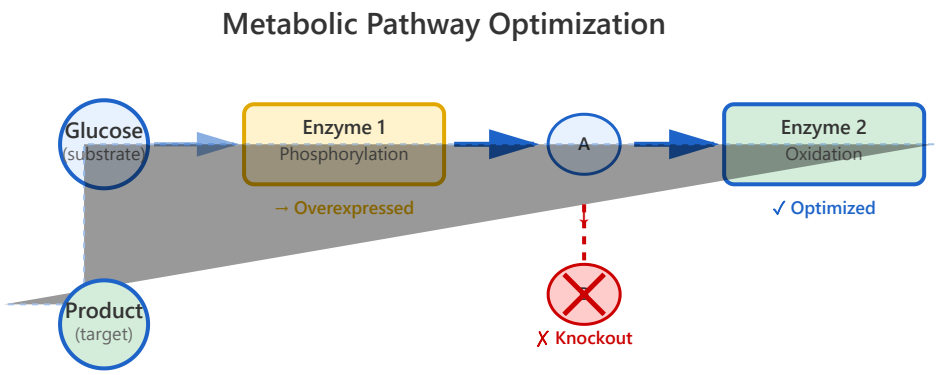
Metabolic Engineering



AI-Guided Metabolic Engineering

Flux Balance Analysis (FBA)
- Identify bottlenecks
- Predict knockouts
- Optimize expression levels

Machine Learning Models
- Enzyme activity prediction
- Strain design
- Growth prediction

Flux optimization
Balance metabolic flow

Enzyme engineering
Improve catalytic efficiency

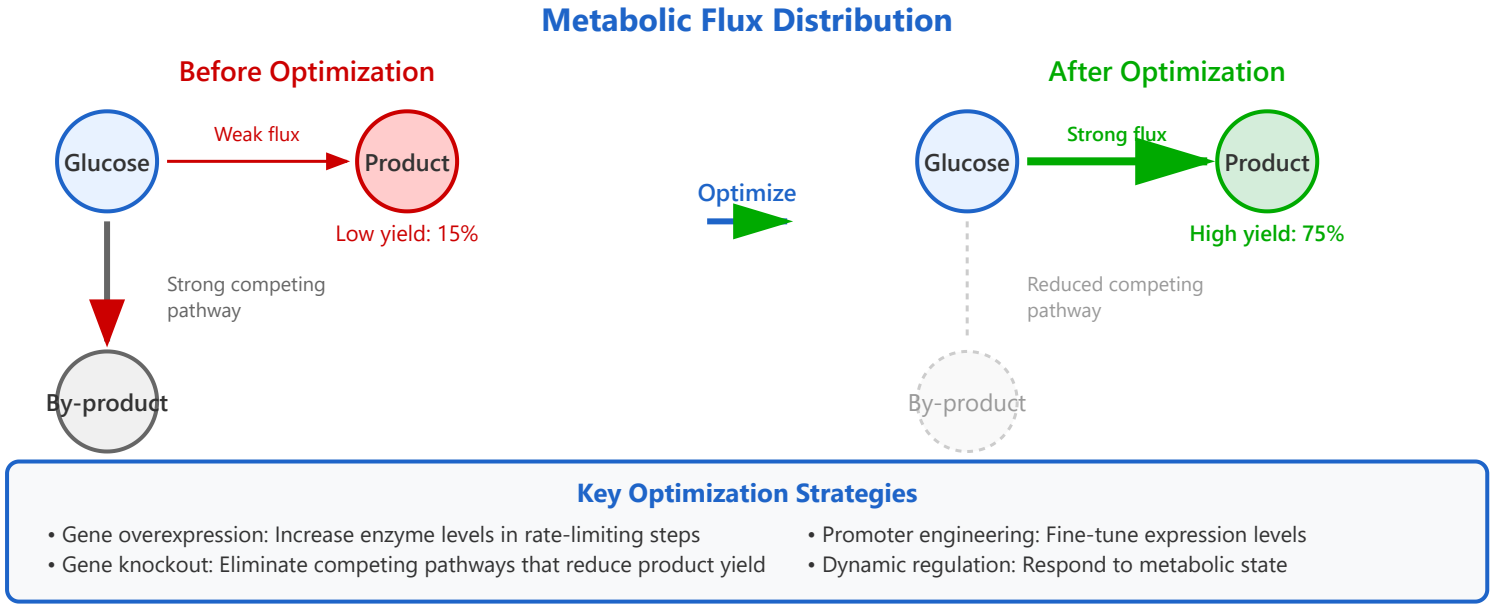
Pathway design
Novel biosynthetic routes

Strain optimization
Host organism engineering

Scale-up prediction
Lab → production modeling

1. Flux Optimization

Flux optimization involves redistributing metabolic flux through cellular pathways to maximize the production of desired compounds. This strategy focuses on balancing the flow of metabolites to prevent bottlenecks and accumulation of toxic intermediates while maximizing product yield.



Real-World Example: 1,3-Propanediol Production

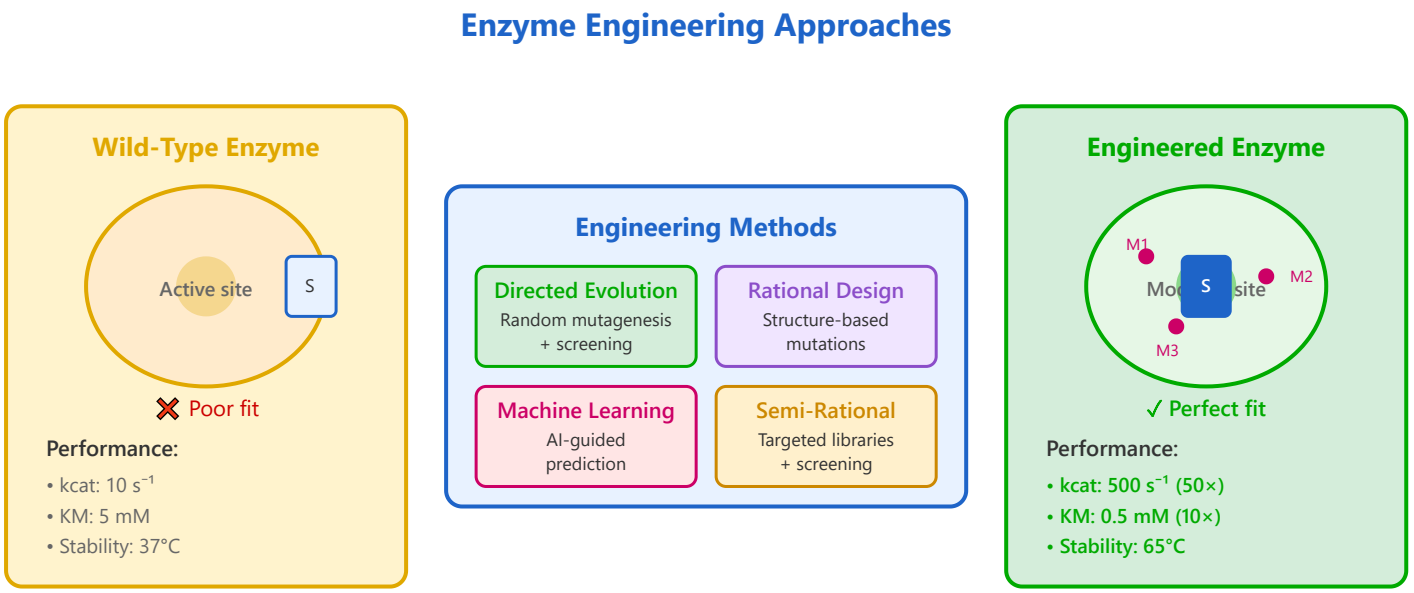
DuPont engineered E. coli to produce 1,3-propanediol (a polymer precursor) from glucose. By redirecting flux from glycerol to 1,3-propanediol through overexpression of dhaB and dhaT genes while knocking out competing pathways, they achieved industrial-scale production with yields exceeding 130 g/L, making it commercially viable for producing Sorona® polymer.

- Computational tools like Flux Balance Analysis (FBA) predict optimal gene modifications
- 13C metabolic flux analysis experimentally validates flux distributions

- Balancing growth rate and production rate is critical for industrial applications

2. Enzyme Engineering

Enzyme engineering focuses on improving the catalytic properties of enzymes through directed evolution, rational design, or computational methods. This approach enhances reaction rates, substrate specificity, stability, and overall pathway efficiency.



Real-World Example: Artemisinic Acid Production

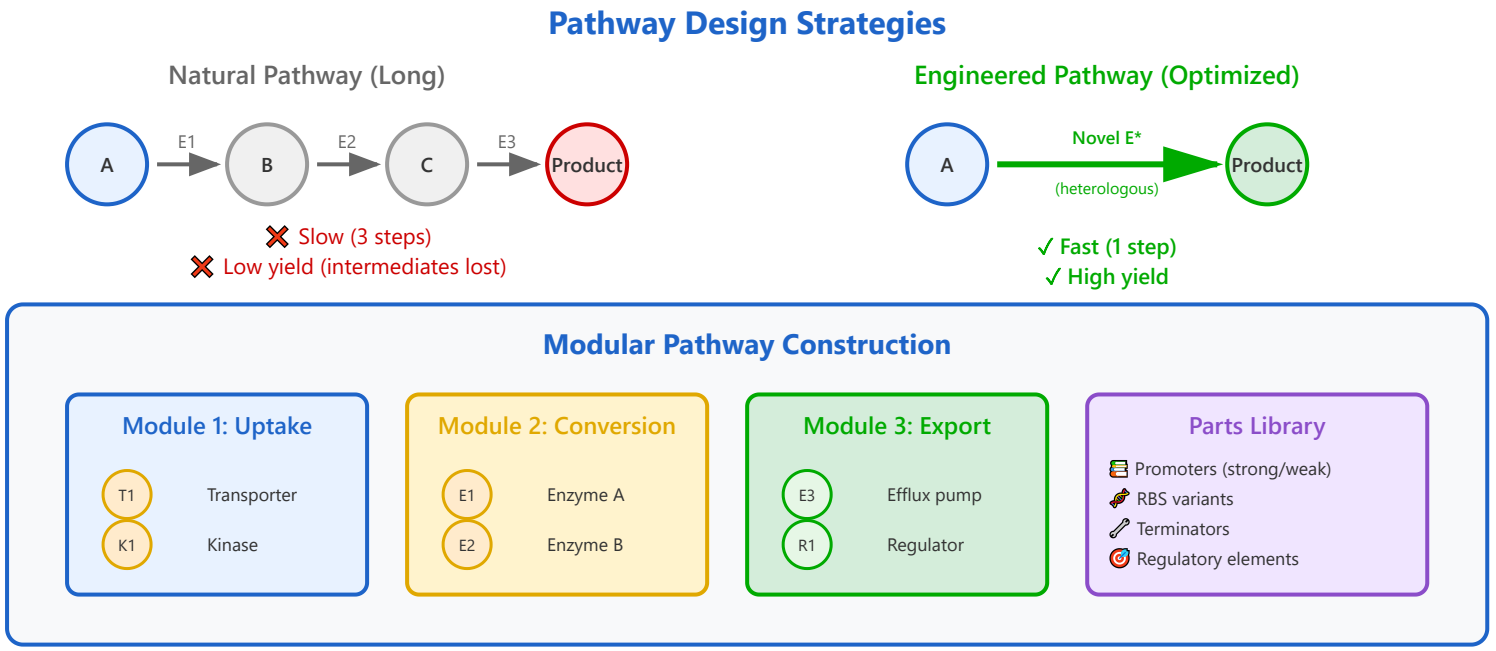
Researchers at UC Berkeley and Amyris engineered cytochrome P450 enzymes in yeast to produce artemisinic acid, a precursor to the anti-malaria drug artemisinin. Through directed evolution and rational design, they improved enzyme activity by over 200-fold, enabling economical semi-synthetic production of this life-saving medication that was previously only available from plant extraction.

- AlphaFold and RoseTTAFold enable structure-guided enzyme design
- Deep learning models predict beneficial mutations with >80% accuracy

- Combination approaches (rational + directed evolution) often yield best results
- Enzyme thermostability is crucial for industrial bioprocesses

3. Pathway Design

Pathway design involves constructing novel biosynthetic routes by introducing heterologous genes or creating entirely synthetic pathways. This strategy enables production of compounds not naturally made by the host organism or improves efficiency through shortened pathways.



Real-World Example: Taxol (Paclitaxel) Precursor Production

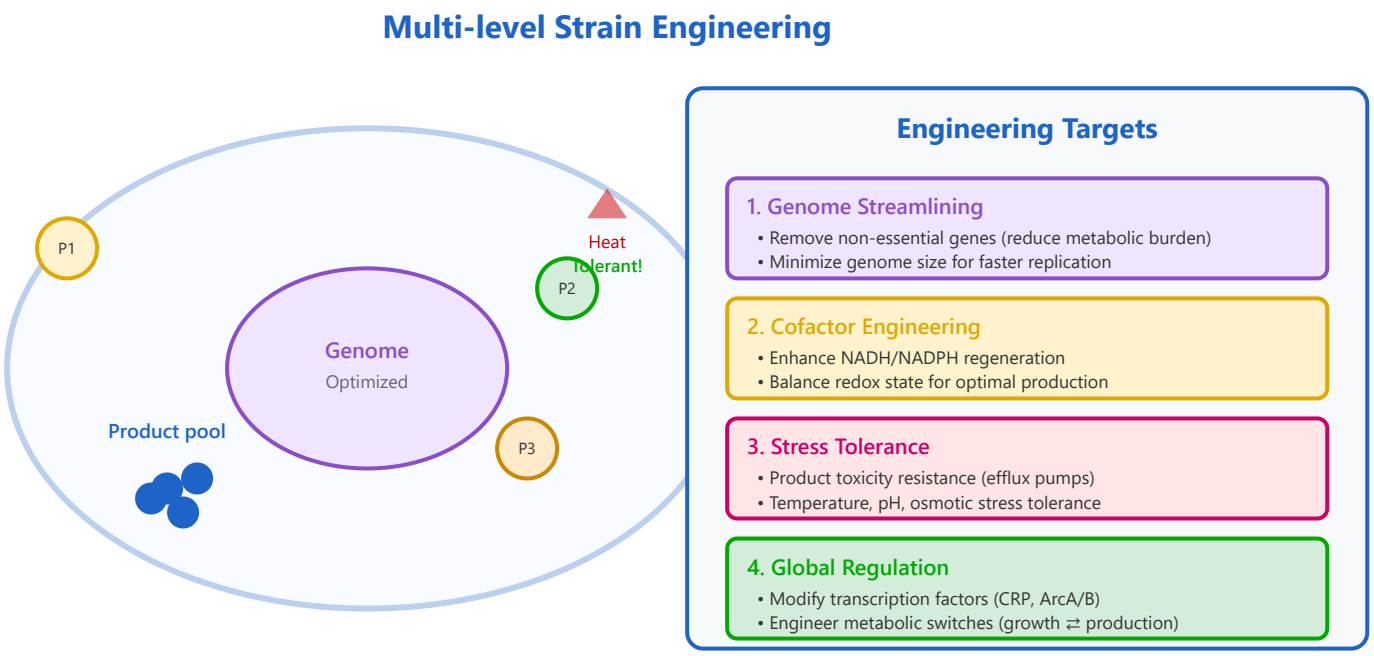
Researchers assembled a 13-gene heterologous pathway in yeast to produce taxadiene, a key precursor to the anti-cancer drug Taxol. This pathway combined genes from Pacific yew trees with engineered yeast enzymes, creating a biosynthetic route that would be impossible in any natural organism. The modular approach allowed rapid optimization of each pathway segment independently.

- Retrosynthetic analysis identifies optimal enzymatic routes to target molecules

- Codon optimization ensures proper expression of heterologous genes
- Enzyme scaffolding and spatial organization reduce intermediate diffusion
- Standardized genetic parts (BioBricks, MoClo) accelerate pathway construction

4. Strain Optimization

Strain optimization encompasses the systematic improvement of host organisms to enhance production capabilities. This includes modifying cellular machinery, improving tolerance to stress, and engineering global regulatory networks for optimal biosynthesis.



Real-World Example: Tolerance-Enhanced E. coli for Biofuel Production

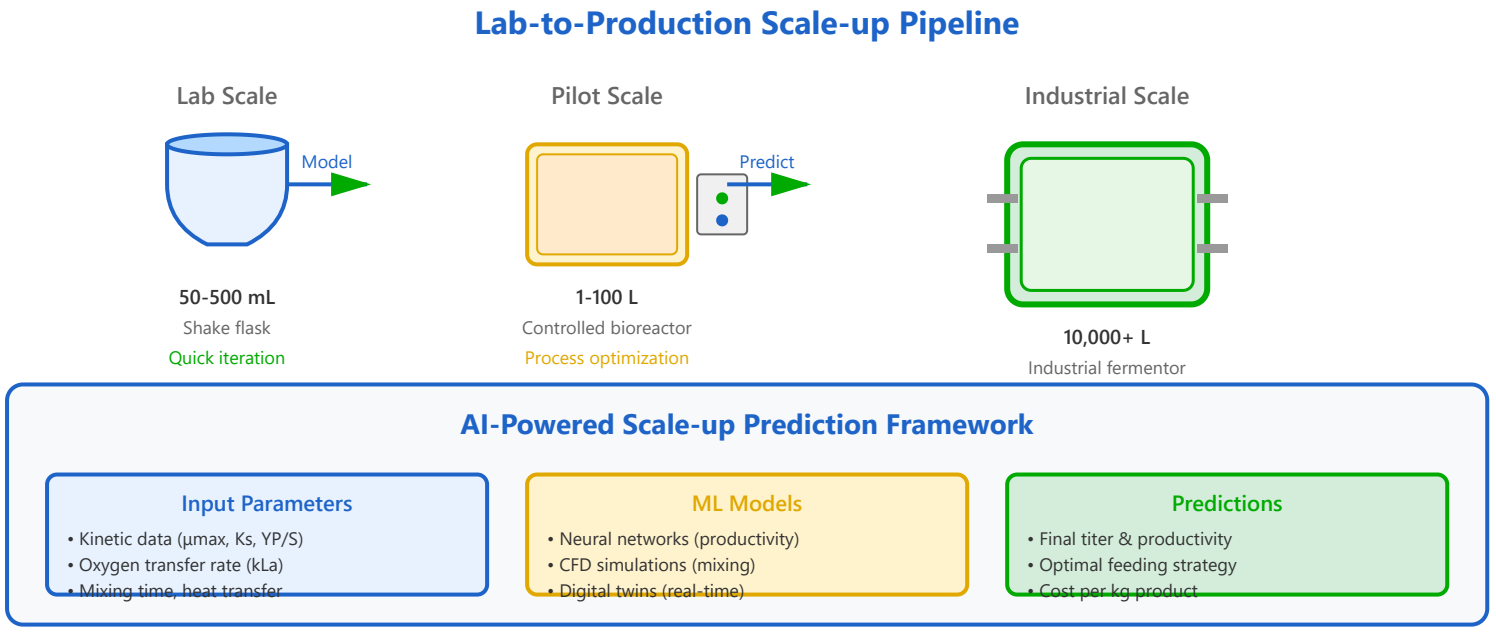
Scientists at MIT engineered E. coli for advanced biofuel production by incorporating multiple modifications: deletion of 15 non-essential gene clusters to reduce metabolic burden, overexpression of stress response genes (groESL, dnaKJ), and modification of membrane composition to tolerate high concentrations of toxic alcohols and fatty acids. The optimized strain showed 3-fold improvement in titer and could operate continuously in fed-batch fermentation for over 200 hours.

- Adaptive laboratory evolution (ALE) complements rational engineering for complex phenotypes

- CRISPR interference (CRISPRi) enables fine-tuning of gene expression without knockouts
- Systems biology approaches identify hidden bottlenecks in cellular metabolism
- Chassis organisms (E. coli, S. cerevisiae, B. subtilis) each offer unique advantages

5. Scale-up Prediction

Scale-up prediction uses computational models and machine learning to forecast how laboratory-scale bioprocesses will perform at industrial scale. This strategy reduces costly trial-and-error in pilot plants and enables rapid translation from bench to commercial production.



Real-World Example: Ginkgo Bioworks Scale-up Platform

Ginkgo Bioworks developed a machine learning platform that predicts fermentation performance at 10,000L+ scale from 96-well plate data. By training models on thousands of fermentation runs, their system predicts titer within 15% accuracy and recommends optimal media composition and feeding strategies. This reduced their time from strain design to commercial production from 18 months to under 6 months for multiple products including fragrance molecules and pharmaceutical intermediates.

- Computational Fluid Dynamics (CFD) models predict mixing and mass transfer at scale

- Machine learning on historical data outperforms mechanistic models for complex systems
- Digital twins enable real-time process optimization during production
- Scale-down simulators validate predictions before committing to large-scale runs
- Techno-economic analysis (TEA) integrated with models guides cost-effective scale-up