



NATIONAL TECHNICAL UNIVERSITY OF ATHENS  
SCHOOL OF CHEMICAL ENGINEERING

# Analysis of steady-state and dynamic properties of human ovarian cancer cell metabolism

A diploma thesis of  
**Ilias Toumpe**

Supervisors  
**Prof. Vassily Hatzimanikatis**  
École Polytechnique Fédérale de Lausanne  
**Prof. Boudouvis Andreas**  
National Technical University of Athens

Athens, September 2022

# Acknowledgments

The present thesis was developed in collaboration with the Laboratory of Computational Systems Biotechnology, EPFL, under the supervision of Professor Vassily Hatzimanikatis and his lab members.

During my stay at EPFL, the computational resources, research guidance, and academic conversations were of utmost significance in successfully completing this thesis and providing a mature closure to my undergraduate years. Therefore, I express my deepest gratitude to Professor Vassily Hatzimanikatis for giving me this opportunity and helping me uncover and interact with the fascinating fields of metabolic engineering and systems biotechnology.

This work would not have been possible without the additional support of Professor Andreas Boudouvis, professor at the School of Chemical Engineering NTUA, who was a co-supervisor on behalf of NTUA. He has been an inspiring academic figure throughout my undergraduate studies and I would like to thank him for his assistance and trust all these years.

Furthermore, I am immensely grateful for the constant encouragement and guidance of Lecturer Ljubisa Miskovic. His support allowed me to explore new research topics that became integral parts of the finalized thesis and he taught me how to approach the challenges of the project strategically.

Moreover, the presented work would not have been the same without the time and efforts of Dr. Maria Masid Barcon and Bharath Narayanan. They helped me with educational pieces of training and careful strategizing and equipped me with critical thinking that proved necessary for the completion of the different parts of this work.

I would also like to thank Zeno Karl Schindler Foundation for providing me with the Master Thesis Grant, which funded my stay in Switzerland. The contribution to the project was pivotal and significantly elevated the quality of the thesis.

I am also very thankful to my friends from Heraklion, Athens and my extended iGEM family for being part of my professional and personal life. With their help, I found the drive to pursue many goals outside my comfort zone and filled my undergraduate years with many great memories.

Lastly, I would like to acknowledge the immense support of my family. They have given me the means to work on my childhood dreams, research and science, and I thank them for always being a pillar of encouragement in my everyday life. This work and all that is to come, I owe it to you.

Good thing I discovered how to code in parallel!



# Abstract

This thesis aimed to build dynamic and steady-state models that describe the physiology of human ovarian cancer cells. A reduced constraint-based metabolic model derived from the Recon3D GEM was used along with thermodynamic, transcriptomic, metabolomic, and fluxomic data to examine the steady-state properties of ovarian cancer metabolism. To build the kinetic models, the directionalities of all reactions should be predefined. While the integrated data and other manually added constraints significantly reduced the number of bidirectional reactions, many remained. A modified ACHR sampling algorithm was used to sample feasible steady-state profiles and the average sample point was used to extract directionalities for the remaining bidirectional reactions. Further thermodynamic curations and modal analysis implementation resulted in the production of steady sets of kinetic parameters that could also operate in physiologically relevant timeframes. Possible drug targets were proposed through Metabolic Control Analysis and further examined through simulations of enzyme perturbations. Enzymes Hexokinase (HEX1) and ATPM (ATP maintenance requirement reaction) significantly affected the cell's survival, and Long-chain-fatty-acid—CoA ligase (FACOAL1821) showcased similar results, although to a lesser extent. Lastly, the model responses uncovered reaction r0474, which represents one of the many catalyzed reactions of the enzyme Ribonucleoside diphosphate reductase, as a promising drug target for mitigating cancer proliferation. All these enzymes are connected with deregulated metabolic pathways in cancer cells and have been suggested as interesting points for treatment development. Overall, the methods used and developed allowed for the production of the first large-scale kinetic model of ovarian cancer. Using this model, it is possible to identify drug targets for cancer elimination, perform simulations of dynamic responses of the cancer cell's metabolic network as well as quantify the effects of candidate drug targets on cancer proliferation. Using such models can assist the study of specific types of cancer cells' physiology and accelerate the drug development process. Physiology-specific drug targets can be suggested, and early estimations of the toxicity response and effective drug dosages can be calculated.

*DISCLAIMER: Ovarian cancer can occur in all people with internal reproductive organs. However, the scientific literature has not made substantial efforts to separate the cases regarding the patients' gender status (e.g., assigned female at birth, trans people). This thesis will use inclusive terminology to present statistics and facts indicated in the literature review, as long as it does not convey misinformation.*

# Περιγραφή

Ο καρκίνος αποτελεί την κύρια αιτία θανάτου παγκοσμίως, αντιπροσωπεύοντας σχεδόν 10 εκατομμύρια θανάτους το 2020. Υπολογίζεται ότι τα επόμενα χρόνια, ο αριθμός των περιπτώσεων καρκίνου θα αυξηθεί κατά 60%. Ο όρος «καρκίνος» περιγράφει μια ομάδα ασθενειών που προκαλούν τα κύτταρα σε οποιοδήποτε μέρος του σώματος να υποστούν αλλαγές και να αναπτύσσονται ανεξέλεγκτα. Οι περισσότεροι τύποι καρκινικών κυττάρων δημιουργούν όγκους, οι οποίοι μπορούν να εισβάλουν στους υγιείς ιστούς και να προκαλέσουν βλάβη. Επιπλέον, τα καρκινικά κύτταρα έχουν την δυνατότητα να αποσπώνται από την αρχική περιοχή ανάπτυξης καρκίνου και να μετακινούνται σε άλλα μέρη του σώματος όπου μπορούν να συνεχίσουν να εξαπλώνονται. Αυτό το φαινόμενο ονομάζεται μετάσταση και οι εκτεταμένες μεταστάσεις είναι η κύρια αιτία θανάτου από καρκίνο.

Ο καρκίνος των ωοθηκών είναι η πιο θανατηφόρα μορφή γυναικολογικών καρκίνων. Η εμφάνιση καρκίνου των ωοθηκών δεν είναι πολύ συχνή, αλλά συνδέεται με κακή ιατρική πρόγνωση. Κατά μέσο όρο, μόνο το 40% των ασθενών επιβιώνει πέντε χρόνια μετά τη διάγνωση. Το φαινόμενο αυτό οφείλεται στην αργοπορημένη διάγνωση της νόσου για πολλούς ασθενείς, με αποτέλεσμα τα ποσοστά θνησιμότητας να είναι αυξημένα.

Πολλές χημειοθεραπείες και φαρμακευτικές αγωγές έχουν προταθεί για την καταπολέμηση του καρκίνου και μερικές είναι εξειδικευμένες για την καταπολέμηση του καρκίνου των ωοθηκών. Ωστόσο η διαδικασία ανάπτυξης φαρμάκων είναι μακροχρόνια και κοστοβόρα με μερικές εκτιμήσεις να κυμαίνονται στα 1.4 με 2.5 δισεκατομμύρια δολάρια. Για την μείωση του κόστους αλλά και του χρόνου ανάπτυξης, έχουν αναπτυχθεί μια πληθώρα από υπολογιστικές μεθόδους που στοχεύουν στην ανακάλυψη υποσχόμενων φαρμακευτικών στόχων και αρχικών φαρμακευτικών ουσιών.

Ωστόσο οι τωρινές μέθοδοι που χρησιμοποιούνται για την πρόταση πιθανών φαρμακευτικών στόχων για τον καρκίνο παρέχουν λίγες πληροφορίες που περιορίζονται στην εύρεση γονιδίων ενδιαφέροντος (που έχουν συνήθως υποστεί κάποια απορρύθμιση) καθώς και την επίδραση που θα έχει η αφαίρεση τους στην μόνιμη κατάσταση της αναπτυσσόμενης καρκινικής βιομάζας.

Η ανάπτυξη κινητικών μοντέλων που περιγράφουν το μεταβολισμό του εξεταζόμενου κυττάρου μπορεί να παρέχει μια πληθώρα επιπρόσθετων πληροφοριών. Οι δυναμικές προσομοιώσεις του μεταβολικού δικτύου μπορούν να αποκαλύψουν στοιχεία σχετικά με την χρονική κλίμακα του ανασταλτικού φαινομένου ενός φαρμάκου, καθώς και πιθανές τοξικολογικές αποκρίσεις που μπορεί να συμβούν πριν από τον μετριασμό της κυτταρικής ανάπτυξης. Επιπλέον, χρησιμοποιώντας κατάλληλες μεθόδους, τα κινητικά μοντέλα μπορούν να εμπλουτιστούν με multiomics πειραματικά δεδομένα προσφέροντας έτσι εξειδικευμένες πληροφορίες για την φαινοτυπική απόκριση του προσομοιωμένου κυττάρου.

Η παρούσα διπλωματική εργασία στοχεύει να συμβάλει στα αρχικά στάδια ανάπτυξης και έρευνας φαρμάκων για τον καρκίνο των ωοθηκών του ανθρώπου, με την ανάλυση μόνιμης κατάστασης και δυναμικών ιδιοτήτων του μεταβολισμού των καρκινικών κυττάρων ωοθήκης. Για τον σκοπό αυτό, κατασκευάστηκε ένα προσαρμοσμένο μεταβολικό μοντέλο, που βασίστηκε στο ανθρώπινο

μεταβολικό μοντέλο μεγάλης κλίμακας Recon3D. Το μοντέλο μόνιμης κατάστασης δημιουργήθηκε με την χρήση του redHUMAN workflow, που στοχεύει στην χρήση των αντιδράσεων που συμμετέχουν στα μεταβολικά υποσυστήματα ενδιαφέροντος. Το μοντέλο περιέχει 1724 αντιδράσεις και 622 μεταβολικές ενώσεις. Για την καλύτερη αντιπροσώπευση της φυσιολογίας των καρκινικών κυττάρων της ωοθήκης, έγινε κατάλληλη ενσωμάτωση πειραματικών δεδομένων και άλλων μαθηματικών περιορισμών. Συγκεκριμένα μέσω της μεθόδου TFA το μοντέλο διαμορφώθηκε ώστε να περιλαμβάνει πληροφορίες για την ενέργεια Gibbs των αντιδράσεων και τις συγκεντρώσεις των μεταβολιτών σε σταθερή κατάσταση. Επιπλέον μέσω της μεθόδου REMI έγινε προσθήκη transcriptomic δεδομένων που εξήχθησαν από Wild Type (WT) και BRCA1 μεταλλαγμένα καρκινικά κύτταρα ωοθηκών. Τέλος, έγινε προσθήκη metabolomic και fluxomic δεδομένων καθώς και περιορισμών ελάχιστης ροής και ελαχιστοποίησης του αθροίσματος των ροών. Όλες οι παραπάνω τροποποιήσεις στο μοντέλο μείωσαν σημαντικά τον επιτρεπτό χώρο λύσεων και μετέτρεψαν πολλές αμφίδρομες αντιδράσεις σε μονόδρομες. Ωστόσο, για την κατασκευή των κινητικών μοντέλων, οι κατευθύνσεις όλων των αντιδράσεων θα πρέπει να είναι προκαθορισμένες. Για τον σκοπό αυτό δημιουργήθηκε μια τροποποιημένη εκδοχή του ACHR αλγορίθμου, ενός αλγορίθμου που χρησιμοποιεί Monte Carlo random walks για τη δειγματοληψία ασυμπτωτικά κατανομμένων σημείων στο επιτρεπτό χώρο λύσεων του μοντέλου. Τα δείγματα ροών μόνιμης κατάστασης χρησιμοποιήθηκαν για τον υπολογισμό ενός μέσου δείγματος, από το οποίο εξήχθησαν οι κατευθύνσεις των αντιδράσεων που απέμειναν.

Με βάση αυτό το τροποποιημένο μοντέλο, έγινε η παραγωγή κινητικών μοντέλων του προσομοιωμένου μεταβολικού δικτύου. Για την δημιουργία αυτών, ορίστηκε αρχικά το είδος ρυθμού αντίδρασης για κάθε αντίδραση με βάση την στοιχειομετρία της. Στην συνέχεια κάνοντας χρήση δειγμάτων μόνιμης κατάστασης του μοντέλου καθώς και της μεθόδου ORACLE, επιλέχθηκαν τιμές κινητικών παραμέτρων που είναι σύμφωνες με τα παραπάνω δείγματα μόνιμης κατάστασης. Με αυτό τον τρόπο μπορούν να παραχθούν μια πληθώρα από κινητικά μοντέλα με διαφορετικές τιμές στις κινητικές παραμέτρους τους, αλλά η ευστάθεια τους δεν είναι σίγουρη. Κάνοντας χρήση της Ιακωβιανής μήτρας του διαφορικού συστήματος, τα μοντέλα μπορούν να εξεταστούν ως προς την ευστάθεια τους γύρω από το σημείο σταθερής κατάστασης με το οποίο κατασκευάστηκαν (το πραγματικό μέρος όλων των ιδιοτιμών πρέπει να είναι αρνητικό). Επιπλέον, τα μοντέλα επιλέχθηκαν έτσι ώστε να λειτουργούν σε χρονικά πλαίσια με βιολογική σημασία, δηλαδή γρηγορότερα από τον χρόνο διπλασιασμού του κυττάρου. Αυτό επιτεύχθηκε με τροποποιήσεις στις παραμέτρους του μοντέλου σταθερής κατάστασης έπειτα από την εξαγωγή πληροφοριών από τους τροπικούς τελεστές των μοντέλων.

Το έργο της διπλωματικής εργασίας ολοκληρώνεται με την χρήση μεθόδων για την εύρεση πιθανών φαρμακευτικών στόχων στο μεταβολισμό του καρκίνου των ωοθηκών. Συγκεκριμένα, χρησιμοποιήθηκε η τεχνική MCA, η οποία μπορεί να παράγει εκτιμήσεις για τις αλλαγές σε συγκεντρώσεις και ροές που προκαλούνται από διαταραχές στις κινητικές παραμέτρους. Με αυτό τον τρόπο προτάθηκαν ένζυμα και αντιδράσεις του μεταβολικού δικτύου που αποτελούν υποσχόμενους φαρμακευτικούς στόχους. Για την ποσοτική εξέταση της επίδρασης τους πραγματοποιήθηκαν προσομοιώσεις ενζυμικών αναστολών με την εφαρμογή διαταραχών στην τιμή του μέγιστου ρυθμού αντίδρασης ( $v_{max}$ ) και την ολοκλήρωση του διαφορικού συστήματος έως την νέα μόνιμη κατάσταση. Η επιλογή της παραμέτρου  $v_{max}$  βασίζεται στο γεγονός ότι η τιμή της είναι γραμμικά ανάλογη της συγκέντρωσης του συνολικά διαθέσιμου ενζύμου και έτσι μπορεί να γίνει άμεσα εμφανής η επίδραση στην παραγωγή της βιομάζας.

Τα ένζυμα που εξετάστηκαν εκτενώς ήταν η εξοκινάση HEX1, ATPM (αντίδραση συντήρησης ATP), λιγάση μακράς αλυσίδας λιπαρού οξέος - CoA FACOAL1821 και το r0474 το οποίο αντιπροσωπεύει μία από τις πολλές καταλυόμενες αντιδράσεις του ενζύμου διφωσφορική αναγωγή ριβονουκλεοζιτών. Τα HEX1 και ATPM επηρέασαν σημαντικά την επιβίωση του κυττάρου και το FACOAL1821 παρουσίασε παρόμοια αποτελέσματα, αν και σε μικρότερο βαθμό. Τέλος, οι αποκρίσεις των μοντέλων αποκάλυψαν ότι το ένζυμο r0474 είχε υποσχόμενα αποτελέσματα, ως φαρμακευτικός στόχος για τον μετριασμό του πολλαπλασιασμού του καρκίνου. Όλα τα ένζυμα συνδέθηκαν και με την βιβλιογραφία καθώς ανήκουν σε μεταβολικά μονοπάτια τα οποία έχουν καταγραφεί ως απορυθμισμένα, κατά την εξέταση καρκινικών κυττάρων, και για αυτό τον λόγο έχουν προταθεί ως ενδιαφέροντα σημεία στοχευμένης θεραπευτικής αγωγής.

Συνολικά, οι μέθοδοι που χρησιμοποιήθηκαν και αναπτύχθηκαν επέτρεψαν την παραγωγή του πρώτου, μεγάλης κλίμακας κινητικού μοντέλου για τον καρκίνο των ωοθηκών. Χρησιμοποιώντας αυτό το μοντέλο, είναι δυνατός ο εντοπισμός φαρμακευτικών στόχων για την εξάλειψη του καρκίνου, η εκτέλεση προσομοιώσεων δυναμικών αποκρίσεων του μεταβολικού δικτύου του καρκινικού κυττάρου καθώς και η ποσοτικοποίηση των επιπτώσεων των υποψήφιων φαρμακευτικών στόχων στον πολλαπλασιασμό του. Η χρήση τέτοιων μοντέλων μπορεί να βοηθήσει στη μελέτη συγκεκριμένων τύπων φυσιολογίας καρκινικών κυττάρων και να επιταχύνει τη διαδικασία ανάπτυξης φαρμάκων. Μπορούν επίσης, να προταθούν εξειδικευμένοι φαρμακευτικοί στόχοι για τη φυσιολογία των μελετώμενων καρκινικών κυττάρων και μπορούν να γίνουν πρώιμες εκτιμήσεις της τοξικής απόκρισης του κυττάρου αλλά και των αποτελεσματικών δόσεων του φαρμάκου.

# Table of contents

<b>1. INTRODUCTION.....</b>	<b>1</b>
1.1 CANCER.....	1
1.2 OVARIAN CANCER.....	1
1.3 RISK FACTORS.....	2
1.4 CURRENT DRUGS USED IN CANCER AND OVARIAN CANCER .....	4
1.5 DRUG DEVELOPMENT CYCLE.....	5
1.6 <i>IN SILICO</i> METHODS .....	6
1.7 THE GOAL OF THIS THESIS .....	7
<b>2. METHODS .....</b>	<b>8</b>
2.1 STEADY-STATE METABOLIC MODELS .....	8
2.2 CONSTRAINT-BASED MODELING OF METABOLISM .....	9
2.2.1 <i>Flux Balance Analysis</i> .....	9
2.2.2 <i>Thermodynamics-based Flux Analysis</i> .....	10
2.2.3 <i>Basal Fluxes constraint</i> .....	10
2.2.4 <i>Minimization of the sum of fluxes</i> .....	11
2.2.5 <i>REMI</i> .....	11
2.2.6 <i>Variability analysis</i> .....	12
2.3 SAMPLING PROCESS .....	12
2.3.1 <i>ACHR sampling</i> .....	12
2.3.2 <i>Avoiding basal bounds</i> .....	13
2.4 KINETIC MODELS .....	14
2.4.1 <i>Building kinetic models</i> .....	15
2.4.2 <i>Pruning kinetic models</i> .....	15
2.5 MCA AND ODE SIMULATIONS.....	16
2.6 CODE AVAILABILITY .....	17
<b>3. RESULTS .....</b>	<b>18</b>
3.1 BUILDING THE MODEL OF OVARIAN CANCER CELLS .....	18
3.1.1 <i>Model curation</i> .....	18
3.1.2 <i>Kinetic model building</i> .....	21
3.2 DRUG TARGET ANALYSIS .....	23
3.2.1 <i>Drug targets derived from MCA</i> .....	23
3.2.2 <i>Simulations of enzyme perturbations</i> .....	25
<b>4. CONCLUSIONS .....</b>	<b>32</b>
<b>5. BIBLIOGRAPHY .....</b>	<b>34</b>
<b>6. APPENDIX A .....</b>	<b>38</b>
<b>7. APPENDIX B .....</b>	<b>39</b>

# Nomenclature

<b>ACHR</b>	Artificial Centered Hit-and-Run
<b>ADP</b>	Adenosine Diphosphate
<b>AMP</b>	Adenosine Monophosphate
<b>ATP</b>	Adenosine Triphosphate
<b>BDR</b>	Bidirectional Reactions
<b>FBA</b>	Flux Balance Analysis
<b>GCM</b>	Group Contribution Method
<b>GEM</b>	Genome scale metabolic Network
<b>GPR</b>	Gene-Protein-Reaction
<b>LP</b>	Linear Programming
<b>MCA</b>	Metabolic Control Analysis
<b>MILP</b>	Mixed Integer Linear Programming
<b>ORACLE</b>	Optimization and Risk Analysis of Complex Living Entities
<b>OV, OVCA</b>	Ovarian Cancer
<b>PPA</b>	Pentose Phosphate Pathway
<b>REMI</b>	Relative Expression and Metabolomic Integrations
<b>TFA</b>	Thermodynamics-based Flux Analysis
<b>TCA</b>	Tricarboxylic Acid cycle

---

$[X_j]$	Concentration variable of compound j
$N$	Stoichiometric matrix
$v_i$	Flux variable of reaction i
$\Delta_r G_i$	Gibbs free energy of reaction i
$\Delta_r G_i^\circ$	Standard Gibbs free energy of reaction i
$R$	Ideal gas constant
$K_{eq}$	Equilibrium constant
$K_m$	Michaelis-Menten constant
$J$	Jacobian matrix
$\lambda_i$	eigenvalue of Jacobian matrix



# 1. Introduction

## 1.1 Cancer

Cancer is a leading cause of death worldwide, accounting for nearly 10 million deaths in 2020 (approximately one in six deaths). It is estimated that the number of cancer cases will grow by 60% in the years to come<sup>1</sup>. The term “cancer” describes a group of diseases that cause cells in any part of the body to change and grow out of control. Most types of cancer cells create a lump called a tumor (more specifically, malignant tumor, since not all tumors are cancer). Cancerous tumors can invade healthy tissues and cause damage. It is also common for cancer cells to detach from the main cancer site and travel to other parts of the body where they can continue spreading. This phenomenon is called metastasis, and widespread metastases are the primary cause of death from cancer<sup>1</sup>.

As the scientific community strives to tackle this medical problem, it is evident that every type of cancer poses unique challenges and should be treated as a separate disease.

## 1.2 Ovarian cancer

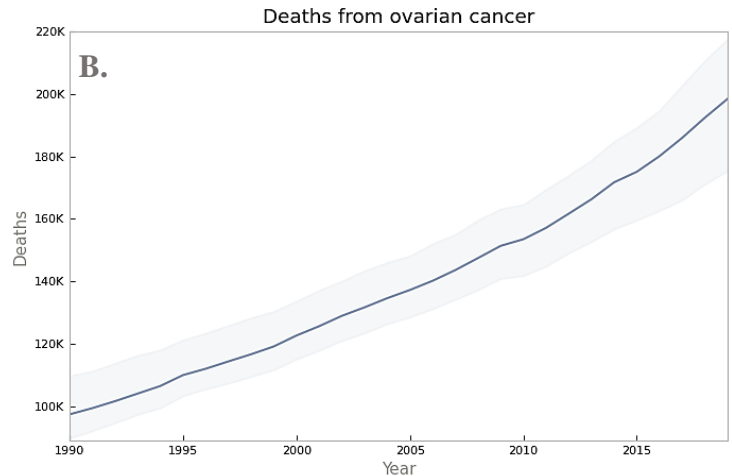
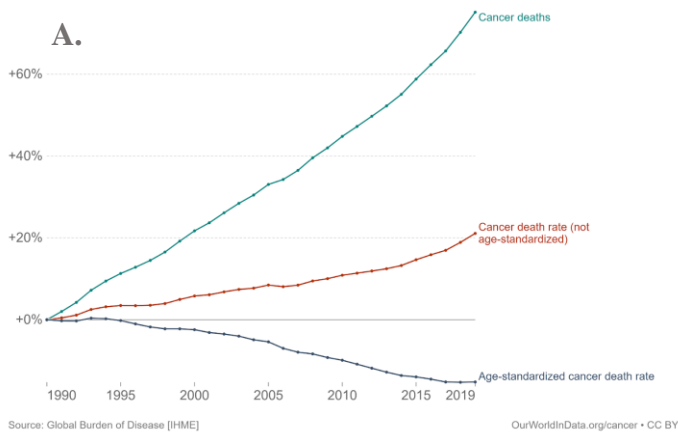
Ovarian cancer (*OC or OVCA*) is the most lethal form of gynecological cancers<sup>2</sup>, with the number of associated deaths steadily increasing over the years (Figure 1.1). The occurrence of ovarian cancer is not very common, but it is linked to poor prognosis. On average, only 40% of patients survive five years after the diagnosis, when taking into account all types and stages of the disease. This phenomenon is observed because most patients are diagnosed when the disease is in its late stages, and as a result, the mortality rates are elevated<sup>3</sup>.

The most common origin of OC is from the fallopian tube, where Serous Tubal Intraepithelial Carcinomas (STICs) shed onto the ovary surface and establish a tumor<sup>4</sup>. Past reports indicated that malignant ovarian cancer tumors originate from metastatic activities of primary tumors from the colon, stomach, pancreas and uterus. However, newer studies claim that they do indeed begin in the ovary<sup>4</sup>.

The heterogenous microenvironment of an OC tumor contains tumor and non-tumor cells, including immune cells. Since no anatomical barriers limit the ovaries, OC cells can detach from the primary tumor and induce metastatic effects in the organs in the pelvic cavity, such as the uterus, bladder, rectum and small intestine, and even in other organs like the liver and lung<sup>4</sup>.

The main types of ovarian cancer are epithelial, which account for around 90% of all malignant OC, and the less common stromal OC. Further classification can be based on the different histological subtypes. The differences arise from various sites of origin and specific responses to tumor suppressors, chemotherapeutic treatments, and molecular target drugs. Malignant histological subtypes arising from epithelial cells include high-grade serous ovarian cancer (HGSOC), ovarian clear cell carcinoma (OCCC), endometrioid ovarian cancer (EnOC), mucinous ovarian cancer (MOC), low-grade serous ovarian cancer (LGSOC), and malignant Brenner cell tumors<sup>4</sup>.

Change in three measures of cancer mortality, World, 1990 to 2019  
This chart compares cancer deaths, the cancer death rate, and the age-standardized death rate.



**Figure 1.1** (A) Changes in cancer deaths cases and cancer-death rates worldwide and (B) estimated deaths from ovarian cancer from 1990 to 2019 <sup>5,6</sup>.

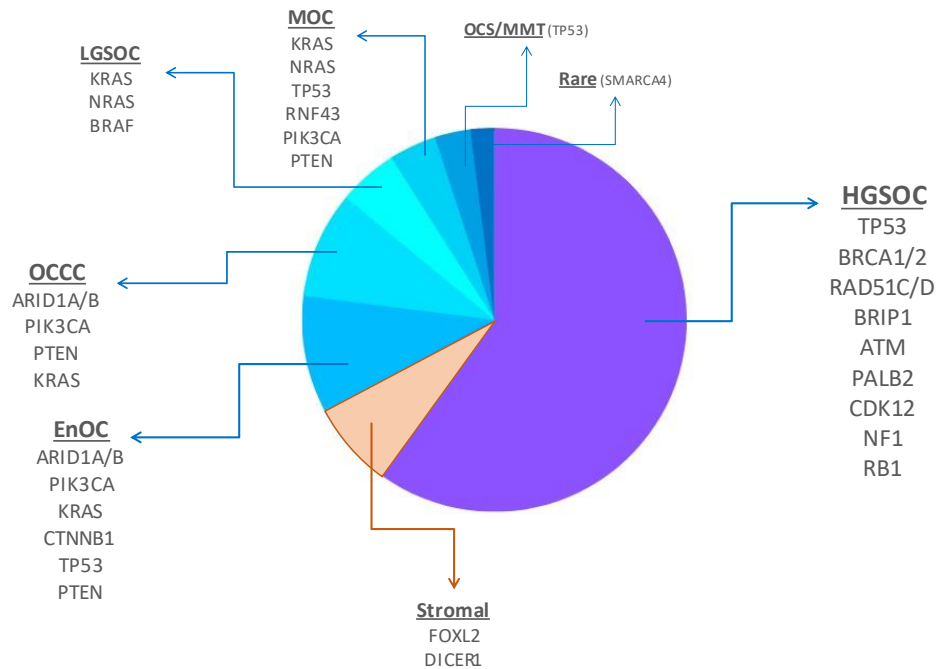
Ovarian carcinosarcomas (OCS), also known as malignant mixed mullerian tumors (MMMT), have epithelial and mesenchymal components. Lastly, Ovarian sex-cord stromal tumors (SCST), the most common of which are granulosa cell tumors (GCT), and the rarer Sertoli–Leydig cell tumors originate from stromal cells<sup>4</sup>.

Common gene mutations leading to development of OC are observed on the tumor suppressor TP53 gene and the DNA repairing BRCA1/BRCA2 genes<sup>4</sup>. Additional mutated genes are associated with RAS/MAPK signaling pathways linked with cell growth and division. The relative percentages of the subtypes with their respective gene mutations are illustrated in Figure 1.2 <sup>4</sup>.

The notion of ovarian cancer has serious psychological effects on patients at all phases of diagnosis, treatment and possible recurrence. Specifically, potential patients with a higher risk of developing OC are more likely to have higher rates of anxiety and depression during the screening processes. If a diagnosis occurs, these symptoms are raised, at which point psychological interventions may be required. Patients have reported intense anxiety and depression, physical discomfort, and functional disability during the treatment period. These effects may be due to the reduced levels of estrogen and androgens that occur as a result of surgeries and the stressors induced on the body by chemotherapeutic treatments. Many psychological symptoms remain even after post-treatment survival periods and are accompanied by anxiety of family members getting the disease or the chance of recurrence<sup>7</sup>.

### 1.3 Risk factors

In general, people with certain genetic dispositions and habitual activities are more likely to develop cancer. These characteristics are called risk factors and vary based on the type of cancer. Research has shown that epithelial ovarian cancer is associated with many risk factors listed below<sup>8</sup>:



**Figure 1.2** Types of ovarian cancer histotypes and their known gene mutations. 90% of the diagnosed ovarian cancers are in one of the seven epithelial histotypes, the most common being HGSOE. Stromal ovarian cancer originates from stromal cells and showcases different properties but is less common<sup>4</sup>.

- Age is an important factor in developing OC, as half of all cases are found in patients 63 years of age or older.
- Obesity is a potential risk factor, as it has been linked to a higher chance of developing cancer. However, there is still no clear correlation with ovarian cancer.
- Anyone who had not had a full-term pregnancy before age 35 or never carried a pregnancy to term is at a higher risk of developing OC.
- Hormone therapy after menopause is linked with an increased risk of having OC.
- Having a family history of ovarian, breast or collateral cancer can also increase the chances of OC. Moreover, specific gene mutations, such as on BRCA1 and BRCA2 genes, can either be inherited or induced by environmental factors.
- Similarly, if the patient used to have breast cancer, the risk of developing OC is higher.

On the flip side, there are documented factors that are possibly linked with a smaller chance of developing OC. These include:

- Having a full-term pregnancy at an age before 26.
- Having multiple full-term pregnancies.
- Breastfeeding.
- Using birth control pills or other forms of birth control.

## 1.4 Current drugs used in cancer and ovarian cancer

The chemotherapeutic treatments are designed around specific weaknesses or metabolic features of the tumor cells. These cellular characteristics might be common in all types of cancer, which leads to the design of “general-purpose” drugs, or they can be specific to the type of cancer and the patient's genetic profile, leading to a more customized chemotherapy. Most developed drugs inhibit a specific signaling pathway, which is pivotal for the tumor's survival and growth<sup>2</sup>.

### Cisplatin/paclitaxel combination

One of the earliest chemotherapeutic medication for OC is the cisplatin/paclitaxel combination. Cisplatin is able to bind to DNA strands and prohibit the repair of damaged DNA. This effect impacts all types of cells but is more detrimental to the quickly proliferating cancer cells. Similarly, paclitaxel targets tubulin and effectively does not allow the formation of spindles, a vital factor in the cell division stage<sup>9,10</sup>.

### Antiangiogenic therapies

Angiogenesis is the process by which the organism forms new blood vessels from pre-existing vessels. It is an integral part of tumor growth as it ensures that the cancer cells have adequate levels of nutrients and oxygen. Angiogenesis is crucial for the persistence of epithelial OC, and many antiangiogenic therapies have been proposed. Most drugs target the growth factor (VEGF), a signaling protein for promoting new blood cells. Examples include bevacizumab, aflibercept, nintedanib, pazopanib, sunitinib, sorafenib, and cediranib<sup>2</sup>. These drugs can also target other enzymes, which further inhibits the production of new blood cells. Different studies have examined the effects of these targets in different types of OC with mixed outcomes.

### Poly (adenosine diphosphate [ADP]-ribose) polymerase inhibitors

The PARP enzymes coordinate the base-excision repair pathways and are essential to the cellular DNA repair system. Inhibiting the PARP enzymes results in many single-strand breaks, translating into double-strand breaks when the cell replicates. Such mistakes are taken care of by homologous recombination with the help of BRCA proteins. However, as stated before, many patients with OC have a genetic mutation in BRCA1/2 genes and, as a result, display targeted sensitivity to treatment with PARP inhibitors. This phenomenon of two otherwise nonlethal changes resulting in tumor cell death is called synthetic lethality and is extremely useful for providing targeted therapy. Drugs used in clinical trials include olaparib, veliparib, rucaparib, and niraparib.

### Other drug targets

The ERBB/HER-family of receptor tyrosine kinases play a key role in cell growth and survival. The dysregulation of associated enzymes like EGFR, HER2, HER3, and HER4 can affect tumor growth significantly; however, recent clinical trials had discouraging results<sup>2</sup>.

Folate receptors have been recorded to be highly over-expressed in non-mucinous OC. New clinical studies are trying to target those receptors like Folate Receptor-A, with encouraging results.

The activation of the insulin-like growth factor (IGF) system initiates a surge of downstream signaling events, which facilitate tumorigenesis by promoting cellular proliferation, angiogenesis, invasion, and metastatic potential and inhibiting apoptosis. Recent studies are trying to inhibit this signaling pathway, and new data are pending.

Other drugs are developed for specific types of OC targeting the gene mutations commonly associated with them, like KRAS, BRAF and PTP53 genes<sup>2</sup>.

## 1.5 Drug development cycle

The process of developing a drug and getting it into the market contains many challenges. On average, it is estimated that for every 10,000 structures synthesized during drug design, 500 will reach animal testing, ten will reach phase I clinical trials, and only one will reach the marketplace. Additionally, over the past decades, the costs linked to a new drug's research, development and marketing approval have been steadily increasing<sup>11</sup>.

Based on 106 randomly selected new drugs from a survey of ten pharmaceutical firms, it is estimated that the average out-of-pocket cost per approved new compound is almost \$1.4 billion (2013 dollars). Capitalizing out-of-pocket costs to the point of marketing approval at a real discount rate of 10.5% yields a total pre-approval cost estimate of over \$2.5 billion (2013 dollars)<sup>12</sup>. The general process followed during the development and distribution of the drug is as follows<sup>11</sup>:

### Drug discovery and optimization

Based on a disease's metabolic and genetic characteristics, suitable drug targets are picked for which lead compounds are produced. The drug targets can be enzymes, proteins and nucleic acids, and the lead compounds might come from natural products, existing drugs, and chemically similar compounds to the enzyme's substrate. There are many ways to discover suitable drug targets and lead compounds, from *in vitro* and *in vivo* assays to computational screening methods. An iterative optimization phase follows the discovery of a promising drug candidate to enhance some of the compounds' characteristics, such as the activity and specificity, minimization of side effects and ADME (*absorption, distribution, metabolism, and excretion*) properties.

### Preclinical and Clinical trials

Once the drug has been optimized, it has to go through rigorous preclinical and clinical trials that are significantly more expensive than the previous steps. Preclinical trials involve animal testing to estimate the drug's effective dose (ED) and lethal dose (LD), pharmacological properties, and form of production.

Clinical trials take 5-7 years to complete with hundreds of volunteer patients and are the bulk cost of the whole procedure. They are carried out in 4 phases which are not necessarily successive, but they can often run in parallel and change their goals based on the general results. The involved volunteers are both healthy patients and patients with the targeted disease. The purpose of the different phases is to examine the drug's effectiveness and record any short-term and long-term side effects. In Phase IV, the final phase, the drug is released to the market, and everyone prescribed to take it can report its effectiveness and any unexpected side effects.

## 1.6 *In silico* methods

As discussed, the drug development market includes high monetary risks, so it is thus vital for a drug to have some first strong indications of therapeutic abilities. Computational tools can provide some first estimations about the effective drug targets of a particular disease and prune through databases of chemical compounds to find some early drug candidates. The use of such *in silico* methods helps in cutting down the costs by discarding strategies that would otherwise need experimental results to validate while at the same time being time-efficient due to their standardized procedures.

A method used to predict drug targets in cancer cells is to examine the specific metabolic networks constituting the physiology of the targeted cells. Possible ways of achieving this are by simulating different regulatory genes and their interactions<sup>13</sup> or simulating the active metabolic subsystems in the cell<sup>14</sup>.

Gene regulatory network inference is a method that explores the different regulatory connections of genes in a specific biological setting. The method can assess the strength of interactions through experimental data and uncover deregulation of critical biological processes that are not present in normal cells. In the context of cancer, pinpointing the critical genes in a gene regulatory network (e.g., the ones with most interactions) can produce an early estimate of drug targets that can exclusively target cancer cells<sup>13</sup>.

A similar method aims to detect the different metabolic alterations that lead to cancer-specific metabolism, which can help develop drugs that can reduce cell proliferation in cancer cells without affecting normal cells. These metabolic rewirings can be associated with mutated genes, and an outline of possible drug targets can be produced based on gene essentiality methods. Further validation can be yielded by simulating gene knockouts of the essential genes and examining how the metabolic system is affected<sup>14</sup>.

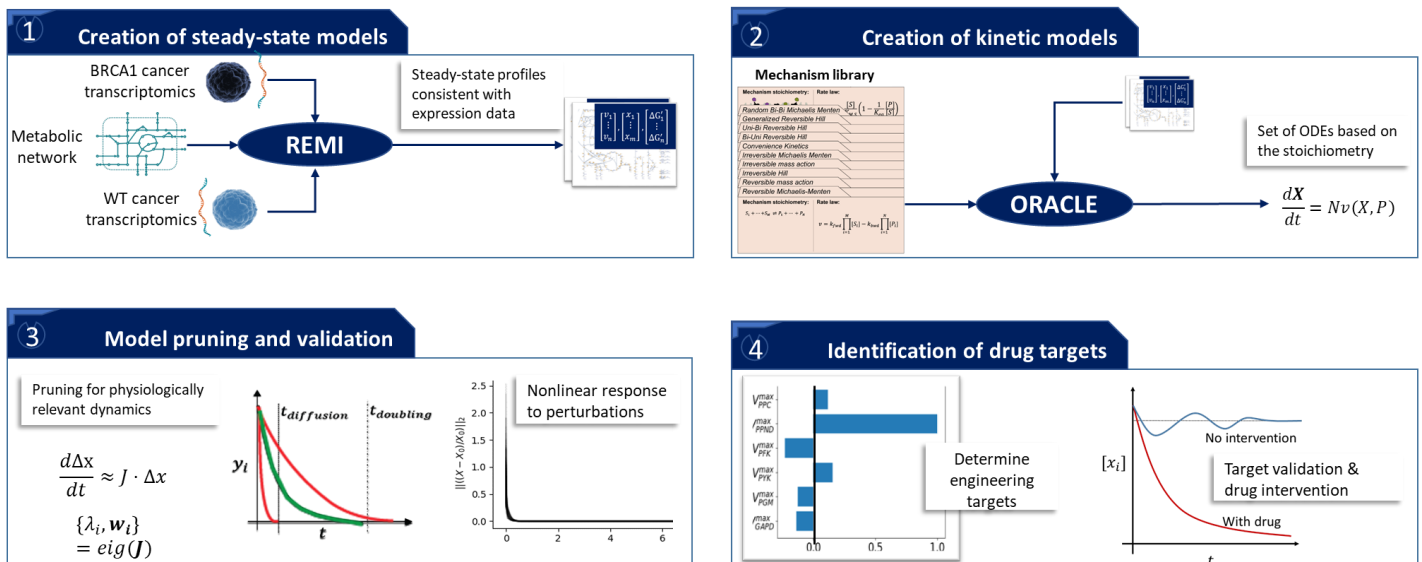
Such methods are accurate at estimating which genes have to be targeted to inhibit cancer spread. In addition, gene knockout simulations can provide an early estimation of the effect of drugs on the physiology of the cell (where essentially, the drug will inhibit a particular gene or its expressed enzyme/protein). However, gene regulatory networks can only provide binary information i.e., if a gene is a possible drug target or not. At the same time, gene essentiality methods only provide steady-state flux profiles, disregarding the dynamic responses that occur when the original perturbation (drug administration/gene knockout) is imposed.

Knowing the system's transient response can provide information about the time scale of the inhibitory effect, as well as possible toxicological responses that might happen before cell growth mitigation. Furthermore, dynamic simulations of the metabolic network can exhibit additional information about the phenotypic response to the proposed changes. The dynamic simulation is achievable through kinetic modeling based on a model enriched with a complete set of multiomics data.

## 1.7 The goal of this thesis

The present thesis aims to assist in the early stages of drug development and research, specifically for ovarian cancer, by creating steady-state and kinetic models of metabolic networks that can represent the physiology of OC cells. A graphical overview of the thesis is shown in Figure 1.3. Newly developed methods around the production of constraint-based models and kinetic analysis will be used to capture thermodynamic data and multiomics data of OC cells. By implementing omics data, steady-state samples of the metabolic network will be more representative of OC physiology. Still, due to the large scale of the developed constrained-based models, there is high uncertainty in reaction directionalities, which hinders the reconstruction of kinetic models. This uncertainty will be tackled by effectively sampling the solution space and deriving statistical information about the system's capabilities.

Additionally, the kinetic models will be built around steady-state samples to better capture dynamics consistent with the cell's physiology. More steps must be completed to ensure that the kinetic models used will operate in timescales that can be considered physiologically relevant. The final kinetic models should be able to reproduce the dynamic responses of the metabolic network when a perturbation is imposed. Ultimately, the model size and complexity will reveal the elaborate responses to the enzymatic inhibitions imposed by simulated drugs. Such research endeavors aim to explore the metabolic network of ovarian cancer cells and propose metabolic drug targets which differ from the signaling and antiangiogenic approaches currently used.



**Figure 1.3** Overview of the thesis: Integration of omics data and generation of steady-state metabolic models of two ovarian cancer phenotypes (1), Creation of kinetic models of the WT cancer phenotype (2), Model pruning and validation (3), and Identification of targets for therapeutic interventions (4).

## 2. Methods

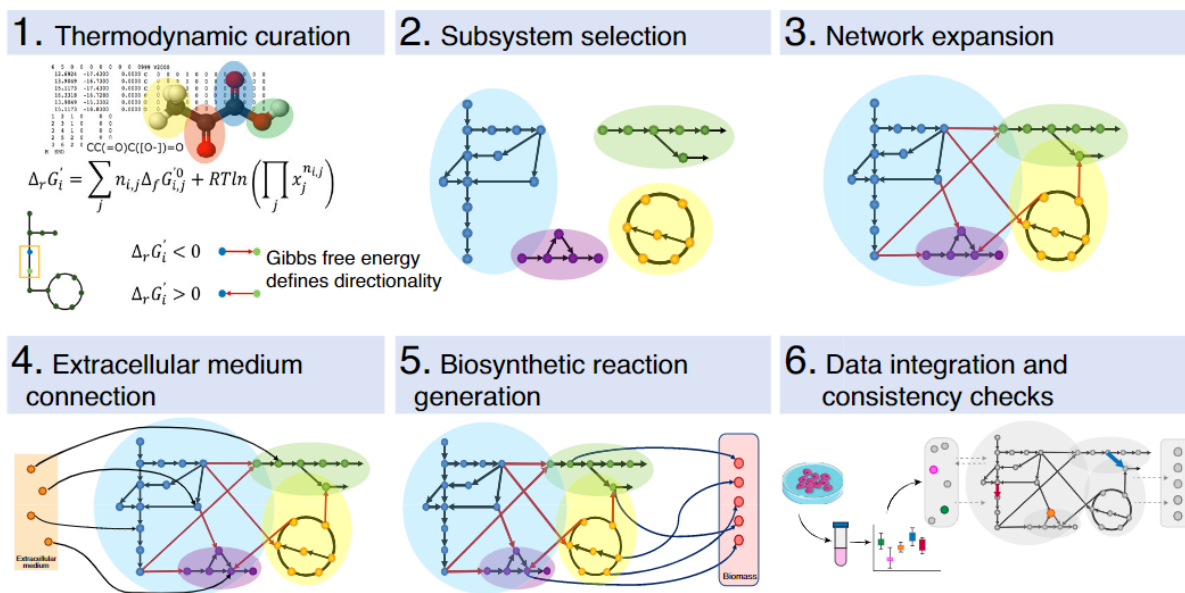
### 2.1 Steady-state metabolic models

Genome Scale Metabolic models, abbreviated as GEMs, are a stoichiometric representation of the metabolism of different organisms. They portray all known biochemical reactions taking place inside the modeled cells, which are based on functional annotation of known gene sequences and the corresponding gene-protein-reaction relationships (GPR) <sup>15</sup>.

An extensive catalog of GEMs has been developed for different types of cells like bacteria, yeast and plants. Many genome-scale models have also been developed for human metabolism, with the most recent one, Recon3D, containing 13,543 reactions involving 4,140 metabolites <sup>16</sup>.

For this thesis, a reduced version of the Recon3D GEM describing the physiology of ovarian cancer cells has been developed through the redHUMAN workflow <sup>17</sup>. The redHUMAN workflow begins with a thermodynamic curation of the directionality of reactions through estimated Gibbs free energies of compounds and reactions. Next, a set of subsystems relevant to the simulated metabolism are selected and are subsequently connected with a network expansion using reactions from the GEM. The produced core metabolic network is then connected with the extracellular medium components, and any essential pathways for the production of biomass building blocks are added. Lastly, the reduced model is enriched with omics data and is verified through consistency checks. Figure 2.1 presents the above workflow schematically.

The final reduced model for ovarian cancer contains 1724 reactions and 622 metabolites, and its characteristics are further analyzed in the *Results* section of the thesis.



**Figure 2.1** Schematic of the redHUMAN workflow<sup>17</sup>.

## 2.2 Constraint-based modeling of metabolism

### 2.2.1 Flux Balance Analysis

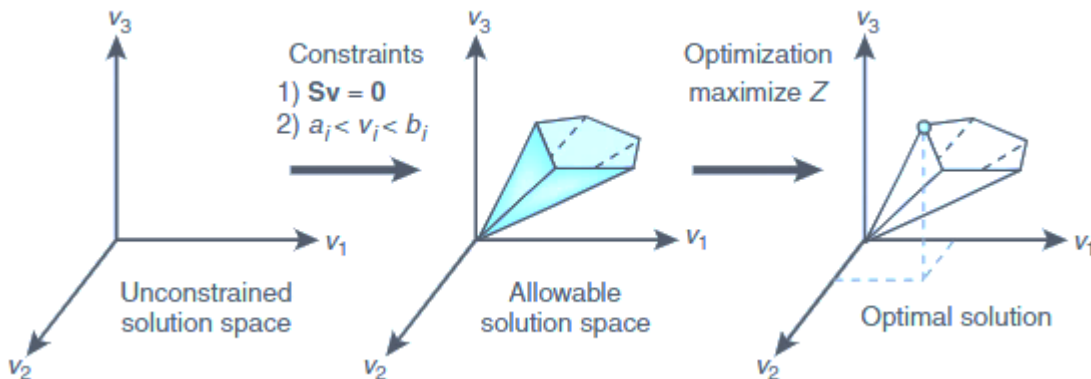
Many tools have been developed over the years that can simulate the metabolic network of cells with known reactions. The earliest constraint-based modeling method is Flux Balance Analysis (FBA) which uses linear programming to analyze flux distributions under the assumption of quasi-steady-state<sup>18</sup>. Based on this assumption, the net production rate of all metabolites in the model is zero and can be represented as:

$$\frac{dX}{dt} = N\mathbf{v} = 0 \quad (2.1)$$

where  $X$  is an  $m$ -length vector representing the metabolite concentrations in the model,  $N$  an  $m \times r$  stoichiometric matrix for the  $r$  reactions and  $m$  metabolites in the model, and  $\mathbf{v}$  is a  $r$ -length vector of the steady-state fluxes of the  $r$  reactions.  $N$  is a sparse matrix, as most metabolites do not participate in many reactions. Furthermore, there are practically always more reactions than metabolites meaning that the system described in 2.1 is always underdetermined. Even after integrating constraints that force the system to have physiologically relevant reaction fluxes, the allowable solution space is still vast. For that reason, it is common to assume an objective (optimization) function that describes the phenotype of the modeled cell:

$$Z = \mathbf{c}\mathbf{v} \quad (2.2)$$

where  $Z$  is the objective function and  $\mathbf{c}$  is  $1 \times r$  vector containing the weights for each reaction describing their contribution to  $Z$ . Examples of objective functions are the maximization of growth rate or the maximization of ATP or lactate production. With the use of an objective function, FBA can calculate an optimal solution (Fig 2.2).



**Figure 2.2** How different constraints in FBA lead to an optimal solution of the modeled metabolic network<sup>18</sup>.

## 2.2.2 Thermodynamics-based Flux Analysis

While FBA can assess the feasible flux distributions of the systems and can integrate fluxomic data from the modeled physiology of the cell, it fails to prove the thermodynamic feasibility of the solution space and lacks the ability to integrate metabolomic data. Thermodynamics-based Flux Analysis (TFA) builds on the FBA framework in order to address the above weaknesses<sup>19</sup>. This is done through a Mixed Integer Linear Programming (MILP) formulation that can calculate thermodynamically feasible fluxes and concentration profiles. The additional constraints are:

$$\text{(Flux coupling)} \quad v_i^F - z_i^F K \leq 0 \quad | \quad v_i^B - z_i^B K \leq 0 \quad | \quad v_i^F, v_i^B \geq 0 \quad (2.3)$$

$$\text{(Directionality coupling)} \quad z_i^F + z_i^B \leq 1 \quad (2.4)$$

$$\text{(Gibbs free energy of reaction)} \quad \Delta_r G_i = (\Delta_r G_i^\circ)_{pH,I,T,\Delta\psi} + RT \sum_{j=1}^m \ln [X_j] \quad (2.5)$$

$$\text{(Thermodynamic feasibility)} \quad \Delta_r G_i + z_i^F K - K \leq 0 \quad | \quad \Delta_r G_i + z_i^B K - K \leq 0 \quad (2.6)$$

Equations 2.3 split the net flux of a reaction into forward and backward variables ( $v_i^F$  and  $v_i^B$ ) and equation 2.4 ensures that only one of these variables is active. Here  $K$  is a parameter with an arbitrary high value (so that it will not constrain the upper bound of the fluxes' range). Equation 2.5 calculates the Gibbs free energy of the reaction based on the standard Gibbs free energy ( $(\Delta_r G_i^\circ)_{pH,I,T,\Delta\psi}$ ) and the concentrations ( $[X_j]$ ) of the metabolites participating in the reaction. Lastly, equation 2.6 couples the value of Gibbs free energy with the directionality of the reaction. This formulation allows the integration of metabolomic data as constraints, which outline the physiologically relevant ranges of the metabolite concentrations. Note that the concentration variable must be in logarithmic form ( $\ln [X_j]$ ) in order to respect the linear programming formulation.

The standard Gibbs free energy value is calculated from the standard Gibbs free energy of formation  $\Delta_f G_j^\circ$ , of the individual components of the reaction. These values are estimated by the Group Contribution Method (GCM)<sup>20</sup> and corrected for biophysical conditions in the cell such as ionic strength and pH value, membrane potential and temperature<sup>19</sup>.

Based on this new MILP formulation that allows for binary variables, several new constraints can be added based on assumptions about the cell's regulation and physiology and help constrain the system further.

## 2.2.3 Basal Fluxes constraint

No matter the cell's physiology, it can be assumed that there will always be some enzyme concentration for each reaction inside the cell that produces a basal metabolic flux. In order to reflect that on the model formulation, a new constraint can be added:

$$v_i^F + v_i^B \geq \varepsilon \quad (2.7)$$

Based on 2.7, every reaction will have a flux value of at least some basal tolerance  $\varepsilon$ , regardless of the reaction directionality. This constraint allows for every reaction to carry at least a basal flux through the metabolic network. In the current thesis the basal tolerance will be equal to  $10^{-6}$  mol/h.

## 2.2.4 Minimization of the sum of fluxes

A hypothesis often employed in constraint-based modeling, is based on the idea that cells operate as efficiently as possible in order to achieve optimal growth conditions<sup>21</sup>. One way to integrate this in the above modeling formulation is by adding a constraint that minimizes the sum of fluxes:

$$\sum_{i=1}^r v_i^F + v_i^B \leq Z_{msf} \quad \text{where } Z_{msf} = \min (\sum_{i=1}^r v_i^F + v_i^B) \quad (2.8)$$

By setting the sum of all fluxes as an objective function  $Z_{msf}$ , it is possible to calculate a minimum allowable value that can then be added as a constraint. This way the solution space can be scaled down even further.

## 2.2.5 REMI

The REMI method (Relative Expression and Metabolomic Integrations) further expands on the ability to integrate thermodynamics and metabolomics on GEMs by allowing the addition of gene expression as constraints<sup>22</sup>. The core assumption that this method operates on is that the relative changes in gene expression between two conditions correlate with the resulting differential flux profiles. As such two, versions of the model (wild-type and mutant) can be linked together in order to add constraints that satisfy information from transcriptomic data:

$$v_i^{mutant} \geq \varepsilon B_i \mid v_i^{wild} \geq \varepsilon B_i \quad (2.9)$$

$$v_i^{mutant} \geq p v_i^{wild} - \sigma_i \mid v_i^{mutant} \leq p v_i^{wild} + \sigma_i \quad (2.10)$$

$$\varepsilon(1 - B_i) \leq \sigma_i \leq \varepsilon + (1 - B_i)K \quad (2.11)$$

Transcriptomic data can indicate if a gene is upregulated or downregulated when comparing expression data between the two cell states. This is reflected on proportional changes in flux values of the reactions regulated by these genes. Equations 2.9 ensure that upregulated or downregulated fluxes always carry a basal flux  $\varepsilon$ . The binary variable  $B_i$  dictates which deregulations are actively imposed. Equations 2.10 enforce a reaction ratio  $p$  between the two fluxes (first inequality refers to upregulated fluxes and the second one to downregulated fluxes). The slack variable  $\sigma_i$  has a very small value when the deregulation is imposed (equation 2.11) so as to not affect the reaction ratio constraints.

The need for the binary variables  $B_i$  stems from the fact that in practice, not all reaction ratios can be imposed simultaneously. In order to find the maximum number of imposed constraints, consistent with the data, an objective function is used:

$$Z_{consistency} = \max (\sum_{i=1}^n B_i) \quad (2.12)$$

Where  $n$  refers to the number of added upregulated and downregulated constraints. A similar approach can be used to find alternative optimal solutions with different groups of active constraints.

## 2.2.6 Variability analysis

All the added constraints in both LP and MILP formulations are changing the solution space such that many variables cannot operate on the whole range that was initially imposed. Variability analysis is a method that aims to find the feasible bounds for a given variable in the model. The type of variable can change depending on the method used such as fluxes for FBA and concentrations and Gibbs free energies for TFA. This is done by setting an objective function for each variable separately and subsequently solving the system to find the maximum and minimum allowable values. The applications of the method include assessing the effect of a constraint on the bounds of a variable and testing if an otherwise bidirectional reaction became unidirectional.

## 2.3 Sampling process

Most biochemical reactions can operate in both directions, giving cells the flexibility to respond swiftly to environmental changes. However, when building a kinetic model from the metabolic system, it is necessary to determine the reactions' directionality. Generally, there will be  $2^k$  possible states of the metabolic network when  $k$  bidirectional reactions are part of the system. These states of the system are named Flux Directionality Profiles (FDPs) and it is crucial to pick a representative in order to build the kinetic model. Integrating data from various sources and different biological levels reduces the number of bidirectional reactions (BDRs). Nevertheless, even after integrating data, there will typically be several bidirectional reactions for whom the directionality should be determined. If the number of BDRs is small, it is common to manually curate the directionalities based on literature data about each biochemical reaction. One approach is to exhaustively enumerate all the possible states of the metabolic network and choose the one that best fits certain predetermined criteria (e.g., stability of kinetic models, variability of solutions space). However, this method is impractical when the number of BDRs is high. To circumvent this, sampling methods can be used that sample an extensive set of steady-state profiles in the solution space. The average of these sample points can be used to extract average directionalities for the remaining BDRs of the system.

### 2.3.1 ACHR sampling

Artificial Centered Hit-and-Run (ACHR) is a sampling algorithm that utilizes Monte-Carlo random walks to pick asymptotically distributed points from the solution space<sup>23</sup>. The sampler requires that the sampled space is convex, meaning that only LP formulations are allowed. A number of convergence criteria can be used in order to assess the convergence of the sampling process<sup>24</sup>. The algorithm used contains the following steps:

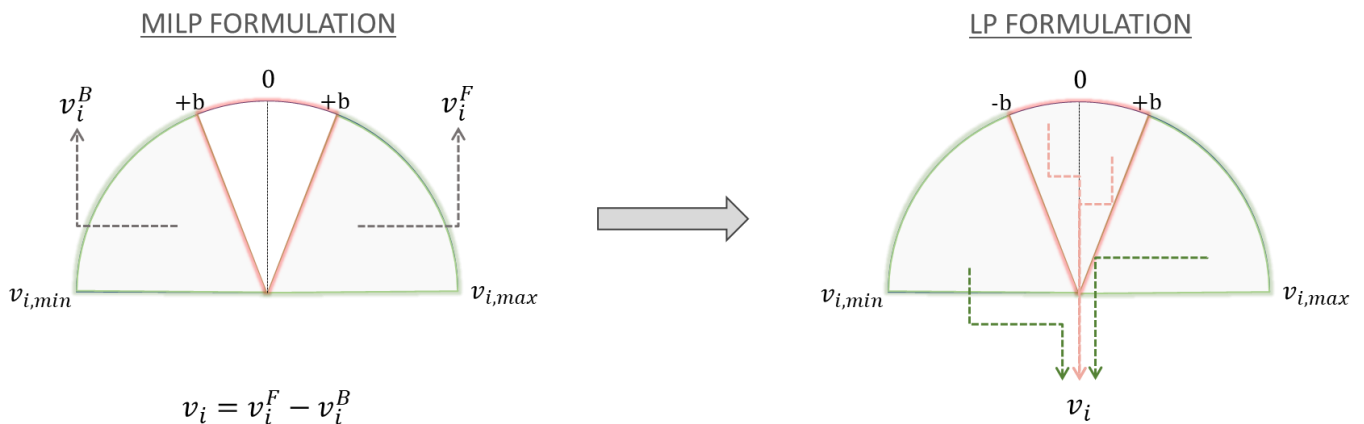
1. Sample warmup points, which are steady-state solutions in the edges of the solution space. A minimum of  $2x$  the number of reactions, warmup points are required for which every reaction is sequentially maximized and minimized. Any further warmup points are derived by using an objective function with randomly chosen weights for the  $c$  vector.
2. Find the center point of the sampled warmup points and begin Monte-Carlo random walks on the lines connecting the warmup points and the center point. Note that every point on these lines, respects equation 2.1.

- The random walk consists of a number of thinning steps that separate the actual sampled points. When a new sample point is picked, it is used to recalculate the center point. The process continues until the number of requested sample points is chosen.

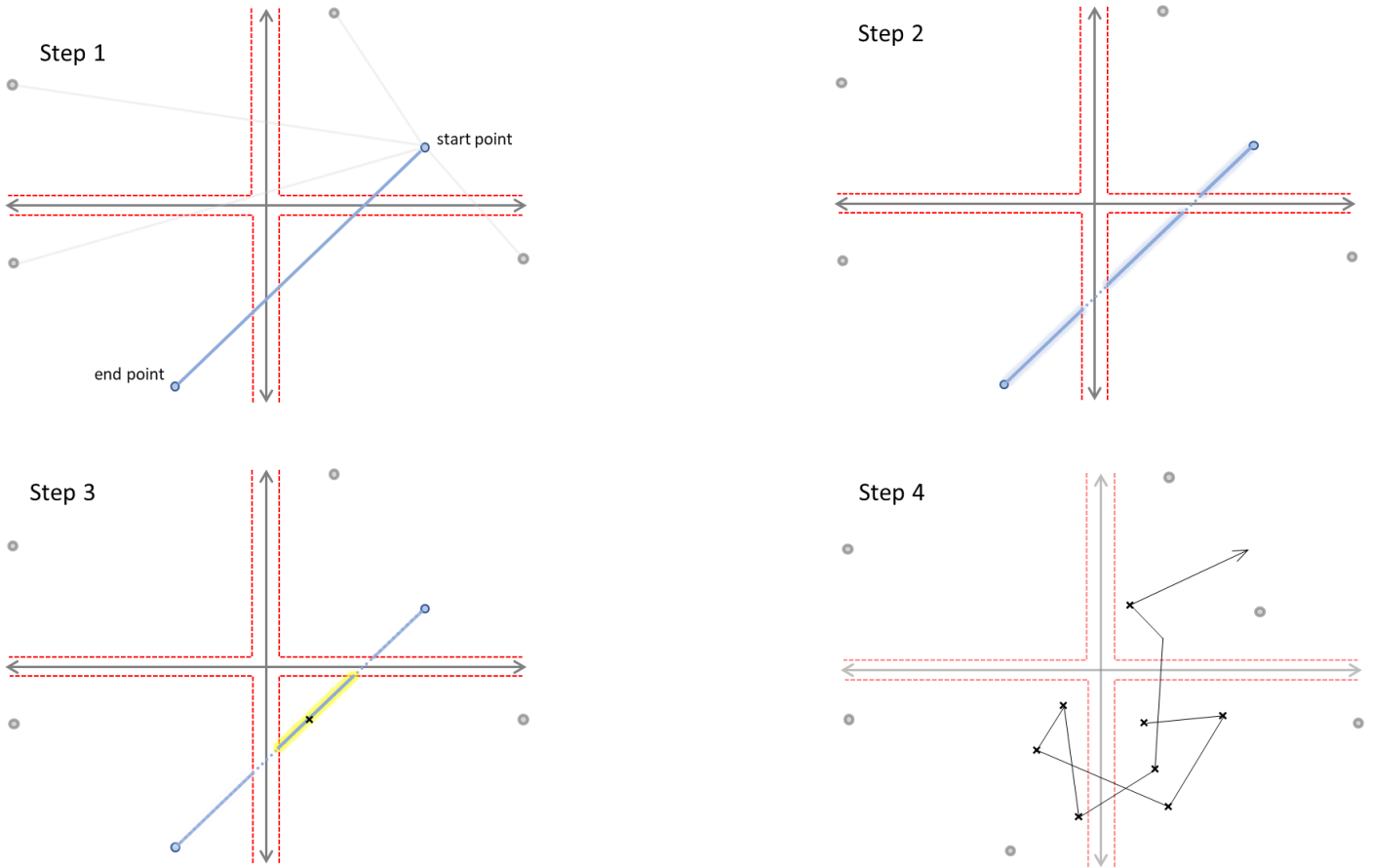
### 2.3.2 Avoiding basal bounds

As stated above, the model must be in LP formulation to perform sampling with the ACHR sampler. In the current thesis, the MILP formulation of the model was simplified to an LP version by keeping only the variables of flux and assigning bounds to them based on variability analysis performed on the MILP formulation. However, flux variables may have feasible bounds from negative to positive values since they are not split into their forward and backward counterparts. This particular case is observed when a reaction can still operate in both directionalities despite all the imposed constraints (BDR). These bidirectional reactions pose a problem because they do not reflect the basal bound constraint imposed on the MILP formulation (Figure 2.3). Sampling with this version of the model will result in steady-state samples containing fluxes of bidirectional reactions that do not satisfy the basal tolerance. A workaround is introducing these bounds to the sampler so that the random walk cannot sample in the respective areas. In order to achieve this, several changes are required:

- Warmup points are sampled using the model in the MILP formulation. The points are then confirmed as feasible in the LP version of the model.
- Since the ACHR sampler uses lines between two points to choose the next point randomly, it is possible that these lines intersect basal bounds. The line parts that are inside basal bounds should be discarded, as sampling in them will yield unfeasible sample points. To avoid this, an additional algorithm was developed that segments the line into parts that do not intersect basal bounds and randomly chooses one of them to continue the random walk on (see Figure 2.4).



**Figure 2.3** In MILP formulation, variables  $v_i^F$ ,  $v_i^B$  can have positive, bigger than the basal tolerance values and only one is active at a time (based on eq. 2.3). The net flux feasible range is  $[v_{i,min} \ v_{i,max}]$  (green ranges) but basal bound constraints exclude the values that are smaller than the basal tolerance (red range). However, when imposing these bounds in the LP formulation, the flux variable is not confined by basal bound constraints and can thus have values that are smaller than the basal tolerance.



**Figure 2.4** Step 1: randomly choose a direction for the next sample point. Step 2: if the line crosses basal bounds (red areas) perform a segmentation and only keep the parts that satisfy basal tolerance for all flux variables. Step 3: randomly pick one of the allowed segments and sample a point from it. Step 4: perform the above process iteratively until the number of requested sample points have been reached. All of these points respect the basal tolerance constraint.

These changes can increase the computation time of the algorithm by an order of magnitude, however, they ensure that the new sample points are feasible in the MILP formulation. Additionally, since all the sampled points satisfy the basal flux tolerance, the average values calculated for the fluxes of the bidirectional reactions should also be larger than the basal tolerance.

## 2.4 Kinetic models

Kinetic models of metabolism capture the transient state of the system based on a set of ordinary differential equations of mass balances and reaction kinetic formulas:

$$\frac{dX}{dt} = Nv(X, p) \quad (2.13)$$

Equation 2.13 is similar to 2.1, only now the fluxes  $v_i$  are time dependent functions. Specifically,  $v(X, p)$  represent reaction rate laws that have as input metabolite concentrations  $X$  and kinetic

parameters  $p$ . A summary of the reaction kinetics used is documented on Appendix A. These reaction rate laws are assigned based on the different stoichiometric coefficients of the respective reactions and the kinetic parameters are estimated using the ORACLE framework<sup>25</sup>.

### 2.4.1 Building kinetic models

The ORACLE framework produces sets of kinetic parameters around steady-state samples of concentrations and fluxes. In particular, ORACLE estimates the following kinetic parameters:

$$K_{eq,i} = -RT\Delta_r G_i \quad (2.14)$$

$$\sigma_{i,j} = \frac{[X_j]}{K_{M,i,j} + [X_j]} \text{ where } \sigma_{i,j} \in [0, 1] \quad (2.15)$$

$$v_{i,steady\ state} = v_{max,i} * f(X, p) = v(X, p) \quad (2.16)$$

Here,  $\sigma_{i,j}$  is a saturation constant that is sampled between 0 and 1 to more easily calculate the values for  $K_{M,i,j}$  parameters.  $K_{eq,i}$  is the equilibrium constant that can be calculated by the standard Gibbs free energy and  $v_{max,i}$  is the maximum velocity of the rate law of the reaction, that can be calculated when taking into consideration the sampled steady-state flux value. The  $f(X, p)$  is the saturation term which changes depending on the rate law of the reaction.

### 2.4.2 Pruning kinetic models

The sampling of kinetic parameters described in the previous section produces kinetic models based on steady-state samples of concentrations and fluxes. However, the sets of kinetic parameters sampled will not always satisfy local stability. The criteria used to check for local stability are based on the eigenvalues of the Jacobian matrix of the ODE system. Specifically, assuming small perturbations in the concentrations  $[X_j]$  of the system, the differential equations can be linearized with a Taylor expansion:

$$\frac{d(\delta X_j)}{dt} = \sum_{k=1}^m \frac{\partial(Nv)}{\partial X_k} \delta X_j + \frac{1}{2} \sum_{k=1}^m \frac{\partial^2(Nv)}{\partial X_k^2} \delta X_j^2 + \dots = \sum_{k=1}^m \frac{\partial(Nv)}{\partial X_k} \delta X_j + \text{error} \propto (\delta X_j^2)$$

$$\text{or in matrix form } \frac{d(\delta X)}{dt} \cong J\delta X \quad (2.14)$$

If the Jacobian matrix  $J$  is diagonalizable, the above equation can be integrated and return:

$$\delta X_j(t) = \sum_{i=1}^n c_i b_i \exp(\lambda_i t) \quad (2.15)$$

Here,  $c_i$  are unknown constants determined by the initial perturbations,  $b_i$  denotes the eigenvectors and  $\lambda_i$  the eigenvalues of the Jacobian matrix. Based on equation 2.15, the kinetic model can be considered asymptotically stable if and only if the real part of the eigenvalues of the Jacobian matrix are negative. When sampling a new set of kinetic parameters, ORACLE performs the above stability check by calculating all the eigenvalues of the system and checking the sign of their real part.

The sampled kinetic parameters can be further pruned by requiring physiologically relevant time scales. While eigenvalues with a negative real part ensure linear stability, the system must be able to return to the steady-state profile in a timescale that is practical for the metabolic network of the system (when a perturbation is imposed). One way to quantify this constraint is to assume that the metabolic processes operate with a timescale several times faster than the doubling time of the cell. For a cell culture growing in the exponential phase, the doubling time can be calculated by:

$$\frac{dB}{dt} = \mu B \Rightarrow B(t) = \exp(\mu t) + \ln(B_o) \rightarrow t_d = \frac{\ln 2}{\mu} \quad (2.16)$$

Where  $B$  is the cell concentration,  $\mu$  is the growth rate constant and  $t_d$  is the doubling time of the cell. For the purposes of this thesis, the set of kinetic parameters is further pruned to meet physiologically relevant timescales. The criteria enforced is that the maximum eigenvalue of the Jacobian matrix must be larger than three times the doubling time of the simulated cell. This condition is based on the idea that aperiodic systems will settle to approximately 95% of the steady-state value after three dominant time constants.

Lastly, the kinetic models can be examined on their robustness to nonlinear perturbations. Specifically, by applying random perturbations to the initial steady-state concentrations, the kinetic models can be assessed based on their ability to dampen the perturbations in less than one doubling time cycle. In the current thesis, this is done by applying a 10% random perturbation to each concentration 50 times, integrating the ODE system and checking if the Euclidean norm of the normalized concentrations has reached 95% of their steady-state values before the doubling time. The sets of kinetic parameters that satisfy these criteria for all 50 perturbation runs are considered robust and are used in the later stages of the thesis.

## 2.5 MCA and ODE simulations

The pruned sets of kinetic parameters can be used to examine proposed drug targets and their effects on biomass production and the system at large. Additionally, using the Metabolic Control Analysis method (MCA), new drug targets can be proposed based on the simulated physiology of ovarian cancer cells. MCA uses control coefficients to estimate the concentration and flux responses to a specific parameter perturbation:

$$C_{p_k}^{x_j} = \frac{d \ln[x_j]}{d \ln p_k}, C_{p_k}^{v_i} = \frac{d \ln v_i}{d \ln p_k} \quad (2.17)$$

For example, in the case of this thesis, MCA calculates the control coefficient values for the biomass flux variable. These calculations are done for each pruned kinetic model, and an average control coefficient is produced for each kinetic parameter. The control coefficients are then ranked by absolute value, and the top ranked are examined. MCA is a valuable tool for uncovering complex interactions inside the simulated metabolic network.

Simulating enzyme perturbations is a way to quantify the MCA estimations as well as test already proposed drug targets. The perturbation of the enzyme's concentration can be imposed on one or more reactions of the metabolic network simultaneously. In practice, the perturbation is applied by altering the  $v_{max}$  value of the respective reaction rate laws (since  $v_{max} = k_{cat}[E_{tot}]$ ). The ODE

system is then integrated until it reaches a new steady-state. The change in the  $v_{max}$  is similar to the effect of an inhibiting ligand on the enzyme's activity or a gene knockdown. It is even possible to simulate gene knockouts by setting the  $v_{max}$  value to zero. Ideally, for promising drug targets, the calculated biomass flux should decrease or even get nullified, signaling cell death.

## 2.6 Code availability

The mathematical formulations described in section 2.2 have been integrated into the current thesis through the Python package PyTFA and the MATLAB toolbox matTFA<sup>26</sup>. The kinetic analysis results were produced with the use of the SKiMpy python package. The computations were performed on a desktop with a 36-core Intel® Xeon(R) Gold 6254 CPU @ 3.10GHz and 93 GB of RAM. All the code used in the current thesis, along with the required toolboxes and packages, can be found on the GitHub page: [https://github.com/Ilias719/LS\\_kinetic\\_OC](https://github.com/Ilias719/LS_kinetic_OC).

## 3. Results

### 3.1 Building the model of Ovarian Cancer cells

#### 3.1.1 Model curation

The model used for the current thesis is a reduced version of Recon3D GEM produced with the redHUMAN workflow and was provided by the EPFL LCSB lab. It consists of 614 metabolites participating in 1724 reactions. The data integrated into the model were:

- Thermodynamic constraints for 998 reactions (section 2.2.2).
- Transcriptomic data for 332 reactions. The gene expression data were extracted from wild-type and BRCA1 mutated ovarian cancer cells<sup>27</sup> and were integrated into the model if the reported fold change in gene expression was at least 1.3.
- Fluxomic data for 67 reactions<sup>27</sup>.

After the integration of the above data, the model contained 488 BDRs, of which 72% were transport reactions, 23% were part of the intracellular metabolism and synthesis pathways, and 5% accounted for boundary reactions which are necessary for maintaining the mass balances of the extracellular metabolites within the network .

As discussed in the Methods section, in order to build sets of kinetic models, all reactions should be constrained to operate in one direction. For this reason, a series of curation steps were followed to reduce the number of BDRs. After each imposed step, a variability analysis of the remaining BDRs was performed.

Firstly, basal flux constraints were imposed on the model, reducing the BDRs to 472. The number of BDRs was further reduced to 456 by minimizing the sum of fluxes. Next, a new objective function was temporarily used to describe the cancer cell's physiology more precisely and to examine if it will further decrease the number of BDRs. This objective function maximizes biomass production, ATP synthesis, glutamine uptake and lactate production. The reactions in the model involved in this objective function are shown below:

$$Z_{physiology} = \max (F_{biomass} + ATPS4mi + PGK + PYK + EX_{gln\_L\_e} + EX_{lac\_L\_e})$$

This objective function is based on experimental observations of cancer physiology that include elevated levels of energy requirements, high lactate production (Warburg Effect) and glutamine uptake. By enforcing this objective function and doing variability analysis for the rest of the BDRs, their respective ranges were further constrained while also closely describing cancer cell physiology. After all the modifications, 450 BDRs remained, of which 347 are transport reactions, 18 are exchange reactions, 25 are nucleotide interconversion reactions and 60 are part of intracellular metabolism and synthesis pathways. The latter group was manually curated in order to pick a directionality leaving 390 BDRs. The imposed curations are documented in Appendix B.

The above steps considerably reduced the number of bidirectional reactions; however, it is difficult to accurately choose a directionality for the rest of them, especially since almost 90% are transport reactions. For example, many bidirectional transport reactions occur between the extracellular medium and the cytosolic compartment of the cell, where two or more amino acids are interchanged through the compartments. It is difficult to manually select a directionality for these types of reactions as it leads to a bias towards the preferred compound inside the cell. Additionally, it is known that cells utilize these transport reactions to dynamically change the concentration of amino acids based on their needs and environmental changes. The same argument can be made for nucleotide interconversion reactions, where it is difficult to assess the importance of converting one nucleotide to another.

Sampling a set of feasible steady-state profiles of fluxes can provide a statistical approach for acquiring directionalities for the rest of the BDRs. Using the modified ACHR sampler (section 2.3.2), a set of 1 million sample points was produced from an LP version of the model. The number of warmup points was 10000, and the number of thinning steps was set to 500. The 1 million steady-state flux values were averaged, and the mean point was used to extract the directionality for each of the remaining BDRs.

The final version of the model is showcased in Figure 3.1. Specifically, core parts of the metabolic network are shown, including the glycolysis pathway, pentose phosphate pathway, tricarboxylic acid cycle, and urea cycle. While there are only 73 reactions displayed, they can depict the attempt of the enforced constraints to better describe the physiology of the cell. The reported values of the fluxes are the average of 1000 steady-state samples of the final FDP.

The glycolysis pathway is shown to carry relatively high amounts of carbon flux most of which ends up secreted as lactate. Parts of the pentose phosphate pathway aid in this carbon flux and also produce ribulose 5-phosphate, a precursor for the production of nucleotides. The urea cycle includes fluxes that overall lead to the secretion of urea outside the cell.

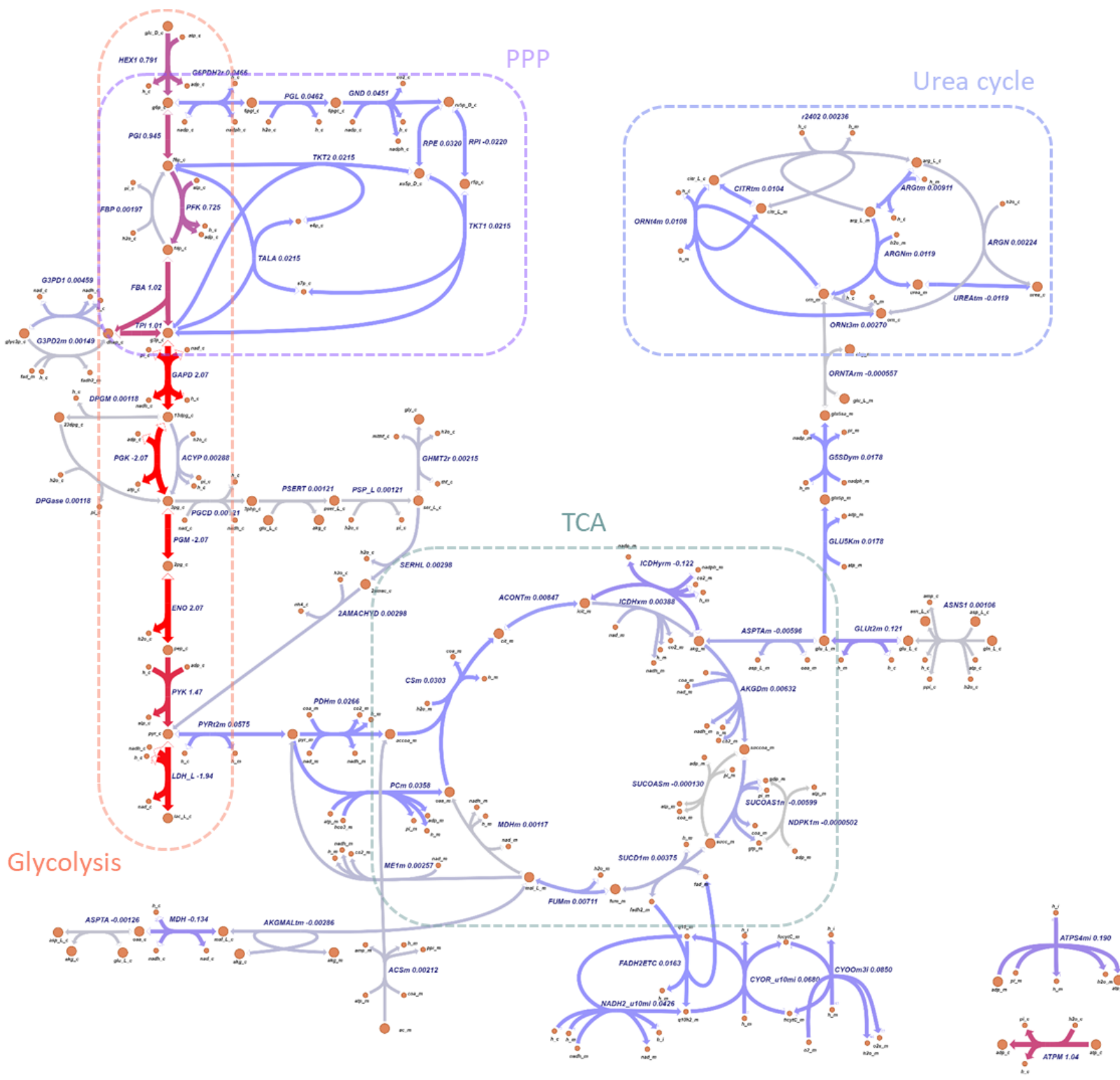
An important pool of metabolites in the metabolic network of any cell is the adenylate molecules i.e., ATP, ADP, and AMP. More specifically, it has been reported that the concentrations of these metabolites are always in a specific ratio called adenylate energy charge, which is calculated by:

$$ec = \frac{[ATP] + \frac{1}{2}[ADP]}{[ATP] + [ADP] + [AMP]} \quad (3.1)$$

The energy charge can range from 0.7 all the way to 0.95, with human cells being closer to the upper end of the range<sup>28</sup>. In order to better depict this property of the metabolic network, the model was enriched with ATP/ ADP/ AMP metabolomic data from *Homo Sapiens*<sup>29</sup>. The bounds of the concentrations were relaxed by 20%, to not constrain the system too much. It is worth noting that these relaxations result in high and low energy charge values that were inside the above provided range. The concentrations imposed are reported in Table 3.1.

**Table 3.1** Metabolomic data from *Homo Sapiens*.

Metabolite	Concentration (mol/L)
ATP	$4.67 \times 10^{-3}$
ADP	$5.69 \times 10^{-4}$
AMP	$4.23 \times 10^{-5}$



**Figure 3.1** Core reactions of the metabolic network. The pathways included are Glycolysis, Pentose Phosphate Pathway (PPP), Tricarboxylic Acid cycle (TCA) and Urea cycle. The values displayed next to the reaction's name are an average of 1000 sampled steady-state flux profiles, in mol/h. Negative values indicate that the reaction operates in the opposite direction to the one that was assigned by default in the model. The color intensity of the fluxes goes from light grey to blue to red as the absolute value increases. The map was made using the Escher Toolbox<sup>30</sup>.

### 3.1.2 Kinetic model building

The structure of the kinetic model was built by assigning reaction rate laws based on the stoichiometry of each reaction. As shown in Table 3.2, the vast majority of reactions were assigned the Generalized Reversible Hill mechanism followed by Reversible Michaelis-Menten. The resulting kinetic model must be populated with values for all the kinetic parameters involved in the rate laws. To achieve this, the following steps were done:

1. 1000 steady-state samples of fluxes, concentrations, and thermodynamic variables were produced.
2. For each steady-state profile, 100 sets of kinetic parameters were generated that were consistent with the associated steady-state profile and resulted in a linearly stable system (real part of eigenvalues is negative).
3. The 100,000 resultant kinetic models were pruned based on their adherence to physiologically relevant timeframes.

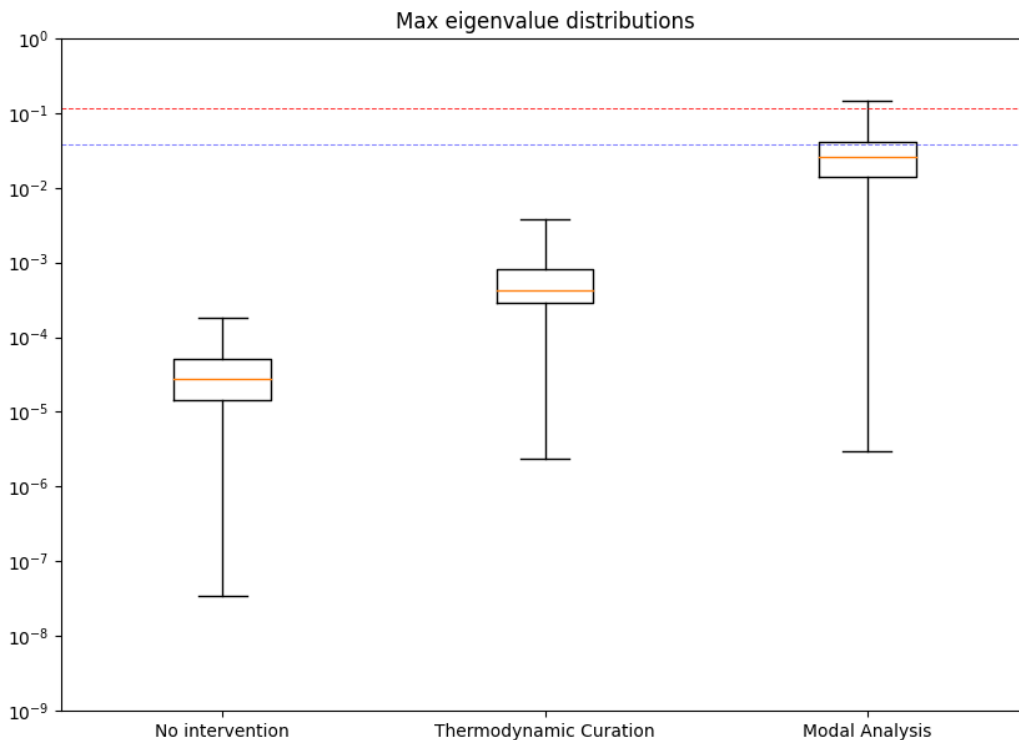
For the last step, based on equation 2.16, it is possible to calculate that for a growth rate of  $0.026 \text{ hour}^{-1}$ , the doubling time of the simulated cell is a little over 26 hours. As such, for the purposes of this thesis, those models were chosen whose dominant time constant was at least three times faster than the doubling time of the cell (i.e., the maximum eigenvalue of the Jacobian matrix should be smaller than  $-3/26 \approx -0.11$ ).

In the reduced model of the cell, there were 610 reactions for which thermodynamic data was not available. In the first run, before generating steady-state samples, these reactions were assigned standard Gibbs free energy  $\Delta_r G^\circ$  ranges from -10 kcal/mol to 10 kcal/mol. No other intervention was done, but when following the above steps, the kinetic models produced did not have physiologically relevant dynamics. Even the fastest sets of kinetic parameters had dominant time constants that were three orders of magnitude slower than the cell's doubling time (Figure 3.2). A series of changes were made to the constraints imposed at the steady-state level to explore what contributes to the slow dynamics.

Firstly, all reactions that facilitate a diffusion transport of compounds were enforced to operate close to thermodynamic equilibrium. This step ensured that diffusion phenomena did not limit the whole metabolic network. Additionally, all other reactions were set to operate away from thermodynamic equilibrium to diversify the equilibrium constants' values. Indeed, as shown in Figure 3.2, these changes resulted in a faster response of the metabolic network, but it was still two orders of magnitude slower than the desirable eigenvalue boundary.

**Table 3.2** Types of rate law mechanisms and the number of reactions reported for each of them.

<b>Mechanism</b>	<b>Number of reactions</b>
Generalized Reversible Hill	1048
Reversible Michaelis-Menten	293
Convenience	136
Uni-Bi Reversible Hill	90
Irreversible Michaelis-Menten	27
Bi-Uni Reversible Hill	9



**Figure 3.2** Max eigenvalue boxplot distributions of the stable sets of kinetic parameters. The y-axis is in log form with positive values for easier understanding. The blue dashed line represents the doubling time of the cell, while the red dashed line represents three times the doubling time. The orange lines in the boxplots show the median value.

To achieve physiologically relevant dynamics, the model's variables should be curated in order to guide the sampling process towards faster kinetic models. Modal analysis can quantify the timeframe of the network by linearizing the ODE system and performing a linear transformation to identify groups of metabolites that evolve together. In the case of constrained-based models, these pools of metabolites are associated with one of the eigenvalues of the Jacobian matrix. By performing modal analysis, it is possible to connect the eigenvalues that do not achieve the physiologically relevant timeframe with groups of metabolites. It is safe to assume that these metabolites are evoking the slowness in the system. So, the model was tweaked to enforce lower concentration ranges for the targeted metabolites and higher flux ranges for the reactions they participate in. These changes exclude these metabolites from being the rate-limiting step at some point in the evolution of the system. The modal analysis results returned 342 metabolites that were linked with slow dynamics, for which a manual curation of their concentration bounds and connected flux ranges was performed. This resulted in an overall improvement of the dynamics which led to 9 sets of kinetic parameters with dominant time constants that were at least three times faster than the doubling time of the cell (Figure 3.2).

These 9 kinetic models were further pruned based on their robustness to nonlinear perturbations. Specifically, by imposing small random perturbations to the concentrations of the system, the sets of kinetic parameters can be examined as to their ability to return to steady-state efficiently (in less than one doubling time cycle). The steps followed in Methods section 2.4.2 returned 6 sets of kinetic parameters that can be considered robust to nonlinear perturbations. Moving on to drug target analysis, the results discussed were derived from these 6 kinetic models.

## 3.2 Drug target analysis

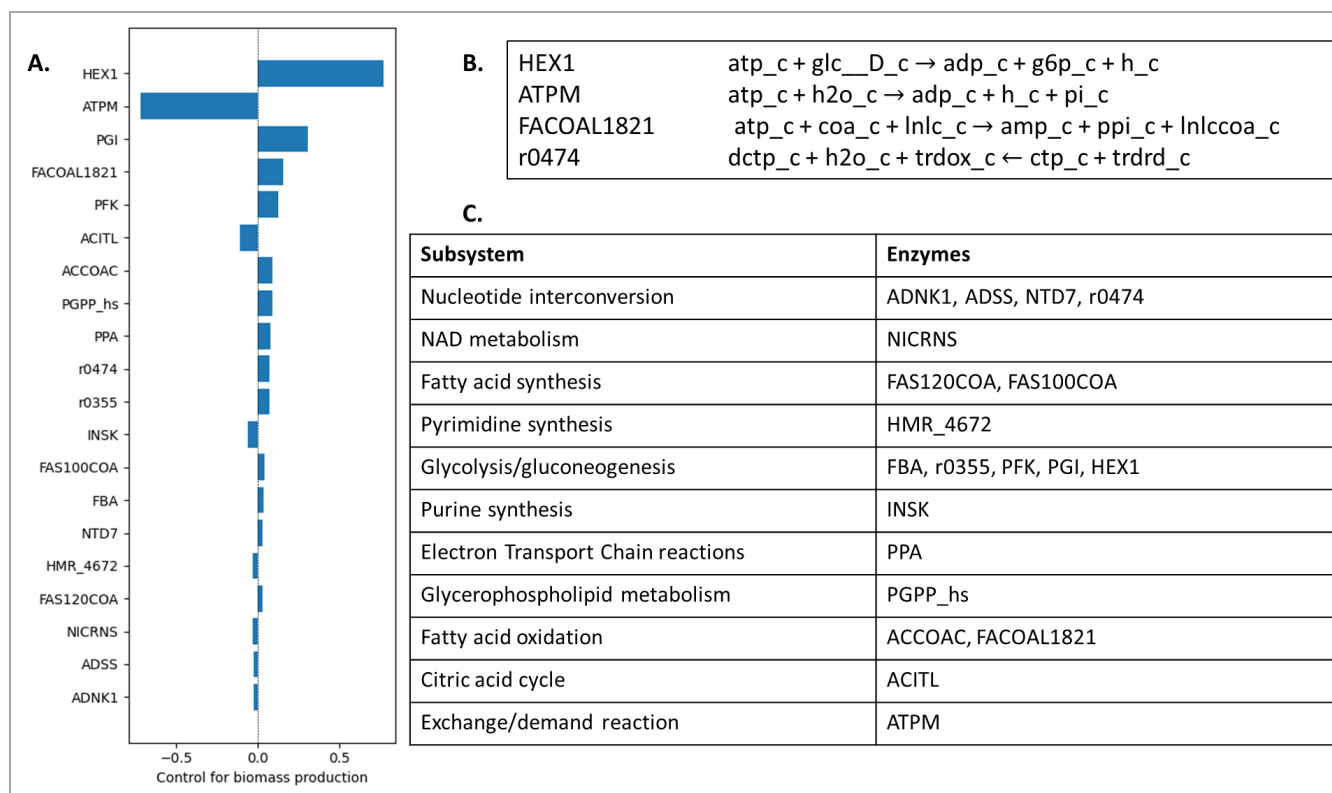
### 3.2.1 Drug targets derived from MCA

Metabolic Control Analysis was performed on each of the 6 kinetic models. The control coefficients for biomass production were calculated, and the results were averaged over the 6 kinetic models and ranked by absolute value. The top 20 control coefficients are shown in Figure 3.3A.

Enzyme HEX1 (Hexokinase: responsible for converting glucose to glucose-6-phosphate) has the highest control coefficient value, which is positive, meaning that an increase in its corresponding reaction flux will result in an increase in biomass production. The opposite effect is displayed in ATPM (ATP maintenance requirement), which holds the second highest absolute value control coefficient, and an increase in the reaction's flux will decrease the biomass flux value. In total, there are 13 reactions with a positive control coefficient and 7 with a negative control coefficient. Reactions that rank lower than the top 20 are not considered since their control coefficient value is insufficient to derive meaningful observations.

The top 20 reactions are from a group of 11 subsystems (Figure 3.3C). The one with the most impactful control coefficients is the glycolysis/glycogenesis pathway containing 5 reactions. Three subsystems (nucleotide interconversion, pyrimidine synthesis, purine synthesis) are directly linked with the metabolism and synthesis of nucleotides. Additionally, another three subsystems (fatty acid synthesis, fatty acid oxidation, glycerophospholipid metabolism) are linked with the metabolism of lipids and fatty acids. These results indicate that biomass production is heavily impacted by changes in reactions that are part of synthesis pathways for some mandatory compounds in the growth and doubling process of the cell. Specifically, the biomass production reaction, used to simulate the growth of the cells in the model, contains amino acids, glucose compounds, dNTPs, and several lipids and the examined reactions are linked with one of those key compounds.

Recent literature indicates agreement with the above findings. Regarding glycolysis, it is well established that cancer cell metabolic phenotypes are dependent on this pathway, mainly for energy production, regardless of the presence of oxygen<sup>31</sup>. Cancer cells produce the required levels of cellular energy through increased glucose uptake and aerobic glycolysis, even though this is a less efficient pathway (compared to oxidative phosphorylation in mitochondria). Additionally, glycolysis is used to produce vital metabolic intermediates for the biosynthesis of essential macromolecules, which serve as building blocks for tumor proliferation and growth. Such compounds include NADPH, which is stated to provide chemoresistance by regulating the redox status of the cells, and ribose-5-phosphate, an essential precursor of nucleic acids and lipids. Lastly, many glycolytic enzymes have been linked with crucial non-glycolytic processes, such as HEX1 which is involved in transcriptional regulation and phosphorylation of histones. Many drug targets from the glycolysis pathway have been experimentally examined, with many of them reaching Phase III clinical trials<sup>31</sup>.



**Figure 3.3** A. The 20 enzymes with the top control coefficients by absolute value calculated by MCA. B. Reaction formula of the enzymes that are further examined in section 3.2.2. C. List of groups of enzymes and the metabolic subsystems they participate.

Interestingly, enzymes HEX1 and PFK (Phosphofruktokinase), which MCA suggested, have already been studied as the reactions catalyzed by them are rate-limiting in the glycolysis pathway. Other notable enzymes are glucose transporters for blocking the uptake of glucose, GAPD for the deregulation of NADPH production and PK (Pyruvate kinase) and LDH (Lactate dehydrogenase). While experimental trials have shown promising results for many of these enzymes, a significant drawback is that the inhibitors used also affect other types of cells leading to concerns about toxicities. Despite that, antiglycolytic agents may prove a pivotal part of combination therapies, where the parallel use with chemotherapeutic drugs has been shown to lead to lower drug resistance<sup>31</sup>.

Moving on to lipids, there is an abundance of different categories and subsequent compounds that contribute to different tasks in the cell, such as membrane formation, energy storage and production (through fatty acid oxidation), and signaling pathways that contribute to tumorigenesis. Of the many enzymes that participate in these tasks, there are a few that showcase deregulation in malignant tumor cells and are of importance when it comes to inhibiting cancer physiology. Two of those enzymes are Acetyl-CoA carboxylase and Fatty Acid Synthase, both of which were highlighted by MCA as crucial to biomass production (enzymes ACCOAC and FAS120COA / FAS100COA, respectively). They have been experimentally tested for cancer inhibition with promising results. Additional deregulated enzymes are Stearoyl-CoA desaturase and the cholesterol synthesizing enzymes HMGCR and SM. Since lipid metabolism is closely regulated by intracellular oncogenic signaling, it is challenging to propose simple treatment strategies that

inhibit fatty acid or cholesterol synthesis, especially since differences in dietary lipids may disrupt the results. A more combinatorial and specific targeting of the lipid metabolism is speculated to achieve pivotal results<sup>32</sup>.

One more metabolic subsystem of importance is nucleotide metabolism. It is complex and diverse and includes de novo synthesis of nucleotides as well as salvage pathways. Nucleotide metabolism is regulated by key metabolic enzymes as well as allosteric effects by different nucleotide substrates. An upregulated nucleotide metabolism is essential in cell proliferation, a finding which has led to the development of many inhibitory drugs. They account for nearly 20% of all developed drugs for cancer treatment and can be classified into three main categories; purine analogs, pyrimidine analogs, and metabolic enzymatic inhibitors. Apart from the last category, nucleotide analog drugs have a more general inhibitory effect as they often target more than one enzyme. Interestingly, dysregulated nucleotide metabolism has been also found to alter the immune microenvironment and affect the host immune response<sup>33</sup>. However, such cell interactions are beyond the scope of this thesis. Regarding the MCA results, 4 enzymes execute nucleotide interconversions, HMR\_4672 (dCTP Deaminase) produces dUTP and INSK (Inosine kinase) produces inosine monophosphate. The sensitivity of the biomass production to these enzymes might imply physiology-specific irregularities, providing a way for more targeted inhibition of nucleotide metabolism. However, due to the high uncertainty involved with the directionality of many nucleotide interconversions, these results might occur from the specific state of the metabolic network. Further examinations with different alterations of the reaction directionalities (different FDPs) could show the possible uniqueness and importance of these results.

The above comparison with literature data, showcases the validity of the model's estimations. Based on the fact that three subsystems of interest and particular enzymes were correctly estimated, it is likely that other proposed drug targets that were not mentioned in the literature are also promising drug targets.

### 3.2.2 Simulations of enzyme perturbations

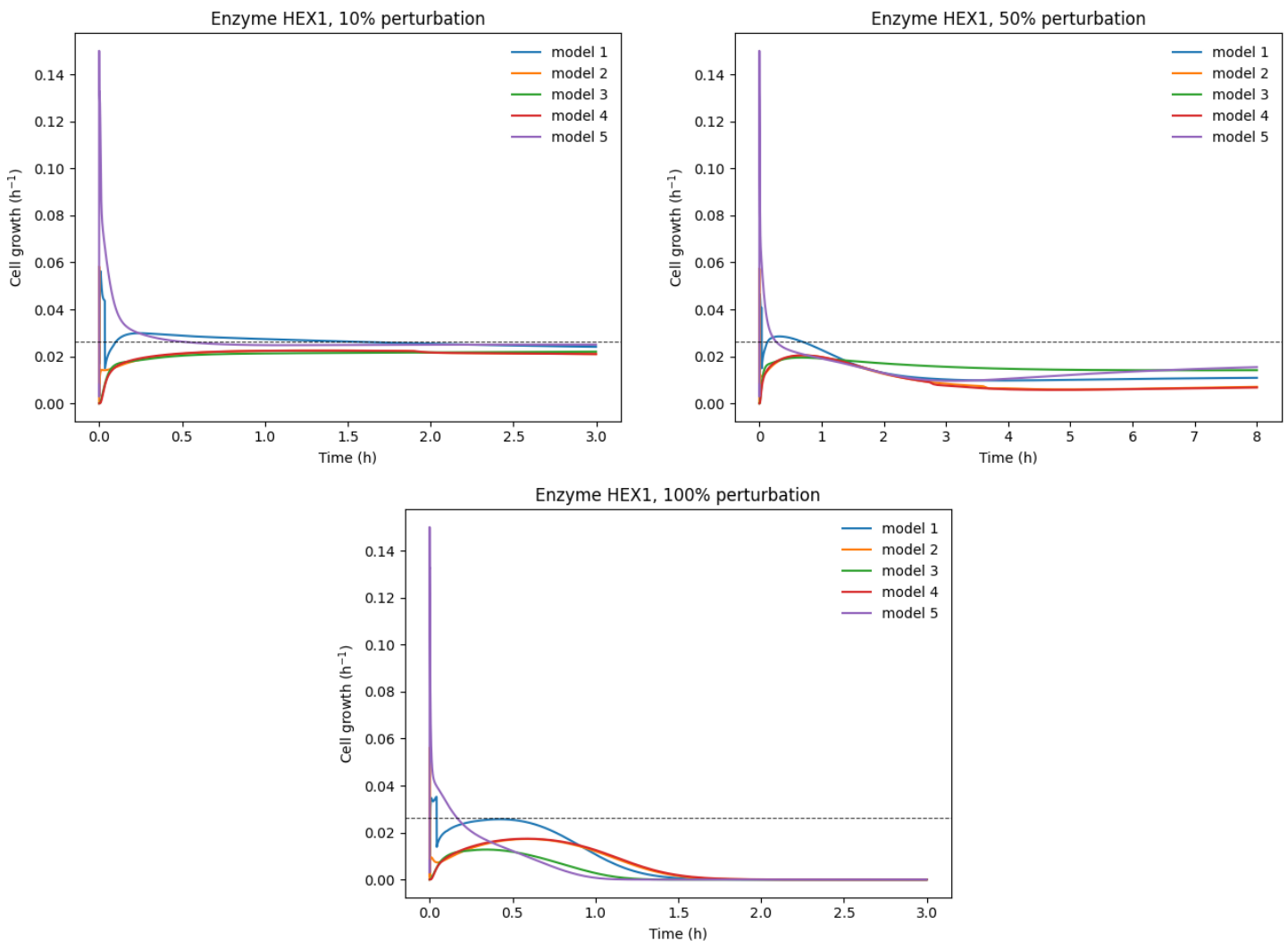
For the purposes of this thesis, each enzyme with the highest control coefficient from the three above discussed metabolic subsystems was examined (Figure 3.3B). Enzymes HEX1, FALCOAL1821 (Long-chain-fatty-acid—CoA ligase), r0474 (Ribonucleoside diphosphate reductase dCTP-forming) along with ATPM (for its potential to deregulate the elevated levels of the cell's ATP) were subjected to enzyme perturbations by changing their  $v_{max}$  value by 10%, 50% and 100%. This change in the  $v_{max}$  value was positive for enzymes with a negative control coefficient and vice versa, in order to achieve a decrease in biomass production. A 100% perturbation indicates gene knockout in enzymes with a negative control coefficient and double the enzyme activity in enzymes with a positive control coefficient.

It should be noted that in every perturbation run, regardless of the enzyme, one of the six models was highly sensitive to the changes imposed and always returned a biomass production flux of zero. This may indicate that the particular set of kinetic parameters was sensitive to numerical errors or that this state of the kinetic model portrayed a very unstable physiology. As such, it is excluded from any further analysis and figures.

## HEX1

Figure 3.4 portrays the results from all three enzyme perturbations. As expected, the perturbations led to new, lower steady-state value for the biomass flux. By increasing the intensity of the perturbation, the effect is more significant. Differences are also observed between the five models which showcase diverse responses to the perturbation and distinct steady-state profiles.

A 10% enzyme perturbation ( $v_{max} = 0.9v_{max,original}$ ) does not have significant effect to the biomass flux with the largest deviation calculated being only 21%. The effect is more prominent at 50% perturbation, where the deviations range from 41% all the way to 74%. This comes to show the inhibitory effect that HEX1 has on the biomass production. Finally, a 100% perturbation, simulating complete inhibition of HEX1 enzyme or a relevant gene knockout, leads to cell death for all models. Thus, the metabolic model is heavily dependent on HEX1 establishing its gene as essential for the survival of the cell.



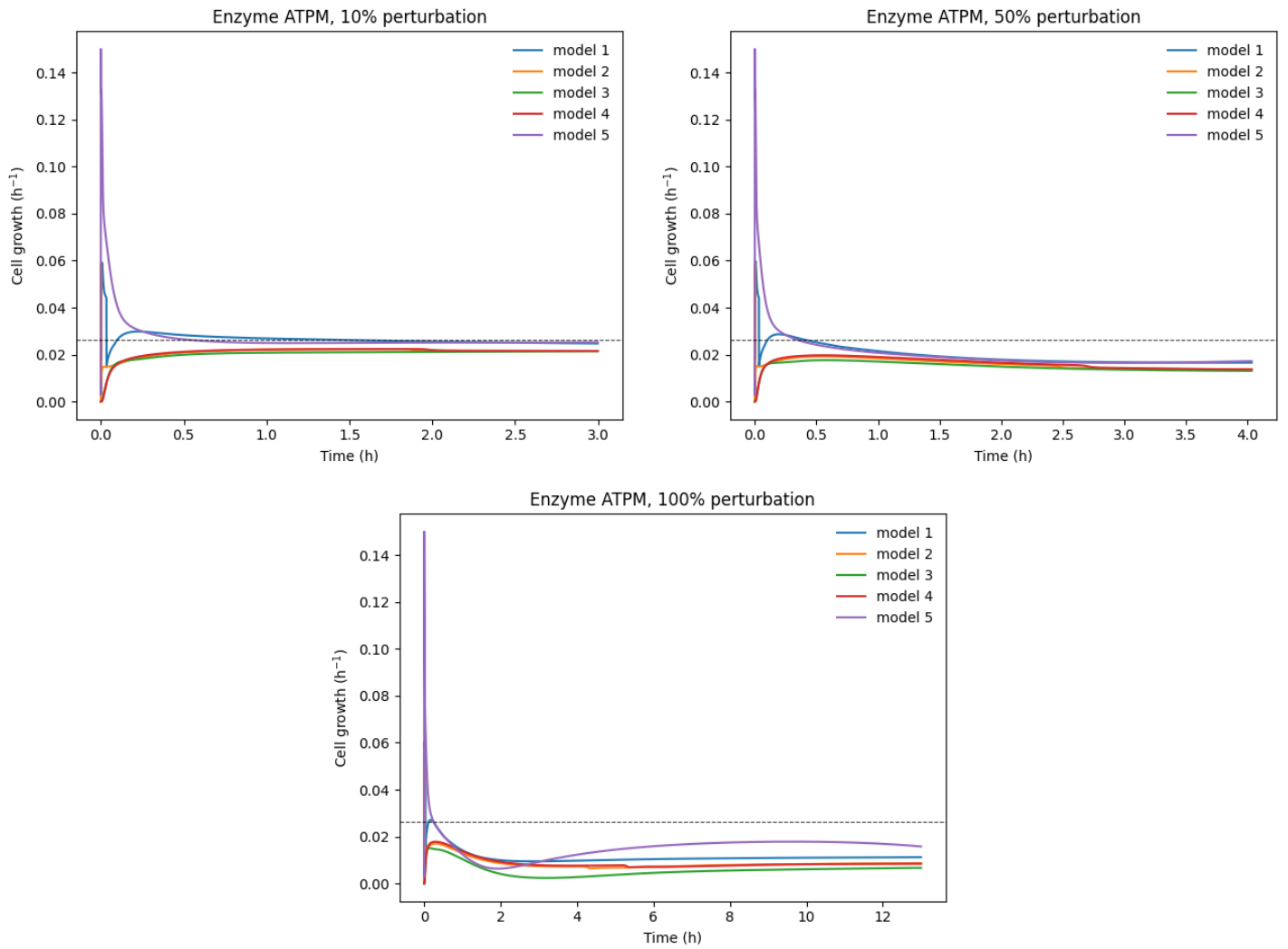
**Figure 3.4** Time graphs of cell growth for each model and each enzyme perturbation of HEX1. The grey dashed line indicates the original biomass flux value at steady-state.

Regarding the time scale of the simulated system, it can be observed that reaching a new steady-state occurs in a couple of hours, which is much faster than the doubling time of the cell. The maximum time needed is 8 hours (for 50% perturbation), and the minimum is less than 2 hours (for 100% perturbation). It is also worth mentioning that all models reached steady-state in very similar timeframes. Information about the effect's timescale is vital for drug development experimental trials.

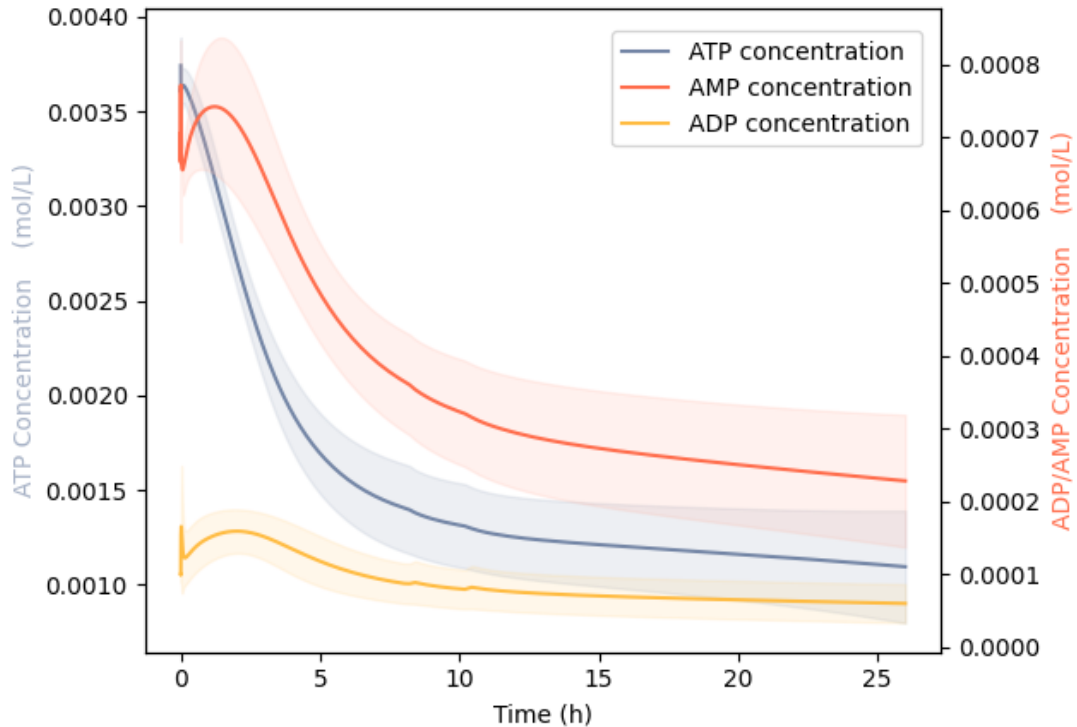
Another advantage of large-scale kinetic modeling is any information regarding the system's dynamic response to such changes. In this particular case, all dynamic responses showcase an initial increase in biomass production, which in some cases surpasses the original steady-state value. This effect is exaggerated with models 1 and 5, which quickly reach values higher than  $0.14 \text{ h}^{-1}$ , a concerning, one order of magnitude increase. Such high initial deviations indicate drastic responses of the cell's metabolic network and may be linked with toxic responses of the tumor or aggravation of ovarian cancer. These observations should be taken into consideration when developing a drug and deciding on drug dosages, as toxicity levels may inhibit the patient's successful treatment. An extensive collection of kinetic models could give a more statistically significant dynamic response that would better encapsulate the drawbacks of such perturbations.

### ATPM

ATPM, while not belonging to one of the metabolic pathways discussed above, had the second highest control coefficient value and the highest negative value. This means that an increase in the enzyme concentration will lead to a decrease in biomass production which may be because the ATP levels of the system will be depleted (Figure 3.3B). The simulated perturbations of  $v_{max} = 1.1v_{max,original}$  &  $v_{max} = 1.5v_{max,original}$  &  $v_{max} = 2v_{max,original}$  are showcased in Figure 3.5. Similar observations can be made as in HEX1 results, with the biomass flux steadily decreasing as the perturbation increases. In the case of a 100% perturbation, the deviations from the original biomass flux value range from 39% to 74%. Since, in the case of ATPM, the  $v_{max}$  value is increased, further perturbations e.g., 200%, 1000%, can be imposed, and a further decrease in biomass production is expected. Since the time evolutions of the system contain information about all concentrations and fluxes, it is possible to observe how the ATP/ ADP/ AMP levels are affected. Figure 3.6 showcases the concentrations of these metabolites for a 100% perturbation. As expected, all metabolites suffer a decrease in their concentration values (ATP 71%, ADP 66%, and AMP 40% relative decrease). The destabilization of the adenylate levels has significant effects on the metabolic network, which forces the cell to operate at slower rates.



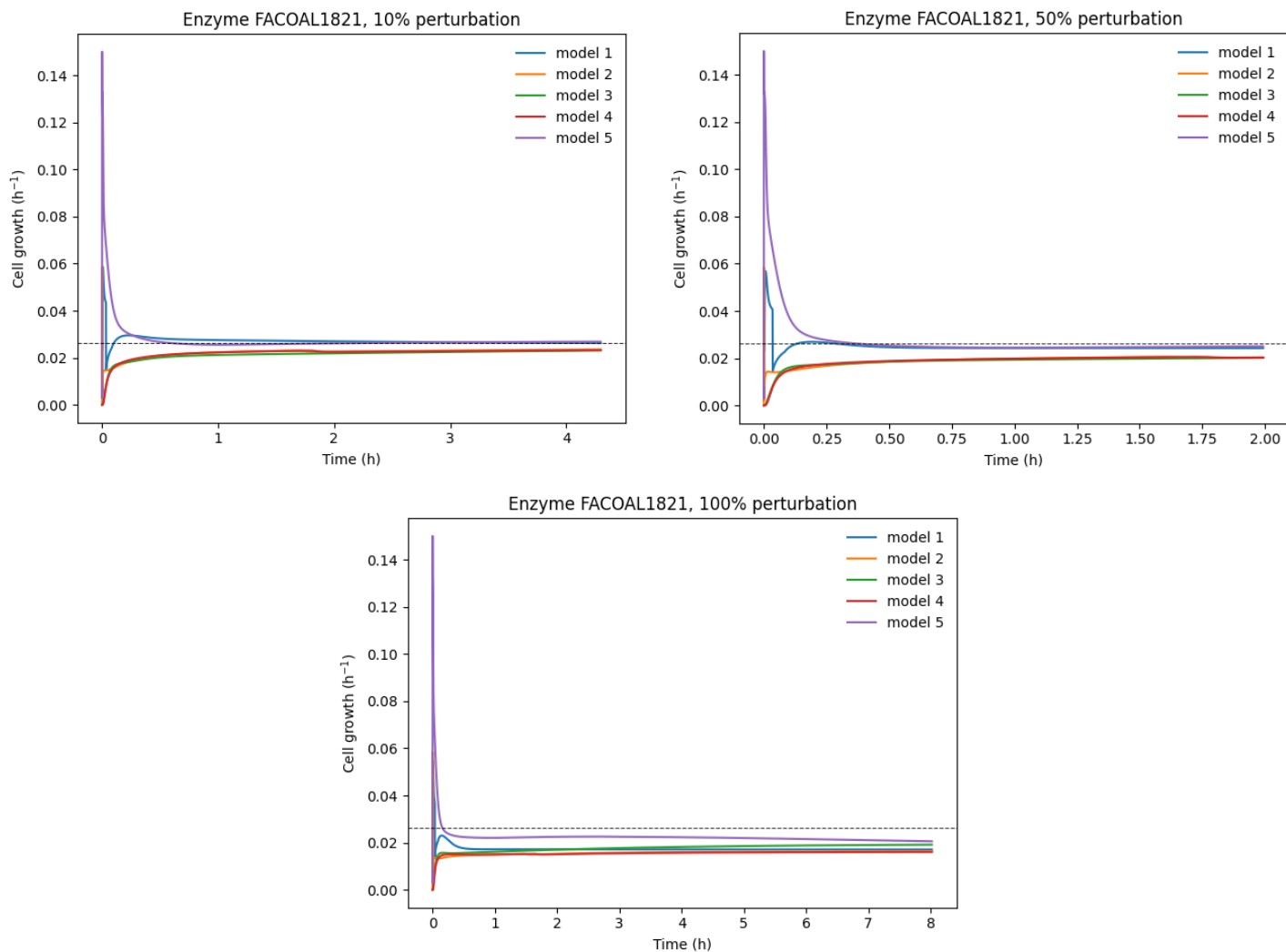
**Figure 3.5** Time graphs of cell growth for each model and each enzyme perturbation of ATPM. The grey dashed line indicates the original biomass flux value at steady-state.



**Figure 3.6** Concentration levels for ATP, ADP and AMP when the system is imposed to 100% ATPM enzyme perturbation. The filler lines represent the standard deviation between the 5 kinetic models. ATP has a different axis than ADP/AMP due to two orders of magnitude differences in concentration values.

### FACOAL1821

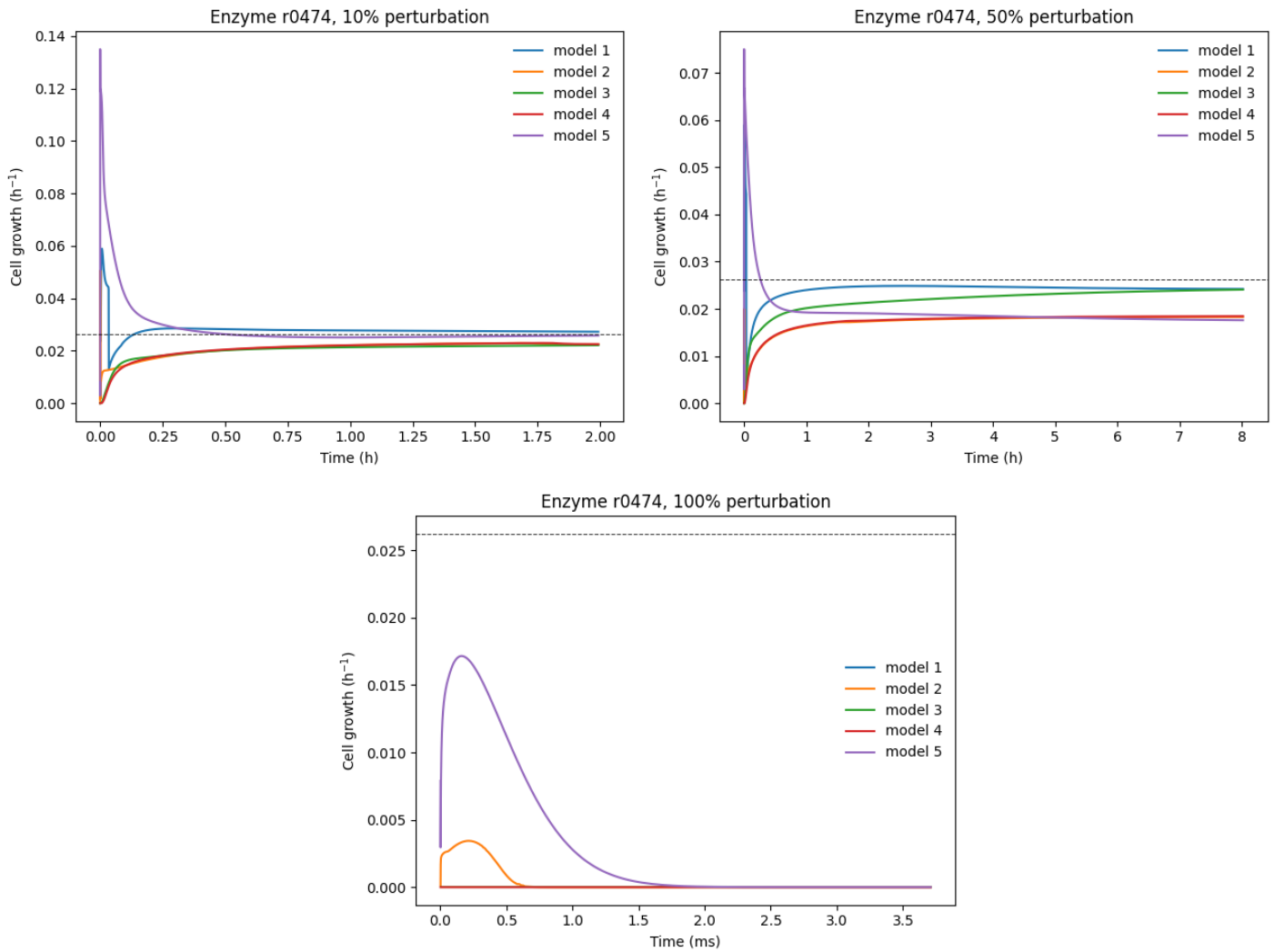
To study the effects of a deregulated lipid metabolism on the cell's proliferation, FACOAL1821 was chosen to be perturbed. The results follow similar patterns as the previous two cases and are portrayed in Figure 3.7. However, it is evident that the effects on biomass production are less significant. For a 10% perturbation, biomass production decreases by only 5% and for the gene knockout simulation, the decrease percentiles range from 25% to 39%. These results indicate that the control coefficient values can also estimate how impactful the perturbation effect is. While HEX1 and ATPM had high control coefficients and drastically affected biomass production, FACOAL1821, whose control coefficient value is less than half of theirs, did not impact biomass production to the same magnitude.



**Figure 3.7** Time graphs of cell growth for each model and each enzyme perturbation of FACOAL1821. The grey dashed line indicates the original biomass flux value at steady-state.

### r0474

Lastly, r0474 is part of the nucleotide metabolism and has a negative control coefficient. The time evolutions of the simulated perturbations are shown in Figure 3.8. This enzyme gives intriguing results since it would be expected that the impact on biomass production would be even more trivial than FACOAL1821's effect (6 spots down in MCA results). While a 10% perturbation does not significantly affect cellular growth, 50% perturbation displays a 21% average decrease in biomass production and 100% perturbation leads to cell death. Furthermore, cell death happens in a matter of milliseconds. For these reasons, r0474 can be considered essential for the cell's survival. Additionally, this case shows that MCA can provide a set of the candidate drug targets but that further simulations are required to investigate the complex interactions inside the metabolic networks and confirm *in silico* these targets.



**Figure 3.8** Time graphs of cell growth for each model and each enzyme perturbation of r0474. The grey dashed line indicates the original biomass flux value at steady-state. The results are shown in different timescales (h, ms) depending on the intensity of the perturbation.

## 4. Conclusions

This work represents, based on the writer's known literature, the first kinetic model of a metabolic network of this scale and was used to reliably describe the physiology of ovarian cancer cells. Thermodynamic, transcriptomic, metabolomic, and fluxomic data were integrated into kinetic models that can simulate metabolic responses of ovarian cancer's metabolism. The directionality for some remaining bidirectional reactions was decided based on a modified ACHR sampling algorithm. A set of stable sets of kinetic parameters were sampled that could also operate in physiologically relevant timeframes. This was achieved after imposing thermodynamic curations and information extraction from modal analysis results. The sets of kinetic parameters were used to find possible drug targets through Metabolic Control Analysis, and four enzymes of interest were further examined. Simulations of enzyme perturbations indicated that the produced kinetic models could give vital information about the dynamic response of the system as well as the impact on steady-state biomass production. Enzymes HEX1 (Hexokinase) and ATPM (ATP maintenance requirement) significantly affected the cell's survival and FALCOAL1821 (Long-chain-fatty-acid—CoA ligase) showcased similar results, although to a lesser extent. Particular attention should be given to r0474 (enzyme Ribonucleoside diphosphate reductase dCTP-forming), which delivered promising outcomes, despite its relatively small control coefficient. A ligand that explicitly targets this enzyme and not the whole nucleotide metabolism, as it is done with most developed drugs, may provide a physiology-specific treatment.

Overall, the methods used and developed for the current thesis produced large-scale metabolic models. The results highlight the importance of using such methods, which unveiled groups of enzymes that are likely to serve as drug targets, intricate dynamic responses of the whole network and quantifiable estimations of the overall effects on cancer proliferation. Using such models can assist the study of the physiology of specific types of cancer cells and can accelerate the drug development process. Physiology-specific drug targets can be suggested, and early estimations of the toxicity response and effective drug dosages can be calculated.

Of course, even though this thesis has demonstrated results that confirm the robustness of the methods used, there is still room for improvement and further research exploration. Firstly, it is critical to sample more sets of kinetic parameters with the desired characteristics and use them to acquire statistically more diverse simulations. While this will increase the computation time proportionally to the added number of kinetic models, it has the potential to deliver greater overall results. Specifically, more kinetic models will strengthen the statistical significance of the results and rule out outliers, such as one of the six kinetic models that were sensitive to any perturbation. In general, metabolic responses reported consistently within the population of models may prove to be more reliable predictors of cancer physiology. Secondly, it would be very intriguing to explore different states of the metabolic system by picking a different FDP. Since the uncertainty for many reactions' directionality is high, it would be interesting to explore the uniqueness of the thesis' findings. Other sets of FDPs and their subsequent kinetic models can produce similar or unexpected outcomes. This does not mitigate the presented findings, since they are backed up by relevant literature, but it serves as a way to explore the complex interactions inside the cell. Lastly, the methods developed for the enzyme perturbation simulations allow multiple inhibitions to happen simultaneously. This might be a vital tool for further investigation of suggested treatments.

It can aid in exploring the combinatorial drug targeting approach by providing vital information on how newly discovered drug targets respond when introduced in parallel with existing chemotherapeutic treatment.

In order to combat each type of cancer as a different disease, researchers will need to combine experimental data and *in silico* experiments to minimize the cost, time, and risk associated with the drug development process. In that regard, the models and methods developed within this project pave a way for advancing researchers' understanding of cancer metabolism and allow for the development of models for further research of other types of cancers.

## 5. Bibliography

- (1) WHO Cancer fact-sheet. <https://www.who.int/news-room/fact-sheets/detail/cancer> (accessed 2022-08-01).
- (2) Coward, J. I. G.; Middleton, K.; Murphy, F. New Perspectives on Targeted Therapy in Ovarian Cancer. *Int J Womens Health* **2015**, *7*, 189–203. <https://doi.org/10.2147/IJWH.S52379>.
- (3) Jayde, V.; Boughton, M. The Diagnostic Journey of Ovarian Cancer: A Review of the Literature and Suggestions for Practice. *Contemporary Nurse*. April 2012, pp 5–17. <https://doi.org/10.5172/conu.2012.41.1.5>.
- (4) Yee, C.; Dickson, K. A.; Muntasir, M. N.; Ma, Y.; Marsh, D. J. Three-Dimensional Modelling of Ovarian Cancer: From Cell Lines to Organoids for Discovery and Personalized Medicine. *Frontiers in Bioengineering and Biotechnology*. Frontiers Media S.A. February 10, 2022. <https://doi.org/10.3389/fbioe.2022.836984>.
- (5) Hannah Ritchie; Max Roser. *Causes of death*. <https://ourworldindata.org/causes-of-death> (accessed 2022-08-01).
- (6) *2019 Global Burden of Disease study database*. <https://vizhub.healthdata.org/gbd-results/> (accessed 2022-08-01).
- (7) Hamilton, A. B. *Psychological Aspects of Ovarian Cancer*; 1999; Vol. 17.
- (8) American Cancer Society. *Ovarian Cancer Risk Factors*. <https://www.cancer.org/cancer/ovarian-cancer/causes-risks-prevention/risk-factors.html> (accessed 2022-08-01).
- (9) Hyun, M.-S.; Joon Shin, S.; Ho Chun, S.; Ok Kim, K.; Kyoung Kim, M.; Hee Lee, K.; Soo Hyun, M.; Hwa Bae, S.; Mo Ryoo, H.; Rok Do, Y.; Young Kwon, K.; Suk Song, H. *The Efficacy of Paclitaxel and Cisplatin Combination Chemotherapy for the Treatment of Metastatic or Recurrent Gastric Cancer: A Multicenter Phase II Study*; 2005; Vol. 20.
- (10) Wang Dong; J. Lippard Stephen. CELLULAR PROCESSING OF PLATINUM ANTICANCER DRUGS. **2005**. <https://doi.org/10.1038/nrd1691>.
- (11) Graham Patrick. *An Introduction to Medicinal Chemistry*; 2017.
- (12) DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W. Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs. *J Health Econ* **2016**, *47*, 20–33. <https://doi.org/10.1016/J.JHEALECO.2016.01.012>.

- (13) Madhamshettiwar, P. B.; Maetschke, S. R.; Davis, M. J.; Reverter, A.; Ragan, M. A. Gene Regulatory Network Inference: Evaluation and Application to Ovarian Cancer Allows the Prioritization of Drug Targets. *Genome Med* **2012**, *4*, 41. <https://doi.org/10.1186/gm340>.
- (14) Pacheco, M. P.; Bintener, T.; Ternes, D.; Kulms, D.; Haan, S.; Letellier, E.; Sauter, T. Identifying and Targeting Cancer-Specific Metabolism with Network-Based Drug Target Prediction. *EBioMedicine* **2019**, *43*, 98–106. <https://doi.org/10.1016/J.EBIOM.2019.04.046>.
- (15) Zhang, C.; Hua, Q. Applications of Genome-Scale Metabolic Models in Biotechnology and Systems Medicine. *Front Physiol* **2016**, *6* (JAN), 413. <https://doi.org/10.3389/FPHYS.2015.00413/XML/NLM>.
- (16) Thiele, I.; Swainston, N.; Fleming, R. M. T.; Hoppe, A.; Sahoo, S.; Aurich, M. K.; Haraldsdottir, H.; Mo, M. L.; Rolfsson, O.; Stobbe, M. D.; Thorleifsson, S. G.; Agren, R.; Bölling, C.; Bordel, S.; Chavali, A. K.; Dobson, P.; Dunn, W. B.; Endler, L.; Hala, D.; Hucka, M.; Hull, D.; Jameson, D.; Jamshidi, N.; Jonsson, J. J.; Juty, N.; Keating, S.; Nookaew, I.; le Novère, N.; Malys, N.; Mazein, A.; Papin, J. A.; Price, N. D.; Selkov, E.; Sigurdsson, M. I.; Simeonidis, E.; Sonnenschein, N.; Smallbone, K.; Sorokin, A.; van Beek, J. H. G. M.; Weichart, D.; Goryanin, I.; Nielsen, J.; Westerhoff, H. v.; Kell, D. B.; Mendes, P.; Palsson, B. O. A Community-Driven Global Reconstruction of Human Metabolism. *Nat Biotechnol* **2013**, *31* (5), 419–425. <https://doi.org/10.1038/nbt.2488>.
- (17) Masid, M.; Ataman, M.; Hatzimanikatis, V. Analysis of Human Metabolism by Reducing the Complexity of the Genome-Scale Models Using RedHUMAN. *Nat Commun* **2020**, *11* (1), 1–12. <https://doi.org/10.1038/s41467-020-16549-2>.
- (18) Orth, J. D.; Thiele, I.; Palsson, B. O. What Is Flux Balance Analysis? *Nat Biotechnol* **2010**, *28* (3), 245–248. <https://doi.org/10.1038/nbt.1614>.
- (19) Henry, C. S.; Broadbelt, L. J.; Hatzimanikatis, V. Thermodynamics-Based Metabolic Flux Analysis. *Biophys J* **2007**, *92* (5), 1792–1805. <https://doi.org/10.1529/biophysj.106.093138>.
- (20) Jankowski, M. D.; Henry, C. S.; Broadbelt, L. J.; Hatzimanikatis, V. Group Contribution Method for Thermodynamic Analysis of Complex Metabolic Networks. *Biophys J* **2008**, *95* (3), 1487–1499. <https://doi.org/10.1529/BIOPHYSJ.107.124784>.
- (21) Lewis, N. E.; Hixson, K. K.; Conrad, T. M.; Lerman, J. A.; Charusanti, P.; Polpitiya, A. D.; Adkins, J. N.; Schramm, G.; Purvine, S. O.; Lopez-Ferrer, D.; Weitz, K. K.; Eils, R.; König, R.; Smith, R. D.; Palsson, B. Omic Data from Evolved E. Coli Are Consistent with Computed Optimal Growth from Genome-Scale Models. *Mol Syst Biol* **2010**, *6* (1), 390. <https://doi.org/10.1038/MSB.2010.47>.

- (22) Pandey, V.; Hadadi, N.; Hatzimanikatis, V. Enhanced Flux Prediction by Integrating Relative Expression and Relative Metabolite Abundance into Thermodynamically Consistent Metabolic Models. *PLoS Comput Biol* **2019**, *15* (5), 1–23. <https://doi.org/10.1371/journal.pcbi.1007036>.
- (23) Kaufman, D. E.; Smith, R. L. Direction Choice for Accelerated Convergence in Hit-and-Run Sampling. <https://doi.org/10.1287/opre.46.1.84> **1998**, *46* (1), 84–95. <https://doi.org/10.1287/OPRE.46.1.84>.
- (24) Fallahi, S.; Skaug, H. J.; Alendal, G. A Comparison of Monte Carlo Sampling Methods for Metabolic Network Models. *PLoS One* **2020**, *15* (7), e0235393. <https://doi.org/10.1371/JOURNAL.PONE.0235393>.
- (25) Miskovic, L.; Hatzimanikatis, V. Production of Biofuels and Biochemicals: In Need of an ORACLE. *Trends Biotechnol* **2010**, *28* (8), 391–397. <https://doi.org/10.1016/j.tibtech.2010.05.003>.
- (26) Salvy, P.; Fengos, G.; Ataman, M.; Pathier, T.; Soh, K. C.; Hatzimanikatis, V. PyTFA and MatTFA: A Python Package and a Matlab Toolbox for Thermodynamics-Based Flux Analysis. *Bioinformatics* **2019**, *35* (1), 167–169. <https://doi.org/10.1093/bioinformatics/bty499>.
- (27) Bruand, M.; Barras, D.; Mina, M.; Ghisoni, E.; Morotti, M.; Lanitis, E.; Fahr, N.; Desbuisson, M.; Grimm, A.; Zhang, H.; Chong, C.; Dagher, J.; Chee, S.; Tsianou, T.; Dorier, J.; Stevenson, B. J.; Iseli, C.; Ronet, C.; Bobisse, S.; Genolet, R.; Walton, J.; Bassani-Sternberg, M.; Kandalaft, L. E.; Ren, B.; McNeish, I.; Swisher, E.; Harari, A.; Delorenzi, M.; Ciriello, G.; Irving, M.; Rusakiewicz, S.; Foukas, P. G.; Martinon, F.; Dangaj Laniti, D.; Coukos, G. Cell-Autonomous Inflammation of BRCA1-Deficient Ovarian Cancers Drives Both Tumor-Intrinsic Immunoreactivity and Immune Resistance via STING. *Cell Rep* **2021**, *36* (3). <https://doi.org/10.1016/j.celrep.2021.109412>.
- (28) de La Fuente, I. M.; Cortés, J. M.; Valero, E.; Desroches, M.; Rodrigues, S.; Malaina, I.; Martínez, L. On the Dynamics of the Adenylate Energy System: Homeorhesis vs Homeostasis. *PLoS One* **2014**, *9* (10), e108676. <https://doi.org/10.1371/JOURNAL.PONE.0108676>.
- (29) Park, J.; Rubin, S.; Xu, Y.-F.; Amador-Noguez, D.; Fan, J.; Shlomi, T.; Rabinowitz, J. Metabolite Concentrations, Fluxes and Free Energies Imply Efficient Enzyme Usage. *Nat Chem Biol* **2016**. <https://doi.org/10.1038/nchembio.2077>.
- (30) King, Z. A.; Dräger, A.; Ebrahim, A.; Sonnenschein, N.; Lewis, N. E.; Palsson, B. O. Escher: A Web Application for Building, Sharing, and Embedding Data-Rich

Visualizations of Biological Pathways. *PLoS Comput Biol* **2015**, *11* (8), e1004321. <https://doi.org/10.1371/JOURNAL.PCBI.1004321>.

- (31) Ganapathy-Kanniappan, S.; Geschwind, J. F. H. Tumor Glycolysis as a Target for Cancer Therapy: Progress and Prospects. *Molecular Cancer* *2013* *12:1* **2013**, *12* (1), 1–11. <https://doi.org/10.1186/1476-4598-12-152>.
- (32) Bian, X.; Liu, R.; Meng, Y.; Xing, D.; Xu, D.; Lu, Z. Lipid Metabolism and Cancer. *Journal of Experimental Medicine* **2021**, *218* (1). <https://doi.org/10.1084/JEM.20201606/211616>.
- (33) Wu, H. liang; Gong, Y.; Ji, P.; Xie, Y. fan; Jiang, Y. Z.; Liu, G. yu. Targeting Nucleotide Metabolism: A Promising Approach to Enhance Cancer Immunotherapy. *Journal of Hematology & Oncology* *2022* *15:1* **2022**, *15* (1), 1–21. <https://doi.org/10.1186/S13045-022-01263-X>.

## 6. Appendix A

Rate law formulations used for the purposed of the kinetic model building. Note that these are some examples and that the model contains many different stoichiometries, so the respective rate law formulations change accordingly.

Reaction	Mechanism	Explicit rate law formulation
$A \rightleftharpoons B$	Reversible Michaelis-Menten	$v_r = \frac{v_{max,r} \left( \frac{[A]_r}{K_A^M} \left( 1 - \frac{[B]_r}{K_{eq} [A]_r} \right) \right)}{1 + \frac{[A]_r}{K_A^M} + \frac{[B]_r}{K_B^M}}$
$A + B \rightleftharpoons C + D$	Generalized reversible Hill	$v_r = \frac{v_{max,r} \frac{[A]_r [B]_r}{K_A^M K_B^M} \left( 1 - \frac{[C]_r [D]_r}{K_{eq} [A]_r [B]_r} \right) \left( \frac{[C]_r}{K_C^M} + \frac{[D]_r}{K_D^M} \right)^{hr-1}}{\left( 1 + \left( \frac{[A]_r}{K_A^M} + \frac{[C]_r}{K_C^M} \right)^{hr} \right) \left( 1 + \left( \frac{[B]_r}{K_B^M} + \frac{[D]_r}{K_D^M} \right)^{hr} \right)}$
$A \rightleftharpoons B + C + D$	Convenience	$v_r = \frac{v_{max,r} \frac{[A]_r}{K_A^M} \left( 1 - \frac{[B]_r [C]_r [D]_r}{K_{eq} [A]_r} \right)}{\left( 1 + \frac{[A]_r}{K_A^M} \right) + \left( 1 + \frac{[B]_r}{K_B^M} \right) \left( 1 + \frac{[C]_r}{K_C^M} \right) \left( 1 + \frac{[D]_r}{K_D^M} \right) - 1}$
$A \rightleftharpoons B + C$	Uni-Bi Reversible Hill	$v_r = \frac{v_{max,r} \frac{[A]_r}{K_A^M} \left( 1 - \frac{[B]_r [C]_r}{K_{eq} [A]_r} \right) \left( \frac{[A]_r}{K_A^M} + \frac{[B]_r [C]_r}{K_B^M K_C^M} \right)^{hr-1}}{1 + \left( \frac{[A]_r}{K_A^M} + \frac{[B]_r}{K_B^M} \right)^{hr} + \left( \frac{[A]_r}{K_A^M} + \frac{[C]_r}{K_C^M} \right)^{hr} + \left( \frac{[A]_r}{K_A^M} + \frac{[B]_r [C]_r}{K_B^M K_C^M} \right)^{hr} - 2 \left( \frac{[A]_r}{K_A^M} \right)^{hr}}$
$A \rightarrow B$	Irreversible Michaelis-Menten	$v_r = \frac{v_{max,r} \left( \frac{[A]_r}{K_A^M} \right)}{1 + \frac{[A]_r}{K_A^M}}$

## 7. Appendix B

*Bidirectional intracellular reactions along with the Subsystem they belong to, their reaction formula and the manually imposed directionality.*

Variable ID	Reaction name	Subsystem	Formula	Directionality
10805	ASPTA	Alanine and aspartate metabolism	akg_c + asp_L_c <=> glu_L_c + oaa_c	Reverse
9984	r0615	Arginine and proline metabolism	nad_c + 4hpro_LT_c <=> 3 h_c + nadh_c + 1p3h5c_c	Forward
9986	r0617	Arginine and proline metabolism	nadp_c + 4hpro_LT_c <=> 3 h_c + nadph_c + 1p3h5c_c	Reverse
11143	r0783	Cholesterol metabolism	nadp_r + lthstrl_r <=> 2 h_r + nadph_r + chlstol_r	Forward
9879	SUCOAS1m	Citric acid cycle	coa_m + gtp_m + succ_m <=> pi_m + gdp_m + succoa_m	Reverse
9935	r0081	Citric acid cycle	akg_m + ala_L_m <=> glu_L_m + pyr_m	Reverse
10797	ACONT	Citric acid cycle	cit_c <=> icit_c	Forward
10856	FUM	Citric acid cycle	h2o_c + fum_c <=> mal_L_c	Forward
10978	MDH	Citric acid cycle	nad_c + mal_L_c <=> 2 h_c + nadh_c + oaa_c	Reverse
10873	MTHFC	Folate metabolism	h2o_c + methf_c <=> 10fthf_c + 2 h_c	Forward
10874	MTHFD	Folate metabolism	nadp_c + mlthf_c <=> nadph_c + methf_c	Forward
10990	r0792	Folate metabolism	nad_c + 5mthf_c <=> 2 h_c + nadh_c + mlthf_c	Forward
11001	MTHFD2	Folate metabolism	nad_c + mlthf_c <=> nadh_c + methf_c	Forward
9514	ALATA_L	Glutamate metabolism	akg_c + ala_L_c <=> pyr_c + glu_L_c	Reverse
9690	GLUDxm	Glutamate metabolism	h2o_m + nad_m + glu_L_m <=> 2 h_m + akg_m + nadh_m + nh4_m	Forward
9691	GLUDym	Glutamate metabolism	h2o_m + nadp_m + glu_L_m <=> 2 h_m + nadph_m + akg_m + nh4_m	Forward
9929	r0021	Glutathione metabolism	nad_c + 2 gthrd_c <=> 2 h_c + nadh_c + gthox_c	Forward
9930	r0022	Glutathione metabolism	nad_m + 2 gthrd_m <=> 2 h_m + nadh_m + gthox_m	Forward
9846	PSSA1_hs	Glycerophospholipid metabolism	ser_L_c + pchol_hs_c <=> chol_c + ps_hs_c	Reverse
9971	r0480	Glycerophospholipid metabolism	amet_c + ethamp_c <=> 2 h_c + ahcys_c + HC00718_c	Reverse
10857	G3PD1	Glycerophospholipid metabolism	nad_c + glyc3p_c <=> 2 h_c + nadh_c + dhap_c	Forward
11108	CDIPTr	Glycerophospholipid metabolism	inost_c + cdpdag_hs_c <=> h_c + cmp_c + pail_hs_c	Forward
11111	PCHOLP_hs	Glycerophospholipid metabolism	h2o_c + pchol_hs_c <=> h_c + chol_c + pa_hs_c	Reverse
11112	RE3301C	Glycerophospholipid metabolism	h2o_c + ps_hs_c <=> h_c + ser_L_c + pa_hs_c	Reverse
9981	r0553	Glycine, serine, alanine, and threonine metabolism	h2o_c + nadp_c + betald_c <=> 3 h_c + nadph_c + glyb_c	Forward
10863	GHMT2r	Glycine, serine, alanine, and threonine metabolism	ser_L_c + thf_c <=> h2o_c + gly_c + mlthf_c	Forward
11067	HMR_7748	Glycolysis/gluconeogenesis	M01389_c <=> g6p_c	Reverse
11068	HMR_7749	Glycolysis/gluconeogenesis	M01389_c <=> f6p_c	Forward
9977	r0525	Lysine metabolism	h2o_m + nad_m + sacrcp_L_m <=> 2 h_m + nadh_m + glu_L_m + L2aadp6sa_m	Forward
10989	r0392	Miscellaneous	h2o_c + nad_c + glyald_c <=> 3 h_c + nadh_c + glyc_R_c	Reverse
9951	r0249	Pentose phosphate pathway	r5p_r <=> ru5p_D_r	Reverse
9996	r0784	Pentose phosphate pathway	nad_c + xytl_c <=> 2 h_c + nadh_c + xyly_L_c	Forward
10719	RPEc	Pentose phosphate pathway	2 ru5p_D_c <=> 2 xu5p_D_c	Forward
10859	G6PDH2r	Pentose phosphate pathway	nadp_c + g6p_c <=> 2 h_c + nadph_c + 6pgl_c	Forward
10896	PPM	Pentose phosphate pathway	r1p_c <=> r5p_c	Reverse
10909	RPE	Pentose phosphate pathway	ru5p_D_c <=> xu5p_D_c	Forward

10910	RPI	Pentose phosphate pathway	r5p_c <=> ru5p_D_c	Reverse
10913	TALA	Pentose phosphate pathway	g3p_c + s7p_c <=> f6p_c + e4p_c	Forward
10916	TKT1	Pentose phosphate pathway	xu5p_D_c + r5p_c <=> g3p_c + s7p_c	Forward
10917	TKT2	Pentose phosphate pathway	xu5p_D_c + e4p_c <=> g3p_c + f6p_c	Forward
10925	r0570	Pentose phosphate pathway	2dr1p_c <=> 2dr5p_c	Reverse
11099	RE3273C	Phosphatidylinositol phosphate metabolism	h2o_c + pail_hs_c <=> h_c + pa_hs_c + inost_c	Reverse
10904	PUNP5	Purine catabolism	pi_c + ins_c <=> hxa_c + r1p_c	Reverse
9586	D3AIBTm	Pyrimidine catabolism	pyr_m + 3aib_D_m <=> 2mop_m + ala_L_m	Reverse
9848	PYNP2r	Pyrimidine catabolism	pi_c + uri_c <=> ura_c + r1p_c	Forward
9972	r0483	Pyrimidine catabolism	akg_m + 3aib_D_m <=> 2mop_m + glu_L_m	Forward
10622	RE0453C	Pyrimidine synthesis	datp_c + dtdp_c <=> dadp_c + dttp_c	Leave BDR
10633	RE1530C	Pyrimidine synthesis	dgtp_c + duri_c <=> h_c + dgdp_c + dump_c	Leave BDR
10642	RE2954C	Pyrimidine synthesis	h_c + pep_c + dtdp_c <=> pyr_c + dttp_c	Leave BDR
9997	r0788	Sphingolipid metabolism	amet_c + HC00718_c <=> 2 h_c + ahcys_c + HC01842_c	Reverse
9998	r0789	Sphingolipid metabolism	amet_c + HC01842_c <=> 2 h_c + ahcys_c + cholp_c	Reverse
9959	r0381	Taurine and hypotaurine metabolism	h2o_c + nad_c + hyptaur_c <=> 2 h_c + nadh_c + taur_c	Forward
10630	RE1233C	Tryptophan metabolism	akg_c + Lkynr_c <=> h2o_c + glu_L_c + kynate_c	Forward
10795	ACACT1r	Tryptophan metabolism	2 accoa_c <=> coa_c + aacoa_c	Forward
10987	TYRTA	Tyrosine metabolism	akg_c + tyr_L_c <=> 34hpp_c + glu_L_c	Forward
9669	GACMTRc	Urea cycle	amet_c + gudac_c <=> 2 h_c + ahcys_c + creat_c	Forward
9819	ORNTArm	Urea cycle	akg_m + orn_m <=> glu_L_m + glu5sa_m	Reverse
10627	RE0691C	Urea cycle	4abutn_c <=> h2o_c + CE1944_c	Forward
10636	RE1897C	Urea cycle	nad_c + spmd_c <=> 2 h_c + nadh_c + CE5586_c	Reverse
10637	RE1898C	Urea cycle	CE5586_c <=> 13damp_c + CE1944_c	Reverse