Biophysical Chemistry, 33 (1989) 217–226
Elsevier

BIOCHE 01356

Curvature of the light-responsive element of the plant gene *rbc*S-3A encoding the small subunit of ribulose-1,5-bisphosphate carboxylase

S. Cacchione, M. Savino and G. Tecce

Dipartimento di Genetica e Biologia Molecolare, Università di Roma 'La Sapienza', P. le A. Moro 5, 00185 Roma, Italy

Received 7 November 1988
Revised manuscript received 26 January 1989
Accepted 6 February 1989

Ribulose-1,5-bisphosphate carboxylase; Light-responsive element; Nucleotide sequence; Gene, *rbc*S-3A; DNA curvature

The possible curvature in the light-responsive element (LRE) of the plant gene which codes for the small subunit of ribulose-1,5-bisphosphate carboxylase has been derived from the nucleotide sequence of the LRE on the basis of a theoretical method previously developed in our research group. The oligonucleotides corresponding to regions of higher curvature in the profile, taking also into account biochemical data, have been synthesized and their electrophoretic mobilities have been measured after ligation. Retardation effects have been found, increasing with length of the multimers, in good agreement with theoretical prediction. These results suggest that the curvature of LRE could be the structural determinant in the control of *rbc*S-3A gene photoactivation.

1. Introduction

Light plays the main role in the development of plants. The photoregulation of plant genes depends mainly on the interaction of electromagnetic radiation of defined frequency and intensity with the molecule of phytochrome, that acts as photoreceptor.

The gene family, *rbc*S, which codes for the small subunit of ribulose-1,5-bisphosphate carboxylase, a protein active in photosynthesis, shows a large increase in the quantity of transcript in the presence of light. Analyses by deletions of the *rbc*S genes in transgenic plants have led to the localization of nucleotide sequences, upstream of the *rbc*S genes. which confer phytochrome-induced transcription [1].

More recently, nuclear extracts, used in gel-retardation assays and DNase I footprinting experiments, have permitted the identification of a protein factor that specifically interacts with short sequences in a 240 bp DNA fragment upstream of the *rbc*S-3A gene. This region is known to function as the light-responsive element (LRE) [2].

We have carried out an analysis of the possible curvature in LRE, since it has been recently demonstrated that this property of DNA plays an important role in the recognition of specific nucleotide sequences by proteins. Crothers and co-workers [3] have shown that a DNA sequence upstream of the *lac* operon becomes bent on binding to the catabolite activator protein (CAP); it would not be surprising to find that other specific protein factors recognize their target DNA sequences on the basis of potential or actual curvature.

Our analysis has taken advantage of a theoretical method recently developed by De Santis et al. [4–7] to predict DNA curvature from the nucleotide sequence. Oligonucleotides, corresponding to the peaks of the theoretical curvature profile of the sequence under investigation, were synthesized

Correspondence address: M. Savino, Dipartimento di Genetica e Biologia Molecolare, Università di Roma 'La Sapienza', P. le A. Moro 5, 00185 Roma, Italy.

taking also into account the data derived from biochemical analyses of the complexes between this DNA region and specific protein factors [2]. The electrophoretic mobility of the synthesized oligomers was measured after ligation up to, at least, 15 repeating units.

With this method the local curvature of 10 or 20 nucleotides is repeated n times in the relative polynucleotide, obtained by ligation, and gives rise to a global superstructure, allowing measurement by gel-retardation assay [8,9].

2. Materials and methods

Polynucleotide kinase and T_4 DNA ligase were purchased from Boehringer; all nucleoside β3-cyanoethylphosphoroamidites were from Biosearch.

Oligodeoxyribonucleotides were synthesized on a Biosearch model 8600 DNA synthesizer and purified on a 20% polyacrylamide gel in the presence of 7 M urea, followed by high salt (0.1 M ammonium bicarbonate) elution from a Sephadex G-50 column.

Polydeoxyribonucleotides were synthesized according to the following procedure. 2 μg of each oligomer was 5′-labelled with 5 μCi [γ-^{32}P]ATP (5 μCi pmol^{-1}; Amersham) in 10 μl solution, at 37°C for 15 min. The reaction solution contained 70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT) and 3 U polynucleotide kinase. After labelling with [γ-^{32}P]ATP, 50 nmol unlabelled ATP and an additional 2 U kinase were added, and the reaction volume was increased to 15 μl. Kinasing with unlabelled ATP was continued in the same buffer and at the same temperature as above for a period of 1 h. Two kinased oligonucleotides of complementary sequence (except for single-strand residues at the 5′-ends) were mixed, heated to 55°C and slowly cooled to 0°C to form hybrids. 1 U T_4 DNA ligase was added to 10 μl hybridized mixture and the reaction volume was adjusted to 20 μl: the ligation mixture contained the same concentration of Tris-HCl, $MgCl_2$ and DTT as the kinasing step, plus 2.3 mM unlabelled ATP. The ligation reaction was allowed to proceed on ice overnight, thereafter being quenched by addition of EDTA (pH 8) to 25 mM final concentration.

Ligated products were run on non-denaturing 10% polyacrylamide gels (mono/bisacrylamide ratio = 29 : 1; 90 mM Tris-borate, 2.5 mM EDTA, pH 8.3) until the bromophenol blue dye had migrated 25 cm. The applied voltage was 4 V cm^{-1}. Electrophoresis was carried out at room temperature or at 4°C.

3. Results

3.1. Theoretical curvature profile and synthetic oligonucleotides as probes

Recently, a theoretical method was developed by De Santis et al. [4–7] to translate the fluctuations of DNA base sequences in elements of the DNA superstructure by summing up the pertinent elements of a complex matrix representing the local deviations of the 16 dinucleotide double-helical fragments from the standard B-DNA, obtained by energy-conformational calculations. Many natural DNA sequences were analysed using this method and their curvature was represented by a pair of diagrams where both the modulus and relative phase of the curvature vector per turn, C_n, were reported as function of the sequence number, n. C_n represents the curvature vector, namely, the angular deviation of the perpendicular at the n-th base-pair plane; it corresponds to the deviation in direction of the helical axis after n nucleotide steps. According to the method of De Santis et al. [7], C_n is calculated as follows:

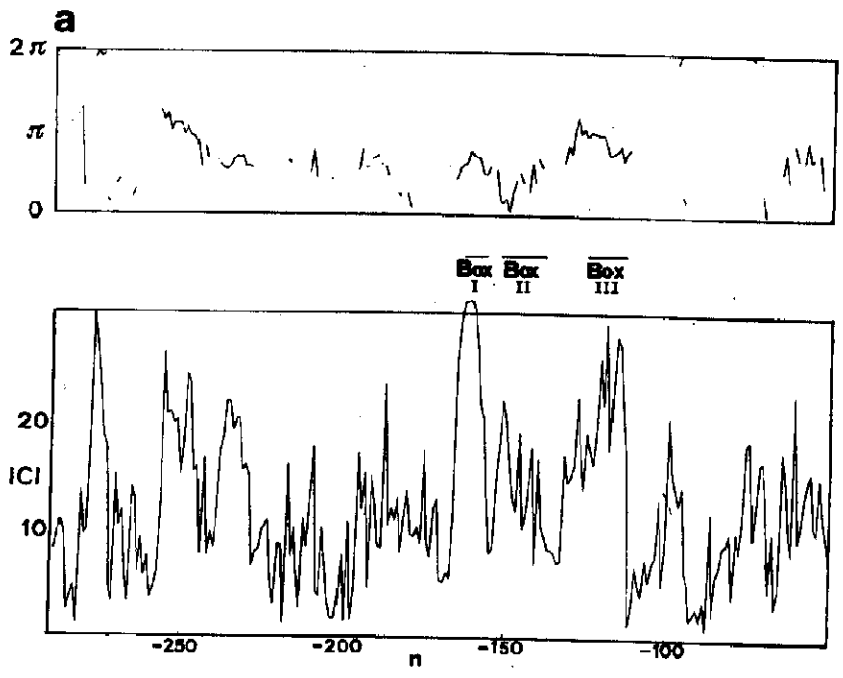
$$C(n,\nu) = \nu^0/\nu \sum_{j=n}^{n+\nu} d_j \exp\left[2\pi i(j-1)/\nu^0\right]$$

where $d_j = r_j - it_j$ (r, roll; t, tilt), ν^0 represents the periodicity of B-DNA, d_j is Fourier transform of the local structural fluctuations along the sequence and ν is chosen to be about a multiple of ν^0.

Fig. 1a illustrates the curvature profile, reported as the modulus of the curvature vector $|C|$, and the relative phase of the 240 bp element that has recently been shown [1,2] to control the

photoactivation, via phytochrome, of one of the genes which code for the small subunit of the ribulose-1,5-bisphosphate carboxylase, *rbc*S-3A, in *Pisum sativum*. The profile appears complex, but four regions with a remarkably high curvature emerge, corresponding to the sequences −280/−270, −260/−240, −165/−155 and −125/−115 from the start of transcription, respectively.

In a first attempt, we decided to limit our analysis to the DNA region between −170 and −110, where the theoretical curvature profile shows two maxima. Further, Green et al. [2] have recently shown that this region is capable of inter-



b

rbcS-3A

-291

ACAAAATAAA AATATCGACA TAACCACCAT CACACATTTA CACTCTTCAC

ATGAAAAGAT AAGATCAGTG AGGTAATATC CACATGGCAC TGTCCTATTG

GTGGCTTATG ATAAGGCTAG CACACAAAAT TTCAAATCTT GTGTGGTTAA

TATGGCTGCA AACTTTATCA TTTTCACTAT CTAACAAGAT TGGTACTAGG

CAGTAGCTAA GTACCACAAT ATTAAGACCA TAATATTGGA

-52

Fig. 1. (a) Curvature profile and relative phase of the 240 bp upstream regulatory region of the ribulose-1,5-bisphosphate carboxylase (*rbc*S) 3A gene from pea plants. The bars indicate the experimental localization of the nucleotide sequences, that interact with a protein, having probably regulatory functions [2]. (b) Nucleotide sequence of the 240 bp DNA fragment.

Table 1

Sequences examined (relative to *rbc*S-3A regulatory region)

Only one strand of the duplex is shown. The duplexes of the decamers (except NBH, which is self-complementary) were constructed with 2-bp protruding 5′-ends, while those of the 20-mers have 3-bp protruding 5′-ends.

Name	Sequence (5′-3′)
BI	AATTTCAAAT (−164/−155)
BI/II	AATCTTGTGT (−157/−148)
BII	GTGTGGTTAA (−151/−142)
B▼I/II	CAAAATTTCAAATCTTGTGT (−167/−148)
MII/III	TGCAAACTTT (−135/−126)
BIII	ATCATTTTCA (−125/−116)
B▼III	TGCAAACTTTATCATTTTCA (−135/−116)
B▼II	CTTGTGTGGTTAATATGGCT (−154/−135)
N(ot)B(ent)	TCCTATTGGTGGCTTATGAT (−199/−180)
A_5N_5 [8]	GGCAAAAACG
N(ot)B(ent)-H(agerman) [9]	GGGTCGACCC

acting specifically with a protein factor probably having a regulatory function. In fact, the complexes between DNA fragments belonging to this region and protein nuclear extracts show inhibition of DNase I cleavage and band shift with respect to naked DNA on gel electrophoresis.

To check for the presence of the two remarkably curved regions mentioned above about $n = -160$ and -120 and of a region with a lower curvature between $n = -150$ and -130, we synthesized and then ligated up to at least 15 monomers, 10-mers and 20-mers corresponding to these regions. The superstructure of these multimers has been investigated by gel electrophoresis retardation assay, according to the method first developed by Crothers et al. [8] and Hagerman [9].

The examined sequences are reported in table 1. We selected the sequences to be synthesized, with the taking into account of both the theoretical curvature profile and the results of biochemical analysis [2] that led to the localization within the region from −170 to −110 of three sequences called BoxI–III, which appear to be necessary to gene photoactivation. The three boxes correspond

to the nucleotides underlined in the sequence reported in fig. 1b.

The five synthesized 10-mers reported in table 1 correspond to the two maxima of the theoretical curvature profile and to sequences in between for which the analysis predicts a noticeably lower curvature. They also correspond to BoxI–III and to two sequences between BoxI and II and between BoxII and III, respectively. However, it should be taken into account that the sequences derived from biochemical analysis contain a different number of bases from 10, namely, BoxI (7 bp), BoxII (14 bp) and BoxIII (12 bp), whereas it is necessary to synthesize 10-mers or 20–21-mers to obtain polymers with the repetition period of B-DNA (about 10.5). The four 20-mers have been planned as follows: B$^{\blacktriangledown}$III contains 10 nucleotides of the sequence between BoxII and III and 10 nucleotides of BoxIII: B$^{\blacktriangledown}$I/II contains BoxI and a part of BoxII; B$^{\blacktriangledown}$II contains the entire BoxII sequence and six adjacent nucleotides; NB comprises 20 nucleotides of the region upstream with respect to the three boxes.

3.2. Correlation between curvature profile and gel-electrophoretic mobility

The 10- and 20-mers were subsequently ligated to obtain multimer distributions. Also, the sequences A_5N_5 and NBH [8,9] were synthesized and ligated for use as standards. In fact, it was previously shown that the sequence NBH presented normal electrophoretic behaviour whereas the sequence A_5N_5 demonstrated marked retardation that was interpreted as being due to large curvature of the corresponding monomer. Fig. 2 shows a typical polyacrylamide gel, where the anomalous electrophoretic behaviour of some sequences with respect to others is evident. The migration patterns obtained from this and similar gels are illustrated by plotting the ratio R of the apparent number of base-pairs (BP_{app}) to the true number of base-pairs (BP_{seq}) as a function of the degree of polymerization, N.

The results on multimers obtained from the 10-mers listed in table 1 are collected in fig. 3. Apparent bending is substantial for BI and BIII, and decreases on going from BoxIII to II as

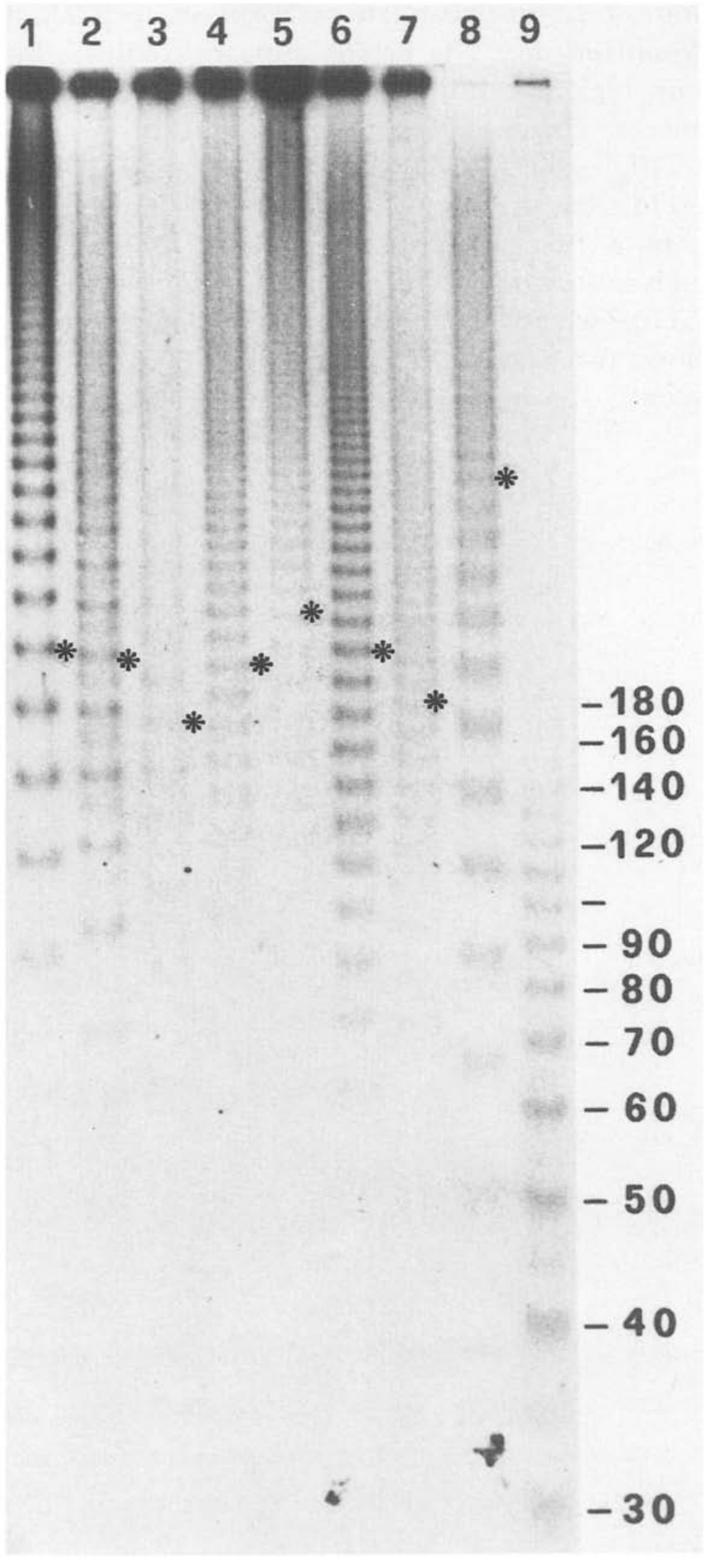


Fig. 2. Autoradiogram of the multimer series (derived from ligation of nine monomers reported in table 1) electrophoresed on a non-denaturing 10% polyacrylamide gel at room temperature. (1) B$^{\blacktriangledown}$III, (2) B$^{\blacktriangledown}$I/II, (3) NB, (4) MII/III, (5) BIII, (6) BI, (7) BI/II, (8) A_5N_5, (9) NBH. Asterisks indicate multimers of 160 bp sequence length.

shown by the decrease in retardation of MII/III in comparison with that of BIII. However, the apparent bending of MII/III is still considerable as compared with the sequence BII which gives rise to only slightly anomalous behaviour on electrophoresis. It is interesting to note the small but evident increase in retardation in passing from BoxII to I, shown by the sequence BI/II with respect to BII.

From these results the correlation between the trend of the theoretical curvature profile and the mobility on gel electrophoresis of synthetic multimers appears sufficiently high for one to consider the general trend of the theoretical curvature profile of fragment $-170/-110$ as being reliable on an experimental basis. Some minor discrepancies result from a quantitative comparison; in particular, the predicted curvature for BI is slightly greater than that of BIII, while the reverse of this trend is shown by electrophoretic analysis. However, one should bear in mind that the agreement between

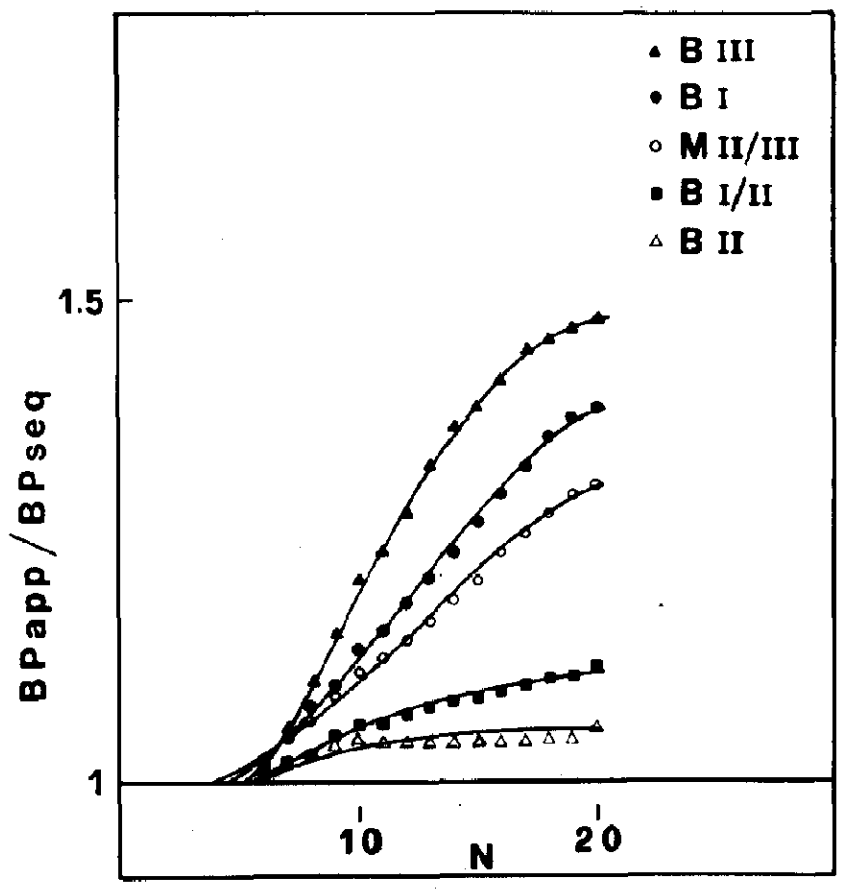


Fig. 3. Mobilities of duplex polymers derived from 10-mers reported in table 1, represented as R, the ratio of the apparent number of base-pairs (BP_{app}) to the true number of base-pairs (BP_{seq}), are plotted as a function of the degree of polymerization, N. The ratio BP_{seq}/BP_{app} is determined from comparison with the electrophoresis marker. Gels were run at room temperature.

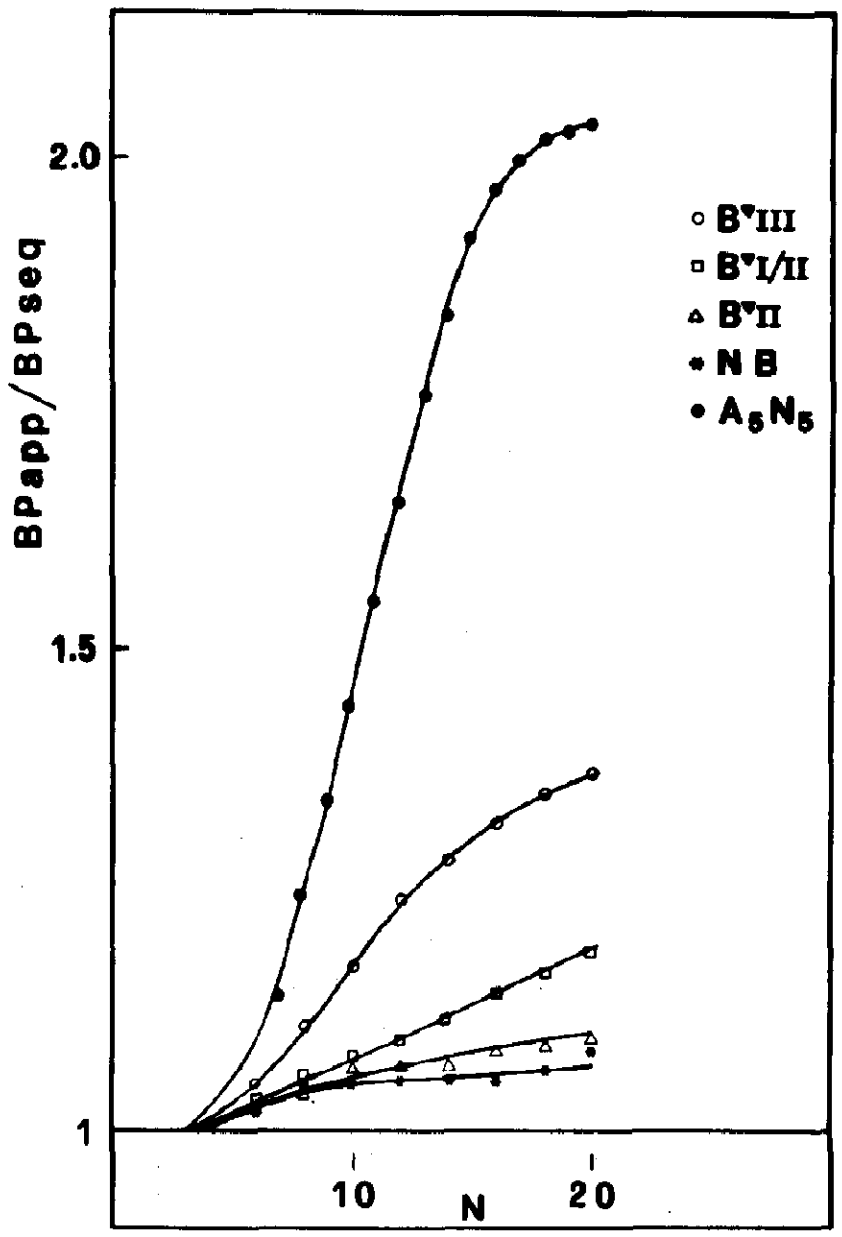


Fig. 4. Mobilities of duplex polymers derived from 20-mers reported in table 1, represented as R, are plotted as a function of N. R is determined from comparison with the electrophoresis marker. Gels were run at room temperature.

theoretical and experimental results could be influenced by the terminal nucleotides. This last consideration prompted us to investigate the behaviour of multimers deriving from ligation of 20-mers, since in this case the terminal effects should be decreased.

The results are reported in fig. 4. The behaviour of A_5N_5 is also shown for the sake of comparison. The multimers, derived from 20-mers, display behaviour in agreement with the theoretical curvature profile, namely, large retardation effects, for B▼III and B▼I/II, corresponding to the two peaks of the theoretical curvature profile, and low R values in the case of B▼II and NB.

The sequence corresponding to peak B▼I/II has a lower curvature with respect to that corre-

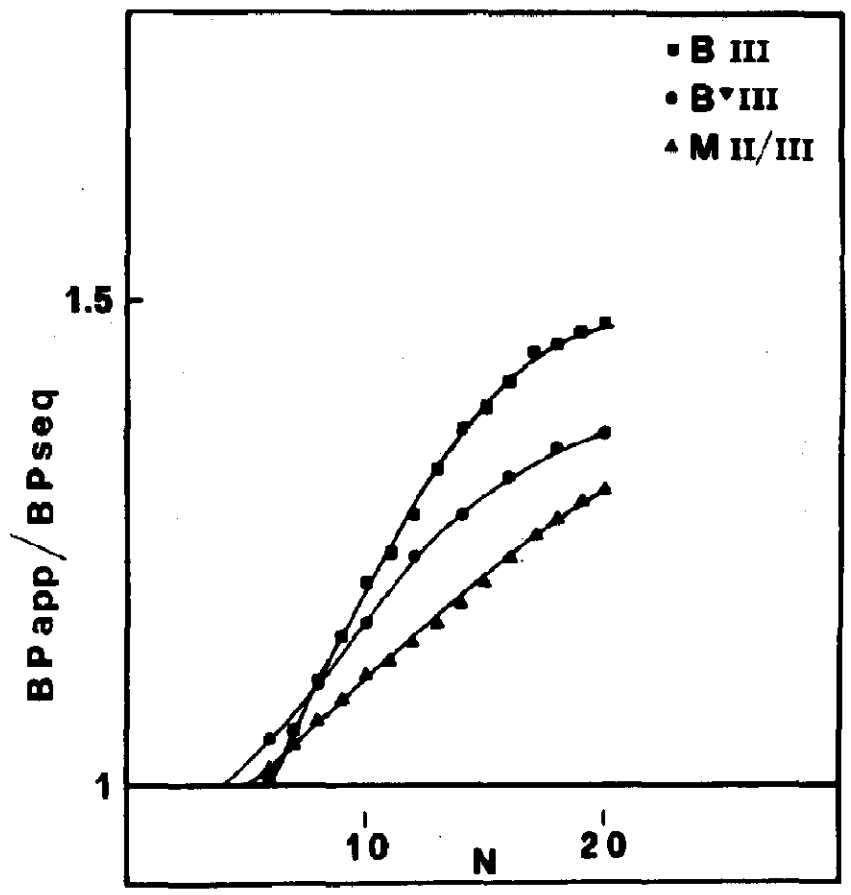


Fig. 5. Comparison between the mobilities of multimers of the 20-mer B▼III and of the two decamers MII/III and BIII that represent the consecutive half-lengths of 20-mer from 5′.

sponding to peak B▼III; this feature could derive from the deep minimum contained in B▼I/II.

3.3. Phasing of monomers

The curvature of a sequence derives from the repetition of elements in phase with the helix screw and it is obviously of interest to verify experimentally the theoretical phase profile. In fig. 1 the trend of the phase angle shows a large variation in the region from $n = -170$ to -110. However, the points corresponding to the 10-mer centred on $n = -131$ and to the 10-mer centred on $n = -120$ have equal phase angles, namely, 158°. To confirm this prediction we have compared (see fig. 5) the gel electrophoretic behaviour of multimers derived from the 20-mer B▼III and those derived from the 10-mers BIII and MII/III. It has been found that R_{15} of B▼III is approximately equal to $(R_{15}\text{BIII} + R_{15}\text{MII/III})/2$. This result demonstrates that the two adjacent 10-mers are in phase and that the curvature vector as well as the modulus is verifiable by measurements on suitable oligonucleotides.

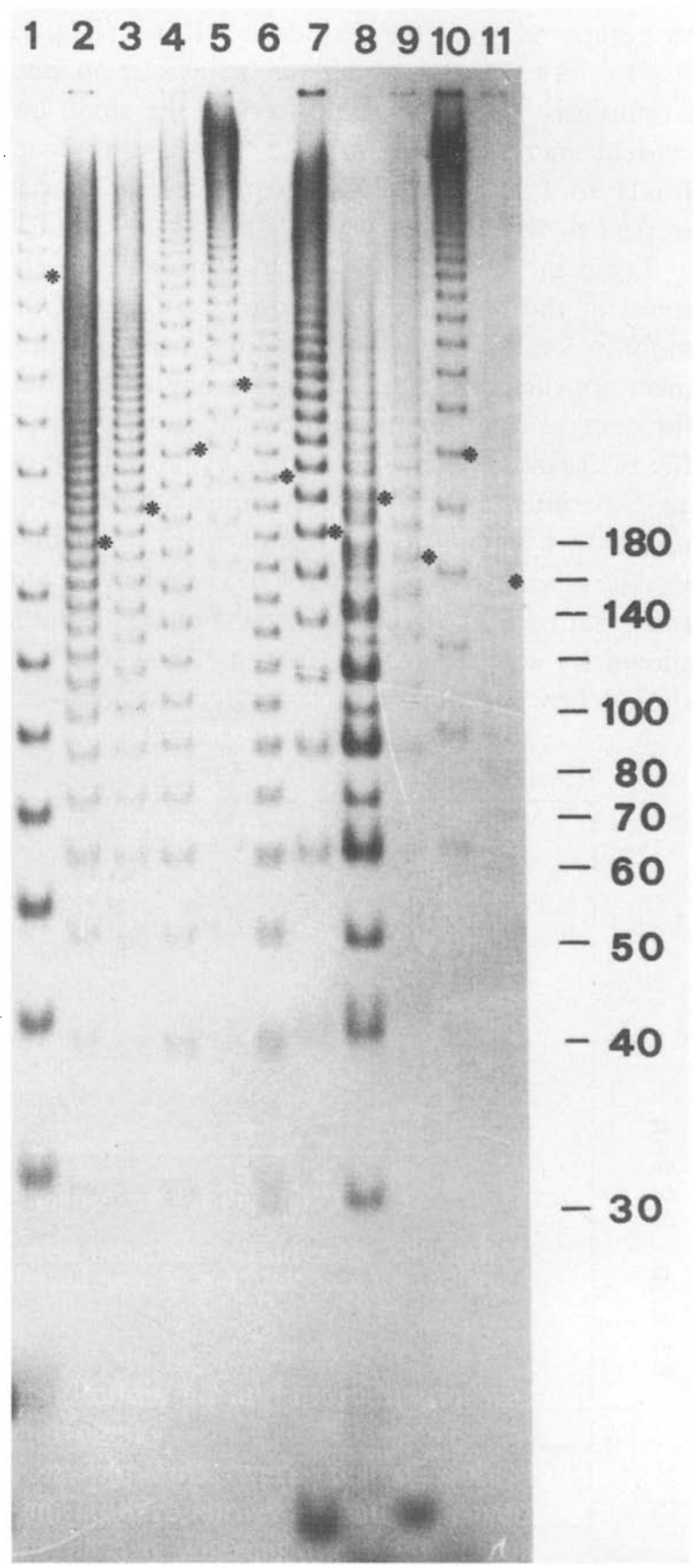


Fig. 6. Autoradiogram of the multimer series (derived from ligation of the monomers reported in table 1) electrophoresed on a non-denaturing 10% polyacrylamide gel at 4°C. (1) A_5N_5, (2) BII, (3) BI/II, (4) BI, (5) BIII, (6) MII/III, (7) B▼II, (8) B▼I/II, (9) NB, (10) B▼III, (11) NBH. Asterisks indicate multimers of 160 bp sequence length.

3.4. Temperature

It has been reported that in natural curved DNA fragments, as well as in synthetic bent sequences, the gel mobility is considerably influenced by temperature [8,10].

Fig. 6 depicts the mobilities of multimers on a gel run at 4°C and greater retardation effects are clearly evident with respect to those shown in fig. 2. In figs. 7a and 8a, the respective retardation of multimers derived from 20-mers and 10-mers is reported. Figs. 7b and 8b demonstrate the relative changes in R values with temperature, expressed as the ratio R^4/R; it can be seen that the highly anomalous series of molecules A_5N_5 become even more so at lower temperature (R_{15} at 4°C = 2.7: R_{15} at 20°C = 1.9), in agreement with the results previously reported by Koo et al. [8].

The multimers referring to the regulatory region of *rbc*S-3A in all cases show a greater degree of anomaly on gel electrophoresis at 4°C with various extents of increase for different sequences. It appears possible for one to conclude that the bending generally increases at lower temperatures with different trends for the various sequences examined. A possible explanation is that the temperature-induced change in superstructure is due to an influence on the DNA twist angle [11,12]. Taking into account these results, namely, the possibility that sequences that are actually bent

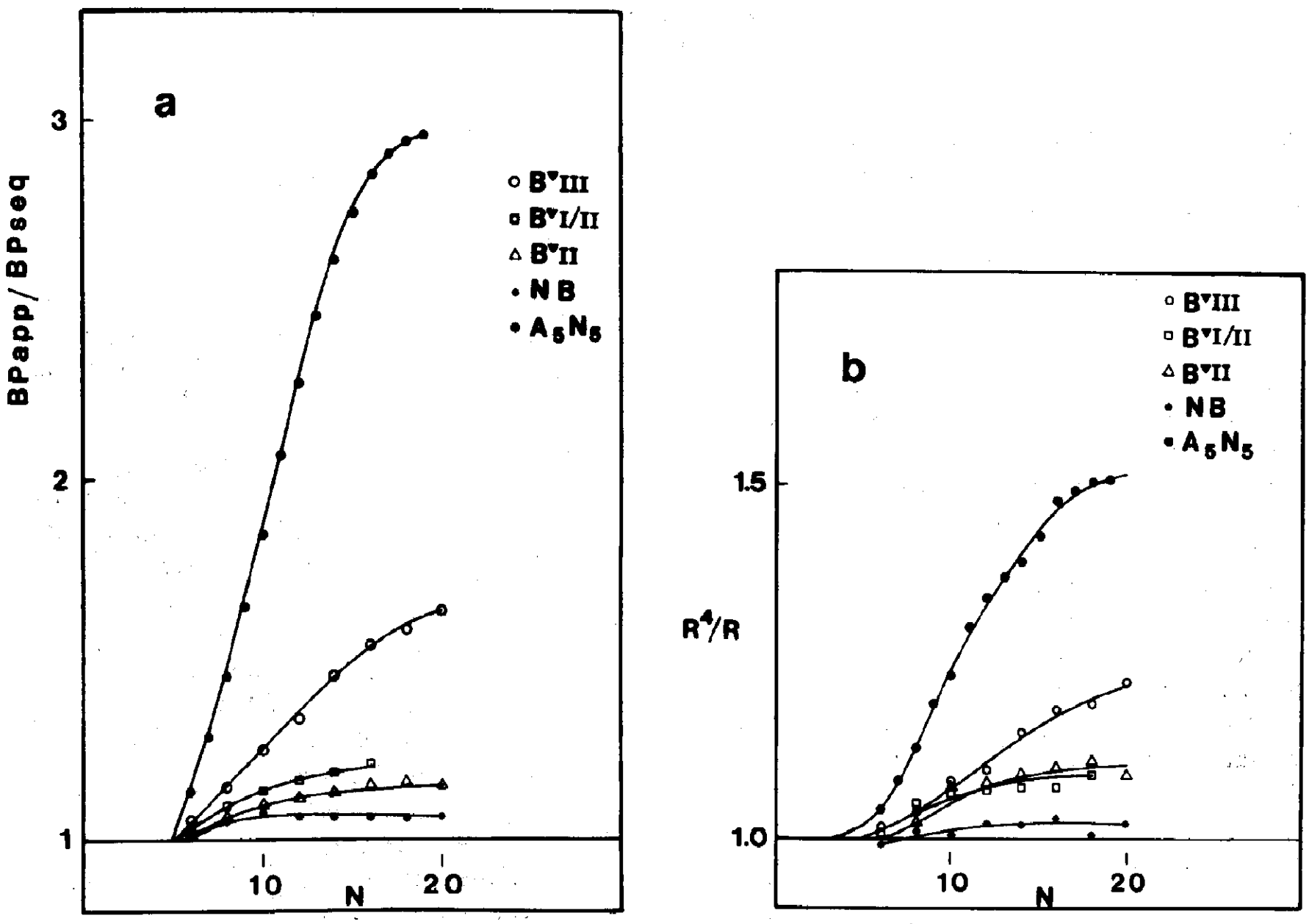


Fig. 7. (a) Effect of lowered temperature on the mobility of multimers derived from 20-mers reported in table 1; (b) relative changes in values of retardation ratios at 4°C with respect to room temperature, shown as the ratio R^4/R.

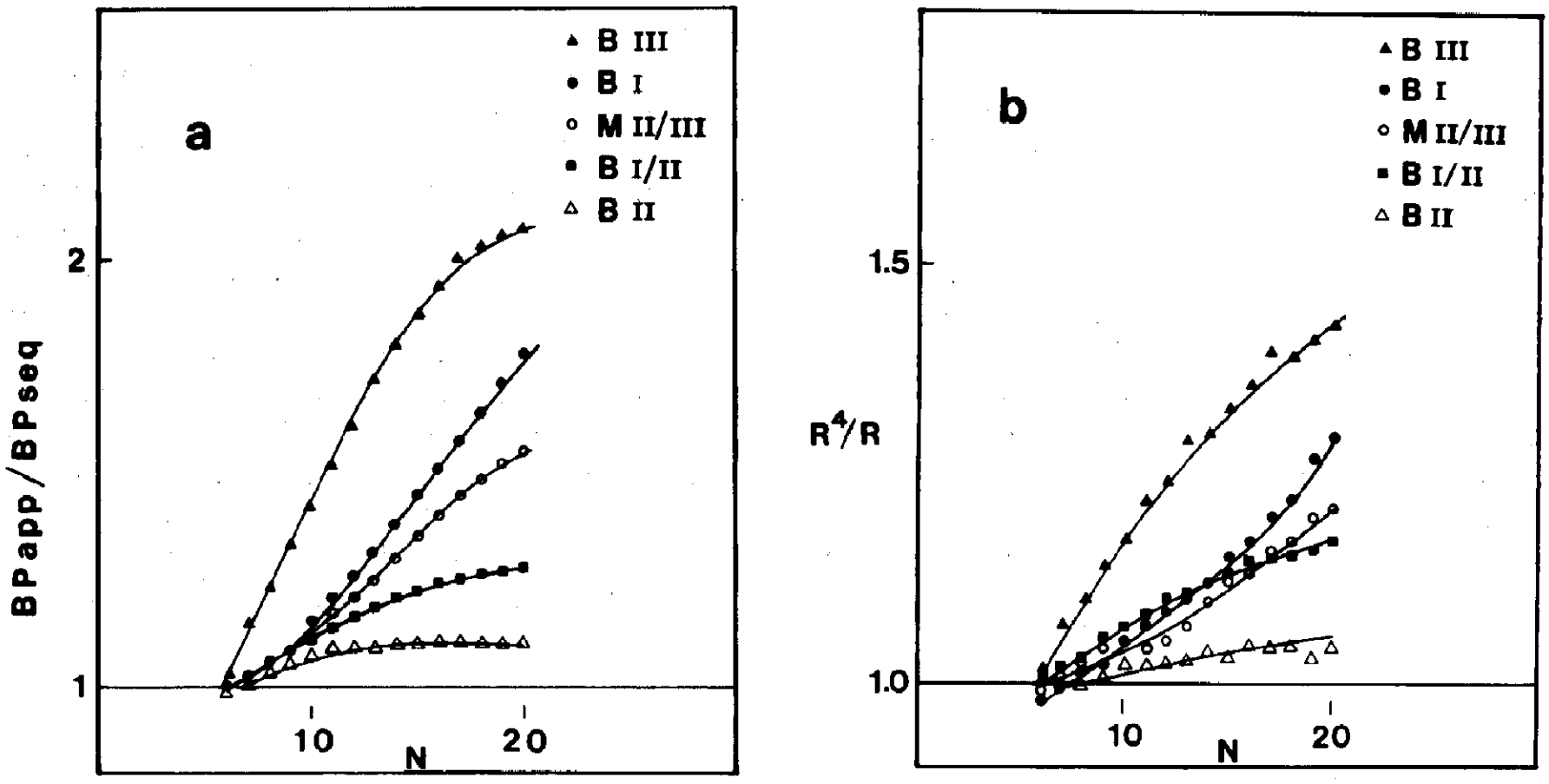


Fig. 8. (a) Effect of lowered temperature on the mobility of multimers derived from 10-mers reported in table 1; (b) relative changes in R^4/R.

increase in curvature at lower temperatures, it can be suggested that specific interactions with proteins could also have the same effect.

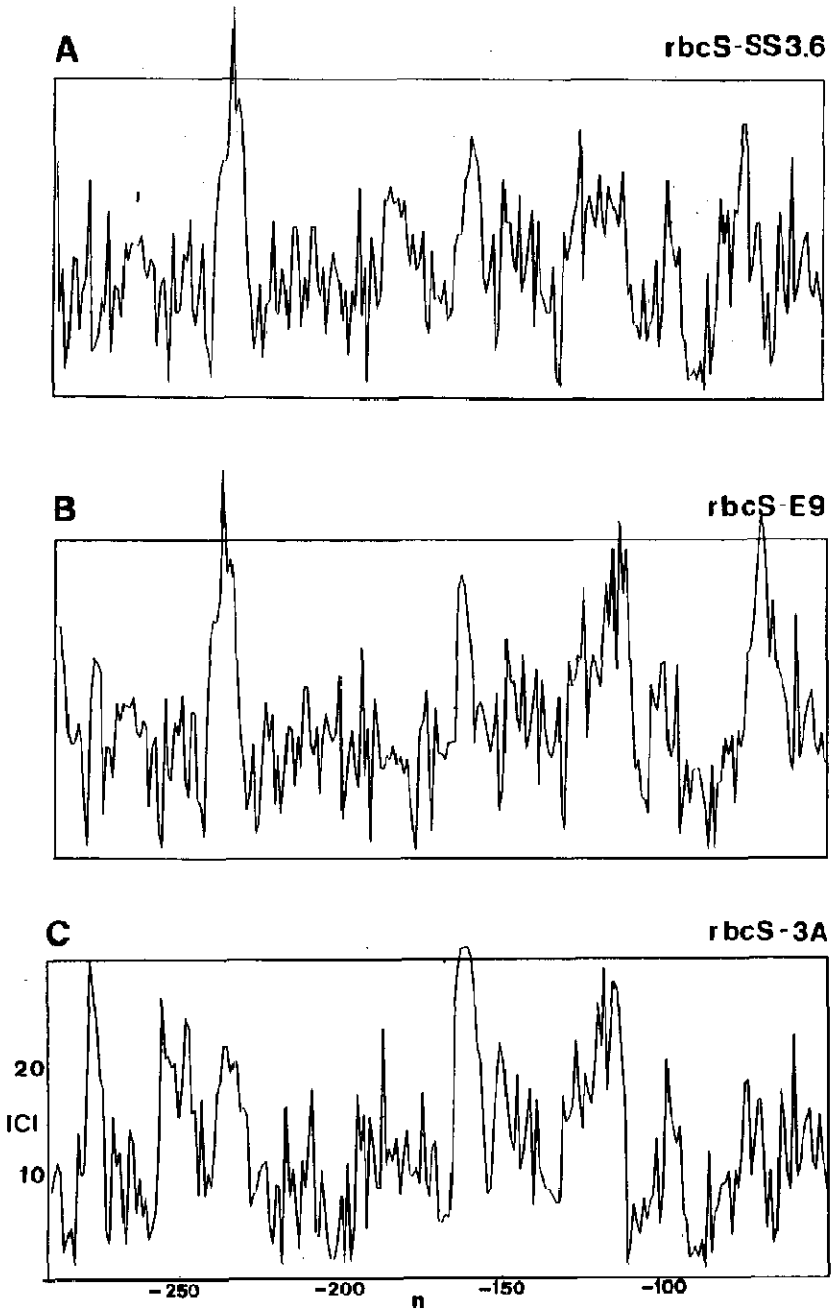
4. Discussion

The present results generally support the theoretical method developed by De Santis et al. [4–7] for prediction of the curvature profile of a DNA fragment from a nucleotide sequence on the basis of energy-conformational calculations. The testing of the most prominent features of the predicted curvature profile by the measurement of a global physicochemical property, such as the electrophoretic mobility of multimers of synthetic monomers, seems rather satisfactory.

It is worth noting that the predicted phase of curvature, for the examined sequences, also shows fairly good agreement with experimental findings. However, we are aware that a larger number of cases must be examined in order to evaluate the general range of applicability for this method of prediction of DNA superstructures. In addition, one should also take into account that to obtain a quantitative correlation between the predicted curvature and that derived from the retardation ratio in the examined sequence, the contribution of the terminal basis to the curvature must be negligible.

In the case of the upstream regulatory region of *rbc*S-3A, we wish to report here an example of the possible applications, which relates to the modulation of genetic expression in the multigenic family of *rbc*S.

The sequences of the DNA fragments that determine the photoactivation of three *rbc*S genes (*rbc*S-3A, *rbc*S-E9 and *rbc*S-SS 3.6) in pea plants have been determined and it has been shown that although the three sequences share a large degree of chemical homology (see fig. 9) the quantity of transcripts displays large variations [13]. Namely, in mature green leaves *rbc*S-3A accounts for about 40% of the total *rbc*S transcript, whereas together *rbc*S-E9 and *rbc*S-SS 3.6 contribute about 7%. The theoretical curvature profiles of the regulatory regions of these three genes are reported in fig. 9. In all cases the profiles show definite peaks, but considering the region between −170 and −110 interesting differences emerge. Namely, in the case



a

b

```
-170

ACACAAAA.. TTTCAAATCT TGTGTGGTTA ATATGGCTGC AAACTTTATC

ACACAAAACT TTTCAA.TCT TGTGTGGTTA ATATGACTGC AAAGTTTATC

ACACACAACT TTTCAA.TCT TGTGTGGTTA ATATGGCTGC AAAGTTTATC

ATTTTCACTA TCTAACAAGA TT    rbcS-3A

ATTTTCACAA TCCAACAA.A CT    rbcs-E9

ATTT.CACAA TCTAACAAGA TT    rbcS-SS 3.6

                    -101
```

Fig. 9. (a) Curvature profiles of (A) *rbc*S-3A, (B) *rbc*S-E9 and (C) *rbc*S-SS 3.6 genes in pea plants. (b) Alignment of the relative nucleotide sequences: dots reflect gaps created in the aligning procedure. The differences with respect to the *rbc*S-3A gene are underlined.

of *rbc*S-3A, two peaks are shown centred on −160 and −110 that correspond to boxes I/II and III derived from biochemical analyses. In the same region the profile of *rbc*S-E9 is similar to that of *rbc*S-3A except for the slightly lower $|C|$ corresponding to BoxI and the shift of about 7 bp of the peak corresponding to BoxIII. As concerns *rbc*S-SS 3.6, the profile shows lower values of the curvature in the positions of BoxI and III, that in this case are scarcely different from the intermediate region. These results show that in the case of a significant degree of chemical homology as well, the curvature profile can be noticeably different, due to a few mutations. In conclusion, it seems tempting to suggest that if the DNA regions examined owe their role to the interaction with a regulatory protein (as appears to be proved by biochemical analyses), the curvature of B-DNA could be the structural determinant for specific protein binding.

Acknowledgements

Thanks are due to P. De Santis and A. Palleschi for many helpful discussions and to A. Tufillaro and R. Gargamelli for technical assistance in experiments and drawings. This research has been financially supported by MPI 60% – Grandi Progetti d'Ateneo, University of Rome I – and by Fondazione 'Istituto Pasteur – Fondazione Cenci-Bolognetti'. S.C. has been supported by an ENI fellowship.

References

1 R. Fluhr, C. Kuhlemeier, F. Nagy and N.-H. Chua, Science 232 (1986) 1106.
2 P.J. Green, S.A. Kay and N.-H. Chua, EMBO J. 6 (1987) 2543.
3 H.-N. Liu-Johnson, M.R. Gartenberg and D.M. Crothers, Cell 47 (1986) 995.
4 P. De Santis, S. Morosetti, A. Palleschi, M. Savino and A. Scipioni, in: Structures and dynamics of nucleic acids, proteins and membranes, eds. E. Clementi and S. Chin (Plenum, New York, 1986) p. 31.
5 P. De Santis, S. Morosetti, A. Palleschi, M. Savino and A. Scipioni, in: Biological and artificial intelligence systems, eds. E. Clementi and S. Chin (ESCOM, Leiden, 1988) p. 143.

6 P. De Santis, S. Morosetti, A. Palleschi, M. Savino and A. Scipioni, in: Proceedings of workshop on molecular structure, dynamics and recognition, eds. N. Niccolai and P. Neri (1989) in the press.
7 P. De Santis, A. Palleschi, M. Savino and A. Scipioni, Biophys. Chem. 32 (1988) 305.
8 H.-S. Koo, H. Wu and D.M. Crothers, Nature 320 (1986) 501.
9 P.J. Hagerman, Nature 321 (1986) 449.
10 J.C. Marini, P.N. Effron, T.C. Goodman, C.K. Singleton, R.D. Wells, R.M. Wartell and P.T. Englund, J. Biol. Chem. 259 (1984) 8974.
11 R.E. Depew and J.C. Wang, Proc. Natl. Acad. Sci. U.S.A. 72 (1975) 4275.
12 S. Diekmann, Nucleic Acids Res. 15 (1987) 247.
13 C. Kuhlemeier, P. Green and N.-H. Chua, Annu. Rev. Plant Physiol. 38 (1987) 221.
14 C. Kuhlemeier, R. Fluhr, P. Green and N.-H. Chua, Genes Dev. 1 (1987) 247.
15 E. Serfling, M. Jasin and W. Schaffner, Trends Genetic 1, (1985) 224.