It is hard to overestimate the potential of biotechnology to address some of the most pressing challenges of the 21st century. These challenges include feeding a growing global population, improving healthcare and access to therapies, and mitigating the impacts of climate change.

Considering the “data-rich” nature of biological experimentation and heavy reliance on such data, no wonder why artificial intelligence (AI) is playing a crucial role in advancing synthetic biology by facilitating the design, construction, and optimization of biological systems. Some of the key ways AI is used to advance synthetic biology include:

**CRISPR-based gene editing**

Artificial intelligence plays a significant role in enhancing gene editing using CRISPR-Cas9 technology. CRISPR-Cas9 is a powerful gene-editing tool that allows researchers to make precise edits in the genome by adding, deleting, or altering sections of the DNA sequence. However, one of the challenges in CRISPR-Cas9 technology is the prediction and minimization of off-target effects, which are unintended modifications in DNA sequences that are similar to the intended target site.

AI can help address this issue by analyzing vast amounts of genomic data to predict potential off-target effects and their likelihood, thus guiding researchers towards more accurate and efficient gene editing. Machine learning algorithms can be trained on large datasets of genomic sequences and CRISPR-Cas9 cutting profiles to predict off-target effects based on the similarities between the target and off-target sequences, as well as other factors like chromatin accessibility.

Moreover, AI can assist in identifying optimal target sites for CRISPR-Cas9 editing by analyzing the genomic context, functional annotations, and potential off-target sites. This enables researchers to select target sites with minimal off-target risks and higher editing efficiency.
Another aspect where AI can contribute to CRISPR-Cas9 technology is the optimization of guide RNA (gRNA) design. The gRNA is a crucial component in the CRISPR-Cas9 system, responsible for guiding the Cas9 nuclease to the target DNA sequence. AI algorithms can be employed to analyze sequence features, predict gRNA binding efficiency, and suggest optimal gRNA sequences for a specific target, improving the overall gene editing efficiency and specificity.

In this context, an interesting company is Synthego, which is a provider of CRISPR genome engineering solutions, using machine learning algorithms to analyze and predict optimal guide RNA (gRNA) designs, minimizing off-target effects and maximizing editing efficiency.

Another company, Inscripta is a gene editing technology company developing the Onyx platform, a fully automated benchtop instrument for high-throughput gene editing. Their advanced algorithms and machine learning models allow for the rapid design, optimization, and execution of CRISPR experiments, streamlining the process and improving results.

**Gene sequence optimization**

Gene sequence optimization leverages AI algorithms, specifically machine learning and deep learning models, to analyze large amounts of genetic data and determine the ideal gene sequences for targeted biological functions. By doing so, researchers can engineer synthetic genes with increased efficiency and stability.

One prominent company in this field is Benchling, which offers a cloud-based platform for life science research, enabling scientists to design, edit, and simulate gene sequences using machine learning algorithms. Another example is DNA2.0, now known as ATUM, which uses its proprietary GeneGPS™ technology to design genes optimized for expression in any host organism. Twist Bioscience is also at the forefront of gene sequence optimization, harnessing silicon-based DNA synthesis to generate optimized, high-quality genes for various applications.

**Protein design**
Protein design and engineering utilize AI to forecast the structure and function of proteins according to their amino acid sequences, empowering scientists to create innovative proteins with specific characteristics. This process can result in the creation of new enzymes, therapeutics, and biomaterials. A key player in this domain is DeepMind, the company behind AlphaFold, an AI system that accurately predicts protein structures based on amino acid sequences, revolutionizing protein structure prediction. Rosetta by the Institute for Protein Design is another powerful computational tool for protein structure prediction and design, enabling the creation of custom proteins. Additionally, Zymergen employs AI-driven techniques to design and engineer novel proteins, with applications in industries such as agriculture, healthcare, and materials science.

Finally, Meta, the company behind Facebook and Instagram, unveiled the ESM Metagenomic Atlas, which contains the structures of over 600 million putative proteins. Although the Meta AI algorithm, ESMFold, may not be as precise as DeepMind's AlphaFold, it boasts a faster processing time. This increased speed stems from the algorithm's protein structure prediction method, which employs a language model trained on sequence data, or the order of amino acids in the linear chains constituting proteins.

As a result of its accelerated processing capability, researchers were able to predict the 600 million structures within a mere 2-week timeframe, utilizing a network of roughly 2,000 graphics processing units.

**Metabolic pathway engineering**

Metabolic pathway engineering harnesses AI to pinpoint and enhance metabolic pathways for generating specific compounds, including biofuels, pharmaceuticals, and various chemicals. By examining the interplay among genes, enzymes, and metabolites, AI algorithms can forecast the most effective routes for synthesizing target molecules.

One pioneering company in this field is Ginkgo Bioworks, which utilizes AI-driven techniques to engineer microbes for producing valuable compounds with industrial applications. Another notable example is Inscripta, which offers the Onyx™ platform for designing and engineering microbial strains to optimize metabolic pathways for targeted molecule production. Additionally, Zymergen employs AI and automation to engineer microbes and optimize metabolic pathways, with applications spanning agriculture,
healthcare, and materials science.

Automated experiment design

AI can help optimize experimental designs in synthetic biology, predicting which experiments are most likely to yield valuable results and guiding researchers in making more informed decisions.

A leader in this area is Synthace, which offers the Antha platform, cloud-based software for designing, simulating, and executing biological research workflows using machine learning algorithms. Another important player is Arzeda, a company that employs AI and computational protein design to construct custom genetic circuits for various industrial applications.

TeselaGen is deploying the first solution that incorporates state-of-the-art AI-enabled information technology for biotechnology as a secure enterprise operating system. Powered by the Synthetic Evolution® machine learning engine, TeselaGen standardizes, analyzes, and integrates data from various sources, and allows design high throughput, high content experiments.

To wrap things up, AI has the power to bring about significant changes in synthetic biology, helping us tackle some of today's most urgent problems. However, it's important to remember that this collaboration between disciplines comes with its own set of challenges. The complicated nature of biological systems, the limitations of the data we have, the struggle to validate models, and the need for cooperation between various fields all present hurdles to overcome.

- ATUM
- Arzeda
- DeepMind
- Ginkgo Bioworks
- Inscripta
- Synthace
- Synthego
- TeselaGen
- Twist Bioscience
- Zymergen