Medicon Medical Sciences
Volume 8 Issue 4 April 2025
Article Type: Research Article
ISSN: 2972-2721

Ab Initio Whole Cell Kinetic Model of *Yarrowia lipolytica* CLIB122 (yliYKY24)

Keng Yao Yeo[1,2], Madhunisha Arivazhagan[1,2], Ashmitha Senthilkumar[1,2], Tanisha Saisudhanbabu[1,2], Minh Anh Le[1,2], Travina BS Wong[1,2], Victor R Lukianto[1,2] and Maurice HT Ling[1,2,3]*

[1]*School of Life Sciences, Management Development Institute of Singapore, Singapore*

[2]*Department of Applied Sciences, Northumbria University, United Kingdom*

[3]*HOHY PTE LTD, Singapore*

***Corresponding Author:** Maurice HT Ling, School of Life Sciences, Management Development Institute of Singapore, Singapore; Department of Applied Sciences, Northumbria University, United Kingdom; HOHY PTE LTD, Singapore.

Received: March 04, 2025; **Published:** April 05, 2025

DOI: 10.55162/MCMS.08.281

Abstract

Yarrowia lipolytica is an oleaginous yeast with potential probiotics properties that may be engineered for various industrial applications and the presence of mathematical models (genome-scale models or GSMs, and kinetic models or KMs) will be a valuable tool to guide engineering approaches. To date, there are several GSMs for *Y. lipolytica* but no whole-cell KM. Here, we present a whole cell simulatable KM of Y. lipolytica, yliYKY24, based on the genome of *Y. lipolytica* CLIB122. The resulting model consists of 931 metabolites, 703 enzymes with corresponding transcriptions and translations, and 856 enzymatic reactions; which can be a baseline model for incorporating other cellular and growth processes, or as a system to examine cellular resource allocations necessary for engineering.

Introduction

Yarrowia lipolytica is an aerobic oleaginous yeast commonly found in oil-rich habitats [1]. It is capable of storing large amount of lipids [2], and can produce a variety of organic acids [3]. Generally recognized as safe [4], *Y. lipolytica* is currently used in the food industry as flavour enhancers [5], and considered for its potential probiotics properties [6, 7]. It also possess potential industrial applications; such as, bio-oil production [8], wastewater treatment [9], and platform chemical production [10]. These often require metabolic engineering or genetic engineering [11, 12].

Mathematical modelling is an important tool in metabolic engineering [13] as it can predict biological phenotypes under metabolic perturbations, which can be used to guide engineering approaches [14, 15]. Genome-scale models (GSMs, also known as constraint-based models) and kinetic models (KMs) are the two main modelling approaches [16, 17]. GSMs are steady-state stoichiometric models which lacks enzymatic regulation [18] and is difficult to add genes (transgenes) into the system as its original purpose is to evaluate changes in native gene expression on its metabolism [19, 20]. On the other hand, kinetic models (KMs) can have regulation and is much easier to add transgenes. Furthermore, KMs can predict both rates and yield of metabolites [21] while GSMs are primarily for rates. At the same time, large-scale KMs; such as, whole cell KMs; offer more intricate details [22] and have important applications [23, 24] compared to smaller-scale KMs; such as, KMs of pathways. There are a number of GSMs for Y. lipolytica; for example, iMK735 [25], and iYli21 [26]. However, there is no whole cell KM of *Y. lipolytica* to date.

Therefore, this study presents the construction a KM of *Y. lipolytica* CLIB122 using *ab initio* approach by identifying enzymes from its genome [27], and linking to its corresponding reaction from KEGG [28]. The result is a whole cell KM of *Y. lipolytica* CLIB122, named as yliYKY24 using the nomenclature proposed by Cho and Ling [29], which consists of 931 metabolites, 703 enzymes with corresponding transcriptions and translations, and 856 enzymatic reactions.

Methods

Reactome Identification

Enzymatic genes were identified from the genome of *Yarrowia lipolytica* CLIB122 (NCBI RefSeq GCF_000002525.2) using the process adapted from previous studies [30, 31]. Briefly, each enzymatic gene was identified as a presence of complete Enzyme Commission (EC) number directly from Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Database for *Y. lipolytica* CLIB122 (Figure 1) or indirectly from UniprotKB and GenBank via the coding sequence's protein ID or locus tag. Each EC number is then mapped into reaction IDs via KEGG Enzyme Database for Enzyme Nomenclature [28].

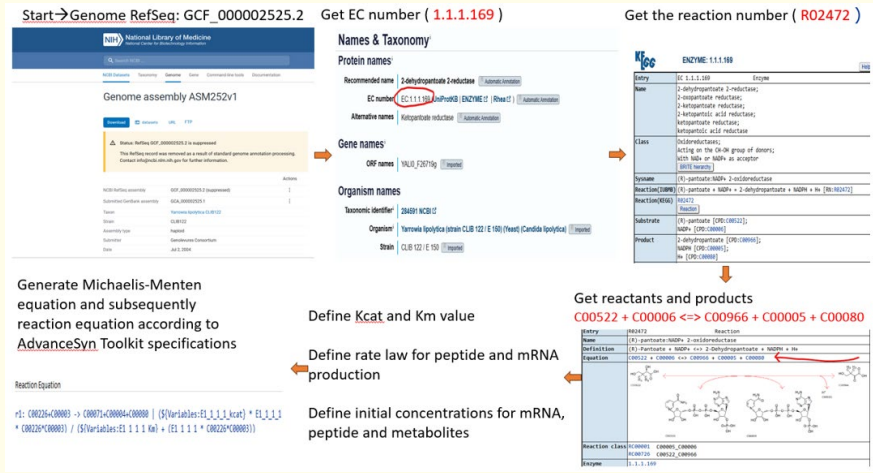

Figure 1: Flowchart of reactome identification and model development.

Model Development

The production of each enzyme was modelled as a pair of ODEs - an ODE for transcription, and an ODE for translation. Given that each *E. coli* cell is has about 750 active units of RNA polymerase (BioNumbers 106199 [32] and 111676 [33]) polymerizing 22 ribonucleotides per second (BioNumbers 104109) [34]. With 339.5 Daltons per ribonucleotide, implying the total mRNA synthesis rate at 5600 kDa per second. One Dalton is 1.66054e-24 gram; hence 5600 kDa per second is 9.3e-18 grams per second. An *E. coli* cell is about 0.7 cubic micrometres [35] or 7e-16 litres with 4225 protein-coding genes (BioNumbers 105443) [36], the total mRNA synthesis rate can be estimated at 2.92 uM per protein-coding genes per second. The average lifespan of mRNA transcripts is 1.79 minutes (BioNumbers 107666) [37] or 107.56 seconds; therefore, 0.93% degraded per second. Therefore, the rate law for mRNA concentration can be written as d[mRNA]/dt = (0.00292 - 0.0093[mRNA]) mM per second. Similarly, the median protein synthesis in mammalian cell culture is 1000 peptides per mRNA transcript per hour (BioNumbers 106382) [38], which equates to 0.278 peptides per mRNA transcripts per second; and the average protein degradation rate for *E. coli* is about 1 percent per hour (BioNumbers 109924) [39], which equates to 0.00000278 per second; the rate law for peptide concentration can be written as d[peptide]/dt = (0.278[mRNA] - 0.00000278[peptide]) uM per second. The reactome was modelled as a set of ordinary differential equations (ODEs) where each ODE

represented one metabolite concentration [30, 40]. The turnover number of enzyme (Kcat) and Michaelis-Menten constant (Km) were set at 13.7 per second and 1 millimolar respectively; which were the median values estimated by Bar-Even et al. [41]. The model was written in AdvanceSyn Model Specification [16].

Model Simulation

The constructed model was tested for simulatability using AdvanceSyn Toolkit [16]. Initial concentrations of all mRNA and enzymes were set to 0 mM. Initial concentrations of all metabolites were set to 1 mM except the following which were set to 1000 mM: (i) C00001 (water), (ii) C00002 (ATP), (iii) C00003 (NAD+), (iv) C00004 (NADH), (v) C00005 (NADPH), (vi) C00006 (NADP+), (vii) C00007 (Oxygen), (viii) C00008 (ADP), (ix) C00009 (Phosphate), (x) C00011 (Carbon dioxide), (xi) C00014 (Ammonia), (xii) C00025 (Glutamate), (xiii) C00031 (Glucose), (xiv) C00041 (L-Alanine), (xv) C00047 (L-Lysine), (xvi) C00049 (L-Aspartate), (xvii) C00062 (L-Arginine), (xviii) C00064 (L-Glutamine), (xix) C00065 (L-Serine), (xx) C00078 (L-Tryptophan), (xxi) C00079 (L-Phenylalanine), (xxii) C00082 (L-Tyrosine), (xxiii) C00097 (L-Cysteine), (xxiv) C00123 (L-Leucine), (xxv) C00135 (L-Histidine), (xxvi) C00148 (L-Proline), (xxvii) C00152 (L-Asparagine), (xxviii) C00183 (L-Valine), (xxix) C00188 (L-Threonine) and (xxx) C00407 (L-Isoleucine).The model was simulated using fourth-order Runge-Kutta method [42, 43] from time zero to 3600 seconds with timestep of 0.1 second, and the concentrations of metabolites were bounded between 0 millimolar and 1000 millimolar. The simulation results were sampled every 2 seconds.

Results and Discussion

The genome of *Yarrowia lipolytica* CLIB122 has 7144 coding sequences, with 1838 EC numbers. Out of 1838 EC numbers, 703 are unique and used to identify 856 reactions consisting of 931 metabolites. The resulting model, denoted as yliYKY24, was simulated using AdvanceSyn Toolkit [16]. The presence of simulation results (Figure 2) suggests that the model is free from syntax error.

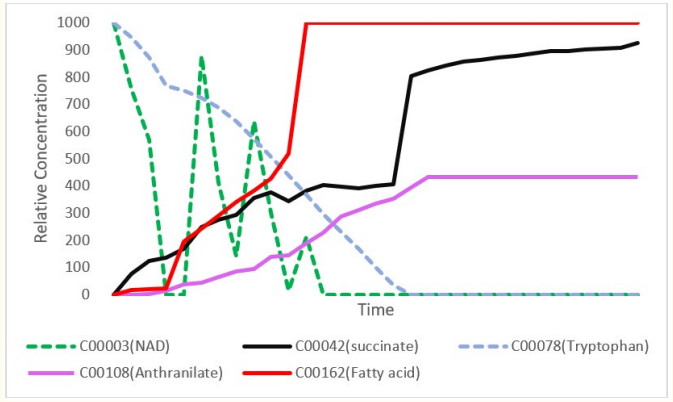

Figure 2: Selection of simulation results.

Our simulation result shows a gradual increase in anthranilate's concentration with a corresponding drop in tryptophan's concentration, indicating some form of feedback inhibition by end-product tryptophan on anthranilate precursor. This type of feedback inhibition is seen commonly in many biochemical processes, including tryptophan inhibition on anthranilate synthase [44]. NAD and fatty acid concentration also shows an inverse relationship, which might be explained by NAD involvement in beta-oxidation of fatty acid; hence, the increase in fatty acid concentration along with NAD depletion. However, the simulation results cannot be taken at face value since all enzyme kinetics (turnover number and Michaelis-Menten constant) are kept the median levels [41], which implies that

all enzymes are producing at same rate and with the sample activity. Hence, the current model represents a whole cell structural model that contains all the necessary kinetic parameters and can be simulated without errors.

Therefore, we present a simulatable whole cell KM of *Y. lipolytica* CLIB122, which can be a base template for incorporating other cellular and growth processes [45-47] or as a system to examine cellular resource allocations [24, 48-50] as it contains comparable number of metabolites, and reactions (931 metabolites and 856 reactions) compared to iMK735 (1111 metabolites and 1136 reactions) [25]. Future work can extend the number of metabolites and reactions to that of iYli21 [26] - 1868 metabolites and 2285 reactions.

Conclusion

In this study, we present an *ab initio* whole cell kinetic model of *Y. lipolytica* CLIB122, yliYKY24; comprising of 931 metabolites, 703 enzymes with corresponding transcriptions and translations, and 856 enzymatic reactions.

Supplementary Materials

Reaction descriptions and model can be download from https://bit.ly/yliYKY24.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Jach ME and Malm A. "Yarrowia lipolytica as an Alternative and Valuable Source of Nutritional and Bioactive Compounds for Humans". Molecules (Basel, Switzerland) 27.7 (2022): 2300.
2. Niehus X., et al. "Engineering Yarrowia lipolytica to Enhance Lipid Production from Lignocellulosic Materials". Biotechnology for Biofuels 11.1 (2018): 11.
3. Fickers P, Cheng H and Sze Ki Lin C. "Sugar Alcohols and Organic Acids Synthesis in Yarrowia lipolytica: Where Are We?". Microorganisms 8.4 (2020): 574.
4. Bilal M., et al. "Yarrowia lipolytica as an Emerging Biotechnological Chassis for Functional Sugars Biosynthesis". Critical Reviews in Food Science and Nutrition 61.4 (2021): 535-552.
5. Zinjarde SS. "Food-Related Applications of Yarrowia lipolytica". Food Chemistry 152 (2014): 1-10.
6. Keskin P, Kılıç Kanak E and Öztürk Yılmaz S. "Assessment of the Probiotic Properties of Yarrowia lipolytica Isolated from Cold-Pressed Olive Oil". Microorganisms 12.9 (2024): 1905.
7. Reyes-Becerril M, Alamillo E and Angulo C. "Probiotic and Immunomodulatory Activity of Marine Yeast Yarrowia lipolytica Strains and Response Against Vibrio parahaemolyticus in Fish". Probiotics and Antimicrobial Proteins 13.5 (2021): 1292-1305.
8. Beopoulos A., et al. "Yarrowia lipolytica as a Model for Bio-Oil Production". Progress in Lipid Research 48.6 (2009): 375-387.
9. Mohiuddin O., et al. "Bioremediation of Waste by Yeast Strains". Electronic Journal of Biotechnology 69 (2024): 30-42.
10. Fu J., et al. "Reprogramming Yarrowia lipolytica Metabolism for Efficient Synthesis of Itaconic Acid from Flask to Semipilot Scale". Science Advances 10.32 (2024): eadn0414.
11. Ling MH., et al. "Yarrowia lipolytica as a Valorization Biofactory". SynBioSG Conference 2024 (Matrix, Biopolis, Singapore) (2024).
12. Lukianto VR., et al. "A Systematic Review (Before 31 August 2024) on the Applications of Yarrowia lipolytica". Acta Scientific Microbiology 8.3 (2025): 58-65.
13. Khanijou JK., et al. "Metabolomics and Modelling Approaches for Systems Metabolic Engineering". Metabolic Engineering Communications 15 (2022): e00209.
14. Gudmundsson S and Nogales J. "Recent Advances in Model-Assisted Metabolic Engineering". Current Opinion in Systems Biology 28 (2021): 100392.

15. Strutz J., et al. "Metabolic Kinetic Modeling Provides Insight into Complex Biological Questions, but Hurdles Remain". Current Opinion in Biotechnology 59 (2019): 24-30.
16. Ling MH. "AdvanceSyn Toolkit: An Open Source Suite for Model Development and Analysis in Biological Engineering". MOJ Proteomics & Bioinformatics 9.4 (2020): 83-86.
17. Richelle A., et al. "Towards a Widespread Adoption of Metabolic Modeling Tools in Biopharmaceutical Industry: A Process Systems Biology Engineering Perspective". npj Systems Biology and Applications 6.1 (2020): 6.
18. Srinivasan S, Cluett WR and Mahadevan R. "Constructing Kinetic Models of Metabolism at Genome-Scales: A Review". Biotechnology Journal 10.9 (2015): 1345-1359.
19. Gu C., et al. "Current Status and Applications of Genome-Scale Metabolic Models". Genome Biology 20.1 (2019): 121.
20. O'Brien EJ, Monk JM and Palsson BO. "Using Genome-scale Models to Predict Biological Capabilities". Cell 161.5 (2015): 971-987.
21. Prabhu S., et al. "Derivative-Free Domain-Informed Data-Driven Discovery of Sparse Kinetic Models". Industrial & Engineering Chemistry Research 64.5 (2025): 2601-2615.
22. Lucido A., et al. "Multiscale Mathematical Modeling in Systems Biology: A Framework to Boost Plant Synthetic Biology". Plants 14.3 (2025): 470.
23. Carrera J and Covert MW. "Why Build Whole-Cell Models?". Trends in Cell Biology 25.12 (2015): 719-722.
24. Thornburg ZR., et al. "Fundamental Behaviors Emerge from Simulations of a Living Minimal Cell". Cell 185.2 (2022): 345-360. e28.
25. Kavšček M., et al. "Optimization of Lipid Production with a Genome-Scale Model of Yarrowia lipolytica". BMC systems biology 9 (2015): 72.
26. Guo Y., et al. "Dissecting Carbon Metabolism of Yarrowia lipolytica Type Strain W29 Using Genome-Scale Metabolic Modelling". Computational and Structural Biotechnology Journal 20 (2022): 2503-2511.
27. Yu DS., et al. "Complete Genome Sequence of the Probiotic Bacterium Bifidobacterium bifidum Strain BGN4". Journal of Bacteriology 194.17 (2012): 4757-4758.
28. Okuda S., et al. "KEGG Atlas mapping for global analysis of metabolic pathways". Nucleic Acids Research 36(Web Server issue) (2008): W423-W426.
29. Cho JL and Ling MH. "Adaptation of Whole Cell Kinetic Model Template, UniKin1, to Escherichia coli Whole Cell Kinetic Model, ecoJC20". EC Microbiology 17.2 (2021): 254-260.
30. Kwan ZJ., et al. "Ab Initio Whole Cell Kinetic Model of Stutzerimonas balearica DSM 6083 (pbmKZJ23)". Acta Scientific Microbiology 7.2 (2024): 28-31.
31. Arivazhagan M., et al. "Ab Initio Whole Cell Kinetic Model of Bifidobacterium bifidum BGN4 (bbfMA24)". Acta Scientific Nutritional Health 9.1 (2025): 42-45.
32. Müller-Hill B. The lac Operon: A Short History of a Genetic Paradigm (Berlin, Germany) (1996).
33. Churchward G, Bremer H and Young R. "Transcription in Bacteria at Different DNA Concentrations". Journal of Bacteriology 150.2 (1982): 572-581.
34. Gray WJ and Midgley JE. "The Control of Ribonucleic Acid Synthesis in Bacteria. The Synthesis and Stability of Ribonucleic Acid in Rifampicin-Inhibited Cultures of Escherichia coli". The Biochemical Journal 122.2 (1971): 161-169.
35. Kubitschek HE. "Cell Volume Increase in Escherichia coli After Shifts to Richer Media". Journal of Bacteriology 172.1 (1990): 94-101.
36. Hu P., et al. "Global Functional Atlas of Escherichia coli Encompassing Previously Uncharacterized Proteins". PLoS biology 7.4 (2009): e96.
37. So L-H., et al. "General Properties of Transcriptional Time Series in Escherichia coli". Nature Genetics 43.6 (2011): 554-560.
38. Schwanhäusser B., et al. "Corrigendum: Global Quantification of Mammalian Gene Expression Control". Nature 495.7439 (2013): 126-127.
39. Maurizi MR. "Proteases and Protein Degradation in Escherichia coli". Experientia 48.2 (1992): 178-201.

40. Murthy MV., et al. "UniKin1: A Universal, Non-Species-Specific Whole Cell Kinetic Model". Acta Scientific Microbiology 3.10 (2020): 04-08.
41. Bar-Even A., et al. "The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters". Biochemistry 50.21 (2011): 4402-4410.
42. Ling MH. "COPADS IV: Fixed Time-Step ODE Solvers for a System of Equations Implemented as a Set of Python Functions". Advances in Computer Science: an International Journal 5.3 (2016): 5-11.
43. Yong B. "The Comparison of Fourth Order Runge-Kutta and Homotopy Analysis Method for Solving Three Basic Epidemic Models". Journal of Physics: Conference Series 1317 (2019): 012020.
44. Bhagat AK., et al. "Photoswitching of Feedback Inhibition by Tryptophan in Anthranilate Synthase". ACS synthetic biology 11.8 (2022): 2846-2856.
45. Ahn-Horst TA., et al. "An Expanded Whole-Cell Model of E. coli Links Cellular Physiology with Mechanisms of Growth Rate Control". npj Systems Biology and Applications 8.1 (2022): 30.
46. Chagas M da S., et al. "Boolean Model of the Gene Regulatory Network of Pseudomonas aeruginosa CCBH4851". Frontiers in Microbiology 14 (2023): 1274740.
47. Hao T., et al. "Reconstruction of Metabolic-Protein Interaction Integrated Network of Eriocheir sinensis and Analysis of Ecdysone Synthesis". Genes 15.4 (2024): 410.
48. Bianchi DM., et al. "Toward the Complete Functional Characterization of a Minimal Bacterial Proteome". The Journal of Physical Chemistry B 126.36 (2022): 6820-6834.
49. Sun G., et al. "Cross-Evaluation of E. coli's Operon Structures via a Whole-Cell Model Suggests Alternative Cellular Benefits for Low- Versus High-Expressing Operons". Cell Systems 15.3 (2024): 227-245.e7.
50. Choi H and Covert MW. "Whole-cell modeling of E. coli confirms that in vitro tRNA aminoacylation measurements are insufficient to support cell growth and predicts a positive feedback mechanism regulating arginine biosynthesis". Nucleic Acids Research 51.12 (2023): 5911-5930.

Volume 8 Issue 4 April 2025