Advances in Colloid and Interface Science 232 (2016) 49–56

Contents lists available at ScienceDirect

Advances in Colloid and Interface Science

journal homepage: www.elsevier.com/locate/cis

Historical perspective

Nucleic acid polymeric properties and electrostatics: Directly comparing theory and simulation with experiment

Adelene Y.L. Sim

*Bioinformatics Institute (A*STAR), 30 Biopolis Street #07-01, Matrix, Singapore 138671*

ARTICLE INFO

Available online 9 October 2015

Keywords:
Nucleic acids
Polymer theory
Polyelectrolytes
Ion atmosphere
Poisson–Boltzmann theory
Benchmarking

ABSTRACT

Nucleic acids are biopolymers that carry genetic information and are also involved in various gene regulation functions such as gene silencing and protein translation. Because of their negatively charged backbones, nucleic acids are polyelectrolytes. To adequately understand nucleic acid folding and function, we need to properly describe its i) polymer/polyelectrolyte properties and ii) associating ion atmosphere. While various theories and simulation models have been developed to describe nucleic acids and the ions around them, many of these theories/simulations have not been well evaluated due to complexities in comparison with experiment. In this review, I discuss some recent experiments that have been strategically designed for straightforward comparison with theories and simulation models. Such data serve as excellent benchmarks to identify limitations in prevailing theories and simulation parameters.

Contents

1. Introduction . . . 49
2. Nucleic acids as polymers . . . 50
   2.1. Polymer models applied to nucleic acids . . . 50
   2.2. Biological relevance of different length scales . . . 50
3. Nucleic acids are polyelectrolytes . . . 50
   3.1. Ion atmosphere around nucleic acids . . . 50
   3.2. Poisson–Boltzmann theory . . . 51
   3.3. Beyond Poisson–Boltzmann theory . . . 51
4. Need for well-designed experiments to benchmark theory/simulation . . . 51
   4.1. Double-stranded DNA polymer properties . . . 51
   4.2. Single-stranded nucleic acid behavior . . . 52
   4.3. Quantifying the ion atmosphere . . . 53
   4.4. Simple hairpin pulling . . . 53
5. Conclusions and outlook . . . 54
Acknowledgements . . . 54
References . . . 54

1. Introduction

Nucleic acids (NAs) play fundamental roles in biology. Deoxyribonucleic acid (DNA) stores the genetic code that is passed through generations, while ribonucleic acid (RNA) actively participates in many gene regulation mechanisms [1], some of which we are only just beginning to uncover.

NAs have to fold and pack into specific conformations in order to function properly. DNA famously folds into a B-form double helix [2] held together by base-pairing (hydrogen-bonding) and base-stacking (aromatic ring stacking) interactions. The double helix further coils around histones to form a tightly packed nucleosome core particle [3]. Such a highly ordered structure allows a large amount of DNA to be stored within a small volume. Tight genome packaging (both in the form of DNA and RNA) also occurs in viruses [4]. RNA often exists as a single chain, but like DNA, it also forms base-pairing interactions to fold into helices (albeit RNA folds into A-form helices) that in turn

E-mail address: adelenes@bii.a-star.edu.sg.

http://dx.doi.org/10.1016/j.cis.2015.09.008
0001-8686/

coaxially stack and pack closely relative to each other [5–7]. These specific three-dimensional tertiary folds render RNA the ability to carry out ligand sensing [8,9], catalysis [10–12], translation [13,14], and other gene-regulating [15–17] roles.

2. Nucleic acids as polymers

2.1. Polymer models applied to nucleic acids

NA folding and packing depends on its ability to bend and distort in the presence and absence of external activation; this in turn has significant implications on NA biology such as nucleosome positioning [18]. Because the structure and flexibility of an NA molecule are affected by its bases that preferentially stack and base-pair, general models for polymers might not be appropriate for NAs. For example, even though the wormlike chain [19] and self-avoiding random walk [20] models have been applied to long double-stranded (ds) DNA (see also [21] and references therein), a recent study suggests that dsDNA, even at more than 48 kilo base-pairs, does not exhibit ideal polymer scaling behavior [22]. Experimental studies of long single-stranded (ss) DNA also indicate the need to go beyond standard polymer models [23–26], although the freely jointed chain theory appeared to previously match single molecule pulling experiments [19].

Even if polymer theories appropriately describe ds- and ssNA properties, naturally occurring NA actually does not always exist as purely ds or ss; that is, an NA could be heterogeneous in its physical properties. NA (in particular RNA) preferentially folds upon itself to form helical regions connected by ss portions (also known as junctions); this branched structure (an RNA usually comprises multiple junctions) dictates NA conformation and dynamics [27–29], but is not well described by simple homogeneous polymer theory.

2.2. Biological relevance of different length scales

An additional complicating factor is that NA structure is hierarchical: nucleotides form base-pairs that in turn stack neatly to form helices; these helices pack relative to each other and/or around neighboring proteins to form tertiary and/or quaternary structures. Depending on the desired structural information, the length scale of the problem varies, often falling within the finite-length regime. For instance, protein-DNA interactions typically occur around length scales less than 50 base-pairs [30]. This means that the typical infinite length assumption in polymer theory and/or simplification of NA structure (e.g. DNA helices as rigid cylinders) is not always applicable. In fact, as will be discussed subsequently, at such small finite-lengths (less than about 100 base-pairs), dsDNA appears to deviate from the WLC model and shows substantially more flexibility than expected [31–35], although this observation of increased flexibility seems to be somewhat controversial [36–38].

Short ssDNA strands also show sizable deviation from standard polymer models [26]. In particular, ssDNA exhibits noticeable sequence-specific effects that are typically not accounted for in standard polymer theory; some sequence effects have to be additionally incorporated (e.g. [39]). Simplifications of ssDNA as "beads-on-a-string" also appeared to be inadequate in describing ssDNA scaling properties (Fig. 1A–C, adapted from reference [40]), suggesting that sequence and geometry might need to be considered when developing a more robust polymer model suitable for short ssDNA.

3. Nucleic acids are polyelectrolytes

3.1. Ion atmosphere around nucleic acids

Besides intrinsic molecular properties like bond length, angles, and torsional rigidity, NA's physical polymeric properties are confounded by its intrinsic charge. NAs are naturally occurring highly charged

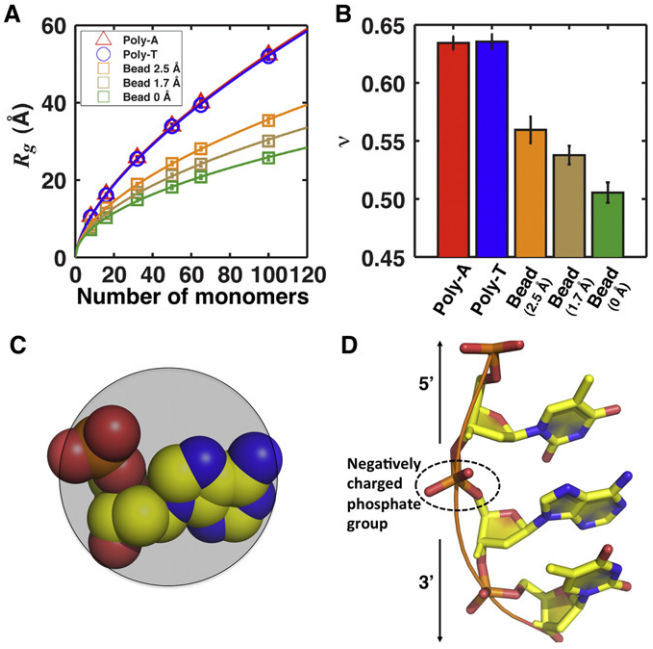

Fig. 1. (A) Simulated variation in radius of gyration (R_g) with number of monomers (N). The goal was to identify the scaling law at small length scales, so only excluded volume was considered. Poly-deoxyadenine (poly-A) and poly-deoxythymine (poly-T) clearly behave differently from beads-on-a-string (three different bead radii of 0, 1.7 Å, and 2.5 Å were used). (B) The scaling exponents ν derived from (A), where $R_g \propto N^{\nu}$. Because of the size and shape of the nucleotides, the beads-on-a-string model does not adequately describe poly-A and poly-T. (C) The geometry of adenine is not well represented by a single spherical bead. (D) Nucleic acids have negatively charged phosphate groups in their backbones. Molecular images were made using PyMOL [131].

polyelectrolytes because each phosphate group on the backbone of a nucleotide carries a single negative charge (Fig. 1D). Packing of NAs into confined spaces is highly energetically unfavorable: the electrostatic repulsion of folding a ~400 nucleotide RNA is about 600 kcal/mol in the absence of neutralizing counterions [41,42]! Therefore, understanding NA structure, folding and compaction require elucidation of how neutralizing counterions affect NA polymer properties. (This is much less critical in the case of proteins since they are weakly charged – if at all – in comparison.) That is, we need to understand the polyelectrolyte properties of NA.

Beyond comprehending polymer/polyelectrolyte properties of NA, we need to quantify how ions associate around NA, because they implicate NA folding kinetics and energetics [43,44]. Ions orient around NA structures, forming an ion atmosphere [42,45] rich in positively charged counterions but excluded of negatively charged coions [46] (Fig. 2). Ions

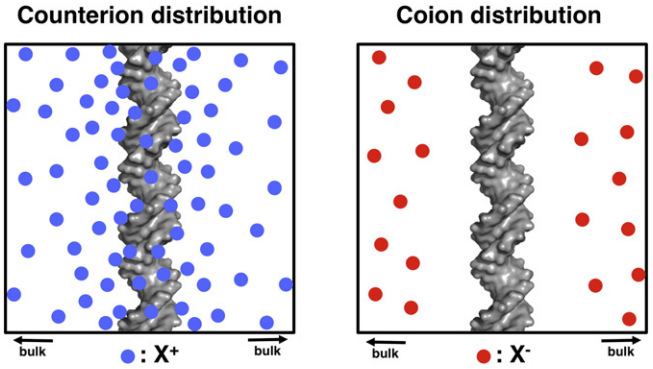

Fig. 2. Schematic distribution of counterions and coions in the ion atmosphere around B-form DNA. There is an excess (depletion) of counterions (coions) close to the DNA compared to in bulk.