APPLICATION NOTE

# Rapid Improvement of Enzyme Performance with LYNX

BM-AN0003-RevA

biomatter

# LYNX Enzyme Optimization™)

**LYNX Enzyme Optimization™** is a breakthrough solution for enzyme improvement with industry-leading speed and performance.

Built on Biomatter's proprietary, experimentally validated **AI- and Physics-based** Intelligent Architecture™ platform.

Designed to provide **simultaneous** property improvements for enzyme activity, stability and production yield as well as redesigning enzyme sequence without loss of performance.

With no prior experimental data required, LYNX has successfully transformed enzymes for numerous Biomatter partners in a **single design round**, testing fewer than 30 variants the lab.

## I. Unlock the Power of High-Performance Enzymes with LYNX

Synthetic Biology is among the most important tools we have to tackle the many challenges facing humans in the coming decades. Despite the significant advances that have been made over the last half a century, we remain far away from realising the full potential of biology due to the high complexity of the task at hand.
Generative AI has transformed multiple industries by moving beyond human intuition to solve problems. Biomatter pioneered the use of generative AI in enzyme engineering, revolutionizing it from an unpredictable, slow, and resource-intensive process into an efficient and reliable design task. As one of the leaders in generative protein design, **Biomatter** has created a cutting-edge solution for **enzyme optimization.**

**LYNX Enzyme Optimization™** delivers remarkable enzyme performance improvements with incomparable speed and efficacy. Beginning with just a single sequence as an input, the unique blend of generative AI & physics-driven models does all of the engineering without the need for prior experimental data. LYNX Enzyme Optimization™ is built to improve enzyme **activity**, enhance **stability**, increase enzyme **yield** as well as **redesign entire sequences** of enzymes without compromising function. More importantly, any combination of these **properties** can be optimized simultaneously. The package is enzyme-agnostic and designed to work regardless of enzyme size, oligomeric state or origin.

Design with LYNX takes around 5 days, and it is built from the ground up to deliver optimized enzymes in a single round of engineering. Our scoring and selection system reduces the number of **variants** that need to be tested experimentally to **fewer than 30**, saving both time and resources. Finally, our customers retain full ownership of the generated IP.

**Table 1.** Key features of LYNX Enzyme Optimization™

| |
|---|
| Improve diverse enzyme properties, including activity, stability, yield, or gain FTO |
| Take full ownership of the generated enzyme IP |
| No experimental data required, only the sequence of the starting enzyme |
| Enzyme-agnostic |
| Only a handful of enzyme designs need to be tested (10-30 enzymes) |
| Designs obtained through LYNX cover a large sequence space |

## II. How does LYNX Work?

**LYNX Enzyme Optimization™** is powered by a proprietary generative AI and physics-based modeling technology that Biomatter pioneered and has continued to develop since 2018. Over the last five years, we have continuously improved our platform by using experimental testing to guide the development of new algorithms that in turn improve our predictions. These proprietary AI- and physics-based models work in tandem to generate and predict the performance of novel enzyme sequences, streamlining the process of enzyme optimization.
Thanks to the extensively validated combination of algorithms for sequence generation and scoring we can enhance enzymes in a single step without any need for prior experimental data. As a result, LYNX is **the fastest and most reliable method** to optimize enzymes of any type.

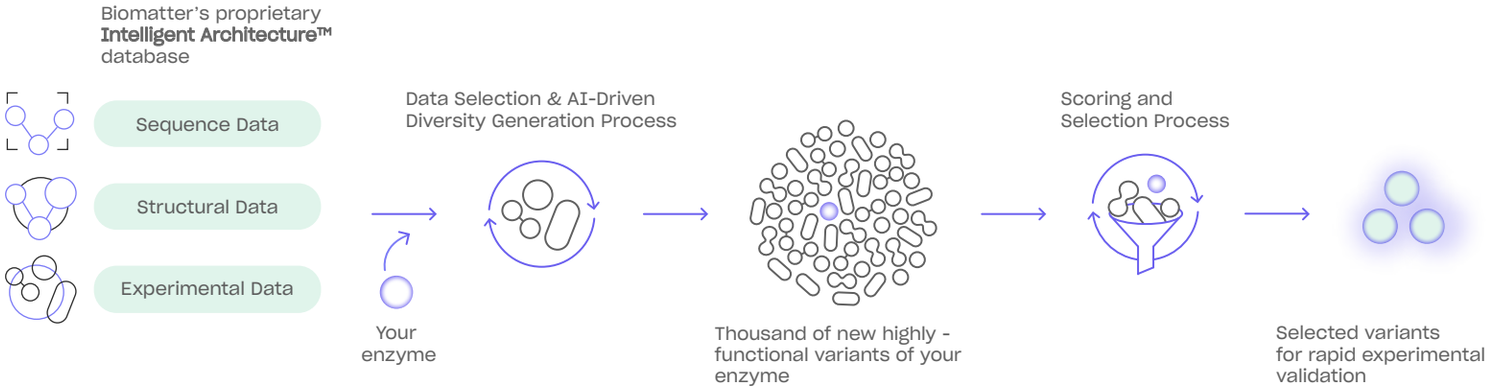



**Figure 1.** An overview of how LYNX works. Data from public and private databases, including sequence and structural data, as well as optional experimental data from previous engineering rounds are used to train our models. The models then generate novel enzyme sequence candidates, score and rank them to select the most promising sequences for testing.

### 1. Training AI models

**Biomatter's pioneering** research on generative AI was showcased in the publication of [the world's first functional enzymes entirely designed by AI.](#) Since then, we have worked to continuously refine our models by introducing new algorithms, enhancements to the model architecture, training and dataset preparation methodologies, all of which are **tested and validated experimentally.** Now, trained on billions of data points, our models offer a robust protein design framework for enzyme optimization, with the option to be further refined by incorporating experimental data when available.

### 2. Sequence generation

Beginning with a single sequence of the starting-point enzyme, AI models generate a diverse array of novel enzymes. The generative nature of **LYNX** enables the introduction of insertions and deletions, along with dozens of mutations across the entire enzyme sequence. This approach facilitates the discovery of epistatic interactions that would be impossible to access through classical enzyme engineering approaches. The result is a set of **highly diverse synthetic enzyme variants** that enables us to probe a large sequence space and gain insights into the sequence-function relationships for different enzyme properties.

### 3. Scoring and Selection

The generated enzyme variants are scored using a **combination of AI- and Physics-driven algorithms** to identify the most promising candidates based on the optimization objectives. The scoring process assesses **multiple enzyme properties**, including stability, solubility, expression level, activity, structure packing, and sequence-structure compatibility.

During the selection, enzyme variants with the best combination of the desired features are prioritized. We also evaluate the diversity of the designed enzymes to enable a thorough exploration of the enzyme sequence space.

### 4. Testing

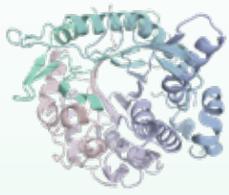
Based on the scoring, we select **10 to 30 optimized enzyme candidates**, which significantly reduces the number of variants that need to be characterized experimentally. In most cases, this number is sufficient to find significantly improved enzyme variants in the first design round. This extensively validated combination of sequence generation and scoring is what enables us to develop improved enzymes in a single step without the need of prior experimental data. As a result, the development time of improved enzymes is remarkably fast.

# III. Examples of Enzyme Optimization with LYNX

To showcase the effectiveness of the **LYNX Enzyme Optimization™**, we have assessed its performance in real-world enzyme optimization scenarios. These case studies span a range of enzyme optimization objectives, from improving a single property to optimizing multiple enzyme features simultaneously to re-designing enzyme sequences in order to gain freedom to operate. In all cases, the results were obtained from a **single LYNX design round** and without using any prior experimental data. The results represent real values measured through experimental characterization, not computational predictions, and include the data from all the tested variants without omitting designs that did not perform well.

## Improving melting temperature by 35°C

Beta - Glucosidase
(Prokaryotic, 448 amino acids)



LYNX75Y
Best performing enzyme variant in the LYNX campaign

- **25°C increased melting temperature**
- 7.5x improved activity
- 71% identity to starting point

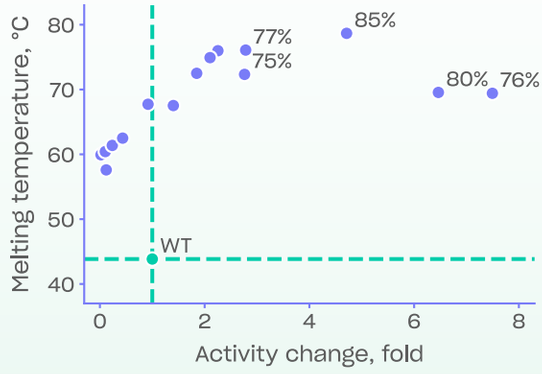



**Figure 2.** The activity and stability measurements for the designed beta-glucosidase enzymes. The starting enzyme is depicted in teal, and the designed enzyme variants are shown in violet. The percentages represent the sequence identity of the designed enzymes compared to the starting point.

**CONTEXT**

**Enzyme stability** is of crucial importance for industrial applications. Stable enzymes have a longer shelf-life and functional lifespan, they are suitable for high-temperature application, can function in a wide pH range or in the presence of solvents or inhibitors. Improving the stability of enzymes can lead to cost savings by reducing the need for tightly controlled reaction conditions or frequent enzyme replacement. Moreover, engineering new enzyme functions (such as improved activity or specificity) can lead to the introduction of destabilizing mutations, resulting in poor expression or misfolding of the enzyme, which is why it is advisable to begin enzyme engineering campaigns with a stable protein.

**Beta-glucosidase** is an industrial enzyme used in the conversion of cellulosic biomass. In this high-temperature process, steam is used to make substrates accessible for enzymatic hydrolysis during saccharification, and therefore enzyme thermostability is essential for the application. We applied **LYNX** to beta-glucosidase, focusing on the high-temperature **stability** improvements. In addition to stability, we also focused on preserving the enzyme **activity** which is often lost during enzyme design campaigns.

The starting enzyme was a 448-amino-acids monomer of prokaryotic origin. AI models were used to generate 8 000 designs that were ranked and scored to select the final 15 computationally designed enzymes. The selected designs were expressed and tested experimentally to evaluate the enzyme activity and stability.

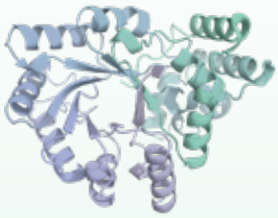
**RESULTS**

The stability of the enzymes was measured using the melting point assay. The melting point of the starting enzyme was measured at 45°C. Out of the LYNX generated candidates, the most thermostable enzyme variant exhibited a **melting temperature of 80°C (Figure 2)**. Among the 13 remaining enzymes, ten exhibited both activity and stability above that of the wild type. The most thermostable design had a sequence identity of 85% compared to the starting enzyme, translating to a total of 67 mutations introduced.

We also assessed enzyme activity alongside stability to confirm that our designs avoided any activity-stability trade-offs. The outcomes showed not only enhanced thermostability but also up to an eightfold increase in activity. Essentially, LYNX effectively improved beta-glucosidase's stability for high-temperature use without compromising, and in fact, it significantly boosted its activity.

## Improving the activity by 11-fold

### Ketoreductase (326 amino acids)



**LYNX84P**
Best performing enzyme variant in the LYNX campaign

- 11x improved activity
- 15°C increased melting temperature
- 97% identity to starting point

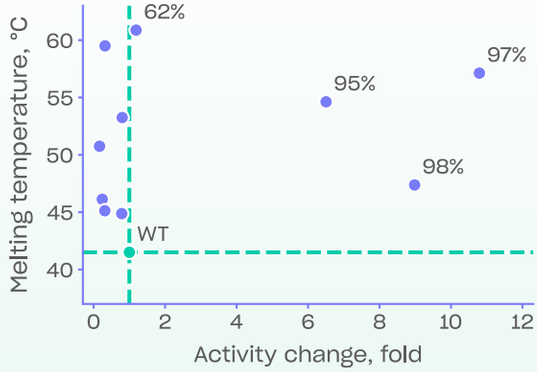



**Figure 3.** The activity and stability measurements for the designed ketoreductase enzymes. The starting enzyme is depicted in teal and designed enzyme variants are shown in violet. The percentages represent the sequence identity of the designed enzymes compared to the starting point.

**CONTEXT**
Enzymes are used as catalysts for chemical reactions across many industries. Engineering enzymes to have higher **activity** can significantly speed up these reactions, translating into improved process economics. This is particularly crucial in industries like pharmaceuticals and biofuels, where the rate of reaction can directly influence production timelines and costs. However, it's important to also consider the enzyme's stability while working on activity improvement to avoid negative stability-activity trade-offs. With LYNX we can optimize **multiple enzyme properties** simultaneously to avoid trade-offs and minimize development timelines.

To demonstrate the simultaneous improvement in **activity and stability** with LYNX, we selected a ketoreductase enzyme capable of reducing the Wieland-Miescher ketone (WMK) substrate. WMK serves as a building block for the synthesis of over 50 natural products with potential applications in the pharmaceutical industry. It has been identified as an intermediate in the retrosynthesis of promising active pharmaceutical ingredients (APIs) with anticancer, antimicrobial, antiviral, anti-neurodegenerative, and immunomodulatory activities. A crucial step in the synthesis of these products is the reduction of WMK to 5-hydroxy-4a-methyl-4,4a,5,6,7,8-hexahydronaphthalen-2(3H)-one, which can be enabled by the WMK ketoreductase.

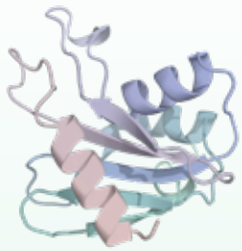
Previously, we used the Intelligent Architecture™ platform to identify enzymes capable of this reduction. This custom enzyme design package The enzyme discovery campaign has provided a diverse set of enzymes from distinct evolutionary sequence clusters capable of performing the desired reaction. We chose one of those sequences to further improve its efficiency for industrial applications. The starting ketoreductase was a monomeric NADH-dependent 326-amino-acid enzyme of a bacterial origin. We applied LYNX to create a set of novel ketoreductases with the goal of improving the enzymatic activity. Initially, LYNX was used to generate 12 000 candidate WMK ketoreductases. Following the scoring and selection process, **10 final enzyme variants** were selected for the experimental characterization.

**RESULTS**
All the designed enzymes were soluble and expressible. Moreover, three of them exhibited significantly improved activity, with the most active enzyme displaying an 11-fold increase in activity compared to the wild-type starting point (**Figure 3**).

Besides the retained enzyme expression level and improved activity, the stability of all the characterized enzyme designs was also retained or improved. This showcases that the LYNX-based enzyme optimization can be used to target the specific properties of interest while ensuring that the other properties are retained or improved at the same time. The application of LYNX not only circumvented the activity-stability trade-off but also significantly improved the stability of the enzyme, even in the most active designs.

# Redesigning enzyme sequences without loss of performance

## Cytidine Deaminase (136 amino acids)



**LYNX71R**
Best performing enzyme variant in the LYNX campaign

- 71% identity to starting point
- 3x improved activity

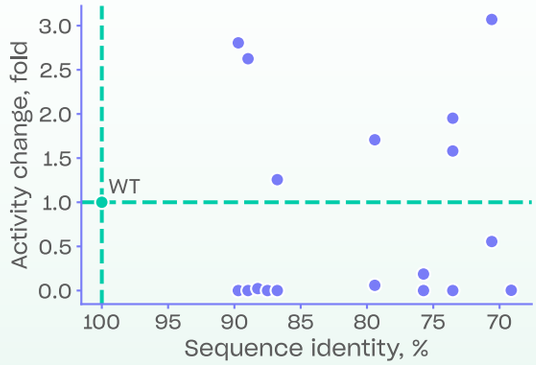



**Figure 4**. The activity and identity of the designed cytidine deaminases. The starting enzyme is depicted in teal and novel enzyme variants are shown in violet.

**CONTEXT**
Due to the highly competitive nature of the biotechnology industry, many relevant enzymes are protected by patents. **Redesigning enzymes** allows companies to develop and market their products without infringing on the IP rights of incumbent technologies. This not only avoids costly legal disputes but also fosters a more diverse and competitive market. However, simply changing the coding sequence of the enzyme is often insufficient to circumvent existing IP constraints. **LYNX** can make substantial changes to the enzyme sequence space, while retaining performance.

**LYNX enzyme redesign** can create novel enzyme variants that have up to 40% divergent sequence compared to the starting enzyme, which is generally sufficient to gain freedom to operate. Additionally, the redesign process can enhance IP protection of the enzyme by broadening the patented sequence space for more comprehensive coverage. The designed enzymes are unique and differ from their naturally occurring counterparts, which can help strengthen the novelty aspect of the patent claims.

An added benefit of enzyme redesign is that these synthetic enzymes can be used as alternative starting points for enzyme engineering campaigns. Sequence-diversified enzymes allow researchers to navigate around the local minima in the protein fitness landscape which are sometimes encountered in natural proteins. Redesigned enzymes are pre-optimized in terms of stability, which can cut down the number of engineering rounds required to obtain the desired performance characteristics.

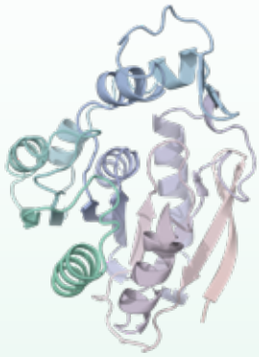
Following the popularity of CRISPR/Cas gene editing systems, the use of cytidine deaminase (CDA) in conjunction with Cas9 for base editing applications has been on the rise in recent years. However, the broad adoption and use of CDA systems has faced obstacles due to existing IP that restricts the use of cytidine deaminases with identity exceeding 90% relative to the patented enzyme sequence. Therefore, our objective was to apply LYNX to redesign the starting CDA enzyme sequence to create a set of synthetic variants with sequences identity below 90% to while ensuring that the enzyme is fully functional.

**RESULTS**
The starting CDA was a homotetrameric enzyme, composed of four 136-amino-acid monomers. LYNX was used to generate 11 000 new enzyme variants, which were scored to select 18 enzyme designs with sequence identity below 90%. Seven of these enzymes had improved activity compared to that of the starting enzymes, with up to three-fold improvement.

**The percent identity of the seven best designs ranged from 90 to 71%,** which translates to 14 to 40 amino acid mutations per monomer (**Figure 4**). Considering the monomer's small size, this enzyme has a high density of functional amino acids. Consequently, an exceptionally high proportion of potential mutations could critically impact the enzyme's proper functioning. We were able to mitigate this concern thanks to the model's ability to efficiently navigate the functional sequence space and avoid adverse changes. In the case of CDA, this translated to changing 30% of the enzyme sequence without negatively affecting other properties.

## Improving the yield by 3.5-fold

### Esterase (261 amino acids)



**LYNX35Y**
Best performing enzyme variant in the LYNX campaign

- 3.5x improved purification yield
- 5x improved activity
- 8°C increased melting temperature
- 72% identity to starting point.

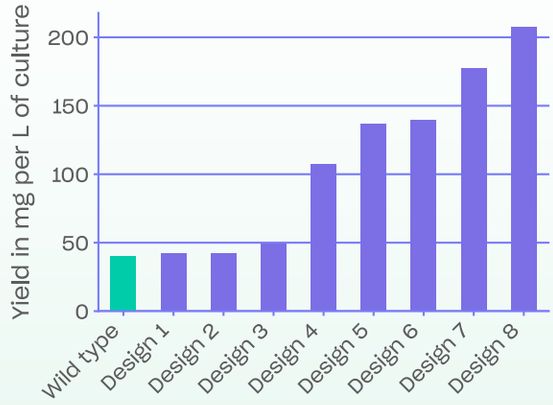



**Figure 5.** The production yields of the starting enzyme (teal) and the four improved enzymes (violet) that also demonstrated retained or improved activity and stability.

**CONTEXT**
The ability to express enzymes in host organisms at high yields is fundamental for commercial manufacturing of enzymes as products. Engineering enzymes for better expression ensures that they can be produced at scale, efficiently and cost-effectively. This is particularly important for industries where enzymes are used in large quantities, such as in detergents. Improving the yields of enzymes is also important for R&D initiatives that can be hindered by the amount of enzyme available.

The following **LYNX** case focuses on a chemical industry relevant esterase with a unique substrate scope. A key limitation of this enzyme's industrial applicability was its low yield in E. coli production strain, which persisted even after optimization of the production strain, culture, and process. We approached this challenge by focusing on optimizing the esterase amino acid sequence to improve the enzyme yield in E. coli. All improvements in the enzyme yield were assessed in the same production strain, vector and conditions, without further modification.

The starting esterase was a 261-amino-acid monomeric enzyme of prokaryotic origin that was poorly expressed in E. coli. LYNX optimization scheme was applied to generate 8 000 novel sequences and select 20 final enzyme designs for experimental characterization. Other properties like stability and activity were also evaluated to ensure that the produced enzyme designs remained active and stable.

**RESULTS**
Out of the 20 tested enzymes, five had at least a **two-fold improved production yield**. Out of the five enzymes with significantly improved production yield, four enzyme variants had either retained the activity and stability or improved it. The characterized enzyme variants also retained their desired substrate promiscuity. We identified that the yield improvement for these enzyme designs came from increased solubility levels. The best expressed enzyme reached a yield of 178 mg/L compared to the yield of 46 mg/L for the starting enzyme (**Figure 5**).

LYNX application to improve the production of an industrially important esterase enzyme generated a nearly four-fold improvement in enzyme yield. The enzyme production yield was enhanced without adversely affecting other enzyme properties. In fact, the optimization improved enzyme solubility and potentially resolved other fundamental issues, such as enzyme misfolding, short half-life, and aggregation. By prioritizing enzyme design over optimization of production conditions, we achieved substantial enhancement in enzyme yield.

# Try LYNX Enzyme Optimization™ for yourself today!

Contact us for a free feasibility evaluation

partnering@biomatter.ai

biomatter