (104) reported that a single, a nucleosome core present on a 1216bp fragment was preferentially located in a curved region of the molecule. The direction of curvature of the positioning sequence of core 2 is, unfortunately, not known, and cannot be easily determined by classical circular permutation studies (278) because of the paucity of unique restriction sites in the H1-H4 fragment. As such one may only conclude that curvature of the positioning sequence of core 2 may contribute to the translational placement of this core.

Other interesting sequences incorporated into core 2 include a 32bp run of d(A-G).d(C-T) preceded by a run of $d(G)_{11}.d(C)_{11}$. The fractional variation in occurrence of the trinucleotides AGA and GAG in core sequences (210) suggests little angular preference, and as such the AG-run is expected to be isotropically flexible. One of the core fragments sequenced by Satchwell *et al.* (210) consisted almost entirely of d(A-G).d(C-T), and Jayasena and Behe (114) reported successful reconstitution of poly d(A-G).poly d(C-T) into nucleosome cores. The incorporation of this poly(purine).poly(pyrimidine) stretch into core 2 is therefore not surprising.

Jayasena and Behe (114) also repcrted the successful reconstitution of poly(dG).poly(dC) into nucleosome cores by high salt exchange reconstitution. This is in contrast to the results of Rhodes (204) and Simpson and Künzler (219) who reported the inability of the homopolymer to be reconstituted by salt dialysis. However, as suggested by Rhodes (204), and discussed in section 1.4.3, the formation

of a poly(dG) tetrad (182) at the high initial salt concentration employed in the reconstitution method probably excluded proper core formation. In this regard Satchwell *et al.* (210) reported that extended runs of $d(G)_n.d(C)_n$ were not statistically excluded from core DNA, and the appearance of the oligo(dG).oligo(dC) run in the DNA of core 2 is therefore not unusual.

It was not possible to assign the locations of the dyad for cores 1 and 3 from the available experimental evidence. However, if both cores are bound to approximately 146bp of DNA, the locations of the single borders identified at positions 93 and 248, suggests that the dyad axes of cores 1 and 3 should be placed in the regions of positions 10-30 and 310-340 respectively.

Examining these regions in the figure relating the likelihood of rotational placement to the position of the core dyad (see figure 7.1), it is difficult to identify features compatible with the single rotational frame observed for core 1, and the mixture of rotational frames differing by 5bp for core 3. The gross signal in the relevant areas appears more erratic in the case of core 1 than that of core 3. However, at positions 307 and 321, minima, removed by 14bp, appear in the signal (see figure 7.1). If, in a large population of DNA molecules, the dyad axis of core 3 is located at either of these positions, it is expected that the experimentally observed mixture of rotational placements differing by approximately 5bp will be visible (see section 5.2).

The shift in the translational position of core 3 following insertion of a 4bp insert after position 295 in pHP2R is intriguing. It was shown that this modification resulted in the invasion of the translational frame of core 2 by core 3, as judged by the appearance of split peaks in the two

terminal helical turns of core 2. The position of the 4bp insert was approximately 2,5 helical turns from the center of the possible range of locations of the dyad of core 3, but did not contain any of the trinucleotides shown to be preferentially located at the dyad (250), and is therefore not expected to define a new translational setting from sequence content alone. The position of the insert is clearly insufficient to constitute a translational signal for the right-hand side border of core 3. It was suggested in section 6.3 that the rotation of the DNA molecule on either side of position 295 caused the better alignment of anisotropically flexible sequences over an extended length of the DNA of core 3, allowing the core extended translational freedom in the absence of strong translational determinants. As such, the misalignment of the anisotropically flexible sequences in the absence of the insert after position 295 may have constituted a translational determinant. In order to test this proposal, the likelihood of a rotational placement *versus* sequence position was calculated in the region from position 158 to 482 in pHP2RR which corresponds to dyad location from position 220 to 420. The result is presented in figure 7.5.

It is clear from this figure that the anisotropic flexibility over 124 base-steps centered in the range from positions 230 to 270 has changed as a result of the insert after position 295 (compare figure 7.5 and 7.1). In particular, the amplitude in the oscillation of the signal has increased in this region in figure 7.5, suggesting that the direction of preferential bending of the DNA is more clearly defined as a result of the relative rotation of the DNA molecule on either side of position 295. In order to test whether that change in anisotropic flexibility has indeed produced a better long range alignment of sequences with similar bending preferences at a period of 10,2bp, the location of the minima in the range from 220 to 350 (at a

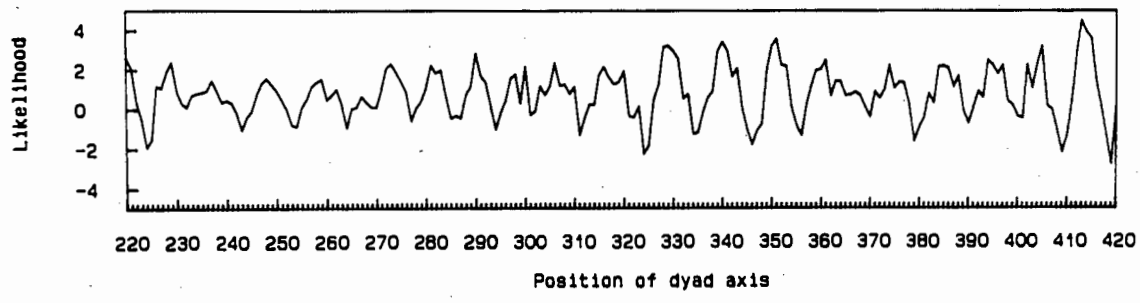

Figure 7.5. Predicted rotational positions of nucleosome cores on pHP2RR. The likelihood of the location of the core dyad at different positions in the sequence of pHP2 is indicated as the likelihood times RT (0,616 at room temperature) and varies from -4 for a favorable rotational setting to +4 for an unfavorable one (after Drew and Calladine (55)) See text for detail.

period of approximately 10bp) in figures 7.1 and 7.5 were plotted as a *modulus* of 10,2. If the minima occur at average spacings of approximately 10,2bp, indicating a similar direction of anisotropic flexibility over an extended length of the DNA molecule, the *modulus* of these minima are expected to be located on a horizontal line. If, however, the minima are located more erratically, the *modulus* of the minima is expected to drift from a line or display sharp deflections. The plot of the minima *modulus* 10,2 in figures 7.1 and 7.5 are presented in figure 7.6.

In figure 7.6 it is seen that the molecule that was not rotated on either side of position 295 (squares in figure 7.6) display a gradual change in the predicted preferential rotational position of cores proceeding from position 220 to 350 as indicated by the deflection of the line drawn through the datapoints. In the case of the rotated molecule (circles in figure 7.6), the datapoints appear clustered around the line without deflections. In particular, datapoints on either side of the range can be connected through several of the datapoints in the center of the range. This suggests that the predicted rotational position of cores occur within narrow limits on the same side of the DNA helix, whereas, in former case, the predicted rotation positions include all sides of the helix in the range tested. It would therefore appear that if a core which is rotationally placed in the region of position 310-340 is shifted in the direction of core 2 while the orientation of the core relative to the DNA helix is maintained, the direction of anisotropic flexibility of the right-hand section of the core DNA would differ from that encountered in the left-hand sections. In contrast, on the molecule rotated on either side of position 295, the stretches of DNA encountered on the left-hand side of the shifting core would more closely match that on the

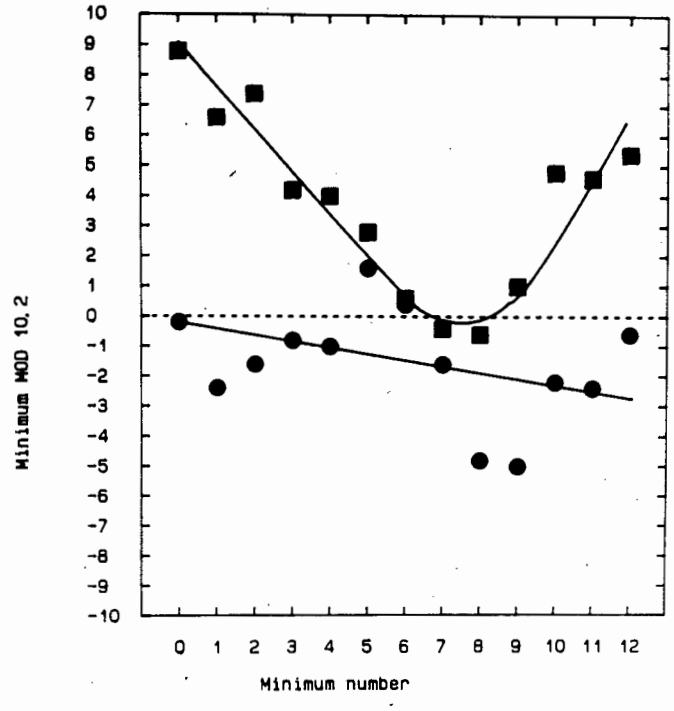

Figure 7.6. Change in the predicted rotational frame of nucleosome cores on pHP2 and pHP2RR. The locations of 13 minima present in the rotational preference plots in the regions from 220 to 350 in pHP2 (squares) and pHP2RR (circles) are shown as the sequence position of the minimum *modulus* 10,2 on the ordinate *versus* the number of the minimum (counting from the first minimum present after position 220) on the abscissa. The curves indicate the tendency of the two datasets, and were not fitted to the datapoints.

right-hand side. This suggests that the change in the translational positioning of core 3 in the rotated molecule may indeed be due to the alignment of sequences with similar bending properties on either side of position 295, imparting wider translational freedom to core 3 in the absence of other strong translational signals inherent in the DNA sequence. This argument clearly holds only for one of the two rotational placements assumed by core 3 (see section 5.2.2). Due to the expected superimposition of the maxima in the probability of DNase I cleavage at positions 228_{rr} and 238_{rr}, it is not possible to deduce from the available experimental evidence whether the translational frame for core 3 present in both rotational settings (see section 6.3) shifted.

7.3 Negative supercoiling and nucleosome core positioning:

It was shown above that the position assumed by core 2 can be explained by the sequence of the DNA molecule in terms of the known rotational determinants of nucleosome core positioning (37, 55, 56, 59, 246, discussed in section 1.4.4). It was further shown that core 2 assumed identical translational and rotational settings on linearized molecules (section 5.2.2 and 5.2.3), circular molecules in the absence of negative superhelical stress (section 5.2.4), circular molecules in the presence of negative superhelical stress (section 5.3), and on molecules where sections of the rotational frames of adjacent cores were rotated relative to that of core 2, both on linearized and on circular molecules in the absence of negative superhelical stress (sections 6.2 and 6.3). These result clearly indicate that core 2 is very stably positioned relative to the underlying DNA sequence, and no shift to an alternative rotational and/or translational setting was discernable in the range of geometric and topological constraints experimentally studied. Although the position of core 3 was often less

clear-cut (by virtue of it being a less-well positioned core), no change in the positioning frame of this core was similarly detectable as a result of modifications of the geometric or topological properties of the DNA molecule. These results indicate, as suggested by the analysis of the geometric properties of the free and reconstituted supercoiled DNA molecule (see sections 5.5.2.1 and 5.5.2.2), that the positioning of nucleosome cores are not influenced by intermediate levels of negative supercoiling. It was also found in several repetitions (>10) of the determination of the position of core 2 on linearized and supercoiled molecules, that the rotational position of this core was not enhanced as a result of the circular constraint of the supercoiled plasmid. These results suggest that random fluctuations in the shape and unrestrained level of negative superhelical stress in chromatin loops will not cause nucleosomes to occupy alternative setting. This conservation of the positioning frames is expected to be essential for regulatory nucleosomes where the exposure of specific sequences to DNA-binding proteins are required.

There is evidence that the positioning of core 2 may be biologically equally significant. It was found that binding of a protein (FAG), present in embryonic nuclear extracts of *Psammechinus angulosus* embryos, protected the sugar phosphate backbone of bases at positions 170-203 (Watson) and 166-191 (Crick) from DNase I cleavage (J P Hapgood, unpublished results). It is not currently known whether the sequence specificity of the protein is conferred by contacting the base-pairs in the major groove, or by recognition of the structure of the $d(G)_{11}.d(C)_{11}$ run, shown to be A-DNA like (151). If recognition by either of these mechanisms proceed from the major groove side, the protein should still be able to recognise its binding site on the surface of core 2, where the position of the dyad axis would place the major groove at base-pairs 179-181, shown by

methylation interference studies to be important in the protein binding, away from the surface of the octamer. Preliminary DNase I protection results of a ternary complex between DNA, histone octamer and FAG indeed indicated that FAG still bound to its recognition sequence in the reconstituted DNA molecule (J P Hapgood, unpublished results). These results suggest that the positioning of core 2 may be genetically essential, similar to the nucleosomes present on the *PHO5* promoter region (65) and the mouse mammary tumor virus long terminal repeat (194). The absence of a shift in this core as a response to the geometry and topology of the DNA molecule suggests that protein binding to DNA is not regulated by the differential exposure of recognition sites on cores via changes in geometry and topology of chromatin fibres *in vivo*.

7.4 Negative supercoiling and nucleosome core formation:

Unlike the positioning of nucleosome cores, the formation of nucleosome cores was clearly shown to be responsive to the topology of the DNA molecule (see sections 4.3 and 4.4). It was also shown that the increase in the affinity of the octamer for negatively supercoiled DNA can be accounted for by the decrease in the level of unconstrained negative superhelicity upon core formation (see section 4.6). Negri *et al* (172) and Goulet *et al.* (87) presented evidence which showed that the affinity of the octamer for DNA decreased if the formation of the core resulted in the introduction of nett unrestrained positive superhelical stress into small circular DNA molecules. It was shown that transcription *in vivo* resulted in the formation of twin supercoiled domains (82), suggested to reflect viscous drag on the processing transcription complex and nascent RNA chain, causing the rotation of the right-handed DNA molecule (136), thereby introducing a positively supercoiled region ahead of the transcription complex, and a negatively supercoiled domain

behind it. Although the question of octamer displacement by transcription remains contentious (138, 139 and 140), the twin-supercoiled domain suggests a neat model whereby cores in front of the transcription complex are destabilised, causing them to dissociate from the DNA, and reassociate with high affinity with the negatively supercoiled region behind the transcription complex. During the preparation of this manuscript similar proposals were published by Pederson and Morse (186), and by Clark and Felsenfeld (43). Clark and Felsenfeld (43) demonstrated that the migration of the octamer from positively to negatively supercoiled plasmids did, indeed, occur *in vitro*, but showed that this migration required an ionic strength of 800mM NaCl. The migration of the octamer from negatively supercoiled plasmids saturated with nucleosome cores to naked negatively supercoiled plasmids was shown in section 4.5 to require similar high ionic strengths. Clearly, as pointed out by Clark and Felsenfeld (43), these high ionic strengths appear unrealizable *in vivo*, and the inability of high concentrations of poly[L-glutamate] to affect octamer migration (see section 4.5) suggests that the presence of acidic proteins, such a nucleoplasmin (131), may not be sufficient to support the migration of the octamer from positively to negatively supercoiled domains. Clark and Felsenfeld (43) proposed that the phosphorylated C-terminal tail domain of RNA polymerase or acetylation of the octamer may facilitate the shortdistance migration of the octamer within the same DNA coil. The proposed mechanism, however, awaits experimental confirmation.

7.5 Relation to other work:

Two research papers, raising the question of the effect of negative supercoiling on nucleosome core positioning, appeared recently.

The first, by Almouzni *et al.* (7), was published during the final stages of experimental work, and cited an unpublished observation by D J Clark and A P Wolffe that the position assumed by a nucleosome core on a fragment of the somatic 5S rRNA gene from *Xenopus* was identical in a linear and negatively supercoiled molecule. No detail on the quantitation of the cores on the molecule or the level of negative supercoiling tested was given, but at face value this result is in agreement with those reported above in chapters 5 and 6.

The second paper, by Piña *et al.* (193), appeared during the preparation of this manuscript, and reported the positioning of nucleosome cores on linear and circularized 212bp and 492bp molecules containing sections from the long terminal repeat of the mouse mammary tumor virus. It was shown (193) that the nucleosome core occupied identical positions on both fragments irrespective of the topology of the molecule. It is interesting to compare the assigned sites of maximal DNase I cleavage in the reconstituted 212bp circular fragment with the sites on the reconstituted linear fragment reported in earlier studies by both Piña *et al.* (194) and Perlmann and Wrange (189). In the study by Piña *et al.* (194), where a 208bp linear fragment was employed, sites of maximum cleavage, proceeding from the vicinity of the assigned dyad axis towards the core terminus, were assigned to positions -133, -145, -155, -165, -176, -185 and -195. Perlmann and Wrange (189), employing a 199bp linear fragment, reported sites of maximum DNase I cleavage at positions -133, -146, -156, -165, no site, -185 and -195. It is clear that the assigned sites of maximum DNase I cleavage in the reconstituted fragments employed in the two studies agree well. In the 212bp circular fragment ($\Delta Lk=-2$), where the nucleosome core positioned in a setting identical to that on the linearized molecule (193), the sites of maximal DNase I cleavage were assigned to positions

-133, -145, -155, -168, -177, -188 and -197. If the last four values in the circular fragment are compared with those in the linearized fragments, it appears that the local helical twist of the terminal three helical turns in the circular fragment has decreased in response to the unwound nature of the supercoiled molecule. If all the excess negative superhelical stress in the reconstituted circular fragment was confined to the linker areas, the specific linking difference in these regions would have equaled -0,156. Similar reductions in the local helical twist of the terminal regions of core 2 reconstituted onto supercoiled pHP2 were not found (see section 5.3). However, in the experiment reported in section 5.3 the specific linking difference in the linker areas equaled -0,04. This degree of negative superhelical stress may have been insufficient to cause terminal unwinding of the core DNA.

Thus, the results of Piña *et al.* (193), obtained at higher negative specific linking differences, are in perfect agreement with the results and conclusions on negative supercoiling and core positioning presented above.

CHAPTER 8

8 References.

1) Adrian, M., ten Heggeler-Bordier, B., Wahli, W., Stasiak, A. Z., Stasiak, A. & Dubochet, J. (1990). Direct visualization of supercoiled DNA molecules in solution. *EMBO J.* 9, 4551-4554.

2) Aggarwal, A. N., Rodgers, D. W., Drottar, M., Ptashne, M. & Harrison, S. C. (1988). Recognition of a DNA operator by the repressor of Phage 434: a view at high resolution. *Science*, 242, 899-907.

3) Akima, H. (1970). A new method of interpolation and smooth curve fitting based on local procedures. *J. Ass. Comp. Masch.* 17, 589-609.

4) Albright, S. C., Nelson, P. P. & Garrard, W. T. (1979). Histone molar ratios among different electrophoretic forms of mono- and dinucleosomes. *J. Biol. Chem.* 254, 1065-1073.

5) Allan, J., Hartman, P. G., Crane-Robinson, C. & Aviles, F. X. (1980). The structure of histone H1 and its location in chromatin. *Nature (London)*, 288, 675-679.

6) 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.

7) Almouzni, G., Méchali, M. & Wolffe, A. P. (1990). Competition between transcription complex assembly and chromatin assembly on replicating DNA. *EMBO J.* 9, 573-582.

8) Ambrose, C., Rajadhyakska, A., Lowman, H. & Bina, M. (1989). Locations of nucleosomes on the regulatory region of simian virus 40 chromatin. *J. Mol. Biol.* 209, 255-263.

9) Anderson, P. & Bauer, W. (1978). Supercoiling in closed circular DNA: dependence upon ion type and concentration. *Biochemistry*, 17, 594-601.

10) Ausio, J., Borochov, N., Seger, D. & Eisenberg, H. (1984). Interaction of chromatin with NaCl and $MgCl_2$. Solubility and binding studies, transition to and characterization of the higher-order structure. *J. Mol. Biol.* 177, 373-398.

11) Ausio, J., Seger, D. & Eisenberg, H. (1984). Nucleosome core particle stability and conformational change. Effect of temperature, particle and NaCl concentrations, and crosslinking of histone H3 sulfhydryl groups. *J. Mol. Biol.* 176, 77-104.

12) Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1987). *Current protocols in molecular biology*, John Wiley and Sons, New York.

13) Bates, D. L., Butler, P. J. G., Pearson, E. C. & Thomas, J. O. (1981). Stability of higher-order structure of chicken-erythrocyte chromatin in solution. *Eur. J. Biochem.* 119, 469-476.

14) Bauer, W. R., Crick, F. H. C. & White, J. H. (1980). Supercoiled DNA. *Sci. Amer.* 243, 100-113.

15) Bavykin, S. G., Usachenko, S. I., Lishanskaya, A. I., Shick, V. V., Belyavsky, A. V., Undritsov, I. M., Strokov, A. A., Zalenskaya, I. A. & Mirzabekov, A. D. (1985). Primary organization of nucleosomal core particles is invariable in repressed and active nuclei from animal, plant and yeast cells. *Nucl. Acids Res.* 13, 3439-3459.

16) Benyajati, C. & Worcel, A. (1976). Isolation, characterization, and structure of the folded interphase genome of *Drosophila melanogaster*. *Cell*, 9, 393-407.

17) Berg, O. G. & von Hippel, P. H. (1988). Selection of DNA binding sites by regulatory proteins. *Trends Biochem. Sci.* 13, 207-211.

18) Bergman, L. W. & Kramer, R. A. (1983). Modulation of chromatin structure associated with derepression of the acid phosphatase gene of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 258, 7223-7227.

19) Bernardi, G. (1971). Chromatography of nucleic acids on hydroxyapatite columns. *Meth. Enz.* 21, 95-111.

20) Blaho, J. A., Larson, J. E., McLean, M. J. & Wells, R. D. (1988). Multiple DNA secondary structures in perfect inverted repeat inserts in plasmids, right-handed B-DNA, cruciforms, and left-handed Z-DNA. *J. Biol. Chem.* 263, 14446-14455.

21) Bloom, K. S. & Carbon, J. (1982). Yeast centromere DNA is unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell*, 29, 302-317.

22) Boles, T. C., White, J. H. & Cozzarelli, N. R. (1990). Structure of plectonemically supercoiled DNA. *J. Mol. Biol.* 213, 931-951.

23) Bonne-Andrea, C., Wong, M. L. & Alberts, B. M. (1990). *In vitro* replication through nucleosomes without histone displacement. *Nature (London)*, 343, 719-726.

24) Borowiec, J. A. & Gralla, J. D. (1985). Supercoiling response of the *lac* p^s promoter *in vitro*. *J. Mol. Biol.* 184, 587-598.

25) Borowiec, J. A. & Gralla, J. D. (1987). All three elements of the *lac* p^s promoter mediate its transcriptional response to DNA supercoiling. *J. Mol. Biol.* 195, 89-97.

26) Borowiec, J. A., Zhang, L., Sasse-Dwright, S. & Gralla, J. D. (1987). DNA supercoiling promotes formation of a bent repression loop in *lac* DNA, *J. Mol. Biol.* 196, 101-111.

27) Boulikas, T., Wiseman, J. M. & Garrard, W. T. (1980). Points of contact between histone H1 and the histone octamer, *Proc. Natl. Acad. Sci., U.S.A.* 77, 127-131.

28) Brady, G. W., Foos, D. & Benham, C. J. (1985). *Biopolymers*, 23, 2963-2966.

29) Brady, G. W., Satkowski, M., Foos, D. & Benham, C. J. (1987). Environmental influences on DNA superhelicity. The effect of ionic strength on superhelical conformation in solution. *J. Mol. Biol.* 195, 185-191.

30) Brill, S. J. & Sternglanz, R. (1988). Transcription-dependent DNA supercoiling in yeast DNA topoisomerase mutants. *Cell*, 54, 403-411.

31) Broyles, S. S. & Pettijohn, D. E. (1986). Interaction of the *Escherichia coli* HU protein with DNA. Evidence for formation of nucleosome-like structures with altered DNA helical pitch. *J. Mol. Biol.* 187, 47-60.

32) Burkhoff, A. M. & Tullius, T. D. (1987). The unusual conformation adopted by the adenine tracts in kinetoplast DNA. *Cell*, 48, 935-943.

33) Butler, P. J. G. (1980). Changes in chromatin folding in solution. *J. Mol. Biol.* 140, 505-529.

34) Butler, P. J. G. (1983). The folding of chromatin. *CRC Critical Reviews in Biochemistry*, 15, 57-91.

35) Butler, P. J. G. (1984). A defined structure of the 30nm chromatin fibre which accommodates different nucleosomal repeat lengths. *EMBO J.* 3, 2599-2604.

36) Caffarelli, E., Leoni, L., Sampaolese, B., Savino, M. (1986). Persistence of cruciform structure and preferential location of nucleosomes on some regions of pBR322 and ColE1 DNAs. *Eur. J. Biochem.* 156, 335-342.

37) Calladine, C. R. & Drew, H. R. (1986). Principles of sequence-dependent flexure of DNA. *J. Mol. Biol.* 192, 907-918.

38) Calladine, C. R. (1982). Mechanisms of sequence-dependent stacking of bases in B-DNA. *J. Mol. Biol.* 161, 343-352.

39) Calladine, C. R., Drew, H. R. & McCall, M. J. (1988). The intrinsic curvature of DNA in solution. *J. Mol. Biol.* 201, 127-137.

40) Camerini-Otero, R. D. & Felsenfeld, G. (1977). Supercoiling energy and nucleosome formation: the role of the arginine-rich histone kernel. *Nucl. Acids Res.* 4, 1159-1181.

41) Castell, S. E. & Halford, S. E. (1989). DNA supercoiling determines the activation energy barrier for site specific recombination by Tn*21* resolvase. *Nucl. Acids Res.* 17, 7045-7058.

42) Chao, M., Gralla, J. D. & Martinson, H. G. (1979). DNA sequence directs placement of histone cores on restriction fragments during nucleosome formation. *Biochemistry*, 18, 1068-1074.

43) Clark, D. J. & Felsenfeld, G. (1991). Formation of nucleosomes on positively supercoiled DNA. *EMBO J.* 10, 387-395.

44) Clark, D. J. & Kimura, T. (1990). Electrostatic mechanisms of chromatin folding. *J. Mol. Biol.* 211, 883-896.

45) Collis, C. M., Molloy, P. I., Both, G. W. & Drew, H. R. (1989). Influence of the sequence-dependent flexure of DNA on transcription in *E.coli*. *Nucl. Acids Res.* 17, 9447-9468.

46) Cooney, M., Czernaszewicz, G., Postel, E. H., Flint, S. J. & Hogan, M. E. (1988). Site-specific oligonucleotide binding represses transcription of the human *c-myc* gene *in vitro*. *Science*, 241, 456-459.

47) Costanzo, G., Di Mauro, E., Salina, G. & Negri, R. (1990). Attraction, phasing and neighbour effects of histone octamers on curved DNA. *J. Mol. Biol.* 216, 363-374.

48) Cozzarelli, N. R., Boles, T. C. & White, J. H. (1990). Primer on the topology and geometry of DNA supercoiling. In *DNA Topology and its Biological Effects* (Cozzarelli, N. R. & Wang, J. C., eds), pp 139-184, Cold Spring Harbor Laboratory Press, New York.

49) Crick, F. H. C. (1976). Linking numbers and nucleosomes. *Proc. Natl. Acad. Sci., U.S.A.* 73, 2639-2643.

50) Depew, R. E. & Wang, J. C. (1975). Conformational fluctuations of DNA helix. *Proc. Natl. Acad. Sci., U.S.A.* 72, 4275-4279.

51) Diekmann, S. & Wang, J. C. (1985). On the sequence determinants and flexibility of the kinetoplast DNA fragments with abnormal gel electrophoretic mobilities. *J. Mol. Biol.* 186, 1-11.

52) Diekmann, S. (1987). DNA curvature. In *Nucleic Acids and Molecular Biology* (Eckstein, F. and Lilley, D.M.J., eds), vol 1, pp 138-156, Springer-verlag, Berlin.

53) Dodgson, J. B. & Wells, R. D. (1977). Action of single strand specific nucleases on model DNA heteroduplexes of defined size and sequence. *Biochemistry*, 16, 2374-2379.

54) Dong, F., Hansen, J. C. & van Holde, K. E. (1990). DNA and protein determinants of nucleosome positioning on sea urchin 5S rRNA gene sequences *in vitro*. *Proc. Natl. Acad. Sci., U.S.A.* 87, 5724-5728.

55) Drew, H. R. & Calladine, C. R. (1987). Sequence specific positioning of core-histones on an 860bp DNA. Experiment and Theory. *J. Mol. Biol.* 195, 143-173.

56) Drew, H. R. & McCall, M. J. (1987). Structural analysis of a reconstituted DNA containing three histone octamers and histone H5. *J. Mol. Biol.* 197, 485-511.

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

58) Drew, H. R. & Travers, A. A. (1984). Structural junctions in DNA: the influence of flanking sequences on nuclease digestion specificities. *Nucl. Acids Res.* 13, 4445-4467.

59) Drew, H. R. & Travers, A. A. (1985). DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* 186, 773-790.

60) Drew, H. R., McCall, M. J. & Calladine, C. R. (1988). Recent studies of DNA in the crystal. *Ann. Rev. Cell. Biol.* 4, 1-20.

61) Drew, H. R., McCall, M. J. & Calladine, C. R. (1990). New approaches to DNA in the crystal and in solution. In *DNA Topology and its Biological Effects* (Cozzarelli, N.R., Wang, J.C., eds), pp 1-56, Cold Spring Harbor Laboratory Press, New York.

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

63) Esposito, F. & Sinden, R. R. (1987). Supercoiling in prokaryotic and eukaryotic DNA: changes in response to topological perturbations of plasmids in *E.coli* and S-V40 *in vitro*, in nuclei and in CV-1 cells. *Nucl. Acids Res.* 15, 5105-5124.

64) Estruch, F., Pérez-ortín, J. E., Matallana, E. & Franco, L. (1989). *In vivo* assembly of chromatin on pBR322 sequences cloned into yeast plasmids. *Plasmid*, 21, 113-119.

65) Fascher, K. D., Schmitz, J. & Hörz, W. (1990). Role of *trans*-acting proteins in the generation of active chromatin at the *PHO5* promoter in *S.cerevisiae*. *EMBO J.* 9, 2523-2528.

66) Felsenfeld, G., McGhee, J. D. (1986). Structure of the 30nm chromatin fibre. *Cell*, 44, 375-377.

67) Financsek, I., Tora, L., Kelemen, G. & Hidvégi, E. J. (1986). Supercoil induced S1 hypersensitive sites in the rat and human ribosomal RNA genes. *Nucl. Acids Res.* 14, 3263-3277.

68) Finch, J. T. & Klug, A. (1976). Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci., U.S.A.* 73, 1897-1901.

69) 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.

70) 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. Mol. Biol.* 190, 619-633.

71) Fox, K. & Waring, M. (1984). DNA structural variations produced by actinomycin and distamycin as revealed by DNase I footprinting. *Nucl. Acids Res.* 12, 9271-9285.

72) Frank-Kamenetskii, M. D. (1990). DNA supercoiling and unusual structures. In *DNA Topology and its Biological Effects* (Cozzarelli, N. R. & Wang, J. C., eds), pp 185-215, Cold Spring Harbor Laboratory Press, New York.

73) Fuller, F. B. (1971). The writhing number of a space curve. *Proc. Natl. Acad. Sci., U.S.A.* 68, 815-819.

74) Garner, M. M. & Felsenfeld, G. (1987). Effect of Z-DNA on nucleosome placement. *J. Mol. Biol.* 196, 581-590.

75) Gasser, S. M. & Laemmli, U. K. (1986). Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. *Cell*, 46, 521-530.

76) Gasser, S. M. & Laemmli, U. K. (1986). The organization of chromatin loops: characterization of a scaffold attachment site. *EMBO J.* 5, 511-518.

77) Gasser, S. M. (1988). Nuclear scaffold and higher-order folding of eukaryotic DNA. In *Architecture of Eukaryotic Genes* (Kahl, G., ed), pp 461-471, VCH Verlagsgesellschaft, Weinheim, FRG.

78) Gellert, M. (1981). DNA topoisomerases. *Ann. Rev. Biochem.* 50, 879-910.

79) Gellert, M., O'Dea, M. H., Itoh, T. & Tomizawa, J. (1976). Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase. *Proc. Natl. Acad. Sci., U.S.A.* 73, 4474-4478.

80) Gerchman, S. E. & Ramakrishnan, V. (1987). Chromatin higher-order structure studied by neutron scattering and scanning transmission electron microscopy. *Proc. Natl. Acad. Sci., U.S.A.* 84, 7802-7806.

81) 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. Natl. Acad. Sci., U.S.A.* 72, 1843-1847.

82) Giaever, G. N. & Wang, J. C. (1988). Supercoiling of intracellular DNA can occur in eukaryotic cells. *Cell*, 55, 849-856.

83) Giaever, G., Snyder, L. & Wang, J. C. (1988). DNA supercoiling *in vivo*. *Biophys. Chem.* 29, 7-15.

84) Glaser, R. L., Thomas, G. H., Siegfried, E., Elgin, S. C. R. & Lis, J. T. (1990). Optimal heat-induced expression of the *Drosophila hsp26* gene requires a promoter sequence containing $(CT)_n.(GA)_n$ repeats. *J. Mol. Biol.* 211, 751-761.

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

86) Gottschling, D. E. & Cech, T. R. (1984). Chromatin structure of the molecular ends of *Oxytrichia* macronuclear DNA: phased nucleosomes and a telomeric complex. *Cell*, 38, 501-510.

87) Goulet, I., Zivanovic, Y. & Prunell, A. (1988). Chromatin reconstitution on small DNA rings. I. *J. Mol. Biol.* 200, 253-266.

88) Gralla, J. D. (1985). Rapid "footprinting" on supercoiled DNA. *Proc. Natl. Acad. Sci., U.S.A.* 82, 3078-3081. (G20)

89) 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. Mol. Biol.* 185, 461-478.

90) Greyling, H. J., Schwager, S., Sewell, B. T. & von Holt, C. (1983). The identity of conformational states of reconstituted and native histone octamer. *Eur. J. Biochem.* 137, 221-226.

91) Gross, D. S. & Garrard, W. T. (1988). Nuclease hypersensitive sites in chromatin. *Ann. Rev. Biochem.* 57, 159-197.

92) Gupta, G., Sarma, M. H. & Sarma, R. H. (1988). On the question of DNA bending: two dimensional NMR studies on $d(GTTTTAAAAC)_2$ in solution. *Biochemistry*, 27, 7907-7919.

93) Hagerman, P. J. (1984). Evidence for the existence of stable curvature of DNA in solution. *Proc. Natl. Acad. Sci., U.S.A.* 81, 4632-4636.

94) Hagerman, P. J. (1985). Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. *Biochemistry*, 24, 7033-7037.

95) Hagerman, P. J. (1986). Sequence-directed curvature of DNA. *Nature (London)*, 321, 449-450.

96) Hagerman, P. J. (1990). Sequence-directed curvature of DNA. *Ann. Rev. Biochem.* 59, 755-781.

97) Han, M., Kim, U. J., Kayne, P. & Grunstein, M. (1988). Depletion of histone H4 and nucleosomes activates the *PHO5* gene in *Saccharomyces cerevisiae*. *EMBO J.* 7, 2221-2228.

98) Haniford, D. B. & Pulleyblank, D. E. (1985). Transition of a cloned $d(AT)_n.d(AT)_n$ tract to a crufiform *in vivo*. *Nucl. Acids Res.* 13, 4343-4363. (H5)

99) Hansen, J. C., Ausio, J., Stanik, V. H. & van Holde, K. E. (1989). Homogeneous reconstituted oligonucleosomes, evidence for salt-dependent folding in the absence of histone H1. *Biochemistry*, 28, 9129-9136.

100) Hansen, J. C., van Holde, K. E. & Lohr, D. (1991). The mechanism of nucleosome assembly onto oligomers of the sea-urchin 5S DNA positioning sequence. *J. Biol. Chem.* 266, 4276-4282.

101) Hanvey, J. C., Klysik, J. & Wells, R. D. (1988). Influences of DNA sequence on the formation of non-B right-handed helices in oligopurine.oligopyrimidine inserts in plasmids. *J. Biol. Chem.* 263, 7386-7396.

102) Hanvey, J. C., Shimizu, M. & Wells, R. D. (1989). Intramolecular triplexes in supercoiled plasmids. II Effect of base composition and non-central interruptions on formation and stability. *J. Biol. Chem.* 264, 5950-5956.

103) Horowitz, D. S. & Wang, J. C. (1984). Torsional rigidity of DNA and length dependence of the free energy of DNA supercoiling. *J. Mol. Biol.* 173, 75-91.

104) Hsieh, C. H. & Griffith, J. D. (1988). The terminus of SV40 DNA replication and transcription contains a sharp sequence-directed curve. *Cell*, 52, 535-544.

105) Htun, H. & Dahlberg, J. E. (1989). Topology and formation of triple-stranded H-DNA. *Science*, 24, 1571-1576.

106) Htun, H., Lund, E. & Dahlberg, J. E. (1984). Human U1 RNA genes contain an unusually sensitive nuclease S1 cleavage site within the conserved 3' flanking region. *Proc. Natl. Acad. Sci., U.S.A.* 81, 7288-7292.

107) Hunter, W. N., D'Estaintot, B. L. & Kennard, O. (1989). Structural variation in a deoxyoctanucleotide d(CTCTAGAG), implications for protein-DNA interaction. *Biochemistry*, 28, 2444-2451.

108) Iacono-Connors, L. & Kowalski, D. (1986). Altered DNA conformations in the gene regulatory region of torsionally-stressed SV40 DNA, *Nucl. Acids Res.* 14, 8949-8962.

109) Ichimura, S., Mita, K. & Zama, M. (1982). Essential role of arginine residues in folding of deoxyribonucleic acid into nucleosome cores. *Biochemistry*, 21, 5329-5334.

110) Izaurralde, E., Mirkovitch, J. & Laemmli, U. K. (1988). Interaction of DNA with nuclear scaffolds *in vitro*. *J. Biol. Chem.* 200, 111-125.

111) Jackson, D. A., Dickinson, P. & Cook, P. R. (1990). The size of chromatin loops in HeLa cells, *EMBO J.* 9, 567-571.

112) Jackson, V. (1990). *In vivo* studies of the dynamics of histone-DNA interaction: evidence for nucleosome dissolution during replication and transcription and a low level of dissolution independent of both. *Biochemistry*, 29, 719-731.

113) Jarman, A. P. & Higgs, D. R. (1988). Nuclear scaffold attachment sites in the human globin gene complexes. *EMBO J.* 7, 3337-3344.

114) Jayasena, S. D. & Behe, M. J. (1989). Competitive nucleosome reconstitution of polydeoxynucleotides containing oligoguanosine tracts. *J. Mol. Biol.* 208, 297-306.

115) Jayasena, S. D. & Behe, M. J. (1989). Nucleosome reconstitution of core-length poly(dG).poly(dC) and poly(rG-dC).poly(rG-dC). *Biochemistry*, 28, 975-980.

116) Johnson, R. A. & Walseth, T. F. (1979). The enzymatic preparation of [α-^{32}P]ATP, [α-^{32}P]GTP, [γ-^{32}P]cAMP and [γ-^{32}P]cGMP, and their use in the assay of adenylate and guanylate cyclases and cyclic nucleotide phosphodiesterases. In *Advances in Cyclic Nucleotide Research* (Brooker, G., Greengard, P. & Robison, G. A., eds), vol 19, pp 135-167, Raven Press, New York.

117) Kefalas, P., Gray, F. & Allan, J. (1988), Precise nucleosome positioning in the promoter of the chicken β^A globin gene. *Nucl. Acids Res.* 16, 501-517.

118) Kinniburgh, A. J. (1989). A *cis*-acting transcription element of the *c-myc* gene can assume an H-DNA conformation. *Nucl. Acids Res.* 17, 7771-7778.

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

120) Klug, A. & Travers, A. A. (1989). The helical repeat of nucleosome-wrapped DNA. *Cell*, 56, 10-11.

121) Klug, A., Rhodes, D., Smith, J., Finch, J. T. & Thomas, J. O. (1980). A low resolution structure of the histone core of the nucleosome. *Nature (London)*, 287, 509-516.

122) Kolb, A. & Buc, H. (1982). Is DNA unwound by the cyclic AMP receptor protein? *Nucl. Acids Res.* 10, 473-485.

123) Koo, H. S. & Crothers, D. M. (1988). Calibration of DNA curvature and a unified description of sequence directed bending. *Proc. Natl. Acad. Sci., U.S.A.* 85, 1763-1767.

124) Kornberg, R. (1981). The location of nucleosomes in chromatin: specific or statistical? *Nature (London)*, 292, 579-580.

125) Kornberg, R. D. & Stryer, L. (1988). Statistical distribution of nucleosomes: nonrandom locations by a stochastic mechanism. *Nucl. Acids Res.* 16, 6677-6690.

126) Kornberg, R. D. (1977). Structure of chromatin. *Ann. Rev. Biochem.* 46, 931-954.

127) Krämer, H., Amouyal, M., Nordheim, A. & Müller-Hill, B. (1988). DNA supercoiling changes the spacing requirement of two *lac* operators for DNA loop formation with *lac* repressor. *EMBO J.* 7, 547-556.

128) Lambert, S. F. & Thomas, J. O. (1986). Lysine-containing DNA-binding regions on the surface of the histone octamer in the nucleosome core particle. *Eur. J. Biochem.* 160, 191-201.

129) Langmore, J. P. & Schutt, C. (1980). The higher order structure of chicken erythrocyte chromosomes *in vivo*. *Nature (London)*, 288, 620-622.

130) Laskey, R. A. & Earnshaw, W. C. (1980). Nucleosome assembly. *Nature (London)*, 286, 763-767.

131) Laskey, R. A., Honda, B. M., Mills, A. D. & Finch, J. T. (1978). Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA, *Nature (London)*, 275, 416-420.

132) Laundon, C. H. & Griffith, J. D. (1988). Curved helix segments can uniquely orient the topology of supertwisted DNA. *Cell*, 52, 545-549.

133) Levene, S. D. & Zimm, B. H. (1989). Understanding the anomalous electrophoresis of bent DNA molecules: a reptation model. *Science*, 245, 396-399.

134) Linsey, G. G. & Thompson, P. (1989). Isolation and characterisation of a 167bp core particle isolated from stripped chicken erythrocyte chromatin. *Bioch. Biophys. Act.* 1009, 257-263.

135) Linxweiler, W. & Hörz, W. (1985). Reconstitution experiments show that sequence specific histone-DNA interactions are the basis for nucleosome phasing on mouse satellite DNA. *Cell*, 42, 281-290.

136) Liu, L. F. & Wang, J. C. (1987). Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci., U.S.A.* 84, 7024-7027.

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

138) Lorch, Y., LaPointe, J. W. & Kornberg, R. D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell*, 49, 203-210.

139) Lorch, Y., LaPointe, J. W. & Kornberg, R. D. (1988). On the displacement of histones from DNA by transcription. *Cell*, 55, 743-744.

140) Losa, R. & Brown, D. D. (1987). A bacteriophage RNA polymerase transcribes *in vitro* through a nucleosome core without displacing it. *Cell*, 50, 801-808.

141) Losa, R., Omari, S. & Thoma, F. (1990). Poly(dA).poly(dT) rich sequences are not sufficient to exclude nucleosome formation in a constitutive yeast promoter. *Nucl. Acids Res.* 18, 3495-3502.

142) Lutter, L. C. (1978). Kinetic analysis of deoxyribonuclease I cleavages in the nucleosome core: evidence for a DNA superhelix. *J. Mol. Biol.* 124, 391-420.

143) Lutter, L. C. (1989). Thermal unwinding of simian virus 40 transcription complex DNA. *Proc. Natl. Acad. Sci., U.S.A.* 86, 8712-8716.

144) Lyamichev, V. I., Mirkin, S. M. & Frank-Kamenetskii, M. D. (1985). A pH-dependent structural transition in the homopurine-homopyrimidine tract in superhelical DNA. *J. Biomolec. Struct. Dyn.* 3, 327-338.

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

146) Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

147) Manning, G. S. (1978). The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Quart. Rev. Biophys.* 11, 179-246.

148) Margot, J. B. & Hardison, R. C. (1985). DNase I and nuclease S1 sensitivity of the rabbit β1 globin gene in nuclei and in supercoiled plasmids. *J. Mol. Biol.* 184, 195-210.

149) Marini, J., Levene, S., Crothers, D. M. & Englund, P. T. (1982). Bent helical structure in kinetoplast DNA. *Proc. Natl. Acad. Sci., U.S.A.* 79, 7664-7668.

150) Maxson, R., Cohn, R., Kedes, L. & Mohun, T. (1983). Expression and organization of histone genes. *Ann. Rev. Gen.* 17, 239-277.

151) McCall, M. J., Brown, T. & Kennard, O. (1985). The crystal structure of d(GGGGCCCC): a model for poly(dG).poly(dC). *J. Mol. Biol.* 183, 385-396.

152) McClarin, J. A., Frederick, C. A., Wang, B. C., Greene, P., Boyer, H. W., Grable, J. & Rosenberg, J. M. (1986). Structure of the DNA-*Eco* RI endonuclease recognition complex at 3Å resolution, *Science*, 234, 1526-1541.

153) McGhee, J. D. & Felsenfeld, G. (1980). The number of charge-charge interactions stabilizing the ends of nucleosome DNA. *Nucl. Acids Res.* 8, 2751-2769.

154) McGhee, J. D. & Felsenfeld, G. (1982). Reconstitution on nucleosome core particles containing glycosylated DNA. *J. Mol. Biol.* 158, 685-698.

155) McGhee, J.D. & Felsenfeld, G. (1980). Nucleosome structure. *Ann. Rev. Biochem.* 49, 1115-1156.

156) McGhee, J.D., Nickol, J.M., Felsenfeld, G. & Rau, D.C. (1983). Higher order structure of chromatin: orientation of nucleosomes within the 30nm solenoid is independent of species and spacer length. *Cell*, 33, 831-841.

157) McGhee, J.D., Rau, D.C., Charney, E. & Felsenfeld, G. (1980). Orientation of the nucleosome within the higher order structure of chromatin. *Cell*, 22, 87-96.

158) McGhee, J. D. & von Hippel, P. H. (1974). Theoretical aspects of DNA-protein interactions: co-operative and non-cooperative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.* 86, 469-489.

159) Menzel, R. & Gellert, M. (1987). Modulation of transcription by DNA supercoiling: a deletion analysis of the *Escherichia coli gyrA* and *gyrB* promoters. *Proc. Natl. Acad. Sci., U.S.A.* 84, 4185-4189.

160) Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A. & Frank-Kamenetskii, M. D. (1987). DNA H form requires a homopurine-homopyrimidine mirror repeat. *Nature (London)*, 330, 495-497.

161) Mirkovitch, J., Gasser, S.M. & Laemmli, U.K. (1988). Scaffold attachment of DNA loops in metaphase chromosomes. *J. Mol. Biol.* 200, 101-109.

162) Mirkovitch, J., Mirault, M. & Laemmli, U.K. (1984). Organisation of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell*, 39, 223-232.

163) Mirzabekov, A. D., Shick, V. V., Balyavsky, A. V. & Bavykin, S.G. (1987). Primary organization of nucleosome core particles of chromatin: sequence of histone arrangement along DNA, *Proc. Natl. Acad. Sci., U.S.A.* 75, 4184-4188.

164) Mirzabekov, A.D., Pruss, D.V. & Ebralidse, K.K. (1990). Chromatin superstructure-dependent crosslinking with DNA of the histone H5 residues Thr1, His25 and His62. *J. Mol. Biol.* 211, 479-491.

165) Mizuuchi, K. & Mizuuchi, M. (1979). *Cold Spring Harbor Symp. Quant. Biol.* 43, 1111-1114.

166) Monro, D. M. (1980). Interpolation methods for surface mapping. *Compu. Program. Biomed.* 11, 145-157.

167) Morse, R. H. (1989). Nucleosomes inhibit both transcription initiation and elongation by RNA polymerase III *in vitro*. *EMBO J.* 8, 2343-2351.

168) Morse, R. H., Pederson, D. S., Dean, A. & Simpson, R. T. (1987). Yeast nucleosomes allow thermal untwisting of DNA. *Nucl. Acids Res.* 15, 10311-10330.

169) Murray, K., Vidali, G. & Neelin, J. M. (1968). Stepwise removal of histones from chicken erythrocyte nucleoprotein. *Biochem. J.* 107, 207-215.

170) Muzard, G. Théveny, B. & Révet, B. (1990). Electron microscopy mapping of pBR322 DNA curvature. Comparison with theoretical models. *EMBO J.* 9, 1289-1298.

171) Nadeau, J. G. & Crothers, D. M. (1989). Structural basis for DNA bending. *Proc. Natl. Acad. Sci., U.S.A.* 86, 2622-2626.

172) Negri, R., Costano, G., Venditti, S. & Di Mauro, E. (1989). Linkage reduction allows reconstitution of nucleosomes on DNA microdomains. *J. Mol. Biol.* 207, 615-619.

173) 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 (London)*, 330, 221-226.

174) Nelson, T., Wiegand, R. & Brutlag, D. (1981). Ribonucleic acid and other polyanions facilitate chromatin assembly *in vitro*. *Biochemistry*, 20, 2594-2601.

175) Neubauer, B., Linxweiler, W. & Hörz, W. (1986). DNA engineering shows that nucleosome phasing on the African Green Monkey α-satellite is the result of multiple additive histone-DNA interactions. *J. Mol. Biol.* 190, 639-645.

176) Nickol, J. & Martin, R. G. (1983). DNA stem-loop structures bind poorly to histone octamer cores. *Proc Natl. Acad. Sci., U.S.A.* 80, 4669-4673.

177) Nickol, J., Behe, M. & Felsenfeld, G. (1982). Effect of the B-Z transition in poly(dG-^{5}mdC).poly(dG-^{5}mdC) on nucleosome formation. *Proc. Natl. Acad. Sci., U.S.A.* 79, 1771-1775.

178) Noll, M. & Kornberg, R. D. (1977). Action of micrococcal nuclease on chromatin and the location of histone H1. *J. Mol. Biol.* 109, 393-404.

179) Ostrander, E. A., Benedetti, P. & Wang, J. C. (1990). Template supercoiling by a chimera of yeast *GAL4* protein and phage T7 RNA polymerase. *Science*, 249, 1261-1265.

180) Otwinowski, Z., Schevitz, R. W., Zhang, R. G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F. & Sigler, P. B. (1988). Crystal structure of *trp* repressor/operator complex at atomic resolution. *Nature (London)*, 335, 321-329.

181) Pabo, C. O., Krovatin, W., Jeffrey, A. & Sauer, R. T. (1982). The N-terminal arms of λ-repressor wrap around the operator DNA. *Nature (London)*, 298, 441-443.

182) Panyutin, I. G., Kovalsky, O. I., Budowsky, E. I., Dickerson, R. E., Rikhirev, M. E. & Lipanov, A. A. (1990). G-DNA: a twice-folded DNA structure adopted by single-stranded oligo(dG) and its implications for telomeres. *Proc. Natl. Acad. Sci., U.S.A.* 87, 867-870.

183) Paulson, J. R. & Laemmli, U. K. (1977). The structure of histone depleted metaphase chromosomes. *Cell*, 12, 817-828.

184) Peck, L. J. & Wang, J. C. (1981). Sequence dependence of the helical repeat of DNA in solution. *Nature (London)*, 292, 375-378.

185) Peck, L. J. & Wang, J. C. (1983). Energetics of B-to-Z transition in DNA. *Proc. Natl. Acad. Sci., U.S.A.* 80, 6206-6210.

186) Pederson, D. S. & Morse, R. H. (1990). Effect of transcription of yeast chromatin on DNA topology *in vivo*. *EMBO J.* 9, 1873-1881.

187) Pederson, D. S., Shupe, K., Gorovsky, M. A. (1984). Changes in chromatin structure accompany modulation of the rate of transcription of 5S ribosomal genes in *Tetrahymena*. *Nucl. Acids Res.* 12, 8489-8507.

188) Pennings, S., Muyldermans, S., Meersseman, G. & Wyns, L. (1989). Formation, stability and core histone positioning of nucleosomes reassembled on bent and other nucleosome-derived DNA. *J. Mol. Biol.* 207, 183-192.

189) Perlmann, T. & Wrange, Ö. (1988). Specific glucocorticoid receptor binding to DNA reconstituted in a nucleosome. *EMBO J.* 7, 3073-3079.

190) Pettijohn, D. E. & Pfenninger, O. (1980). Supercoils in prokaryotic DNA restrained *in vivo*. *Proc. Natl. Acad. Sci., U.S.A.* 77, 1331-1335.

191) Pettijohn, D. E. (1988). Histone-like proteins and bacterial chromosome structure. *J. Biol. Chem.* 263, 12793-12796.

192) Phi-Van, L. & Strätling, W. H. (1988). The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. *EMBO J.* 7, 655-664.

193) Piña, B., Barenttino, D., Truss, M. & Beato, M. (1990). Structural features of a regulatory nucleosome. *J. Mol. Biol.* 216, 975-990.

194) Piña, B., Brüggemeier, U. & Beato, M. (1990). Nucleosome positioning modulates accessibility of regulatory proteins to the mouse mammary tumor virus promoter. *Cell*, 60, 719-731.

195) Poljak, L. G., Gralla, J. D. (1987). Competition for formation of nucleosomes on fragmented SV40 DNA: a hyperstable nucleosome forms on the termination region. *Biochemistry*, 26, 295-303.

196) Pruss, G. J. & Drlica, K. (1989). DNA supercoiling and prokaryotic transcription. *Cell*, 56, 521-523.

197) Pulleyblank, D. E., Shure, M., Tang, D., Vinograd, J. & Vosberg, H. P. (1975). Action of nicking-closing enzyme on supercoiled and non-supercoiled closed circular DNA: formation of a Boltzmann distribution of topological isomers. *Proc. Natl. Acad. Sci., U.S.A.* 72, 4280-4284.

198) Rahmouni, A. R. & Wells, R. D. (1989). Stabilization of Z DNA *in vivo* by localized supercoiling. *Science*, 246, 358-363.

199) Ramsay, N. (1986). Deletion analysis of a DNA sequence that positions itself precisely on the nucleosome core. *J. Mol. Biol.* 189, 179-188.

200) 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.

201) Rau, D. C., Gellert, M., Thoma, F. & Maxwell, A. (1987). Structure of the gyrase-DNA complex as revealed by transient electric dichroism. *J. Mol. Biol.* 193, 555-569.

202) Retief, J. D. (1986). Poly(glutamic acid) promoted assembly of nucleosome cores on the histone gene quintet of *Psammechinus miliaris*. Ph.D. thesis. University of Cape Town.

203) Retief, J. D., Sewell, B. T. & von Holt, C. (1987). Nucleosome cores assembled *in vitro* occupy two preferred frames flanking the histone H1 gene from *Psammechinus miliaris*. *Biochemistry*, 26, 4449-4453.

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

205) Rhodes, D. (1985). Structural analysis of a triple complex between the histone octamer, a *Xenopus* gene for 5S rRNA and transcription factor IIIA. *EMBO J.* 4, 3473-3483

206) Rhodes, D. & Klug, A. (1980). Helical periodicity of DNA determined by enzyme digestion. *Nature (London)*, 286, 573-578.

207) Rhodes, D. & Klug, A. (1981). Sequence-dependent helical periodicity of DNA. *Nature (London)*, 292, 378-380.

208) Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D. & Klug, A. (1984). *Nature (London)*, 311, 532-537.

209) Ryoji, M. & Worcell, A. (1984). Chromatin assembly in *Xenopus* oocytes: *in vivo* studies. *Cell*, 37, 21-32.

210) Satchwell, S. C., Drew, H. R. & Travers, A. A. (1986). Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.* 191, 659-675.

211) Satchwell, S. C. & Travers, A. A. (1989). Asymmetry and polarity of nucleosomes in chicken erythrocyte chromatin. *EMBO J.* 8, 229-238.

212) Selleck, S. B., Elgin, S. C. R. & Cartwright, I. L. (1984). Supercoil-dependent features of DNA structure at *Drosophila* locus 67B1. *J. Mol. Biol.* 178, 17-33.

213) 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. Mol. Biol.* 166, 183-201.

214) Shea, R. G., Ng, P. & Bischofberger, N. (1990). Thermal denaturation profiles and gel mobility shift analysis of oligodeoxynucleotide triplexes. *Nucl. Acids Res.* 18, 4859-4866.

215) Shure, M., Pulleyblank, D. E. & Vinograd, J. (1977) The problems of eukaryotic and prokaryotic DNA packaging and *in vivo* conformation posed by superhelix density heterogeniety. *Nucl. Acids Res.* 4, 1183-1205.

216) Siegfried, E. Thomas, G. H., Bond, U. M. & Elgin, S. C. R. (1986). Characterization of a supercoil-dependent S1 sensitive site 5' to the *Drosophila melanogaster hsp26* gene. *Nucl. Acids Res.* 14, 9425-9444.

217) Simpson, R. T. (1986). Nucleosome positioning *in vivo* and *in vitro*. *Bioessays*, 4, 172-176.

218) Simpson, R. T. (1990). Nucleosome positioning can affect the function of a *cis*-acting DNA element *in vivo*. *Nature (London)*, 343, 387-389.

219) Simpson, R. T. & Künzler, P. (1979). Chromatin and core particles formed from the inner histones and synthetic polydeoxyribonucleotides of defined sequence. *Nucl. Acids Res.* 6, 1387-1415.

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

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

222) Sinden, R. R., Broyles, S. S. & Pettijohn, D. E. (1983). Perfect palindromic *lac* operator DNA sequence exist as a stable cruciform in supercoiled DNA *in vitro* but not *in vivo*. *Proc. Natl. Acad. Sci., U.S.A.* 80, 1797-1801.

223) Sinden, R. R., Carlson, J. O. & Pettijohn, D. E. (1980). Torsional tension in the DNA double helix measured with trimethylpsoralen in living *E. coli* cells: analogous measurements in insect and human cells. *Cell*, 21, 773-783.

224) Sinden, R. R. & Kochel T. J. (1987) Reduced 4,5′,8-trimethylpsoralen cross-linking of left-handed Z-DNA stabilized by DNA supercoiling. *Biochemistry*, 26, 1343-1350.

225) Spengler, S. J., Stasiak, A. & Cozzarelli, N. R. (1985). The stereostructure of knots and catenanes produced by phage λ integrative recombination: implications for mechanism and DNA structure. *Cell*, 42, 325-334.

226) Srinivasan, A. R., Torres, R., Clark, W., Olson, W. K. (1987). Base sequence effects in double helical DNA. I. Potential energy estimates of local base morphology. *J. Biomolec. Struct. Dynam.* 5, 459-496.

227) Stein, A. (1979), DNA folding by histones: the kinetics of chromatin core particle reassembly and the interaction of nucleosomes with histones. *J. Mol. Biol.* 130, 103-134.

228) Stein, A., Whitlock, J. P. & Bina, M. (1979). Acidic polypeptides can assemble both histones and chromatin *in vitro* at physiological ionic strength. *Proc. Natl. Acad. Sci., U.S.A.* 76, 5000-5004.

229) Stellwagen, N. C. (1983). Anomalous electrophoresis of deoxyribonucleic acid restriction fragments on polyacrylamide gels. *Biochemistry*, 22, 6186-6193.

230) Suau, P., Bradbury, E. M. & Baldwin, J. P. (1979). Higher-order structures of chromatin in solution. *Eur. J. Biochem.* 97, 593-602.

231) Suck, D. & Oefner, C. (1986). Structure of DNase I at 2,0Å resolution suggests a mechanism for binding to and cutting DNA. *Nature (London)*, 321, 620-625.

232) Thoma, F. (1986). Protein-DNA interactions and nuclease-sensitive regions determine nucleosome positioning in yeast plasmid chromatin. *J. Mol. Biol.* 190, 177-190.

233) Thoma, F. (1988). Mechanism of nucleosome positioning tested in minichromosomes of yeast *Saccharomyces cerevisiae*. In *Architecture of Eukaryotic Genes* (Kahl, G., ed), pp 269-280, VCH Verlagsgesellschaft, FRG.

234) Thoma, F. (1988). The role of histone H1 in nucleosomes and chromatin fibres. In *Architecture of Eukaryotic Genes* (Kahl, G., ed), pp 163-185, VCH Verlagsgesellschaft, FRG.

235) Thoma, F., Bergman, L. W. & Simpson, R. T. (1984). Nuclease digestion of circular TRP1ARS1 chromatin reveals positioned nucleosomes separated by nuclease sensitive regions. *J. Mol. Biol.* 177, 715-733.

236) Thoma, F., Koller, T. & Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J. Cell. Biol.* 83, 403-427.

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

238) Thoma, F. & Zatchej, M. (1988). Chromatin folding modulates nucleosome positioning in yeast minichromosomes. *Cell*, 55, 945-953.

239) Thomas, J. O. (1983). Chromatin structure and superstructure. In *Eukaryotic Genes. Their Structure, Activity and Regulation* (McLean, N., Gregory, S. P. & Flavell, R. A., eds), pp 9-30, Butterworths, London.

240) Thomas, J. O. & Butler, P. J. G. (1980). Size-dependence of a stable higher-order structure of chromatin. *J. Mol. Biol.* 144, 89-93.

241) Thomas, J. O. & Wilson, C. M. (1986). Selective radiolabelling and identification of a strong nucleosome binding site on the globular domain of histone H5. *EMBO J.* 5, 3531-3537.

242) Thomas, T. J. & Bloomfield, V. A. (1983). Chain flexibility and hydrodynamics of the B and Z forms of poly(dG-dC).poly(dG-dC). *Nucl. Acids Res.* 11, 1919-1930.

243) Travers, A. A. (1987). DNA bending and nucleosome positioning. *Trends Biochem. Sci.* 12, 108-112.

244) Travers, A. A. (1988). Protein-induced DNA bending. In *Nucleic Acids and Molecular Biology* (Eckstein, F. & Lilley, D. M. J., eds), vol 2, pp 136-148, Springer-Verlag, Berlin.

245) Travers, A. A. & Klug, A. (1987). The bending of DNA in nucleosomes and its wider implications. *Phil. Trans. R. Soc. Lond.* B317, 537-561.

246) Travers, A. A. & Klug, A. (1990). Bending of DNA in nucleoprotein complexes. In *DNA Topology and Its Biological Effect* (Cozzarelli, N. R. & Wang, J. C., eds), pp 57-106, Cold Spring Harbor Laboratory Press, New York.

247) Tsao, Y. P., Wu, H. Y. & Liu, L. F. (191989). Transcription-driven supercoiling of DNA: direct biochemical evidence from *in vitro* studies. *Cell*, 56, 111-118.

248) Tucker, P. W., Hazen, E. E., Cotton, J. R. & Cotton, F. (1979). Staphylococcal nuclease reviewed: a prototypic study in contemporary enzymology. *Molec. Cell. Biochem.* 23, 67-86.

249) Tullius, T. D. & Dombroski, B. A. (1986). Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to λ repressor and *Cro* protein. *Proc. Natl. Acad. Sci., U.S.A.* 83, 5469-5473.

250) Turnell, W. G., Satchwell, S. C. & Travers, A. A. (1988). A decapeptide motif for binding to the minor groove of DNA. A proposal. *FEBS Lett.* 232, 263-268.

251) Uberbacher, E. C. & Bunick, G. J. (1989). Structure of the nucleosome core particle at 8Å resolution. *J. Biomol. Struct. Dynam.* 7, 1-18.

252) Ulanovsky, L. Bodner, M., Trifonov, E. N. & Choder, M. (1986). Curved DNA: design, synthesis, and circularization. *Proc. Natl. Acad. Sci., U.S.A.*, 83, 862-866.

253) Ulanovsky, L. E. & Trifonov, E. N. (1987). Estimation of wedge components in curved DNA, *Nature (London)*, 326, 720-722.

254) Van Holde, K. E. (1988). *Chromatin*, Springer-Verlag, New York.

255) Viswamitra, M. A., Kennard, O., Jones, P. C., Sheldrick, G. M., Salisbury, S., Falvello, L. & Shakked, Z. (1978). DNA double helical fragment at atomic resolution. *Nature (London)*, 273, 687-688.

256) Vojtísková, M., Mirkin, S., Lyamichev, V., Voloshin, O. & Frank-Kamenetskii, M. (1988). Chemical probing of the homopurine.homopyrimidine tract in supercoiled DNA at single nucleotide resolution. *FEBS*, 234, 295-299.

257) Vojtísková, M. & Palecek, E. (1987). Unusual protonated structure in the homopurine.homopyrimidine tract of supercoiled and linearized plasmids recognized by chemical probes. *J. Biomolec. Struct. Dynam.* 5 , 283-296.

258) Volkaert, G, Devleeschouwer, E., Blocker, H. & Frank, R. (1984). A novel type of cloning vectors for ultrarapid chemical degradation sequencing of DNA. *Gene Anal. Tech.* 1, 52-59.

259) Wang, A. H. J., Fujii, S., van Boom, J. H., van der Marel, G. A., van Boeckel, S. A. & Rich, A. (1982). Molecular structure of r(GCG)d(TATACGC): a DNA-RNA hybrid helix joined to a double-helical DNA. *Nature (London)*, 299, 601-604.

260) Wang, J. C. (1979). *Cold Spring Harbor Symp. Quant. Biol.* 43, 29-33.

261) Wang, J. C. (1979). Helical repeat of DNA in solution. *Proc. Natl. Acad. Sci., U.S.A.*, 76, 200-203.

262) Wang, J. C. (1985), DNA topoisomerases. *Ann. Rev. Biochem.* 54, 665-697.

263) Wang, J. C. (1987). Recent studies of DNA topoisomerases. *Biochem. Biophys. Act.* 909, 1-9.

264) Wang, J. C., Peck, L. J. & Becherer, K. (1983). DNA supercoiling and its effects on DNA structure and function. *Cold Spring Harbor Symp. Quant. Biol.*, 47, 85-91.

265) Wells, R. D., Collier, D. A., Hanvey, J. C., Shimiza, M. & Wohlrab, F. (1988). The chemistry and biology of unusual DNA structures adopted by oligopurine.oligopyrimidine sequences. *FASEB J.* 2, 2939-2949.

266) White, J. H. (1969). Self-linking and the Gauss integral in higher dimensions. *Am. J. Math.* 91, 693-728.

267) White, J. H. & Bauer, W. R. (1986). Calculation of the twist and the writhe for representative models of DNA. *J. Mol. Biol.* 189, 329-341.

268) White, J. H. & Bauer, W. R. (1989). The helical repeat of nucleosome-wrapped DNA. *Cell*, 56, 9-10.

269) White, J. H., Cozzarelli, N. R. & Bauer, W. R. (1988). Helical repeat and linking number of surface wrapped DNA. *Science*, 241, 323-327.

270) Widom, J. (1986). Physicochemical studies of the folding of the 100Å nucleosome filament into the 300Å filament. *J. Mol. Biol.* 190, 411-424.

271) Widom, J. (1989). Towards a unified model of chromatin folding. *Ann. Rev. Biophys. Biophys. Chem.* 18, 365-395.

272) Widom, J. Finch, J. T., Thomas, J. O. (1985). Higher-order structure of long repeat chromatin, *EMBO J.* 4, 3189-3194.

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

274) Williams, S. P., Athey, B. D., Muglia, L. J., Schappe, R. S., Gough, A. H. & Langmore, J. P. (1986). Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. *Biophys. J.* 49, 233-248.

275) Wolffe, A. P. & Drew, H. R. (1989). Initiation of transcription on nucleosomal templates. *Proc. Natl. Acad. Sci., U.S.A.*, 86, 9817-9821.

276) Woodcock, C. L. F., Frado, Y. L. L. & Rattner, J. B. (1984). The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J. Cell. Biol.* 99, 42-52.

277) Worcel, A., Gargiulo, G., Jessee, B., Udvardy, A., Louis, C. & Schedl, P. (1983). Chromatin structure of the histone gene complex of *Drosophila melanogaster*. *Nucl. Acids Res.* 11, 421-439.

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

279) Yager, T. D. & van Holde, K. E. (1984). Dynamics and equilibria of nucleosomes at elevated ionic strength. *J. Biol. Chem.* 259, 4212-4222.

280) Yager, T. D., McMurray, C. T. & van Holde, K. E. (1989). Salt-induced release of DNA from nucleosome core particles. *Biochemistry*, 28, 2271-2281.

281) Zahn, K. & Blattner, F. R. (1987). Direct evidence for DNA bending at the lambda replication origin. *Science*, 236, 416-422.

282) Zivanovic, Y., Duband-Goulet, I., Schultz, P., Stofer, E., Oudet, P. & Prunell, A. (1990). Chromatin reconstitution on small DNA rings. III. Histone H5 dependence of DNA supercoiling in the nucleosome. *J. Mol. Biol.* 214, 479-495.

283) Zivanovic, Y., Goulet, I., Revet, B., Le Bret, M. & Prunell, A. (1988). Chromatin reconstitution on small DNA rings. II. DNA supercoiling on the nucleosome. *J. Mol. Biol.* 200, 267-290.

Appendix A.

A. Predicted rotational orientation of a small circular DNA molecule.

The explanation of the sequence directed curvature of DNA in solution in terms of the conformation of individual base-steps or cooperative structural properties of short sequences has been the subject of many experimental and theoretical studies (reviewed by Diekmann (A2) and Hagerman (A4)). In one such study, by Olson and colleagues (A5), a nearest neighbor model was proposed where the preferred helical parameters (roll and tilt; twist was assumed constant at 34°) of the 16 unique dinucleotides were calculated from the electrostatic interaction, London attraction and van der Waal's repulsion of adjacent base-pairs constituting each dinucleotide.

In this appendix, the contribution of only these proposed forces to the preferred conformation of individual dinucleotides is tested by comparing the predicted rotational orientation of a 169bp *tyrT* DNA circle where the sum of the potential energies of the constituent base-steps is at a minimum, with the preferred rotational orientation of the DNA circle determined by experiment (A3).

In the discussion that follows, the helical parameters are defined to conform to the EMBO convention (A1). Note that the helical parameters used by Srinivasan *et al.* (A5) is a private convention; these were modified in this text to similarly comply with the EMBO convention.

The problem was approached by constraining the helical axis of the DNA molecule to a circle in the right-handed reference system of the first base-pair (see figure A.1). It is assumed that the energy of such a DNA conformation is

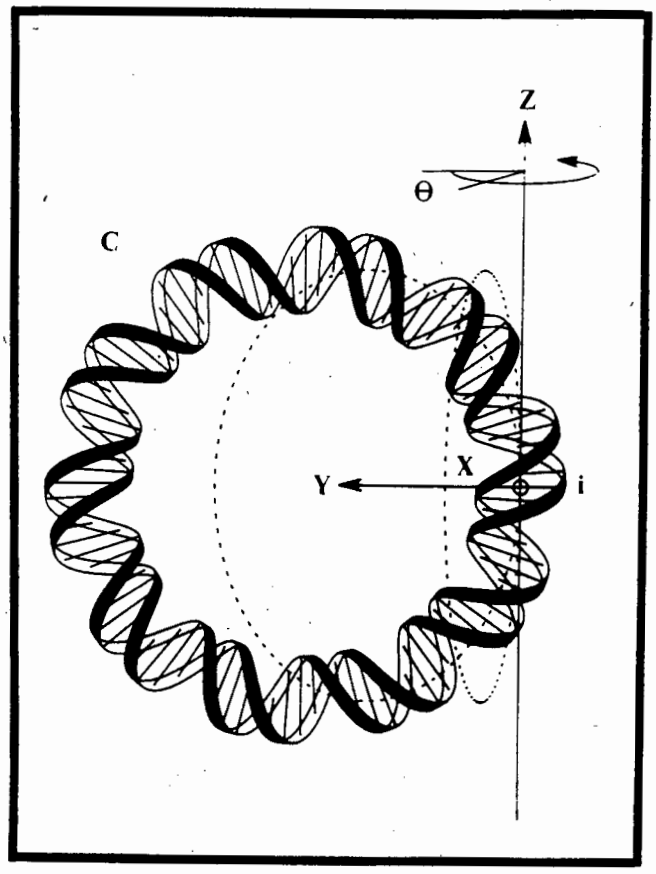

Figure A.1. Schematic representation of a writheless, circular DNA molecule defined in the right-handed orthogonal reference set of base-pair i. The anti-clockwise rotation (θ) of the molecule (out of the level of the page) in the reference set of base-pair i changes the orientation of base-pairs relative to the center of the molecule, and causes different regions of the minor groove to point to the outside of the circle. See text for detail.

equal to the cumulative energies of atomic interaction of individual base-pair dimers. (As a first approximation, the effect of the sugar-phosphate backbone is neglected). The potential energy of a single dimer in turn depends on the base-step conformation in terms of its helical parameters (once again, as an approximation, only roll (ρ), tilt (τ) and twist (Ω) are considered). With the potential energy of each of the 16 possible dimers known in a given conformation range, the energy of the spatially fixed DNA molecule can be calculated by solving for the helical parameters and adding the corresponding energies of the constituent base-pair dimers collectively placing the helix axis of the DNA molecule on the circumference of the circle. The problem is thus reduced to finding the value of roll, tilt, and twist of each base-step in the circular molecule, successively defined in progressive rotational states in the axial set of base-pair i (see Figure A.1).

Referring to Figure A.2, the base-pairs (i,i+1,...) are defined in a right-handed orthogonal reference set of the first base-pair (i), and placed with their centers on the circumference of the circle C at points $P_0,P_1,\ldots$ The base-pair centers are further connected by vectors $v_i,v_{i+1},\ldots,$ defined as normal to the plane of the preceding base-pair (i,i+1,...) and with magnitude equal to the average local rise in B-DNA ($|v_i| = |v_{i+1}| = 3.4$).

Since the vector v_{i+1} (normal to base-pair i+1) terminates at P_2, the conformation of dimer 1 will be such that the local helical axis is constrained to lie on the circle C. The deflection of v_{i+1} from co-linearity with v_i to the point P_2 is due to rotation round the $X(\tau)$, $Y(\rho)$ and $Z(\Omega)$ axes. This sequence of rotations is in the same order chosen by Srinivasan *et al.* (A5) for the calculation of their potential energy maps. To remain compatible with these energy maps and since the resulting dimer conformation

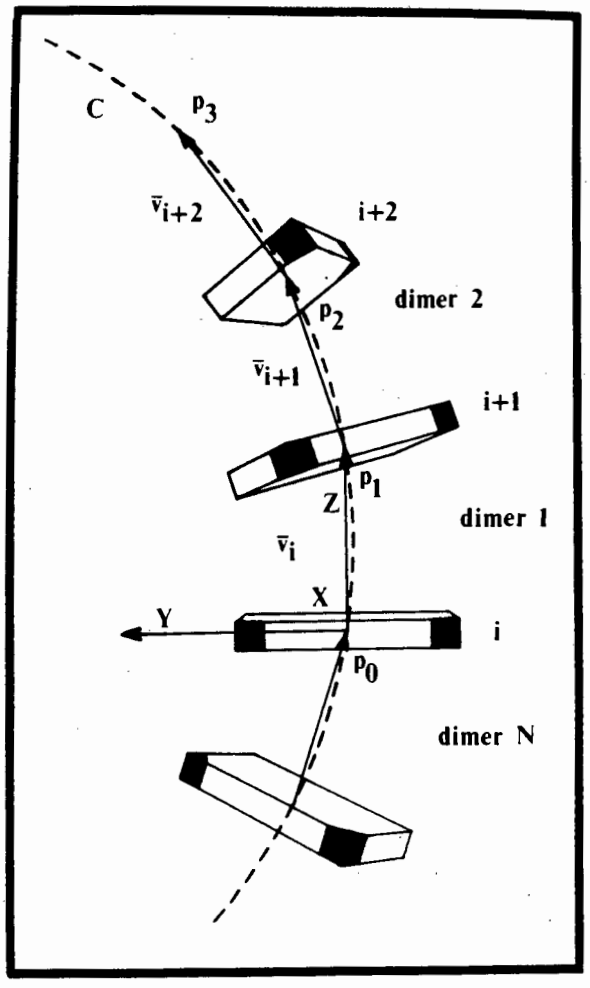

Figure A.2. Schematic representation of individual base-pairs in a circularly constrained molecule. Base-pairs (i, i+1, i+2) are represented by rectangular blocks at positions p_0, p_1 and p_2 on the circle C with the filled corners orientated towards the minor groove side. The normal vectors (v_i, v_{i+1}, v_{i+2}) to these base-pairs are indicated in the figure.

depends on the exact order by which the adjacent base-pairs are rotated, the helical parameters must be solved in exactly the reverse order as indicated in Figure A.3. Thus, given the coordinates of P_2, values for τ, ρ and Ω must be found which, when applied in the bottom set of operations in Figure A.3, gives v'''_{i+1} co-linear with v_i. In attempting the same approach to the next base-step in the sequence, dimer 2, a problem immediately becomes obvious. Since the circle C and therefore points $P_0, P_1, \ldots$ and vectors $v_i, v_{i+1}, \ldots$ are defined in the axial set of base-pair i (which does not correspond to the axial set of base-pair i+1), all points must be transformed to the coordinate system of base-pair i+1 using the calculated values of τ, ρ and Ω before solving for the helical parameters of dimer 2. This is rather tedious, but because the curvature of the helical axis is uniform, it is also true to simply rotate v_{i+1} in a clock-wise sense by the local helical twist in the reference system of base-pair i. The local twist of the duplex is given by the expression:

$$\Omega = (2\pi Lk_0)/N \qquad (A.1)$$

where Lk_0 is the linking number of the writheless molecule, and N is the number of base-pairs in the circular molecule. The clock-wise rotation of v_{i+1} by Ω gives a vector in the axial set of base-pair i which is equivalent to v_{i+2} in the axial set of i+1. In the general case, for axial set i+n-1:

$$v_{i+n} = Z^{-1}(\sigma_n).v_s \qquad \text{for } n=1,2,\ldots,N-1,N \qquad (A.2)$$

where n is the relevant dimer and $Z^{-1}(\sigma_n)$ is a rotation matrix as a function of σ_n which is the accumulated duplex winding angle at v_{i+n} relative to v_i, given by $\sigma_n = \Omega n$. The fullstop (.) indicates matrix multiplication. The vector v_s is a "seed vector" given by:

$$v_s = Z(\theta).X^{-1}(\phi).v_i \qquad (A.3)$$

$Z(\theta)$ and $X^{-1}(\phi)$ are rotation matrixes in terms of θ and ϕ respectively, where θ is the inclination of the circle C relative to the YZ-plane of axial set i as shown in figure A.1, $\phi = 2\pi/N$ and $v_i = (0,0,3.4)$. Once again the fullstop denotes matrix multiplication.

$$v_i \equiv v'''_{i+1} \underset{X(\tau)}{\overset{X^{-1}(\tau)}{\rightleftharpoons}} v''_{i+1} \underset{Y(\rho)}{\overset{Y^{-1}(\rho)}{\rightleftharpoons}} v'_{i+1} \underset{Z(\Omega)}{\overset{Z^{-1}(\Omega)}{\rightleftharpoons}} v_{i+1}$$

Figure A.3. Order of rotation of base-pairs. $X(\tau)$, $Y(\rho)$ and $Z(\Omega)$ represent rotation matrixes of the tilt, roll and twist angles respectively. v_i and v_{i+1} are vectors normal to base-pairs i and i+1 respectively. See text for detail.

Note that the local helical twist (Ω) given by equation A.1 is constant. This is imperative since no unique solution exists for three rotational parameters given only the coordinates of v_{i+n} before and after deflection. With Ω known, the coordinates of v_{i+n} can be calculated before twist was applied, giving vector v'_{i+n} (see Figure A.3),

$$v'_{i+n} = Z(\Omega).v_{i+n} \qquad (A.4)$$

where $Z(\Omega)$ is a rotation matrix in terms of the local twist, and v_{i+n} the vector normal to the plane of base-pair i+n with coordinates as calculated in equation A.2. It further follows from figure A.3 that v'_{i+n} rotated through ρ_n (giving v''_{i+n}) and τ_n will give a vector v'''_{i+n} co-linear with v_i and magnitude 3.4. Because of the sequence of rotations by which the vector v'''_{i+n} is defined (see figure A.3), ρ must be solved before τ. The roll angle cannot simply be found from the scalar product of v''_{i+n} and

v'_{i+n} since not all the coordinates of v''_{i+n} are known. What is known about this vector is that it resulted from a rotation of the vector (0,0,3.4) round the X(tilt) axis; thus $x''_{i+n} = 0$. From the application of the simple trigonometric relation,

$$x'_{i+n}.\cos(\rho_n) + z'_{i+n}.\sin(\rho_n) = 0 \qquad (A.5)$$

it follows that

$$\rho_n = \tan^{-1}(x'_{i+n}/z'_{i+n}) \qquad (A.6)$$

Both the coordinates y''_n and z''_n can then be found with the solved value of ρ_n. Since the coordinates of v'''_n = (0,0,3.4), the value of tilt can be calculated in an identical manner to the above, and is given by:

$$\tau = -\tan^{-1}(y''_{i+n}/z''_{i+n}) \qquad (A.7)$$

Thus, with n = 1 for equations A.2, A.6 and A.7, the values of the helical parameters (τ_1, ρ_1 and Ω_1) for dimer 1 have been found. The next base-step in the sequence (i+1,i+2) can now be tackled by calculating the coordinates of v_{i+2} in the reference system of i+1 with equation A.1 above and the helical parameters solved with equation A.6 and A.7 with n = 2, and so on all the way round the circle up to n = N. Once the helical parameters for all the base-steps in the sequence have been calculated, the potential energies of the individual base-steps can be found by determining the chemical identity (i.e. AG, AA, etc.) of the base-step and looking up the potential energy of the relevant base-step at the calculated conformation on an energy contour map. I have recalculated the energy maps of all 16 DNA dimers at 0.1° intervals for $-5° \leq \rho \geq 5°$ and $-5° \leq \tau \geq 5°$ in terms of London attraction, van der Waal's repulsion and electrostatic

interactions of constituent atoms as explained fully by Srinivasan *et al.* (A5).

The summation of the energies of all the base-steps in the circular sequence gives the potential energy of the DNA circle in that particular rotational orientation. The final object, however, is to find the DNA conformation at the global energy minimum of rotational states of the circularly constrained molecule. The potential energy of the molecule is therefore calculated in different rotational orientations by successively incrementing θ in equation 2 from 0 to 2π. The effect of incrementing θ is a rotation of circle C, on which the helical axis is defined to lie, round the Z-axis. Referring to figure A.1, it is clear that this rotation is in the orthogonal reference system of base-pair i. Since adjacent base-pairs are defined relative to base-pair i, the result of such a rotation is that the base-steps underlying the minor groove pointing towards the center of the circle at $\theta = 0$, will have their minor grooves pointing towards the outside at $\theta = \pi$. It is thus equivalent to rotating the DNA helix round the toroidal axis of the circle. If the potential energy of each rotational state is plotted versus the value of θ for that rotational conformation, inspection of the graph for $0 \leq \theta \leq 2\pi$ would indicate if one rotational state is energetically most favorable. This rotational orientation in then predicted only in terms of atomic interactions of adjacent base-pairs, and comparison with the experimentally determined rotational orientation of the circular DNA molecule would provide a test of the Srinivasan-Olson energy surfaces (A5) in governing the preferred angular orientation of dinucleotides in solution.

To this end the 169bp *tyrT* DNA sequence, for which the rotational orientation in solution was determined by Drew and Travers (A3) was tested with the above algorithm on a microcomputer.

The linking number (Lk_0) was set to 16 yielding a local twist (Ω_n) of 34.1° (from equation A.2) and the energy calculated as explained above at 1° intervals of θ between 0 and 2π. The result is shown in figure A.4.

It is seen in figure A.4 that at a rotational orientation of θ = 297°, the circle is at a minimum potential energy in terms of the conformation of the constituent dinucleotides (according to the energy contour maps of Srinivasan *et al.* (A5)). At this rotational orientation, the minor groove at positions 1, 11, 22, 33, 43, 54, 64, 75, 85, 96, 106, 117, 128, 138, 149 and 159 is orientated towards the outside of the circle. In the study by Drew and Travers (A3) positions 14, 24, 35, 45, 56, 66, 77, 87, 98, 108, 119, 129, and 140 were found to be orientated towards the outside of the circle, judged by the accessibility of DNase I to nucleotides staggered by one bond in the 3′ direction relative to these sites. Although the average correlation between the listed sites is within 2bp, this still corresponds to a 68° difference in the predicted and experimental rotational orientation of the circular DNA molecule. Thus, although the agreement is encouraging, the clear difference suggests that forces other than those used in the calculation of the energy contour maps (A5) must contribute to the angular preferences of dinucleotides.

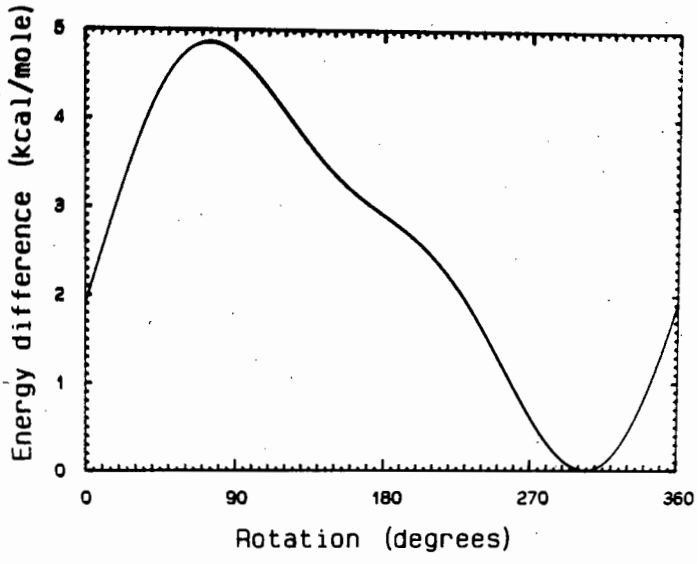

Figure A.4. The potential energy of the 169bp tyrT DNA circle calculated from the energy contour maps of Srinivasan *et al.* (A5) at different rotational orientations as explained in the text.

References cited in Appendix A:

A1) Dickerson, R. E., Bansal, M., Calladine, C. R., Diekmann, S., Hunter, W. N., Kennard, O., von Kitzing, E., Lavery, R., Nelson, H., Olson, W. K., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., Tung, C. S., Wang, A. H. J. & Zhurkin, V. B. (1989). Definitions and nomenclature of nucleic acid structure parameters. *EMBO J.* 8 , 1-4.

A2) Diekmann, S. (1987). DNA curvature. In *Nucleic Acids and Molecular Biology* (Eckstein, F. & Lilley, D. M. J., eds), vol 1, pp138-156, Springer-Verlag, Berlin.

A3) Drew, H. R. & Travers, A. A. (1985). DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* 186, 773-790.

A4) Hagerman, P. J. (1990). Sequence-directed curvature of DNA. *Ann. Rev. Biochem.* 59, 755-781.

A5) Srinivasan, A. R., Torres, R., Clark, W. & Olson, W. K. (1987). Base sequence effects in double helical DNA. I. Potential energy estimates of local base morphology. *J. Biomolec. Struct. Dynam.* 5, 459-496.

Appendix B

B. Program listing of *CUBESPLN*:

```
 10 REM:_____________________________________________
 20 REM:
 30 REM:                    'CUBESPLN'
 40 REM:
 50 REM:          For ACORN Archimedes 310 and up
 60 REM:_____________________________________________
 70 :
 80 MODE0
 90 PROCload_datafile
100 PROCinit
110 :
120 loop%=0
130 WHILE loop%<=datum_points%-1
140   PROCcubic_spline_coefficients
150   PROCinterpolate
160   loop%+=1
170 ENDWHILE
180 :
190 PROCsave_surface
200 :
210 END
220 :
230 DEF PROCload_datafile
240 INPUT"Enter number of scans: "N%
250 INPUT"Enter number of datapoints per scan:"
datum_points%
260 DIM X(N%-1), Y(datum_points%-1,N%-1)
270 FORscan%=0TON%-1
280   PRINT"Enter name of datafile #";scan%+1;": ";:INPUT""
file$
290   PRINT"Enter value associated with scan";scan%+1;":";:
INPUT""X(scan%)
300   filename$=":0.$.DENS_FILES.SMOOTHED.LANE"+file$
310   A=OPENUP filename$
320   FORX%=0TOdatum_points%-1
330     INPUT#A,Y(X%,scan%)
340   NEXT
350   CLOSE#A
360 NEXT
370 CLS
380 FORX%=0TON%-1
390   FORY%=0TOdatum_points%-1
400     IFY%=0:MOVEY%,Y(Y%,X%):ELSEDRAWY%,Y(Y%,X%)
410   NEXT
420 NEXT
430 ENDPROC
440 :
```

```
450 DEF PROCinit
460 I=0.1
470 DIM D(N%-1), F(N%-2), G(N%-2), H(N%-2)
480 M%=(X(N%-1)-X(0))/I
490 DIM P(datum_points%-1,M%)
500 PRINTTAB(30,25)"Calculating ..."
510 ENDPROC
520 :
530 DEF PROCcubic_spline_coefficients
540 D%=1
550 A=FNA_coeff(D%):B=FNB_coeff(D%,A)
560 D(0)=FND_value(0,A,B):D(1)=FND_value(1,A,B)
570 D%=2
580 REPEAT
590   A=FNA_coeff(D%):B=FNB_coeff(D%,A)
600   D(D%)=FND_value(D%,A,B)
610   D%+=1
620 UNTIL D%=N%-1
630 D(N%-1)=FND_value(N%-1,A,B)
640 D%=0
650 REPEAT
660   F(D%)=FNF_coeff(D%)
670   G(D%)=FNG_coeff(D%)
680   H(D%)=FNH_coeff(D%)
690   D%+=1
700 UNTIL D%=N%-1
710 ENDPROC
720 :
730 DEF PROCinterpolate
740 C%=0
750 T=X(0)
760 FORX%=0TOM%-1
770   X=T+X%*I
780   IF X>=(X(C%+1)):C%+=1
790   W=FNW_coeff(X,C%)
800   P(loop%,X%)=FNV_coeff(W,C%)
810 NEXT
820 W=FNW_coeff(X(N%-1),C%)
830 P(loop%,M%)=FNV_coeff(W,C%)
840 ENDPROC
850 :
860 DEF PROCsave_surface
870 VDU7:CLS
880 PRINT"NOTE: The data points are saved at 5 point
intervals"
890 INPUT"Enter filename for surface: "file$
900 filename$=":0.$.CUBESPLN.SURFACE."+file$
910 T=X(0)
920 A=OPENOUT filename$
```

```
 930 FORX%=0TOM%
 940   X_value=T+X%*I
 950   PRINT#A,X_value
 960   FORY%=0TOdatum_points%-1STEP5
 970     PRINT#A,P(Y%,X%)
 980   NEXT
 990 NEXT
1000 CLOSE#A
1010 ENDPROC
1020 :
1030 DEF FNA_coeff(E%)
1040 T=(Y(loop%,E%-1)-Y(loop%,E%))/(X(E%-1)-X(E%))
1050 T-=(Y(loop%,E%)-Y(loop%,E%+1))/(X(E%)-X(E%+1))
1060 =T/(X(E%-1)-X(E%+1))
1070 :
1080 DEF FNB_coeff(E%,U)
1090 T=(Y(loop%,E%-1)-Y(loop%,E%))/(X(E%-1)-X(E%))
1100 =T-U*(X(E%-1)+X(E%))
1110 :
1120 DEF FND_value(E%,T,U)
1130 =2*T*X(E%)+U
1140 :
1150 DEF FNF_coeff(E%)
1160 =D(E%)*(X(E%+1)-X(E%))
1170 :
1180 DEF FNG_coeff(E%)
1190 T=3*Y(loop%,E%+1)-D(E%+1)*(X(E%+1)-X(E%))
1200 =T-3*Y(loop%,E%)-2*F(E%)
1210 :
1220 DEF FNH_coeff(E%)
1230 =Y(loop%,E%+1)-Y(loop%,E%)-F(E%)-G(E%)
1240 :
1250 DEF FNW_coeff(T,E%)
1260 =(T-X(E%))/(X(E%+1)-X(E%))
1270 :
1280 DEF FNV_coeff(T,E%)
1290 =Y(loop%,E%)+F(E%)*T+G(E%)*W^2+H(E%)*W^3
```

Appendix C

C Program listing of *FINDFIT*:

```
 10 REM:_____________________________________________________
 20 REM:
 30 REM:
 40 REM:                        'FINDFIT'
 50 REM:
 60 REM:              For ACORN Archimedes 310 and up
 70 REM:
 80 REM:_____________________________________________________
 90 :
100 MODE12
110 DIM SUP(198,110), LIN(198,110), a(198), b(198),
diff(198), sqr(198), least_sqr(110), min(110)
120 :
130 INPUT"Enter the filename of the reference surface:"file$
140 filename$=":0.$.CUBESPLN.SURFACE."+file$
150 A=OPENUP filename$
160 FOR X%=0TO110
170   INPUT#A,drain
180   FOR Y%=0TO198
190     INPUT#A,SUP(Y%,X%)
200   NEXT
210 NEXT
220 CLOSE#A
230 :
240 INPUT"Enter the filename of the fitting surface: "file$
250 filename$=":0.$.CUBESPLN.SURFACE."+file$
260 A=OPENUP filename$
270 FOR X%=0TO110
280   INPUT#A,drain
290   FOR Y%=0TO198
300     INPUT#A,LIN(Y%,X%)
310   NEXT
320 NEXT
330 CLOSE#A
340 :
350 CLS
360 FORX%=0TO110
370   FORY%=0TO198
380     IF Y%=0:MOVE5*Y%,SUP(Y%,X%):ELSEDRAW5*Y%,SUP(Y%,X%)
390   NEXT
400 NEXT
410 GCOL3
420 FORX%=0TO110
430   FORY%=0TO198
440     IF Y%=0:MOVE5*Y%,LIN(Y%,X%):ELSEDRAW5*Y%,LIN(Y%,X%)
450   NEXT
460 NEXT
470 :
```

```
480 FORZ%=0TO110STEP10
490   FORX%=0TO110
500     FORY%=0TO198
510       a(Y%)=SUP(Y%,X%)
520       b(Y%)=LIN(Y%,Z%)
530     NEXT
540     diff()=a()-b()
550     k_min=-(SUM(diff()))/198
560     diff()+=(k_min)
570     sqr()=diff()*diff()
580     least_sqr(X%)=SUM(sqr())
590   NEXT
600   min=least_sqr(0)
610   FORX%=0TO110
620     IFleast_sqr(X%)<min:min=least_sqr(X%):min(Z%)=X%
630   NEXT
640 NEXT
650 :
660 INPUT"Enter filename for best fit data: "file$
670 filename$="$.leastsq.res_"+file$
680 A=OPENOUT filename$
690 FORX%=0TO110STEP10
700   PRINT#A,min(X%)
710   PRINTX%;" ";min(X%)
720 NEXT
730 CLOSE#A
740 END
```