Volume 300, number 2, 175–178 FEBS 10865 March 1992

A theoretical method to predict DNA permutation gel electrophoresis from the sequence

D. Boffelli[a], P. De Santis[a], A. Palleschi[a], G. Risuleo[b] and M. Savino[b]

[a]*Dipartimento di Chimica, Università di Roma 'La Sapienza', P. le A. Moro 5, 00185 Roma, Italy* and [b]*Dipartimento di Genetica e Biologia Molecolare, Università di Roma 'La Sapienza' P. le A. Moro 5, 00185 Roma, Italy*

Received 15 January 1992

The gel electrophoretic permutation assays of DNA fragments experimentally investigated by different authors were theoretically reproduced using our theoretical model of sequence-dependent curvature. The general pattern of agreement obtained suggests that our method can be usefully adopted as an alternative to the experimental assay, in particular where the lack of a sufficient number of unique restriction sites in the fragment prevents the correct localization of the main bend site.

Theoretical prediction; DNA curvature; DNA permutation; Gel electrophoresis

1. INTRODUCTION

After the detection of sequence dependent curvature in K-DNA, a kinetoplast DNA fragment of *L. tarentolae* [1], many DNA tracts were shown to be intrinsically curved, as indicated by their anomalously slow electrophoretic mobility through polyacrylamide gels [1–12].

This superstructural aspect of the double helix has, in fact, been found in biologically relevant DNA tracts implicated in regulation, recombination and replication processes, and seems to be involved in the phasing of nucleosomes [13].

While it appears to be relatively easy to detect the presence of curvature in DNA fragments from their electrophoretic anomalies, it is, however, more difficult to locate bend sites. Several years ago Wu and Crothers [3] devised an ingenious electrophoretic assay to map the bending site (where a unique maximum curvature is present) in a DNA fragment. The technique involves digesting tandem dimer fragments at unique restriction endonuclease sites and plotting the gel electrophoretic retardation of the fragments against their base-pair position. The possible bend site should correspond to the interpolated minimum of this retardation plot.

Curvature sites of several DNA tracts have been identified using this technique, and permutation gel electrophoresis has become a powerful tool to investigate the implications of curvature in biologically relevant DNA fragments when a suitable set of single restriction sites is present. The technique, however, can fail where two or more bend sites in contrast of phase are present in the DNA tract.

A few years ago we theoretically investigated the physical basis of the sequence-dependent curvature in DNA and found that the nearest-neighbour differential interaction between base pairs could suffice to explain the origin of superstructures in DNAs [14–19], as originally suggested by Trifonov [20,21].

Among these investigations, we found electrophoretic retardation to be linearly correlated in first approximation with the dispersion of curvature (σ^2) along the sequence. This is proved by the good agreement (correlation factor = 0.98) between σ^2 and the corresponding electrophoretic retardation (ranging in value between 1.0 and 3.3), of about 450 multimeric synthetic DNA oligomers which differ in periodicity, sequence and length (50–200 bp) [16–19].

This success prompted us to theoretically reproduce the permutation assays of DNA tracts experimentally investigated, even if their lengths might overcome the molecular complexity of the synthetic DNAs investigated and the electrophoretic conditions used (percentage of acrylamide and temperature) were not standardized. In fact it must be considered that in permutation gel electrophoresis only the trend of retardation against the restriction sites along the sequence is required for locating the main bend site.

2. EXPERIMENTAL

Modulus and phase of the differential curvature were calculated per turn of B DNA as the angular deviation of the double helical axis on a DNA tract of g bp around sequence number n:

Correspondence address: P. De Santis, Dipartimento di Chimica, Università di Roma 'La Sapienza', P. le A. Moro 5, 00185 Roma, Italy.

Published by Elsevier Science Publishers B.V.

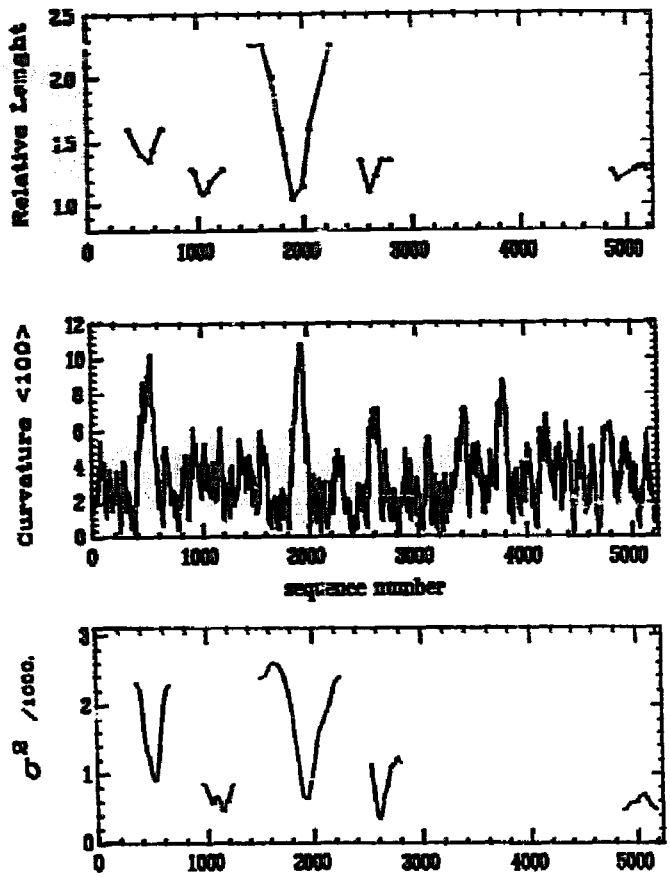

Fig. 1. General pattern of the permutation assays of SV40 fragments investigated by Milton and Gesteland [11] compared with the curvature diagram of the whole sequence calculated over recurrent tracts of 100 bp, as well as with the pattern of theoretical assays.

$$C(n) = v^\circ/g \sum_{s=n-g/2}^{n+g/2} d(s) \exp(2\pi is/v^\circ)$$

where: g is the integration grid, $v^\circ = 10.4$ is the periodicity of B-DNA, and $d(s) = (\rho - i\tau)_s$ is the deviation of the dinucleotide step at sequence number s from the canonical B-DNA, as theoretically evaluated by conformational energy calculations.

Curvature was evaluated over 3 recurrent turns of B-DNA (g = 31 bp) to reduce the noise and reported versus the sequence number n as 'curvature function'. This diagram allows an easy visualization of curvature sites along DNA corresponding to the maxima of the curvature profile.

Starting from this representation of the superstructural features of DNA it is easy to reproduce a suitable three-dimensional representation of the DNA superstructure as a writhing of the helix axis, assigning the modulus of curvature as a bending angle and the phase increment as a torsional angle around the corresponding local helical axis.

The dispersion σ^2 of the curvature in a given fragment of DNA was calculated as the second moment of the angular deviations of the local helical axis (practically the perpendicular to each base pair plane) from their average direction.

This corresponds in the complex plane representation to:

$$\sigma^2 = < C(n) \mid C(n) > - < C(n) > < C(n) >$$

where:

$$C(n) = v^\circ/g \sum_{s=1}^{n} d(s) \exp(2\pi is/v^\circ)$$

is the integral curvature between the first and the n-th nucleotide step (n is the total number of bp in the considered tract).

We found σ^2 to be linearly correlated with gel electrophoretic retardation, quantified as the ratio between apparent and real lengths. This allows the easy calculation of an analog of the permutation assay by plotting σ^2 for the cyclic permuted sequences.

3. RESULTS AND DISCUSSION

We have adopted our theoretical model to predict the experimental permutation assays of several DNA fragments (180–750 bp in length) investigated by different authors. The experimental points were obtained from the corresponding illustrations of the original papers by visual interpolation.

A general pattern of agreement is obtained, in particular as the common trends of σ^2 and retardation profiles are concerned, whereas the absolute values are not always reproduced because of the lack of standardized experimental conditions (percentage of acrylamide and temperature), as well as of the high molecular weight which plausibly amplifies the approximation defects of the model. Nevertheless, the positions of the curved regions along the sequence were correctly predicted in all the cases.

Fig. 1 show the five tracts corresponding to some anomalous fragments of restricted SV40 investigated by Milton and Gesteland [11]. The trends of experimental and theoretical permutation assays are reported along the sequence. In the same figure the positions of the bent regions are represented in the curvature diagram, averaged over recurrent tracts of 100 bp. It is easy to observe that the minima of permutation assays correspond to the maxima of the curvature diagram which also shows the presence of some regions of SV40 DNA not considered by the cited authors.

Fig. 2 illustrates the case of the curved DNA at the upstream activation sites (UAS) of α-cell-specific genes in yeast, investigated by Inokuchi et al. [12] using circular permutation analysis of the 513 bp fragment, which contains two UAS regions around the main curvature site. The two diagrams of the right panel show the direct comparison between the relative mobility and the change of the dispersion of curvature

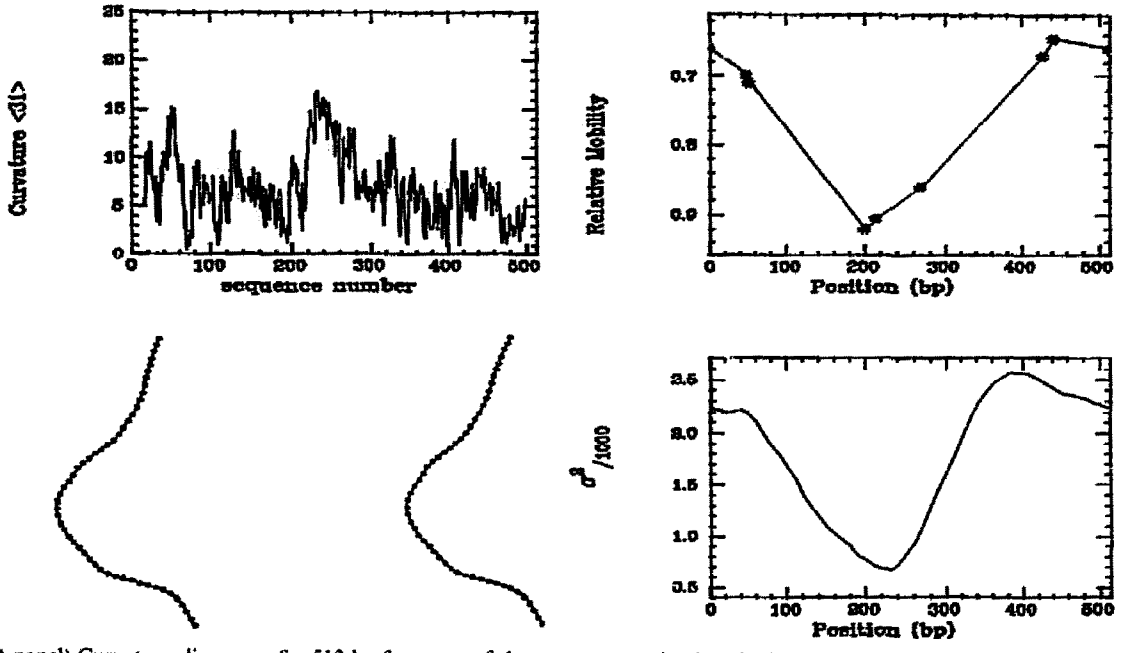

Fig. 2. (Left panel) Curvature diagram of a 513 bp fragment of the upstream activation site in yeast and the corresponding stereoprojections of the writhing of the helix axis. (Right panel) Comparison between experimental [12] and theoretical permutation assays.

resulting from the cyclic permutation of the sequence. The trend of the two diagrams is practically the same and the position of the minimum corresponds to the center of the curved tract as represented in the stereoprojection as well as to the main maximum of the curvature diagram, both reported at the left panel of the figure.

Fig. 3 shows the case of the ori region of λ bacteriophage 178 bp fragment investigated by Zahn and Blatter [5]. In spite of the lack of restriction sites over more than two thirds of the sequence, the authors localized the maximum of the curvature at about $n \approx 51$. Such a position in fact should be shifted by one turn of a double helix, as clearly shown by the theoretical diagrams.

Finally, Fig. 4 illustrates the case an electrophoreti-

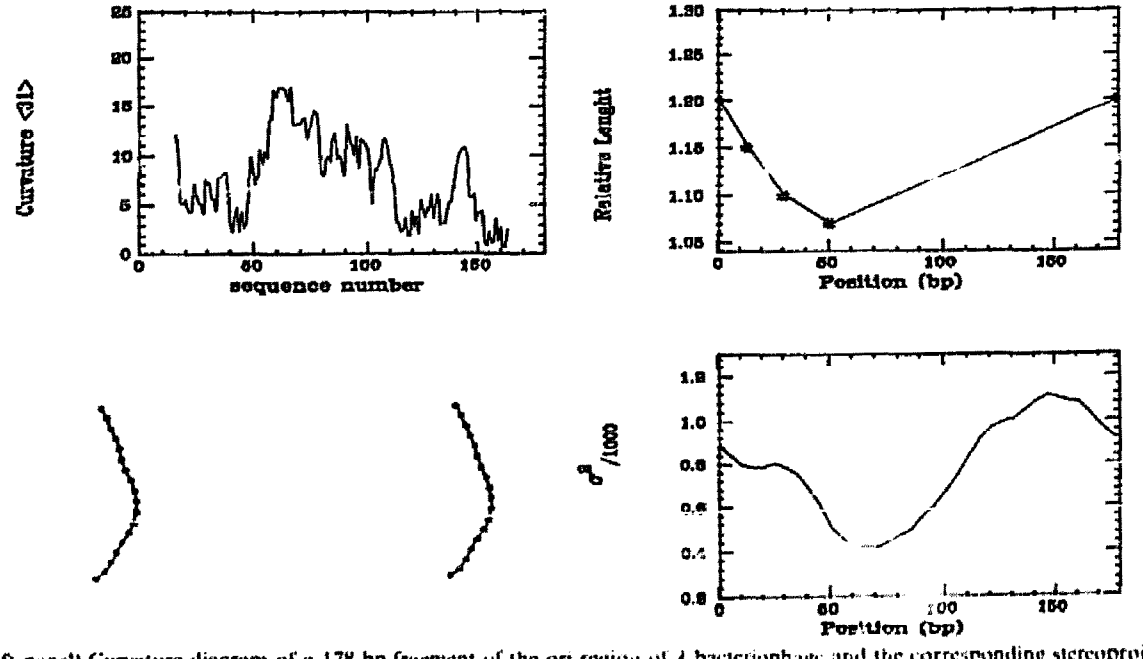

Fig. 3. (Left panel) Curvature diagram of a 178 bp fragment of the ori region of λ bacteriophage and the corresponding stereoprojection of the writhing of the helix axis. (Right panel) Comparison between experimental [5] and theoretical permutation assays.

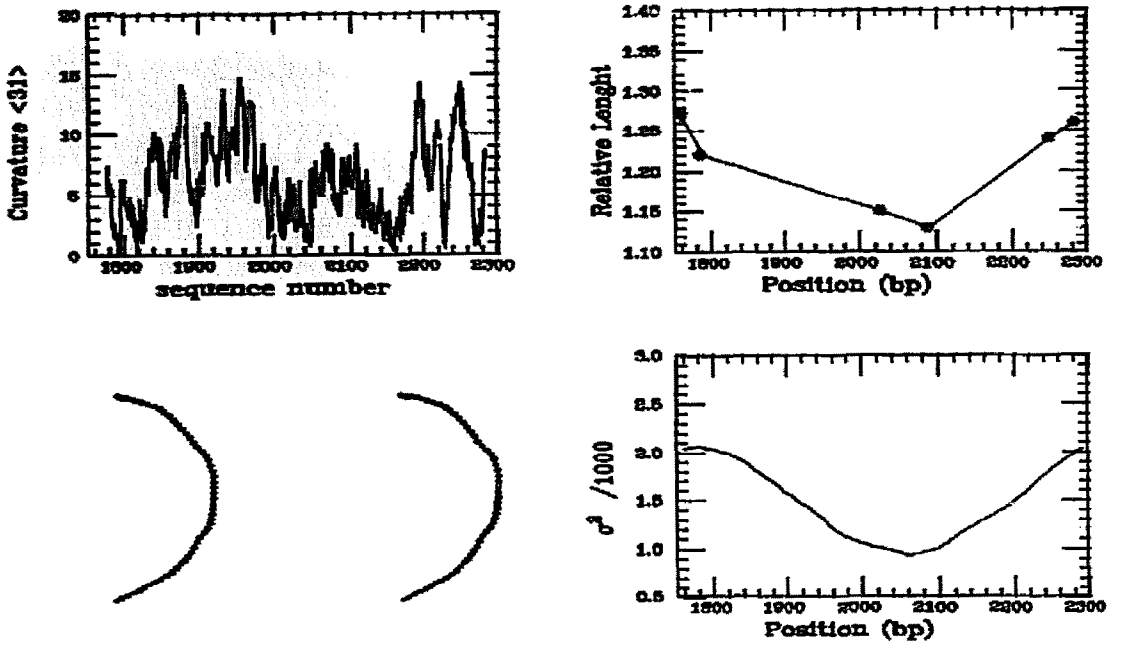

Fig. 4. (Left panel) Curvature diagram of a 537 bp fragment of the Polyoma virus and the corresponding stereoprojection of the writhing of the helix axis. (Right panel) Comparison between experimental [22] and theoretical permutation assays.

cally anomoulous tract of Polyoma virus (1761–2298 tract) investigated in our laboratory [22]. The fragment was restricted with *Acc*I, *Pvu*II, *Sph*I, *Ava*II and *Taq*I after dimerization and cloning in pUC18. Also in this case, in spite of the rather uniform curvature as shown by the stereoprojection, the retardation gel permutation diagram is theoretically reproduced.

Apart from the cases illustrated in this paper, we analyzed all the permutation assays appearing in the literature and we generally found satisfactory results. However, a deeper analysis of experimental data shows that in most cases the limited number of unique restriction sites bars the possibility of precisely locating curvature sites. In some cases the interpolation of experimental data locates the hypothetical curvature centre on an incorrect DNA region with a consequent misunderstanding of the sequences responsible for DNA curvature.

As a general conclusion, the main result of this paper is that our theoretical method is a valid alternative to permutation gel electrophoresis in order to locate the curved regions of DNA fragments.

Acknowledgements: The financial support of CNR-Progetto Finalizzato Chimica Fine II and of Istituto Pasteur-Fondazione Cenci Bolognetti is gratefully acknowledged.

REFERENCES

[1] Marini, J.C., Levene, S.D., Crothers, D.M. and Englund, P.T. (1982) Proc. Natl. Acad. Sci. USA 79, 7664–7668.
[2] Rees, W. and Landy, A. (1982) J. Mol. Biol. 156, 523–530.
[3] Wu, H.M. and Crothers, D.M. (1984) Nature 308, 509–512.
[4] Bossi, L. and Smith, D.M. (1984) Cell 39, 643–652.
[5] Zahn, K. and Blattner, F.R. (1985) Nature 317, 451–453.
[6] Koepsel, R.R. and Khan, S.A. (1986) Science 233, 1316–1319.
[7] Deb, S., DeLucia, A.L., Koff, A., Tsui, S. and Tegtmayer, P. (1986) Mol. Cell Biol. 6, 4578–4584.
[8] Snyde, M., Buchman, A.R. and Davis, R.W. (1986) Nature 324, 87–89.
[9] Gourse, R.L., de Boer, H.A. and Homura, M. (1986) Cell 44, 197–205.
[10] McAllister, C.F. and Achberger, E.C. (1988) J. Biol. Chem. 263, 11743–11749.
[11] Milton, D.L. and Gesteland, R.F. (1988) Nucleic Acids Res. 16, 3931–3949.
[12] Inokuchi, K., Nakayama, A. and Hishinuma, F. (1988) Nucleic Acids Res. 16, 6693–6711.
[13] Boffelli, D., De Santis, P., Palleschi, A. and Savino, M. (1991) Biophys. Chem. 39, 127–136.
[14] De Santis, P., Morosetti, S., Palleschi, A. and Savino, M. (1986) in: Structures and Dynamics of Nucleic Acids, Proteins and Membranes (E. Clementi and S. Chin eds.) pp. 31–49, Plenum Press, New York.
[15] De Santis, P., Morosetti, S., Palleschi, A., Savino, M. and Scipioni, A. (1988) in: Biological and Artificial Intelligence Systems (E. Clementi and S. Chin eds.) pp. 143–161, ESCOM Science, New York.
[16] De Santis, P., Palleschi, A., Savino, M. and Scipioni, A. (1988) Biophys. Chem. 32, 305–317.
[17] De Santis, P., Gallo, G., Palleschi, A., Savino, M. and Scipioni, A. (1989) J. Mol. Liquids 41, 291–304.
[18] Cacchione, S., De Santis, P., Foti, D., Palleschi, A. and Savino, M. (1989) Biochemistry 28, 8706–8713.
[19] De Santis, P., Palleschi, A., Savino, M. and Scipioni, A. (1990) Biochemistry 29, 9269–9273.
[20] Trifonov, E.N. and Sussmann, J.L. (1980) Proc. Natl. Acad. Sci. USA 77, 3816–3820.
[21] Trifonov, E.N. (1980) Nucleic Acids Res. 8, 4041–4053.
[22] De Santis, P., Leoni, L., Palleschi, A., Risuleo, G. and Savino, M. (1991) Biochemistry Int. 24, 709–719.