Correlation of intrinsic DNA curvature with DNA property periodicity

Andrei Gabrielian**, Sándor Pongor*

International Centre for Genetic Engineering and Biotechnology (ICGEB), Area Science Park, Padriciano 99, 34012 Trieste, Italy

Received 27 June 1996; revised version received 19 July 1996

Abstract Twelve di- and trinucleotide parameter sets representing various structural, thermodynamic or bendability-related properties of DNA were tested in the prediction of DNA curvature applying Fourier analysis on curved and straight, A/T-type or G/C-type DNA sequence motifs. The best predictions were obtained with a new consensus bendability scale created by combining a nucleosome-based and a deoxyribonuclease I-based parameter set. Geometry calculations on the same sequences showed that the helical parameters derived from NMR structures can correctly predict curvature, as distinct from the parameters derived from X-ray crystallographic analysis.

Key words: DNA structure; DNA curvature; Sequence periodicity; Fourier analysis

1. Introduction

Sequence-dependent DNA bending plays a crucial role in many biological events such as packaging, transcription, site-specific recombination and DNA replication [1]. It is important therefore to predict both intrinsically curved regions and flexible, easy-to-bend sites in DNA sequences. There are a number of models and algorithms available (see e.g. [2,3] for references) but there is still no complete consensus on either the nature of intrinsic DNA curvature or the sequence-dependent parameters suitable for its prediction.

The central concept of intrinsic DNA curvature is helical phasing [1]: Repeating a uniform structural deformation at about every 10 or 10.5 basepairs in a sequence (i.e. 'in phase' with the helical repeat of B-DNA) can result in a macroscopic curvature of the DNA trajectory. By far the best characterized elements involved in curvature are phased A/T tracts. A/T-tract-based curvature can be correctly predicted by most algorithms [2,4], moreover it can be simply recognized by visual inspection or by pattern recognition methods. Also, it was shown that DNA curvature can be caused by the periodic repeats of other motifs, such as GGGCCC [5], as well as by such elements as nicks, gaps [6] and, to some extent, unpaired loops [7]. These examples lead us to suppose that a helically phased repetition of a certain property (such as those related to stability or flexibility of DNA) and not only a given sequence pattern might be a sufficient basis for intrinsic curvature.

The aim of the present work was to determine the correlation of experimentally known curvature with the periodicity of various sequence-dependent properties of DNA. For this purpose we chose 11 representative sets covering various structural, thermodynamic and bendability-related properties. We compare these parameters on a selected set of 'difficult DNA motifs' [5,8,9] that are known to cause problems in predicting curvature from sequence. We show that helical roll angles deduced from NMR measurements, but not the ones based on X-ray crystallographic data, can correctly predict curvature in these examples, and present a consensus bendability scale which is especially suitable for this purpose.

2. Materials and methods

2.1. Sequence-dependent parameters and calculation of property profiles

Twelve di- and trinucleotide property scales were used (Tables 1 and 2). Six of them are helical roll angles [10–15] that are known to reflect the major deviations of the helix axis in DNA structures from the ideal straight line [3]. Three scales are free energies (ΔG) that characterize the relative stability of B-DNA with respect to the transitions to A, Z or coil forms, respectively [16–18]. The two bendability scales are based on fractional occurrences of trinucleotides in bent or bendable regions, derived from deoxyribonuclease I (DNase I) cutting [19] and nucleosome positioning experiments [2,20], respectively. All scales were taken from the literature and applied in their original form. A consensus bendability scale was calculated as the average of the last two scales.

The property profiles were constructed by dividing the sequence into overlapping di- or trinucleotides and assigning a corresponding scale value from Table 1 or Table 2, respectively. For the uniformity of the further calculations, we constructed test sequences of 210 nucleotides from the short repeat sequences given in Table 3.

2.2. Fourier analysis

For the analysis of periodicity in a DNA parametric profile, we used a sensitive and robust variant of the discrete Fourier transform

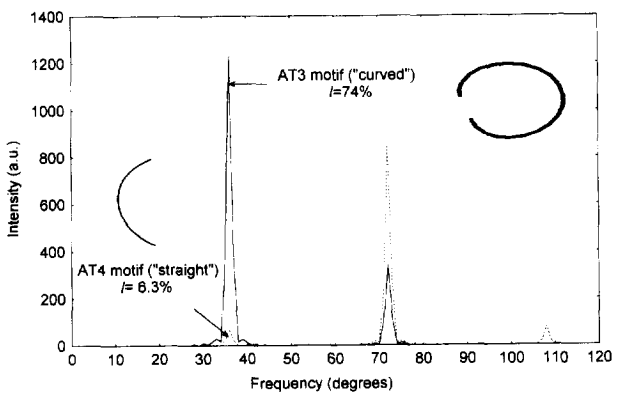

Fig. 1. Examples of Fourier spectra. Fourier spectrum of the sequence AT3 $(AAAATTTTGC)_n$ (continuous line) and of the sequence AT4 $(TTTTAAAAGC)_n$ (dotted line) calculated according to Eq. 1 using the consensus bendability scale (Table 2). The discrimination index D is calculated according to Eq. 3. The sequences are taken from [9]. Representations of the DNA path were calculated from the helical roll, tilt and twist angles of Ulyanov and James [10] using standard matrix transformation technique [14].

*Corresponding author. Fax: (39) (40) 226 555.
E-mail: pongor@icgeb.trieste.it

**Permanent address: Institute of Molecular Biology, Russian Academy of Sciences, Vavilov St. 32, 117984 Moscow, Russian Federation.

0014-5793/96/$12.00 © 1996 Federation of European Biochemical Societies. All rights reserved.
PII S0014-5793(96)00855-1