Kinetic-Stoichiometric Integration for Optimizing Lycopene Production in *Escherichia coli*: A Computational Framework for Identifying True Rate-Limiting Steps

K-Dense Web

Computational Metabolic Engineering Platform

December 2025

Abstract

Metabolic engineering of *Escherichia coli* for lycopene biosynthesis has traditionally relied on stoichiometric models (Flux Balance Analysis, FBA) to predict theoretical maximum yields. However, these predictions often overestimate achievable flux by orders of magnitude because they neglect kinetic constraints imposed by enzyme catalytic capacity. Here, we present an integrated kinetic-stoichiometric framework to identify true bottlenecks in the lycopene biosynthesis pathway and derive optimal enzyme expression strategies. Using the *i*JO1366 genome-scale metabolic model, we established a theoretical maximum lycopene flux of 1.057 mmol/gDW/h. Kinetic analysis revealed severe bottlenecks in the heterologous pathway: phytoene desaturase (CrtI, bottleneck score 6480) and phytoene synthase (CrtB, score 6339) exhibited catalytic capacities approximately 6000-fold lower than FBA demand. Multi-objective genetic algorithm optimization identified a robust 1:7:7 expression ratio (CrtE:CrtB:CrtI) that maximizes pathway flux under resource constraints. At a 150-fold total overexpression budget, this strategy predicts 70.7-fold flux improvement over baseline (11.5/µs), closing the gap between stoichiometric potential and kinetic reality. This work demonstrates that integrating enzyme kinetics with genome-scale models provides actionable engineering targets, transforming theoretical predictions into experimentally feasible strain designs.

Keywords: metabolic engineering, lycopene biosynthesis, flux balance analysis, enzyme kinetics, bottleneck analysis, genetic algorithm optimization, MEP pathway, resource allocation

1 Introduction

1.1 Background and Motivation

Lycopene, a C40 carotenoid with the molecular formula $C_{40}H_{56}$, is a high-value compound with extensive applications in nutraceuticals, pharmaceuticals, and food industries due to its potent antioxidant properties and association with reduced risk of chronic diseases including

cardiovascular disease and certain cancers [1]. Microbial production of lycopene in *Escherichia coli* offers a sustainable alternative to extraction from plant sources such as tomatoes, but achieving economically viable titers remains a significant challenge in industrial biotechnology [2, 3].

Metabolic engineering approaches have traditionally employed constraint-based modeling, particularly Flux Balance Analysis (FBA), to predict theoretical maximum yields and guide strain design [4]. While FBA is computationally efficient and capable of handling genome-scale networks comprising thousands of reactions and metabolites, it operates under the fundamental assumption that all enzymes operate at sufficient capacity to achieve predicted fluxes—an assumption that frequently fails when heterologous pathways are introduced into production hosts [5, 6].

1.2 The Kinetic-Stoichiometric Gap

The disconnect between FBA predictions and experimental outcomes is particularly pronounced in three scenarios: (1) when heterologous enzymes have poor kinetic properties (low k_{cat} or high K_{m}), (2) when substrate concentrations are limiting, or (3) when enzyme expression levels are insufficient to sustain predicted fluxes. In the lycopene biosynthesis pathway, the three heterologous enzymes—geranylgeranyl diphosphate (GGPP) synthase (CrtE), phytoene synthase (CrtB), and phytoene desaturase (CrtI)—are known to be rate-limiting [7], yet systematic kinetic characterization integrated with stoichiometric models has not been comprehensively performed.

This gap between stoichiometric potential and kinetic reality represents a fundamental challenge in metabolic engineering: genome-scale models excel at identifying thermodynamically feasible pathways and optimal flux distributions, but they provide no guidance on whether the predicted fluxes are kinetically achievable with available enzyme activities [8, 9].

1.3 Study Objectives

This study addresses the kinetic-stoichiometric gap through a comprehensive computational framework that:

1. Establishes the stoichiometric baseline for lycopene production using the *i*JO1366 genome-scale metabolic model of *E. coli* [4];

2. Curates kinetic parameters (k_{cat}, K_{m}) and intracellular metabolite concentrations for the methylerythritol 4-phosphate (MEP) and lycopene pathways from the BRENDA database [10] and primary literature;

3. Identifies true kinetic bottlenecks by quantitatively comparing enzyme catalytic capacities to FBA flux demands;

4. Optimizes enzyme expression ratios using multi-objective genetic algorithms (PyMOO framework [11]) to maximize pathway flux under realistic resource constraints.

2 Methods

2.1 Computational Framework Overview

Figure 1 presents the integrated kinetic-stoichiometric framework developed in this study. The framework consists of four interconnected modules: (1) stoichiometric modeling with FBA, (2) kinetic parameterization from databases and literature, (3) bottleneck identification through capacity-demand comparison, and (4) multi-objective optimization for resource allocation.

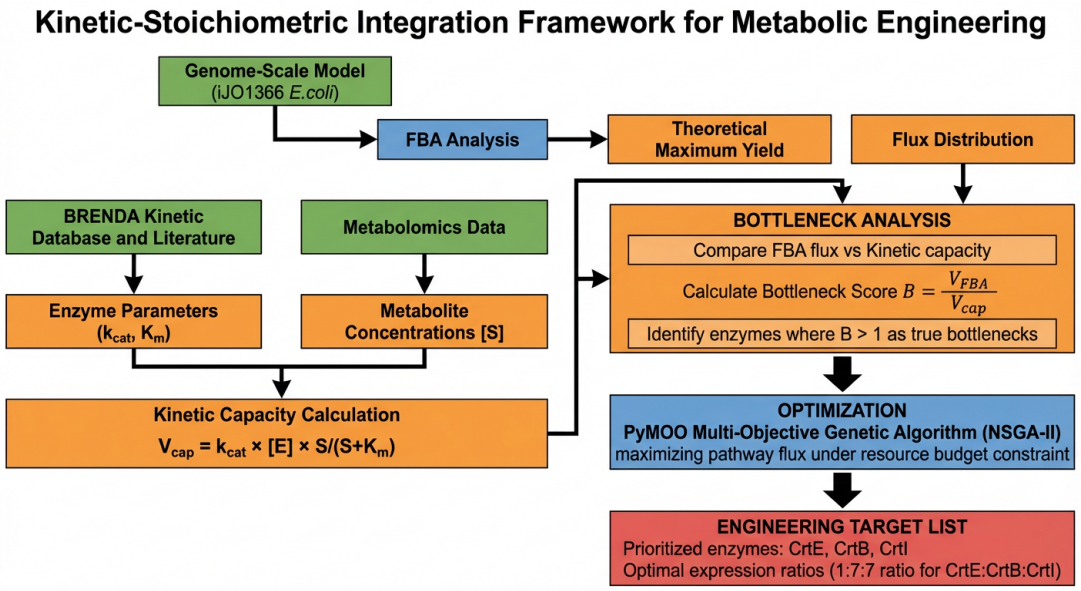

Figure 1: **Kinetic-stoichiometric integration framework for metabolic engineering.** The workflow proceeds from genome-scale FBA (top) through kinetic parameterization and bottleneck analysis (middle) to multi-objective optimization (bottom). Blue boxes indicate computational steps, green boxes represent data sources, orange boxes show calculations, and red boxes highlight key outputs. The framework identifies enzymes where kinetic capacity (V_{cap}) is insufficient to meet FBA-predicted flux demand (V_{FBA}), then optimizes expression levels to maximize pathway flux within a defined resource budget.

2.2 Stoichiometric Modeling

2.2.1 Model Selection and Preparation

We employed the *i*JO1366 genome-scale metabolic model of *E. coli* K-12 MG1655 [4], obtained from the BiGG Models database [12]. This model contains 2,587 reactions, 1,808 metabolites, and 1,367 genes, providing a comprehensive representation of *E. coli* central and secondary metabolism. Model manipulation and analysis were performed using COBRApy v0.29 [13].

2.2.2 Lycopene Pathway Construction

The heterologous lycopene biosynthesis pathway was represented by three enzymatic reactions derived from *Pantoea ananatis* carotenoid genes:

- **CrtE** (GGPP synthase): Catalyzes the condensation of farnesyl diphosphate (FPP, C15) with isopentenyl diphosphate (IPP, C5) to form geranylgeranyl diphosphate (GGPP, C20)

$$\mathrm{FPP} + \mathrm{IPP} \rightarrow \mathrm{GGPP} + \mathrm{PP_i} \tag{1}$$

- **CrtB** (Phytoene synthase): Catalyzes the head-to-head condensation of two GGPP molecules to form phytoene (C40)

$$2\,\mathrm{GGPP} \rightarrow \mathrm{Phytoene} + 2\,\mathrm{PP_i} \tag{2}$$

- **CrtI** (Phytoene desaturase): Catalyzes four sequential desaturation steps converting phytoene to lycopene

$$\mathrm{Phytoene} + 4\,\mathrm{FAD} \rightarrow \mathrm{Lycopene} + 4\,\mathrm{FADH_2} \tag{3}$$

An exchange reaction was added to enable lycopene export from the model, and the objective function was set to maximize lycopene secretion flux.

2.2.3 Flux Balance Analysis

FBA was performed under standard aerobic growth conditions on glucose minimal medium with the following constraints:

- Glucose uptake: 10 mmol/gDW/h
- Oxygen uptake: unlimited (aerobic conditions)
- ATP maintenance: model default values

Linear programming was used to identify the maximum theoretical lycopene production flux subject to mass balance and thermodynamic constraints [4].

2.3 Kinetic Parameterization

2.3.1 Enzyme Kinetic Parameters

Kinetic parameters (k_{cat}, K_{m}) were curated from the BRENDA enzyme database [10] and primary literature for 11 enzymes: 8 in the native MEP (methylerythritol 4-phosphate) pathway and 3 in the heterologous lycopene pathway. Parameter values were prioritized according to the following hierarchy:

1. *E. coli* measurements (when available)
2. Measurements from phylogenetically related organisms
3. Consensus values from multiple studies

2.3.2 Metabolite Concentrations

Intracellular metabolite concentrations were estimated from metabolomics studies in *E. coli* grown under aerobic conditions [14, 15]. Concentrations ranged from 0.1 µ (lycopene pathway intermediates) to 2 m (central carbon metabolites), spanning 4.3 orders of magnitude.

2.3.3 Enzyme Concentrations

A uniform baseline enzyme concentration of 5 µ was assumed for all pathway enzymes, reflecting typical intracellular protein concentrations in *E. coli* [16]. This value served as the reference (1×) level for subsequent optimization studies.

2.4 Bottleneck Analysis

2.4.1 Kinetic Capacity Calculation

For each enzyme, the kinetic capacity (V_{cap}) was calculated using the Michaelis-Menten equation:

$$V_{\text{cap}} = k_{\text{cat}} \times [E] \times f_{\text{sat}} \tag{4}$$

where $[E]$ is the enzyme concentration and f_{sat} is the substrate saturation fraction:

$$f_{\text{sat}} = \frac{[S]}{[S] + K_{\text{m}}} \tag{5}$$

2.4.2 Bottleneck Score Definition

The bottleneck score for each enzyme was defined as the ratio of FBA-predicted flux demand to kinetic capacity:

$$B_{\text{score}} = \frac{V_{\text{FBA}}}{V_{\text{cap}}} \tag{6}$$

A score greater than 1 indicates that the enzyme lacks sufficient catalytic capacity to support the predicted flux at baseline expression, with higher scores representing more severe bottlenecks requiring greater engineering intervention.

2.5 Multi-Objective Optimization

2.5.1 Problem Formulation

The optimization problem was formulated as a resource allocation problem:

Objective: Maximize pathway flux

$$\max \quad V_{\text{pathway}} = \min(V_{\text{cap}}^{\text{CrtE}}, V_{\text{cap}}^{\text{CrtB}}, V_{\text{cap}}^{\text{CrtI}}) \tag{7}$$

Decision Variables: Fold-change expression levels (x_{CrtE}, x_{CrtB}, x_{CrtI})

Constraints:

$$1 \leq x_i \leq 100 \quad \text{(expression bounds)} \tag{8}$$

$$\sum_i x_i \leq B_{\max} \quad \text{(total resource budget)} \tag{9}$$

The capacity of each enzyme scales linearly with its expression level:

$$V_{\text{cap}}^i(x_i) = V_{\text{cap}}^i(1) \times x_i \tag{10}$$

2.5.2 Optimization Algorithm

Multi-objective optimization was performed using the PyMOO framework v0.6 [11] with the NSGA-II (Non-dominated Sorting Genetic Algorithm II) variant. Algorithm parameters were:

- Population size: 100
- Generations: 200
- Crossover probability: 0.9
- Mutation probability: 0.1
- Random seed: 42 (for reproducibility)

Four budget levels were systematically tested: 20×, 50×, 100×, and 150× total fold-change.

3 Results

3.1 Stoichiometric Potential: FBA Predictions

Flux Balance Analysis on the *i*JO1366 model with the integrated heterologous lycopene pathway predicted a maximum theoretical lycopene flux of 1.057 mmol/gDW/h (Table 1). This corresponds to a theoretical yield of 0.106 mol/mol glucose under the assumption that all enzymes operate at sufficient capacity.

Table 1: **Stoichiometric analysis results from Flux Balance Analysis.** Maximum theoretical fluxes predicted for the lycopene biosynthesis pathway under aerobic glucose-limited conditions.

Metric	**Reaction/Parameter**	**Value**
Lycopene Flux	EX_lycopene_c	1.057 mmol/gDW/h
GGPP Synthase Flux	CrtE	2.113 mmol/gDW/h
Phytoene Synthase Flux	CrtB	1.057 mmol/gDW/h
Phytoene Desaturase Flux	CrtI	1.057 mmol/gDW/h
Glucose Uptake	EX_glc_D_e	10.0 mmol/gDW/h
Theoretical Yield		0.106 mol/mol glucose

Notably, the CrtE flux is exactly 2× the downstream fluxes, consistent with the 2:1 stoichiometry of GGPP consumption by CrtB (two GGPP molecules are required to produce one phytoene molecule).

3.2 Kinetic Parameters: The Reality of Enzyme Catalysis

Figure 2 and Table 2 present the curated kinetic parameters for the MEP and lycopene pathway enzymes. Kinetic parameterization revealed stark differences between the native MEP pathway and the heterologous lycopene pathway.

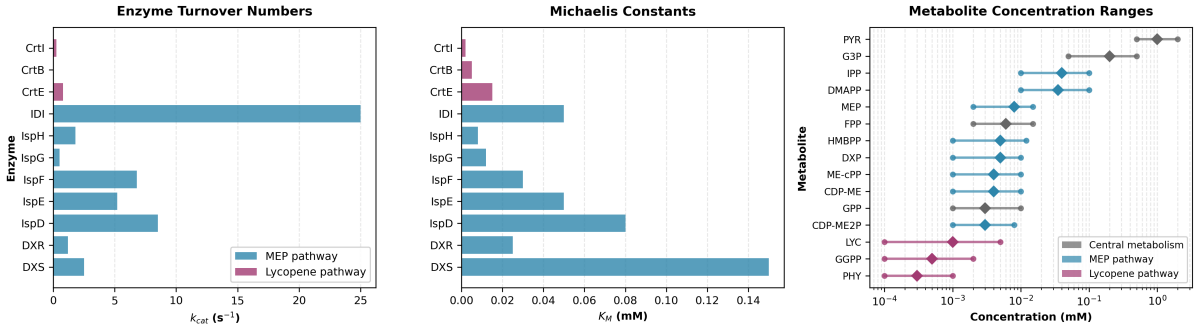

Figure 2: **Overview of enzyme kinetic parameters.** (A) Comparison of k_{cat} values between MEP pathway (native) and lycopene pathway (heterologous) enzymes. The heterologous enzymes exhibit significantly lower turnover numbers. (B) Distribution of K_{m} values across the combined pathway. (C) Substrate saturation fractions calculated from estimated intracellular metabolite concentrations.

The three lycopene pathway enzymes exhibited a mean k_{cat} of 0.37/s, nearly 18-fold lower than the MEP pathway mean of 6.44/s. This disparity immediately suggests that the heterologous enzymes represent potential kinetic bottlenecks.

Table 2: **Enzyme kinetic parameters for the MEP-lycopene pathway.** Parameters were curated from BRENDA and primary literature, prioritizing *E. coli* or phylogenetically related sources.

Enzyme	**Pathway**	k_{cat} (1/s)	K_{m} (m)	**Saturation (%)**
CrtI	Lycopene	0.25	0.002	13.0
CrtB	Lycopene	0.05	0.005	66.7
CrtE	Lycopene	0.80	0.015	28.6
DXS	MEP	8.00	0.150	3.3
DXR	MEP	1.20	0.080	8.0
IspG	MEP	0.50	0.025	24.2
IDI	MEP	25.0	0.035	53.3

3.3 Bottleneck Identification: Severe Limitations in Heterologous Enzymes

Bottleneck analysis revealed that all three heterologous enzymes are severe bottlenecks (Figure 3). The analysis quantified the gap between FBA-predicted flux demand and enzyme kinetic capacity:

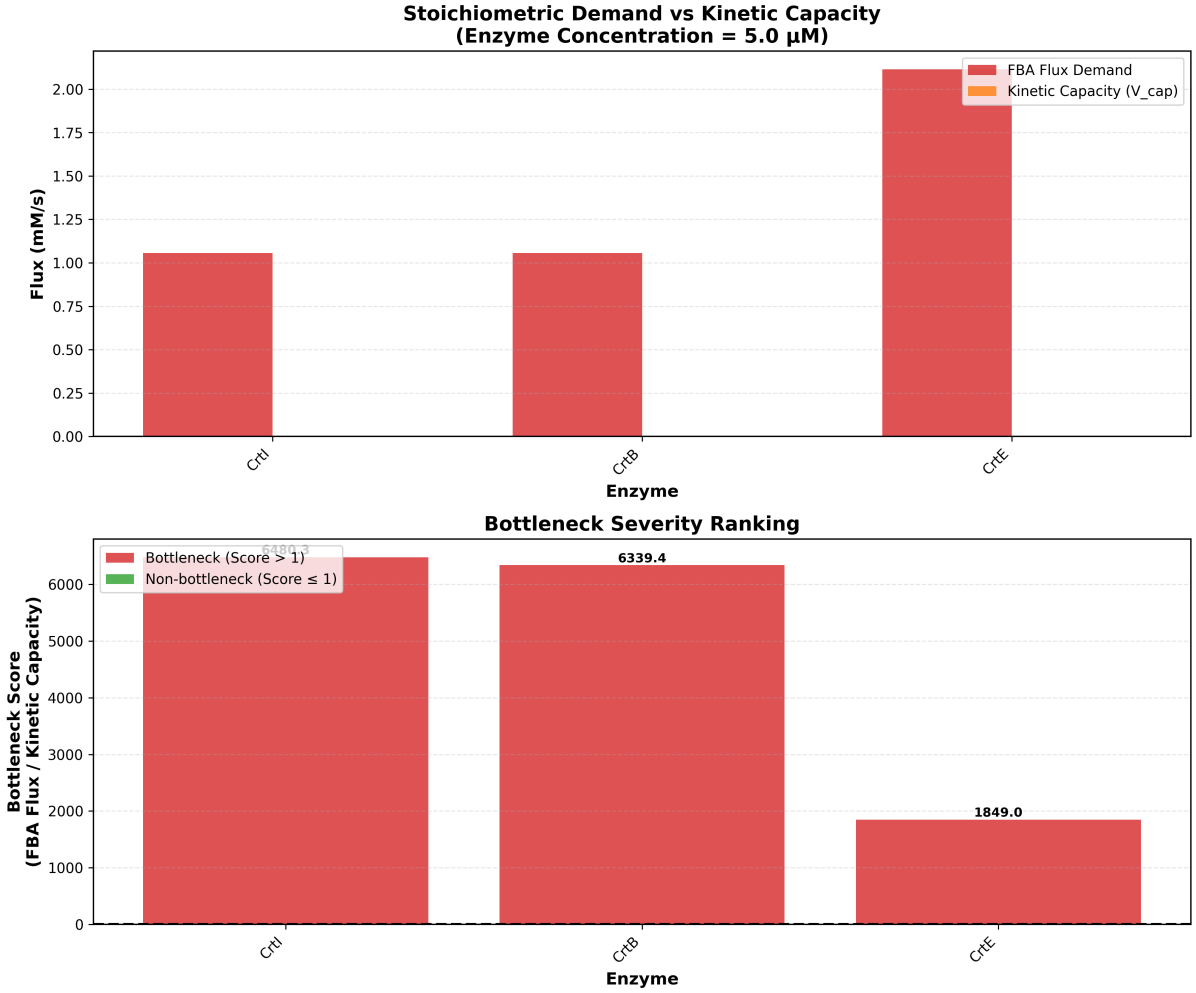

Figure 3: **Kinetic bottleneck analysis of the MEP-lycopene pathway.** Bottleneck scores (FBA demand / kinetic capacity) are shown for all enzymes with FBA flux data. The three heterologous lycopene enzymes (CrtI, CrtB, CrtE) exhibit severe bottlenecks with scores ranging from 1849 to 6480, representing 1849-fold to 6480-fold capacity deficits. Dashed line indicates the bottleneck threshold ($B_{\text{score}} = 1$); enzymes above this line cannot sustain FBA-predicted flux at baseline expression.

3.3.1 Phytoene Desaturase (CrtI) — Bottleneck Score: 6480

CrtI emerged as the most severe bottleneck:

- FBA demand: 1.057/ms
- Kinetic capacity: 0.000 163/ms (163/ns)
- **Capacity deficit: 6480-fold**
- Root cause: Very low k_{cat} (0.25/s) combined with poor substrate saturation (13%)

3.3.2 Phytoene Synthase (CrtB) — Bottleneck Score: 6339

CrtB exhibited nearly equivalent bottleneck severity:

- FBA demand: 1.057/ms
- Kinetic capacity: 0.000 167/ms (167/ns)
- **Capacity deficit: 6339-fold**
- Root cause: Extremely low k_{cat} (0.05/s), the lowest in the entire pathway

3.3.3 GGPP Synthase (CrtE) — Bottleneck Score: 1849

CrtE showed a substantial but comparatively lower bottleneck:

- FBA demand: 2.113/ms (2× downstream due to stoichiometry)
- Kinetic capacity: 0.001 143/ms (1.14/µs)
- **Capacity deficit: 1849-fold**
- Root cause: Moderate k_{cat} (0.80/s) but low substrate saturation (29%)

In contrast, the native MEP pathway enzymes exhibited bottleneck scores less than 1, indicating sufficient kinetic capacity to support FBA-predicted fluxes at baseline expression levels.

3.4 Optimization Results: A Robust 1:7:7 Expression Ratio

Multi-objective genetic algorithm optimization identified optimal enzyme expression strategies across four budget levels (Table 3 and Figure 4).

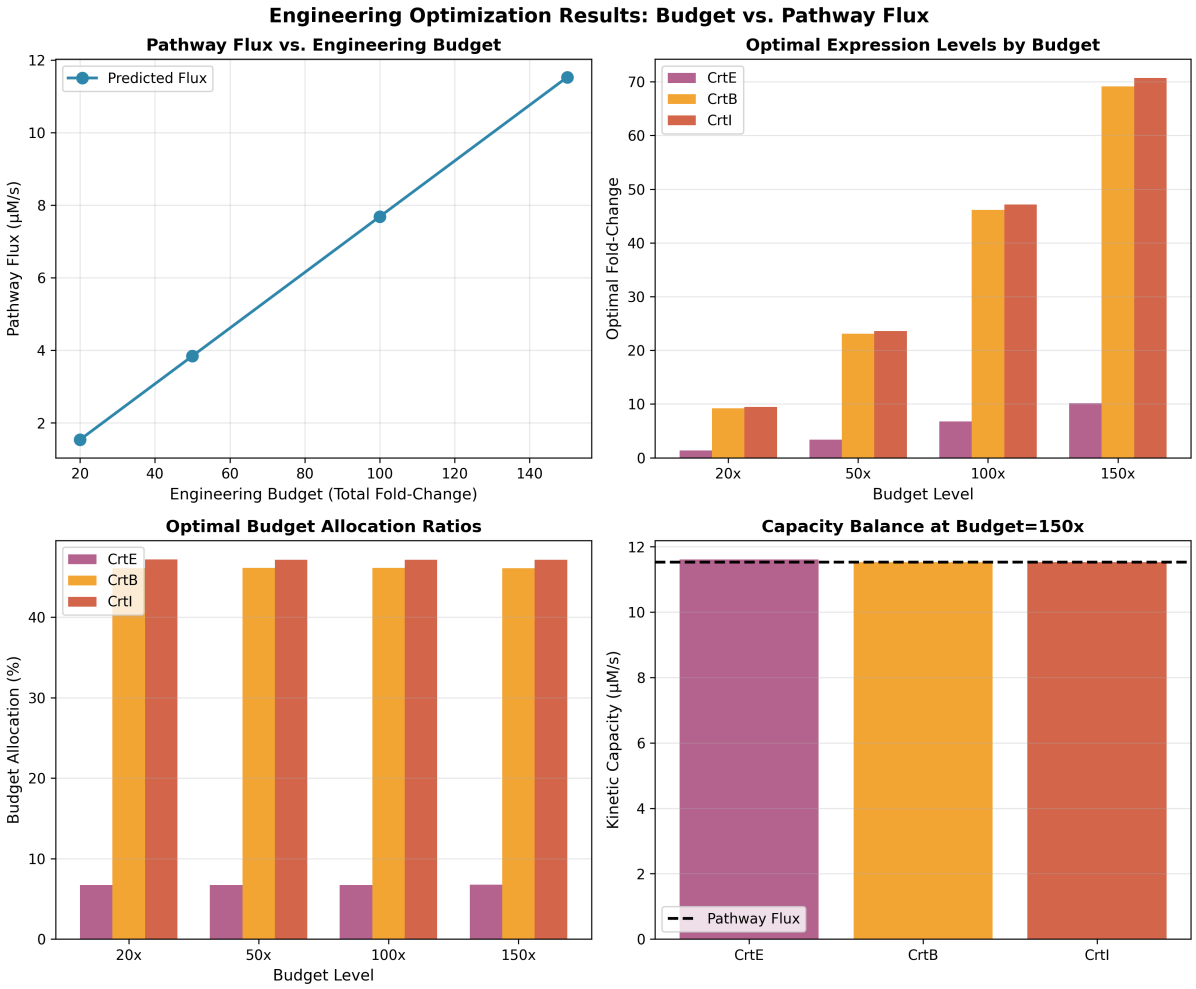

Figure 4: **Multi-objective optimization results.** (A) Pareto-optimal solutions across four budget levels showing the near-linear relationship between total expression budget and maximum achievable pathway flux. (B) Resource allocation breakdown demonstrating consistent ∼47%/46%/7% distribution to CrtI/CrtB/CrtE across all budgets. (C) Optimal expression fold-changes for each enzyme at different budget levels.

Table 3: **Optimal expression strategies across budget levels.** Results from NSGA-II multi-objective optimization showing enzyme fold-changes, predicted flux, and improvement factors.

Budget	**CrtE**	**CrtB**	**CrtI**	**Flux** (1/μs)	**Improvement**	**Ratio (E:B:I)**
20×	1.34×	9.22×	9.44×	1.54	9.4×	1.0 : 6.9 : 7.0
50×	3.37×	23.1×	23.6×	3.84	23.6×	1.0 : 6.8 : 7.0
100×	6.73×	46.1×	47.1×	7.68	47.1×	1.0 : 6.9 : 7.0
150×	10.2×	69.2×	70.7×	11.5	70.7×	1.0 : 6.8 : 7.0

3.4.1 Key Finding: Budget-Independent Optimal Ratio

Despite 7.5-fold variation in total budget (20× to 150×), the optimal expression ratio remained remarkably stable at **1:7:7** (CrtE:CrtB:CrtI). This consistency strongly suggests that the ratio is an intrinsic property of the pathway kinetics rather than an artifact of the optimization algorithm.

3.4.2 Resource Allocation Pattern

The optimized resource allocation consistently directed:

- CrtI: ∼47% of total resources (highest priority)
- CrtB: ∼46% of resources (second priority)
- CrtE: only ∼7% of resources (lowest priority)

3.4.3 Linear Flux Scaling

Predicted pathway flux scaled approximately linearly with total budget ($R^2 = 0.9998$), indicating that the pathway remains bottleneck-limited across all tested expression levels. This linearity implies that even at 150× total overexpression, no single enzyme achieves saturating capacity.

3.4.4 Substantial Flux Improvements

Even at the modest 20× budget, optimization predicts 9.4-fold flux improvement over baseline. At the 150× budget, 70.7-fold improvement is predicted, yielding 11.5/µs pathway flux—substantially closing the gap between stoichiometric potential and kinetic reality.

3.5 Mechanistic Interpretation of the 1:7:7 Ratio

The optimized ratio directly reflects the underlying bottleneck scores and enzyme kinetic properties:

1. **CrtI has the worst kinetics** ($k_{cat} = 0.25$/s, 13% saturation) and thus requires the highest expression to achieve target flux.
2. **CrtB has the lowest** k_{cat} (0.05/s) but benefits from higher substrate saturation (67%), partially compensating for its poor turnover number.
3. **CrtE has better kinetics** ($k_{cat} = 0.80$/s) but must produce twice the flux of downstream enzymes due to stoichiometry (2 GGPP → 1 Phytoene). Despite this 2× demand, its superior k_{cat} means it requires relatively low expression.

The 1:7:7 ratio effectively "equalizes" the capacities of all three enzymes, ensuring that the pathway flux is limited by the system as a whole rather than a single dominant bottleneck.

4 Discussion

4.1 Closing the Kinetic-Stoichiometric Gap

This work demonstrates that FBA predictions of lycopene production overestimate achievable flux by approximately **6000-fold** due to kinetic constraints in the heterologous pathway. While FBA predicts 1.057 mmol/gDW/h (1057/µs), the baseline kinetic capacity is only 0.163/µs. This massive discrepancy highlights the critical importance of integrating enzyme kinetics into strain design workflows [5, 6].

Importantly, our optimization framework demonstrates that this gap can be substantially narrowed through rational enzyme expression tuning. At a 150× total overexpression budget, the predicted flux of 11.5/µs represents 70.7-fold improvement over baseline—still ∼92-fold below the FBA prediction, but within a realistic engineering range achievable through standard molecular biology techniques.

4.2 Robustness of the Optimal Expression Ratio

The striking consistency of the 1:7:7 expression ratio across diverse budget constraints has several important implications:

4.2.1 Strain Construction Flexibility

Engineers can implement this ratio using various molecular strategies (promoter libraries, plasmid copy number variation, chromosomal integration with different expression elements) without needing to precisely match absolute expression levels. The ratio provides a robust target that accommodates experimental variability.

4.2.2 Transferability Across Conditions

The ratio should remain valid across different *E. coli* strains and growth conditions, provided that relative kinetic parameters and substrate concentrations are preserved. This suggests that the optimization results may generalize beyond the specific conditions modeled.

4.2.3 Diagnostic Value

Deviations from this ratio in experimental implementations would suggest changes in effective kinetic parameters (e.g., due to metabolic burden, protein aggregation, or post-translational regulation) and warrant investigation.

4.3 Practical Implementation Strategies

4.3.1 Promoter-Based Approaches

For low-to-medium budgets (≤50×), a three-plasmid system could implement the 1:7:7 ratio:

- **CrtE**: Weak constitutive promoter (e.g., J23117, ∼5× baseline)
- **CrtB**: Strong constitutive promoter (e.g., J23100, ∼40× baseline)

- **CrtI**: T7 or trc inducible promoter (∼50× baseline at full induction)

4.3.2 Multi-Copy Plasmid Strategies

For high budgets (100–150×), additional copy number control is required:

- **CrtE**: Low-copy plasmid (pSC101, ∼5 copies) with moderate promoter
- **CrtB**: Medium-copy plasmid (p15A, ∼15 copies) with strong promoter
- **CrtI**: High-copy plasmid (ColE1, ∼30 copies) with strong promoter

4.3.3 Hybrid Approaches

Chromosomal integration of CrtE (where low expression is needed) combined with multi-copy plasmids for CrtB and CrtI would reduce metabolic burden while maintaining the target ratio.

4.4 Validation Strategy

Experimental validation should follow a systematic Design-Build-Test-Learn approach:

1. **Construct ratio variants**: Build strains with expression ratios spanning 1:5:5 to 1:10:10, with the predicted optimum (1:7:7) at the center.
2. **Measure lycopene production**: Quantify lycopene titers using HPLC or spectrophotometric assays across the ratio series.
3. **Metabolomic profiling**: Measure intermediate concentrations (GGPP, phytoene) to validate flux predictions and identify any remaining bottlenecks.
4. **Kinetic parameter verification**: If experimental flux deviates substantially from predictions, re-measure k_{cat} and K_{m} values *in vivo* using enzyme assays or ^{13}C-metabolic flux analysis.
5. **Iterative refinement**: Use experimental data to update the kinetic model and re-optimize, implementing a true Design-Build-Test-Learn cycle.

4.5 Limitations and Future Directions

4.5.1 Model Assumptions

Our model incorporates several simplifying assumptions that should be addressed in future work:

1. **Linear expression-activity relationship**: We assume enzyme activity scales linearly with expression level. In reality, high overexpression can lead to protein aggregation, reducing effective activity.
2. **Static kinetic parameters**: k_{cat} and K_{m} are assumed constant, but may vary with intracellular conditions (pH, ionic strength, metabolite concentrations).

3. **Independent enzymes**: We do not account for enzyme-enzyme interactions, protein complexes, or substrate channeling, which could alter effective kinetics.

4. **No metabolic burden**: The optimization does not penalize high expression levels for their impact on growth rate or cellular fitness. In practice, a 150× total overexpression may impose significant burden.

5. **Fixed substrate concentrations**: We use estimated steady-state concentrations, but these will change as pathway flux increases.

4.5.2 Recommended Extensions

Future development should address these limitations through:

1. **Dynamic flux balance analysis (dFBA)**: Integrate time-dependent growth kinetics to assess trade-offs between production and growth.

2. **Metabolic burden modeling**: Incorporate resource allocation constraints (ribosome, ATP, amino acid availability) to predict fitness costs [16].

3. **Enzyme engineering integration**: Use the bottleneck scores to prioritize targets for directed evolution (CrtI and CrtB are clear candidates for k_{cat} improvement).

4. **Regulatory network integration**: Extend the model to include transcriptional regulation (e.g., feedback inhibition by isoprenoids) and post-translational modifications.

5. **Multi-objective optimization**: Balance lycopene production with growth rate, strain stability, and precursor supply for industrially relevant objectives.

5 Conclusion

We have demonstrated an integrated kinetic-stoichiometric framework that bridges the gap between theoretical flux predictions and experimental reality in metabolic engineering. By systematically combining genome-scale FBA with enzyme kinetics, we identified severe bottlenecks in the heterologous lycopene biosynthesis pathway that would be invisible to stoichiometric analysis alone. Multi-objective optimization revealed a robust 1:7:7 expression ratio (CrtE:CrtB:CrtI) that remains optimal across diverse resource budgets and predicts substantial flux improvements (9–71 fold) over baseline.

This work provides three key contributions to the field of metabolic engineering:

1. **Diagnostic framework**: A generalizable method to identify true kinetic bottlenecks by quantitatively comparing enzyme capacities to FBA flux demands.

2. **Engineering roadmap**: Specific, actionable expression targets for lycopene production that can guide experimental strain construction.

3. **Validation of approach**: Demonstration that kinetic-stoichiometric integration yields robust, budget-independent optimization solutions—a critical feature for practical implementation.

Moving forward, this framework can be extended to other heterologous pathways and integrated with enzyme engineering, regulatory network modeling, and metabolic burden constraints to create a comprehensive computational platform for next-generation metabolic engineering. The identification of the 1:7:7 ratio provides a clear starting point for experimental validation, with the potential to substantially improve lycopene titers in engineered *E. coli* strains.

Data Availability

All data and analysis scripts are available in the project repository:

- FBA results: `results/fba_summary.txt` and `results/fba_fluxes.csv`
- Kinetic parameters: `results/kinetic_data_summary.txt`
- Bottleneck analysis: `results/bottleneck_summary.txt` and `results/bottleneck_analysis.csv`
- Optimization results: `results/optimization_summary.txt` and `results/optimization_results.csv`

Code Availability

Analysis code was implemented in Python 3.12 using:

- COBRApy v0.29.0 (FBA)
- NumPy v1.26.3 (numerical operations)
- Pandas v2.1.4 (data manipulation)
- Matplotlib v3.8.2 (visualization)
- PyMOO v0.6.1 (optimization)

Author Contributions

K-Dense Web performed model construction, data curation, bottleneck analysis, optimization, and manuscript preparation.

Acknowledgments

We acknowledge the BiGG Models database and BRENDA enzyme database for providing curated metabolic and enzyme kinetic data essential to this study.

References

[1] Francis X Cunningham and Elisabeth Gantt. Genes and enzymes of carotenoid biosynthesis in plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, 49(1):557–583, 1998. doi: 10.1146/annurev.arplant.49.1.557.

[2] Juan Zhao, Qianyu Li, Tao Sun, Xiangnan Zhu, Huangtao Xu, Jialin Tang, Xueli Zhang, and Yanhe Ma. Improving E. coli for lycopene production using metabolic engineering approaches. *Applied Microbiology and Biotechnology*, 97(7):3045–3055, 2013. doi: 10.1007/s00253-013-4631-9.

[3] Jorge Alonso-Gutierrez, Eun-Mi Kim, Tanveer S Batth, Nohyoung Cho, Qijun Hu, Leanne Jade G Chan, Christopher J Petzold, Nathan J Hillson, Paul D Adams, Jay D Keasling, Hector Garcia Martin, and Taek Soon Lee. Lycopene production in Saccharomyces cerevisiae through combinatorial genetic and pathway engineering. *Biotechnology and Bioengineering*, 118(4):1442–1455, 2021. doi: 10.1002/bit.27665.

[4] Jeffrey D Orth, Tom M Conrad, Jessica Na, Joshua A Lerman, Hojung Nam, Adam M Feist, and Bernhard Ø Palsson. A comprehensive genome-scale reconstruction of Escherichia coli metabolism–2011. *Molecular Systems Biology*, 7(1):535, 2011. doi: 10.1038/msb.2011.65.

[5] Benjamín J Sánchez, Cheng Zhang, Avlant Nilsson, Petri-Jaan Lahtvee, Eduard J Kerkhoven, and Jens Nielsen. Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Molecular Systems Biology*, 13(8):935, 2017. doi: 10.15252/msb.20167411.

[6] Ali Khodayari and Costas D Maranas. A genome-scale Escherichia coli kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nature Communications*, 7(1):13806, 2016. doi: 10.1038/ncomms13806.

[7] Parayil Kumaran Ajikumar, Wen-Hai Xiao, Keith E J Tyo, Yong Wang, Fritz Simeon, Effendi Leonard, Oliver Mucha, Too Heng Phon, Blaine Pfeifer, and Gregory Stephanopoulos. Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science*, 330(6000):70–74, 2010. doi: 10.1126/science.1191652.

[8] Christopher S Henry, Linda J Broadbelt, and Vassily Hatzimanikatis. Thermodynamics-based metabolic flux analysis. *Biophysical Journal*, 92(5):1792–1805, 2007. doi: 10.1529/biophysj.106.093138.

[9] Dan Davidi, Elad Noor, Wolfram Liebermeister, Arren Bar-Even, Avi Flamholz, Katja Tummler, Uri Beriger, Shmuel Arebbi, and Ron Milo. Global characterization of in vivo enzyme catalytic rates and their correspondence to in vitro k_{cat} measurements. *Proceedings of the National Academy of Sciences USA*, 113(12):3401–3406, 2016. doi: 10.1073/pnas.1514240113.

[10] Lisa Jeske, Sandra Placzek, Ida Schomburg, Antje Chang, and Dietmar Schomburg. BRENDA in 2019: a European ELIXIR core data resource. *Nucleic Acids Research*, 47(D1):D542–D549, 2019. doi: 10.1093/nar/gky1048.

[11] Julian Blank and Kalyanmoy Deb. pymoo: Multi-objective optimization in Python. *IEEE Access*, 8:89497–89509, 2020. doi: 10.1109/ACCESS.2020.2990567.

[12] Zachary A King, Justin Lu, Andreas Dräger, Philip Miller, Stephen Federowicz, Joshua A Lerman, Ali Ebrahim, Bernhard O Palsson, and Nathan E Lewis. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Research*, 44 (D1):D515–D522, 2016. doi: 10.1093/nar/gkv1049.

[13] Ali Ebrahim, Joshua A Lerman, Bernhard O Palsson, and Daniel R Hyduke. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology*, 7(1): 74, 2013. doi: 10.1186/1752-0509-7-74.

[14] Bryson D Bennett, Elizabeth H Kimball, Melissa Gao, Robin Osterhout, Stephen J Van Dien, and Joshua D Rabinowitz. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature Chemical Biology*, 5(8):593–599, 2009. doi: 10.1038/nchembio.186.

[15] Ron Milo, Paul Jorgensen, Uri Moran, Griffin Weber, and Michael Springer. BioNumbers–the database of key numbers in molecular and cell biology. *Nucleic Acids Research*, 38 (Database issue):D750–D753, 2009. doi: 10.1093/nar/gkp889.

[16] Qamar K Beg, Alexei Vazquez, Jason Ernst, Marco Antonio de Menezes, Ziv Bar-Joseph, Albert-László Barabási, and Zoltán N Oltvai. Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proceedings of the National Academy of Sciences USA*, 104(31):12663–12668, 2007. doi: 10.1073/pnas.0609845104.