Scaling of the average crossing number in equilateral random walks, knots and proteins

A. Dobay[†], Y. Diao[‡], Jacques Dubochet[†] and A. Stasiak[†]

[†] *Laboratory of Ultrastructural Analysis*

University of Lausanne

Lausanne, CH 1015, Switzerland

[‡] *Department of Mathematics*

University of North Carolina at Charlotte

Charlotte, NC 28223, USA

ABSTRACT

We compare here the scaling behaviour of the mean average crossing number $\langle \mathrm{ACN} \rangle$ of equilateral random walks in linear and closed form with the corresponding scaling observed in natural protein trajectories. We have shown recently that the scaling of $\langle \mathrm{ACN} \rangle$ of equilateral random walks of length n follows the relation $\langle \mathrm{ACN} \rangle = \frac{3}{16} n \ln n + bn$ and that a similar result holds for equilateral random polygons [13]. Furthermore, our earlier numerical studies indicated that when random polygons of length n are divided into individual knot types, the $\langle \mathrm{ACN}(\mathcal{K}) \rangle$ for each knot type $\mathcal{K}$ can be described by a function of the form $\langle \mathrm{ACN}(\mathcal{K}) \rangle = a(n - n_0) \ln(n - n_0) + b(n - n_0) + c$ where a, b and c are constants depending on $\mathcal{K}$ and n_0 is the minimal number of segments required to form $\mathcal{K}$ [13]. Here we analyze in addition natural protein structures and observe that the relation $\langle \mathrm{ACN} \rangle = \frac{3}{16} n \ln n + bn$ also describes accurately the scaling of $\langle \mathrm{ACN} \rangle$ of protein backbones.

Keywords: Random Walks, Random Polygons, Crossing Number, Average Crossing Number, Equilateral Random Walks, Equilateral Random Polygons, Protein Structures, Protein Backbones.

1 Introduction

Random walks are frequently used to model the behaviour of polymers at thermodynamic equilibrium [15, 16, 18]. The simplest but also the most fundamental type of

random walks is represented by chains composed of freely jointed segments of equal length (equilateral) where the individual segments have no thickness. Such random walks are known as ideal random walks and are used to model the behaviour of polymers under so-called theta conditions where polymer segments neither attract nor repel each other [26]. The behaviour of ideal random walks is thoroughly researched and it is well established, for example, that such measures of overall dimensions of ideal random walks like the average end to end distance or the average radius of gyration scale with the number of segments n as nn^{ν} where $\nu = 0.5$ [15, 16, 18]. Although the overall dimensions provide important information about the modelled polymers, frequently additional characteristics of polymers are of interest. Among these characteristics are such that can tell whether a polymer is knotted or can measure the extent of polymer entanglement. In the case of circular polymers their knot type can be rigorously determined [1, 23]. However, the information about the knot type tells us only the minimal crossing number that one could see upon elimination of all nugatory crossings. The determination of the knot type does not tell us, for example, how many nugatory crossings there are in the analyzed polymer trajectories. In addition, in the case of linear polymers it is difficult to define their knotedness [24, 31]. There are no difficulties, however, with defining a directional crossing number that corresponds to the perceived number of crossings that can be observed on an orthogonal projection of a given non-perturbed trajectory in open or cyclised form. To be independent of the choice of a particular projection, the average crossing number ACN is defined as the average of directional crossing numbers over all possible orthogonal projections of a given rigid or momentary trajectory. Here we will consider $\langle \mathrm{ACN} \rangle$, i.e. the average of ACN over the whole statistical ensemble of ideal random walks (or polygons) with a given number of segments. Of course $\langle \mathrm{ACN} \rangle$ corresponds also the time averaged ACN of randomly fluctuating polymer in solution. $\langle \mathrm{ACN} \rangle$ provides an interesting measure of physical behaviour of knotted polymers since it shows a strong correlation with the experimentally observed speed of electrophoretic migration of knotted DNA molecules of the same size but of various knot types [29]. Furthermore, $\langle \mathrm{ACN} \rangle$ correlates well with the expected sedimentation coefficient of different types of DNA knots formed on the same size DNA molecules [32]. In the case of protein chains, $\langle \mathrm{ACN} \rangle$ provides an interesting measure of their compactness [4] and several studies were devoted to investigate how $\langle \mathrm{ACN} \rangle$ in proteins scales with the length of polypeptide chain [2, 3, 5, 6, 19]. Another scaling aspect of ACN was discussed in the case of ideal geometric representations of knots [7, 8, 9, 10, 21]. In this paper we compare the scaling of $\langle \mathrm{ACN} \rangle$ in random walks in open and closed forms with $\langle \mathrm{ACN} \rangle$ scaling determined for protein backbones.

2 Simulation methods

The average crossing number of a random walk or polygon W can be calculated by the modified Gauss formula

$$\frac{1}{4\pi} \int_W \int_W \frac{|(\dot{\gamma}(t), \dot{\gamma}(s), \gamma(t) - \gamma(s))|}{|\gamma_(t) - \gamma(s)|^3} dtds, \tag{1}$$

where γ is the arclength parameterization of W. However, the application of this formula leads to problems when analyzed configurations have some non-consecutive

segments very close to each other. For this reason our numerical determination for the ACN of simulated random chains or of protein backbones is based on detecting the number of crossing points in numerous projections of the analyzed trajectories. We calculated the number of crossings in individual projections and then averaged over 50 randomly chosen directions of projections to obtain a good approximation of the actual ACN value for a given trajectory. To generate random equilateral walks (open walks) we first created a set of random unit vectors with the same origin that uniformly sampled the surface of the unit sphere. The vectors were then joined sequentially while maintaining their original directions. To generate random equilateral polygons (closed walks) we followed the approach of Klenin et al., [22]. To construct a 100 segment long random polygon, for example, we first create a set of 50 random unit vectors uniformly sampling the surface of a unit sphere. Subsequently, we add to this set another 50 unit vectors that are opposite to the original set. This procedure assures that the sum of the 100 vectors is zero and that the trajectory obtained by any random permutation of sequential joining of all 100 vectors will be always closed. To eliminate the correlation between anti-parallel vectors the entire set of vectors was de-correlated by multiple rotations of randomly chosen pairs of vectors around their respective sum vectors. Finally, all 100 randomized vectors were sequentially joined in a random order to create an equilateral random polygon. Knot types of the resulting random polygons were recognized by calculation of their HOMFLY polynomials [17]. The trajectories traced by backbones of proteins were entered as segmented chains where each segment corresponded to the distance between sequential α-carbons in a given protein structure taken from Protein Data Bank (PDB). For good statistics it is important to analyze a maximal number of "independent" proteins and eliminate from the data set all "duplicates" i.e. more than one entries defining the same protein structure under somewhat different crystallization conditions or a very similar protein structure from another organism [30]. The set of filtered and thus "independent" protein trajectories consisting of 2230 entries from PDB was kindly provided by William Taylor.

3 Results

3.1 $\langle$ACN$\rangle$ scaling in linear and closed random walks.

Figure 1(A) shows the $\langle$ACN$\rangle$ values obtained in numerical simulations of ideal random walks in a linear (open) form and in a closed form. We have analyzed walks with up to 1000 segments and each of the $\langle$ACN$\rangle$ data points was obtained by averaging the ACN values from 105 independent random configurations of open or closed random walks of the corresponding size. To fit the data points we have used the analytically predicted scaling function $an \ln n + bn$ [13], leaving the two parameters a and b free. In both cases (closed and open walks) we have obtained an excellent fit (with the correlation coefficient $R = 1$) where the prefactor a was practically equal to the analytically predicted 3/16 ($3/16 \pm 0.00065$ and $3/16 \pm 0.00105$, respectively). Therefore, we have proceeded with another fit where the prefactor a was set to 3/16, leaving just one free parameter b. These fits are presented in Figure 1 and it is visible that the fitted functions pass almost perfectly through all the data points. These results therefore

confirm earlier theoretical prediction that the $\langle \mathrm{ACN} \rangle$ of open and closed random walks scales with the number of segments n as $\frac{3}{16} n \ln n + bn$. Comparing the $\langle \mathrm{ACN} \rangle$ of open and closed random walks of the same chain length one notices that closed walks have

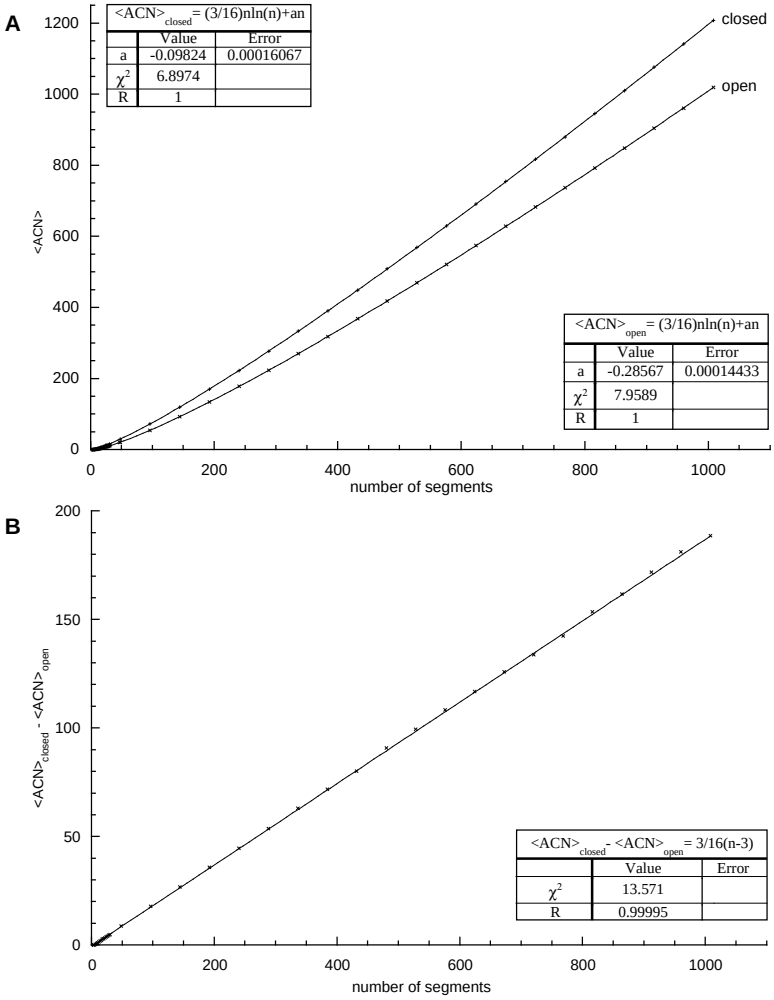

Figure 1: Comparison of the mean average crossing number values $\langle \mathrm{ACN} \rangle$ for corresponding chain lengths of closed and open ideal random walks (equilateral random polygons and equilateral random walks). The standard deviation is about the size of the data points. The values of the correlation coefficient (R) and chi-squared test (χ^2) are given in the inset. A. The $\langle \mathrm{ACN} \rangle$ values obtained in numerical simulations of closed and open random walks are marked as data points and the fitting functions are listed. The analyzed sample sizes of simulated configurations were bigger than 10^5 for each chain length. B. The difference of $\langle \mathrm{ACN} \rangle$ between closed and open ideal random walks of the same length.

higher $\langle ACN \rangle$ than open walks. This result is a consequence of the fact that closed walks have smaller overall dimensions than open walks of the same chain length and this increases the number of perceived crossings in a random projection. As can be expected from the scaling function, the difference of the $\langle ACN \rangle$ between closed and open random walks of the same chain length n can be described by the linear relation. In fact the difference of $\langle ACN \rangle$ between closed and open walks with n segments can be accurately described by the formula $\langle ACN \rangle_{\text{closed}} - \langle ACN \rangle_{\text{open}} = \frac{3}{16}n - c$ where c is the $\langle ACN \rangle$ of three segment-long random walks. Figure 1(B) shows that in closed chains each segment contribution to $\langle ACN \rangle$ is bigger by 3/16 than the corresponding contribution of each segment in open chains.

3.2 $\langle ACN \rangle$ scaling in the individual knot types.

The $\langle ACN \rangle$ data points for closed walks presented in Figure 1 were obtained for all closed random walks of a given chain i.e. the averaging was performed over configurations forming unknots and different types of knots. It is possible however to divide closed walks into individual knot types and to analyze how the $\langle ACN \rangle$ scales when a given type of knot increases its chain length. Figure 2 illustrates how the $\langle ACN \rangle$ values scale with the chain length of random polygons representing various types of knots. It is clearly visible that random configurations of more complex knots have higher $\langle ACN \rangle$ values than the random configurations of simpler knots. This is the consequence of the fact that more complex knots have smaller overall dimensions than less complex knots with the same chain length [14, 25, 27] and therefore in an average projection the chain has to cross with itself more frequently in more complex knots. Interestingly, the $\langle ACN \rangle$ scaling of individual knot types can be accurately fitted with a formula $\langle ACN(\mathcal{K}) \rangle = a(n - n_0)ln(n - n_0) + b(n - n_0) + c$ where n_0 is the minimal number of segments needed to form a given knot [11] and a, b and c are free parameters (the function is defined for $n > n_0$). As can be seen in Figure 2 the prefactor a decreases with the complexity of knots indicating that the contribution of $n \ln n$ part of the formula decreases with the complexity of the knot. However, the contribution of the linear component increases with the complexity of knots. The constant c also increases with the complexity of the knot and its value is close to the ACN of ideal geometric representation of a given knot [21, 28]. For short chain lengths, the difference between the $\langle ACN \rangle$ of random configurations of a given knot type and the $\langle ACN \rangle$ of random configurations of unknots with the same chain length is well approximated by the actual ACN value of ideal (ropelength minimizing) geometric representation of a given knot type [21]. For example, for 6 and 14 segment-long random knots the difference between $\langle ACN \rangle$ of trefoil knots and unknots amounts to 4.14 and 5.21, while the ACN of an ideal trefoil is close to 4.26. This relation between the ACN of ideal knots and the $\langle ACN \rangle$ of random walks with relatively small chain length was noted earlier [21]. However, as the length of the analyzed random walks increases the $\langle ACN \rangle$ values of random knots of different types diverge from each other as observed in [20].

Comparing the fitted parameters of $\langle ACN \rangle$ scaling profiles of individual knot types with these obtained for all closed walks grouped together irrespectively of their knot type (see Figure 2) it is striking that that the $n \ln(n)$ part has a much weaker contribution in the case of individual knot types while the linear contribution is much

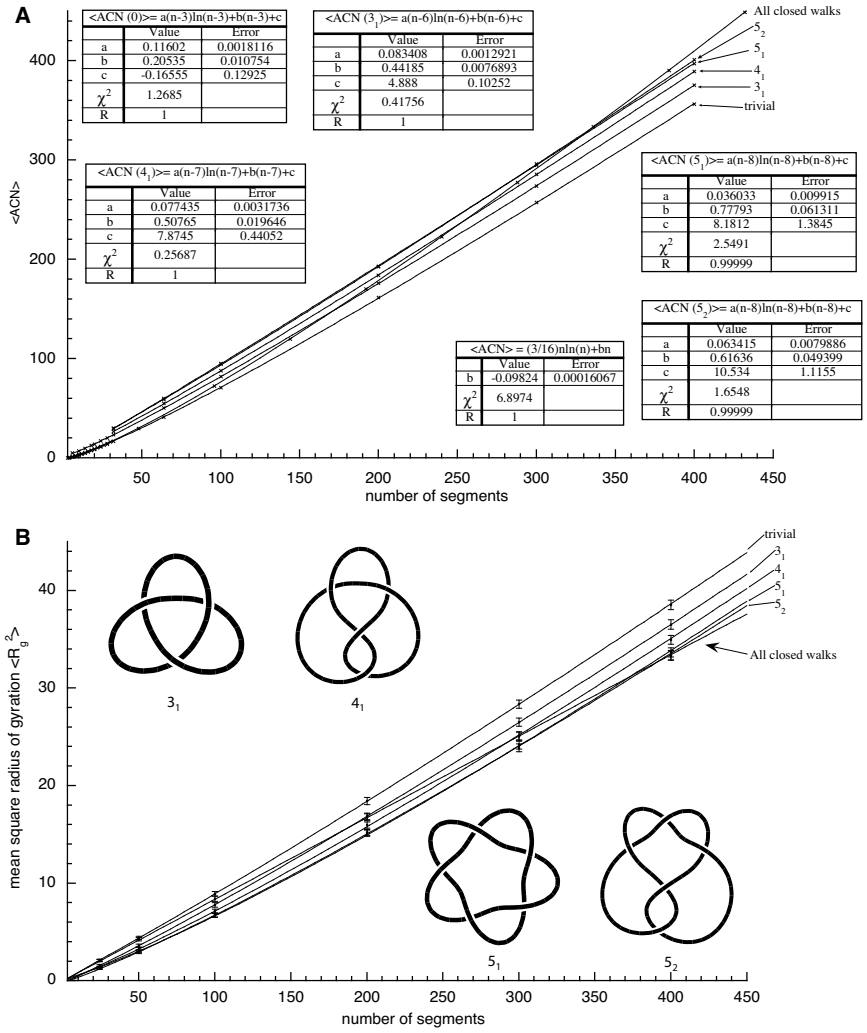

Figure 2: Effect of the topology on the average crossing number values $\langle \mathrm{ACN} \rangle$ for different types of closed random walk. The standard deviation is about the size of the data points. A. The $\langle \mathrm{ACN} \rangle$ values obtained in numerical simulations of closed random walks representing different knot types (trivial, 3_1, 4_1, 5_1 and 5_2) and all closed walks grouped together. The data points are marked and the fitting functions are listed. The statistical sets for different knots and different chain size were not the same. Highest quality data are for unknots with $2,319,455$ configurations analyzed in total and the poorest data set was that of 5_1 knots with $11,406$ configurations analyzed in total. Note that $\langle \mathrm{ACN}(\mathcal{K}) \rangle$ profiles for individual knot types intersect with the $\langle \mathrm{ACN} \rangle$ profile for all closed random chains of the same length. B. The profiles for radius of gyration $\langle \mathrm{R_g}^2 \rangle$ of individual random knots intersect with the $\langle \mathrm{R_g}^2 \rangle$ profile of all closed random walks. The points of intersections define an $\mathrm{R_g}$-based equilibrium length (the data here are taken from [14]).

stronger. Consequently, the $\langle \mathrm{ACN}(\mathcal{K}) \rangle$ profiles of random polygons of each knot type $\mathcal{K}$ show a lower growth rate than the $\langle \mathrm{ACN} \rangle$ profile of all closed walks grouped together independently of their knot type.

3.3 The equilibrium length of a knot.

Comparing the $\langle \mathrm{ACN}(\mathcal{K}) \rangle$ profiles for random polygons of knot type $\mathcal{K}$ with that of all closed walks grouped together, it is visible that (with the exception of the unknots) the $\langle \mathrm{ACN}(\mathcal{K}) \rangle$ profiles intersect with the profile for all closed walks (see Figure 2). This is due to the fact that individual knot types show a smaller growth rate than all closed walks grouped together (see the discussion above), while each individual knot type (with an exception of the unknots) initially has higher $\langle \mathrm{ACN}(\mathcal{K}) \rangle$ values than the $\langle \mathrm{ACN} \rangle$ values for the ensemble of all closed walks. The more complex the knot, the later its $\langle \mathrm{ACN}(\mathcal{K}) \rangle$ profile intersects with the $\langle \mathrm{ACN} \rangle$ profile of all closed walks. The points of intersection between profiles $\langle \mathrm{ACN}(\mathcal{K}) \rangle$ and $\langle \mathrm{ACN} \rangle$ determine $\langle \mathrm{ACN} \rangle$-based equilibrium length of the corresponding knot type $\mathcal{K}$. Below the equilibrium length a given knot shows an excess of $\langle \mathrm{ACN} \rangle$ as compared with the $\langle \mathrm{ACN} \rangle$ of all possible walks realized with the same chain length. Therefore if one would cut a knot realized with a chain shorter than its equilibrium length, let it equilibrate and then promote reclosure of the ends, one would observe a tendency to form simpler knots than the starting knot type. Above the equilibrium length, the situation reverses. If one would cut a knot realized with the chain length longer than its equilibrium length, one would observe after a reclosure of the chain its tendency to form more complex knots. At the equilibrium length, however, a knot would show no tendency to decrease or increase its $\langle \mathrm{ACN} \rangle$ after cutting, equilibration and reclosure. Interestingly, the equilibrium lengths determined by the intersections of $\langle \mathrm{ACN} \rangle$ profiles of individual knot types with $\langle \mathrm{ACN} \rangle$ profile of all closed walks grouped together practically coincide with the equilibrium lengths determined by intersections of the corresponding profiles of the mean radius of gyration $\langle \mathrm{R_g} \rangle$ (see Figure 2). So for example, the equilibrium length for trefoils and figure of eight knots based on measurements of $\langle \mathrm{ACN} \rangle$ amounted to 176 ± 10 and 258 ± 10 segments, respectively, while the equilibrium length of these knots based on measurements of $\langle \mathrm{R_g} \rangle$ [14] amounted to 174 ± 14 and 270 ± 17 segments, respectively (the bigger error range for $\mathrm{R_g}$-based equilibrium length is caused by a smaller statistical set of the data used to determine $\langle \mathrm{R_g} \rangle$).

3.4 Scaling of $\langle \mathrm{ACN} \rangle$ in natural protein structures.

In protein structures one can follow their backbones and analyze trajectories determined by the positions of sequential C_α atoms (α-Carbons). There has been a substantial interest in studies of the entanglement of protein backbones including studies of $\langle \mathrm{ACN} \rangle$ scaling in proteins [2, 3, 5, 6, 19]. Most frequently a simple power law is used to fit the experimental data and to conclude about the general scaling behavior of proteins [2, 3, 5] but $n \ln n$ scaling was also considered [19]. As we have discussed above the $\frac{3}{16} n \ln n + bn$ scaling of $\langle \mathrm{ACN} \rangle$ was proven to apply to random walks that behave like chains in which segments neither attract nor repulse each other (see Figure 1). In such chains their overall dimensions such as their radius of gyration grow

with the size of the chain as $\langle R_g \rangle = an^{\nu}$ where $\nu = 0.5$. Can the same type of scaling behavior also apply to protein backbones that are known to behave as collapsed walks with the exponent ν being close to 1/3? On the other hand it was postulated that in the maximally tightly packed, i.e. collapsed, walks $\langle$ACN$\rangle$ should scale according to a simple power law where the entanglement exponent B should be equal to 4/3 [6, 7, 8, 9, 10, 12]. Figure 4 presents a dot plot where ACN values measured for backbones of 2230 independent proteins are plotted against the number of $C_\alpha - C_\alpha$ segments in the respective protein structures. All points are fitted using either the formula $\langle \text{ACN} \rangle = \frac{3}{16} n \ln n + bn$ or the formula $\langle \text{ACN} \rangle = an^{4/3}$. Notice that in both cases we just have one free parameter and that in both cases the fit is excellent. Although the formula $\langle \text{ACN} \rangle = \frac{3}{16} n \ln n + bn$ gives a slightly better fit one can certainly say that both formulas describe very accurately the scaling behavior of $\langle$ACN$\rangle$ in natural proteins.

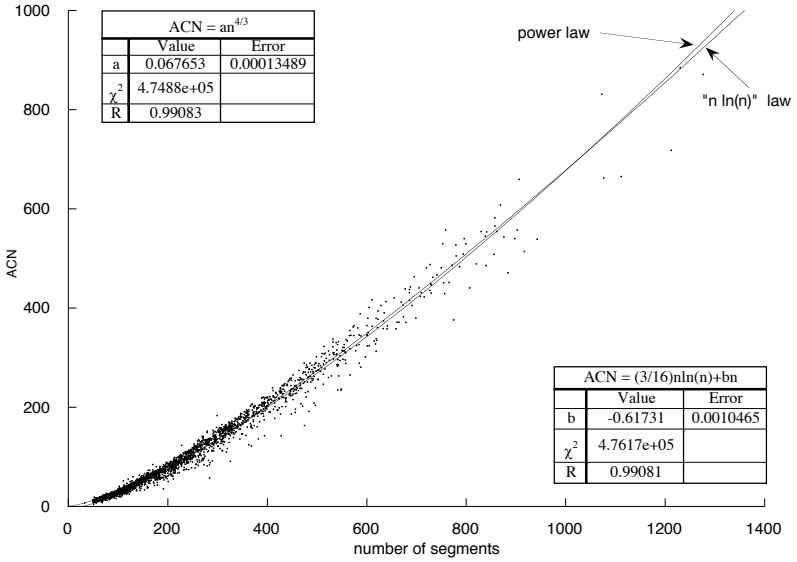

Figure 3: Scaling of ACN in natural protein structures. ACN values calculated for the backbones of 2230 independent proteins were fitted using two formulas: $\langle \text{ACN} \rangle = \frac{3}{16} n \ln n + bn$ and $\langle \text{ACN} \rangle = an^{4/3}$. The quality of the fit and the fitted value of respective free parameter are indicated in the insets.

4 Conclusions

We have compared here the scaling of ACN in ideal random walks in which segments neither repulse nor attract each other with the ACN scaling in backbones of proteins, which are known to behave like collapsed walks. It is known that the overall dimensions,

like the average radius of gyration or the average end-to-end distance, in ideal random walks scale with the chain length n like $n^{\frac{1}{2}}$ while overall dimensions in collapsed walks scale like $n^{\frac{1}{3}}$. Despite these fundamental differences of the scaling behavior of overall dimensions in these two types of walks the scaling of ACN seems to be quite similar in the both types of walks and can be accurately described by the formula $\langle \mathrm{ACN} \rangle = \frac{3}{16} n \ln n + bn$ where the free parameter b amounts to ca. -0.29 in the case of ideal random walks and -0.62 in the case of protein backbones. While the relation $\langle \mathrm{ACN} \rangle = \frac{3}{16} n \ln n + bn$ was analytically proven to apply to ideal random walks [13] there is no formal proof that the same relation should also hold for the collapsed walks. Therefore one should take into account that $\langle \mathrm{ACN} \rangle$ scaling in protein backbones can be also accurately described by a power law relation $\langle \mathrm{ACN} \rangle = an^{4/3}$ where the 4/3 exponent was predicted theoretically for tightly packed walks (). More statistical data will be needed especially extending to higher sizes of proteins to tell which scaling form applies better to $\langle \mathrm{ACN} \rangle$ of protein backbones. The relation $\langle \mathrm{ACN} \rangle = \frac{3}{16} n \ln n + bn$ applies also to ideal random polygons but the linear term is less important than in the case of ideal random walks and the parameter b amounts to ca. -0.1. Interestingly, the difference between $\langle \mathrm{ACN} \rangle$ for closed and open ideal random walks with the same number of segments n follows a linear relation $\langle \mathrm{ACN} \rangle_{\mathrm{closed}} - \langle \mathrm{ACN} \rangle_{\mathrm{open}} = \frac{3}{16} n + c$. We have also analyzed the scaling of $\langle \mathrm{ACN} \rangle$ for individual knot types K and observed that in each case the observed relation can be described by a formula $\langle \mathrm{ACN}(\mathcal{K}) \rangle = a(n - n_0) \ln(n - n_0) + b(n - n_0) + c$, where n_0 is the minimal number of equilateral segments needed to form the given knot type $\mathcal{K}$ and where the constant c approximates the ACN value of the ideal knot of type $\mathcal{K}$. The $\langle \mathrm{ACN}(\mathcal{K}) \rangle$ profiles show slower growth rates than the corresponding $\langle \mathrm{ACN} \rangle$ profiles of all closed walks grouped together. As the complexity of the knot type $\mathcal{K}$ increases, the coefficient a in our fitting formula decreases while the coefficient b increases. This scaling behavior of ACN in individual knot types causes the $\langle \mathrm{ACN} \rangle$ scaling profiles of each knot (with exception of unknot) to intersect the scaling profile of $\langle \mathrm{ACN} \rangle$ of all closed walks grouped together. The points of intersection define the equilibrium length of a given knot type i.e. the chain length at which a given knot type is not under- or over-knotted and would therefore show no tendency to increase or decrease its $\langle \mathrm{ACN} \rangle$ upon closure, equilibration and re-closure.

Acknowledgments. We thank G. Arteca, R. Kusner, K. Millet and W. Taylor for discussions. We also thank W. Taylor for providing us with the filtered set of coordinates of independent protein structures. This work was supported by Swiss National Science Foundation grants 31-68151.02 and 3100A0-103962 and Y. Diao is partially supported by NSF Grant #DMS-0310562. A preliminary version of this paper was presented at the conference 'Knots, random walks and biomolecules' (co-organized by J. H. Maddocks & A. Stasiak and sponsored by the Bernoulli Centre of the Swiss Federal Institute of Technology in Lausanne) held in Les Diablerets, Switzerland during 14-17 July 2003.

References

[1] C. C. Adams, The Knot Book. W.H. Freeman and Company, New York, 1994.

[2] G. A. Arteca. Overcrossing spectra of protein backbones: characterization of three-dimensional molecular shape and global structural homologies Biopolymers, **33**, 1829-1841, 1993.

[3] G. A. Arteca. Scaling behavior of some molecular shape descriptors of polymer chains and protein backbones Phys. Rev. E, 2417-2428, 1994.

[4] G. A. Arteca. Scaling regimes of molecular size and self-entanglements in very compact proteins Physical Review. E. Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics, **51**, 2600-2610, 1995.

[5] G. A. Arteca. Self-similarity in entanglement complexity along the backbones of compact proteins Phys. Rev. E, **56**, 4516-4520, 1997.

[6] G. A. Arteca. Analytical estimation for the entanglement complexity of a bond network J. Chem. Inf. Comp. Sci., **42**, 326-330, 2002.

[7] G. Buck. Four-thirds power law for knots and links Nature, **392**, 238-239, 1998.

[8] G. Buck and J. Simon. Thickness and crossing number of knots Topology and its Applications, **69**, 1-12, 1999.

[9] J. Cantarella, R. B. Kusner and J. M. Sullivan. Tight knot values deviate from linear relations Nature, **392**, 237-238, 1998.

[10] J. Cantarella, R. B. Kusner and J. M. Sullivan. On the minimum ropelength of knots and links Invent. math., **150**, 257-286, 2002.

[11] J. A. Calvo and K. C. Millett, in Ideal Knots, Eds. A. Stasiak, V. Katritch and L.H. Kauffman World Scientific, Singapore, 107-128, 1998.

[12] Y. Diao and C. Ernst. The complexity of lattice knots Topology and its Applications, **90**, 1-9, 1998.

[13] Y. Diao, A. Dobay, R. B. Kusner, K. Millett and A. Stasiak. The average crossing number of equilateral random polygons Journal of Physics A: Mathematical and General, **36**, 11561-11574, 2003.

[14] A. Dobay, J. Dubochet, K. Millett, P.-E. Sottas and A. Stasiak. Scaling behavior of random knots Proc. Natl Acad. Sci. USA, **100**, 5611-5615, 2003.

[15] M. Doi and S. F. Edwards, The Theory of Polymer Dynamics. Oxford University Press, Oxford, 1986.

[16] P. J. Flory, Principles of Polymer Chemistry. Cornell University Press, Ithaca, NY, 1953.

[17] P. Freyd, D. Yetter, J. Hoste, W. Lickorish, K. Millett and A. Ocneau. A new polynomial invariant of knots and links. Bull. Amer. Math. Soc., **12**, 239-246, 1985.

[18] P. G. de Gennes, Scaling Concepts in Polymer Physics. Cornell University Press, Ithaca, N.Y., 1979.

[19] P. Grassberger. Opacity and entanglement of polymer chains J. Phys. A: Math. Gen., **34**, 9959-9963, 2001.

[20] J.-Y. Huang and P.-Y. Lai. Crossings and writhe of flexible and ideal knots Phys. Rev. E, **63**, 21506, 2001.

[21] V. Katritch, J. Bednar, D. Michoud, R. G. Scharein, J. Dubochet and A. Stasiak. Geometry and physics of knots. Nature, **384**, 142-145, 1996.

[22] K. V. Klenin, A. V. Vologodskii, V. V. Anshelevich, D. A.M. and M. D. Frank-Kamenetskii. Effect of excluded volume on topological properties of circular DNA J. Biomol. Struct. Dyn., **5**, 1173-1185, 1988.

[23] M. A. Krasnow, A. Stasiak, S. J. Spengler, F. Dean, T. Koller and N. R. Cozzarelli. Determination of the absolute handedness of knots and catenanes of DNA Nature, **304**, 559-560, 1983.

[24] M. L. Mansfield, Are there knots in proteins? Nat Struct Biol, **1**, 213-214, 1994.

[25] E. Orlandini, M. C. Tesi, E. J. Janse Van Rensburg and S. G. Whittington. Asymptotics of knotted lattice polygons J. Phys. A: Math. Gen., **31**, 5935-5967, 1998.

[26] D. Poland and H. A. Scheraga, Theory of Helix-Coil Transitions in Biopolymers. Academic Press, New York, 1970.

[27] M. K. Shimamura and T. Deguchi. Anomalous finite-size effect for the mean-squared gyration radius of Gaussian random knots. J. Phys. A: Math. Gen., **35**, 241-246, 2002.

[28] A. Stasiak, J. Dubochet, V. Katritch and P. Pieranski, in Ideal Knots, Eds. A. Stasiak, S. Katritch and L.H. Kauffman World Scientific, Singapore, 1-19, 1998.

[29] A. Stasiak, V. Katritch, J. Bednar, D. Michoud and J. Dubochet. Electrophoretic mobility of DNA knots Nature, **384**, 122, 1996.

[30] W. R. Taylor. Multiple sequence threading: an analysis of alignment quality and stability J. Mol. Biol., **269**, 902-943, 1997.

[31] W. R. Taylor. A deeply knotted protein structure and how it might fold Nature, **406**, 916-919, 2000.

[32] A. Vologodskii, N. Crisona, B. Laurie, P. Pieranski, V. Katritch, J. Dubochet and A. Stasiak. Sedimentation and electrophoretic migration of DNA knots and catenanes J. Mol. Biol., **278**, 1-3, 1998.