

Three-dimensional models for homogeneous DNA near denaturation

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2005 J. Phys.: Condens. Matter 17 7755

(<http://iopscience.iop.org/0953-8984/17/50/001>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.252.86.83

The article was downloaded on 28/05/2010 at 07:05

Please note that [terms and conditions apply](#).

Three-dimensional models for homogeneous DNA near denaturation

Gabriel F Calvo¹ and Ramón F Alvarez-Estrada²

¹ Grup de Física Teòrica, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

² Departamento de Física Teórica I, Facultad de Ciencias Físicas, Universidad Complutense de Madrid, 28040, Madrid, Spain

Received 8 July 2005

Published 2 December 2005

Online at stacks.iop.org/JPhysCM/17/7755

Abstract

Motivated by the melting transition of DNA, we study genuinely three-dimensional models for two interacting open, flexible and homogeneous macromolecular chains, bound or unbound to each other, at thermal equilibrium from about room temperature up to about the denaturation temperature (T_{un}). In each chain, angular constraints on bond angles (due to covalent bonding) determine monomers: each monomer contains n_e nucleotides and has an effective length d_e . These monomers could remain practically unaltered for temperatures in a range above and below T_{un} , down to 300 K. Estimates for n_e and d_e are provided and justified. Upon proceeding from Quantum Mechanics to the classical limit and using suitable large-distance approximations (partly, due to those monomer configurations), we get a generalization of Edwards' model, which includes effective potentials between monomers. The classical partition function for the two-chain system is reduced to an integral of a generalized and discretized two-chain Green's function. We analyse conditions for the denaturing transition. The fact that each single chain is an extended one-dimensional system modifies their mutual global interaction, in comparison with typical potentials between nucleotides: this is simply illustrated by computing a global effective potential between the two chains. Applications for Morse potentials are presented. Our models seem to be physically compatible with some previous one-dimensional ones and could allow us to efficiently extend the latter to three spatial dimensions.

1. Introduction

DNA constitutes an essential piece of our understanding about what life is [1]. Indeed, one of its more fascinating features is its ability to encode the information for the synthesis of amino acids (that build proteins) in a remarkably compact fashion. It is a macromolecule consisting of two helical chains or strands (in a doubled helix or plectonemic structure) formed by sequences

of nucleotides. Each nucleotide is a compound of three elements: a phosphate group, a sugar ring and one of four possible bases: adenine (A) and guanine (G) (purine bases), cytosine (C) and thymine (T) (pyrimidine bases). In order to comprehend why these two chains give rise to a double helix, it is necessary to analyse the interactions between the nucleotides. Along each chain, nucleotides interact via the phosphate groups and the sugar rings (forming fairly rigid covalent bonds). Rotations around the bonds are possible within each strand, but the separation between consecutive attachment points of the bases and the angles between neighbouring bonds are, nevertheless, well defined. Additional interactions between the bases of successive nucleotides also contribute to the stacking along the axis of the helix. Besides these intra-chains interactions there exist inter-chain interactions that are also responsible for the plectonemic topology. The two strands are linked owing to the tendency of bases to assemble in pairs through hydrogen bonds obeying Chargaff's rule: A–T which are linked by two hydrogen bonds and G–C which are linked by three hydrogen bonds (being therefore more robust) [2]. The energies involved in the hydrogen bonds are, typically, at least one order of magnitude smaller than those of covalent bonds.

Of primary importance for biological activity is the denaturation (or melting, see [3–22] and references therein) of DNA, that is, the separation of the two double helix strands (by thermally induced breaking of the hydrogen bonds). Denaturation is a necessary step prior to processes such as replication and transcription. During denaturation, the stability of the DNA double helix depends on the relative content of G–C and A–T pairs: the larger the fraction of G–C pairs, the higher the required melting temperature T_{un} (which is of the order of $T_{\text{un}} \simeq 360$ K). Less intuitive is that, although, in the cell nucleus, enzymes contribute to help separating the two strands, there is evidence [18, 23] that specific nucleotide sequences may favour the formation of bubbles (denaturated localized sites). We shall concentrate exclusively on thermal denaturation. The latter is a purely physical process, without direct biological reference *a priori*. However, its study provides, at least, a quite useful validity test for DNA modelling. We emphasize that, wherever we refer to denaturation throughout this article, we shall mean, really, thermal denaturation.

In order to acquire an approximate picture of the above class of processes that occur in DNA or more generic macromolecules it seems natural, at least in principle, to start from a quantum mechanical approach (see, for instance, [24–26]). The nucleotides (or, more generally, the atoms) are subject to dominant effective potentials. Vibrational degrees of freedom vary rather rapidly and should be treated quantum-mechanically. For the temperatures involved ($300 \text{ K} \leq T \leq 360 \text{ K}$, and somewhat above) and for the physically meaningful states, relative distances (bond lengths) from any nucleotide to its nearest neighbours along each chain are approximately constant [27, 28], and in many cases those quantized vibrations appear to decouple approximately from slower individual internal rotations about bonds [3, 24–26]. Rotational configurations of nucleotides may be, typically, in an appreciable number of excited rotational states. For T 's about or somewhat above room temperature, the slower rotational degrees of freedom can be approximated, in many cases, through classical statistical mechanics. That pattern has yielded effective quantum and classical Hamiltonians and partition functions, depending only on the relevant slowly varying degrees of freedom [29, 30].

The present work reports non-trivial generalizations of those models for a system of two interacting open macromolecular chains (in which the basic constituents are nucleotides) in three-dimensional space for the above-mentioned temperatures. Our models will also incorporate, approximately, angular constraints on bond angles. At a later stage, we concentrate on T 's about the melting temperature T_{un} , in which a bound double-stranded macromolecular chain breaks into the two single chains, while each of the latter retains its basic structure. Several physically important effects (interactions in each single chain contributing to produce

helical structures, interactions among nucleotides which are not in similar positions in both chains, etc) will be treated below in a simplified way. Other important effects (inhomogeneities in the chains, the detailed structure of nucleotides, influence of sugars, solvent, etc) will be disregarded. To grasp the relevance of those and other effects, see [4, 5, 14, 31, 32]. Regarding excluded-volume effects in DNA, they have been considered in [12, 33–37], while arguments for disregarding them in certain cases and for assessing them in others have also been given [4]. Our framework allows, in principle, for long-range interactions.

The contents of this work are as follows. Section 2 presents succinctly an approach to the two macromolecular chain system, and a classical partition function $Z_{C,e}$, which, in turn, is reduced to a simpler one, Z_R . In section 3: (a) a lower bound $Z_{R,1}$ is obtained for Z_R , (b) $Z_{R,1}$ (through approximations like angular constraints, averages over suitable monomers and Gaussian approximations) is expressed in terms of a formal Green's function, depending on suitable effective potentials between monomers: all that generalizes Edwards' model [31, 32, 38]. Conditions for the existence of a (double-stranded) bound structure and estimates for T_{un} , both with continuum approximations and in the discretized case, are also discussed in section 3. Section 4 introduces a global effective potential between both chains. Section 5 contains the conclusions and discussions. The work is organized so that sections 2–5 could be read, in essence, independently from derivations, which are given, in outline, in suitable appendices. Thus, appendix A collects some useful properties for Morse potentials [39]. Appendix B outlines, starting from Quantum Mechanics and generalizing non-trivially [29, 30], the derivation of $Z_{C,e}$ and Z_R . Appendices C and D summarize approximate computations regarding $Z_{C,e}$ and useful averages over monomers (incorporating constraints on bond angles). Based upon appendices B–D, the approximations leading from $Z_{R,1}$ to the generalized Edwards model of section 3 are outlined in appendix E. Appendix F provides the approximations leading to the estimates of T_{un} for the discretized case in section 3.

2. Interaction potentials and partition function

Our approach starts by considering a system of two open macromolecular chains in thermodynamical equilibrium at absolute temperature T (from about 300 K up to about the denaturation temperature $T_{un} \simeq 360$ K), based upon Quantum Mechanics, in three spatial dimensions. Each chain is formed by a large number N of nucleotides (about 10^{10} for DNA). The mass and the position vector of the i th nucleotide ($1 \leq i \leq N$) along the r th chain ($r = 1, 2$) are $M_i^{(r)}$ and $\mathbf{R}_i^{(r)}$. Let the overall centre of mass (CM) of the two-chain system and that of the r th chain be given by the vectors \mathbf{R}_{CM} and $\mathbf{R}_{CM}^{(r)}$, respectively [30]. The relative position vector between both $\mathbf{R}_{CM}^{(r)}$ is $\mathbf{y} = \mathbf{R}_{CM}^{(2)} - \mathbf{R}_{CM}^{(1)}$. The relative ('bond') vectors along the r th chain are $\mathbf{y}_j^{(r)} = \mathbf{R}_{j+1}^{(r)} - \mathbf{R}_j^{(r)}$, $j = 1, \dots, N - 1$. For given r , one has

$$\mathbf{R}_i^{(r)} = \mathbf{R}_{CM}^{(r)} + \sum_{j=1}^{N-1} \tilde{\alpha}_{i,j}^{(r)} \mathbf{y}_j^{(r)}, \quad (1)$$

where $(\sum_{s=1}^N M_s^{(r)}) \tilde{\alpha}_{i,j}^{(r)} = \sum_{h=1}^j M_h^{(r)}$ and $-\sum_{h=j+1}^N M_h^{(r)}$, for $j = 1, \dots, i - 1$ and $j = i, \dots, N - 1$, respectively. We suppose that the region in which both chains move is a sphere of suitably large radius R_0 . By assumption, electronic degrees of freedom have already been dealt with from the outset (say, *à la* Born–Oppenheimer), so that the constituents (nucleotides) of one macromolecular chain are subject to dominant effective potentials (disregarding non-dominant contributions like effective potential vectors [40, 41]). A qualitatively adequate effective interaction between two nucleotides is the Morse potential $V_M = V_M(y)$, where $y (\geq 0)$ denotes the nucleotide separation,

$$V_M(y) = D \{ \exp[-2\alpha(y - d)] - 2 \exp[-\alpha(y - d)] \}. \quad (2)$$

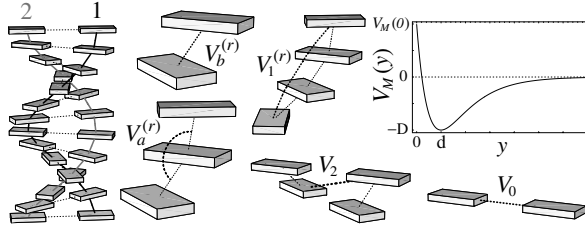


Figure 1. Left: block-model of the DNA double helix structure. Blocks represent nucleotides. Possible interactions between nucleotides considered in the present work are schematically displayed (dark dotted lines connect the interacting nucleotides) with their corresponding potentials (see the text). The right-upper inset plots the dependence of the Morse potential with the nucleotide separation.

D is the dissociation energy corresponding to the equilibrium distance d and α^{-1} represents the range of the potential (see figure 1 for the potential profile). Some useful results for $V_M(y)$ are summarized in appendix A.

The interaction between the two macromolecular chains is described in terms of an effective quantum Hamiltonian operator, \tilde{H}_Q , whose precise form is provided in appendix B. \tilde{H}_Q includes the total potential $\sum_{r=1}^2 V_b^{(r)} + \sum_{r=1}^2 V_a^{(r)} + V_I$, which depends on relative distances and displays overall rotational invariance. In short, the interactions $V_b^{(r)}$ and $V_a^{(r)}$ continue to be important in a range of temperatures above denaturation, while the effect of V_I becomes negligible just above melting. We shall characterize all of them below.

2.1. Intra-chain potentials

The potentials $V_b^{(r)}$ and $V_a^{(r)}$ describe strong covalent bonds and give rise, respectively, to the r th macromolecular chain (say, to its primary structure). In practice, $V_b^{(r)}$ and $V_a^{(r)}$ essentially yield fixed bond lengths and constrained bond angles, respectively. Recall that in the denaturation of DNA (in the range of T s considered here) no covalent bonds are broken [2]. We accept that the magnitude of $|V_b^{(r)}|$ is larger than that of $|V_a^{(r)}|$ which, in turn, is larger than that of $|V_I|$ (to be characterized later). The geometrical meaning of these potentials is depicted in figure 1.

Let $y_j^{(r)} = |\mathbf{y}_j^{(r)}|$. In what follows, $V_b^{(r)}$ will consist of a sum of Morse potentials (2) as

$$V_b^{(r)} = \sum_{j=1}^{N-1} V_{M,j}^{(r)}(y_j^{(r)}), \quad (3)$$

with corresponding equilibrium distances $d_j^{(r)}$ (for the j th bond length), and parameters $D_{0,j}^{(r)}$ and $\alpha_j^{(r)}$. It will be convenient to introduce the vibrational frequencies $\omega_{0,j}^{(r)} = \alpha_j^{(r)} \sqrt{2D_{0,j}^{(r)} A_{jj}^{(r)}}$, where the inverse-mass constants $A_{jj'}^{(r)}$ ($j, j' = 1, \dots, N-1$) are defined by

$$A_{jj'}^{(r)} = \begin{cases} \frac{1}{M_j^{(r)}} + \frac{1}{M_{j+1}^{(r)}} & j = j', \\ -\frac{1}{M_j^{(r)}} & j' = j \pm 1, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Also, $\mathbf{y}_j^{(r)} = y_j^{(r)} \mathbf{u}_j^{(r)}$, with $\mathbf{u}_j^{(r)} = (\cos \varphi_j^{(r)} \sin \theta_j^{(r)}, \sin \varphi_j^{(r)} \sin \theta_j^{(r)}, \cos \theta_j^{(r)})$, in spherical coordinates. We denote the sets $(\theta_1^{(r)}, \dots, \theta_{N-1}^{(r)})$ and $(\varphi_1^{(r)}, \dots, \varphi_{N-1}^{(r)})$, with $r = 1, 2$, by θ and φ , respectively.

Potential $V_a^{(r)}$ describes the angular constraints

$$V_a^{(r)} = \sum_{j=1}^{N-2} v_{j,j+1}^{(r)} (|\mathbf{y}_j^{(r)} + \mathbf{y}_{j+1}^{(r)}|). \quad (5)$$

We will assume that the nearest-neighbour potentials $v_{j,j+1}^{(r)} (|d_j^{(r)} \mathbf{u}_j^{(r)} + d_{j+1}^{(r)} \mathbf{u}_{j+1}^{(r)}|)$ (to be discussed in appendix B) have a unique deep minimum, thus constraining $\mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)}$ to some fixed value $\beta_j^{(r,0)}$.

We suppose that all $D_{0,j}^{(r)}$ are of similar orders of magnitude and that the same holds for all $\omega_{0,j}^{(r)}$ and for all $v_{j,j+1;\max}^{(r)} (= \max |v_{j,j+1}^{(r)}|)$. We assume that $D_{0,j}^{(r)}$ are larger than all $v_{j,j+1;\max}^{(r)}$ and that all $v_{j,j+1;\max}^{(r)}$ are larger than all $\hbar\omega_{0,j}^{(r)}$. Moreover, we focus on the case where $k_B T$ is less than or, at most, of the order of $\hbar\omega_{0,j}^{(r)}$ and

$$\hbar\omega_{0,j}^{(r)} \gg \frac{\hbar^2 A_{jj}^{(r)}}{(d_j^{(r)})^2}. \quad (6)$$

Typical values employed in various analysis of macromolecules and of DNA are consistent with the above assumptions for $T \simeq 300$ K and somewhat above [3, 5, 14, 42]. For instance (with $0.6 \text{ kcal mol}^{-1} \simeq 0.025 \text{ eV}$, corresponding to $T \simeq 300$ K), $D_{0,j}^{(r)}$ about $100 \text{ kcal mol}^{-1}$ (or larger), $d_j^{(r)} \simeq 1$ to a few \AA , $\alpha_j^{(r)} \simeq$ a few \AA^{-1} , $\hbar\omega_{0,j}^{(r)}$ about $3\text{--}10 \text{ kcal mol}^{-1}$, and for atomic masses typical in organic chemistry [14]. On the other hand, all $v_{j,j+1;\max}^{(r)}$ are supposed to be somewhere between 12 and $100 \text{ kcal mol}^{-1}$ [42].

2.2. Inter-chain potentials and residual interactions

The potential V_I describes residual (weaker) interactions among nucleotides in each chain (not included in $V_b^{(r)} + V_a^{(r)}$) and the binding between both chains. Having this in mind, we decompose it as

$$V_I = V_0 + \sum_{r=1}^2 V_1^{(r)} + V_2, \quad (7)$$

where

$$V_0 = \sum_{i=1}^N v_{0,i} (|\mathbf{R}_i^{(1)} - \mathbf{R}_i^{(2)}|), \quad (8)$$

$$V_1^{(r)} = \sum_{j=1}^{N-1-\xi'} \sum_{\xi \geq 2}^{\xi'} v_{1;j,\xi}^{(r)} \left(\left| \sum_{j'=j}^{j+\xi} \mathbf{y}_{j'}^{(r)} \right| \right), \quad (9)$$

$$V_2 = \sum_{j \neq j'} v_{2;j',j} (|\mathbf{R}_{j'}^{(1)} - \mathbf{R}_j^{(2)}|). \quad (10)$$

The summations in (10) run over j and j' , with $j \neq j'$, $|j - j'| \leq \xi''$. ξ' and ξ'' are integers. V_0 (V_2) is the total potential due to nucleotides which have complementary (non-complementary) locations in both chains: both V_0 and V_2 vanish very quickly as both chains separate. $\sum_{r=1}^2 V_1^{(r)} + V_2$ may be comparable to (but, eventually, somewhat smaller than) V_0 . Potential V_0 (accounting for hydrogen bonds) plays a very important role in giving rise to the bound double-stranded structure: each $v_{0,i} (|\mathbf{R}_i^{(1)} - \mathbf{R}_i^{(2)}|)$ may be supposed to have, qualitatively at least, the form of another Morse potential (with some parameter D very different

from those involved in $V_b^{(r)}$ and $V_a^{(r)}$. The residual potential $V_1^{(r)}$ among nucleotides in the r th chain, which is not included in $V_a^{(r)} + V_b^{(r)}$, contributes to produce additional angular constraints and further structures in the r th chain. Regarding V_2 , it will also give rise to the latter effects and play a role in the binding between both single chains. One may consider that the terms $\sum_{r=1}^2 V_1^{(r)} + V_2$ will account for stacking and, thus, to the stability of secondary (helical-like) structures and, eventually, also for some longer-range (e.g. excluded-volume-like and supercoiling) effects. It seems more convenient to treat $v_{1;j,\xi}^{(r)}$, $v_{2;j,j'}^{(r)}$ and $v_{0;i}$ on a similar footing. Our subsequent analysis will hold for rather general V_I . The geometrical interpretation of potentials V_0 , V_1 , and V_2 is schematically shown in figure 1.

We shall employ ($q = |\mathbf{q}|$)

$$v_{0;i}(|\mathbf{y}|) = \int \frac{d^3\mathbf{q}}{(2\pi)^3} \tilde{v}_{0;i}(q) \exp[i\mathbf{q} \cdot \mathbf{y}]. \quad (11)$$

Although \mathbf{q} should be regarded as a discretized wavevector as the volume is not infinite ($R_0 < +\infty$), no essential error is made if we treat the former as a continuous variable. And so on for $v_{1;j,\xi}^{(r)}$ and $v_{2;j,j'}^{(r)}$, with Fourier transforms $\tilde{v}_{1;j,\xi}(q)$ and $\tilde{v}_{2;j,j'}(q)$ respectively. It is interesting to compare the (three-dimensional) structure in (7) with those of the (essentially one-dimensional) interaction potentials employed in [7, 16–18]. In short, V_0 in (8) corresponds to V in [7], while $\sum_{r=1}^2 V_1^{(r)} + V_2$ in (9) and (10) corresponds to the additional stacking interaction between adjacent base pairs, as envisaged in [16–18] (in addition to the interaction potential V from [7]): further qualitative comparison between V_0 and $\sum_{r=1}^2 V_1^{(r)} + V_2$ with [16–18] will be made in sections 3.4 and 5.

We suppose that all $v_{0;i;\max}$ ($=\max|v_{0;i}(|\mathbf{y}|)|$), $v_{1;j,\xi;\max}$ ($=\max|v_{1;j,\xi}(|\mathbf{y}|)|$) and $v_{2;j,j';\max}$ ($=\max|v_{2;j,j'}(|\mathbf{y}|)|$) are smaller than any $v_{j,j+1;\max}^{(r)}$. The effect of $v_{1;j,\xi}^{(r)}$ becomes quite small near the denaturation transition and above. The possibility that $\hbar\omega_{0,j}^{(r)}$ be about (even perhaps a bit smaller than) $v_{0;l;\max}$ has been entertained in [14], but the case $\hbar\omega_{0,j}^{(r)} > v_{0;l;\max}$ has also been considered [13, 43, 44]. With typical values employed [3, 5, 14, 42], $v_{0;i;\max}$ may be, on the average, about 4–12 kcal mol⁻¹ (consistent with [14]) ($v_{1;j,\xi;\max}$ and $v_{2;2;j,j';\max}$ may be similar). However, smaller values for $v_{0;i;\max}$ (about 1.2 or 1.8 kcal mol⁻¹) have also been used in recent analyses [16–18]. The inverse ranges of $v_{0;i}$, $v_{2;j,j'}$ and $v_{1;j,\xi}$ may be expected to be about a few Å⁻¹, in agreement with [5, 7, 13, 16–18]. The above general properties of V_0 , $\sum_{r=1}^2 V_1^{(r)}$ and V_2 will suffice to obtain effective approximations for all of them.

2.3. Classical partition function

Using the above assumptions and generalizing non-trivially the results of [29, 30], a new variational computation (outlined in appendix B) leads to a model for a three-dimensional microscopic system formed by two open macromolecular chains, characterized by an effective quantum partition function $Z_{Q,e}$, in terms of slowly varying degrees of freedom. As a consequence, each single macromolecular chain behaves as an extended and flexible bound structure (with fixed bond lengths $d_j^{(r)}$ and angles $\beta_j^{(r,0)}$), which is a thermodynamical system in thermal equilibrium at T . As T is not much larger than 300 K and as $D_{0,j}^{(r)}$ and $v_{j,j+1;\max}^{(r)}$ are larger than $k_B T$ (so that the strong covalent bonding remains practically unaffected [2]), each single chain does not break itself into its nucleotide constituents but it retains its individuality as an extended bound system. On the other hand, the two macromolecular chains, interacting with each other in three-dimensional space, could be either forming a two-chain bound state (dominating as $T < T_{\text{un}}$), or an unbound one (dominating as $T > T_{\text{un}}$), with equation (6) being

always respected. Next, we turn to the classical limit assuming also

$$k_B T \gg \frac{\hbar^2 A_{jj}^{(r)}}{(d_j^{(r)})^2}, \tag{12}$$

which holds for T about 300 K (and somewhat above). Then, quantum operators and statistics can be approximated by classical ones. In the classical limit, $Z_{Q,e}$ becomes the effective (dimensionless) classical partition function for the two-chain system, $Z_{C,e}$ (see appendix B). We also regard \mathbf{y} and its associated momentum as classical variables, since each single chain is a very extended and massive object. We define $\mathbf{d}\Omega = \prod_{r=1}^2 \mathbf{d}\Omega^{(r)}$, with $\mathbf{d}\Omega^{(r)} \equiv \prod_{j=1}^{N-1} d\varphi_j^{(r)} d\theta_j^{(r)} \sin\theta_j^{(r)}$. Let $A^{(r)}$ denote the tridiagonal matrix of order $(N - 1) \times (N - 1)$ with elements $A_{ij}^{(r)}$ and determinant $\det A^{(r)}$. After some calculations (see appendix B for details), $Z_{C,e}$ becomes $[M_{\text{red}}^{-1} = \sum_{r=1}^2 (\sum_{i=1}^N M_i^{(r)})^{-1}]$

$$Z_{C,e} = \left[\frac{k_B T}{2\pi\hbar^2} \right]^N \left[\frac{M_{\text{red}} k_B T}{2\pi\hbar^2} \right]^{3/2} \left\{ \prod_{r=1}^2 \frac{[\prod_{j=1}^{N-1} (d_j^{(r)})^2]}{(\det A^{(r)})^{3/2}} \right\} Z_R, \tag{13}$$

with

$$Z_R = \prod_{r=1}^2 \int d^3\mathbf{y} \int \mathbf{d}\Omega \frac{F^{(r)}}{(\Delta^{(r)})^{1/2}} \exp\left[-\frac{V_I(\mathbf{y}; \theta, \varphi)}{k_B T}\right], \tag{14}$$

and $F^{(r)} \equiv \prod_{j=1}^{N-2} F_j^{(r)} (|d_j^{(r)} \mathbf{u}_j^{(r)} + d_{j+1}^{(r)} \mathbf{u}_{j+1}^{(r)}| - d_{j,j+1}^{(r)})$. The function $F_j^{(r)}$, arising from $V_a^{(r)}$, is strongly peaked at $|d_j^{(r)} \mathbf{u}_j^{(r)} + d_{j+1}^{(r)} \mathbf{u}_{j+1}^{(r)}| \simeq d_{j,j+1}^{(r)}$, where $(d_{j,j+1}^{(r)})^2 = (d_j^{(r)})^2 + (d_{j+1}^{(r)})^2 + 2d_j^{(r)} d_{j+1}^{(r)} \beta_j^{(r,0)}$. Let $\beta_j^{(r,0)}$ be close to +1, which is approximately the case for DNA. Two forms for $F_j^{(r)}$ are given in appendix B, depending on whether the degrees of freedom involved in $V_a^{(r)}$ are treated classically or quantum-mechanically. It will be gratifying to see that both treatments of $V_a^{(r)}$ and, so, both $F_j^{(r)}$ s lead essentially to the same model in section 3, for the T s considered here. $V_I(\mathbf{y}; \theta, \varphi)$ denotes the restriction of V_I when $y_j^{(r)} = d_j^{(r)}$, for any $r = 1, 2$ and $j = 1, \dots, N - 1$. If $V_I(\mathbf{y}; \theta, \varphi) \equiv 0$, $Z_R = [\prod_{r=1}^2 Z_R^{(r)}] [\int d^3\mathbf{y}]$, with $Z_R^{(r)} \equiv \int \mathbf{d}\Omega^{(r)} [\Delta^{(r)}]^{-1/2} F^{(r)}$ (reduced single-chain partition function). $\Delta^{(r)}$ is a determinant: a study of $[\Delta^{(r)}]^{-1/2}$ for large N for each open single chain and some approximations for $Z_R^{(r)}$ are summarized in appendix C. Equations (13) and (14) constitute the main result of this section (and of the computations in appendix B). Thus far, no long-distance approximations have been performed.

In what follows, we shall restrict ourselves to homogeneous chains (that is, we suppose that all $d_j^{(r)} = d$, $M_i^{(r)} = M_0$, $d_{j,j+1}^{(r)} = d_{1,2}$, $\beta_j^{(r,0)} \simeq \beta^{(0)}$ for any r and j , and that $v_{0;i} = v_0$, $v_{1;l,\xi}^{(r)} = v_{1;\xi}$ and $v_{2;j,j'} = v_{2;|j-j'|}$, for any j, j'), and we shall consider the case of B-DNA. For B-DNA, the helix pitch is about 34 Å, the transverse diameter of the double helix (say, some average distance between both chains in the bound double-stranded structure) is about 20 Å and the rotation angle from one nucleotide to the next in each strand is about 36° (10 nucleotides per turn or pitch), so $\beta_j^{(r,0)} \simeq \beta^{(0)} \simeq 0.8$ [3]. Then, we estimate $d \simeq 7.2 \text{ \AA} (\simeq [3.4^2 + 2 \times 10^2 (1 - \beta_j^{(r,0)})]^{1/2})$ for B-DNA. Let us discuss the homogeneity assumptions. Those for $d_j^{(r)}$ and $\beta_j^{(r,0)}$ seem reasonable. The loss of accuracy implied by replacing the nucleotide masses $M_i^{(r)}$ by some average M_0 will be, more or less, similar to those for other approximations to be carried out later (thus, leaving aside phosphate groups and sugar rings, we simply notice that the ratios of the masses of A, C, G and T over the hydrogen mass

are, roughly, 135, 111, 151 and 114, respectively). The homogeneity assumptions regarding potentials may imply a larger loss of accuracy, since, for instance, the potential energy for the G–C interaction is appreciably larger than the A–T one [14, 18].

3. Generalized Edwards model

In order to arrive at the class of models for analysing the denaturation transition, further approximations will be required, successively, for Z_R (equation (14)). The first approximation will be based upon the introduction of suitable monomers and a thermodynamical inequality, and it will be provided in section 3.1. The subsequent approximations involve certain lengthy calculations: estimation of determinants; treatment of angular constraints and averages over monomers (described in appendices C and D), and Gaussian long-distance approximations (in appendix E). After all these approximations, the main result will be given in equations (19)–(21) and discussed in section 3.2. The remaining sections 3.3 and 3.4 will be devoted to discussions, estimates and applications to the denaturation transition.

3.1. Monomers and thermodynamical inequality

In order to formulate the first approximation, it will be convenient to introduce, first, suitable monomers. Let n_e be an integer, about some tens (say, the number of nucleotides in a few turns of the double helix in DNA). For the r th chain, we shall introduce the monomer vectors $\mathbf{z}_l^{(r)}$, $l = 1, \dots, L$ with $L = (N - 1)/n_e$ (N and L being large): $\mathbf{z}_l^{(r)}$ is associated to the vector $d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)}$ from the $n_e(l - 1) + 1$ nucleotide to the $n_e l + 1$ one, and the set of all $\mathbf{z}_l^{(r)}$ extends along the whole r th chain. We choose n_e as the smallest integer such that, with respect to the statistical average defined by $Z_R^{(r)}$, all $d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)}$, as l varies, be statistically independent from one another (for given r), at least approximately: as discussed in appendix D, $20 \leq n_e \leq 40$ may be not unreasonable. We denote the set of all $\mathbf{z}_l^{(r)}$, $l = 1, \dots, L$, by $[\mathbf{z}^{(r)}]$ and the set formed by $[\mathbf{z}^{(1)}]$ and $[\mathbf{z}^{(2)}]$ by $[\mathbf{z}]$. We shall define, for any function $f = f(\theta, \varphi)$, the partial average

$$\langle f \rangle(\mathbf{y}; [\mathbf{z}]) \equiv \left[\prod_{r=1}^2 Z_R^{(r)}([\mathbf{z}^{(r)})] \right]^{-1} \int d\Omega \left[\prod_{r=1}^2 [\Delta^{(r)}]^{-1/2} F^{(r)} \right] f(\mathbf{y}; \theta, \varphi) \\ \times \prod_{r=1}^2 \prod_{l=1}^L \delta^{(3)} \left(\mathbf{z}_l^{(r)} - d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)} \right). \quad (15)$$

The new $Z_R^{(r)}([\mathbf{z}^{(r)})]$ is defined as

$$Z_R^{(r)}([\mathbf{z}^{(r)})] = \int d\Omega^{(r)} [\Delta^{(r)}]^{-1/2} F^{(r)} \prod_{l=1}^L \delta^{(3)} \left(\mathbf{z}_l^{(r)} - d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)} \right), \quad (16)$$

where $\delta^{(3)}$ denotes the three-dimensional Dirac delta function. Then, equation (14) is recast as

$$Z_R = \int d^3 \mathbf{y} \left[\prod_{r=1}^2 \prod_{l=1}^L d^3 \mathbf{z}_l^{(r)} Z_R^{(r)}([\mathbf{z}^{(r)})] \right] \left\langle \exp \left[-\frac{V_I}{k_B T} \right] \right\rangle(\mathbf{y}; [\mathbf{z}]). \quad (17)$$

At this stage, we employ, for any given \mathbf{y} and $\mathbf{z}^{(r)}$, the inequality [45] $\langle \exp[-(k_B T)^{-1} V_I] \rangle(\mathbf{y}; [\mathbf{z}]) \geq \exp[-(k_B T)^{-1} \langle V_I \rangle(\mathbf{y}; [\mathbf{z}])]$, where $\langle V_I \rangle(\mathbf{y}; [\mathbf{z}])$ is given through

equation (15), with $f = V_I$. Consequently

$$Z_R \geq \int d^3\mathbf{y} \left[\prod_{r=1}^2 \prod_{l=1}^L d^3\mathbf{z}_l^{(r)} Z_R^{(r)}([\mathbf{z}^{(r)})] \right] \exp\left[-\frac{\langle V_I \rangle(\mathbf{y}; [\mathbf{z}])}{k_B T}\right] \equiv Z_{R,1}. \quad (18)$$

The bound (18) embodies the main result of this subsection.

3.2. Two-chain Green's function

Let d_e be an effective length, which takes into account the net effect of constraints on bond angles, as implied by equation (5) for the r th chain: with DNA values, we take $d_e \simeq 2.2 n_e^{1/2}$ nm. A justification of this formula, based upon estimating determinants and the effect of angular constraints, is given in appendices C and D (see, in particular, equation (D.1)). For $20 \leq n_e \leq 40$, it follows $10 \leq d_e \leq 14$ nm. Approximations for $Z_R^{(r)}([\mathbf{z}^{(r)})]$ and $\langle V_I \rangle(\mathbf{y}; [\mathbf{z}])$ (in equation (18)) are outlined in appendices D and E (based upon monomer averaging and Gaussian approximations). They allow us to approximate $Z_{R,1}$ (equation (18)) as

$$Z_{R,1} \simeq Z_{R,1,ap} = \left(\frac{4\pi R_0^3}{3}\right)^{-1} \left[\prod_{r=1}^2 Z_R^{(r)} \right] \int \left[\prod_{r=1}^2 d^3\mathbf{R}_{L+1}^{(r')} d^3\mathbf{R}_1^{(r')} \right] G(L). \quad (19)$$

One has $\mathbf{z}_l^{(r)} = \mathbf{R}_{l+1}^{(r')} - \mathbf{R}_l^{(r')}$, so the vectors $\mathbf{R}_l^{(r')}$ and $\mathbf{R}_{l+1}^{(r')}$ are associated to the origin and to the end of the l th monomer. $G(L) \equiv G(\mathbf{R}_{L+1}^{(1)'}, \mathbf{R}_{L+1}^{(2)'}, \mathbf{R}_1^{(1)'}, \mathbf{R}_1^{(2)'}; L)$ (with dimension $length^{-6}$) is the two-chain Green's function for the (discretized) double chain with a discrete 'time' (l) and three effective interactions. $G(L)$ will allow for global effective interactions among monomers. To fix the ideas, we shall restrict $G(L)$ to a somewhat simplified, but far from trivial, case in which effective interactions occur only between complementary monomers in both chains and among nearest-neighbouring monomers (so that ξ' in (9) and $|j - j'|$ in (10) are limited adequately). Let $W_G(\mathbf{z}_l^{(r)}; \delta l) = [3/(2\pi d_e^2 \delta l)]^{3/2} \exp[-3(\mathbf{z}_l^{(r)})^2/(2d_e^2 \delta l)]$, ($\delta l = 1$) denote the Gaussian distribution for the l th monomer. Then, $G(L)$ is

$$G(L) = \prod_{r=1}^2 \prod_{l'=2}^L \prod_{l=1}^L \prod_{l''=1}^L \int d^3\mathbf{R}_{l'}^{(r')} W_G(\mathbf{R}_{l+1}^{(r')} - \mathbf{R}_l^{(r)'}; \delta l) [1 + \sigma_{2;l}(T)] \\ \times \exp\left\{-\frac{1}{k_B T} \sum_{r=1}^2 \left[v_{1;0,e}^{(r)}(0) + v_{1;1,e}^{(r)}(|\mathbf{z}_l^{(r)}|) \right] \right\} f_{ad}, \quad (20)$$

where $\sigma_{2;l}(T) = \sigma_{2;l}(T; \mathbf{R}_{l+1}^{(1)'}, \mathbf{R}_l^{(1)'}, \mathbf{R}_{l+1}^{(2)'}, \mathbf{R}_l^{(2)'})$ is defined through

$$1 + \sigma_{2;l}(T) = \exp\left\{-\frac{1}{k_B T} \left[\frac{1}{2} \sum_{n=0}^1 v_{0,e}(|\mathbf{R}_{l+n}^{(1)' } - \mathbf{R}_{l+n}^{(2)' }|) \right. \right. \\ \left. \left. + v_{2;1,e}(|\mathbf{R}_{l+1}^{(1)' } - \mathbf{R}_l^{(2)' }|) + v_{2;1,e}(|\mathbf{R}_{l+1}^{(2)' } - \mathbf{R}_l^{(1)' }|) \right] \right\}. \quad (21)$$

One has $f_{ad} = \exp[-(2k_B T)^{-1}(v_{0,e}(|\mathbf{R}_{L+1}^{(1)' } - \mathbf{R}_{L+1}^{(2)' }|) + v_{0,e}(|\mathbf{R}_1^{(1)' } - \mathbf{R}_1^{(2)' }|))]$. The effective interactions are: (a) for two (complementary) monomers at the same positions in the different chains $v_{0,e}(|\mathbf{R}_{l+n}^{(1)' } - \mathbf{R}_{l+n}^{(2)' }|)$ (equation (E.5)); (b) $v_{1;0,e}^{(r)}(0)$ for one monomer with itself, while $v_{1;1,e}^{(r)}(|\mathbf{z}_l^{(r)}|)$ for two neighbour monomers in the same (r)th chain (equation (E.7)); (c) for two (non-complementary) monomers at different positions in the different chains $v_{2;1,e}(|\mathbf{R}_{l+1}^{(1)' } - \mathbf{R}_l^{(2)' }|)$ and $v_{2;1,e}(|\mathbf{R}_{l+1}^{(2)' } - \mathbf{R}_l^{(1)' }|)$ (equation (E.8)). In short, the product of W_G s in equation (20) is an approximation for the product of $Z_R^{(r)}([\mathbf{z}^{(r)})]$ s in equation (18). Similarly, $\exp[-\frac{\langle V_I \rangle(\mathbf{y}; [\mathbf{z}])}{k_B T}]$

in equation (18) has been approximated by $\exp\{-(k_B T)^{-1} \sum_{r=1}^2 [v_{1;0,e}^{(r)}(0) + v_{1;1,e}^{(r)}(|\mathbf{z}_l^{(r)}|)]\} f_{ad}$ times the product of $(1 + \sigma_{2;l}(T))$ s in equation (20). The derivation of the approximate forms of $Z_{R,1,ap}$ and $G(L)$ (which follow from equations (E.5) and (E.7)–(E.9)) are valid, at least, for length scales larger than or about d_e . Equation (19) generalizes Edwards' model [31, 32, 38, 46]. Since $v_{0,e}$, $v_{1;n,e}^{(r)}$ and $v_{2;|l-l',e}$ are translationally invariant, centre-of-mass dependences factor out in $Z_{R,1,ap}$ in general, and they cancel out $(4\pi R_0^3/3)^{-1}$. The reliability of employing the Gaussian approximations in equation (20) for scales shorter than d_e has been discussed in appendix E.

Our $v_{0,e}(|\mathbf{R}_{l+n}^{(1)'} - \mathbf{R}_{l+n}^{(2)'}|)$ can be considered as a three-dimensional counterpart of the one-dimensional interaction potential V considered in [7]. Also, $v_{1;1,e}^{(r)}(|\mathbf{z}_l^{(r)}|)$, $v_{2;1,e}(|\mathbf{R}_{l+1}^{(1)'} - \mathbf{R}_l^{(2)'}|)$ and $v_{2;1,e}(|\mathbf{R}_{l+1}^{(2)'} - \mathbf{R}_l^{(1)'}|)$ appear to play a role, for three dimensions, similar to that of the additional one-dimensional interaction used in [16–18].

3.3. Continuum approximation

Continuum approximations in the analysis of macromolecular chains are well documented [38, 46]. Their application, in our framework, to $G(L)$ and $Z_{R,1,ap}$ may be subject to some criticisms, as we shall see, although they may also have some methodological interest. By invoking [38, 46], we shall regard l as a continuous real variable ($0 \leq l \leq L$) in $G(L)$ and $Z_{R,1,ap}$. For the l th pair of monomers (one per chain), we introduce the centre of mass $\mathbf{R}_{CM,l}'' (= (\mathbf{R}_l^{(1)'} + \mathbf{R}_l^{(2)'})/2 \neq \mathbf{R}_{CM})$ and the relative position $\mathbf{z}'_l (= \mathbf{R}_l^{(2)'} - \mathbf{R}_l^{(1)'} \neq \mathbf{y})$. Then, in the continuum approximation, $G(L)$ factorizes $G(L) \simeq \exp[-(k_B T)^{-1} L \Sigma] W_G(\mathbf{R}_{CM,L+1}'' - \mathbf{R}_{CM,1}''; L) \cdot G_{rel}^{(0)}(\mathbf{z}'_{L+1}; \mathbf{z}'_1; L)$. Here, $\Sigma = 2 \sum_{n=0}^1 v_{1;n,e}^{(1)}(0)$ is the only remainder of $\sum_{r=1}^2 V_1^{(r)}$ and $W_G(\mathbf{R}_{CM,L+1}'' - \mathbf{R}_{CM,1}''; L) = [3/(\pi d_e^2 L)]^{3/2} \exp[-3(\mathbf{R}_{CM,L+1}'' - \mathbf{R}_{CM,1}'')^2 / (d_e^2 L)]$. $G_{rel}^{(0)}(\mathbf{z}'_{L+1}; \mathbf{z}'_1; L)$ is a relative Green's function, two monomers at the same position (l) in the different chains interact through $v_{02,e} = v_{02,e}(|\mathbf{z}'_l|) = v_{0,e}(|\mathbf{z}'_l|) + 2v_{2;1,e}(|\mathbf{z}'_l|)$ (allowing for cancellations among $v_{0,e}$ and $v_{2;1,e}$ s). Two criticisms regarding the continuum approximation in our framework already show up at this level: (1) somehow, one extends even more the use of the Gaussian approximations for scales shorter than d_e , and (2) one includes the interaction of two (non-complementary) monomers into the effective one, $v_{02,e}$, between complementary monomers. Anyway, we go on tentatively and connect with the analysis in [32, 38, 46], which leads us to replace $G_{rel}^{(0)}(\mathbf{z}'; \mathbf{z}'_1; L)$ by $\sum_{\lambda} \exp(-\lambda L) [f_{\lambda}(z')/z'] [f_{\lambda}(z'_1)^*/z'_1]$, where $z' \equiv |\mathbf{z}'|$. The (normalized) eigenfunction f_{λ} with its associated eigenvalue λ and $Z_{R,1,ap}$ fulfill

$$\left[-\frac{d_e^2}{6} \frac{d^2}{dz'^2} + \frac{v_{02,e}(z')}{k_B T} \right] f_{\lambda}(z') = \lambda f_{\lambda}(z'), \quad (22)$$

$$Z_{R,1,ap} \simeq \left[\prod_{r=1}^2 Z_R^{(r)} \right] \sum_{\lambda} \exp[-\lambda L] \left| \int_0^{+\infty} dz' z' f_{\lambda}(z') \right|^2. \quad (23)$$

For two macromolecular chains bound (bo) to each other, the allowed values of λ are real and negative and labelled by a discrete integer, n_{bo} , constituting a discrete and finite set $\lambda(n_{bo})$. If the chains are unbound to each other, all allowed values of λ sweep the whole infinite real interval $0 \leq \lambda < +\infty$. So, \sum_{λ} means the discrete sum $\sum_{n_{bo}}$ plus the integral $\int_0^{+\infty} d\lambda$. Bounds about the total number (N_{bo}) of bound states of quantum-mechanical radial equations exist [47] and can be applied to equation (22). Then, $N_{bo} \leq (2/\pi d_e)[6/(k_B T)]^{1/2} \int_0^{+\infty} dz' |v_{02,e}(z')|^{1/2}$. Then, T_{un} is such that, for $T \geq T_{un}$, equation (22) has only solutions with $0 \leq \lambda < +\infty$ but no bound-state solution. Then, the above bound

on N_{bo} yields $(2/\pi d_e)[6/(k_B T_{\text{un}})]^{1/2} \int_0^{+\infty} dz' |v_{02,e}(z')|^{1/2} \simeq 1$. Let $v_{02,e}(z')$ be attractive in a suitably large region, α_{DS}^{-1} be the size of the (larger) region where $v_{02,e}(z')$ is non-negligible and D_{DS} be some suitable average of $|v_{02,e}|$ (in either of those regions). A rough estimate of $N_{\text{bo}} \simeq 1$ yields $(2/\pi d_e)[6/(k_B T_{\text{un}})]^{1/2} \alpha_{DS}^{-1} D_{DS}^{1/2} \simeq 1$.

Next, let $v_{02,e}(z')$ be approximated by an effective Morse potential, so that equation (22) can be solved exactly (appendix A). Then, we choose $v_{02,e}(z')$ as given by the right-hand side of equation (2), with effective parameters D_{DS} , d_{DS} , and α_{DS} , with $2 < \exp(d_{DS}\alpha_{DS})$ (so that $v_{02,e}(z')$ is repulsive for small z'). With $2/(\hbar^2 A_{ii}^{(r)})$ replaced by $6/(k_B T d_e^2)$, the bound-state eigenvalues of equation (22) are given by equation (A.2) which, in turn, can be approximated by equation (A.3) for estimates. Then, equation (A.3) yields $\lambda(n_{\text{bo}}) \simeq -\alpha_{DS}^2 [n_{\text{bo}} + 2^{-1}(1-k)]^2$, for $0 \leq n_{\text{bo}} \leq 2^{-1}(k-1)$ with $k = [(24D_{DS})/(\alpha_{DS}^2 k_B T d_e^2)]^{1/2}$. The latter implies the absence of bound states for $k \leq 1/2$, that is, $k_B T_{\text{un}} = [24D_{DS}][\alpha_{DS}^2 d_e^2]$ (independent of d_{DS}), which resembles the above estimate based upon $N_{\text{bo}} \simeq 1$. The above formulae for $\lambda(n_{\text{bo}})$ and T_{un} , based on the Morse potential, agree formally with those obtained previously in [32], except for our use of d_e instead of the bond length d and our interpretation of D_{DS} . Let $n_e = 40$ and $d_e \simeq 14$ nm. It is not easy to compute D_{DS} . However, for a rough order of magnitude estimate, let $D_{DS} \simeq \lambda n_e \times 6$ kcal mol $^{-1}$, $\alpha_{DS} \simeq 2.8$ Å $^{-1}$. The value 6 kcal mol $^{-1}$ seems to be a reasonable estimate for $v_{0;i;\text{max}}$ (as it is >0.6 kcal mol $^{-1}$, corresponding to room temperature). In the present estimate, we also use 6 kcal mol $^{-1}$ for $v_{2;j,j';\text{max}}$, tentatively. The factor λn_e tries to account for the combined effect of the factor n_e , contained in $v_{0,e}$ (see equation (E.5)), and of the factor $2n_e^2$, due to $2v_{2;1,e}$ (see equation (E.8)); $\lambda = 1$ would correspond to $v_{2;1,e} = 0$. Then, we obtain $k_B T_{\text{un}} \simeq \lambda \times 0.04$ kcal mol $^{-1}$. Clearly, for $\lambda \simeq 1$ we would get an unreasonably low T_{un} , while for $\lambda \simeq 20$, we get the correct order, namely, $T_{\text{un}} \simeq 370$ K. Of course, if α_{DS} decreases adequately, one may reach the right T_{un} with $\lambda < 20$. These results may hold, qualitatively, for other potentials. Then, to get the adequate T_{un} in the continuum approximation: (a) for larger α_{DS} , we would need $v_{2;1,e} \neq 0$ and, physically, it would appear that $v_{2;1,e}$ should be attractive (instead of repulsive); (b) for smaller α_{DS} , far smaller $v_{2;1,e}$ would be required; and (c) $v_{1;n,e}^{(r)}$ would appear to play no role regarding T_{un} . Compare with [5, 7, 13, 32, 43]. Another criticism of the continuum approximation is that the above (a), (b) and (c) will differ with the consequences to be drawn when both chains are discretized (which, in turn, will appear more adequate), as we shall see in section 3.4.

Finally, we use the continuum approximation for a methodological discussion. We write $Z_{R,1,\text{ap}} \simeq Z_{R,1,\text{ap}}^{(\text{bo})} + Z_{R,1,\text{ap}}^{(\text{ub})}$, where $Z_{R,1,\text{ap}}^{(\text{bo})}$ ($Z_{R,1,\text{ap}}^{(\text{ub})}$) is the contribution to equation (23) due to all bound (bo) and all unbound (ub) states, respectively. For instance, $Z_{R,1,\text{ap}}^{(\text{bo})}$ ($Z_{R,1,\text{ap}}^{(\text{ub})}$) is the contribution to equation (23) due to $\lambda(n_{\text{bo}}) < 0$ ($\lambda \geq 0$). The case of physical interest corresponds to both large R_0 and large N . $Z_{R,1,\text{ap}}^{(\text{ub})}$ increases proportionally, more or less, to $\int d^3\mathbf{y}$ as R_0 does. $Z_{R,1,\text{ap}}^{(\text{bo})}$ remains finite regarding $\int d^3\mathbf{y}$, while it increases exponentially with N , as N does. There is no reasonable relationship between the magnitudes of $Z_{R,1,\text{ap}}^{(\text{bo})}$ and $Z_{R,1,\text{ap}}^{(\text{ub})}$, either if N and $\int d^3\mathbf{y}$ grow independently or if some thermodynamical limit argument (say, imposing the finiteness of $N/[\int d^3\mathbf{y}]^\gamma$, for some given exponent γ) is invoked. To solve this possible ambiguity, it appears natural to regard that $Z_{R,1,\text{ap}}^{(\text{bo})}$ ($Z_{R,1,\text{ap}}^{(\text{ub})}$) yields, though equation (13), the classical partition function, under all approximations involved, for all bo (ub) states. A similar discussion would apply for $Z_{C,e}$.

3.4. $G(L)$ with discrete l

We shall go beyond the continuum approximation and discuss the influence of the fact that both chains are discretized. With discrete l , the analysis is far more complicated,

and we shall limit ourselves to the non-trivial case of equations (20) and (21). One can regard $\sigma_{2;l}(T)$ as a perturbation and expand $G(L)$ into a sum of products of them times $W_G \exp[-(k_B T)^{-1} \sum_{r=1}^2 [v_{1;0,e}^{(r)}(0) + v_{1;1,e}^{(r)}(|\mathbf{z}_l^{(r)}|)]]$ s. The resulting perturbative expansion for $G(L)$ can be interpreted with diagrams, by extending standard perturbation techniques in field theory [46]. We shall analyse the convergence of that perturbative sum and estimate T_{un} . For $\sigma_{2;l}(T) = 0$, one has $Z_{R,1,\text{ap}} = Z_{R,1,\text{ap},0} = (4\pi R_0^3/3)$. With $\sigma_{2;l}(T) \neq 0$, we consider $Z_{R,1,\text{ap}} - Z_{R,1,\text{ap},0} \equiv Z_{R,1,\text{ap},\text{int}}$, arising from the perturbative sum. We estimate $Z_{R,1,\text{ap},\text{int}}$ in appendix F. One gets the following approximate expression for large L ($\nu_0 = [(4\pi d_e^2)/3]^{3/2}$):

$$Z_{R,1,\text{ap},\text{int}} \simeq \left[\prod_{r=1}^2 Z_R^{(r)} \right] \nu_0 \Lambda, \quad (24)$$

where $\ln \Lambda \simeq L g_0$,

$$g_0 = \frac{x \ln x}{1+x} + \ln[1+x^{-1}], \quad (25)$$

and

$$\begin{aligned} x = & \frac{1}{\nu_0(4\pi R_0^3/3)} \int \left[\prod_{r=1}^2 \prod_{l=1}^2 d^3 \mathbf{R}_l^{(r')} \right] \left[\prod_{r=1}^2 W_G(\mathbf{R}_2^{(r')} - \mathbf{R}_1^{(r')}; \delta l) \right] \\ & \times \sigma_{2;1}(T; \mathbf{R}_2^{(1')}, \mathbf{R}_1^{(1')}, \mathbf{R}_2^{(2')}, \mathbf{R}_1^{(2')}) \\ & \times \exp \left\{ -\frac{1}{k_B T} \sum_{r=1}^2 \left[v_{1;0,e}^{(r)}(0) + v_{1;1,e}^{(r)}(|\mathbf{z}_l^{(r)}|) \right] \right\}. \end{aligned} \quad (26)$$

$\sigma_{2;1}$ is given by equation (21) (allowing for cancellations among $v_{0,e}$ s and $v_{2;1,e}$ s). The number of pairs of monomers in the double chain which, on the average, are bound to each other is $n_{\text{max}} \simeq L/[1+x^{-1}]$. The number of pairs of monomers which are not bound to each other is $(L - n_{\text{max}}) \simeq L/[1+x]$. Notice that $g_0 > 0$ for any $x > 0$ ($g_0 \rightarrow +\infty$ for $x \rightarrow 0$, $g_0 \rightarrow +\infty$ for $x \rightarrow +\infty$ and g_0 has a unique minimum, equal to $\ln 2$, for $x = 1$). The bound double-stranded and unbound structures correspond, respectively, to $(L - n_{\text{max}})/L$ and $n_{\text{max}}/L \ll 1$. Equation (26) yields, roughly,

$$x \simeq \frac{1}{(\alpha_{DS} d_e)^3} \exp \left\{ -\frac{2}{k_B T} \left[v_{1;0,e}^{(1)}(0) + v_{1;1,e}^{(1)}(\alpha_{21}^{-1}) \right] \right\} \left\{ \exp \left[\frac{D_{DS}}{k_B T} \right] - 1 \right\}. \quad (27)$$

Now, assume $n_e = 40$ and $d_e \simeq 14$ nm. Also, let α_{DS}^{-1} be some average of the sizes of the regions where $v_{0,e}$ and $v_{2;1,e}$ are non-negligible, $D_{DS} (\simeq |v_{0,e}(\alpha_{DS}^{-1}) + 2v_{2;1,e}(\alpha_{DS}^{-1})|)$, and let α_{21}^{-1} be the range of $v_{1;1,e}^{(1)}$. Taking into account our d_e and α_{DS}^{-1} , we estimate $\alpha_{DS} d_e \simeq 390/n$, where $n = 1$ for $\alpha_{DS} \simeq 2.8 \text{ \AA}^{-1}$, and a reasonable range may be $1 \leq n \leq 10$. Then, we estimate T_{un} with the condition that x be of order 1. Using this, (27) can be roughly approximated by

$$\frac{D_{DS} - 2[v_{1;0,e}^{(1)}(0) + v_{1;1,e}^{(1)}(\alpha_{21}^{-1})]}{6 - \ln n} \simeq k_B T_{\text{un}}. \quad (28)$$

The model also makes sense in a range of T s above T_{un} , provided that $x > 0$ ($x = 0$ for $v_{0,e} = v_{2;1,e} = 0$). Let $n_e \times 6 \text{ kcal mol}^{-1}$ be a reasonable estimate for $|v_{0,e}(\alpha_{DS}^{-1})|$. This and equation (28) would provide an unreasonably large value for T_{un} . Then, in order for (28) to produce an acceptable T_{un} , we need that $v_{2;1,e}(\alpha_{DS}^{-1})$ and $v_{1;0,e}^{(1)}(0) + v_{1;1,e}^{(1)}(\alpha_{21}^{-1})$ provide some adequate repulsion, so as to cancel out part of the attractive $v_{0,e}(\alpha_{DS}^{-1}) (< 0)$. Thus, unlike what was met in the continuum approximation, to get the adequate T_{un} in the actual discrete- l estimates: (a) $v_{1;n,e}^{(r)}$ may, eventually, play a role; (b) the combined effect of both $v_{2;1,e} \neq 0$

and $v_{1;n,e}^{(r)}$, $r = 1, 2$, would appear to be repulsive (say, to be positive); (c) the influence of the inverse range α_{DS}^{-1} , described by $-\ln n$, seems somewhat limited.

These qualitative conclusions (a) and (b) would seem to receive additional support from the following argument. In real double-stranded B-DNA (below T_{un} , down to 300 K), there is a persistent length (d_{pers}) about 50 nm, amounting to 150 nucleotides [3, 4]. We recall [3, 28, 42, 46] and note that d_{pers} and d_e could play, approximately and respectively, the roles of d_e and d when the interactions in V_I are taken into account (recall equation (D.1) and the arguments in appendices B–D). Following [3, 42], one could argue that $d_{pers} \simeq [150/n_e]^{1/2} d_e F$, where F is some factor which would arise from V_I . The case $F > 1$ (< 1) would correspond to some net repulsion (attraction) contained in V_I . With $d_{pers} \simeq 50$ nm and the values of d_e obtained, one estimates that $F > 1$. Then, some effective interactions due to V_I would give rise to some repulsion: such interactions would reasonably be $v_{2;1,e} \neq 0$ and $v_{1;n,e}^{(r)}$ (while $v_{0,e} \neq 0$ would provide attraction). Furthermore, the above qualitative conclusions (a) and (b) for three dimensions seems to be consistent with the inclusion of the additional positive interaction employed in the one-dimensional model of [16–18]. Consistently with this, and more generally, we recall that the interactions between phosphates (belonging to either the same or different chains) in DNA are repulsive. *A priori*, and at least partly, $v_{1;n,e}^{(r)}$ and $v_{2;1,e}$ should account somehow for those phosphate interactions and, hence, be repulsive. Such a repulsive character appears to be consistent with and to be confirmed by the above qualitative conclusion (b), implied by the analysis of the denaturation temperature.

4. Global effective potentials

The aim of this section is simply to analyse, in an alternative, simple and qualitative way, the influence of the extension of each single chain on their mutual global interaction. For that purpose, we shall apply equations (15)–(17) with all $\delta^{(3)}(\mathbf{z}_l^{(r)} - d \sum_{j=n_{eff}^{(l-1)+1}}^{n_{eff}^{(l)}} \mathbf{u}_j^{(r)})$ and $\int d^3 \mathbf{z}_l^{(r)}$ omitted. Then, equation (17) becomes

$$\begin{aligned} Z_R &\geq z_R \left[\prod_{r=1}^2 Z_R^{(r)} \right], \\ z_R &= \int d^3 \mathbf{y} \exp \left[-\frac{\langle V_I \rangle(\mathbf{y})}{k_B T} \right]. \end{aligned} \tag{29}$$

Clearly, for $V_I \equiv 0$, $z_R = \int d^3 \mathbf{y}$, that is, Z_R increases with $\int d^3 \mathbf{y}$. We shall regard the global effective potential $\langle V_I \rangle(\mathbf{y}) = \langle V_0 \rangle(\mathbf{y}) + \sum_{r=1}^2 \langle V_1^{(r)} \rangle + \langle V_2 \rangle(\mathbf{y})$, given by equations (15), (16) (without $\mathbf{z}_l^{(r)}$ dependences), as a sort of upper bound on $\langle V_I \rangle(\mathbf{y}; [\mathbf{z}])$. It may suffice to concentrate the discussion on $\langle V_0 \rangle(\mathbf{y})$. We recall equations (1) and (8). Hence,

$$\langle V_0 \rangle(\mathbf{y}) = \sum_{i=1}^N \int \frac{d^3 \mathbf{q}}{(2\pi)^3} \tilde{v}_0(q) \exp(-i\mathbf{q} \cdot \mathbf{y}) \left\langle \exp \left[id \sum_{r=1}^2 \sum_{j=1}^{N-1} (\mathbf{q} \cdot \mathbf{u}_j^{(r)}) \tilde{\alpha}_{i,j;12}^{(r)} \right] \right\rangle(\mathbf{y}), \tag{30}$$

with $\tilde{\alpha}_{i,j;12}^{(1)} = \tilde{\alpha}_{i,j}^{(1)}$, $\tilde{\alpha}_{i,j;12}^{(2)} = -\tilde{\alpha}_{i,j}^{(2)}$ and $\tilde{v}_{0;i}(q) = \tilde{v}_0(q)$ for a homogeneous chain. For generic \mathbf{y} , $\langle V_0 \rangle(\mathbf{y})$ seems qualitatively different from $v_0(|\mathbf{y}|)$, due to $\langle \exp[i \sum_{r=1}^2 \sum_{j=1}^{N-1} (\mathbf{q} \cdot \mathbf{y}_j^{(r)}) \tilde{\alpha}_{i,j;12}^{(r)}] \rangle(\mathbf{y})$. One has $\int d^3 \mathbf{y} \langle V_0 \rangle(\mathbf{y}) = \sum_{l=1}^N \tilde{v}_0(q = 0)$. So, in some average sense, $\langle V_0 \rangle(\mathbf{y})$ increases with N . For a Morse potential with parameters $D'_{0,DS}$, α'_{DS} and d'_{DS} , we have

$$\tilde{v}_0(q) = \frac{2D'_{0,DS} \alpha'_{DS} \exp(d'_{DS} \alpha'_{DS})}{\pi^2} \left[\frac{\exp(d'_{DS} \alpha'_{DS})}{(q^2 + 4\alpha'^2_{DS})^2} - \frac{1}{(q^2 + \alpha'^2_{DS})^2} \right]. \tag{31}$$

For small q , $\langle \exp[i d \sum_{r=1}^2 \sum_{j=1}^{N-1} (\mathbf{q} \cdot \mathbf{u}_j^{(r)}) \tilde{\alpha}_{l,j}^{(r)}] \rangle(\mathbf{y})$ can be approximated by Gaussian distributions in \mathbf{q} , as outlined at the end of appendix E. Then

$$\langle V_0 \rangle(\mathbf{y}) = \langle V_0 \rangle(y) \simeq \frac{n_e}{(2\pi)^3} \int d^3 \mathbf{q} \exp[-i \mathbf{q} \cdot \mathbf{y}] \tilde{v}_0(q) \tilde{\beta}(q), \quad (32)$$

with

$$\tilde{\beta}(q) = \sum_{l=1}^L \exp[-\frac{1}{2} q^2 \tilde{\rho}_l d_e^2], \quad (33)$$

and $\tilde{\rho}_l = [6l^2 - 6(L+1)l + (L+1)(2L+1)]/9L$.

Based on a numerical integration of equation (32), we have found an excellent approximate closed-form expression for $\langle V_0 \rangle(\mathbf{y})$ when $L \gg 1$. It reads as

$$\langle V_0 \rangle(y) = \frac{27n_e \tilde{v}_0(0)}{2\pi^{3/2} d_e^3 L^{1/2}} \exp\left[-\frac{27y^2}{4d_e^2 L}\right]. \quad (34)$$

The restriction to small q implies that equation (34) is reliable for suitable length scales $y \geq d_e$. We conclude that the global interaction given by equation (34) (which embodies the interaction of all complementary pairs of nucleotides) is, qualitatively, rather different from the Morse potential between two complementary single nucleotides. For completeness, detailed expressions for $\langle V_1^{(r)} \rangle$ and $\langle V_2 \rangle(\mathbf{y})$ are provided at the end of appendix E.

5. Conclusions and final comments

We have obtained classical partition functions $Z_{C,e}$ for a class of simplified models for two interacting open homogeneous macromolecular chains in three-dimensional space, at thermal equilibrium from about room temperature up to about the denaturing temperature (T_{un}). Our analysis has employed some general empirical data of B-DNA (with bond length $d = 7.2 \text{ \AA}$). We have paid special attention to the crucial role of angular constraints on bond angles for each single chain, determined by the covalent bonding. Those constraints determine monomers in each chain. Each monomer contains n_e nucleotides, has effective length $d_e \simeq 2.2n_e^{1/2} \text{ nm}$, and could remain practically unaltered for temperatures in a range from about 300 K and up to somewhat above T_{un} . Reasonable values for n_e may range between $\simeq 20$ and 40 nucleotides. Weaker interactions (stacking and, eventually, various longer-range interactions amounting to some excluded-volume effects, supercoiling, etc), denoted as V_l , are allowed: they would be responsible for the double-stranded structure below T_{un} . The change of scale length from d to d_e could be regarded as some sort of renormalization: through it, V_l becomes an effective interaction among monomers.

The actual computation leading to $Z_{C,e}$ is interesting for, at least, the following reasons.

- (i) It has a methodological interest, as it starts from Quantum Mechanics and displays the sequence of arguments and approximations (variational computations with Morse potentials for each chain, classical limit, inclusion of angular constraints on bond angles, etc) yielding the generalized models.
- (ii) It leads to the same effective quantum and classical models and, hence, to the same general $Z_{C,e}$ as in previous works using harmonic-oscillator-like potentials [29, 30, 52], thereby confirming that the models follow from some general properties of $\sum_{r=1}^2 (V_b^{(r)} + V_a^{(r)})$ (but not from their details).
- (iii) It could suggest improvements of the variational computations for higher excited states with Morse-like potentials.

Through subsequent Gaussian approximations for each single chain, which are suggested, at least, partly by the monomers and may be adequate for distances $\geq d_e$ for rather general V_I , $Z_{C,e}$ is reduced formally to a suitable two-chain Green's function $G(L)$ (with $L = (N-1)/n_e$) for certain field-theoretic-like polymer models, which generalize Edwards' model. $G(L)$ contains effective potentials among monomers belonging both to the same and to different chains. $Z_{C,e}$ has been analysed in some detail. The reliability of integrating with Gaussian approximations in equation (20) for scales shorter than d_e has been discussed in appendix E. At least, non-Gaussian corrections for those scales should become decreasingly small as T increases towards T_{un} and somewhat above. While we believe that $Z_{C,e}$ continues to be reliable as T decreases down to about room temperature, it is open whether the same holds regarding the Gaussian approximation. Anyway, the computation leading to a generalized Edwards model from first principles could have methodological interest, as it may suggest corrections for the Gaussian approximations involved, for distances $\leq d_e$, as $T_{\text{un}} - T$ increases.

The discretized models in sections 3.2 and 3.4 (limited to effective interactions among complementary and nearest-neighbour non-complementary monomers) could provide a natural framework to extend the one-dimensional models (implying long-range cooperative effects at melting) in [7] and [16–18] to three spatial dimensions. This seems supported by the physical compatibility between the one-dimensional ones and our models, as discussed in the present work. Such a qualitative consistency would mean that our models in sections 3.2 and 3.4 would also contain some cooperative effects.

Some estimates of T_{un} are presented and discussed, based on continuum and discretized treatments of $G(L)$. The fact that $d_e > d$ ($=7.2$ Å, the bond length), the number n_e of nucleotides per monomer and various competing effective interactions could provide an adequate estimate for T_{un} .

A global effective potential (more properly, an upper bound thereof) between the two chains has been considered (for suitable length scales $\geq d_e$). It illustrates, in an independent way, that the spatial (one-dimensional) extension of each chain modifies the mutual global interaction between both chains, in comparison with typical potentials between nucleotides. These differences are displayed qualitatively for Morse potentials between complementary nucleotides.

Acknowledgments

We thank Professor J L F Abascal for an interesting discussion. RFA-E acknowledges the financial support from the Spanish Ministerio de Educación y Ciencia (Project FPA2004-02602). He is an Associate Member of Instituto de Biocomputación y Física de Sistemas Complejos, Universidad de Zaragoza, Zaragoza, Spain.

Appendix A. Morse potential: some useful results

The Schrödinger equation, with mass $[A_{jj}^{(r)}]^{-1}$ and the Morse potential $V_M(y)$ (see equation (2)), can be solved exactly for zero angular momentum states [39, 47]. Its bound-state solutions are summarized here. Let $\phi_M(y) = \chi_M(y)/y$ denote the radial eigenfunction with zero angular momentum and energy $E < 0$, where $\chi_M(y)$ satisfies the boundary conditions $\chi_M = 0$ at $y = 0$ and $+\infty$. Then, the bound-state solution of the Schrödinger equation is

$$\phi_M(y) = \frac{1}{y} \sqrt{\frac{\alpha}{N}} \exp\left[-\frac{z}{2}\right] z^{Ap} M(Ap + 1/2 - A\sqrt{D}, 2Ap + 1; z), \quad (\text{A.1})$$

where $z = 2A\sqrt{D}\exp[-\alpha(y-d)]$, $A = [2/(\alpha^2 A_{jj}^{(r)} \hbar^2)]^{1/2}$, $p = \sqrt{-E}$, and N is a normalizing constant, while $M(a, b; z)$ is Kummer's standard form of the confluent hypergeometric series [48]. The energy levels $E = E_{M,n} < 0$ (labelled with the integer n) are obtained from the zeros of the equation

$$M(Ap + 1/2 - A\sqrt{D}, 2Ap + 1; 2A\sqrt{D}\exp(\alpha d)) = 0. \quad (\text{A.2})$$

In the application yielding equation (13) (see appendix B), $2A\sqrt{D}\exp(\alpha d)$ is large. Therefore, with the help of asymptotic expansions for $M(a, b; z)$ [48], the finite number of discrete energy levels $E_{M,n}$ are

$$E_{M,n} = -D + \hbar\omega_0 \left(n + \frac{1}{2}\right) - \frac{\hbar^2\omega_0^2}{4D} \left(n + \frac{1}{2}\right)^2, \quad (\text{A.3})$$

$0 \leq n \leq (k-1)/2$ and $\omega_0 = \alpha\sqrt{2DA_{jj}^{(r)}}$. Their corresponding radial wavefunctions adopt the form

$$\phi_{M,n}(y) = \frac{1}{y} \sqrt{\frac{\alpha}{N_n}} \exp\left[-\frac{z}{2}\right] z^{(k-2n-1)/2} L_{k-n-1}^{k-2n-1}(z), \quad (\text{A.4})$$

where $k = 4D/\hbar\omega_0$ and $L_s^q(z)$ are the generalized Laguerre polynomials of degree s and order q [48]. The normalization constants in equation (A.4), N_n , are expressed by

$$N_n = [\Gamma(k-n)]^2 \sum_{s=0}^n \frac{\Gamma(k-2n+s-1)}{s!} \quad (\text{A.5})$$

($\Gamma(x)$ is the gamma function). In order to simplify the variational calculations in appendix B, we shall notice the following useful property:

$$|\phi_{M,n}(y)|^2 \rightarrow \Lambda_n \frac{\delta(y-d)}{d^2}, \quad (\text{A.6})$$

when $\alpha \rightarrow +\infty$ and $D \rightarrow +\infty$, while the dimensionless ratio $\sqrt{2D}/(\hbar\alpha[A_{jj}^{(r)}]^{1/2})$ remains constant (say, when the Morse potential is very deep). δ denotes the Dirac delta function.

Then, ω_0 diverges like $\sqrt{2A_{jj}^{(r)}}[\alpha^{-1}\sqrt{D}]^{\alpha^2}$. The factors Λ_n are given by

$$\Lambda_n = \begin{cases} 1 & n = 0, \\ \frac{1 + 2 \sum_{j=1}^n \binom{n}{j} S_j^{k-n} S_j^{k-1} + \sum_{j,l=1}^n \binom{n}{j} \binom{n}{l} S_j^{k-n} S_l^{k-n} S_{j+l}^{k-1}}{(n!)^2 k^{(2n+1-k)} \sum_{j=0}^n \frac{1}{j!} S_{2n-j}^{k-1}} & 1 \leq n \leq \frac{k-1}{2}, \end{cases} \quad (\text{A.7})$$

where $S_j^a = \prod_{l=1}^j (a-l)$.

An important feature of the Morse potential is that it is a bounded potential, i.e., it contains a finite-number spectrum, so it can account both for bound as well as unbound states. This clearly distinguishes it from other non-bounded potentials such as the harmonic oscillator potential.

Appendix B. Variational calculations with Morse potentials

An effective quantum Hamiltonian operator, $H_{Q,1}$, for two interacting macromolecular chains has been discussed in [30]. In terms of \mathbf{R}_{CM} , all $\mathbf{y}_j^{(r)}$ and \mathbf{y} , $H_{Q,1}$ becomes $H_{Q,1} = -(\hbar^2/2M_{CM})\nabla_{\mathbf{R}_{CM}}^2 + \tilde{H}_Q$, with $(M_{CM} = \sum_{r=1}^2 \sum_{i=1}^N M_i^{(r)} = \sum_{r=1}^2 M^{(r)})$

$$\begin{aligned} \tilde{H}_Q = & -\frac{\hbar^2}{2M_{red}} (\nabla_{\mathbf{y}})^2 - \sum_{r=1}^2 \frac{\hbar^2}{2M_1^{(r)}} (\nabla_{\mathbf{y}_1^{(r)}})^2 - \sum_{r=1}^2 \sum_{j=2}^{N-1} \frac{\hbar^2}{2M_j^{(r)}} [\nabla_{\mathbf{y}_j^{(r)}} - \nabla_{\mathbf{y}_{j-1}^{(r)}}]^2 \\ & - \sum_{r=1}^2 \frac{\hbar^2}{2M_N^{(r)}} (\nabla_{\mathbf{y}_{N-1}^{(r)}})^2 + \sum_{r=1}^2 V_b^{(r)} + \sum_{r=1}^2 V_a^{(r)} + V_I. \end{aligned} \quad (\text{B.1})$$

The quantum partition function Z_Q is [50] $Z_Q = \text{Tr}\{\exp[-(k_B T)^{-1} \tilde{H}_Q]\}$, where Tr denotes the usual trace operation over all states (in which nucleotides may be bound or unbound to other nucleotides, constituting chains or not). The degrees of freedom associated to \mathbf{R}_{CM} have been factored out in Z_Q . In a first variational computation, the larger (covalent-bond) potential $V_{M,j}^{(r)}$ (in $V_b^{(r)}$) will be dealt with quantum-mechanically. The smaller (covalent-bond) potential $v_{j,j+1}^{(r)}$ (in $V_a^{(r)}$) will be treated in two alternative ways. In a first treatment, we suppose that its (slower) rotational degrees of freedom, related to $|\mathbf{y}_j^{(r)} + \mathbf{y}_{j+1}^{(r)}|$, could be in an appreciable number of excited rotational states (instead of being just in the ground and in the first few excited states). For that first variational computation, the (normalized) variational wavefunctions are chosen as $\Phi_\xi = \left[\prod_{r=1}^2 \prod_{j=1}^{N-1} \phi_{j,M,n(j)}^{(r)}(y_j^{(r)}) \right] \psi(\mathbf{y}; \theta, \varphi)$. For the r th chain, $\phi_{j,M,n(j)}^{(r)}$ are, as j and $n(j)$ vary, the (normalized) wavefunction for the ground state and the first few excited ones for the Morse potential (embodied in $V_b^{(r)}$), given by equation (A.4) for $n = n(j)$, all with energies ($E_{j,M,n(j)}^{(r)}$) well below zero energy: they would correspond, in the harmonic-oscillator-like approximation for the Morse potential, to the ground state and to the first few excited vibrational-like states [49]. All higher excited vibrational-like states are either unoccupied or no longer represent physically meaningful states (as harmonic-oscillator-like vibrational potentials cease to be reasonable approximations). $\psi(\mathbf{y}; \theta, \varphi)$ are arbitrary (orthonormalized) wavefunctions. We perform a new variational computation, generalizing non-trivially previous computations [29, 30]. We evaluate $(\Phi_\xi, \tilde{H}_Q \Phi_\xi)$, using equation (6) and the assumptions preceding it. Related variational calculations carried out previously with harmonic-oscillator potentials [29, 30] have set the pattern for the actual more complicated and lengthier one with Morse potentials, the details of which will be omitted. Applying Peierls' inequality [45, 50], we get $Z_Q \geq z(\sum_{r=1}^2 V_b^{(r)}) Z_{Q,e}$ with $z(\sum_{r=1}^2 V_b^{(r)}) = \sum_{n(1)} \cdots \sum_{n(N-1)} \exp[-(k_B T)^{-1} \sum_{r=1}^2 \sum_{i=1}^{N-1} E_{i,M,n(i)}^{(r)}]$. Here, $E_{i,M,n(i)}^{(r)}$ stands for the energy (<0) of the $n(i)$ th bound state for the Morse potential (equation (A.3)) associated to the i th bond in the r th single chain. On the other hand, $Z_{Q,e} = \text{Tr}\{\exp[-(k_B T)^{-1} \tilde{H}_{Q,e}]\}$. Morse potentials and harmonic-oscillator-like potentials yield different values for $z(\sum_{r=1}^2 V_b^{(r)})$. The effective (Hermitian) quantum Hamiltonian $\tilde{H}_{Q,e}$ reads

$$\begin{aligned}
 \tilde{H}_{Q,e} = & -\frac{\hbar^2}{2M_{\text{red}}} (\nabla_{\mathbf{y}})^2 + \sum_{r=1}^2 \frac{1}{2M_1^{(r)}} \left(\frac{\mathbf{e}_1^{(r)}}{d_1^{(r)}} \right)^2 + \sum_{r=1}^2 \sum_{j=2}^{N-1} \frac{1}{2M_j^{(r)}} \left(\frac{\mathbf{e}_j^{(r)}}{d_j^{(r)}} - \frac{\mathbf{e}_{j-1}^{(r)}}{d_{j-1}^{(r)}} \right)^2 \\
 & + \sum_{r=1}^2 \frac{1}{2M_N^{(r)}} \left(\frac{\mathbf{e}_{N-1}^{(r)}}{d_{N-1}^{(r)}} \right)^2 + \sum_{r=1}^2 V_a^{(r)}(\theta, \varphi) + V_I(\mathbf{y}; \theta, \varphi). \tag{B.2}
 \end{aligned}$$

The (Hermitian) operators $\mathbf{e}_j^{(r)}$, $r = 1, 2$, $j = 1, \dots, N-1$, are given in [30]. The actual variational computations with Morse potentials yield the same $\tilde{H}_{Q,e}$ as in [30] (except for the inclusion of $\sum_{r=1}^2 V_a^{(r)} + \sum_{r=1}^2 V_1^{(r)}$ here). Some misprints in [30] are corrected in equation (B.2). The wavefunctions $\psi(\mathbf{y}; \theta, \varphi)$ have been taken as the complete set of all orthonormal eigenfunctions of $\tilde{H}_{Q,e}$ (describing one chain bound or unbound to the other chain): this justifies the trace instruction Tr in $Z_{Q,e}$. We have got the same $Z_{Q,e}$ for all variational Φ_ξ , due to the structures chosen for the latter. The property (A.6) has been employed in order to approximate the $(y_j^{(r)})$ s in equation (B.1) by the $(d_j^{(r)})$ s in (B.2). Were (A.6) not to have been employed, then the structure of (B.2) (in particular, its quadratic dependence on $\nabla_{\mathbf{y}}$ and on all operators $\mathbf{e}_j^{(r)}$ s) would still hold, but $(d_j^{(r)})$ would be replaced by effective lengths $(d_{j,\text{eff}}^{(r)})$: for the typical parameters involved in $V_b^{(r)}$, as discussed in section 2.1, to

approximate the $(d_{j,\text{eff}}^{(r)})$ s by the standard bond lengths $(d_j^{(r)})$ s does not appear unreasonable. These facts seem to be physically consistent with the restriction to the $\phi_{j,M,n(j)}^{(r)}$ s corresponding solely to the ground state and the first few excited ones. $Z_{Q,e}$ describes a model for a microscopic system formed by two open macromolecular chains, in which the bond lengths in each single chain are given constants $(d_j^{(r)})$. A simple estimate indicates that the uncertainty of $y_j^{(r)}$ in all those allowed $\phi_{j,M,n(j)}^{(r)}$ (as $n(j)$ sweeps the ground state and the first few excited ones) is small compared to $d_j^{(r)}$, by virtue of (6) and the assumptions preceding it.

We now turn to the classical limit, assuming also equation (12). Then, $\tilde{H}_{Q,e}$ becomes, in the classical limit

$$H_C = \frac{1}{2M_{\text{red}}}(\pi_C)^2 + \sum_{r=1}^2 \frac{1}{2M_1^{(r)}} \left(\frac{\mathbf{a}_{1,C}^{(r)}}{d_1^{(r)}} \right)^2 + \sum_{r=1}^2 \sum_{j=2}^{N-1} \frac{1}{2M_j^{(r)}} \left(\frac{\mathbf{a}_{j,C}^{(r)}}{d_j^{(r)}} - \frac{\mathbf{a}_{j-1,C}^{(r)}}{d_{j-1}^{(r)}} \right)^2 + \sum_{r=1}^2 \frac{1}{2M_N^{(r)}} \left(\frac{\mathbf{a}_{N-1,C}^{(r)}}{d_{N-1}^{(r)}} \right)^2 + \sum_{r=1}^2 V_a^{(r)}(\theta, \varphi) + V_I(\mathbf{y}; \theta, \varphi), \quad (\text{B.3})$$

where $\mathbf{e}_j^{(r)}$ have become the classical variables: $\mathbf{a}_{j,C}^{(r)} = -\mathbf{u}_{\theta_j}^{(r)} P_{\theta_j}^{(r)} - \mathbf{u}_{\varphi_j}^{(r)} (P_{\varphi_j}^{(r)} / \sin \theta_j^{(r)})$. The unit vectors $\mathbf{u}_{\theta_j}^{(r)}$ and $\mathbf{u}_{\varphi_j}^{(r)}$ are given in [30]. In turn, $P_{\theta_j}^{(r)}$, $P_{\varphi_j}^{(r)}$ are the classical momenta canonically conjugate to $\theta_j^{(r)}$ and $\varphi_j^{(r)}$, $j = 1, \dots, N-1$. The classical vector π_C is the classical limit of $-i\hbar \nabla_{\mathbf{y}}$. In the classical limit, $Z_{Q,e}$ becomes the corresponding classical partition function for the system of two open chains, $Z_{C,e}$, which is Gaussian in all classical momenta $P_{\theta_j}^{(r)}$, $P_{\varphi_j}^{(r)}$ and π_C . Upon performing all those Gaussian integrations, one gets equations (13) and (14), with $F_j^{(r)} = (k_B T / 2\pi\hbar^2)^{1/2} \exp[-(k_B T)^{-1} v_{j,j+1}^{(r)}]$ [30]. At this stage, one assumes that $v_{j,j+1}^{(r)}(|d_j^{(r)} \mathbf{u}_j^{(r)} + d_{j+1}^{(r)} \mathbf{u}_{j+1}^{(r)}|)$ has a unique somewhat deep minimum for $\mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)} \simeq \beta_j^{(r,0)}$, which constrains the bond angle between $\mathbf{u}_j^{(r)}$ and $\mathbf{u}_{j+1}^{(r)}$.

One could argue that those rotational degrees of freedom associated to $v_{j,j+1}^{(r)}$ could give rise to frequencies not appreciably smaller than $k_B T / \hbar$. Then, the former should also be treated quantum-mechanically and they would be in the ground and in some excited states. In the second variational computation, the degrees of freedom for $\sum_{r=1}^2 V_b^{(r)} + \sum_{r=1}^2 V_a^{(r)}$ are treated both quantum-mechanically on the same footing, by generalizing [29] and [52] and assuming that $v_{j,j+1}^{(r)}(|\mathbf{y}_j^{(r)} + \mathbf{y}_{j+1}^{(r)}|)$ is another Morse potential having a sharp minimum for $|\mathbf{y}_j^{(r)}| = d_j^{(r)}$ together with $\mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)} = \beta_j^{(r,0)}$. The (normalized) variational wavefunctions are now chosen as $\Phi_\xi = \left[\prod_{r=1}^2 \prod_{j=1}^{N-1} \phi_{j,M,n(j)}^{(r)}(y_j^{(r)}) \prod_{j'=1}^{N-2} \phi_{j',M,n(j')}^{(r)}(|\mathbf{y}_{j'}^{(r)} + \mathbf{y}_{j'+1}^{(r)}|) \right] \psi(\mathbf{y}; \theta, \varphi)$, $\phi_{j,M,n(j)}^{(r)}(y_j^{(r)})$ being as before, while $\phi_{j',M,n(j')}^{(r)}(|\mathbf{y}_{j'}^{(r)} + \mathbf{y}_{j'+1}^{(r)}|)$ denotes the (normalized) wavefunction for the ground state and the excited ones considered for the Morse potential $v_{j',j'+1}^{(r)}(|\mathbf{y}_{j'}^{(r)} + \mathbf{y}_{j'+1}^{(r)}|)$. One finds $Z_Q \geq z(\sum_{r=1}^2 V_b^{(r)}) z(\sum_{r=1}^2 V_a^{(r)}) Z'_{Q,e}$. The factor $z(\sum_{r=1}^2 V_a^{(r)})$ depends on the energies of the ground state and the excited ones considered for the Morse potential $v_{j',j'+1}^{(r)}(|\mathbf{y}_{j'}^{(r)} + \mathbf{y}_{j'+1}^{(r)}|)$, in the same way as $z(\sum_{r=1}^2 V_b^{(r)})$ does on those for $V_b^{(r)}$. The new effective quantum partition function $Z'_{Q,e}$ is given in [52]. The last step in this second variational computation is the transition to the classical limit, which is more difficult than in the first variational one. In fact, one transition to the classical limit for one chain also appeared in [52], which led from $Z'_{Q,e}$ to another effective classical $Z'_{C,e}$, which has a shortcoming: its integrand is not rotationally invariant, in an explicit way. Let $d_j^{(r)} = d (> 0)$ and $M_i^{(r)} = M_0$ (homogeneous chains). A new result here is that, at least for $\beta_j^{(r,0)}$ close to +1 ($\beta_j^{(r,0)} \simeq 0.8$ for any r and j corresponding

approximately to B-DNA [3]), we can approximate $Z'_{C,e}$ by another $Z''_{C,e}$ which contains an integrand which is, certainly, rotationally invariant in an explicit way (using [52]). In fact, one recasts $Z'_{C,e}$ (given in [52]) as a Gaussian integral over all classical momenta, realizes that, for $\beta_j^{(r,0)} = 1 - \delta\beta$ ($\delta\beta$ being positive and small compared to +1), one has $\mathbf{u}_{\theta_j}^{(r)} \cdot \mathbf{u}_{\theta_{j+1}}^{(r)} \simeq -\mathbf{u}_{\theta_{j+1}}^{(r)} \cdot \mathbf{u}_{\theta_j}^{(r)}$ and $\mathbf{u}_{\theta_j}^{(r)} \cdot \mathbf{u}_{\theta_{j+1}}^{(r)} \simeq +1$, employs them to perform several cancellations and, carrying out the remaining Gaussian integrations over classical momenta, one readily gets $Z''_{C,e}$. It is easily seen that $Z''_{C,e}$ has essentially the same form as $Z_{C,e}$ (equations (13) and (14)), now with $F_j^{(r)} = \lambda_0(2\delta\beta)^{1/2}(dM_0^{1/2})^{-1}\delta(\mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)} - \beta_j^{(r,0)})$. Here, λ_0 is a dimensionless constant (coming from the Gaussian integrations), which is irrelevant for our purposes, and $\delta(\mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)} - \beta_j^{(r,0)})$ denotes Dirac's delta function. All these complete the derivation of equations (13) and (14).

Appendix C. Properties of the single-chain determinant

In this appendix we shall analyse $\Delta^{(r)}$ and, using it and the effect of the constraints on bond angles (say, of $F_j^{(r)}$ s), we will estimate $Z_R^{(r)}$. We shall also introduce an angular average, which will be very useful later in appendices D and E. We have

$$\Delta^{(r)} \equiv \det \left[\mathbf{u}_j^{(r)} ((A^{(r)})^{-1})_{jj'} \mathbf{u}_{j'}^{(r)} \right]. \quad (\text{C.1})$$

Here, $\mathbf{u}_j^{(r)} ((A^{(r)})^{-1})_{jj'} \mathbf{u}_{j'}^{(r)}$ is a $(N-1) \times (N-1)$ -order symmetric matrix with elements

$$\mathbf{u}_j^{(r)} ((A^{(r)})^{-1})_{jj'} \mathbf{u}_{j'}^{(r)} = \begin{cases} \frac{1}{M^{(r)}} \sum_{k=1}^j \sum_{l=j+1}^N M_k^{(r)} M_l^{(r)} & j = j', \\ \frac{(\mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j'}^{(r)})}{M^{(r)}} \sum_{k=1}^j \sum_{l=j'+1}^N M_k^{(r)} M_l^{(r)} & j < j', \\ \frac{(\mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j'}^{(r)})}{M^{(r)}} \sum_{k=1}^{j'} \sum_{l=j+1}^N M_k^{(r)} M_l^{(r)} & j > j'. \end{cases}$$

The matrix $\mathbf{u}_j^{(r)} ((A^{(r)})^{-1})_{jj'} \mathbf{u}_{j'}^{(r)}$ is rotationally invariant but it is not tridiagonal. So, $\Delta^{(r)}$ is rotationally invariant and depends on all angles between every pair of bonds i and j . We shall suppose that all $d_j^{(r)} = d$ and $M_i^{(r)} = M_0$ (homogeneous chain). A study of $\Delta^{(r)}$, generalizing [30], yields the following. (i) $[\Delta^{(r)}]^{-1/2}$ takes on its dominant contributions (its maximum) when all $(\mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j'}^{(r)})^2 = 1$, ($j, j' = 1, \dots, N-1$)

$$\max[\Delta^{(r)}]^{-1/2} = \sqrt{\frac{N}{M_0^{N-1}}}. \quad (\text{C.2})$$

(ii) The ratio of the maximum over the minimum of $[\Delta^{(r)}]^{-1/2}$ increases with N as $(2/N)3^{(N-2)/2}$. (iii) In such domains where all $\mathbf{u}^{(r)}_j \cdot \mathbf{u}^{(r)}_{j'}$ are close to ± 1 , one can regard $\Delta^{(r)}$ as an even function of all $\mathbf{u}^{(r)}_j \cdot \mathbf{u}^{(r)}_{j+1}$, with all the remaining $\mathbf{u}^{(r)}_j \cdot \mathbf{u}^{(r)}_{j'}$, $j \neq j' \pm 1$ set equal to ± 1 (so that $\Delta^{(r)}$ becomes independent on all $\theta_j^{(r)}$), as a first approximation. One easily gets $\det A^{(r)} = N/M_0^{N-1}$. A necessary condition yielding the minimum of $[\Delta^{(r)}]^{-1/2}$ is that the bonds be oriented in space in such a way as to maximize the number of them that are perpendicular to the others. Furthermore, knowing a configuration with $N-1$ nucleotides for which $[\Delta^{(r)}]^{-1/2}$ attains its minimum, then one can find new configurations with N nucleotides

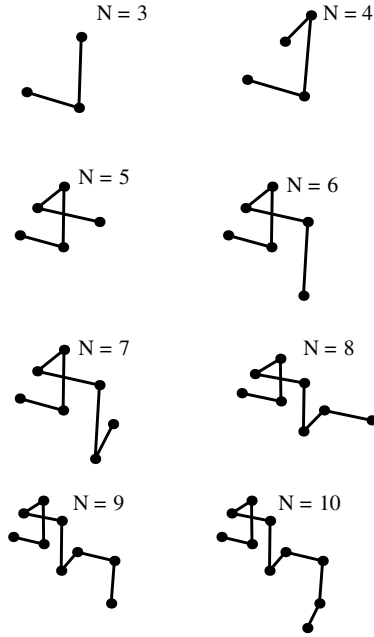


Figure C.1. Configurations for an open chain of N atoms for which $[\Delta_{3,N-1}]^{-1/2}$ attains its minimum. Notice that, given a configuration with N atoms for which $[\Delta_{3,N-1}]^{-1/2}$ is a minimum, then one can find new configurations for $N+1$ atoms for which $[\Delta_{3,N}]^{-1/2}$ reaches its minimum. The procedure consists in adding to one of the ends of the chain of N atoms another bond which is orthogonal to the last two bonds corresponding to the end where it is attached.

for which $[\Delta^{(r)}]^{-1/2}$ also reaches its minimum. This corresponds to imposing that the N th bond be perpendicular to the last two bonds of the chain end where the N th bond is attached, as displayed in figure C.1. Then, the minimum of $[\Delta^{(r)}]^{-1/2}$ is

$$\min[\Delta^{(r)}]^{-1/2} = \frac{N}{2} \sqrt{\frac{N}{3^{N-2} M_0^{N-1}}}. \quad (\text{C.3})$$

From equations (C.2) and (C.3), the above statement (ii) follows.

Using the previous results, we now estimate the reduced single-chain partition function $Z_R^{(r)} = \int d\Omega^{(r)} [\Delta^{(r)}]^{-1/2} F^{(r)}$, with $F^{(r)} = \prod_{j=1}^{N-2} F_j^{(r)}$. Two forms for $F_j^{(r)}$ have been discussed in appendix B. Let $\beta_j^{(r,0)} = 1 - \delta\beta$, $\delta\beta$ being small ($\simeq 0.2$ for DNA). We change variables in the integral for $Z_R^{(r)}$, by taking $\varphi_1^{(r)}, \theta_j^{(r)}, j = 1, \dots, N-1$ and $\beta_{j,j+1}^{(r)} \equiv \mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)}$, $j = 1, \dots, N-2$ as new integration variables. For given $\theta_{N-2}^{(r)}$ and $\beta_{N-3,N-2}^{(r)}$, we approximate the integrals over the whole ranges of $\beta_{N-2,N-1}^{(r)}$ and $\theta_{N-1}^{(r)}$, by three integrals over suitably restricted intervals, as follows. $\theta_{N-1}^{(r)}$ is restricted to vary only in three angular intervals: two with small size $\delta\theta_{\max}$ about $\theta_{N-2}^{(r)}$ and $\pi - \theta_{N-2}^{(r)}$ (associated to the maxima of $[\Delta^{(r)}]^{-1/2}$ and to $\beta_{N-2,N-1}^{(r)} \simeq \pm 1$), and another one about small $\beta_{N-2,N-1}^{(r)}$ (associated to the minima of $[\Delta^{(r)}]^{-1/2}$). The integrals in the two domains about the minima and about $\theta_{N-1}^{(r)} \simeq \pi - \theta_{N-2}^{(r)}$ are neglected, precisely due to $F_{N-2}^{(r)}$ (in any of the two forms discussed in appendix B), as are the contributions from the remaining angular domains. Regarding the integration about the maximum at $\theta_{N-1}^{(r)} \simeq \theta_{N-2}^{(r)}$ and the associated one over $\beta_{N-2,N-1}^{(r)}$ (which also includes the

contribution of the strong peak of $F_{N-2}^{(r)}$, we follow appendix B of [30]. Those integrations are estimated to yield $I\delta\theta_{\max}(\max[\Delta^{(r)}]^{-1/2})^{1/(N-2)}$, with $I = \int d\epsilon' (2\epsilon')^{-1/2} F_{N-2}^{(r)}$, performed in the small interval $0 \leq \epsilon' \leq \delta\beta$, where $F_{N-2}^{(r)}$ reaches its maximum. We iterate the procedure regarding the integration over $\beta_{N-2,N-3}^{(r)}\theta_{N-2}^{(r)}$ (for fixed $\theta_{N-3}^{(r)}, \beta_{N-3,N-4}^{(r)}$) and so on. Then, we estimate, for the r th single chain,

$$Z_{\mathbb{R}}^{(r)} \simeq 4\pi [2I\delta\theta_{\max}]^{N-2} (\max[\Delta^{(r)}]^{-1/2}). \quad (\text{C.4})$$

Let $\langle \mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)} \rangle = [Z_{\mathbb{R}}^{(r)}]^{-1} \int d\Omega^{(r)} \mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)} [\Delta^{(r)}]^{-1/2} F^{(r)}$. For the T s considered in this paper, $\langle \mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)} \rangle \simeq \beta_j^{(r,0)}$ is essentially a non-vanishing constant ($\simeq 0.8$ for DNA): it is independent of j, r and, practically, also of T (above and below the denaturing transition, down to 300 K). The conclusion is consistent with the above analysis of $[\Delta^{(r)}]^{-1/2}$ and holds for both $F_j^{(r)}$ treated in appendix B. The same $\langle \mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)} \rangle$ will be used in appendices D and E.

Appendix D. Averaging over monomers

This appendix will be devoted to characterize physically the monomers introduced in section 3 and their statistical properties, using the angular average $\langle \mathbf{u}_j^{(r)} \cdot \mathbf{u}_{j+1}^{(r)} \rangle$ discussed at the end of appendix C. In turn, those monomers and their averages will play a crucial role in the Gaussian approximations of appendix E. We choose n_e as the smallest integer such that, with respect to the above $\langle \cdot \rangle$, all $d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)}$, as l varies, be statistically independent from one another (for given r), at least approximately. For that purpose, notice that $\langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+s}^{(r)} \rangle = \langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+1}^{(r)} \rangle^s$, for $s = 2, 3, \dots$ [42], and $\langle (d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)})^2 \rangle = d_e^2 + n_e d^2 x_d$, with

$$d_e = n_e^{1/2} d \sqrt{\frac{1 + \langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+1}^{(r)} \rangle}{1 - \langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+1}^{(r)} \rangle}}, \quad (\text{D.1})$$

and $x_d = -[2\langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+1}^{(r)} \rangle / n_e] [(1 - \langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+1}^{(r)} \rangle^{n_e}) / (1 - \langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+1}^{(r)} \rangle)^2]$ [28, 46, 51]. For DNA [3] ($\langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+1}^{(r)} \rangle \simeq 0.8$ and $d \simeq 7.2 \text{ \AA}$), one has $(1 + \langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+1}^{(r)} \rangle) / (1 - \langle \mathbf{u}_l^{(r)} \cdot \mathbf{u}_{l+1}^{(r)} \rangle) \simeq 9$ which leads to the useful estimate: (A) $d_e / n_e^{1/2} \simeq 2.2 \text{ nm}$. We choose n_e so that x_d is adequately small. If $20 \leq n_e \leq 40$, x_d ranges between $\simeq 2$ and $\simeq 1$. For $n_e \simeq 20$ and $n_e \simeq 40$, one has $d_e \simeq 10 \text{ nm}$ and $d_e \simeq 14 \text{ nm}$, respectively. One has

$$d^2 \sum_{j=n_e(l-1)+1}^{n_e l} \sum_{j'=n_e(l'-1)+1}^{n_e l'} \langle (\mathbf{u}_j^{(r)})_{\alpha} (\mathbf{u}_{j'}^{(r)})_{\beta} \rangle \simeq 3^{-1} d_e^2 \delta_{l,l'} \delta_{r,r'} \delta_{\alpha,\beta}, \quad (\text{D.2})$$

α and β denoting three-dimensional Cartesian components and the δ s being Kronecker deltas. This important approximation embodies the use of $F_j^{(r)}$ (in any of its two forms).

If $N-1 = Ln_e$, one also has $\langle (\mathbf{R}_N^{(r)} - \mathbf{R}_1^{(r)})^2 \rangle \simeq Ld_e^2$.

We shall replace, for given r ($=1, 2$) and l ($=1, \dots, L$), the n_e vectors $d\mathbf{u}_j^{(r)}$ ($j = n_e(l-1)+1, \dots, n_e l$) by the monomer vectors $\mathbf{z}_l^{(r)}$ and other $n_e - 1$ vectors ($\mathbf{w}_{l,j}^{(r)}$), which will not appear at equations (19), (20) and (21) but will be necessary for intermediate stages. An analogy with mechanics will be useful. We regard those $d\mathbf{u}_j^{(r)}$ as the position vectors of nucleotides having equal masses (say, M_0) and express them in terms of their ‘centre of mass’ $d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)}$ (associated to $\mathbf{z}_l^{(r)}$) and $n_e - 1$ ‘relative’ position vectors [53], which will be the $\mathbf{w}_{l,j}^{(r)}$ s (their choice not being unique for $n_e > 2$). For given l and $j = n_e(l-1)+1, \dots, n_e l$,

one has

$$\mathbf{u}_j^{(r)} d = n_e^{-1} d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)} + \sum_{j'=1}^{n_e-1} v'_{j,l,j'} \mathbf{w}_{l,j'}^{(r)}, \quad (\text{D.3})$$

where $v'_{j,l,j'}$ are dimensionless coefficients (in which no M_0 dependence appears, as it cancels out). Thus: (i) for $n_e = 2$, $\mathbf{w}_{l,1}^{(r)} = d(\mathbf{u}_{2l-1}^{(r)} - \mathbf{u}_{2l}^{(r)})$; (ii) for $n_e = 3$, $\mathbf{w}_{l,1}^{(r)} = d(\mathbf{u}_{3l-2}^{(r)} - \mathbf{u}_{3l}^{(r)})$ and $\mathbf{w}_{l,2}^{(r)} = d[\mathbf{u}_{3l-1}^{(r)} - 2^{-1}(\mathbf{u}_{3l-2}^{(r)} + \mathbf{u}_{3l}^{(r)})]$; (iii) for $n_e = 4$, $\mathbf{w}_{l,1}^{(r)} = d(\mathbf{u}_{4l-2}^{(r)} - \mathbf{u}_{4l-3}^{(r)})$, $\mathbf{w}_{l,2}^{(r)} = d(\mathbf{u}_{4l}^{(r)} - \mathbf{u}_{4l-1}^{(r)})$ and $\mathbf{w}_{l,3}^{(r)} = 2^{-1}d[(\mathbf{u}_{4l}^{(r)} + \mathbf{u}_{4l-1}^{(r)}) - (\mathbf{u}_{4l-2}^{(r)} + \mathbf{u}_{4l-3}^{(r)})]$. And so on for $5 \leq n_e$. Those equations and (D.3) yield immediately the coefficients $v'_{j,l,j'}$. We consider $\langle d \sum_{j=n_e(l-1)+1}^{n_e l} (\mathbf{u}_j^{(r)} \cdot \mathbf{w}_{l',j'}^{(r)}) \rangle$, for any allowed l, l', j' as n_e increases. One sees easily that all those statistical averages vanish exactly for $n_e = 2, 3$, and that some of them do not for $n_e \geq 4$. However, as all $\mathbf{w}_{l',j'}$ are differences of $d\mathbf{u}_{j'}^{(r)}$ s, for $n_e \geq 4$ (and, at least, if $\beta^{(0)}$ is close to +1), one has the following property: (B) all $\langle d \sum_{j=n_e(l-1)+1}^{n_e l} (\mathbf{u}_j^{(r)} \cdot \mathbf{w}_{l',j'}^{(r)}) \rangle$ are, on the average, smaller than d_e^2 (by about one order of magnitude), independently on the choice of $\mathbf{w}_{l',j'}^{(r)}$.

Notice that, for $M_i = M_0$, $\tilde{\alpha}_{i,j}^{(r)} = N^{-1}j$ and $-N^{-1}(N-j)$, for $j = 1, \dots, i-1$ and $j = i, \dots, N-1$, respectively. Using them, one has: (C) $n_e^{-1}(\sum_{j'=n_e(l'-1)+1}^{n_e l'} \tilde{\alpha}_{i,j'}^{(r)})$ for $i = n_e l$, $l = 1, \dots, L$, and $l' = 1, \dots, L-1$ can be approximated by the leading contribution $\tilde{\alpha}_{l,l'}^{(r)}$ ($=L^{-1}l'$ and $-L^{-1}(L-l')$, for $l' = 1, \dots, l-1$ and $l' = l, \dots, L-1$, respectively). That approximation is quite reliable for $l \neq l'$ (say, for $i \neq j'$, with $j' = n_e(l'-1) + 1, \dots, n_e l'$). It could be improved for $l = l'$ by adding corrections (of order n_e^{-1}), which could give rise to fractions of monomers, but they lie outside our scope here.

The most useful results in this appendix are equations (D.2), (D.3) and the properties (A), (B) and (C).

Appendix E. Gaussian approximations

In this appendix, we shall be mostly concerned with providing suitable approximations for $Z_R^{(r)}([\mathbf{z}^{(r)}])$ s and for $\exp[-\langle V_l \rangle(\mathbf{y}; [\mathbf{z}]) / k_B T]$ in equation (18). Upon carrying out the analysis, we shall employ various results from appendices C and D. Based upon those approximations, by the end of the appendix we will also provide some additional information regarding the global effective potentials discussed in section 4.

We start with $Z_R^{(r)}([\mathbf{z}^{(r)}])$ (equation (16)). Its Fourier transform is

$$\tilde{Z}_R^{(r)}([\mathbf{k}^{(r)}]) = \int d\mathbf{\Omega}^{(r)} [\Delta^{(r)}]^{-1/2} F^{(r)} \prod_{l=1}^L \exp \left[i\mathbf{k}_l^{(r)} \cdot \left(d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)} \right) \right], \quad (\text{E.1})$$

where $[\mathbf{k}^{(r)}]$ denotes the set of all $\mathbf{k}_1^{(r)}, \dots, \mathbf{k}_L^{(r)}$ and $[\mathbf{k}]$ will represent the set formed by $[\mathbf{k}^{(1)}]$ and $[\mathbf{k}^{(2)}]$. We expand the exponential in equation (E.1) into powers of all $\mathbf{k}_l^{(r)}$ s (for fixed r) up to and including second order. The rotational invariance of $[\Delta^{(r)}]^{-1/2}$ and of $F_l^{(r)}$ will be useful below, upon performing angular integrations. Thus, linear terms in $\mathbf{k}_l^{(r)}$ s do not contribute, as their coefficients vanish (after carrying the angular integrations). Then, $\tilde{Z}_R^{(r)}([\mathbf{k}^{(r)}]) \simeq Z_R^{(r)}[1 - 2^{-1} \sum_{l=1}^L \sum_{l'=1}^L \sum_{\alpha,\beta=1}^3 (\mathbf{k}_l^{(r)})_\alpha (\mathbf{k}_{l'}^{(r)})_\beta (d^2 \sum_{j=n_e(l-1)+1}^{n_e l} \sum_{j'=n_e(l-1)+1}^{n_e l} (\mathbf{u}_j^{(r)})_\alpha (\mathbf{u}_{j'}^{(r)})_\beta)]$. We make use of equation (D.2) (the very existence of monomers). By approximating the

resulting polynomial quadratic in $\mathbf{k}_l^{(r)}$ s by an exponential (Gaussian approximation) and performing the inverse Fourier transform, we get

$$Z_{\mathbf{R}}^{(r)}([\mathbf{z}^{(r)}]) \simeq Z_{\mathbf{R}}^{(r)} \prod_{l=1}^L W_G(\mathbf{z}_l^{(r)}; \delta l), \quad (\text{E.2})$$

with $W_G(\mathbf{z}_l^{(r)}; \delta l) = [3/(2\pi d_c^2 \delta l)]^{3/2} \exp[-3(\mathbf{z}_l^{(r)})^2/(2d_c^2 \delta l)]$ ($\delta l = 1$). Thus, the product of W_G s in equation (20) is an approximation for the product of $Z_{\mathbf{R}}^{(r)}([\mathbf{z}^{(r)}])$ s in equation (18).

We now turn to $\langle V_I \rangle(\mathbf{y}; [\mathbf{z}])$. Let $g_I \equiv [\prod_{r=1}^2 Z_{\mathbf{R}}^{(r)}([\mathbf{z}^{(r)}])] \langle V_I \rangle(\mathbf{y}; [\mathbf{z}])$, using equation (15) with $f = V_I$ and equation (7) for V_I . Then, we have $g_I = g_0 + \sum_{r=1}^2 g_1^{(r)} + g_2$, where g_0 is also given by equation (15) for $f = V_0$ and so on. We shall start with g_0 . We use equations (1), introduce $\tilde{\alpha}_{i,j;12}^{(1)} = \tilde{\alpha}_{i,j}^{(1)}$, $\tilde{\alpha}_{i,j;12}^{(2)} = -\tilde{\alpha}_{i,j}^{(2)}$ and get $\mathbf{R}_i^{(1)} - \mathbf{R}_i^{(2)} = -\mathbf{y} + d \sum_{r=1}^2 \sum_{j=1}^{N-1} \tilde{\alpha}_{i,j;12}^{(r)} \mathbf{u}_j^{(r)}$. We employ (8), (11) and (D.3). Then

$$g_0 = \int \frac{d^3 \mathbf{q}}{(2\pi)^3} \tilde{v}_0(q) \sum_{i=1}^N \exp \left\{ i\mathbf{q} \cdot \left[-\mathbf{y} + n_e^{-1} \sum_{r=1}^2 \sum_{l'=1}^L \mathbf{z}_{l'}^{(r)} \sum_{j'=n_e(l'-1)+1}^{n_e l'} \tilde{\alpha}_{i,j';12}^{(r)} \right] \right\} G_{0;i}(\mathbf{q}; [\mathbf{z}]), \quad (\text{E.3})$$

where

$$G_{0;i}(\mathbf{q}; [\mathbf{z}]) = \int d\Omega \left[\prod_{r=1}^2 [\Delta^{(r)}]^{-1/2} F^{(r)} \prod_{l''=1}^L \delta^{(3)}(\mathbf{z}_{l''}^{(r)} - d \sum_{j''=n_e(l''-1)+1}^{n_e l''} \mathbf{u}_{j''}^{(r)}) \right] \times \exp \left\{ i\mathbf{q} \cdot \left[\sum_{r=1}^2 \sum_{l'=1}^L \sum_{j''=n_e(l'-1)+1}^{n_e l'} \sum_{j''=1}^{n_e-1} \tilde{\alpha}_{i,j'';12}^{(r)} v'_{j''',l',j''} \mathbf{w}_{l',j''}^{(r)} \right] \right\}. \quad (\text{E.4})$$

We consider the Fourier transform $\tilde{G}_{0;i}(\mathbf{q}; [\mathbf{k}])$ of $G_{0;i}(\mathbf{q}; [\mathbf{z}])$ in all variables $[\mathbf{z}]$. We expand $\tilde{G}_{0;i}(\mathbf{q}; [\mathbf{k}])$ up to and including second order in \mathbf{q} and all $\mathbf{k}_1^{(r)}, \dots, \mathbf{k}_L^{(r)}$ and invoke the rotational invariance of $[\Delta^{(r)}]^{-1/2}$ and of $F^{(r)}$, like for $\tilde{Z}_{\mathbf{R}}^{(r)}([\mathbf{k}^{(r)}])$. Linear terms in \mathbf{q} and $[\mathbf{k}]$ do not contribute. We approximate the quadratic terms, use equation (D.2) and estimate, upon employing the properties in appendix D (in particular, property (B) in that appendix), that all contributions due to $\mathbf{w}_{l',j''}^{(r)}$ s (and to $(\mathbf{w}_{l',j''}^{(r)})_{\alpha}(\mathbf{w}_{l',j''}^{(r)})_{\beta}$, $\alpha, \beta = 1, 2, 3$) are subdominant and small for large L . We approximate the quadratic polynomial by an exponential and perform the inverse Fourier transform. Then $G_{0;i}(\mathbf{q}; [\mathbf{z}]) \simeq \prod_{r=1}^2 [Z_{\mathbf{R}}^{(r)} \prod_{l=1}^L W_G(\mathbf{z}_l^{(r)}; \delta l)]$. We use equations (E.2)–(E.4) and, through property (C) in appendix D, the leading contributions embodied in $\tilde{\alpha}_{i,l'}^{(1)}$ s. Then, we get, for large L , $\langle V_0 \rangle \simeq \langle V_0 \rangle_e$ where

$$\langle V_0 \rangle_e = \sum_{l=1}^{L+1} v_{0,e}(|\mathbf{R}_l^{(1')} - \mathbf{R}_l^{(2')}|), \quad (\text{E.5})$$

$$\mathbf{R}_l^{(1')} - \mathbf{R}_l^{(2')} \equiv -\mathbf{y} + \sum_{l'=1}^L \mathbf{z}_{l'}^{(1)} \tilde{\alpha}_{l,l'}^{(1')} - \sum_{l'=1}^L \mathbf{z}_{l'}^{(2)} \tilde{\alpha}_{l,l'}^{(2')}, \quad (\text{E.6})$$

with $l = 1, \dots, L+1$ in equation (E.6). We have $v_{0,e}(|\mathbf{y}|) = (2\pi)^{-3} \int d^3 \mathbf{q} n_e \tilde{v}_0(q) \exp[i\mathbf{q} \cdot \mathbf{y}]$. Notice that $\mathbf{z}_l^{(r)} = \mathbf{R}_{l+1}^{(r')} - \mathbf{R}_l^{(r')}$. At this stage, we recall the centre-of-mass (CM) vector \mathbf{R}_{CM} and we introduce the new vectors $\mathbf{R}_l^{(r')}$, $l' = 1, \dots, L+1$, $r = 1, 2$. They are related to the set formed by \mathbf{R}_{CM} , \mathbf{y} and all $\mathbf{z}_l^{(r)}$ through equation (E.6), $\mathbf{R}_l^{(r')} = \mathbf{R}_{\text{CM}}^{(r')} + \sum_{l'=1}^L \tilde{\alpha}_{l,l'}^{(r')} \mathbf{z}_{l'}^{(r)}$ and $\mathbf{R}_{\text{CM}}^{(r')} = \mathbf{R}_{\text{CM}} + (-1)^r \mathbf{y}/2$. Equation (E.5) includes global interactions among full monomers. This completes the approximate analysis for V_0 , yielding $\langle V_0 \rangle_e$. Similar approximations can be applied to $g_1^{(r)}$ and g_2 , corresponding to $V_1^{(r)}$ and V_2 , respectively, and we shall

limit ourselves to giving the final results. We introduce two integers $n(21)$ and $n(22)$ as $0 \leq n(21) \leq \xi'/n_e < n(21) + 1$ and $1 \leq n(22) \leq \xi''/n_e < n(22) + 1$. Then,

$$\langle V_1^{(r)} \rangle \simeq \langle V_1^{(r)} \rangle_e([\mathbf{z}^{(r)}]) = \sum_{l=1}^{L+1} \sum_{n=0}^{n(21)} v_{1,n,e}^{(r)}(|\mathbf{R}_{l+n}^{(r')} - \mathbf{R}_l^{(r')}|), \quad (\text{E.7})$$

with $v_{1,n,e}^{(r)}(|\mathbf{y}|) = (2\pi)^{-3} \int d^3\mathbf{q} n_e^2 \tilde{v}_{1;n,e}(q) \exp[i\mathbf{q} \cdot \mathbf{y}]$. In turn, $\tilde{v}_{1;n,e}(q) = n_e^{-1} \sum_{\xi=nn_e}^{(n+1)n_e} \tilde{v}_{1;\xi}(q)$. Also,

$$\langle V_2 \rangle \simeq \langle V_2 \rangle_e = \sum_{l=1}^L \sum_{l'=l+1}^{l+n(22)} [v_{2;l-l',e}(|\mathbf{R}_{l'}^{(1')} - \mathbf{R}_l^{(2')}|) + v_{2;l'-l,e}(|\mathbf{R}_{l'}^{(2')} - \mathbf{R}_l^{(1')}|)], \quad (\text{E.8})$$

with $1 \leq |l - l'| \leq n(22)$ and $v_{2;|l-l',e}(|\mathbf{y}|) = (2\pi)^{-3} \int d^3\mathbf{q} \tilde{v}_{2;|l-l',e}(q) n_e^2 \exp[i\mathbf{q} \cdot \mathbf{y}]$. One has $\tilde{v}_{2;|l-l',e}(q) = n_e^{-2} \sum_{j,j'} \tilde{v}_{2;|j-j'|}(q)$, the summations being over $n_e(l-1) \leq j < n_e l$ and $n_e(l'-1) \leq j' < n_e l'$. The above discussion regarding the new vectors $\mathbf{R}_{l'}^{(r')}$ in connection with $\langle V_0 \rangle_e$ will hold for (E.7) and (E.8). Also, $\mathbf{R}_{l'}^{(1')} - \mathbf{R}_l^{(2')} = -\mathbf{y} + \sum_{l''=1}^L \mathbf{z}_{l''}^{(1)} \tilde{\alpha}_{l',l''}^{(1')} - \sum_{l''=1}^L \mathbf{z}_{l''}^{(2)} \tilde{\alpha}_{l,l''}^{(2)'}$. For suitably large $n(21)$ and $n(22)$, equations (E.7) and (E.8) would allow us to include long-range effects (excluded-volume, supercoiling, etc). Corrections to equations (E.5), (E.7) and (E.8) for length scales $\geq d_e$ are expected to be smaller (by about one order of magnitude), by virtue of property (B) in appendix D.

The previous approximations for $Z_R^{(r)}([\mathbf{z}^{(r)}])$ (equation (E.2)) and $\langle V_l \rangle(\mathbf{y}; [\mathbf{z}])$ and equation (18) yield

$$Z_{R,1} \simeq \left[\prod_{r=1}^2 Z_R^{(r)} \right] \int d^3\mathbf{y} \left[\prod_{r=1}^2 \prod_{l=1}^L d^3\mathbf{z}_l^{(r)} \right] W_G(\mathbf{z}_l^{(r)}; \delta l) \times \exp[-(k_B T)^{-1} \langle V_l \rangle_e(\mathbf{y}; [\mathbf{z}])] \equiv Z_{\text{red},1,\text{ap}}, \quad (\text{E.9})$$

where $\langle V_l \rangle_e = \langle V_0 \rangle_e + \sum_{r=1}^2 \langle V_1^{(r)} \rangle_e + \langle V_2 \rangle_e$. Notice that $\langle V_l \rangle_e$ is translationally invariant (under the overall spatial translation $\mathbf{R}_l^{(r')} \rightarrow \mathbf{R}_l^{(r')} + \mathbf{a}$, with \mathbf{a} independent on r, l). In equation (E.9), we multiply by $(4\pi R_0^3/3)^{-1} \int d^3\mathbf{R}_{\text{CM}} = 1$ and we transform back from \mathbf{R}_{CM} , \mathbf{y} and all $\mathbf{z}_l^{(r)}$ to all $\mathbf{R}_{l'}^{(r')}$, $l' = 1, \dots, L+1$, $r = 1, 2$. In the latter variables, $\int [\prod_{r=1}^2 \prod_{l=2}^L d^3\mathbf{R}_{l'}^{(r')}] (\prod_{l=1}^L W_G(\mathbf{R}_{l+1}^{(r')} - \mathbf{R}_l^{(r')}; \delta l)) \exp[-(k_B T)^{-1} \langle V_l \rangle_e] \equiv G(\mathbf{R}_{L+1}^{(1)}, \mathbf{R}_{L+1}^{(2)}, \mathbf{R}_1^{(1)}, \mathbf{R}_1^{(2)}; L) = G(L)$ is the Green's function for the double chain, with the discrete monomer index l playing the role of a discretized 'time' ($\delta l = 1$). It includes three effective interactions: (a) between two monomers at the same ($v_{0,e}$) and different ($v_{2;l'-l \neq 0,e}$) positions in the different chains, (b) between different ($v_{1;n,e}^{(r)}$, $n \neq 0$) monomers in the same chain. The terms with $v_{1;0,e}^{(r)}$ (0) represent self-interactions of one monomer. These structures are formally analogous to those found in field theories [46]. They also resemble (but they bear certain differences from) the one found in the study of the osmotic second virial coefficient for polymers [46, 54]. To fix the ideas, we simplify the models characterized by equations (E.5), (E.7) and (E.8), and we treat in section 3 the choice $n(21) = n(22) = 1$ (thereby omitting direct interactions between monomers with longer separations, at this stage). This justifies equations (19)–(21).

The reliability of employing the Gaussian approximations upon integrating in equation (E.9) for scales shorter than d_e has to be discussed. The contribution to $G(L)$ from distances much shorter than d_e could be small since, at least, V_0 certainly (and, rather possibly, $\sum_{r=1}^2 V_1^{(r)} + V_2$ as well) should be adequately repulsive there. Then, the long-distance approximations involved in equation (19) may be adequate for length scales less than about some Å. Then, there remains to discuss the reliability of the Gaussian approximations for length

scales between several Å and, say, 14 nm, where the net effect of $V_0 + \sum_{r=1}^2 V_1^{(r)} + V_2$ should be attractive. In principle, corrections to the Gaussian approximations for distances shorter than d_e would arise from orders higher than quadratic in the expansions of the exponential in equation (E.1), of $\tilde{G}_{0;i}(\mathbf{q}; [\mathbf{k}])$ for V_0 and of its counterparts for $V_1^{(r)}$ and V_2 , into powers of \mathbf{q} and all $\mathbf{k}_1^{(r)}, \dots, \mathbf{k}_L^{(r)}$. It is not easy to estimate the coefficients of the resulting terms of orders higher than quadratic (equation (D.2) may not suffice for that). A systematic review of corrections to the Gaussian approximation for the simpler case of one single classical freely jointed chain appears in [27], but it is not easy to extend it quantitatively to our case. Beyond [27], the approximate validity of the Gaussian approximations for adequately large distances, for single chains with constraints on bond angles and, at least, for $\beta_l^{(r,0)}$ close to +1 and $V_l = 0$ (say, for $Z_R^{(r)}([\mathbf{z}^{(r)})$]s) receives support from the combination of: (i) the argument at the end of appendix B (which stated that $Z''_{C,e}$ had essentially the same form as $Z_{C,e}$); and (ii) [52] (and references therein), where $Z''_{C,e}$ was studied. However, the extension of the last arguments for shorter distances and, moreover, for $V_l \neq 0$ (involving $v_{0;i}, v_{2;j,j'}$ and $v_{1;j,\xi}$ with inverse ranges about a few Å⁻¹) is not easy. Anyway, the average separations among monomers belonging to the different chains should increase, physically, as one approaches denaturation. Then, it appears that the corrections to the Gaussian approximation, for shorter distances, should become decreasingly small as T increases towards T_{un} and somewhat above.

The computation of the global effective potential $\langle V_0 \rangle(\mathbf{y})$ and, hence, of $\langle \exp[i\mathbf{q}d \sum_{r=1}^2 \sum_{j=1}^{N-1} \tilde{\alpha}_{l,j;12}^{(r)} \mathbf{u}_j^{(r)}] \rangle(\mathbf{y})$ in equation (30) follows a similar pattern. We discard all subdominant contributions (from $\mathbf{w}_{l,j',s}^{(r)}$, etc) and restrict ourselves to including $\langle (d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)}) (d \sum_{j=n_e(l-1)+1}^{n_e l} \mathbf{u}_j^{(r)}) \rangle$. The latter gives rise, in the Fourier transform of $\langle V_0 \rangle(\mathbf{y})$, to the factor $\tilde{\beta}(q)$, which appears in equations (32) and (33). Similarly, for $\langle V_1^{(r)} \rangle(\mathbf{y}$ -independent) and $\langle V_2 \rangle(\mathbf{y})$, one gets

$$\langle V_1^{(r)} \rangle \simeq L n_e^2 \sum_{n=0}^{n(21)} \frac{1}{(2\pi)^3} \int d^3 \mathbf{q} \tilde{v}_{1;n,e}^{(r)}(q) \exp\left[-\frac{1}{6} \mathbf{q}^2 d_e^2\right], \quad (\text{E.10})$$

and

$$\langle V_2 \rangle(\mathbf{y}) \simeq \sum_{l=1}^L \sum_{l'(\neq l)} \frac{1}{(2\pi)^3} \int d^3 \mathbf{q} \exp[-i\mathbf{q} \cdot \mathbf{y}] n_e^2 \tilde{v}_{2;|l-l',e}(q) \tilde{\beta}(q)_{l,l',22}, \quad (\text{E.11})$$

where

$$\tilde{\beta}(q)_{l,l',22} = \exp\left[-\frac{1}{4} \mathbf{q}^2 d_e^2 (\tilde{\rho}_l + \tilde{\rho}_{l'})\right], \quad (\text{E.12})$$

with $1 \leq |l - l'| \leq n(22)$, the same $\tilde{v}_{1;n,e}^{(r)}$ and $\tilde{v}_{2;|l-l',e}$ as in equations (E.7) and (E.8) and the same $\tilde{\rho}_l$ as in equation (33).

Appendix F. Estimates for discrete I

We expand $Z_{R,1,ap,int}$ into a perturbative sum of integrals of powers of $\sigma_{2;l}(T)$. Lx (with $x > 0$ given in equation (26)) is essentially, up to overall factors, the contribution of all terms linear in $\sigma_{2;l}(T)$. We recall that: (i) $\int d^3 \mathbf{R}_l^{(r)'} W_G(\mathbf{R}_{l+1}^{(r)'} - \mathbf{R}_l^{(r)'}; \delta l) = 1$; and (ii) the dominant contribution from $W_G(\mathbf{R}_{l+1}^{(r)'} - \mathbf{R}_l^{(r)'}; \delta l)$ (which is $\simeq v_0^{-1}$) comes from a domain with $|\mathbf{R}_{l+1}^{(r)'} - \mathbf{R}_l^{(r)'}|/d_e \leq 1$. Using both (i) and (ii) and $f_{ad} \simeq 1$ (which is the more accurate, the larger L is), we approximate the successive terms (containing integrals of products of $n > 1$

factors like $\sigma_{2,l}(T)$ s in the perturbative sum, by powers of x times combinatorial factors. The perturbative sum becomes readily $[\prod_{r=1}^2 Z_R^{(r)}] \nu_0 \Lambda$, with

$$\Lambda \simeq \sum_{n=1}^L \Lambda(n), \quad (F.1)$$

$$\Lambda(n) = \frac{L!}{n!(L-n)!} x^n,$$

and $L+1 \simeq L$ as L is large. We expect that the above approximation (ii), using the Gaussian approximation for distances shorter than d_e , is more reliable as $T_{un} - T$ decreases. A generic (l th) monomer in one single chain may interact or not with its mate monomer (the l th one) in the other single chain. n would be the number of pairs of monomers in the double chain which, on the average, are bound to each other. $(L-n)$ would count the number of monomer pairs, in the double chain, which are not bound to each other. Then, Λ could be regarded as some sort of partition function for all configurations, with all possible n . We shall study $\ln \Lambda(n)$ as a function of n , when not only L but n and $L-n$ are also large, by using Stirling's approximation [48] for all of them. One finds that $\ln \Lambda(n)$ has, indeed, a maximum ($\ln \Lambda(n_{\max})$) when $n = n_{\max}$. The dominant contribution is

$$\Lambda(n_{\max}) \simeq \frac{\exp[Lg_0 + g_1]}{(2\pi L)^{1/2}}, \quad (F.2)$$

$$g_1 = -\frac{1}{2} \ln \left[\frac{x^{-1}}{(1+x^{-1})^2} \right],$$

n_{\max} and g_0 are given in section 3.2. One also finds that $\Lambda(n) \simeq \Lambda(n_{\max}) \sum_{n=1}^L \exp[-(2L)^{-1}(2+x+x^{-1})(n-n_{\max})^2]$, that is, Λ can be approximated by a Gaussian distribution in n , centred about n_{\max} , with a width δn which increases as $L^{1/2}$ ($L^{-1}\delta n$ decreasing as $L^{-1/2}$ for $L \rightarrow +\infty$). Notice that $(2\pi L)^{-1/2} \sum_{n=1}^L \exp[-(2L)^{-1}(2+x+x^{-1})(n-n_{\max})^2]$ is finite for large L . Equation (F.2), together with n_{\max} , g_0 , and the estimate of $\Lambda(n)$ constitute the main results of this appendix.

References

- [1] Peyrard M 2004 *Nonlinearity* **17** R1
- [2] Lehninger A L, Nelson D L and Cox M M 1993 *Principles of Biochemistry* 2nd edn (New York: Worth Publishers)
- [3] Volkenshtein M V 1983 *Biophysics* (Moscow: Mir)
- [4] Frank-Kamenetskii M D 1997 *Phys. Rep.* **288** 13
- [5] Prohofsky E 1995 *Statistical Mechanics and Stability of Macromolecules* (Cambridge: Cambridge University Press)
- [6] Techera M, Daemen L L and Prohofsky E 1989 *Phys. Rev. A* **40** 6636
- [7] Peyrard M and Bishop A R 1989 *Phys. Rev. Lett.* **62** 2755
- [8] Zhang Y, Zheng W, Liu J and Chen Y 1997 *Phys. Rev. E* **56** 7100
- [9] Lifson S 1964 *J. Chem. Phys.* **40** 3705
- [10] Poland D and Scheraga H A 1970 *Theory of Helix-Coil Transitions in Biopolymers* (New York: Academic)
- [11] Wartell R M and Benight A S 1985 *Phys. Rep.* **126** 67
- [12] Kafri Y, Mukamel D and Peliti L 2000 *Phys. Rev. Lett.* **85** 4988
- [13] Dauxois T, Peyrard M and Bishop A R 1993 *Phys. Rev. E* **47** 684
- [14] Yakushevich L V 2004 *Non-Linear Physics of DNA* 2nd rev. edn (Weinheim: Wiley-VCH)
- [15] Gotoh O 1983 *Adv. Biophys.* **16** 1
- [16] Dauxois T, Peyrard M and Bishop A R 1993 *Phys. Rev. E* **47** R44
- [17] Dauxois T and Peyrard M 1995 *Phys. Rev. E* **51** 4027
- [18] Ares S, Voulgarakis N K, Rasmussen K O and Bishop A R 2005 *Phys. Rev. Lett.* **94** 035504
- [19] Ambjörnsson T and Metzler R 2005 *J. Phys.: Condens. Matter* **17** S1841
- [20] Gompper G and Löwen H 2005 *J. Phys.: Condens. Matter* **17** S1827

- [21] Mohammad-Rafiee F and Golestanian R 2005 *J. Phys.: Condens. Matter* **17** S1165
- [22] Cherstvy A G 2005 *J. Phys.: Condens. Matter* **17** S1363
- [23] Kalosakas G, Rasmussen K Ø, Bishop A R, Choi C H and Usheva A 2004 *Europhys. Lett.* **68** 127
- [24] Go N and Scheraga H A 1969 *J. Chem. Phys.* **51** 4751
- [25] Rallison J M 1979 *J. Fluid. Mech.* **93** 251
- [26] Messiah A 1962 *Quantum Mechanics* vol 2 (Amsterdam: North-Holland)
- [27] Flory P J 1975 *Statistical Mechanics of Chain Molecules* 2nd edn (New York: Wiley-Interscience)
- [28] Grossberg A Y and Khokhlov A R 1994 *Statistical Physics of Macromolecules (AIP Series in Polymers and Complex Materials)* (New York: American Institute of Physics)
- [29] Calvo G F and Alvarez-Estrada R F 2000 *Macromol. Theory Simul.* **9** 585
- [30] Alvarez-Estrada R F and Calvo G F 2002 *Mol. Phys.* **100** 2957
- [31] Lubensky D K and Nelson D R 2002 *Phys. Rev. E* **65** 031917
- [32] Sung W and Jeon J-H 2004 *Phys. Rev. E* **69** 031902
- [33] Carlon E, Orlandini E and Stella A L 2002 *Phys. Rev. Lett.* **88** 198101
- [34] Baiesi M *et al* 2003 *Phys. Rev. E* **67** 021911
- [35] Theodorakopoulos N, Dauxois T and Peyrard M 2000 *Phys. Rev. Lett.* **85** 6
- [36] Causo M S, Coluzzi B and Grassberger P 2000 *Phys. Rev. E* **62** 3958
- [37] Garel T, Monthus C and Orland H 2001 *Europhys. Lett.* **55** 132
- [38] Doi M and Edwards S F 1986 *The Theory of Polymer Dynamics* (Oxford: Oxford University Press)
- [39] Morse P M 1929 *Phys. Rev.* **34** 57
- [40] Herzberg M and Longuet-Higgins H 1963 *Discuss. Faraday Soc.* **35** 77
- [41] Mead C 1980 *Chem. Phys.* **23** 33
- [42] McQuarrie D A 1964 *Statistical Thermodynamics* (New York: Harper and Row)
- [43] Lipniacki T 2001 *Phys. Rev. E* **64** 051919
- [44] Barbi M, Cocco S and Peyrard M 1999 *Phys. Lett. A* **253** 358
- [45] Chaikin P M and Lubensky T C 1995 *Principles of Condensed Matter Physics* (Cambridge: Cambridge University Press)
- [46] Freed K F 1987 *Renormalization Group Theory of Macromolecules* (New York: Wiley)
- [47] Galindo A and Pascual P 1990 *Quantum Mechanics* vol 1 (Berlin: Springer)
- [48] Abramowitz M and Stegun I A (ed) 1964 *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables* (New York: Dover)
- [49] Schroedinger E 1967 *What is Life? The Physical Aspect of the Living Cell* (Cambridge: Cambridge University Press)
- [50] Huang K 1987 *Statistical Mechanics* 2nd edn (New York: Wiley)
- [51] Eyring H 1932 *Phys. Rev.* **39** 746
- [52] Alvarez-Estrada R F and Calvo G F 2004 *J. Phys.: Condens. Matter* **16** S2037
- [53] Messiah A 1961 *Quantum Mechanics* vol 1 (Amsterdam: North-Holland)
- [54] Yamakawa H 1971 *Modern Theory of Polymer Solutions* (New York: Harper and Row)