EPJ E
Soft Matter and Biological Physics
EPJ.org
your physics journal

Eur. Phys. J. E (2013) **36**: 46 DOI 10.1140/epje/i2013-13046-7

Helix-coil transition in terms of Potts-like spins

Artem Badasyan, Achille Giacometti, Rudolf Podgornik,
Yevgeni Mamasakhlisov and Vladimir Morozov

edp sciences Società Italiana di Fisica Springer

Eur. Phys. J. E (2013) **36**: 46
DOI 10.1140/epje/i2013-13046-7

THE EUROPEAN PHYSICAL JOURNAL E

Regular Article

Helix-coil transition in terms of Potts-like spins

Artem Badasyan[1,a], Achille Giacometti[2], Rudolf Podgornik[1,3], Yevgeni Mamasakhlisov[4], and Vladimir Morozov[4]

[1] Department of Theoretical Physics, J. Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia
[2] Dipartimento di Scienze Molecolari e Nanosistemi, Università Ca' Foscari Venezia, Calle Larga S. Marta DD2137, I-30123 Venezia, Italy
[3] Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana - SI-1000 Ljubljana, Slovenia
[4] Department of Molecular Physics, Yerevan State University, A. Manougian Str. 1, 375025, Yerevan, Armenia

Received 25 September 2012 and Received in final form 24 February 2013
Published online: 14 May 2013 –

Abstract. In the spin model of a helix-coil transition in polypeptides a preferred value of spin has to be assigned to the helical conformation, in order to account for different symmetries of the helical *vs.* the coil states, leading thus to the *Generalized Model of Polypeptide Chain* (GMPC) Hamiltonian as opposed to the Potts model Hamiltonian, both with many-body interactions. Comparison of explicit transfer matrix secular equations of the Potts model and the GMPC model reveals that the largest eigenvalue of the Potts model with Δ many-body interactions *coincides* with the largest eigenvalue of the GMPC model with $\Delta - 1$ many-body interactions, indicating the identity of both free energies. In distinction, the second largest eigenvalues in both models do *not coincide*, indicating a different behavior for the spatial correlation length that in its turn defines the width of the helix-coil transition interval. We explore in detail the thermodynamic consequences, resulting from spin models with and without the built-in spin anisotropy, that should indicate which model to favour as a more appropriate description of the equilibrium physical properties pertaining to the helix-coil transition.

1 Introduction

Though effective one-dimensional spin models have only a limited relevance for hard condensed matter systems [1,2], recently discovered examples of low-dimensional quantum systems such as the one-dimensional spin liquids and electrons in quantum wires [3,4], followed by low-dimensional macromolecular soft matter systems, hold promise for revitalising the importance of one-dimensional spin model description. In particular for the latter case, phenomena such as helix-coil transition in polypeptides [5,6] or indeed the stretch-induced B-to-S conformational transition in DNA [7,8] can both be described within a spin model dictionary with the *proviso* that the ordered state is being realized only when spins have a set, preferred direction [9]. Specifically, for the helix-coil transition the description of polypeptide conformations can be reduced to a consideration of a pair of torsional angles, pertaining to each of the peptide units [10]. Ramachandran's plot [11,12] of accessible *vs.* not accessible regions of these two torsional angles shows that helix formation is promoted only when both assume preferred values from a well-defined α-helical lacunae in this plot. When modeled in terms of spins the implication of the Ramachandran's plot is then that a helix can be formed locally only when spins take on a set, preferred value, which is exactly the situation we alluded to above. The situation with mechanical stretch induced B-to-S DNA conformational transformation, where the more compact B-form is reorganised into a more extended S-form of DNA, is in fact similar and has been described within models of the helix-coil variety pertaining to the same model universe [7,8].

In these and similar contexts the preferred direction of spins could be formalised within the Potts-like many-valued spin models that naturally allow for a preferred direction, distinguishing it by a selected spin value. The Potts-like model with preferred spin direction is fundamentally different from the classical Potts variety without any assigned preferred direction and one needs to elucidate in what ways, if at all, they differ between themselves in terms of the predicted equilibrium properties, in order to assess their appropriateness for description of real systems. This is the problem we set ourselves to clarify in this work.

A preferred direction of the spin is not the only feature differentiating these Potts-like models. The range of many-body interactions that one needs to consider is another one. Specifically, while in the case of a polypeptide chain

[a] *Present address*: Materials Research Laboratory, University of Nova Gorica, Vipavska 13, SI-5000 Nova Gorica, Slovenia; e-mail: abadasyan@gmail.com

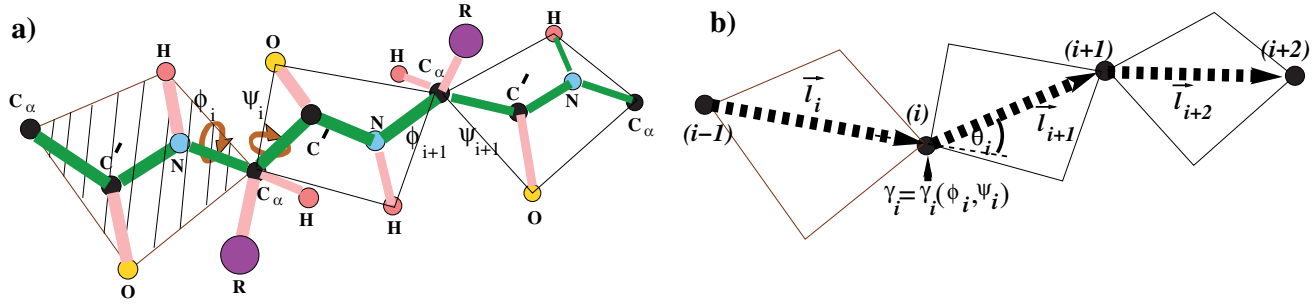

Fig. 1. (Color online) Schematic view of a polypeptide chain in a *trans* conformation. Parallelograms indicate the plane of the virtual peptide bond. a) Polypeptide chain where the main-chain atoms are represented as rigid peptide segments, linked by virtual bonds through the C_α atoms. Each segment has two degrees of freedom due to the rotation around the C_α–C' (torsional angle ϕ) and N–C_α (torsional angle ψ) bonds. R stands for the amino acid residues, while all other atoms have the corresponding chemical labels. b) Coarse-grained representation of a polypeptide chain: the conformation of the i-th repeating unit is described with the help of bond length l_i, bond angle θ_i and a pair of torsional angles ϕ_i, ψ_i.

spin model it is essential that three successive spins all be in a chosen conformation corresponding to a single helix-inducing hydrogen bond (see fig. 1) [5], the description of DNA overstretching implies that up to ten successive spins be engaged in a double helix-engendering hydrogen bond between opposing strands [13]. It is thus clear from both examples that it is necessary to consider some finite range of interactions and thus a finite number of nearest spins, Δ, as being crucial for the local formation of an ordered state —a hydrogen bond in the considered case.

There are thus two distinguishing features of the one-dimensional spin models worthy of further consideration: the existence of a preferred spin orientation pertinent to the ordered state, and the number of nearest spins entering into the local formation of an ordered state. The effects of the latter can be analyzed within the many-valued Potts spin model with an arbitrary but finite range of many-body interactions, while the former provides the basis of *Generalized Model of Polypeptide Chain* (GMPC) that accounts for the preferred spin orientation. GMPC has been formulated several decades ago [14–16] and has been extensively studied specifically in the context of the helix-coil transition [9,17–19]. It was shown that the Zimm-Bragg model [20] and the Lifson-Roig model [21] both correspond to particular cases of the GMPC variety with $\Delta = 2$ and $\Delta = 3$, respectively [9,16]. The Wako-Saito-Munoz-Eaton (WSME) model, widely applied to protein folding (see [22–28]), can also be shown to be related to the GMPC model as will be elaborated later. On the other hand, if no preferred spin value is taken into consideration, the standard one-dimensional Potts model has been solved exactly [1,2] but only with nearest-neighbor interactions $\Delta = 2$, with no known solutions for $\Delta > 2$, and was then applied to a helix-coil transition by Goldstein [29]. Furthermore, models that do not even allow for a spin Hamiltonian description have also been traditionally used to describe the statistical characteristics of the helix-coil transition [6,20,21].

As there obviously exists a spectrum of models that fall within the same class as the GMPC model it seems to be appropriate to know and understand the differences in the equilibrium properties implied by these various models. For this it is necessary to analyse their solutions and compare them. Our first task will therefore be to embark on a detailed study and comparison of the solutions of the GMPC model with preferred spin direction and the classical Potts variety for different values of the range of many-body spin interactions, Δ, by analysing the respective explicit transfer matrix secular equations. Specifically, we will show by a detailed but general analysis that the lowest eigenvalues of the two models coincide but the next higher eigenvalues do not! This non-trivial result has fundamental repercussions for the equilibrium states described by the two models which in general do not coincide.

2 Helix-coil model formulation in terms of spin variables

Statistical description of polypeptide chain conformations involves important coarse-graining on the level of the C_α atoms, because of the planar configurations of the atomic groups ($C^{(\alpha)}_{i-1}$, C_{i-1}, O_{i-1}, N_i) due to specific bond hybridizations. The planar structure of these groups (peptide groups) allows the introduction of virtual bonds, connecting the neighboring asymmetric carbon atoms (fig. 1(a)) [5,30]. The configuration of a polypeptide chain can then be described with the sequence of (virtual) bond vectors $\{\boldsymbol{l}_i\}$ [10,31,32], related to its backbone. In this description the bond lengths $\{l_i\}$, $i = 1 \ldots N-1$, bond (valence) angles $\{\theta_i\}$, $i = 1 \ldots N-2$, and pairs of $\{\phi_i, \psi_i\}$ torsional angles can be associated with each repeating unit. This description can be further simplified by taking into account that bond lengths and angle values usually vary within very narrow intervals (average fluctuation of ± 3–5% at room temperature) and their fluctuations can be ignored [10]. The only relevant variables remaining are thus the $\{\phi_i, \psi_i\}$ torsional angles (fig. 1(b)).

The conformational partition function of the repeating unit can be represented as a finite sum after discretization of the torsional angles (approximation of rotational isomers) [33], opening up a possibility for a spin-based description of the polypeptide conformations. Assume that

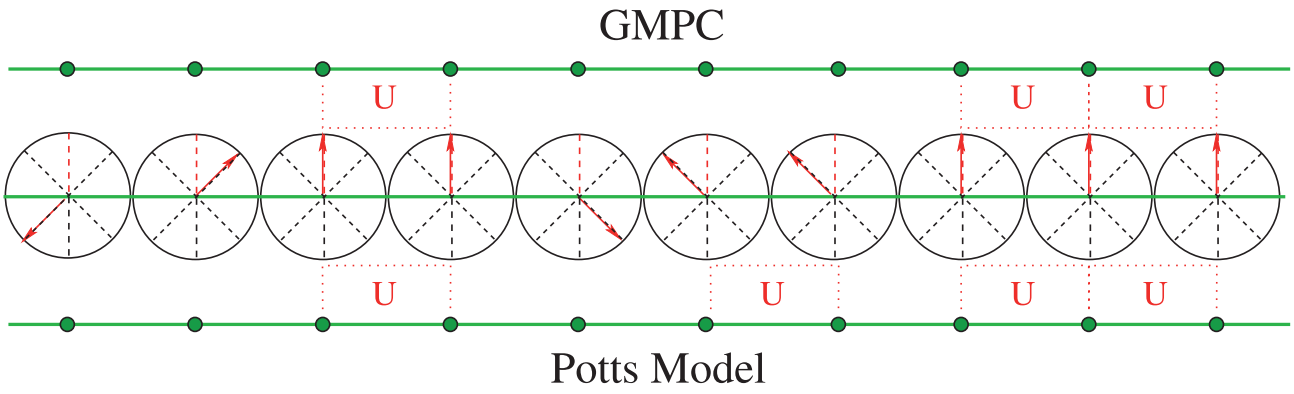

Fig. 2. (Color online) Schematic representation of a 10-mer with spins assigned to each repeating unit. Spins are shown as red arrows with $Q = 8$ possible orientations. (Top) Assigned energy (U) in the nearest-neighbor ($\Delta = 2$) GMPC model (with the red dotted line indicating the preferred orientation of the spin). (Bottom) Potts model as in ref. [29]. The Potts energy assignment results in a higher energy of the sample spin sequence.

spin γ_i describing the conformation of the i-th repeating unit can take one of the $Q(\geq 2)$ values; $\gamma_i = 1$ corresponding to values of the torsional angles $\{\phi_i, \psi_i\}$ from the helical region of the Ramachandran map, while the other $Q - 1$ values correspond to torsional angles from allowed (non-helical) region. The magnitude of Q (number of spin orientations) can be identified with the ratio of the allowed region area *vs.* helical region area on a Ramachandran map. According to the polypeptide chain geometry the equilibrium hydrogen bond formation can be established between the NH and CO groups, separated by three asymmetric carbon atoms [5]; the energy U is associated with every formed hydrogen bond corresponding to a *coupling constant* $W = \exp(U/T)$, where T is the temperature (in units where $k_B = 1$ for the Botzmann constant). One hydrogen bond thus restricts three $\{\phi, \psi\}$ pairs of rotation angles and establishes the structure with screw symmetry (α-helix) [34]. Within the spin language this means that the hydrogen bond fixes three successive spin values along the chain. On the other hand, hydrogen bonds in double-stranded DNA are formed between repeating units on the opposite strands and are approximately perpendicular to the DNA axis. Creation of hydrogen bonds in one pair of opposing bases thus applies restrictions to conformational states of $\sim$ 10 neighbors (on the scale of single-strand Kuhn length) [13]. It thus makes sense to generalize and consider that one hydrogen bond formation restricts arbitrary (but finite) Δ number of spins [16], corresponding to many-body interactions. As helix formation comes at an entropic cost [32], the larger is Δ, the higher is such an entropic cost [16]. The transformation from a coil to a helical conformation is energetically favorable (negative hydrogen bond energy is gained) but entropically unfavorable (the number of micro-states, available for repeating unit in a helical macro-state is decreased, as compared to the coil). As we show below, the compensation of energetic and entropic costs engenders a transition at the temperature corresponding to $\exp(U/T) = Q$.

To summarize, statistical description of the helix-coil transition requires three basic parameters: an energetic parameter, $W = \exp(U/T)$, where U is the energy of a hydrogen bond; an entropic parameter, Q, that stands for the number of spin values; and a geometric parameter, Δ, that describes the many-body geometry of the hydrogen bond formation. The corresponding Hamiltonian can thus be built in terms of the γ_i spins [9,14–16,18,19] and corresponds to the GMPC model if the proper helix formation demands that Δ successive γ's are all in the same preferred conformation, *e.g.*, 1 (see fig. 2, top). In the case of no preferred spin assignment to the helix formation we are then back to the Potts model (see fig. 2, bottom). In what follows we consider both types of spin Hamiltonians and discuss similarities and differences between the ensuing thermodynamics.

3 Generalized model of polypeptide chain in the transfer matrix formalism

3.1 Hamiltonian

The Hamiltonian for the Generalized Model of Polypeptide Chain (GMPC) is defined as

$$-\beta H = J \sum_{i=1}^{N} \delta(\gamma_{i-2}, 1)\delta(\gamma_{i-1}, 1)\delta(\gamma_i, 1), \qquad (1)$$

where $J = U/T$ is the reduced energy of the hydrogen bond. The strength of the hydrogen bond is between the valence bond and the van der Waals interactions. By definition, the energy of hydrogen bond formation is negative. $\delta(a, b)$ is the Kronecker symbol. When generalized to any finite Δ, the above Hamiltonian assumes the form

$$-\beta H = J \sum_{i=1}^{N} \prod_{k=\Delta-1}^{0} \delta(\gamma_{i-k}, 1) = J \sum_{i=1}^{N} \delta_i^{(\Delta)}, \qquad (2)$$

where $\delta_i^{(\Delta)} = \prod_{k=\Delta-1}^{0} \delta(\gamma_{i-k}, 1)$ is the product of Δ Kronecker symbols for neighboring repeating units.

3.2 Transfer matrix approach and characteristic equation

A transfer matrix can be constructed corresponding to the Hamiltonian eq. (2) for the $\Delta = 2$, $\Delta = 3$ and $\Delta = 4$ cases and in fact any finite $1 < Q < \infty$. The algorithm for larger values of Δ is similar to that for the $\Delta = 3$ and $\Delta = 4$ cases.

Starting at $\Delta = 2$ it is straightforward to show that the $Q \times Q$ transfer matrix reads

$$\widehat{G}(2) = \begin{pmatrix} e^J & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{pmatrix}, \tag{3}$$

and contains many $(Q-1)$ identical rows and columns. At $\Delta = 3$ the situation is more complicated, since many-body interactions start to play an important role, and a direct construction of the transfer matrix is not possible in this form. However, the goal can be achieved by using the following preliminary transformation. Instead of the three spin variables $\gamma_{i-2}, \gamma_{i-1}, \gamma_i$, one introduces a pair of two-index variables $\Omega_{i-1} = (\gamma'_{i-1}, \gamma_i)$ and $\Omega_i = (\gamma_{i-1}, \gamma_i)$ and sets to zero all elements of the matrix $\left(Q^2 \times Q^2\right)$ for which $\gamma'_{i-1} \neq \gamma_{i-1}$ [10]. This allows us to write the Hamiltonian in the form

$$\begin{aligned} -\beta H &= J \sum_{i=1}^{N} \delta(\gamma_{i-2}, 1)\delta(\gamma_{i-1}, 1)\delta(\gamma'_{i-1}, 1)\delta(\gamma_i, 1) \\ &= J \sum_{i=1}^{N} \delta(\Omega_{i-1}, 1)\delta(\Omega_i, 1), \end{aligned} \tag{4}$$

with the statistical weight

$$g(\Omega_{i-1}\Omega_i) = e^{-\beta H}\delta(\gamma'_{i-1}\gamma_{i-1}). \tag{5}$$

Then the $(Q^2 \times Q^2)$ transfer matrix reads:

$$\widehat{G}(3) = \begin{pmatrix} e^J & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots\dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots\dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots\dots & 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots\dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots\dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots\dots & 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots\dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots\dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots\dots & 1 & 1 & \dots & 1 \end{pmatrix}. \tag{6}$$

Here again, there are many $(Q^2 - \Delta)$ identical rows and columns. At $\Delta = 4$ there are 4 spin variables and we add another 2, following a similar trick as above, so that a pair of three-index variables reads $\Omega_{i-1} = (\gamma_{i-3}, \gamma'_{i-2}, \gamma'_{i-1})$ and $\Omega_i = (\gamma_{i-2}, \gamma_{i-1}, \gamma_i)$. The statistical weight is then prescribed as

$$g(\Omega_{i-1}\Omega_i) = e^{-\beta H}\delta(\gamma_{i-2}, \gamma'_{i-2})\delta(\gamma_{i-1}, \gamma'_{i-1}). \tag{7}$$

The resulting transfer matrix has dimensions $(Q^3 \times Q^3)$ and is by its structure similar to eq. (6). For larger Δ's it is necessary to group γ's into Ω_{i-1} and Ω_i in a similar way. The procedure can be generalized and the appropriate statistical weight can be written as

$$g(\Omega_{i-1}, \Omega_i) = e^{-H/T} \prod_{k=\Delta-2}^{1} \delta\left(\gamma'_{i-k}, \gamma_{i-k}\right), \tag{8}$$

resulting in a transfer matrix $\widehat{G}(\Delta)$ of dimensions $(Q^{\Delta-1} \times Q^{\Delta-1})$. Since there are $(Q^{\Delta-1} - \Delta)$ identical rows and columns, the characteristic equation for $\widehat{G}(\Delta)$ turns out to be quite simple

$$\begin{aligned} P_{GMPC}(\lambda, W, Q, \Delta) &= \lambda^{Q^{\Delta-1}-\Delta} \times p_{GMPC}(\lambda, W, Q, \Delta) \\ &= 0, \end{aligned} \tag{9}$$

where

$$\begin{aligned} p_{GMPC}(\lambda, W, Q, \Delta) &= \lambda^{\Delta} - (W - 1 + Q)\lambda^{\Delta-1} \\ &\quad + (W-1)(Q-1)\sum_{k=2}^{\Delta} \lambda^{\Delta-k}. \end{aligned} \tag{10}$$

Obviously, there are only Δ non-trivial eigenvalues, so that it is sufficient to consider a transfer matrix of a much smaller $(\Delta \times \Delta)$ size to derive the thermodynamics. Such a matrix has been derived in [14] by elementary transformations of $\widehat{G}(\Delta)$ and has the form

$$\widehat{g}(\Delta) = \begin{pmatrix} W-1 & W-1 & \dots & W-1 & W-1 & W-1 \\ 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & Q \end{pmatrix}. \tag{11}$$

One can construct the corresponding transfer matrix by noting that:

- all elements of the first row are equal to $W-1 = e^J - 1$;
- all elements of the first lower pseudo-diagonal are 1;
- the element (Δ, Δ) is Q;
- all other elements are zero.

Alternative elementary transformations lead to

$$\widehat{g^*}(\Delta) = \begin{pmatrix} W & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & Q-1 \\ 1 & 1 & 1 & \dots & 1 & 1 & Q-1 \end{pmatrix}. \tag{12}$$

Both $\widehat{g}(\Delta)$ and $\widehat{g^*}(\Delta)$ have a much smaller size than $\widehat{G}(\Delta)$ and result in the same characteristic eq. (10). By adding an artificial $\lambda = 1$ root, eq. (10) can be written in a more compact form,

$$p_{GMPC}(\lambda, W, Q, \Delta) = \lambda^{\Delta-1}(\lambda - W)(\lambda - Q) - (W-1)(Q-1) = 0. \quad (13)$$

4 One-dimensional Potts model with many-body interactions in the transfer matrix formalism

4.1 Hamiltonian

Unlike the previous GMCP case, Goldstein's formulation [29] starts from a standard Potts model with Δ many-body interaction without any distinction between spin values as

$$-\beta H = J \sum_{i=1}^{N} \prod_{k=\Delta-1}^{1} \delta(\gamma_{i-k}, \gamma_{i-k-1}). \quad (14)$$

One can notice that for the same Δ many-body interactions the Hamiltonian of Potts model contains the product of $\Delta - 1$ Kronecker symbols instead of exactly Δ as in the case of the GMPC model. This fact has important consequences, as we will show below.

4.2 Transfer matrix approach and characteristic equation

The transfer matrix corresponding to the Hamiltonian eq. (14) can be constructed *seriatim* for $\Delta = 2$, $\Delta = 3$, $\Delta = 4$ and then for any finite $2 < Q < \infty$. The algorithm for larger values of Δ is similar to that for the $\Delta = 3$ and $\Delta = 4$ cases.

At $\Delta = 2$ the $Q \times Q$ transfer matrix reads

$$\widehat{G}(2)_{Potts} = \begin{pmatrix} e^J & 1 & \dots & 1 \\ 1 & e^J & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & e^J \end{pmatrix}. \quad (15)$$

At $\Delta = 3$ we use the same transformation discussed above for implementing the transfer matrix of the GMPC model. Instead of the three spin variables γ_{i-2}, γ_{i-1}, γ_i, we now introduce a pair of two-index variables $\Omega_{i-1} = (\gamma_{i-2}, \gamma'_{i-1})$ and $\Omega_i = (\gamma_{i-1}, \gamma_i)$ and set to zero all elements of the matrix $\left(Q^2 \times Q^2\right)$ for which $\gamma'_{i-1} \neq \gamma_{i-1}$ [10]. This results in the transfer matrix

$$\widehat{G}(3)_{Potts} = \begin{pmatrix}
e^J & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots & 0 & 0 & \dots & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & 1 & \dots & 1 \\
1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & 0 & \dots & 0 & 1 & e^J & \dots & 1 & \dots & 0 & 0 & \dots & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & 1 & \dots & 1 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots & 0 & 0 & \dots & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & 1 & \dots & e^J
\end{pmatrix}. \quad (16)$$

At $\Delta = 4$ there are 4 spin variables and we add another 2, as above. The statistical weight is prescribed according to

$$g(\Omega_{i-1}\Omega_i) = e^{-\beta H_{Potts}} \delta(\gamma_{i-2}, \gamma'_{i-2}) \delta(\gamma_{i-1}, \gamma'_{i-1}). \quad (17)$$

The resulting transfer matrix has dimensions $(Q^3 \times Q^3)$ and is by its structure similar to eq. (16). For larger Δ's it is necessary to group γ's into Ω_{i-1} and Ω_i accordingly. The procedure can be generalized and the statistical weight would be written as

$$g(\Omega_{i-1}, \Omega_i) = e^{-H/T} \prod_{k=\Delta-1}^{1} \delta\left(\gamma'_{i-k}, \gamma_{i-k}\right), \quad (18)$$

resulting in a transfer matrix $\widehat{G}(\Delta)_{Potts}$ of dimensions $(Q^{\Delta-1} \times Q^{\Delta-1})$. This matrix differs from the GMPC case (see eq. (6)) in that all the diagonal elements are multiplied by e^J, while in eq. (6) only the element $(1,1)$ is multiplied by this factor. Here again, there are $a(Q, \Delta) = Q^{\Delta-1} - Q(\Delta - 1)$ identical rows and columns. The corresponding characteristic equation is obtained as a product of three terms and can be rewritten as

$$P_{Potts}(\lambda, W, Q, \Delta) = \lambda^{a(Q,\Delta)} \times p_{Potts}(\lambda, W, Q, \Delta) = 0, \quad (19)$$

or, after the elimination of trivial eigenvalues,

$$p_{Potts}(\lambda, W, Q, \Delta) = p_{GMPC}(\lambda, W, Q, \Delta - 1) \times p_{GMPC}(\lambda, W, 0, \Delta - 1)^{Q-1} = 0, \quad (20)$$

and $p_{GMPC}(\lambda, W, Q, \Delta)$ has been defined in eq. (10). Unfortunately, it is not possible to derive a single compact transfer matrix, as for the GMPC, that would correspond to such characteristic equation. According to eq. (20), the characteristic equation of the Potts model can be expressed in terms of the characteristic polynomials of the GMPC model. We have explicitly checked eq. (19) to be true up to $\Delta = 7$ by direct brute force calculations.

5 Results and discussion

5.1 Comparison of characteristic equations for the GMPC and Potts models

The transfer matrix, being a matrix of statistical weights, is non-negative. Then Frobenius-Perron theorem ensures that there exists a positive, non-degenerate largest eigenvalue λ_1. Upon solving the characteristic equation and assuming cyclic boundary conditions, we can evaluate the partition function as

$$Z(\lambda) = \lim_{N\to\infty} \sum_{i=1}^{\Delta} \lambda_i^N = \lambda_1^N, \tag{21}$$

the corresponding free energy

$$F(\lambda) = -TN \ln \lambda_1, \tag{22}$$

and the spatial correlation length as

$$\xi(\lambda) = \ln^{-1}\left(\frac{\lambda_1}{\lambda_2}\right), \tag{23}$$

where λ_1 is the largest and λ_2 is the second largest eigenvalue, implying of course that the thermodynamics of the model is completely determined by its characteristic equation.

The characteristic equation for the Potts model prompts a comparison with the $\Delta - 1$ particle GMPC model, the first bracket of eq. (20) being exactly the characteristic equation of $\Delta - 1$ particle GMPC (see eq. (13)). In the region of positive temperatures ($W > 1$) this first bracket has two positive roots, while the second bracket has a single, positive and $Q - 1$ times degenerate root.

Let us first study in more details the first bracket on the rhs of eq. (20) (which is eq. (13) with Δ to $\Delta - 1$ rescaling). It has two asymptotes: W and Q. We may rewrite it in the form, suitable for the iterative solution,

$$\lambda = Q + \frac{(\lambda - 1)(W - 1)}{\lambda^{\Delta-2}(\lambda - W) + W - 1}, \tag{24}$$

or, alternatively, as

$$\lambda = W + \frac{(\lambda - 1)(Q - 1)}{\lambda^{\Delta-2}(\lambda - Q) + Q - 1}. \tag{25}$$

An interesting feature of eq. (13), visible from its structure is that it is symmetrical towards the interchanges of W with Q, although the parameters have different physical meanings and different temperature dependencies. This is a reflection of a physical fact that it is the balance between the entropy (described by Q) and the energy (described by W) that determines the free energy of the system and its stability. From eqs. (24) and (25) one can immediately see that at high temperatures ($W \to 1$) the largest eigenvalue is approaching the Q asymptote from above, while at low temperatures ($Q \to 1$), the largest eigenvalue is approaching from above the W asymptote. Taking into account the symmetry of eq. (13), the second eigenvalue at low temperatures approaches the Q asymptote from below and at high temperatures it approaches the W asymptote from below. The point where the two asymptotes cross is the most interesting. According to the above consideration, the largest two eigenvalues resulting from the first bracket on the rhs of eq. (20) at $W = Q$ can be estimated as $Q \pm \epsilon$, where $\epsilon > 0$ is small. Inserting these estimates into the characteristic eq. (13) results in $\epsilon \approx Q^{\frac{4-\Delta}{2}}$, so that $\lambda \approx Q(1 \pm Q^{\frac{2-\Delta}{2}})$. For the second bracket on the rhs of eq. (20) we similarly obtain

$$\lambda = W - \frac{\lambda - 1}{\lambda^{\Delta-1} - 1}, \tag{26}$$

so that W is an asymptote for the second bracket too (which is not surprising, the structure of both equations is similar), but now the largest eigenvalue approaches the W asymptote from below. In the $W = Q$ point it can be estimated as $\epsilon_0 \approx Q^{2-\Delta}$.

Thus three largest eigenvalues from eq. (20) at $W = Q$ are: $\lambda \approx Q(1 \pm Q^{\frac{2-\Delta}{2}})$ (from the first bracket) and $\lambda \approx Q(1 - Q^{1-\Delta})$ (from the second bracket). Only the two largest eigenvalues are important to define the correlation length, so we need to order the solutions. The first is $\lambda \approx Q(1 + Q^{\frac{2-\Delta}{2}})$ and there are two candidates for the second: $Q - \epsilon$ and $Q - \epsilon_0$. Since $\epsilon_0 \approx Q^{2-\Delta}$ is always smaller than $\epsilon \approx Q^{\frac{4-\Delta}{2}}$, the following order for the three largest solutions of eq. (20) is established:

- $\lambda_1(\Delta) = \lambda_1^{GMPC}(\Delta - 1, Q)$ (from the first bracket);
- $\lambda_2(\Delta) = \lambda_1^{GMPC}(\Delta - 1, Q = 0)$ (from the second bracket);
- $\lambda_3(\Delta) = \lambda_2^{GMPC}(\Delta - 1, Q)$ (from the first bracket).

Such ordering resulted from qualitative consideration of eq. (20) at $W = Q$ point. Since both W and Q are the asymptotes and the distance between them increases at points far from $W = Q$, the *ordering by magnitude* for the eigenvalues remains the same. Thus we conclude for any $\Delta > 2$, $Q > 1$, $W > 1$ that, the eigenvalue from the second bracket is in between the two largest from the first brackets. To double check the ordering of eigenvalues, we have numerically solved eq. (20) for several values of Δ and have presented the results in fig. 3. The results of the numerical solution completely agree with the qualitative estimates above.

Two conclusions follow immediately. First, the one-dimensional Potts model with Δ many-body interactions has the same free energy as the $(\Delta - 1)$ many-body GMPC model (since determined by $\lambda_1(\Delta) = \lambda_1^{GMPC}(\Delta - 1, Q)$). All averages that depend on the largest eigenvalue, such as helicity degree or number of junctions [15,16], are therefore the same for both models. Second, correlation lengths are obviously different, since the second largest eigenvalues for the models differ. Indeed, as one can see from fig. 3, while the correlation lengths of both models coincide at low W (high temperatures), they markedly differ at intermediate to high values of W (low temperatures). The correlation length of the Potts model abruptly increases near the temperature, where energetic and entropic pa-

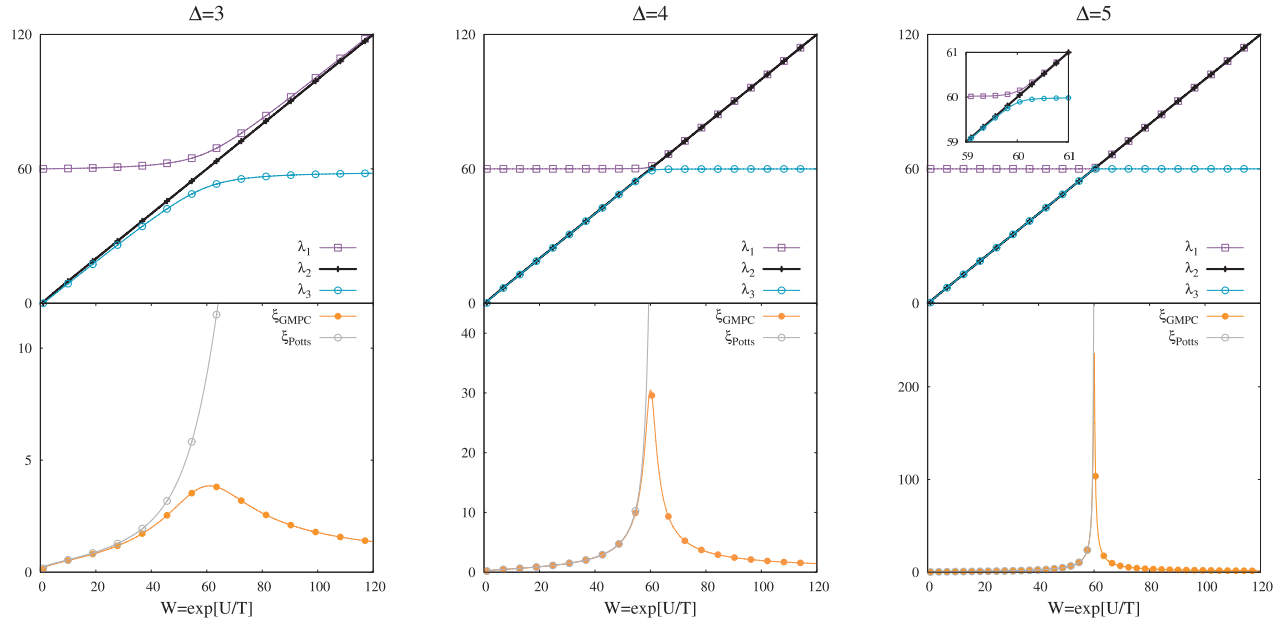

Fig. 3. Color online) (a) Three largest eigenvalues (top) from eq. (20) *vs.* $W = \exp[U/T]$ for $\Delta = 3$ and $Q = 60$. Correlation lengths (bottom) for Δ particle Potts and $\Delta - 1$ particle GMPC models. (b) The same for $\Delta = 4$ and (c) for $\Delta = 5$.

rameters compensate each other ($W = Q$), while the correlation length of the GMPC model passes through a maximum at this point. The monotonic growth of correlations with decreased temperature in the Potts model is a direct consequence of the absence of a preferred spin value.

5.2 Mapping between the Potts model and several biologically inspired models

Since the GMPC model has been shown to include several helix-coil models as particular cases [9,15,16,18,19], the established relation between the Potts and GMPC model is important to understand the mapping between the Potts model and several helix-coil models.

The analysis presented above shows that a Potts model with $\Delta = 2$ [29], formulated on the level of the effective free energy, is equivalent to the one-body (free-particle) GMPC model ($\Delta = 1$). Since the Zimm-Bragg model has been shown as originating from the $\Delta = 2$ GMPC model [9], this means that in order to achieve at least the same level of description as with the Zimm-Bragg classical model, the approach of ref. [29] should be extended to next-nearest-neighbor, three-body, interactions.

The characteristic equation of another model, widely applied for the helix-coil transition description, namely, the Lifson-Roig model can be derived from the GMPC model Hamiltonian with three-body interactions [15]. So a Potts model with $\Delta = 4$ would be necessary to achieve the equivalent description of the free energy.

It is worth noticing that the GMCP model is also related to another interesting model that has been frequently used in the framework of protein folding, namely the Wako-Saito-Munoz-Eaton (WSME) model [22–28]. Unlike the GMPC model, and quite similarly to the Zimm-Bragg model, the WSME model sets out from a phenomenological expression of the free energy. The methodology to pass from the GMPC Hamiltonian model to the corresponding free energy has already been elucidated in ref. [9], and it turns out that the resulting free energy bears strong similarities with the corresponding WSME one — being in fact equivalent for finite-range interactions, apart from an appropriate rescaling in the parameters. As this point appears to be interesting, it will be the subject of a future detailed analysis [35].

5.3 Explanation for the different behavior of correlation lengths of the GMPC and Potts models

From the largest eigenvalue, that determines the temperature dependence of helicity degree, both models proceed from the ordered conformation to the disordered one when the temperature is increased. At high temperatures (disordered conformation) the correlations in both models are equally small (fig. 3), while at intermediate to low temperatures the two correlations differ qualitatively.

Let us consider the behavior of fluctuations of spins in the range of intermediate to low temperatures in more details. The probability of fluctuations reads [36]

$$\omega \sim \exp[\Delta F] = \exp[\Delta E - T\Delta S], \qquad (27)$$

where ΔF is the change of free energy due to fluctuations, ΔE being the corresponding internal energy change and ΔS is the entropy change. It is important to remark that at the same values of correlation length, ΔE is the same

for the two models and the only relevant parameter is the entropy change.

For the Potts model, ΔS is comprised of two parts: a larger negative input stems from restricting spins in the ordered conformation, while a smaller positive input arises due to the degeneracy of the ordered conformation so that many spins can simultaneously flip without changing the order. Longer correlation lengths increase both the negative and positive inputs, keeping the overall entropy change in eq. (27) for the Potts model negative but small in magnitude. Thus there are enough fluctuations at any temperature and the correlation length increases as the temperature decreases.

In the case of the GMPC model the situation with fluctuations is different. Since there is a chosen and fixed value of spin orientation (*e.g.*, 1) corresponding to the ordered conformation, there is only one negative term in the change of entropy in eq. (27). As we can see in fig. 3, at a particular non-zero transition temperature the effect of the entropy sets in: along with the decrease in temperature the correlation length increases to a value corresponding to a very low entropy. Further increase of the correlation length becomes impossible (since it would mean zero fluctuations) and, to keep the fluctuations alive, the system has to decrease the correlation length. This is the physical origin of the maximum of correlation length appearing in the GMPC model.

In several classical models (*e.g.*, the Zimm-Bragg or Lifson-Roig ones) the correlation length is known to have a finite maximum, exactly as in the GMPC model and contrary to the Potts model. In this respect, the description provided by the GMPC model appears to be superior than that given by the Potts model.

In view of the results presented here, it would be very interesting to validate the difference in the correlation function of the two models by a direct comparison with experimental results. Unfortunately, we are not aware of such a study, to the best of our knowledge. On the other hand, the qualitative picture elucidated above strongly suggests the GMPC model to be more suitable to describe the onset of complex correlations close to the helix-coil transition in biopolymers. We hope that our study will stimulate further activities along these lines.

6 Conclusion

The work presented here deals with the description of the helix-coil transition, relevant for protein folding, in the vocabulary of spins, and compares the predictions of two models applied to this problem: the many-body Potts model and the many-body GMPC model. The low dimensionality of the many-body Potts model allowed us to explicitly construct the transfer matrix eq. (16) and to derive the characteristic equation for several finite Δs. The first and second roots of eq. (19) are, accordingly, the largest and second largest eigenvalues of the Potts model. The structure of the characteristic equation of the Potts model suggested that the first largest and third largest roots correspond to the first and second eigenvalues of the GMPC model with $\Delta - 1$ interacting spins. As a result, it was found that the free energy of Δ many-body Potts model is equal to that of the $\Delta - 1$ many-body GMPC model, and therefore the correlation lengths (which depend on the second largest eigenvalue as well) of the two models differ significantly. While the derivation of the secular equation of the many-body Potts model follows relatively standard protocols, it was not previously derived or analyzed. The question then arises of whether two models can be used interchangeably for the description of the helix-coil transition, as is tacitly assumed in the literature. The answer depends upon the required goal.

As a matter of fact, if the main aim is the fit of experimental data, in terms of the helicity degree and the heat capacity (both depending on λ_1), then the two models are essentially identical. However, if the goal is to study the qualitative effects of different external fields on the helix-coil transition, the Potts model is not optimal. Structurally the helical conformation is unique and very different from the coil conformation and translates into a preferred, set spin value. This feature is absent in the Potts model, thus resulting into an improper description of the corresponding correlations. There is one additional argument *against* the application of the Potts model when describing the helix-coil transition. The solvents, unavoidably present *in vivo* where the biopolymers operate, can be effectively regarded as external fields. While the external field can be easily renormalised in the GMPC vocabulary [18], its Potts counterpart becomes very complicated and often analytically untractable. It thus seems that the application of the GMPC model as opposed to the Potts model would be more adequate, as these examples show explicitly. The difference of temperature dependence in the behavior of correlation lengths is a very interesting and highly non-trivial result, and it would be interesting to see if it persists at higher spatial dimensions.

Interest in 2D and 3D spin models surged mainly in order to describe bulk properties of magnetic materials, which show ferromagnetic-paramagnetic (phase) transitions. In this context the study of 1D models, which do not show any phase transition at a finite temperature in the absence of long-range interactions [36], makes no sense. The situation in polymer-related systems is exactly opposite as many are essentially low dimensional, thus making the study of 1D models of extreme relevance. Since no phase transition happens, no universality in behavior is observed, and choosing a 1D model that describes the physical situation in a proper way is of paramount necessity. As a deep and consequential analogy between magnetic and polymer systems is well established and long known, we believe that there are likewise situations in the theory of magnetism, where application of the GMPC model instead of the standard Potts model would lead to a better understanding of the thermodynamic and magnetic properties.

AB and AG acknowledge the support from PRIN-COFIN 2010-2011 2010LKE4CC grant. RP and AB acknowledge ARRS grants P1-0055 and J1-4297.

References

1. R.J. Baxter, in *Exactly Solved Models in Statistical Mechanics* (Academic Press, 1982).
2. F.Y. Wu, Rev. Mod. Phys. **54**, 3720 (1982).
3. A. Imambekov, T.L. Schmidt, L.I. Glazman, Rev. Mod. Phys. **84**, 1253 (2012).
4. V.V. Deshpande, M. Bockrath, L.I. Glazman, Amir Yacoby, Nature **464**, 209 (2010).
5. C. Cantor, T. Schimmel, *Biophysical Chemistry* (Freeman and Co., San Francisco, 1980).
6. D. Poland, H. Scheraga, *The Theory of Helix-Coil Transition* (Academic Press, New York, 1970).
7. A. Ahsan, J. Rudnick, R. Bruinsma, Biophys. J. **74**, 132 (1998).
8. O. Punkkinen, P.L. Hansen, L. Miao, I. Vattulainen, Biophys. J. **89**, 967 (2005).
9. A. Badasyan, A. Giacometti, Y. Sh. Mamasakhlisov, V.F. Morozov, A.S. Benight, Phys. Rev. E **81**, 021921 (2010).
10. P.J. Flory, *Statistical Mechanics of Chain Molecules* (Interscience, New York, 1969).
11. G.N. Ramachandran, G. Ramacrishnan, V. Sasisekharan, J. Mol. Biol. **7**, 96 (1963).
12. V. Renugopalakrishnan, Lakshminarayanan, V. Sasisekharan, Biopolymers **10**, 1159 (1971).
13. V.F. Morozov, E.Sh. Mamasakhlisov, Sh. Hayryan, Chin-Kun Hu, Phys. A **281**, 51 (2000).
14. N.S. Ananikyan, E.Sh. Mamasakhlisov, V.F. Morozov, Z. Phys. Chem. (Leipzig) **27**, 603 (1990).
15. N. Ananikyan, Sh. Hayryan, E. Mamasakhlisov, V. Morozov, Biopolymers **30**, 357 (1990).
16. Sh. Hayryan, E.Mamasakhlisov, V. Morozov, Biopolymers **35**, 75 (1995).
17. J.S. Schrek, J.M. Yuan, Phys. Rev. E **81**, 061919 (2010).
18. A. Badasyan, Sh. A. Tonoyan, Y. Sh. Mamasakhlisov, A. Giacometti, A.S. Benight, V.F. Morozov, Phys. Rev. E **83**, 051903 (2011).
19. A. Badasyan, Sh.A. Tonoyan, A. Giacometti, R. Podgornik, V.A. Parsegian, Y.Sh. Mamasakhlisov, V.F. Morozov, Phys. Rev. Lett. **109**, 068101 (2012).
20. B.H. Zimm, J.K. Bragg, J. Chem. Phys. **31**, 526 (1959).
21. S. Lifson, A. Roig, J. Chem. Phys. **34**, 1963 (1961).
22. H. Wako, N. Saito, J. Phys. Soc. Jpn. **44**, 1931 (1978).
23. H. Wako, N. Saito, J. Phys. Soc. Jpn. **44**, 1939 (1978).
24. V. Munoz, P.A. Thompson, J. Hofrichter, W.A.Eaton, Nature **390**, 196 (1997).
25. V. Munoz, E.R. Henry, J. Hofrichter William, A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **95**, 5872 (1998).
26. V. Munoz, W.A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **96**, 11311 (1999).
27. P. Bruscolini, A. Pelizzola, Phys. Rev. Lett. **88**, 258101 (2002).
28. M. Zamparo, A. Pelizzola, Phys. Rev. Lett. **97**, 068106 (2006).
29. R.E. Goldstein, Phys. Lett. A **104**, 285 (1984).
30. B. Alberts, D. Bray, J. Lewis, K. Roberts, D. Watson, *Molecular Biology of the Cell*, Vol. **1** (Garland Publishing Inc., New York, London, 1983).
31. V.G. Dashevsky, *Conformational Analysis of Macromolecules* (Nauka Publishers, Moscow, 1987) (in Russian).
32. A.Yu. Grosberg, A.R. Khokhlov, *Statistical Physics of Macromolecules* (AIP Press, New York ,1994).
33. M.V. Volkenstein, in *Configurational Statistics of Polymeric Chains* (Wiley Interscience, 1963).
34. L. Pauling, R.B. Corey, H.R. Bronson, Proc. Natl. Acad. Sci. USA **57**, 205 (1951).
35. A. Badasyan, A. Giacometti, R. Podgornik, Y. Mamasakhlisov, V. Morozov, unpublished results.
36. L. Landau and E. Lifshits, *Statistical Physics* (Pergamon, Oxford, 1988).