MATHEMATICS MATTERS

Unravelling the Genetic Code

Understanding how to manipulate tangles of DNA could help us create new treatments for diseases, so mathematicians are working with biologists to explain how our genetic code becomes knotted. Their work combines the latest technology with a centuries-old branch of pure mathematics.

Knots crop up in all sorts of places: shoe laces, computer cables, and even in our DNA. The difference is that while untangling the mess of wires behind your desktop is merely frustrating, knots in DNA molecules are life-threatening. Learning to control these knots could lead to new treatments for genetic disorders and diseases such as cancer or MRSA, so biologists have called in the mathematicians to help unravel the problem.

Mathematicians have studied knots since the 19th century, when Lord Kelvin theorised that atoms were actually circular knots in an invisible substance known as ether. The idea was ultimately proved false, but his work founded a new branch of mathematics known as knot theory. Today, mathematicians such as Dorothy Buck at Imperial College London are using knot theory to understand why DNA becomes knotted, and how our bodies react to the tangles. It is all down to the unique packaging of our genetic code.

The DNA molecule is a double helix that resembles a twisted ladder, with two long strands joined together by a series of short rungs of paired molecules. There are two types of pairs, "A" with "T" and "C" with "G", and the particular order of these molecules forms the genetic code for creating life. DNA replicates by splitting the rungs down the middle, so that each strand contains one half of the genetic code. Since the molecules in the rungs will only join with their opposites, the two strands act as a template for reconstructing the entire genetic code, resulting in two identical copies of DNA that each contain one strand from the original. This process would be simple if our DNA molecules were stored in one long line, but the human genome stretches for nearly two metres, so it is wrapped into tiny parcels that fits inside our cells. It is compact, but not always convenient, as our cells sometimes need to make minor rearrangements to their DNA. To achieve this, specialised enzymes called recombinases cut out small parts of the genetic code and replace them with DNA from elsewhere. We also use these enzymes as molecular scissors for inserting genes in genetic engineering. This procedure, known as recombination, involves repeatedly winding and unwinding the DNA parcel, inevitably creating knots and tangles. They make it impossible for the DNA to replicate, so other enzymes called topoisomerases act to keep the ball organised by tidying up after the recombinases.

> "We don't fully understand what combination of tugs and cuts unpicks a particular DNA knot. The same problem is at the heart of knot theory."

Although biologists have known about these enzymes for some time, we don't fully understand how they work or what combination of tugs and cuts unpicks a particular DNA knot. The same problem is at the heart of knot theory: for a given knot, what set of moves will unknot it? Buck and her colleagues have recently helped narrow down the scope of the problem by showing that DNA can only form a certain family of knots, which allows biologists to get a better picture of how the enzymes function. Buck hopes that this research will help characterise newly discovered recombinases, expanding the biologists' genetic engineering tool kit and allowing them to treat genetic disorders by directly altering damaged DNA.

Buck's work could also help improve existing treatments for a variety of diseases by explaining the mechanisms behind topoisomerases. For example, DNA molecules are often supercoiled, meaning that the double helix structure is itself coiled into a helix, much like the way the spiralled cord of a telephone can twist around itself and become tangled. Simply tugging at a knotted cord often just builds up more coils elsewhere, but letting the handset dangle free can relieve the built up tension. One type of topoisomerase performs a similar job during the DNA replication process, relieving tension in the unwinding molecule by creating a small cut in one strand of the DNA.

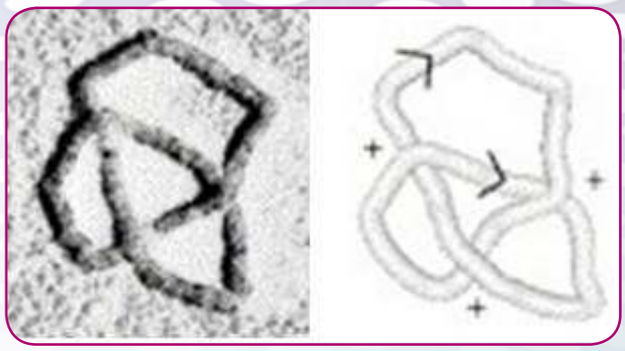
Institute of mathematics & its applications

EPSRC Pioneering research and skills

The situation is described mathematically by an equation relating the tightness of the double helix coil to the extent of supercoiling. These two quantities, called twist and writhe, add together to make the linking number, a value that describes the number of times two strands cross. This relationship explains why untangling both DNA and telephone cords is so difficult: decreasing writhe automatically increases twist, leading to increased tension. Topoisomerases act to reduce the linking number, relieving the extra tension.

These enzymes play an important role in cancer treatment. Cancer is essentially uncontrolled cellular growth caused by genetic errors, so the DNA within cancer cells is constantly replicating. Some chemotherapy treatments work by disabling the straightening topoisomerases, halting tumour growth. Unfortunately, this also affects DNA replication in healthy cells, leading to side effects such as hair loss. A better understanding of how the topoisomerases work within cancer cells could allow specific targeting, sparing enzymes in healthy cells and reducing the harmful side effects.

Another variety of topoisomerase can also act upon circular DNA, which is formed by the double helix looping back on to itself and joining together. This kind of DNA is found in bacteria and also in mitochondria, the small parts of a cell that convert glucose into energy. Unlike in normal DNA, the two strands of circular DNA remain intertwined during replication, resulting in two linked molecules. These circles must be separated before they can be placed into separate cells, so this second variety of topoisomerase cuts through one loop and pulls the other free before repairing the cut. This action also serves to reduce the linking number, simplifying the knot.



Fluoroquinolones, an effective type of antibiotic, work by inhibiting these unlinking enzymes within bacterial cells, preventing them from replicating and killing the bacterial infection. As with all antibiotics, there is a danger that overuse will lead to bacteria developing resistance, rendering the drugs useless against diseases like MRSA. By revealing how the bacterial topoisomerases work, knot theory offers the possibility of developing new variations of fluoroquinolones to attack the enzymes in a variety of ways, reducing the chance of resistance.

Pure mathematics can often seem abstract, but Buck's research demonstrates that it can still have practical applications. The foundations of knot theory were laid over 150 years ago, but this latest collaboration between mathematicians and biologists has only been made possible in the last few decades, thanks to new imaging techniques that help create pictures of knotted DNA. Now that technology has caught up to the mathematics, the powerful techniques of knot theory could one day lead to a medical breakthrough.

TECHNICAL SUPPLEMENT

Knot theory terminology

The mathematical concepts of twist, writhe and linking number apply to all knots and links, not just those found in DNA, but the values they can take in these knots are constrained by the molecule's physical properties. There are roughly 10.5 base pair rungs per complete revolution of a normal DNA double helix, and the twist is the total number of base pairs in the whole molecule divided by this number. The writhe is calculated by tracing along the DNA molecule and counting the crossings; going over a crossing adds one to the total, while going under subtracts one.

Both twist and writhe are dependant on the particular orientation of the knot or link in space, but their sum, the linking number, is always the same. It is calculated in the same way as the writhe, but only the crossings that involve both knots in a link are counted, and this total is then halved to give the linking number.

The two classes of topoisomerases change the linking number by different amounts. The type I topoisomerases that straighten DNA only cut one strand, changing the linking number in steps of one. Type II topoisomerases, which unlink circular DNA, cut both strands of one molecule and thus change the linking number in steps of two.

Predicting DNA topology

Mathematicians often don't work with knots directly, as it can be easier to manipulate their complement. Imagine encasing a knotted portion of string in a plastic sphere, then dissolving the string – the complement is the holey plastic sphere left over.

Dorothy Buck and Erica Flapan used this technique to develop a general model for the knots and links left in DNA after site-specific recombination. By mathematically placing spheres around the recombinase enzymes and the DNA sections they act upon, Buck and Flapan were able to compare the knot complements before and after recombination, then classify all possible knots and links that arose. Their research will help to determine the pathways taken by the enzymes, allowing biologists to better understand newly discovered recombinases.

References

Buck, D. & Flapan, E. (2007) A topological characterization of knots and links arising from site-specific recombination. *Journal of Molecular Biology,* 374 (5), 1186–1199. DOI: 10.1016/j.jmb.2007.10.016

Buck, D. (2009) DNA Topology. In: Buck, D. and Flapan, E. (eds.) *Proceedings of Symposia in Applied Mathematics, Volume 66: Applications of Knot Theory.* American Mathematical Society. pp 47-82.

EPSRC Grants

Reference: EP/G039585/1
Title: DNA Knotting and Linking: Applications of 3-Manifold Topology to DNA-Protein Interactions

Reference: EP/H031367/1
Title: SANDPIT : Knots and Evolution - Topologically Driven Integrase Mutagenesis

The IMA would like to thank Dr. Dorothy Buck, Imperial College London for her help in the preparation of this document