# Rknots, an *R* package for the topological analysis of knotted biological polymers

**Federico Comoglio**[1*], **Maurizio Rinaldi**[2]

1. Department of Biosystems Science and Engineering, Swiss Federal Institute of Technology (ETH) Zurich, Basel, Switzerland
2. Department of Chemical, Food, Pharmaceutical and Pharmacological Sciences (DiSCAFF), University of Piemonte Orientale "A. Avogadro", Novara, Italy

*Contact author: federico.comoglio@bsse.ethz.ch

**Keywords:** **Rknots**, knotted polymers, computational knot theory, topology, proteins

The topological study of biological polymers has led to important insights into their structural properties and evolution Virnau et al. (2006). From a topological point of view polymers can be naturally modeled as sequences of 3D points, i.e. open polygonal paths. Their closure generates classical objects in topology called knots. Boosted by Taylor's work Taylor (2000), the characterization of knotted proteins is a subject of increasing interest in both experimental and computational biology due to the close structure-function relationship and reproducible entangled folding. The exponential growth of the total number of structures deposited into the Protein Data Bank (PDB) requires dedicated computational high-throughput methods able to deal with a large amount of data. These methods combine a structure reduction scheme of a protein backbone model with the computation of a knot invariant. Being easy to compute, the Alexander polynomial represents the current default choice Kolesov et al. (2007). Although this choice is primarily supported by the fact that the knots so far identified in proteins are of the simplest types, the Alexander polynomial is not able to discern knots chirality and its limited power is not optimal to develop a generalized framework aiming to describe topological properties of 3D structures.

Recently, we developed a topological framework for the computation of the HOMFLY polynomial Freyd et al. (1985), a more powerful invariant able to define knots chirality. By screening the entire PDB, we obtained an up-to-date table of knotted proteins that also includes two newly detected right-handed trefoil knots in recently deposited protein structures Comoglio and Rinaldi (2011).

However, the application of our framework is not limited to proteins and in order to provide an open-source package to the scientific community working in the field, we developed and here we present **Rknots**, an *R* package that includes functions and utilities to process and topologically explore polymers along with dedicated utilities for working with protein structures. Among the implemented functions, this first release includes PDB entry import or fetching (exploiting the flexibility of the *R* package **bio3d**), structure reduction via the Alexander-Briggs algorithm Alexander and Briggs (1926) and the MSR (Minimal Structure Reduction algorithm) Comoglio and Rinaldi (2011), topological invariants and linking number computation, 3D and knot diagram graphics function along with a knots and links dataset repository and utilities.

We welcome external contributions to **Rknots** in order to further extend the package functionalities and provide a general purpose tool to identify knots in three-dimensional structures.

## References

Alexander, J.-W. and G.-B. Briggs (1926). On types of knotted curves. *Ann of Math 28*, 562–586.

Comoglio, F. and M. Rinaldi (2011). A topological framework for the computation of the homfly polynomial and its application to proteins. *PLoS ONE in the press*.

Freyd, P., D. Yetter, J. Hoste, W.-B.-R. Lickorish, K. Millett, and A. Ocneanu (1985). A new polynomial invariant of knots and links. *Bull Amer Math Soc (NS) 12*, 239–246.

Kolesov, G., P. Virnau, M. Kardar, and L.-A. Mirny (2007). Protein knot server: detection of knots in protein structures. *Nucleic Acid Res 35*, W425–W428.

Taylor, W.-R. (2000). A deeply knotted protein structure and how it might fold. *Nature 406*, 916–919.

Virnau, P., L.-A. Mirny, and M. Kardar (2006). Intricate knots in proteins: Function and evolution. *PLoS Comput Biol 2*, e122.