BIOINFORMATICS APPLICATIONS NOTE Vol. 18 no. 7 2002
Pages 1011–1012

TangleSolve: topological analysis of site-specific recombination

*Yuki Saka and Mariel Vázquez**

Mathematics Department, U.C. Berkeley, Berkeley, CA 94720-3840, USA

Received on November 29, 2001; revised on January 22, 2002; accepted on January 29, 2002

ABSTRACT

Summary: TangleSolve is a program for analysing site-specific recombination using the tangle model. The program offers an easy-to-use graphical user interface and a visualization tool. Biologists working in topological enzymology can use this program to compute and visualize site-specific recombination mechanisms that accommodate their experimental data. TangleSolve can also prove useful as a teaching aid for mathematical biology and computational molecular biology courses.
Availability: http://bio.math.berkeley.edu/TangleSolve/
Contact: mariel@math.berkeley.edu

DNA manipulations such as cutting and rearranging play a central role in many biological processes such as viral DNA integration into, or excision from, a host genome. Enzymes such as site-specific recombinases can fill the role of molecular scissors. Their local action is characterized by a simple process where two short regions of DNA, called recombination sites, are put close together, cut at a cleavage site, have their DNA ends exchanged, and are rejoined in a new recombinant form. Despite the apparent simplicity of the reaction in the close vicinity of the recombination sites, some site-specific recombinases acting on circular substrates yield spectra of knotted and/or linked (catenated) molecules, thus suggesting a non-trivial global component of the enzymatic event that is affected by the geometry and topology of the DNA substrate.

Knotted and linked products of recombination can be analysed using the mathematical theory of knots and tangles. In the tangle model (reviewed in Sumners *et al.*, 1995) the topology and geometry of the DNA that are relevant to the recombination reaction are analysed using several tangles, which represent comparatively localized spatial compartments that contain two DNA strands (illustrated as boxes in Figure 1C). Tangles, knots and links (Figure 1A') are topological objects. This mathematical tool translates the site-specific recombination event into a system of two or more tangle equations (Figure 1C). Under appropriate assumptions, and using tools in low-dimensional topology (Ernst and Sumners, 1990) the system can be solved numerically, thereby specifying, in a mathematically rigorous fashion, all possible mechanisms that account for given experimental data.

TangleSolve is an interactive application which implements the mathematical tangle model, and offers an easy-to-use graphical interface for analysing and visualizing recombination mechanisms. It is a Java standalone program and web-based applet. The user interface consists of three panes: input, selection, and display. The user specifies whether or not processive recombination, involving multiple consecutive reactions, occurs. The experimental data is entered in the input pane (Figure 1A). The user is asked to enter all the available topological information about substrate and recombination products. Information can be submitted using the mathematical notation of the specific knot or link type, or by selecting substrate and product(s) from a knot table (Figure 1A'), or simply by entering their crossing number(s). The program then solves the system(s) of tangle equations corresponding to the experimental data. Output solutions are listed in the selection pane (Figure 1B). The user can select one of the solutions that will then be viewed in the display pane (Figure 1C).

The solution algorithm is based on recently developed results in low dimensional topology (reviewed in Darcy, 2001 and in Vázquez, 2000). The experimental data consist of a list of two or more knot types. Most knots and links seen in biology are 4-plats (defined as closed braids in four strands in Figure 1A'). TangleSolve assumes that substrates and products of recombination are 4-plats; that the two recombination sites are enclosed inside a parental tangle, visualized as a box with two horizontal arcs (tangle (0), Figure 1C); and that upon recombination the parental tangle is converted into a recombinant tangle consisting of a horizontal row of twists, while the other intervening tangles remain unchanged. All solutions to the equations that are rational, or sums of rational, tangles are found. Rational tangles are the most appropriate to model enzymatic actions since they are easily created and 'unwound' by smooth deformations of the arcs.

*To whom correspondence should be addressed.

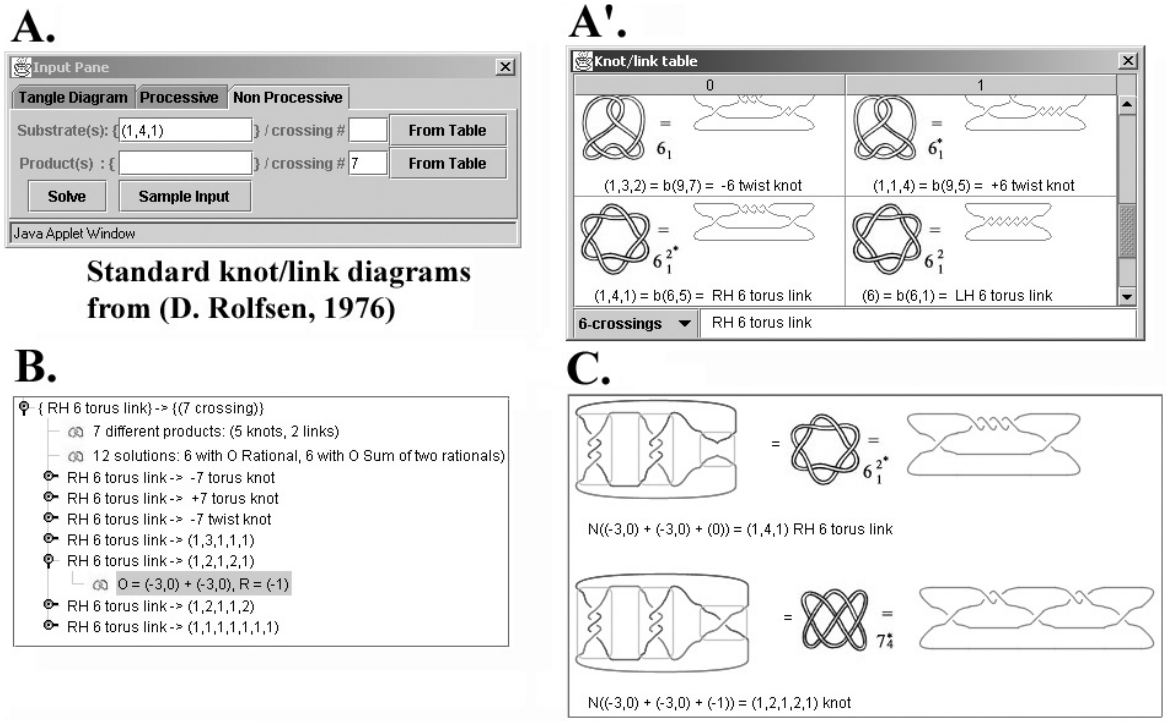

Fig. 1. Java applet for tangle analysis of site-specific recombination. **A.** Input pane where the experimental data are entered as lists of knots or links. **A'.** A table of all 4-plat knots and links with up to 10 crossings is available. The 4-plat diagrams were generated using TangleSolve, the standard knot diagrams were adapted from (Rolfsen, 1976) with the author's permission. The user can select substrates/products from the table. **B.** The selection pane displays the list of all the solutions computed by the tangle model. **C.** The display pane presents a specific solution, representing local enzyme action and global topology of DNA substrate and products, in diagrammatic form.

The figure illustrates the use of the software with a concrete example. Xer is a site-specific recombination system that, when acting on right-handed 6-crossing torus links, denoted mathematically by its Conway vector (1,4,1) or the corresponding Conway symbol b(6,5) (illustrated in panel A'), produces a 7-crossing product of unknown knot type (Bath *et al.*, 1999); no processive recombination has been observed for the Xer reaction. The data are inserted as illustrated in panel A. The solutions to the system are computed and listed (panel B). The solution highlighted in panel B, is represented in panel C by means of tangle and knot diagrams.

Summarizing, TangleSolve computes solutions to systems of tangle equations for processive or non-processive recombination. All solutions that are rational or sums of rational tangles are obtained for given substrate/product 4-plats. The solutions are displayed as knot and tangle diagrams. TangleSolve may allow the tangle model to be efficiently used in the design of experiments. More detailed information about TangleSolve, including a formal mathematical description of the tangle model, can be found at http://bio.math.berkeley.edu/TangleSolve/.

ACKNOWLEDGMENTS

Research was funded by the URAP program (YS) and by NSF grant DMS-9971169 (MV). We wish to thank R.K. Sachs and J. Arsuaga for helpful comments, and D. Rolfsen for allowing us to use his knot table.

REFERENCES

Darcy,I.K. (2001) Biological distances on DNA knots and links: applications to Xer recombination. Knots in Hellas '98, Vol.2 (Delphi). *J. Knot Theory Ramifications*, **10**, 269–294.

Sumners,D.W., Ernst,C., Cozzarelli,N.R. and Spengler,S.J. (1995) Mathematical analysis of the mechanisms of DNA recombination using tangles. *Quarterly Reviews of Biophysics*, **28**, 253–313.

Rolfsen,D. (1976) *Knots and Links*. Publish or Perish, Houston.

Ernst,C. and Sumners,D.W. (1990) A calculus for rational tangles: applications to DNA recombination. *Math. Proc. Camb. Phil. Soc.*, **108**, 489–515.

Vázquez,M. (2000) Tangle analysis of site-specific recombination: Gin and Xer systems, Ph.D. Dissertation, Mathematics, Florida State University.

Bath,J., Sherratt,D.J. and Colloms,S.D. (1999) Topology of Xer recombination on catenanes produced by lambda Integrase. *J. Mol. Biol.*, **289**, 873–883.