The EMBO Journal vol.14 no.8 pp.1812–1818, 1995

Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides

Ivan Brukner, Roberto Sánchez, Dietrich Suck[1] and Sándor Pongor[2]

International Centre for Genetic Engineering and Biotechnology (ICGEB), Area Science Park, Padriciano 99, 34012 Trieste, Italy and [1]Biological Structures and Biocomputing Division, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, 69012 Heidelberg, Germany

[2]Corresponding author

Communicated by A.Falaschi

Structural parameters characterizing the bending propensity of trinucleotides were deduced from DNase I digestion data using simple probabilistic models. In contrast to dinucleotide-based models of DNA bending and/or bendability, the trinucleotide parameters are in good agreement with X-ray crystallographic data on bent DNA. This improvement may be due to the fact that the trinucleotide model incorporates more sequence context information than do dinucleotide-based descriptions.

Key words: DNA curvature/DNA flexibility/DNA–protein interaction/DNase I/trinucleotides

Introduction

Sequence-dependent bending of DNA can be inherent in the DNA molecule in a stress-free state or it can be induced by ligands (Travers and Klug, 1990). DNA binding proteins generally bend DNA toward the major groove (imposing positive roll angles), irrespective of whether they bind to the major or to the minor groove. At present there is little structural data on DNA bending and/or bendability at the level of atomic resolution, due to the fact that these can be best obtained from crystallographic analysis.

Predictive models of sequence-dependent DNA bending are based, generally speaking, on assigning parameters to motifs within DNA. The current models use parameters which are assigned to the 10 independent dinucleotide steps. The parameters themselves are determined using various experimental approaches, such as: (i) gel-mobility analysis (Bolshoy *et al.*, 1991); (ii) a combination of gel-mobility data with X-ray crystallography (Calladine *et al.*, 1988) or energy minimization (De Santis *et al.*, 1990); (iii) interactions of DNA with other agents, such as nucleosomes (Drew and Travers, 1985; Calladine and Drew, 1986; Satchwell *et al.*, 1986), catabolic activator protein (Gartenberg and Crothers, 1988) or DNase I (Brukner *et al.*, 1990).

Dinucleotide models represent the simplest form of nearest neighbour descriptions. Extension of these models to tri- or tetranucleotides is considered to be an important improvement, since this may account for the well-known fact that dinucleotides are bent and/or bendable differently in different sequential contexts (Hunter, 1993; Poncin *et al.*, 1992). On the other hand, since there are 32 independent trinucleotides and 136 tetranucleotides, such an extension would need a much larger set of experimental data than the one currently available. In this paper we show, based on a large amount of experimental data, that bovine pancreatic deoxyribonuclease I (DNase I) digestion profiles can be used to obtain realistic DNA bending propensity parameters of trinucleotides.

DNase I is considered to be a good molecular probe of DNA bending and/or bendability, since all DNase I–DNA complexes solved so far show that productive binding of DNase I requires DNA to be bent (Lahm and Suck, 1991; Weston *et al.*, 1992). DNase I interacts with a 6 bp contact surface of the minor groove and bends the DNA molecule away from the enzyme, toward the major groove (positive roll) (Lahm and Suck, 1991; Weston *et al.*, 1992; Suck, 1994). Therefore, base sequences that are flexible or inherently bent towards the major groove should be more accessible to DNase I cleavage. A qualitative connection between DNase I digestion and DNA structural properties was recognized quite early (Lomonossoff *et al.*, 1984) and correlation with minor groove width (Drew and Travers, 1984), bending and/or bendability (Drew and Travers, 1985; Hogan *et al.*, 1989; Brukner *et al.*, 1990; Lahm and Suck, 1991; Weston *et al.*, 1992; Suck, 1994) and looping (Hochschild and Ptashne, 1986; Lavigne *et al.*, 1994) have been suggested by several authors. In this paper we try to extract quantitative and sequence-specific bending and/or bendability parameters for trinucleotides, directly from the DNase I digestion data.

Results

Description of the model

The model is based on the assumptions that: (i) DNase I interacts with the window of 6 nt around the cleaved bond; (ii) one single structural parameter $p(a)$ of the trinucleotides, constituting the enzyme–DNA contact surface, will influence the cutting rate (this is an obvious simplification, since local effects, such as specific residue contacts between the enzyme and the DNA molecule, are not considered); (iii) the bending propensity $p(a)$ of each trinucleotide contributes independently to the probability of DNase I cutting, P_w. The model thus assumes that the contribution of one element (trinucleotide) does not depend on any other element being present or absent in the window around the cut. So P_w for the 6 nt window can be written as the product of the n different and assumedly independent $p(a)_i$ probabilities:

$$P_w = \prod_{i=1}^{4} p(a)_i \qquad (1)$$

Table I. DNA bending and/or bendability parameters as revealed by DNase I: parameters of trinucleotide steps[a]

Trinucleotide step	No. of occurrences in dataset	DNase I-derived trinucleotide parameter (ln *p*)
AAT/ATT	89	−0.280
AAA/TTT	278	−0.274
CCA/TGG	45	−0.246
AAC/GTT	81	−0.205[b]
ACT/AGT	77	−0.183[b]
CCG/CGG	73	−0.136
ATC/GAT	112	−0.110
AAG/CTT	110	−0.081
CGC/GCG	84	−0.077
AGG/CCT	101	−0.057
GAA/TTC	117	−0.037
ACG/CGT	84	−0.033
ACC/GGT	87	−0.032
GAC/GTC	81	−0.013
CCC/GGG	141	−0.012
ACA/TGT	52	−0.006[b]
CGA/TCG	84	−0.003
GGA/TCC	71	0.013
CAA/TTG	74	0.015[b]
AGC/GCT	35	0.017
GTA/TAC	83	0.025
AGA/TCT	127	0.027
CTC/GAG	102	0.031[b]
CAC/GTG	55	0.040
TAA/TTA	99	0.068[b]
GCA/TGC	34	0.076
CTA/TAG	64	0.090
GCC/GGC	57	0.107
ATG/CAT	71	0.134[b]
CAG/CTG	61	0.175
ATA/TAT	80	0.182
TCA/TGA	127	0.194

[a]Parameters are given in arbitrary units. High values indicate bending towards the major groove. The parameters were determined using a system of equations given in equation (2); the linear correlation coefficient was 0.60 for a total of 709 data points. The reproducibility of the densitometric readings was characterized by, on average, a 10% error. In the parameter values, this variation results in an average error of 0.002 (SD, standard deviation of the mean). Those values having a higher SD value (between 0.003 and 0.004) are marked. The stability of the numerical model is shown by the fact that an estimated error of 20% would result in an average error of only 0.005 in the parameter values, as determined by computer simulations. The fact that there was no overall change in either the resulting parameter values or in the ranking order of the trinucleotides makes us believe that the triplet parameters presented are close to or identical with the absolute minimum of the model (Press *et al.*, 1992).
[b]Parameters which were found to be the least stable with respect to variations in the data (SD between 0.003 and 0.004).

Equating P_w with the experimentally determined frequencies of cleavage, F_w, leads to a linear system of equations of the form

$$\ln F_w = \sum_{i=1}^{4} \ln p(a)_i \qquad (2)$$

which can then be solved by the least squares method. In this work, we use a new set of 709 DNase I digestion data points for pUC18 and T7 A1 promoter DNA (determined in triplicate) in order to derive trinucleotide parameters. The densitometric readings (which are proportionate to observed DNase I cutting frequencies) were then substituted into equation (2), which was subsequently solved to give ln $p(a)$ parameters of trinucleotides (Table I).

Until now we have not mentioned the possible physical meaning of the *p* parameters. As DNase I can, in principle, cut at sites that are either *a priori* bent toward the major groove or flexible, the *p* parameters may incorporate these two properties, without necessarily distinguishing between the static and dynamic components.

Comparison with data obtained by X-ray crystallography

The trinucleotide bending and/or bendability parameters were compared with experimental X-ray crystallography data obtained on bent DNA, using parameter versus sequence plots (Figure 2). The examples include inherently curved DNA structures (Figure 2I–K) as well as protein–DNA complexes with bent DNA (Figure 2A–H). In both cases there is an obvious qualitative agreement between the published roll angles (or the direction of groove distortion) and the trinucleotide parameters; regions with major groove compression (positive roll angles) show high parameter values, while non-bent regions or motifs showing compression of the minor groves (negative roll angles) show low values. Only in one example, that of the DNA complex with the TATA box binding protein TBP (Kim and Burley, 1994; Figure 2C), is there an obvious disagreement between the predicted and experimental values. This might be due to the fact that DNA in the TBP–DNA complex is not only doubly kinked, but also substantially unwound. This is quite atypical compared with the other structures. These examples show that DNA bending propensity, extracted from DNase I digestion profiles, and distortions, seen by X-ray crystallography, are clearly correlated.

Evaluation of direction of curvature using helical phasing experiments

Oligonucleotides were designed in order to test the relative direction of curvature in GGC/GCC, CGG/CCG and CCA/TGG trinucleotide motifs. These motifs were chosen because the trinucleotide model predicts their bending propensities in a way that contradicts the current dinucleotide-based descriptions (see also Discussion). The full sequence context of the tested sequences was designed to follow the sequence-dependent bending propensity of trinucleotides presented in Table I. The principle of this technique is as follows: oligonucleotides containing curved elements which are spaced by a half-helical turn and bend DNA in the same direction will have higher mobility (μ−) than those oligonucleotides in which the spacing is a full helical turn (μ+) (Figure 3A). Using a reference motif, this method estimates the relative direction of the curvature of the elements which are part of the test sequence. If there is no intrinsic curvature in the test sequence, then μ− and μ+ will be independent of the phasing and equal.

All designed oligonucleotides contain A tracts (as a reference motif) which are differentially spaced from the tested elements (Figure 3A and B). One should note from Figure 3 that the outcome of this experiment ranks the curvature of the tested elements in the following manner: GGCCC versus CCGGG (Figure 3B, gc) and CCAAT versus CTCAC motifs (Figure 3B, ca). The μ−:μ+ ratio was significantly different from unity in both examples (Figure 3D, black and white bars), clearly showing that

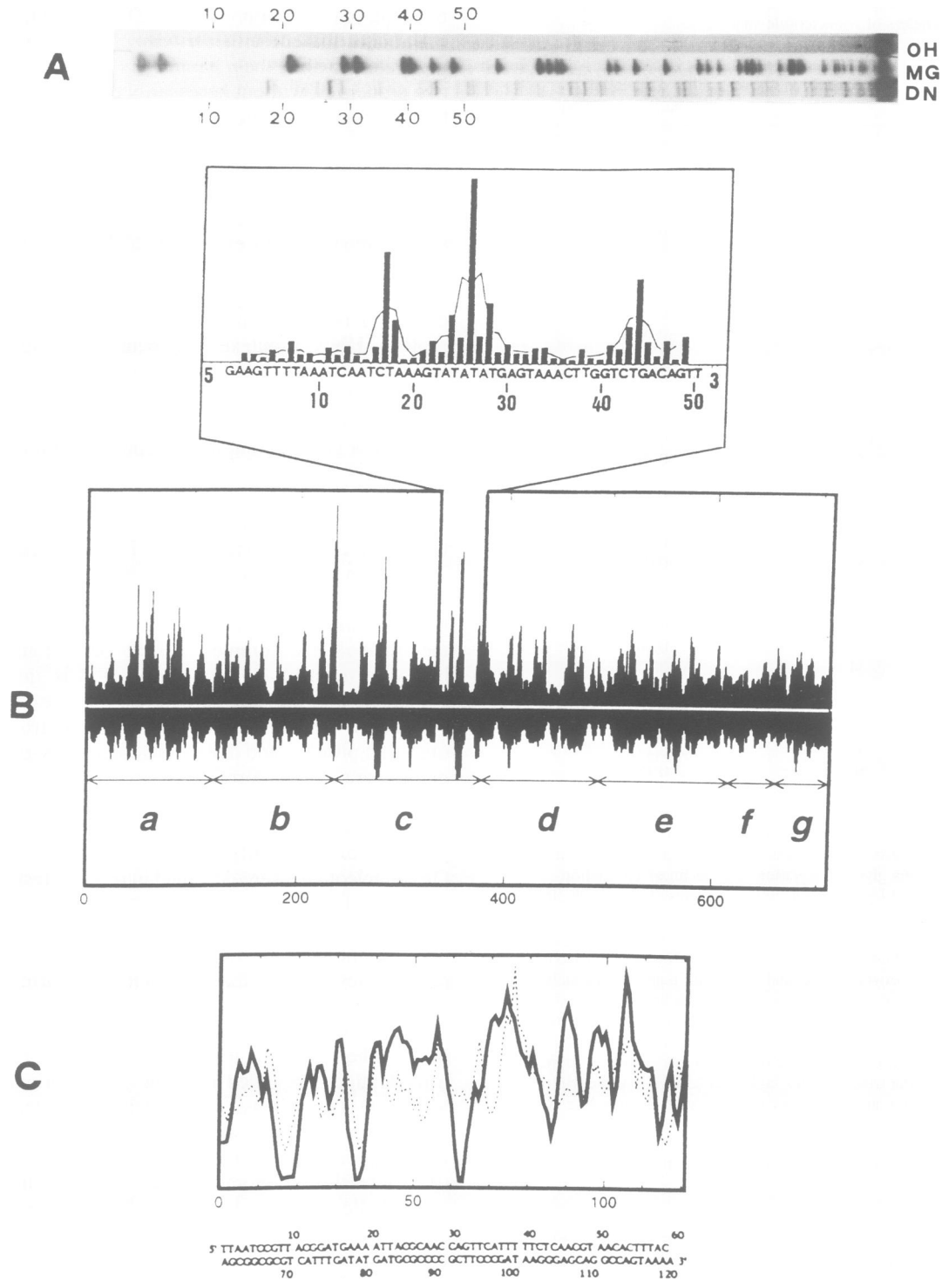

Fig. 1. (**A**) A 'typical' DNase I digestion pattern (line DN); hydroxyl radical (line OH) and Maxam–Gilbert (line MG) reactions were used to assign the cleaved bond to the sequence. Numbers refer to the sequence fragment shown in the magnification in (B). (**B**) Comparison of measured (upper) DNase I cutting frequencies with those predicted by the trinucleotide model (lower values). Origin of DNA, pUC18 restriction fragments; numbers denote site position in databank (Vector Bank 8): *a*, 2686–48; *b*, 1376–1529; *c*, 1529–1780; *d*, 2233–2534; *e*, 399–524; *f*, 245–399 and bacteriophage T7 A1 promoter DNA (EMBL databank); *g*, 80–150. Magnification: measured values (columns) and three bond average (continuous line) corresponding to the digestion profile shown in (A). The relative frequency of cleavage for each P–O3′ bond was determined, transformed to the three bond running average (full line) and presented as a function of the DNA sequence. The predicted values were calculated according to equation (2), using the parameters in Table I. (**C**) DNase I digestion data of Drew and Travers (1984), which were not included into the models (thick line), are presented together with the predictions based on the trinucleotide step (dotted line) model. The numbering corresponds to the DNA sequence written below.

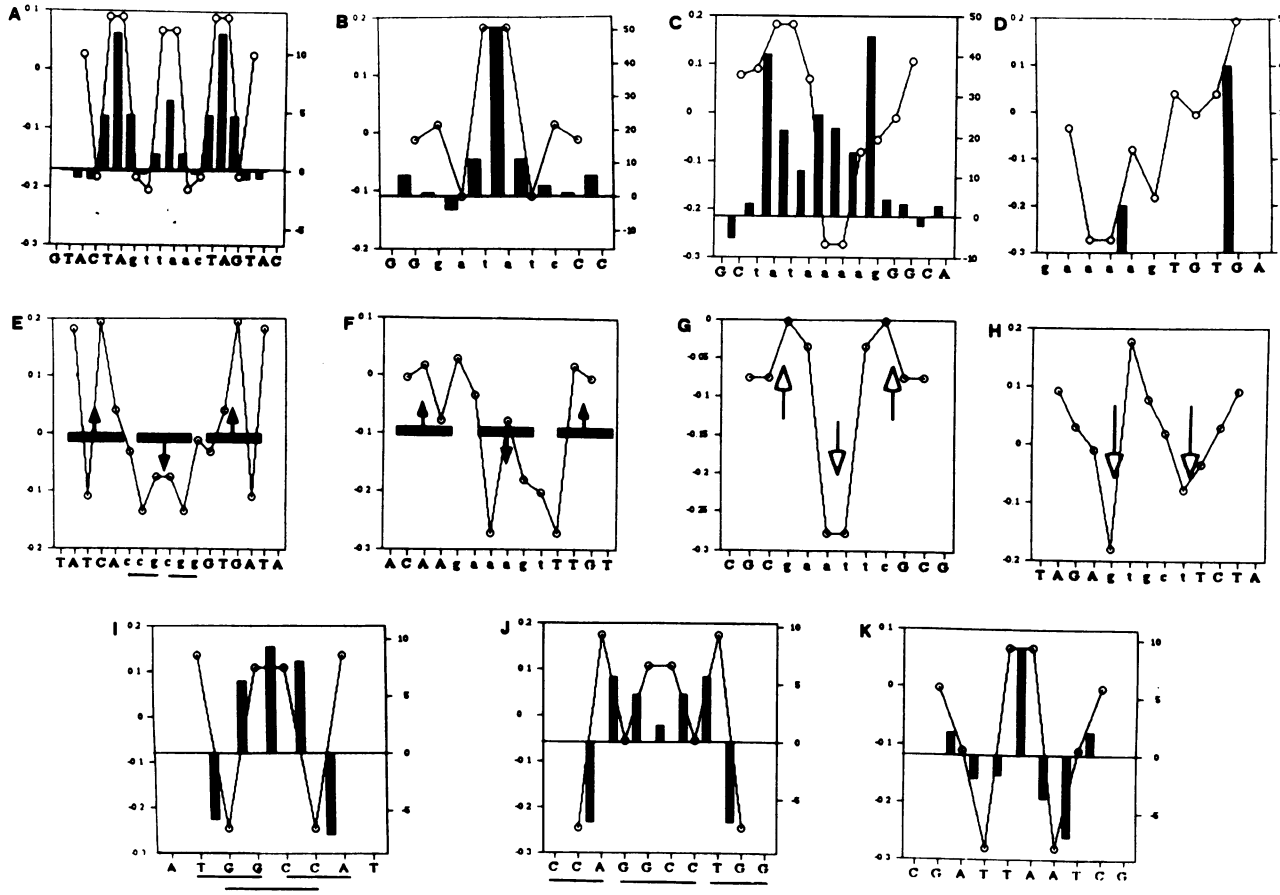

Fig. 2. Comparison of the predicted DNA bending and/or bendability with three-dimensional data. (A–H) Protein–DNA complexes with strong protein-induced bending: (**A**) *trp* repressor (Shakked *et al.*, 1994); (**B**) *Eco*RV–DNA complex (Winkler *et al.*, 1993); (**C**) TATA box binding protein TBP (Kim and Burley, 1994); (**D**) catabolic activator protein CAP–DNA complex (Schultz *et al.*, 1991); (**E**) λ Cro–operator complex (Brennan *et al.*, 1990); (**F**) the phage 434 O_R2–R1-69 complex (Shimon and Harrison, 1993); (**G**) *Eco*RI–DNA complex (McClarin *et al.*, 1986); (**H**) Arc repressor–operator complex (Raumann *et al.*, 1994). (I–K) DNA crystal structures with inherent bending: (**I**, **J**) GGCC-containing oligonucleotides (Grzeskowiak *et al.*, 1993); (**K**) a TA-containing oligonucleotide (Quintana *et al.*, 1992). *x*-axis, DNA sequence. Segments in which the minor groove faces the bound protein are written in lower case. *y*-axis, predicted bending and/or bendability values; trinucleotide values (continuous line). Roll angles, determined by X-ray crystallography, are shown by bars and are indicated on the right hand side. In cases where the roll angles were not given in the original papers, kinks (white arrows) or regions of groove compression (underlined black arrows) are indicated, as originally given by the authors.

the test sequence has intrinsically curved elements. Finally, deducing from the helical phasing experiment, we demonstrated that the major groove-directed curvature is higher for GGCCC than for CCGGG and lower for CCAAT then for CTCAC sequence elements. This exactly confirms our predictions based on bending and/or bendability revealed by DNase I, presented in Table I.

The experiments shown in Figure 3 were carried out in the presence of 10 mM Mg^{2+} ions. When Mg^{2+} was left out of the electrophoresis buffers, few or no detectable differences were found between the respective pairs (data not shown).

Discussion

What is the difference between this trinucleotide model and the current dinucleotide models? In the first place, the trinucleotide model includes more context information: instead of a single value (e.g. AA/TT) describing a particular dinucleotide, we now have several trinucleotide values in which the dinucleotide is present (i.e. AAA/TTT, CAA/TTG, GAA/TTC, TAA/TTA, AAC/GTT, AAG/CTT and AAT/ATT). Of the previous models, only the nucleosome packing data (of Calladine and Drew, 1986) incorporate some context dependence.

A second, more important, difference is that, in contrast to the dinucleotide models, the trinucleotide model is in reasonable agreement with X-ray crystallographic data (Figure 2), even though the X-ray data were not included in the model. It is worth noting that the qualitative agreement between the trinucleotide data and the experimental roll angles is reasonably good, even though the protein–DNA complexes used as examples represent a variety of different binding mechanisms. This is in keeping with the notion that bending propensity is an intrinsic property of DNA, which can be exploited by different DNA binding proteins.

Thirdly, there are significant differences in the prediction of individual motifs. (i) It is supposed by all dinucleotide models that CC and CA can be bent toward the major groove. CCA, however, frequently has negative roll angles in crystals (underlined in Figure 2I and J), which is correctly shown by the low trinucleotide parameter values. (ii) It is considered that GC can be bent towards the minor

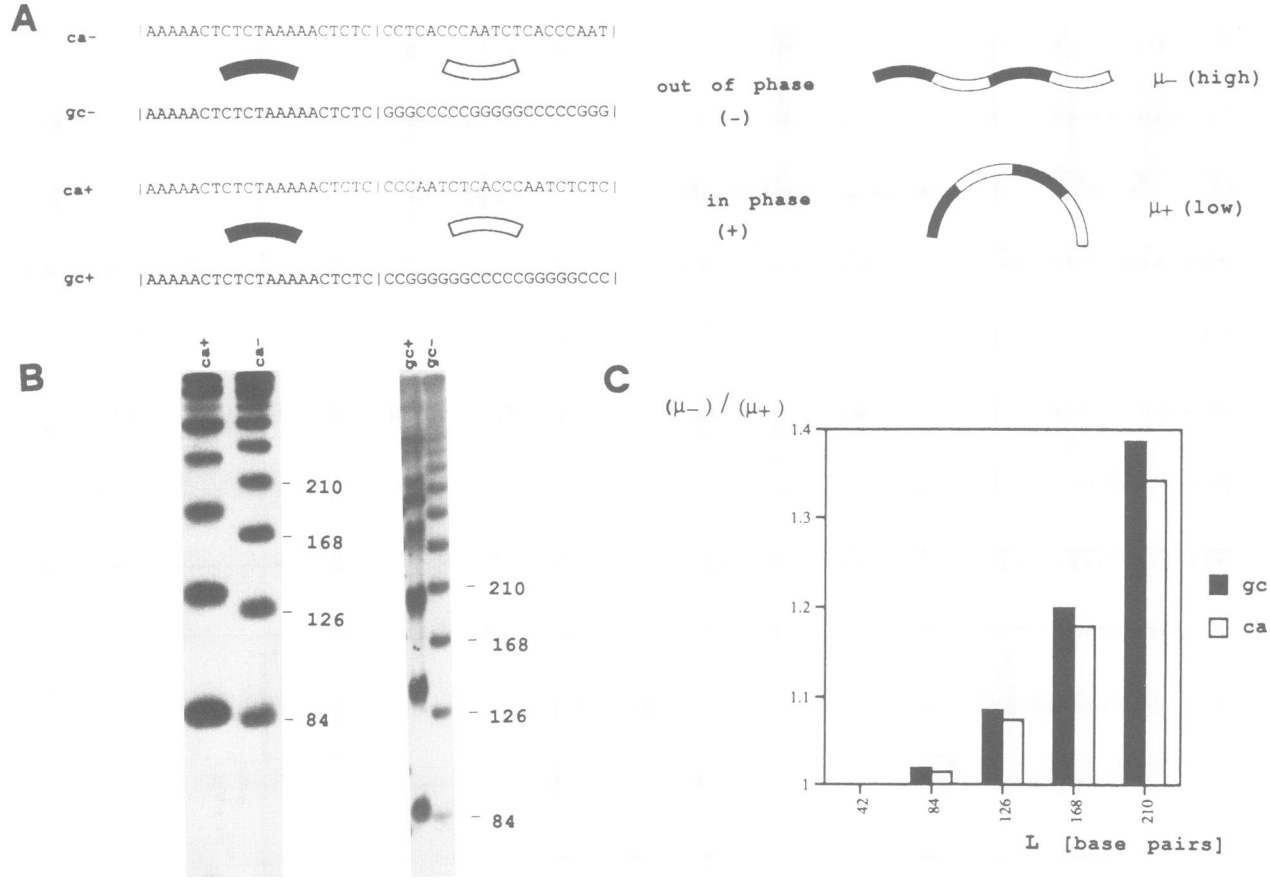

Fig. 3. Relative directions of curvature of tested sequence elements determined by helical phasing with A tracts. (**A**) Design of DNA molecules with 360° (in phase) and 180° (out of phase) angular distances between sequence elements to be tested (white arc) and A tract-containing DNA, used as reference elements (black arc). Note that the angular distance between test elements (180°) is the same in all molecules. (**B**) Dependence of gel mobility (μ) on the phasing between sequence elements of unknown direction of curvature. Lanes gc and ca are ligation multimers characterized by in phase (+) and out of phase (−) helical distance of tested elements from the A tracts. (**C**) The ratio between gel mobility μ−/μ+ of gc (black bars) and ca (white bars) multimers is plotted against linker length (L). Exerimental error in determination of μ is ±0.03. Macroscopic curvature is produced only if there is an appropriate helical phasing between curved sequence elements. This logic allows us to deduce that the major groove-directed curvature of GGGCCC is higher than CCCGGG and that CTCAC is higher than CCAAT.

groove by the dinucleotide models (except that of Calladine and Drew, 1986). In contrast, it is predicted that GGC is bent towards the major groove by the trinucleotide model, in agreement with X-ray crystallographic data (underlined in Figure 2I and J). (iii) The dinucleotide models predict that CG can be bent toward the major groove. CCG, on the other hand, has a low trinucleotide parameter, which is in agreement with the fact that CCG/CGG is found in a region of minor groove compression in the λ Cro−DNA complex (underlined in Figure 2E).

Perhaps the most intriguing of these differences are: (i) the opposite bending and/or bendability directions predicted for GCC and CGG elements respectively and (ii) the low bending propensity of the CA step in the context of the CCA motif. According to the purine/pyrimidine rule (Barber and Zhurkin, 1990) or the dinucleotide models of Bolshoy *et al.* (1991) and De Santis *et al.* (1990), bending/bendability features of CCA, GGC and CCG sequence elements show exactly the opposite trend from our trinucleotide model. We tested the curvature of GGCCC versus the CCGGG element and that of the CA step in the context of a CCA motif, using helical phasing experiments (Zinkel and Crothers, 1987) in the presence of Mg^{2+} ions (Diekmann, 1987; Shlyakhtenko *et al.*, 1990; Brukner *et al.*, 1994). DNA fragments which contain tested elements are placed at varying distances from the A tract (Figure 3). It was found that the major groove-directed curvature of GGC is higher than that of CCG and that the CCA trinucleotide has a bending propensity similar to that of the A tracts, i.e. low propensity for bending toward the major groove. Thus, both results are in firm agreement with the trinucleotide data.

The advantage of the approach used here lies in the fact that the experimental dataset (DNase I digestion data) can be expanded relatively easily. It has to be noted that the current dataset does not yet allow determination of the 136 tetranucleotide parameters. It should also be noted that the model is not meant for quantitative prediction of DNase I digestion profiles, since it does not consider local effects, such as specific enzyme−substrate interactions. In principle, local structural features recognized by DNase I can be captured by using sequence patterns. Recently, Herrera and Chaires (1994) determined such patterns, using a different and rather limited dataset. These patterns, however, show little correlation with our experimental data; only seven of the 20 preferred cleavage sites and none of the 20 most poorly cleaved sites found in our experimental dataset contain the patterns suggested by

Herrera and Chaires (1994). While it may be possible, in principle, to separate local and global effects, at present these influences are incorporated as a bias into the $p(a)$ parameters of the present model. In spite of these limitations, the trinucleotide model presented here seems to show, in a correct manner, the sequence-specific differences in DNA bending and/or bendability and offers a convenient way to predict the bending propensity of DNA using simple parameter versus sequence plots. We feel that the improvement is partly due to the trinucleotide-based representation and partly due to the larger dataset, since a previous dinucleotide model based on a smaller set of DNase I data (Brukner *et al.*, 1990) did not detect the differences which are now correctly shown by the present trinucleotide model.

The physical meaning of the trinucleotide parameters can be summarized as follows: for efficient cleavage, DNase I has to distort a particular sequence to a conformation that is optimal for cutting; the energy required for this purpose will be low for sequences which are (i) flexible (i.e. able to bend) or (ii) inherently bent to a conformation which is near to that required for cleavage. The trinucleotide parameters capture, at least in principle, these two properties at the same time, but do not in themselves enable us to distinguish between them. A discrimination between flexible and inherently curved sequence elements is possible only with additional experiments, like the helical phasing studies described here. In the future it will be possible to derive more accurate models, such as those based on tetranucleotides or on applying 'position-dependent' weighting schemes within the window. This will require, however, a larger dataset than the one currently available.

Materials and methods

DNase I digestion

pUC18 (Boehringer Mannheim, Mannheim, Germany) fragments (see Figure 1) and phage T7 A1 promoter DNA (kindly provided by Dr H.Heumann, Max-Planck Institute of Biochemistry, Martinsried, Germany) were dephosphorylated and end-labelled as described (Drew and Travers, 1984). After end-labelling, restriction cleavage was performed to remove the labelled nucleotide from one end of the DNA, which was purified on 8% polyacrylamide gels (Drew and Travers, 1984). DNA was subjected to digestion with DNase I and hydroxyl radicals, under limited digestion conditions, as described (Brukner *et al.*, 1993). These conditions produce on average 0.16 cuts/molecule, the obtained cleavage intensities therefore reflect the affinity of the cleavaging agent for different parts of the DNA. The fragments resulting from DNA cleavage and the Maxam–Gilbert G reaction (Maxam and Gilbert, 1980) were separated on 8% polyacrylamide denaturing gels containing 8 M urea. The cleavage patterns were recorded using autoradiography. Quantitative evaluation was carried out by scanning image plates in a Molecular Dynamics PhosphorImager System.

Data processing and calculations

The procedure was similar to that described before (Brukner *et al.*, 1990). Briefly, the densitometric readings (which are proportional to observed DNase I cutting frequencies) were determined in triplicate and averaged. The raw data obtained in this manner were corrected for cuts that occurred between a cut and the radioactive label and smoothed by a three bond averaging procedure (Drew and Travers, 1984; Brukner *et al.*, 1990). This was done in order to diminish local fluctuations caused by individual structural features recognized by DNase I. The value of stagger between the complementary strands was two to the 5′ direction (Brukner *et al.*, 1990).

The smoothed values were then substituted into equation (2), which was subsequently solved by the singular value decomposition least squares method (Press *et al.*, 1992) to give the ln $p(a)$ parameter of the trinucleotides (Table I).

The stability of the model was tested with computer simulation experiments. Random error terms with an expected average value corresponding to ±5, ±10, ±20, ±30 and ±40% of the measured values were generated and added to the raw data and the ln $p(a)$ parameter was calculated with the simulated dataset. This procedure was repeated 20 times and the average and standard deviation of the parameters were calculated. For the per cent error values listed, the standard deviation values (averaged for all parameters) were found to be 0.002, 0.004, 0.007, 0.012 and 0.016 respectively.

Helical phasing experiments

Oligonucleotides were synthesized using an Applied Byosystems 381A Oligonucleotide Synthesizer and purified by 8% denaturing polyacrylamide gel electrophoresis. The purified oligonucleotides (0.6 μg) were 5′-end-labelled as described (Brukner *et al.*, 1993), complementary strands were mixed, heated to 90°C and slowly cooled to form hybrids. The ligation reactions were carried out as described (Brukner *et al.*, 1993) and the ligated products were run on 8% polyacrylamide gels, in the presence of 10 mM $MgCl_2$, in both gel and gel running buffers (Diekmann, 1987). The applied voltage was 5 V/cm and the electrophoresis was carried out at 22°C.

Acknowledgements

The authors thank Professor A.Falaschi (ICGEB) for help and advice throughout this project and Ms Suzanne Kerbavcic for her help in preparing the manuscript. Drs Horace Drew, Richard Lavery, Ana Savic, Andrew Travers and Victor Zhurkin provided valuable suggestions on the work. The help of Drs Miklós Cserzö and Mensur Dlakic with a previous version of the least squares program is gratefully acknowledged. D.S. is suppoorted by a grant from Deutsche Forschungsgemeinschaft. I.B. and R.S. are recipients of an ICGEB fellowship.

References

Barber,A.M. and Zhurkin,V.B. (1990) CAP binding sites reveal pyrimidine–purine pattern characteristic of DNA bending. *J. Biomol. Struct. Dynam.*, **8**, 213–232.

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

Brennan,G.R., Roderick,S.L., Takeda,Y. and Matthews,B.W. (1990) Protein–DNA conformational changes in the crystal structure of a λ Cro–operator complex. *Proc. Natl Acad. Sci. USA*, **87**, 8165–8169.

Brukner,I., Jurukovski,V. and Savic,A. (1990) Sequence-dependent structural variations of DNA revealed by DNase I. *Nucleic Acids Res.*, **18**, 891–894.

Brukner,I., Dlakic,M., Savic,A., Susic,S., Pongor,S. and Suck,D. (1993) Evidence for opposite groove-directed curvature of GGGCCC and AAAAA tracts. *Nucleic Acids Res.*, **21**, 1025–1029.

Brukner,I., Susic,S., Dlakic,M., Savic,A. and Pongor,S. (1994) Physiological concentration of magnesium ions induces a strong macroscopic curvature in GGGCCC-containing DNA. *J. Mol. Biol.*, **236**, 26–32.

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

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

De Santis,P., Palleschi,A. Savino,M. and Scipioni,A. (1990) Validity of the nearest-neighbour approximation in the evaluation of the electrophoretic manifestations of DNA curvature. *Biochemistry*, **29**, 9269–9273.

Diekmann,S. (1987) Temperature and salt dependence of the gel migration anomaly of curved DNA fragments. *Nucleic Acids Res.*, **15**, 247–265.

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

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

Gartenberg,M.R. and Crothers,D.M. (1988) Sequence determinants of CAP-induced bending and protein binding affinity. *Nature*, **333**, 824–829.

Grzeskowiak,K., Goodsell,D.S., Kaczor-Grezskowiak,M., Cascio,D. and Dickerson,R.E. (1993) Crystallographic analysis of CCAAGCTTGG

and its implications for bending in B-DNA. *Biochemistry*, **32**, 8923–8931.
Herrera,J.E. and Chaires,J.B. (1994) Characterization of preferred deoxyribonuclease I cleavage sites. *J. Mol. Biol.*, **236**, 405–411.
Hochschild,A. and Ptashne,M. (1986) Homologous interactions of lambda repressor and lambda Cro with the lambda operator. *Cell*, **44**, 681–687.
Hogan,M.E., Roberson,M.W. and Austin,R.H. (1989) DNA flexibility variation may dominate DNase I cleavage. *Proc. Natl Acad. Sci. USA*, **86**, 9273–9277.
Hunter,C.A. (1993) Sequence-dependent DNA structure: the role of base stacking interactions. *J. Mol. Biol.*, **230**, 1025–1054.
Kim,J.L. and Burley,S.K. (1994) 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nature Struct. Biol.*, **1**, 638–653.
Lahm,A. and Suck,D. (1991) DNase I-induced DNA conformation 2Å structure of a DNase I–octamer complex. *J. Mol. Biol.* **222**, 645–667.
Lavigne,M., Kolb,A., Yermian,E. and Buc,H. (1994) CRP fixes the rotational orientation of covalently closed DNA molecules. *EMBO J.*, **13**, 4983–4990.
Lomonossoff,G.P., Butler,P.J.G. and Klug,A. (1984) Sequence dependent variation in the conformation of DNA. *J. Mol. Biol.*, **149**, 745–760.
Maxam,A.M. and Gilbert,W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavage. *Methods Enzymol.*, **65**, 499–560.
McClarin,J.A., Frederick,C.A., Wang,B.-C., Greene,P., Boyer,H.W., Grable,J. and Rosenberg,J.M. (1986) Structure of the DNA–*Eco*RI endonuclease recognition complex at 3 Å resolution. *Science*, **234**, 1526–1541.
Poncin,M., Hartmann,B. and Lavery,R. (1992) Conformation sub-states in B-DNA. *J. Mol. Biol.*, **226**, 775–794.
Press,W.H., Teukolsky,S.A., Vetterling,W.T. and Flannery,B.P. (1992) In *Numerical Recipes in C: The Art of Scientific Computing*, 2nd Edn. Cambridge University Press, Cambridge, UK, pp. 676–680.
Quintana,J.R., Grzeskowiak,K., Yanagi,K. and Dickerson,R.E. (1992) Structure of a B-DNA decamer with a central TA step: CGATTAATCG. *J. Mol. Biol.*, **225**, 379–395.
Raumann,B.E., Rould,M.A., Pabo,C.O. and Sauer,R.T. (1994) DNA recognition by β-sheets in the Arc repressor–operator crystal structure. *Nature*, **367**, 754–757.
Satchwell,S.C., Drew,H.R. and Travers,A.A. (1986) Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.*, **191**, 639–659.
Schultz,S.C., Shields,G.C. and Steitz,T.A. (1991) Crystal structure of a CAP–DNA complex: the DNA is bent by 90°. *Science*, **253**, 1001–1007.
Shakked,Z., Guzikevich-Guerstein,G., Frolow,F., Rabinovich,D., Joachimiak,A. and Sigler,P.B. (1994) Determinants of repressor/operator recognition from the structure of the *trp* operator binding site. *Nature*, **368**, 469–473.
Shimon,L.J.W. and Harrison,S.C. (1993) The phage 434 OR2/R1-69 complex at 2.5 Å resolution. *J. Mol. Biol.*, **232**, 826–838.
Shlyakhtenko,L.S., Lyubchenko,Y.L., Chernov,B.K. and Zhurkin,V.B. (1990) Influence of temperature and ionic strength on electrophoretic mobility of synthetic DNA fragments. *Moleculyrnaya Biologiya*, **24**, 79–95.
Suck,D. (1994) DNA recognition by DNase I. *J. Mol. Recognition*, **7**, 65–70.
Travers,A.A. and Klug,A. (1990) In Cozzarelli,N.R. and Wang,J.C. (eds), *DNA Topology and its Biological Effects*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 57–106.
Weston,S.A., Lahm,A. and Suck,D. (1992) X-ray structure of the DNase I–d(GGTATCC) complex at 2.3 Å resolution. *J. Mol. Biol.* **226**, 1237–1256.
Winkler,F.K. *et al.* (1993) The crystal structure of *Eco*RV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.*, **12**, 1781–1795.
Zinkel,S.S. and Crothers,D.M. (1987) DNA bend direction by phase sensitive detection. *Nature*, **328**, 178–181.

Received on November 28, 1994; revised on January 25, 1995