




Marginally compact phase and ordered ground states in a model polymer with side spheresTatjana Škrbić ^{*}*Department of Physics and Institute for Fundamental Science, University of Oregon, Eugene, Oregon 97403, USA
and Dipartimento di Scienze Molecolari e Nanosistemi, Università Ca' Foscari di Venezia, Campus Scientifico, Edificio Alfa,
Via Torino 155, 30170 Venezia Mestre, Italy*Trinh Xuan Hoang [†]*Institute of Physics, Vietnam Academy of Science and Technology, 10 Dao Tan, Ba Dinh, Hanoi 11108, Vietnam*Achille Giacometti [‡]*Dipartimento di Scienze Molecolari e Nanosistemi, Università Ca' Foscari di Venezia, Campus Scientifico, Edificio Alfa, Via Torino 155,
30170 Venezia Mestre, Italy
and European Center for Living Technologies, Ca' Bottacin, Dorsoduro 3911, Calle Crosera, 30123 Venezia, Italy*Amos Maritan [§]*Dipartimento di Fisica e Astronomia, Università di Padova and INFN, Via Marzolo 8, 35131 Padova, Italy*Jayanth R. Banavar ^{||}*Department of Physics and Institute for Fundamental Science, University of Oregon, Eugene, Oregon 97403, USA*

(Received 4 May 2021; accepted 14 June 2021; published 26 July 2021)

We present the results of a quantitative study of the phase behavior of a model polymer chain with side spheres using two independent computer simulation techniques. We find that the mere addition of side spheres results in key modifications of standard polymer behavior. One obtains a marginally compact phase at low temperatures; the structures in this phase are reduced in dimensionality and are ordered, they include strands assembled into sheets and a variety of helices, and at least one of the transitions on lowering the temperature to access these ordered states is found to be first order. Our model serves to partially bridge conventional polymer phases with biomolecular phases.

DOI: [10.1103/PhysRevE.104.L012501](https://doi.org/10.1103/PhysRevE.104.L012501)

Our principal goal is a careful quantitative computational analysis of a simple chain model of tethered spheres with side spheres attached to the main chain spheres. Our motivation for attaching side spheres comes from proteins, the amazing molecular machines of life. The side spheres merely restrict the conformational space that the chain can explore. In spite of this somewhat innocuous role, we demonstrate that there are significant and surprising effects on the phase diagram and the nature of the ground-state conformations. We note that our paper is not about proteins but rather about a simple standard model of polymer physics.

Some small globular proteins fold in an all-or-nothing manner under physiological conditions. The transition from the unfolded to a folded state is akin to a first-order phase transition, albeit for a finite-size system [1]. The folding is rapid as well as reproducible in that a given protein folds into the same native state structure upon unfolding and refolding.

The native state structures of proteins are modular and are built up of building blocks of helices and zigzag strands assembled into almost planar sheets [2]. The simplest model of a homopolymer (made up of just one type of monomer) chain of tethered spheres does *not* account for many of these observations. For a generic range of attractive interaction, mimicking the mediating influence of the water, we obtain a continuous transition from a high-temperature expanded coil phase to a globular phase at the so-called θ temperature [3] via an ideal coil state. On lowering the temperature further we find indications of a second transition in our computer simulations into a more compact unstructured globular phase. There are no building blocks of helices or sheets, nor is there any evidence of crystalline ordering, e.g., tethers passing through a face-centered-cubic (fcc) crystalline arrangement [4]. There is no rapid, let alone reproducible, folding into a specific ground state. Motivated by proteins, our principal goal here is to quantitatively study the effects of modifying the standard homopolymer chain model in the simplest possible way by adding side spheres to the main chain spheres.

Proteins are made up of 20 types of naturally occurring amino acids, each with a distinctive side chain. Incorporating this feature would make the problem inherently complex because now the details of the distinct monomer types along

^{*}tskrbic@uoregon.edu[†]hoang@iop.vast.ac.vn[‡]achille.giacometti@unive.it[§]amos.maritan@unipd.it^{||}Corresponding author: banavar@uoregon.edu

with their interactions become a necessary part of the story. Here instead we build a bridge between the homopolymer and protein behaviors by incorporating a simple attribute of proteins into the polymer model. We seek to study the impact of side spheres (attached to all but the two end main spheres) sticking out in the negative normal direction in the local Frenet coordinate frame [5]. We find a marginally compact phase at low temperatures between the compact globular phase and a less well-packed restricted coil phase upon varying the side sphere size. The marginally compact phase is the analog (on the side sphere size axis, not the temperature axis), for chain molecules, of the liquid crystal phase [4], which is a sensitive phase of matter with applications in displays, mood rings, and sensors. The vicinity of the marginally compact phase to other phases confers sensitivity to a chain poised in it. The structures are ordered and the transition to the marginally compact phase upon lowering the temperature is first order. This is not unexpected because the symmetry of the unstructured coil phase is distinct from that of the ordered marginally compact phase [4]. We find ground-state structures including a planar sheet made up of zigzag strands along with a variety of helices and dual helices. We find, generally, that the limitations imposed by our study of moderate length chains as well as difficulties associated with lack of equilibration are greatly reduced as the side sphere size increases. Even though the simple addition of side spheres already introduces many proteinlike features, we find clear differences between the marginally compact structures and protein native state structures, underscoring that other essential features yet need to be incorporated into simple chain models to describe the amazing molecular machines of life.

The tangent homopolymer chain comprises N spherical beads of diameter σ (set equal to 1) tethered into a chain. Consecutive beads are kept at a fixed distance $b = \sigma$ and nonconsecutive beads are not allowed to overlap, resulting in excluded volume [6]. Solvent effects are incorporated by including a short-range square-well attraction between beads separated in sequence by at least three. The range of attraction of the square well, R_c/σ , measured in units of the sphere diameter, is a free parameter.

The phase diagram as a function of the reduced temperature (measured in units of the attractive potential, which can also be set equal to 1) has been studied over the years by many different groups using various methods [7]. It has a high-temperature swollen phase also called the coil phase, where the chain is in a relatively open stretched conformation, which is dominated by entropy. We find two continuous transitions upon lowering the temperature. The first is a transition to a collapsed globule phase at the so-called θ temperature and the second is to an unstructured compact globule phase. At low temperatures, the minimization of the energy or, equivalently, the maximization of the number of attractive contacts is the dominating factor resulting in the compact globule phase (for a chain of moderate length approximately equal to 100) with many degenerate conformations having no discernible order and a high number of attractive contacts. There are two unresolved issues of importance: First, what might happen for a much larger system, e.g., might a highly compact crystalline arrangement (such as a Hamiltonian walk on a fcc lattice) be the true ground state, and second, could the tether-

ing of the spheres thwart equilibration, i.e., is one observing glassy behavior characterized by low-lying metastable minima? Indeed, there is some evidence in careful simulations [7] suggesting some murkiness in reliably identifying the ground state.

We generalize the simple polymer model by mounting on each main chain sphere (except the first and last) tangent side spheres in the negative normal direction. The side spheres provide steric constraints on both main chain spheres and other side spheres and reduce the conformational space. In our analysis, the side spheres do not interact except that they are hard and are therefore not allowed to overlap with each other or with the main chain spheres. A partial penetrability would result in an effective smaller side sphere size but would not change the results qualitatively. The side sphere diameter, measured in units of the main chain sphere diameter σ_{SC}/σ , is the second and final parameter of the model. We do not impose any attraction between beads separated in sequence by two in order to not suppress the bond bending angles θ artificially. The conformation of a chain of constant bond length is fully specified by two angle variables at each location [8], the bond bending angle θ and the torsion angle μ , the angle between successive binormals of a Frenet local coordinate system.

There have been previous insightful studies of heteropolymer lattice protein models with side chains (in contrast, our work here deals with an off-lattice homopolymer), which have noted enhanced cooperative folding [9] due to the denser packing in the interior of the structure. These studies did not fix the side chain location in the Frenet coordinate system of the main chain residues like we do here. Our own earlier studies [10] of homopolymers with side spheres mostly dealt with overlapping main chain spheres (the overlap was overtly introduced to replace the spurious spherical symmetry associated with individual main chain spheres with an axial symmetry befitting a chain). Furthermore, in the earlier work, the attractive potential was extended to pairs of main chain spheres separated by two along the chain (unlike three in our studies here), which promoted the occurrence of artificially small bending angles θ . Finally, here we use the machinery for the quantitative characterization of the variety of helical geometries observed in the simulations using the (θ, μ) variables.

Our Monte Carlo simulations were carried out using two complementary methods: microcanonical Wang-Landau (WL) simulations [11] and replica exchange (RE) (or parallel tempering) canonical simulations [12]. The WL simulations entail the filling of consecutive energy histograms to derive the density of states $g(E)$. The acceptance probability is chosen to promote moves exploring less populated energy states seeking to flatten energy histograms over the course of the simulations, thus finding lower-energy states. In all cases, 28–30 levels of iterations were carried out with a flatness criterion in each iteration of at least 80%, ensuring convergence of the $g(E)$ allowing for thermodynamic quantities to be calculated. The RE approach entails running canonical simulations in parallel at M different temperatures T_i , $i = 1, 2, \dots, M$. Each simulation can be thought of as a replica, or a system copy in thermal equilibrium. The key advantage is the possibility of swapping replicas at different temperatures without affecting the equilibrium condition at each temperature, permitting rapid equilibration even when there

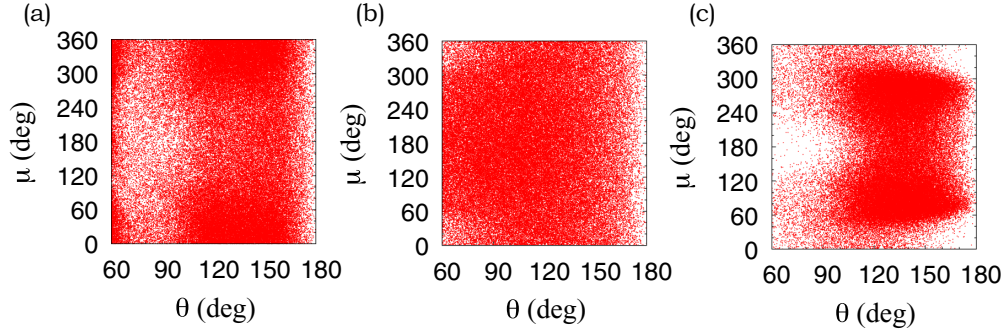


FIG. 1. Cross plot of θ and μ showing local structure in three unstructured phases of a standard polymer chain, where θ is the bond-bending angle and μ is the torsion angle. (a) Low-temperature compact globular phase in the absence of side spheres for a chain of length $N = 60$. (b) and (c) Infinite-temperature coil phase for $N = 20$ for a chain with (b) no side spheres and (c) large side spheres of size $\sigma_{SC}/\sigma = 2.8$. All three panels show 100 000 points.

is a rugged landscape. Both methods employ standard local moves including crankshaft, reptation, and an end point along with the nonlocal pivot move. The results using both methods are completely consistent with each other and each method is useful to benchmark the other.

We now proceed to discuss our results. Figure 1 shows the distinct characteristics of chain conformations in the compact globular (zero-temperature possibly nonequilibrium phase maximizing attractive interactions while respecting steric constraints) phase, the coil (infinite-temperature, high-entropy) phase, and a restricted coil phase in (θ, μ) cross plots. Figures 1(a) and 1(b) are for chains with no side spheres, whereas

the chain in Fig. 1(c) has large side spheres with $\sigma_{SC}/\sigma = 2.8$. There is a tendency, in the ground state, for a chain to bend maximally (favor the smallest sterically permitted value of θ) and avoid planar conformations (by exhibiting a local minimum at $\mu = 180^\circ$ in contrast to a local maximum in the coil phase).

We monitored the specific heat as a function of temperature for three different chain lengths in order to understand the nature of all the phase transition(s). As an example, Fig. 2(a) depicts the specific heat versus temperature plot for $R_c/\sigma = 1.6$ in the absence of side spheres, with the inset showing the canonical energy probability distributions at three

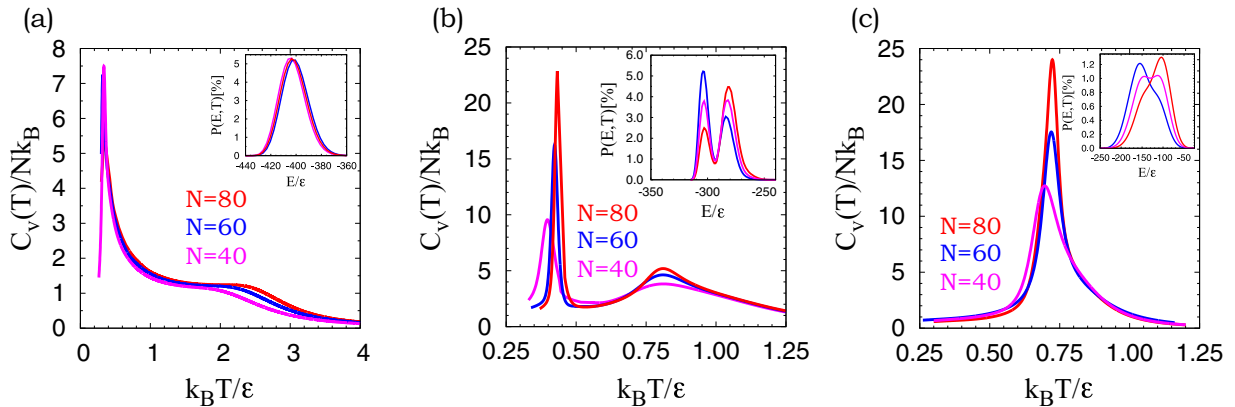


FIG. 2. Specific heat per bead (C_v/Nk_B) as a function of reduced temperature ($T^* = k_B T/\epsilon$) for chains of length $N = 40, 60$, and 80 . (a) Specific heat curve for $R_c/\sigma = 1.6$ and $\sigma_{SC}/\sigma = 0$; the ground state is an unstructured compact globule. The specific heat curve suggests the existence of two phase transitions. The transition at the higher temperature, between the coil and the collapsed globule (the θ transition), is signaled by the shoulder. The more pronounced lower-temperature peak is suggestive of a second continuous transition between two unstructured phases, the collapsed globule and the compact globule. (b) Specific heat curve for $R_c/\sigma = 1.6$ and $\sigma_{SC}/\sigma = 1.0$; the ground state is the (7,45) precessing helix [shown in Fig. 3(c)]. (c) Specific heat curve for $R_c/\sigma = 1.6$ and $\sigma_{SC}/\sigma = 1.5$; the ground state here is the (1,5) dual helix [shown in Fig. 3(d)]. In all cases, the inset shows the canonical probability distribution of the energy in the vicinity of the transition to the ground state (the lower transition temperature when there are two transitions) for the longest chain studied ($N = 80$) at three temperatures: the transition temperature corresponding to the peak in the specific heat (purple curve), a temperature 1% below the transition temperature (blue curve), and a temperature 1% above the transition temperature (red curve). The numbers of attractive contacts in our simulations are 443, 316, and 259 for $\sigma_{SC}/\sigma = 0, 1$, and 1.5 , respectively. The key message is that the low-temperature transition in the absence of side spheres appears to be continuous, whereas the addition of large enough side spheres results in a structured helical ground state and a first-order transition to it upon lowering the temperature. The low-temperature first-order transition weakens upon increasing the side sphere size [compare (c) versus (b)] because of the decrease in conformational entropy in the restricted coil phase.

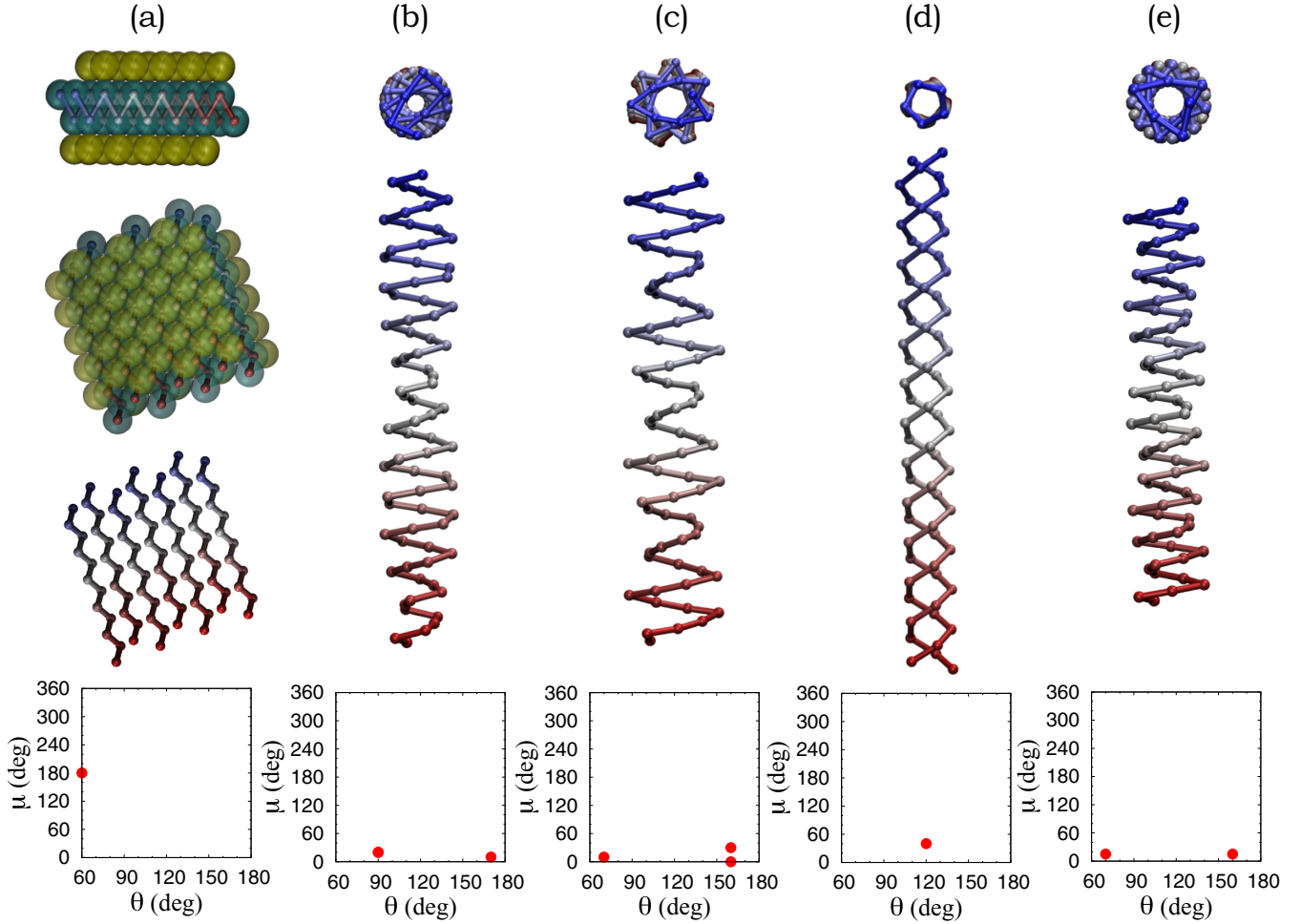


FIG. 3. Sketches of several ordered structures in the marginally compact phase of a tangent chain for different attraction ranges and side chain diameters. The structures are mathematical idealizations of those observed in the simulations having the same number of attractive contacts. Small variations around the idealized structures are permitted without any change in energy (the number of contacts). The helical structures shown are robust and reproducible over a range of chain lengths with two different simulation techniques and are likely ground states in the thermodynamic limit. The sheet structure is not observed in our simulations of a short chain but is mathematically constructed to maximize the number of contacts for the parameters shown. The side spheres are explicitly shown only for the sheet. For the helices, the side spheres stick out tangentially. In all cases, there are no steric clashes. The ground state for (a) $R_c/\sigma = 1.2$ and $\sigma_{SC}/\sigma = 1.0$ is a (1,2) sheet with six contacts per interior sphere and a repeat of (θ, μ) values of $(60^\circ, 180^\circ)$; for (b) $R_c/\sigma = 1.4$ and $\sigma_{SC}/\sigma = 1.0$, a (7,39) helix with $(5 + 5 + 8 + \dots)$ contacts per main chain sphere (degenerate with the sheet and a dual helix) corresponding to repeat (θ, μ) values of $(90^\circ, 20^\circ)$, $(90^\circ, 20^\circ)$, and $(170^\circ, 10^\circ)$; for (c) $R_c/\sigma = 1.6$ and $\sigma_{SC}/\sigma = 1.0$, a (7,45) helix with $(6 + 10 + 10 + \dots)$ contacts and a repeat of (θ, μ) values of $(60^\circ, 180^\circ)$, $(160^\circ, 30^\circ)$, and $(160^\circ, 0^\circ)$; for (d) $R_c/\sigma = 1.6$ and $\sigma_{SC}/\sigma = 1.5$, a dual helix with six contacts per bulk main chain sphere and a repeat of (θ, μ) values of $(120^\circ, 40^\circ)$; and for (e) $R_c/\sigma = 1.8$ and $\sigma_{SC}/\sigma = 1.0$, a (7,38) helix with $(12 + 16 + \dots)$ contacts and a repeat of (θ, μ) values of $(70^\circ, 15^\circ)$ and $(160^\circ, 15^\circ)$. The notation used is (number of approximate full turns per period of the helix, number of main chain spheres per period). The sheet picture does not depict the turns connecting the zigzag strands and the double helix picture does not show the single turn.

temperatures in the vicinity of the lower temperature continuous transition. Our studies become more reliable and the transition becomes first order on introducing large enough side spheres. The temperature dependence of the specific heat is shown for $\sigma_{SC}/\sigma = 1$ [Fig. 2(b)] and $\sigma_{SC}/\sigma = 1.5$ [Fig. 2(c)]. The ground state for the former case is a precessing helix, whereas the ground state for the latter case is a structured dual helix [see Figs. 3(c) and 3(d)]. For the chain lengths we have studied and likely in the infinite chain length limit, the presence of side spheres greater than or equal to $\sigma_{SC}/\sigma = 0.9$ does not allow for folding of the helix or the dual helix onto itself to create a hairpin structure and avail

of more attractive main sphere interactions. The scaling of the specific heat peak with system size and the two-peaked structure in the energy probability distribution in the vicinity of the transition to the ground state are both signatures of a first-order transition, albeit for a finite-size chain. Note that the first-order transition is weakened upon increasing the side sphere size from 1 to 1.5 (Fig. 2 insets).

Figure 4 shows the phase diagram in the temperature-side sphere size plane for a fixed range of attractive interaction $R_c/\sigma = 1.6$. The region of the phase diagram for side sphere size less than 0.9 seems robust (independent of Monte Carlo technique and different runs), yet is possibly unreliable be-

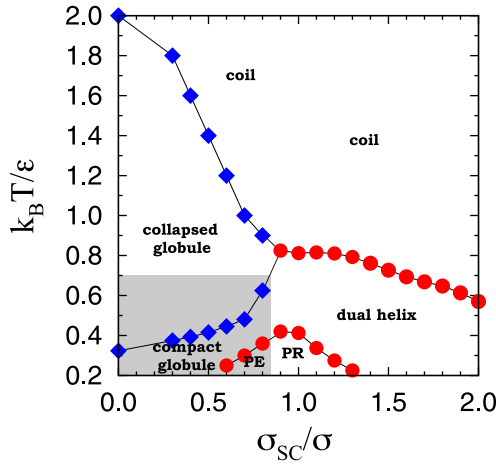


FIG. 4. Phase diagram in the reduced temperature: side chain diameter plane for the tangent polymer chain of length $N = 60$. The blue diamonds indicate continuous transitions whereas the red circles are first-order transitions. Both parallel tempering and Wang-Landau simulations yield consistent results. The shaded region indicates a part of the phase diagram where the relatively modest chain length and potential equilibration issues make the results possibly unreliable. In the unshaded region, the error estimates are smaller than the size of the points. The coil is the canonical high-temperature phase. For small side spheres, there are two globule phases, a collapsed globule at intermediate temperatures (with the familiar θ transition between the coil and the collapsed globule) as well as a distinct low-temperature phase of a compact globule (with a second continuous transition between the compact and collapsed globule phases). The penetrated helix phase (PE) has, as its ground state, a single or a dual helix with a large enough radius to allow penetration of a rod (an essentially straight chain segment) within it (see Fig. 5), the precessing helix phase (PR) has a helical ground state with a period of 45 main chain spheres yielding a rotation of approximately seven turns [see Fig. 3(c)], and the dual helix is comprised of two symmetric helices connected by a turn [see Fig. 3(d) (turn not shown)].

cause of the glassy behavior and finite size effects (see the shaded region in Fig. 4). One observes both continuous (blue diamonds) and first-order (red circles) transitions. For no side spheres, one observes two continuous phase transitions. Around a value of $\sigma_{SC}/\sigma = 0.6$, there is a distinct feature in the ground state with the unstructured compact globule giving way to an ordered ground state. The higher transition is still a continuous coil to unstructured globule transition, whereas the lower transition between the globule and the structured ground state is first order because of the distinct symmetries of the two phases. This situation persists until $\sigma_{SC}/\sigma = 0.8$, where there is a steep increase in the coil-globule transition temperature along with changes in the nature of the structured ground state. Starting at $\sigma_{SC}/\sigma = 0.9$, the results of the simulations become reliable. The globule is replaced by a structured dual helix at intermediate temperatures, and the ground state becomes a precessing helix [see Fig. 3(c)]. The lower transition temperature diminishes to zero on increasing the side sphere size and there is just a single coil-double helix weakening first-order transition for σ_{SC}/σ greater than 1.4 (Fig. 2). The transition temperature continues to decrease as the side sphere size increases; attractive contacts are harder to make in any sustained manner. Ultimately, for large enough

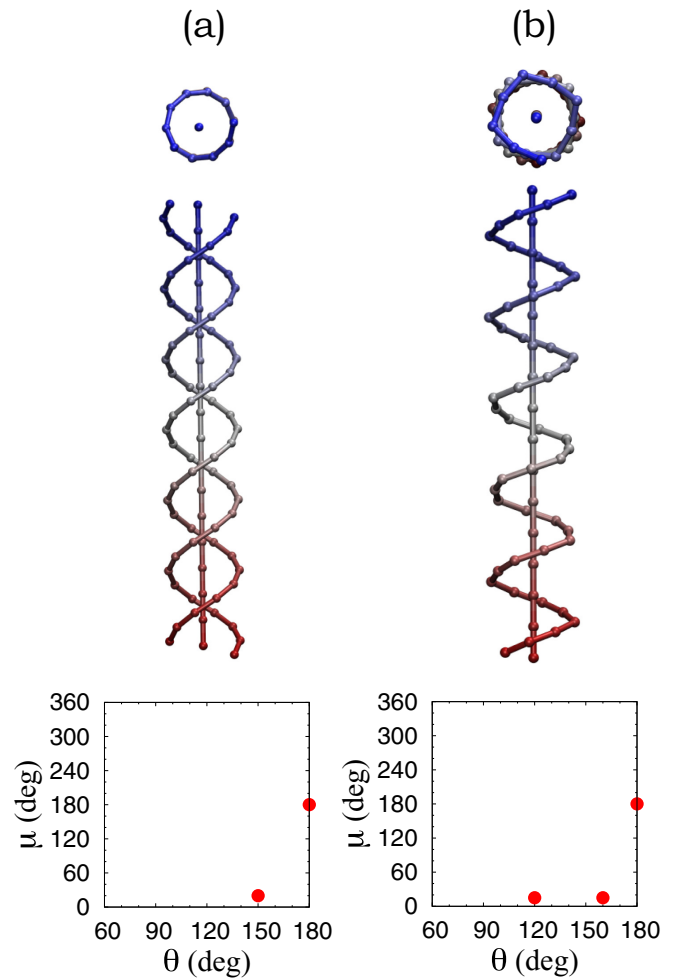


FIG. 5. Mathematical idealizations of two structures, penetrated helical structures, corresponding to a single rod (an essentially straight chain segment) within (a) a dual helix and (b) a single helix, for $R_c/\sigma = 1.6$ and (a) $\sigma_{SC}/\sigma = 0.7$ and (b) $\sigma_{SC}/\sigma = 0.8$. The point corresponding to both θ and μ equal to 180° denotes the rod, whereas the small- μ values correspond to the helical jackets. These structures are seen consistently in our computer simulations, they can be sustained in the long chain limit, and they are probably ground-state structures.

side sphere size, even at low temperatures, one obtains a coil phase with a greatly reduced phase space because of the steric constraints imposed by the large side spheres [Fig. 1(c)].

Figures 3 and 5 show some of the ordered ground-state structures in the marginally compact phase along with their (θ, μ) plots. Even the small sampling of structures shown here present a beauty and richness not previously observed, to the best of our knowledge, in polymer models or for that matter in standard biological systems. The idealized structures (whose topology of attractive contacts is faithfully realized in computer simulations, albeit with some variations) include a planar sheet of zigzag strands, a variety of distinct precessing helices [with repeating (θ, μ) angles], a dual helix, a helix, and a dual helix with a straight chain segment (rod) penetrating it. The precessing helices are in fact two or three uniform helices intercalated between each other. Our quanti-

tative analysis is facilitated by the (θ, μ) description of the chain conformations.

We find that one obtains a compact globule to coil transition in two distinct ways: First, for the standard case of no side spheres, the globule to coil continuous transition occurs on increasing the temperature; second, at zero temperature, on increasing the size of the side spheres, the globule phase switches to the coil phase through a sequence of first-order transitions passing through structured ground states that we refer to as marginally compact states. They are marginally compact because they are more compact than the coil phase structures but less so than the globular compact phase structures.

A tangent chain of spheres has problems associated with spurious symmetries. A spherical monomer looks the same when viewed from any direction. However, when strung along a chain, then this spherical symmetry becomes spurious. A chain has a tangent direction at each location and thus the isotropy is lost. Attaching side spheres in the negative normal direction is one way to overtly move away from this fake isotropy of the constituent spheres. One might imagine the compact globule and the somewhat open entropically favored coil states as analogs of the crystal and liquid phases. The marginally compact phase then is the analog of the liquid crystal phase [4]. First, the marginally compact structures lie in the vicinity of transitions to other phases and, second, they arise on breaking the spurious symmetry in the model associated with spherical monomers constituting a uniaxial chain. Liquid crystals are very sensitive to the right types of perturbations and one would expect that to hold as well for the structures in the marginally compact phase. Further, the marginally compact conformation has a Goldilocks compactness between that of a dense globule (in which the conformation seeks to maximize the number of attractive contacts somewhat akin to a sphere surrounded by a dozen spheres in a face-centered-cubic lattice) and a coil conformation (in which a few chance contacts are made but the conformation is a typical self-avoiding walk). Finally, the ordered marginally compact structures are all reduced-dimensional structures (topologically one-dimensional helices or two-dimensional sheets), just as liquid crystals are not ordered in the same manner in all three directions. Striking examples of marginally compact conformations presented in our paper are dimensionally reduced conformations such as a two-dimensional sheet and one-dimensional helices and dual helices. Interestingly, these conformations tend to be ordered promoting a first-order phase transition between a disordered higher-temperature phase and this ordered marginally compact phase.

We have presented a quantitative study of the phase behavior of a tangent polymer chain with side spheres. Our results

demonstrate that the mere addition of side spheres results in key modifications of the model behavior. As is observed in our simulations, proteins [2] fold in an all-or-nothing fashion akin to the behavior at a first-order transition; their structures are made up of reduced-dimensional building blocks including topologically one-dimensional helices and two-dimensional sheets. Proteins are stable yet sensitive and are able to carry out a dizzying array of functions. In this regard, they are reminiscent of liquid crystals. Despite the encouraging results reported here, the simple model we have studied is still missing some essential ingredients for faithfully describing proteins. The symptoms of these missing elements are that the structures we obtain are not those found in proteins. Unlike the rich variety of helices found here, the α helix is characterized by (θ, μ) values around $(91.3 \pm 2.2^\circ, 49.7 \pm 3.9^\circ)$ [13], which we do not find in our studies here. Also, the sheet structure in proteins is qualitatively different from that found in our model; the key difference is that adjacent strands in proteins are in phase with each other unlike the out-of-phase packing, which maximizes the number of pairwise contacts here. In addition, we do not observe coexistence of helices and sheets in our model here, which is a crucial ingredient for obtaining the diversity of protein folds. A powerful hint regarding the missing ingredient comes from recent work [13] that used a tubelike description to derive the building blocks of protein structures from first principles with no adjustable parameters. We look forward to studies incorporating a tube description along with side chains to further understand proteins in a simplified manner. More generally, our work here ought to be of relevance for understanding the phase behavior of polymers [6].

We are indebted to Pete von Hippel for his warm hospitality. This project received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 894784. The contents reflect only the authors' view and not the views of the European Commission. Support from the University of Oregon (through a Knight Chair to J.R.B.), Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant No. 103.01-2019.363 (T.X.H.), University of Padova through "Excellence Project 2018" of the Cariparo foundation (A.M.), MIUR PRIN-COFIN2017 Soft Adaptive Networks Grant No. 2017Z55KCW, and COST action CA17139 (A.G.) is gratefully acknowledged. The computer calculations were performed on the Talapas cluster at the University of Oregon. T.Š. carried out the calculations with significant help from T.X.H. and under the guidance of J.R.B. J.R.B. wrote the paper. All authors helped interpret the results and reviewed the manuscript.

The authors declare that they have no conflict of interest.

[1] C. B. Anfinsen and H. A. Scheraga, in *Experimental and Theoretical Aspects of Protein Folding*, edited by C. B. Anfinsen, J. T. Edsall, and F. M. Richards, *Advances in Protein Chemistry* (Academic Press, 1975), Vol. 29, pp. 205–300; R. Zwanzig, Two-state models of protein folding kinetics, *Proc. Natl. Acad. Sci. USA* **94**, 148 (1997); C. Maffi, M.

Baiesi, L. Casetti, F. Piazza, and P. De Los Rios, First-order coil-globule transition driven by vibrational entropy, *Nat. Commun.* **3**, 1065 (2012); A. Badasyan, M. Valant, J. Grdadolnik, and V. N. Uversky, The finite size effects and two-state paradigm of protein folding, *Int. J. Mol. Sci.* **22**, 2184 (2021).

- [2] A. M. Lesk, *Introduction to Protein Science: Architecture, Function and Genomics* (Oxford University Press, Oxford, 2004); I. Bahar, R. L. Jernigan, and K. A. Dill, *Protein Actions: Principles and Modeling* (Garland Science, New York, 2017).
- [3] P.-G. de Gennes, Collapse of a polymer chain in poor solvents, *J. Phys. (Paris) Lett.* **36**, 55 (1975); I. M. Lifshitz, Y. A. Grosberg, and A. R. Khokhlov, Some problems of the statistical physics of polymer chains with volume interaction, *Rev. Mod. Phys.* **50**, 683 (1978).
- [4] P. Chaikin and T. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 2000).
- [5] R. D. Kamien, The geometry of soft materials: A primer, *Rev. Mod. Phys.* **74**, 953 (2002).
- [6] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1979); A. R. Khokhlov, A. Y. Grosberg, and V. S. Pande, *Statistical Physics of Macromolecules*, AIP Series in Polymers and Complex Materials (American Institute of Physics, New York, 1994); M. Rubinstein and R. H. Colby, *Polymer Physics* (Oxford University Press, Oxford, 2003).
- [7] Y. Zhou, M. Karplus, J. M. Wichert, and C. K. Hall, Equilibrium thermodynamics of homopolymers and clusters: Molecular dynamics and Monte Carlo simulations of systems with square-well interactions, *J. Chem. Phys.* **107**, 10691 (1997); M. P. Taylor, W. Paul, and K. Binder, Phase transitions of a single polymer chain: A Wang-Landau simulation study, *ibid.* **131**, 114907 (2009); All-or-none proteinlike folding transition of a flexible homopolymer chain, *Phys. Rev. E* **79**, 050801(R) (2009); Applications of the Wang-Landau algorithm to phase transitions of a single polymer chain, *Polym. Sci. Ser. C* **55**, 23 (2013); On the polymer physics origins of protein folding thermodynamics, *J. Chem. Phys.* **145**, 174903 (2016).
- [8] T. Škrbić, A. Maritan, A. Giacometti, and J. R. Banavar, Local sequence-structure relationships in proteins, *Protein Sci.* **30**, 818 (2021).
- [9] A. Kolinski and J. Skolnick, Discretized model of proteins. I. Monte Carlo study of cooperativity inhomopolypeptides, *J. Chem. Phys.* **97**, 9412 (1992); M. H. Hao and H. A. Scheraga, Monte Carlo simulation of a first-order transition for protein folding, *J. Phys. Chem.* **98**, 4940 (1994); Statistical thermodynamics of protein folding: Sequence dependence, **98**, 9882 (1994); A. Kolinski, W. Galazka, and J. Skolnick, On the origin of the cooperativity of protein folding: Implications from model simulations, *Proteins* **26**, 271 (1996); D. K. Klimov and D. Thirumalai, Cooperativity in protein folding: From lattice models with side chains to real proteins, *Fold. Des.* **3**, 127 (1998); V. V. Vasilevskaya, P. G. Khalatur, and A. R. Khokhlov, Conformational polymorphism of amphiphilic polymers in a poor solvent, *Macromolecules* **36**, 10103 (2003).
- [10] J. R. Banavar, M. Cieplak, T. X. Hoang, and A. Maritan, First-principles design of nanomachines, *Proc. Natl. Acad. Sci. USA* **106**, 6900 (2009); T. Škrbić, A. Badasyan, T. X. Hoang, R. Podgornik, and A. Giacometti, From polymers to proteins: The effect of side chains and broken symmetry on the formation of secondary structures within a Wang-Landau approach, *Soft Matter* **12**, 4783 (2016); T. Škrbić, T. X. Hoang, and A. Giacometti, Effective stiffness and formation of secondary structures in a protein-like model, *J. Chem. Phys.* **145**, 084904 (2016); T. Škrbić, T. X. Hoang, A. Maritan, J. R. Banavar, and A. Giacometti, The elixir phase of chain molecules, *Proteins* **87**, 176 (2019); Local symmetry determines the phases of linear chains: A simple model for the self-assembly of peptides, *Soft Matter* **15**, 5596 (2019); T. Škrbić, J. R. Banavar, and A. Giacometti, Chain stiffness bridges conventional polymer and bio-molecular phases, *J. Chem. Phys.* **151**, 174901 (2019).
- [11] F. Wang and D. P. Landau, Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States, *Phys. Rev. Lett.* **86**, 2050 (2001).
- [12] R. H. Swendsen and J. S. Wang, Replica Monte Carlo Simulation of Spin-Glasses, *Phys. Rev. Lett.* **57**, 2607 (1986).
- [13] T. Škrbić, A. Maritan, A. Giacometti, G. D. Rose, and J. R. Banavar, Building blocks of protein structures: Physics meets biology, *Phys. Rev. E* **104**, 014402 (2021).