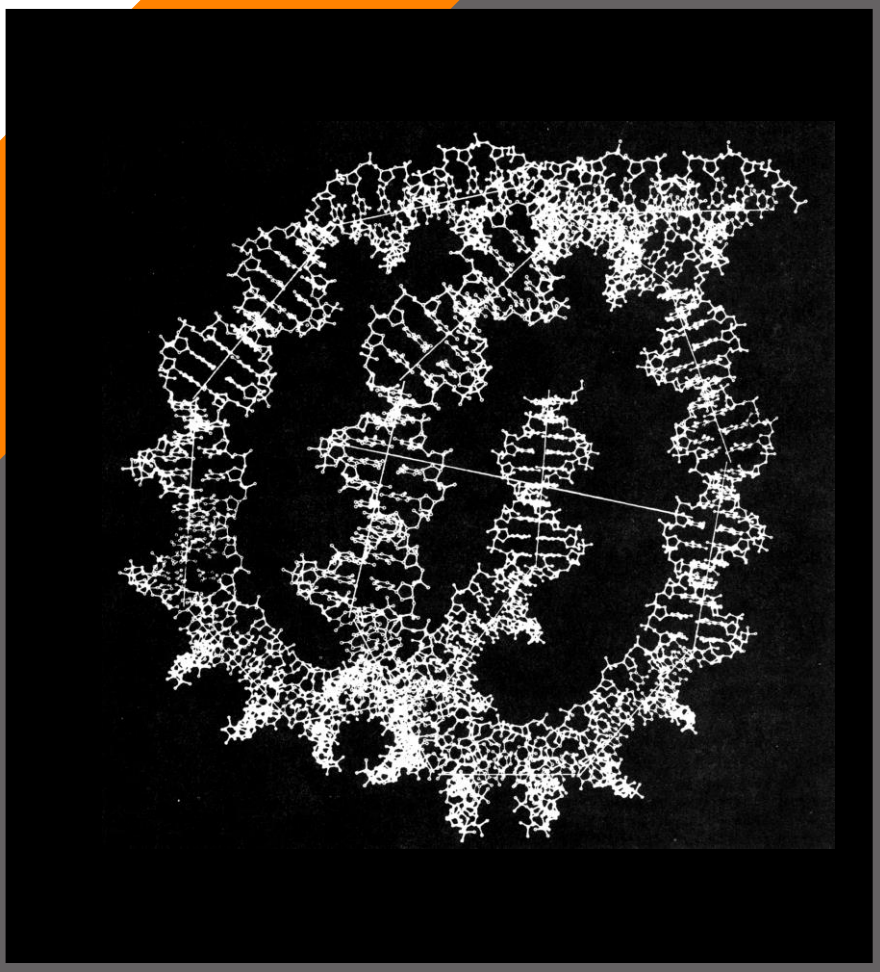


ORGANIZATION OF DNA IN CHROMATIN

Rather than bending uniformly along its length nucleosomal DNA is proposed to consist of multiple segments of B- and A- DNA held together by kinks when forming its left handed toroidal superhelical structure.

Henry M. Sobell

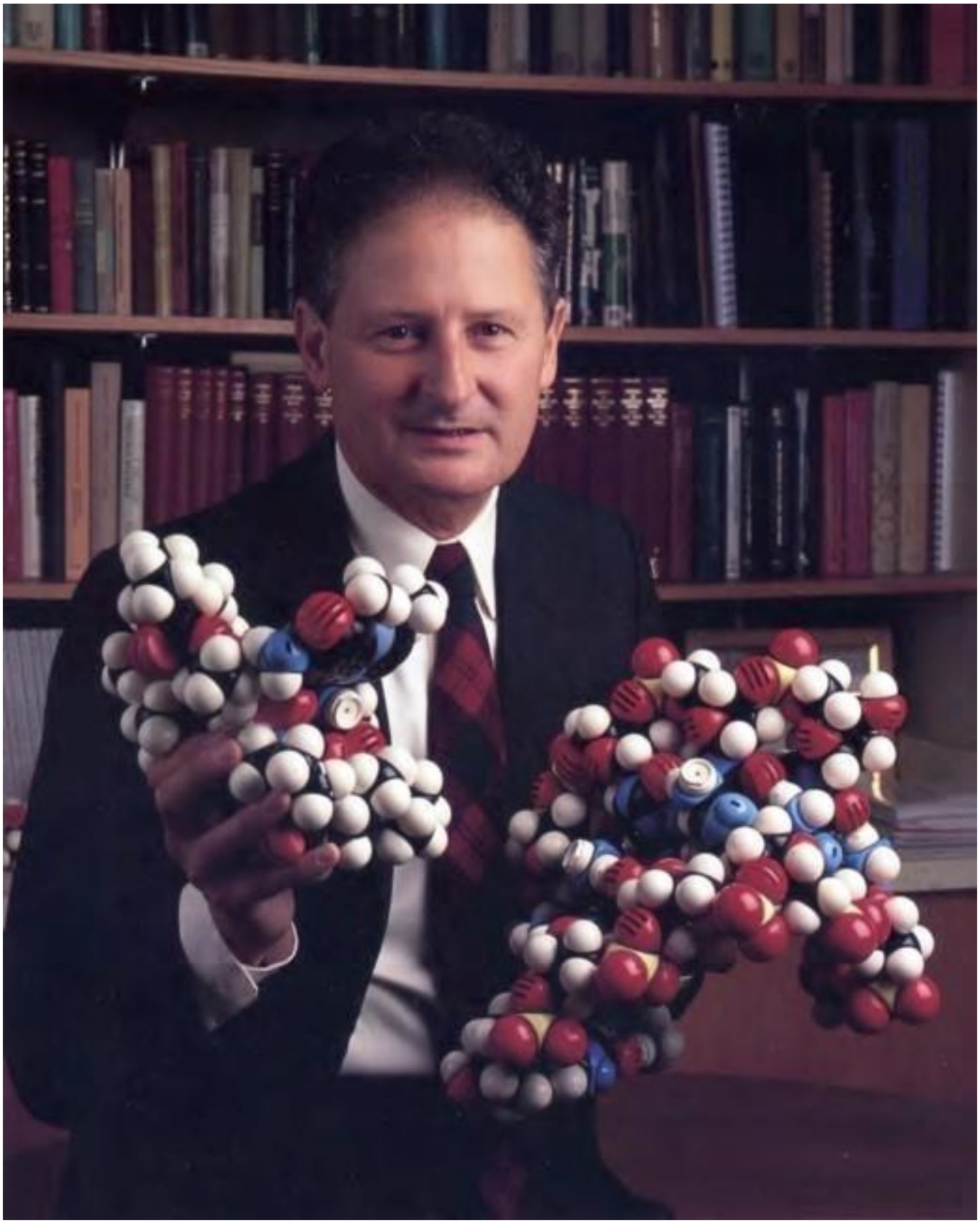
Boffin Access

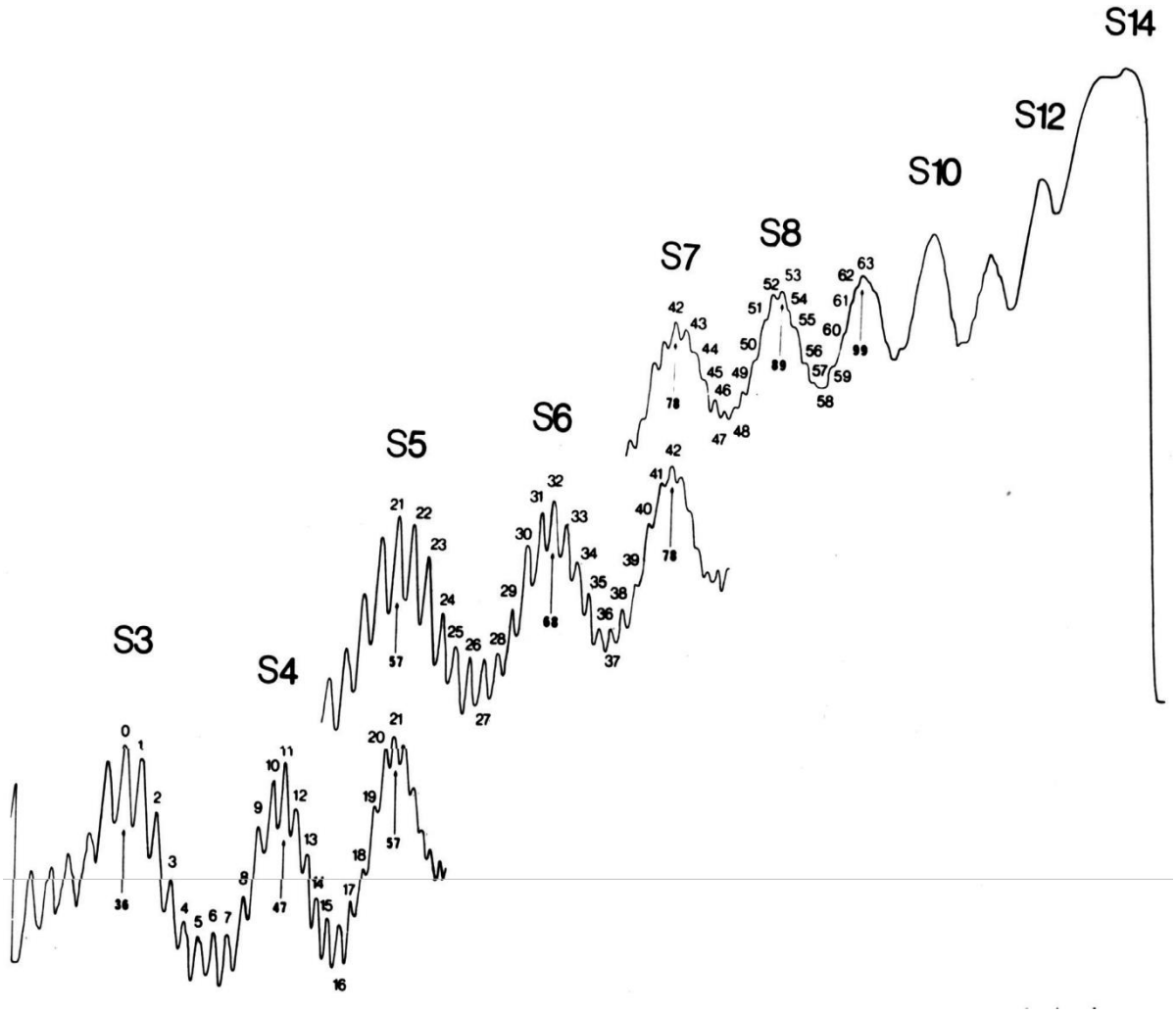
Title: ORGANIZATION OF DNA IN CHROMATIN

Author: Henry M. Sobell

Publisher: Boffin Access Limited

Published Date: JUNE-2018





This densitometer tracing shows an electrophoretic pattern of oligonucleotides arising after pancreatic DNase I digestion of a collection of DNA molecules with random sequences, each containing 147 base-pairs and labelled at their 5' ends with radioactive phosphorous -- in the presence of calcium phosphate crystals. The tracing shows a series of maxima spaced 10.5 base-pairs apart; these, in addition, containing finer peaks that differ by one nucleotide. Similar patterns have been obtained with nucleosomal DNA when it is complexed to histone octamer cores in nucleosomes.

Rather than bending uniformly along its length, the author proposes these observations to reflect the presence of a left-handed toroidal superhelical structure being composed of multiple segments, each containing 10 base pairs of B- DNA or 11 base-pairs of A- DNA, these being held together by "mixed-puckered kinks". In both cases, probability considerations predict cutting patterns to be symmetrically distributed around integral multiples of 10.5 base-pairs along DNA, the relative magnitudes of the surrounding peaks in these patterns being governed by the binomial distribution.

Preface

This paper demonstrates that -- provided there were an equal probability that either 10 base- pairs of B-DNA or 11 base-pairs of A-DNA exist within any given segment of the left handed toroidal superhelix in nucleosomal DNA, these being connected together by "mixed-puckered" kinks in its structure -- then a population of such *aperiodic* structures is expected to give rise to the *periodic* cutting-patterns observed experimentally. This would be true for naked DNA molecules immobilized on a calcium-phosphate crystalline surface as well, provided they also form left- handed toroidal superhelices under such conditions. In both cases, probability considerations predict cutting patterns to be symmetrically distributed around integral multiples of 10.5 base-pairs along DNA – the relative magnitudes of the surrounding peaks in these patterns being governed by the *binomial-distribution*. We describe the possible implications this model has in understanding the higher-order structural organization of DNA in chromatin.

Table of Contents

1. Introduction

Although it is now widely accepted that nucleosomal DNA forms a left-handed toroidal superhelix when winding around the nucleosome core particle, little is known about the flexibility in its structure that allows this to happen.

Earlier, I put forward a kinked model to understand how DNA is organized within the nucleosome [1, 2-5]. The model assumed nucleosomal DNA to be in its B-form, separated by 'mixed-puckered kinks' every 10 base-pairs. Here, I present a modification to this model, this being necessary to explain important additional experimental information uncovered several years after the model was proposed [6-10].

The modified model proposes that if there were an equal probability that both 10 base-pairs of B-DNA or 11 base-pairs of A-DNA exist within any given segment of the left handed toroidal superhelical structure -- these being connected together by mixed-puckered kinks -- then a population of such aperiodic structures is expected to give rise to the periodic cutting-patterns observed experimentally. This would be true for naked DNA molecules immobilized on a calcium- phosphate crystalline surface as well, provided they also formed left-handed toroidal superhelices under these conditions. *In both cases, probability considerations predict cutting patterns to be symmetrically distributed around integral multiples of 10.5 base-pairs along DNA, the relative magnitudes of the surrounding peaks in these patterns being governed by the binomial distribution.*

An important prediction made by this revised model concerns the number of base-pairs present within any given nucleosome. The model predicts this number to vary (i.e., lying between 140 and 154 base-pairs; however, having the highest probability that it contains 147 base-pairs), the magnitudes of the surrounding peaks being governed by the binomial distribution.

The reader should consult PREMELTONS IN DNA, A Unifying Polymer Physics Concept to Understand DNA Physical Chemistry and Molecular Biology, Explanatory Publications (2009) for further discussion of DNA structural and physical chemistry relevant to the model presented here [11].

2. Presence of Anisotropic Flexibility in DNA Structure

We begin by reviewing the properties of space-filling CPK models of DNA, which reveal the presence of a highly directional (i.e., anisotropic) flexibility in its structure. This anisotropy becomes evident upon examining these models in the absence of its aluminum helix-axis, whose presence imparts an artificial rigidity to its structure.

See Figure 1.

If one begins by bending DNA towards its wide groove direction (i.e., down the 2-fold symmetry axis lying *between* adjacent base-pairs, this being perpendicular to its helix axis), base-pairs begin to "roll" on each other's van der Waals surfaces, this being accompanied by small but systematic alterations in the torsional angles defining the sugar-puckering, base-sugar orientation and sugar-phosphate linkages in DNA. These alterations permit DNA to bend into its wide groove, eventually resulting in the formation of the mixed puckered kink [i.e., C3' *endo* (3'-5') C2' *endo*] visible from the narrow groove.

CPK MODELS ILLUSTRATING THE ANISOTROPIC FLEXIBILITY OF DNA

H.M. Sobell, C. -C. Tsai, S.C. Jain and S.G. Gilbert, J. Mol. Biol. 114, 333-365 (1977)

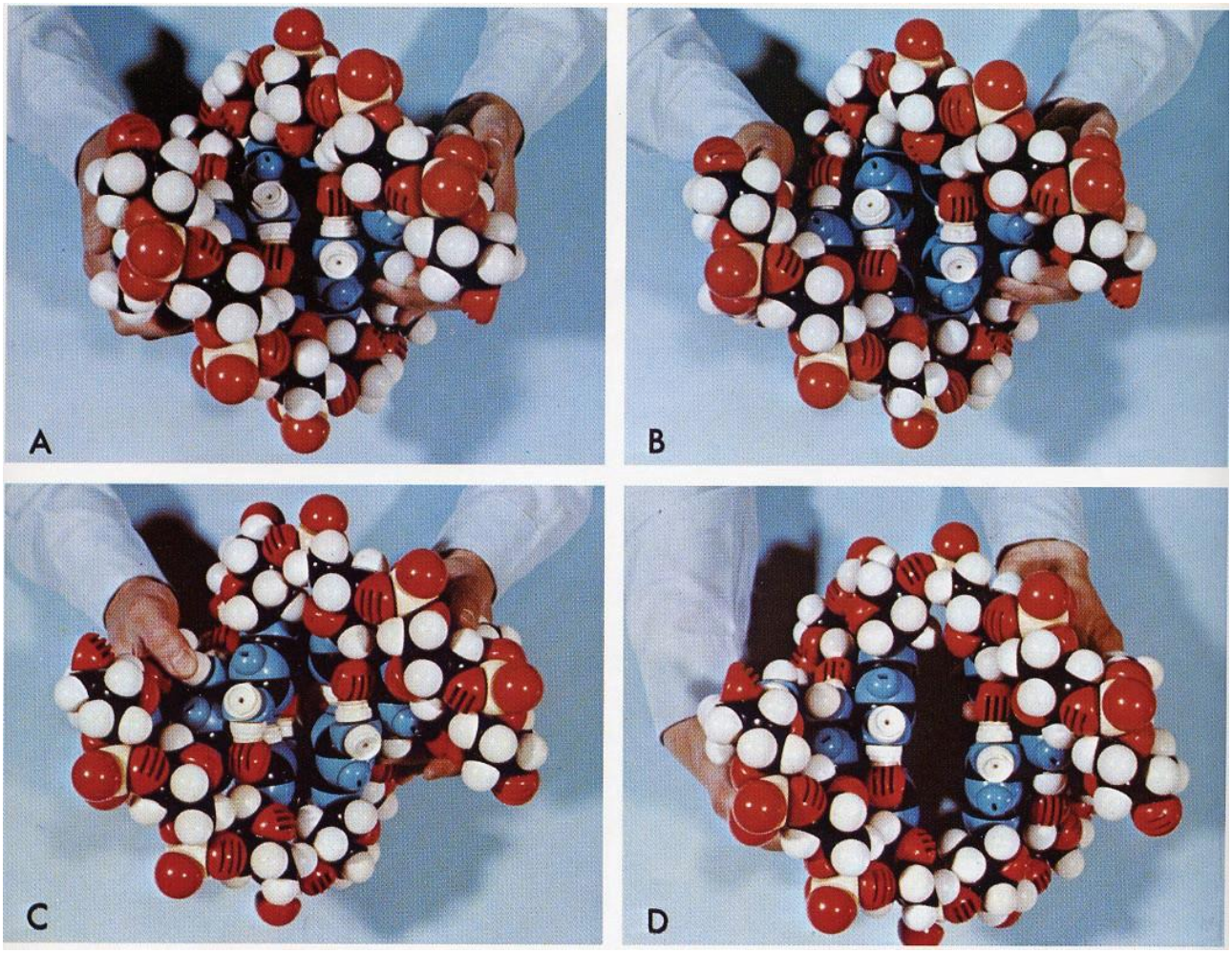


Figure 1: Space-filling Corey-Pauling-Koltun (CPK) molecular models of DNA used to demonstrate the anisotropic flexibility that is present down its 2-fold axis of symmetry that passes between adjacent base-pairs.

A. B- DNA viewed from the narrow groove direction down its 2-fold axis, passing between adjacent base-pairs.

B. B- DNA flexed into its major-groove direction, resulting in a "roll angle" of about 15 degrees between base-pairs, accompanied by a flattening of the sugar-pucker on the 5' side of the base-paired dinucleotide-structure, combined with a small decrease in its *chi* torsional angle.

C. The appearance of the 'mixed-puckered kink', characterized by the formation of a *C3' endo* (3 - 5') *C2' endo* mixed sugar-puckering pattern -- surrounded by B- DNA on either site. Adjacent base-pairs are now partially unstacked, forming a "roll-angle" of about 40 degrees and unwound -12 degrees, relative to B- DNA.

D. The flexible-hinge that results, allowing B- DNA to straighten and adjacent base-pairs to separate an additional 3.4 Angstroms. This localized conformational change is accompanied by additional unwinding, the total now being -26 degrees.

ORGANIZATION OF DNA IN CHROMATIN

H.M. Sobell, C.C- Tsai, S.G. Gilbert, S.C. Jain and T.D. Sakore, Proc. Natl. Acad. Sci. USA 73, 3068-3072 (1976)

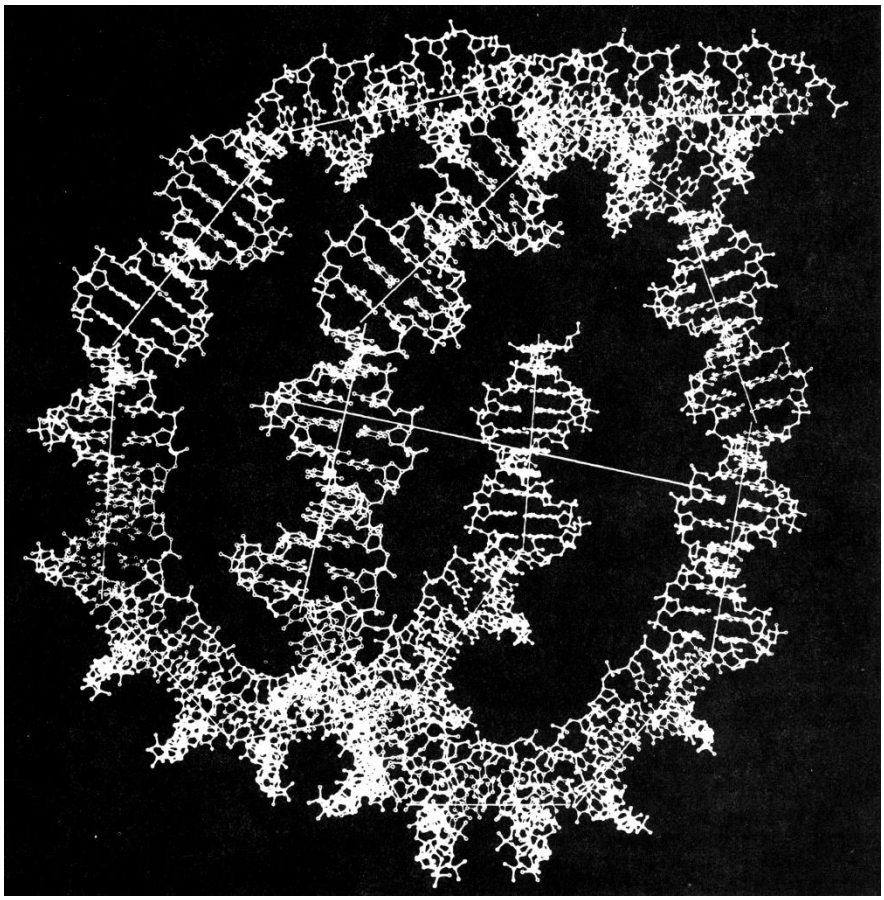


Figure 2: The original model put forward to understand the organization of DNA within the nucleosome – the 'mixed-puckered kink' appears every 10-base-pairs in B-DNA as it winds around the nucleosome histone-core to relieve the bending strain-energy that would otherwise result while forming this left-handed toroidal superhelical structure. The superhelix is generated from this ten base-pair containing asymmetric unit, by a twist of -41.1 degrees and translation along the superhelix axis of 5.26 Angstroms. The structure has a diameter of 100 Angstroms and contains about one and one-half turns per 140 base-pairs. The long central-line indicates the superhelix-axis – the length shown is 90 Angstroms. The information in this figure should be correlated with the information in the previous figure, Figure 1.

If, however there were an equal probability that either 11 base-pairs of A-DNA or 10 base-pairs of B-DNA exist within any given segment of the left-handed toroidal superhelical structure shown above, a population of such aperiodic structures can give rise to the periodic cutting-patterns observed experimentally (it is interesting to note in this regard that an A-DNA allomorph is known that contains 11 base-pairs per turn in 31.0 Angstroms (12)). This would be true for naked DNA molecules immobilized on a calcium-phosphate crystalline-surface as well, provided they also form left-handed toroidal superhelices under such conditions. In both cases, probability considerations predict cutting patterns to be symmetrically distributed around integral multiples of 10.5 base-pairs along DNA, the relative magnitudes of the surrounding peaks in these patterns being governed by the binomial-distribution. I now propose this modified model to understand how DNA is organized within the nucleosome.

A

$(a + b)^1 = a + b$ [1 1]

$(a + b)^2 = a^2 + 2ab + b^2$ [1 2 1]

$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ [1 3 3 1]

$(a + b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$ [1 4 6 4 1]

$(a + b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$ [1 5 10 10 5 1]

$(a + b)^6 = a^6 + 6a^5b + 15a^4b^2 + 20a^3b^3 + 15a^2b^4 + 6ab^5 + b^6$ [1 6 15 20 15 6 1]

B

1 1

1 2 1

1 3 3 1

1 4 6 4 1

1 5 10 10 5 1

1 6 15 20 15 6 1

---

1 7 21 35 35 21 7 1

1 8 28 56 70 56 28 8 1

1 9 36 84 126 126 84 36 9 1

1 10 45 120 210 252 210 120 45 10 1

1 11 55 165 330 462 462 330 165 55 11 1

1 12 66 220 495 792 924 792 495 220 66 12 1

1 13 78 286 715 1287 1716 1716 1287 715 286 78 13 1

1 14 91 364 1001 2002 3003 3432 3003 2002 1001 364 91 14 1

2^n = 2 4 8 16 32 64 128 256 512 1,024 2,048 4,096 8,192 16,384

n = 1 2 3 4 5 6 7 8 9 10 11 12 13 14

Table 1: A. Algebraic explanation of the binomial theorem, and the origin of the binomial coefficients B. Pascal's triangle, indicating how these binomial coefficients can be extended indefinitely, without having to carry out further algebra.

The presence of this 'mixed puckered" kink connecting segments containing 10 base-pairs of B-DNA or 11 base-pairs of A-DNA causes nucleosomal DNA to bend and unwind, giving rise to a left-handed toroidal superhelical structure very similar to that originally proposed (i.e., an A-DNA allomorph is known that contains 11 base-pairs per turn in 31.0 Angstroms (12)).

See Figure 2.

I will next explain how this modified model explains the 10.5 base-pair periodicity observed in electrophoretic patterns of nucleosomal DNA after digestion with the pancreatic DNase I. We first begin with a brief review of the binomial theorem.

3. A Brief Review of the Binomial Theorem

The binomial expression (a + b) raised to any power n (where n = 1, 2, 3 ...) leads to a series of polynomials, each having numerical coefficients defined as the binomial coefficients (see Table 1A).

These coefficients can be arranged in the form of a triangle (known as Pascal's triangle), which permit their values to be readily extrapolated to any value of n, without having to carry out further algebra (see Table 1B).

Thus, for example:

(n = 1) [1 1] leads to (n = 2) [1 2 1], since 1 + 1 = 2

(n = 2) [1 2 1] leads to (n = 3) [1 3 3 1], since 1 + 2 = 3 2 + 1 = 3

(n = 3) [1 3 3 1] leads to (n = 4) [1 4 6 4 1], since 1 + 3 = 4 3 + 3 = 6 3 + 1= 4

(n = 4) [1 4 6 4 1] leads to (n = 5) [1 5 10 10 5 1], since 1 + 4 = 5 4 + 6 = 10 6 + 4 = 10 4 + 1 = 5

and so on.

Once Pascal's triangle has been calculated, the binomial coefficients can be normalized as follows:

Thus, for example:

(n = 1) [1 1] $2^1 = 2$ (1/2 1/2), or [0.5000 0.5000]

(n = 2) [1 2 1] $2^2 = 4$ (1/4 2/4 1/4), or [0.2500 0.5000 0.2500]

(n = 3) [1 3 3 1] $2^3 = 8$ (1/8 3/8 3/8 1/8), or [0.1250 0.3750 0.3750 0,1250]

(n = 4) [1 4 6 4 1] $2^4 = 16$ (1/16 4/16 6/16 4/16 1/16), or [0.06250.2500 0.3750 0.2500 0.0625]

and so on.

These normalized values of the binomial coefficients have been tabulated in Table 2, their sums being 1.0000 for each value of n. For reasons we next discuss, they will be used in the analysis which follows.

0.5000 0.5000

0.2500 0.5000 0.2500

0.1250 0.3750 0.3750 0.1250

0.0625 0.2500 0.3750 0.2500 0.0625

0.0312 0.1562 0.3125 0.3125 0.1562 0.0312

0.0156 0.0936 0.2343 0.3125 0.2343 0.0936 0.0156

0.0078 0.0546 0.1640 0.2734 0.2734 0.1640 0.0546 0.0078

0.0039 0.0312 0.1093 0.2187 0.2734 0.2187 0.1093 0.0312 0.0039

0.0019 0.0175 0.0703 0.1640 0.2460 0.2460 0.1640 0.0703 0.0175 0.0019

0.0009 0.009 0.0439 0.1171 0.2050 0.2460 0.2050 0.1171 0.0439 0.0097 0.0009

0.0004 0.0053 0.0268 0.0805 0.1611 0.2255 0.2255 0.1611 0.0805 0.0268 0.0053 0.0004

0.0002 0.0029 0.0161 0.0537 0.1208 0.1933 0.2255 0.1933 0.1208 0.0537 0.0161 0.0029 0.0002

0.0001 0.0015 0.0095 0.0349 0.0872 0.1571 0.2094 0.2094 0.1571 0.0872 0.0349 0.0095 0.0015 0.0001

0.0000 0.0008 0.0055 0.0222 0.0610 0.1221 0.1832 0.2094 0.1832 0.1221 0.0610 0.0222 0.0055 0.0008 0.0000

Table 2: Normalized coefficients obtained from Table 1B, as described in the text. Note that the sum of the decimals in each horizontal row is 1.0000.

4. Application to Nucleosomal DNA

Figure 3 summarizes the size and composition of DNA segments predicted after limited digestion of the (kinked) left-handed toroidal superhelical form in nucleosomal DNA by the pancreatic DNase I enzyme. We will begin by assuming this molecule to have 147 base-pairs, and to be end-labeled at one of two 5' phosphate termini with P_{32} (indicated by the asterisks).

If the pancreatic DNase I begins by cleaving a kink after the first base-paired segment of B-DNA (i.e., $*B_{10}$), or after the first base-paired segment of A-DNA (i.e., $*A_{11}$), oligonucleotides having chain lengths of 10 or 11 will appear with relative frequencies [1 1]. If on the other hand, the pancreatic DNase 1 begins by cleaving a kink after the second base-paired segment containing B-DNA (i.e., $*B_{10}$ B_{10}), or after the second base-paired segment containing both B- and A-DNA (i.e., $*B_{10}$ A_{11} or $*A_{11}$ B_{10}), or after the second base-paired segment containing A-DNA (i.e., $*A_{11}$ A_{11}), oligonucleotides having chain lengths of 20, 21 and 22 will appear with relative frequencies [1 2 1].

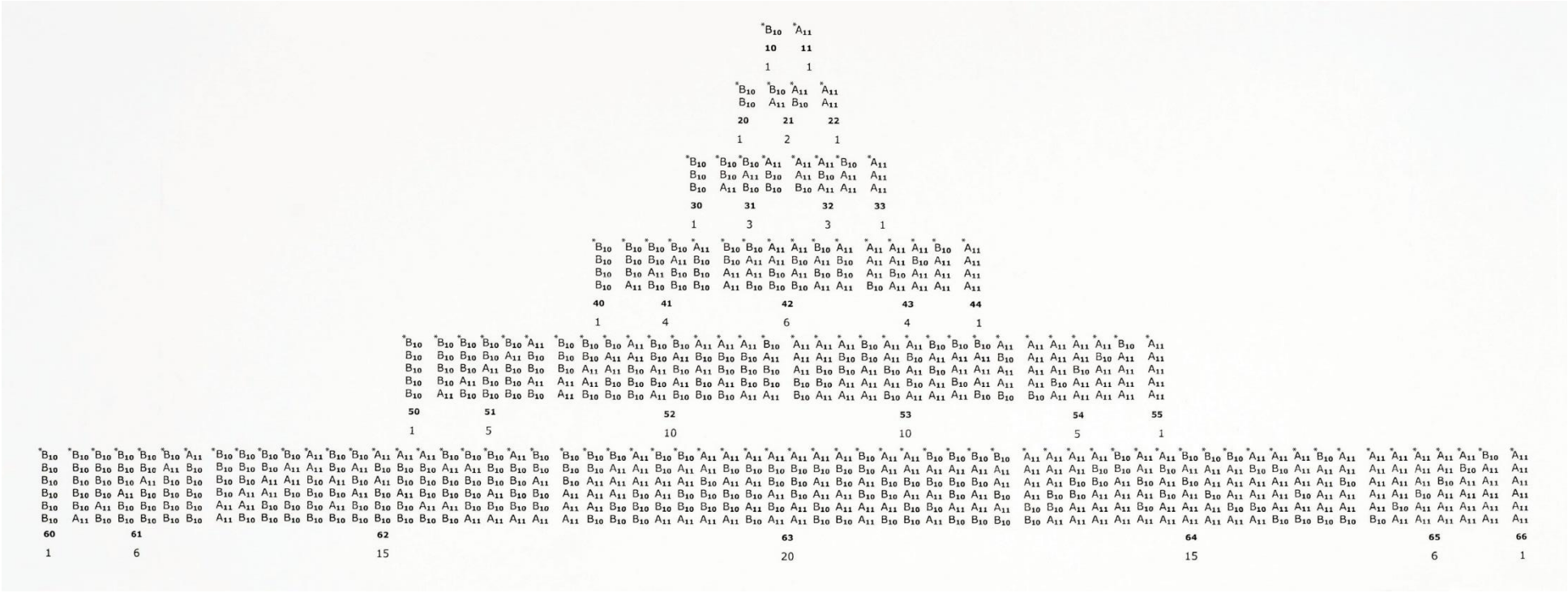


Figure 3: See text for discussion

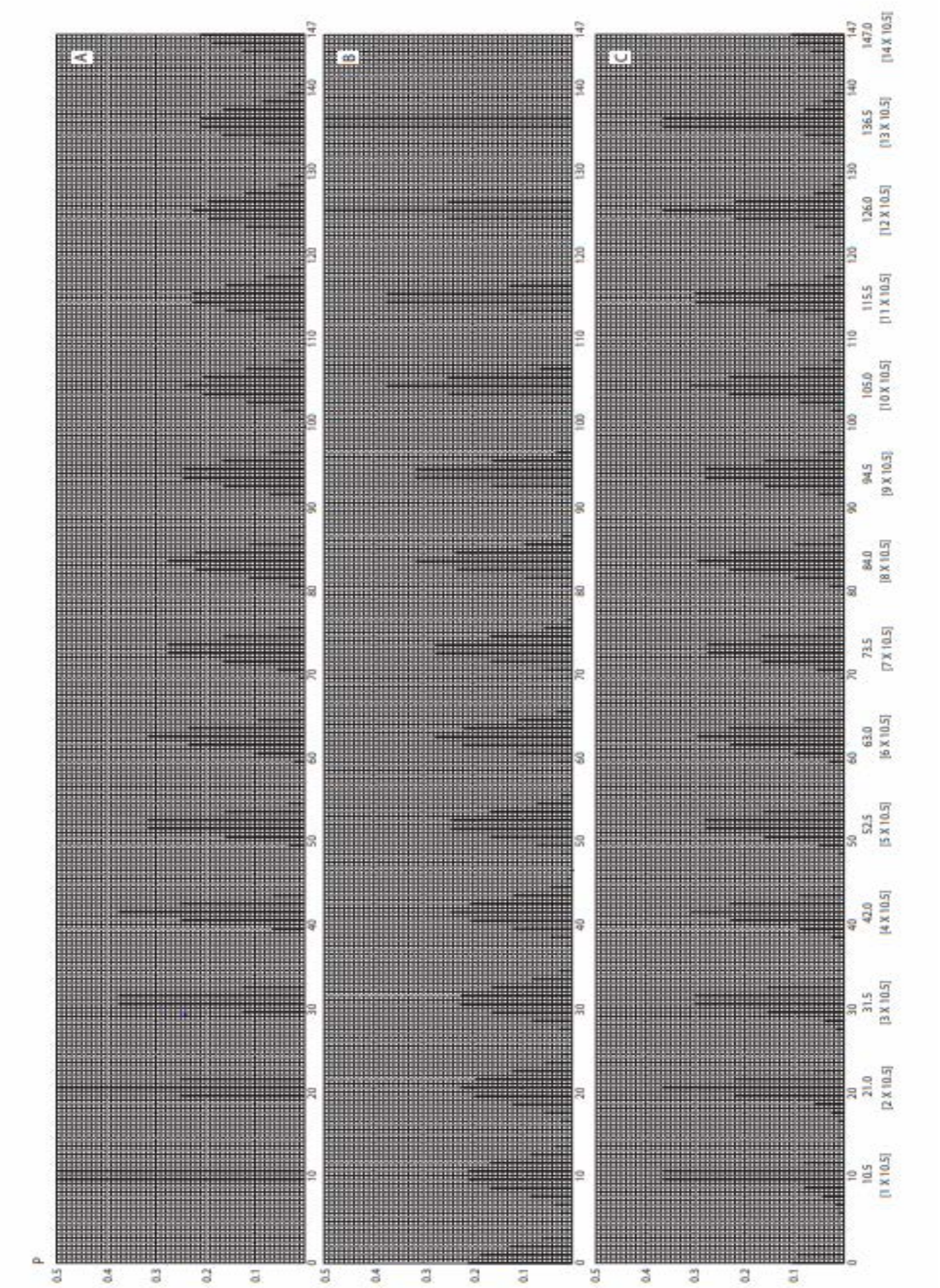


Figure 4: See text for discussion

HELICAL PERIODICITY OF DNA ON AND OFF THE NUCLEOSOME AS PROBED BY NUCLEASES

A. Klug, L.C. Lutter and D. Rhodes, Cold Spring Harbor Symposium of Quantitative Biology 47, 285-292 (1983)

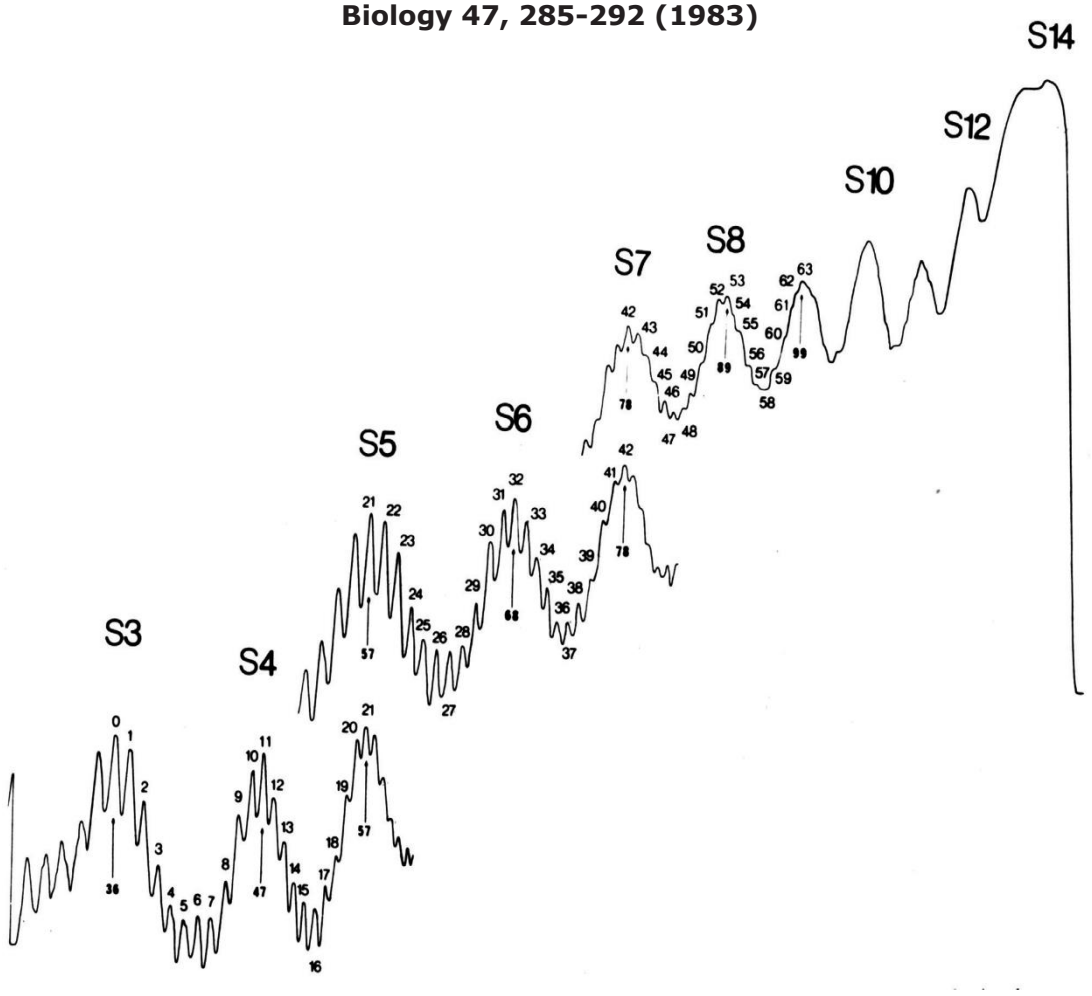


Figure 5: This densitometer tracing shows an electrophoretic pattern of oligonucleotides arising after pancreatic DNase I digestion of a collection of DNA molecules with random sequences, each containing 147 base-pairs and labelled at their 5' ends with radioactive phosphorous -- in the presence of calcium phosphate crystals. The tracing shows a series of maxima spaced 10.5 base-pairs apart; these, in addition, containing finer peaks that differ by one nucleotide. Similar patterns have been obtained with nucleosomal DNA when it is complexed to histone-octamer cores in nucleosomes.

Rather than bending uniformly along its length, the author proposes these observations to reflect the presence of a left-handed superhelical structure being composed of multiple segments, each containing 10 base-pairs of B-DNA or 11 base-pairs of A-DNA, these being held together by 'mixed-puckered kinks'. In both cases, probability considerations predict cutting patterns to be symmetrically distributed around integral multiples of 10.5 base-pairs along DNA, the relative magnitude of the surrounding peaks in these patterns being governed by the binomial distribution.

The fit is not perfect, however, since these experimental results reveal the existence of additional fine peaks between maxima differing by one nucleotide. These can be explained as arising from a limited number of left-handed toroidal superhelical structures that begin with shorter (or longer) B- and A-DNA end-segments. This would have the effect of shifting the patterns (shown in Figure 4A and B) to the left and to the right, causing each maximum in Figure 4C to broaden. This end-effect can explain the presence of additional finer peaks between maxima which differ by one nucleotide.

Finally, if the pancreatic DNase begins by cleaving a kink appearing after the third base-paired segment of B- DNA (i.e., *B_{10} B_{10} B_{10}), or after the third base-paired segment containing both B- and A-DNA (i.e., *B_{10} B_{10} A_{11}, *B_{10} A_{11} B_{10}, *A_{11} B_{10} B_{10}), or after the third base-paired segment containing a different combination and permutation of both B- and A-DNA (i.e., *A_{11} A_{11} B_{10}, *A_{11} B_{10} A_{11}, *B_{10} A_{11} A_{11}), or after the third base-paired segment containing A-DNA (i.e., *A_{11} A_{11} A_{11}), oligonucleotides having chain lengths of 30, 31, 32 and 33 will appear with relative frequencies [1 3 3 1]. Likewise, it can be readily verified that oligonucleotides having chain lengths of 40 through 44 will appear with relative frequencies [1 4 6 4 1]; chain lengths of 50 through 55 will appear with relative frequencies [1 5 10 10 5 1]; and chain lengths of 60 through 66 will appear with relative frequencies [1 6 15 20 15 6 1].

And so on.

From this it is clear that the size and composition of the nucleosomal DNA segments generated by pancreatic DNase 1 cleavage in this example obey the binomial distribution. We shall now show that the normalized values of the binomial coefficients define the probability that kinks appear at specific locations determined by n, these kinks being recognized and subsequently cleaved by the pancreatic DNase I enzyme.

Figure 4 summarizes these predictions, *both* ends of the 147 base-paired fragments having been end-labeled with P_{32} in this example. The ordinate shows the probability (i.e., the normalized values of the binomial coefficients) that kinks appear at specific locations (determined by n) and then cleaved by the nuclease, while the abscissa shows its chain length position. Since the pancreatic DNase is expected to cause double stranded breaks each time it recognizes a kink, two different P_{32} labeled fragments will be generated from a single nucleolytic cutting event. These can be detected by gel electrophoresis as two separate single-stranded fragments, one arising from the 'Watson side', the other arising from the 'Crick side'. For simplicity, we have separated these to show them both individually (Figures 4A and B), and in combination (Figure 4C), the latter graph having been rescaled for comparison.

Figure 5 shows a densitometer tracing of an electrophoretic pattern of oligonucleotides arising after pancreatic DNase I digestion of a collection of DNA molecules with random sequences, each containing 147 base-pairs and labeled at their 5' ends with radioactive phosphorous in the presence of calcium phosphate crystals [10]. The tracing shows a series of maxima spaced 10.5 base-pairs apart -- these, in addition, containing finer peaks that differ by one nucleotide. Similar patterns have been reported with nucleosomal DNA when it is complexed to histone-octamer cores of nucleosome [6-9]. The data in Figure 5 can readily be compared with that predicted in Figure 4C. In both cases, cutting patterns are symmetrically distributed around integral multiples of 10.5 base-pairs along DNA, the relative magnitudes of the surrounding peaks in these patterns being explained by the binomial distribution. The fit is not perfect, however, since the experimental results reveal the existence of additional finer peaks between maxima differing by one nucleotide. These can be explained as arising from a small number of left-handed toroidal superhelical structures having shorter (or longer) B- or A- DNA end-segments. This would have the effect of shifting the patterns (in Figures 4A and B) to the left and to the right, causing each maximum in Figure 4C to broaden. This end-effect can explain the presence of the additional finer peaks between maxima differing by one nucleotide just mentioned.

5. A Model to Understand the Higher-Order Solenoidal Folding of DNA in Chromatin

The presence of mixed-puckered kinks connecting B_{10} with A_{11} DNA segments within the nucleosome suggests the likelihood that this same basic structural motif continues to be used when forming the 100 Angstrom fiber. Based on their detailed electron microscopic studies of chromatin, Finch and Klug (1976) and Worcel (1978) have proposed the 100 Angstrom fiber to be transformed (i.e., in the presence of the H-1 histone) into a more compact left-handed solenoidal superhelix, this having a pitch of 110 Angstroms and a diameter of about 300 Angstroms (13, 14). We have explored the possibility that such a solenoid arises due to the flexibility present in the mixed-puckered kink to fold internucleosomal DNA into this form.

Figure 6 shows the results of these calculations. The model shown is a 5.8-fold solenoid, obtained by varying both theta and alpha in the central three kinks within the 60 base-pair connecting region (akin to that shown in Figure 1, panels C and D). The connecting region remains on the outside of the solenoid, with two-fold symmetry being maintained.

Structures in this class can readily form the right pitch (110 Angstroms) and the right diameter (300 Angstroms) to fit the electron microscopic data.

An important prediction which this model makes concerns the ability of intercalative drugs and dyes to bind tightly within these connecting regions.

Lawrence and Daune (1976) have described the existence of a limited number of tight binding sites for ethidium in native chromatin, these having binding constants two orders of magnitude greater than for naked DNA. These sites disappear when H1 histones are removed, suggesting a correlation with the solenoidal structure described here [15].

In addition, Shen and Hearst (1978) describe psoralin cross-linking with high probability every 200 base-pairs in D. melanogaster nuclear-DNA [16]. Our model would predict these sites to occur with highest frequency at integral multiples of 10.5 base pairs (again, these surrounding locations obeying the binomial distribution), and to be highly clustered within the 60 base-paired connecting regions. Finally, Cartwright and Elgin (1982, 1983) have shown that both the intercalator, 1, 10-phenanthroline-copper (I), and the micrococcal nuclease cut DNA every 200 base-pairs to liberate the basic subunit structure of chromatin, the nucleosome [17, 18].

These data strongly suggest higher-energy mixed-puckered kinks to be present within the internucleosomal connecting regions, enabling nucleosomes to form its higher-order solenoidal structure.

6. Experimental Predictions

The model makes a number of testable predictions.

1) The model predicts the number of base-pairs within any given nucleosome to be variable (i.e., lying between 140 and 154 base-pairs; however, having the highest probability that it contains 147 base-pairs), the magnitudes of the surrounding peaks being governed by the binomial distribution.

A MODEL TO UNDERSTAND THE HIGHER-ORDER SOLENOIDAL FOLDING OF DNA IN CHROMATIN

H.M. Sobell, B. S. Reddy, K.K. Bhandary, S.C. Jain, T.D. Sakore and T.P. Seshadri, Cold Spring Harb. Symp. Quant. Biol. 42, 87-102 (1978)

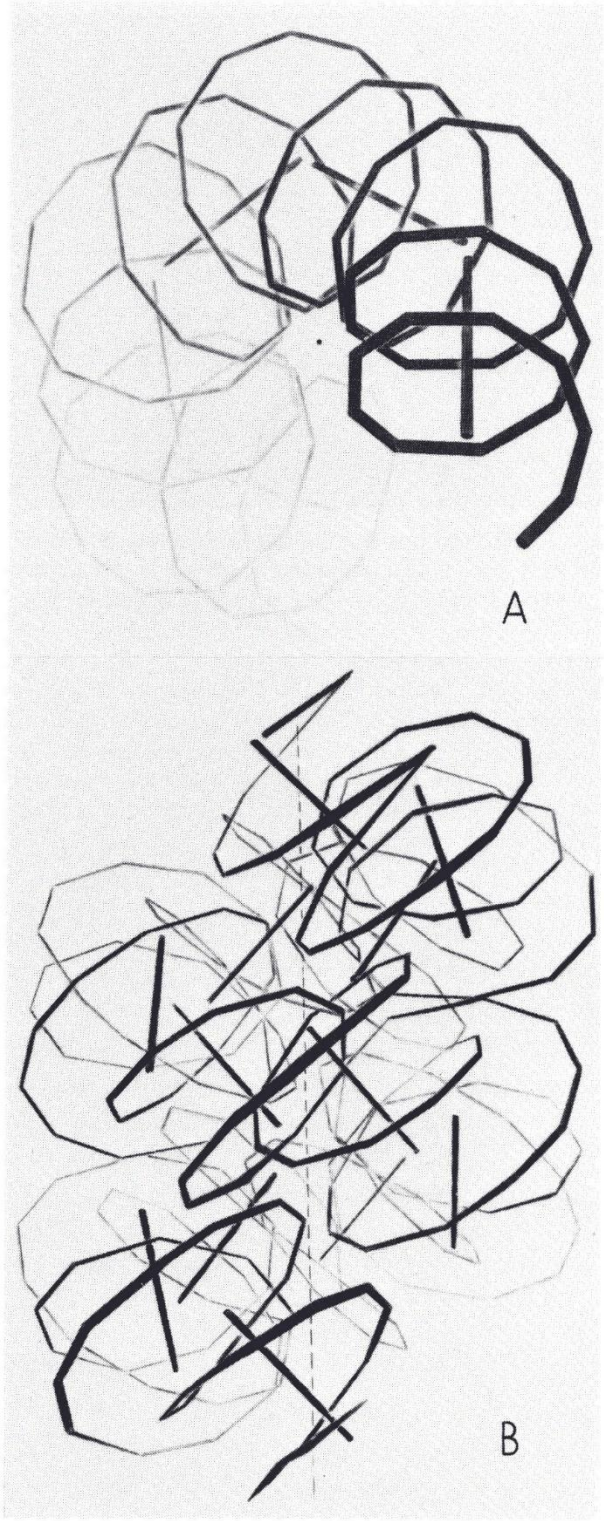


Figure 6: A model to understand the higher-order solenoidal folding of DNA in chromatin. See text for discussion

2) The model predicts irehdiamine and dipyrandium (but not ethidium) -- two steroidal diamines that bind by partial intercalation to the lower-energy form of the kink in nucleosomal DNA (11) -- to be competitive inhibitors of the pancreatic DNase I enzyme. Their presence is predicted to suppress the appearance of the cutting patterns observed in Figure 5.

3) The model further predicts ethidium (but not irehdiamine or dipyrandium) to competitively inhibit the micrococcal-nuclease and the chemical-nuclease, 1, 10 phenanthroline-coppers (I), from cleaving hypersensitive-sites that exist between nucleosomes in whole chromatin. These regions are proposed to contain higher-energy mixed-puckered kinks in their internucleosomal regions when forming the higher-order solenoidal structure.

7. Appendix

This section describes a soliton mechanism for thermal DNA melting, having been inadvertently omitted from my earlier book, PREMELTONS IN DNA (2009) (11). Premeltons are examples of "kink-antikink bound-states" (i.e., or equivalently, "breather solitons") in DNA, these arising spontaneously within the early melting regions of DNA to nucleate site-specific DNA-melting. Their presence allows one to understand how drugs and dyes intercalate into DNA, and also to understand how nucleases such as the micrococcal-nuclease and the pancreatic DNase 1 enzymes recognize and cleave early melting regions in naked DNA molecules. The reader is referred to this book for further detail.

Energies necessary for the formation of the premelton in DNA molecules come from Brownian motion, excited by solvent collisions at normal (i.e., kT) energies. We envision DNA in solution to be continually bombarded by solvent collisions along its length. Although at first glance, one might expect the average excitation force to be zero, this is not the case in the microscopic domain, where DNA is continually experiencing unbalanced forces (i.e., the Brownian force). Since the collision cross-sectional area of DNA is small (i.e., the diameter is about 20 Angstroms), relatively few solvent molecules impinge on its surface in short time intervals. Moreover, the flexibility of DNA is highly anisotropic. Due to this anisotropy, most solvent collisions are expected to have little effect on DNA structure, exciting only small amplitude normal mode motions in functional groups. These are expected to damp readily through solvent interactions. There are, however, small windows of collisions that deform DNA nonlinearly. We believe these collisions to hit DNA from both wide and narrow groove directions, striking DNA along dyad axes located between adjacent base pairs. Such collisions give rise to nonlinear pulses in DNA (also called solitons, or solitary excitations), which move along the polymer chain with a velocity significantly less than the speed of sound. These contain a modulated beta-alternation in sugar puckering along both polynucleotide chains, and are nontopological -- that is, although these excitations unwind DNA, this is counterbalanced by right-handed superhelical writhing to keep the linking invariant. Energies stored and transmitted by such intrinsic locally coherent excitations can travel considerable distances along DNA with minimal dissipative loss, since they are largely internal to the polymer structure. In addition, such nonlinear pulses remain "robust", since the nonlinearity present in the sugar-pucker conformations acts to minimize dispersion-effects. The shape of the energy density profile accompanying low energy solitary excitations is expected to be sensitive to the nucleotide base sequence in DNA. This is because different DNA regions contain different base-stacking energies, and have, therefore, different intrinsic flexibilities. Energy-density profiles of traveling solitary excitations are expected to sharpen

up (i.e., the leading edge of the excitation traveling more slowly than the trailing edge) within regions that start out by being more flexible than other regions. This acts to deform DNA structure, and to enhance the lifetime of these excitations. This increases the probability that, with increasing temperature, still larger excitations can form from the coalescence of additional solitary excitations that arrive in these regions. The appearance of these larger excitations deforms DNA regions still further and gives rise to even greater flexibility in their most central regions. Eventually, with increasing temperature, premeltons arise that contain hyperflexible (liquid-like) beta-DNA cores, surrounded by phase-boundaries termed "kink" and "antikink". Premeltons have longer lifetimes, and, with increasing temperature, become nucleation-centers for DNA-melting. Traveling solitary excitations, originating from solvent collisions at earlier times and from more remote locations, enter the premelton, their energies being trapped within the core. This energy continues to enlarge beta-DNA core-regions and increases the amplitude of breather-motions, leading to DNA "breathing". Finally, with increasing temperature, permanently melted single-stranded DNA regions appear within the premelton. Their appearance signals the onset of DNA melting. We have called these larger structural solitons, "meltons". With increasing temperature; meltons continue to trap energy from entering solitary excitations that have arisen elsewhere (and at earlier times) along B-DNA. Their energies are used to separate kink from antikink while lengthening internally melted single-stranded DNA-regions. Finally, with increasing temperature, meltons coalesce and all DNA becomes single-stranded. Single-stranded DNA is extremely flexible (i.e., entropic), and can be likened to a gas-like phase. In summary then, my thermal mechanism predicts the existence of three structural phases for DNA: the Watson-Crick A- or B- forms (solid), the hyperflexible beta-DNA form (liquid), and the highly entropic single-stranded melted-DNA form (gas). The beta-DNA phase within the premelton is predicted to play a key role in nucleating DNA-melting.

8. References

1) Crick FHC, A Klug. Kinky Helix. *Nature*. 1975; **255**: 530-533.

2) Sobell HM, CC Tsai, SG Gilbert, SC Jain, TD Sakore. Organization of DNA in Chromatin. *Proc. Natl. Acad. Si. USA.* 1976; **73**: 3068-3072.

3) Sobell HM, CC Tsai, SC Jain, SG Gilbert. Visualization of drug-nucleic acid interactions at atomic resolution: II. Structure of an Ethidium/Dinucleoside Monophosphate Crystalline Complex, ethidium:5-iodocytidylyl (3′–5′) guanosine. *J Mol Biol. 1977; **114**: 317-331.*

4) Sobell HM, BS Reddy, KK Bhandary, SC Jain, TC Sakore, TP Seshadri. Conformational Flexibility in DNA Structure as Revealed by Structural Studies of Drug Intercalation and its Broader Implications in Understanding the Organization of DNA in Chromatin. *Cold Spring Harbor Symposia on Quantitative Biology*. 1978; **42**: 87-102.

5) Sobell HM, TD Sakore, SC Jain, A Banerjee, KK Bhandary, BS Reddy, ED Lozansky. Beta-Kinked DNA – a Structure That Gives Rise to Drug Intercalation and DNA Breathing – and its Wider Significance in Determining the Premelting and Melting Behavior of DNA. *Cold Spring Harbor Symposia on Quantitative Biology*. 1983; **47**: 293-314.

6) Lutter LC. Precise location of DNase I cutting sites in the nucleosome core determined by high resolution gel *electrophoresis. Nucleic Acids Res.* 1979; **6**: 41-56.

7) Prunell A, RD Kornberg, L Lutter, A Klug, M Levitt, FHC Crick. Periodicity of Deoxyribo-nuclease Digestion of Chromatin. *Science*. 1979; **204**: 855-858.

8) Lutter LC. DNase II digestion of the nucleosome core: precise locations and relative exposures of sites. *Nucleic Acids Res.* 1981; **9**: 4251-4266.

9) Klug A, L Lutter, The helical periodicity of DNA on the nucleosome. *Nucleic Acids Res.* 1981*;* **9**: 4267-4283.

10) Klug, A, LC Lutter and R Rhodes. Helical Periodicity of DNA On and Off the Nucleosome as Probed by Nucleases. *Cold Spring Harbor Symposia on Quantitative Biology.* 1983; **47**: 285-292.

11) Sobell HM. PREMELTONS IN DNA, Library of Congress Control Number: 2009942069, ISBN: 978-0-615-33828-6, Explanatory Publications, Lake Luzerne, New York (2009).

12) Arnott S, R Chandrasekaran, IH Hall, LC. Puigjaner, JK Walker et al. DNA Secondary Structures: Helices, Wrinkles, and Junctions. *Cold Spring Harbor Symp Quant Biol* 1983; **47**: 53-65.

13) Finch JT, A Klug. Solenoidal Model for Superstructure in Chromatin. *Proc Natl Acad Sci.* 1976; **73**: 1897-1901.

14) Worcel A. Molecular Architecture of the Chromatin Fiber, *Cold Spring Harbor Symp Quant Biol*. 1978: **42:** 313-324.

15) Lawrence JJ, M Daune. Ethidium bromide as a probe of conformational heterogeneity of DNA in chromatin. The role of histone H1. *Biochemistry* 1976; ***15***: 3301-3307.

16) Shen CKJ, JE Hearst. Chromatin Structures of Main-band and Satellite DNAs in Drosophila Melanogaster Nuclei as Probed by Photochemical Cross-linking of DNA with Trioxsalen. *Cold Spring Harbor Symp Quant Biol.* 1978; **42**: 179-189.

17) Cartwright IL, SCR Elgin. Analysis of chromatin structure and DNA sequence organization: use of the 1,10-phenanthroline-cuprous complex. *Nucleic Acids Res*.1982; **10**: 5835-5852.

18) Elgin SCR, IL Cartwright, G, Fleischmann, K Lowenhaupt, MA Keene. Cleavage Reagents as Probes of DNA Sequence Organization and Chromatin Structure: Drosophila Melanogaster Locus 67B1. *Cold Spring Harbor Symp Quant Biol*. 1983; **47**: 529-538.

As the DNA in our genome comprises about three billion base-pairs, with each base-pair separated by a third of a nanometer -- its total length is about a meter -- all of which residing within a compact form known as *chromatin*. At first guess, one might think the DNA to be wound up like a ball of yarn, but chromatin turns out to be a more complex structure, DNA being organized into a hierarchical series of superhelices.

Counting the right-handed double-helix as the first stage in the hierarchical ordering, the second consists of 147 base- pairs wound around the outside of nucleosomes as a left-handed toroidal-superhelix containing one and three-quarter turns. Each nucleosome contains two pairs each of four different histones, small positively charged basic proteins called H2A, H2B, H3 and H4 spatially related by two-fold symmetry. Adjacent nucleosomes remain connected together by linker DNA, additional DNA (variable in length, but generally between 50 to 60 base- pairs) that exists between nucleosomes, resulting in the formation of an extended 100 Angstrom fiber.

In the presence of an additional histone (H1), DNA is known to undergo a still higher level of compaction, organizing itself into a solenoidal superhelical structure having a diameter of about 300 Angstroms. This 300 Angstrom fiber can readily be seen by electron microscopy and, almost certainly, the unraveling of its structure foreshadows still further complex structural features of chromatin to be discovered in future years.

In order for DNA to be organized into this hierarchical series of superhelices, there must be a source of flexibility in DNA structure that allows this to happen.

Earlier, the author put forward a kinked model to understand how DNA is organized within the nucleosome. The model assumed nucleosomal DNA to be in its B- form, separated by 'mixed-puckered kinks' every 10 base-pairs. In this book, he presents a modification to this model; this being necessary to explain important additional experimental information uncovered several years after the model was proposed. The modified model proposes that if there were an equal probability of both 10 base-pairs of B- DNA or 11 base-pairs of A- DNA existing within any given segment of the left-handed toroidal superhelical structure – these being connected together by 'mixed-puckered kinks' – then a population of such aperiodic structures can be expected to give rise to the periodic cutting-patterns observed experimentally. This would be true for naked DNA molecules immobilized on a calcium-phosphate crystalline surface as well, provided they also formed left-handed toroidal superhelices under these conditions.

In both cases, probability considerations predict cutting patterns to be symmetrically distributed around integral multiples of 10.5 base-pairs along DNA, the relative magnitudes of the surrounding peaks in these patterns being governed by the binomial distribution – the proof of which is presented in this book.

Henry M. Sobell was born in Los Angeles, California November 7, 1935, and grew up in Brooklyn, New York, where he attended Brooklyn Technical High School (1948-1952), Columbia College (1952-1956) and the University of Virginia School of Medicine (1956-1960). Instead of practicing clinical medicine, Dr. Sobell then went to the Massachusetts Institute of Technology, Cambridge, Massachusetts, to join Professor Alexander Rich in the Department of Biology (1960-1965) where, as a Helen Hay Whitney Postdoctoral Fellow, he learned the technique of single-crystal X-ray analysis. Dr. Sobell then joined the Chemistry Department at the University of Rochester, College of Arts and Sciences, and was then jointly appointed to the Department of Biophysics at the University of Rochester School of Medicine and Dentistry, becoming a full professor in both departments (1965-1993).

He is now retired and living in the Adirondacks in Lake Luzerne, New York.