Available online at www.sciencedirect.com

SciVerse ScienceDirect

Physics of Life Reviews 10 (2013) 41–67

PHYSICS of LIFE reviews

www.elsevier.com/locate/plrev

Review

Sequence-dependent collective properties of DNAs and their role in biological systems ☆

Pasquale De Santis [a,b], Anita Scipioni [a,b,*]

[a] *Dipartimento di Chimica, Università di Roma La Sapienza, P.le A. Moro, 5, I-00185 Roma, Italy*
[b] *Istituto Pasteur – Fondazione Cenci Bolognetti, Italy*

Received 17 January 2013; accepted 22 January 2013

Available online 23 January 2013

Communicated by E. Di Mauro

Abstract

DNA actively interacts with proteins involved in replication, transcription, repair, and regulation processes inside the cell. The base sequence encodes the dynamics of these transformations from the atomic to the nanometre scale length, and over higher spatial scales. In fact, although an important part of the DNA informational content acts locally, it exerts its functions as collective properties of relatively long sequences and manifests as static and dynamic curvature. Physical models that explore different aspects of DNA collective properties associated to such superstructural properties encoded in the sequence will be reviewed. The B-DNA periodicity operates as band-pass-filter; only the local physical–chemical variance associated to the sequence, in phase with the helical periodicity, sums up and reveals at higher scale. In this light, the gel electrophoresis behaviour of DNAs, the nucleosome thermodynamic stability and positioning along genomes were interpreted and discussed. Finally, a part of this review is reserved to describe the ability of some inorganic crystal surfaces to recognize and stabilize certain DNA tracts with peculiar sequences. The collective superstructural properties of DNAs could be involved in the selective interaction between DNA sequence and particular crystal surfaces. It may be conceived that sequences strongly adsorbed on surface could nucleate and expand bits of information in primeval DNA (and/or RNA) chains, early characterized by random sequences, since more protected against the physical–chemical injuries by the environment, and therefore involved in the evolution of their informational content.

Keywords: DNA static curvature; DNA dynamic curvature; Gel electrophoretic retardation; Nucleosome stability; Nucleosome mapping; DNA recognition

1. Introduction

The substantial homogeneity and the consequent conformational degeneracy of the nucleotide residues along DNA double helix have allowed great progress in the knowledge of the molecular mechanisms that control the functional

☆ Authors are available to process any DNA sequences to calculate DNA curvature and flexibility, nucleosome stability and positioning. Requests to A. Scipioni (anita.scipioni@uniroma1.it).

* Corresponding author at: Dipartimento di Chimica, Università di Roma La Sapienza, P.le A. Moro, 5, I-00185 Roma, Italy. Tel.: +39 0649913228; fax: +39 0649913228.

E-mail addresses: pasquale.desantis@uniroma1.it (P. De Santis), anita.scipioni@uniroma1.it (A. Scipioni).

1571-0645/$ – see front matter
http://dx.doi.org/10.1016/j.plrev.2013.01.004

organization of the genome. However, the classic image of DNA, commonly represented as a straight and uniform double helix, is presently just an icon. The DNA chain, considered as a simple repository of the gene information in the past, is a very complex polymorphic macromolecule, which plays a relevant part in the management of the informational content of the nucleotide sequence. Actually, every function of DNA, as transcription, replication and recombination, is ruled by deviations from the monotonous regularity of the canonical B-DNA structure. Inside the cell, DNA actively interacts with the proteins involved in replication, transcription, repair, and regulation processes. During these processes, the DNA transforms between packed and unpacked architectures, like that of chromatin or other higher-order structures morphing into shapes with structural spikes alternative to the canonical B-form in connection with biological events as in replication bubbles, hairpins or in G-quadruplex structures. The base sequence encodes the dynamics of these transformations from the atomic to the nanometre scale length, and over higher spatial scales. In fact, an important part of the DNA informational content is not localized on the codon regions but is related to collective features of relatively long tracts of sequence and manifested as static and dynamic curvature. Such superstructural features are intrinsic properties of the sequence and are recognized and amplified by protein binding. This is particularly evident in the case of sequence dependent DNA–histone octamer association, the nucleosome, which governs the packaging and the architecture of the eukaryotic genome as well as the gene regulation. The topological constraints integrate and amplify at higher scale the local structural changes. In fact, collective properties of DNA are related to the topological invariance of the linking number in actual or virtual circular DNAs. Often, binding or releasing of regulatory proteins changes the local twisting number of DNA with a consequent variation of the writhing number and the global shape of DNA topological domains that assure the topological invariance of the linking number. This interplay transforms local sequence-dependent DNA properties into global collective properties and vice versa. This large-scale structural amplification influences DNA high-scale functions by integrating sequence-dependent curvature effects and local binding of different factors.

We have investigated the effects of the sequence on modelling the superstructural properties of DNA by integrating the theoretically evaluated slight structural [1–3] and electronic features [4] of the different dinucleotide steps along the sequence. The helical periodicity, which characterizes the B-DNA structure, operates as band-pass-filter; only the local perturbations associated to the sequence, in phase with the helical periodicity, sum up and provide effects at higher scale translating the sequence in superstructural features of DNA [1–3,5–13]. The deviations from the intrinsic sequence-dependent superstructure are modelled adopting a first order elasticity approach [14–19].

Such a theoretical model allows predicting in excellent agreement with the experiments (i) the electrophoretic manifestations of the DNA curvature of synthetic and biologically relevant DNA tracts [14], (ii) the thermodynamic constants of the sequence-dependent circularization reactions of DNA tracts [15], (iii) the sequence dependent writhing transitions from relaxed to supercoiled circular forms also when proteins-induced bending or unwinding are involved [16], (iv) the thermodynamics stability constants of the histone octamer associated to various DNA tracts, different for length and sequence [17–19]. Importantly, the quantitative evaluation of nucleosome stability provides a physical support for predicting the nucleosome positioning along eukaryotic genomes [20].

The collective properties of DNAs could be involved in the evolution of their informational content based on prebiotic chemical processes as suggested by the ability of inorganic surfaces to recognize and stabilize DNA tracts with peculiar sequence [4,21]. In fact, we suggested that the selective interaction between DNA sequence and some inorganic crystal surfaces could nucleate and expand bits of information in primeval DNA (and/or RNA) chains, early characterized by random sequences. In fact, sequences strongly adsorbed on surface could be more protected against the physical–chemical injuries by the environment and therefore selected as “DNA phenotypes”. The sequence gaps originated by the cleaving of less protected tracts could be repaired following two-dimensional diffusion on the inorganic surface of remaining tracts.

In this paper we describe the physical models that explore different aspects of DNA collective properties associated to relatively long tracts of sequence as resulting from effective integration of local contributions.

1.1. DNA curvature

The tendency of some dinucleotide steps to be repeated along the eukaryotic genomes with DNA helical periodicity was first evidenced by Trifonov and Sussman [5]. Such a feature was interpreted as a manifestation of sequence-dependent deformational anisotropy of the chromatin DNA, which facilitates its smooth folding in the nucleosomes. In fact, the perfect base pairing of the DNA double helix restricts the structural variance in the biological conditions to