Emergence of secondary motifs in tubelike polymers in a solvent

Chiara Poletto,[1] Achille Giacometti,[2] Antonio Trovato,[1,3] Jayanth R. Banavar,[4] and Amos Maritan[1,3,5]
[1]*Dipartimento di Fisica, Università di Padova, via Marzolo 8 I-35131 Padova*
[2]*Dipartimento di Chimica Fisica, Università di Venezia, Calle Larga S. Marta DD 2137, I-30123 Venezia, Italy*
[3]*CNISM, Unità di Padova, Via Marzolo 8, I-35131 Padova, Italy*
[4]*Department of Physics, 104 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*
[5]*Sezione INFN, Università di Padova, I-35131 Padova, Italy*
(Received 7 December 2007; published 18 June 2008)

We study the effects of two kinds of interactions in tubelike polymers and demonstrate that they result in the formation of secondary motifs. The first has an entropic origin and is a measure of the effective space available to the solvent. The second arises from solvophobic interactions of the solvent with the polymers and leads to an energy proportional to the contact surface between the tube and solvent particles. The solvent molecules are modeled as hard spheres and the two interactions are considered separately with the solvent density affecting their relative strength. In addition to analytical calculations, we present the results of numerical simulations in order to understand the role played by the finite length of short polymers and the discrete versus continuum descriptions of the system in determining the preferred conformation.

DOI: 10.1103/PhysRevE.77.061804 PACS number(s): 82.35.Lr, 05.70.−a, 61.20.−p, 87.15.kr

The study of entropic effects has a venerable history in physics. Onsager [1] studied the nature of an entropically driven isotropic-nematic transition of a fluid composed of infinitely thin hard rods. The addition of a small amount of polymers to a colloidal suspension induces an effective attraction between the colloidal particles which can lead to flocculation through the Asakura-Oosawa (AO) mechanism [2]. Our focus in this paper is on studying the optimal conformations of a tube subject to compaction through its interactions with the solvent. Our principal results are that the tube adopts helical and planar conformations with the latter resulting from the discreteness of the tube.

Our results are useful for understanding proteins [3,4]. In spite of their complexity, proteins fold into a limited number [5] of evolutionarily conserved structures [6] made up of helices and almost planar sheets. It has been proposed that the geometries of protein folds originate from the common attributes of all proteins that can be encoded in a simple geometrical model of a flexible tube of nonzero thickness [7–9]. Proteins occupy the marginally compact phase of a tubelike polymer [9–13] which is rich in secondary motifs and has a relatively low structural degeneracy compared to the generic compact polymer phase. While it is well recognized that the formation of secondary structure arises from backbone-backbone hydrogen bonds [14,15], the solvent is known to play a fundamental role in promoting the folding process [3]. The importance of the solvation free energy and of entropy-induced interactions has been increasingly recognized in the last few years [16–19].

In a system of hard spheres, the excluded volume constraint can be enforced by forbidding configurations in which the distance between pairs of sphere centers is smaller than the sphere diameter. In the continuum limit, a self interacting tubelike polymer cannot have two-body interactions [20]. The self-avoidance of a flexible tube of thickness Δ (the tube radius) can be enforced through a three-body potential [7,20,21] which forbids conformations in which the radius of a circle through an arbitrary triplet of points on the axis is smaller than the tube thickness. A solvophobic polymer in a solvent, such as a protein *in vivo*, adopts a self-avoiding conformation in which the contact surface is minimized, an effect that may be captured by considering [9] an effective solvent induced (free-) energy proportional to the contact surface between the tube and the solvent particles. Furthermore, due to the discrete nature of the solvent one expects entropy-induced effective interactions of the AO type [2,16,18] as investigated by Snir and Kamien [17] and more recently by Hansen-Goos *et al.* [19] for a tubelike polymer in helical conformations. The AO type of interaction yields an effective (free-) energy proportional to the volume excluded to the solvent.

It was noted recently by Hansen-Goos *et al.* [19,22] that Hadwiger's theorem of integral geometry [23], under rather general conditions, allows one to write the solvation free energy $F_{\rm sol}$ of the tube in a solvent as

$$F_{\rm sol} = PV_{\rm exc} + \sigma\Sigma_{\rm acc} + kC + \bar{k}X, \tag{1}$$

where $V_{\rm exc}$, $\Sigma_{\rm acc}$, C, and X are the volume excluded to the solvent, the area accessible to the solvent and integrated mean and Gaussian curvatures of the accessible surface, respectively. P and σ are the solvent pressure and the planar surface tension, respectively, whereas k and $\bar{k}$ represents the two bending rigidities; the properties of the solvent enters through these coefficients and can hence be computed separately. The AO approximation amounts to setting $\sigma=k=\bar{k}=0$ and assuming an ideal gas solvent $P=P_{\rm id}$ corresponding to an infinitely diluted system. The principal aim of our letter is to provide exact expressions for the excluded volume and the contact surface of a tubelike polymer in a solvent of hard spheres and to investigate the effects of these terms both analytically and by Monte Carlo simulations in determining the optimal conformations of short polymers.

In a continuum description, the axis of a tube is described by the curve $\mathbf{R}(s)$ parametrized by its arc length s. The Frenet reference frame at location s is given in terms of the unit vectors $\{\hat{\mathbf{T}}(s),\hat{\mathbf{N}}(s),\hat{\mathbf{B}}(s)\}$, the tangent, the normal, and

the binormal, respectively [24]. The curvature at location s, $\kappa(s)$ (the inverse radius of curvature), is defined in terms of one of the Frenet-Serret equations $d\hat{\mathbf{T}}(s)/ds=\kappa(s)\hat{\mathbf{N}}(s)$ and $\hat{\mathbf{B}}(s)=\hat{\mathbf{T}}(s)\times\hat{\mathbf{N}}(s)$.

The position vector of a point inside the tube of length L is

$$\mathbf{r}(s,\rho,\theta)=\mathbf{R}(s)+\rho[\cos\ \theta\hat{\mathbf{N}}(s)+\sin\ \theta\hat{\mathbf{B}}(s)], \tag{2}$$

where $0\leq\theta\leq 2\pi$, $0\leq\rho\leq\Delta$, and $0\leq s\leq L$. The volume element in the above curvilinear coordinate system is given by $dV=\rho|\chi(s,\rho,\theta)|dsd\rho d\theta$, where $\chi(s,\rho,\theta)\equiv 1-\kappa(s)\rho\cos\theta$. Let ϵ be the radius of the solvent particles assumed to be spherical. For a straight tube the excluded volume is simply $V_{\mathrm{exc},S}=\pi L\Delta_\epsilon^2$, where $\Delta_\epsilon\equiv\Delta+\epsilon$, since the volume inaccessible to the centers of the solvent particles is a tube with thickness Δ_ϵ. In order to calculate the excluded volume when the tube curls and different regions of the tube come close to each other, we have to consider points which are shared by more than one circular section of radius Δ_ϵ associated with distinct points of the tube axis. We avoid multiple counting by introducing the integer function $n(s,\rho,\theta)$ which gives the number of circular cross sections of radius Δ_ϵ, centered at distinct points of the tube axis, containing the point $\mathbf{r}(s,\rho,\theta)$ where $\rho\leq\Delta_\epsilon$. A point, $\mathbf{r}(s,\rho,\theta)$, belongs to the circular section associated with $\mathbf{R}(s')$ if the vector distance $\mathbf{d}_{s,\rho,\theta}(s')=\mathbf{r}(s,\rho,\theta)-\mathbf{R}(s')$ is perpendicular to the tangent at s', $\hat{\mathbf{T}}(s')$, and its modulus is less than Δ_ϵ. Note that $n(s,\rho,\theta)\geq 1$ because $s'=s$ trivially satisfies the above conditions. The excluded volume is thus given by

$$V_{\mathrm{exc}}=\int_{\mathcal{D}}dV\frac{1}{n(s,\rho,\theta)}, \tag{3}$$

where $\mathcal{D}$ is the domain $0\leq\theta\leq 2\pi$, $0\leq\rho\leq\Delta_\epsilon$, and $0\leq s\leq L$. The same integer function can be used to obtain an expression for the contact surface on setting the ρ coordinate equal to the radius Δ of the tube

$$\Sigma_{\mathrm{cont}}=\int_{\mathcal{D}}dV\delta(\rho-\Delta)\Theta[1-n(s,\Delta_\epsilon,\theta)]. \tag{4}$$

Here $\Theta(x)$ is the step function equal to 1 if $x\geq 0$ and 0 otherwise. It is not difficult to show that Eq. (4) is equivalent to the formula presented in Ref. [9], for both an infinite length tube and a finite one. In the latter case hemispheres needs to be attached to both ends while computing $n(s,\rho,\theta)$, as is the case in our numerical simulations. In a swollen configuration corresponding to the absence of overlap, $n(s,\rho,\theta)=1$ for all (s,ρ,θ) and $\kappa(s)<\Delta_\epsilon^{-1}$ for all s, and Eqs. (3) and (4) reproduce the correct results $V_{\mathrm{exc},S}=\pi L\Delta_\epsilon^2$ and $\Sigma_{\mathrm{cont},S}=2\pi\Delta L$. It is illuminating to rewrite the difference $V_{\mathrm{exc},S}-V_{\mathrm{exc}}$ as

$$\Delta V\equiv V_{\mathrm{exc},S}-V_{\mathrm{exc}}=V_{\mathrm{ov}}-2V_{\chi<0} \tag{5}$$

where we have introduced the overlap volume

$$V_{\mathrm{ov}}=\int_{\mathcal{D}}dV\frac{n(s,\rho,\theta)-1}{n(s,\rho,\theta)} \tag{6}$$

which is the part of the excluded volume with nonvanishing overlap, i.e., with $n(s,\rho,\theta)>1$, and the volume related to the highly curved regions given by $\kappa(s)>\Delta_\epsilon^{-1}$,

$$V_{\chi<0}=\int_{\mathcal{D}}dV\Theta[-\chi(s,\rho,\theta)]. \tag{7}$$

As a result, the difference ΔV between the excluded volumes in a swollen and a compact conformation is not simply equal to the overlap volume V_{ov} as in the usual Asakura-Oosawa mechanism involving rigid spheres [2]. There is no change in shape of the colloidal particle whereas the tube polymer undergoes a shape change upon folding. This difference has some interesting consequences. Suppose, for instance, that the tube is twisted and connected at the two ends to form a torus/donut of thickness Δ and radius R. If $R>\Delta_\epsilon$ one clearly has $\Delta V=0$. If $R\leq\Delta_\epsilon$, a straightforward calculation shows that $V_{\mathrm{ov}}=V_{\chi<0}$ so that $\Delta V=-\Phi(R/\Delta_\epsilon)4\pi\Delta_\epsilon^3/3$ becomes negative with the function

$$\Phi(x)=\left(1+\frac{1}{2}x^2\right)\sqrt{1-x^2}-\frac{3}{2}x\arccos x \tag{8}$$

positive in the interval $0\leq x\leq 1$. The swollen configuration is entropically more favorable compared to a tight donut. This counterintuitive result stems from the fact that the curvature yields both inflated and deflated volume elements. In the absence of overlap, the two effects balance against each other yielding $\Delta V=0$. When $\chi<0$, the same balance can be achieved only by allowing deflated "negative" volumes. An imbalanced inflated volume is then present, which may overturn the excluded volume reduction due to the presence of overlap, as in the donut case.

Consider an infinite helix of radius R and pitch P. The translational invariance along the tube axis simplifies the calculations, and we can consider both the excluded volume and the contact surface per unit length. Figure 1 shows the values of $c=P/R$ for the helices which minimizes V_{exc}/L, as the solvent molecule radius changes—the analogous results for Σ_{cont}/L exhibits similar behavior. When ϵ is below a threshold value, the optimal configuration corresponds to the ideal helix [7,11,17,19,25] which simultaneously minimizes the local and nonlocal radius of curvature [21]. For an ideal helix, $c=c^*=2.5126\cdots$ which is within 3% of the value for α helices in proteins [7]. Interestingly, a similar optimality condition is also found for double strand DNA [26]. These observations suggest that α helices might originate in general optimization processes transcending the large differences in the amino acid sequences of different proteins. As the solvent sphere size ϵ increases beyond the threshold, Fig. 1 shows that the optimal value of c decreases from its value in the ideal helical structure. Our results confirm those obtained in previous work [17,19]. For the excluded volume interaction, the helix maintains its ideal pitch-to-radius ratio c^* up to a critical value $\epsilon^*\simeq 0.0835\Delta$. Above this threshold, the ideal helix is no longer the optimal conformation—even though the excluded volume along the central axis increases, there is

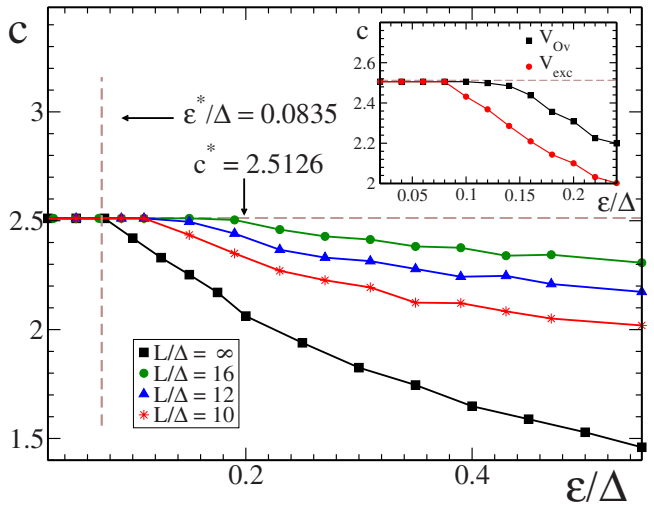

FIG. 1. (Color online) Plot of the pitch-to-radius value $c=P/R$ for the optimal helix on minimizing the excluded volume as a function of the ratio ϵ/Δ, for different helix lengths L. Below a given threshold value (shown for the $L=\infty$ case) of the solvent radius, the ideal helix conformation is obtained. Our results were obtained by first setting the value of the pitch P so that two different turns of the helix have minimum distance for a fixed R [17]. Second, we identify the value of c which minimizes V_{exc}/L for a given ϵ and these values are plotted against each other. An analogous calculation for Σ_{cont}/L shows a similar trend toward optimal finite helices having a greater c with respect to the infinite helix. In the inset, the different behaviors obtained on considering the minimization of the excluded volume and the maximization of the overlap volume is shown for the $L=\infty$ case.

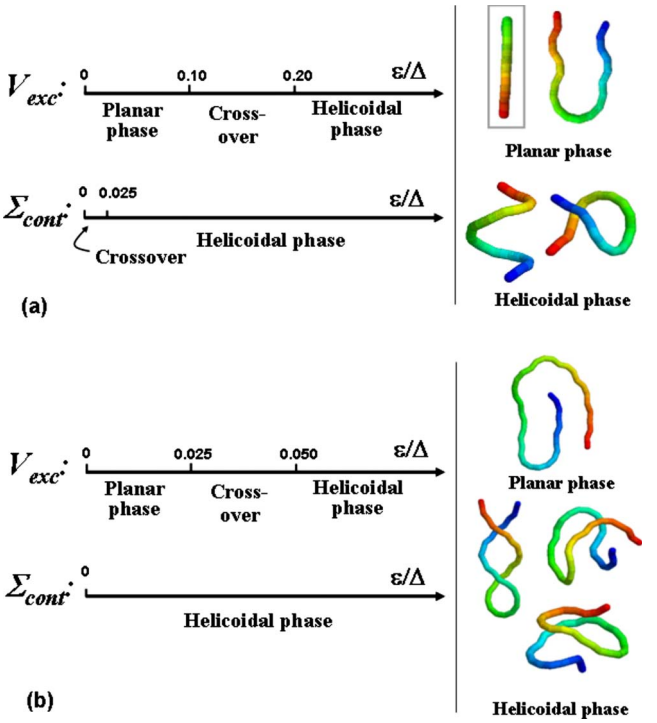

FIG. 2. (Color online) Conformations adopted by tubes of length (a) $L=10\Delta$ and (b) $L=20\Delta$ subject to either the excluded volume or contact surface prescriptions for promoting compaction. The values of ϵ/Δ are 0.005, 0.025, 0.05, 0.10, 0.20, and 0.50. The resulting conformations for the shorter tube are either saddles or helices for $\epsilon/\Delta \gtrsim 0.20$, while for smaller ϵ, after a crossover phase, the hairpin becomes the ground state—the planarity of the structure is highlighted on the left of (a). For the longer tube, the planar phase consist of β-sheet-like structures and the helicoidal phase is characterized by double helices, saddles and irregular helices. The planar phase disappears for the contact surface case. All these conformations are akin to those found in Ref. [11].

a better overlap between subsequent turns thus decreasing the excluded volume. The inset of Fig. 1 shows the results maximizing the overlap rather than minimizing the excluded volume. The optimal value c^* is maintained up to $\epsilon^* \simeq 0.1\Delta$, a value larger than that obtained in the minimum excluded volume case. This mirrors the situation with the donut: for large curvature, there is a nonzero, albeit small, difference between ΔV and V_{ov} as dictated by Eq. (5). An exact analytical computation for the contact surface [27] yields an identical pattern with a smaller critical value $\epsilon^*=0.04627\Delta$ for the solvent radius. Again, this is in perfect agreement with the value given in Ref. [19] using a different method. Our method allows one to treat the case of finite helices efficiently using numerical integration. Figure 1 shows the optimal value of c for helices of varying length L. Short helices show a tendency to remain ideal for higher values of the solvent radius.

When the tube is very long or there are many independent tubes, one might expect that the helical conformation is destabilized. For many long tubes, an optimal configuration which minimizes both the contact surface and the excluded volume is likely a regular arrangement of parallel tubes forming a hexagonal arrangement somewhat reminiscent of the Abrikosov phase in type-II superconductors [28]. We have verified that $N \geq 4$ infinitely long tubes have a lower excluded volume per unit length with respect to the optimal helix (for a single tube one should also take into account the $N-1$ turns).

In order to investigate the possibility for other kinds of optimal conformations for a short tube as well as the role of discrete versus continuum descriptions of the tube, we resort to Monte Carlo simulations using the formulae given above for both the excluded volume and the contact surface. We seek to identify the conformations which minimize V_{exc} or Σ_{cont} as the solvent radius ϵ and the polymer length L are varied (the thickness Δ is kept fixed because the results depend only on the dimensionless ratios ϵ/Δ and L/Δ). Our Monte Carlo simulations combine crankshaft and pivot moves of a homopolymer with a Metropolis annealing schedule [29] following initializing the chain in an arbitrary extended conformation. We have considered two different polymer lengths, $L=10\Delta$ and $L=20\Delta$, fixing the bond length $b=\Delta/2$. The results are summarized in Fig. 2—the optimal structures obtained on varying ϵ are shown—while the values of $\Delta V=V_{\mathrm{exc},S}-V_{\mathrm{exc}}$ and $\Delta\Sigma=\Sigma_{\mathrm{cont},S}-\Sigma_{\mathrm{cont}}$ reached are reported in Table I. Figure 2(a) displays the conformations obtained with $L=10\Delta$. For both the excluded volume and contact surface, two different "phases" emerge: at small ϵ, a planar phase, in which the ground state has a typical planar hairpin structure, is found; at large ϵ, a "helicoidal" phase is found in which helices and saddles dominate, with compa-

TABLE I. Values of $\Delta V/(L\Delta^2)$ and $\Delta\Sigma/(L\Delta)$ related to the optimal structures, for the tested values of ϵ/Δ and L/Δ.

	$L=10\Delta$		$L=20\Delta$	
ϵ/Δ	$\Delta V/(L\Delta^2)$	$\Delta\Sigma/(L\Delta)$	$\Delta V/(L\Delta^2)$	$\Delta\Sigma/(L\Delta)$
0.005	0.049	0.62	0.053	0.98
0.025	0.061	0.80	0.067	1.2
0.050	0.079	0.98	0.098	1.5
0.100	0.12	1.2	0.18	1.9
0.200	0.24	1.6	0.39	2.1
0.500	0.77	2.1	1.3	2.9

rable V_{exc} and Σ_{cont}. The two different regimes are separated by a crossover region characterized by the coexistence of all these structures. Note that the excluded volume effect exhibits a stronger propensity for the planar phase. This is evident for $L=20\Delta$ shown in Fig. 2(b)—the contact surface planar phase disappears already at $\epsilon/\Delta=0.005$, the lowest value tested. The helicoidal phase in this case consist of double helices, saddles and irregular helices, with turns of different lengths. Our simulations suggest the existence of an energy barrier between the two classes of conformations (hairpin versus helix or saddle) in the crossover region. The appearance of planar structures is a consequence of the discrete nature of the polymer (nonzero bond length) which plays a crucial role especially at small ϵ.

In summary, tubelike polymers with naturally arising solvent induced interactions exhibit low free energy conformations with secondary motifs. This suggests that the secondary motifs commonly found in bio-polymers such as proteins and DNA have a common and fundamental origin which transcends chemical detail. When the discrete nature of the polymer dominates, planar conformations akin to the β sheet in proteins emerge along with helical conformations as the optimal ones. In contrast, when the solvent molecule size is sufficiently large so that the discrete nature of the polymer can be neglected, the continuum approximation is valid and helical conformations dominate with single and double helix conformations.

We are grateful to T. X. Hoang for stimulating discussions. This work was supported by PRIN Grant No. 2005027330 in 2005 and INFN.

[1] L. Onsager, Ann. N.Y. Acad. Sci. **51**, 627 (1949).
[2] S. Asakura and F. Oosawa, J. Chem. Phys. **22**, 1255 (1954); S. Asakura and F. Oosawa, J. Polit. Sci. **33**, 183 (1958); A. Vrij, Pure Appl. Chem. **48**, 471 (1976); C. N. Likos, Phys. Rep. **348**, 267 (2001).
[3] A. V. Finkelstein and O. B. Ptitsyn, *Protein Physics* (Academic, New York, 2002).
[4] A. R. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York, 1998).
[5] C. Chothia, Nature (London) **357**, 543 (1992); C. Chothia and A. V. Finkelstein, Annu. Rev. Biochem. **59**, 1007 (1990).
[6] M. Denton and C. Marshall, Nature (London) **410**, 417 (2001); C. Chothia, J. Gough, C. Vogel, and S. A. Teichmann, Science **300**, 1701 (2003).
[7] A. Maritan, C. Micheletti, A. Trovato, and J. R. Banavar, Nature (London) **406**, 287 (2000).
[8] J. R. Banavar and A. Maritan, Rev. Mod. Phys. **75**, 23 (2003); J. R. Banavar and A. Maritan, Annu. Rev. Biophys. Biomol. Struct. **36**, 261 (2007).
[9] J. R. Banavar, T. X. Hoang, J. H. Maddocks, A. Maritan, C. Poletto, A. Stasiak, and A. Trovato, Proc. Natl. Acad. Sci. U.S.A. **104**, 17283 (2007).
[10] T. X. Hoang, A. Trovato, F. Seno, J. R. Banavar, and A. Maritan, Proc. Natl. Acad. Sci. U.S.A. **101**, 7960 (2004).
[11] J. R. Banavar, T. X. Hoang, A. Maritan, F. Seno, and A. Trovato, Phys. Rev. E **70**, 041905 (2004).
[12] D. Marenduzzo, A. Flammini, A. Trovato, J. R. Banavar, and A. Maritan, J. Polym. Sci., Part B: Polym. Phys. **43**, 650 (2005).
[13] T. R. Lezon, J. R. Banavar, and A. Maritan, J. Phys.: Condens. Matter **18**, 847 (2006).
[14] L. Pauling, R. B. Corey, and H. R. Branson, Proc. Natl. Acad. Sci. U.S.A. **37**, 205 (1951); L. Pauling and R. B. Corey, *ibid.* **37**, 729 (1951).
[15] D. Eisenberg, Proc. Natl. Acad. Sci. U.S.A. **100**, 11207 (2003); D. Poland and H. A. Scheraga, *Theory of Helix-Coil Transition in Biopolymers* (Academic, New York, 1970).
[16] Y. Harano and M. Kinoshita, Biophys. J. **89**, 2701 (2005).
[17] Y. Snir and R. D. Kamien, Science **307**, 1067 (2005); Y. Snir and R. D. Kamien, Phys. Rev. E **75**, 051114 (2007).
[18] M. Kinoshita, Chem. Eng. Sci. **61**, 2150 (2006).
[19] H. Hansen-Goos, R. Roth, K. Mecke, and S. Dietrich, Phys. Rev. Lett. **99**, 128101 (2007).
[20] J. R. Banavar, O. Gonzalez, J. H. Maddocks, and A. Maritan, J. Stat. Phys. **110**, 35 (2003).
[21] O. Gonzalez and J. H. Maddocks, Proc. Natl. Acad. Sci. U.S.A. **96**, 4769 (1999).
[22] Note that the definitions of the accessible surface area Σ_{acc} and the contact surface Σ_{cont} are slightly different. The present

work, as well as Ref. [9], deal with the contact surface. Nevertheless, numerical simulations considering both definitions yield the same results.

[23] D. A. Klain and G. C. Rota, *Introduction to Geometric Probability*, edited by L. Lincee (Cambridge University Press, Cambridge, 1999).

[24] H. S. M. Coxeter, *Introduction to Geometry* (Wiley, New York, 1989); R. Kamien, Rev. Mod. Phys. **74**, 953 (2002).

[25] S. Przybyl and P. Pieranski, Eur. Phys. J. E **4**, 445 (2001).

[26] A. Stasiak and J. H. Maddocks, Nature (London) **406**, 251 (2000).

[27] A. Trovato, H. Gerlach, and J.H. Maddocks (in preparation).

[28] M. Tinkham, *Introduction to Superconductivity* (McGraw-Hill, New York, 1996).

[29] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, Science **220**, 671 (1983).