Letters in Mathematical Physics **55:** 239–247, 2001.

Predicting Optimal Lengths of Random Knots

AKOS DOBAY,[1] PIERRE-EDOUARD SOTTAS,[1,2], JACQUES DUBOCHET[1] and ANDRZEJ STASIAK[1]
[1] *Laboratory of Ultrastructural Analysis, University of Lausanne, 1015 Lausanne, Switzerland. e-mail: {akos.dobay, jacques.dubochet, andrzej.stasiak}@lau.unil.ch*
[2] *Center for Neuromimetic Systems, Swiss Federal Institute of Technology, EPFL-DI, 1015 Lausanne, Switzerland. e-mail: pierre-edouard.sottas@epfl.ch*

(Received: 23 January 2001)

Abstract. In a thermally fluctuating long linear polymeric chain in a solution, the ends, from time to time, approach each other. At such an instance, the chain can be regarded as closed and thus will form a knot or rather a virtual knot. Several earlier studies of random knotting demonstrated that simpler knots show a higher occurrence for shorter random walks than do more complex knots. However, up to now there have been no rules that could be used to predict the optimal length of a random walk, i.e. the length for which a given knot reaches its highest occurrence. Using numerical simulations, we show here that a power law accurately describes the relation between the optimal lengths of random walks leading to the formation of different knots and the previously characterized lengths of ideal knots of a corresponding type.

Mathematics Subject Classifications (2000). 82D60, 46N30, 57M25, 82B41, 92C05.

Key words. knots, polymers, scaling laws, DNA, random walks, biophysics.

A random walk can frequently lead to the formation of knots and it has been proved that, as the walk becomes very long, the probability of forming nontrivial knots upon closure of such a walk tends to one [1, 2]. Many different simulation approaches were used to study random knotting [3–7]. Probably the most fundamental one is by simulation of ideal random chains where each segment of the chain is of the same length and has no thickness [4, 8]. In ideal random chains, the neighboring segments are not correlated with each other and thus show the average deflection angle of 90°. Ideal random chain behavior is interesting from the physical point of view as it reflects the statistical behavior of long polymer chains in the so-called melt phase and in θ solvents where the excluded volume effect vanishes [8]. Highly diluted polymer chains in θ solvents are unlikely to interact with each other and therefore, upon circularization, will form mainly knots rather than links. In thermally fluctuating long linear polymers, the ends of the same chain can, from time to time, come into a close vicinity of each other. This can lead to a cyclization of the polymer whereby the end closure frequently traps a nontrivial knot on the chain. By studying knotting in simulated ideal random chains, we can gain insight into knotting of real polymer chains in θ solvents and in the dense melt phase frequently used for the preparation of such synthetic polymeric materials as fabrics, paints, or adhesives [9]. However, ideal chains do not reflect the behavior of real polymer chains in good solvent. Intramolecular interactions cannot be neglected in these conditions, but can be well approximated by introducing an effective diameter. When such a constraint is introduced into simulated chains, one can also model knotting of polymers in good solvents like, for example, knotting of DNA molecules in typical reaction buffers used for biochemical experiments [4]. Our simulations can be adjusted to both situations and we shall present here the results for random chains with and without an effective diameter.

Several earlier studies of random knotting showed that simpler knots reach a maximum of their occurrence for a shorter length of random walks than is required for the formation of more complex knots [5, 6, 10]. In considering the equilibrium ensemble of closed walks, these studies showed that the relative frequency of the occurrence of each type of knot first increases with the length of the chain, then passes through a maximum and finally decreases exponentially at very long chains. However, these earlier studies did not attempt to establish a relationship between the type of knot and the optimal length of the random walk leading to the maximal occurrence of this knot. If we consider a thermally fluctuating polymer with ends that can stick to each other with energy much smaller than kT, then, from time to time, these ends will stay in contact for a short period and at that moment the polymer will form a trivial or nontrivial knot. In this study, we characterize statistical ensembles of fluctuating linear polymers in order to find specific lengths (expressed in the number of statistical segments) at which a given type of knot or rather a virtual knot reaches its highest occurrence.

Recently, we have characterized ideal geometric configurations of knots corresponding to the shortest trajectories of flexible cylindrical tubes with uniform diameter to form a given knot [11]. The ratio of length to diameter of the tube forming an ideal configuration of a given knot is topologically invariant and here we call it the length of ideal knots. Ideal knots turned out to be good predictors of statistical behavior of random knots. So, for example, the writhe of the ideal configuration of a given knot was equal to the average writhe of a thermally fluctuating polymer forming a given random knot [11]. We also showed that electrophoretic migrations of various types of knotted DNA molecules of the same molecular weight or their expected sedimentation constants were almost proportional to the length of the corresponding ideal knots [12, 13]. Therefore we decided to check whether the length of ideal knots is related to the length of ideal random chains for which different knots reach their highest occurrence. To this aim, we used the following simulation procedure. 2×10^9 random walks of 170 segments were started and each time the growing end approached the starting end at a distance smaller than the length of one segment, the configuration was saved, upon which the walk was continued for the remaining number of steps. Each vector (segment) of the chain was randomly chosen from uniformly distributed vectors pointing from the center to the surface of the unit sphere. Thus, some of the random walks showed one or more approaches of the growing and starting ends and we collected 2×10^9 random walks for every number of segments between 5 and 170. Each saved configuration with nearby ends was then closed with a connecting segment and the type of the formed knot was determined by the calculation of its Alexander polynomial [7, 14–16].

For random linear walks to efficiently form different knots, a compromise has to be met between the length optimizing their closure and the length which is sufficient to form a knot of a given type. The present analysis differs from earlier studies [4, 5, 10] where the statistics were based only on equilibrium knotting of closed walks. In our case, we consider the formation of knots through the approach of the terminal segments of linear chains. Therefore, not only closed chains, but also linear chains are taken into account in our statistics.

Figure 1 shows the occurrence profiles of different knots with up to six crossings as a function of the length of the random walk which leads to the formation of these knots. It is visible that trefoil knots show their highest occurrence for 25 ± 1 segments, while 4_1 knots form most frequently for 42 ± 1 segments. The formation of more complicated knots happens much less frequently than simpler knots, therefore in the insert in Figure 1 a change of scale is applied to better visualize the occurrence of more complicated knots. We observed that the obtained probability values for different knots can be well fitted with the function

$$P_k(N) = a(N - N_0)^b \exp\left(\frac{-N^c}{d}\right), \tag{1}$$

where, for each knot, a, b and d are free parameters, c is an empirical constant equal to 0.18, N_0 is the minimal number of segments required to form a given type of knot [17] without the closing segment, and N is the number of segments in the walk. Our fitting function was adapted from Katritch *et al.* [18] but modified to take into account the probability of cyclization. Table I lists the positions of maximal occurrence for the analyzed types of knots. In order to concentrate on the position of the maximum for different knots and not on their actual probability values, we decided to present probability profiles for each knot upon normalizing them by assigning a value 1 to the respective maximum of probabilities. Figure 2 presents normalized probability profiles for the analyzed knots. It is visible that different knots with the same minimal number of crossing now show quite similar types of profiles (e.g. knots 5_1 and 5_2), whereby the differences in the position of maximum between knots with a different minimal number of crossings can be easily perceived. It may be surprising that we observed here such a short optimal length for analyzed knots, while earlier studies showed that several hundred segments are needed to observe maximum occurrence of a given knot among closed walks of a given size [5, 10, 19]. This is simply due to the fact that our system takes into account the probability of cyclization.

In Figure 3, we show the relation between the optimal length of random knots and the length of the corresponding ideal knots. This relation is well approximated by a

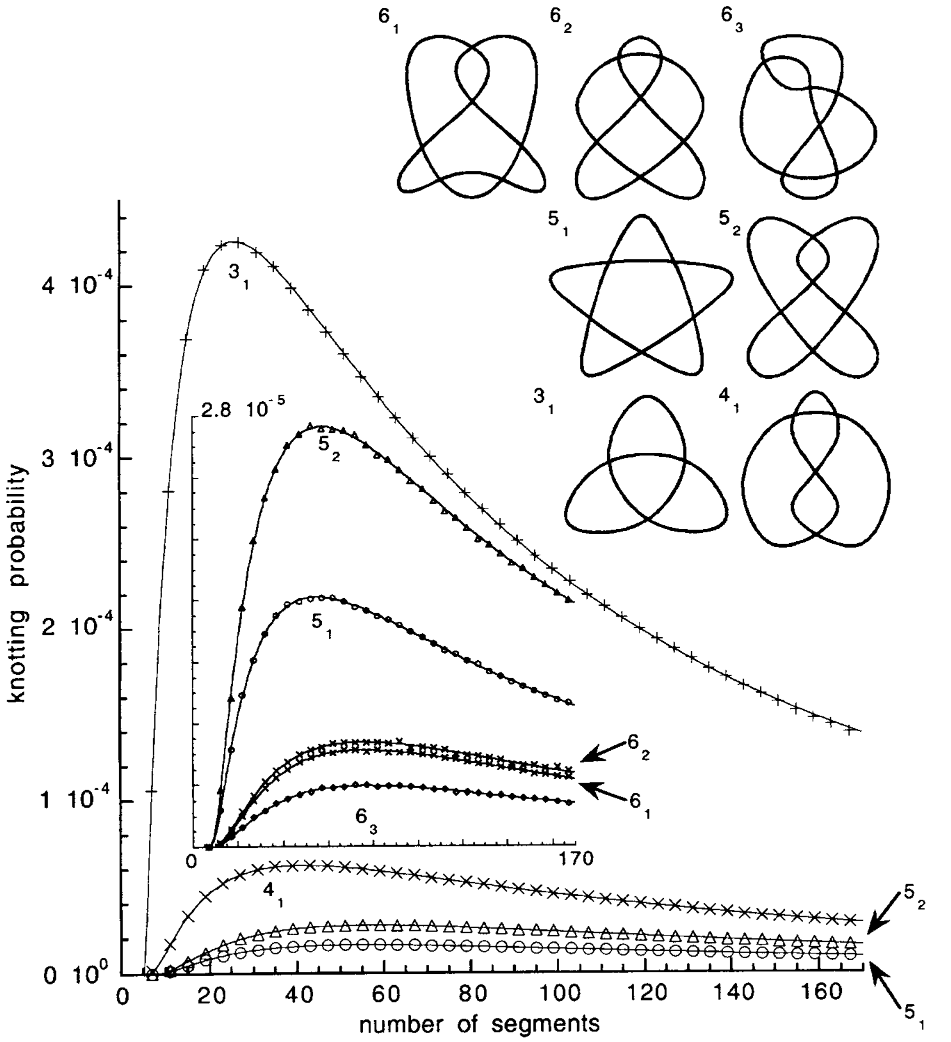

Figure 1. Probability of forming a given knot among all random walks of a given size is plotted as a function of the number of segments in the walk. Note the change of the scale between the main panel and the insert. Diagrams of the corresponding knots are drawn to visualize the differences between analyzed types of knots. The notations accompanying the diagrams correspond to those in standard tables of knots [21], where the main number indicates the minimal number of crossings possible for this knot type and the index indicates the tabular position amongst the knots with the same minimal crossings number. The diagrams represent left-handed and right-handed knots at the same time and therefore do not show the underlying and overlying segments, but since these knots are alternating, it is easy to recover the proper topology by applying a rule that under- and over-passages alternate. Formed knot types were recognized by computing their Alexander polynomial. Since Alexander polynomials do not distinguish between left-handed and right-handed knots of the same type, we have to group them together and therefore the diagrams of the knots do not show the handedness. This polynomial has sometimes the same value for different knots like, for example, knots 6_1 and 9_{46} [22]. However, within groups of knots with the same Alexander polynomial, more complicated knots have such a low occurrence that their effect on the position of the maximum of the simplest knot within the group can be neglected.

Table I. Optimal sizes O_s of random walks (in number of segments) leading to the formation of corresponding knots, the length/diameter ratio L_D values of ideal configurations of these knots K_n [23] and the values of the parameters of the function (1) used in the fits of the observed probabilities (see Figure 1). The presented data are limited to knots with up to 7 crossings since obtained by us, statistics for more complex knots is less good.

K_n	O_s	L_D	a	b	d	N_0
3_1	25 ± 1	16.33	$(1.84 \pm 0.01) \times 10^{-1}$	1.57 ± 0.01	0.165 ± 0.001	5
4_1	42 ± 1	20.99	$(0.45 \pm 0.01) \times 10^{-1}$	2.24 ± 0.01	0.134 ± 0.001	6
5_1	54 ± 2	23.55	$(1.28 \pm 0.02) \times 10^{-2}$	2.65 ± 0.01	0.121 ± 0.001	7
5_2	56 ± 2	24.68	$(2.31 \pm 0.04) \times 10^{-2}$	2.77 ± 0.01	0.118 ± 0.001	7
6_1	74 ± 2	28.30	$(0.78 \pm 0.03) \times 10^{-2}$	3.75 ± 0.02	0.095 ± 0.001	7
6_2	75 ± 2	28.47	$(0.74 \pm 0.03) \times 10^{-2}$	3.67 ± 0.02	0.096 ± 0.001	7
6_3	76 ± 2	28.88	$(0.39 \pm 0.02) \times 10^{-2}$	3.69 ± 0.02	0.097 ± 0.001	7
7_1	89 ± 3	30.70	$(5.91 \pm 0.69) \times 10^{-4}$	3.95 ± 0.06	0.093 ± 0.001	8
7_2	92 ± 3	32.41	$(1.72 \pm 0.16) \times 10^{-3}$	4.33 ± 0.05	0.086 ± 0.001	8
7_3	92 ± 3	31.90	$(9.43 \pm 0.85) \times 10^{-4}$	4.03 ± 0.05	0.092 ± 0.001	8
7_4	97 ± 3	32.53	$(5.55 \pm 0.67) \times 10^{-4}$	4.25 ± 0.06	0.087 ± 0.001	8
7_5	97 ± 3	32.57	$(1.32 \pm 0.09) \times 10^{-3}$	4.24 ± 0.04	0.089 ± 0.001	8
7_6	98 ± 3	32.82	$(1.71 \pm 0.14) \times 10^{-3}$	4.36 ± 0.04	0.086 ± 0.001	8
7_7	95 ± 3	32.76	$(8.82 \pm 0.96) \times 10^{-4}$	4.31 ± 0.06	0.087 ± 0.001	8

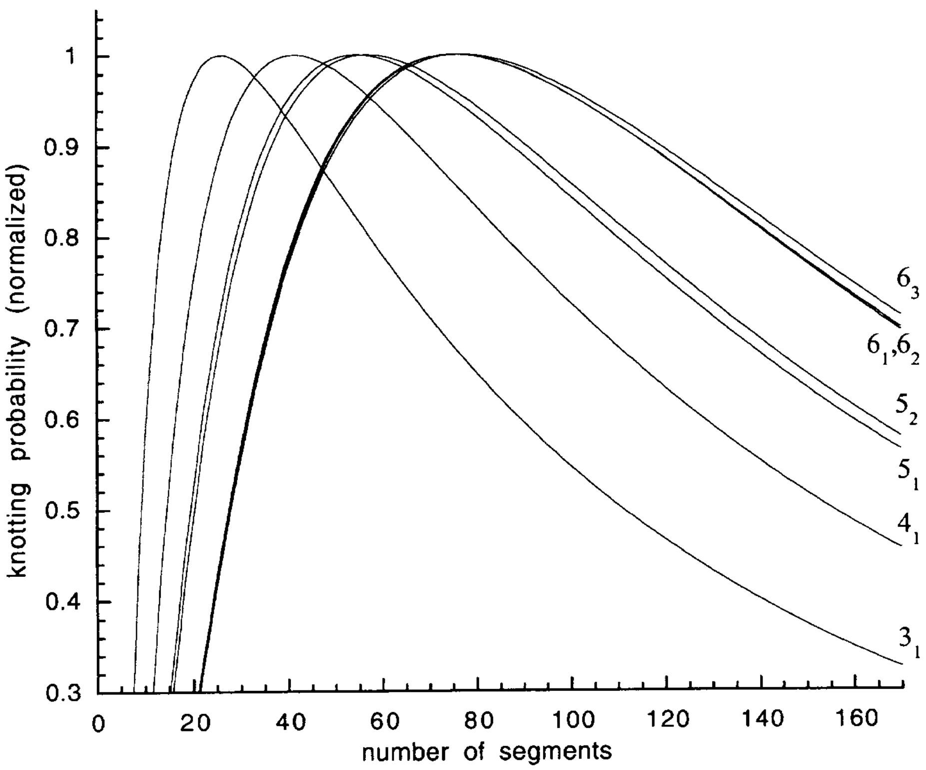

Figure 2. Normalized probability profiles for the analyzed knots.

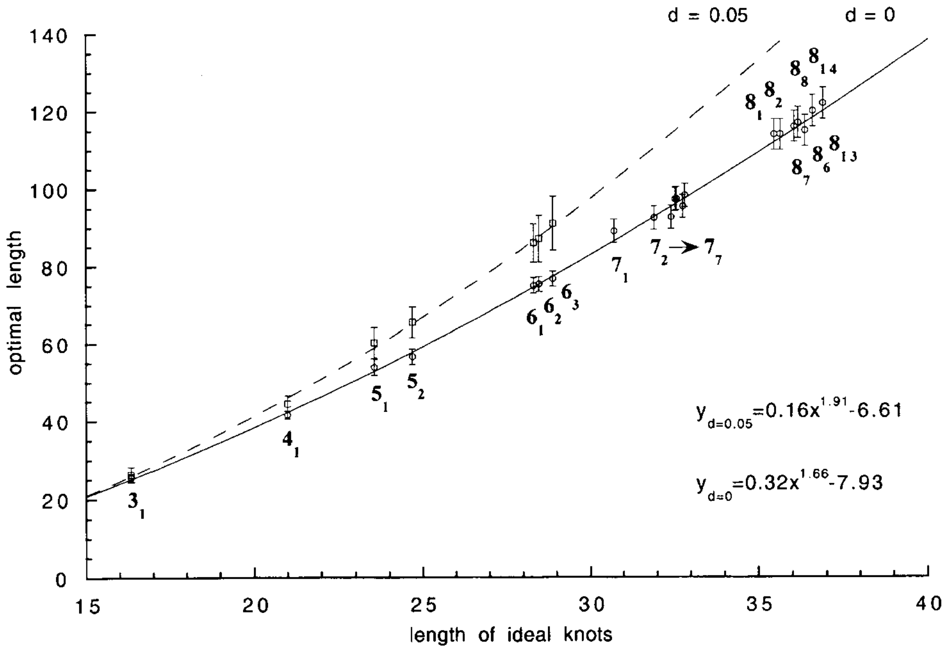

Figure 3. Relationship between the length of the ideal geometric representations of knots [23] and positions of maximal occurrence for the corresponding random knots. The lower curve: the optimal length of random knots with an effective diameter set to zero. The simulation data for the knots with up to seven crossings were fitted with a power law function and the best fit curve was extrapolated. Data points for eight crossing knots for which we obtained good statistics coincide with the extrapolated curve. The upper curve: data points of maximal occurrence of knots for random chains with an effective diameter set to 0.05 to the segment length. In both cases, a power law function adequately describes the relation between the optimal length of the random knots and the length of ideal knots of a given type. Best-fit parameters for both cases are indicated.

power law function. Upon fitting the free parameters of this function into the simulation data obtained for the knots with up to 7 crossings, we decided to check if, by knowing the length of ideal configurations of more complicated knots, we can predict positions of the maximum of occurrence for the corresponding random knots. As the statistics of random knotting becomes poor for knots with an increasing crossing number, we limited verifications of our predictions to these knots with eight crossings which, at their maxima of occurrence, were represented more than 500 times out of 2×10^9 random walks with a given number of segments. Analysis of our simulation data (Figure 3) positively verified our predictions for optimal sizes of random walks leading to the formation of these knots.

As already mentioned, ideal random chains have no thickness and this causes them to reflect the behavior of polymers in the melt phase where thin polymers have practically no exclusion volume [7, 8]. However, when polymers are suspended in a good solvent, like DNA in an aqueous solution, the exclusion volume of polymers becomes significant and this strongly decreases the probability of forming knots [7]. It was observed that the higher the effective diameter of the polymer, the lower the probability of forming knots by random cyclization [4, 7, 19]. We decided therefore to investigate whether a power law relation between the length of ideal knots and the optimal length of randomly knotted chains also holds for chains with an exclusion volume. To this aim, from our original set of 2×10^9 ideal random walks for every segment length from 5 to 100, we selected the walks which never showed a closer approach between any pair of nonneighboring segments than the considered effective diameter (terminal segments of the chain are considered as neighboring ones). Subsequently, we analyzed all configurations with approached ends for the types of formed knots and calculated the probabilities of various knots among all random chains which fulfilled the criteria of a given effective diameter. We observed that, as the effective diameter grows, the probability of forming various knots decreases and the positions of the maximum move toward longer chains. Figure 3 (dashed line) shows the relation between the length of ideal knots and the optimal length of the corresponding random knots formed by chains with the effective diameter being set to 0.05 of the segment length. The effective diameter 0.05 corresponds to this of diluted solutions of DNA molecules in about 100 mM NaCl [4]. In the case of DNA, each segment in the random chain corresponds to 300 base pair long region [20]. It is visible that the data can be again approximated by a power law function. The fact that the lengths of ideal knots show a correlation with the optimal sizes of the corresponding random knots formed by chains with a given effective diameter, provides another example that ideal knots are good predictors of the physical behavior of real knots [11].

Post factum, it might seem to be obvious that knots requiring a greater length of rope to tie them should require a longer random walk to reach their highest occurrence. However, until recently the minimal length of rope required to tie a given knot was not known. In addition, the relation between the optimal length of random walk producing a given knot and the length of an ideal knot had not yet been proposed in the literature. On the other hand, a simple expectation would dictate that the shorter the length of ideal knot, the higher the probability of its formation. So, for example, trivial knots are more frequent than trefoils and these are more frequent than 4_1 knots. However, this does not hold for 5_1 and 5_2 knots. Ideal knot 5_1 is slightly shorter than ideal 5_2 knot (which is consistent with the optimal size of the random walks leading to the formation of corresponding knots), but 5_2 knot formation by random walks is circa twice more frequent than the formation of a 5_1 knot. Therefore, the values of random knot probabilities (in contrast to the positions of the maxima) are not related by a relatively simple growing function to the values of length of the corresponding ideal knots.

What can be the possible applications resulting from the determination of the optimal size of knots? For chemical cyclization of polymer chains, we can use a linear polymer of a specific length and thus promote the formation of a given type of knot. Materials with interesting properties could be formed in this way.

Acknowledgements

We thank Alexander Vologodskii and Vsevolod Katritch for making available their routine for the calculation of Alexander polynomials. We also thank Piotr Pieranski for numerous discussios. This work was supported by the Swiss National Science Foundation (31-58841.99 and 3152-061636.00).

References

1. Sumners, D. W. and Whittington, S. G.: Knots in self avoiding walks, *J. Phys. A: Math. Gen.* **21** (1988), 1689–1694.
2. Matthews, R.: Knotted rope: a topological example of Murphy's law, *Math. Today* **33** (1997), 82–84.
3. Frank-Kamenetskii, M. D., Lukashin, A. V. and Vologodskii, A. V.: Statistical mechanics and topology of polymer chains, *Nature* **258** (1975), 398–402.
4. Rybenkov, V. V., Cozzarelli, N. R. and Vologodskii, A. V.: The probability of DNA knotting and the effective diameter of the DNA double helix, *Proc. Nat. Acad. Sci. U.S.A.* **90** (1993), 5307–5311.
5. Deguchi, T. and Tsurusaki, K.: A statistical study of random knotting using the Vassiliev invariants, *J. Knot Theory Ramifications* **3** (1994), 321–353.
6. Janse van Rensburg, E. J., Orlandini, E. Sumners, D. W., Tesi, M. C. and Whittington, S. G.: *J. Knot Theory Ramifications* **6** (1997), 31–44.
7. Koniaris, K. and Muthukumar, M.: Knottedness in ring polymers, *Phys. Rev. Lett.* **66** (1991), 2211–2214.
8. de Gennes, P. G.: *Scaling Concepts in Polymer Physics*, Cornell Univ. Press, Ithaca, 1979.
9. Alper, J. and Nelson, G. L.: *Polymeric Materials: Chemistry for the Future*, Amer. Chem. Soc., Washington, DC, 1989.
10. Orlandini, E., Janse van Rensburg, E. J., Tesi, M. C. and Whittington, S. G.: Entropic exponents of knotted lattice polygons, In: S. G. Whittington and D. W. Sumners (eds). *Topology and Geometry in Polymer Science*, Springer, New York, 1998, pp. 9–21.
11. Katritch, V., Bednar, J., Michoud, D., Scharein, R. G., Dubochet, J. and Stasiak, A.: Geometry and physics of knots, *Nature (London)* **384** (1996), 142–145.
12. Stasiak, A., Katritch, V., Bednar, J., Michoud, D. and Dubochet, J.: Electrophoretic mobility of DNA knots, *Nature (London)* **384** (1996), 122.
13. Vologodskii, A., Crisona, N., Laurie, B., Pieranski, P., Katritch, V., Dubochet, J. and Stasiak, A.: Sedimentation and electrophoretic migration of DNA knots and catenanes, *J. Molec. Biol.* **278** (1998), 1–3.
14. Alexander, J. W.: Topological invariants of knots and links, *Trans. Amer. Math. Soc.* **30** (1928), 275–306.
15. Frank-Kamenetskii, M. D. and Vologodskii, A. V.: Topological aspects of polymer physics: theory and its biophysical applications, *Soviet Phys. Usp.* **24** (1981), 679–696.
16. Adams, C. C.: *The Knot Book*, W. H. Freeman, New York, 1994.
17. Calvo, J. A. and Millett, K. C.: Minimal edge piecewise linear knots, In: A. Stasiak, V. Katritch and L. H. Kauffman (eds), World Scientific, Singapore, 1998, pp. 107–128.
18. Katritch, V., Olson, W. K., Vologodskii, A., Dubochet, J. and Stasiak, A.: Tightness of random knotting, *Phys. Rev. E* **61** (2000), 5545–5549.
19. Deguchi, T. and Tsurusaki, K., In: S. Suzuki (ed.) *Lectures at Knots 96*, World Scientific, Singapore, 1997, p. 95.

20. Vologodskii, A. V.: *Topology and Physics of Circular DNA*, Physical approaches to DNA, CRC Press, Boca Raton, 1992.
21. Rolfsen, D.: *Knots and Links*, Publish or Perish, Berkeley, CA, 1976.
22. de la Harpe, P.: Introduction to knot and link polynomials, In: A. Amann, L. Cederbaum and W. Gans (eds), *Fractals, Quasicrystals, Chaos, Knots and Algebraic Quantum Mechanics*, Kluwer Acad. Publ., Dordrecht, 1988, pp. 233–263.
23. Stasiak, A., Dubochet, J., Katritch, V. and Pieranski, P.: Ideal knots and their relation to the physics of real knots, In: A. Stasiak, V. Katritch and L. H. Kauffman (eds), *Ideal Knots*, World Scientific, Singapore, 1998, pp. 1–19.