STUART CROSBY
Evonetix Ltd, Cambridge, United Kingdom

Unlocking synthetic biology through DNA synthesis

KEYWORDS: DNA synthesis, error correction, gene, nucleotide, oligo, synthetic biology, biotechnology.

ABSTRACT

DNA synthesis is an essential part of synthetic biology, providing access to custom-made oligonucleotides (oligos) and double-stranded DNA, offering the potential to design new genes from scratch for an exciting range of applications, including pharmaceuticals, renewable energy and data storage. Synthesis usually proceeds via sequential chemical addition of nucleotide bases to a growing strand on a solid-phase support, before cleavage and hybridisation to generate double-stranded DNA. These methods are invaluable in research but become increasingly time-consuming and error prone as more bases are added.

The anticipation of a synthetic biology revolution has created huge demand for high-quality oligonucleotides and double-stranded DNA, resulting in several companies developing novel approaches to DNA synthesis. This article will outline the fundamentals behind DNA synthesis, including some recent advances in the field which have sought to overcome caveats by reducing synthesis costs and increasing sequence fidelity. In addition to this there will be some discussion of the potential for DNA synthesis to deliver solutions to current global challenges.

INTRODUCTION

Synthetic Biology

Synthetic biology is a term used to describe an approach by which novel artificial biological pathways, organisms or devices are designed and constructed (1). Today, many would argue that synthetic biology is at the cusp of enabling a multitude of major breakthroughs in the production of, for example, renewable chemicals, biofuels, pharmaceuticals and food ingredients. Already synthetic biology is positively impacting many of our present and future needs in medicine including the development of novel vaccines (e.g. antigen genes or nucleic acid vaccines), gene therapy (with its requirement for synthetic DNA), CRISPR-Cas9, and other cellular engineering techniques. For example, the field of immuno-oncology has yielded chimeric antigen receptor technology (CAR-T), a cancer cell therapy which engineers a patient's own immune cells (T-cells) to recognise and attack cancer cells and this approach has already shown significant benefits for certain patients (2).

Synthetic biology is enabled by the vast amounts of DNA sequence information available from a huge array of organisms, and the availability of genomic and metagenomic tools to understand the potential functionality of such information. However, the true scale of its impact relies on the continued development of its underlying technologies, including bioinformatics, the identification of suitable host organisms and reliable methods of designing biosynthetic pathways, and importantly, DNA synthesis.

An Introduction to DNA Synthesis

DNA synthesis is a powerful technique that requires the joint expertise of chemists, biologists and materials scientists, offering researchers the power to better understand and manipulate the fundamental building blocks of life.

In nature, genomic DNA is replicated by 'unzipping' the double-stranded helix with DNA helicase and building up complementary strands recruiting nucleotides with DNA polymerase; these strands later come together to form a new double helix. Biochemists can replicate this process with the same enzymes in vitro by carrying out the polymerase chain reaction (PCR)—a technique that won its creator Kary Mullis the Nobel Prize for Chemistry in 1993, which he shared with Michael Smith for the development of site-directed mutagenesis (3). These complementary methods allow scientists to rapidly produce many copies of genes and introduce into them specific mutations which allow alternative genes to be engineered. The development of these molecular cloning approaches, which became fundamental tools for molecular biology, represented a critical development in the emergence of synthetic biology (4), allowing researchers to build modified genes and proteins in order to better understand the relationship between chemical structure and biological function.

Molecular cloning techniques form the foundation of much of modern research, and most bioscience labs are equipped with their own PCR machine. Molecular cloning relies on the synthesis of short DNA sequences to insert into host organisms for an array of potential uses—a process which is typically outsourced to specialist companies and remains inaccessible to conventional researchers, although companies such as UK-based Evonetix are looking to change this (Figure 1). Some scientists are interested in going beyond modifying existing genes and want to create DNA for broad-ranging downstream applications. This would allow us to take advantage of DNA's potentially useful physicochemical properties (5), (6), in addition to enabling the creation of entirely novel genes, proteins and materials.
One particularly intriguing example is 'DNA origami', which exploits the molecule's hydrogen bonding motif to fold DNA strands into specific shapes, held together by short complementary 'staple' strands (7). By synthesising oligos with carefully engineered sequences, it is possible to fold DNA into shapes ranging from two-dimensional 'smiley faces' to three-dimensional wires and cages (7). This could enable the creation of nanoparticles with complex phase chemistry that are bio-compatible and stable in a physiological environment—with potential applications in drug delivery, electronics and materials science, among other areas (7-9).

Furthermore, DNA synthesis can be used to design structural variants of proteins, e.g. biotherapeutics, to enhance their properties (8). This may be achieved through structure-based design of modified DNA sequences, or by synthesising large libraries of sequences, and selecting for those that achieve the desired phenotypic outcome. The latter approach of directed evolution essentially mimics natural selection and requires methods that can synthesise DNA in a high-throughput manner to reach its full potential.

Some biopharmaceutical companies are combining DNA synthesis with computational design to engineer entirely new proteins.

Figure 1. A benchtop DNA writer, currently in development, aims to place DNA synthesis in the hands of every researcher (Credit: Evonetix).

For example, the development of a novel enzyme to degrade gluten, a component of wheat that triggers an autoimmune response in individuals with Coeliac disease (10), and the generation of protein variants of phenylalanine hydroxylase (PAH) with gastrointestinal stability to treat phenylketonuria, a rare, autosomal recessive disorder (11).

Biotechnology companies are also aiming to use DNA synthesis to help to develop alternative, biology-based methods for industries that typically use petroleum-based products as inputs and produce carbon emissions as outputs. Biofuels are generally produced through the microbial digestion of plant material to generate useful products such as ethanol. However, most plants contain a large proportion of lignocellulosic biomass, which is difficult to break down into digestible sugars (12). To help address this, one research group genetically engineered poplar trees, introducing chemically labile ester bonds into the lignin that made up their cell walls (13). The introduction of these bonds did not affect the growth of these trees but made their lignin much more amenable to enzymatic digestion, demonstrating the potential to address a major limiting factor in biofuel production.

As illustrated through the examples above, synthetic biology is capable of delivering new solutions for global healthcare, agriculture, manufacturing, and environmental challenges. However, DNA synthesis has not made the same progress as had DNA sequencing. In the same way that DNA sequencing underwent a revolution in the mid-2000s resulting in massive parallelisation and reduction in cost, a comparable breakthrough in DNA synthesis would catalyse many opportunities in these areas.

Principles of DNA Synthesis

The current paradigm of DNA synthesis uses phosphoramidite chemistry to build single-stranded oligos, which are later annealed to generate double-stranded DNA (usually 5-50 oligos provide the raw substrate for constructing 200–3,000bp). The phosphoramidite-based methods for DNA synthesis were first developed in the 1980s by Caruthers et al (14) and have since remained largely unchanged (5), (15). Some methods did, however, predate this, including Robert Letsinger's successful attempt to synthesise DNA on styrene supports in 1965 (16).

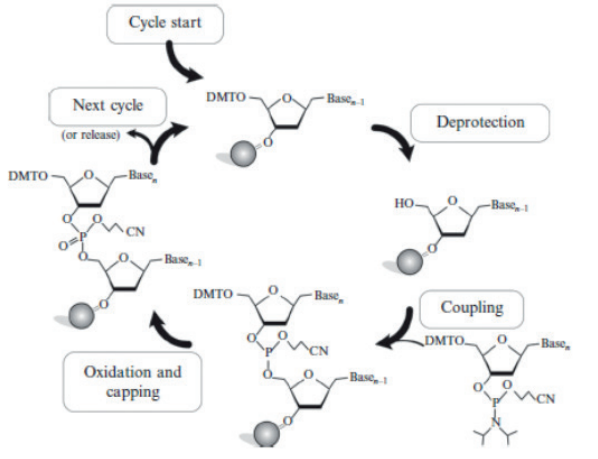

Figure 2. The phosphoramidite method: the current gold standard for oligo synthesis (15).

The typical phosphoramidite process relies on a four-step cycle (Figure 2) that adds nucleotides one-by-one to a growing single strand, attached to a solid support of controlled pore glass or polystyrene beads. In most commercial DNA synthesisers, the solid support is packed into a continuous flow column or reaction well, into which the reagents needed to build the sequence are added.Nucleotides are added in a 3'-5' manner, with the 5'-hydroxy (OH) group lying at the end of the chain and the 3'-end attached to the surface via a cleavable linker. The 5' terminus must be kept protected to prevent uncontrolled addition until the next nucleotide is ready to be added, and the 5'-protecting group must be removed under conditions which are orthogonal to the cleavage of the completed oligo from the surface at the end of the synthetic process. Once the protecting group is removed under acidic conditions, the exposed 5'-OH can then couple to the next nucleotide, which is attached to a phosphorous (P)(III) group to provide electrophilic reactivity. Before proceeding further, any unreacted 5'-OH groups are capped to prevent the growth of chains missing a nucleotide (shortmers), producing truncated strands which are easier to remove. At this stage, the unstable P(III) is also oxidised to P(V). Once this is complete, the next nucleotide in the sequence can be added, and the process continues.

Following stepwise synthesis of the desired oligo, the product is cleaved from the solid support and the remaining exocyclic P(V) protecting groups are removed in a single step, using a strong base such as ammonium hydroxide. To make a DNA product from these individual oligos, they must be purified to remove errors and then ligated to give the complementary strands which are hybridised to generate double-stranded DNA.

A more recent advancement and alternative to the method described above is microarray oligo synthesis. This approach was pioneered by Affymetrix, whose researchers developed methods for spatially localized polymer synthesis on surfaces using light-activated chemistries. Using standard mask-based photolithographic techniques, they selectively deprotected photolabile nucleoside phosphoramidites (15). Others have since simplified this procedure, for example by eliminating the photolithography masks and using programmable micromirror devices to precisely control the light-based chemistries (17). Others are now modifying the phosphoramidite cycle to become thermally controllable. By using a thermosensitive protecting group at the 5'-position, base addition is thermally controlled, and researchers can distinguish between reaction sites which otherwise see exactly the same chemical conditions. Although this method remains under development, preliminary results have indicated that the temperature control achieved is highly accurate, allowing the researchers to carry out oligo synthesis on a scale which is compatible with standard laboratory analysis such as LC-MS.

Limitations of current methods

The typical phosphoramidite method (Figure 2) has some inherent limitations causing it to become a bottleneck to scale for larger gene synthesis and genome assembly projects. These issues will be challenging to overcome without a fundamental rethink of the chemistry involved.

Firstly, the process is not amenable to extremely high-throughput applications, as the number of strands that can be sequenced in parallel is fairly low. This is because, in conventional DNA synthesisers, every reaction well remains in the same chemical environment. Therefore, the number of different sequences that can be produced at any one time is directly determined by the number of reaction wells available.

Secondly, the likelihood of introducing errors increases as more nucleotides are added. Even though each nucleotide coupling proceeds with around 99% accuracy, the yield of the pure DNA sequence decays by the accuracy of each coupling raised to the power of the number of nucleotides added. This means that by 100 base pairs, the typical yield is only 37 % (5) and purification to remove the error-containing

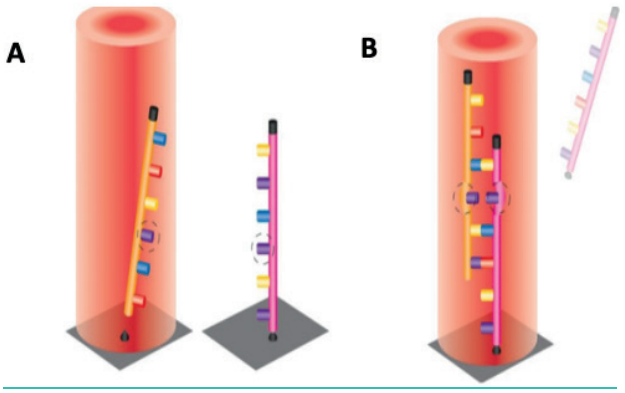

Figure 3. Schematic of the thermal error-correction process. (A) Oligos are heated to selectively remove them from the surface. (B) Oligos hybridise to complementary strands – these sites are heated, causing strands with significant mismatch to dissociate.

oligos is time-consuming and laborious. As most genes comprise tens of thousands of base pairs, it is therefore necessary to synthesise and purify hundreds of short oligos separately, then stitch them together using DNA ligase (6).

In addition, the yield of DNA synthesised can be affected by depurination, particularly at adenosine sites, which can occur during acidic detritylation. This can cause a decline in yield of full-length oligos by promoting the cleavage of the oligo phosphate backbone during the removal of the remaining nucleobase and backbone protecting groups following the final synthesis cycle (15). Single-base deletions, resulting from either failure to remove dimethoxytrityl or combined inefficiencies in the coupling and capping steps, are also a large source of errors for purified oligos (20).

The use of microarrays has proven popular as a cheap source of oligos. Oligos produced in this way are synthesised at femtomolar scale and require reduced amounts of reagent, thus being 2–4 orders of magnitude cheaper than conventional approaches. Costs range from $0.00001–0.001 per nucleotide, depending on length, scale and platform (15). However, there are several challenges when using microarray methods for gene synthesis, and the issue of high error rates remain (21).

Enabling significantly higher-fidelity double-stranded DNA synthesis in comparison to current approaches, where high error rates represent a major limitation, is a significant challenge for modern technologies. Thermal control of oligo synthesis and assembly offers significant advantages as it allows errors to be detected and removed during assembly, as opposed to post-synthesis and purification (Figure 3). Using this method, DNA mismatching is easily identified as once oligos are annealed, the site is heated to a precise, sequence-dependent temperature at which dissociation will be faster for imperfect matches, allowing the population of oligos to be purified by discarding imperfect oligos. This allows for longer sequences with low error rates, preventing bad sequences from diluting the population of correct molecules (6).

CONCLUDING REMARKS

Synthetic biology has the potential to deliver a wide range of new products to several industries, such as renewable chemicals, biofuels and pharmaceuticals. Crucial for the achievement of many of these opportunities is the further development of synthetic DNA synthesis, which today is still largely based upon a method developed last century.

While companies have made great progress in upscaling DNA synthesis, there remain challenges in achieving further scale, greater accuracy and reduced cost. In response to growing market pressures, new approaches are being developed, including a method of thermally controlled synthesis, which could enable the successful synthesis of even the most difficult target sequences in the not-so-distant future. Commercialisation of such technologies and the continued development of other efficient DNA synthesis techniques will further enhance the availability of increasingly inexpensive synthetic DNA.

REFERENCES AND NOTES

1. Synthetic biology – does industry get it? Conference report. Royal Society (2017). Available at: https://royalsociety.org/-/media/events/2017/02/tof-synbio/synthetic-biology-post-conference-report.PDF
2. June, C. H., O'Connor, R. S., et al., Science 359(23), 1361–1365 (2018).
3. The Swedish Academy of Sciences. Decisive progress in gene technology through two new methods: the polymerase chain reaction (PCR) method and site-directed mutagenesis. (1993). Available at: https://www.nobelprize.org/prizes/chemistry/1993/press-release/
4. Cameron, D., Bashor, C. and Collins. Nature Reviews Microbiology, 12(5), 381-390 (2019).
5. Extance, A. Chemistry World, (April 2018). Available at: https://www.chemistryworld.com/features/step-by-step-synthesis-of-dna/3008753.article
6. Kukwikila, M., Gale, N., et al., Nature Chemistry, 9(11), 1089-1098 (2017).
7. Hu, Y. and Niemeyer, C. Advanced Materials, 31(26), 1-18 (2019).
8. Brears, T. European Biopharmaceutical review, (Spring 2019)
9. Udomprasert, A. and Kangsamaksin, T. Cancer Science, 108(8), 1535-1543 (2017).
10. PVP Biologics. (2017). https://www.takeda.com/newsroom/newsreleases/2020/takeda-acquires-pvp-biologics-following-results-of-a--phase-1-study-of-tak-062-kuma062-for-the-treatment-of-celiac-disease/ (last checked on Apr. 1st 2020)
11. Codexis, Biotherapeutics https://codexis.com/industries/biotherapeutics/ (last checked on Apr. 1st 2020)
12. Marguet, P., Balagadde, F., et al., Journal of The Royal Society Interface, 4(15), 607-623 (2007).
13. Wilkerson, C., Mansfield, S., et al., Science, 344(6179), 90-93 (2014).
14. Caruthers, M. H., Beaucage, S. L., et al., Gene Amplification and Analysis, 3, 1–26 (1983).
15. Hughes, R., Miklos, A., et al., Methods in Enzymology, 9(1), 277-309 (2011).
16. Robert L. Letsinger and V. Mahadevan. Journal of the American Chemical Society, 879(15), 3526-3527 (1965)
17. Gao X., et al., Nucleic Acids Research, 29(22), 4744–4750 (2001).
18. SynBioBeta: https://synbiobeta.com/building-dna-on-a-chip-a-novel-approach-to-gene-synthesis/ (last checked on Apr. 1st 2020)
19. Chmielewski, M. Tetrahedron Letters, 53(6), 666-669 (2012).
20. Binkowski, B.F., Richmond, K.E., et al., Nucleic Acids Research, 33(6), e55 (2005).
21. Ma S., Tang N., et al., Current Opinions in Chemical Biology, 16(3-4), 260–267 (2012). ■

ABOUT THE AUTHOR

*Dr **Stuart Crosby** has over 20 years' research experience working in the fields of synthetic chemistry and drug discovery. He is now Head of Synthetic Chemistry at Evonetix Ltd, where he is developing the chemistry that underpins Evonetix's third generation DNA synthesis platform.*