A chain is knotted if it does not disentangle after being pulled from Web servers for protein knots analysis (Kolesov et al., 2009). The topological study of knotted biological polymers is an active informal introduction to the subject). However, the vast majority of deals with closed structures (see Supplementary Material for an underlying framework cannot be generalized to cope with more complex structures and it cannot be customized.

In this note we present Rknots, an R package combining a generalized framework for the topological analysis of knotted biopolymers with the benefits of R programming. Different structures can be analyzed with a simple syntax and methods have been implemented accounting for modularity. Rknots requires a standard R installation and depends on bio3d (Rent et al., 2009) available from http://mcrcammon.acsd.edu/~bgrant/bio3d/, rgl [Adler, D. and Murdoch, D. (2011) rgl: 3D visualization device system (OpenGL)] and r sympy [Grothendieck, G. et al. (2010) rSymPy: R interface to SymPy computer algebra system. R package version 0.2-1.1]. In the following we will provide an overview of the available methods. A case study on a knotted protein will be used as an example. Additional examples can be found in the package manual and vignette at http://cran.r-project.org/web/packages/Rknots

Structural bioinformatics

Rknots: topological analysis of knotted biopolymers with R

Federico Comoglio and Maurizio Rinaldi

1 Department of Biosystems Science and Engineering, ETH Zurich, Basel, CH-4058, Switzerland. 2 Life Science Zurich PhD Program on Systems Biology and 3 Dipartimento di Scienze del Farmaco di Novara, University of Piemonte Orientale “A. Avogadro”, Novara, 28100, Italy

ABSTRACT

Motivation: Rknots is a flexible R package providing tools for the detection and characterization of topological knots in biological polymers. The package is well documented and provides a simple syntax for data import and preprocessing, structure reduction, topological analyses and 2D and 3D visualization. Remarkably, Rknots is not limited to protein knots and allows researchers from interdisciplinary fields to analyze different topological structures and to develop simple yet fully custom pipelines.

Availability: Rknots is distributed under the GPL-2 license and is available from the CRAN (the Comprehensive R Archive network) at http://cran.r-project.org/web/packages/Rknots

Supplementary Information: Supplementary data are available at Bioinformatics online.

Received on December 24, 2011; revised on February 27, 2012; accepted on March 31, 2012

1 INTRODUCTION

The topological study of knotted biological polymers is an active interdisciplinary field of research (Bölinger et al., 2008) and Marenduzzo et al., 2009). In polymers, knots influence both material properties and polymer chain dynamics (Koniaris and Muthukumar, 2009). Informs, knots are closed curves in 3D and links are collections of non-intersecting knots. Hence, knot theory deals with closed structures (see Supplementary Material for an informal introduction to the subject). However, the vast majority of biological or synthetic polymers are open chains. In this context, the definition of knot is relaxed and transferred to open curves. A chain is knotted if it does not disentangle after being pulled from both ends.

Nature tends to avoid knots. Knotted protein backbones are rare (Lavlo, 2003; Virnau et al., 2006) and the physical mechanisms governing their formation is largely unknown. Intriguing experimental results unraveling protein knots properties have recently been reported (Savoy et al., 2011; Virnau et al., 2013) and in certain viral capsids there are evidences of mechanisms preventing knotted DNA formation (Burnier et al., 2008). In order to understand structural properties of knotted polymers, rigorous yet simply to use, generalized computational methods are required. Web servers for protein knots analysis (Kolesov et al., 2009; Lai et al., 2007) nicely accomplished this for proteins. However, the underlying framework cannot be generalized to cope with more complex structures and it cannot be customized.

In this note we present Rknots, an R package combining a generalized framework for the topological analysis of knotted biopolymers with the benefits of R programming. Different structures can be analyzed with a simple syntax and methods have been implemented accounting for modularity. Rknots requires a standard R installation and depends on bio3d (Rent et al., 2009) available from http://mcrcammon.acsd.edu/~bgrant/bio3d/, rgl [Adler, D. and Murdoch, D. (2011) rgl: 3D visualization device system (OpenGL)] and r sympy [Grothendieck, G. et al. (2010) rSymPy: R interface to SymPy computer algebra system. R package version 0.2-1.1]. In the following we will provide an overview of the available methods. A case study on a knotted protein will be used as an example. Additional examples can be found in the package manual and vignette at http://cran.r-project.org/web/packages/Rknots

2 APPROACH

Proteins can be loaded in .pdb format from the file system or by fetching the Protein Data Bank (PDB; Berman et al., 2000) and they undergo a dedicated preprocessing (see Supplementary Material). Coordinates are then stored in the S4 Knot class (see the package vignette for details) and the following workflow applies afterwards. First, open chains are closed and a knot diagram is obtained through the here proposed principal component analysis projection (PCAP) algorithm (see Supplementary Material). Figure 1B-D illustrates the results of a simulation on 1000 proteins sampled randomly from the PDB in comparison to a set of generic projections or to standard projection (see Supplementary Material for details). Second, the structure is reduced to the minimal set of points topologically equivalent to the original structure, by applying a reduction algorithm. This step removes structural redundancies, speeding up downstream computations. Two structure reduction algorithms, Alexander-Briggs (Alexander and Briggs, 1926) and MSR (Comoglio and Rinaldi, 2011) are implemented in Rknots. The former has been proposed for the first time in polymers knot theory by Koniaris and Muthukumar (Koniaris and Rinaldi, 2011). We also provide an extension of this algorithm for links.

Finally, knots can be detected in the minimal structure by computing a topological invariant, generally a polynomial. Protein knots discovered so far are among the simplest types (Bölinger et al., 2010; Virnau et al., 2006). The Alexander polynomial is sufficient to...
The following example illustrates a typical Rknots session for the Supplementary Material for details.

The open chain is closed and projected using the closeAndProject function. The knot diagram can then be visualized with plot:

```r
chain <- closeAndProject(chain)
plot(chain, lwd = 1.5)
```

Then, the knot type is determined by computing a polynomial invariant with the computeInvariant function. By setting invariant = "HOMFLY", the HOMFLY polynomial is returned:

```r
computeInvariant(chain, invariant = "HOMFLY")
```

Finally, further information on the knot type can be obtained with the function getKnotType (see the package vignette for details).

## 4 Conclusions

Rknots is the first package providing generalized tools for the study of knotted biopolymers. The modularity of the package allows integration in custom pipelines. We encourage contributions from other members of the research community.

### Funding

This work was partially funded by the Italian Ministry of Education, University and Research (MIUR).

### Conflict of Interest

None declared.

### References

A detailed list of references can be found in the package vignette. Résumé

A straightforward method to analyze knots in biopolymers is the computation of their complexity. The HOMFLY polynomial is a powerful tool for this purpose, as it can detect all possible knot types. In this study, we have applied this method to analyze the complexity of the protein 2k0a. The results show that this protein is knotted, and the HOMFLY polynomial is a useful tool for this purpose.

### Technical Details

The Rknots package is available on CRAN and can be installed using the command `install.packages("Rknots")`. The package vignette provides detailed instructions on how to use the package.

For more information, please visit the package website at https://rkn.cat.

---

**Authors:**


**Acknowledgments:**

This work was supported by the Italian Ministry of Education, University and Research (MIUR).