Clareus Scientific Medical Sciences
Volume 3 Issue 1 January 2026
DOI: 10.70012/CSMS-03-015
ISSN: 3064-8017

Ab Initio Whole Cell Kinetic Model of *Streptococcus pneumoniae* NCTC 7465 (spnLHP26)

Citation: Maurice HT Ling., et al. "*Ab Initio* Whole Cell Kinetic Model of *Streptococcus pneumoniae* NCTC 7465 (spnLHP26)". Clareus Scientific Medical Sciences 3.1 (2026): 02-07.

Article Type: Research Article

Received: December 26, 2025

Published: January 26, 2026

Leesha Haarshiny Perumal[1,2], Atoshi Abirami RajKumar[1,2], Cheryl Kai Ning Kang[1,2], Sragvi Verma[1,2], Diya Nanthakumarvani[1,2], Shafeeqa Abul-Hasan[1,2] and Maurice HT Ling[2,3,4]*

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

[2]*Management Development Institute of Singapore, Singapore*

[3]*Newcastle Australia Institute of Higher Education, University of Newcastle, Australia*

[4]*HOHY PTE LTD, Singapore*

***Corresponding Author:** Maurice HT Ling, Management Development Institute of Singapore, Singapore; Newcastle Australia Institute of Higher Education, University of Newcastle, Australia; HOHY PTE LTD, Singapore.

Abstract

Streptococcus pneumoniae is a pathogen able to utilize various carbon sources to produce polysaccharide capsule to avoid the immune system, leading to responsible for many life-threatening infections. Kinetic models may be used to examine *S. pneumoniae*'s carbon utilization but there are no whole-cell kinetic models of *S. pneumoniae* to date. In this study, we construct a whole-cell kinetic model based on *S. pneumoniae* NCTC 7465 using its annotated genome. The resulting model, spnLHP26, consists of 460 enzymes catalysing 836 reactions involving 163 metabolites; which may be suitable as a baseline draft model to examine virulence-associated metabolic susceptibilities of *S. pneumoniae*.

Keywords: Whole-cell model; Kinetic model; Differential equations; AdvanceSyn Toolkit

Introduction

Streptococcus pneumoniae is one of the Gram-positive pathogens causing life-threatening infections [1]; such as, pneumonia, meningitis, and septicaemia that prevail among young children, elderly, or immunocompromised persons [2]. More than 90 serotypes have been determined, and each serotype differs in the structure of the capsule and the possible virulence, demonstrating the capability of the bacterium to avoid the host immunity and to adapt rapidly to the changing host environment [3]. One of the key factors in studying pneumococcal metabolism is that it is directly related to virulence [4-6]. Unlike other pathogens, *S. pneumoniae*'s polysaccharide capsule [7], which can be synthesized from

various carbon sources [8], is the most important virulence factor [9] as it enables it to avoid the immune system [10].

Although kinetic modelling has applications in a number of fields in systems biology and metabolic engineering, such models to learn the underlying metabolic responses to the pathogen like *S. pneumoniae*, their responses to nutrient limitations, and how their metabolic networks reorganize to deal with antimicrobial stresses [11]. Kinetic models can also show the possible drug targets, which are highly controlled in their flux or sensitive to perturbation by measuring enzyme regulation and changes in metabolites [12, 13].

However, there is no whole cell kinetic models of *S. pneumoniae* to date. Hence, this study aims to construct a KM of *S. pneumoniae* NCTC7465 using *ab initio* approach by identifying enzymes from its published genome, and identifying the corresponding reaction from KEGG [14]. The result is a whole cell KM of *S. pneumoniae* NCTC7465, named as spnLHP26 using the nomenclature proposed by Cho and Ling [15], which consists of 1014 metabolites, 460 enzymes with corresponding transcriptions and translations, and 836 enzymatic reactions.

Materials and Methods

Identification of Reactome

The annotated genome of *Streptococcus pneumoniae* NCTC7465 [16] (NCBI RefSeq assembly GCF_001457635.1; NCBI GenBank Accession NZ_LN831051.1) was used as source to identify enzymatic genes using the process previously described [17-19]. Briefly, each enzymatic gene was identified as a presence of complete Enzyme Commission (EC) number in the GenBank record and mapped into reaction IDs via KEGG Ligand Database for Enzyme Nomenclature [14]. For example, EC 1.1.1.23 (https://www.genome.jp/entry/1.1.1.23) catalyses reactions R01158, R01163, and R03012; where the substrates and products of each rection can be identified.

Model Development

The model was developed using the principles described in Sim et al. [20]. Drawing from BioNumbers, *E. coli* carries about 3000 RNA polymerases (BioNumbers 106199) [21], with roughly 25% engaged in transcription (BioNumbers 111676) [22]. At an elongation rate of 22 nt/s (BioNumbers 104109) [23] and a nucleotide mass of 339.5 Da, this corresponds to ~5600 kDa of RNA synthesized each second, or 9.3×10^{-18} g/s. When distributing this across the 0.7 cubic micrometres cell volume [24] and 4225 protein-coding genes (BioNumbers 105443) [25], the resulting transcription rate is ~2.92 micromolar per gene per second. With a mean transcript life of 107.56 s (BioNumbers 107666) [26] (0.93% decay per second), the mRNA rate law becomes d[mRNA]/dt = 0.00292 - 0.0093[mRNA]. Translation proceeds at about 1000 peptides per transcript per hour (0.278/s) (BioNumbers 106382) [27], while protein degradation occurs at 1%/h (2.78×10^{-6}/s) (BioNumbers 109924) [28], giving: d[peptide]/dt = 0.278[mRNA] - 0.00000278[peptide]. Each metabolite in the reactome was then represented as an ODE [18, 29], using median kinetic constants (kcat = 13.7 s^{-1}; Km = 1 mM) [30] and written according to the AdvanceSyn Model Specification [31].

Model Simulation

The constructed model was tested for simulatability using AdvanceSyn Toolkit [31]. 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) C00011 (Carbon Dioxide), (IX) C00014 (Ammonia), (X) C00025 (L-Glutamate), (XI) C00031 (D-Glucose), (XII) C00037 (Glycine), (XIII) C00041 (L-Alanine), (XIV) C00047 (L-Lysine), (XV) C00049 (L-Aspartate), (XVI) C00064 (L-Glutamine), (XVII) C00065 (L-Serine), (XVIII) C00073 (L-Methionine), (XIX) C00097 (L-Cysteine), (XX) C00133 (D-Alanine), (XXI) C00138 (Reduced ferredoxin), (XXII) C00139 (Oxidized ferredoxin), (XXIII) C00148 (L-Proline), (XXIV) C00248 (Lipoamide), (XXV) C00579 (Dihydrolipoamide), (XXVI) C00662 (Reduced adrenal ferredoxin), (XXVII) C00667 (Oxidized adrenal ferredoxin), (XXVIII) C01352 (FADH2), (XXIX) C01847 (FMNH2), (XXX) C02953 (7,8-Dihydrobiopterin), (XXXI) C02972 (Dihydrolipoylprotein), (XXXII) C03170 (Trypanothione disulfide), (XXXIII) C03451 (S-D-Lactoylglutathione), (XXXIV) C03541 (Tetrahydrofolyl-[Glu](n)), (XXXV) C03688 (Apo-[acyl-carrier protein]), (XXXVI) C03880 (N-Substituted aminoacyl-tRNA), (XXXVII) C03939 (Acetyl-[acp]), (XXXVIII) C04088 (Stearoyl-[a-

cyl-carrier protein]), (XXXIX) C04152 (rRNA containing N1-methylguanine), (XL) C04153 (rRNA containing N2-methylguanine), (XLI) C04157 (N1-Methylguanine in tRNA), (XLII) C04161 (tRNA containing a thionucleotide), (XLIII) C04419 (Carboxybiotin-carboxyl-carrier protein), (XLIV) C04574 (di-trans,poly-cis-Undecaprenyl diphosphate), (XLV) C04618 ((3R)-3-Hydroxybutanoyl-[acyl-carrier protein]), (XLVI) C04619 ((3R)-3-Hydroxydecanoyl-[acyl-carrier protein]), (XLVII) C04620 ((3R)-3-Hydroxyoctanoyl-[acyl-carrier protein]), (XLVIII) C04633 ((3R)-3-Hydroxypalmitoyl-[acyl-carrier protein]), (XLIX) C04688 ((3R)-3-Hydroxytetradecanoyl-[acyl-carrier protein]), (L) C04735 (Apo-[acetyl-CoA carboxylase]), (LI) C05167 (alpha-Amino acid), (LII) C05684 (Selenite), (LIII) C05928 ((6R)-10-Formyltetrahydropteroyldiglutamate), (LIV) C05929 (10-Formyl-THF-polyglutamate), (LV) C06424 (Tetradecanoic acid), (LVI) C08362 ((9Z)-Hexadecenoic acid), (LVII) C09332 (THF-L-glutamate), (LVIII) C11434 (2-C-Methyl-D-erythritol 4-phosphate), (LIX) C11435 (4-(Cytidine 5'-diphospho)-2-C-methyl-D-erythritol), (LX) C11437 (1-Deoxy-D-xylulose 5-phosphate), (LXI) C11475 (DNA containing guanine), (LXII) C11477 (Sugar), (LXIII) C11907 (dTDP-4-oxo-6-deoxy-D-glucose), (LXIV) C13378 (alpha, beta-Dihydroxyethyl-TPP), (LXV) C15585 (myo-Inositol phosphate), (LXVI) C15587 (Purine), (LXVII) C16310 (3-Hexenal), (LXVIII) C16348 (cis-3-Chloro-2-propenal), (LXIX) C16538 (1,5-Anhydro-D-mannitol), (LXX) C16551 (Alcophosphamide), (LXXI) C16565 (Aminopropylcadaverine), (LXXII) C16586 (2-Phenyl-1,3-propanediol monocarbamate), (LXXIII) C16587 (3-Carbamoyl-2-phenylpropionaldehyde), (LXXIV) C16595 (4-Hydroxy-5-phenyltetrahydro-1,3-oxazin-2-one), (LXXV) C16596 (5-Phenyl-1,3-oxazinane-2,4-dione), (LXXVI) C16614 (Thiopurine S-methylether), (LXXVII) C16633 (5-Fluorouridine), (LXXVIII) C16635 (5'-Deoxy-5-fluorocytidine), (LXXIX) C17234 (2-Aminobut-2-enoate), (LXXX) C18091 (Ethylnitronate), (LXXXI) C18796 ((2R)-2-Hydroxy-2-methylbutanenitrile), (LXXXII) C18902 (Methylselenic acid), (LXXXIII) C19081 ((4S)-Limonene-1,2-epoxide), (LXXXIV) C19909 (N-Acetyl-alpha-neuraminate), (LXXXV) C19910 (N-Acetyl-beta-neuraminate), (LXXXVI) C20258 ((2S,4S)-4-Hydroxy-2,3,4,5-tetrahydrodipicolinate), (LXXXVII) C20463 (Purine deoxyribonucleoside), (LXXXVIII) C20904 (2-Iminopropanoate), (LXXXIX) C20905 (2-Iminobutanoate), (XC) C21748 (5-Fluorouridine diphosphate), (XCI) C21749 (5-Fluorouridine triphosphate), (XCII) C21750 (5-Fluorodeoxyuridine diphosphate), (XCIII) C21751 (5-Fluorodeoxyuridine triphosphate), (XCIV) C22288 (5-Deoxy-D-ribose), (XCV) C22337 (D-Ribulose 1-phosphate), (XCVI) C22382 (Aceneuramic acid). The model was simulated using the fourth-order Runge-Kutta method [32, 33] 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 annotated genome of *S. pneumoniae* NCTC7465 consists of 2189 genes, including 1931 protein coding sequences. 460 unique EC numbers consisting of 836 enzymatic reactions involving 1014 metabolites were identified and developed into a model based on AdvanceSyn Model Specification [31]. In addition, 920 ODEs acting as placeholder for enzyme transcriptions and translations were added.

The spnLHP26 model was run through the AdvanceSyn Toolkit [31], and the existence of valid simulation traces (Figure 1) indicates that the model is syntactically sound and internally consistent as previously argued [17, 19, 34-38]. Considering the density of interactions encoded in a whole-cell kinetic model, error-free execution is a strong indication of structural robustness. The preferential use of L-Alanine over D-glucose may arise from the uniform application of median kinetic constants [39], which are used here purely for structural testing rather than biological accuracy. These values inevitably distort the relative magnitudes of metabolic fluxes. Nevertheless, the resulting model serves as a functional and extensible whole-cell kinetic representation of *S. pneumoniae* NCTC7465. It is well positioned for examining cellular resource allocations [40-44] or strain-specific adaptation [45].

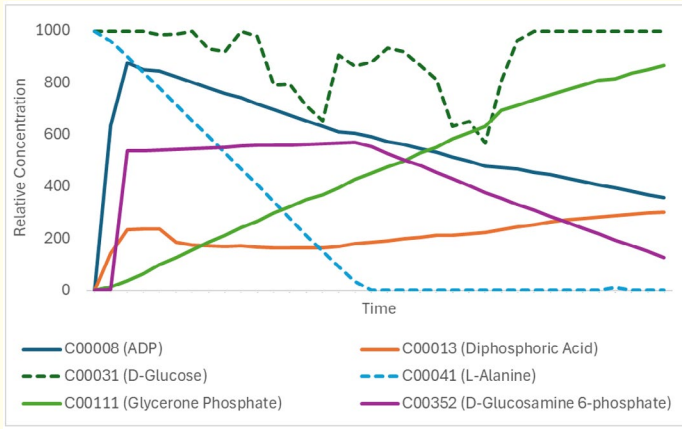

Figure 1: Selection of Simulation Results.

Conclusion

We present an *ab initio* whole cell kinetic model of *Streptococcus pneumoniae* NCTC7465, spnLHP26; comprising of 1014 metabolites, 469 enzymes with corresponding transcriptions and translations, and 836 enzymatic reactions.

Supplementary Materials

Reaction descriptions and model can be download from https://bit.ly/spnLHP26.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

The authors wish to thank the institute, Management Development Institute of Singapore, for its support towards this work. The cost of publication fees was borne by the authors.

References

1. Brooks LRK and Mias GI. "Streptococcus pneumoniae's Virulence and Host Immunity: Aging, Diagnostics, nd Prevention". Frontiers in Immunology 9 (2018): 1366.
2. Lynch J and Zhanel G. "Streptococcus pneumoniae: Epidemiology, Risk Factors, and Strategies for Prevention". Seminars in Respiratory and Critical Care Medicine 30.02 (2009): 189-209.
3. Geno KA., et al. "Pneumococcal Capsules and Their Types: Past, Present, and Future". Clinical Microbiology Reviews 28.3 (2015): 871-899.
4. Leonard A and Lalk M. "Infection and Metabolism - Streptococcus pneumoniae Metabolism Facing the Host Environment". Cytokine 112 (2018): 75-86.
5. Obolski U., et al. "The Metabolic, Virulence and Antimicrobial Resistance Profiles of Colonising Streptococcus pneumoniae Shift after PCV13 Introduction in Urban Malawi". Nature Communications 14.1 (2023): 7477.
6. Im H., et al. "Alterations in Nutrient Availability in the Lungs During Streptococcus pneumoniae-Induced Pneumonia". Infection

and Immunity 93.11 (2025): e0038025.

7. Paton JC and Trappetti C. "Streptococcus pneumoniae Capsular Polysaccharide". Microbiology Spectrum 7.2 (2019).
8. Werren JP., et al. "Carbon Source-Dependent Capsule Thickness Regulation in Streptococcus pneumoniae". Frontiers in Cellular and Infection Microbiology 13 (2023): 1279119.
9. Mitchell AM and Mitchell TJ. "Streptococcus pneumoniae: Virulence Factors and Variation". Clinical Microbiology and Infection 16.5 (2010): 411-418.
10. Gao S., et al. "Bacterial Capsules: Occurrence, Mechanism, and Function". NPJ Biofilms and Microbiomes 10.1 (2024): 21.
11. Sertbas M and Ulgen KO. "Genome-Scale Metabolic Modeling for Unraveling Molecular Mechanisms of High Threat Pathogens". Frontiers in Cell and Developmental Biology 8 (2020): 566702.
12. Ohno S, Uematsu S and Kuroda S. "Quantitative Metabolic Fluxes Regulated by Trans-Omic Networks". Biochemical Journal 479.6 (2022): 787-804.
13. Toumpe I., et al. "The Dawn of High-Throughput and Genome-Scale Kinetic Modeling: Recent Advances and Future Directions". ACS Synthetic Biology 14.8 (2025): 2906-2919.
14. Okuda S., et al. "KEGG Atlas mapping for global analysis of metabolic pathways". Nucleic Acids Research 36 (2008): W423-W426.
15. 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.
16. Sharma S., et al. "Antimicrobial Susceptibility Pattern and Serotype Distribution of Streptococcus pneumoniae Isolates From a Hospital-Based Study in Chandigarh, North India". Cureus 14.1 (2022): e21437.
17. Yeo KY., et al. "Ab Initio Whole Cell Kinetic Model of Yarrowia lipolytica CLIB122 (yliYKY24)". Medicon Medical Sciences 8.4 (2025): 13-18.
18. Kwan ZJ., et al. "Ab Initio Whole Cell Kinetic Model of Stutzerimonas balearica DSM 6083 (pbmKZJ23)". Acta Scientific Microbiology 7.2 (2024): 28-31.
19. Maiyappan S., et al. "Four Ab Initio Whole Cell Kinetic Models of Bacillus subtilis 168 (bsuLL25) 6051-HGW (bshSM25), N33 (bsuN33SS25), FUA2231 (bsuGR25)". Journal of Clinical Immunology & Microbiology 6.2 (2025): 1-6.
20. Sim BJH, Tan NTF and Ling MHT. "Multilevel Metabolic Modelling Using Ordinary Differential Equations". Encyclopedia of Bioinformatics and Computational Biology (Second Edition), eds Ranganathan S, Cannataro M, Khan AM (Elsevier, Oxford) (2025): 491-498.
21. Müller-Hill B. "The lac Operon: A Short History of a Genetic Paradigm (Berlin, Germany)" (1996).
22. Churchward G, Bremer H and Young R. "Transcription in Bacteria at Different DNA Concentrations". Journal of Bacteriology 150.2 (1982): 572-581.
23. 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.
24. Kubitschek HE. "Cell Volume Increase in Escherichia coli After Shifts to Richer Media". Journal of Bacteriology 172.1 (1990): 94-101.
25. Hu P., et al. "Global Functional Atlas of Escherichia coli Encompassing Previously Uncharacterized Proteins". PLoS biology 7.4 (2009): e96.
26. So L-H., et al. "General Properties of Transcriptional Time Series in Escherichia coli". Nature Genetics 43.6 (2011): 554-560.
27. Schwanhäusser B., et al. "Corrigendum: Global Quantification of Mammalian Gene Expression Control". Nature 495.7439 (2013): 126-127.
28. Maurizi MR. "Proteases and Protein Degradation in Escherichia coli". Experientia 48.2 (1992): 178-201.
29. Murthy MV., et al. "UniKin1: A Universal, Non-Species-Specific Whole Cell Kinetic Model". Acta Scientific Microbiology 3.10 (2020): 04-08.
30. Bar-Even A., et al. "The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters". Biochemistry 50.21 (2011): 4402-4410.

31. Ling MH. "AdvanceSyn Toolkit: An Open Source Suite for Model Development and Analysis in Biological Engineering". MOJ Proteomics & Bioinformatics 9.4 (2020): 83-86.
32. 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.
33. 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.
34. Saisudhanbabu T., et al. "Ab Initio Whole Cell Kinetic Model of Limosilactobacillus fermentum EFEL6800 (lfeTS24)". EC Clinical and Medical Case Reports 8.4 (2025): 01-04.
35. Arivazhagan M., et al. "Ab Initio Whole Cell Kinetic Model of Bifidobacterium bifidum BGN4 (bbfMA24)". Acta Scientific Nutritional Health 9.1 (2025): 42-45.
36. Senthilkumar A., et al. "Ab Initio Whole Cell Kinetic Model of Lactobacillus acidophilus NCFM (lacAS24)". Journal of Clinical Immunology & Microbiology 6.1 (2025): 1-5.
37. Wong TB., et al. "Ab Initio Whole Cell Kinetic Models of Escherichia coli BL21 (ebeTBSW25) and MG1655 (ecoMAL25)". Scholastic Medical Sciences 3.2 (2025): 01-04.
38. Ambel WB., et al. "UniKin2 - A Universal, Pan-Reactome Kinetic Model". International Journal of Research in Medical and Clinical Science 3.2 (2025): 77-80.
39. Bar-Even A., et al. "The Moderately Efficient Enzyme: Futile Encounters and Enzyme Floppiness". Biochemistry 54.32 (2015): 4969-4977.
40. 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.
41. Thornburg ZR., et al. "Fundamental Behaviors Emerge from Simulations of a Living Minimal Cell". Cell 185.2 (2022): 345-360. e28.
42. 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.
43. 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.
44. 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.
45. Babonis LS., et al. "Evo-Inspired Engineering of Radical Phenotypes and Emergent Traits". Integrative And Comparative Biology 65.2 (2025): 198-215.