and endoperoxides by the two reactions; one or the other method is definitely preferred for many cases.

On the other hand, the range of hydrocarbon structural types for which CRCC oxygenations have been shown to occur is extremely narrow. Reaction to give dioxetanes from tetraalkylolefins has not been detected unless 1,1-bis-α-branched alkyl substituents are present, yet too much steric hindrance also prevents oxygenation. The diene to endoperoxide reaction requires some steric hindrance for oxygenation to compete effectively with CRCC Diels–Alder reaction but is more tolerant of steric hindrance than is singlet dioxygen addition, as shown by the successful formation of **37**. Considerably more work defining what hydrocarbon structural types are necessary to observe CRCC reactions is clearly necessary. Thermodynamic constraints appear to seriously limit the reactions which are plausible. Because peroxide cation radicals oxidize near 2.3 V vs. SCE, chain oxygenations of hydrocarbons to peroxides will not work for most mono- or dialkylolefins. O_2 and starting diene (as in the CRCC Diels–Alder reaction[44]) are not the only components possible for reaction with a cation radical, but the necessity for the product being harder to oxidize than the starting hydrocarbon severely limits the compounds which can be added in chain fashion.

σ–π interactions are unusually large in cation radicals, which was argued to lead to the unusually large olefin face selectivities observed in dioxetane formation from **9** and *syn*-**6**. Such large effects caused by fairly long range electronic interactions are both interesting theoretically and would be of practical importance if they can be predicted and controlled. We believe that a great deal of interesting chemistry remains to be discovered for cation radicals.

It is a pleasure to thank my particularly able co-workers who studied the CRCC oxygenation reaction in Madison, Dr. R. Akaba, D. L. Kapp, and M. F. Teasley. We thank the donors of the Petroleum Research Fund, administered by the American Chemical Society, the Wisconsin Alumni Research Foundation, and the National Science Foundation for partial financial support of this work, as well as the Amoco Oil Company for fellowship support for D.L.K. and M.F.T. during parts of this work.

DNA Flexing, Folding, and Function

HENRYK EISENBERG

Polymer Department, The Weizmann Institute of Science, Rehovot 76100, Israel

Received January 5, 1987 (Revised Manuscript Received May 8, 1987)

Huge Size Genome DNA

To a biased chemist DNA certainly is the most interesting macromolecule created by either nature or man. Not until the beginning of the second half of the twentieth century was the functional role of DNA as carrier of biological inheritance clearly established. In 1953 Watson and Crick,[1] in proposing the guanine-cytosine (G-C), adenine-thymine (A-T) base paired structure for double-helical DNA, note that "it has not escaped our notice that the specific base pairing postulated immediately suggests a possible copying mechanism for the genetic material." One page in *Nature (London)*, stating the basic model and its far-reaching biological implications, set the course of the life sciences for decades to follow. DNA can be likened to a linear computer tape in which important information transmitted by heredity is encoded as sets of consecutive bases (A, C, G, T) inscribed in long sequences; the well-known "genetic code" is composed of base triplets, each coding for a given amino acid. Base triplets also exist for signaling the start or the termination of a given message.[2] DNA replication[3] ensures the transmittance of the genetic message to subsequent generations, whereas transcription to messenger RNA (mRNA) eventually leads to translation into a large number of proteins. Proteins in turn constitute the "chemical" machinery of the organism, some of them strongly interacting with DNA in regulating its structure and function, in particular the replication and transcription processes.

DNA as a macromolecule is basically different from the globular proteins with which it is intimately connected in the circle of life sketched above. Globular proteins are moderate in size (extending over distances of about 3–10 nm) and are composed of one or more polypeptide chains (around 10–100 kg/mol (kDa) molar mass). Following synthesis on the ribosome they fold into globular structures and stay this way throughout their life cycle, maintaining active and regulatory binding sites. DNA, on the other hand, is not active in a folded form, though it must maintain essentially folded structures, as will presently become apparent. In contrast to the proteins DNA is of much larger size.[3] Thus, for instance, the DNA of a lower order organism, the *E. coli* bacterium, codes for about 4000 proteins and contains about 4×10^6 base pairs (bp) in the Watson–Crick double-helical structure. The molar mass of one bp is, on the average, 662 g/mol. Stretched out linearly, this DNA would have a molecular length of 1.36 mm, while the diameter of the molecule is only about 2.5 nm.

Henryk Eisenberg was born in 1921 and educated at the Hebrew University in Jerusalem. He has been at the Polymer Department of the Weizmann Institute since 1949. In 1973 he was a Fogarty Scholar in Residence at the National Institutes of Health, Bethesda, MD. Research interests include nucleic acid and chromatin structure, structure of proteins, and protein–nucleic acid interactions in extreme halophilic bacteria, polyelectrolytes, thermodynamics of multicomponent systems, ultracentrifugation, scattering of light, X-rays, and neutrons.

(1) Watson, J. D.; Crick, F. H. C. *Nature (London)* **1953**, *171*, 737.
(2) Stryer, L. *Biochemistry*, 2nd ed.; Freeman: San Francisco, 1981.
(3) Kornberg, A. *DNA Replication*; Freeman: San Francisco, 1980.

0001-4842/87/0120-0276$01.50/0

The DNA of a higher organism, such as man, for instance, is much longer, a human chromosome containing between 4.8×10^7 and 2.4×10^8 bp, corresponding to a stretched-out length of 1.6–8.2 cm. It is quite clear that molecules of this size cannot be accommodated into the restricted space of a biological cell or cell nucleus without considerable folding. DNA compaction, as we shall see, is usually accomplished with the help of proteins, and the resulting structures possess sufficient lability in active regions to allow local unwinding for either replication or transcription. In eukaryotic cell nuclei chromatin, the cellular DNA complexed with core histones and other proteins, strongly interacts with the nuclear matrix, to form the cell's chromosomes.[4] Structural stability and lability interplay strongly to ensure RNA and protein synthesis as well as cellular replication.

Once forcibly released from its nuclear protective environment, DNA encounters great difficulty to maintain itself in unbroken native form. Early solution work on DNA was undertaken with highly purified and protein-free samples of trout sperm, herring sperm, and calf thymus DNA.[5,6] These were believed to have molar masses of about $6–7 \times 10^6$ g/mol (about 9–10 kbp, kbp = kilobase pairs) though it was later shown to correspond more closely to about 2×10^7 g/mol (about 30 kbp).[7,8] This is much smaller than the native genome size, and it soon became apparent that high molar mass DNA breaks even by slow pipetting in the course of the purification procedures.[9,10] The largest DNA samples that could be obtained unbroken, once this was realized, were the T-even or T-odd phage DNA samples whose molar masses were unequivocally established after prolonged sophisticated experimentation.[11] T4 phage DNA, freed of proteins, was shown to have a molar mass of 1.12×10^8 g/mol, and it could be broken in the middle of the chains into halves, and quarters, by careful syringing at well-defined rates of shear. Twenty years ago, and before restriction enzymes were discovered, this was used as a tool for genetic research.[9,10]

Considering that it was utterly impossible to isolate and purify unbroken eukaryotic DNA, and also that the classical physical methods of determining molar masses failed as the size of the DNA increased, it became a challenging problem to determine whether a total genome, or a single chromosome of a eukaryotic organism, is formed by a single, or by a multiplicity of chains. Bruno Zimm and his colleagues in a series of elegant investigations[12–14] proved the former hypothesis to be true.

In viscous flow deformable macromolecules undergo continuous expansion and contraction following tension and compression in the process of flow. When the flow is stopped, the deformed particles reassume random configurations as the result of thermal motion. Cells of *Drosophila*, bacteriophage, and bacteria were lysed in situ, in the measuring vessel of a low-speed rotating viscoelastometer, and the recoil time was determined after stoppage of flow. The recovery of the deformation of the large DNA molecules, deformed in slow shearing flow, could thus be determined. Even if some of the large DNA molecules have been broken in the mild handling to which they were being exposed, the longest relaxation time measured in a superposition of exponential decays determines a quantity associated with the size of the intact molecules that can be calculated on the basis of Zimm's theoretical treatment.

The experimental studies and theoretical analysis of Zimm and colleagues were confirmed in a no less elegant approach of Feher and his colleagues.[15] Light scattering is a method of choice for determining the mass, size, and shape of macromolecules in solution;[16] however, as the size of the macromolecules increases considerably beyond the size of wavelength λ of the light, recourse to increasingly lower angles of scattering is not sufficient to compensate for the large size of the object under study. Classical elastic light scattering, even with the excellent optics now provided by focused laser light sources, is thus limited to DNA molecules not larger than about 2.5×10^7 g/mol.[17,18] A modern variation of light scattering appropriate for the study of very large molecules is fluorescence correlation spectroscopy.[15] The fluorescence intensity scattered from a small number of particles in a small scattering volume can be followed as a function of time as molecules leave and enter the scattering volume. Correlation analysis relates the fluctuating quantity to the number of particles, or the molecular size, if the absolute amount in the volume is known. Unfortunately, for the large particles of interest the fluctuations are so slow as to render the experiment impracticable. The trick was that rather than to follow the process with time, to rotate slowly the sample and thereby look at a large number of scattering volumes at identical times, which was a cunning application of the ergodic principle of statistical mechanics. Needless to say, the results of Weissmann et al.[15] confirmed the earlier analysis of Zimm and colleagues.[12–14]

In a number of recent publications[19,20] a general method was described for preparing intact nuclear DNA. Cook and his colleagues had previously shown that essentially intact DNA could be isolated in a pipettable form from only those cells (e.g., HeLa cells) that possess a nuclear "cage" robust enough to protect it. Yet most primary diploid cells possess weak cages that break easily, releasing the DNA, which is then broken in shear. In the current work Cook[19] "encapsulates living cells in agarose microbeads by homogenizing an agarose phase containing cells in molten agarose with an immiscible phase of liquid pa-

(4) Nelson, W. G.; Pienta, K. J.; Barrack, E. R.; Coffey, D. S. *Annu. Rev. Biophys. Biophys. Chem.* **1986**, *15*, 457.
(5) Eisenberg, H. *J. Polymer Sci.* **1957**, *25*, 257.
(6) Cohen, G.; Eisenberg, H. *Biopolymers* **1966**, *4*, 429.
(7) Harpst, J.; Krasna, A.; Zimm, B. H. *Biopolymers* **1968**, *6*, 585, 595.
(8) Eisenberg, H. In *Basic Principles in Nucleic Acid Chemistry*; Ts'o, P. O. P., Ed.; Academic: New York, 1974; Vol. 2, p 171.
(9) Burgi, E.; Hershey, A. D. *J. Mol. Biol.* **1961**, *3*, 458.
(10) Levinthal, C.; Davison, P. D. *J. Mol. Biol.* **1961**, *3*, 676.
(11) Schmid, C. W.; Rinehart, F. P.; Hearst, J. E. *Biopolymers* **1971**, *10*, 883.
(12) Kavenoff, R.; Klotz, L. C.; Zimm, B. H. *Cold Spring Harbor Symp. Quant. Biol.* **1974**, *38*, 1.
(13) Uhlenhopp, E. L.; Zimm, B. H.; Cummings, D. J. *J. Mol. Biol.* **1974**, *89*, 689.
(14) Zimm, B. H. In *Methods in Enzymology, Enzyme Structure*; Vol. 117, Part J; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic: New York, 1985; p 94.
(15) Weissmann, M.; Schindler, H.; Feher, G. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 2776.
(16) Eisenberg, H. In *Procedures in Nucleic Acid Research*; Cantoni, G. L., Davies, D. R., Eds.; Harper and Row: New York, 1971; Vol. 2, p 137.
(17) Harpst, J. A. *Biophys. Chem.* **1980**, *11*, 295.
(18) Harpst, J. A.; Sobell, E. S. *Biophys. J.* **1983**, *41*, 288a.
(19) Cook, P. R. *EMBO J.* **1984**, *3*, 1837.
(20) Jackson, D. A.; Cook, P. R. *EMBO J.* **1985**, *4*, 913.

raffin; on cooling, suspended agarose droplets gel into microbeads. The number of cells enclosed in the microbeads can be varied. The pores in the beads are large enough to allow free exchange of quite big molecules but no chromosomal DNA. Therefore, encapsulated cells can now be lysed and extracted to yield essentially pure DNA which is completely protected from shear." This method allows for a variety of cell components to be selectively removed by the use of salts, solvents, detergents, or enzyme action. With care, the nucleic acid components remain supercoiled and attached to points of anchor. Structural features are maintained and can be subjected to detailed investigation. Specific enzymes and other components may be reintroduced into the nuclear domain, allowing precise study of nucleo–protein interactions.

A powerful analysis of the size of DNA molecules over a wide range of sizes can be made by gel electrophoresis in poly(acrylamide) or agarose gels.[21] Gel electrophoresis is used extensively for both proteins in sodium dodecyl sulfate, an ionic denaturing detergent, and for nucleic acids. It is worth considering the principle behind it in a qualitative way.

In distinction to gel chromatography with two phase beaded gels, both poly(acrylamide) and agarose gels are continuous macromolecular networks, the pore size of which can be changed considerably by adjusting the gel concentration. In the electrophoretic motion brought about by the application of the electric field, the mobility (u) of the macromolecular particles is proportional to the number of charges, and this in turn is proportional to the molar mass (M). Steady motion in the field results in interaction with molecular friction (f), and u becomes proportional to M/f. Small fragments of DNA, not exceeding 100–200 bp in size, can be approximated by rods, for which $f \sim M$ in aqueous solution, nonencumbered by the gel. In solution u is therefore independent of M, and the efficient size separation of small DNA fragments, differing by a single bp, is due entirely to the interaction with the gel. Gel concentration thus becomes a major experimental variable in determining efficient separations with respect to size. Larger pieces of DNA coil (cf. below), and the electrophoretic behavior becomes more complex.[22-24] In gel-free medium $f \sim R$, the radius of the coil, and for Gaussian coils, $R \sim M^{1/2}$; thus $u \sim M/f \sim M^{1/2}$. Free diffusion is not possible in the restricted pores of the gel and is replaced by a mechanism of motion, reptation, characteristic of concentrated macromolecular systems.[25] The motion is a simile to the motion of a snake or a worm in a curved tunnel. In reptation $f \sim M^3/R^2$; for Gaussian coils $R^2 \sim M$, thus $f \sim M^2$ and $u \sim 1/M$. This is a first approximation to the observation that the mobility of DNA decreases with increasing M.

Gel electrophoresis of DNA becomes inefficient with increasing M, and an upper limit of separation corresponds to about 50 kbp. In an extremely interesting recent development[26] this limitation has been overcome by an approach based on understanding the viscoelastic properties of very large DNA fragments.[23] This was the effect used by Zimm and colleagues[12-14] to determine the molar mass of genome size DNA. For DNA of intermediate size the relaxation to random configuration is almost instantaneous, and only when the molar mass increases considerably do these times become measurable. To markedly increase the resolution of gel electrophoresis, the usual, constant, electric fields were replaced by pulsed fields with variable pulse length and frequency.[26] While the very large DNA molecule is pulsed with a high electric field, it will deform and then relax with a time constant characteristic of its molecular size as the field is removed. As the deformation affects its motional ability, it is immediately seen that a valuable scale expansion parameter is introduced for separating molecules of large size for whom the separation efficiency is not sensitive enough to M. Separation of large size DNA has now been moved into the 2000-kpb range, and large defined fragments of the human genome can be separated, on the way toward the total characterization of this genome.

Phase DNA, Plasmids, and Restriction Fragments

As already mentioned, early structural work on DNA suffered from the fact that most of the DNA samples investigated were random fragments of unequal size. Interesting results, in particular in equilibrium sedimentation of DNA in CsCl gradients, were later obtained with intact, uniquely sized bacteriophage DNA.[27] Molar masses may be derived from the half-width of Gaussian bands in such experiments without any knowledge of partial volumes, interaction parameters, and their dependence on pressure, by means of an internal calibration procedure, in which the shift of the center of the band is determined following isotope substitution (^{15}N for ^{14}N) in the sample.[28,29] For bacteriophage T4 DNA $M = (113 \pm 6) \times 10^6$ g/mol, for T5 $M = (68.7 \pm 6) \times 10^6$ g/mol, and for T7 $M = (24.8 \pm 0.4) \times 10^6$ g/mol; cf. ref 30 (Table 5.5) for a comparison of values of M by various experimental approaches. Controlled double-strand breakage effected by shear or ultrasonic treatment made feasible recombination studies based on the original observation of Marmur and Doty[31] that DNA strands separated by heat can be renatured, and this led to the disclosure of significant information on genome structure and function. The original proof by Meselson and Stahl[32] of semiconservative DNA replication was based on equilibrium sedimentation in CsCl gradients and isotope substituted growth media and did not require the use of accurately sized DNA.

A turning point in DNA studies and understanding came with the discovery of plasmid DNAs, which are bacterial circular DNA molecules. These replicate independently from the main bacterial genome, execute specific functions, and contain between about 2 and 200

(21) *Gel Electrophoresis of Nucleic Acids*; Rickwood, D., Hames, B. D., Eds.; IRL Press: London, 1982.
(22) Lumpkin, O. J.; Dejardin, P.; Zimm, B. H. *Biopolymers* **1985**, *24*, 1573.
(23) Slater, G. W.; Noolandi, J. *Biopolymers* **1986**, *25*, 431.
(24) Hurley, I. *Biopolymers* **1986**, *25*, 539.
(25) deGennes, P.-G. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca, NY, 1979.
(26) Schwartz, D. C.; Cantor, C. R. *Cell* **1984**, *37*, 67.
(27) Schmid, C. W.; Hearst, J. E. *Biopolymers* **1971**, *10*, 1901.
(28) Eisenberg, H. *Biopolymers* **1967**, *5*, 681.
(29) Cohen, G.; Eisenberg, H. *Biopolymers* **1968**, *6*, 1077.
(30) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I., Jr. *Physical Chemistry of Nucleic Acids*; Harper and Row: New York, 1974.
(31) Marmur, J.; Doty, P. *J. Mol. Biol.* **1962**, *5*, 109.
(32) Meselson, J.; Stahl, E. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 671.

kbp.[33] Electron microscopy revealed that plasmids are circular (no free ends) and supercoiled from natural torsional constraints, with occasional relaxed circles due to accidental single-strand breaks. The usefulness of restriction enzymes is their ability to produce specific cuts in DNA molecules, leading to specific and accurately sized DNA fragments;[34] these can be either investigated or used, in combination with other DNA fragments, in genetic engineering technology.[35] A host of additional powerful enzymes such as gyrases and topoisomerases[36,37] have become indispensable in manipulating DNA in the test tube or in understanding the role of supercoiling in DNA replication, recombination, and transcription.

Following earlier work on the non-Newtonian viscosity of DNA solutions[5] and evaluation of the hydration and partial volume properties of DNA[29] consistent with the interpretation of the behavior of multicomponent solutions,[38,39] we realized the importance of characterizing a well-defined plasmid DNA.[40] *Col*E1 DNA was studied by light scattering, sedimentation, viscosity, and ultracentrifugation in the supercoiled form I, the circular relaxed form II, and the linear form III. The latter was obtained by linearizing form I by use of *Eco*RI restriction enzyme for which *Col*E1 has a single restriction site. For the molecular weight of *Col*E1 DNA an average of $(4.3 \pm 0.05) \times 10^6$ g/mol was obtained from light scattering, diffusion, and sedimentation.[40-42] This corresponds well to the value derived from the sequence, 6646 bp or 4.40×10^6 g/mol, for which a definitive value has now been given.[43]

Double-helical DNA is a stiff macromolecule, the flexibility of which is expressed by the persistence length a of the Kratky–Porod persistent or wormlike chain;[44] a loosely relates to the persistence of direction along the macromolecular chain. A physical basis for the rather abstract concept of a is based on the elastic properties of a curving chain,[45] leading to $a = b/kT$, where b is an elastic force constant. For a rigid rod $a = \infty$, as the initial direction along the rod persists forever. For a Gaussian coil a is numerically equal to one-half the length of the Kuhn statistical chain element.[8] For chains without excluded volume a can be analytically calculated from the radius of gyration of the molecule R_g, experimentally determined from the angular dependence of scattered light. In Figure 1, we have simulated the coiling behavior of wormlike coils with persistence length a = 60, 50, and 40 nm. It is immediately obvious that nucleosome core length DNA,[46,47] 146 bp, behaves almost like a rigid rod in solution, whereas *Col*E1 DNA, a moderate size DNA, behaves like a Gaussian coil. By quasi-elastic laser light scattering from DNA solutions, it is possible to discern dynamic properties such as internal motion, in addition to translational diffusion[48,49] though a definitive understanding of the relaxation processes has not yet been achieved.[48-50]

Table I.
Comparison of Persistence Lengths, a, Corrected for Excluded Volume by Different Methods

[Na⁺], M	R_g,[a] nm	d,[b] nm	α[c]	a, nm
0.007	244.8	18.9	1.072	77.8[d]
		17.2	1.06	82.0[e]
		18.9		72.5[f]
				68.4[g]
0.05	210.2	7.4	1.055	57.5[d]
		6.1	1.04	60.0[e]
		7.4		55.8[f]
0.20	186.0	4.4	1.040	45.7[d]
		3.5	1.04	46.5[e]
		4.4		44.1[f]
				40.8[g]
1.0	159.5	2.95	1.045	32.7[d]
		2.33	1.04	33.0[e]
		2.95		32.3[f]
				30.3[g]

[a] Radius of gyration from light-scattering data of Borochov et al.[41] [b] Effective cylinder diameter. [c] Calculated linear expansion parameter. [d] As calculated by Post.[51] [e] As reported by Manning.[52] [f] Determined by Stigter.[53] [g] As reported by Kam et al.[48]

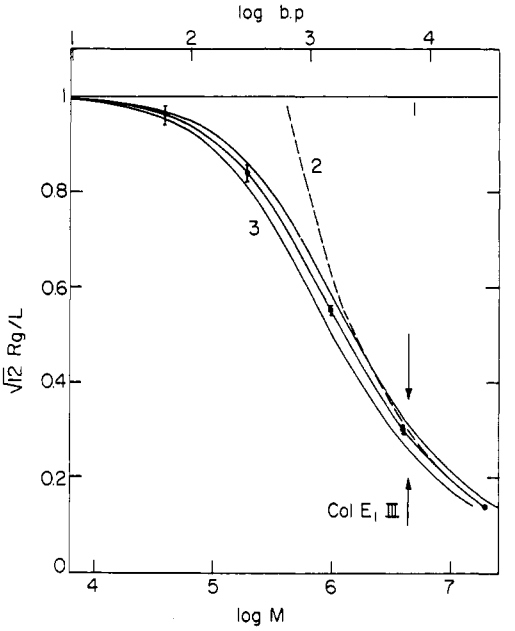

Figure 1. Coiling ratio $12R_g/L$ of NaDNA chains against log M (or log number of base pairs, bp): curve 1, rigid rod; curve 2, Gaussian coil with statistical element $2a$ equal to 100 nm; set of curves 3, Kratky–Porod wormlike chains without excluded volume, for (in descending order) persistence lengths a equal to 60, 50, and 40 nm. Experimental points have been simulated for a = 50 nm, assuming a combined error in R_g and L of ±2%; a repeat length 0.34 nm/bp and M/L = 1950 g/(mol nm) have been assumed for the B form of NaDNA. Reprinted with permission from ref 48. Copyright 1981 Wiley.

(33) Smith, H. O. *Science* **1979**, *205*, 455.
(34) Roberts, R. J. *Gene* **1978**, *4*, 183.
(35) Old, R. W.; Primrose, S. B. *Principles of Gene Manipulation*, 3rd ed.; Blackwell: Oxford, 1985.
(36) Fisher, L. M. *Nature (London)* **1984**, *307*, 686.
(37) Gellert, M.; Menzel, R.; Mizuuchi, K.; O'Dea, M. H.; Friedman, D. I. *Cold Spring Harbor Symp. Quant. Biol.* **1983**, *62*, 763.
(38) Casassa, E. F.; Eisenberg, H. *Adv. Protein Chem.* **1964**, *19*, 287.
(39) Eisenberg, H. *Biological Macromolecules and Polyelectrolytes in Solution*; Clarendon: Oxford, 1976.
(40) Voordouw, G.; Kam, Z.; Borochov, N.; Eisenberg, H. *Biophys. Chem.* **1978**, *8*, 171.
(41) Borochov, N.; Eisenberg, H.; Kam, Z. *Biopolymers* **1981**, *20*, 231.
(42) Eisenberg, H. In *Biophysics—Nucleic Acids*; Landolt Börnstein New Series; Saenger, W., Ed.; Springer-Verlag: West Berlin, in press.
(43) Chan, P. T.; Ohmori, H.; Tomizawa, J.; Lebowitz, J. *J. Biol. Chem.* **1985**, *260*, 8925.
(44) Kratky, O.; Porod, G. *Recl. Trav. Chim. Pays-Bas* **1949**, *68*, 1106.
(45) Landau, L.; Lifshitz, E. *Statistical Physics*; Wiley-Interscience: New York, 1958; p 478.
(46) Richmond, T. J.; Finch, J. T.; Rushton, B.; Rhodes, D.; Klug, A. *Nature (London)* **1984**, *311*, 532.
(47) Ausio, J.; Seger, D.; Eisenberg, H. *J. Mol. Biol.* **1984**, *176*, 77.
(48) Kam, Z.; Borochov, N.; Eisenberg, H. *Biopolymers* **1981**, *20*, 2671.
(49) Soda, K.; Wada, A. *Biophys. Chem.* **1984**, *20*, 185.

Light scattering measurments were undertaken[41,48] with form III of unique size *Col*E1 DNA over a wide range of salt concentrations to determine the radius of gyration (R_g), taking into account corrections for excluded-volume of finite size chains.[48,51-53] A brief summary of our data is given in Table I adapted from ref 51. The values of a at high salt concentrations were confirmed in solutions of LiCl in which the virial coefficients, expressing the excluded volume, vanish.[54]

The flexibility of short monodisperse restriction fragments comprising as few hundred base pairs has also been determined at low salt concentrations by electrooptical measurements from rotational diffusion and relaxation.[55,56] The values of a, derived from these measurements, are believed not to change[55] or to change little[56] with increase in NaCl concentration above 1 mM, in distinction to our light scattering results (Table I) on larger plasmid DNA. High precision is required as can be appreciated by noting the uncertainty in a produced by a ±2% variation of molecular dimensions for short pieces of DNA, as simulated in Figure 1. Higher valency ions such as Mg^{2+}, spermine, spermidine, and $[Co(NH_3)_6]^{3+}$ on the other hand induce strong bending of even short DNA fragments,[56] leading to values of a similar as observed by us at very high concentrations of monovalent salt.[41,48] This may be related to toroidal condensation of DNA with polyvalent cations,[57] a strong indication that a natural trajectory of DNA, unperturbed by thermal fluctuations, may be bent, rather than linear. In protein-induced DNA bending, which is of major biological significance, λ-repressor binds cooperatively to sites separated by integral turns of the DNA helix; the DNA forms a loop allowing the two repressors to touch.[58] Nucleosome core DNA, complexed with core histones, is bent considerably,[46] quite unlike its expected stretched-out configuration in solution (Figure 1).

DNA need not necessarily be in the classical right-handed B helical form of Watson and Crick,[1] in which it is usually found in nature. In recent years a left-handed helical Z form has been discovered and characterized on a molecular level[59] by X-ray diffraction of crystals of a guanine-cytosine (G-C) oligomer. A favorite pursuit today is "Z DNA in search of a biological role".[60] The B to Z transition is favored by GC sequences and multimolar salt concentrations which may be considerably reduced if trivalent ions such as cobalt salts referred to above are used and, if in addition, cytosines following guanines are methylated.[61] Methylation of DNA is important in gene control.[62] An outstanding review concerning the role of ionic interactions in DNA solutions is by Record,[63] based on the central ion condensation concept in polyelectrolytes, elaborated by Manning.[64]

As noted above, the tacit assumption, based on the low-resolution fiber diagrams and molecular model building which led to the original Watson and Crick B structure, that the hydrogen-bonded base pairs stack in parallel layers, leading in unperturbed DNA to straight-chain trajectories over short distances, may require modification. Molecular details have emerged from recent high-resolution X-ray diffraction analysis of DNA oligomer molecular crystals of defined sequence, such as a B-form dodecamer[65] or an A-form octamer.[66] A surprising discovery of increasing interest derives from observations[67,68] that kinetoplast DNA fragments migrate anomalously slowly in electrophoresis gels. It could indeed be established that this effect was caused by inherent DNA curvature, rather than by increased flexibility. Curved DNA seems to be of wider significance than originally assumed, and a number of examples as well as implications have been discussed in a recent review.[69]

A "wedge" model of adenine pairs repeated at 10–11-bp intervals was proposed to explain inherently curved DNA.[70,71] Other explanations have been advanced,[72] and the matter is still undecided. It should be possible, however, to reach a decisive explanation after some recent experiments with synthetic curved DNA. As is well-known, sufficiently long DNA stretches can be linked in circles, and Shore et al.[73] have shown that the highest circularization efficiency of nonconstrained DNA is for linear chains of between 366 and 1361 bp. Ulanovsky et al.[74] synthesized oligomers of a 21-bp precursor which was designed to be naturally curved. It was expected that these ligated oligomers could form circles which would then close easily, primarily because of inherent curvature coupled to flexibility and thermal motion. The authors found that the fraction of circular products peaked sharply around 126 bp, well below the natural circularization optimum[73] for B-form DNA. The "wedge" hypothesis was criticized recently by Hagerman,[75,76] who showed that oligomers $(CA_4T_4G)_N$ and $(GA_4T_4C)_N$ moved abnormally slowly in electrophoretic gels, but oligomers $(CT_4A_4G)_N$ and $(GT_4A_4C)_N$, which Hagerman thought should move equally slowly purely on symmetry considerations, behaved quite normally. Ulanovsky and Trifonov[77] estimate from these very data the roll and tilt components of the wedge formed by adenine pairs, in their models, to be 8.4° and 2.3°, respectively, essentially confirming

(50) Schurr, J. M.; Schmitz, R. S. *Annu. Rev. Phys. Chem.* **1986**, *37*, 271.
(51) Post, C. B. *Biopolymers* **1983**, *22*, 1087.
(52) Manning, G. S. *Biopolymers* **1981**, *20*, 1751.
(53) Stigter, D. *Macromolecules* **1982**, *15*, 635.
(54) Borochov, N.; Eisenberg, H. *Biopolymers* **1984**, *23*, 1757.
(55) Hagerman, P. J. *Biopolymers* **1981**, *22*, 1503.
(56) Porschke, D. *J. Biomol. Struct. Dyn.* **1986**, *4*, 373.
(57) Gosule, L. C.; Schellman, J. A. *J. Mol. Biol.* **1978**, *121*, 311.
(58) Hochschild, A.; Ptashne, M. *Cell* **1986**, *44*, 681.
(59) Wang, A. H.-J.; Quigley, G. J.; Kolpak, F. J.; Crawford, J. L.; van Boom, J. H.; van der Marel, G.; Rich, A. *Nature (London)* **1979**, *282*, 680.
(60) Marx, J. L. *Science* **1985**, *230*, 794.
(61) Behe, M.; Felsenfeld, G. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 1619.
(62) Felsenfeld, G.; McGhee, J. D. *Nature (London)* **1982**, *296*, 602.
(63) Record, M. T., Jr.; Anderson, C. F.; Lohmann, T. M. *Q. Rev. Biophys.* **1978**, *11*, 103.
(64) Manning, G. S. *Q. Rev. Biophys.* **1978**, *11*, 179.
(65) Kopka, M. L.; Fratini, A. V.; Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1983**, *163*, 129.
(66) Shakked, Z.; Rabinovich, D.; Cruse, W. R. I.; Egert, E.; Kennard, O.; Sola, G.; Salisburg, S. A.; Vismamitra, M. A. *Proc. R. Soc. London, Ser. B* **1981**, *213*, 479.
(67) Simpson, L.; Simpson, A. M.; Kidane, G.; Livingston, L.; Spithill, T. W. *Am. J. Trop. Med. Hyg.* **1980**, *29*(5) Suppl., 1053.
(68) Marini, J. C.; Levene, S. D.; Crothers, D. M.; Englund, P. T. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 7664.
(69) Trifonov, E. N. *CRC Crit. Rev. Biochem.* **1985**, *19*, 89.
(70) Trifonov, E. N.; Sussman, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 3816.
(71) Trifonov, E. N. *Nucleic Acids Res.* **1980**, *8*, 4041.
(72) Lilley, D. *Nature (London)* **1986**, *320*, 487.
(73) Shore, D.; Langowski, J.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4833.
(74) Ulanovsky, L.; Bodner, M.; Trifonov, E. N.; Choder, M. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 862.
(75) Hagerman, P. J. *Biochemistry* **1985**, *24*, 7033.
(76) Hagerman, P. J. *Nature (London)* **1986**, *321*, 449.
(77) Ulanovsky, L.; Trifonov, E. N. *Nature (London)* **1987**, *326*, 720.

the predictions of the wedge model. The vectorial sum of these components, 8.7°, is obtained from the circularization experiments mentioned above,[74] thus providing a realistic estimate of the inherent curvature of DNA with repeating runs of adenine and a simple demonstration of its very existence.

Partial Volumes and DNA Hydration

Hydration is an important contributor in maintaining the stability of the double helix in either the A, B, or Z form, as has been pointed out recently[78] on the basis of high-resolution X-ray diffraction studies locating water molecules in DNA oligomer crystals.[65,66] For many years DNA hydration has been studied by solution methods, leading to operational definitions of this elusive concept. Thus, hydration values may depend, in addition to the strength of hydrogen bonding, localization, or volume change, on the experimental method used for its determination, be it a hydrodynamic, thermodynamic, spectroscopic, NMR, X-ray, or neutron diffraction or scattering approach. For a critical discussion I refer to the work of Kopka et al.,[65] who have determined ordered water structure around a B-form DNA dodecamer and have compared their results to literature data.

I would like to discuss here briefly an approach to hydration based on thermodynamic solution considerations developed by us over many years.[38,39] A density increment $(\partial\rho/\partial c_2)_\mu$ at constant chemical potential μ of low molecular weight components diffusible through a semipermeable membrane can be determined directly or derived from ultracentrifugation; c_2 is the concentration (in g/mL) of the macromolecular component.

In a three-component system, comprising water, component 1, and a low molecular weight ionic or nonionic component 3, in addition to the macromolecular component 2, an appropriately modified Archimedes buoyancy expression is

$$(\partial\rho/\partial c_2)_\mu = (1 - \bar{v}_2\rho^\circ) + \xi_i(1 - \bar{v}_i\rho^\circ) \qquad (1)$$

where i is either 1 or 3, $\bar{v}$ are partial specific volumes (mL/g), ρ° is the solvent density, and ξ_i is a "preferential" interaction parameter (grams of component i per gram of component 2), indicating interaction of component 2 with both components 1 and 3; ξ_1 or ξ_3 can be calculated if the density increments and partial volumes in eq 1 are known[79,80] (Figure 2). The formal parameters ξ_i may be related to model parameters B_i, and in particular

$$\xi_1 = B_1 - B_3/w_3 \qquad (2)$$

where w_3 is the weight molality (grams of component 3 per gram of component 1) and B_i define an equivalent particle with which are "associated" B_1 and B_3 (grams of components 1 and 3, respectively, per gram of component 2). Equation 2 provides a convenient means for determining both "hydration" B_1 and B_3 (Figure 3). It was possible to derive[80] from earlier data of NaDNA in NaCl and of CsDNA in CsCl solutions[79] values of 3.7 water molecules per nucleotide for the former and 5.9 water molecules per nucleotide for the latter, an average of five water molecules, with an uncertainty of one.

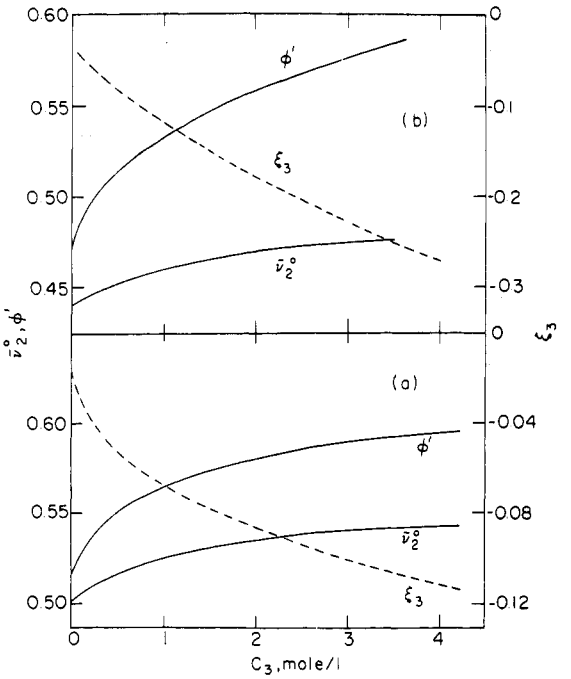

Figure 2. Specific volumes ($\bar{v}_2$ and ϕ') and preferential interaction parameters ξ_3 in DNA solutions: (a) NaDNA in NaCl; (b) CsDNA in CsCl. Reprinted with permission from ref 8. Copyright 1974 Academic.

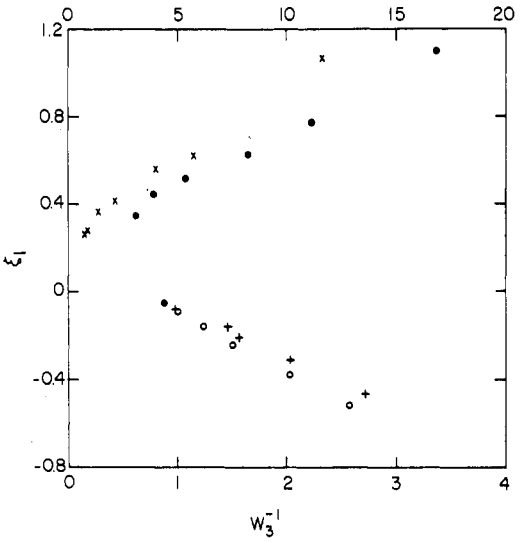

Figure 3. Preferential interaction (hydration) parameter, ξ_1, as a function of reciprocal salt concentration, w_3 (g of salt/g of water), for (○) bovine serum albumin and (+) aldolase solutions in GuHCl (lower scale) and (●) NaDNA in NaCl and (×) CsDNA in CsCl solution (upper scale) (from ref 80).

This compares well with the more recent value of five water molecules, with an uncertainty of one, estimated from the X-ray diffraction of the B dodecamer.[65] Water molecules which are "localized" in the X-ray crystallography sense thus appear similar in considerations relating to volume exclusion in a thermodynamic sense.

Similarly to the mass density constrast expressed in eq 1, it is possible to derive expressions for scattering density contrasts from either light, X-ray, or neutron scattering. In particular, for the last case we derive[81]

$$(\partial\rho_N/\partial c_2)_\mu = (b_2 - \bar{v}_2\rho_N^\circ) + \xi_i(b_i - \bar{v}_i\rho_N^\circ) \qquad (3)$$

where ρ_N° is the scattering length density (in cm^{-2}) of

(78) Saenger, W.; Hunter, W. N.; Kennard, O. *Nature (London)* **1986**, *324*, 385.
(79) Cohen, G.; Eisenberg, H. *Biopolymers* **1968**, *6*, 1077.
(80) Reisler, E.; Haik, Y.; Eisenberg, H. *Biochemistry* **1977**, *16*, 1490.
(81) Eisenberg, H. *Q. Rev. Biophys.* **1981**, *14*, 141.

the solvent, $(\partial\rho_N/\partial c_2)_\mu$ is the scattering length density increment, and b_2 and b_i are scattering lengths per gram of component. These quantities are defined by the known chemical compositions of the components. Equations 1 and 3 can now be solved for $\bar{v}_2$ and ξ_i for identical solvent components without further assumptions. Good results were achieved by the combination of mass and neutron data of halophilic enzymes[82] because b_1 is very different from b_2 or b_3; in fact, it is of opposite sign. We have now applied[42] these considerations to recent neutron scattering data of Lederer et al.[83] on a 130-bp NaDNA fragment of known sequence, containing the strong promoter A1 of the *E. coli* phage T7, and calculate $\bar{v}_2 = 0.503$ in 0.1 M NaCl, in good agreement with well-proven mass measurements, which are rather difficult to perform at high concentrations of salt (Figure 2).

Concluding Remarks

In this work I have summarized structural and biophysical studies on DNA since the discovery of the double helix in 1953, stressing aspects with which I have been personally involved. In our continuing quest for understanding huge size genomic DNA, a most significant advance was effected by the discovery of restriction enzymes, leading to the production of well-defined DNA fragments, useful for physical studies as well as for molecular cloning and sequencing. Powerful tools such as gel electrophoresis, particularly pulsed field gradient gel electrophoresis, X-ray and neutron diffraction, and two-dimensional nuclear manetic resonance, not discussed in this Account, will enable molecular sorting and study of sequence-dependent local structure. Discussion of chromatin, the ultimate folding-function form of DNA,[84-86] could become the topic of a future Account.

My own contribution to this exciting field is overshadowed by that of many outstanding, intellectual contemporaries quoted in this work. It has been a privilege to contribute to an effort which has led to extraordinary advances in our understanding of the processes of life. I acknowledge with gratitude the efforts of my colleagues, quoted in joint works, who have shared my tribulations and excitement. Work of the author was supported by grants from the Israel–U.S. Binational Foundation, Jerusalem, Israel, and from the Minerva Foundation, Munich.

(82) Zaccai, G.; Wachtel, E.; Eisenberg, H. *J. Mol. Biol.* **1986**, *190*, 97.
(83) Lederer, H.; May, R. P.; Kjems, J. K.; Baer, G.; Heumann, H. *Eur. J. Biochem.* **1986**, *161*, 191.
(84) Widom, J.; Klug, A. *Cell* **1985**, *43*, 207.
(85) Felsenfeld, G.; McGhee, J. D. *Cell* **1986**, *44*, 375.
(86) Greulich, K. O.; Wachtel, E.; Ausio, J.; Seger, D.; Eisenberg, H. *J. Mol. Biol.* **1987**, *193*, 709.

Vicarious Nucleophilic Substitution of Hydrogen

MIECZYSŁAW MĄKOSZA* and JERZY WINIARSKI

Institute of Organic Chemistry, Polish Academy of Sciences, Warsaw, Poland

Received October 30, 1986

Nucleophilic substitution in aromatic rings is usually limited to replacement of halogens or other nucleofugal groups. This process can proceed via a variety of mechanisms: addition–elimination (S_NAr), single-electron transfer ($S_{RN}1$), formation of arynes, transition-metal catalysis, etc. In spite of the abundance of mechanistic schemes and the rich possibilities emerging therefrom, no general process for nucleophilic replacement of hydrogen was known until the end of the 1970s. Although there are quite a few earlier reports on reactions which represent nucleophilic replacement of hydrogen, in monographs and textbooks one could often find a message that nucleophilic replacement of hydrogen is rarely observed and that these examples are no more than specific cases.[1]

Professor Mieczysław Mąkosza, born in 1934 in Poland, received his M.Sc. from the University of Leningrad USSR in 1956, his Ph.D. in 1963 from the Technical University, Warsaw, and his DSc (habilitation) in 1967. In 1971, he was a postdoctoral fellow with Prof. G. A. Russell at Iowa State University. Since 1979 he has been director of the Institute of Organic Chemistry, Polish Academy of Sciences, in Warsaw. His main scientific interests include methodology of organic synthesis, reaction of carbanions and halocarbenes in two-phase systems, and nucleophilic aromatic substitution.

Dr. Jerzy Winiarski, born 1953 in Warsaw, received his M.Sc. in 1977 and his Ph.D. in 1981 from Technical University, Warsaw (with Prof. M. Mąkosza). In 1984 he was a postdoctoral fellow with Prof. J. F. Bunnett at University of California, Santa Cruz. Since 1985 he has been a research associate at the Institute of Organic Chemistry, Polish Academy of Sciences, in Warsaw.

Several years ago, we started a project aimed at elaboration of a method of direct nucleophilic replacement of hydrogen in electrophilic arenes, mainly nitroarenes, in reactions with carbanions. In this endeavor, our guiding philosophy was as follows: the nucleophilic substitution of halogen, e.g., in *p*-chloronitrobenzene, proceeds via addition of a nucleophile to the ring carbon atom bearing the halogen, followed by departure of the halide anion. Since activation of the ring by the nitro group is responsible for the addition, similar addition to the ring carbon atom bearing hydrogen is also possible. It is well-known that polynitroarenes form relatively stable adducts with a variety of nucleophiles and that addition to carbon atoms bearing hydrogen is faster than to those bearing other substituents, including halogens.[2] Extrapolation of this rule to mononitroarenes leads to the assumption that, in most cases of nucleophilic aromatic substitution of halogen or other nucleofugal groups, the initial process is the fast and reversible addition of the nucleophiles to carbon atoms

(1) Miller, J. *Aromatic Nucleophilic Substitution*; Elsevier: New York, 1968. De Boer, T. J.; Dirks, J. P. In *Chemistry of the Nitro and Nitroso Group*; Feuer, H., Ed.; Interscience: New York, 1969.
(2) Strauss, M. J. *Chem. Rev.* **1970**, *70*, 667. Terrier, A. *Chem. Rev.* **1982**, *82*, 77. Artamkina, O. A.; Egorov, M. P.; Beletskaya, I. P. *Chem. Rev.* **1982**, *82*, 4[illegible].

0001-4842/87/0120-0282$01.50/0