CONTEMPORARY MATHEMATICS

746

Topology and Geometry of Biopolymers

AMS Special Session
Topology of Biopolymers
April 21–22, 2018
Northeastern University, Boston, Massachusetts

Erica Flapan
Helen Wong
Editors

AMS American Mathematical Society

Topology and Geometry of Biopolymers

CONTEMPORARY MATHEMATICS

746

Topology and Geometry of Biopolymers

AMS Special Session
Topology of Biopolymers
April 21–22, 2018
Northeastern University, Boston, Massachusetts

Erica Flapan
Helen Wong
Editors

AMS American Mathematical Society

2010 *Mathematics Subject Classification.* Primary 57M25, 57M27, 05C10, 92C05, 92C40, 92D20, 92E10, 82D60, 82B41, 65C05.

Library of Congress Cataloging-in-Publication Data

Names: AMS Special Session on Topology of Biopolymers (2018 : Boston, Mass.), author. | Flapan, Erica, 1956- editor. | Wong, Helen, 1978- editor.

Title: Topology and geometry of biopolymers : AMS Special Session on Topology of Biopolymers, April 21-22, 2018, Northeastern University, Boston, Massachusetts / Erica Flapan, Helen Wong, editors.

Description: Providence, Rhode Island : American Mathematical Society, [2020] | Series: Contemporary mathematics, 0271-4132 ; volume 746 | Includes bibliographical references.

Identifiers: LCCN 2019040106 | ISBN 9781470448400 (paperback) | ISBN 9781470454562 (ebook)

Subjects: LCSH: Biopolymers–Mathematics–Congresses. | Topology–Congresses. | Knot theory–Congresses. | AMS: Manifolds and cell complexes {For complex manifolds, see 32Qxx} – Low-dimensional topology – Knots and links in $S3$ {For higher dimensions, see 57Q45}. | Manifolds and cell complexes {For complex manifolds, see 32Qxx} – Low-dimensional topology – Invariants of knots and 3-manifolds. | Combinatorics {For finite fields, see 11Txx} – Graph theory {For applications of graphs, see 68R10, 81Q30, 81T15, 82B20, 82C20, 90C35, 92E10, 94C15} – Planar graphs; geometric and topological aspect | Biology and other natural sciences – Physiological, cellular and medical topics – Biophysics. | Biology and other natural sciences – Physiological, cellular and medical topics – Biochemistry, molecular biology. | Biology and other natural sciences – Genetics and population dynamics – Protein sequences, DNA sequences. | Biology and other natural sciences – Chemistry {For biochemistry, see 92C40} – Molecular structure (graph-theoretic methods, methods of differential topology, etc.). | Statistical mechanics, structure of matter – Applications to specific types of physical systems – Polymers. | Statistical mechanics, structure of matter – Equilibrium statistical mechanics – Random walks, random surfaces, lattice animals, etc. [See also 60G50, 82C41]. | Numerical analysis – Probabilistic methods, simulation and stochastic differential equations {For theoretical aspects, see 68U20 and 60H35} – Monte Carlo methods.

Classification: LCC QP801.B69 A47 2018 | DDC 572–dc23

LC record available at https://lccn.loc.gov/2019040106

Contents

Preface

Biopolymers are long flexible molecules which are produced by living organisms. The families of polynucleotides (including DNA and RNA) and polypeptides (including proteins) are especially important in basic cellular functions. Because of their flexibility, biopolymers can have complex topological conformations including knots, links, and other types of entanglements, which may affect how they interact with other molecules inside the cell. However, the exact role of such topological features and how they occur is not well understood. Techniques from topology, geometry, dynamics, probability, and statistics, in addition to simulations, and experimentation have all been used to identify and analyze topologically complex biopolymers. These approaches not only provide evidence that could help to explain how and why such features occur but might also pave the way to future uses of topologically complex biopolymers in treating disease or in developing new biodegradable products. By bringing together articles about the topology and geometry of DNA and proteins from many different perspectives, we hope to give readers a sense of the complexity and interconnections within this exciting and fast-moving field.

Part I. The Topology and Geometry of DNA

The first half of this volume is devoted to recent results about the topology and geometry of DNA and how they affect DNA behavior. We begin with an article by Jonathan M. Fogg and Lynn Zechiedrich reporting on experimental results about DNA supercoiling. In particular, they used electron cryo-tomography to explore the 3-dimensional structure of DNA minicircles each of whose length is only 336 base pairs. They also simulated how introducing sequences with a higher propensity to bend affects the 3-dimensional geometry of such minicircles. Their results show unexpected sequence-dependent and supercoiling-dependent structural changes in DNA which may play a role in the development of gene therapy for treating disease.

Next there is an article by Pengyu Liu, Ryan Polischuk, Yunan Diao, and Javier Arsuaga also presenting results about DNA minicircles. However, in this case, they study kinetoplast DNA which has the form of chainmail made up of thousands of minicircles linked together. The authors use numerical simulations to compare the density of minicircles in the kinetoplast DNA with that of a "random knotting" model that they develop. The paper concludes that minicircle density is the most important factor in network formation, with minicircle orientation restrictions playing a significant role as well. Their results may be useful in targeted drug delivery systems.

The third article, written by Alex Kasman and Brenton LeMesurier, studies the relationship between the geometry of a DNA molecule and the amino acid

sequences it encodes. The authors use Monte Carlo and Gaussian sampling methods to approximate a multiple integral that represents both the amino acid sequences and the geometry of a DNA molecule. Their results imply the rather surprising result that the genetic code allows for less freedom in the geometry than what would be expected due to random chance. They offer interesting speculations as to what this result might say about the evolution of the genetic code.

This is followed by an article by Tetsuo Deguchi and Erica Uehara which uses simulations to evaluate the probability that a DNA chain of a given length contains a specific knot. In particular, they represent DNA molecules as self-avoiding polygons made up of rigid segments. They find that as the volume around one segment of the chain increases, the probability of obtaining a trefoil knot increases while that of obtaining more complex prime knots decreases. Furthermore, they observe that as the number of segments in a self-avoiding polygon goes to infinity, the probability that a knotted polygon contains a given prime knot as a summand approaches 1.

The article by Dorothy Buck and Danielle O'Donnol considers knotting of replication intermediates. They explain that while type II topoisomerases normally facilitate replication by performing crossing changes to unknot, unlink, or relieve torsional stress on a DNA molecule, topoisomerases can also accidentally cause knotted intermediates. In order to better understand how and why this occurs, the article models replication intermediates by θ-curves and uses the unknotting number together with six biologically justified restrictions to show that only ten knot types can be produced as intermediates in this way.

In the last article on DNA, Allison H. Moore and Mariel Vazquez use the topological operation of band surgery to model site-specific recombination. The article argues that if no such surgery is possible to get from one given knot to a second given knot, then site-specific recombination will not cause DNA containing the first knot to change into the second knot. The authors draw on techniques from 3-manifold topology to impose strong restrictions on when there is a band surgery between a given pair of knots. In addition to this topological approach, the article uses numerical simulation of knots in a cubic lattice to determine whether a band surgery between two knots is possible and, if so, what is the relative frequency of such a surgery.

Part II. The Topology and Geometry of Proteins

In contrast with DNA, biologists did not realize that proteins could contain knots until recently. As more and more knotted proteins are identified, understanding the possible function of knots in proteins and how such proteins fold becomes increasingly important. We begin the second half of this volume with a survey article written by Sophie Jackson on why proteins might knot. In particular, since the folding rates of certain deeply knotted proteins are slow compared with unknotted proteins, some scientists have hypothesized that such proteins must have an evolutionary advantage over faster folding unknotted structures. In this review, the evidence for and against this theory is discussed, including how a knot within a protein may affect the thermodynamic, kinetic, mechanical, and cellular stability of the protein.

The second article in Part II of the volume, written by Ana Nunes and Patrícia Faísca, presents a survey of lattice Gō models that have been used to better understand how proteins fold into a knotted native state. The article includes a discussion

of how the results of such models fit together with experimental, theoretical, and other computational results on the folding of knotted proteins. Furthermore, the authors explain some folding properties of knotted proteins and present possible functional advantages of knots in proteins.

The next article, by Dimos Goundaroulis, Julien Dorier, and Andrzej Stasiak, explains how knotoids can be used to identify, analyze, and characterize knots in open protein chains. In particular, a knotoid is a diagram of an open knotted chain where the endpoints are fixed in the plane or sphere of projection. Using knotoids to represent proteins avoids the problem of having to decide which algorithm to use to close the endpoints of a knotted protein chain in order to characterize the knot. In addition to discussing how to use probability distributions of knotoids to measure entanglement of open chains, the paper explains how to incorporate intra-chain protein bonds in the analysis of entanglement. Finally, the article introduces their software package Knoto-ID which, so far, is the only computational tool that uses knotoids to analyze the topology of open curves.

In the fourth article in Part II of the volume, Kenneth C. Millett explains how to adapt Gauss linking numbers, which have played a role in analyzing the topological aspects of DNA molecules, to identify and quantify entanglements in proteins. In particular, the method introduced in this paper can be used to analyze local entanglements that resemble knotting or slipknotting. This makes it possible to compare the topological and geometric properties of entanglements in unknotted proteins to those in knotted proteins. Furthermore, the paper explains how this can be done with or without considering intra-chain bonds. Finally, the paper surveys the software that is currently available for analyzing linking and entanglement of protein structures.

The last paper in this volume, written by Eleni Panagiotou and Kevin W. Plaxco, explains how various topological and geometric properties of unknotted proteins can be used to predict folding rates. In particular, they consider the Gauss linking integral, the torsion, the average crossing number, and the number of sequence-distant contacts of certain proteins and show how all of these factors can cause folding rates to increase or decrease. Among their many results, they show that folding rates decrease as the writhe and torsion of the proteins becomes more negative.

This volume is based on an AMS special session on topology of biopolymers that we organized at Northeastern University in Boston, April 21–22, 2018. We very much enjoyed organizing the special session, listening to all of the talks, and putting together this collection of articles. We hope that readers will find the diversity of approaches to studying the topology and geometry of biopolymers in this volume as stimulating as we have. Finally, we want to thank the AMS for giving us the opportunity to organize such an interdisciplinary session and for inviting us to put together this volume.

Acknowledgments

Erica Flapan was supported in part by NSF grant DMS-1607744, and Helen Wong was supported in part by NSF grants DMS-1510453, DMS-1841221.

Selected Published Titles in This Series

746 **Erica Flapan and Helen Wong, Editors,** Topology and Geometry of Biopolymers, 2020

740 **Alina Bucur and David Zureick-Brown, Editors,** Analytic Methods in Arithmetic Geometry, 2019

739 **Yaiza Canzani, Linan Chen, and Dmitry Jakobson, Editors,** Probabilistic Methods in Geometry, Topology and Spectral Theory, 2019

738 **Shrikrishna G. Dani, Surender K. Jain, Jugal K. Verma, and Meenakshi P. Wasadikar, Editors,** Contributions in Algebra and Algebraic Geometry, 2019

737 **Fernanda Botelho, Editor,** Recent Trends in Operator Theory and Applications, 2019

736 **Jane Hawkins, Rachel L. Rossetti, and Jim Wiseman, Editors,** Dynamical Systems and Random Processes, 2019

735 **Yanir A. Rubinstein and Bernard Shiffman, Editors,** Advances in Complex Geometry, 2019

734 **Peter Kuchment and Evgeny Semenov, Editors,** Differential Equations, Mathematical Physics, and Applications, 2019

733 **Peter Kuchment and Evgeny Semenov, Editors,** Functional Analysis and Geometry, 2019

732 **Samuele Anni, Jay Jorgenson, Lejla Smajlović, and Lynne Walling, Editors,** Automorphic Forms and Related Topics, 2019

731 **Robert G. Niemeyer, Erin P. J. Pearse, John A. Rock, and Tony Samuel, Editors,** Horizons of Fractal Geometry and Complex Dimensions, 2019

730 **Alberto Facchini, Lorna Gregory, Sonia L'Innocente, and Marcus Tressl, Editors,** Model Theory of Modules, Algebras and Categories, 2019

729 **Daniel G. Davis, Hans-Werner Henn, J. F. Jardine, Mark W. Johnson, and Charles Rezk, Editors,** Homotopy Theory: Tools and Applications, 2019

728 **Nicolás Andruskiewitsch and Dmitri Nikshych, Editors,** Tensor Categories and Hopf Algebras, 2019

727 **André Leroy, Christian Lomp, Sergio López-Permouth, and Frédérique Oggier, Editors,** Rings, Modules and Codes, 2019

726 **Eugene Plotkin, Editor,** Groups, Algebras and Identities, 2019

725 **Shijun Zheng, Marius Beceanu, Jerry Bona, Geng Chen, Tuoc Van Phan, and Avy Soffer, Editors,** Nonlinear Dispersive Waves and Fluids, 2019

724 **Lubjana Beshaj and Tony Shaska, Editors,** Algebraic Curves and Their Applications, 2019

723 **Donatella Danielli, Arshak Petrosyan, and Camelia A. Pop, Editors,** New Developments in the Analysis of Nonlocal Operators, 2019

722 **Yves Aubry, Everett W. Howe, and Christophe Ritzenthaler, Editors,** Arithmetic Geometry: Computation and Applications, 2019

721 **Petr Vojtěchovský, Murray R. Bremner, J. Scott Carter, Anthony B. Evans, John Huerta, Michael K. Kinyon, G. Eric Moorhouse, and Jonathan D. H. Smith, Editors,** Nonassociative Mathematics and its Applications, 2019

720 **Alexandre Girouard, Editor,** Spectral Theory and Applications, 2018

719 **Florian Sobieczky, Editor,** Unimodularity in Randomly Generated Graphs, 2018

718 **David Ayala, Daniel S. Freed, and Ryan E. Grady, Editors,** Topology and Quantum Theory in Interaction, 2018

717 **Federico Bonetto, David Borthwick, Evans Harrell, and Michael Loss, Editors,** Mathematical Problems in Quantum Physics, 2018

For a complete list of titles in this series, visit the
AMS Bookstore at **www.ams.org/bookstore/conmseries/**.

CONM 746

Topology of Biopolymers • Flapan et al., Editors

This book contains the proceedings of the AMS Special Session on Topology of Biopolymers, held from April 21–22, 2018, at Northeastern University, Boston, MA.

The papers cover recent results on the topology and geometry of DNA and protein knotting using techniques from knot theory, spatial graph theory, differential geometry, molecular simulations, and laboratory experimentation. They include current work on the following topics: the density and supercoiling of DNA minicircles; the dependence of DNA geometry on its amino acid sequence; random models of DNA knotting; topological models of DNA replication and recombination; theories of how and why proteins knot; topological and geometric approaches to identifying entanglements in proteins; and topological and geometric techniques to predict protein folding rates.

All of the articles are written as surveys intended for a broad interdisciplinary audience with a minimum of prerequisites. In addition to being a useful reference for experts, this book also provides an excellent introduction to the fast-moving field of topology and geometry of biopolymers.

AMS
www.ams.org

ISBN 978-1-4704-4840-0
9 781470 448400
CONM/746

AMS