

Statistical physics of RNA folding

M. Müller

Laboratoire de Physique Théorique et Modèles Statistiques, Université Paris–Sud, Bâtiment 100, F-91405 Orsay, France

(Received 14 June 2002; revised manuscript received 12 December 2002; published 27 February 2003)

We discuss the physics of RNA as described by its secondary structure. We examine the static properties of a homogeneous RNA model that includes pairing and base stacking energies as well as entropic costs for internal loops. For large enough loop costs the model exhibits a thermal denaturation transition which we analyze in terms of the radius of gyration. We point out an inconsistency in the standard approach to RNA secondary structure prediction for large molecules. Under an external force a second-order phase transition between a globular and an extended phase takes place. A Harris-type criterion shows that sequence disorder does not affect the correlation length exponent while the other critical exponents are modified in the glass phase. However, at high temperatures, on a coarse-grained level, disordered RNA is well described by a homogeneous model. The characteristics of force-extension curves are discussed as a function of the energy parameters. We show that the force transition is always second order. A reentrance phenomenon relevant for real disordered RNA is predicted.

DOI: 10.1103/PhysRevE.67.021914

PACS number(s): 87.14.Gg, 87.15.-v, 64.60.-i

I. INTRODUCTION

In this paper we discuss the equilibrium statistical mechanics of RNA or single-stranded DNA as described by their secondary structure (base pairing pattern). We mainly concentrate on homogeneous polymers with uniform interactions between monomers which will be shown to capture well the physics of random disordered sequences at sufficiently high temperatures on a coarse-grained level.

We generalize the results of de Gennes' pioneering paper [1] on (homogeneous) periodic dAT polymers (sequences ATAT . . .) by including entropic penalties for internal loops in order to account, to some extent, for self-avoidance effects. In the case where the latter are large, we predict a thermal denaturation transition that manifests itself in the scaling behavior of the radius of gyration. The scaling found in the low-temperature phase is smaller than $N^{1/3}$, where N is the number of monomers of the polymer. This signals an inconsistency for large N , since the monomer density in three-dimensional space becomes increasingly large with N . Excluded volume effects can, therefore, not be neglected in the secondary structure prediction of large molecules.

Recently diverse micromanipulation techniques have been developed that allow to monitor the response of single biomolecules, RNA or ssDNA, in particular [2–4], to an external force. These experiments have raised considerable interest in the theoretical study of force-extension characteristics of biomolecules. Within our model, force-extension curves can easily be obtained upon coupling an external force to the extremities of the polymer. The molecule undergoes a thermodynamic phase transition of second order that separates the globular collapsed state from an extensive phase containing a large number of small globules [5]. We characterize the associated critical behavior and study in how far it is modified by the introduction of sequence randomness. Using a Harris-type criterion we argue that the correlation length remains unaffected by the disorder irrespective of temperature while other critical exponents maybe modified. Numerical results indicate that at higher temperatures disordered models belong to the same universality class as the homogeneous

model, while at low temperatures the collapsed phase becomes glassy and the critical behavior changes.

In a recent article [6] the authors claim that the introduction of base stacking energies instead of base pairing energies may change the order of the force-induced phase transition. However, they were misled by the appearance of a sharp first-order-like crossover that occurs at a higher force than the continuous phase transition described above. While the latter is almost entirely of entropic nature, reflecting the large space of energetically equivalent secondary structures, the crossover is governed by the competition between the pairing energy and the energy gained from opening and stretching it. The sharpness of the crossover results from the cooperativity due to the base stacking energy that favors long helices of stacked pairs.

The secondary structure of RNA

RNA is a linear polymer made up of four types of nucleotides, A, C, G and U. In single-stranded DNA, U is replaced by T. In solution with a sufficiently high ionic concentration to screen the charge of the phosphate backbone, the single RNA strand has a tendency to fold back onto itself to form local double helices of Watson-Crick base pairs (A-U and G-C) between complementary substrands of the base sequence. The entropy loss due to a bound helix is compensated for by the pairing energy due to the 2 (in A-U) or 3 (in C-G) hydrogen bonds of the base pairs and, more importantly, the stacking energy which is gained by the expulsion of water molecules between the hydrophobic parts of neighboring stacked base pairs.

The set of all base pairings in the RNA molecule determines its secondary structure. The typical scale of pairing and stacking energies is considerably larger than the energy scale associated with the tertiary structure, i.e., the spatial arrangement of the RNA molecule (see Ref. [7], and references therein). This separation of energy scales is at the basis of the usual paradigm to split the RNA-folding problem into the analysis of the base pairing pattern and a subsequent determination of the tertiary structure. The set of all pairing

patterns considered as secondary structures is further restricted by discarding all pairings between different loops, so-called pseudoknots. Such structures lead to knotted configurations if the helices between the loops are sufficiently long to intertwine. While knots are prevented in nature by the linear transcription process from DNA to RNA, short helices between loops can occur, in principle, but they are found to constitute only a minor fraction of all base pairings. They are thus considered as elements of the tertiary interactions that can be neglected when determining the secondary structure. If we number the bases in the sequence as $i = 1, \dots, N$ according to their position in the strand, the above constraint can be formalized by forbidding the coexistence of two (ordered) base pairs (i_1, j_1) , (i_2, j_2) in the secondary structure with either $i_1 < i_2 < j_1 < j_2$ or $i_2 < i_1 < j_2 < j_1$.

As we mentioned above, the separation of energy scales breaks down for large molecules, and the folding problem is complicated by the highly nonlocal condition that the secondary structure must have a realization in 3D. However, for intermediate degrees of polymerization N the classical approach is expected to work well as is witnessed by the success of secondary structure prediction tools [8,9] that are based on the above assumptions.

In the following we will start from the usual paradigm and concentrate on the secondary structures excluding pseudoknots and other tertiary interactions. In Sec. II we discuss the statistical properties of homogeneous RNA working directly in the abstract phase space of secondary structures. This allows us to take into account systematically some excluded volume effects that reduce significantly the available configuration space of interior loops, and to gain some insight into the thermal denaturation of RNA. The second part of the paper is devoted to the response of RNA to an external force. In Sec. III the critical behavior at the force-induced opening transition is characterized and the effect of sequence disorder on the phase transition is discussed. Section IV deals with force-extension curves in the thermodynamic limit and its properties as a function of the energy parameters and temperature. We show that the phase transition is always of second order, but can be masked by a subsequent first-order-like crossover when the cooperativity of the pairing behavior due to the stacking energy is high.

II. FOLDING OF HOMOGENEOUS RNA

In this section we neglect all effects due to sequence specificity. Instead we consider an RNA model where any two bases can form a bond, their pairing affinity being independent of the bases. This exactly solvable model describes the physics of “homogeneous” RNA/ssDNA strands GCGCGC... or ATATAT... [1], renormalized on the level of dimers. We will provide some evidence that random base sequences are also well described by a homogeneous model, at least at sufficiently high temperatures, if one switches to a more coarse-grained description where the monomers of the model correspond to short subunits of the single strands rather than to real bases.

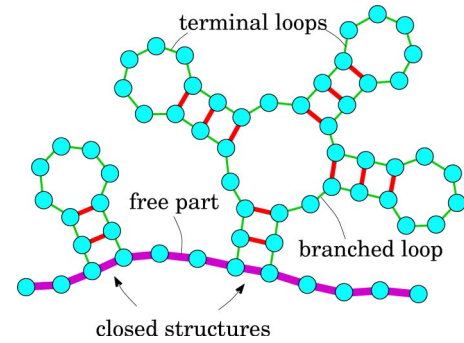


FIG. 1. Elements of the secondary structure. The unpaired bases between the two closed structures constitute the free part (thick line) of the chain. The structure on the right contains a branched loop with $l=3$ unpaired bases and $m'=m+1=4$ outgoing stems. The contour length L of the loop is taken to be the number of backbone elements in the loop, i.e., $L=l+m'$. Complementary substrands that directly fold back onto themselves form hairpins which end in terminal loops containing at least t bases.

A. The model

Following the empiric rules established by Tinoco's group [10,11] we consider three different terms in the free energy of a given secondary structure (cf. Fig. 1 for illustration of the notions): Each base pair contributes the pairing free energy f_{pair} that we normalize with respect to the completely denatured chain where all bases are unpaired. This takes into account the mean pairing free energy (bond enthalpy and entropy cost for localization) as well as the stacking energy with the neighboring pair. In this way we count an excess stacking energy at one end of the helices which we have to compensate for by a free energy cost $-f_{\text{stack}}$.

In the following these free energies will appear in the form of the temperature dependent parameters

$$\eta = \exp(\beta f_{\text{stack}}) \quad (1)$$

and

$$s = \exp(-\beta f_{\text{pair}}), \quad (2)$$

where β is the inverse of the temperature T . Under biological conditions (ionic strength and pH as in a living cell) $\eta \ll 1$, reflecting the importance of the stacking energy as compared to the binding energy of the hydrogen bonds. As we will see later, this is responsible for the high degree of cooperativity in the denaturation transition.

The last contribution is an entropic cost for each closed internal loop which accounts for the reduced phase space available to the loop with respect to an unconstrained string containing the same number of bases. We assume this part of the cost function to depend only on the length L of the loop and the number m' of stems connected to it. The reduction of phase space gives rise to a free energy contribution of the form $\beta^{-1} \ln[\phi(L; m')]$. Specific expressions for ϕ will be discussed later.

When describing homogeneous RNA, we should add two further constraints on the secondary structures to be considered: First, terminal loops (the loops at the end of a hairpin)

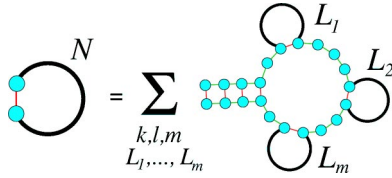


FIG. 2. Schematic representation of the recursion relation for the partition function Z_N^c of closed structures used to obtain Eq. (3). The sum on the rhs is over the length k of the terminal helix, the number of unpaired bases l in the first loop and the number m and sizes L_i ($i = 1, \dots, m$) of closed structures attached to the loop. The shown configuration corresponds to $k=4$, $l=7$, $m=3$. The loop contributes an entropic cost $\phi(L; m')$, where $m' = m+1$ is the number of outgoing stems and L is the contour length of the loop, $L = l + m'$.

have to contain a minimal number t of unpaired bases ($t = 3$ from experiments), since the bending rigidity of single-stranded RNA is finite and the typical distance of a hydrogen bond l_b is about 3–4 times larger than the base distance l in the backbone. Secondly, stacks of less than three base pairs are unstable, and we thus require a helix to have a minimal length of $n=3$ bases.

These two conditions do not make sense for the description of real disordered base sequences on a coarse-grained level. The natural values to be taken in this case are $t=0$ and $n=1$.

B. The partition function

We denote by Z_N^c the partition function of a *closed* RNA molecule with N bases whose ends are required to form a helix. We can easily obtain a recursion relation for Z_N^c (cf. Fig. 2): The closed secondary structure terminates in a helix containing $k \geq n$ base pairs. It is followed by a first loop which contains $l \geq 0$ unpaired bases and $m \geq 0$ closed substructures, containing $L_i \geq 2n + t$ bases, respectively. The arrangement of the free bases and substructures within the loop underlies no constraints and gives rise to a combinatoric factor. The loop contributes an entropic cost $\phi(L; m' = m + 1)$, where we take the length L to be given by the number of backbone elements it contains, i.e., $L = l + m + 1$. Finally, the sum over all configurations can be decomposed as

$$Z_N^c = \eta \sum_{k \geq n} s^k \sum_{\{l, m\}} \sum_{L_1 \geq 2n+t} \dots \sum_{L_m \geq 2n+t} \times \delta \left(2k + l + \sum_{i=1}^m L_i - N \right) \times \binom{m+l}{m} \frac{1}{\phi(l+m+1; m+1)} \prod_{i=1}^m Z_{L_i}^c. \quad (3)$$

In the sum over l (number of unpaired bases) and m (number of outgoing stems) the following pairs have to be excluded: ($m=0, l < t$) to prevent terminal loops smaller than t , and ($m=1, l=0$) to avoid double counting of structures.

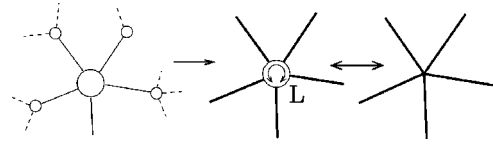


FIG. 3. The effective entropy cost for internal loops of length L can be obtained by idealizing the environment of the loop as m outgoing rods (here $m=5$) and comparing the scaling expression for its configurational entropy with that of a starlike polymer with m rays.

The structure of Eq. (3) suggests to study the generating function $\Xi_c(\zeta) \equiv \sum_{N=2n+t}^{\infty} Z_N^c \zeta^N$ of the partition function. Taking the discrete Laplace transform we obtain

$$\Xi_c(\zeta) = \eta \frac{(s\zeta^2)^n}{1 - s\zeta^2} \sum_{\{m, l\}} \frac{\Xi_c(\zeta)^m \zeta^l}{\phi(l+m+1; m+1)} \binom{m+l}{m}. \quad (4)$$

To proceed we have to make an assumption about the loop cost function $\phi(L; m)$. Neglecting loop costs altogether corresponds to putting $\phi(L; m) = 1$, which yields

$$\Xi_c(\zeta) = \eta \frac{(s\zeta^2)^n}{1 - s\zeta^2} \left(\frac{1}{1 - \Xi_c(\zeta) - \zeta} - \Xi_c(\zeta) - \frac{1 - \zeta^t}{1 - \zeta} \right). \quad (5)$$

We will use this simple model in Sec. IV to discuss the general shape of the force-extension characteristics. We expect it to describe well the low-temperature regime where large internal loops are negligible. In order to describe denaturation we should, however, use a more realistic loop cost function.

C. Denaturation

If RNA were an ideal chain without self-interaction, the entropic cost of a closed loop would just derive from the probability of a three-dimensional random walk to return to the origin, and thus $\phi(L; m) \propto L^{3/2}$ for large values of L . This corresponds to the case discussed in Refs. [1,5] where the authors start from real space recursion relations treating the single strands as ideal chains. If one considers the loops as self-avoiding walks, forgetting about the stems that are connected to them, one is lead to use $\phi(L; m) \propto L^{3\nu_{SAW}}$, with the wandering exponent $\nu_{SAW} = 0.588$ (in 3D) characteristic of a self-avoiding walk. (This is the form used for large interior loops in the Zuker algorithm [12].) Clearly, this is too simple since the branches attached to the loop have a non-negligible effect on the conformational degrees of freedom of the loop and one should consider a more sophisticated form of the loop cost.

A generalization of these entropy cost functions can be obtained from the results of Duplantier and co-workers [13,14] for the configurational entropy of a network with given topology. In order to find the scaling of the effective entropy cost of an internal loop as a function of its size we consider the secondary structure as a treelike network of helices, linked by internal loops. Let us single out one internal loop with m branches that we idealize as outgoing rods (see

Fig. 3). In the limit of loop sizes much smaller than the extension of the rods the scaling of the entropy cost for the internal loop follows from comparison of the expressions for the configurational entropy of a starlike network with m rays $\Gamma_{\text{star}}(m)$ with that of a small loop with m attached branches $\Gamma_{\text{loop}}(L;m)$. The latter scale as $\Gamma_{\text{star}}(m) \sim N^{\gamma_{\text{star}}(m)-1}$ and $\Gamma_{\text{loop}}(L;m) \sim N^{\gamma_{\text{loop}}(m)-1} g(L/N)$, where N is the typical length scale of the attached rods (the remainder of the network), L is the length of the loop, and $g(x)$ is a scaling function. In the limit $L/N \rightarrow 0$ the scaling of two expressions should coincide which requires $g(x) \sim x^{\gamma_{\text{loop}}(m) - \gamma_{\text{star}}(m)}$. The effective loop cost function then follows as $\phi(L;m) \sim \Gamma_{\text{loop}}(L;m)/\Gamma_{\text{star}}(m) \sim a(m)(L)^{\nu(m)}$, where $\nu(m) = \gamma_{\text{star}}(m) - \gamma_{\text{loop}}(m)$. The renormalization group results from Refs. [13,14] yield $\nu(m) = 3\nu_{\text{SAW}} - m\sigma_3 + \sigma_m$, where σ_k is the exponent related to the renormalization of a vertex with k legs. (See also Refs. [15–17] for a completely analogous reasoning in the closely related problems of DNA denaturation and studies of “slip-linked” polymers.) While good estimates are available via ϵ expansion for small values of m [13,14] the thermodynamics of denaturation is essentially determined by the behavior of the cost for loops with many attached stems, i.e., by $\nu(m)$ for large m , about which very little is known. We can proceed, however, without knowing an exact expression for $\nu(m)$. Instead we will illustrate the general condition (8) below with a discussion of the *ad hoc* forms $\phi(L;m) = a(m)(L)^{\nu(m)}$ with $\nu(m) = \nu^* = \text{const}$, and $\nu(m) = \nu_0 + m\nu_1$, the first one being a reasonable approximation for the case that $\nu(m)$ saturates at ν^* for large values of ν , the second one assuming that each branch contributes a further entropic constraint on the loop conformations, as suggested by the term $m\sigma_3$ above. The prefactor $a(m)$ is assumed to be a moderate function of m that does not grow exponentially.

The asymptotic behavior of the partition function Z_N^c can be derived from the generating function $\Xi_c(\zeta)$ without performing the full inverse Laplace transform. It is given by $Z_N^c \sim \zeta_*^{-N}/N^\alpha$, where ζ_* is the smallest value of ζ at which $\Xi_c(\zeta)$ becomes nonanalytic. The leading finite size corrections in the form of the preexponential factor $1/N^\alpha$ are determined by the nature of that nonanalyticity.

There are only two possible singularities for $\Xi_c(\zeta)$: Taking successive derivatives of Eq. (4) one can check that all derivatives $d^k \Xi_c(\zeta)/d\zeta^k$ exist, unless the partial derivatives of both sides of Eq. (4) with respect to Ξ_c are equal, i.e.,

$$1 = \eta \frac{(s\zeta_*^2)^n}{1 - s\zeta_*^2} \sum_{m,l} \frac{m \Xi_c(\zeta_*)^{(m-1)} \zeta_*^l}{\phi(l+m+1; m+1)} \binom{m+l}{m}. \quad (6)$$

In turn this condition is sufficient to ensure a nonanalyticity of $\Xi_c(\zeta)$. Writing $\Xi_c(\zeta) = \Xi_c(\zeta_*) - \delta$ and expanding Eq. (4) for small $(\zeta_* - \zeta)$ one finds $\delta^2 \sim (\zeta_* - \zeta)$. Hence the singularity of $\Xi_c(\zeta)$ is approached as

$$\Xi_c(\zeta) = \Xi_c(\zeta_*) - \text{const}(\zeta_* - \zeta)^{1/2} + O(\zeta_* - \zeta), \quad (7)$$

which leads to finite size corrections of the form $Z_N^c \sim \zeta_*^{-N}/N^{3/2}$. This result is central to the discussion of the

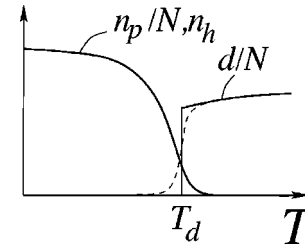


FIG. 4. The denaturation transition is best described by the typical distance d between bases within the secondary structure whose scaling changes from $N^{1/2}$ in the globular state at low temperatures to N in the necklacelike state above T_d . (The dashed curve indicates d/N for a molecule with finite N . The transition becomes sharp only in the thermodynamic limit.) In contrast, the fraction of paired bases n_p or the relative number of helices n_h/N only exhibit a crossover at the denaturation temperature, but never drop to zero.

critical behavior at the force-induced denaturation transition. Note that the preexponential factor is essentially independent of the specific choice of the model, in particular, it does not depend on the shape of the loop cost function ϕ .

The second possible singularity corresponds to the double sum on the right-hand side of Eq. (4) being evaluated at its radius of convergence. To analyze this situation in more detail we suppose that the loop penalty assumes the form $\phi(L;m) = a(m)L^{\nu(m)}$. According to Hadamard's formula, the radius of convergence is determined by

$$\lim_{L \rightarrow \infty} \left[\zeta_*^L \sum_{m=0}^L \frac{(\Xi_c(\zeta_*)/\zeta_*)^m}{a(m+1)(L+1)^{\nu(m+1)}} \binom{L}{m} \right]^{1/L} = 1. \quad (8)$$

In the case where $\nu(m)$ grows at most sublinearly in m , i.e., $\nu(m)/m \rightarrow 0$, as $m \rightarrow \infty$, the sum can be estimated by its saddle point, and one finds the condition $\Xi_c(\zeta_*) + \zeta_* = 1$. In the same way, $\nu(m) = \nu_0 + \nu_1 m$ (with $\nu_1 \leq 1$) leads to the condition $\zeta_* = 1$, but, since the sum diverges when $\zeta_* \rightarrow 1$, the singularity is ruled out in this case.

The singularity given by Eq. (6) is always the smallest one at low temperatures. The system undergoes a thermodynamic phase transition when the singularity determined by Eq. (8) crosses the first one as a function of temperature. This can only occur if the first derivative with respect to Ξ_c of the double sum in Eq. (4) stays finite on approaching the radius of convergence from below, that is, for $\Xi_c(\zeta_*) + \zeta_* \rightarrow 1$. This requires $\lim_{m \rightarrow \infty} \nu(m) \geq 2$. As we will show below the corresponding phase transition is associated to thermal denaturation. In all other cases our model does not exhibit a phase transition but only a crossover whose sharpness depends both on the loop cost function and the stacking as described by the parameter η .

D. Radius of gyration

Let us now characterize the thermodynamic properties of RNA. Most observables can be obtained as appropriate derivatives of the free energy per base which is simply related to ζ_* via $f = \beta^{-1} \ln(\zeta_*)$. For example, the fraction of paired bases is given by $n_p = \partial \ln(\zeta_*)/\partial \ln(s)$ and the average num-

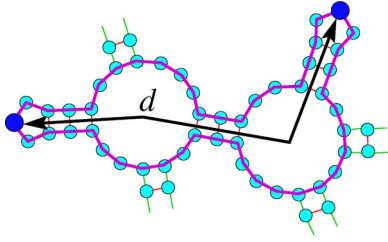


FIG. 5. Definition of the distance d between bases in terminal loops: The unique shortest path through helices and loops from one terminal loop to another allows us to define d as the sum of the lengths of the helices and half of the contour lengths of the encountered loops, including the two terminal loops. For the bases in the figure one has $d = 3/2 + 3 + 15/2 + 3 + 15/2 + 3 + 3/2$.

ber of helices is $n_h = N \partial \ln(\zeta_*) / \partial \ln(\eta)$. Evaluating these derivatives in the two possible phases one finds that n_p and n_h/N are both finite at all temperatures and thus do not provide a good order parameter for the phase transition (see Fig. 4). The (small) extensive number of pairings even in the high-temperature phase is due to accidental pairings of bases that are close to each other within the linear RNA strand. This result is independent of the details of the model such as the minimal number of bases in a hairpin loop, entropy costs, and energy parameters.

A better choice of observable is the average distance between two bases that belong to two different terminal loops of the secondary structure. This quantity is a measure for the diameter of the molecule and distinguishes the compact globular phase from the denatured loose phase. We define the distance between two bases as the length of the shortest path linking them through the secondary structure, see Fig. 5. This path is a succession of loops and helices, its length being given as the sum of the lengths of the helices and half the contour lengths of the loops.

Alternatively, we can consider again closed secondary structures constrained to terminate in a helix. It is easy to see that the same structural information as encoded by the distance between terminal bases is captured by the average distance of terminal bases from the closing base pair $(1, N)$. The latter quantity is more convenient to compute, however.

Let us denote by $n(d; \mathcal{S})$ the number of bases in terminal loops at distance d from the pair $(1, N)$. We define the Boltzmann weighted sum over secondary structures on a closed strand with N bases

$$C_N(d) = \sum_{\mathcal{S}} Z_N^c(\mathcal{S}) n(d; \mathcal{S}), \quad (9)$$

for which we can write down a recursion relation in the same spirit as in Eq. (3): The sum over all secondary structures can be decomposed into a sum over the length $k \geq n$ of the terminal helix starting at $(1, N)$, and the first loop. The latter can either be a terminal loop containing $N - 2k$ unpaired bases, or there can be further closed structures connected to it. In the second case, we single out the closed structure X whose terminal loops we want to consider, and denote by m_1, l_1 and m_2, l_2 the number of closed structures and un-

paired bases to either side of X in the loop. The $m_1 + m_2$ closed structures just contribute the product of their partition functions, together with appropriate combinatorial factors for their arrangement within the loop, whereas the structure X contributes the Boltzmann sum $C_{L_X}(d')$ for a distance d' that is reduced with respect to d by the length k of the terminal helix and half of the loop contour $m_1 + l_1 + m_2 + l_2 + 2$. This results in the recursion

$$C_N(d) = \sum_{k \geq n} \eta s^k \left[\frac{(N - 2k) \delta(d - N/2 - 1/2)}{\phi(N - 2k + 1)} + \sum_{m_1, l_1 \geq 0} \sum_{m_2, l_2 \geq 0} \sum_{L_X \geq 2n + t} \sum_{\substack{L_i \geq 2n + t; \\ i = 1, \dots, m_1 + m_2}} \binom{m_1 + l_1}{m_1} \times \binom{m_2 + l_2}{m_2} \delta \left(2k + l_1 + l_2 + \sum_{i=1}^{m_1 + m_2} L_i + L_X - N \right) \times \frac{C_{L_X}[d - k - (m_1 + l_1 + m_2 + l_2 + 2)/2]^{m_1 + m_2}}{\phi(l_1 + m_1 + l_2 + m_2 + 2)} \prod_{i=1}^{m_1 + m_2} Z_{L_i}^c \right]. \quad (10)$$

Here we used the form $\phi(L; m) \equiv \phi(L)$ for simplicity. Passing to the Laplace transform with respect to both variables N and d ,

$$C(\zeta, p) = \sum_{N \geq 2n + t} \sum_{d \geq n + t/2} C_N(d) \zeta^N e^{-pd}, \quad (11)$$

the equation is easily solved,

$$C(\zeta, p) = \frac{\eta (s \zeta^2 e^{-p})^n e^{-p/2} g_\phi^{(t)}(\zeta e^{-p/2})}{1 - s \zeta^2 e^{-p} - \eta (s \zeta^2 e^{-p})^n g_\phi'([\zeta + \Xi_c(\zeta)] e^{-p/2})}. \quad (12)$$

We have introduced the functions $g_\phi^{(t)}(x) = \sum_{N \geq t} N x^N / \phi(N + 1)$ and $g_\phi(x) = \sum_{N \geq 1} x^N / \phi(N + 1)$.

Note that $C(\zeta, p = 0)$ has two possible singularities, the vanishing of the denominator and the singularity in g_ϕ' that occurs when $\zeta + \Xi_c(\zeta) = 1$. The first singularity is associated with the globular phase, the vanishing of the denominator being equivalent to condition (6) in the case $\phi(L; m) = \phi(L)$. The singularity related to $g_\phi'(x \rightarrow 1)$ governs the denatured phase. The denaturation transition occurs when the two singularities cross which can only happen if $g_\phi'(1)$ is finite.

Let us now calculate the mean distance from bases in terminal loops to the free part. This is given by the logarithmic derivative of $C_N(p)$ with respect to p ,

$$\langle d \rangle = \partial_p C_N(p) |_{p=0} / C_N(p=0), \quad (13)$$

which can be evaluated by inverse Laplace transform of $\partial_p C(\zeta, p) |_{p=0}$ and $C(\zeta, p=0)$ with respect to ζ . Note that $\partial_p C_N(p) |_{p=0}$ and $C_N(p=0)$ have the same leading expo-

nential asymptotics since their smallest singularities are the same, but their finite size corrections ζ_*^{-N}/N^α differ and will determine the scaling of $\langle d \rangle$ as a function of N .

Let us now analyze the case of a loop cost function with asymptotics $\phi(L) \propto L^\nu$ and $\nu > 2$ which implies a phase transition. In the low-temperature phase, the denominator in Eq. (12) vanishes like $(\zeta_* - \zeta)^{1/2}$ as ζ approaches ζ_* from below [cf. Eq. (7)] which gives rise to $\alpha = 1/2$ for the back transform of $C(\zeta, p=0)$. On the other hand, the derivative $\partial C(\zeta, p)/\partial p|_{p=0}$ diverges like $(\zeta_* - \zeta)^{-1}$ which yields $\alpha = 0$, whence $\langle d \rangle \sim N^{1/2}$.

The situation changes in the high-temperature phase where the denominator remains finite, but the generating function develops a (leading) singular part $C_{\text{sing}}(\zeta, p=0) \sim (\zeta_* - \zeta)^{\nu-2}$ giving rise to $\alpha = \nu - 1$. Correspondingly, the singular part of the derivative behaves as $(\zeta_* - \zeta)^{\nu-3}$, which implies $\alpha = \nu - 2$, and therefore $\langle d \rangle \sim N$. The high-temperature phase is characterized by typical distances of order N from terminal bases down to the free part. This is what one expects for an essentially free, noncollapsed chain [but constrained to be paired at the ends $(1, N)$]. The secondary structure is rather trivial in this case, consisting essentially of one big loop with small structures attached to it. The low-temperature scaling $\langle d \rangle \sim N^{1/2}$, however, indicates the collapse to a globular state with a rich branched secondary structure.

We mention that the latter scaling can easily be derived in the absence of loop costs from the “mountain height” representation of secondary structures [18] where the distance of terminal bases to the free part scales like the height of the mountain representation. This in turn is the typical excursion of a random walk of N steps in one dimension h , constrained by $h > 0$ and $h(1) = h(N) = 0$ which is known to scale as $N^{1/2}$. There is, however, no simple equivalent of the above phase transition in the mountain height or random walk picture, since the loop cost translates into an awkward nonlocal energy term.

The distance d is a structural property of the treelike skeleton of the secondary structure. In order to relate it to the real diameter of the molecule we assume that the helices and parts of internal loops connecting a terminal loop to the free part essentially realize a (constrained) random walk in space. If we assumed the random walk to be ideal, the radius of gyration of RNA molecules would follow from the above findings as $R_g \sim d^{\nu_{\text{RW}}} \sim N^{1/4}$ with the wandering exponent for ideal random walks $\nu_{\text{RW}} = 1/2$. This has already been obtained in Ref. [1]. However, the random walks should at least be considered as self-avoiding, having a larger wandering exponent ν_{SAW} , and correspondingly $R_g \sim N^{\nu_{\text{SAW}/2}}$. In addition, there are also constraints from the presence of the remainder of the molecule which have the tendency to increase the value of the wandering exponent. This effect is difficult to estimate, but we may obtain some idea about its importance by considering the two-dimensional case: The wandering exponent for self-avoiding walks is known exactly as $\nu_{\text{SAW}} = 3/4$ [19]. On the other hand, we need to know how the Euclidean distance d_{eu} between two points on a branched polymer (a coarse grained version of the secondary structure) scales with the length of the shortest path between

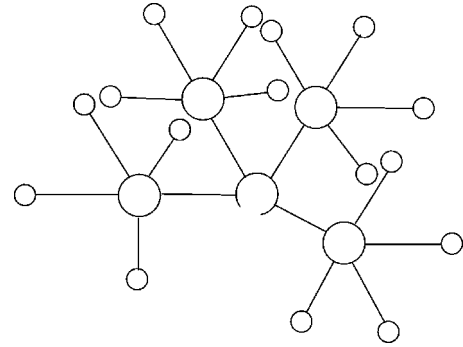


FIG. 6. Schematic “fractal” secondary structure. The large circles represent internal loops, the rods symbolize helices and the small circles hairpin loops. By choosing various complementary G-C sequences on the helices and A’s in the terminal loops, e.g., it is easy to design sequences that the structure prediction algorithms predict to fold as depicted. The resulting structure is extremely dense in real space and must be discarded as an admissible folding.

them, the so-called chemical distance d_{ch} . The exponent ν_{ch} defined by $d_{\text{eu}} \sim d_{\text{ch}}^{\nu_{\text{ch}}}$ is exactly known for the special case of a space-filling branched polymer, where it takes the value $\nu_{\text{ch}} = 4/5$ [20]. It is to be expected that the exponent is slightly smaller in the case of a branched polymer of arbitrary density and thus almost equals ν_{SAW} . This suggests that the wandering behavior is only weakly affected by the presence of branches connected to a self-avoiding walk, and the approximation of ν_{ch} by ν_{SAW} is quite good.

E. Discussion

If we assume that also in the three-dimensional case the presence of side branches does not increase substantially the wandering exponent from its value for the self-avoiding walk and suppose $R_g \sim N^{\nu_{\text{SAW}/2}}$ we encounter a consistency problem in the thermodynamic limit: the monomer density in space diverges as $N/R_g^3 \sim N^{1-3\nu_{\text{SAW}/2}}$ in the collapsed phase. This problem is common to all models considered above, irrespective of the existence of a denaturation transition. (It does not occur if the side branches have a much stronger effect than in two dimensions and increase the wandering exponent beyond $2/3$.) It reflects the fact that (local) entropy cost functions for interior loops are not sufficient to take into account global spatial constraints. The model customarily used in RNA prediction will thus be inconsistent for sufficiently large molecules in that it neglects excluded volume effects, deferring them to the subsequent analysis of the tertiary structure. This separation is, however, only justified as long as typically obtained secondary structures can easily be accommodated in space.

In order to demonstrate that the standard RNA structure predicting programs (Zuker’s mfold [12], Vienna package [9]) are indeed limited by their neglect of excluded volume effects, we have used them to determine the folding of RNA sequences that were deliberately designed to form fractal secondary structures (see Fig. 6). These were constructed starting from a short terminal helix that ends in a loop with two closed structures attached to it. Each of them again starts

with a short helix that ends in a branched loop with two outgoing stems, and so fourth in a self-similar way. The density of such treelike or starlike structures grows exponentially with their radius. Even with a modest number of bases, one can design sequences for which the structure prediction indeed yields the desired pairing pattern that can hardly be accommodated in space, and should therefore be ruled out.

While this is not a problem for small RNA molecules such as transfer RNA or ribosomal RNA, the natural sizes of messenger RNA are on the order of several thousand bases which is likely in the regime where excluded volume effects play an important role for the folding, and the usual structure prediction algorithms based solely on the base pairing pattern will likely fail. The inclusion of an effective loop costs to take into account the reduced available phase space might help to take into account those effects to some extent, but as we have seen they fail to cure the problem completely in the low-temperature phase.

At this point, it is worth mentioning that although neither loop costs nor the topological condition on the absence of pseudoknots are able to avoid (too) dense secondary structures, the situation would be much worse if no topological constraints were introduced in the model at all, that is, if all base pairings were allowed, irrespective of the resulting entanglement of the structure. It is rather obvious that a generic base pairing pattern obtained in such a model could not be accommodated in space: Let us consider a base pair and the two strands to which it belongs. If there is no constraint on the pairing behavior of the nearby bases within these strands, they can be paired to completely different parts of the chain which are then all forced into the same spatial region. The topological constraint forbidding pseudoknots weakens this tendency, since the substrand embraced by the given base pair is only allowed to interact with itself, which reduces largely the possibilities of spatial entanglement.

The above observations lead to the conclusion that typical structures for large molecules are determined by a competition between favorable base pairings and the requirement that the resulting secondary structure can be accommodated in space. This will result in rather densely packed and entangled spatial arrangements of RNA that we expect to exhibit very slow dynamics and glassiness due to the inevitable spatial hindrance to pass from one favorable folded state to another. This is indeed observed in folding experiments on large ribozymes, where several misfolded states compete with the correctly folded native state [21,22]. Thirumalai and Woodson [23] propose a “kinetic partitioning mechanism” to describe this type of slow dynamics, according to which a fraction of all molecules fold directly to the ground state, while the remaining molecules remain in metastable misfolded state until they find a pathway via the transition state ensemble to the native state. Glassiness may also arise purely on the level of secondary structure [18,24–29], where topological constraints (in the form of backbone connectivity or constraints on pseudoknots) introduce a weak frustration in the system and establish a multiplicity of metastable valleys in phase space. For small molecules this has been shown to lead to slow dynamics [30]. In large molecules the jamming of the spatial arrangement of energetically favorable second-

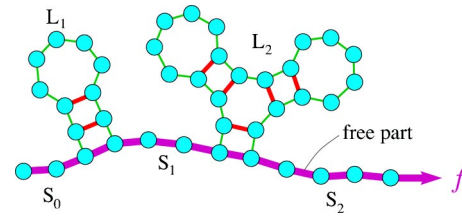


FIG. 7. RNA under an external force. The force pulls on the free part of the chain that can be subdivided into single-stranded portions containing S_i unpaired bases ($i=0, \dots, m$). Those are separated by closed structures, containing L_i bases ($i=1, \dots, m$). The free length derives from the terminating bonds of the closed structures [contributing $\zeta_b^{-1}(f)$ to the partition function] and the backbone elements in the single-stranded parts [contributing $\zeta_{ss}^{-1}(f)$ in the approximation of uncorrelated monomers].

ary structures probably plays an equally important role. It will be a major challenge to understand the interplay of secondary and tertiary structure and its relation with the closely related problem of protein folding.

III. RESPONSE OF RNA TO AN EXTERNAL FORCE

A. The partition function with force

In this section we will extend our formalism to treat problems with an external force. In experiments, the RNA molecule is usually fixed on one end, while the other end is manipulated by optical tweezers, magnetic beads in an inhomogeneous field, or the cantilever of an atomic force microscope. The extension of the molecule is monitored as a function of the applied force, or the position is imposed and the average force needed to maintain the position is measured. Here we will concentrate on the situation where the force is fixed while the extension is subject to thermal fluctuations.

To include the effect of an external force we have to add the term $-\vec{F} \cdot (\vec{r}_N - \vec{r}_1)$ to the energy of the system, where \vec{r}_i denotes the spatial position of the i th base. The force only acts on the free part of the chain, see Fig. 7 and we can rewrite the additional term as

$$-\vec{F} \cdot (\vec{r}_N - \vec{r}_1) = - \sum_{i=1}^{l_b-1} \vec{F} \cdot (\vec{r}_{b(i+1)} - \vec{r}_{b(i)}), \quad (14)$$

where $\{b(i)\}$ is the ordered list of all bases in the free part, and $b(l_b) = N$. There are two types of contributions to the sum: Terms with $b(i+1) = b(i) + 1$ correspond to successive bases in the backbone of the RNA, while terms with $b(i+1) > b(i) + 1$ correspond to the paired bases terminating the closed structure between bases $b(i)$ and $b(i+1)$. For simplicity we consider both the distance l of bases within the backbone and the distance l_b of covalently bonded bases as fixed. We treat the terminating hydrogen bonds of closed structures as free joints that are inserted between single strands of unpaired bases. They contribute a factor of $1/\zeta_b(f) = \sinh(\beta f l_b) / (\beta f l_b)$ to the partition function. For the single strands in between we will restrict ourselves to the simple model of a freely jointed chain. If one wants to fit to experimental data [5], one should, however, consider a more

realistic description involving correlations of the monomers on the scale of the persistence length ($l_p \approx 3l$). At high forces, bond elasticities should also be included. Both modifications can straightforwardly be taken into account in the formalism below.

The partition function $Z_N^f(f)$ including the external force can easily be obtained once the partition function $Z_N^c(f)$ of closed structures is known. We may decompose the sum over all structures into a sum over the number m and sizes L_i ($i = 1, \dots, m$) of closed structures in the free part, and the lengths S_i ($i = 0, \dots, m$) of the single-stranded segments linking them:

$$\begin{aligned} Z_N^f(f) &= \sum_{m \geq 0} \sum_{S_0, \dots, S_m \geq 0} \sum_{L_1, \dots, L_m \geq 0} \\ &\times \delta \left(S_0 + \sum_{i=1}^m (L_i + S_i) - N \right) Z_{S_0}^{ss} \\ &\times \prod_{i=1}^{m-1} \left[\frac{Z_{L_i}^c Z_{S_{i+1}}^{ss}(f)}{\zeta_b(f)} \right] \frac{Z_{L_m}^c Z_{S_m}^{ss}(f)}{\zeta_b(f)}. \end{aligned} \quad (15)$$

Here we have introduced the partition function $Z_N^{ss}(f)$ of a single-stranded segment with N backbone elements under the force f . The corresponding generating function follows from a discrete Laplace transform as

$$\begin{aligned} \Xi_f(\zeta; f) &\equiv \sum_N Z_N^f(f) \zeta^N \\ &= \frac{\Xi_{ss}^2(\zeta; f) \zeta}{(\Xi_{ss}(\zeta; f) - 1) \left(1 - \frac{\Xi_c(\zeta)}{\zeta_b(f)} \frac{\Xi_{ss}(\zeta; f) - 1}{\zeta} \right)}. \end{aligned}$$

The partition function is again found by inverting the Laplace transform, in particular, the free energy derives from the logarithm of the smallest singularity of $\Xi_f(\zeta; f)$. Apart from the singularities of $\Xi_c(\zeta)$ that we discussed in Sec. II C, Ξ_f also has a pole singularity $\zeta(f)$ when the denominator in Eq. (16) vanishes,

$$1 = \frac{\Xi_c(\zeta(f))}{\zeta_b(f)} \frac{\Xi_{ss}(\zeta(f); f) - 1}{\zeta(f)}. \quad (16)$$

Let us now fix the temperature. The singularity deriving from Ξ_c then takes the force-independent value $\zeta_* = \zeta_*(T)$. A phase transition occurs at the critical force $f_c(T)$, where $\zeta(f)$ crosses $\zeta_*(T)$. For larger forces, the force-extension characteristics follow from $\langle L(f)/N \rangle = -\beta^{-1} \partial \ln[\zeta(f)] / \partial f$ in the thermodynamic limit. Here, $L(f)$ denotes the projection of the end to end distance of the molecule onto the direction of the force.

In the following we are interested in two aspects of the force-extension curve. First, we will examine in detail the critical behavior around f_c . Later, we address the temperature dependence of $f_c(T)$ and an associated reentrance phenomenon slightly below denaturation. Finally, in Sec. IV we

will discuss in detail the dependence of the force-extension characteristics on the parameters s and η .

B. The critical behavior around f_c

In the following we treat the single-stranded parts as freely jointed chains, whose generating function is given by

$$\Xi_{ss}(\zeta; f) = 1 / [1 - \zeta / \zeta_{ss}(f)], \quad (17)$$

where $\zeta_{ss}(f) = \sinh(\beta fl) / (\beta fl)$. Furthermore, we only consider temperatures below denaturation.

In the vicinity of the (finite) critical force we may restrict ourselves to the relevant singularity structure of $\Xi(\zeta; f)$ and expand the denominator in Eq. (16) to lowest nontrivial order around f_c and ζ_* . Using Eq. (7) we find

$$\Xi_f(\zeta; f) \approx \frac{B}{(1 - \zeta / \zeta_*)^{1/2} - A(f - f_c)}, \quad (18)$$

or, on substituting $\zeta \equiv e^{-s}$ and $\zeta_* \equiv e^{-s_*}$,

$$\Xi_f(s; f) \approx \frac{B}{(s - s_*)^{1/2} - A(f - f_c)}. \quad (19)$$

A and B are slowly varying functions of f and ζ that we replace by their values at the critical point, $A \equiv A(f_c, \zeta_*)$ and $B \equiv B(f_c, \zeta_*)$. The (continuous) inverse Laplace transform of Eq. (19) is explicitly known, and we obtain the partition function in the transition region as

$$Z_N^f(f) = \frac{B e^{s_* N}}{\sqrt{\pi N}} \psi(A(f - f_c) N^{1/2}), \quad (20)$$

where $\psi(x) = 1 + \sqrt{\pi} x \exp(x^2) \operatorname{erfc}(-x)$. The force-extension characteristics follow immediately as $\beta L(f) = AN^{1/2} (\ln \psi)' [A(f - f_c) N^{1/2}]$. In the asymptotic regimes of the scaling variable $x = A(f - f_c) N^{1/2}$ one obtains the expansions

$$\beta L(f) \approx \begin{cases} 2A^2(f - f_c)N \left[1 + \frac{1}{2x^2} + O\left(\frac{e^{-x^2}}{x^3}\right) \right], & x \gg 1 \\ AN^{1/2} \sqrt{\pi} \left[1 + \frac{4 - \pi}{\sqrt{\pi}} x + O(x^2) \right], & |x| \ll 1 \\ \frac{2}{f_c - f} \left[1 - \frac{3}{2x^2} + O\left(\frac{1}{x^4}\right) \right], & x \ll -1. \end{cases} \quad (21)$$

Sufficiently above the critical force ($x \gg 1$), the extension grows linearly with $f - f_c$ and scales as the system size. The chain organizes in the kind of necklace: the number of closed structures in the free chain is proportional to N , their average size being finite. In the low force regime ($x \ll 1$) the chain is collapsed, but its extension diverges as $1/(f_c - f)$ upon approaching the critical point.

There are two critical exponents characterizing this phase transition: At the critical force ($x \ll 1$) the extension obeys a

power law $L \sim N^\delta$ with $\delta = 1/2$. The second exponent is related to the characteristic length scale in the problem, $N_c \propto (f - f_c)^{-\nu}$, where $\nu = 2$, as one can read off from the form of the scaling variable. Below we will see that N_c can be understood as a correlation length.

C. Correlations and length scales at $f \gg f_c$

Slightly above the critical force, the typical number of bases in a closed structure is given by $l_{\text{typ}} = N/n_{\text{cs}}(f)$, where $n_{\text{cs}}(f)$ is the number of closed structures in the free part. It has the same critical behavior as the extension, i.e., $n_{\text{cs}}(f \gg f_c) \propto N(f - f_c)$, and thus, $l_{\text{typ}} \propto 1/(f - f_c)$. This is surprising since the characteristic length scale $N_c \propto (f - f_c)^{-2}$ diverges much faster.

To understand the meaning of N_c let us introduce the indicator function η_i which equals 1 if base i belongs to the free part, and 0 otherwise. The correlation function $\langle \eta_i \eta_j \rangle$ is simply obtained as the ratio between the partition function with bases i and j constrained to be free, and the total partition function,

$$\langle \eta_i \eta_j \rangle = \frac{Z_{i-1}^f Z_{j-1-(i+1)}^f Z_{N-(j+1)}^f}{Z_N^f}. \quad (22)$$

Using Eq. (20) we obtain the connected correlation function as

$$\begin{aligned} \frac{\langle \eta_i \eta_j \rangle - \langle \eta_i \rangle \langle \eta_j \rangle}{\langle \eta_i \rangle \langle \eta_j \rangle} &= \frac{Z_{j-i-2}^f Z_N^f}{Z_{N-(i+1)}^f Z_{j-1}^f} - 1 \\ &\approx \frac{e^{-(j-i-2)[A(f-f_c)]^2}}{4\sqrt{\pi}(j-i-2)^{3/2}[A(f-f_c)]^3}, \end{aligned} \quad (23)$$

where the last approximation is valid in the scaling regime above the critical force for $i, N-j, N \gg (j-i) \gg [A(f-f_c)]^{-2}$. The quantity $N_c = [A(f-f_c)]^{-2}$ clearly appears as the correlation length beyond which the pairing behavior becomes essentially independent. To see why the correlation length is much larger than typical closed structures, let us look at the probability distribution of the sizes of the latter.

Suppose that a closed structure starts at base i . The probability that it is paired to the base $j = i + l + 1$ is given by

$$P(l) = \frac{Z_{i-1}^f Z_l^c Z_{N-(i+l+2)}^f}{\sum_{l' > 0} Z_{i-1}^f Z_{l'}^c Z_{N-(i+l'+2)}^f} \propto \frac{\exp[-lA^2(f-f_c)^2]}{l^{3/2}}, \quad (24)$$

from which we recover the expectation value for the structure size $l_{\text{typ}} = \langle l \rangle = \sum_l l P(l) \propto 1/(f - f_c)$. On the other hand, we can calculate the fraction $\chi(l_*)$ of bases that belong to closed structures of size at least l_* ,

$$\begin{aligned} \frac{\sum_{l=l_*}^{\infty} l P(l)}{\sum_{l=1}^{\infty} l P(l)} &\approx \frac{\int_{l_*[A(f-f_c)]^2}^{\infty} x^{-1/2} \exp(-x) dx}{\int_0^{\infty} x^{-1/2} \exp(-x) dx} \\ &= \text{erfc}[A(f-f_c)l_*^{1/2}]. \end{aligned} \quad (25)$$

A finite fraction of all bases thus belong to structures of size $O((f-f_c)^{-2})$ which sets the scale of the correlation length N_c . The vast majority of closed structures is much smaller, however.

D. The critical behavior with sequence disorder

After having understood the critical behavior in the homogeneous case, it is natural to ask whether disorder in the form of sequence inhomogeneities and varying pairing affinities between the bases is a relevant perturbation for the force-induced phase transition. In Refs. [31,32] the authors studied the force-induced unzipping of DNA and found the presence of disorder to significantly alter the critical behavior with respect to that of a homogeneous double strand. In RNA, the disorder effects are less pronounced, since the two opening transitions are not really of the same nature. In DNA, essentially all base pairs are broken up at the transition and the double strand becomes denatured. The force always acts only on the single base pair closing the yet unzipped double helix. In RNA, however, the transition occurs at a point where the entropy of large secondary structures and the free energy gain from the extension of the chain compete in a quite subtle manner, the base-pairing energies playing a less important role at the critical force. Furthermore, already in the critical region the force acts in parallel on a large number of globular structures aligned along the free part of the chain, which averages out the effect of disorder to some extent [33,34].

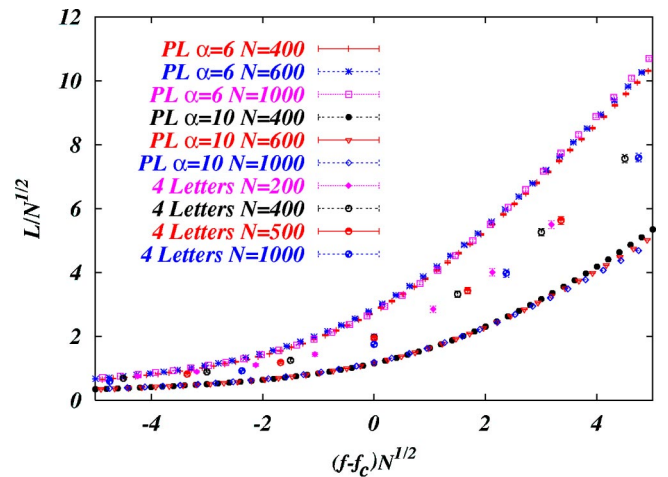


FIG. 8. Scaling plot of force-extension curves for disordered models (power law distributions with $\alpha = 6$ and $\alpha = 10$, and the four letters model) at high temperature ($T = 0.6$). For better visibility, the extension of the four letters model has been multiplied by 1.5.

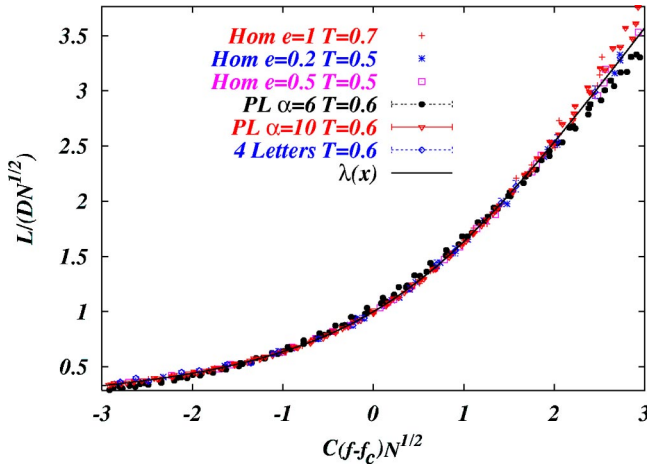


FIG. 9. Force-extension curves from Fig. 8 and for homogeneous models at different temperatures T and pairing energies e . All curves superpose with the analytical prediction (21) $L(f) = DN^{1/2}\lambda[C(f-f_c)N^{1/2}]$ upon rescaling with model-dependent factors C , D . The error bars are smaller than the symbol sizes.

In the case of RNA will be interested in the low-temperature regime where we can simplify the model by neglecting the loop cost function, i.e., putting $\phi \equiv 1$. Furthermore, we replace the pairing and stacking free energies by simple (temperature-independent) pairing energies e_{ij} between the bases i and j . This does not change the critical behavior at the force transition in the homogeneous case ($e_{ij} \equiv e$), and we checked numerically that a disordered model with pure stacking energies leads qualitatively to the same results as the pairing model.

As in earlier work [18,29,35], we consider different types of disordered models. The most natural one starts from RNA made of the four base species $b_i \in \{A, C, G, U\}$. The pairing energies will then depend on the sequence via $e_{ij} = E(b_i, b_j)$, where E is a symmetric 4×4 matrix. We used the simple matrix $E(C, G) = E(G, C) = -3$, $E(A, U) = E(U, A) = -2$ (Watson-Crick pairs) and $E(G, U) = E(U, G) = -1$ (wobble pairs), and $E = +\infty$ for all other pairs. Alternatively we considered more abstract random coupling models, where the e_{ij} are independent variables taken from a distribution $P(e)$. In the following we focus onto the two cases where $P(e)$ is Gaussian or has power law tails decaying like $|e|^{-\alpha}$, respectively, both being centered on a negative value.

The numerical evaluation of force-extension characteristics for these types of models is straightforward using the $O(N^3)$ recursion relation as introduced in Refs. [36,37] to compute the partition function Z_N^f exactly for a given realization of the disorder (see Refs. [33], for a related investigation, and [34] for a more thorough discussion of the effects of disorder and the low temperature behavior). In Fig. 8 we show scaling plots of the disorder-averaged force-extension characteristics for several disordered models at temperatures well above the glass transition temperature [18,29]. The data collapse in the critical regime was obtained optimizing the critical force in the scaling ansatz $L(f) = N^{1/2}\lambda[(f-f_c)N^{1/2}]$ supposing that the critical exponents are the same

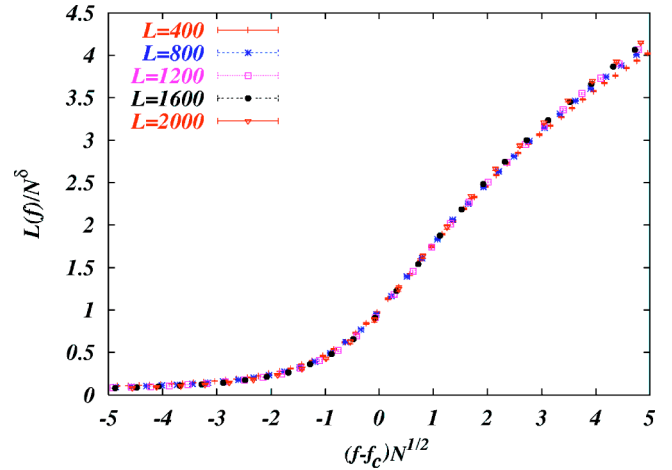


FIG. 10. Scaling plot of force-extension curves for the Gaussian model at $T=0$. The critical exponent is modified by the disorder to $\delta \approx 0.7$.

as in the homogeneous case ($\delta = 1/\nu = 1/2$). The scaling works well for the model with four letters and for the random coupling models with a Gaussian probability distribution, or with $P(e) \sim e^{-\alpha}$ and $\alpha > 4$. However, the data obtained for distributions $P(e) \sim e^{-\alpha}$ with $\alpha \leq 4$ (not shown) cannot be collapsed satisfactorily even when allowing for different critical exponents. This is due to the dominance of some rare but very strong couplings as will be explained in the following subsection.

By rescaling the axes of the plots with model-dependent metric factors C and D , $L(f) = DN^{1/2}\lambda[C(f-f_c)N^{1/2}]$, one can perfectly superpose the scaling functions for the different models with that of the homogeneous model, as shown in Fig. 9. This indicates that, above the glass transition temperature, disorder is irrelevant for the force-induced phase transition in the sense that all sufficiently short-ranged disordered models fall into the same universality class as homogeneous RNA. The latter suggests that the behavior of disordered RNA at high temperatures is well captured by a coarse-grained homogeneous description with renormalized parameters. The effect of disorder is washed out by thermal fluctuations which allow for a large number of secondary structures to be explored, so that the large entropy of secondary structures (with approximately the same number of bonds) dominates the physics.

The situation is different at low temperatures where the molecule is restricted to a small number of favorable foldings. The disorder-averaged force-extension curves at zero temperature are shown in Fig. 10 where the data collapse has been achieved with the general ansatz $L(f) = N^\delta \lambda[(f-f_c)N^{1/\nu}]$, optimizing for f_c , δ , and ν . As we will discuss in the next paragraph, the correlation length exponent $\nu = 2$ stays unchanged with respect to the homogeneous case. However, the exponent δ is modified ($\delta = 0.7$) [34].

E. Harris-type criterion for the relevance of disorder

The relevance of disorder for a phase transition can often be judged by applying Harris' criterion according to which

disorder is relevant if the specific heat exponent $\alpha = d\nu - 2$ is positive. Plugging in $d = 1$, since the sequence of bases is one dimensional, and using the correlation length exponent $\nu = 2$, we are led to conclude that disorder is marginal for the force transition.

In order to derive this result, we start from the homogeneous model where the pairing behavior is correlated up to the scale $\xi \sim |f - f_c|^{-2}$. The introduction of disorder will locally modify the value of the critical force, $f_c \rightarrow f_c + \Delta f_c$ whereby this only makes sense as long as one considers substrands that are large compared to the bare correlation length. To estimate the typical fluctuations $\Delta f_c(\xi)$ for regions of size ξ we observe that the opening transition is mainly of entropic nature (see Sec. IV) and results from the competition between the gain in stretching energy and the decrease of the number of possible secondary structures when the chain changes from a globular state to a necklace with a larger number of closed substructures.

The effect of the pairing energies comes only as a perturbation. It is thus reasonable to expect that $\Delta f_c(\xi)$ scales like the fluctuations of the average binding energy per base for secondary structures that are restricted to a substrand of length ξ . This implies $\Delta f_c(\xi) \sim \xi^{-1/2}$.

Locally, the correlation length is modified according to

$$\xi \sim |f - f_c - \Delta f_c(\xi)|^{-2}. \quad (26)$$

The scaling $\Delta f_c(\xi) \sim \xi^{-1/2} \sim (f - f_c)$ corresponds just to the limiting case, for which the critical force is still uniquely defined and the exponent $\nu = 2$ remains unchanged. This reflects the marginality as predicted by the vanishing of $d\nu - 2$.

The above considerations are wrong if the disorder distribution has large tails in which case very large though rare couplings may dominate the secondary structure pattern. For the random coupling model with power law tails $P(e) \sim |e|^{-\alpha}$ we can easily find a lower bound on α below which disorder will significantly alter the behavior of the model, rendering it even non-self-averaging. The energy fluctuations will scale like the energy of rare favorable secondary structures. We may estimate the latter by considering a ‘‘greedy’’ algorithm that constructs a secondary structure by choosing iteratively the base pair with the most negative energy available while respecting the topological constraints imposed by the pairs already chosen. There are $N(N-1)/2$ pairing energies available for the first step, and the best among them will scale as $N^{2/\alpha}$. There will be of the order of $\ln(N)$ further choices that lead to comparable energies, while in later stages the pairing energies will be significantly smaller. We thus expect the disorder induced energy (fluctuations) to scale at least as $\Delta E(N) \sim N^{2/\alpha} \ln(N)$, or $\Delta f_c(N) \sim N^{2/\alpha - 1} \ln(N)$. Thus, the fluctuations dominate for $\alpha \leq 4$. Indeed, we did not succeed in collapsing the numerical data for $\alpha = 3, 4$.

F. Why is the force induced transition of the second order?

It is rather unusual to find a continuous phase transition in force-extension experiments. The closely related globule-coil

transition in polymers is of the first order which is a consequence of the large finite size corrections to the free energy of the globular phase of a chain with N elements, $F_{\text{gl}}(N) = f_{\text{gl}}N + aN^{2/3}$. The term $aN^{2/3}$ takes into account solvent effects at the surface. Such corrections are essentially absent in the (extensive) random coil phase, since all monomers are more or less in a similar environment; but the free energy depends on the external force f since the structure is extensible, $F_{\text{coil}}(N) = Nf_{\text{coil}}(f)$. At the force where the extensive parts of the free energy become equal, $f_{\text{gl}} = f_{\text{coil}}(f)$, a discontinuous transition from the globular to the random coil phase takes place [38,39]. Mathematically, the first-order nature of the force transition is reflected by an essential rather than algebraic singularity in the Laplace transform of the partition function in the globular phase.

In our model for RNA a surface term in the globular phase is absent since solvation energies and surface effects are part of the tertiary interactions that are far less important than the base pairing (at least for small and intermediate sizes). The finite size corrections of the free energy in the globular phase are only of order $\ln(N)$. Thus, a subdivision of the chain into a necklace of globules is less costly than in the presence of surface effects. At the critical force, this leads to a continuous crossover from a single large globule to a necklace containing an extensive number of smaller globules which takes place over a force window decreasing as $N^{-1/2}$. In a two-dimensional homogeneous model of the globule-coil transition [40], the authors found the force transition to be continuous, too, which can be traced back to the absence of surface energies that grow polynomially with the system size.

It is worth mentioning that the thermodynamic phase transition would be absent in our model if the secondary structures were not allowed to contain multibranching loops, but were limited to single hairpins (with possible alignment gaps). Instead, there would only be an opening crossover. Although the continuous phase transition is an otherwise robust feature of all models irrespective of the details of the pairing and stacking rules, it critically depends on the topological constraints.

IV. DISCUSSION OF FORCE-EXTENSION CURVES IN THE THERMODYNAMIC LIMIT

In a recent paper [6] the authors claim that the inclusion of large stacking energies in the model renders the force transition first order in contrast to the second-order transition found in a model with only pairing energies. This has resulted from an erroneous analysis of a system of equations for generating functions that are real space analogs of our Eqs. (16) and (4). The authors were misled by a sharp force-induced denaturation crossover that masks the true thermodynamic transition at a smaller force where the extension begins to grow only very slowly as a function of force.

Before we discuss the general properties of force-extension curves, let us recall the parameters entering the model: Helices are required to contain at least n base pairs, and terminal loops closing a hairpin consist of at least t unpaired bases. The distance between monomers in the back-

bone is denoted by l , while l_b is the average distance between covalently bound bases. If we describe homogeneous RNA, we should use the empirically determined values $l_b \approx 4l$, $n=3$ for the minimal helix length and $t=3$ for the minimal hairpin loop. However, if the model is used to describe disordered RNA on a coarse-grained level, natural values are $l_b \ll l$, $n=1$, and $t=0$.

The base pair interactions are described by two parameters, $s(T)$ and $\eta(T)$, accounting for the pairing and stacking energy per base pair, and the cost for the initiation of a helix, respectively. As mentioned earlier, under physiological salt conditions the cooperativity parameter $\eta(T)$ is very small and thus favors the formation of long helices. The parameter $s(T) = \exp(-\beta f_{\text{pair}})$ is large at sufficiently low temperatures, but approaches $s \approx 1$ in the denaturation regime. In the following discussion we will consider $\eta(T)$ to be small throughout and $s(T)$ to be large at low temperatures, while approaching 1 around the denaturation temperature T_d . We distinguish the three temperature regimes, dropping the explicit temperature dependence of s and η : $s \gg 1$, $1 \gg s - 1 \gg \eta^{1/3}$, and $0 < s - 1 \ll \eta^{1/3}$. The case $s < 1$ corresponds to denatured RNA which is not of interest for force-extension studies.

For analytical simplicity we use the model without loop cost function, $\phi \equiv 1$, see Eq. (5). This is expected to be justified in the low-temperature regime $s \gg 1$, as well as at high forces, while the results for $s \approx 1$ at low force have to be taken with some care.

For the details of the calculations the reader should refer to the Appendix.

A. Critical force and reentrance

In order to discuss the thermodynamic limit of force-extension curves, we need the free energy per base $\phi(f)$ as a function of the force. It follows via $\zeta(f) \equiv \exp[-\beta\phi(f)]$ from Eq. (16) that we rewrite as

$$\zeta_{\text{ss}}(f) = \zeta(f) + \frac{\Xi_c(\zeta(f))}{\zeta_b(f)}. \quad (27)$$

Here we treat the single strands linking the closed structures as freely jointed chains, whose free energy per base is related to $\zeta_{\text{ss}}(f)$ via $\zeta_{\text{ss}}(f) \equiv \exp[-\beta\phi_{\text{ss}}(f)]$. The free energy per base in the globular phase ϕ_* is determined from the singularity $\zeta_* = \exp(-\beta\phi_*)$ of Ξ_c .

The chain begins to open when both free energies are equal, i.e., when $\phi_* = \phi(f)$, and the critical force is determined by the equation $\zeta_* = \zeta(f_c)$.

In the low-temperature regime ($s \gg 1$), the critical force depends on the parameter t for the minimal length of terminal loops. For $t \geq 1$ we find

$$f_c(T) \approx \frac{t}{4(l_b + l)} |f_{\text{pair}}(T)|. \quad (28)$$

The dependence on t is due to the fact that each hairpin terminates in a loop with at least t unpaired bases. The corresponding loss in energy is very important at low tempera-

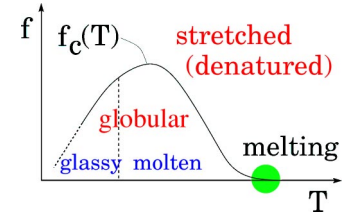


FIG. 11. Phase diagram of disordered RNA as a function of temperature and the force. The molecule undergoes a continuous opening transition at a critical force $f_c(T)$ that we predict to be nonmonotonic as a function of temperature. This gives rise to a reentrance phenomenon at fixed forces in a certain interval. At low temperature the system is in a glassy phase characterized by a small number of low-lying metastable states. At higher temperatures, RNA is in a molten state, that behaves in essentially the same way as a homopolymer. There is a thermal denaturation transition if the entropic penalties for loops are sufficiently large. Otherwise there is simply a crossover, and $f_c(T)$ never really vanishes.

tures and limits the thermally accessible phase space to rather elongated hairpinned structures with few branchings. This manifests itself in the critical force being proportional to the pairing free energy, almost as in the unzipping of DNA.

For $t=0$ the situation is different, since there is no energy cost associated to a hairpin, and therefore the available phase space of secondary structures is still large, even at low temperatures. The phase transition is governed by the competition between the force, trying to increase the number of closed structures in the free part, and the entropy that favors one big closed structure. The equation for the critical force reduces to

$$2\zeta_{\text{ss}}(f_c)\zeta_b(f_c) \approx 1, \quad (29)$$

and thus $f_c(T) \propto T/l$, almost independently of s . This reflects the purely entropic origin of the critical force sufficiently below denaturation.

Clearly, $t=0$ corresponds to an unphysical situation if the monomers in our model are interpreted as nucleotides. However, if we regard the homogeneous model as a coarse-grained description of a disordered base sequence, the monomers in the model stand for short substrands with an average affinity to pair with other substrands. The frustration in the secondary structure of disordered RNA necessarily leads to gaps in the base pairing that are usually larger than the minimal length of terminal loops. It is therefore unnecessary to impose a constraint on the terminal loops, i.e., we may safely put $t=0$ in this case. At the same time, the minimal length n of helices and the length parameters l and l_b have to be renormalized appropriately, as we indicated earlier.

In the denaturation regime, $s \approx 1$, the critical force becomes small. Independently of t , it decreases as

$$f_c(T) = \begin{cases} O((s-1)^{1/2}), & 1 \gg s-1 \gg \eta^{1/3} \\ O(\eta^{1/6}), & s-1 \ll \eta \end{cases} \quad (30)$$

on approaching denaturation.

For $t=0$ an interesting reentrance phenomenon occurs: We have seen that in this case $f_c(T)$ is an increasing function

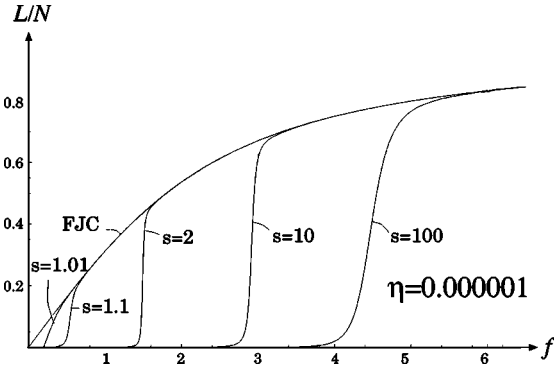


FIG. 12. At large cooperativity, the slope of the f - l curves at the transition is proportional to η . The phase transition is masked by the subsequent crossover whose steepness scales as $\eta^{-1/2}$, independently of s .

of temperature at sufficiently low temperatures, which is due to the entropic nature of the critical point. The large value of the binding parameter s merely forces the dominant secondary structure to have all bases paired, but does not influence the critical behavior otherwise. This picture does, however, not apply near denaturation where the base pairs are only loosely bound. Rather, Eq. (30) shows that f_c decreases essentially to zero in the denaturation regime. This is a consequence of the vanishing of the pair binding free energy ($\propto s-1$) which competes with the free energy gained from stretching. The latter grows as f^2 at low forces, and thus, the phase transition takes place on a force scale of the order of $(s-1)^{1/2}$ which vanishes at the denaturation transition.

The different behavior at low and high temperatures implies that the critical force reaches a maximum somewhere below the denaturation temperature which gives rise to the following reentrance phenomenon (see Fig. 11): If one fixes the external force at a value smaller than the maximum of $f_c(T)$ and then decreases the temperature, starting in the denatured regime, the molecule will collapse into the globular state when it crosses the critical line for the first time. However, it will reenter the stretched phase again at lower temperature when the second crossing occurs. Since the transitions are of second order this behavior would be seen as a (continuous) breathing of the molecule upon changing the

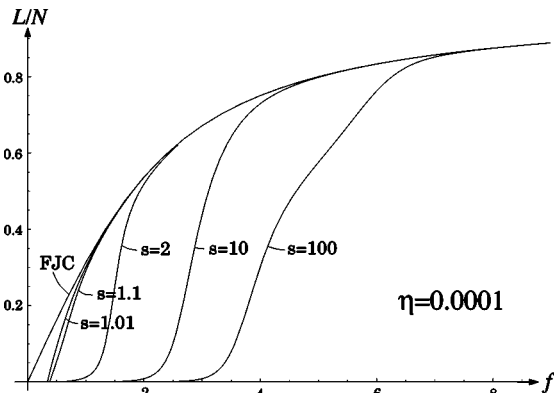


FIG. 13. At intermediate cooperativity the critical point and the crossover merge for sufficiently small values of s .

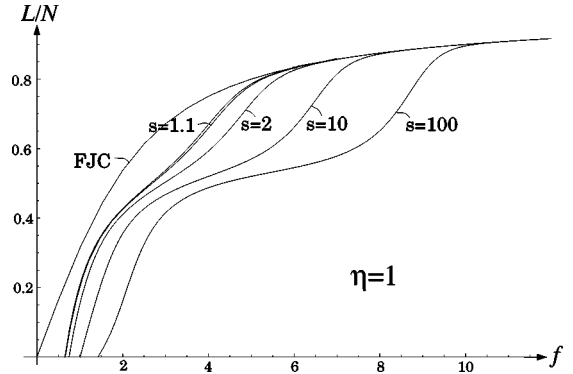


FIG. 14. In the absence of cooperativity, the large closed structures first open up to form smaller units. The plateau corresponds to necklaces of hairpins of the smallest possible size. Those structures are disrupted only at a higher force. This behavior is an artifact of very strong pairing energies.

temperature. We expect this effect to be relevant for disordered RNA sequences, where $t=0$ applies as we argued above. Such a behavior has been seen in numerical simulations of protein unfolding [41]. A similar effect was also predicted in the form of cold denaturation in DNA unzipping [42,43].

B. The opening crossover above f_c

In the following, we discuss several features of the force-extension curves in the thermodynamic limit, in particular, we will derive how the linear slope above the critical point and the characteristics of the subsequent crossover depend on the energy parameters of the model. The results are illustrated in Figs. 12, 13, and 14 where we plot force-extension curves for various pairs of s and η . The structural parameters are those appropriate for homogeneous RNA ($n=3$, $t=3$, and $l_b=4l$).

As we discussed in Sec. III, the extension $L(f)$ of the molecule slightly above the critical force grows like $N(f-f_c)$. The prefactor and the range of validity of the linear regime can be calculated from an expansion of Eq. (27) around the critical point. In the different temperature regimes, we find

$$L(f)/N(f-f_c) \sim \begin{cases} O(\eta s^{-1/4}), & s \gg 1 \\ O(\eta/(s-1)^4), & 1 \gg s-1 \gg \eta^{1/3} \\ O(\eta^{-1/3}), & s-1 \ll \eta^{1/3} \end{cases} \quad (31)$$

valid within a force window

$$f-f_c = \begin{cases} O(1), & s \gg 1 \\ O(s-1), & 1 \gg s-1 \gg \eta^{1/3} \\ O(\eta^{1/3}), & s-1 \ll \eta^{1/3}. \end{cases} \quad (32)$$

Note that at low temperatures ($s \gg 1$), the extension is suppressed by a factor of η (see Fig. 12). This is a consequence of the cooperativity in the system, i.e., the tendency to form long helices and large structures. Their size increases typically as η^{-1} , so that necklaces composed of such struc-

tures are very short. The effect is even enhanced for $t > 0$ where it is favorable to form longer helices in order to reduce the energy loss in terminal loops. The extension grows linearly with the force over a range of order $O(1)$, which is comparable to the critical force itself. This is in contrast to the case near denaturation where the extent of the linear regime is much smaller than f_c and the coefficient of the linear term can be appreciably large.

At higher forces the force-extension characteristics becomes nonlinear, and finally, the majority of base pairings opens up in a sharp crossover. This happens at a force f_* where the free energy of bases in the free part of the chain equals half of the base pairing free energy, $\zeta_{ss}(f_*) \approx s^{-1/2}$. Note that this is the analog of the critical force in DNA unzipping which is of a different nature than f_c . Beyond f_* , the molecule behaves essentially as a freely jointed chain.

The crossover is most interesting at low temperatures, $s \gg 1$. In the Appendix it is shown to take place within a force window determined by $\rho(f) \equiv [\eta s^{-(1+t)/2} / \zeta_b(f)]^{1/2} / \zeta_{ss}(f) - s^{-1/2} = O(1)$. Its width scales like

$$\Delta f \sim (\eta s^{1+(l_b/l-t-1)/2})^{1/2} \approx (\eta s)^{1/2}, \quad (33)$$

where we assumed $l_b/l \approx t+1$ for homogeneous RNA. This expresses the fact that the smallest possible hairpin loop [with length $l(t+1)$] will almost be a direct, stretched bridge between the bases of the adjacent pair (at distance l_b). The crossover is very sharp if the cooperativity, as measured by η^{-1} , is high.

Below the crossover [$\rho(f) \ll 1$] we find the extension to be small as compared to that of a freely jointed chain subject to the same force,

$$\frac{L(f)}{N} \approx \frac{L_{\text{FJC}}(f)}{N} \rho(f). \quad (34)$$

Above the crossover [$\rho(f) \gg 1$], the role played by closed structures is negligible, and the force-extension curve approaches that of a freely jointed chain,

$$\frac{L(f)}{N} \approx \frac{L_{\text{FJC}}(f)}{N} \left[1 + O\left(\frac{1}{\rho(f)}\right) \right]. \quad (35)$$

Near denaturation ($s \approx 1$), the situation is less interesting since both the phase transition and the final opening of the secondary structures take place at very small forces. Still, there is a discernible crossover in the case $s-1 \gg \eta^{1/3}$ that occurs at a force $f_* \approx f_c + O(s-1)$ while its width scales like $\Delta f \sim [\eta/(s-1)]^{1/2}$. In the limit $s-1 \ll \eta^{1/3}$, however, no crossover can be seen since all characteristic force scales behave as $\eta^{1/3}$.

V. CONCLUSION

We have studied several aspects of RNA folding on the level of the secondary structure. We concentrated on a homogeneous model that we expect to describe random sequences on a coarse-grained level as well, provided that the parameters are appropriately renormalized. This point of view is strongly supported by the finding that, at sufficiently

high temperatures, disorder is an irrelevant perturbation for the force-induced opening transition in the sense that scaling functions for disordered sequences superpose perfectly with the analytical curve calculated in the homogeneous case.

Our model takes into account base pairing and stacking energies, as well as entropic costs for loops. The latter have been shown to give rise to a thermodynamic denaturation transition if they are sufficiently strong. However, even when loop penalties are included, the dominant secondary structures in the collapsed phase of large molecules are too dense to be accommodated in three-dimensional space. Even at moderate, biologically relevant sizes, it is questionable whether the usual neglect of excluded volume (and other tertiary) interactions in the prediction of secondary structure is justified.

The force-induced unfolding has been studied in detail. We have characterized the second-order phase transition separating a globular from a necklacelike extensive phase. A correlation length diverging like $(f-f_c)^{-2}$ has been identified as the typical size of the largest closed structures that appear in the necklace above f_c . The critical scaling of the correlation length remains (marginally) unchanged upon introduction of disorder, as follows from a Harris-type criterion. However, in the low temperature glassy phase the other critical exponents are modified, in contrast to the high-temperature phase which belongs to the same universality class as a homogeneous polymer. This difference of the temperature regimes manifests itself as well in the force-extension characteristics. In the present paper, we have restricted ourselves to the discussion of the homogeneous or high-temperature case. The jumplike force-extension curves in the glassy regime are discussed in Ref. [34].

The second-order phase transition in force-induced unfolding has been shown to be a very robust feature of the model. In particular, it is independent of the specific pairing and stacking energies and further structural parameters. It occurs at a critical force that we predict to be a nonmonotonic function of temperature in the case of disordered sequences. This gives rise to a reentrance phenomenon when the temperature is varied at constant force.

When the stacking energy is large, the pairing behavior is highly cooperative. This can render the second-order phase transition almost invisible, since the extension grows very slowly as a function of $f-f_c$. On the other hand, a very sharp first-order-like crossover occurs at the (higher) force where essentially all base pairs are disrupted.

ACKNOWLEDGMENTS

It is a pleasure to thank Florent Krzakala, Marc Mézard, and Andrea Montanari for many helpful discussions. I am indebted to Marc Mézard for carefully reading the manuscript. The LPTMS is a Unité de Recherche de l'Université Paris XI associée au CNRS.

APPENDIX

In this appendix we outline in more detail how to obtain the characteristics of the force-extension curves of Sec. IV.

1. The singularities of the partition function

Let us start from Eq. (27),

$$\zeta_{ss}(f) = \zeta(f) + \frac{\Xi_c(\zeta(f))}{\zeta_b(f)}, \quad (\text{A1})$$

$$\Xi_c(\zeta) = \frac{1}{2} \left(1 - \zeta - \frac{\eta(s\zeta^2)^n(1-\zeta^t)/(1-\zeta)}{1-s\zeta^2 + \eta(s\zeta^2)^n} \right) - \sqrt{\frac{1}{4} \left(1 - \zeta - \frac{\eta(s\zeta^2)^n(1-\zeta^t)/(1-\zeta)}{1-s\zeta^2 + \eta(s\zeta^2)^n} \right)^2 - \frac{\eta(s\zeta^2)^n \zeta^t}{1-s\zeta^2 + \eta(s\zeta^2)^n}}. \quad (\text{A2})$$

The free energy per base of the globular phase is related to the singularity ζ_* of Ξ_c via $\zeta_* = \exp[-\beta\phi_c]$. ζ_* is given by the vanishing of the square root in Eq. (A2). In the low-temperature regime, $s \gg 1$, we find

$$\zeta_* \approx \begin{cases} s^{-1/2} \left(1 - \frac{3\eta}{2} \right), & t=0 \\ s^{-1/2} \left(1 - \frac{\eta}{s^{1/4}} \right), & t=1 \\ s^{-1/2} \left(1 - \frac{2\eta}{s^{1/2}} \right), & t=2 \\ s^{-1/2} \left(1 - \frac{\eta}{s^{1/2}} \right), & t \geq 3, \end{cases} \quad (\text{A3})$$

while near the denaturation, $s-1 \ll 1$, we obtain

$$\zeta_* \approx \begin{cases} s^{-1/2} \left(1 - \frac{8\eta}{(s-1)^2} \right), & 1 \gg s-1 \gg \eta^{1/3} \\ 1 - (2\eta)^{1/3}, & s-1 \ll \eta^{1/3}. \end{cases} \quad (\text{A4})$$

We recall that we assume $\eta \ll 1$ throughout.

2. The critical force

The critical force f_c has to be determined from the crossing of the two singularities, $\zeta(f_c) = \zeta_*$, or more explicitly,

$$\zeta_{ss}(f_c) = \zeta_* + \Xi_c(\zeta_*)/\zeta_b(f_c), \quad (\text{A5})$$

where we can use the fact that the square root in Eq. (A2) vanishes at ζ_* ,

$$\Xi_c(\zeta_*) = \sqrt{\frac{\eta(s\zeta_*^2)^n \zeta_*^t}{1-s\zeta_*^2 + \eta(s\zeta_*^2)^n}}. \quad (\text{A6})$$

In the low-temperature regime, $s \gg 1$, we have to distinguish different values of t . In the case $t \geq 1$ we find approximately

$$\zeta_{ss}(f_c(T)) \approx \zeta_* + \frac{\zeta_*^{t/2}}{\zeta_b(f_c(T))}. \quad (\text{A7})$$

which determines the pole singularity $\zeta(f)$ as a function of the force for $f > f_c$. The functional form of the generating function $\Xi_c(\zeta)$ for closed structures can be obtained from Eq. (5) as

We can neglect the first term on the right-hand side for $t < 2(l_b/l+1)$ which is always satisfied for homogeneous RNA, where we have $t \approx l_b/l-1$.

Recalling the definition $s = \exp(\beta|f_{\text{pair}}(T)|)$ and approximating $\zeta_{ss}(f) \equiv \beta l f / \sinh(\beta l f) \approx \exp(-\beta l f)$ and $\zeta_b(f) \equiv \beta l_b f / \sinh(\beta l_b f) \approx \exp(-\beta l_b f)$ at low temperatures, we find

$$f_c(T) \approx \frac{t}{4(l_b+l)} |f_{\text{pair}}(T)|. \quad (\text{A8})$$

The behavior for $t=0$ is different since all bases can be paired, even those at the end of a hairpin. In that case, Eq. (A5) reduces to

$$2\zeta_{ss}(f_c)\zeta_b(f_c) \approx 1, \quad (\text{A9})$$

implying $f_c(T) \propto T/l$, almost independently of s . This underlines the purely entropic origin of the critical force sufficiently below denaturation.

In the regime $s \approx 1$ the critical force is small, decreasing as

$$f_c(T) = \begin{cases} O((s-1)^{1/2}), & 1 \gg s-1 \gg \eta^{1/3} \\ O(\eta^{1/6}), & s-1 \ll \eta \end{cases} \quad (\text{A10})$$

on approaching the denaturation. This follows simply from an expansion of Eq. (A5) at low forces.

3. The linear regime above f_c

In the thermodynamic limit, the force-extension curve starts off linearly from zero extension at f_c . To obtain the slope of the curve, we expand Eq. (A1) around the critical point. In particular, we have to expand the square root in Eq. (A2) at ζ_* ,

$$\Xi_c(\zeta) \approx \frac{1}{2} \left(1 - \zeta - \frac{\eta(s\zeta^2)^n}{1-s\zeta^2 + \eta(s\zeta^2)^n(1-\zeta^t)/(1-\zeta)} \right) - \sqrt{A(\zeta_* - \zeta) + B(\zeta_* - \zeta)^2}. \quad (\text{A11})$$

As long as we can neglect $B(\zeta_* - \zeta)^2$ with respect to $A(\zeta_* - \zeta)$, i.e., for $\zeta_* - \zeta(f) \ll A/B$, we may obtain $\zeta(f)$ approximately from [see Eq. (A1)]

$$\begin{aligned} \zeta_{ss}(f)\zeta_b(f)|_{f_c}^f &\approx (f-f_c)[\zeta_{ss}\zeta_b]'(f_c) \\ &\stackrel{!}{\approx} \Xi_c(\zeta)|_{\zeta_*}^{\zeta(f)} \approx -[A(\zeta_* - \zeta(f))]^{1/2}, \end{aligned} \quad (\text{A12})$$

and thus,

$$\zeta(f) \approx \zeta_* - \frac{(f-f_c)^2}{A} [\zeta_{ss}\zeta_b]''(f_c). \quad (\text{A13})$$

The extension in the linear scaling regime above f_c then follows via $L(f) = -N\beta^{-1} \partial \ln[\zeta(f)] / \partial f \sim N(f-f_c)$. In the different temperature regimes, the coefficient of $N(f-f_c)$ scales like

$$L(f)/(N(f-f_c)) \sim \begin{cases} O(\eta s^{-t/4}), & s \gg 1 \\ O(\eta/(s-1)^4), & 1 \gg s-1 \gg \eta^{1/3} \\ O(\eta^{-1/3}), & s-1 \ll \eta^{1/3}. \end{cases} \quad (\text{A14})$$

The linear regime extends up to forces determined by $\zeta_* - \zeta(f) \approx A/B$, from which we obtain the force windows

$$f-f_c = \begin{cases} O(1), & s \gg 1 \\ O(s-1), & 1 \gg s-1 \gg \eta^{1/3} \\ O(\eta^{1/3}), & s-1 \ll \eta^{1/3}. \end{cases} \quad (\text{A15})$$

4. The nonlinear regime and the crossover

At higher forces the force-extension curve becomes nonlinear, but the typical size of closed structures is still large, so that the second term on the right-hand side of Eq. (A1) can not be neglected. Only in a later stage the base pairs open up completely, and the molecule becomes a freely jointed chain.

In the further discussion of the characteristics of the force-extension curve we will restrict ourselves to (real) homogeneous RNA with $1 < t \approx l_b/l - 1$. We will only treat the case $s \gg 1$ in some detail, the case $s \approx 1$ can be treated analogously.

Beyond the linear regime, we can solve for $\zeta(f)$ using the following approximation: The generating function for closed structures can be simplified by expanding the square root in Eq. (A2) with respect to the second term

$$\Xi_c(\zeta(f)) \approx \frac{\eta(s\zeta^2)^n \zeta^t}{(1-\zeta)[1-s\zeta^2 + \eta(s\zeta^2)^n] - \eta(s\zeta^2)^n \frac{1-\zeta^t}{1-\zeta}}. \quad (\text{A16})$$

Equation (27) for $\zeta(f)$ then reduces to a quadratic equation up to terms of the order of $\sqrt{\eta}$, except in a narrow region around the force f_* where the opening crossover takes place. The crossover force f_* is approximately determined by $\zeta_{ss}(f_*) = s^{-1/2}$.

Before proceeding let us estimate the width of the crossover. Considering the right-hand side of Eq. (A1), we note

that the closed structures are important as long as the variation of $\Xi_c(\zeta(f))/\zeta_b(f)$ with force dominates that of $\zeta(f)$. The crossover to the freely jointed chain regime occurs when both variations become comparable. Let us therefore write $\zeta(f) = s^{-1/2}(1 - \epsilon(f))$ and determine the value ϵ_X at which the correction $s^{-1/2}\epsilon_X$ begins to dominate the term $\Xi_c(\zeta(f))/\zeta_b(f)$. Using Eq. (A16) and the fact that η/ϵ_X will be small, we find $\Xi_c \approx s^{-t/2} \eta/\epsilon_X$. Equating this to $s^{-1/2}\epsilon_X \zeta_b(f)$, we find

$$s^{-1/2}\epsilon_X \equiv s^{-1/2} - \zeta(f) \sim [\eta s^{-(1+t)/2} / \zeta_b(f)]^{1/2}. \quad (\text{A17})$$

More quantitatively, one finds that for forces such that $\zeta_{ss}(f) - s^{-1/2} \gg [\eta s^{-(1+t)/2} / \zeta_b(f)]^{1/2}$, $\zeta(f)$ is given by

$$\zeta(f) \approx s^{-1/2} \left(1 - \frac{\eta s^{-t/2}}{2\zeta_b(f)[\zeta_{ss}(f) - s^{-1/2}]} \right). \quad (\text{A18})$$

The extension follows from a logarithmic derivative,

$$\frac{L(f)}{N} \approx \frac{L_{\text{FJC}}}{N} \frac{\eta s^{-t/2} \zeta_{ss}(f)}{\zeta_b(f)[\zeta_{ss}(f) - s^{-1/2}]^2}, \quad (\text{A19})$$

which is small as compared to that of a freely jointed chain at the same force.

For forces such that $\zeta_{ss}(f) - s^{-1/2} \ll -[\eta s^{-(1+t)/2} / \zeta_b(f)]^{1/2}$ closed structures play a negligible role. To the same degree of approximation as before one finds

$$\zeta(f) \approx \zeta_{ss}(f) \left[1 + O\left(\frac{\eta s^{-(1+t)/2}}{\zeta_b(f)[\zeta_{ss}(f) - s^{-1/2}]} \right) \right], \quad (\text{A20})$$

and the force-extension curve joins that of a freely jointed chain

$$\frac{L(f)}{N} \approx \frac{L_{\text{FJC}}(f)}{N} \left[1 + O\left(\frac{\eta s^{-(1+t)/2}}{\zeta_b(f)[\zeta_{ss}(f) - s^{-1/2}]} \right) \right]. \quad (\text{A21})$$

The force window over which the crossover takes place derives from Eq. (A17) and scales like

$$\Delta f \sim (\eta s^{1+(l_b/l-t-1)/2})^{1/2}. \quad (\text{A22})$$

For the case near denaturation, $s \approx 1$ the calculations are analogous and yield a crossover around $f_* \approx f_c + O(s-1)$ with a width scaling like $\Delta f \sim [\eta/(s-1)]^{1/2}$ for $s-1 \gg \eta^{1/3}$. In the limit $s-1 \ll \eta^{1/3}$, however, all characteristic force scales behave as $\eta^{1/3}$, and a separation into different regimes does not make sense.

- [1] P.-G. de Gennes, *Biopolymers* **6**, 715 (1968).
- [2] S.B. Smith, Y. Cui, and C. Bustamante, *Science* **271**, 795 (1996).
- [3] B. Maier, D. Bensimon, and V. Croquette, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12002 (2000).
- [4] U. Bockelmann *et al.*, *Biophys. J.* **82**, 1537 (2002).
- [5] A. Montanari and M. Mézard, *Phys. Rev. Lett.* **86**, 2178 (2001).
- [6] H. Zhou and Y. Zhang, *J. Chem. Phys.* **114**, 8694 (2001).
- [7] I. Tinoco and C. Bustamante, *J. Mol. Biol.* **293**, 271 (1999).
- [8] M. Zuker and D. Sankoff, *Bull. Math. Biol.* **46**, 591 (1984).
- [9] I.L. Hofacker *et al.*, *Monatsch. Chem.* **125**, 167 (1994).
- [10] I. Tinoco, O.C. Uhlenbeck, and M.D. Levine, *Nature (London)* **230**, 362 (1971).
- [11] I. Tinoco *et al.*, *Nature New Biology* **246**, 40 (1973).
- [12] D.H. Mathews, J. Sabina, M. Zuker, and D.H. Turner, *J. Mol. Biol.* **288**, 911 (1999).
- [13] B. Duplantier, *Phys. Rev. Lett.* **57**, 941 (1986).
- [14] L. Schäfer, C. von Ferber, U. Lehr, and B. Duplantier, *Nucl. Phys. B* **374**, 473 (1992).
- [15] Y. Kafri, D. Mukamel, and L. Peliti, *Phys. Rev. Lett.* **85**, 4988 (2000).
- [16] Y. Kafri, D. Mukamel, and L. Peliti, *Eur. Phys. J. B* **27**, 135 (2002).
- [17] R. Metzler *et al.*, *Phys. Rev. E* **65**, 061103 (2002).
- [18] R. Bundschuh and T. Hwa, *Phys. Rev. E* **65**, 031903 (2002).
- [19] B. Nienhuis, *Phys. Rev. Lett.* **49**, 1062 (1982).
- [20] A. Coniglio, *Phys. Rev. Lett.* **62**, 3054 (1989).
- [21] S.A. Woodson, *Nat. Struct. Biol.* **7**, 349 (2000).
- [22] R. Russell, I.S. Millett, S. Doniach, and D. Herschlag, *Nat. Struct. Biol.* **7**, 367 (2000).
- [23] D. Thirumalai and S.A. Woodson, *Acc. Chem. Res.* **29**, 433 (1996).
- [24] P.G. Higgs, *Phys. Rev. Lett.* **76**, 704 (1996).
- [25] R. Bundschuh and T. Hwa, *Europhys. Lett.* **59**, 903 (2002).
- [26] A. Pagnani, G. Parisi, and F. Ricci-Tersenghi, *Phys. Rev. Lett.* **84**, 2026 (2000).
- [27] A. Pagnani, G. Parisi, and F. Ricci-Tersenghi, *Phys. Rev. Lett.* **86**, 1383 (2001).
- [28] A. Hartmann, *Phys. Rev. Lett.* **86**, 1382 (2001).
- [29] F. Krzakala, M. Mézard, and M. Müller, *Europhys. Lett.* **57**, 752 (2002).
- [30] H. Isambert and E.D. Siggia, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6515 (2000).
- [31] D.K. Lubensky and D.R. Nelson, *Phys. Rev. Lett.* **85**, 1572 (2000).
- [32] D.K. Lubensky and D.R. Nelson, *Phys. Rev. E* **65**, 031917 (2002).
- [33] U. Gerland, R. Bundschuh, and T. Hwa, *Biophys. J.* **81**, 1324 (2001).
- [34] M. Müller, F. Krzakala, and M. Mézard, *Eur. Phys. J. E* **9**, 67 (2002).
- [35] E. Marinari, A. Pagnani, and F. Ricci-Tersenghi, *Phys. Rev. E* **65**, 041919 (2002).
- [36] M. Waterman, *Adv. Math. Suppl. Studies* **1**, 167 (1978).
- [37] J.S. McCaskill, *Biopolymers* **29**, 1105 (1990).
- [38] P.L. Geissler and E.I. Shakhnovich, *Phys. Rev. E* **65**, 056110 (2002).
- [39] P.L. Geissler and E.I. Shakhnovich, *Macromolecules* **35**, 4429 (2002).
- [40] F.S.D. Marenduzzo and A. Maritan, *J. Phys. A* **35**, L233 (2002).
- [41] D.K. Klimov and D. Thirumalai, *J. Phys. Chem. B* **105**, 6648 (2001).
- [42] E. Orlandini *et al.*, *J. Phys. A* **34**, L751 (2001).
- [43] D. Marenduzzo *et al.*, *Phys. Rev. Lett.* **88**, 028102 (2002).