Proc. Indian natn. Sci Acad. **B70** No. 2 pp 239-254 (2004)

DNA Bending and Curvature: A 'Turning' Point in DNA Function?

ADITI KANHERE[1] and MANJU BANSAL[1,2*]
[1]*Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560 012, India*
[2]*Institute of Bioinformatics and Applied Biotechnology, ITPL, Bangalore-560 066, India*

(Received on 12 April 2004; Accepted after revision on 20 May 2004)

The genomic DNA in the cell is tightly packed and is involved in various functions like transcription, replication and recombination. As a response to the functional requirements as well as different environments in the cell, DNA is known to undergo a number of structural variations from the canonical, uniform, right-handed double-helical B-DNA structure (Watson & Crick 1953). One such structural characteristic of DNA, which is the focus of this review, is DNA curvature and bending. A curved DNA is defined as a double stranded DNA with a nearly smoothly curved helical axis. In the last twenty years, the significance of DNA curvature and bending in different biological functions has been elucidated. A number of experiments have shown presence of curved DNA in many promoter regions and the importance of DNA curvature in the process of transcription has become clearly evident. Many analyses suggest that the curved DNA segments in the promoter region act as transcriptional signals and thus play a role in regulation of gene expression. Various aspects of DNA curvature and bending, related to transcription, replication and nucleosome formation have been reviewed previously, but there have been many new developments in this field in recent days. In addition, with the availability of a large amount of genomic data, an understanding of the structural characteristics of DNA and their functional role has become even more important. In the post-genomic era, when gene prediction has emerged as a very great challenge, identifying additional structural signals will help in improving the success rate of existing algorithms and new efforts are being made to include "curvature signals" in these programs. In this review, we aim to provide an overview of the recent developments in the field of DNA curvature and bending, in both '*in vitro*' as well as '*in vivo*' research. The '*in silico*' predictions and their importance in post-genome era are also discussed.

Key Words: DNA bending, DNA curvature, Dinucleotide parameters, Trinucleotide parameters, Transcription

Introduction

The helical axis of a double stranded DNA molecule is very often deflected from its linear path to take up a bent or a curved structure. In this article we refer to the DNA molecule as being curved, if it shows nearly smooth curvature over a significant length, while it is described as being bent if there is a sharp bend (or kink) that spans only a few base pairs. Both bent and curved DNA is known to play an important role in many protein-DNA interactions and influence vital biological processes such as transcription, replication, recombination and chromatin organization. DNA bending is also known to play an important role in many protein-DNA interactions. The present understanding of the ability of DNA to bend or curve, either as an intrinsic property or under the influence of external factors, has come from various experimental studies using techniques such as X-ray crystallography, NMR, electron microscopy, gel electrophoresis, cyclization kinetics, etc. as well as from different theoretical modeling and database analysis studies. The immensity of the subject makes it difficult to cover all the aspects in detail; hence in this review we have attempted to give a broad overview of DNA curvature and bending. We particularly focus on developments relating to the potential utility of this property of DNA structure in improving various prediction tools being used in genomics. In addition, this review is centered on sequence dependent DNA curvature and does not explore some related areas such as stiffness and persistence length of random sequence DNA. Various aspects of DNA curvature and bending have been reviewed in the past (Hagerman 1990, Bansal et al. 1991, Perez-Martin et al. 1994 Perez-Martin & de Lorenzo 1997, Williams & Maher 2000) and we intend to provide a fresh overview of this subject and avoid overlap with earlier publications.

**Corresponding address*: E-mail: mb@mbu.iisc.ernet.in, mbansal@ibab.ac.in; Tel: 91-80-2293 2534, 91-80-2841 0029 Fax: 91-80-2360 0535

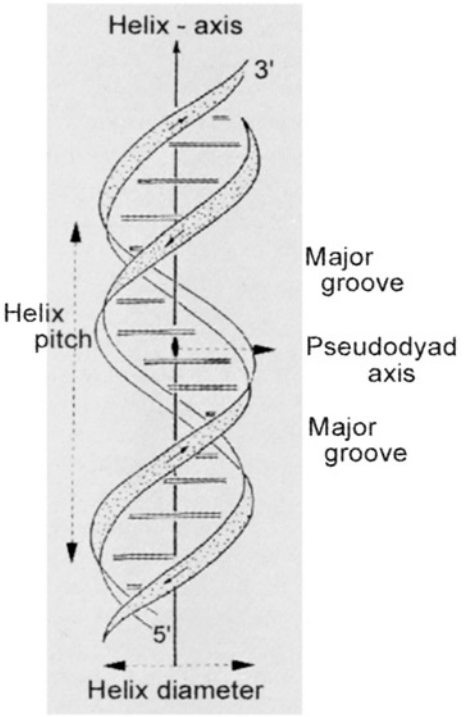

Figure 1: A schematic diagram of the Watson-Crick double helix. Horizontal bars represent the base pairs and ribbons running in opposite directions represent the sugar-phosphate backbones of the two chains, related by a twofold rotation axis perpendicular to the helix axis. The 5′ and 3′ ends are labeled for the ascending strand. The helix axis, as well as pitch, diameter and the major and the minor grooves of the double helix have been indicated. (Reproduced from paper of Ghosh and Bansal (2003) with copyright permission from IUCr).

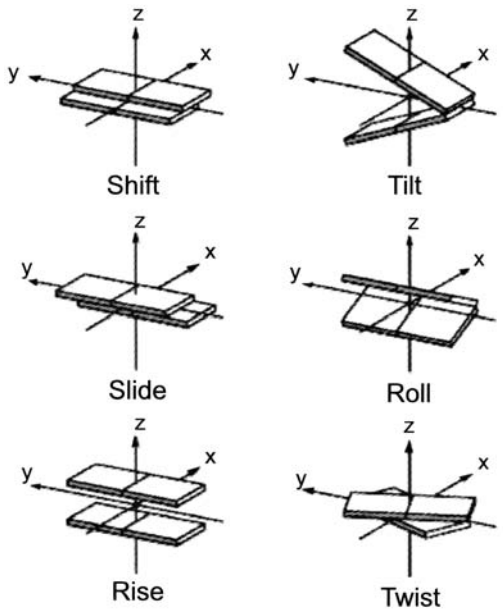

Figure 2: Schematic diagram showing the translational and rotational parameters (Reproduced from Lu and Olson 2003), which describe a dinucleotide step geometry as suggested by the Cambridge convention (Dickerson et al. 1989).

Historical Perspective

In the two decades, immediately following the postulation of the right- handed double helical structure (figure 1) of B-form DNA (Watson & Crick 1953) many other polymorphs of DNA were discovered, using x-ray fiber diffraction studies. In 1953 itself, Rosalind Franklin had shown that another form of DNA called the 'A-form' also exists and depending on the water content, DNA can inter-convert between the A and B forms (Franklin 1953a, Franklin 1953b). Two other related structures of DNA were also characterized *viz.* 'C' (Marvin et al. 1961) and 'D' (Davies & Baldwin 1963) forms. These early studies were followed by other fiber diffraction studies of synthetic polynucleotides with well-defined sequences and these revealed several other variants of DNA structure (reviewed in Leslie et. al. 1980, Chandrasekaran & Arnott 1989).

The first crystal structure for an oligonucletide was solved only in 1978, for the tetramer sequence d(pApTpApT) (Viswamitra et al. 1978). Though this tetramer did not form a double helical structure, within a couple of years a 12-mer oligonucleotide structure showing a complete helical turn of B-DNA was solved (Wing et al. 1980). The discovery of left handed DNA around the same period (Wang et al. 1979, Crawford et al. 1980) came as a surprise to biologists who thought that DNA can only be right handed. In the next 10 years, roughly 1979-1988, many more crystal structures were solved and a lot of efforts were put in to draw up the rules that governed sequence-structure relationship. It became evident that different base sequences display variation in dinucleotide step parameters (figure 2) and hence, in local DNA structure. Thus towards the end of 1980's, considerable evidence had accumulated that the DNA structure is much more complicated than first anticipated and the sequence dependent structural polymorphism of DNA could play an important role in various biological processes. Subsequent studies have confirmed this and it is now clear that the DNA molecule, depending on its sequence, responds to its environment by twisting, turning, looping-out and stretching, which leads to a large number of structural forms (for recent reviews see Bansal 2003, Ghosh & Bansal 2003).

On another front, the discovery of the tightly bound complex of histone proteins and DNA, termed as nucleosome, had been made and researchers speculating about the structure of the DNA folded around the histone core. To explain the packing of DNA, in the limited space corresponding to nucleosomal dimensions, models of 'kinked' DNA (Crick & Klug 1975) or 'bent 'DNA (Selsing et al. 1979)

were proposed. This heralded the beginning of another decade of discoveries of novel DNA conformations such as kinked and curved duplex, as well as quadruplex and cross-over junction structures.

Although some examples of protein induced DNA deformations were reported quite early, the finding of well-defined static curvature in the kinetoplast DNA of trypanosomes (Marini et al. 1982, Marini et al. 1984) really boosted interest in curved DNA structures, arising due to an intrinsic characteristic of the sequence. In these studies it was found that a restriction fragment of kinetoplast DNA (K-DNA) from *Leishmania tarentolae* has unusual physical properties *viz.* a restriction fragment has an apparent size of 450 bp on a 1% agarose gel, but migrates as though it is 1,380 bp in length, on a 12% polyacrylamide gel. The electric dichroism studies, in addition to the gel electrophoresis and gel filtration studies, suggested that this molecule has an unusually compact B-DNA structure. Since the experiments ruled out the possibility of modification of the molecule, it was concluded that these properties could only be due to some feature intrinsic to the kinetoplast DNA sequence. Analysis of the kinetoplast DNA sequence revealed that there is a periodic occurrence of ApA and TpT dinucleotides and Marini et al. attributed the unusual DNA structure to these periodicities. Finally, Wu and Crothers successfully identified the bending locus of K-DNA and found that the sequence around this bending locus contains a striking pattern of periodically repeating $(dA)_{5\text{-}6}$ tracts, separated by four to six base pairs of G+C rich sequences (Wu & Crothers 1984).

'A-tracts'

Next few years were marked by identification of curved conformations in sequences that are involved in important functions such as transcription, replication, recombination and chromatin organization. Most of the natural sequences exhibiting curvature had A-tracts (a stretch containing consecutive A nucleotides) in phase with the DNA double helix repeat of 10-11 residues per turn. It was proposed that each A-tract produces a small bend in the helix axis and repetition of the A-tract elements in phase with the DNA helical repeat, results in their coherent addition leading to an overall curvature. The initial experimental efforts were, hence, based on solution studies of synthetic oligo/poly nucleotides containing A-tracts. The most significant contribution came from the studies done by Donald Crothers and his colleagues at Yale University (Koo & Crothers 1988, Haran et al. 1994). They synthesized a series of polynucleotides containing A-tracts of variable lengths, located in phase with the DNA helix repeat. Most of the synthetic polymers with phased A-tracts showed retardation of mobility on PAGE, a hallmark of DNA curvature. Based on these studies many theoretical models were developed to explain the phenomenon of A-tract related DNA bending. The two most debated models, which tried to explain this, were 1) the 'AA wedge model' (figure 3b), (Trifonov & Sussman 1980) and 2) the 'junction model' (figure 3c) (Levene & Crothers 1983).

Based on systematic modelling studies, Selsing et al. had earlier postulated the presence of a bend at a junction between two contiguous stretches of DNA with different conformations (figure 3c), such as A- and B- DNA (Selsing et al. 1979). The Levene-Crothers' 'junction model', which suggested that the runs of A-nucleotides or A-tracts adopt a non-B conformation, with base pairs at an inclination to the helix axis, is essentially similar to this model. Since the base pairs in the flanking regions, with B-DNA conformation,

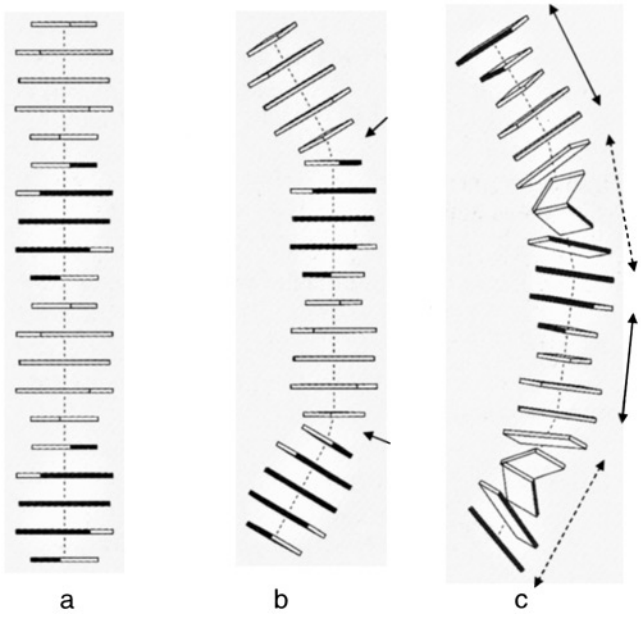

Figure 3: The extent of DNA bending can be described in different ways. (a) A schematic showing 20 base pairs (as horizontal bars) in two complete helical turns of a linear fragment of B-DNA. This can undergo distortion so as to have an effective bending angle of ~ 60°, which can be explained either by the AA wedge model or the junction model. (b) A representative AA wedge model: two AA wedges (with roll angle of ~ 30°) located 10 bp apart, add-up to give rise to a DNA structure bent by ~ 60°. The two AA wedges are indicated by arrows. (c) A junction model: the DNA molecule is bent at the junction between two regions, with different DNA conformations. Stretches with B-DNA like conformation (indicated by continuous double-headed arrow) are shown alternating with stretches of DNA wherein the dinucleotide steps have a ~12° roll angle (or base pairs are inclined at ~ 20° to the helix axis), which are indicated using dotted arrows. Three such junctions also bend the DNA molecule by ~60°. These schematic models have been generated using '3DNA' software (Lu and Olson 2003).

are nearly perpendicular to the helix axis, there is a marked change in the direction of the helix axis at the junction of the two regions. On the other hand, Trifonov-Sussman's 'AA wedge model' suggested that DNA is essentially B-like throughout, but an AA/TT dinucleotide step would open to form 'wedges' by a combination of tilt and roll movements (figure 2) of the two base pairs. This hypothesis was based on the observation that in highly curved nucleosomal DNA, AA/TT dinucleotides appear with the same periodicity as the helical repeat. According to the 'AA wedge model', the cumulative effect of these periodic wedges results in a curved DNA molecule (figure 3b).

Structural studies using techniques of X-ray crystallography and NMR were also carried out in order to understand these properties of DNA. But unfortunately, the availability of structural information did not make things more clear or conclusive. The solution and crystallographic studies provided conflicting results. Many NMR studies on A-tract containing DNA support the idea that A-tracts have a negative inclination in accord with the Junction model (MacDonald et al. 2001). Crystallographic evidence, on the other hand, indicated that A-tracts do not have such an inclination (Dickerson et al. 1994). It has been suggested that the effect of organic solvents, which are routinely used to induce DNA crystallization, is responsible for the discrepancy (Haran et al. 2004). Though the discussion on the exact nature of 'A-tract' contribution to DNA bending still goes on, new experimental data, which showed that sequence elements other than A-tracts also contribute to deflection of the helix axis, has diluted the debate.

General Wedge Models

The first blow to the concept that oligo(dA)-oligo(dT) runs, approximately in-phase with the helix repeat were responsible for DNA curvature, came from an experiment which showed that the polymer $(GTTTTAAAAC)_n$ migrates with normal mobility on a gel, whereas a very similar sequence $(GAAAATTTTC)_n$ migrates with retarded mobility and thus appears to be strongly curved (Hagerman 1986). Both the sequences have runs of As but one of them appears to be straight and the other curved. Additional reports (Abagyan et al. 1990, Milton et al. 1990a, Milton et al. 1990b) showed that not only A-tracts but the flanking sequences also make important contribution towards the curvature of DNA. Finally, it became evident that some nucleic acid sequences, entirely lacking in AA dinucleotides, can also take up a curved structure (Brukner et al. 1991). These developments highlighted the need for a more general model for describing the bending of the helical axis in DNA.

Thus new models, which not only take into consideration geometry of the AA dinucleotide but the geometries of all the 16 dinucleotide steps, were proposed. The currently used dinucleotide models are based on direct experimental data from the analysis of synthetic oligomers using gel electrophoresis (Bolshoy et al. 1991), NMR (Ulyanov & James 1995, Gabrielian & Pongor 1996) and X-ray crystallography (Bansal et al. 1995, Gorin et al. 1995, El Hassan & Calladine 1995). In these models, the geometries of dinucleotides are described in terms of the three rotational parameters *viz.* roll, tilt and twist angles (Dickerson et al. 1989, Olson et al. 2001).

By analyzing high-resolution oligonucleotide crystal structure data (Bansal et al. 1995, Bansal 1996), a dinucleotide model (CS) for DNA bending has been developed in our laboratory. The CS model is based on mean values of dinucleotide step parameters (figure 2) calculated from the crystal structure data (table 1). Since the compilation of our original dataset, only a few oligonucleotide crystal structures with novel sequences have been published. The CS parameters can be upgraded when some more data on novel sequences becomes available. The striking feature of the model is the absence of large roll and tilt for the AA/TT dinucleotide step. Another unique feature of this parameter set is that it includes two distinctly different arrangements of CA/TG steps corresponding to the

Table 1 Dinucleotide model (CS) based on mean values of dinucleotide parameters calculated from high resolution B-DNA crystal structures (Reproduced from Bansal 1996)

	Data points	Roll*	Tilt*	Twist*
CA/TG(BI)*	7	5.1	-0.31	31.03
GG/CC	10	5.02	-1.83	32.4
AG/CT	6	4.3	2.68	29.46
CG/CG	52	3.5	0.15	34.1
TA/TA	9	2.94	0.04	39.94
AA/TT	49	2.6	-1.66	35.58
AC/GT	10	-0.7	-0.15	33.29
AT/AT	30	-1.79	0.21	32.49
GA/TC	25	-2.31	-0.88	38.55
GC/GC	29	-6.49	-0.18	38.61
CA/TG(BII)*	12	-7.5	0.68	47.62

This parameter set includes two distinctly different arrangements of CA/TG steps corresponding to the B-I and B-II backbone conformations. The roll and tilt values corresponding to the B-II conformation are used when the CA/TG step is either preceded by a pyrimidine or followed by a purine. The twist value used for CA/TG steps is the mean of the values of B-I and B-II *viz.* 39.3°

B-I and B-II conformations. The roll and tilt values corresponding to the B-II conformation are used when the CA/TG step is either preceded by a pyrimidine or followed by a purine, e.g. in CCA, TCA, CAA or CAG and their self-complementary sequences involving the TG sequence. Recently we have carried out an assessment of three different dinucleotide step parameters (Kanhere & Bansal 2003). These parameters were evaluated quantitatively for their ability to correctly predict the experimentally determined curvature for a large set of nucleic acid sequences containing A-tracts as well as GC-rich motifs. Our study showed that these dinucleotide parameters with some further refinement can be used to correctly predict sequence-dependent curvature in genomic sequences (Nagaich et al. 1994, Bansal 1996, Kanhere & Bansal 2003).

Software available for Generation and Curvature Analysis of DNA Structures

In order to evaluate the dinucleotide and trinucleotide models, computer programs for generation and analysis of non-linear DNA molecules are needed. Our laboratory has developed an algorithm (Bhattacharya & Bansal 1988, Bansal et al. 1995), which predicts DNA structural trajectory for a given sequence. This program uses a set of local doublet parameters, viz. tilt, roll, twist, shift, slide and rise defined in the dinucleotide model selected according to the user's choice. The DNA structure of a given sequence is generated by applying corresponding rotational angles and translational parameters (figure 2). It is pertinent to mention here that only rotational parameters (roll, tilt and twist) affect the bending of a DNA molecule. These generated structures can be visualized using any of the available molecular visualization softwares. Representative DNA structures, for the synthetic $d(CAAAATTTTG)_n$ and $d(CTTTTAAAAG)_n$ oligomers (Hagerman 1986) and kinetoplast DNA (Linial & Shlomai 1988), generated using our in-house software 'NUCGEN' are shown in figure 4. Our program also allows the estimation of the extent of curvature in terms of different measures such as: (1) d/l_{max} (where 'd' is the end-to-end distance and 'l_{max}' is the actual path traced by a DNA molecule) (2) I_{max}/I_{min} (ratio of largest component of moments of inertia to smallest component of the moments of inertia) (3) Radius of curvature (the radius of circle fitted to the base pair centers) (4) angle between successive local helix axis.

Some other groups (Tan et al. 1988, De Santis et al. 1988, Lavery 1988, Shpigelman et al. 1993, Goodsell & Dickerson 1994, Carter & Tung 1996, Dlakic & Harrington 1998, Vlahovicek & Pongor 2000) have also developed similar programs, which are based on dinucleotide and/or trinucleotide models but differ

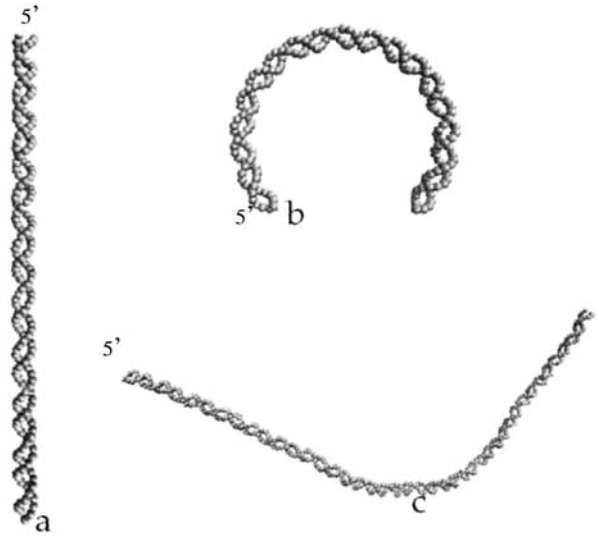

Figure 4. The DNA structures generated using the in-house software 'NUCGEN' (Bansal et al. 1995) for: (a) an almost straight 96-mer with repetitive sequence $d(CTTTTAAAAG)_n$ (b) a 96-mer with the repetitive sequence $d(CAAAATTTTG)_n$ showing large curvature (Hagerman 1986) and (c) a 211 base pair long kinetoplast DNA fragment (Linial and Shlomai 1988). In case of all three structures only the C1' atoms are shown and the 5' end is indicated.

Table 2. *Summary of softwares available for generation and curvature analysis of DNA structures*

Name	Availability	Are the coordinates given?	Measures of curvature	Can dinucleotide parameters be given?	Can trinucleotide parameters be given?	Reference
CURVATURE	edward.trifonov@weizmann.ac.il (E. Trifonov) bolshov@research.haifa.ac.il (A Bolshov)	✓[a]	Curvature (in curvature units)	✓[b]	–	Shpigelman *et al* 1993
BEND	http://www.scripps edu/pub/goodsell/research/bend/	–	Curvature (in degrees/turn)	✓	✓	Goodsell and Dickerson 1994
NUCGEN	mb@mbu.iisc.ernet.in (M Bansal)	✓	d/l_{max}, I_{max}/I_{min}, r,	✓	–	Bansal *et al* 1995
NAMOT2	http://www t10.lanl.gov/namot/	✓	–	✓	–	Carter and Tung 1996
DIMOD	http://www-personal.unich edu/~mensur/software.html, mensur@unich.edu (M. Dlakic)	✓	End-to-end distance	✓	✓	Dlakic and Harrington 1998
Model.it	http://www.icgeb.trieste it/dna/	✓	Curvature (in degrees/turn)	✓	✓	Vlahovicek and Pongot 2000

[a] Only phosphates and base pair centers.
[b] In this case, wedge and direction angles are used instead of roll and tilt as in other cases.
[c] For details see the section 'Software available for generation and curvature analysis of DNA structures'

in algorithmic details. Some of these programs are available on request while some others can be used 'online' (summarized in table 2). It should be noted that the models produced by almost all the above softwares (except 'Model.it' and 'JUMNA' by Vlahovicek & Pongor 2000, Lavery 1988 respectively) may or may not have energetically optimum values of bond lengths, bond angles, torsion angles and van der Waals contacts, especially those related to the backbone atoms. Therefore, in certain cases, energy refinement of these DNA models may be required.

Static Curvature vs. Bendability

Depending on sequence, bending can be an inherent property of a DNA molecule or can be induced by environmental factors. In other words, even in the absence of any external force, some DNA molecules can adopt a static curved structure, while some others can be bent under the influence of environmental factors such as proteins. In all cases of DNA bending, the minor grooves and major grooves of the DNA double helix, which are facing inwards need to be compressed. Some DNA molecules can easily adopt this distortion related to bending and hence can easily take up a bent conformation. Bendability of a DNA molecule can be defined as the ease with which the molecule can be made to bend in any direction. A bendable molecule can exist in a number of different conformations and the equilibrium can be shifted towards one particular conformation in the presence of an external force, such as binding of a protein or some other ligand. This can also be due to other reasons such as asymmetric charge neutralization of the phosphate groups (Williams & Maher 2000). This differs from the case of static curvature, where DNA exists as a curved molecule due to the structural properties intrinsic to its sequence and hence, is a relatively rigid conformation. It is difficult to make a distinction between curvature and bendability, experimentally (Rivetti et al. 1998 Cognet et al. 1999, Zuccheri et al. 2001, Scipioni et al. 2002) but it is necessary to understand the difference between the two, especially considering their relevance in protein-DNA interaction. Three representative examples, where protein induces different extent of DNA bending, are shown in figure 5.

Bendability Parameters

Presently, only two models are available which are used to predict the DNA bending propensity of various sequences. These models are based on trinucleotide parameters calculated from analysis of (1) DNase I digestion data and (2) Nucleosome-forming DNA sequences.

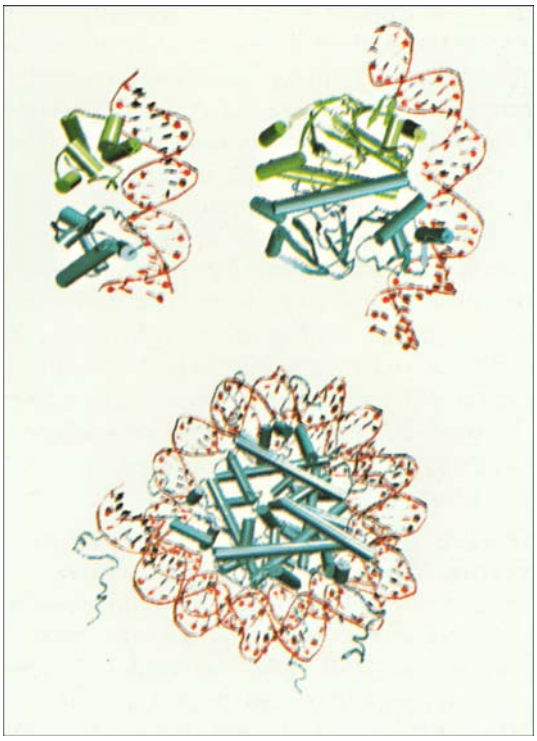

Figure 5:.Three representative structures illustrating different extents of protein induced DNA bending. DNA is shown in red and protein subunits are shown in cyan and green. a) No significant bending: Structure of phage 434 Cro repressor-operator complex (PDB code: 3cro, Mondragon and Harrison 1991). b) Strong bending: CAP dimer bound to DNA (PDB code: 1cgp, Schultz et al. 1991) c) Strong bending: Structure of the nucleosome core particle (PDB code: 1eqz, Harp et al. 2000) where a 146-mer DNA is wound around the histone protein core. All the figures have been generated using the molecular visualization software VMD (Humphrey et al. 1996)

DNase I Sensitivity Based Model

DNase I is an enzyme, which binds to DNA from minor groove side and bends it towards the major groove. The crystal structures of DNase I-DNA complexes have shown that the binding of DNase I to DNA requires DNA to be bent. The DNA molecules, which are flexible and hence capable of bending easily as well as those that are inherently bent towards the major groove should be more prone to cutting with DNase I. Based on this assumption, the cutting frequency of various trinuclotides present in any DNA sequence is used as a measure of its inherent bending/bendability. Based on an analysis of 709 DNase I cutting frequencies, the bending/bendability parameters for 32 trinucleotides (table 3) have been determined (Brukner et al. 1995). It should be noted, however, that DNase I based trinucleotide parameters do not differentiate between bending and bendability.

Table 3. *DNA bending and/or bendability parameters based on DNase I cutting frequencies* (after Brukner et al. 1995)

Trinicleotide Step	DNase I-derived trinucleotide parametres	
AAT/ATT	-0.280	Least bent bendable towards major groove
AAA/TTT	-0.274	
CCA/TGG	-0.246	
AAC/GTT	-0.205	
ACT/AGT	-0.183	
CCG/CGG	-0.136	
ATC/GAT	-0.110	
AAG/CTT	-0.081	
CGC/GCG	-0.077	
AGG/CCT	-0.057	
GAA/TTC	-0.037	
ACG/CGT	-0.033	
ACC/GGT	-0.032	
GAC/GTC	-0.013	
CCC/GGG	-0.012	
ACA/TGT	-0.006	
CGA/TCG	-0.003	
GGA/TCC	0.013	
CAA/TTG	0.015	
AGC/GCT	0.017	
GTA/TAC	0.025	
AGA/TCT	0.027	
CTC/GAG	0.031	
CAC/GTG	0.040	
TAA/TTA	0.068	
GCA/TGC	0.076	
CTA/TAG	0.090	
GCC/GGC	0.107	
ATG/CAT	0.134	
CAG/CTG	0.175	
ATA/TAT	0.182	Most bent bendable towards major groove
TCA/TGA	0.194	

Nucleosome Positioning Based Model

Though nucleosomes are known to form on a variety of DNA sequences, they seem to prefer some sequences to others. Since structural and mechanical properties of DNA are sequence dependent, it is thought that the bending ability of a DNA sequence around the histone is the determining factor for nucleosome positioning. Based on these observations, Satchwell et al. analyzed a number of DNA molecules from chicken erythrocyte nucleosomes and identified the regularities in trinucleotide arrangements (Satchwell et al. 1986). It was observed that there are sequence dependent preferences that correlate most closely with the rotational orientation of the DNA molecule with respect to the protein surface. It was found that certain trinucleotides, such as ApApA/TpTpT and ApApT/ApTpT prefer their minor grooves to face towards the histone core, while some others, like GpGpC/GpCpC and ApGpC/GpCpT prefer to bind with their minor grooves facing outwards. A few of the trinucleotides did not show any such preference (table 4). Subsequently, based on the assumption that flexible sequences can be in any rotational orientation and rigid sequences are restricted to only a particular rotational orientation, these trinucleotide preferences have been used as a measure of flexibility and bendability (Goodsell & Dickerson 1994).

Functional Importance of DNA Bending and Curvature

A) Role of DNA Bending and Curvature in Transcription

Many experimental evidences show that bent DNA structure is closely related to transcriptional activity. Experimental studies on prokaryotes as well as eukaryotes, have shown the presence of such structures in the vicinity of transcription start sites. This kind of

Table 4. *Trinucleotide bendability parameters derived from nucleosomal positioning preferences* (after Satchwell et al. 1986)

Trinicleotide Step	Percentage variation in occurrences	
GCC/GGC	45	Minor groove faces outwards away from histone core
TCG/CGA	31	
AGC/GCT	25	
CGC/GCG	25	
ATG/CAT	18	
CAC/GTG	17	
CCC/GGG	13	
GCA/TGC	13	
ACT/AGT	11	
CTC/GAG	8	
ACC/GGT	8	
CCA/TGG	8	
ACG/CGT	8	
AGG/CCT	8	
GAC/GTC	8	
TCA/TGA	8	
ATC/GAT	7	
ACA/TGT	6	
AAG/CTT	6	
CCG/CGG	2	No rotational preference
CAG/CTG	2	
GGA/TCC	5	
GTA/TAC	6	
AAC/GTT	6	
AGA/TCT	9	
CAA/TTG	9	
GAA/TTC	12	
ATA/TAT	13	
CTA/TAG	18	
TAA/TTA	20	Minor groove faces outwards away from histone core
AAT/ATT	30	
AAA/TTT	36	

structure is either intrinsic to the sequence or is induced by proteins involved in transcriptional regulation.

The first report on the presence of intrinsically bent DNA structure in the promoter region came from a study on tRNA operon (Bossi & Smith 1984). They showed the existence of an unusual DNA conformation, probably involving DNA bending, preceding the tRNA operon promoter of *Salmonella*, located at around -70 position with respect to the transcription start site. The transcriptional defect of the -70 mutants suggested a role for the unusual conformation in promoter function. Over the last two decades, many bacterial promoters known to have intrinsic bends have been reported and some of them are listed in table 5. In general, it is thought that the curvature facilitates DNA-protein interaction and promotes binding of proteins, thus modulating transcription. The protein-DNA contacts are known to have specific interactions, involving the specific bases and amino acids, as well as indirect readouts, which involves non-specific protein contacts with the sugar-phosphate backbone of DNA. The DNA curvature can play a role in maximizing both the nonspecific and specific contacts between DNA and proteins. It is also thought that structural features like bending and bendability can act as a signal for the binding of proteins, which do not have specific binding sites.

Early studies, which addressed the importance of DNA curvature in DNA-protein interaction, involved deletion or modification of curved DNA and observing its effect on the transcriptional efficiency. One such elegant experiment was carried out by Bracco et al. involving the binding site of CAP protein (figure 5b), which is known to induce DNA bending (Bracco et al. 1989). They introduced synthetic and naturally occurring curved DNA sequences upstream of a partially truncated *gal* promoter, devoid of its -35 region and of its CAP binding site. In these hybrid promoters, transcription can not only be restored but some of these hybrid promoters are in fact more efficient than the wild type promoter. On the other hand, control experiments performed using similar sequences without any curvature produced weak promoters. In a related study, it was shown that appropriately phased DNA bending sequences, in place of the CAP binding site, upstream from the *lac* promoter could increase the rate of specific transcription initiation by roughly tenfold (Gartenberg & Crothers 1991). A significant number of proteins other than CAP are known which bend DNA and are involved in modulating transcription. It has been shown that the CAP protein can be effectively replaced with another protein (IHF), which bends the DNA to a similar degree as CAP (Dethiollaz et al. 1996).

The CAP induced DNA bending (figure 5b) plays a key role in regulation of different promoters. The CAP regulated promoters can be categorized into three

Table 5. *Examples of promoters reported to have static DNA curvature intrinsic to their sequence (shown in reverse chronologicalorder)*

Gene Name	Organism	Distance of the bent site from TSS	Reference
Prokaryotic			
virF	*Shigella flexneri*	-137	Prosseda et al. (2004)
per-fdx	*Clostridium perfringens*	-43	Kaji et al. (2003)
Streptokinase	*streptococcus equisimilis* H 46A	-98	Malke et al. (2000)
aprE	*Bacillus subtilis*	-103	Jan et al. (2000
nifLA	*Klebsiella pneumoniae*	-95	Cheemna et al (1999)
GyrA	*Streptococcus pneumoniae*	-23	Balas et al (1998)
appY	*Escherichia coli*	-350 (from start codon)	Atlung et ai. (1996)
rrnB PI	*Escherichia coli*	-110	Gaal et al. (1994)
ompF	*Escherichia coli* K-12	-101 to -71	
Eukaryotic			
pabA2	*Microcystic aeruginosa* K-81	-161	Agrawal et al. (2003)
Adenylate kinase (AKY2)	*Saccharomyces cerevisiae*	-62	Angermayr et al. (2002)
estrogen recept alpha (ERalpha	Human	-848	Li et al. (2002)
DNA Polymerase (delta)	*Plasmodium falciparum*	-1838 to -1312	Porter (2002)
GALI, GAL-10	Yeast	-150, -224	Bash et al. (2001)
EIA	Human adenovirus type -2	-298 to -218	Ohyama (1996)

classes, depending on the position of the CAP binding site with respect to the transcription start site. In case of class-I promoters, the CAP binding site is located at base pair -61, which is upstream of the RNA polymerase (RNAP) binding site, whereas in case of class II promoters the CAP binding site is located at -41, which overlaps the RNAP binding site. At class I and class II CAP-dependent promoters, transcription activation involves protein-protein interactions between CAP and the RNAP. The RNA polymerase is capable of distinguishing between class-I and class-II promoter sites based on differential contacts with CAP, which is in turn a result of different degrees of DNA bending induced by CAP at different promoters (Pyles & Lee 1998, Lin & Lee 2003). It has been proposed that the geometry of a DNA-protein complex, which is a reflection of differential bending, plays a major role in determining the molecular mechanism of gene transcription from class I and class II promoters. Class III CAP-dependent promoters require multiple activator molecules for full transcription activation and the CAP induced bending of DNA is mainly involved in juxtaposing different transcriptional factors located on neighboring sites. Examples include the ansB promoter (Scott et al. 1995), the araBAD promoter (Zhang & Schleif 1998), the malE promoter (Richet 1996), and the uhpT promoter (Chen & Kadner 2000).

Recently, the yeast MAD-box transcription factor Mcm1p mediated DNA bending has been revealed to be involved in maintaining local promoter architecture and permitting formation of ternary promoter complexes (Lim et al. 2003). Other proteins like integration host factor (IHF) are also known to bend DNA and favor contact between upstream DNA bound activator and the downstream RNA polymerase (Engelhorn & Geiselmann 1998). Using atomic force microscopy, a sharp DNA bend induced by binding of IHF is directly observed to facilitate the contact between RNA polymerase bound to the promoter sequence and protein attached to the upstream regulatory regions (Seong et al. 2002).

Thus both sequence dependent intrinsic curvature, as well as that induced by DNA binding proteins, plays a significant role in the process of transcription. The role of protein induced DNA bending in transcription has been discussed in more detail in an earlier review (Perez-Martin & de Lorenzo 1997).

B)Role of DNA Bending and Curvature in Chromatin Organization

As mentioned earlier, DNA bending is also known to play a role in the chromatin organization by influencing nucleosome formation in eukaryotes and by facilitating binding of proteins involved in organizing the genome in prokaryotes. In a nucleosome, DNA coils around the histone octamer (figure 5c) to form a highly compact structure. Although nucleosome formation is overall a sequence independent phenomenon, it has been well established that on some sequences nucleosomes are formed with relative ease (Thastrom et al. 2004). It has been speculated that it is the bendability and presence of curvature in different sequences, which influence ease of nucleosome formation. The role of intrinsic curvature in nucleosome formation is still an open question. Originally it was hypothesized that intrinsically curved DNA, characterized by a periodic occurrence of AA/TT dinucleotide steps, have a high propensity to form nucleosome (Satchwell et al. 1986). But over the last several years, much data has accumulated that supports the idea that static curvature in fact destabilizes nucleosomes (Shrader & Crothers 1989, Shrader & Crothers 1990). It has also been recently demonstrated that the handedness of the curvature decides whether a DNA sequence will attract nucleosomes or not (Nishikawa et al. 2003).

Though nucleus is absent in prokaryotes, nucleic acid is confined to a particular region termed as nucleoid, which serves an analogous function to the nucleus of a eukaryotic cell. There is clear evidence showing that most of the nucleoid associated proteins favorably interact with curved DNA. One of the prominent examples is that of H-NS protein, which is known to be involved in regulation of a number of genes and is known to preferentially bind to curved DNA (Dame et al. 2001). A number of H-NS responsive promoters have indeed shown regions of intrinsic DNA curvature upstream or downstream of the transcription start site. In other words, the available data support the idea that presence of such a curved element in the promoter region makes the corresponding gene available for modulation by H-NS. The protein expression profiling of H-NS-deficient mutant showed increased expression of many genes compared to wild type, confirming its role as a repressor (Hommais et al. 2001). Other nucleoid-associated proteins like CbpA, HU, IciA, StpA are also known to bind curved DNA while some others like FIS are known to induce bends in DNA.

Statically curved DNA elements are also known to be present at many replication origins. Presence of static curvature at the origin of replication was first reported in case of phage λ origin of replication (Zahn & Blattner 1985). Subsequently static curvature has been shown to be present in replication origins of plasmids, viruses, bacteria as well as eukaryotes (table 6). Many of these lie in the vicinity of transcription start site of one or the other gene and it

Table 6. *Examples of replication origins, which are known to show static DNA curvature (shown in reverse chronological order)*

Origin of Replication	Reference
M13	Lu et al. (2003)
Broad host range plasmid RK2	Doran et al. (1998)
Escherichia coli	Polaczek et al. (1997)
replication origin (ori) of leading strand (ori-H) of rat mitochondrial genome	Gadaleta et al. (1995)
Wheat dwarf virus	Suarez-Lopez et al. (1995)
DARC146	Timchenko et al. (1994)
Staphylococcus aureus plasmid pT181	Henriquez et al. (1993)
origin of replication of chicken alpha-globulin gene domains	Kraevskii et al. (1992)
Methylomonas clara plasmid pBE-2	Kues and Stahl (1992)
origin gamma of R6K	Kelley and Bastia (1991)
Chlamydomonas reinharditii chloroplast DNA	Hsieh et al. (1991)
dihydrofolate reductase of CHO cells	Caddle et al. (1990)
human mitochondrial L-strand replication origin	Welter et al. (1989)
Plasmid pLS1	Perez-Martin et al. (1988)
simian virus 40	Hertz et al. (1987)
Plasmid pSC101	Stenzel et al. (1987)
Yeast autonomously replicating sequence (ARSI)	Snyder et al. (1986)
bacteriophage lambda origin	Zahn and Blattner (1985)

is hence possible that these curved regions are also involved in transcriptional regulation of these genes.

In Silico Analysis of Genomic Sequences

With successful completion of genome sequencing of many organisms, ranging from viruses to human, new challenges are being faced in extracting useful information from the sequence data. One of the most central problems is that of gene prediction and annotation. Various computational and bioinformatics techniques are utilized for prediction of genes in the genome data. The existing gene prediction programs are not accurate. Moreover there are discrepancies in the genes predicated by different algorithms. This is mainly because the task of gene prediction is not straightforward e.g., there are genes nested within other genes, genes with overlap, alternate splice variants etc. In addition, most gene prediction programs search for ORFs and hence, miss out the non-coding RNA genes. Identification of some additional signals will help in improving existing gene prediction algorithms. The importance of DNA curvature in transcription suggests that identification of such curvature signals will prove to be an important input for gene prediction algorithms. With availability of large number of genome sequences, the curvature predication studies are gathering momentum.

Several *'in silico'* analyses have shown that the overall curvature level varies from one genome to another. Differences in sequence dependent curvature levels of prokaryotic (*E. coli* genomic fragments, *M. genitalium, H. influenzae, M. jannaschii*), viral (adenovirus 2, equine herpes virus 1), phage (M13, λ), eukaryotic (*S. cerevisiae and* human genomic fragments) and mitochondrial genomes have been reported (Gabrielian et al. 1997). Long, highly curved segments, similar to artificially designed curved DNA, are apparently absent from the genomes. It is found that prokaryotic and phage genomes appear to have a consistently higher frequency of curved DNA compared to the other genomes. In another example, the differences in overall curvature level of *Escherichia coli* and human genome sequence were revealed (Gabrielian et al. 1999-2000). The overall curvature was found to be lower in human genome sequences than in *Escherichia coli*. It was suggested that this difference in curvature properties of human and *Escherichia coli* sequences is due to the differences in their transcription schemes, such as presence of nucleosomes and different types of transcription factors. A detailed analysis of prokaryotic genomes (Bolshoy & Nevo 2000) showed that depending on curvature distribution, they could be divided into two groups. The first group indicated a substantial fraction of promoters characterized by intrinsic DNA curvature located within or upstream of the promoter region while this peculiar DNA curvature distribution was missing in the second group. Remarkably, all bacteria of the first group were mesophilic, whereas many prokaryotes of the second group were hyperthermophilic. It was therefore hypothesized that DNA curvature plays a biological role in gene

regulation in mesophilic as opposed to hyperthermophilic prokaryotes. Recent evidences, which show the role of curvature in thermoregulation in *Escherichia coli* O157:H7 (Prosseda et al. 2004) and *Shigella* (Yoon et al. 2004), support the above observation.

In another study, the regulatory regions of genes belonging to different orthologous groups (COG) were analyzed (Jauregui et al. 2003). COGs possessing a significant number of genes with curvature signals were further studied and conserved properties were found in several cases. This study provides evidence for the conservation of curvature signals in putative regulatory regions of several archaeal and eubacterial genomes. Other computational studies involving curvature calculation in ribosomal promoters as well as in RpoD promoters also suggest presence of conserved structure in these regions (Petersen et al. 2003).

Differences are not only observed across genomes from different organisms but also in different regions of a particular genome. Gabrielian et al. (Gabrielian et al. 1999-2000) divided the *Escherichia coli* genome into promoter, coding and noncoding regions and estimated the mean curvature in the different regions, as well as in random sequences. They found that mean curvature in promoter sequences is maximum while coding regions are least curved. Concentration of the curved DNA fragments in the intergenic regions rather than in the ORFs has also been observed in *Bacillus subtilis* genome (Tosato et al. 2003). In a similar effort (Pedersen et al. 2000), tools have been constructed to show the variations in structural properties along the genome sequence in the form of colour-coded wheels (http://www.cbs.dtu.dk/services/GenomeAtlas/). These "structural atlases" are useful for the discovery of interesting features at a gross level that may then be investigated in more depth using statistical methods. Interestingly, this analysis also revealed that practically in all the eubacterial and archaeal genomes investigated, there is a trend for promoter DNA being more curved, less flexible and less stable than DNA in coding regions and in intergenic DNA without promoters. In addition, around 20 regions which have extreme structural properties could be extracted and the authors suggested that these may function as topological domain boundaries for efficient organization of plectonemically supercoiled DNA.

There are other studies, which have correlated the amino acid composition and codon usage with the curvature of the genome (Jauregui et al. 1998, Jauregui et al. 2000). These results support the idea of a possible selection toward characteristic curvature of genomes. Recently, it has also been shown that DNA curvature plays a role in shaping the overall distribution of mutational hotspots along the p53 gene, which in turn play an important role in lung cancer (Lewis & Parry 2004).

Concluding Remarks and Future Directions

The perception that DNA is a monotonously uniform molecule has changed over the years. The significance of structural variations in DNA, such as curvature, in many biological functions has become very clear. But much more progress remains to be made, both on experimental as well as computational front, in characterizing the finer details. Though importance of curved DNA in the process of transcription is evident, it is still unclear at exactly which step does DNA curvature play a role? Based on a few examples, it is speculated that there may be a relation between presence of curved element and strength of a promoter. But additional experimental data is required to prove the universality of the proposal. More experiments, which can give direct proof, will also help in providing greater insight into the mechanisms involved.

The crystal and solution structure data on oligonucleotide sequences is still quite sparse and mostly redundant. The biased nature of crystal structure dataset is reflected in dinucleotide frequencies, with certain dinucleotides (AA/TT, CG) being highly frequent while others (AC/GT, AG/CT) being sparse. There is a need to determine the atomic structures of DNA fragments with novel and more variable sequences, which will help in developing better models for more reliable structure predictions. On the other hand, there is considerable structural data being accumulated in case of DNA-protein complexes. Sequence dependent DNA bending in protein-DNA complexes has been analyzed previously (Dickerson & Chiu 1997) but since then there have been many new additions. To gain a better insight into protein induced DNA bending it is necessary to critically analyze these DNA-protein complexes. This will also help in distinguishing intrinsic bending from bendability or flexibility of DNA.

Experimental studies on DNA curvature in genomic DNA have indicated its importance in gene regulation but such additional inputs on the exact nature and basis of this feature of DNA structure will help to make it an integral component of gene prediction algorithms.

Acknowledgements

The authors wish to thank the Department of Biotechnology, India for financial support to this work.

References

Abagyan R A, Mironov V N, Chernov B K, Chuprina V P and Ulyanov A V 1990 Electrophoretic behavior of $d(GGAAAAAAGG)_n$, $d(CCAAAAAACC)_n$, and $(CCAAAAAAGG)_n$ and implications for a DNA bending model; *Nucleic Acids Res.* **18** 989-992

Agrawal G K, Asayama M and Shirai M 2003 Two distinct curved DNAs upstream of the light-responsive psbA gene in a cyanobacterium; *Biosci. Biotechnol. Biochem.* **67** 1817-1821

Angermayr M, Oechsner U, Gregor K, Schroth G P and Bandlow W 2002 Transcription initiation in vivo without classical transactivators: DNA kinks flanking the core promoter of the housekeeping yeast adenylate kinase gene, AKY2, position nucleosomes and constitutively activate transcription; *Nucleic Acids Res.* **30** 4199-4207

Asayama M, Hayasaka Y, Kabasawa M, Shirai M and Ohyama A 1999 An intrinsic DNA curvature found in the cyanobacterium Microcystis aeruginosa K-81 affects the promoter activity of rpoD1 encoding a principal sigma factor; *J. Biochem.* (Tokyo) **125** 460-468

Atlung T, Sund S, Olesen K and Brondsted L 1996 The histone-like protein H-NS acts as a transcriptional repressor for expression of the anaerobic and growth phase activator AppY of Escherichia coli; *J. Bacteriol.* **178** 3418-3425

Bageshwar U K, Raina R, Choudhury N R and Das H K 1998 Analysis of upstream activation of the vnfH promoter of Azotobacter vinelandii; *Can. J. Microbiol.* **44** 405-415

Balas D, Fernandez-Moreira E and De La Campa A G 1998 Molecular characterization of the gene encoding the DNA gyrase A subunit of Streptococcus pneumoniae; *J. Bacteriol.* **180** 2854-2861

Bansal M 1996 Structural variations observed in DNA crystal structures and their implications for protein-DNA interaction; in Biological Structure and Dynamics, Proceedings of the Ninth Conversation.(Vol. I) pp 121-134 eds R H Sarma and M H Sarma (New York: Adenine Press)

_____ 2003 DNA structure: revisiting the Watson-Crick double helix; *Curr. Sci.* **85** 1556-1563

_____, Bhattacharya D and Mohanty D 1991 DNA bending: A natural consequence of base sequence dependent variability; in *Molecular Conformation and Biological Interactions* pp 347-362 eds P Balaram and S Ramaseshan (Bangalore: Indian Academy of Sciences)

_____, _____and Ravi B 1995 NUPARM and NUCGEN: software for analysis and generation of sequence dependent nucleic acid structures; *Comput. Appl. Biosci.* **11** 281-287

Bash R C, Vargason J M, Cornejo S, Ho P S and Lohr D 2001 Intrinsically bent DNA in the promoter regions of the yeast GAL1-10 and GAL80 genes; *J. Biol. Chem.* **276** 861-866

Bhattacharya D and Bansal M 1988 A general procedure for generation of curved DNA molecules; *J. Biomol. Struct. Dyn.* **6** 93-104

Bolshoy A and Nevo E 2000 Ecologic genomics of DNA: upstream bending in prokaryotic promoters; *Genome Res.* **10** 1185-1193

Bolshoy A, McNamara P, Harrington R E and Trifonov E N 1991 Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles; *Proc. Natl. Acad. Sci.* USA **88** 2312-2316

Bossi L and Smith D M 1984 Conformational change in the DNA associated with an unusual promoter mutation in a tRNA operon of Salmonella; *Cell* **39** 643-652

Bracco L, Kotlarz D, Kolb A, Diekmann S and Buc H 1989 Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*; *EMBO J.* **8** 4289-4296

Brukner I, Jurukovski V, Konstantinovic M and Savic A 1991 Curved DNA without AA/TT dinucleotide step; *Nucleic Acids Res.* **19** 3549-3551

_____, Sanchez R, Suck D and Pongor S 1995 Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides; *EMBO J.* **14** 1812-1818

Caddle M S, Dailey L and Heintz N H 1990 RIP60, a mammalian origin-binding protein, enhances DNA bending near the dihydrofolate reductase origin of replication; *Mol. Cell. Biol.* **10** 6236-6243

Chandrasekaran R and Arnott S 1989 The structures of DNA and RNA helices in oriented fibers; in *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology* (Vol. VII/1b) pp 31-170 ed. W Saenger (Berlin: Springer-Verlag).

Cheema A K, Choudhury N R and Das H K 1999 A- and T-tract-mediated intrinsic curvature in native DNA between the binding site of the upstream activator NtrC and the nifLA promoter of Klebsiella pneumoniae facilitates transcription; *J. Bacteriol.* **181** 5296-5302

Chen Q and Kadner R J 2000 Effect of altered spacing between uhpT promoter elements on transcription activation; *J. Bacteriol.* **182** 4430-4436

Crawford J L, Kolpak F J, Wang A H, Quigley G J, van Boom J H, van der Marel G and Rich A 1980 The tetramer d(CpGpCpG) crystallizes as a left-handed double helix; *Proc. Natl. Acad. Sci.* USA **77** 4016-4020

Crick F H and Klug A 1975 Kinky helix; *Nature* **255** 530-533

Dame R T, Wyman C and Goosen N 2001 Structural basis for preferential binding of H-NS to curved DNA; Biochimie **83** 231-234

Cognet J A, Pakleza C, Cherny D, Delain E and Cam E L 1999 Static curvature and flexibility measurements of DNA with microscopy. A simple renormalization method, its assessment by experiment and simulation. *J. Mol. Biol.* 1999 **285** 997-1009

Da Rocha P S and Bertrand H 1995 Structure and comparative analysis of the rDNA intergenic spacer of Brassica rapa. Implications for the function and evolution of the Cruciferae spacer; *Eur. J. Biochem.* **229** 550-557

Davies D R and Baldwin R L 1963 X-ray studies on two synthetic DNA copolymers; *J. Mol. Biol.* **6** 251-255

Delabre M L, Pasero P, Marilley M and Bougis P E 1995 Promoter structure and intron-exon organization of a scorpion alpha-toxin gene; *Biochemistry* **34** 6729-6736

De Santis P, Palleschi A, Savino M and Scipioni A. 1988 A theoretical model of DNA curvature; *Biophys. Chem.* **32** 305-317

Dethiollaz S, Eichenberger P and Geiselmann J 1996 Influence of DNA geometry on transcriptional activation in Escherichia coli; *EMBO J.* **15** 5449-5458

Dickerson R E and Chiu T K 1997 Helix bending as a factor in protein/DNA recognition; *Biopolymers* **44** 361-403

Dickerson R E, Goodsell D S and Neidle S 1994 "...the tyranny of the lattice..."; *Proc. Natl. Acad. Sci. USA* **91** 3579-3583

Dickerson R E, Bansal M, Calladine C R, Diekmann S, Hunter W N, Kennard O, Kitzing E V, Lavery R, Nelson H C M, Olson W K, Saenger W, Shakked Z, Skelnar H, Soumpasis D M, Tung C S, Wang A H J and Zhurkin V B 1989 Definations and nomenclature of nucleic acid structure parameters; *J. Mol. Biol.* **205** 787-791

Dlakic M and Harrington R E 1998 DIAMOD: display and modeling of DNA bending; *Bioinformatics* **14** 326-331

Doran K S, Konieczny I and Helinski D R 1998 Replication origin of the broad host range plasmid RK2. Positioning of various motifs is critical for initiation of replication; *J. Biol. Chem.* **273** 8447-8453

el Hassan M A and Calladine C R 1995 The assessment of the geometry of dinucleotide steps in double-helical DNA; a new local calculation scheme; *J. Mol. Biol.* **251** 648-664

Engelhorn M and Geiselmann J 1998 Maximal transcriptional activation by the IHF protein of Escherichia coli depends on optimal DNA bending by the activator; *Mol. Microbiol.* **30** 431-441

Fiorini A, Basso LR Jr, Paco-Larson M L and Fernandez M A 2001 Mapping of intrinsic bent DNA sites in the upstream region of DNA puff BhC4-1 amplified gene; *J. Cell Biochem.* **83** 1-13

Franklin R E and Gosling R G 1953a Molecular configuration in sodium thymonucleate; *Nature* **171** 740-741

_____and Gosling R G 1953b The structure of sodium thymonucleate fibers. I. The influence of water content; *Acta. Crystallogr.* **758** 16-17

Frey P M, Scharer-Hernandez N G, Futterer J, Potrykus I and Puonti-Kaerlas J 2001 Simultaneous analysis of the bidirectional African cassava mosaic virus promoter activity using two different luciferase genes; *Virus Genes.* **22** 231-242

Fukuda Y 1999 Characterization of matrix attachment sites in the upstream region of a tobacco chitinase gene; *Plant Mol. Biol.* **39** 1051-1062

Gaal T, Rao L, Estrem ST, Yang J, Wartell R M and Gourse R L 1994 Localization of the intrinsically bent DNA region upstream of the E.coli rrnB P1 promoter; *Nucleic Acids Res.* **22** 2344-2350

Gabrielian A and Pongor S 1996 Correlation of intrinsic DNA curvature with DNA property periodicity; *FEBS Lett.* **393** 65-68

_____, Vlahovicek K and Pongor S 1997 Distribution of sequence-dependent curvature in genomic DNA sequences; *FEBS Lett.* **406** 69-74

Gabrielian A E, Landsman D and Bolshoy A 1999-2000 Curved DNA in promoter sequences; *Silico Biol.* **1** 183-196

Gadaleta G, D'Elia D, Saccone C and Pepe G 1995 A 67-kDa protein binding to the curved DNA in the main regulatory region of the rat mitochondrial genome; *Gene* **160** 229-234

Gartenberg M R and Crothers D M 1991 Synthetic DNA bending sequences increase the rate of in vitro transcription initiation at the Escherichia coli lac promoter; *J. Mol. Biol.* **219** 217-230

Ghosh A and Bansal M 2003 A glossary of DNA structures from A to Z; *Acta. Crystallogr. D* **59** 620-626

Goodsell D S and Dickerson R E 1994 Bending and curvature calculations in B-DNA; *Nucleic Acids Res.* **22** 5497-5503

Gorin A A, Zhurkin V B and Olson W K 1995 B-DNA twisting correlates with base-pair morphology; *J. Mol. Biol.* **247** 34-48

Goyard S and Bertin P 1997 Characterization of BpH3, an H-NS-like protein in Bordetella pertussis; *Mol. Microbiol.* **24** 815-823

Hagerman P J 1986 Sequence-directed curvature of DNA; *Nature* **321** 449-450

_____1990 Sequence-directed curvature of DNA; *Annu. Rev. Biochem.* **59** 755-781

Haran T E, Kahn J D and Crothers D M 1994 Sequence elements responsible for DNA curvature; *J. Mol. Biol.* **244** 135-143

_____, Cohen I, Spasic A, Yang K and Mohanty U 2004 Characteristics of migration patterns of DNA oligomers in gels and the relationship to the question of intrinsic DNA bending; *J. Am. Chem. Soc.* **126** 2372-2377

Harp J M, Hanson B L, Timm D E and Bunick G J 2000 Asymmetries in the nucleosome core particle at 2.5 A resolution; *Acta. Crystallogr. D* 2000 **56** 1513-1534

Henriquez V, Milisavljevic V, Kahn J D and Gennaro M L 1993 Sequence and structure of cmp, the replication enhancer of the Staphylococcus aureus plasmid pT181; *Gene* **134** 93-98

Hertz G Z, Young M R and Mertz J E 1987 The A+T-rich sequence of the simian virus 40 origin is essential for replication and is involved in bending of the viral DNA; *J. Virol.* **61** 2322-2325

Hommais F, Krin E, Laurent-Winter C, Soutourina O, Malpertuy A, Le Caer J P, Danchin A and Bertin P 2001 Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS; *Mol. Microbiol.* **40** 20-36

Hsieh C H, Wu M and Yang J M 1991 The sequence-directed bent DNA detected in the replication origin of Chlamydomonas reinhardtii chloroplast DNA is important for the replication function; *Mol. Gen. Genet.* **225** 25-32

Hsu L M, Giannini J K, Leung T W and Crosthwaite J C 1991 Upstream sequence activation of Escherichia coli argT promoter in vivo and *in vitro*; *Biochemistry* **30** 813-822

Huang K J, Schieberl J L and Igo M M 1994 A distant upstream site involved in the negative regulation of the Escherichia coli ompF gene; *J. Bacteriol.* **176** 1309-1315

Humphrey W, Dalke A and Schulten K 1996 VMD - Visual Molecular Dynamics; *J. Molec. Graphics.* **14** 33-38

Jan J, Valle F, Bolivar F and Merino E 2000 Characterization of the 5' subtilisin (aprE) regulatory region from Bacillus subtilis; *FEMS Microbiol. Lett.* **183** 9-14

Jauregui R, Bolivar F and Merino E 2000 Relationship between whole proteome aminoacid composition and static DNA curvature; *Microb. Comp. Genomics.* **5** 7-15

_____, O'Reilly F, Bolivar F and Merino E 1998 Relationship between codon usage and sequence-dependent curvature of genomes; *Microb. Comp. Genomics.* **3** 243-253

Jauregui R, Abreu-Goodger C, Moreno-Hagelsieb G, Collado-Vides J and Merino E 2003 Conservation of DNA curvature signals in regulatory regions of prokaryotic genes; *Nucleic Acids Res.* **31** 6770-6777

Jordi B J, Fielder A E, Burns C M, Hinton J C, Dover N, Ussery D W and Higgins C F 1997 DNA binding is not sufficient for H-NS-mediated repression of proU expression; *J. Biol. Chem.* **272** 12083-12090

Kaji M, Matsushita O, Tamai E, Miyata S, Taniguchi Y, Shimamoto S, Katayama S, Morita S and Okabe A 2003 A novel type of DNA curvature present in a Clostridium prfringens ferredoxin gene: characterization and role in gene expression; *Microbiology* **149** 3083-3091

Kanhere A and Bansal M 2003 An assessment of three dinucleotide parameters to predict DNA curvature by quantitative comparison with experimental data; *Nucleic Acids Res.* **31** 2647-2658

Kelley W L and Bastia D 1991 Conformational changes induced by integration host factor at origin gamma of R6K and copy number control; *J. Biol. Chem.* **266** 15924-15937

Kneidl C, Dinkl E and Grummt F 1995 An intrinsically bent region upstream of the transcription start site of the rRNA genes of Arabidopsis thaliana interacts with an HMG-related protein; *Plant Mol. Biol.* **27** 705-713

Koo H S and Crothers D M 1988 Calibration of DNA curvature and a unified description of sequence-directed bending; *Proc. Natl. Acad. Sci. USA* **85** 1763-1767

Kraevskii V A, Mikhailov V S and Razin S V 1992 Stably bent and low-melting DNA sequences at the origin of replication of chicken alpha-globulin gene domains; *Mol. Biol. (Mosk.)* **26** 1011-1021

Kues U and Stahl U 1992 The replication origin of the Methylomonas clara plasmid pBE-2; *DNA Seq.* **3** 153-165

Lavery R 1988 Junctions and bends in nucleic acids: a new theoretical modeling approach in *Structure and Expression*, (Vol. 3: *DNA Bending and Curvature*). pp. 191-211 eds W K Olson, R H Sarma, M H Sarma and M Sundaralingam (New York: Adenine Press)

Leslie A G, Arnott S, Chandrasekaran R and Ratliff R L 1980 Polymorphism of DNA double helices; *J. Mol. Biol.* **143** 49-72

Levene S D and Crothers D M 1983 A computer graphics study of sequence-directed bending in DNA; *J. Biomol. Struct. Dyn.* **1** 429-435

Lewis P D and Parry J M 2004 In silico p53 mutation hotspots in lung cancer; Carcinogenesis (in press)

Li S J and Cronan J E Jr. 1993 Growth rate regulation of Escherichia coli acetyl coenzyme A carboxylase, which catalyzes the first committed step of lipid biosynthesis; *J. Bacteriol.* **175** 332-340

Li X M, Onishi Y, Kuwabara K, Rho J Y, Wada-Kiyama Y, Sakuma Y and Kiyama R 2002 Ligand-dependent transcriptional enhancement by DNA curvature between two half motifs of the estrogen response element in the human estrogen receptor alpha gene; *Gene* **294** 279-290

Lim F L, Hayes A, West A G, Pic-Taylor A, Darieva Z, Morgan B A, Oliver S G and Sharrocks A D 2003 Mcm1p-induced DNA bending regulates the formation of ternary transcription factor complexes; *Mol. Cell. Biol.* **23** 450-461

Lin S H and Lee J C 2003 Determinants of DNA bending in the DNA-cyclic AMP receptor protein complexes in Escherichia coli; *Biochemistry* **42** 4809-4818

Linial M and Shlomai J 1988 A unique endonuclease from Crithidia fasciculata which recognizes a bend in the DNA helix. Specificity of the cleavage reaction; *J. Biol. Chem.* **263** 290-297

Lu X J and Olson W K 2003 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures; *Nucleic Acids Res.* **31** 5108-5121

Lu Y, Weers B D and Stellwagen N C 2003 Analysis of the intrinsic bend in the M13 origin of replication by atomic force microscopy; *Biophys. J.* **85** 409-415

MacDonald D, Herbert K, Zhang X, Pologruto T, Lu P and Polgruto T 2001 Solution structure of an A-tract DNA bend; *J. Mol. Biol.* **306** 1081-1098

Malke H, Steiner K, Gase K and Frank C 2000 Expression and regulation of the streptokinase gene; *Methods* **21** 111-124

Marini J C, Levene S D, Crothers D M and Englund P T 1982 Bent helical structure in kinetoplast DNA; *Proc. Natl. Acad. Sci. USA* **79** 7664-7668

_____, Effron P N, Goodman T C, Singleton C K, Wells R D, Wartell R M and Englund P T 1984 Physical characterization of a kinetoplast DNA fragment with unusual properties; *J. Biol. Chem.* **259** 8974-8979

Marvin D A, Spencer M, Wilkins M H and Hamilton L D 1961 The molecular configuration of deoxyribonucleic acid. III. X-ray diffraction study of the C form of the lithium salt; *J. Mol. Biol.* **3** 547-565

McAllister C F and Achberger E C 1989 Rotational orientation of upstream curved DNA affects promoter function in Bacillus subtilis; *J. Biol. Chem.* **264** 10451-10456

Milton D L, Casper M L and Gesteland R F 1990a Saturation mutagenesis of a DNA region of bend. Base steps other than ApA influence the bend; *J. Mol. Biol.* **213** 135-140

Milton D L, Casper M L, Wills N M and Gesteland R F 1990b Guanine tracts enhance sequence directed DNA bends; *Nucleic Acids Res.* **18** 817-820

Mondragon A and Harrison S C 1991 The phage 434 Cro/OR1 complex at 2.5 Å resolution; *J. Mol. Biol.* **219** 321-334

Nachaliel N, Melnick J, Gafny R and Glaser G 1989 Ribosome associated protein(s) specifically bind(s) to the upstream activator sequence of the E. coli rrnA P1 promoter; *Nucleic Acids Res.* **17** 9811-9822

Nagaich A K, Bhattacharyya D, Brahmachari S K and Bansal M 1994 CA/TG sequence at the 5' end of oligo(A)-tracts strongly modulates DNA curvature; *J. Biol. Chem.* **269** 7824-7833

Nishikawa J, Amano M, Fukue Y, Tanaka S, Kishi H, Hirota Y, Yoda K and Ohyama T 2003 Left-handedly curved DNA regulates accessibility to *cis*-DNA elements in chromatin; *Nucleic Acids Res.* **31** 6651-6662

Ohyama T 1996 Bent DNA in the human adenovirus type 2 E1A enhancer is an architectural element for transcription stimulation; J. Biol. Chem. **271** 27823-27828

_____ and Hirota Y 1993 A possible function of DNA curvature in transcription; *Nucleic Acids Symp. Ser.* **29** 153-154

Olson W K, Bansal M, Burley S K, Dickerson R E, Gerstein M, Harvey S C, Heinemann U, Lu X J, Neidle S, Shakked Z, Sklenar H, Suzuki M, Tung C S, Westhof E, Wolberger C and Berman H M 2001 A standard reference frame for the description of nucleic acid base-pair geometry; *J. Mol. Biol.* **313** 229-237

Onishi Y and Kiyama R 2001 Enhancer activity of HS2 of the human beta-LCR is modulated by distance from the key nucleosome; *Nucleic Acids Res.* **29** 3448-3457

Ostrovsky de Spicer P, O'Brien K and Maloy S 1991 Regulation of proline utilization in *Salmonella typhimurium*: a membrane-associated dehydrogenase binds DNA *in vitro*; *J. Bacteriol.* **173** 211-219

Pedersen A G, Jensen L J, Brunak S, Staerfeldt H H and Ussery D W 2000 A DNA structural atlas for *Escherichia coli*; *J. Mol. Biol.* **299** 907-930

Perez-Martin J and Espinosa M 1994 Correlation between DNA bending and transcriptional activation at a plasmid promoter; *J. Mol. Biol.* **241** 7-17

_____and de Lorenzo V 1997 Clues and consequences of DNA bending in transcription; *Annu. Rev. Microbiol.* **51** 593-628

_____, del Solar G H, de la Campa A G and Espinosa M 1988 Three regions in the DNA of plasmid pLS1 show sequence-directed static bending; *Nucleic Acids Res.* **16** 9113-9126

Petersen L, Larsen T S, Ussery D W, On S L and Krogh A 2003 RpoD promoters in Campylobacter jejuni exhibit a strong periodic signal instead of a -35 box; *J. Mol. Biol.* **326** 1361-1372

Polaczek P, Kwan K, Liberies D A and Campbell J L 1997 Role of architectural elements in combinatorial regulation of initiation of DNA replication in *Escherichia coli*; *Mol. Microbiol.* **26** 261-275

Porter M E 2002 Positive and negative effects of deletions and mutations within the 5' flanking sequences of *Plasmodium falciparum* DNA polymerase delta; *Mol. Biochem. Parasitol.* **122** 9-19

Prosseda G, Falconi M, Giangrossi M, Gualerzi C O, Micheli G and Colonna B 2004 The virF promoter in Shigella: more than just a curved DNA stretch; *Mol. Microbiol.* **51** 523-537

Pyles E A and Lee J C 1998 *Escherichia coli* cAMP receptor protein-DNA complexes. 2. Structural asymmetry of DNA bending; *Biochemistry* **37** 5201-5210

Richet E 1996 On the role of the multiple regulatory elements involved in the activation of the *Escherichia coli* malEp promoter; *J. Mol. Biol.* **264** 852-862

Rivetti C, Walker C and Bustamante C 1998 Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility; *J. Mol. Biol.* **280** 41-59

Robbe-Saule V, Schaeffer F, Kowarz L and Norel F 1997 Relationships between H-NS, sigma S, SpvR and growth phase in the control of spvR, the regulatory gene of the Salmonella plasmid virulence operon; *Mol. Gen. Genet.* **256** 333-347

Satchwell S C, Drew H R and Travers A A 1986 Sequence periodicities in chicken nucleosome core DNA; *J. Mol. Biol.* **191** 659-675

Schulz A, Langowski J and Rippe K 2000 The effect of the DNA conformation on the rate of NtrC activated transcription of Escherichia coli RNA polymerase sigma-54 holoenzyme; *J. Mol. Biol.* **300** 709-725

Schultz S C, Shields G C and Steitz T A. 1991 Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees; Science **253** 1001-1007

Scipioni A, Anselmi C, Zuccheri G, Samori B and De Santis P 2002 Sequence-dependent DNA curvature and flexibility from scanning force microscopy images; *Biophys. J.* **83** 2408-2418

Scott S, Busby S and Beacham I 1995 Transcriptional coactivation at the ansB promoters: involvement of the activating regions of CRP and FNR when bound in tandem; *Mol. Microbiol.* **18** 521-531

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

Seong G H, Kobatake E, Miura K, Nakazawa A and Aizawa M 2002 Direct atomic force microscopy visualization of integration host factor-induced DNA bending structure of the promoter regulatory region on the Pseudomonas TOL plasmid; *Biochem. Biophys. Res. Commun.* **291** 361-366

Shpigelman E S, Trifonov E N and Bolshoy A 1993 CURVATURE: software for the analysis of curved DNA; *Comput. Appl. Biosci.* **9** 435-440

Shrader T E and Crothers D M 1989 Artificial nucleosome positioning sequences; *Proc. Natl. Acad. Sci. USA* **86** 7418-7422

Shrader T E and Crothers D M1990 Effects of DNA sequence and histone-histone interactions on nucleosome placement; *J. Mol. Biol.* **216** 69-84

Snyder M, Buchman A R and Davis R W 1986 Bent DNA at a yeast autonomously replicating sequence; *Nature* **324** 87-89

Stenzel T T, Patel P and Bastia D 1987 The integration host factor of Escherichia coli binds to bent DNA at the origin of replication of the plasmid pSC101; *Cell* **49** 709-717

Suarez-Lopez P, Martinez-Salas E, Hernandez P and Gutierrez C 1995 Bent DNA in the large intergenic region of wheat dwarf geminivirus; *Virology* **208** 303-311

Tan R K, Prabhakaran M, Tung C S and Harvey S C 1988 AUGUR: a program to predict, display and analyze the tertiary structure of B-DNA; *Comput. Appl. Biosci.* **4** 147-151

Thastrom A, Bingham L M and Widom J 2004 Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning; *J. Mol. Biol.* **338** 695-709

Timchenko N A, Luniak V V and Luniak A N 1994 Segments initiating replication of the dihydrofolate reductase domain and the DNA fragment DARC146 interact with the same nuclear proteins; *Mol. Biol. (Mosk.)* **28** 865-874

Timchenko T, Bailone A and Devoret R 1996 Btcd, a mouse protein that binds to curved DNA, can substitute in Escherichia coli for H-NS, a bacterial nucleoid protein; *EMBO J.* **15** 3986-3992

Tosato V, Gjuracic K, Vlahovicek K, Pongor S, Danchin A and Bruschi C V 2003 The DNA secondary structure of the Bacillus subtilis genome; *FEMS Microbiol. Lett.* **218** 23-30

Trifonov E N and Sussman J L 1980 The pitch of chromatin DNA is reflected in its nucleotide sequence. *Proc. Natl. Acad. Sci. USA* **77** 3816-3820

Troup B, Jahn M, Hungerer C and Jahn D 1994 Isolation of the hemF operon containing the gene for the Escherichia coli aerobic coproporphyrinogen III oxidase by in vivo complementation of a yeast HEM13 mutant; *J. Bacteriol.* **176** 673-680

Tung C S, Carter E S 2nd 1994 Nucleic acid modeling tool (NAMOT): an interactive graphic tool for modeling nucleic acid structures; *Comput. Appl. Biosci.* **10** 427-433

Ulyanov N B and James T L 1995 Statistical analysis of DNA duplex structural features; *Methods Enzymol.* **261** 90-120

Viswamitra M A, Kennard O, Jones P G, Sheldrick G M, Salisbury S, Favello L and Shakked Z 1978 DNA double helical fragment at atomic resolution; *Nature* **273** 687-688

Vlahovicek K and Pongor S 2000 Model.it: building three dimensional DNA models from sequence data; *Bioinformatics* **16** 1044-1045

Wang A H, Quigley G J, Kolpak F J, Crawford J L, van Boom J H, van der Marel G and Rich A 1979 Molecular structure of a left-handed double helical DNA fragment at atomic resolution; *Nature* **282** 680-686

Watson J D and Crick F H 1953 Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid; *Nature* **171** 737-738

Welter C, Dooley S, Zang K D and Blin N 1989 DNA curvature in front of the human mitochondrial L-strand replication origin with specific protein binding; *Nucleic Acids Res.* **17** 6077-6086

Williams L D, Maher L J 3rd 2000 Electrostatic mechanisms of DNA deformation; *Annu. Rev. Biophys. Biomol. Struct.* **29** 497-521

Wing R, Drew H, Takano T, Broka C, Tanaka S, Itakura K and Dickerson R 1980 Crystal structure analysis of a complete turn of B-DNA; *Nature* **287** 755-758

Wu H M and Crothers D M 1984 The locus of sequence-directed and protein-induced DNA bending; *Nature* **308** 509-513

Yamashino T, Kakeda M, Ueguchi C and Mizuno T 1994 An analogue of the DnaJ molecular chaperone whose expression is controlled by sigmas during the stationary phase and phosphate starvation in *Escherichia coli; Mol. Microbiol.* **13** 475-483

Yoon J W, Minnich S A, Ahn J S, Park Y H, Paszczynski A and Hovde C J 2004 Thermoregulation of the Escherichia coli O157:H7 pO157 ecf operon and lipid A myristoyl transferase activity involves intrinsically curved DNA; *Mol. Microbiol.* **51** 419-435

Zahn K and Blattner F R 1985 Sequence-induced DNA curvature at the bacteriophage lambda origin of replication; *Nature* **317** 451-453

Zhang X and Schleif R 1998 Catabolite gene activator protein mutations affecting activity of the araBAD promoter; *J. Bacteriol.* **180** 195-200

Zuccheri G, Scipioni A, Cavaliere V, Gargiulo G, De Santis P, Samori B. 2001 Mapping the intrinsic curvature and flexibility along the DNA chain; *Proc. Natl. Acad. Sci. USA.* **98** 3074-3079