Can we model DNA at the mesoscale? Comment on: Fluctuations in the DNA double helix: A critical review

Michel Peyrard, Thierry Dauxois

Can we model DNA at the mesoscale ?
Comment on: *Fluctuations in the DNA double helix: A critical review.*

Michel Peyrard[1] and Thierry Dauxois[1]
[1] *Ecole Normale Supérieure de Lyon, Laboratoire de Physique CNRS, 46 allée d'Italie, 69364 Lyon Cedex 7, France*
(Dated: April 8, 2015)

This review includes a discussion on the theoretical approaches intending to provide a quantitative analysis of DNA fluctuations, but, as this topic is very broad, it concentrates on two particular cases: the two-state Poland-Scheraga model [3] and the Peyrard-Bishop-Dauxois (PBD) mesoscopic approach which describes the fluctuations of a base pair in terms of a single variable [4], the stretching of the pair. The main point of the authors, emphasized in the conclusion, is that any mesoscopic model does not make sense, and as molecular dynamics cannot reach the time scales needed to investigate an event so rare as base-pair opening, only the two-state model remains. We think that such a sharp judgment can be questioned.

Discarding all molecular dynamics studies is certainly unjustified. It is true that the simulations face difficulties owing to the very large number of atoms that have to be studied, even for a short DNA fragment, due to the solvent, and owing to the gap in time scales between small vibrational motions and the full opening of a base pair. However, molecular dynamics is more than brute force calculations of time trajectories. It can be used for a clever sampling of the phase space allowing a calculation of the free energy pathways for base pair openings, taking into account the solvent and counterions [5], giving in turn precious information for modeling DNA at an intermediate scale.

In their assessment of mesoscopic models, the authors of the review adopt the same kind of narrow view, which leads them to two fundamental misunderstandings: first, as in the case of molecular dynamics, the review puts emphasis on the dynamics, which is not at all the best way to derive useful results from a model, and second the authors do not seem to realize that, in a mesoscopic model, the "potentials" are actually containing the missing microscopic degrees of freedom in some effective way so that a statement such as "the model considers DNA as if it where in vacuum" is completely wrong. Let us comment on these two points.

Dynamics versus statistical physics. In contrast to the two-state model, for a mesoscopic model such as PBD it is easy to write equations of motions. This does not mean that they have to be used to derive the most useful results that can be obtained with such a model. Very few experiments are actually able to access time-dependent properties. Most of them probe equilibrium properties, such as the average opening at a given temperature (i.e. the melting curve) or local equilibrium probabilities. Even the studies discussed in the review, which determined the time scale of the opening of the base pairs, were actually based on chemical equilibria. The theory of these measurements is obtained much more accurately with statistical physics than by extracting probabilities from dynamical calculations. For rare events a dynamical approach can get particularly unreliable [6]. For the PBD model, statistical physics calculations can be made very efficiently because the model does not include long–range interactions, in contrast to the entropic contribution of the two-state model. Provided the sequence is properly included in the parameters, i.e. one goes beyond the uniform model described in the review, accurate melting profiles for long sequences can easily be deduced from a statistical physics calculations [7] whereas getting them from dynamical simulations involving several tens of thousands of base pairs would be hopeless.

Potentials in a mesoscopic model. A mesoscopic model attempts to describe the properties of a system by considering some degrees of freedom only. This does not mean that the other degrees of freedom are totally ignored. Even molecular dynamics at the atomic scale is a "mesoscopic" model because it studies the positions of the atoms without specifying the state of the electrons as an ab-initio simulation would do. Nobody would think that molecular dynamics describes nuclei without electrons! Indeed the potentials implicitly depends on the electronic degrees of freedom. The same is true, at a larger scale, in a model of DNA which only describes the stretching of the base pairs. The other degrees of freedom, including the effect of the solvent or ions, enter in the "potential" parameters. Actually the word "potentials" for the functions of the simplified Hamiltonian is not appropriate. One should have in mind free energies, or what is called "potential of mean force" in some molecular dynamics approaches [8]. This is why for instance the ionicity of the solution is taken into account when the PBD model is used to quantitatively describe DNA melting curves [7]. Therefore a mesoscopic model certainly does not describe "DNA in vacuum", although good effective potentials may be difficult to obtain. Basic considerations can be used to evaluate some general features, such as the dissociation energy of the base pairs in the presence of the counterions, but a systematic elimination of the microscopic degrees of freedom is a hard task. This explains why the model parameters have evolved from the first comparison with experiments for short DNA chains [9], to more recent applications to long complex sequences [7].

As effective potentials are complex, their choice may depend on the problem of interest. To fit a melting curve, an appropriate description of the energy levels for the closed and open states and of the entropic effects, introduced through a nonlinear stacking term, is sufficient. If one intends to use the model for dynamical properties such as the evaluation of the open–state lifetime, the effective potential determination is more demanding. As pointed in the review, it is true that the potentials used in the original version of the PBD model lead to completely incorrect results for the lifetime of the open state if the model is used in dynamical simulations. They have been modified for that purpose [10]. Nevertheless it is important to realize that small amplitude motions of DNA *are not overdamped* otherwise one would not observe the Raman peaks characteristic of the vibration of the molecule. This is because

water does not easily penetrate inside the base pair stack although it may form hydrogen bonds in the groves. Only when the amplitude of the fluctuations becomes so large that the bases get fully embedded in water, does the damping increase strongly, as stated in the review. However other phenomena are even more important to increase the lifetime of the open state: on the one hand out-of-the-stack bases have more degrees of freedom than stacked bases, i.e. more entropy and, on the other hand, they can form hydrogen bonds with the solvent. Both effects lower the free energy of the open state, leading to an effective barrier for reclosing. This barrier does not significantly change the statistical properties of the model although it makes the denaturation transition a bit sharper [10], but it drastically changes its dynamics. When it is included the lifetime of the open state increases by several orders of magnitude so that it approaches the observed values [10]. Further improvements would still be needed for a quantitative description of DNA dynamics, in particular the introduction of an extended Langevin simulation method taking into account the variation of the damping with the amplitude of the fluctuations.

In addition to modeling, this review on DNA fluctuations also discusses experiments. Although our comment focuses on modeling, we would nevertheless like to point out that experimental studies of DNA fluctuations can still reveal unexpected results as shown by a recent investigation using Guanine radical chemistry that demonstrated that the large fluctuations of an AT-rich region may affect the configurations of bases up to about 10 base pairs away [1] even at physiological temperatures. This should not be a surprise because the experiments, discussed in the review, which confirmed the very low opening probability of base pairs [2], also show that the activation enthalpy for the opening, which stays roughly constant in the $0 - 25°$ C range, changes significantly above 25° C. A straightforward extrapolation of the data of Fig. 12 of [2b] suggests an increase of the opening rate by a factor of 10 between 25° C and 40° C. As a result the role of DNA fluctuations could even be more important than suggested in the review, particularly when specific sequences are concerned.

The main message that we would like to convey in this comment is that there is not one good model and one bad model for DNA. The reality is more complex and depends on the questions of interest. As pointed out in the review, the two-state model is successful and its numerous empirical parameters are now well under control, so that it is widely used. A mesoscopic model, such as the PBD model, attempts to build the description from a microscopic view by lumping many degrees of freedom in a few variables at a larger scale. The interest is that, at least in principle, the parameters of its effective potentials can be quantitatively derived, although it is not a trivial task. Results from an exploration of the phase space by molecular dynamics could be used [5]. When statistical physics is used, both the two-state and the PBD model give similar melting curves for complex DNA sequences. However when one is interested in some peculiar properties that depend on the magnitude of the fluctuations, a model going beyond a two-state picture is required. This is for instance the case in the analysis of the temperature dependence of the broadening of some Bragg peaks in neutron diffraction [11]. The width of the peak associated to the base stacking reflects the correlation length of the double helix. It increases sharply at the denaturation transition, when the stack breaks into short segments, destroying spatial coherence. However prior to the full denaturation, the width of the peak also depends on the magnitude of the transverse fluctuations of the base pairs, which can be estimated from a statistical physics study of the PBD model [11].

The PBD model poses interesting questions in nonlinear dynamics, and, as noticed in the review, it is rather easy to handle by mathematicians or physicists. This is why it had many "followers" (to use the word of the review) and some of them extrapolated to biology conclusions drawn in conditions which are not relevant for DNA (no thermal bath, inadequate parameters). This has led to statements such as the prediction of solitons moving in DNA, which are easy to criticize. However, exactly as one would not reject cars because bad drivers have killed people with them, one should not discard the modeling of DNA at a mesoscale. It provides further insights and it is likely that the future lies in a combination of models at various scales, as done recently in an hybrid approach which mixed molecular dynamics and mesoscale modeling [12] to study DNA flexibility.

---

[1] Cuesta-López S, Menoni H, Angelov D, Peyrard M. Guanine radical chemistry reveals the effect of thermal fluctuations in gene promoter regions. Nucl. Acids Res. 2011; 39:5276-5283.

[2] Gueron M, Kochoyan M, Leroy J-L. A single mode of DNA base-pair opening drives imino proton exchange. Nature 1987; 328, 89-92 ; Leroy J-L, Kochoyan M, Huynh-Dinh T, Gueron M. Characterization of Base-pair Opening in Deoxynucleotide Duplexes using Catalyzed Exchange of the Immino Proton. J. Mol. Biol. 1988; 200:223-238.

[3] D. Poland D, Scheraga HA. Phase transitions in one-dimension and the Helix-Coil Transition in Polyamino-Acids. J. Chem. Phys. 1966; 45:1456-1463; Poland D, Scheraga HA. Occurence of a phase transition in nucleic acid models. J. Chem. Phys. 1966; 45:1464-1469.

[4] Peyrard M, Bishop AR. Statistical Mechanics of a Nonlinear Model for DNA Denaturation. Phys. Rev. Lett. 1989; 62:2755-2758 ; Dauxois T, Peyrard M, Bishop AR. Entropy driven DNA denaturation. Phys. Rev. E 1993; 47:R44-R47.
[5] Giudice E, Várnai P, Lavery R. Base pair opening within B-DNA: free energy pathways for GC and AT pairs from umbrella sampling simulations. Nucleic Acid Research 2003; 31:1434-1443.
[6] van Erp TS, Cuesta-Lopez S, Hagmann J-G, Peyrard M. Can one predict DNA Transcription Start Sites by studying bubbles? Phys. Rev. Lett. 2005; 95:218104-1-4
[7] Theodorakopoulos N. Melting of genomic DNA: Predictive modeling by nonlinear lattice dynamics. Phys. Rev. E 2010; 82:021905-1-4.
[8] Garcia AE, Soumpasis DM. Harmonic vibrations and thermodynamic stability of a DNA oligomer in monovalent salt solution. Proc. Natl. Acad. Sci. USA 1989; 86:3160-3164.
[9] Campa A, Giansanti A. Experimental tests of the Peyrard-Bishop model applied to the melting of very short DNA chains. Phys. Rev. E 1998; 58:3585-3588.
[10] Peyrard M, S. Cuesta López S, James G. Nonlinear analysis of the dynamics of DNA breathing. J. Biol. Phys. 2009; 35:73-89
[11] Wildes A, Theodorakopoulos N, Valle-Orero J, Cuesta-López S, Garden J-L, Peyrard M. Structural correlations and melting of B-DNA fibers. Phys. Rev. E 2011; 83:061923-1-11.
[12] Meyer S, Jost D, Theodorakopoulos N, Peyrard M, Lavery R, Everaers R. Temperature Dependence of the DNA Double Helix at the Nanoscale: Structure, Elasticity, and Fluctuations. Biophys. J. 2013; 105:1904-1914.