

ARTICLE

Received 16 May 2012 | Accepted 30 Jul 2012 | Published 18 Sep 2012

DOI: 10.1038/ncomms2055

First-order coil-globule transition driven by vibrational entropy

Carlo Maffi¹, Marco Baiesi², Lapo Casetti³, Francesco Piazza⁴ & Paolo De Los Rios¹

By shifting the balance between conformational entropy and internal energy, polymers modify their shape under external stimuli, such as changes in temperature. Prominent among such transformations is the coil-globule transition, whereby a polymer can switch from an entropy-dominated coil conformation to a globular one, governed by energy. The nature of the coil-globule transition has remained elusive, with evidence for both continuous and discontinuous transitions, with the two-state behaviour of proteins as an instance of the latter. Theoretical models mostly predict second-order transitions. Here we introduce a model that takes into consideration hitherto neglected features common to any polymer. We show that a first-order phase transition smoothly appears as a function of the model parameters. Our results can relieve part of the conflicts between theory and experiments in the field of protein folding, in the wake of recent studies tracing back the remarkable properties of proteins to basic polymer physics.

¹ Laboratory of Statistical Biophysics, SB ITP, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015, Lausanne, Switzerland. ² Dipartimento di Fisica e Astronomia, Università di Padova, and INFN, Sezione di Padova Via Marzolo 8, Padova 35131, Italy. ³ Dipartimento di Fisica e Astronomia and CSDC, Università di Firenze and Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Firenze, via G. Sansone 1, Sesto Fiorentino (FI) I-50019, Italy. ⁴ Department of Physics of Engineering, University of Orléans and Centre de Biophysique Moléculaire, CBM-CNRS UPR 4301, Rue Charles Sadron, Orléans 45071, France. *Correspondence and requests for materials should be addressed to P.D.L.R. (email: paolo.delosrios@epfl.ch).

The coil-globule transition (CGT)¹ is one of the most important phenomena exhibited by polymeric systems; polymers collapse into compact globules in a bad solvent, that is, a solvent that is unable to screen the mutual attraction between monomers, whereas they expand to coil-like conformations in a good solvent, where the only relevant interaction is the repulsive volume exclusion between monomers². The CGT lies behind well-established technologies based on responsive polymers³, as well as emerging applications to drug delivery⁴ and microfluidic devices and sensors⁵, and it sets the logical framework to understand protein folding⁶, possibly the most important natural occurrence of polymer collapse. Because of such widespread and crucial experimental and technological relevance, the CGT has been under careful theoretical scrutiny for the past three decades. In fact, there is a general agreement on the bad and good solvent sides of the transition. Instead, understanding the behaviour at the transition point represents a greater challenge. In his seminal work, de Gennes^{7,8} built upon the approach of Flory⁹ to describe the CGT as the transition of a tricritical magnetic system, inferring that it had to be a continuous transition. Although a large number of experiments have confirmed such prediction, there has also been evidence for discontinuous transitions^{10–12}, findings that have been to date difficult to reconcile with the theory. Moreover, and more fundamentally from a conceptual point of view, proteins are known to undergo in most cases a two-state kind of transitions, where the folded (native in the biochemistry language) and the unfolded states coexist at the transition. This is very much akin to what is expected for discontinuous, first-order phase transitions, that are indeed characterized by phase coexistence. Although the special (non-random) heteropolymeric features of proteins have a crucial role in protein thermodynamics¹³, it must be noted that no simple protein model with designed sequence exhibits full-fledged coexistence at the transition. It becomes thus necessary to understand how the continuous transition predicted by de Gennes^{7,8} can be turned into a discontinuous one as shown by proteins^{6,14}, and whether it is possible to do so. Two paths have been previously proposed to that goal. In an early attempt^{15,16}, and more recently in a paper by Doniach *et al.*¹⁷, polymers were endowed with a stiffness penalizing bent conformations, managing to elicit a new phase transition taking place within the globular phase and discriminating between amorphous and ordered globules. Such phase transition turned out to be similar to a first-order liquid–solid transition. Yet, the usual CGT retained its continuous nature. Only for very stiff polymers the collapse would take place directly toward ordered globules, thus changing the CGT order. A different strategy relies on the fine-tuning of the range of non-bonded interactions, so that within the globular phase, some rearrangements could lead from amorphous to ordered globules¹⁸. Again, this is a discontinuous transition that, for a carefully chosen interaction range, could pre-empt the CGT, which would otherwise remain continuous. Both these approaches, nonetheless, are unlikely to be relevant for synthetic polymers and proteins, that are very flexible¹⁹ and equipped with interactions of many different types and ranges, unlikely to be so fine-tuned, as it is required for switching the order of the CGT.

The approach we follow here is based instead on a critical analysis of traditional models of polymers. Taking into account hitherto neglected terms in the approximation of realistic Hamiltonians, we identify a single key parameter whose variation can result in a switch of the transition from continuous to discontinuous. Furthermore, we highlight how the subtle interplay between such parameter and the polymer length controls the emergence of the distinguishing features of first-order phase transitions, in particular phase coexistence.

Results

The self-interacting polymer model. Polymers are made of monomers strung together in long chains. Monomers can be simple, as in polyethylene (one carbon atom with two side

hydrogens), or complex, as amino acids in proteins. The covalent nature of the backbone interactions enforces almost fixed distances between consecutive monomers, whereas their mutual bending and torsional (dihedral) angles are restrained to discrete sets of values. For example, in polyethylene, the bond length is 0.153 nm, the angle between two bonds is about 112°, and three consecutive bonds are prevalently arranged in *trans*, *gauche₊* or *gauche₋* conformations²⁰. As mentioned above, the interactions between non-consecutive monomers are governed by two contrasting terms: short-range repulsive volume exclusion, which forbids any two monomers to occupy the same region of space, and medium-range attractive forces such as hydrogen-bonds, hydrophobic and van der Waals interactions. Because of the tight constraints on monomer–monomer distances and relative angles, polymers mostly populate only a discrete set of conformations. These features are captured by a lattice description, where consecutive monomers occupy nearest-neighbour lattice sites and two monomers cannot lie at the same site. As a consequence, it is a deeply rooted belief that lattices can be used to explore the properties of real polymers²¹. In particular, by including an energy term that favours pairs of non-consecutive monomers to occupy neighbouring sites (usually called contacts), lattice polymers exhibit a CGT belonging to the same universality class as predicted by de Gennes^{22,23}. The resulting model has been referred to as the Interacting Self-Avoiding Walk (ISAW). To improve upon the ISAW and identify the missing ingredients that may change the order of the CGT, it is necessary to review how the lattice embedding of real polymers is formally justified.

The discrete set of conformations accessible to real polymers and captured by the lattice description are the inherent structures²⁴ (the minima) of a potential function that depends on the full set of coordinates in continuous space, $V(\{\vec{r}_i\})$, where \vec{r}_i is the three-dimensional position vector of the i -th monomer. The full partition function of a polymer of N beads is thus

$$Z_N = \frac{1}{\Lambda^{3N}} \int \prod_{i=1}^N d\vec{r}_i e^{-\beta V(\{\vec{r}_i\})} \quad (1)$$

where Λ is the usual de Broglie thermal wavelength. If, in the spirit of a Gaussian approximation, only the dominant contribution of the set $\{\gamma\}$, corresponding to the conformations where $V(\{\vec{r}_i\})$ attains its minimal value, is taken into account, equation (1) becomes

$$Z_N \simeq \frac{1}{\Lambda^{3N}} \sum_{\{\gamma\}} e^{-\beta E(\gamma)} = \frac{1}{\Lambda^{3N}} \sum_{\{\gamma\}} e^{\beta \varepsilon n(\gamma)}. \quad (2)$$

The last term of equation (2) corresponds to the usual ISAW partition function; the energy is simply given by the number $n(\gamma)$ of non-bonded contacts of conformation γ , each contributing an energy $-\varepsilon$ (from here on, $\varepsilon=1$ and $k_B=1$). Here we focus on the next term in the Gaussian approximation of equation (1), which accounts for the fluctuations of the polymer around the minima²⁵

$$\begin{aligned} Z_N &= \frac{1}{\Lambda^{3N}} \sum_{\gamma} \left[e^{\beta n(\gamma)} \int \prod_{i=1}^N d(\Delta\vec{r}_i) e^{-\beta \frac{1}{2} \Delta\vec{r}^T \hat{H}(\gamma) \Delta\vec{r}} \right] \\ &= \frac{1}{\Lambda^{3N}} \sum_{\gamma} e^{\beta n(\gamma)} w(\gamma) \end{aligned} \quad (3)$$

$\hat{H}(\gamma)$ is the Hessian matrix of the potential evaluated at the minimum corresponding to conformation γ , and $\Delta\vec{r}_i$ is the displacement of the i -th monomer from its equilibrium position in γ . Up to a multiplicative factor independent of the configuration, $w(\gamma) = (\prod_{i=1}^{3N-6} \lambda_i)^{-1}$, (in the minima six eigenvalues vanish, corresponding to rigid

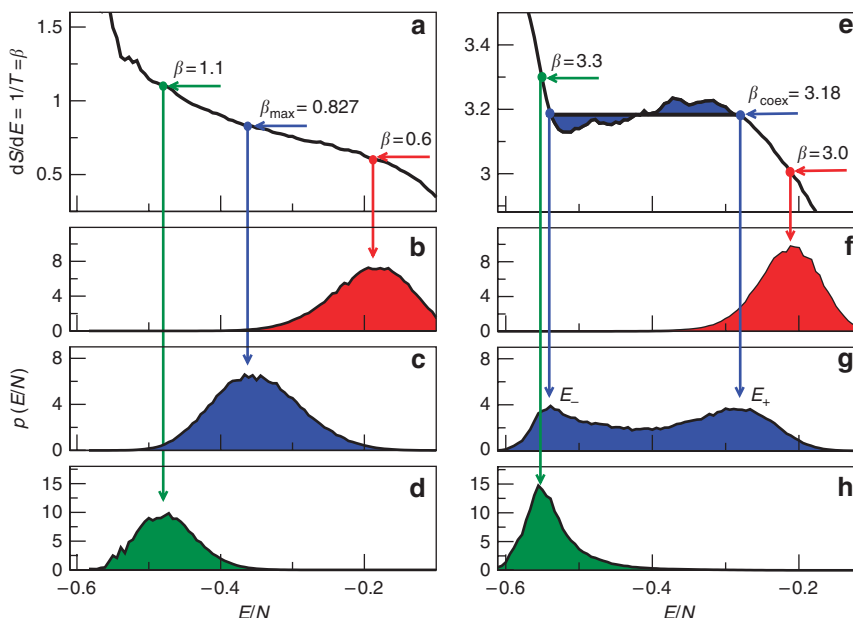


Figure 1 | Model thermodynamics. The microcanonical behaviour of polymers according to the ISAW (left column) and VISAW (right column) for $N=180$. **(a)** The derivative of the entropy with respect to the energy, proportional to the number of contacts, corresponds to the inverse temperature $\beta=1/T$. The slope is always negative, corresponding to a physical positive specific heat; the energy probability distribution is sampled at temperatures **(b)** higher than, **(c)** equal to and **(d)** lower than the temperature of the maximum of the specific heat T_{\max} , and in all cases it is found to be unimodal. **(e)** A positive slope corresponds to a microcanonical negative specific heat, which is not physical in the canonical ensemble and needs, therefore, the Maxwell construction to be avoided; the energy probability distribution at temperatures **(f)** higher than and **(h)** lower than T_{coex} are unimodal, whereas at a temperature equal to T_{coex} **(g)** the coexistence of two phases of low and high energy is marked by the presence of two distinct peaks at E_- and E_+ .

rotations and translations). As $-k_B T \ln[w(\gamma)]$ is the vibrational free energy associated with conformation γ and the internal energy is trivially $3Nk_B T/2$ for all conformations, $w(\gamma)$ is seen to embody the effects of vibrational entropy of conformation γ .

In our potential energy, the inherent structures correspond to conformations that can be embedded on the tetrahedral lattice. With this choice, bending angles between consecutive bonds have an equilibrium value of $\theta_0 = \cos^{-1}(-1/3) \simeq 109^\circ$, whereas four-body dihedral angles can take the three different equilibrium values $\psi_{t,g\pm} = -\pi, \pm(\pi/3)$. The minimum of the contact energy between non-consecutive monomers is set at a distance, which is precisely the lattice spacing l ($l=1$ henceforth). The complete potential then reads

$$\begin{aligned}
 V(\{\vec{r}_i\}) = & \sum_{i=1}^{N-1} \frac{1}{2} k_{bb} (|\vec{r}_{i+1} - \vec{r}_i| - 1)^2 \\
 & + \sum_{i=1}^{N-2} \frac{1}{2} k_\theta (\theta_{i+2,i} - \theta_0)^2 \\
 & + \sum_{i=1}^{N-3} \frac{1}{2} k_{\psi_{t,g}} (\psi_{i+3,i} - \psi_{t,g\pm})^2 \\
 & + \sum_{i,j} \Delta_{i,j} \left[-1 + \frac{1}{2} k (|\vec{r}_i - \vec{r}_j| - 1)^2 \right]
 \end{aligned} \tag{4}$$

where k_{bb} , k_θ , $k_{\psi_{t,g}}$ and k are the stiffness constants of the bonds between consecutive monomers, of the angles between consecutive triplets, of the dihedral angles between consecutive quadruplets and of the non-bonded nearest-neighbour forces, respectively. The parameter $k_{\psi_{t,g}}$ further distinguishes between trans and gauche conformations (see Methods). The contact matrix $\Delta_{i,j}$ is equal to 1 if non-consecutive ($j \neq i \pm 1$) monomers i and j are nearest neighbours,

and it is zero otherwise. The Hessian matrix of the energy function (equation (4)) is then obtained after expressing the bending and dihedral angles using Cartesian coordinates and expanding to second order in the displacements $\{\Delta \vec{r}_i\}$. We have then compared the behaviour of the traditional ISAW equation (2) model and of the present scheme, that we have termed, for brevity, Vibrating ISAW, henceforth VISAW, equation (3).

Thermodynamics. The thermodynamic behaviour of the ISAW model confirms that the associated CGT is continuous in three dimensions, as already found previously²². The microcanonical entropy is $S(E) = \ln \mathcal{N}(E)$, where $\mathcal{N}(E)$ is the number of conformations with a given energy E , and its derivative with respect to E yields the inverse temperature, $\beta = 1/T = dS(E)/dE$. Figure 1a–d show that β is a monotonously decreasing function of the energy, with no signs of phase coexistence. The picture changes dramatically in the VISAW model. Consistently with our canonical simulations (see Methods), we define the VISAW microcanonical entropy as $S(E) = \ln[\sum_\gamma w(\gamma)]$, where the sum runs over all the conformations γ characterized by an energy E . As for the ISAW, here E refers to the contact part only of the energy. Figure 1e–h reveal that the derivative of $S(E)$ is no longer monotonic when vibrations are included in the picture. The region where β grows with E , which would correspond to a negative microcanonical specific heat, is canonically forbidden and is one of the hallmarks of a discontinuous transition with coexistence between two different phases^{26–28}. The coexistence temperature can be extracted from the $\beta(E)$ curve using the customary Maxwell construction; this value $\beta_{\text{coex}} \simeq 3.18$ agrees well with the temperature of the (canonical) specific heat maximum $\beta_{\max} = 3.20(2)$.

The difference between Fig. 1a and e, in particular, points to the vibrational entropy as the culprit for the change in the order of the transition. As a matter of fact, in the context of solid–solid transformations, it had already been observed that the specific

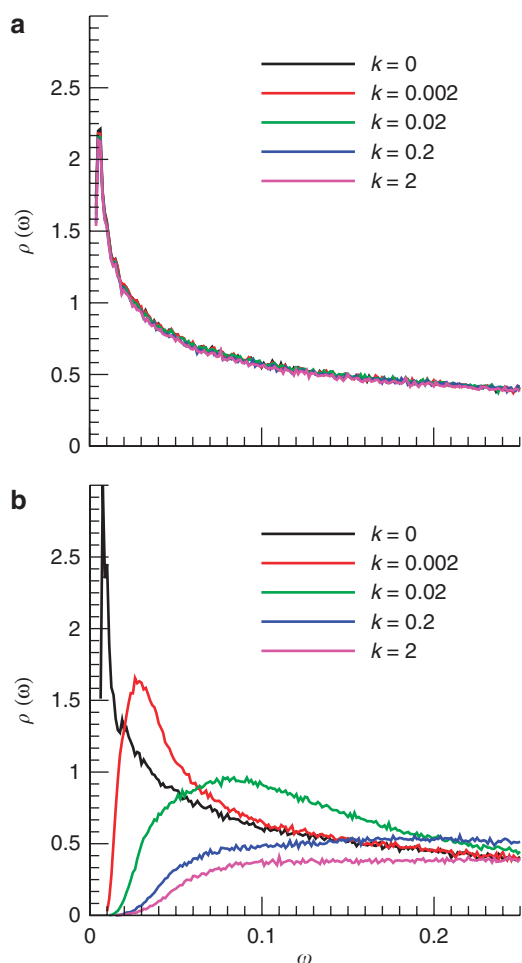


Figure 2 | Vibrational density of states. Low-frequency part of the density of vibrational states as a function of the frequency $\omega = \sqrt{\lambda}$ (the λ s are the eigenvalues of the Hessian matrix) for polymers of length $N=100$ and for various values of the contact stiffness parameter k : **(a)** coil-like conformations (5 contacts); **(b)** compact conformations (number of contacts between 52 and 59). The absence of a relevant dependence on k for coil conformations contrasts with the dramatic effect that k has on compact conformations.

heat is enhanced at the transition if the high-temperature phase is softer, that is, richer in low-frequency vibrational modes²⁹, than the low-temperature one. In the present case, the scenario is even more dramatic because the coil and globule phases are completely different from a topological point of view; coils are low-dimensional, typically soft structures, whereas globules are higher-dimensional and stiffer conformations. More precisely, the difference in rigidity between coil and globule conformations is controlled by the elastic constant k associated with non-bonded interactions. Coil conformations, featuring few contacts, are characterized by vibrational spectra rather insensitive to changes in k (Fig. 2a). By converse, the spectra of globular conformations crucially depend on k ; for $k=0$, only backbone potential terms (bond-length, bending and torsional angles) are present, so that all conformations have a vibrational spectrum akin to those of coils (Fig. 2b); as k increases, the elastic terms transverse to the backbone become progressively more important, making the spectra more similar to what is expected for a compact solid, with a dramatic depletion of low-frequency modes, the ones that contribute the most to the entropy.

Three consequences proceed from the above argument. First, the discontinuous nature of the phase transition should be magnified by increasing values of k , as the depletion of low-frequency modes becomes more pronounced (Fig. 2b). Second, the discontinuous nature of the phase transition should equally become more evident for increasing values of the polymer length N , at constant values of k , because the vibrational modes of longer polymers are progressively softer, hence the low-frequency region more populated and thus the depletion induced by $k \neq 0$ more relevant. Third, for $k=0$, the ISAW and the VISAW should fall in the same universality class, because the entropy weight is roughly the same for coil and globule conformations (Fig. 2a,b for $k=0$), and is, to a good approximation, a simple multiplicative constant for the partition function.

To prove the above inferences, we have analysed the energy probability distribution in these three scenarios. For a fixed length ($N=180$), the histogram does indeed evolve from single to two-state as k increases (Fig. 3a–d). The evolution of the density of states for a fixed k and growing lengths (Fig. 3e–h) reveals that phase coexistence emerges as the length is increased, as argued above. To obtain a complete picture, we finally inspected how the density of states changes with length as $k=0$; no coexistence appears, and the energy histogram approaches, as argued above, one of the traditional tricritical model, which thus defines the universality class for $k=0$ (Fig. 3i–l).

Scaling properties at the transition. The non-analytic part of the extensive free energy of self-interacting polymers in the proximity of the transition is, by scaling arguments³⁰, of the form

$$F(t, N) = F(tN^\phi) \quad (5)$$

where $t = (T - T_C)/T_C$ is the reduced temperature (T_C being the transition temperature) and ϕ is the crossover exponent. Thus, the specific heat, which is the second derivative of the free energy per monomer with respect to temperature, scales at the transition ($t=0$) as $C_V \sim N^{2\phi-1}$. As a consequence of the mean-field nature of tricritical transitions in three dimensions, $\phi = 1/2$, implying that $C_V \sim \text{const}$ (with expected logarithmic corrections)^{8,23,31}. By converse, in the presence of a first-order phase transition, the crossover exponent is $\phi = 1$ and $C_V \sim N$, by standard finite-size arguments^{32,33}. Interestingly, the inverse relation $t \sim 1/N$ (consistent with $\phi = 1$) has been observed for a collection of proteins of various lengths, strengthening the case for a first-order-like protein-folding transition³⁴. Our simulations confirm such predictions (Fig. 4a); the specific heat of the ISAW model barely grows with an effective crossover exponent $\phi = 0.56$, compatible with the expected $\phi = 1/2$ dressed by logarithmic corrections and/or finite-size effects. When instead fluctuations around the energy minima are taken into account, for $k=2$ the specific heat grows more than linearly with N . As the extensivity of the energy sets $\phi = 1$ as the upper bound for the crossover exponent, we interpret superlinearity as a transient behaviour due to the increasing separation of the two energy peaks corresponding to the coexisting phases (see Figs 1g and 3e–h). Once the distance between the two peaks has stabilized, the crossover exponent attains its asymptotic value $\phi = 1$. Calorimetry, thus, confirms the discontinuous nature of the transition of the present model.

The metric properties of polymers at the transition are also of fundamental interest. Polymers below the transition are compact globules and their radius of gyration R scales with their length as $R \sim N^{1/3}$. Much above the transition temperature, the only relevant inter-monomer interaction is excluded volume and $R \sim N^{v_{\text{SAW}}}$, where $v_{\text{SAW}} \approx 0.587$ is the exponent of three-dimensional self-avoiding walks³⁵. Once again, due to the tricritical nature of the transition, characterized by mean-field values of the exponents in three dimensions, at the transition point ISAW polymers behave as

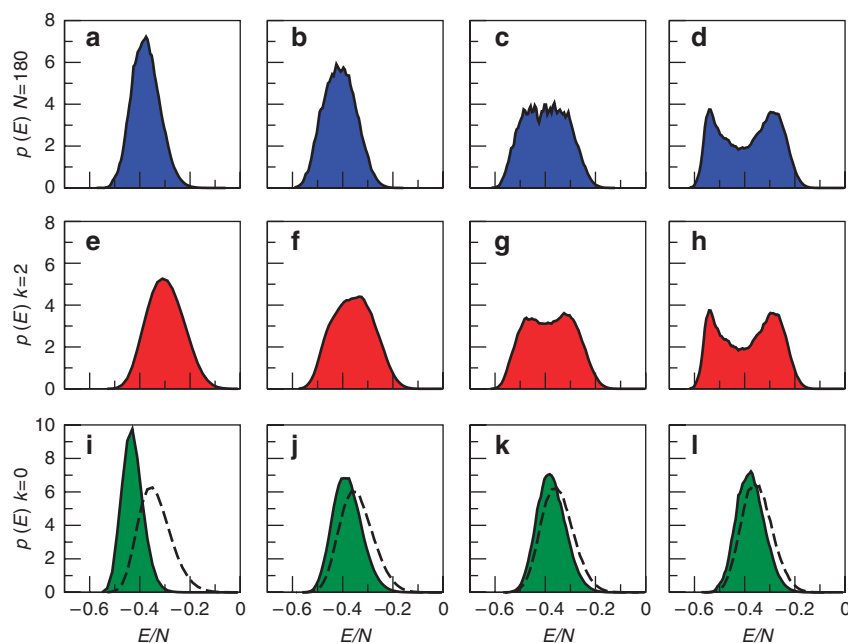


Figure 3 | Evolution of the energy probability distribution of the VISAW. The first row portrays the evolution of the energy (per monomer) probability distribution for a fixed polymer length, $N=180$, and varying values of k (panel **a**: $k=0$, **b**: 0.02, **c**: 0.2 and **d**: 2). As it can be clearly seen, the distribution progressively broadens until two peaks become clearly discernible, signalling phase coexistence. The second row reports the evolution of the distribution for a fixed value of k ($k=2$) and progressively longer polymers (panel **e**: $N=50$, **f**: 100, **g**: 140 and **h**: 180). As for the first row, the distribution becomes broader and, eventually, characterized by two peaks. The behaviour of the VISAW model for $k=0$ and increasing lengths (**i-l**, same length sequence as for the second row) approaches the one of the standard ISAW model (dashed line), suggesting that the ISAW model and the VISAW model for $k=0$ belong to the same universality class.

ideal chains, with the radius of gyration R scaling as the square root of the number of monomers ($R \sim N^{1/2}$). The behaviour of $R/N^{1/2}$ as a function of N at the ISAW CGT transition is shown in Fig. 4b, confirming the theoretical predictions. Two regimes are instead clearly discernible in the metric properties of the VISAW model at the coexistence temperature; polymers of small length ($N \leq 100$) behave as random coils ($R \sim N^{1/2}$), whereas longer polymers ($N \geq 140$) scale as SAWs (inset in Fig. 4b). These results are a consequence of the discontinuous nature of the transition, where no divergences are expected. Indeed, the critical behaviour of the ISAW model conforms to the usual scaling at continuous transitions; the ideal metric properties encompass a number of monomers that scales as $t^{-1/\phi}$, as posited by equation (5), diverging as the whole polymer is taken up at $t \rightarrow 0$. By converse, at a discontinuous transition, the critical region remains of finite size, so that the ideal behaviour that we observe can extend only up to finite lengths, and must cross over to the dominant non-critical scaling, which in the present case is represented by the self-avoiding exponent $\nu \approx 0.587$ (inset of Fig. 4b). To provide a more complete picture, we have also analysed the scaling properties of the VISAW model for $k=0$ (Fig. 4a,b), confirming that in the absence of vibrations of non-bonded contacts, VISAW polymers remain in the ISAW universality class.

Discussion

Ever since de Gennes showed that, for the ISAW model, the CGT is continuous, identifying the features capable of changing its order has remained one of the major challenges in polymer physics. In the years, several different modifications have been proposed, such as switching from flexible to semi-flexible polymers^{15–17}, or suitably tuning the range of the non-bonded attractive interactions¹⁸. These previous attempts have never really succeeded in changing the order of the CGT, which is instead simply pre-empted by a first-order transition under special conditions, such as strongly reduced

flexibility of the chain backbone or very narrow-range interactions. As such scenarios do not really apply to proteins and synthetic polymers, whether any basic polymer model could produce the desired discontinuous transition has remained an unanswered issue.

The VISAW model that we presented here is a natural generalization of the ISAW, stemming from basic considerations about the structure of a realistic potential compatible with the ISAW approximations. The inclusion of fluctuations around inherent structures in the calculation of the partition function is coherent with modern developments in condensed matter physics²⁵ and provides a clearer picture of the CGT. The change in the order of the transition is brought about by the force constant k of non-bonded contacts, because of its depleting effect of low-frequency modes for globular conformations. It is crucial here to stress that it is not the absolute flexibility of the polymers that determines the transition order change, but rather the softness difference between globule and coil conformations.

Our results also shed light on some difficulties that have hindered the assessment of the CGT order from the experimental point of view, and on the problems at reconciling the two-state behaviour of proteins with fundamental polymer physics. Indeed, very soft, non-bonded contacts (small k values in the VISAW model) would not manifest the discontinuity of the transition if not for extremely long polymers. By converse, stiff contacts (high k values) would induce phase coexistence for shorter polymers. In this respect, proteins stand out because the interactions between adjacent amino acids include, most of the time, hydrogen-bond interactions as in secondary structures (α -helices and β -sheets). The hydrogen bond is rather rigid, as captured by coarse-grained models where its stiffness can be up to tens of times larger than for van der Waals or hydrophobic interactions³⁶. As a consequence, at variance with other polymers with softer interaction, such as polyethylene, polymethylene or polystyrene, proteins might be more likely to exhibit two-state

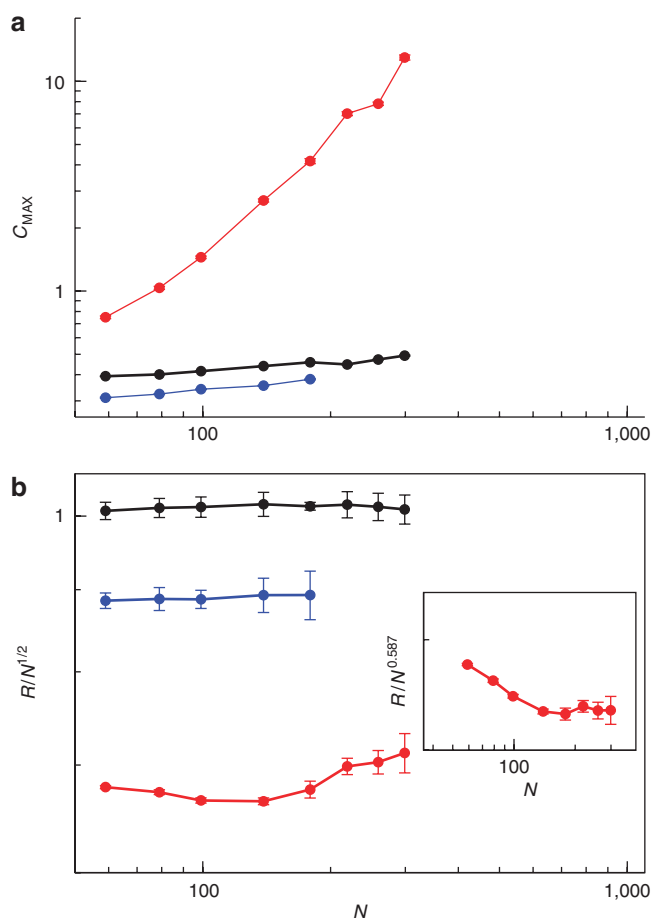


Figure 4 | Scaling properties at the coil-globule transition. (a) Log-log plot of the maximum of the specific heat, C_{MAX} as a function of the polymer length N for the ISAW model (black circles) and for the VISAW model in the two cases $k=0$ (blue circles) and $k=2$ (red circles). The VISAW-specific heat has been reduced by a factor to be plotted with the much lower one of the ISAW without compressing too much the latter. (b) Log-log plot of the ratio of the average radius of gyration, R , to the expected ideal scaling behaviour, $N^{1/2}$, as a function of the polymer length. The data for the ISAW and VISAW ($k=0$) models have been taken at the transition point, identified as the temperature T_C where the curves $R(T)/N^{1/2}$ intersect, rather than as the temperature of the maximum of the specific heat, which is affected by considerable finite size effects. Importantly, in both cases, a temperature T_C is well identified from the data (not shown). In the case of the VISAW ($k=2$), the data have been taken at the temperature of the maximum of the specific heat, as in this case no clear temperature can be found where the curves $R(T)/N^{1/2}$ intersect, and because the data show, especially for longer polymers, that the transition temperature identified using the Maxwell construction (Fig. 1e) coincides with the one of the maximum of the specific heat. The inset shows the behaviour of $R(N)^{0.587}$ as a function of polymer length to highlight that for longer lengths VISAW polymers behave as self-avoiding chains. Data keys in panel **b** are the same as in panel **a**.

behaviour already at intermediate lengths. At the same time, the limited length of proteins clearly hinders the possibility to unambiguously distinguish discontinuous from continuous transitions^{6,37}. Moreover, it is important to recall that the non-random heterogeneity of the amino acid sequence is likely to contribute to the sharpness of the transition in proteins¹³.

The VISAW model that we have introduced here contributes thus to a recent emerging scenario where salient features of proteins,

such as the two-state behaviour and the presence of secondary structures, would be attributed to previously unexplored general properties of simpler homopolymers^{38,39}.

The results shown here inevitably raise several issues. First and foremost, it is crucial to discuss the relations of the present model with other approaches that have been used in the past to study the CGT. In particular, both lattice and off-lattice models have lent support to the continuous transition framework. In a previous section, it has been shown how the VISAW model builds upon the usual approximations leading to lattice models, pushing them one step further toward lattice polymers closer to the real ones. The connection with off-lattice schemes is instead more delicate. Indeed, most off-lattice approaches describe polymers as freely jointed chains of beads, with no bending and, most importantly, no torsional potentials^{23,40}. A polymer whose backbone is described by a potential similar to the one of equation (4) can be considered as fully flexible in the following scenarios: either some re-scaling is implicitly invoked, such that the connection between consecutive beads actually captures several chemical monomers that are in a random coil conformation, or the energy barrier between torsional minima is much lower than the thermal energy at the CGT, in which case polymers do not feel any torsional restriction. Neither assumption, yet, is likely to apply to real polymers; the re-scaling procedure leading to a freely jointed description cannot be applied equally well to the coil and to the globule phases, as they are conformationally different both at the global and local scales; as well, it has been shown that realistic torsional potentials do strongly confine conformations at room temperature^{41,42}, which is close to the CGT for real polymers, implying that the barriers between trans and gauche_± conformations are significantly larger than the thermal energy. As a consequence, we argue that a flexible, torsion-free polymer model is not necessarily a good model to describe the CGT of real polymers. Furthermore, it is not easy to draw conclusions about the connections between lattice VISAW and off-lattice torsion-free polymers, as the latter cannot be recovered as a limit of the former. Rather a full-fledged off-lattice study would be needed, with the energy barrier between torsional minima as an extra parameter.

The present work has also some more theoretical, and in some sense fundamental, implications. As $k=0$ still belongs to the ISAW universality class but $k \neq 0$ does not, our generalization has actually introduced a new parameter that is relevant, in a renormalization group perspective, because it leads away from the usual tricritical point. Further simulations for small values of k , thus close to the ISAW CGT, could also uncover the scaling relation between k , t and N , leading to a free energy of the form $F(N^{\phi_t} t, N^{\phi_k} k)$ with ϕ_t the usual crossover exponent. Unfortunately, the heavy computational costs implied by the need to diagonalize the Hessian matrix at each Monte-Carlo step is presently hindering our capability to analyse, with reliable statistics, the large span of lengths that would be necessary to assess such subtle scaling features. This, as many other analyses, will be needed in the future to fully understand the properties of this new model and to unveil its connections with the behaviour of real polymers and proteins.

Methods

Details of the model. In this work, we adopted non-dimensional units. In particular, the energy of non-bonded monomer contacts is $\epsilon=1$ and the Boltzmann constant $k_B=1$. Temperature is expressed in units of ϵ/k_B . The lattice spacing is set to $l=1$, and the stiffness constants of the potential equation (4) are expressed in units of energy. The stiffness constants of the elastic potential equation (4) are set to $k_{bb}=150$, $k_0=1$, $k_{\psi_t}=1$ and $k_{\psi_g}=0.6$. The non-bonded interaction constant k is varied in the range (0,2). The stiffness of the bond between two consecutive monomer pairs, k_{bb} , is much larger than the others, consistently with other coarse-grained off-lattice force fields^{36,43,44} where the bond is rigid ($k_{bb}=\infty$). This choice reflects the almost fixed length of a covalent bond. The ratio between the trans (k_{ψ_t}) and gauche (k_{ψ_g}) stiffness constants is such that the eigenvalues of the corresponding Hessian sub-matrices are equal. In this way the trans and the gauche

direction are individually associated with the same weight and we avoid introducing any chain semi-flexibility that was already treated in previous models^{15–17}.

The sum over consecutive quadruplets (third term of equation (4)) can be expressed more explicitly as three sums over the trans, gauche₊ and gauche₋ quadruplets:

$$\sum_{i=1}^{N-3} \frac{1}{2} k_{\psi_{t,g}} (\psi_{i+3,i} - \psi_{t,g_{\pm}})^2 = \sum_{\text{trans}} \frac{1}{2} k_{\psi_t} (\psi_{i+3,i} - \psi_t)^2 + \sum_{\text{gauche}_+} \frac{1}{2} k_{\psi_g} (\psi_{i+3,i} - \psi_{g_+})^2 + \sum_{\text{gauche}_-} \frac{1}{2} k_{\psi_g} (\psi_{i+3,i} - \psi_{g_-})^2$$

where each quadratic term corresponds to the second-order expansion of a potential that can be written as⁴⁵

$$V(\psi_{i+3,i}) = \sum_k a_k \cos(k\psi_{i+3,i}) \quad (6)$$

with the a_k coefficients chosen so that $V(\psi_{i+3,i})$ has three minima of the same value in $\psi_{i+3,i} = \pi, \pm\pi/3$ with second derivative $V''(\pi) = k_{\psi_t}$ and $V''(\pm\pi/3) = k_{\psi_g}$. The number of a_k coefficients depends then on the number of further conditions that are imposed on $V(\psi_{i+3,i})$.

Details of the simulations and of data analysis. The model was investigated on the tetrahedral lattice by means of dynamic Monte Carlo with Pivot moves^{46,47}. Each new proposed configuration (γ_{new}) is accepted, if self-avoiding, with probability

$$p(\gamma_{\text{old}} \rightarrow \gamma_{\text{new}}) = \min \left(1, e^{\beta[n(\gamma_{\text{new}}) - n(\gamma_{\text{old}})]} \frac{w(\gamma_{\text{new}})}{w(\gamma_{\text{old}})} \right). \quad (7)$$

The Multiple Histogram Method technique⁴⁸ was used to merge data obtained from simulations at different temperatures and fixed length. The error bars were estimated with the bootstrap method. The derivative of the microcanonical entropy $\beta = dS(E)/dE$ (Fig. 1) was calculated by means of a low-noise Lanczos differentiator⁴⁹.

References

- Stockmayer, W. H. Problems of the statistical thermodynamics of dilute polymer solution. *Makromolekul. Chem.* **35**, 54–74 (1960).
- De Gennes, P.-G. *Scaling Concepts in Polymer Physics* (Cornell University, 1979).
- Gil, E. S. & Hudson, S. M. Stimuli-responsive polymers and their bioconjugates. *Prog. Polym. Sci.* **29**, 1173–1222 (2004).
- Schmaljohann, D. Thermo- and pH-responsive polymers in drug delivery. *Adv. Drug Deliver. Rev.* **58**, 1655–1670 (2006).
- Minko, S. Responsive polymer brushes. *Polym. Rev.* **46**, 397–420 (2006).
- Ziv, G., Thirumalai, D. & Haran, G. Collapse transition in proteins. *Phys. Chem. Chem. Phys.* **11**, 83–93 (2009).
- De Gennes, P.-G. Collapse of a polymer-chain in poor solvents. *J. Phys. Lett. Paris* **36**, L55–L57 (1975).
- De Gennes, P.-G. Collapse of a flexible polymer chain-II. *J. Phys. Lett. Paris* **39**, L299–L301 (1978).
- Flory, P. J. The configuration of real polymer chains. *J. Chem. Phys.* **17**, 303–310 (1949).
- Vshivkov, S. A. & Safronov, A. P. The conformational coil-globule transition of polystyrene in cyclohexane solution. *Macromol. Chem. Phys.* **198**, 3015–3023 (1997).
- Tiktupulo, E. I. *et al.* Domain coil-globule transition in homopolymers. *Macromolecules* **28**, 7519–7524 (1995).
- Tiktupulo, E. I., Bychkova, V., Ricka, J. & Ptitsyn, O. B. Cooperativity of the coil-globule transition in a homopolymer - microcalorimetric study of poly(*n*-isopropylacrylamide). *Macromolecules* **27**, 2879–2882 (1994).
- Ramanathan, S. & Shakhnovich, E. Statistical-mechanics of proteins with evolutionary selected sequences. *Phys. Rev. E* **50**, 1303–1312 (1994).
- Kaya, H. & Chan, H. S. Simple two-state protein folding kinetics requires near-Levinthal thermodynamic cooperativity. *Proteins* **52**, 510–523 (2003).
- Kolinski, A., Skolnick, J. & Yaris, R. Monte-Carlo simulations on an equilibrium globular protein folding model. *Proc. Natl Acad. Sci. USA* **83**, 7267–7271 (1986).
- Kolinski, A., Skolnick, J. & Yaris, R. The collapse transition of semiflexible polymers - a Monte-Carlo simulation of a model system. *J. Chem. Phys.* **85**, 3585–3597 (1986).
- Doniach, S., Garel, T. & Orland, H. Phase diagram of a semiflexible polymer chain in a theta solvent: Application to protein folding. *J. Chem. Phys.* **105**, 1601–1608 (1996).
- Taylor, M. P., Paul, W. & Binder, K. All-or-none proteinlike folding transition of a flexible homopolymer chain. *Phys. Rev. E* **79**, 050801 (2009).
- Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. & Gaub, H. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109–1112 (1997).
- Munk, P. & Aminabhavi, T. M. *Introduction to Macromolecular Science* (Wiley-Interscience, 2002).
- Vanderzande, C. *Lattice Models of Polymers* (Oxford University, 1998).
- Zifferer, G. Monte carlo simulation of tetrahedral chains, 6. Linear and star-branched polymers near to theta-conditions. *Macromol. Chem. Theor.* **2**, 653–671 (1993).
- Grassberger, P. & Hegger, R. Simulations of 3-dimensional theta-polymers. *J. Chem. Phys.* **102**, 6881–6899 (1995).
- Stillinger, F. H. & Weber, T. A. Inherent structure in water. *Phys. Rev. A* **25**, 978–989 (1982).
- Wales, D. J. *Energy Landscapes* (Cambridge University, 2003).
- Wales, D. J. & Berry, R. S. Coexistence in finite systems. *Phys. Rev. Lett.* **73**, 2875–2878 (1994).
- Wales, D. J. & Doye, J. P. K. Coexistence and phase-separation in clusters - from the small to the not-so-small regime. *J. Chem. Phys.* **103**, 3061–3070 (1995).
- Hernandez-Rojas, J. & Gomez Llorente, J. M. Microcanonical versus canonical analysis of protein folding. *Phys. Rev. Lett.* **100**, 258104 (2008).
- Morris, J. R. & Gooding, R. J. Vibrational entropy effects at a diffusionless 1st-order solid-to-solid transition. *Phys. Rev. B* **43**, 6057–6067 (1991).
- Brak, R., Owczarek, A. L. & Prellberg, T. A scaling theory of the collapse transition in geometric cluster-models of polymers and vesicles. *J. Phys. A Math. Gen.* **26**, 4565–4579 (1993).
- Duplantier, B. Tricritical polymer-chains in or below 3 dimensions. *Europhys. Lett.* **1**, 491–498 (1986).
- Binder, K. & Landau, D. P. Finite-size scaling at 1st-order phase-transitions. *Phys. Rev. B* **30**, 1477–1485 (1984).
- Challa, M. S. S., Landau, D. P. & Binder, K. Finite-size effects at temperature-driven 1st-order transitions. *Phys. Rev. B* **34**, 1841–1852 (1986).
- Li, M. S., Klimov, D. K. & Thirumalai, D. Finite size effects on thermal denaturation of globular proteins. *Phys. Rev. Lett.* **93**, 268107 (2004).
- Clisby, N. Accurate estimate of the critical exponent ν for self-avoiding walks via a fast implementation of the pivot algorithm. *Phys. Rev. Lett.* **104**, 055702 (2010).
- Han, W. & Wu, Y.-D. Coarse-grained protein model coupled with a coarse-grained water model: Molecular dynamics study of polyalanine-based peptides. *J. Chem. Theory. Comput.* **3**, 2146–2161 (2007).
- Klimov, D. K. & Thirumalai, D. Linking rates of folding in lattice models of proteins with underlying thermodynamic characteristics. *J. Chem. Phys.* **109**, 4119–4125 (1998).
- Maritan, A., Micheletti, C., Trovato, A. & Banavar, J. R. Optimal shapes of compact strings. *Nature* **406**, 287–290 (2000).
- Banavar, J. R., Hoang, T. X., Maddocks, J. H., Maritan, A., Poletto, C., Stasiak, A. & Trovato, A. Structural motifs of biomolecules. *Proc. Natl Acad. Sci. USA* **104**, 17283–17286 (2007).
- Rubio, A., Freire, J., Clarke, J., Yong, C. & Bishop, M. Characterization of the theta-state and transition curves of off-lattice 3-dimensional chains. *J. Chem. Phys.* **102**, 2277–2281 (1995).
- Launne, T., Neelov, I. & Sundholm, F. Conformational study of substituted p-phenylene terephthalates with molecular mechanics and Metropolis Monte Carlo methods. *Polymer* **40**, 2313–2321 (1999).
- Karayiannis, N., Mavrantzas, V. & Theodorou, D. Detailed atomistic simulation of the segmental dynamics and barrier properties of amorphous poly(ethylene terephthalate) and poly(ethylene isophthalate). *Macromolecules* **37**, 2978–2995 (2004).
- Smith, G. D. & Yoon, D. Y. Equilibrium and dynamic properties of polymethylene melts from molecular-dynamics simulations. 1. N-tridecane. *J. Chem. Phys.* **100**, 649–658 (1994).
- Honeycutt, J. D. & Thirumalai, D. The nature of folded states of globular proteins. *Biopolymers* **32**, 695–709 (1992).
- Smith, G. D., Jaffe, R. L. & Yoon, D. Y. Conformational characteristics of poly(tetrafluoroethylene) chains based upon ab-initio electronic-structure calculations on model molecules. *Macromolecules* **27**, 3166–3173 (1994).
- Zifferer, G. Monte-carlo simulation of tetrahedral chains. 1. Very long (athermal) chains by pivot algorithm. *Macromolecules* **23**, 3166–3172 (1990).
- Madras, N. & Sokal, A. The pivot algorithm: a highly efficient Monte Carlo method for the self-avoiding walk. *J. Stat. Phys.* **50**, 109–186 (1988).
- Ferrenberg, A. & Swendsen, R. H. Optimized monte carlo data analysis. *Phys. Rev. Lett.* **63**, 1195–1198 (1989).
- Savitzky, A. & Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627–1639 (1964).

Acknowledgements

C.M. thanks A. Trovato and F. Seno for enlightening discussions. C.M. and P.D.L.R. thank the Swiss National Science Foundation for financial support.

Author contributions

C. M. performed simulations and analysed the results. C.M., L.C., F.P. and P.D.L.R. designed the model. All authors wrote the manuscript.

Additional information

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Maffi, C. *et al.* First-order coil-globule transition driven by vibrational entropy. *Nat. Commun.* 3:1065 doi: 10.1038/ncomms2055 (2012).