Sequence Controlled Self-Knotting Colloidal Patchy Polymers

Ivan Coluzza,^{1,*} Peter D. J. van Oostrum,² Barbara Capone,¹ Erik Reimhult,² and Christoph Dellago¹

¹Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria

²Department of Nanobiotechnology, University of Natural Resources and Life Sciences Vienna, Muthgasse 11, 1190 Vienna, Austria (Received 30 October 2012; published 11 February 2013)

Knotted chains are a promising class of polymers with many applications for materials science and drug delivery. Here we introduce an experimentally realizable model for the design of chains with controllable topological properties. Recently, we have developed a systematic methodology to construct self-assembling chains of simple particles, with final structures fully controlled by the sequence of particles along the chain. The individual particles forming the chain are colloids decorated with mutually interacting patches, which can be manufactured in the laboratory with current technology. Our methodology is applied to the design of sequences folding into self-knotting chains, in which the end monomers are by construction always close together in space. The knotted structure can then be externally locked simply by controlling the interaction between the end monomers, paving the way to applications in the design and synthesis of active materials and novel carriers for drugs delivery.

DOI: 10.1103/PhysRevLett.110.075501

PACS numbers: 81.16.Dn, 05.10.Ln, 81.05.Zx, 87.15.Cc

Molecular knots are found in nature mainly in DNA and protein molecules [1–4]. The function and the evolutionary advantage behind the origin of natural knots in biological systems is still controversial; however, they receive increasing interest from the scientific community for their peculiar structural and enzymatic properties [5-10]. In materials science the synthesis of knotted molecules has seen a significant effort of chemists who aim for the production of complex knots or interconnected molecules that would assemble into larger intertwined macromolecules [11,12]. In polymer science the study of knots is of high interest, especially for the synthesis of molecules to be used for drug delivery [13–17] as polymer architecture is an essential parameter to control the diffusion of drugs into the body [18–21]. In particular, the cyclization of a protein based pain killer, normally administered intravenously, resulted in a highly stable variant that could resist the gastrointestinal digestion process [22]. However, to this date the synthesis of knotted molecules is a challenging task, especially if a specific internal structure of the knot is required.

Recently, we introduced [23] a methodology to design patchy colloidal chains capable of folding into given target structures. The strength of our methodology is in the ability to design a large variety of folded chain structures by merely selecting the sequence of a simple, experimentally realizable, small set of particle types in a string, reminiscent of how the amino acid sequence uniquely determines protein structures [24]. Here we further demonstrate the robustness of this theoretical framework and apply it to the design of experimentally realizable *self-knotting* patchy chains. The knots are generated by the spontaneous folding of chains of particles designed to collapse into target conformations with the two end segments in close proximity. The foldable chains are based on a small set of anisotropically interacting modular subunits that can be realized as patchy colloids (see Fig. S1 in the Supplemental Material [25]). In contrast to our previous work [23], the patchy particles are decorated with a single patch. We demonstrate that single-patch particles, which have the advantage of easier experimental implementation, retain the designable folding of the more complex, amino acid-like, two-patch particles in our previous work [23]; however, the reduction to a single patch dramatically changes the typical chain secondary structure and topology [26]. Single-patch chains tend to form helical secondary structures and this in turn seems to drive the spontaneous formation of knotted conformations (see Fig. S2 in the Supplemental Material [25]). We further demonstrate the design and folding of a knotted chain, for which the structure can be locked or released by switching the cross-linking between the end monomers of the chain. It is shown how locking the knot renders the folded structure more stable to, e.g., heat induced unfolding.

For our model we considered chains of particles that are connected at their poles and decorated on their surface with a single patch placed on the equator along which it is free to rotate. The patches give rise to a directional, attractive interaction between the spheres in the string. In addition, the colloidal particles interact pairwise via a hard core and an attractive or repulsive finite range potential with a strength that depends on the types of the interacting particles (see the Supplemental Material [25]). Each particle is assigned a particular type (letter), which denotes its interactions with the other particles of the chain. We show that both with a highly reductionist and experimentally realizable two letter alphabet distinguishing only between neutral (N) and attractive particles (A), and a larger alphabet made of 20 different letters (see the Supplemental Material [25]), it is possible to design sequences of particle types for which the ground state configuration of the patchy polymer is unique.

Experimentally, a sufficient control over the surface properties of patchy particles to correspond to our model can be obtained by existing synthetic protocols [27,28]. Altogether, many techniques for the production of colloidal particles with anisotropic interactions are presently available [29–33]. Particle strings have been assembled from both nanoparticles [34–36] and microparticles [37–45].

The first step we need to perform, in order to construct patchy colloids that fold into knots, is to identify a set of structures that can be used as targets for the design. It is expected that not all structures are equally easy to design [23]. Thus, possible target structures are first identified with a Monte Carlo simulation that we called SEEK (see the Supplemental Material [25] and Ref. [23]). For a selected number of typical target structures we then design the optimal sequence to fold into that structure. These simulations we called DESIGN (see the Supplemental Material [25] and Ref. [23]). Finally, we study the folding behavior of these example chains by looking at their specific configurational free energy landscape with an algorithm that we named FOLDING (see the Supplemental Material [25] and Ref. [23]).

We use this procedure to design a self-knotting sequence both of a relatively simple patchy polymer of length N = 20 and an alphabet of M = 2 different patchy particles as well as of a longer patchy polymer of length N = 50 and M = 20, i.e., with an alphabet of 20 letters.

In Fig. 1, we show the free energy, generated using the SEEK scheme, as a function of the number Q_S of spheresphere contacts and the number Q_P of patch-patch contacts obtained for the short and long chains respectively. Both free energy surfaces in Fig. 1 show a large basin that dominates over the rest of the Q_S - Q_P space. Although these free energy minima are not necessarily global minima, several simulations started from different initial conditions vielded identical results indicating that these minima are accessible from a large portion of configuration space. A remarkable property of all observed configurations was the propensity of the chain to wrap on itself creating knottedlike structures (see insets of Fig. 2). Such a preference for knotted structures we believe originates from the tendency of single-patch chains to create helical secondary structures [26] (see Fig. S2 in the Supplemental Material [25]) that under the pressure of the isotropic interactions spontaneously form bundles where knotted topologies have a high probability of forming. If the ends of the chain are sufficiently close to each other, the chain could be externally locked together by a stable bond, such that the overall chain topology would be maintained independently of the intrachain interaction strength, effectively locking the knot. Such string configurations will be of particular importance for applications because they will retain a large portion of their native structure also at conditions far away from the ones at which the original folding occurred [22]. On each of the calculated free energy surfaces we selected a configuration from the global minimum for N = 20 (Min₂₀T =0.4) and N = 50 (Min₅₀T = 0.4). In addition, for longer



FIG. 1 (color online). Color-coded free energy surfaces $F(Q_S, Q_P)/k_B T_{\text{Ref}}$ for (a) (N = 20, M = 2) and (b) (N = 50, M = 20) plotted as a function of the number of contacts Q_S among the spheres and the number of contacts Q_P between the patches at temperature T = 0.4. The simulations were performed by altering the structure of the chain as well as by mutating the identities of the particles; structures near the free energy minimum are expected to be the most designable ones, because they are observed most frequently. Both plots show a broad free energy minimum indicated by the blue cross. On (b) we selected 4 additional locations of interest that correspond to the compact configuration Compact (diamond), highly connected structure Many P Bonds (square), the position of the free energy minimum at T = 0.8 MinT = 0.8 (triangle), and the knotted structure Knot (inverted triangle).

chains [see Fig. 1(b)] we chose four additional types of target: namely one configuration with the free energy minimum at a higher temperature (MinT = 0.8), one configuration with a large number Q_P of patch bonds (Many *P* bonds), one configuration that is more compact (large value of Q_S , Compact), and one knotted structure (Knot) selected as Min₅₀T = 0.4 to be close to the global minimum but with the additional constraint that the end monomers are designed to be closer than 3 particle radii (no other spheres can place itself in between). The structure Knot was identified to consist of two concatenated 3 foil knots (the simplest possible knot) using the tools developed by Tubiana *et al.* [4,46].

As we have shown in Ref. [23], SEEK is reliable to find viable targets, but the sequences generated during the simulations are not necessarily good folders. Hence, we determined the optimal sequence that folds into the 20 particle chain target structure, marked with the blue cross in Fig. 1, by applying the DESIGN procedure. The folding properties



FIG. 2 (color online). Folding free energy landscape $F(\text{DRMSD})/k_B T_{\text{Ref}}$ of the designed patchy polymer (a) $Min_{20}T = 0.4$ and (b) $Min_{50}T = 0.4$ as a function of the distance root mean square displacement DRMSD from the target structure. The target structures are shown in the right insets and the structures corresponding to the free energy minima are shown in the top left insets. We have painted each segment type with a different color to highlight the sequence along the chains (for M = 2 neutral particles are colored in gray while the attractive ones are colored in green). In (a) we plot the free energy for the short N = 20 chain scenario at T = 0.4 below the folding temperature T_F (red continuous curve), at T = 0.9 close to T_F (blue dot-dashed line), and at T = 1.0 above T_F (black dashed line). The plot in (b) shows the folding of the long chain with N = 50 at T = 0.6 below the folding temperature T_F (red continuous curve), at T = 0.9 close to T_F (blue dot-dashed line), and at T = 2.0 above T_F (black dashed line).

of the optimal sequence were then analyzed with the FOLDING algorithm. In Fig. 2 we plot the folding free energy profile $F(\text{DRMSD})/k_BT_{\text{Ref}}$ as a function of the distance root mean square displacement of the interparticle distance (DRMSD) (see definition in the Supplemental Material [25]) between the target structure and each sampled configuration. The plot shows that at low temperatures there is a funneled profile with a global minimum very close to the target structure (DRSMD = $0.2R_{\text{HC}}$). We then designed the

 $Min_{50}T = 0.4$ structure as well, and in Fig. 2 we plot the folding free energy $F(DRMSD)/k_BT_{Ref}$ of the resulting sequence. The computed free energy again converged to a profile with a clear global minimum at small values of DRMSD $\approx 0.3 R_{\rm HC}$, indicating that the equilibrium configurations of the system were mainly concentrated close to the target structure. As the FOLDING simulations were performed at various temperatures we could estimate the folding temperature $(T_F \approx 0.9 \text{ for } \text{Min}_{20}T = 0.4 \text{ and}$ $T_F \approx 1$ for $\dot{\text{Min}}_{50}T = 0.4$) to be the temperature for which the folded states and unfolded states preferred at higher temperatures have approximatively the same free energy value. Because of the definition of DRMSD (see the Supplemental Material [25]), the smaller the value the fewer are the possible structures that can have this value of DRMSD. Ultimately, it is expected that at finite temperatures the free energy minimum is located near but not exactly at DRSMD = 0 due to thermal fluctuations. The local minima (DRMSD $\approx 0.9R_{\rm HC}$ for N = 20, and DRMSD $\approx 1.5 R_{\rm HC}$ for N = 50) indicate that, due to the intricate topology, the chain remains trapped for relatively short times in non-native configurations, and the barrier reflects the partial unfolding necessary to unravel the misfolded structure. Hence, the funneled landscape with single minimum implies that both an ensemble of arrested structures and misfolded structures are less stable compared to the desired configuration. We can therefore conclude that the SEEK procedure is able to identify a suitable target for the design, and DESIGN is capable of selecting at least one sequence with good folding properties.

An important characteristic of the structures $Min_{20}T =$ 0.4 and $Min_{50}T = 0.4$ is that they are neither representative of the most compact structures (large values of Q_S in Fig. 1) nor of the configurations with the largest amount of patch bonds (large values of Q_P). Hence, we compared the refolding properties of sequences optimized to fold into compact structures (Compact), structures with many patch bonds (Many P Bonds) and the minimum free energy structure at the higher temperature T = 0.8 (MinT = 0.8). The results confirm what was already observed for twopatch polymers [23], namely that highly compact structures will tend to produce so-called golf-hole landscapes, where the global energy minimum is so sensitive to fluctuations that it is rarely observed (see the Supplemental Material [25]). Our results therefore indicate that for a structure to be designable it should have many patch-patch interactions, while a more compact structure not only disfavors refolding but also disrupts the designability.

The knotted Knot structure was chosen among the compact structures observed with the SEEK algorithm for the long chain scenario with N = 50 and M = 20. The condition imposed on Knot that the chain ends are in close proximity resulted in a more compact structure than $Min_{50}T = 0.4$. However, the large number of patch-patch bonds should guarantee designability (see Fig. S2 [25] and Ref. [23]). As above, we used the DESIGN procedure to determine a sequence that folds into the target structure Knot. After that, we tested the folding properties of this string by using the FOLDING procedure. In Fig. 3 we plot the free energy $F(DRMSD)/k_BT_{Ref}$ for the chain with free ends (named Free in the legend). This free energy profile indicates that the chain at lower temperatures is indeed capable of folding to the target structure with a remarkable precision (DRMSD $\approx 0.2R_{HC}$). There are no major barriers in this free energy profile which indicates that the folding is not hampered by dynamical arrest. It is important to stress that in Knot the end monomers are very close together ($\sim 2.4R_{HC}$), a feature that is observed in the vast majority of the configurations sampled for this design below the folding temperature.

Next we considered the scenario in which the two end monomers are joined by a standard particle-particle bond (see the Supplemental Material [25]) after the target structure has been reached. Experimentally this can easily be implemented by functionalizing the end segments with reactive groups (e.g., by coating patchy particles) for direct chemical cross-linking or cross-linking through a short flexible polymer segment. The effect of this bond is to lock the topology of the chain into the knotted configuration regardless of the strength of the interactions between the monomers. Hence, the chain cannot change its



FIG. 3 (color online). Free energy $F(\text{DRMSD})/k_B T_{\text{Ref}}$ as a function of the DRMSD for the designed patchy polymer Knot and Knot_{Closed} obtained for T = 0.8 and T = 2.0, respectively, below and above the folding temperature $T_F \approx 1.0$ of the Knot protein. We have painted each sphere type with a different color to highlight the sequence along the chains. The little padlocks positioned next to end monomers indicate either that ends are linked with a chain backbone bond (locked) or free (open). Below T_F both chains have a minimum of the free energy around $DRMSD = 0.2R_{HC}$ so close to the target structure (DRMSD = $0R_{\rm HC}$) that the typical snapshots representative of the structure in the free energy minimum (left and right in the top insets) are virtually identical to the target structure (represented in the bottom inset). Above the folding temperature $T_F \approx 1$ of the Knot chain, the closed chain Knot_{Closed} still folds near the native structure (DRMSD = $0.75R_{\rm HC}$) and maintains the same knotted topology of the native structure (see rightmost inset).

topology or completely unfold even at higher temperatures (or disruptive ambient conditions, similarly to synthetic cyclization of proteins as in Refs. [22,47]). We will refer to the closed chain as Knot_{Closed}. The stabilization effect of the locking is proven by the free energy calculations done with a modified version of the FOLDING algorithm where we did not allow for the chain to change its topological state (such a noncrossing condition was not necessary when the ends of the chain were free). In Fig. 3 we plot the free energy $F(DRMSD)/k_BT_{Ref}$ profiles above and below the folding temperature for the closed chain (named Closed in the legend). The profiles not only show that at low temperature (below the folding temperature $T_F \approx 1$ measured for the free end chain) the folded state of the locked chain is much more stable than the unlocked chain, but also that the native structure is in large portion retained even significantly above the folding temperature of the free end chain (see right inset of Fig. 3).

In summary, we introduced a strategy for the design of knotted chains of patchy particles and used it to create a lockable knotted polymer model in which the structure of the knots and the distance between the end monomers are controlled by the designed sequence. The stabilizing effect of the locked ends has many implications for the possible application of the designed chains. One can, e.g., imagine the design of chains where the colloidal particles contain or are coated with a drug that, in either case, is then protected by the knotted structure [22], and before the release of the lock the chain will maintain the globular shape making it more difficult for the body to expel it [20]. Similarly, building blocks for supramolecular materials with complex geometry and topology can be synthesized and then stabilized both for nanoscale and microscale colloids, where the supramolecular assembly or application takes place under demanding conditions that would otherwise denature the functional structure.

It is important to stress that the procedure presented here does not depend on the particular choice of the interactions among the spheres or among the patches. The conformational space is accessed simply by controlling the sequence of a small set of universally usable subunits in a patchy colloidal polymer. Different choices of patch distribution and of the interaction potentials are possible and are currently being explored. Inspired by our previous work [26] and Tubiana's method [46], we are also currently implementing a new version of the SEEK algorithm capable of screening the designability of a large number of knot topologies as a function of chain length and alphabet sizes, with the long term goal of isolating arbitrary knotted structures for further design.

Altogether, our results on patchy polymers provide strong evidence that previous insights obtained for proteins are of a very general nature, namely that the constraints arising from the chain topology and directional interactions are sufficient conditions for sequence design [48]. In the case of proteins, the particular geometry induced by the peptide bonds and the network of hydrogen bonds are responsible for the typical protein motifs observed in nature. The similarity of our system to proteins gives an idea of the diversity of possible self-assembling materials that could be fabricated. In the artificial particle chains studied here, these monomer interactions are generalized to those provided by the patches placed on the colloidal particle. However, generalized to colloidal particles, the number, the interaction strength, and the position of the patches can be controlled experimentally over a broad range of parameters. This allows for great freedom to develop self-assembling systems with designed properties in tailor-made conformations. Finally, in this Letter we have shown that the analogy of patchy polymers with proteins extends to the greater stability against unfolding reported for cyclic polypeptides [22].

We would like to thank Ronald Zirbs, Cristiano de Michele, Francesco Sciortino, and Mark Miller for fruitful discussions. We would also like to thank Luca Tubiana and Christian Micheletti for providing the topological characterization of the Knot structure. Finally we would like to thank an anonymous referee for suggesting intriguing ideas to expand this research field. We acknowledge support from the Austrian Science Fund (FWF) within the SFB ViCoM (F 41) and the European Science Foundation Research Network, "Exploring the physics of small devices." All simulations presented in this Letter were carried out on the Vienna Scientific Cluster (VSC).

*ivan.coluzza@univie.ac.at

- S.A. Wasserman and N.R. Cozzarelli, Proc. Natl. Acad. Sci. U.S.A. 82, 1079 (1985).
- [2] J. D. Griffith and H. A. Nash, Proc. Natl. Acad. Sci. U.S.A. 82, 3124 (1985).
- [3] O. Lukin and F. Vögtle, Angew. Chem., Int. Ed. Engl. 44, 1456 (2005).
- [4] A. Rosa, M. Di Ventra, and C. Micheletti, Phys. Rev. Lett. 109, 118301 (2012).
- [5] C. Liang and K. Mislow, J. Am. Chem. Soc. 116, 11189 (1994).
- [6] W. Taylor, Nature (London) 406, 916 (2000).
- [7] W. R. Taylor and K. Lin, Nature (London) 421, 25 (2003).
- [8] H.-X. Zhou, J. Am. Chem. Soc. 125, 9280 (2003).
- [9] L. Chiche, A. Heitz, J.-C. Gelly, J. Gracy, P. T. Chau, P. T. Ha, J.-F. Hernandez, and D. Le-Nguyen, Curr. Protein Pept. Sci. 5, 341 (2004).
- [10] L. Cascales and D. J. Craik, Org. Biomol. Chem. 8, 5035 (2010).
- [11] P. W. K. Rothemund, Nature (London) 440, 297 (2006).
- [12] E. Orlandini and S. G. Whittington, Rev. Mod. Phys. 79, 611 (2007).
- [13] J. Grünewald and M. A. Marahiel, Microbiol. Mol. Biol. Rev. 70, 121 (2006).
- [14] J. Matsoukas, V. Apostolopoulos, E. Lazoura, G. Deraos, M.-T. Matsoukas, M. Katsara, T. Tselios, and S. Deraos, Curr. Med. Chem. 13, 2221 (2006).
- [15] H. Kolmar, Expert Rev. Mol. Diagn. 10, 361 (2010).
- [16] C.J. White and A.K. Yudin, Nat. Chem. 3, 509 (2011).
- [17] A. Roxin and G. Zheng, Future Med. Chem. 4, 1601 (2012).
- [18] L. Y. Qiu and Y. H. Bae, Pharm. Res. 23, 1 (2006).

- [19] N. Nasongkla, B. Chen, N. Macaraeg, M. E. Fox, J. M. J. Fréchet, and F. C. Szoka, J. Am. Chem. Soc. 131, 3842 (2009).
- [20] M. E. Fox, F. C. Szoka, and J. M. J. Fréchet, Acc. Chem. Res. 42, 1141 (2009).
- [21] B. Chen, K. Jerger, J. M. J. Fréchet, and F. C. Szoka, J. Control. Release 140, 203 (2009).
- [22] R.J. Clark, J. Jensen, S.T. Nevin, B.P. Callaghan, D.J. Adams, and D.J. Craik, Angew. Chem., Int. Ed. Engl. 49, 6545 (2010).
- [23] I. Coluzza, P. D. J. van Oostrum, B. Capone, E. Reimhult, and C. Dellago, Soft Matter 9, 938 (2012).
- [24] I. Coluzza, PLoS ONE 6, e20853 (2011).
- [25] See Supplemental Material at http://link.aps.org/ supplemental/10.1103/PhysRevLett.110.075501 for details about the model and methodology.
- [26] I. Coluzza and C. Dellago, J. Phys. Condens. Matter 24, 284111 (2012).
- [27] A.B. Pawar and I. Kretzschmar, Macromol. Rapid Commun. 31, 150 (2010).
- [28] E. Bianchi, R. Blaak, and C. N. Likos, Phys. Chem. Chem. Phys. 13, 6397 (2011).
- [29] Q. Chen, S. C. Bae, and S. Granick, Nature (London) 469, 381 (2011).
- [30] Y.-S. Cho, G.-R. Yi, J.-M. Lim, S.-H. Kim, V.N. Manoharan, D.J. Pine, and S.-M. Yang, J. Am. Chem. Soc. 127, 15968 (2005).
- [31] D. M. Andala, S. H. R. Shin, H.-Y. Lee, and K. J. M. Bishop, ACS Nano 6, 1044 (2012).
- [32] J. Zhang, J. Jin, and H. Zhao, Langmuir 25, 6431 (2009).
- [33] A. Böker, J. He, T. Emrick, and T. P. Russell, Soft Matter 3, 1231 (2007).
- [34] J.-C. Wang, P. Neogi, and D. Forciniti, J. Chem. Phys. 125, 194717 (2006).
- [35] S.A. Claridge, H.W. Liang, S.R. Basu, J.M.J. Fréchet, and A.P. Alivisatos, Nano Lett. 8, 1202 (2008).
- [36] J. Zheng, P.E. Constantinou, C. Micheel, A. P. Alivisatos, R. A. Kiehl, and N. C. Seeman, Nano Lett. 6, 1502 (2006).
- [37] E. M. Furst, C. Suzuki, M. Fermigier, and A. P. Gast, Langmuir 14, 7334 (1998).
- [38] S. Biswal and A. Gast, Phys. Rev. E 68, 021402 (2003).
- [39] A. Yethiraj and A. van Blaaderen, Nature (London) 421, 513 (2003).
- [40] L. Cohen-Tannoudji, E. Bertrand, L. Bressy, C. Goubault, J. Baudry, J. Klein, J.-F. Joanny, and J. Bibette, Phys. Rev. Lett. 94, 038301 (2005).
- [41] R. Dreyfus, J. Baudry, M. L. Roper, M. Fermigier, H. A. Stone, and J. Bibette, Nature (London) 437, 862 (2005).
- [42] D. Zerrouki, J. Baudry, D. Pine, P. Chaikin, and J. Bibette, Nature (London) 455, 380 (2008).
- [43] B. Madivala, J. Fransaer, and J. Vermant, Langmuir 25, 2718 (2009).
- [44] B. Peng, E. van der Wee, A. Imhof, and A. van Blaaderen, Langmuir 28, 6776 (2012).
- [45] H.R. Vutukuri, A.F. Demirörs, B. Peng, P.D.J. van Oostrum, A. Imhof, and A. van Blaaderen, Angew. Chem., Int. Ed. Engl. 51, 11249 (2012).
- [46] L. Tubiana, E. Orlandini, and C. Micheletti, Prog. Theor. Phys. Suppl. 191, 192 (2011).
- [47] J.S. Davies, Journal of Peptide Science 9, 471 (2003).
- [48] I. Coluzza, H. G. Muller, and D. Frenkel, Phys. Rev. E 68, 046703 (2003).