Chemically controlled unfolding of a RNA-like polymer model

C. Buzano

Dipartimento di Scienze Applicate e Tecnologia (DISAT), Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy

M. Pretti

Consiglio Nazionale delle Ricerche-Istituto dei Sistemi Complessi (CNR-ISC), Dipartimento di Scienze Applicate e Tecnologia (DISAT), Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy

We consider a lattice polymer model of the two-tolerant type (i.e., a random walk allowed to visit lattice bonds at most twice), in which doubly visited bonds yield an attractive energy term (pairing energy). Such a model has been previously proposed as a rough, nonspecific description of the RNA folding mechanism. Indeed, the model predicts, besides the usual theta collapse, an extra transition to a low-temperature fully paired state. In the current work, we propose an extension of the model, in which a "micromolecular" chemical species can bind the polymer and locally forbid segment pairing. We investigate equilibrium thermodynamics in the grand-canonical picture, at the level of a Bethe approximation, which is, a refined mean-field technique, equivalent to the exact solution on a random-regular graph. The general trend we observe is that expected from the mechanism implemented in the model (increasing micromolecule concentration favors unfolding and lowers the transition temperature), but the resulting phase diagram turns out to be remarkably interesting and rich.

I. INTRODUCTION

In the framework of statistical mechanics, the description of essential properties of macromolecules usually relies on their universality, that is, on the fact that such properties are, to a certain degree, independent of microscopic details of the model [1–3]. For instance, a linear homopolymer in good solvent can be simply described as a self-avoiding walk on a regular lattice (i.e., a random walk which is not allowed to visit lattice sites more than once), whereas an effective monomer-monomer attractive interaction allows one to take into account the quality of the solvent and predict the well-known coil-globule (collapse) transition.

In the last few years, a number of studies have been addressed to investigate whether and how the microscopic interaction details may affect qualitative features of the model, such as the nature (e.g., order, universality class) or the very existence of a collapse transition. For example, it has been pointed out that an attractive interaction between neighbor lattice bonds (rather than sites) visited by the walk may give rise to an extra transition to a low-temperature anisotropic phase [4–8]. Moreover, different transition scenarios may be observed in models which relax the self-avoidance constraint (allowing for multiply occupied lattice sites), if immediate self-reversal of the walk is allowed or not [9–12].

Such and similar effects have been sometimes investigated even in the absence of a specific physical system to be described. Nevertheless, a special attention has been devoted to models which relax the self-avoidance constraint by allowing for doubly visited lattice bonds (rather than sites) because of the speculated possibility of mimicking in this way, at a coarse-grained level, base pairing which takes place in DNA or RNA. This last kind of models, including indeed different possible variants, is generally denoted as "two-tolerant" [13–18] and have been investigated by different techniques, such as exact enumerations [13], Monte Carlo simulations [14,15], scaling arguments [15], and approximate semianalytical techniques [16–18].

In principle, RNA is a long linear heteropolymer, composed of four different monomers or bases (adenine, cytosine, guanine, and uracil), such that certain base pairings (namely, adenine-uracil, guanine-cytosine, and guanine-uracil) are energetically favored by the formation of hydrogen bonds [19]. Therefore, neglecting the heterogeneity of the sequence and assigning an attractive (contact) energy to any possible base pair is an extreme simplification, which, of course, cannot explain properties of specific RNA sequences. Nevertheless, it is tempting to identify the low-temperature fully paired phase appearing in these models as a rough description of the RNA native state, and the corresponding phase transition with the RNA melting.

Quite recently, a lattice polymer model of the two-tolerant type has been endowed with a suitable thermodynamic parametrization [20], along the lines of the well-known Turner model [21], in order to account for the specificity of RNA base pairing. The cited work demonstrates the possibility of enhancing the standard tools for RNA secondary structure prediction by coarse-grained information about the three-dimensional structure, thus enabling also the prediction of tertiary contacts. Of course, the resulting model is not amenable to simple analytical treatments, such as the one presented in the current paper.

In a previous work [16], one of us investigated a (homogeneous) two-tolerant polymer model by means of the Bethe approximation, i.e., a refined mean-field technique, corresponding to the exact solution on the so-called Bethe lattice. Here we consider an extension of that model, aimed at describing a long polymer chain in a solution containing also a certain amount of identical short polymers. As far as the long chain is concerned, this is precisely the two-tolerant polymer of the previous paper [16], whereas the short polymers are assumed to have no internal structure, and to

fill each one a single lattice bond. Due to these features, in the following we shall simply denote the short polymers as *micromolecules*, in contrast to the long polymer chain. We introduce a special (micromolecule-polymer binding) interaction term, which influences the local ability of the long polymer to form double strands. Namely, we assume that, when the same lattice bond is occupied by a micromolecule and by a polymer segment (i.e., a micromolecule binds the polymer), no other base pairing can occur in the neighbor lattice bonds. We also consider an excluded-volume constraint to be satisfied by the micromolecules. Still in the framework of the Bethe approximation, we investigate the changes in the phase diagram, with respect to the "pure polymer" model, depending on the micromolecule concentration and the strength of the binding interaction. It turns out that the former is the more relevant parameter.

The paper is organized as follows. In Sec. II we define the model in full detail, and in Sec. III we report the Bethe lattice calculation. In Sec. IV we present and discuss the thermodynamic properties of the model, whereas Sec. V contains some concluding remarks. Further details of the analytical calculations are reported in two appendices.

II. THE MODEL

As mentioned in the Introduction, we consider a lattice model of a linear polymer in a mixture of solvent and micromolecules.

The pure polymer model is equivalent to the one studied in a previous paper [16], i.e., a two-tolerant lattice walk. Here *two-tolerant* means that each lattice bond can be visited (at most) twice. A doubly visited lattice bond represents pairing between polymer segments, which is assumed to yield an attractive energy $-\beta$ (pairing energy). Some additional constraints are as follows: (1) two paired segments cannot be consecutive along the polymer chain; (2) if a lattice site is visited more than once, every extra visit, besides the first one, must be associated to pairing in some incident lattice bond. All these constraints are local in nature, as they can be imposed at the level of each "star" cluster (by this term, we shall denote the set of lattice bonds incident to a given site). The allowed configurations of a star cluster for the pure polymer model are summarized in the first two stages of Table I. A simple inspection reveals that these configurations mimic the building blocks of typical motifs appearing in RNA structures (hairpin loops, internal loops, bulge loops, multibranch loops) [21]. The possibility of multibranch structures that are shrunk to a single site is also allowed. Note that the model studied in Ref. [16] takes into account also a *stacking* effect, that is, an extra energy term $-\gamma$ favoring longer double-stranded regions. In the current paper we shall consider only the basic version of the model and discard this term.

A micromolecule can bind any polymer segment, provided the latter is not yet paired to another polymer segment, and this process is assumed to yield an attractive energy $-\epsilon$ (binding energy). Micromolecules impose some additional strong constraints as well, namely, a micromolecule occupying a given lattice bond forbids both the occurrence of segment-segment pairing and the presence of other micromolecules on nearby lattice bonds (i.e., all lattice bonds incident to

TABLE I. Allowed configurations of a star cluster (leftmost column) and corresponding occupation numbers m_i, n_i for each bond $i = 0, \ldots, k$ (other columns). Thin lines denote lattice bonds (the graphical representation is limited to four bonds); thick solid lines denote polymer segments; thick dashed lines denote micromolecules. The top rows report pure-polymer configurations (partially unpaired configurations in the first stage, and fully paired configurations in the second stage). The last three rows report extra configurations including one micromolecule (but no more), admitted in the current model. Configurations obtained by bond permutations are allowed as well. The all-empty-bonds configuration is also allowed but not illustrated.

conf.	m_0, n_0	m_1, n_1	m_2, n_2	m_3, n_3	$\ldots$	m_k, n_k
	$0, 1$	$0, 1$	$0, 0$	$0, 0$	$\ldots$	$0, 0$
	$0, 1$	$0, 1$	$0, 2$	$0, 0$	$\ldots$	$0, 0$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$0, 1$	$0, 1$	$0, 2$	$0, 2$	$\ldots$	$0, 2$
	$0, 2$	$0, 2$	$0, 0$	$0, 0$	$\ldots$	$0, 0$
	$0, 2$	$0, 2$	$0, 2$	$0, 0$	$\ldots$	$0, 0$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$0, 2$	$0, 2$	$0, 2$	$0, 2$	$\ldots$	$0, 2$
	$1, 0$	$0, 0$	$0, 0$	$0, 0$	$\ldots$	$0, 0$
	$1, 0$	$0, 1$	$0, 1$	$0, 0$	$\ldots$	$0, 0$
	$1, 1$	$0, 1$	$0, 0$	$0, 0$	$\ldots$	$0, 0$

the end-sites of the given bond). Even these constraints are local and can be imposed at the level of star clusters. The allowed configurations of a star cluster in the presence of a micromolecule are displayed in the last stage of Table I. According to the constraints, only one micromolecule and a single-stranded chain stretch appear.

We consider a grand-canonical formulation of the problem: Chemical potentials μ and μ' are associated with a polymer segment and a micromolecule, respectively, while the solvent chemical potential is conventionally assumed to be zero.

A configuration of the system is defined by specifying the number of polymer segments and/or micromolecules on each lattice bond. Therefore, we define, for the ith lattice bond, two configuration variables: $m_i = 0, 1$ (micromolecule

occupation number) and $n_i = 0, \ldots, 2 - m_i$ (segment occupation number). Note that the upper bound of the segment occupation number depends on the micromolecule occupation number, because the presence of a micromolecule on a given lattice bond implies that no more than one polymer segment can occupy the same bond. As mentioned above, further constraints must be satisfied at the level of star clusters. As far as the segment occupation numbers n_i are concerned, these constraints are precisely the same as those of the pure polymer model [16]. Taking into account the micromolecules, we have to impose the following extra conditions: (1) the total number of micromolecules on each star cluster can just be 0 or 1; (2) in the latter case, the total number of polymer segments on the star cluster can be at most 2. The polymer constraints themselves impose that this number can thus be either 0 or 2, and, in the latter case, that the segments must be unpaired [16].

Assuming a coordination number $k + 1$, the Hamiltonian can be formally written in the same way as the pure polymer model's, namely,

$$\tilde{\mathcal{H}} = \sum_{\{i_0, \ldots, i_k\}} \tilde{H}_{m_{i_0}, n_{i_0}, \ldots, m_{i_k}, n_{i_k}} + \sum_i \tilde{h}_{m_i, n_i}, \quad (1)$$

where the former sum runs over all star clusters $\{i_0, \ldots, i_k\}$, and the latter over all bonds i. Single-bond energy terms $\tilde{h}_{m,n}$ take into account pairing and binding energies, and chemical potential contributions. They can be defined as follows:

$$\tilde{h}_{0,0} = 0, \quad (2)$$

$$\tilde{h}_{0,1} = -\mu, \quad (3)$$

$$\tilde{h}_{0,2} = -\beta - 2\mu, \quad (4)$$

$$\tilde{h}_{1,0} = -\mu', \quad (5)$$

$$\tilde{h}_{1,1} = -\epsilon - \mu - \mu'. \quad (6)$$

Note that for $m = 0$ (i.e., no micromolecules), we recover the single-bond term of the pure polymer model, namely, $\tilde{h}_{0,n} = h_n$, where h is defined in Eqs. (2)–(4) of Ref. [16] (with $\gamma = 0$). Many-bond ("star") terms $\tilde{H}_{m_{i_0}, n_{i_0}, \ldots, m_{i_k}, n_{i_k}}$ take into account the star-cluster constraints, assigning infinite energy to forbidden configurations and zero energy to allowed configurations. We have

$$\tilde{H}_{m_{i_0}, n_{i_0}, \ldots, m_{i_k}, n_{i_k}} = H_{n_{i_0}, \ldots, n_{i_k}} + H'_{m_{i_0}, n_{i_0}, \ldots, m_{i_k}, n_{i_k}}, \quad (7)$$

where H is the corresponding term for the pure polymer model (defined in Ref. [16]), and H' is a "perturbation" term, implementing the constraints associated to the presence of a micromolecule. The pure polymer term is defined as 0, if the set of segment occupation numbers $\{n_{i_0}, \ldots, n_{i_k}\}$ corresponds to one of the allowed pure-polymer configurations (see Table I), and ∞ otherwise. Moreover, according to the above discussion, the perturbation term can be written as

$$H'_{m_{i_0}, n_{i_0}, \ldots, m_{i_k}, n_{i_k}} = \begin{cases} 0, & \text{if} \quad \sum_{i=0}^k m_i = 0; \\ 0, & \text{if} \quad \sum_{i=0}^k m_i = 1 \\ & \text{and} \quad \sum_{i=0}^k n_i = 0, 2; \\ \infty, & \text{otherwise.} \end{cases} \quad (8)$$

III. THE BETHE APPROXIMATION

The Bethe (or *cavity*) approximation [22,23] may be generally regarded as the exact solution on the so-called *Bethe lattice*, the latter being defined as a *random-regular graph* (i.e., a random graph with fixed coordination number). In the thermodynamic limit, this kind of lattice is locally treelike [24], since the length of close paths diverges as the logarithm of the number of sites. Therefore, thermodynamic equilibrium states can be determined in a heuristic manner, assuming that the lattice structure is actually that of a tree graph, together with a self-consistency requirement, which leads to a recursion equation [25] (basically as in Bethe's original paper [26]). It is then possible to show that the thermodynamic states computed along these lines coincide with the minima of a variational free energy functional, called *Bethe free energy* [27], though actually derived for the first time by Guggenheim [28].

Note that, as far as polymer models are concerned, there is an extra approximation, due to the fact that the connectivity constraints of the polymer chain are taken into account only at a local level, so that, shortly speaking, a single self-interacting polymer is replaced by an ensemble of interacting polymers. In spite of such an approximation, several examples have demonstrated that this kind of calculation generally yields qualitatively reliable results (more accurate than pure mean field) for the model phase diagram [5,8,29]. For instance, let us mention the case of the "attracting-bond" self-avoiding walk [5,8], whose phase diagram exhibits several analogies with that of our (pure-polymer) model. Namely, in that case, the phase diagram has been first predicted by the Bethe approximation [5] and subsequently confirmed (at least in the two-dimensional case) by a transfer matrix approach [8].

On the other hand, let us stress the fact that the Bethe approximation necessarily fails in the description of loop statistics, because, as previously mentioned, the Bethe lattice is mostly treelike, whereas closed lattice paths exhibit a diverging length (logarithmic in the system size). This is the reason why in Ref. [16] we explicitly introduced, as a partial cure for this difficulty, the multibranched configurations shown in Table I. However, since loop statistics turn out to have a deep impact on the thermodynamics of RNA folding (as has been clearly pointed out in a nice paper by Einert and coworkers [30]), we can conclude that special cautions are in order for the interpretation of our results in terms of real RNA properties.

In its general form, the recursion equation associated with (the Bethe approximation of) the current model is very similar to that of the pure polymer model and can be deduced in an analogous way. In the treelike assumption, if a given star term of the Hamiltonian (1) is removed, the system splits into $k + 1$ disconnected branches. Thus, one can ideally study the Boltzmann distribution of each isolated branch (let us consider, for instance, the one depicted in Fig. 1 with labeled bonds) and compute the probability distribution w_{m_0, n_0} of the configuration variables m_0, n_0 in the "root" bond (the 0 bond in Fig. 1). These probabilities, which have to satisfy the normalization condition

$$\sum_{m_0=0}^1 \sum_{n_0=0}^{2-m_0} w_{m_0, n_0} = 1, \quad (9)$$
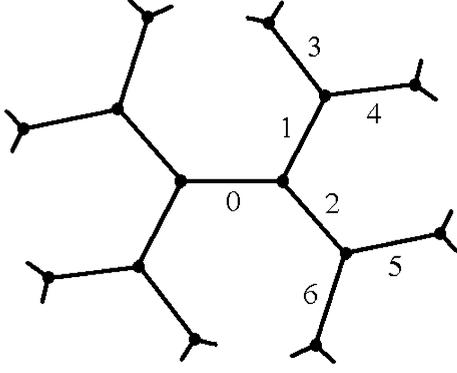
FIG. 1. Sketch of a Bethe lattice with $k = 2$.

are usually denoted as *normalized partial partition functions* [25]. In the thermodynamic limit, and in the hypothesis of a homogeneous thermodynamic state, the subbranches attached to the root bond should be equivalent to "primary" one. Such a self-consistency requirement yields the recursion equation

$$w_{m_0,n_0} = q^{-1}e^{-\tilde{h}_{m_0,n_0}} \sum_{m_1=0}^{1} \sum_{n_1=0}^{2-m_1} w_{m_1,n_1} \cdots$$
$$\cdots \sum_{m_k=0}^{1} \sum_{n_k=0}^{2-m_k} w_{m_k,n_k} e^{-\tilde{H}_{m_0,n_0,m_1,n_1,\ldots,m_k,n_k}}, \quad (10)$$

where q is a normalization factor, the sum runs over configuration variables of bonds attached to the 0 bond (m_1,n_1 and m_2,n_2, in the example of Fig. 1), and the energy terms $\tilde{h}_{m_0,n_0}$ and $\tilde{H}_{m_0,n_0,\ldots,m_k,n_k}$ are understood to incorporate the inverse temperature. Note that this equation is quite general: The explicit equations characterizing the current model can be obtained by taking into account the specific energy terms described in the previous section, which is done in Appendix A. The recursion equation can be solved numerically by a simple iterative (fixed-point) algorithm. All equilibrium properties of the system can be derived from the knowledge of the partial partition functions.

First, we can compute the joint probability distribution $p_{m,n}$ of micromolecule and segment occupation numbers m,n for a generic lattice bond. Consistently with the treelike hypothesis, if a given lattice bond is cut, the system splits into two disconnected branches. Therefore, the configuration variables associated to each given bond "feel" the effect of two independent branches. We thus obtain

$$p_{m,n} = z^{-1}e^{\tilde{h}_{m,n}} w_{m,n}^2, \quad (11)$$

where

$$z = \sum_{m=0}^{1} \sum_{n=0}^{2-m} e^{\tilde{h}_{m,n}} w_{m,n}^2 \quad (12)$$

provides normalization. Note that, according to Eq. (10), the partial partition function w_{m_0,n_0} incorporates the Boltzmann factor of the single-bond energy term $\tilde{h}_{m_0,n_0}$ associated with the root bond. Accordingly, the exponential factor in Eq. (11) is meant to avoid double counting.

The average number of polymer segments per lattice bond, which in the following we shall denote as segment density, can

be evaluated as

$$\rho = \sum_{m=0}^{1} \sum_{n=0}^{2-m} n p_{m,n} = p_{0,1} + 2p_{0,2} + p_{1,1}. \quad (13)$$

Similarly, the average number of micromolecules per lattice bond (micromolecule density) is

$$\rho' = \sum_{m=0}^{1} \sum_{n=0}^{2-m} m p_{m,n} = p_{1,0} + p_{1,1}. \quad (14)$$

Two more properties, useful to characterize the model behavior, are the fraction of paired polymer segments

$$\phi = 2p_{0,2}/\rho, \quad (15)$$

and the fraction of polymer segments bound to a micromolecule

$$\phi' = p_{1,1}/\rho. \quad (16)$$

As in the pure polymer model, the segment density ρ and the fraction of paired segments ϕ act as order parameters for our system: These two quantities are sufficient to define the phase diagram. Let us note, however, that in this case we need the two extra parameters ρ' and ϕ' for a full characterization of the normalized partial partition functions, the latter being indeed a set of five values $w_{0,0}, w_{0,1}, w_{0,2}, w_{1,0}, w_{1,1}$, only four of which are independent (because of the normalization condition).

The grand-canonical free energy (grand-potential) per lattice bond, ω, can be easily computed as a function of the normalization constants q [of the recursion equation (10)] and z [of Eq. (12)]:

$$\omega = -\frac{2\ln q - (k-1)\ln z}{k+1}. \quad (17)$$

A derivation of this expression, relying on the variational Bethe free energy, was reported in the previous paper [16]. In the presence of multiple solutions (i.e., fixed points) of the recursion equation, which corresponds to the occurrence of coexistence phenomena, the knowledge of the grand potential allows us to discriminate the thermodynamically stable phase, and therefore to determine first-order transitions. Conversely, second-order transitions can be better detected by analyzing the *stability* of the solutions. The latter is a rather technical issue, which we discuss in Appendix B.

IV. RESULTS

Let us first recall that, in the grand-canonical formulation of a polymer problem, the segment chemical potential μ controls the average chain length. Taking the self-avoiding walk model as a reference [3], we expect to observe a phase transition at which (for increasing μ) the average length diverges. Such a transition may be identified with the thermodynamic limit of a single polymer chain. The system can also be described in terms of the segment density ρ, the finite-length phase being characterized by $\rho = 0$, and the infinite-length phase by $\rho > 0$ (dense phase). In the limit of μ tending to the transition from above, the properties of the dense phase approach those of a single chain, and, in particular, the segment density ρ is a measure of the chain compactness. Therefore, a continuous transition represents an extended (coil) state, whereas a

discontinuous transition represents a collapsed (globule) state. For a self-avoiding walk with attractive monomer-monomer interaction, lowering the temperature drives the system from the former to the latter regime, giving rise to the so-called Θ collapse. For that model (sometimes simply denoted as Θ model), the transition between the two regimes is continuous, which corresponds to a tricritical point (Θ point) in the grand-canonical phase diagram.

As described in Sec. II, the current model is completely specified by four energy parameters, namely, the pairing energy β, the binding energy ϵ, and the segment and micromolecule chemical potentials μ, μ'. According to our notation, the energy parameters are normalized to temperature. Therefore, defining the pairing interaction as energy unit, $1/\beta$ plays the role of temperature, whereas μ/β and μ'/β can be regarded as the "true" chemical potentials. All the results, reported in the following, have been obtained for equal pairing and binding energies ($\epsilon = \beta$). Indeed, we have observed that the qualitative predictions of the model are quite robust, with respect to even rather large variations of the binding energy parameter.

The lattice is characterized (in the framework of the Bethe approximation) by the single parameter k. We set $k = 11$ for consistency with the previous paper [16], in which this choice was meant to reproduce the coordination number of a face-centered cubic lattice ($k + 1 = 12$).

The model basically exhibits three different phases, similar to those of the pure polymer model [16]: a zero-segment-density phase (O), an "ordinary" dense phase (I), and a fully paired dense phase (II). The O phase corresponds to the finite-length phase of the Θ model, as it is characterized by $\rho = 0$. This implies that both the fraction of paired segments ϕ and the fraction of segments bound to a micromolecule ϕ' are 0/0 undetermined forms. Conversely, the micromolecule density ρ' takes a finite value, which depends on the corresponding chemical potential μ'. As shown in Appendix A, the (inverse) analytical relationship between these quantities can be determined as

$$\mu' = \ln \frac{\rho'_\circ(1 - \rho'_\circ)}{[1 - (k+1)\rho'_\circ]^2}, \qquad (18)$$

where the subscript reminds that this equation is specific for the O phase. We can identify this phase with a simple micromolecule solution. The I phase is characterized by $0 < \rho < 2$ and $0 < \phi < 1$; i.e., it is a dense phase with a less-than-unit fraction of paired segments. Finally, the II phase is characterized by $0 < \rho < 2$ and $\phi = 1$; i.e., it is a phase in which all (but a subextensive number) of segments are paired.

Hereafter we report a sequence of phase diagrams in the segment chemical potential vs temperature plane. The micromolecule chemical potential is controlled by the micromolecule concentration of the O phase via Eq. (18), and each diagram corresponds to a fixed value of $\rho'_\circ$. This parametrization is motivated by the fact that the concentration has a clearer experimental counterpart, with respect to the chemical potential.

Figure 2 displays the phase diagram for $\rho'_\circ = 0.02$. This value can be thought of as a low-concentration regime, because the phase diagram is qualitatively similar to that of the pure polymer model ($\rho'_\circ = 0$). The transition line between the O
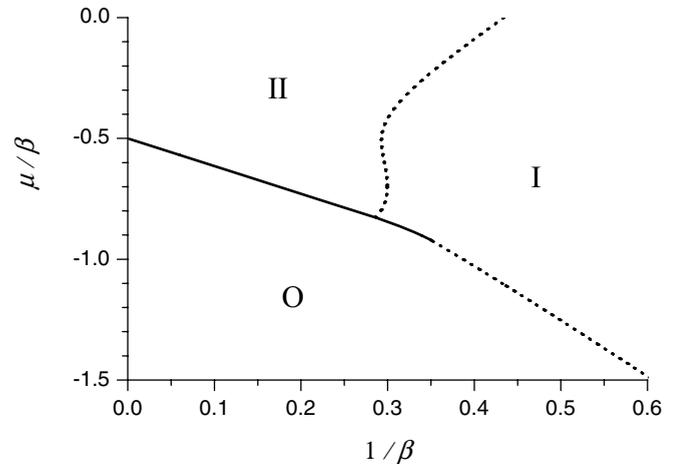
FIG. 2. Segment chemical potential vs temperature (μ/β vs $1/\beta$) phase diagram for $\epsilon/\beta = 1$ and $\rho'_\circ = 0.02$. Solid and dashed lines denote first and second order transitions, respectively. The different phases are denoted by their own tags: O, I, II (see the text).

and I phases is partially first and partially second order. As mentioned in Appendix B, the second-order portion can also be determined analytically as

$$\mu = \ln \frac{\sqrt{(1 - 2\rho'_\circ)^2 + 4\rho'_\circ(1 - \rho'_\circ)e^\epsilon} - (1 - 2\rho'_\circ)}{2k\rho'_\circ e^\epsilon}, \qquad (19)$$

where ϵ is proportional to the inverse temperature β ($\epsilon = \beta$ in the special case under investigation). The first-order and second-order regimes are separated by a Θ-like point. Indeed, as already mentioned in Ref. [16], it has been established that, for lattice polymer models of the two-tolerant type, this point exhibits special critical exponents [13], different from those of the ordinary Θ model. Nonetheless, since this point still corresponds to a continuous collapse (and exponent differences cannot be detected at the level of Bethe approximation), we shall speak of a Θ point as well in the following. The transition line between the O and II phases is always first order, whereas, in the dense region ($\rho > 0$), a second-order transition line separates the I and II phases. The latter line terminates in a critical end point at the O phase boundary.

Upon increasing the micromolecule concentration, the phase diagram undergoes dramatic changes. Figure 3 refers to the case $\rho'_\circ = 0.03$. The Θ point disappears, so that the continuous part of the O-I transition terminates in a critical end point. The latter corresponds to a discontinuous transition in the single-chain limit, as discussed in more detail below. The first-order portion of the O-I transition partially remains and partially becomes a branch within the I phase, which separates two regions at different densities, and terminates in a critical point. The lower temperature part of the phase diagram is qualitatively unchanged, as the II phase is bounded by a continuous transition toward the I phase and by a first-order transition toward the O phase.

For even larger micromolecule concentration, the critical end point of the O-I transition line moves toward lower temperature. At $\rho'_\circ = 0.04$ (Fig. 4) this point lies along the first-order boundary of the II phase, so that the first-order
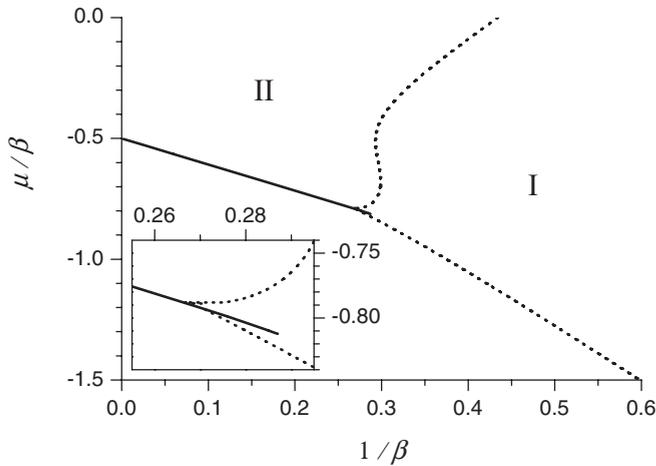
FIG. 3. Segment chemical potential vs temperature (μ/β vs $1/\beta$) phase diagram for $\epsilon/\beta = 1$ and $\rho'_\circ = 0.03$. Lines and tags as in Fig. 2 (the O phase region is left blank). The inset displays the region around the two critical end points.

portion of the O-I transition disappears, being replaced by a direct first-order I-II transition. At $\rho'_\circ = 0.06$ (Fig. 5) the critical O-I line extends down to zero temperature, so that a direct O-II transition no longer exists.

The consequences of all these changes to the single-chain limit are discussed in the remainder of this section. Let us recall that such a limit is the relevant one to characterize the properties of a single two-tolerant polymer chain interacting (in the hypotheses of the model) with an assembly of micromolecules in solution. In particular, we shall discuss the temperature behavior of the fractions of paired and bound polymer segments, ϕ and ϕ' (computed for μ tending to the O phase boundary from above), for fixed values of the micromolecule concentration $\rho'_\circ$. Of course, it is meaningful to consider the same concentration values, for which we have previously reported the grand-canonical phase diagrams.

Figure 6 refers to the case $\rho'_\circ = 0.02$, which, as previously mentioned, can be regarded as a low-concentration regime,
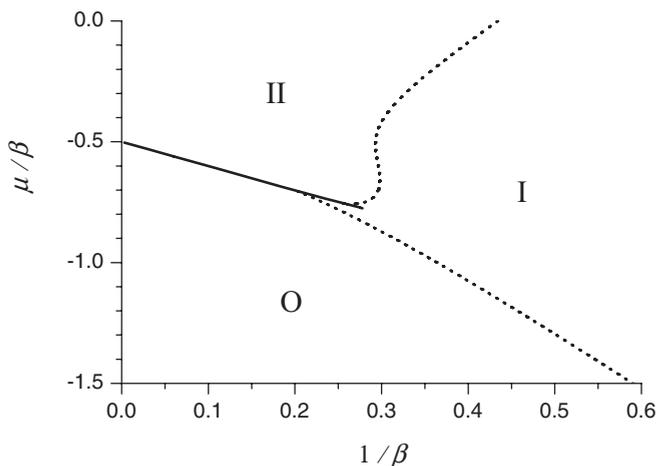
FIG. 4. Segment chemical potential vs temperature (μ/β vs $1/\beta$) phase diagram for $\epsilon/\beta = 1$ and $\rho'_\circ = 0.04$. Lines and tags as in the previous figures.
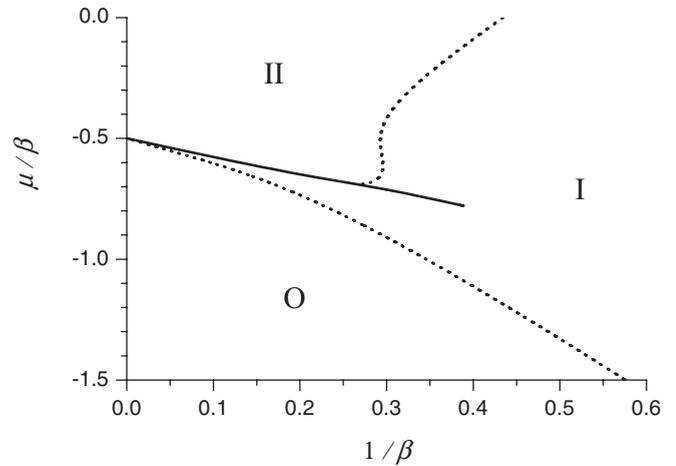
FIG. 5. Segment chemical potential vs temperature (μ/β vs $1/\beta$) phase diagram for $\epsilon/\beta = 1$ and $\rho'_\circ = 0.06$. Lines and tags as in the previous figures.

in which the phase diagram is qualitatively similar to that of the pure polymer model. The Θ temperature and the critical-end-point temperature turn out to decrease upon increasing the micromolecule concentration. We can see that the fraction ϕ of paired segments is rigorously zero above the Θ temperature. In this region, which we can identify with the *coil* state, the polymer chain behaves like an ordinary self-avoiding walk without self-interaction. Conversely, the fraction ϕ' of segments bound to micromolecules increases upon decreasing temperature, since binding is energetically favored. Upon decreasing temperature below the Θ point (collapse transition), ϕ begins to increase, whereas ϕ' inverts the previous trend and begins to decrease. Of course, these effects are strictly related to each other, but their interplay is likely to be nontrivial. Indeed, on the one hand the polymer
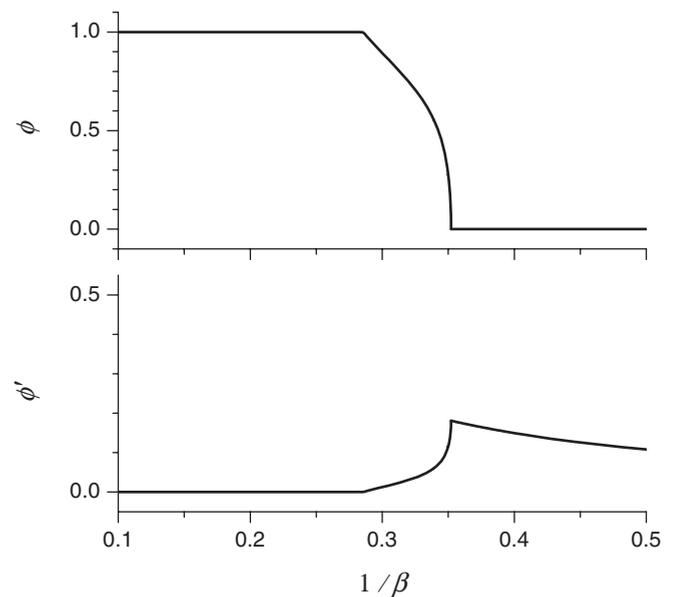
FIG. 6. Fraction of paired segments (ϕ, top panel) and fraction of bound segments (ϕ', bottom panel) as a function of temperature ($1/\beta$) in the single-chain limit for $\epsilon/\beta = 1$ and $\rho'_\circ = 0.02$.

collapse is energetically favored by the pairing mechanism and is entropically disfavored. On the other hand, the polymer collapse requires micromolecules release, and the latter process is energetically *disfavored* but entropically *favored*. We can denote this regime as *molten globule* state. Upon further decreasing temperature, at the critical end-point of the I-II transition ("folding" transition), ϕ reaches the saturation value $\phi = 1$, whereas ϕ' vanishes. We shall denote the latter regime as *folded globule* state, or even simply *paired* state, since all polymer segments (but a subextensive number) are paired (and therefore cannot be bound to micromolecules). In this state, we can imagine the polymer as a branched double chain, such that it resembles, with due cautions, a native RNA structure. In fact, this phase does not correspond to a unique configuration, but rather it can be defined as a statistical ensemble of different structures, as has been pointed out in Ref. [16].

The temperature behavior of the segment and micromolecule densities, ρ and ρ', is qualitatively very similar to that of ϕ and ϕ', respectively. In the coil state, independently of temperature, we have $\rho = 0$ and $\rho' = \rho'_\circ$, as all thermodynamic properties must coincide with those of the O phase. In the molten state, ρ increases and ρ' decreases, upon decreasing temperature. Finally, in the paired state, ρ and ρ' take respectively their maximum and minimum value. The fact that the density values of the paired state are rigorously independent of temperature (though they still depend on the micromolecule concentration of the O phase) is likely to be an artifact of the Bethe approximation. It is noticeable that the micromolecule concentration in the paired state is lower than in the O phase, namely, $\rho' < \rho'_\circ$ (see Fig. 7). We can interpret this fact, recalling that, in the grand-canonical description, an ideal experiment of a single (collapsed) polymer chain in the solution is represented by the coexistence of a very small amount (i.e., a tiny bubble) of dense phase coexisting with the bulk zero-segment-density phase. According to this rough but effective picture, we can argue that the micromolecule concentration "in the bubble" (i.e., in the vicinity of the collapsed polymer chain) must be lower than the overall concentration of the solution. Furthermore, we observe that the micromolecule density in the paired state is higher than that predicted by the pure polymer model (see Fig. 7). The latter effect can roughly be rationalized by an argument of the following kind: In the presence of micromolecules, the polymer chain can fold only by diminishing the binding

FIG. 8. The same as Fig. 6, for $\rho'_\circ = 0.03$. A dotted line denotes the abrupt jump in the order parameters, occurring along the O phase boundary. A thin solid line denote the order parameters values along the first-order branch within the I phase.

FIG. 7. Segment density (ρ, left panel) and micromolecule density (ρ', right panel) of the low-temperature paired state as a function of the micromolecule density of the O phase $\rho'_\circ$, for $\epsilon/\beta = 1$. A dash-dotted line in the right panel represents the condition $\rho' = \rho'_\circ$.

probability, which can be realized by staying in a more compact state.

Let us now analyze the changes induced to the previous picture by larger micromolecule concentrations. Figure 8 refers to the case $\rho'_\circ = 0.03$. We can see that the continuous collapse transition is replaced by a discontinuous transition, which corresponds to the critical end-point of the O-I transition line. As a result, the temperature range of the molten globule state is noticeably reduced. For completeness, we also report in Fig. 8 the values of the order parameters ϕ and ϕ' for the two type-I coexisting phases, along the first-order transition branch terminating in a critical point (see Fig. 3). Let us recall that this branch lies completely within a dense phase region, so that it is in fact irrelevant to the single-chain limit.

Figure 9 displays the situation for $\rho'_\circ = 0.04$. In this case, the intermediate-temperature molten globule state disappears, giving rise to a direct discontinuous transition from the coil state at high temperature to the folded globule state at low temperature. In the grand-canonical phase diagram (Fig. 4) the latter transition still corresponds to the critical end point of the O-I transition line, which yet now lies along the first-order boundary of the II phase. Let us note that, in this case, the curve associated to the first-order branch in the dense phase region exhibits a kink, corresponding to the critical end point of the I-II transition line.

Finally, Fig. 10 refers to $\rho'_\circ = 0.06$. According to the corresponding grand-canonical phase diagram (Fig. 5), in this case the O phase boundary is always a second-order transition, which, in the single-chain limit, means that the polymer chain is in the coil state ($\rho = 0$) at all temperatures, and folding cannot take place. The micromolecule density ρ' exhibits a monotonically increasing trend, upon decreasing temperature. The value 0.5 is an upper-bound for this quantity, because of the excluded volume constraint.
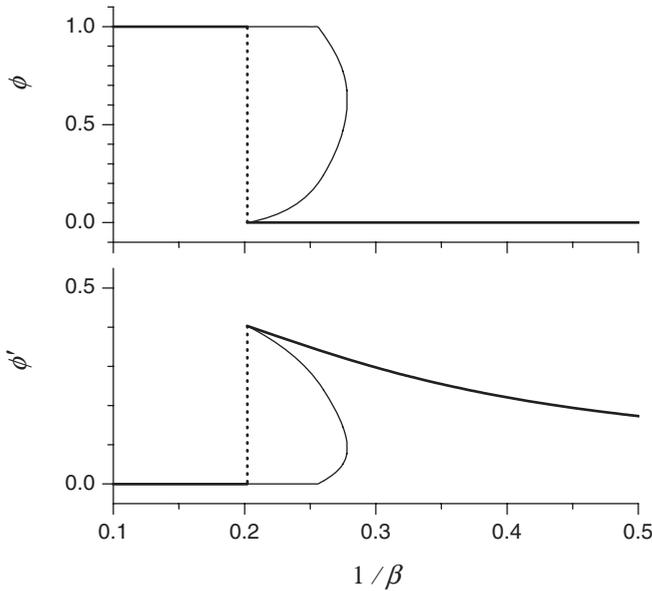
FIG. 9. The same as Fig. 8, for $\rho'_\circ = 0.04$.

FIG. 11. Phase diagram in the single-chain limit (temperature $1/\beta$ vs micromolecule density $\rho'_\circ$). Dotted and solid lines denote continuous and discontinuous transitions, respectively.

A unified picture of the different transitions occurring in the single-chain limit is displayed in Fig. 11. Moving on constant-concentration lines in this diagram, we recover the different (temperature-driven) transition sequences, described in the previous figures. The same kinds of transitions may be observed as concentration-driven transitions, at constant temperature. Nevertheless, it turns out that a transition scenario like that of Fig. 6 (in which *both* the coil-molten and the molten-paired transitions are continuous) could never be driven by the micromolecule concentration, because the temperature region of the Θ-like (i.e., continuous) coil-molten transition does not admit a paired state at low concentration. Furthermore, we can observe that the molten-paired transition is much less sensitive than the coil-molten transition, with respect to the micromolecule concentration. As a result, a

two-transition scenario may be observed only in a relatively narrow temperature range, whereas, in the largest temperature range, below a certain temperature, the transition scenario is to be expected in the form of a single discontinuous coil-paired transition, like that of Fig. 9.

Finally, let us report some results about the role of the lattice connectivity parameter k. As far as the single-chain phase diagram is concerned, Fig. 12 displays a comparison between the case $k = 11$, studied so far, and the case $k = 5$, meant to reproduce the simple cubic lattice ($k + 1 = 6$). We find it useful to display the phase diagrams after a suitable rescaling of both control parameters. Temperature is rescaled by the Θ temperature of the pure polymer model $1/\beta_\Theta$, where, according to Eq. (13) of Ref. [16],

$$\beta_\Theta = \ln \frac{k^2}{2k - 1}. \tag{20}$$

FIG. 10. The same as previous figures, for $\rho'_\circ = 0.06$.
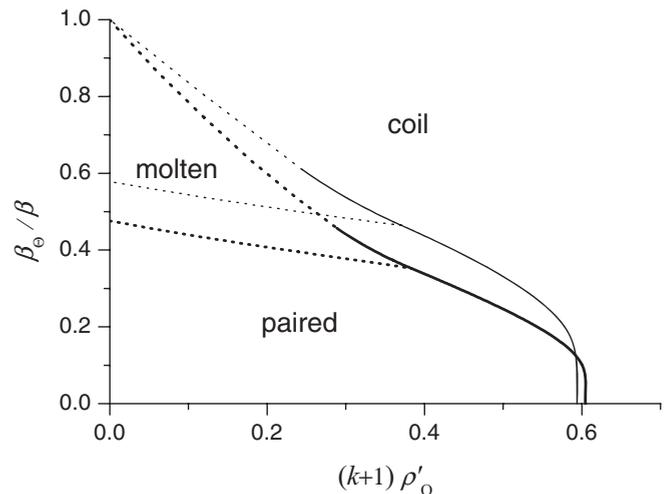
FIG. 12. Rescaled phase diagram in the single-chain limit (temperature β_Θ/β vs micromolecule density $(k+1)\rho'_\circ$) for $k = 5$ (thick lines) and $k = 11$ (thin lines). Dotted and solid lines denote continuous and discontinuous transitions, respectively.

Micromolecule concentration is rescaled by its upper limit value $1/(k+1)$, which descends from Eq. (A8). It turns out that, even though fine effects can be observed, the qualitative features of the phase diagram are basically unaffected by the connectivity parameter. In particular, it is noticeable that the (rescaled) concentration value that marks a change in the unfolding transition scenario (from continuous paired-molten to discontinuous paired-coil) is practically insensitive with respect to k. The same holds for the (concentration-driven) paired-coil transition, in the low-temperature region.

V. SUMMARY AND CONCLUSIONS

We have studied a lattice polymer model of the two-tolerant type, with an extra "micromolecular" chemical species, which is able to bind polymer segments and prevent segment-segment pairing. The equilibrium phase diagram of the model has been worked out in detail, in the framework of the Bethe-lattice approximation.

Let us only briefly mention that, in formulating the current model, we have been inspired by the binding mechanism of microRNA molecules (or similar RNA fragments) to messenger RNA, which has been recently recognized to play a key role for inhibiting messenger RNA functions in the framework of gene regulation networks [31–33]. The above process is still poorly understood, with respect to RNA folding, but it appears to be even more characterized by specific details (for instance, the limited number of specific binding sites along the RNA chain, the formation of microRNA-protein complexes, and so on). All these details determine the dynamical evolution of the system, leading eventually to translation repression and/or messenger RNA degradation. Unfortunately, all these phenomena cannot be captured by a purely thermodynamic and homogeneous model, like the one studied in this paper, so that we cannot draw a significant correspondence between the model and the real biochemical system.

We have nonetheless observed interesting qualitative effects, which do not appear to be predictable on the basis of simple heuristic arguments, and that may be of interest in the more general framework of polymer physics. In particular, we have observed that the special binding mechanism implemented by the model, associated to a local hindrance to the formation of double-stranded regions, turns out to induce a radical change in the nature of the collapse transition. In the grand-canonical picture, such a change corresponds to a switch from the Θ point scenario to the critical end-point scenario. Furthermore, the temperature of the collapse transition turns out be much more sensitive to the micromolecule concentration than the folding transition itself. As a result, we have that the possible scenarios of conformational transitions, driven by micromolecule concentration, always involve a discontinuous collapse, and, for most temperatures, a direct discontinuous coil-paired transition, without an intermediate molten state. In this sense, the micromolecule concentration provides a very effective mechanism to control the folding-unfolding transition.

Apart from questions related to the interpretation of the model in terms of a real physical system, we also have to recall that our results are in fact based on an approximate analytical technique. Even though the Bethe approximation

has often been demonstrated to yield reliable results for the phase diagram of polymer models (as mentioned in Sec. III), further investigations by different methods would be in order, to confirm the transition scenarios observed in the current work.

APPENDIX A: RECURSION EQUATIONS

In this Appendix, we describe the recursion equation (10) (corresponding indeed to a set of five simultaneous equations) in the explicit form, obtained by introducing the specific energy terms of the model. Moving the normalization factor q and the single-bond energies $\tilde{h}_{m,n}$ to the left-hand sides, and remembering Eqs. (2)–(6), we obtain

$$
\begin{aligned}
q w_{0,0} = {} & (w_{0,0} + w_{0,2})^k + \binom{k}{2} w_{0,1}^2 (w_{0,0} + w_{0,2})^{k-2} \\
& - \binom{k}{1} w_{0,2} w_{0,0}^{k-1} + \binom{k}{1} w_{1,0} \\
& \times \left[w_{0,0}^{k-1} + \binom{k-1}{2} w_{0,1}^2 w_{0,0}^{k-3} \right] \\
& + \binom{k}{1}\binom{k-1}{1} w_{1,1} w_{0,1} w_{0,0}^{k-2},
\end{aligned} \tag{A1}
$$

$$
\begin{aligned}
q e^{-\mu} w_{0,1} = {} & \binom{k}{1} w_{0,1}(w_{0,0} + w_{0,2})^{k-1} \\
& + \binom{k}{1}\binom{k-1}{1} w_{1,0} w_{0,1} w_{0,0}^{k-2} + \binom{k}{1} w_{1,1} w_{0,0}^{k-1},
\end{aligned} \tag{A2}
$$

$$
\begin{aligned}
q e^{-\beta-2\mu} w_{0,2} = {} & (w_{0,0} + w_{0,2})^k + \binom{k}{2} w_{0,1}^2 (w_{0,0} + w_{0,2})^{k-2} \\
& - w_{0,0}^k,
\end{aligned} \tag{A3}
$$

$$
q e^{-\mu'} w_{1,0} = w_{0,0}^k + \binom{k}{2} w_{0,1}^2 w_{0,0}^{k-2}, \tag{A4}
$$

$$
q e^{-\epsilon-\mu-\mu'} w_{1,1} = \binom{k}{1} w_{0,1} w_{0,0}^{k-1}. \tag{A5}
$$

In order to give a physical explanation of these equations, let us recall that their right-hand sides have to take into account all the allowed configurations of k lattice bonds sharing one site with another bond (root bond), whose configuration m,n is fixed for each given equation. In Eqs. (A1)–(A3) the fixed configurations are $m = 0$ (no micromolecule) and, respectively, $n = 0, 1, 2$ (bond occupied by 0, 1, 2 polymer segments). These equations are a "perturbation" of the corresponding Eqs. (B4)–(B6) of the pure polymer model [16] (with $\gamma = 0$). Equation (A1) has two extra terms (the last ones), which refer to configurations with one bond occupied by a micromolecule alone or bound to a polymer segment, respectively. The remaining $k - 1$ bonds must be, in the former case, either all empty or two occupied by unpaired polymer segments and $k - 3$ empty, and, in the latter case, one occupied by an unpaired polymer segment and $k - 2$ empty. Equation (A2) has also two extra terms, which refer to configurations with one bond occupied by a micromolecule alone or bound to a polymer segment, respectively. The remaining $k - 1$ bonds must be, in the former case, one occupied by an unpaired

polymer segment and $k-2$ empty, and, in the latter case, all empty. Equation (A3) has no extra terms, because in this case the root bond is occupied by two paired polymer segments, which implies that no micromolecule can occupy a bond in the same star cluster. In Eqs. (A4) and (A5) the fixed configurations are $m=1$ and $n=0,1$, meaning that the root bond is occupied by a micromolecule alone or bound to a polymer segment, respectively. Accordingly, the terms on the right-hand sides refer to configurations with, in the former case, either all empty bonds or two bonds occupied by unpaired polymer segments and $k-2$ empty, and, in the latter case, one occupied by an unpaired polymer segment and $k-1$ empty.

It is easy to verify that the simultaneous equations written above always admit a special solution such that

$$w_{0,1} = w_{0,2} = w_{1,1} = 0. \tag{A6}$$

According to Eq. (11), the latter condition obviously imply

$$p_{0,1} = p_{0,2} = p_{1,1} = 0, \tag{A7}$$

whence, by Eq. (13), a vanishing segment density $\rho = 0$. This solution, denoted as O phase, represents a mixture of only solvent and micromolecules, which is nonetheless not completely trivial, because of the excluded volume constraint. As discussed in the text, for each given lattice site, at most one of its incident bonds may be occupied by a micromolecule. From a graph-theoretical point of view, such a constraint coincides with the so-called *matching* condition [34]. As far as the micromolecule density is concerned, we expect therefore the following upper bound:

$$\rho' < \frac{1}{k+1}. \tag{A8}$$

With the condition (A6), the recursion equations simplify to

$$q w_{0,0} = w_{0,0}^k + k w_{1,0} w_{0,0}^{k-1}, \tag{A9}$$

$$q w_{1,0} = e^{\mu'} w_{0,0}^k, \tag{A10}$$

which, together with the normalization condition

$$w_{0,0} + w_{1,0} = 1, \tag{A11}$$

determine the unknowns, $w_{0,0}$, $w_{1,0}$, and q, as a function of the micromolecule chemical potential μ' and the lattice parameter k. According to Eqs. (11), (12), and (14), the micromolecule density ρ' can be written as

$$\rho' = p_{1,0} = \frac{e^{-\mu'} w_{1,0}^2}{w_{0,0}^2 + e^{-\mu'} w_{1,0}^2}. \tag{A12}$$

By simple algebra, it is possible to eliminate $w_{0,0}$, $w_{1,0}$, and q from Eqs. (A9)–(A12), yielding an explicit expression of ρ' as a function of μ' and k, namely,

$$\rho' = \frac{1}{k+1+\frac{2k}{\sqrt{1+4ke^{\mu'}}-1}}, \tag{A13}$$

from which the upper-bound (A8) follows directly. Of course, Eq. (A13) holds only in the O phase. Eq. (18) of the text is easily derived by inverting (A13).

APPENDIX B: CONTINUOUS TRANSITIONS

As mentioned in the text, the numerical technique we employ to solve the recursion equations is a simple fixed-point method. Given a tentative set of partial partition functions $w_{0,0}, \ldots, w_{1,1}$ (which we collectively denote as w in the following formulas), a new estimate of each function can be computed, according to Eq. (10), by an expression of the form

$$\hat{w}_{m,n}(w) = \frac{e^{-\tilde{h}_{m,n}} f_{m,n}(w)}{q(w)}, \tag{B1}$$

where the functions $f_{m,n}(w)$ coincide with the right-hand sides of Eqs. (A1)–(A5), and

$$q(w) = \sum_{m=0}^{1} \sum_{n=0}^{2-m} e^{-\tilde{h}_{m,n}} f_{m,n}(w) \tag{B2}$$

is the normalization factor.

A second-order phase transition is characterized by the fact that a minimum of the free energy becomes a saddle point, so that the corresponding thermodynamic equilibrium state is no longer stable. In the framework of (the recursive approach to) the Bethe lattice approximation, the thermodynamic (in)stability is reflected in the (in)stability of the fixed point of the recursion equations. Therefore, in order to determine second-order transitions with good precision, it is convenient to analyze the eigenvalues of the Jacobian matrix associated with the recursion equations and the conditions in which any eigenvalue equals unity.

According to Eqs. (B1) and (B2), the elements of such a Jacobian matrix can be written as

$$\frac{\partial \hat{w}_{m,n}(w)}{\partial w_{r,s}} = \frac{e^{-\tilde{h}_{m,n}}}{q(w)} \frac{\partial f_{m,n}(w)}{\partial w_{r,s}} - \frac{\hat{w}_{m,n}(w)}{q(w)} \frac{\partial q(w)}{\partial w_{r,s}}, \tag{B3}$$

where the pairs m,n and r,s represent, respectively, the row and column indices, and

$$\frac{\partial q(w)}{\partial w_{r,s}} = \sum_{m=0}^{1} \sum_{n=0}^{2-m} e^{-\tilde{h}_{m,n}} \frac{\partial f_{m,n}(w)}{\partial w_{r,s}}. \tag{B4}$$

The derivatives of $f_{m,n}(w)$ can be easily computed from Eqs. (A1)–(A5).

As far as the transition lines are concerned, we can locate them numerically as zeroes of $\det(J-I)$, where J is the Jacobian matrix and I the identity matrix. This can be safely done, because the peculiar form of the recursion equations allows us to follow a given solution (fixed point) even in the parameter region where it becomes unstable. More specifically, if one starts with tentative partition functions satisfying Eq. (A6) (which characterizes the O phase), the new estimates, computed by Eqs. (B1) and (B2), turn out to satisfy rigorously the same condition. The same holds for the II phase, characterized by the condition

$$w_{0,1} = w_{1,1} = 0. \tag{B5}$$

In principle, this kind of calculation could also be performed analytically, but only for the O-I line have we been able to obtain an acceptably simple form. Namely, Eq. (19) of the text is obtained plugging Eqs. (A6), (A9)–(A11), and (18) into the determinant equation $\det(J-I)=0$.

Let us finally note that, very close to a second order transition (i.e., when some eigenvalue of the Jacobian matrix is very close to unity), the convergence of the fixed-point procedure becomes quite slow. Therefore, if one primes the procedure without zero elements, the partial partition functions that are expected to vanish according to Eqs. (A6) or (B5) (depending on the phase under investigation) remain sensibly different from zero for a large number of iterations. We have exploited this fact to compute numerically the single-chain limits of ϕ and ϕ', which, from the analytical point of view, give rise to 0/0 undetermined forms.

[1] P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 1979).

[2] J. des Cloiseaux and G. Jannink, *Polymers in Solution: Their Modeling and Structure* (Oxford University Press, Oxford, 1990).

[3] C. Vanderzande, *Lattice Models of Polymers* (Cambridge University Press, Cambridge, 1998).

[4] D. P. Foster and F. Seno, J. Phys. A: Math. Gen. **34**, 9939 (2001).

[5] C. Buzano and M. Pretti, J. Chem. Phys. **117**, 10360 (2002).

[6] C. Buzano and M. Pretti, Mol. Cryst. Liq. Cryst. **398**, 23 (2003).

[7] J. Krawczyk, A. L. Owczarek, and T. Prellberg, J. Stat. Mech. (2007) P09016.

[8] D. P. Foster, J. Phys. A **40**, 1963 (2007).

[9] J. Krawczyk, T. Prellberg, A. L. Owczarek, and A. Rechnitzer, Phys. Rev. Lett. **96**, 240603 (2006).

[10] P. Serra and J. F. Stilck, Phys. Rev. E **75**, 011130 (2007).

[11] T. J. Oliveira, J. F. Stilck, and P. Serra, Phys. Rev. E **80**, 041804 (2009).

[12] T. J. Oliveira and J. F. Stilck, J. Stat. Mech. (2011) P01026.

[13] E. Orlandini, F. Seno, A. L. Stella, and M. C. Tesi, Phys. Rev. Lett. **68**, 488 (1992).

[14] M. Baiesi, E. Orlandini, and A. L. Stella, Phys. Rev. Lett. **91**, 198102 (2003).

[15] P. Leoni and C. Vanderzande, Phys. Rev. E **68**, 051904 (2003).

[16] M. Pretti, Phys. Rev. E **74**, 051803 (2006).

[17] R. A. Zara and M. Pretti, J. Chem. Phys. **127**, 184902 (2007).

[18] D. P. Foster and C. Pinettes, Phys. Rev. E **79**, 051108 (2009).

[19] R. F. Gesteland, T. R. Cech, and J. F. Atkins, editors, *The RNA World* (Cold Spring Harbor Laboratory Press, Woodbury, 2005).

[20] D. Jost and R. Everaers, J. Chem. Phys. **132**, 095101 (2010).

[21] D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner, J. Mol. Biol. **288**, 911 (1999).

[22] M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* (World Scientific, Singapore, 1994).

[23] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).

[24] C. Domb, Adv. Phys. **9**, 245 (1960).

[25] M. Pretti, J. Stat. Phys. **111**, 993 (2003).

[26] H. A. Bethe, Proc. R. Soc. London A **150**, 552 (1935).

[27] R. Kikuchi, Phys. Rev. **81**, 988 (1951).

[28] E. A. Guggenheim, Proc. R. Soc. London A **148**, 304 (1935).

[29] S. Lise, A. Maritan, and A. Pelizzola, Phys. Rev. E **58**, R5241 (1998).

[30] T. R. Einert, P. Näger, H. Orland, and R. R. Netz, Phys. Rev. Lett. **101**, 048103 (2008).

[31] W. Filipowicz, S. N. Bhattacharyya, and N. Sonenberg, Nat. Rev. Genet. **9**, 102 (2008).

[32] M. Ghildiyal and P. D. Zamore, Nat. Rev. Genet. **10**, 94 (2009).

[33] D. Jost, A. Nowojewski, and E. Levine, BMB Rep. **44**, 11 (2011).

[34] L. Zdeborová and M. Mézard, J. Stat. Mech. (2006) P05003.