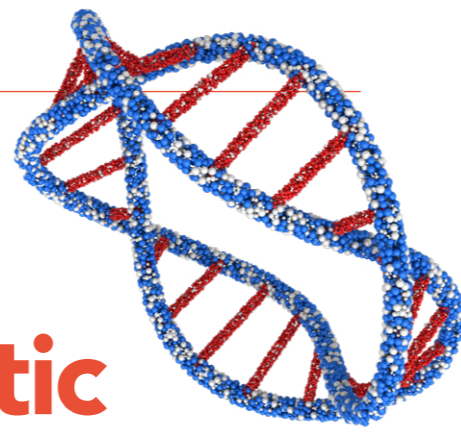


Producing the first functioning synthetic DNA using click chemistry

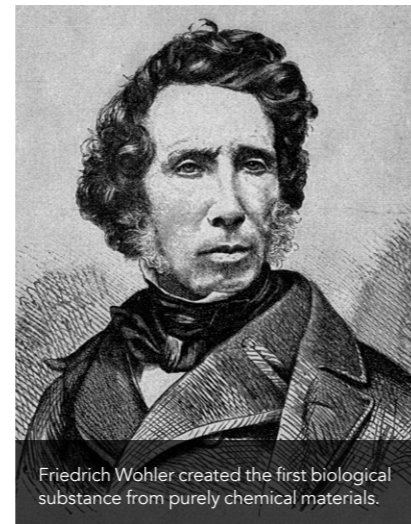


Click chemistry revolutionised our ability to create custom proteins, so could it do the same for DNA? Professor Tom Brown at the University of Oxford and Professor Afaf El-Sagheer at the University of Oxford and Suez University are performing truly groundbreaking research into the application of 'click chemistry' bioconjugation techniques to DNA. Their technique allows the creation of DNA and RNA constructs larger than any other chemical synthetic route and is more flexible than the use of enzymes. This is made possible by a novel chemical linkage which is biocompatible with DNA-manipulating cellular machinery in E coli and human cells.

Every so often an observation, experiment or discovery changes more than just scientific understanding. Darwin, Copernicus, and Newton's breakthroughs, while scientific, also had significance for philosophy, religion and metaphysics.

Perhaps Friedrich Wöhler's 1828 finding – that urea can be made from potassium cyanate and ammonium chloride – should also be included in this list. Urea, in his own words an 'animal substance', was the first to be created without any biological starting materials, signalling the end for the doctrine of vitalism.

In the decades since Wöhler, the list of biochemicals created using purely chemical means has grown longer, and the molecules themselves have grown larger and more complex. Notable early advancements include glucose,



Friedrich Wöhler created the first biological substance from purely chemical materials.

vitamin B12 and insulin; more recently larger molecules have become possible, including chemotherapies, antibiotics, and even proteins. But despite our progress, there is still no effective chemical synthesis of probably the most famous biomolecule of them all: Long DNA.

The modern technique for constructing DNA sequences involves using biological enzymes to ligate (join together) short synthetic sections of DNA, called 'oligonucleotides' or 'oligos'. Multiple oligos can be assembled into genes and plasmids.

Oligos are almost always produced using the solid-phase phosphoramidite method. The error rate in this method is reasonably low – enough to create oligos typically around 100 bases long. However, at around the 150 bases mark, errors start to become a problem. As a result, long segments of DNA require large numbers of oligos, requiring complex biochemical techniques.

Morphart Creation/Shutterstock



DNA is made up of bases which match in pairs. These base pairs form the 'rungs' of the ladder-like structure.

CHEMICAL LIGATION

Since 2007, Afaf El-Sagheer and Tom Brown have been pioneering an alternative approach to the chemical synthesis of large DNA and RNA strands. Their premise is that longer oligos could be created from shorter oligos using chemical ligation rather than enzymatic ligation.

But this is no mean feat: it's not possible for the chemical 'links' to be exact replicas of the phosphoester linkage found in natural DNA. Instead, the chemical links in these DNA replicas need to be 'biocompatible' – indistinguishable to natural enzymes from the real thing.

Finding a chemical reaction which produces a biocompatible chemical linkage is a major milestone in El-Sagheer and Brown's work. The researchers outline in their 2009 research how they use 'click chemistry', a set of chemical reactions known to efficiently form bonds between biomolecules and other molecules, a process known as bioconjugation.

CuAAC

El-Sagheer and Brown began working on developing a variation of the 'copper(I)-catalyzed azide-alkyne cycloaddition' (CuAAC) to find a method which is suitable for ligating oligomers. CuAAC is the 'classic' click chemistry reaction, described as "the premier example of a

El-Sagheer and Brown's triazole linkage is a unique example of a biocompatible artificial DNA linkage formed using chemical synthesis.

click reaction" by Sharpless, a pioneer of click chemistry techniques. CuAAC creates a triazole linkage – a five-membered ring of atoms containing two carbon and three nitrogen atoms. Initial experiments showed that it was possible to combine various oligomers using a version of CuAAC which joined two 'T' nucleotides with a triazole linkage.

These oligomers were then 'amplified' using PCR – the biochemical process used to replicate a DNA sequence. In PCR, the oligomer molecules containing the triazole linkage need to be 'read' by DNA polymerase, the key enzyme in the replication process. If the triazole linkage is biocompatible, then the PCR process should produce copies of the oligomer sequence. These resulting copies would not contain the triazole linkage.

LIMITATIONS OF CuAAC

However, this version of the CuAAC reaction had its limitations. While the DNA triazole linkage is read by DNA polymerase in PCR, the process is slow, and results in the deletion of one of the two nucleotides around the triazole. The

T-triazole-T sequence was converted to a single 'T' in the PCR product. In other words, the linkage is not a very good mimic of the phosphodiester linkage found in natural DNA. In addition, producing oligos with the correct precursors on each end for the ligation step was difficult to accomplish.

Although one base was missing, the remarkable achievement was that the PCR process, for the first time, successfully amplified triazole-link-containing oligomers.

DEVELOPING BETTER LINKAGES

With the promise of biocompatibility, El-Sagheer and Brown began redesigning similar linkages, applying their technique to both DNA and RNA.

In 2010 work, they demonstrated that a redesigned triazole linkage could be used to synthesise ribozymes, which are a type of cellular machinery formed from RNA. RNA is a similar molecule to DNA in many ways – it's a polymer composed of nucleic acids – but it does have some key differences which make long RNA

Pablo Joannopoulos/Shutterstock



Some oligonucleotides are used as fluorescent biomarkers which can be useful in medical diagnosis.

chains even more difficult to synthesise than DNA. Crucially, RNA contains an extra hydroxyl group at the sugar, which can lead to unwanted side reactions unless it is protected during solid phase synthesis and deprotection. This hydroxyl group is also a 'steric hindrance', meaning it physically gets in the way of molecules approaching the bond formation site.

As a result, RNA chains longer than 50 nucleotides are difficult to produce in a pure state, which doesn't come near to the length of biologically functional RNA constructs like ribozymes and riboswitches.

Using the CuAAC reaction, El-Sagheer and Brown were able to cross-link separate strands of oligomers together to form two RNA constructs, both almost 100 nucleotides in length. Both the constructs – a hairpin ribozyme with cross-links, and a hammerhead ribozyme with the improved linkage at the active site – functioned like their natural equivalents.

After establishing the biocompatibility of the new linkage in RNA in 2010, this improved linkage was used to ligate three 100-base-long oligomers to construct a DNA molecule of 300 bases in length, and amplified using PCR. Analysis of the resulting sequence confirmed the biocompatibility of the



The team tested their method using the bacteria *E. coli*.

new triazole linkage: the DNA was accurately copied, with nucleotides on both sides of the triazole linkage included in the oligomer copies.

IN VIVO STUDIES

However, the true test of biocompatibility is within cells. In a crucial collaboration with Ali Tavassoli and Pia Sanzone at Southampton University, oligomers were ligated using the CuAAC to form a triazole-containing piece of DNA known to code for antibiotic resistance, which was inserted into the DNA of *E. coli* cells.

Overnight incubation on agar plates containing ampicillin revealed exactly what the team was hoping for: the plates with the triazole DNA added contained 96.5% of the colonies; in the negative control it was only 1.6%. In other words, *E. coli* cells given artificial DNA were shown to be resistant to the antibiotic – and the triazole link is biocompatible with *E. coli* cellular machinery.

Repeating the test on *E. coli* cells deficient in DNA repair enzymes gave a similar survival rate, demonstrating that the success wasn't down to the linkage being 'corrected' by *E. coli*. The link is truly biocompatible with DNA-reading enzymes.

MEETING THE DEMANDS OF CRISPR

El-Sagheer and Brown's triazole linkage is a unique example of a biocompatible artificial DNA linkage formed using chemical synthesis. Almost 200 years on from Wöhler's synthesis of urea, the enzyme-free synthesis of DNA – or at least a functional mimic of it – is now possible, using click chemistry.

One area where the technique could excel is in the emerging and promising technology of CRISPR gene editing. The technique relies on creating RNA sequences called single guide RNAs (sgRNAs). When CRISPR-Cas9 leaves the laboratory and enters diagnostic and therapeutic use, sgRNAs will be in high demand. But current use of enzymes to produce a library of sgRNAs is both time-consuming and difficult to scale.

These short sections of RNA are formed of two parts. Firstly, a section about 80 nucleotides long which binds to the Cas9 protein, and is the same in all sgRNA molecules. Secondly, a section about 20 nucleotides long which is unique to the sequence of DNA near where the edit should be made. This guides the Cas9 protein to the correct part of the DNA molecule.

In their 2019 research published in *Nature Communications*, the group detailed how the CuAAC reaction can be used to produce a library of sgRNAs in a single tube – by ligating the two parts of sgRNA oligomers using click chemistry. The advantage of using this technique is that the Cas9-binding oligomer could be produced at a large scale, meaning only the 20-base oligomer needs to be tailored, and the two halves can be ligated.

The resulting sgRNA allowed gene editing in cells with no unexpected off-target effects. Perhaps CRISPR will be where the triazole linkage leaves its mark on DNA chemistry.

Behind the Research



Professor Afaf El-Sagheer



Professor Tom Brown

E: afaf.el-sagheer@chem.ox.ac.uk T: +44 1865 275175 E: tom.brown@chem.ox.ac.uk T: +44 1865 275413
W: <http://www.browngroupnucleicacidsresearch.org.uk/>

Research Objectives

Professors El-Sagheer and Brown investigate the chemical synthesis of long DNA and RNA strands.

Detail

Tom Brown
Department of Chemistry
University of Oxford
Chemistry Research Laboratory
12 Mansfield Road
Oxford OX1 3TA
UK

Afaf El-Sagheer
Department of Chemistry
University of Oxford
Chemistry Research Laboratory
12 Mansfield Road
Oxford OX1 3TA
UK
and
Chemistry Branch
Department of Science and Mathematics
Faculty of Petroleum and Mining
Engineering
Suez University
Suez 43721
Egypt

Bio
Tom studied for his chemistry degree and PhD at Bradford University. He was appointed as a Lecturer then Professor at Edinburgh University before moving to Southampton University in 1995 then to Oxford University in 2013. He is co-founder of three Biotech companies (Oswel, Primer Design, ATDBio). He has published over 400 papers and patents.

Afaf studied chemistry at Suez Canal University (Egypt) and did her PhD at Southampton University with Professor John Mellor then moved back to Egypt to become a lecturer. She was promoted to Associate Professor in 2009 then Professor in 2014 at Suez University. Afaf has been on a long sabbatical as an academic visitor at Southampton and Oxford universities since 2005. She has over 150 publications.

Afaf and Tom's research interests are in oligonucleotide synthesis and its applications in chemistry, biology and medicine (diagnostics and therapeutics).

Funding
Biotechnology and Biological Sciences Research Council (BB/J001694/2 and BB/R008655/1)

Collaborators
• Ali Tavassoli
• Pia Sanzone
• Lapatrada Taemaitree
• Arun Shivalingam
• Agnes Tyburn

200 years on from Wöhler's synthesis of urea, the enzyme-free synthesis of long DNA – or at least a functional mimic of it – is now possible.



The team's CuAAC reaction could help produce key elements needed for CRISPR gene editing technology.

References

Taemaitree, L., Shivalingam, A., El-Sagheer, A.H. and Brown, T. (2019). *Nat. Commun.* 10, 1610.
El-Sagheer, A.H. and Brown, T. (2017). *Chem. Commun.* 53, 10700.
El-Sagheer, A.H. and Brown, T. (2012). *Accounts of Chemical Research.* 45 (8), 1258-1267.
El-Sagheer, A.H. et al. (2011). *PNAS.* 108 (28), 11338-11343.
El-Sagheer, A.H. and Brown, T. (2010). *PNAS.* 107 (35), 15329-15334.
El-Sagheer, A.H. and Brown, T. (2009). *J. Am. Chem. Soc.* 131, 3958-3964.

Personal Response

What would it take for this procedure to become a commercially available product?

/// The basic research has been carried out and patents have been filed. The next step is for an industrial partner to work on applications that could benefit from the approach. This could be for example in CRISPR/Cas but there are many other potential areas where click-ligation could offer advantages, such as modified genes, especially with epigenetic modification which can play a vital role in gene therapy. ///