

1

Stereochemical Topology

Stereochemistry is the study of the three-dimensional structure of molecules, and topology is the study of those properties of geometric objects that are invariant under continuous transformations. It is not obvious that these two fields have anything in common. In fact, not long ago there was little communication between researchers in these two areas. Prior to forty years ago, analyzing the topological properties of existing molecular structures was not very difficult, because as topological objects, the graphs of all the molecular structures known at the time could be deformed into a plane. Thus, understanding the stereochemistry of a molecule only required the evaluation of its geometry and not its topology. Recently, knots and links and other molecules have been synthesized whose structures and properties come from their topology as well as their geometry. The chemical motivation for the synthesis of such topologically interesting structures is the desire to synthesize new types of molecules that might have truly unusual properties, as well as the hope that in trying to create such unique molecules, new methods of synthesis will be developed along the way. These new types of molecules are often large enough that they no longer have the rigidity that is characteristic of small molecules, so understanding their deformations is an important part of understanding their structure. In addition to purely synthetic molecules, Liang and Mislow (1994b, 1995) have discovered that knots and links and other nonplanar graphs can occur naturally in proteins. Also, molecular biologists have found that knotted and linked DNA can exist in nature (Clayton & Vinograd, 1967; Hudson & Vinograd, 1967; Liu, Depew, & Wang, 1976) and can be manipulated in the lab (Seeman, 1999). Thus now both chemistry and molecular biology have something to gain by understanding the topology of graphs embedded in three-dimensional space.

Although many applications of topology to chemistry have recently gained attention, it is worth noting that chemistry also made an important contribution to topology over a century ago in the formation of the field of knot theory.

In the 1880s, scientists believed that all of space was filled with a substance called *ether*. William Thomson (a.k.a. Lord Kelvin) theorized that atoms were vortex rings in the ether; rings that were knotted and linked in different ways represented different elements (Kelvin, 1904). In order to try to create a periodic table of the elements, the physicist Peter Guthrie Tait created tables that divided knots into classes in such a way that a knot in one class could not be deformed into a knot in another class (Tait, 1898). By the turn of the century, it was clear that Kelvin's theory of atomic structure was wrong. Thus chemists were no longer interested in classifying knots. However, topologists had become interested in the study of knotted closed curves in space, and the field of knot theory has blossomed since then into an important subfield of three-dimensional topology that is closely tied to many other areas of topology. As we shall see, many of the ideas of knot theory are now the basis for some of the applications of topology to stereochemistry.

This chapter provides an overview of some of the topological theorems that have been proven, as well as some of the open problems that have arisen as a result of applying topological techniques to the study of complex molecular structures. Rigorous definitions or proofs are not given in this chapter; rather the focus is on providing a visual and conceptual understanding. The results that are mentioned here, as well as other results, will be proven in detail in Chapters 2 through 6. Chapter 7 is devoted entirely to the topology of DNA.

Examples of Topologically Complex Molecules

In this section we will give a very brief overview of some of the topologically complex molecules that have been synthesized. By a *graph* we shall mean a finite collection of vertices together with disjoint edges connecting pairs of vertices, with the requirement that there is at most one edge between any pair of vertices and every edge has two distinct vertices. Figure 1.1 illustrates the types of edges that are forbidden in a graph. An *embedded graph* is a graph that exists in a specific position in three-dimensional space, whereas an *abstract graph* is a graph that is considered to be independent of any particular embedding in three-dimensional space. The *molecular bond graph* of a molecule is a model of the molecule, which is represented by a graph embedded in three-dimensional space where the vertices represent atoms or collections of atoms and the edges represent chemical bonds or chains of atoms. Such a molecular bond graph is often labeled to indicate what atom or atoms each vertex represents, and to indicate which edges represent double bonds and which represent single bonds.

A graph in three-dimensional space is considered to be topologically complex if it cannot be deformed into a plane, even under the assumption of complete

Cambridge University Press

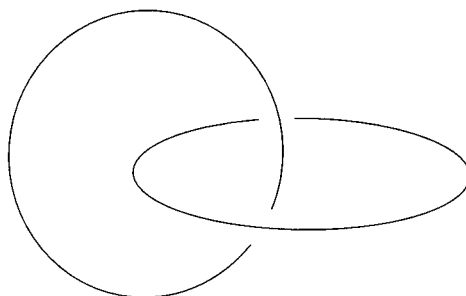
978-0-521-66482-0 - When Topology Meets Chemistry: A Topological Look at Molecular Chirality

Erica Flapan

Excerpt

[More information](#)*Examples of Topologically Complex Molecules*

3

**Figure 1.1.** A graph is not permitted to have these types of edges**Figure 1.2.** The first molecular link to be synthesized had the form of a Hopf link

flexibility. If we imagine that the graph of a typical molecule is made out of rubber, it can easily be deformed into a plane. Of course many molecules are actually too rigid to be deformed into a plane. Nonetheless, any structure that could theoretically be deformed into a plane will not be the object of our study.

The first topologically complex molecules to be synthesized were a pair of linked rings, known in chemistry literature as *catenanes*. Although the individual rings of a catenane are disjoint, they behave together as a single molecule. Starting around 1910, chemists began trying to synthesize a pair of linked rings. However, a molecular ring had to be synthesized that was big enough so that another molecule could pass through the hole in the center in order to create a link. Synthesizing a catenane, and proving that it was such, was finally achieved in 1961 by Frisch and Wasserman (Frisch & Wasserman, 1961). Their catenane was created out of a pair of linked hydrocarbon rings that each contained 34 atoms. A *hydrocarbon* is a compound that is made up entirely of hydrogen and carbon atoms. This molecular link has the structure of the *Hopf link*, which is illustrated in Figure 1.2. This is the simplest possible link, as it is the only link that can be drawn with only two crossings. Even if we imagine that the molecular graph of this link is completely flexible, it cannot be deformed into a plane because the rings are linked and cannot be separated unless one is cut. Although Frisch and Wasserman discussed the possible synthesis of knots, they did not succeed in synthesizing a knot.

Another topologically important molecule is the Simmons–Paquette molecule, which was independently synthesized in 1981 by the laboratories

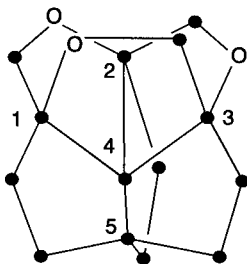


Figure 1.3. The Simmons-Paquette molecule

of Simmons and Maggio (1981) and of Paquette and Vazeux (1981). This molecule is illustrated in Figure 1.3. Throughout the book, when we draw molecular graphs, all unlabeled vertices will represent carbon atoms, and for the sake of simplicity, hydrogen atoms will generally be omitted. In Figure 1.3, the three Os represent oxygen atoms. The numbers on some of the vertices have no chemical significance and will be explained shortly. Like that of the linked rings, even if it were completely flexible, the Simmons–Paquette molecular graph could not be deformed into a plane. However, this graph has the even stronger property that no embedding of the graph can lie in a plane. Thus there is no way to take apart the graph of the Simmons–Paquette molecule into a collection of vertices and edges and reassemble the vertices and edges in a plane in such a way that edges connect the same pairs of vertices as were connected in the original molecular graph and no edges intersect except at their vertices.

A graph with the property that no embedding of it can lie in a plane is said to be *nonplanar*. Nonplanarity is a property of an abstract graph, rather than of the specific way that the graph is embedded in three-dimensional space. For example, the Hopf link, with any number of vertices added to each ring, is a specific embedding of a planar graph, although the particular embedding illustrated in Figure 1.2 cannot be deformed into a plane. We could take apart the vertices and edges and reassemble them as two disjoint circles in the plane. Kuratowski (1930) proved that a graph is nonplanar if and only if it contains either the complete graph on five vertices, K_5 , or the bipartite graph on three vertices, $K_{3,3}$, each with the possible addition of some extra vertices. We illustrate the graphs $K_{3,3}$ and K_5 in Figure 1.4. We have drawn these graphs abstractly, without specifying how they are embedded in space. In particular, these pictures are not meant to indicate that the edges intersect. Rather, no matter how these (or any other) graphs are positioned in three-dimensional space, their edges are required to be disjoint except at the vertices.

A *complete graph* on n vertices, K_n , consists of n vertices together with edges connecting every pair of vertices. A *complete bipartite graph*, $K_{p,q}$, consists

Examples of Topologically Complex Molecules

5

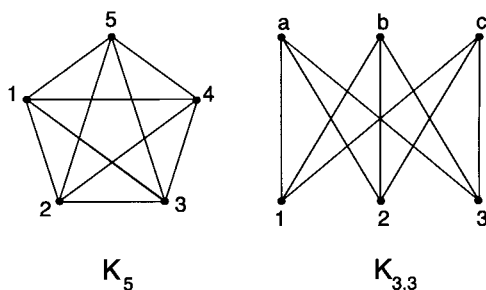


Figure 1.4. Any nonplanar graph contains either $K_{3,3}$ or K_5

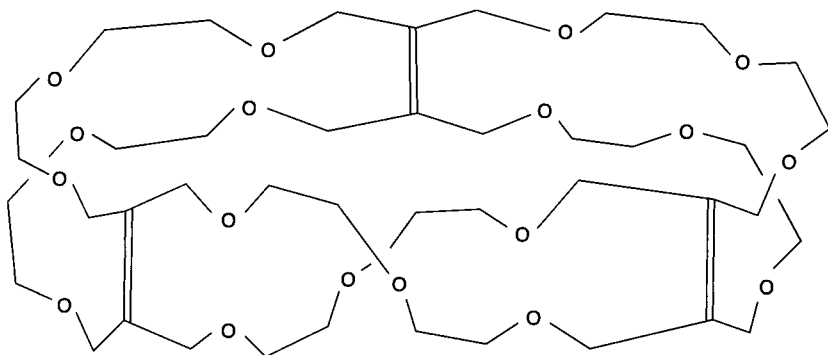


Figure 1.5. A molecular Möbius ladder

of a set of p vertices and a set of q vertices together with edges connecting every vertex in one set to every vertex in the other set. We have numbered five of the vertices in Figure 1.3 to indicate that the Simmons–Paquette graph actually contains a K_5 . The reader should observe that in the Simmons–Paquette molecular graph, there are disjoint paths connecting any numbered vertex to any other numbered vertex. Thus if we were to omit those vertices of the Simmons–Paquette without numbers, we would obtain a complete graph on five vertices. Hence it follows from Kuratowski’s theorem that the Simmons–Paquette graph cannot be embedded in a plane.

In 1982, Walba, Richards, and Haltiwanger (Walba et al., 1982) synthesized the first molecular Möbius ladder with three rungs (see Figure 1.5). This molecule resembles a Möbius strip in which the surface of the strip has been replaced by a ladder. The sides of the ladder represent a polyether chain of 60 atoms that are all carbons and oxygens, and the rungs of the ladder represent carbon–carbon double bonds. Thus the type of edges represented by the sides are chemically quite different from those represented by the rungs. We can see that

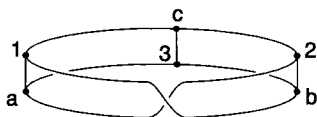


Figure 1.6. A Möbius ladder with three rungs is an embedding of $K_{3,3}$

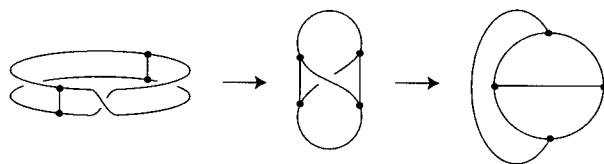


Figure 1.7. A two-rung Möbius ladder can be deformed into a plane

the three-rung Möbius ladder is a nonplanar graph because it contains a $K_{3,3}$ graph. Figure 1.6 represents the structure of a Möbius ladder with three rungs, where the vertices of valence 2 have been omitted and the remaining vertices have been labeled 1, 2, 3 and a , b , c in order to help the reader see that this graph is a $K_{3,3}$.

In 1975, a two-rung Möbius ladder was synthesized by Graf and Lehn (Graf & Lehn; 1975). However, two rungs are not enough to make the graph of a Möbius ladder nonplanar, as we can see from the deformation in Figure 1.7. To get from the first embedded graph to the second, we grab the lower string in the back and pull it down, opening up the boundary of the ladder into a figure eight. Then, we untwist the upper loop and straighten out the picture so that it looks like the third picture in Figure 1.7. Because the graph of the two-rung Möbius ladder can be deformed into a plane, this molecule is not considered to be topologically complex.

Finally, in 1989, a molecular knot was synthesized for the first time by Dietrich-Buchecker and Sauvage (Dietrich-Buchecker & Sauvage, 1989). This was a great achievement, as many chemists had been actively trying to synthesize a knot for over thirty years. Dietrich-Buchecker and Sauvage's molecular knot is made out of 124 atoms and has the form of a *trefoil knot*. We illustrate their molecular knot in Figure 1.8. From a topological point of view, a knot is any embedding of a circle that cannot be deformed into a plane. Although people may say they have knots in their shoelaces, such a knot can be untied by a deformation, and so topologically speaking it is not considered to be a knot. A trefoil knot is the simplest type of knot because it is the only knot with just three crossings, and there cannot exist a knot with fewer than three crossings. The molecular trefoil knot is much smaller than a knotted DNA molecule, so it is

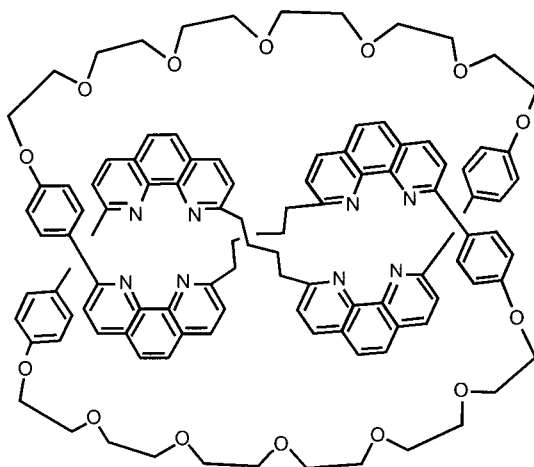


Figure 1.8. The molecular trefoil knot

most likely the smallest knot that has ever physically existed. The line segments inside of the hexagons indicate a specific type of bonding. On a chemical level, such hexagons represent benzene rings. The top and bottom arcs of the knot are molecular chains made of oxygens and carbons.

Like the Hopf link, this molecular graph has the property that it cannot be deformed into a plane; however, as an abstract graph it is planar because a different embedding of this graph looks like a planar circle with hexagons attached. This planar embedding can be seen in Figure 1.9.

The molecules that we have listed above are not the only topologically complex molecules that have been synthesized in the past four decades. Rather, we have selected these examples in order to introduce the reader to the history and variety of structures that have been synthesized. Liang and Mislow (1994a) have found many other molecules whose graphs contain $K_{3,3}$, K_5 , or both. In addition, Liang and Mislow (1994b, 1995a) have found that many proteins naturally contain the nonplanar graphs $K_{3,3}$ or K_5 graphs, as well as knots or links.

Stereoisomers

Walba (1983) coined the term *stereochemical topology* to refer to the synthesis, characterization, and analysis of molecular structures that are topologically non-trivial. Two stereochemical questions that interest topologists are recognizing when one embedding of a graph cannot be deformed to another embedding of the graph and evaluating those properties of embedded graphs that are preserved by deformations.

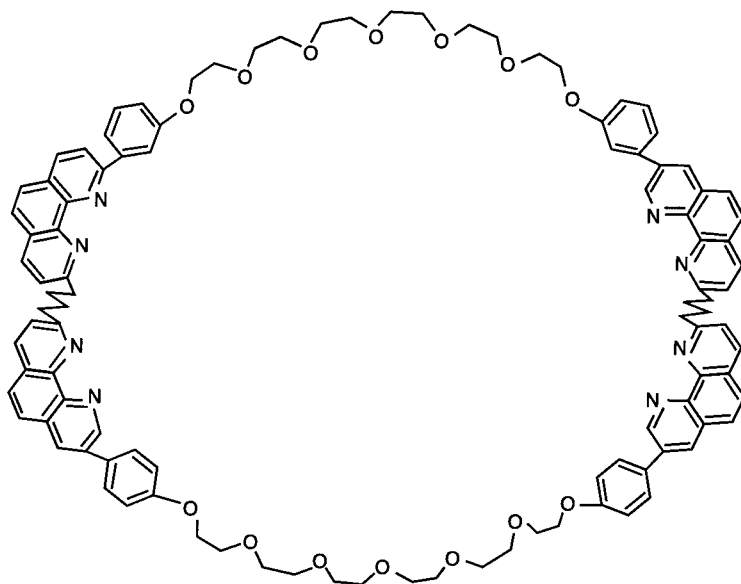


Figure 1.9. An unknotted embedding of the graph of the trefoil knot molecule

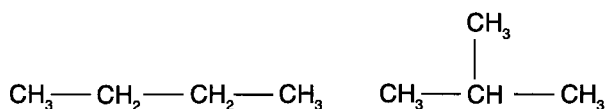


Figure 1.10. A pair of structural isomers

For a given molecule, chemists are interested in enumerating other molecules that are structurally related to it. Such molecules are called the *isomers* of the original molecule. From a chemical, geometric, and topological point of view, we can define three classes of isomers.

1. The *structural isomers* of a given molecule are those molecules that have the same molecular formula but are represented by different molecular bond graphs. In particular, a molecular formula provides a list of all of the atoms contained in a molecule, but it does not specify in what arrangement these atoms are bonded together.

In Figure 1.10, we illustrate a pair of structural isomers called *butane* and *isobutane*, which have different graphs but which both have the molecular formula C_4H_{10} .

Stereoisomers

9

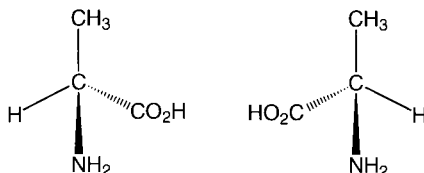


Figure 1.11. A pair of rigid stereoisomers

2. The *rigid stereoisomers* of a given molecule are those molecules that have the same abstract graph as the given molecule, but as rigid objects, one could not be picked up and superimposed on the other.

Figure 1.11 illustrates an L-alanine molecule on the left and a D-alanine molecule on the right, which are rigid stereoisomers. These two molecules are mirror images of each other, and the L and the D are there in order to distinguish the two mirror images. The dark triangular segments in Figure 1.11 indicate those edges that are coming out of the plane of the paper toward the reader, the dashed segments indicate those edges that go back behind the plane of the paper, and the ordinary line segments indicate edges that lie in the plane of the paper. In three-dimensional space, the vertices of these graphs lie at the corners of a regular tetrahedron. The reader should observe that if these graphs were flexible, we could deform one to the other, however, because they are rigid graphs, one cannot be picked up and put in the place of the other. For example, if we turn over the graph on the left to try to obtain the graph on the right, then the CH₃ and the H will be in the right place but the NH₂ will now be behind the page and the HO₂C will be in front of the page.

3. The *topological stereoisomers* of a given molecule are those molecules that have the same abstract graph as the given molecule, but as embedded graphs one cannot be deformed to the other.

Figure 1.12 illustrates a right-handed trefoil molecule (on the right) and a left-handed trefoil molecule (on the left). If you move along the right-handed trefoil knot with your right hand, then you can follow the twists, with your right thumb representing the overcrossings and your right fingers representing the undercrossings. In contrast, if you move along the left-handed trefoil knot with your left hand, then you can follow the twists, with your left thumb representing the overcrossings and your left fingers representing the undercrossings. In 1914, Max Dehn proved that a trefoil knot cannot be deformed to its mirror image (Dehn, 1914). We can use this result to prove that the graphs of the right-handed trefoil molecule and the left-handed trefoil molecule are topological

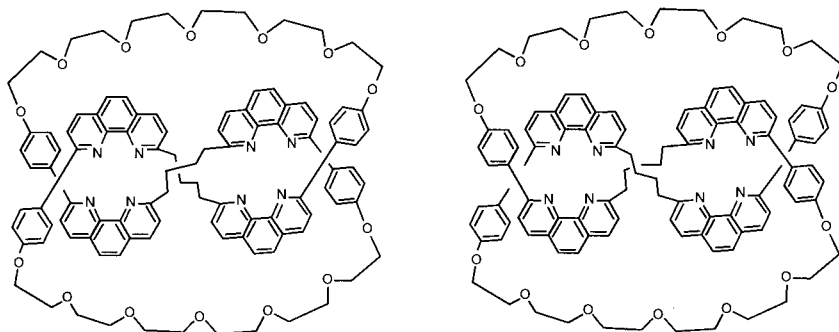


Figure 1.12. A pair of topological stereoisomers

stereoisomers. Although this may seem to follow immediately from Dehn's result, to prove it in complete detail requires more care because the molecular graph contains some hexagons rather than just single edges. In Chapter 2, we will discuss how to rigorously prove that the two trefoil molecular graphs are topological stereoisomers.

Note that the topological stereoisomers of a molecule are a subset of the rigid stereoisomers of the molecule, because a graph that cannot be deformed to another graph certainly cannot be rigidly superimposed on the other graph. Most chemists do not distinguish between rigid and topological stereoisomers. Rather, they define a pair of molecules to be *stereoisomers* if they have the same abstract graph but one cannot chemically change itself into the other. Whether this corresponds to our notion of rigid stereoisomers or to our notion of topological stereoisomers depends on to what extent particular bonds can rotate or bend.

Recognizing when two embeddings of a graph are topological stereoisomers is an extension of the well-known problem in knot theory of recognizing when two knots are distinct. A *knot* is an embedding of a circle in three-dimensional space, and two knots are said to be *distinct* if one cannot be deformed to the other. Similarly, two embeddings of the same graph are said to be *distinct* if one cannot be deformed to the other. For example, the two trefoil knotted molecules illustrated in Figure 1.12 are distinct. If we add a finite number of vertices to any knot, we create a graph. Thus the problem of recognizing when two knots are distinct is actually a special case of the problem of recognizing when a pair of embedded graphs are topological stereoisomers. The classification of knots is a very hard problem, so the classification of embeddings of graphs should be at least as hard a problem in topology. In contrast, the vertices of a graph do provide some extra structure, which can help. Furthermore, molecular graphs