Molecules of the year – 2019: Topological molecular nanocarbons – All-benzene catenane and trefoil knot.

Henry Rzepa ⓘ

Published December 15, 2019

Citation

Rzepa, H. (2019). Molecules of the year – 2019: Topological molecular nanocarbons – All-benzene catenane and trefoil knot. *Henry Rzepa's Blog*. https://doi.org/10.59350/zrnym-67e53
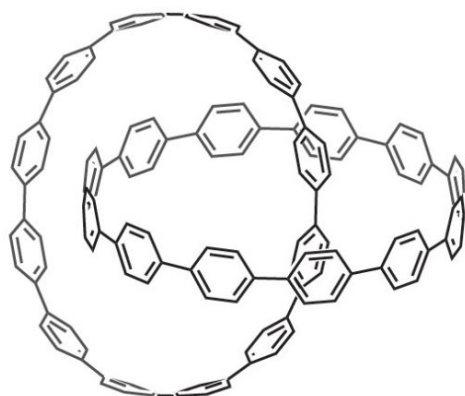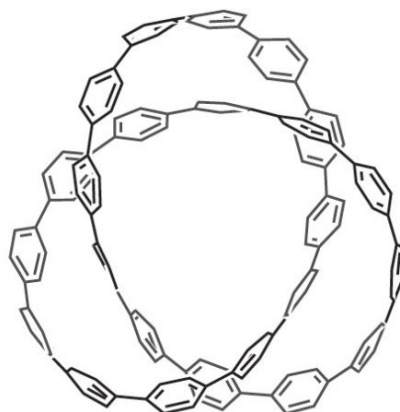
Keywords

Interesting Chemistry

all-benzene catenane **1a**   all-benzene trefoil knot **2**

Copyright

Here is another molecule of the year, on a topic close to my heart, the catenane systems **1** and the trefoil knot **2**[cite]10.1126/science.aav5021[/cite] Such topology is closely inter-twinned with three dimensions (literally) and I always find that the flat pages of a journal are simply insufficient to do them justice. So I set about finding the 3D coordinates.
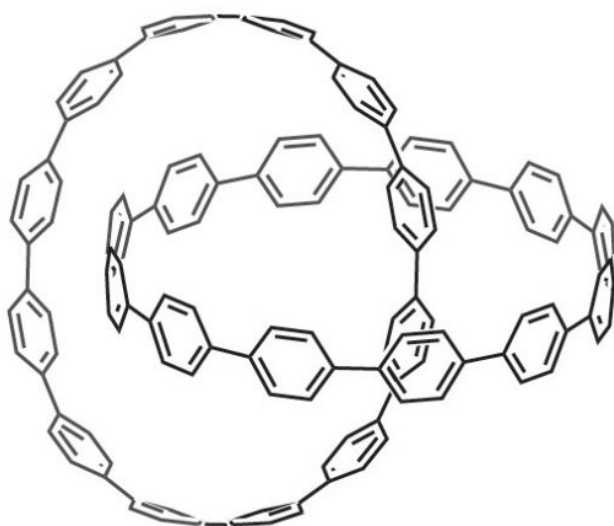
all-benzene catenane **1a**
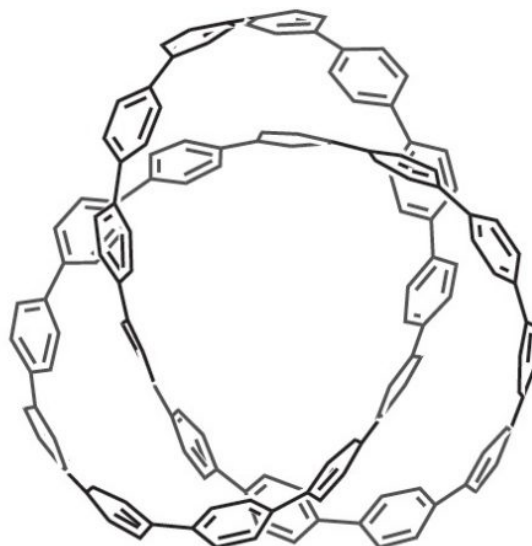
all-benzene trefoil knot **2**

The most obvious place to start is the supporting information. I show below a little snippet of what I found, which is fairly typical of such data in the PDF-based SI documents accompanying most articles.

Table S3.

Cartesian coordinates of optimized structures.

1a (E = -5545.56100209 Hartree)

C	3.832161	5.035441	-0.518998		H	-5.505783	6.838721	-4.953567		C	-6.448347	-1.821358	0.566911
C	2.813845	5.910161	-0.883767		H	-3.145187	7.457052	-5.142584		C	-4.693554	-4.480421	-1.482080
C	2.621798	6.285311	-2.223569		H	-7.584755	7.430370	-0.921461		C	-4.575292	-4.949882	-2.803578
C	3.560847	5.833177	-3.167295		H	-9.300483	5.805544	-0.263430		C	-3.561826	-5.832348	-3.167508
C	4.574485	4.950818	-2.803584		H	-8.180776	3.292570	-3.564700		C	-2.622712	-6.284716	-2.223949

A bit of knowledgeable text-editing is needed to convert these into something that can be displayed as a rotatable 3D model.[‡] For this example, the three-column mode did not actually prove too problematic (but sometimes you have to work very hard to reduce it to the single column mode required for coordinates) and one has to remember to notice and remove the pagination text from the coordinates. Here I include the structure as a static 2D image which when clicked expands to a 3D model. Sadly, I know of no journal that offers up this relatively simple service as part of its "added-value" to the publication processes; it has been possible to do this since 1994![cite]10.1039/C39940001907[/cite],[cite]10.1039/P29950000007[/cite] I also found that the provided coordinates could be symmetrised to D_2.
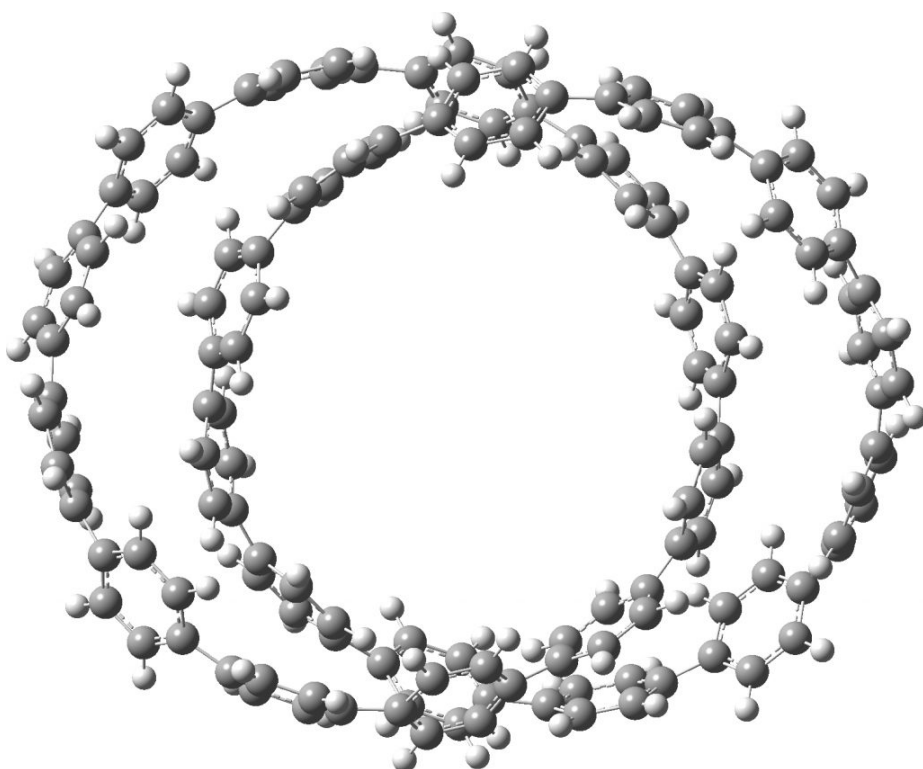
1a, which are here symmetrised to D_2 symmetry.

Molecule **2** posed a new challenge. The coordinates when extracted from the SI had 480 atoms, double that of **1a**. When displayed they overlayed each other. Clearly two sets of 240 atoms was the answer (the first probably in error) only the second set of which displays as a trefoil knot. The coordinates here can be symmetrised to D_3, as appropriate for a trefoil knot.
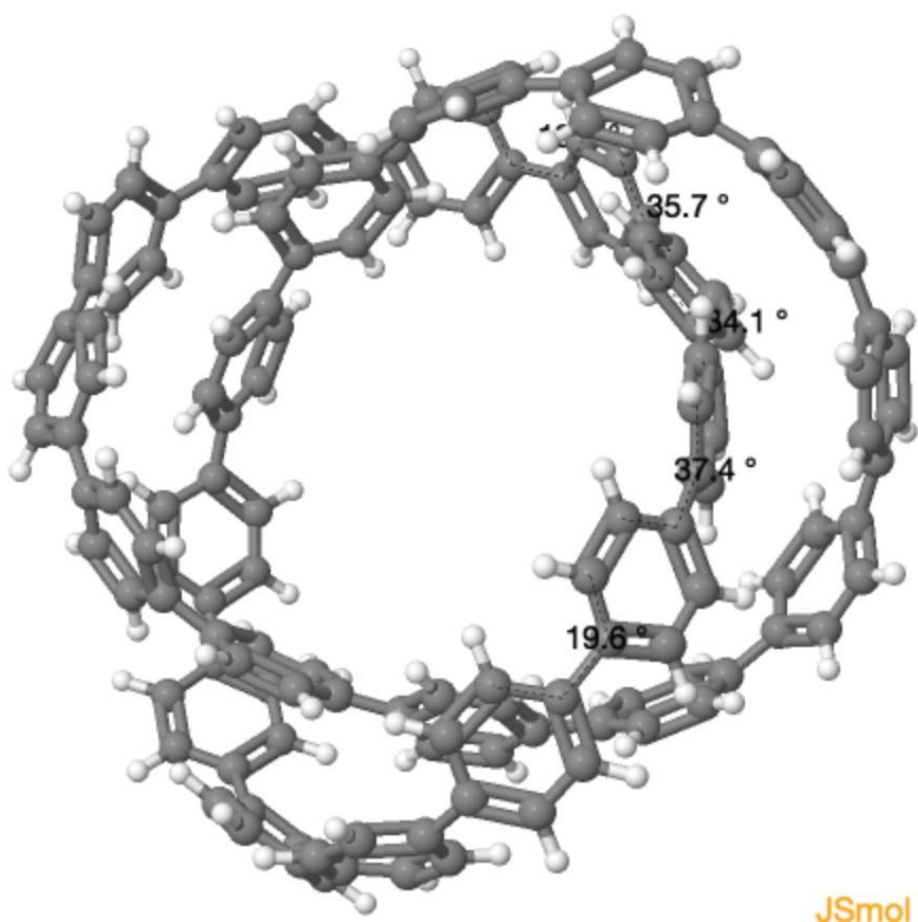
2, which here are symmetrised to D_3 symmetry

The above coordinates are computed using quantum mechanics at the B3LYP-D3/6-31G(d) level. What about the crystallographic coordinates? Here again a little expertise is needed to obtain these.

1. Just after the article acknowledgements, the CCDC identifiers are given as 1860595-7 and 1908693. These values are resolved using www.ccdc.cam.ac.uk/structures/ which allows the download of a CIF file.
2. These files reveal that a number of solvent and other molecules have been occluded into the structures, and for clarity it helps to edit all these out as well as disordered atoms with partial occupancy.
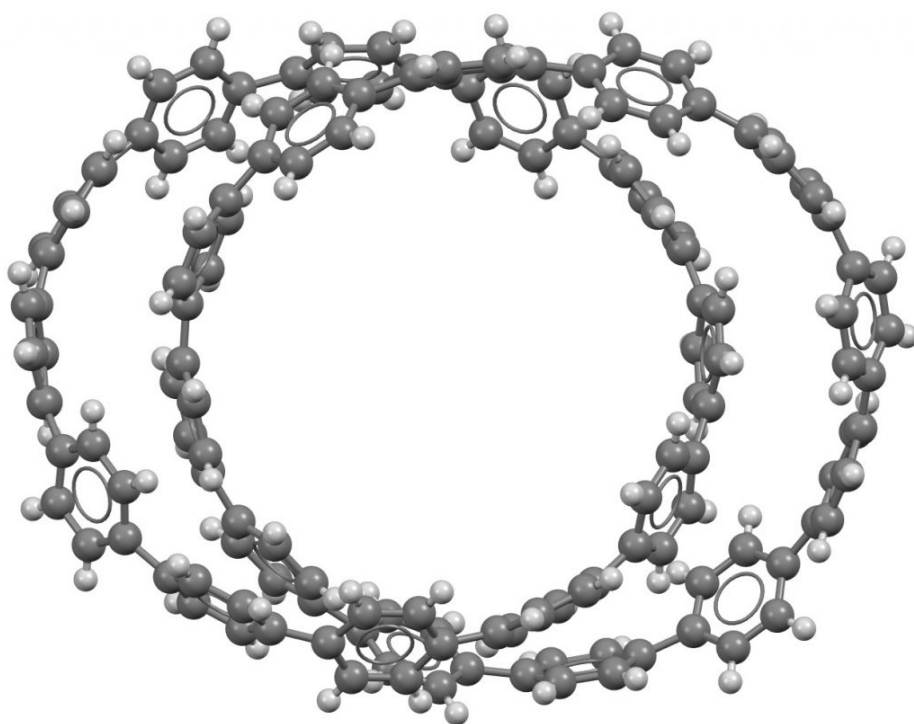
Crystallographic coordinates for 1a. DOI: 10.5517/ccdc.csd.cc20g36t
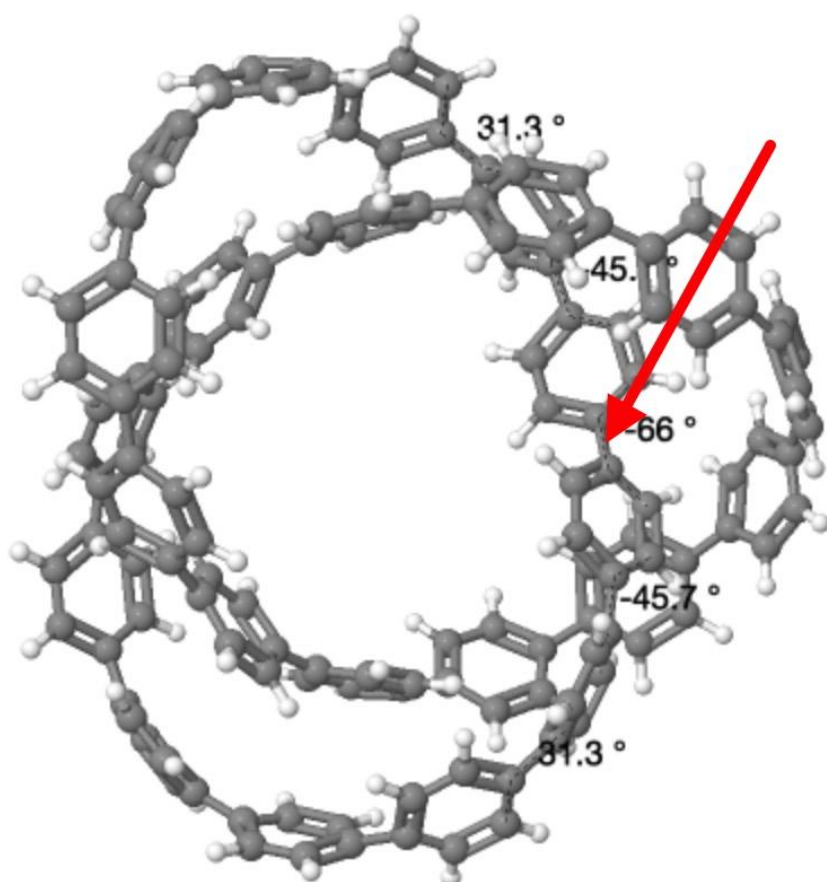
Crystallographic coordinates for 2. DOI: 10.5517/ccdc.csd.cc20g37v

Henry Rzepa's Blog

Each of these experimental structures is also allocated a DOI of its own, which can be accessed from the captions above, if you want to view the un-edited coordinates. I have also made available my coordinates produced for the display here as a FAIR dataset (DOI: 10.14469/hpc/6470), together with metadata such as the InChI descriptors to improve its discoverability.

Having acquired and displayed both the calculated and the measured coordinates, I noticed one oddity. The calculated structure for **2** symmetrises to D_3 symmetry, but the measured crystal structure only to C_2 symmetry. This is due to a "kink" in the twists for the latter coordinates, a single region where the Ar-Ar single bond is twisted by 66°. This kink is absent in the calculated coordinates, where the largest dihedral angle at the Ar-Ar bond is only 37°. Is this effect real? What does it tell us about the conjugations and extended aromaticity of this system? Such twist localisation has been previously noticed in cyclacenes.[cite]10.1039/B005560N[/cite] I think this observation highlights the need to have readily accessible 3D structures of such novel systems, if only to allow readers to spot such apparently anomalies.

So, with a little bit of knowledge and effort, one can indeed proceed from a published article to viewing aspects of the three-dimensional topology of the molecules discussed. I just feel it would be good if these aspects could be better integrated into the article itself, since I suspect that the additional effort and knowledge required to go further is probably not going to appeal to most readers.‡

---

‡Tidying up the PDF cartesian coordinates into a list of atomic number and a set of three coordinates per line of text is relatively simple. To coerce this format into a visualisation program takes more knowledge. Direct conversion to eg a standard molfile is not possible. I instead add a header to the coordinates to make it suitable for visualisation using the Gaussview program. Another good program for handling this is wxMacMolPlot which supports the veritable Xmol XYZ format, but again a correct header at the top of the file is needed for this program to recognise the file. As I noted, only a knowledgeable user would be able to do this, and the average reader is unlikely to go down this road.

https://orcid.org/0000-0002-8635-8390