

Latvijas Lauksaimniecības universitātē  
Informācijas tehnoloģiju fakultātē  
Datoru sistēmu katedra

Latvia University of Agriculture  
Faculty of Information Technologies  
Department of Computer Systems



**Agris Pentjušs**

**STEHIOMETRISKO MODEĻU IZVEIDES UN ANALĪZES  
PROCEDŪRAS METABOLISKAJAI INŽENIERIJAI**

**STOICHIOMETRIC MODELS CREATION AND ANALYSIS  
PROCEDURES FOR METABOLIC ENGINEERING**

**Promocijas darba**

**KOPSAVILKUMS**

Dr.sc.ing. grāda ieguvei Informācijas tehnoloģiju nozarē

**SUMMARY**

of the Thesis for acquiring Doctoral Degree in the field of Information  
Technologies

**Agris Pentjušs**

**Paraksts / Signature**

Jelgava 2016

# INFORMĀCIJA

**Darba izpildes vieta:** Latvijas Lauksaimniecības universitāte  
Informācijas tehnoloģiju fakultāte Datoru sistēmu katedra Lielā iela 2, Jelgava,  
Latvija.

**Eksperimentālā darba izpildes vieta:**

1. Latvijas Lauksaimniecības universitāte Informācijas tehnoloģiju fakultāte  
Datoru sistēmu katedra Lielā iela 2, Jelgava, Latvija.
2. Latvijas Universitāte Mikrobioloģijas un biotehnoloģijas institūts, Kronvalda  
bulvāris 4, Rīga, Latvija.

**Promocijas darba zinātniskais vadītājs:** Dr.oec., Asoc. prof. Aleksandrs Gailums.

**Darbs akceptēts** LLU Informācijas tehnoloģiju fakultātes Datoru sistēmu katedras  
paplašinātajā akadēmiskajā sēdē 2015. gada 29. aprīlī. Protokols Nr.4.

**Oficiālie recenzenti:**

1. Latvijas Lauksaimniecības universitātes profesore, Dr.biol. Ina Alsiņa;
2. Rīgas Tehniskās universitātes profesors, Dr. habil. sc. ing. Zīgurds  
Markovičs;
3. Oxford Brookse universitātes galvenais pētnieks, Dr. Mark Poolman.

**Promocijas darba aizstāvēšana** notiks LLU Informācijas tehnoloģiju nozares  
promocijas padomes atklātā sēdē 2016. gada 11. martā, Jelgavā, Lielajā ielā 2,  
Informācijas tehnoloģiju fakultātes 218. auditorijā, plkst. 11:00.

Ar promocijas darbu var iepazīties LLU Fundamentālajā bibliotēkā, Lielajā ielā 2,  
Jelgavā un tiešsaistē: [http://llufb.llu.lv/promoc\\_darbi.html](http://llufb.llu.lv/promoc_darbi.html).

**Atsaukmes sūtīt** Promocijas padomes sekretārei – Lielajā ielā 2, Jelgava, LV-3001;  
tālrunis 63005621; e-pasts [tatjana.tabunova@llu.lv](mailto:tatjana.tabunova@llu.lv). Atsaukmes vēlams sūtīt skenētā  
veidā ar parakstu.

**Padomes sekretāre** – LLU lektore Mg.paed. Tatjana Tabunova.

# SATURS

PROMOCIJAS DARBA APROBĀCIJA .....	5
IEVADS .....	7
Promocijas darba mērķis un uzdevumi .....	9
Pētījumu metodes .....	9
Zinātniskais jauninājums un praktiskā vērtība .....	10
Pētījuma tēzes .....	10
Promocijas darba struktūra un apjoms .....	10
1. BIOLOĢISKO DATU DAUDZVEIDĪBAS IETVERŠANA	
DATORMODELĪ .....	10
2. STEHIOMETRISKĀS MODELĒŠANAS ANALĪZES ALGORITMI .....	13
<i>Novirzīto</i> algoritmu lietojums .....	13
<i>Nenovirzīto</i> algoritmu lietojums .....	14
3. STEHIOMETRISKĀ MODEĻA PROJEKTĒŠANAS PROCEDŪRAS	
IZSTRĀDE UN IEROBEŽOJUMU IZMANTOŠANA .....	15
4. STEHIOMETRISKO MODEĻU ANALĪZES REZULTĀTU	
VIZUALIZĀCIJAS RĪKA IZSTRĀDE .....	17
5. STEHIOMETRISKO MODEĻU ANALĪZES PROCEDŪRU IZSTRĀDE .....	20
SECINĀJUMI .....	24
Galvenie darba rezultāti .....	24
Secinājumi un attīstības perspektīvas .....	24
LITERATŪRA .....	51

# TABLE OF CONTENTS

APPROBATION OF PHD THESIS .....	29
INTRODUCTION .....	31
The aim and tasks of PhD thesis .....	33
Research methods .....	33
Scientific novelty .....	34
Theses .....	34
PhD thesis structure and volume .....	34
1. INCORPORATION OF BIOLOGICAL DATA DIVERSITY IN COMPUTER MODEL .....	34
2. STOICHIOMETRIC MODELING ANALYSIS ALGORITHMS .....	37
Biased analysis algorithms .....	37
Unbiased analysis algorithms .....	38
3. DEVELOPMENT OF STOICHIOMETRIC MODEL DESIGNING AND CONSTRAINTS APPLICATION PROCEDURE .....	39
4. DEVELOPMENT OF STOICHIOMETRIC MODEL ANALYSIS VISUALIZATION TOOL .....	41
5. DEVELOPMENT OF STOICHIOMETRIC MODEL ANALYSIS ALGORITHMS APPLICATION PROCEDURE .....	43
CONCLUSIONS .....	48
Main results of PhD thesis .....	48
Conclusions and development prospects .....	50
BIBLIOGRAPHY .....	51

## PROMOCIJAS DARBA APROBĀCIJA

Pētījumu rezultāti ir atspoguļoti sekojošās publikācijās:

1. Odzina, I., Pentjuss, A. (2011) In silico analysis of steady state mechanisms of metabolic networks in COBRA Toolbox and FBA-SimVis. In: Proceedings of the 17th International Scientific Conference *Research for Rural Development*, Vol.17, pp. 196-201.
2. Odzina, I., Pentjuss, A. (2012) Algorithm of genome - scale metabolic engineering implementation. In: Proceedings of International Scientific Conference *Engineering for Rural Development*, Jelgava, Latvia, pp. 289-293.
3. Pentjuss, A., Odzina, I., Gailums, A. (2012) Multi agent genome-scale metabolic reconstruction modeling software schema. In: 11th International Scientific Conference on *Engineering for Rural Development*, Jelgava, pp. 189-194.
4. Pentjuss, A., Kalnenieks, U. (2013) Assessment of *Zymomonas mobilis* biotechnological potential in ethanol production by flux variability analysis, *Biosystems and Information Technology*, Vol. 3(1), pp. 1-5.
5. Pentjuss, A., Odzina, I., Kostromins, A., Fell, D.A., Stalidzans, E., Kalnenieks, U. (2013) Biotechnological potential of respiring *Zymomonas mobilis*: a stoichiometric analysis of its central metabolism. *Journal of Biotechnology*, Vol. 165(1), pp. 1-10.
6. Pentjuss, A., Rubenis, O., Bauze, D., Aprupe, L., Lace, B. (2013) Flux variability analysis approach of autism related metabolism in stoichiometric model of mitochondria. In: *Biosystems and Information Technology*, Vol. 2(2), pp. 37-42.
7. Kalnenieks, U., Pentjuss, A., Rutkis, R., Stalidzans, E., Fell, D.A. (2014) Modeling of *Zymomonas mobilis* central metabolism for novel metabolic engineering strategies, *Frontiers in Microbiology*, Vol. 5, pp. 42–49.

Pētījumos iegūtie rezultāti tika prezentēti sekojošās konferencēs:

1. Pentjuss, A. (2012) Algorithm of genome scale metabolic engineering implementation. In: *Applied Information and Communication Technology (AICT 2012)*, April 26-27, 2012, Jelgava, Latvia.
2. Pentjuss, A. (2012) Biotechnological potential of respiring *Zymomonas mobilis*: a stoichiometric analysis of its central metabolism. In: 2nd Conference on *Constraint-based Reconstruction and Analysis (COBRA 2012)*, October 7-9, 2012, Helsingør, Denmark.
3. Pentjuss, A. (2013) Bioķīmisko tīklu modelēšana dzīvos organismos. No: *LU un LMT Datorzinātņu dienas*, 2013. gada 5 -7. augusts, Jelgava, Latvija.
4. Pentjuss, A. (2014) Organisma metabolisma bioķīmisko tīklu analīzes un kalkulācijas metodes. No: Rēzeknes augstskolas Inženieru fakultātes 18.

starptautiskās studentu zinātniski praktiskās konferences *Cilvēks, Vide, Tehnoloģijas*, 2014. gada 23. aprīlis, Rēzekne, Latvija.

5. Pentjuss, A. (2014) Organismu metabolisma bioķīmisko tīklu analīzes metožu vizualizācija. No: Rēzeknes augstskolas Inženieru fakultātes 18. starptautiskās studentu zinātniski praktiskās konferences *Cilvēks, Vide, Tehnoloģijas*, 2014. gada 23. aprīlis, Rēzekne, Latvija.
6. Pentjuss, A. (2014), Central metabolism modeling based metabolic engineering of *Zymomonas mobilis*. In: *2nd Congress of Baltic Microbiologists*, October 2014. gada 16. oktobris, Tartu Igaunija.

Pētījumos iegūtie rezultāti tika prezentēti sekojošos semināros:

1. Pentjuss, A. (2014) Modeling the Entner - Doudoroff Pathway in *Z. mobilis* In: „*110th Latvia University of Agriculture Biosystems group seminar*”, May 20, 2014, Jelgava, Latvia.
2. Pentjuss, A. (2014) Tomāta metabolisma stehiometrijas un kinētikas modelēšana. No: „*108. Latvijas Lauksaimniecības universitātes Biosistēmu grupas semināra*”, 2014. gada 17. janvāris, Jelgava, Latvija.

#### **Dalība ar promocijas darbu saistītajos projektos:**

1. ERAF projekta “Netradicionālo raugu izmantošana bioetanola ražošanas uzlabošanai no inulīnu saturošām izejvielām – biotehnoloģiskā un sistēmbioloģijas pieeja” pētnieks Latvijas Universitātes Mikrobioloģijas un Biotehnoloģijas institūtā (2014.-2015.).
2. „Latvijas starpaugstskolu zinātniskās grupas izveide sistēmbioloģijā” Nr.2009/0207/1DP/1.1.1.2.0/09/APIA/VIAA/128, (2009.–2012.), pētnieks.

## IEVADS

Genotipa-fenotipa saistība ir svarīgs pētīšanas virziens dabaszinātnēs. Problēmas atrisināšana nozīmētu, ka, zinot genotipu (ģenētisko informāciju), ir iespējams paredzēt fenotipu (genotipa determinēto organismu ar tam raksturīgo fizioloģiju) un otrādi. 20. gadsimta otrajā pusē pieeja ģenētikai, biokīmijai un molekulārajai bioloģijai bija orientēšanās uz sarežģīto bioloģisko sistēmu sadalīšanu komponentēs cerībā izprast genotipa-fenotipa saistības. Šīs pieejas sniedza detalizētu izpratni par bioloģiskajām individuālajām komponentēm (metabolīts, reakcija), bet nerisināja sistēmisko mijiedarbību starp bioloģiskajām un vides komponentēm, kas nosaka fenotipu (Price, Qvarnström, & Irwin, 2003).

Mūsdienās tehnoloģiskie sasniegumi ļauj izmantot augstas caurlaidības metodes, lai vispusīgi un vienlaikus raksturotu bioloģiskās komponentes (Lewis, Nagarajan, & Palsson, 2012). Šādu datu veidošanas izmaksas ir strauji samazinājušās, un apstrādājamo datu apjomi ir kļuvuši daudz lielāki, kas ir atļāvuši pētniekiem skatīt un pētīt šūnas vai organismus kā veselas sistēmas mijiedarbības rezultātu (Duarte et al., 2007). Datortehnika un tehnoloģijas ir neizbēgami iesaistītas iepriekš minēto datu ieguvē. Tajā pašā laikā šo datu apjoms nosaka, ka ne tikai to iegūšanu, bet arī to analīzi un interpretāciju vairs nav iespējams veikt bez datortehnoloģijām. Līdz ar to cilvēks deleģē datoriem un tehnoloģijām analīzes un spriešanas funkcijas ar specifisku algoritmu palīdzību, atstājot sev datora sniegto risinājumu un viedokļu pārbaudi ar eksperimentāliem datiem un pārbaudītās likumsakarībās pārbaudītiem testiem.

Datormodelēšana ir viena no tehnoloģijām, kas ļauj analizēt mijiedarbībā esošu elementu veidotu sistēmu. Būtībā, bioloģisku sistēmu datormodelis apkopo publicēto un eksperimentos iegūto informāciju par pētāmo procesu vai objektu un lieto atzītās likumsakarības vienlaicīgi (paralēli). Līdz ar to visa modelī ievietotā informācija un zināšanas tiek pārbaudītas simulācijās. Bieži vien jāatzīst, ka zināšanas ir pretrunīgas un nepieciešami jauni eksperimenti. Modelēšanas eksperimentu (simulāciju) un bioloģisko eksperimentu cikls veido sistēmbioloģijas pieeju – ciklisku progresu procesa izpratnei matemātiskā līmenī (modelī). Izveidotais modelis tālāk izmantojams organisma darbības normalizācijā (tipisks uzdevums medicīnā) vai tā uzlabošanā pēc izvirzītajiem kritērijiem (tipisks uzdevums biotehnoloģijā).

Vielmaiņas jeb metabolisma procesi, iespējams, ir pašlaik vislabāk izpētītie, lielā mērā pateicoties labām iespējām metabolītu koncentrācijas kvantitatīvai mērīšanai un vielas nezūdamības likumam. Šie modeļi joprojām neatbild uz jautājumu par genotipa-fenotipa saistību vēlamajā līmenī, tomēr dod ļoti daudz vērtīgas informācijas par iespējamiem un neiespējamiem organismu stāvokļiem.

Metabolisma modelēšana, līdzīgi kā citu procesu modelēšana, ir iedalāma divos svarīgos posmos:

- 1) **modeļa izstrāde** – darbību kopums, lai iegūtu pietiekami precīzu objekta vai procesa matemātisku atainojumu un eksperimenti ar objektu būtu aizstājami ar datormodelēšanas eksperimentiem;

- 2) **modeļa analīze un/vai optimizācija** – darbību kopums, lai noskaidrotu sistēmiskās īpašības, kas nosaka organisma atbilstību izvirzītajam optimizācijas kritērijam.

Ir dažādas metabolisma modelēšanas tehnikas: stohastiskā kinētiskā, determinētā kinētiskā, stehiometriskā un citas. Starp šīm metodēm stehiometriskajai modelēšanai raksturīgi liela mēroga modeļi, jo to lietošanai ir nepieciešamas tikai no masas nezūdamības likuma viedokļa sabalansētas reakcijas. Stehiometriskā modelēšana var darboties ar modeļiem, kuros ir tūkstošiem reaģentu un tūkstošiem tos saistošo reakciju. Galvenie lietošanas virzieni ir medicīna un biotehnoloģija.

Pašlaik lielākais samērā precīzais stehiometriskais modelis ir cilvēka metabolisma modelis ar vairāk nekā 5000 reaģentiem un 7000 reakcijām starp tiem (Thiele et al., 2013). Caur reakcijas katalizējošajiem fermentiem modeļi ir iesaistīti arī tūkstošiem gēnu. Modelis jau tiek tieši lietots iedzimto vielmaiņas slimību diagnostikas pilnveidošanai, sniedzot arvien jaunas prognozes par slimību izraisītajiem blakusefektiem. Pašlaik attīstās arī jauna pieeja medicīnā klīniskajiem pētījumiem, pēc kuru veikšanas tiek pieļauta kādu medikamentu vai ierīču lietošana veselības aizsardzībā: datamodelēšanā balstītie klīniskie izmēģinājumi (*in silico clinical trials*) (<http://avicenna-isct.org/>). Ar šādu pieeju paveras iespējas agrīni atsijāt neperspektīvos risinājumus, kā arī noteikt pacientu grupas, kurām attiecīgā tehnoloģija nav izmantojama.

Arī biotehnoloģijā stehiometrisko modeļu lietojums sniedz lieliskus rezultātus. Tā populārais organisms *Escherichia coli* tika optimizēts biodegvielas etanola ražošanai (Trinh, Unrean, & Srienc, 2008), biodīzeļa ražošanas blakusprodukta glicerīna pārvēršanai etanolā (Trinh & Srienc, 2009), kā arī aminoskābju ražošanai (Unrean, Trinh, & Srienc, 2010). Minētie risinājumi ir apstiprināti arī praktiskos eksperimentos.

Pastāv arī vidēja un liela mēroga stehiometrisko modeļu izveides un optimizāciju lietošana veiksmes stāstos (Kerkhoven, Lahtvee, & Nielsen, 2014) (Liu et al., 2014).

Līdz ar stehiometrisko modeļu panākumiem un to lielo mērogu šie modeļi ievieš jaunas prasības programmproduktiem, īpaši situācijās, kad darbojas starpdisciplināras komandas, kurās piedalās gan bioloģiskajos, gan matemātikas un skaitļošanas speciālisti.

Reakciju sabalansēšana ir svarīgs, bet sarežģīts process (Thiele & Palsson, 2010), kuru metabolisko modeļu veidotāji bieži vien vienkāršo un optimizācijas rezultātā iegūst risinājumus, kas neatbilst vielas nezūdamības likumam un noved pie nepareiziem secinājumiem.

Cita problēma ir metabolisko tīklu vizualizācija, kas ļautu bioloģiem pietiekoši labi atpazīt procesus un ļautu novērtēt to iespējamo atbilstību realitātei – konkrēta organisma uzvedībai. Ideālā gadījumā vizualizācija ļautu īsā laikā novērtēt procesu, dažādu metabolisko tīklu atzaru mijiedarbību un iespējas ar tiem manipulēt.

Arvien pieaugoša problēma ir arī skaitļošanas jauda. Stehiometriskajos modeļos jēgpilnu uzdevumu skaitļošana var aizņemt mēnešus



(piemēram – elementāro ceļu kalkulācija). Līdz ar to ir nepieciešams uzlabot optimizācijas procedūras, lai stehiometriskie modeļi dotu arvien vairāk precīzas, praktiskajā biotehnoloģijā un medicīnā lietojamas informācijas ar iespējami mazāku skaitļošanas resursu un laika patēriņu.

Ļoti dinamiski attīstās arī bioloģiskās informācijas datu bāzes ar dažādu informāciju: reakcijas kodējošo gēnu esamību organismu genomos, informāciju par reakciju termodinamiskajām īpatnībām, kas nosaka reakcijas virzienu ierobežojumus, informāciju par gēnu saistību ar reakcijas katalizējošajiem gēniem utt.

Lai arī minētajos virzienos jau ir bijuši mēģinājumi uzlabot modelēšanas procesu, vēl ir lielas iespējas sniegtu uzlabot, jo stehiometrisko modeļu izveide un optimizācija ir samērā jauna nozare un turpina strauji attīstīties. Īpaši tas ir attiecināms uz darbībām ar liela mēroga, tajā skaitā genoma mēroga modeļiem apstākļos, kad informācijas daudzums pieaug ļoti strauji, pateicoties eksperimentālajām tehnoloģijām. Šī situācija nosaka pētījuma aktualitāti un formē promocijas darba mērķi un uzdevumus.

### **Promocijas darba mērķis un uzdevumi**

Promocijas darba mērķis ir stehiometrisko datormodeļu balstītas metaboliskās inženierijas uzdevuma procedūru izstrāde skaitļošanas apjoma un manuālā darba samazināšanai.

Pētījuma mērķa sasniegšanai tika izvirzīti šādi uzdevumi:

1. Analizēt matemātiskās likumsakarības starp stehiometrisko modelēšanu un metabolisma molekulārajiem procesiem.
2. Izpētīt stehiometrisko modeļu analīzes algoritmu darbības principus.
3. Uzlabot stehiometrisko modeļu projektēšanas un validācijas procedūru patērētā laika, darbību skaita un manuāli apstrādājamo datu apjoma samazināšanai.
4. Izveidot datu vizualizācijas rīku stehiometrisko modeļu analīzes algoritmiem iegūto rezultātu vizualizācijai metaboliskās inženierijas uzdevumos.
5. Izstrādāt stehiometrisko modeļu analīzes algoritmu procedūras ar samazinātu datu apstrādes apjomu, laiku, manuālo darbu un metaboliskās inženierijas uzdevumos analizējamo risinājumu kopu.

### **Pētījumu metodes**

Aprēķini stehiometriskā modelēšanā pamatā tiek veikti, izmantojot matricās ievadīto informāciju, kas interpretē bioloģiskā organisma vielmaiņas funkcijas. Matricu transponēšana un matricu algebriskās darbības tiek izmantotas, lai atrastu noteikumus, ar kuriem metaboliskais tīkls var atrasties līdzsvara jeb stacionārā stāvoklī.

Lineārās algebras metodes tiek izmantotas, lai no matricā ievadītās informācijas veidotu vektora veidā kopīgo atrisinājumu kopu, kurā var atrast stehiometriskā modeļa visu atrisinājumu kopu.

Eksperimentālo datu validācija stehiometriskā modelī tiek izmantota, salīdzinot eksperimentālos datus ar stehiometriskā modeļa rezultātiem.

Masas nezūdamības likums tiek izmantots reakciju balansēšanā un stacionārā stāvokļa noteikšanai.

Uz ierobežojumiem balstītas modelēšanas principu lietošana.

Vizualizācijas izveidošanai tiek izmantota XML faila struktūra.

### **Zinātniskais jauninājums un praktiskā vērtība**

1. Uzlabota stehiometrisko modeļu projektēšanas procedūra specifiska tipa modeļiem, uzdevumiem un pieejamajiem datiem.
2. Izstrādāts datu vizualizācijas rīks stehiometrisko modeļu reprezentācijas atvieglošanai.
3. Izstrādātas procedūras stehiometrisko modeļu analīzes algoritmiem datu apstrādes apjoma, laika un metaboliskās inženierijas uzdevumu analizējamo atrisinājumu kopas samazināšanai.

### **Pētījuma tēzes**

1. Agrīna ierobežojumu ieviešana stehiometrisko modeļu veidošanā ļauj samazināt validācijas procesā apstrādājamo datu apjomu.
2. Automatizēta moduļa stāvokļa vizualizācija reakciju plūsmu līmenī uzlabo metabolisko ceļu savstarpējās mijiedarbības analīzi.
3. Metaboliskās analīzes algoritmu kombinācijas samazina analizējamo risinājumu telpu un analizējamo risinājumu kopu.

### **Promocijas darba struktūra un apjoms**

Promocijas darbs ir uzrakstīts latviešu valodā, satur anotāciju, ievadu, 5 nodaļas, secinājumus, literatūras sarakstu un 12 pielikumus, kas apkopoti atsevišķā sējumā, tajā skaitā 15 tabulu, 44 attēlus, kopā 145 lappuses. Darbā izmantoti 196 literatūras avoti.

## **1. BIOĻĪSKO DATU DAUDZVEIDĪBAS IETVERŠANA DATORMODELĪ**

Katrs organisma veids (suga) sastāv no unikāla genoma, kurā ir arī kodētas organisma metabolismā iesaistītās reakcijas, metabolīti, to mijiedarbība, funkcionalitāte, to savstarpējā savienojamība.

Bioloģisku sistēmu datorvadībā vadāmās sistēmas datormodelis ir datorvadības izstrādes priekšnosacījums, jo katrai bioloģiskai sistēmai, neatkarīgi no tās mēroga (šūna, audi, organisms), ir sava vadības sistēma (bioloģiskā vadības sistēma - BVS), kas ir ģenētiski determinēta (Stalidzans, 2005).

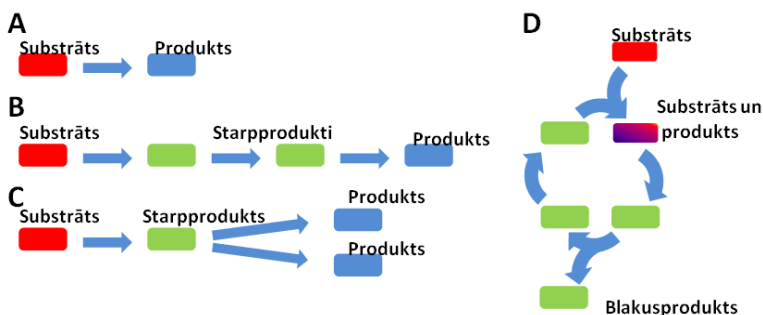
Optimālu datorvadības iespēju projektēšanu un izmantošanu pašlaik kavē tieši BVS nepietiekamais apraksts matemātiski formalizētā veidā datormodelī. BVS datormodelēšanai un optimizācijai ir noteicoša loma organisma sekmīgas

datorvadības projektēšanā un realizācijā. Tā kā katrai vielmaiņas reakcijai atbilst konkrēts gēns, ģenētiski determinēto organismu vielmaiņa būtībā ir BVS, kuras modelēšanai un optimizācijai ir veltīts šis promocijas darbs.

Metabolīti ir vielas, kas nodrošina metabolisma procesus organismā. Metabolītiem ir plašs uzdevumu klāsts, piemēram, kā organisma procesu degviela metabolisma norisei, šūnu, audu un pat orgānu struktūras veidošana. Matemātiski pastāv likumsakarība starp fermentu, metabolītu (substrāta un produkta) mijiedarbību.

Reakcijas plūsma (*metabolic flux* vai *flux*) (Edwards, Covert, & Palsson, 2002) jeb reakcijas ātrums ir matemātisks lielums, kurš parāda, ar kādu ātrumu substrāts tiek pārveidots produktā, izmantojot metaboliskos ceļus un reakcijas, kur šo procesu regulē fermenti.

Metaboliskie ceļi stehiometriskajā modelēšanā tiek veidoti no tajos iesaistītajām reakcijām, kur reakcijas pārveido substrātus produktos (skat. 1. A att.), kā arī starpproduktos. Metaboliskajā ceļā starpproduktu konversija veido nepārtrauktu reakciju virkni (skat. 1. B, C, D att.), ko sauc par metaboliskiem ceļiem. Metaboliskajos ceļos eksistē arī vienādi metabolīti (skat. 1. C, D att.), kurus konvertē dažādas reakcijas, līdz ar to veidojas sasaiste starp šiem ceļiem.



1. att. **Metaboliskā tīkla bioķīmisko ceļu veidu piemēri**

Avots: autora veidots

Katrai metaboliskajai reakcijai ir nepieciešama enerģija, lai fermenti katalizētu (aktivizētu) to. Šo enerģiju sauc par brīvo enerģiju šūnā, organismā. Kad reakcijā notiek process, tā izmaiņas brīvajā enerģijā ir vienādas ar brīvās enerģijas lielumu produktos mīnus brīvās enerģijas lielums substrātos. Matemātiski šo procesu izsaka ar šādu vienādojumu (Welch, 1985):

$$\Delta G = \Delta G_{\text{produkti}} - \Delta G_{\text{substrāti}} \quad (1)$$

$\Delta G$  - ir brīvā enerģija, kas nepieciešama, lai reakcijā uzsāktos konversijas process,

$\Delta G_{\text{produkti}}$  - ir kopējā Gibbs brīvās enerģijas summa produktiem,

$\Delta G_{\text{substrāti}}$  - ir Gibbs brīvās enerģijas summa substrātiem.

Organismā procesi cenšas notikt līdzsvarā, kā rezultātā pēc noteikta laika iestājas līdzsvara stāvoklis, kad nepieciešamā brīvā enerģija organisma, šūnas metabolismam tiek izveidota citos metabolisma procesos. Līdz ar to organismā metaboliskās reakcijas nodrošina stabilitu, nemainīgu metabolismu, kur nemainīgos daudzumos tiek patērēti substrāti un saražoti produkti. Tāpat norit starpproduktu konversija. Stacionārā stāvoklī organismā nemainās reakciju plūsmu vērtības un metabolītu koncentrācijas organismā.

Lai spētu apstrādāt strauji augošos daudzdimensionālos datu apjomus, ir vajadzība izstrādāt sarežģītas datu analīzes metodes. Dažādas metodes ir izstrādātas un lietotas, sākot no stohastiskiem kinētiskiem modeļiem (Golightly & Gillespie, 2013), Beijesa tīkliem (Wilkinson, 2007), stehiometriskiem modeļiem (Lewis et al., 2012) līdz pat grafu un mijiedarbības tīklu izstrādei un analīzei (Le Novère et al., 2009).

Promocijas darbā no vairākām modelēšanas metodēm tiek aplūkota tieši stehiometriskā (uz ierobežojumiem balstītā) modelēšana, kas tiek lietota genoma mērogu stehiometriskā modeļa (GSMR) analīzē. Rekonstruētais genoma mēroga metaboliskais tīkls ietver apvienotu un sistematizētu informāciju par zināmiem metabolītiem un enzimatiskām, spontānām reakcijām šūnā vai organismā.

	ATPGMD	etoh_tx	XI
x_glcD	1	0	0
glcD	0	0	0
x_xylD	0	1	0
xylD	0	0	-1

## 2. att. Stehiometriskās matricas piemērs metaboliskā tīkla reprezentācijai

Avots: autora veidots

Šī ir matrica (skat. 2. att.), kurā stabiņu veidā tiek reprezentēta metabolītu esamība noteiktās reakcijās (apzīmē 1 vai -1). Rindās matricā tiek reprezentētas reakcijas (skat. 2. att.). Pozitīvs skaitlis apzīmē, ka metabolīts tiek saražots reakcijā, negatīvs skaitlis nozīmē, ka metabolīts tiek patērēts šajā reakcijā, savukārt 0 apzīmē to, ka metabolīts nav iesaistīts reakcijā.

$$\begin{bmatrix} dS_1 \\ dt \\ dS_2 \\ dt \\ dS_3 \\ dt \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (2)$$

$S_1, S_2, S_3$ , – reakcijās izmantotie substrāti,

$v_1, v_2, v_3, v_4, v_5$ , - reakcijas ātrumi, ar kādu substrāti tiek pārvērsti produktos.

Formulas kreisajā pusē ir izteiktas fermentu koncentrācijas izmaiņas laikā.

## 2. STEHIOMETRISKĀS MODELĒŠANAS ANALĪZES ALGORITMI

Mūsdienās stehiometriskās modelēšanas analīžu algoritmi tiek iedalīti 2 lielās daļās: *novirzītie* un *nenovirzītie* algoritmi. *Nenovirzītie* algoritmi veic sistemātiskas visu risinājumu pārbaudes, bet *novirzītie* algoritmi meklē vienu no visiem risinājumiem. Promocijas darbā tika apskatīti plūsmu balansa analīzes (*flux balance analysis*, FBA) plūsmas variācijas analīze, (*flux variability analysis*, FVA) (Burgard, Vaidyaraman, & Maranas, 2001), metabolsikā plūsmu analīze (*metabolic flux analysis*, MFA) (Zhao & Shimizu, 2003), kas ir *novirzītie algoritmi*. Savukārt elementārie ceļi (*elementary modes*, EM) (S Schuster, Dandekar, & Fell, 1999), fermentu apakškopas (*enzymatic subsets*, ES) (Pfeiffer, Sánchez-Valdenebro, Nuño, Montero, & Schuster, 1999), ekstrēmie metaboliskie ceļi (*extreme pathway*, EP) (Yeung, Thiele, & Palsson, 2007) analīzes algoritmi tika apskatīti kā *nenovirzītie* algoritmi.

### **Novirzīto algoritmu lietojums**

*Novirzītie* algoritmi izmanto stehiometrisko matricu un bioloģiskos ierobežojumus, lai ar vektoru aprēķiniem analizētu risinājumu telpu. Paši par sevi *novirzītie* algoritmi bez GSMR un tās ierobežojumiem sniedz nepilnīgus un neprecīzus analīzes datus no risinājumu telpas, jo daļa atrisinājumu šajā telpā bioloģiski, fiziski un ķīmiski nespēj dabā pastāvēt. Tiek ieviesti papildu bioloģiskie ierobežojumi, lai izslēgtu no GSMR risinājumu telpas dabā neeksistējošos metabolisma stāvokļus.

Omikas dati ir dažādi eksperimentāli iegūti dati par organismu vai šūnu grupu. Tajos var ietilpt tādi dati kā genomikas (Barrett et al., 2013), proteomikas (Gasteiger, 2003), transkriptomikas (Craigon et al., 2004), metabolomikas (Xia, Mandal, Sinelnikov, Broadhurst, & Wishart, 2012), epigenomikas (Bernstein et al., 2010), plūsmu omikas (Bernstein et al., 2010) dati. Stehiometrisko modeļu analīzē un optimizācijās ir pieļaujams vienlaikus integrēt vairākus omikas datu veidus (Hyduke, Lewis, & Palsson, 2013).

Pastāv vēl viens bioloģisks ierobežojums GSMR – biomasas reakcijas vienādojums (*biomass composition*, BVR) (Gombert, Moreira dos Santos, Christensen, & Nielsen, 2001). Skaitliskās vērtības pie metabolītiem norāda to daudzumu, kas nepieciešamas, lai izveidotu vienu vienību biomasas, kurā tiek ietverti visi metabolītu daudzumi, kas nepieciešami, lai šūna attīstītos, augtu, veiktu dalīšanās procesu un vairotos.

Katrai reakcijai ir noteikts ierobežojumu diapazons plūsmu vērtībām. Šādi ierobežojumi ieviesti GSMR nosaka atrisinājumu telpas ierobežošanu stehiometriskā modelī, kas samazina analīžu algoritmu apstrādājamo atrisinājumu kopu. Pastāv vairāki veidi, kā bioloģiskos datus ievietot modelī kā ierobežojumus:

- substrāta uzņemšanas plūsmas ātrumu ierobežojumu iestatīšana;

- gēnu delēcijas ierobežojumu iestatīšana;
- omikas datu ierobežojumu iestatīšana;
- produkta ražošanas ātrumu plūsmu iestatīšana.

Stehiometriskās modelēšanas analīzes algoritmi pārsvarā modificē reakciju plūsmu robežas GSMR un stehiometriskā modelī ar analīzēm un optimizācijām.

Plūsmu balansa analīzes (FBA) algoritma mērķis ir izvirzīt prognozētu reakciju plūsmu sadalījumu organisma metabolismā, kas atbilst ievadītiem eksperimentālajiem datiem - ierobežojumiem. FBA algoritms tiek lietots, lai noteiktu izveidotā organisma modeļa spējas prognozēt konkrētā organisma fenotipiskās īpašības kā plūsmu sadalījumu, kas tiks iegūtas pie noteiktiem organisma stehiometriskā modelī ieviestajiem ierobežojumiem. (Edwards et al., 2002). Atrisinājuma telpā rezultāts ir izvirzīto kritēriju un ierobežojumu optimums. Pastāv liels modelī neatrasto alternatīvo risinājumu skaits.

Metaboliskās plūsmas analīzes (MFA) algoritma datu kopumi (data sets) (Joyce & Palsson, 2006), kas tiek iegūti plūsmu omikas eksperimentos, tiek izmantoti MFA analīzes algoritmu apstrādes procesā. Metabolisko tīklu plūsmu datu apstrādē un analīzē izmanto FBA. Atšķirībā no FBA analīzes algoritma MFA izmanto eksperimentos iegūtās reakciju plūsmu vērtības, un kā papildu ierobežojumus GSMR ievieto plūsmu robežas konstantās vērtības. Rezultātā atrisinājumu telpa tiek strikti ierobežota un atbilstošu variantu skaits krasi samazinās.

Plūsmu variācijas analīzes (FVA) algoritmam izmanto visus ieviestos GSMR ierobežojumus un izvirzīto mērķa funkciju, lai modelī simulētu katras reakcijas minimālās un maksimālās plūsmu vērtības. FVA var tikt izmantota konkrētam organisma datorsimulāciju fenotipam. Izmainoties kaut vienam ierobežojumam FVA ir jāpārkalkulē, jo analīzes rezultāti var atšķirties.

## ***Nenovirzīto algoritmu lietojums***

*Nenovirzītie algoritmi* tiek lietoti tad, kad nepieciešams izanalizēt nevis kādu no iespējamiem atrisinājumiem, bet gan, kad interesē visa nultās telpas atrisinājumu kopas analīze stehiometriskajā modelī. Pati par sevi nultās telpas vektoru kopa ir vektoru ierobežota kopa atrisinājumu telpā, kurā var atrast jebkuru modeļa izvirzīto atrisinājumu kā stacionāro stāvokli. Pašu atrisinājumu kopu izmanto citos analīzes algoritmos, lai veiktu sistemātiskus un visas nultās telpas atrisinājumu analīzi (Poolman, Sebu, Pidcock, & Fell, 2007). *Nenovirzītie algoritmi* izmanto atrisinājumu telpu ar nultās telpas vektoriem un stehiometrisko matricu, lai veiktu stehiometriskā modeļa optimumu analīzi.

Matemātiski EM nozīmē, ka stacionāro stāvokļu reakciju ātrumu vektoriem (reakciju plūsmu vērtībām) ir jāatrodas stehiometriskās matricas atrisinājumu telpā (Reder, 1988). Tā tiek uzskatīta kā objekts, kuras robežas ierobežo ar nultās telpas vektoriem. Tika definēts, ka EM ir pareiza bāzes vektoru izvēle nultā telpā (Newsholme, Challiss, & Crabtree, 1984), (Fell, 1993).

Visu iespējamo metabolisma stacionāro stāvokļu nultās telpas vektorus aprēķina šādi (Fell, 1992), (Leiser & Blum, 1987) :

$$\frac{dX}{dt} = v \quad (3)$$

v - reakcijas plūsmu ātruma vektors,

X - koncentrācijas vektors.

Stacionārā stāvokļa noteikumu definē algebriskais vienādojums (Stefan Schuster & Hilgetag, 1994) :

$$N * v = 0 \quad (4)$$

N - ir stehiometriskā matrica,

N dimensijas nosaka metabolītu un reakciju daudzums organismā. EM ir mazākā iespējamā metabolītu kopa stacionārā stāvoklī, kur ir vismaz viens substrāts un viens produkts, un nevar būt vairāki stacionārie stāvokļi. EM ir unikāla kopa, kurā nav nevienas citas apakškopas (Pfeiffer et al., 1999).

EP ir visīsākā iespējamā EM atrisinājuma kopa. EP algoritmu izmanto, lai noteiktu metaboliskā tīkla spēju ar mazāko iesaistīto reakciju un fermentu daudzumu nodrošināt metabolismu.

Gēnu ekspressijas, metaboliskās inženierijas vai metabolisko regulāciju pētījumos ir liela interese atrast fermentu grupas, kas visos stacionāros stāvokļos spēj operēt ar vienādu reakciju plūsmu proporciju. Šādu fermentu grupu sauc par fermenta apakškopām (ES). Tas apzīmē fermentu grupu, kas nodrošina metabolisko vadību nevis pār vienu, bet pār vairākiem fermentiem saskaņotā veidā. Stehiometriskā modelēšanā ES lieto tad, ja reakcijas plūsmas izmainās, bet apakškopas plūsmu proporcija nemainās un izpildās 3 noteikumi (Rohwer et al., 1996) :

- pastāv tikai viena neatkarīga reakcijas plūsma, ar kādu tiek patērēts vai saražots metabolīts;
- nepastāv nekādas lineārā algebrā sasaistes starp savienotajiem fermenta apakškopas metabolītiem un no ārējās vides pievadāmiem metabolītiem;
- nepastāv nekādi fermenta vai cita līdzīga veida regulācijas efekti fermentu apakškopas reakcijām un ārpus kopas metaboliskām reakcijām.

### **3. STEHIOMETRISKĀ MODEĻA PROJEKTĒŠANAS PROCEDŪRAS IZSTRĀDE UN IEROBEŽOJUMU IZMANTOŠANA**

Pašlaik dažādu izmēru stehiometrisko modeļu izveidei un projektēšanai pastāv procedūra (Thiele & Palsson, 2010), kura sastāv no pieciem posmiem. Kopā procedūra sastāv no 96 soļiem, kuri būtiski paātrina stehiometriskā modeļa projektēšanas procesu.



### 3. att. Vienkāršota stehiometriskā modeļa projektēšanas procedūra ar pieciem posmiem

avots: autora veidots

Promocijas darba autors, praktiski izstrādājot stehiometriskos modeļus (Pentjuss & Kalnenieks, 2013), (Pentjuss, Odzina, et al., 2013), (Kalnenieks, Pentjuss, Rutkis, Stalidzans, & Fell, 2014), (Pentjuss, Rubenis, Bauze, Aprupe, & Lace, 2013), ievieša uzlabojumus publicētajā projektēšanas procedūrā (skat. 3. att.), lai uzlabotu stehiometriskā modeļa izveides procesu, kā arī samazinātu projektēšanas procesa patērēto laiku un lietotāja manuālo darbu. Galvenie uzlabojumi ir veikti 2. un 4. izveides procedūras posmā. Atšķirība starp oriģināli publicēto procedūru un autora uzlaboto ir tāda, ka 2. posmā lielāks uzsvars tiek veltīts ievadāmās informācijas autentiskuma un informācijas pārbaudei no vairākām datu bāzēm. Līdz ar to 4. projektēšanas procedūras posmā validācijas procesos tiek samazināta šādas informācijas vairākkārtēja cikliskā pārbaude. Salīdzinājumā ar publicētās projektēšanas procedūru, autora izstrādātā stehiometriskā modeļa izveides procedūra samazina lietotāja manuāli veicamās darbības un patērēto laiku projektēšanai kopumā.

2. tabulā apkopotie promocijas darba autora izvirzītie ierobežojumi tika praktiski izpētīti stehiometriskajā modelēšanā un klasificēti kā obligāti, izvēles vai neeksistējoši ierobežojumi konkrētam stehiometrisko modeļu analīzes algoritmam. Kopumā tika secināts, ka katram analīzes algoritma ierobežojumam ir papildus sava pozitīvā ietekme uz analīzes algoritma kalkulācijas apstrādes laiku, izņemot EM, kur katrs papildu ierobežojums palielina kalkulācijas laiku. Pārējie ierobežojumi tiek izmantoti atbilstoši 2. tabulā uzskaitītajām klasifikatora grupām, un to pozitīvā ietekme ir atkarīga no stehiometriskā modeļa lieluma, izvirzītā mērķa, kā arī no stehiometriskā modeļa kvalitātes.



2. tabula. Promocijas darbā pētāmo analīžu metožu apkopojums par ierobežojuma kritērijiem

Algoritms	Augšējā apakšējā reakciju plūsmu vērtības	Termo-dinamika	Transporta reakciju plūsmu vērtības	Substrātu pārvēršanas ātrums	Biomasa vienādojums	Jauna celma izveide, metabolismu funkcionalitāte
<b>FBA</b>	Obligāti	Izvēles	Izvēles	Obligāti	Izvēles	Izvēles
<b>MFA</b>	Obligāti	Izvēles	Obligāti	Obligāti	Izvēles	Izvēles
<b>FVA</b>	Izvēles	Izvēles	Izvēles	Obligāti	Izvēles	Izvēles
<b>EM</b>	Nav	Obligāti	Nav	Nav	Izvēles	Nav
<b>ES</b>	Nav	Obligāti	Nav	Nav	Nav	Nav
<b>EP</b>	Nav	Obligāti	Nav	Nav	Izvēles	Nav

#### 4. STEHIOMETRISKO MODEĻU ANALĪZES REZULTĀTU VIZUALIZĀCIJAS RĪKA IZSTRĀDE

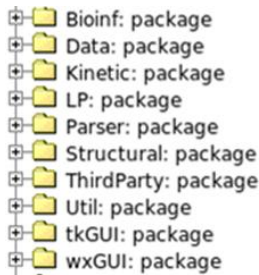
Sistēmbioloģijā arvien vairāk tiek veikti liela apjoma eksperimenti ar organismiem. Tiek iegūta dažāda veida informācija, kas var tikt izmantota modelēšanas, analīzes un simulēšanas procesos. Tomēr ar lielu analizējamo datu apjomu rodas šo datu reprezentācijas problēma. Metaboliskā tīkla attēlu izveidē katrs no tiem ir jāvērtē individuāli, jānosaka mērķis metaboliskā tīkla modeļa analīzei, jāņem vērā metaboliskā tīkla izmēri, analīzes algoritmi. Promocijas darbā tika analizēti 6 stehiometriskās modelēšanas analīzes algoritmi un izveidotas analīzes procedūras dažādu uzlabotu parametru nodrošināšanai, kā arī lietotāja manuālā darba, analīzes algoritmu apstrādes laika un atrisinājumu kopas izmēra samazināšanai. Visu analīzes algoritmu rezultātus var reprezentēt ar automatizēti veidotiem metaboliskā tīkla attēliem un diagrammām.

Promocijas darba izstrādes laikā tika analizēti 54 dažādi modelēšanas rīki, darbības vide, funkcionalitāte, stehiometrisko modeļu vizualizācijas iespējamība, kā arī pielietošana sistēmbioloģijā. Tika meklēts stehiometrisko modeļu modelēšanas un analīzes rīks ar vizualizācijas iespējām. Rezultātā tika atrasti 4 kandidāti ar stehiometrisko modeļu vizualizēšanas funkcijām (Cell illustrator, CellDesigner, CellNetAnalyzer, GLAMM), bet neviens no kandidātiem nespēja nodrošināt promocijas darbā izvirzīto mērķi.

Gala rezultātā promocijas darba autors izvēlējās modelēšanas vidi ScrumPy promocijas darba mērķa sasniegšanai, kas ar dažiem papildinājumiem spēja nodrošināt promocijas darba mērķa sasniegšanu. ScrumPy ir izstrādāta Python programmēšanas vidē, ir brīvpieejas, ietver sevī gan dinamisko, gan stehiometrisko modelēšanu. ScrumPy vidē ir iespējams integrēt lietotājam saprotamā veidā papildus nepieciešamās stehiometriskās modelēšanas algoritmus kā FVA. ScrumPy

modelēšanas rīks ir pieejams uz dažādām Unix, Linux un MAC OS operētājsistēmām un ir brīvi pieejams izstrādes kods.

Lai vizualizācijas rīks varētu automatizēti reprezentēt analīzes algoritmu rezultātus, ir nepieciešams lietotājam manuāli izveidot vizualizācijas skici ar metabolisko tīklu, jo automatizēti no paša sākuma izveidot bioloģisko sistēmu attēlošanas tradīcijām atbilstošu analizējamā metaboliskā tīkla struktūru, turklāt, ar minimālu saišu krustošanos, ir sarežģīts uzdevums pat grafa rīkiem. Līdz ar to par manuālās vizualizācijas skices izveides rīku promocijas darba rezultātu sasniegšanai tika izvēlēts InkScape vektorgrafikas apstrādes rīks, jo tas vislabāk spēj izpildīt iepriekš izvirzītās prasības vizualizācijas skices rīka izvēles kritērijiem stehiometriskās modelēšanas datu un analīzes rezultātu automātiskai reprezentēšanai.



#### 4. att. ScrumPy modelēšanas rīka moduļu klases

avots: autora veidots

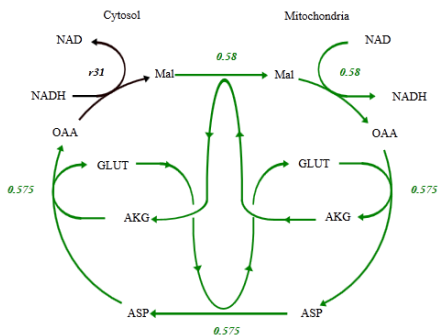
ScrumPy modelēšanas rīka „LP” un „Structural” moduļu klases tika izmantotas (skat. 4. att.), lai pielāgotu vizualizācijas moduli PySVG (<http://codeboje.de/PySvg/>), ar kura palīdzību ScrumPy modelēšanas vidē tika veikta attēla stehiometriskās modelēšanas analīzes metožu rezultātu vizualizācija ar autora izstrādāto rīku.

FVA analīzes algoritmi rezultāti ir reakciju plūsmu diapazoni, kuru vizualizācijai tika izmantots GnuPlot modulis. Šis modulis ScrumPy modelēšanas vidē izveido savienojumu ar GnuPlot (Racine, 2006) vizualizācijas rīku.

```
{'GALSID': 10.0, 'Sucqh_con': 0.0005, 'acet_dehyd': 0.12, 'asp_gluL_shuttle': 0.54500000000000028, 'm_nadph_bio_tx': -0.11, 'm_accoa_bio_tx': -0.015, 'gpl3': 0.005, 'CYT14': -0.09500000000000001, 'g6p_bio_tx': -0.125, 'm_nadp_bio_tx': 0.11, 'ASPAMI_m': 0.54500000000000028, 'ASPAMI_c': 0.54500000000000028, 'UDPEMP': 10.0, 'ATP_diss': 38.682166666666666, 'oaa_bio_tx': -0.05, 'akg_cit_shuttle': -0.09500000000000001, 'GAPD': 39.659166666666664, 'TPI': 19.784166666666664, 'ed5': 0.030}
```

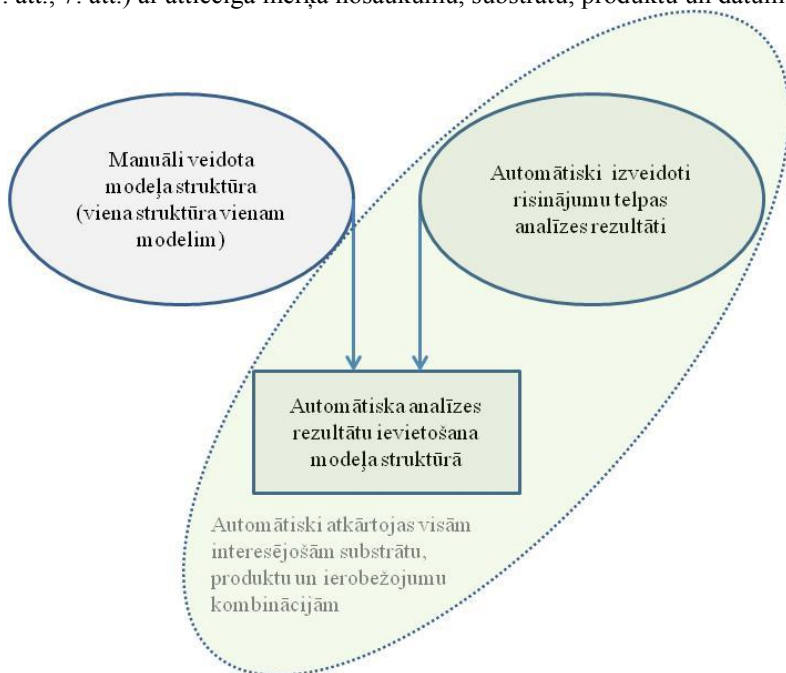
#### 5. att. ScrumPy modelēšanas rīka FBA analīzes algoritma rezultātu interpretācija

avots: autora veidots



6. att. ScrumPy modelēšanas rīka FBA analīzes algoritma rezultātu interpretācija ar promocijas darbā izstrādāto vizualizācijas rīku  
avots: autora veidots

Salīdzinot 5. att. un 6. att., var secināt, ka stehiometrisko modeļu analīžu metožu rezultāti lietotājam ērtāku datu reprezentēšanu nodrošina autora veidotais automatizētais vizualizācijas rīks, kurš īsā laikā var ģenerēt simtiem failu (skat. 6. att., 7. att.) ar attiecīgā mērķa nosaukumu, substrātu, produktu un datumu.



7. att. ScrumPy modelēšanas rīka FBA analīzes algoritma rezultātu interpretācija ar promocijas darbā izstrādāto vizualizācijas rīku  
Lielu izmēru stehiometrisko modeļu apstrādē šāda veida analīzes algoritmu

rezultātu reprezentēšana būtiski atvieglo lietotāja rezultātu apstrādes laiku un manuālo darbu.

## 5. STEHIOMETRISKO MODEĻU ANALĪZES PROCEDŪRU IZSTRĀDE

Analīžu algoritmu procedūru izstrāde promocijas darbā tika veikta ar ScrumPy modelēšanas rīku. Tika papildināta ScrumPy funkcionalitāte ar vizualizācijas iespējām vieglākai stehiometrisku modeļu analīžu algoritmu rezultātu reprezentēšanai, kā rezultātā ar mazāku laika patēriņu tika izstrādātas procedūras stehiometrisku modeļu padziļinātai, ātrākai analīzei ar samazinātu lietotāja manuālo darbu. Procedūras tika veidotas tā, lai tajās gan kombinētu analīžu algoritmus, gan kombinācijām pievienotu mērķa virzītu uzdevumu. Tika noteikts, ka uzlabotu rezultātu iegūšanai tikai konkrēti promocijas darbā izvirzītie analīžu algoritmi var tikt savā starpā kombinēti (skat. 3. tabulu). Promocijas darba izstrādes laikā tika projektēts un izveidots *Z. mobilis* centrālā metabolisma stehiometriskais modelis (Pentjuss, Odzina, et al., 2013) un alternatīvā rauga *K. marxianus* modelis, kā arī pilnveidots cilvēka mitohondrija pilna genoma mēroga stehiometriskais modelis (Pentjuss, Rubenis, et al., 2013). Modeļi attiecīgiem speciālistiem sniedza dažādas hipotēzes organisma metabolisma funkcionalitātē, ko bez tiem uzreiz cilvēks nevarēja novērot. 3. tabula. Promocijas darbā pētāmo analīžu algoritmu apkopojums par ierobežojuma kritērijiem.

Algoritma nosaukums	FBA	MFA	FVA	EM	ES	EP
<b>FBA</b>	---	---	---	---	---	---
<b>MFA</b>	Var	---	---	---	---	---
<b>FVA</b>	Var	Var	---	---	---	---
<b>EM</b>	Nevar	Nevar	Nevar	---	---	---
<b>ES</b>	Nevar	Nevar	Nevar	Var	---	---
<b>EP</b>	Nevar	Nevar	Nevar	Var	Var	---

Kopumā promocijas darbā tika aprakstītas vairākas stehiometriskā modeļa analīžu algoritmu procedūras (skat. 3. tabulu).

ES un EM analīžu algoritmu procedūras pielietošanas laikā kombinētā procedūra nodrošināja par 153 s ātrāku EM kalkulāciju (skat. 8. att.). Līdz ar to tika ietaupīts laiks EM analīzes algoritma kalkulācijās.



8. att. ES un EM analīžu metožu procedūra

avots: autora veidots

FVA un FBA analīzes algoritmu procedūrā tiek izmantoti reakciju diapazoni, kas tiek iegūti, pielietojot FVA analīzes algoritmu. Tie paātrina FBA algoritma kalkulācijas laiku. Tas tiek nodrošināts, ja reakciju diapazona augšējās un apakšējās vērtības tiek ievietotas stehiometriskā modelī un reakcijas tiek pārveidotas par vienvirziena, ja augšējā un apakšējā reakciju diapazona vērtība nemaina zīmi (Pentjuss & Kalnenieks, 2013), (Pentjuss, Rubenis, et al., 2013) (skat. 9. att.).



9. att. **FBA un FVA analīžu metožu procedūra**

avots: autora veidots

FBA un FVA analīžu algoritma procedūrā FVA parāda stehiometriskā modeļa visus alternatīvos risinājumus. Tomēr tas neparāda, kādas reakcijas ar kādām plūsmu vērtībām konkrētajā alternatīvajā stacionārā stāvoklī tiek izmantotas. Līdz ar to atrisinājumu kopā ir daudz stacionāro stāvokļu, tad FBA analīzes algoritma pielietošana samazina atrisinājumu kopas lielumu (Pentjuss, Odzina, et al., 2013) (skat. 10. att.).



10. att. **FVA un FBA analīžu metožu procedūra**

avots: autora veidots

Apvienotā MFA un FVA analīzes algoritmu procedūra palīdz interpretēt alternatīvam stacionāram stāvoklim konkrētas reakciju plūsmu vērtības detalizētā veidā (Pentjuss, Odzina, et al., 2013), (Pentjuss, Rubenis, et al., 2013) (skat. 11. att.).



11. att. **MFA un FVA analīžu metožu procedūra**

avots: autora veidots

MFA, FVA, FBA analīžu algoritmu procedūra ir papildināta ar MFA un FVA algoritmiem, kur organismam ar MFA analīzes algoritmu tiek atrasti vairāki alternatīvie risinājumi, un ar FBA palīdzību šo stacionāro stāvokļu visu reakciju plūsmu vērtības lietotājs var detalizēti izpētīt (skat. 12. att.). Tika pielietots *K. marxianus* stehiometriskā modeļa analīzē sadarbībā ar LUMBI darbiniekiem 2015. gadā.

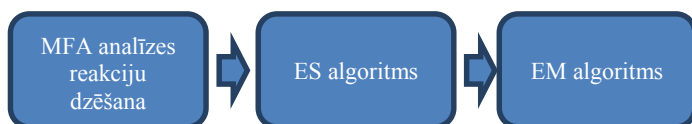


12. att. **MFA, FVA, FBA analīžu metožu procedūra**

avots: autora veidots

Uzlabotā EM analīzes algoritma darbība kopā ar ES, FBA vai FVA analīzes algoritmiem, un ar analīzes mērķa izvērzišanu, kas sastāv kopumā no 6 soļu procedūras. Vispirms ir jāizvirza EM analīzes algoritma mērķis, kādi substrāti un produkti ir iesaistīti metabolismā. Tad ar FVA vai FBA algoritmu sistemātiskā veidā jānoskaidro visi iespējamie blakusprodukti. Pēc tam ir jāizdzēš no stehiometriskā modeļa neiesaistītie produkti. Jāveic ES analīzes algoritma kalkulācija un tikai tad jāpielieto EM analīzes algoritms, lai noskaidrotu visus iespējamus atrisinājumus stehiometriskajā modelī. Piedāvātā procedūra ļauj samazināt EM kalkulācijas laiku, modeļa atrisinājuma kopas lielumu. Promocijas darbā izstrādātā vizualizācija atvieglo EM analīzes rezultātu interpretāciju un reprezentāciju. Procedūra tika pielietota *K. marxianus* stehiometriskā modeļa analīzē sadarbībā ar LUMBI darbiniekiem 2015. gadā, kā arī *Z. mobilis* modeļa apstrādē (Kalnenieks et al., 2014).

Procedūru, kura apkopo MFA analīzes algoritmu ar ES un EM algoritmiem, salīdzinājumā ar iepriekš aprakstīto procedūru, pielieto, lai pētītu kādu specifisku organisma metabolisma stāvokli pēc eksperimentos iegūtiem datiem. MFA analīzes algoritmu rezultātus stehiometriskā modelī nolasa, interpretē un lietotājs nosaka, kuri metabolīti un reakcijas konkrētā metaboliskā stāvoklī netiek iesaistīti. Tās tiek dzēstas stehiometriskā modelī. Atlikušās reakcijas veido ierobežotu metabolisko tīklu, kur ar EM un ES palīdzību tiek pētīti visi iespējamie stacionārie stāvokļi. Šādā veidā tiek aplūkoti alternatīvie stacionārie stāvokļi, kuri visticamāk būtu nākamie iespējamie stacionārie stāvokļi, ko organisma metabolisms sasniegtu. Šo procedūru izmanto, lai nevajadzētu kalkulēt visus stacionāros stāvokļus stehiometriskajā modelī, kas paātrina procesa kalkulācijas laiku, pielietojot tikai EM algoritmu (skat. 13. att.).

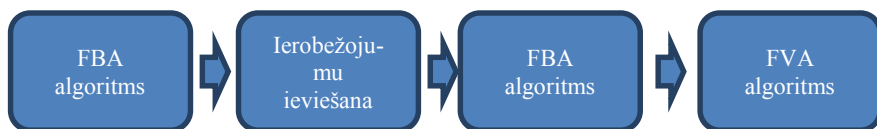


13. att. **Procedūru soļi MFA analīzes algoritma darbības uzlabošanai ar ES, EM un arizvirzītajam mērķim nevajadzīgo reakciju dzēšanu**

avots: autora veidots

FBA analīzes algoritms tiek pielietots, lai atrastu optimuma stāvokli stehiometriskajā modelī. Tomēr pastāv daudz citu alternatīvo risinājumu, kurus FBA nespēj parādīt.

FVA un FBA analīzes algoritma pielietošana kā metožu kombinēšana parādīja, ka ar FVA var iegūt visus alternatīvos risinājumus, bet neparāda konkrētā risinājuma detalizētu stacionāro stāvokli. FBA, FVA analīzes rezultātu dati nodala atrisinājuma kopas daļu, kas ietver sevī specifiskus metabolisma procesus (skat. 14. att.). Abu algoritmu pielietošana kompensē to tūkumus un ļauj lietotājam samazināt nepieciešamo manuālo darbu, samazinot kopējo nepieciešamo analīžu skaitu modelim.

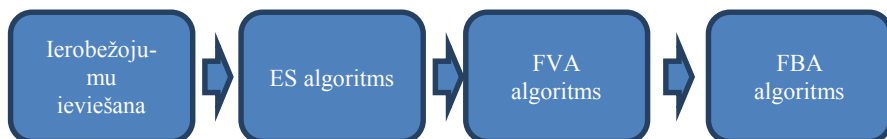


14. att. Procedūru soļi FBA algoritma darbības uzlabošanai ar ierobežojumu ieviešanu modelī un FVA algoritma pielietošanu

avots: autora veidots

FBA analīzes algoritms, izmantojot dažādus ierobežojumus atrisinājuma kopā, atrod optimālo risinājumu ar izvirzīto mērķa funkciju. Savukārt EP analīzes algoritms bez ierobežojumu palīdzības atrod mazāko iespējamo stacionāro stāvokli organismā. FBA algoritms izvirza kritēriju, ka jāatrod optimālais, bet EP izvirza kritēriju, ka jāatrod īsākais jebkāds stacionārais stāvoklis. FBA analīzes algoritmu apvienojot ar EP analīzes algoritmu un izvirzīto mērķi, var iegūt alternatīvo risinājumu, kurš nav optimālais, bet īsākais stacionārais stāvoklis (Pentjuss, Rubenis, et al., 2013) (skat. 13. att.).

Apvienoto EP un ES analīžu algoritmus kopā ar MFA analīžu algoritmu apvienoto procedūru var izmantot, lai samazinātu atrisinājumu kopu organisma metabolisma konkrētam stāvoklim, kā arī samazināt kalkulēšanas laiku. Izdevīgi procedūru izmantot, kad MFA algoritma pielietojums notiek modelim, kuram ir liels skaits eksperimentos nenoteikti reakciju plūsmu dati, īpaši, ja neizmanto biomasas vienādojumu. Savādāk ir jāstrādā ar algoritmu „*reaction correlation coefficient*” (Poolman et al., 2007), kas ir līdzīgs ES algoritma darbības principiem tikai domāts vairāk liela mēroga stehiometriskiem modeļiem ar ieviestu biomasas vienādojumu. Ar ES palīdzību noskaidro, kuras reakcijas ir apvienojamas vienā apakškopā, lai samazinātu kopējo reakciju skaitu. Ja kāda no šīm reakcijām ir ar eksperimentāli mērītām plūsmu vērtībām, tad pārējām reakcijām šajā apakškopā piešķir to pašu vērtību. Šādā veidā var samazināt nezināmo reakciju plūsmu vērtības modelī. Tas samazina alternatīvo risinājumu skaitu, MFA analīzes algoritma kalkulēšanas laiku. Nākamajos soļos, pielietojot EP analīzi, var uzreiz atrast no visiem alternatīviem stacionāriem stāvokļiem īsāko. Ja tiek atzīts EP analīzes algoritma rezultāts par neefektīvu, tad MFA un ES analīzes metožu rezultātus var pielietot tālākā analīzē, izmantojot citus *nenovirzītos* algoritmus (skat. 15. att.).



15. att. Procedūru soļi MFA algoritma darbības uzlabošanai ar ierobežojumu ieviešanu modelī, ES, FVA un FBA analīžu metožu pielietošanu

avots: autora veidots

Iepriekš aprakstītās procedūras ļauj stehiometriskā modelī samazināt analīžu algoritmu kalkulācijas laiku, iegūt dažādus alternatīvos risinājumus, samazināt

atrisinājumu kopu un samazināt manuālo darbu lietotājam ar modeli. Tā uzlabo modeļa analīzes rezultātu precizitāti.

Izmantojot minētās procedūras, ir iegūtas sekojošas bioloģiski nozīmīgas atziņas:

- mikroorganisms *K. marxianus*:
  - spēj ražot sukcinātu no ksilozes ar augstāku iznākumu nekā no laktozes vai heksozēm;
  - Krebsa cikls sukcināta ražošanā var darboties arī bez visām tajā iesaistītām reakcijām;
  - spēj saražot etanolu, acetātu, laktātu un etilacetātu tuvu teorētiskajam iznākuma maksimumam, izmantojot laktozi, glikozi vai inulīnu kā substrātus.
- mikroorganisms *Z. mobilis*:
  - spēj saražot sukcinātu no glicerīna 120 % apmērā, papildus asimilējot ogļskābo gāzi;
  - patērējot glicerīnu, spēj saražot vairāk glicerāta un sukcināta, nekā tas būtu iespējams ar heksozēm kā substrātiem.

Ar autismu saistītā cilvēka mitohondrija *SUCLG2* gēna bojājumā ATP ražošana samazinās par 50,8 %, savukārt ar autismu saistītā *SUCCOAS* gēna bojājumā ATP ražošana samazinājās tikai par 1.0 %.

## SECINĀJUMI

### Galvenie darba rezultāti

Promocijas darbā tika izstrādātas uz stehiometriskajiem datormodeļiem balstītas metaboliskās inženierijas uzdevuma risināšanas procedūras skaitļošanas apjoma un manuālā darba samazināšanai.

1. *Darbā analizēti molekulārās šūnu bioloģijas pamati, enerģijas pārnese principi un stacionārais stāvoklis metabolismā, un tā pielietošana stehiometriskajā modelēšanā matemātisko likumsakarību veidos.*

Metaboliskajos tīklos notiek reakcijas norisei nepieciešamās vielas piegāde un saražotās enerģijas aizvadišana reakcijām. Metabolisma procesi metaboliskajos tīklos virzās uz stacionāriem stāvokļiem, bet ir arī pārejas stāvokļi starp stacionārajiem stāvokļiem. Stacionārajā stāvoklī šūnas metabolismā enerģija tiek saražota un patērēta vienādi.

Stehiometriskajā modelēšanā bija nepieciešams izmantot visus promocijas darbā attēlotos šūnas metabolisma molekulāros procesus, enerģijas pārnese principus un metabolisma līdzsvara stāvokļa īpašības.



Promocijas darbā izpētītas sakarības starp sistēmbioloģijā pastāvošajām modelēšanas metodēm, īpaši sakarības starp reālajiem metabolisma procesiem molekulārajā bioloģijā un dinamiskajā modelēšanā. Līdzsvara stāvokļus metaboliskajā tīklā, ko parāda dinamiskā modelēšana, var aprēķināt, izmantojot lineāro algebru un stehiometrisko matricu.

Lineārās algebras aprēķini ir saistāmi ar šūnas metabolismu, enerģijas pārneši un metabolisma funkcionalitāti līdzsvara stāvoklī stehiometriskajā modelī.

2. *Promocijas darbā analizēti stehiometrisko modeļu analīzes algoritmu darbības principi.*

Stehiometriskajā modelēšanā izmantoti FBA, FVA, MFA, EM, ES, EP analīzes algoritmi.

FBA analīzes algoritms ir *nenovirzītais* analīzes algoritms, kuru lieto, lai aprēķinātu mikroorganisma reakciju plūsmu optimumu stehiometriskajā modelī konkrēta produkta ražošanas metaboliskās inženierijas uzdevuma veikšanai.

FVA analīzes algoritms ir *nenovirzītais* analīzes algoritms, kas parāda optimuma alternatīvos risinājumus stehiometriskajā modelī. To lieto mikroorganisma spējai pielāgoties ārējās vides apstākļu izmaiņu analīzei.

MFA analīzes algoritms ir *nenovirzītais* analīzes algoritms, kuru pielieto, kad ir liels plūsmu omikas eksperimentālo datu apjoms. Stehiometriskajā modelī ar MFA analīzes algoritmu aprēķina optimumu izvīzītajam metaboliskās inženierijas uzdevumam ar mazu alternatīvo risinājumu skaitu.

EM analīzes algoritms stehiometriskajā modelēšanā ir *novirzītais* analīzes algoritms. To pielieto, kad nepieciešams noskaidrot visus alternatīvo modeļu stacionāro stāvokļu risinājumus. EM parasti lieto stehiometriskajā modelī līdz 100 reakcijām, jo lielākiem modeļiem skaitļošanas apjoms strauji palielinās.

ES analīzes algoritms ir *novirzītais* algoritms, kuru lieto, lai noskaidrotu reakcijas, kuras savstarpēji proporcionāli korelē pie dažādu plūsmu izmaiņām stehiometriskā modelī.

EP analīzes algoritms ir *novirzītais* analīzes algoritms, kuru lieto, lai atrastu stehiometriskajā modelī visīsāko metabolisko ceļu produkta ražošanai.

3. *Darba autors ir uzlabojis stehiometrisko modeļu izveides un validācijas procedūru patērētā laika, darbību skaita, manuāli apstrādājamo datu apjoma samazināšanai.*

Autors izstrādājis uzlabotu stehiometrisko modeļu izveides un validācijas procedūru, kura samazina stehiometriskā modeļa izstrādes laiku un manuālo darbu validācijas soļos. Uzlabotā procedūra atšķiras no esošajām ar to, ka ir ieviesti papildu pārbaudes soļi stehiometriskā modeļa izstrādes procesa sākumā, lai validācijas soļos nebūtu vairākas reizes jāpārbauda viena un tā pati informācija.

Promocijas darbā tika noskaidrots, ka *nenovirzītajiem* algoritmiem, lietojot arvien vairāk dažādu ierobežojumu, var panākt analīzes algoritma procesa apstrādes laika samazināšanu. Savukārt no *novirzītajiem* analīzes algoritmiem EM un EP termodinamikas papildu ierobežojums samazina analīzes algoritma procesa apstrādes laiku, bet, ieviešot biomasas ierobežojumu, apstrādes laiks tika palielināts

daudzkārtīgi. ES analīzes algoritmā termodinamikas ierobežojums samazināja analīzes algoritma procesa apstrādes laiku.

4. *Darbā izstrādāts datu vizualizācijas rīks stehiometrisko modeļu analīzes algoritmu rezultātu vizualizācijai metaboliskās inženierijas uzdevumiem.*

Autors noskaidrojis un analizējis 56 sistēmbioloģijā izmantojamās datorizētos resursus, kur neviens no izvēlētajiem resursiem nebija piemērots visu promocijas darba uzdevumu realizācijai.

Tādējādi daļējai promocijas darba uzdevumu realizācijai tika izvēlēts stehiometrisko modeļu analīzes rīks ScrumPy.

Lai pilnvērtīgi veiktu visus promocijas darba uzdevumus, tika izstrādāts jauns vizualizācijas rīks ScrumPy modelēšanas rīkā.

Katram analīzes algoritmam tika pielāgots izstrādātais vizualizācijas rīks, lai lietotājam draudzīgā veidā reprezentētu analīzes algoritmu rezultātus.

5. *Darba autors izstrādājis stehiometrisko modeļu procedūras ar samazinātu datu apstrādes apjomu, laiku un metaboliskās inženierijas uzdevumos analizējamo atrisinājumu kopu.*

Kopumā izstrādātas dažādas procedūras, kur sešās no tām apvienoti dažādi stehiometrisko modeļu analīzes algoritmi noteiktu datu iegūšanai, kas samazina procesa apstrādes laiku, analizējamo datu daudzumu un lietotāja manuālo darbu.

Pārējos procedūru kopumos, kuros apvienoti dažādi stehiometrisko modeļu analīzes algoritmi, izmanto lietotāja specifisku informāciju. Katrs no šiem procedūru kopumiem izvirza uzlabojumus stehiometriskā modeļa analīzes, apstrādes, modelēšanas, optimizācijas, alternatīvu novērtēšanas procesos. Līdz ar to tika uzlabots procesa apstrādes laiks, kā arī samazināts analizējamo datu, alternatīvo risinājumu, lietotāja manuālā darba apjoms.

### **Secinājumi un attīstības perspektīvas**

- Uzlabotai stehiometrisko modeļu izstrādes procedūrai validācijas procesā jāizmanto ierobežots stehiometriskā modeļa apjoms, jo liela izmēra stehiometrisko modeļu izveidē validācijas posms procedūrā tiek bieži cikliski atkārtots.
- Stehiometrisko modeļu analīzes algoritmu izmantošana ar dažādiem papildus ierobežojumiem samazina analīzes algoritmu apstrādes procesa laiku un stehiometriskā modeļa risinājumu kopas apjomu.
- EM analīzes algoritmam katrs ierobežojums palielina skaitļošanas laiku un stehiometriskā modeļa atrisinājumu kopas lielumu.
- Autora izstrādātā automatizētā vizualizācijas rīka pielietošana kopā ar stehiometrisko analīzes algoritmu procedūrām samazina lietotāja iespēju kļūdaini interpretēt starprezultātus procedūru soļos un gala rezultātos.
- Promocijas darbā analizētā stehiometriskā modeļa analīzes algoritmu procedūras var lietot tikai validētam stehiometriskajam modelim, jo katrā procedūrā solī modelim ir jāsasniedz stacionārais stāvoklis.

Kā **attīstības perspektīvas** var iezīmēt vairākus darbības virzienus:

- Promocijas darbā uzlaboto stehiometrisko modeļu izstrādes un validācijas procedūru var pilnveidot un pielāgot signālu tīklu, transkripcijas tīklu, gēnu ekspresijas tīklu izveides vajadzībām.
- Pilnveidot vizualizācijas rīku, lai to varētu izmantot genoma mēroga stehiometriskajiem modeļiem manuāli vai automātiski veidoto metabolisko tīklu attēlošanai.
- Uzlabot un integrēt metabolisko tīklu analīzes algoritmu procedūras tā, lai varētu izmantot dažādu veidu omikas datu integrāciju stehiometriskā modelī.

## PARTICULARS

**Research was executed at:** Latvia University of Agriculture, Faculty of Information Technologies, Department of Computer Systems, Liela 2, Jelgava, Latvia.

**Experimental research was executed at:**

1. Latvia University of Agriculture, Faculty of Information Technologies, Department of Computer Systems, Liela 2, Jelgava, Latvia.
2. University of Latvia Institute of Microbiology and Biotechnology, Kronvalda Boulevard 4, Riga, Latvia.

**Scientific Adviser of the Doctoral Thesis:** Dr.oec. Aleksands Gailums, Associate Professor, Latvia University of Agriculture.

**The thesis was approved** at the expanded academic session of the Department of Computer Systems, Faculty of Information Technologies of the Latvia University of Agriculture on April 29, 2015. Protocol nr. 4.

**Official Reviewers:**

1. Professor of Latvia University of Agriculture, Dr.biol. Ina Alsiņa;
2. Professor of Riga Technical University, Dr. habil. sc. ing. Zigurds Markovics;
3. Main scientific researcher of Oxford Brookes University, Dr. Mark Poolman.

**The defence of the doctoral thesis** will take place at the open session of the Promotion Council in the field of Information Technologies of LUA at 11.00 a.m. on 11 March, 2016, Room 218, Faculty of Information Technologies. Lielā 2, Jelgava.

The thesis can be accessed at the LUA Fundamental Library, Lielā 2, Jelgava, and online at [http://llufb.llu.lv/romoc\\_darbi.html](http://llufb.llu.lv/romoc_darbi.html)

**You are welcome to send your comments**, signed and in a scanned form to secretary of Promotion Council – Liela 2, Jelgava, LV-3001; phone (+371) 63005621; e-mail: [tatjana.tabunova@llu.lv](mailto:tatjana.tabunova@llu.lv).

**Council Secretary:** lecturer, Mg.Paed. Tatjana Tabunova.

## APPROBATION OF PHD THESIS

The research results are presented in the following publications:

1. Kalnenieks, U., **Pentjuss, A.**, Rutkis, R., Stalidzans, E., Fell, D.A. (2014) Modeling of *Zymomonas mobilis* central metabolism for novel metabolic engineering strategies. *Frontiers in microbiology*, Vol. 5, p. 42.
2. Odzina, I., **Pentjuss, A.** (2011) In silico analysis of steady state mechanisms of metabolic networks in COBRA Toolbox and FBA-SimVis. In: Proceedings of the 17<sup>th</sup> International Scientific Conference “*Research for Rural Development*”, Vol. 17/2012(17), pp. 196–201.
3. Odzina, I., **Pentjuss, A.** (2012) Algorithm of genome-scale metabolic engineering implementation. In: Proceedings of International Scientific Conference “*Engineering for rural development*”, Jelgava, Latvia, pp. 289–293.
4. **Pentjuss, A.**, Kalnenieks, U. (2013) Assessment of *Zymomonas mobilis* biotechnological potential in ethanol production by flux variability analysis. In “*Biosystems and Information technology*”, Vol. 3(1), pp. 1–5.
5. **Pentjuss, A.**, Odzina, I., Kostromins, A., Fell, D.A., Stalidzans, E., Kalnenieks, U. (2013) Biotechnological potential of respiring *Zymomonas mobilis*: a stoichiometric analysis of its central metabolism. In: “*Journal of Biotechnology*”, Vol. 165(1), pp. 1–10.
6. **Pentjuss, A.**, Rubenis, O., Bauze, D., Aprupe, L., Lace, B. (2013) Flux variability analysis approach of autism related metabolism in stoichiometric model of mitochondria. In: “*Biosystems and Information Technology*”, Vol. 2(2), pp. 37–42.
7. **Pentjuss, A.**, Odzina, I., Gailums, A. (2012) Multi agent genome-scale metabolic reconstruction modeling software schema. In: 11th International Scientific Conference on „*Engineering for Rural Development*”, Jelgava, pp. 189–194.

The research results were presented at the following conferences:

1. Pentjuss A. (2012) Algorithm of genome scale metabolic engineering implementation. In: „Applied Information and Communication Technology (AICT 2012)”, April 26-27, 2012, Jelgava, Latvia.
2. Pentjuss A. (2012) Biotechnological potential of respiring *Zymomonas mobilis*: a stoichiometric analysis of its central metabolism. In: „2nd Conference on Constraint-based Reconstruction and Analysis (COBRA 2012)”, October 7–9, 2012, Helsingør, Denmark.
3. Pentjuss A. (2013) „Biochemical network modeling in living organisms”, In: LU and LMT computer sciences days, August 5–7, 2013, Ratnieki, Latvia.
4. Pentjuss A. (2014) Biochemical network analysis and calculation methods. In: „18th International Student Scientific Practical Conference of University

of Rezekne, Engineering Faculty “Human, Environment, Technology”, April 23, 2014, Rezekne, Latvia.

5. Pentjuss A. (2014) Constrain based metabolism analysis methods results visualization. In: „18th International Student Scientific Practical Conference of University of RRezekne, Engineering Faculty “Human, Environment, Technology”, April 23, 2014, Rezekne, Latvia.
6. Pentjuss A. (2014) Central metabolism modeling based metabolic engineering of *Zymomonas mobilis*. In: 2nd Congress of Baltic Microbiologists, October 16, 2014, Tartu, Estonia.

The research results were presented at the following seminars:

1. Pentjuss A. (2014) Modeling the Entner – Doudoroff Pathway in *Z. Mobilis* In: „110th Latvia University of Agriculture Biosystems group seminar”, May 20, 2014, Jelgava, Latvia.
2. Pentjuss A. (2014) Stoichiometric and kinetic modeling of Tomato metabolism, In: „108th Latvia University of Agriculture Biosystems Group Seminar”, January 20, 2014, Jelgava, Latvia.

**Participation in the projects related to PhD thesis development:**

1. ERDF project “Untraditional yeasts usage for bioethanol production from inulin containing compounds: biotechnological and systems biological approach”, researcher at Latvia University Microbiology and Biotechnology Institute (2014-2015).
2. „Creation of Latvian Inter-university Research Group in Systems Biology” No.2009/0207/1DP/1.1.1.2.0/09/APIA/VIAA/128, (2009–2012), researcher.

## INTRODUCTION

Genotype – phenotype relationship is important scientific direction in life sciences. Knowing genotype would be possible to predict phenotype and vice versa. In the second half of the 20th century scientists oriented in reductionist way dividing biological systems in smaller parts to get precise and detailed understanding in research of genetic, biochemical and molecular biological problems. This approach gave detailed knowledge about individual biological parts, like metabolites and enzymes, but did not show much progress in understanding systematic interaction between biological and environmental processes, which determines phenotype (Price et al., 2003).

Nowadays technological achievements enabled use of high throughput methods to characterize biological components comprehensively and simultaneously (Lewis et al., 2012). Such data production costs have plummeted and the data processed amounts have become much larger, allowing scientists to explore cells or organisms as one whole system interactions result (Duarte et al., 2007).

Computers and technology are inevitably involved in the above mentioned data extraction and acquisition. At the same time, the amount of data determines that not only extraction and acquisition, but also the analysis and interpretation is no longer possible without computer technology. Thereby, the person delegates to computers and technologies the analysis and reasoning functions with specific algorithms. The function of researcher remains to analyse computer supplied solutions and visualisations, confront them with experimental data and general rules. Computer modeling is one of the technologies that allow analysis of the designed system, where all elements are in interaction. Basically, the computer model of biological system gathers published and experimentally obtained information about the research process or object and applies recognised regularities simultaneously (in parallel). Consequently, the whole model, its parameters, inserted information and knowledge are tested in simulations. Often, it must be recognised that knowledge is controversial and requires new experiments. Modeling experiments (simulations) and biological experiments cycle are part of Systems Biology approach - cyclic progress in the process of understanding in mathematical level (model). The developed model is used for the process normalization in an organism (typical task of medicine) or for the improvement of the specified criteria (typical task of biotechnology). Metabolism, or metabolic processes is probably now best explored biological process largely due to good possibilities for the quantitative measurement of the concentrations of metabolites and application of mass conservation law. These models still do not answer question of the genotype-phenotype relationship at the desired level, but gives valuable information about the possible and impossible states of an organism. Metabolism modeling, similar to other process modeling, can be divided into two important stages:

- **development of model:** set of operations to obtain sufficiently precise mathematical representation of object or process to enable replacement of experimentation with object by computer modeling experiments;
- **analysis and / or optimization of model:** set of operations to determine the systemic features which define the organism compliance with the set optimization criteria.

There are various metabolic modeling techniques: stochastic kinetic, deterministic kinetic, stoichiometric, and others. Among these methods stoichiometric modeling is specific as it can be used for large-scale models, since it is based on the law of mass conservation and balanced reactions. Stoichiometric modeling can work with models with thousands of reagents and reactions. The main application areas of stoichiometric modeling are medicine and biotechnology. Currently, the largest relatively precise stoichiometric model is the stoichiometric model of human metabolism with more than 5,000 metabolites and more than 7,000 reactions (Thiele et al., 2013). Through the reaction catalysing enzymes also thousands of genes become involved in the model. The model is already directly applied to improvement of inborn metabolic disease diagnostic, giving more and more new predictions about the disease and expected side effects. Currently, a new computer simulation-based clinical trials (*in silico* clinical trials) (<http://avicenna-iscet.org/>) approach to medicine clinical trials is under development to reduce the time and resources necessary for the release of new drugs and medical devices on the market. With this approach, possibilities open up for early screening of different medical solutions and identification of patient groups where the proposed treatment should not be applied.

Also in biotechnology stoichiometric model application gives great results. The most popular organism *Escherichia coli* was optimized for bioethanol production (Trinh et al., 2008), conversion of biodiesel by-product glycerol to ethanol (Trinh, Srien, 2009) and amino acid production (Unrean et al., 2010). These model-predicted solutions have been confirmed in practical experiments.

There are also success stories of medium and large-scale stoichiometric model application (Kerkhoven et al., 2014) (Liu et al., 2014).

Consequently, stoichiometric model success and the large scale of these models introduces new requirements on software products, particularly in situations where interdisciplinary team involving both biological and mathematical/computational specialists is working.

Balancing the reaction is important, but complicated process (Thiele, Palsson, 2010), which metabolic model creators often tend to simplify. As a consequence, the optimization result is a solution that does not comply with the mass conservation law and leads to failure in biological experiments.

Another problem is the metabolic network visualization, allowing biologists to recognize the processes and permit an assessment of their potential relevance to the real behavior of particular organism. Ideally, visualization allows to evaluate in



a short time the state of the metabolic network, interaction of different branches and to analyze the manipulation opportunities.

A growing problem is the computational power. Meaningful computing tasks of stoichiometric model analysis can take months (for example – elementary modes calculation). Therefore, it is necessary to improve the optimization procedures to stoichiometric models to get more and more accurate, practical biotechnology and medicine usable information with the least possible computational resources and time consumption.

Bio-information databases with a variety of information: reactions coding genes, new genome sequences of organisms, information about the thermodynamic characteristics of reactions that determine the direction of response limits, etc.

Although in the mentioned directions many studies already have improved the modeling process, there are still great opportunities to improve performance, because the stoichiometric model development and optimization is a relatively new direction and continues to develop rapidly. In particular, this applies to the handling of large-scale, including genome-scale models under circumstances where the amount of information is growing rapidly thanks to experimental technologies. The described situation determines the aim and form objectives and tasks of the PhD thesis.

### **The aim and tasks of PhD thesis**

The aim of thesis is to develop stoichiometric computer models based procedures of metabolic engineering task solving procedures to reduce computation and manual work.

In order to achieve the aim of PhD thesis several tasks were completed:

1. To analyze the mathematical regularities between the stoichiometric modeling and metabolism molecular processes.
2. To research operating principles of stoichiometric model analysis algorithms.
3. To improve stoichiometric model design and validation procedure to reduce time consumption, number of operations and manual processing of data.
4. To develop visualization tool for metabolic engineering tasks to visualize results of stoichiometric model analysis algorithms.
5. To create procedures of stoichiometric model analysis algorithms for reducing number of operations, time consumption, manual work and number of analysed solutions.

### **Research methods**

Stoichiometric model calculations are made in matrix where information about organism metabolic processes is encoded. Transposition of matrices and algebraic operations with matrices are used to find the steady state preconditions of metabolic network.

Linear algebra techniques are used in stoichiometric matrix analysis to find the set of steady state fluxes. Superposition of this set represents all possible steady state solutions for particular stoichiometric model.

Experimental data validation process in stoichiometric model is made comparing experimental data with a stoichiometric model results.

Mass conservation law is used for reaction mass balance and in steady state processes.

Constraint-based modeling and analysis methods are used.

XML files are used for visualization of the results of stoichiometric model analysis algorithms.

### **Scientific novelty**

1. Stoichiometric model design procedure for specific models, objectives and available data.
2. Visualization tool to facilitate representation of stoichiometric model states.
3. Procedures of stoichiometric model analysis algorithms for reduction of spent time, number of operations and manually processed data amount.

### **Theses**

1. Early introduction of constraints in stoichiometric model creation process reduces the amount of operations in the validation process.
2. Automated model state visualization improves metabolic network analysis.
3. Combinations of metabolic analysis algorithms reduce the solution space and set of solutions to be analysed.

### **PhD thesis structure and volume**

The PhD thesis is written in Latvian containing abstract, introduction, 5 chapters, conclusions, bibliography, and 12 annexes combined in additional volume, including 15 tables, 44 figures, and 145 pages in total. 196 literature sources were used.

## **1. INCORPORATION OF BIOLOGICAL DATA DIVERSITY IN COMPUTER MODEL**

Each organism type (species) has unique genome where metabolic reactions, metabolites, their interactions, functionality, and their compatibility are encoded.

Computer model of computer controllable biological systems is a precondition of computer controlled development, since all biological systems irrespectively of their scale (cell, tissue, organism) has a control system (biological control system – BCS) which is genetically determined (Stalidzans, 2005).

Currently, optimum design and use of computer control possibilities is delayed by insufficiently described BCS in a mathematically formalized way in a

computer model. BCS computer modeling and optimization has a key role in successful organism computer designing and implementation. Since a specific gene corresponds to a metabolic reaction, metabolism of genetically determined organisms, basically, is BCS, and its modeling and optimization is the main aim of the PhD thesis.

The metabolites are substances that ensure metabolic processes in an organism. Metabolites have a wide range of tasks, such as fuel of organism metabolic processes for forming cells, tissues and organs.

Mathematically there is a relation between enzyme and metabolites (substrate and product) interaction.

The reaction flux (Edwards et al., 2002) or reaction rate is a mathematical value which describes the speed at which the substrate is converted into product using metabolic pathways and reactions, where this process is regulated by enzyme.

Metabolic pathways in stoichiometric modeling are built from involved reactions where those convert substrates to products (see Figure 1 A). Metabolic conversion of intermediates is forming continuous series of reactions (see Figures 1 B, C, D) known as metabolic pathways. Several metabolic pathways can contain the same metabolite (see Figures 1 C, D), which participates in a variety of reactions, thereby creating the link between these metabolic paths.

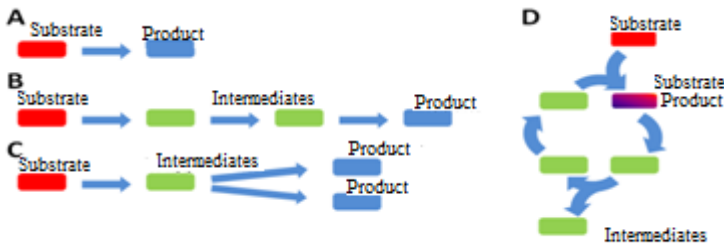


Figure 1. **Metabolic network biochemical pathway examples**

Source: made by author

Each metabolic reaction requires energy to catalyze (activate) it. This energy is minimum energy required to start a chemical reaction. When the reaction is in progress, the variations in the free energy are equal to the free energy value of products minus the amount of free energy substrates. Mathematically, this process is described by the following equation (Welch, 1985):

$$\Delta G = \Delta G_{\text{product}} - \Delta G_{\text{substrate}} \quad (1)$$

$\Delta G$  – free energy required to start reaction conversion process;

$\Delta G_{\text{product}}$  – total product sum of Gibbs free energy;

$\Delta G_{\text{substrate}}$  – total substrate sum of Gibbs free energy.

Metabolism in organism tends to occur in a steady state when all reaction rates are in balance. After certain amount of time all reaction required free energy for steady state processes are established in other metabolism processes in organism. In this case metabolic reactions in organism ensure stable, non-fluctuating metabolism, where substrate consumption and creation of products are constant. Intermediate conversion occurs in the same way. This means that in steady state reaction rates and metabolite concentrations do not change in time.

There is a need to develop complex data analysis techniques to handle the rapidly growing volumes of multi-dimensional data, Different methods have been developed and applied from stochastic kinetic models (Golightly, Gillespie, 2013), Bayesian networks (Wilkinson, 2007), stoichiometric models (Lewis et al., 2012) up to the graph and interaction network development and analysis (Le prevents et al., 2009).

The stoichiometric (constraint-based) modeling, which is used in genome scale stoichiometric model (GSMR) analysis, is discussed in the thesis. The reconstructed genome-scale metabolic network includes integrated and systematic information on certain metabolites and reactions in the cell or organism.

	ATPGMD	etoh_tx	XI
x_glcD	1	0	0
glcD	0	0	0
x_xylD	0	1	0
xylD	0	0	-1

Figure 2. **Example of part of stoichiometric matrix for metabolic network representation**

Source: made by author

This is a matrix (see Figure 2) where the metabolites are represented in columns and metabolites are represented in rows. If metabolite is involved in reaction then it is denoted 1 or -1. If metabolite is not involved in reaction then it is denoted as 0 (see Figure 2). A positive number indicates that the metabolite is produced in the reaction; a negative number means that a metabolite is consumed in the reaction.

$$\begin{bmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (2)$$

$S_1, S_2, S_3$ , – substrates used in reactions,

$v_1, v_2, v_3, v_4, v_5$ , - reaction rates.

The enzyme concentration changes in certain time period is expressed on the left side of the formula.

## 2. STOICHIOMETRIC MODELING ANALYSIS ALGORITHMS

Nowadays stoichiometric modeling analysis methods are divided into two large groups: biased and unbiased algorithms. Unbiased algorithms carry out a systematic examination of all solutions, while biased algorithms are used to analyse one solution from all possible solution set.

In the thesis several analysis methods are described like: the Flux Balance Analysis (FBA) Flux Variability analysis (FVA) (Burgard et al., 2001), Metabolic Flux Analysis (MFA) (Zhao, Shimizu, 2003), which are biased algorithms. Also unbiased algorithms were used: Elementary modes (EM) (Schuster et al., 1999), Enzyme Subset (ES) (Pfeiffer et al., 1999), Extreme Pathway (EP) (Yeung et al., 2007).

### Biased analysis algorithms

Biased algorithms use stoichiometric matrix and biological constraints to analyse solution space with vector calculations. Biased algorithms such as they are without GSMR and additional constraints provide incomplete and inaccurate analysis of the data from the solution space, because not all solutions in the space are biologically, chemically and physically feasible and can occur in organism. Additional and biologically feasible constraints should be introduced to exclude from GSMR solution space unfeasible natural metabolic conditions.

Omics data is different experimentally measured data from the organism or cell group. These may include different kind of data such as genomics (Barrett et al., 2013), proteomics (Gasteiger, 2003) transcriptomics (Craigon et al., 2004), metabolomics (Xia et al., 2012), epigenomics (Bernstein et al., 2010), fluxomics (Bernstein et al., 2010) data. Stoichiometric model analysis and optimization allows integration of multiple omics data types (Hyduke et al., 2013).

Another biological constraint for GSMR is biomass reaction equation (biomass composition, BVR) (Gombert et al., 2001). Numeric values for each metabolite in BVR represent the quantity needed to create a one unit of biomass. BVR incorporates all the metabolites required for cell development, growth, division and reproduction.

Each reaction is constrained in the range of possible Flux values. Such constraints in GSMR decrease solution space, which also improves the stoichiometric modeling analysis algorithms by reducing the processed set of solutions. There are several ways to insert biological data as constraints into a model:

- reaction flux for substrate consumption;
- gene deletion inclusion;
- omics data inclusion;
- reaction flux for product production.

Most stoichiometric modeling methods modify the reaction flux bound values in GSMR.

FBA algorithm aims to predict flux distribution in the organism metabolic model, which corresponds to the entered experimental data – constraints. FBA method seeks optimal solution of model satisfying all constraints values. FBA method is used to predict specific phenotype characteristics for the organism as flux distribution (Edwards et al., 2002). The result in solution space is the criteria and constraints determined optimum in the GSMR. Usually there is a large number of alternative solutions.

MFA algorithm datasets (Joyce, Palsson, 2006), which are obtained in fluxomics experiments are used in MFA analysis algorithm processing.

Unlike to FBA analysis algorithm at MFA analysis in experiments measured reaction rates are used as constant or small diapason values.

As result, the solution space is strictly limited and the number of answers is drastically reduced.

FVA algorithm uses all imposed GSMR constraints and target function to simulate minimal and maximal flux values of model reactions. FVA can be used for computer simulations of a particular organism phenotype. If modifications are made in at least one restriction, FVA analysis results need to be recalculated, because analysis results may change significantly.

## Unbiased analysis algorithms

Unbiased algorithms are applied in cases when it is necessary to see all nullspace solution space instead of only some of the possible solutions.

Nullspace vectors set is a constrained set of solution space, where you can find any possible steady state of stoichiometric model. The solution set itself is used in other methods of analysis to carry out systematic nullspace solutions analysis (Poolman et al., 2007). Unbiased algorithms use solution space with nullspace vectors and the stoichiometric matrix to perform all alternative optimum calculations in stoichiometric model analysis. Mathematically EM means that the steady state reaction rate vectors (reaction flux values) must be in stoichiometric matrix solution space (Reder, 1988). It is considered as an object with boundaries constrained by nullspace vectors. It was defined that EM is correct vector selection in nullspace (Newsholme et al., 1984), (Fell, 1993).

All possible metabolic steady state nullspace vectors are calculated as follows (Fell, 1992), (Leiser, Blum, 1987):

$$\frac{dX}{dt} = v \quad (4)$$

v – reaction flux rate value,  
X – concentration vector,

Steady-state rule defines an algebraic equation (Schuster, Hilgetag, 1994):

$$N * v = 0 \tag{5}$$

N – stoichiometric matrix.

The number of dimensions (N) depends on the number of metabolites and reactions in the stoichiometric model.

EM is the smallest possible set of metabolites at steady state, where is at least one substrate and one product, and there are no multiple steady states. EM is a unique set without other subsets (Pfeiffer et al., 1999).

The EP is the shortest of all possible EM solutions. EP method is used to determine metabolic network capability to convert substrate to products with the least possible number of reactions and enzymes.

In gene expression, metabolic engineering or metabolic regulation studies it is highly interesting to find groups of enzymes that in all steady states operate with the same reaction flow ratio. Such an enzyme group is called enzyme subsets (ES). This is a group of enzymes which provide metabolic control not only of one but for several enzymes group in coordinated manner. In stoichiometric modeling ES is used for enzyme group where the reactions do not change the flux ratio in a subset and fulfills 3 rules (Rohwer et al. 1996):

- there is only one independent reaction flux rate how fast is consumed or produced metabolite;
- there is no linear algebra linkage between the coupled ES internal and external metabolites;
- there is no enzyme or other similar reactions regulatory effects in and outside enzyme subset.

### **3. DEVELOPMENT OF STOICHIOMETRIC MODEL DESIGNING PROCEDURE AND APPLICATION OF CONSTRAINTS**

Currently, for developing and designing stoichiometric models of various size there is a procedure (Thiele, Palsson, 2010) that consists of five parts. In total, the procedure consists of 96 steps which significantly accelerate the process of designing stoichiometric model.

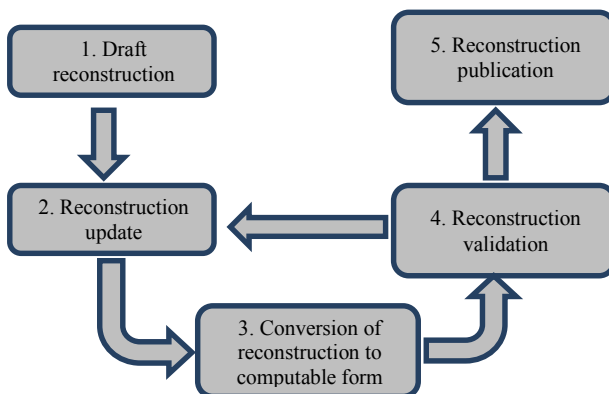


Figure 3. **Simplified stoichiometric model design procedure**

Source: made by author

The author of the thesis during building of stoichiometric models (Pentjuss, Kalnenieks, 2013) (Pentjuss, Odziņas, et al., 2013), (Kalnenieks et al., 2014), (Pentjuss, Rubenis, et al., 2013) developed improvements in the published design procedure (see Figure 3) to improve stoichiometric model creation process as well as reduce the time spent on designing and user's manual work.

The major improvements have been made in Stage 2 and 4 of the design procedure. The difference between the originally published procedure and the procedure improved by the author is in Stage 2 where more attention is paid to the reliability of the entered information and comparison of the information from multiple databases. Therefore, in Stage 4 of design procedure during validation process there is a reduction of repeated cyclical verification of the information. It reduces manual work of the user and the overall time spent for designing stoichiometric models.

In Table 2 the author puts forward within the thesis collected and analysed different constraints of stoichiometric model depending on the modeling algorithm. These constraints were practically applied and classified as mandatory, optional or not applicable for each stoichiometric model analysis algorithm described in the thesis.

In general, it was concluded that each of the analysis algorithm constraints has positive impact on the calculation time, except EM, where each additional constraint increase the calculation time.

Other constraints are used in accordance with classification of listed groups in Table 2, and their positive impact depends on the stoichiometric model size, of posed aim and as well as the quality of the stoichiometric model.



Table 2. **Summary of stoichiometric model analysis methods constraints**

<b>Method</b>	<b>Upper / lower bounds</b>	<b>Thermo-dynamic</b>	<b>Trans-port reaction flux</b>	<b>Substrate consumption rate</b>	<b>Biomass composition reaction</b>	<b>New strain development</b>
<b>FBA</b>	Mandatory	Optional	Mandatory	Mandatory	Optional	Optional
<b>MFA</b>	Mandatory	Optional	Mandatory	Mandatory	Optional	Optional
<b>FVA</b>	Optional	Optional	Optional	Mandatory	Optional	Optional
<b>EM</b>	None	Mandatory	None	None	Optional	None
<b>ES</b>	None	Mandatory	None	None	None	None
<b>EP</b>	None	Mandatory	None	None	Optional	None

#### **4. DEVELOPMENT OF STOICHIOMETRIC MODEL ANALYSIS VISUALIZATION TOOL**

Large-scale experiments with organisms become more frequent in systems biology. There are different types of collected information that can be used for modeling, analysis and simulation processes.

However, large amount of data for analysis raise a representation problem of collected data. For better data interpretation the metabolic network visualization should take into account the defined aim of modeling, dimensions of the metabolic network and applied analysis algorithms. 6 stoichiometric modeling analysis algorithms were applied and analysed in the thesis. Analysis procedures are developed for different improved parameters, reduced manual work for a user and time spent for processing analysis algorithms, as well as for reduction of solution space size. Results of all analysis algorithms can be represented with automatically constructed metabolic network images and diagrams.

In the thesis, 54 different computational tools, their working environment, functionality, visualization capability, and possible use in systems biology were analysed. The aim was to find stoichiometric modeling tool with visualization capabilities. As a result, 4 candidates were found (Cell illustrator, CellDesigner, CellNetAnalyser, GLAMM) but none of them could fulfill the aim of the thesis.

Author chose ScrumPy modeling tool to do all tasks set in the thesis, and with a few additions it was capable of fulfilling the aim of the thesis. ScrumPy is made in Python programming environment and by licensing restrictions for academic purposes it can be used free of charge. In addition, ScrumPy is one environment for dynamic and stoichiometric modeling approach. In a short time period it was possible to integrate additional stoichiometric analysis methods like FVA. ScrumPy tool can be used on different operational systems like MAC OS, UNIX, and Linux.

A user has to manually develop a visualization design with metabolic network in order to allow the visualization tool to automatically represent the results of analysis algorithms, because it is a complicated task even for graphical tools to develop metabolic network structure in a fully automated way that corresponds to traditional representation of biological systems and, moreover, with minimum intersections. Therefore, Inkscape vector graphics editor tool was chosen as a tool for manual development of visualization design to achieve the results of the thesis because it is the most appropriate tool to meet the requirements of the above visualization design tool selection criteria for stoichiometric modeling and automated representation of data analysis results.

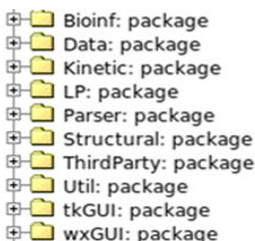


Figure 4. **ScrumPy modeling tool modules**

Source: made by author

ScrumPy modeling tool "LP" and "Structural" class modules (see Figure 4) were used to adjust the visualization module PySVG (<http://codeboje.de/PySvg/>). ScrumPy for drawing the stoichiometric modeling analysis results with user friendly interface.

FVA analysis method results are reaction flux ranges that were visualized using Gnuplot module. This module connects ScrumPy modeling environment to the GNU PLOT (Racine, 2006) visualization tool.

```
{'GALSID': 10.0, 'Sucqh_con': 0.0005, 'acet_dehyd': 0.12, 'asp_gluL_shuttle': 0.54500000000000028, 'm_nadph_bio_tx': -0.11, 'm_accoa_bio_tx': -0.015, 'gpl3': 0.005, 'CYT14': -0.09500000000000001, 'g6p_bio_tx': -0.125, 'm_nadp_bio_tx': 0.11, 'ASPAMI_m': 0.54500000000000028, 'ASPAMI_c': 0.54500000000000028, 'UDPEMP': 10.0, 'ATP_diss': 38.682166666666666, 'oaa_bio_tx': -0.05, 'akg_cit_shuttle': -0.09500000000000001, 'GAPD': 39.65916666666666664, 'TPI': 19.784166666666664, 'ed5': 0.030}
```

Figure 5. **FBA analysis method result interpretation by ScrumPy modeling tool**

Source: made by author

By comparing Figure 5 and Figure 6 it can be concluded that the representation of stoichiometric model analysis methods results in more user friendly interface is provided by the visualization tool created by the author that can generate in short time hundreds of files (see Figure 6, Figure 7) with the corresponding target name, substrate, product, and date.

This type of representation of analysis algorithm results significantly simplifies processing time of analysis results and manual work amount specially using large scale stoichiometric models.

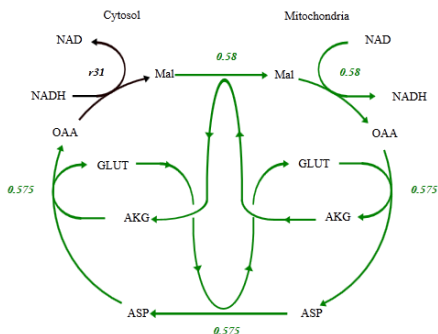


Figure 6. **FBA analysis method result interpretation by ScrumPy modeling tool by the developed visualization tool**

Source: made by author

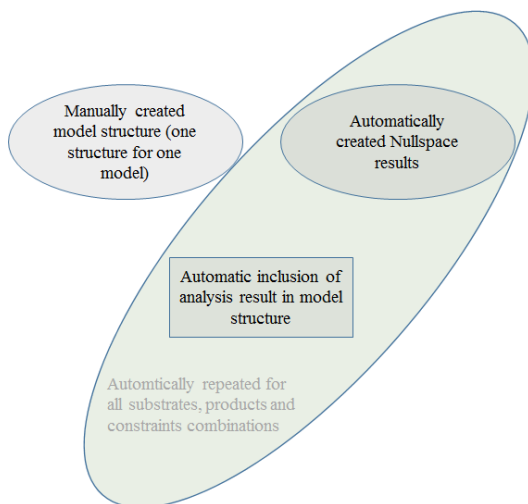


Figure 7. **Interpretation of ScrumPy modeling tool FBA analysis method results by the visualization tool developed within the thesis**

Source: made by author

## 5. DEVELOPMENT OF STOICHIOMETRIC MODEL ANALYSIS PROCEDURE

Analysis algorithms procedures for the thesis were carried out with ScrumPy modeling tool. ScrumPy functionality was extended with visualization capabilities for easier representation of the results of stoichiometric model analysis algorithms. As a result, in shorter period of time analysis algorithm procedures were developed for deeper and faster analysis with reduced manual work. Procedures were developed

in such way so to combine analysis algorithms and include task solution in combinations.

It was determined that only those analysis algorithms which are proposed in the thesis can be combined with each other to improve the obtained results (see Table 3).

During development of the thesis the following models were designed and made: *Z. mobilis* central metabolic stoichiometric model (Pentjuss, Odzina, et al., 2013), alternative yeast *K. marxianus* model and improved human mitochondrial genome scale stoichiometric model (Pentjuss, Rubenis, et al., 2013).

Table 3. **Stoichiometric analysis algorithms constraints summary**

Method	FBA	MFA	FVA	EM	ES	EP
<b>FBA</b>	---	---	---	---	---	---
<b>MFA</b>	Can	---	---	---	---	---
<b>FVA</b>	Can	Cannot	---	---	---	---
<b>EM</b>	Cannot	Cannot	Cannot	---	---	---
<b>ES</b>	Cannot	Cannot	Cannot	Can	---	---
<b>EP</b>	Cannot	Cannot	Cannot	Can	Can	---

In general, several stoichiometric model analysis algorithm procedures (see Table 3) are described in the thesis.

ES and EM analysis algorithms combined procedure provided for 153 fold faster calculations results than EM (see Figure 8). Therefore, it was saving time for calculation process comparing to EM analysis algorithm.



Figure 8. **ES and EM analysis methods procedure**

Source: made by author

In FVA and FBA analysis algorithm procedure the reaction ranges are used, which are obtained using the FVA analysis algorithm. They accelerate the FBA algorithm calculation time. This is done if the current reaction flux range of the upper and lower values are placed in the stoichiometric model as constraints and the reaction is converted to irreversible one, if the upper and lower range value of the reaction is not negative number (Pentjuss, Kalnenieks, 2013) (Pentjuss, Rubenis, et al., 2013) (see Figure 9).



Figure 9. **FBA and FVA analysis methods procedure**

Source: made by author

In FBA and FVA analysis algorithm procedure FVA shows all alternative solutions in stoichiometric model. However, it does not show which reactions and with which flux values are used in any specific alternative solution. Since in this case in the solution space there are many alternative solutions, then FBA analysis algorithm reduces the solution space from FVA analysis algorithm results (Pentjushi, Odziņas, et al., 2013) (see Figure 10).



Figure 10. **FVA and FBA analysis methods procedure**

Source: made by author

Unified procedure of MFA and FVA analysis algorithms helps to interpret specific reaction flux values for alternative steady state in detailed way (Pentjuss, Odzina, et al., 2013) (Pentjuss Rubenis, et al., 2013) (see Figure 11 ).



Figure 11. **MFA and FVA analysis methods procedure**

Source: made by author

MFA, FVA and FBA analysis algorithm procedure is updated with MFA and FVA analysis algorithms, where MFA analysis algorithm is used for organism modeling, to find several alternative solutions. With FBA it is possible to analyse flux values of all steady state reactions in detail (see Figure 12). It was used in stoichiometric model analysis of *K. marxianus* in collaboration with LU MBI (Latvia University Microbiology and Biotechnology Institute) employees in 2015.



Figure 12. **MFA, FVA, FBA analysis methods procedure**

Source: made by author

Improved EM analysis algorithm operation with the ES, FBA or FVA analysis algorithms combined with chosen aim of analysis that consists of a 6 step procedure. First of all, before EM analysis algorithm application there should be posed the aim for substrates consumption and products production. After that by systematically using FVA or FBA analysis algorithm all the possible products must be determined that can be produced in stoichiometric model. When it is determined, then the products not used must be deleted from stoichiometric model solution space. And in the end of procedure ES analysis algorithm has to be calculated and only then EM

analysis algorithm should be run to find all possible solutions in stoichiometric model. The proposed procedure decreases EM calculation time, size of stoichiometric model solution space and the visualization developed in the thesis improves EM analysis results interpretation and representation. This procedure was used in *K. marxianus* central metabolism analysis in collaboration with LUMBI employees in 2015, as well as in processing *Z. mobilis* model (Kalnenieks et al, 2014).

A procedure that combines MFA analysis algorithm and the ES and EM analysis algorithms, compared with the previously described, is used to study a specific condition of the organism metabolism using experimental data. User reads, interprets, and determines which metabolites and reactions are not involved using results of MFA analysis algorithm in this procedure. They are deleted from the stoichiometric model. Remaining reactions in metabolic network are involved in EM and ES analysis to study all steady states. In case of metabolism perturbations these steady states are alternative solutions which could be the closest ones measured in MFA. This procedure is used to decrease EM calculation time and not to seek full solution set steady states, only closest to MFA analysis data (see Figure 13).



Figure 13. **Procedure steps for MFA analysis improvement with ES, EM and reaction deletion not necessary for specific aim**

Source: made by author

FBA analysis algorithm is used to determine the optimum in stoichiometric model. However, there are many other alternative solutions which are not showed by FBA.

FVA and FBA combined procedure shows that the FVA case calculates all the alternative solutions, but does not show the particular steady state values. With FBA, FVA analysis results can be constrained to separate specific solution set (see Figure 14).

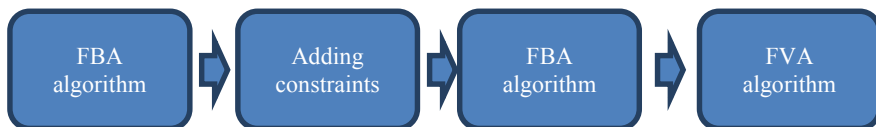


Figure 14. **FBA improving procedure steps with constraints adding and FVA usage**

Source: made by author

FBA analysis algorithm by using different constraints in solution space, finds the optimum solution with the proposed objective function. However, EP without constraints can find the smallest possible steady state solution. FBA relies on criterion that it can find the optimal solution, but the EP relies on that it must find

the shortest any steady state solution. By combining FBA with the EP analysis algorithm and the set modeling aim, the alternative solution can be found, which is not the best one, but the shortest steady state (Pentjuss, Rubenis, et al., 2013) (see Figure 13).

Unified procedure of combined EP and the ES together with MFA can be used to reduce the solution space to smaller and specific one, which helps decrease calculation time. It is beneficial to use the proposed procedure MFA step when experimental data have large set of reactions rates and especially when biomass composition reaction is not included in stoichiometric model. Otherwise, the "Reaction Correlation Coefficient" method must be used (Poolman et al., 2007), which works similarly to ES, but is specialized for large scale stoichiometric models with the biomass composition reaction. Using ES, reactions can be found which correlate with constant ratio. If any of these reactions is experimentally measured, then the other reactions in this subset are assigned with the same value. In this way, it is possible to reduce the unknown reaction upper and lower bounds, which in turn reduces the number of alternative solutions and also reduces MFA analysis method calculation time. In next steps using EP, the shortest alternative solution can be found. If the EP result is recognized as inefficient then MFA and ES results can be used for further analysis using other unbiased algorithms (see Figure 15).

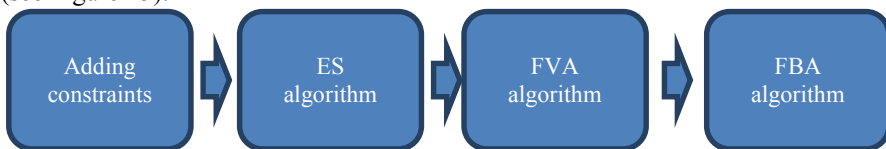


Figure 15. **Steps of procedure for MFA calculations improvement by adding additional constraints and combining ES, FVA and FBA**

Source: made by author

The previously described procedures allow reducing calculation time of the stoichiometric model analysis algorithms, obtaining a variety of alternative solutions, reducing the solution set and manual user work, as well as improving accuracy of the model analysis results.

By using the described procedures, the following biologically relevant conclusions were made:

- microorganism *K. marxianus*:
  - can produce succinate from xylose at a higher result than from lactose or hexose;
  - the Krebs cycle in the case of succinate production can also operate without all reactions involved in it;
  - can produce ethanol, acetate, lactate and ethyl acetate close to the maximum of theoretical yield by using lactose, glucose or inulin as a substrate.
- microorganism *Z. mobilis*:

- can produce succinate from glycerine amounting to 120 per cent with additional assimilation of carbon dioxide;
- using glycerine can produce more glycerate and succinate than it would be possible from hexose as a substrate,
- in the case of autism-related damaged SUCLG2 gene of human mitochondria, the ATP production decreases by 50.8 per cent, but in the case of autism-related damaged SUCCOAS gene – decreased by only 1.0 per cent.

## **CONCLUSIONS**

### **Main results of PhD thesis**

In the PhD thesis procedures for stoichiometric model application in metabolic engineering task were developed to reduce computation and manual work.

1. *In the thesis author analysed molecular cell biology, energy transfer processes and steady state metabolism processes applied in stoichiometric modeling.*

In metabolic networks, the substances necessary for reactions are delivered and the produced energy is transferred to reactions. Mostly metabolic processes in metabolic networks tend to reach steady state solution, but there are also transitional states between steady states. In steady state cell metabolism energy is produced and consumed in equal amount.

For stoichiometric modeling it was necessary to use all molecular cell metabolic processes, energy transfer principles and metabolic steady-state properties described in the thesis.

In the thesis, the relationship between existing methodologies of systems biology modeling were investigated, in particular, the relationship between real metabolic processes of molecular biology and dynamic modeling. Steady states in metabolic network showed by dynamic modeling can be calculated also by linear algebra and stoichiometric matrix approach.

Linear algebra calculations in stoichiometric model are linked with cell metabolism, energy transfer and metabolism functionality at steady state.

2. *Functionality principles of stoichiometric model analysis algorithms were studied in the PhD thesis.*

The author studied the functionality of the FBA, FVA, MFA, EM, ES and EP stoichiometric modeling analysis algorithms.

FBA is unbiased analysis algorithm that is used to calculate reaction flux optimal solution of microorganism in stoichiometric model for metabolic engineering tasks while producing particular product.

FVA is unbiased analysis algorithm that shows alternative solutions of the optimum in stoichiometric model. It is used to analyse full adaption possibility spectrum of a microorganism.



MFA analysis algorithm is unbiased analysis algorithm that is used when a large scale flux omics experimental data is available. In stoichiometric model MFA analysis algorithm calculates the optimum solutions for metabolic engineering tasks put forward by a small number of alternative solutions.

EM analysis algorithm in stoichiometric modeling is biased analysis algorithm. It is used to find all alternative steady-state solutions of a model. Usually, it is applied for stoichiometric models with up to 100 reactions due to combinatorial explosion in calculations.

ES is biased analysis algorithm, which is applied to find reactions which mutually correlate in one proportion at various flux changes in stoichiometric model.

EP is biased analysis algorithm, which is used to find the shortest metabolic pathway in stoichiometric model to produce a product.

- 3. The author has improved the stoichiometric model development and validation procedure reducing spent time, number of operations and manually processed data amount.*

The author developed the improved stoichiometric model development and validation procedure that reduces stoichiometric model development time and manual work amount in validation steps. The improved procedure differs from the existing ones with additional verification steps at the beginning of stoichiometric model development process, which means that the same information in validation steps is not repeatedly checked several times.

In PhD thesis it was proved that unbiased algorithms, using different constraints, can reduce calculation time for stoichiometric modeling analysis. Biased analysis algorithms like EM and EP using additional thermodynamics constraint decrease calculation time, but providing biomass constraint, the calculation time is dramatically increased. In case of ES analysis algorithm, thermodynamics constraint decreased calculation time.

- 4. Data visualization tool for visualizing results of stoichiometric model analysis algorithms was developed.*

The author found and analysed 56 computerized resources, which can be used in systems biology field, but none of the found resources was suitable for all tasks proposed in the thesis.

Therefore, stoichiometric model analysis tool ScrumPy was chosen for partial realisation of PhD tasks. A new visualization tool for ScrumPy was developed in order to fully perform all the tasks proposed in the thesis. Visualization tool was adapted for each analysis algorithm so that analysis results could be represented in a user-friendly way.

- 5. The author of PhD thesis developed stoichiometric model analysis procedures, which reduce the data amount, processing time and solution set of metabolic engineering tasks.*

Different complexity procedures were developed where six of them combined different stoichiometric model analysis algorithms, which decreased analysis process time, processing data amount and user manual work amount.

In other procedures where various stoichiometric model analysis algorithms use not only analysis algorithm combinations, but also user defined specific information. Each of these procedures proposes specific upgrades for stoichiometric model analysis, processing, modeling, optimization and assessment of alternatives. Thus, stoichiometric model analysis time was improved, but amount of processing data, alternative solutions and user manual work was decreased.

### **Conclusions and development prospects**

- Validation process of the upgraded stoichiometric model development procedure uses smaller scale models causing less raw stoichiometric validation data amount, because the larger size of stoichiometric model is created, the longer takes validation phase due to increasing cyclic validation steps.
- Using additional constraints with stoichiometric model analysis methods decrease process calculation and stoichiometric model solution set amount.
- Additional constraints increase computational costs of EM calculation and stoichiometric model solution set amount.
- Visualization tool developed by the author in conjunction with use of stoichiometric model analysis methods procedures reduces the possibility of errors compared to analysis without visualization.
- Stoichiometric model analysis methods procedures described in the PhD thesis can be applied only for validated stoichiometric model, because each procedure step can be applied only in case of steady state.

Several directions can be highlighted as **future development prospects**:

- The improved stoichiometric model development and validation procedure can be adapted to signaling networks, transcriptional networks and gene expression networking needs.
- Visualization tool can be modified for manually or automatically created metabolic network representation of genome-scale stoichiometric models.
- Upgrade and integrate metabolic network analysis methods procedures to be able to use different types of omics experimental data in stoichiometric models.

# LITERATŪRA

## BIBLIOGRAPHY

1. Abedon, S.T. (2009) Kinetics of phage-mediated biocontrol of bacteria. *Foodborne pathogens and disease*, Vol. 6(7), pp. 807–15.
2. Almaas, E., Kovács, B., Vicsek, T., Oltvai, Z.N., Barabási, A.-L. (2004) Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature*, Vol. 427(6977), pp. 839–43.
3. Babur, O., Dogrusoz, U., Demir, E., Sander, C. (2009) ChiBE: Interactive visualization and manipulation of BioPAX pathway models. *Bioinformatics*, 26, pp.429–431.
4. Baena-González, E., Rolland, F., Thevelein, J.M., Sheen, J. (2007) A central integrator of transcription networks in plant stress and energy signalling. *Nature*, Vol. 448(7156), p. 938–42.
5. Barrett, T. et al. (2013) NCBI GEO: archive for functional genomics data sets--update. *Nucleic acids research*, Vol. 41(Database issue), p. D991–5.
6. Bates, J.T., Chivian, D., Arkin, A.P. (2011) GLAMM: Genome-Linked Application for Metabolic Maps. *Nucleic Acids Research*, Vol. 39.
7. Bates, J.T., Chivian, D., Arkin, A.P. (2011) GLAMM: Genome-Linked Application for Metabolic Maps. *Nucleic acids research*, Vol. 39(Web Server issue), p. W400–5.
8. Beard, D.A., Vinnakota, K.C., Wu, F. (2008) Detailed enzyme kinetics in terms of biochemical species: study of citrate synthase. *PLoS one*, Vol. 3(3), p. e1825.
9. Becker, S.A., Palsson, B.O. (2008) Context-specific metabolic networks are consistent with experiments. *PLoS computational biology*, Vol. 4(5), p. e1000082.
10. Belford, P. (2006) Candlestick stock analysis with genetic algorithms. In: *Proceedings of the 8th annual conference on Genetic and evolutionary computation - GECCO '06*, New York, New York, USA: ACM Press, p. 1851.
11. Bergmann, F.T., Vallabhajosyula, R.R., Sauro, H.M. (2006) Computational Tools for Modeling Protein Networks. *Current Proteomics*, 3, pp.181–197.
12. Bernstein, B.E. et al. (2010) The NIH Roadmap Epigenomics Mapping Consortium. *Nature biotechnology*, Vol. 28(10), p. 1045–8.
13. Blazier, A.S., Papin, J.A. (2012) Integration of expression data in genome-scale metabolic network reconstructions. *Frontiers in physiology*, Vol. 3, p. 299.
14. Boerries, M., Eils, R., Busch, H. (2010) Systems Biology.
15. Boiteux, A., Hess, B. (1981) Design of Glycolysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 293(1063), pp. 5–22.
16. Bonarius, H.P.J., Schmid, G., Tramper, J. (1997) Flux analysis of underdetermined metabolic networks: the quest for the missing constraints. *Trends in Biotechnology*, Vol. 15(8), pp. 308–314.
17. Bordbar, A., Feist, A.M., Usaite-Black, R., Woodcock, J., Palsson, B.O., Famili, I. (2011) A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. *BMC systems biology*, Vol. 5, p. 180.
18. Bordbar, A., Lewis, N.E., Schellenberger, J., Palsson, B.Ø., Jamshidi, N. (2010) Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Molecular systems biology*, Vol. 6, p. 422.
19. Bordbar, A., Monk, J.M., King, Z.A., Palsson, B.O. (2014) Constraint-based models predict metabolic and associated cellular functions. *Nature reviews. Genetics*, Vol. 15(2), pp. 107–20.

20. Boyer, P.D. (1998) Energy, Life, and ATP (Nobel Lecture). *Angewandte Chemie International Edition*, Vol. 37(17), pp. 2296–2307.
21. Burgard, A.P., Pharkya, P., Maranas, C.D. (2003) OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. *Biotechnology and Bioengineering*, Vol. 84, pp. 647–657.
22. Burgard, A.P., Vaidyaraman, S., Maranas, C.D. (2001) Minimal reaction sets for *Escherichia coli* metabolism under different growth requirements and uptake environments. *Biotechnology progress*, Vol. 17(5), p. 791–7.
23. Carbonell, P., Fichera, D., Pandit, S.B., Faulon, J.-L. (2012) Enumerating metabolic pathways for the production of heterologous target chemicals in chassis organisms. *BMC systems biology*, Vol. 6, p. 10.
24. Caspi, R. et al. (2014) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic acids research*, Vol. 42(Database issue), p. D459–71.
25. Chandran, D., Sauro, H.M. (2012) Hierarchical modeling for synthetic biology. *ACS Synthetic Biology*, Vol. 1, pp. 353–364.
26. Chaouiya, C. et al. (2013) SBML qualitative models: a model representation format and infrastructure to foster interactions between qualitative modelling formalisms and tools. *BMC systems biology*, Vol. 7, p. 135.
27. Clarke, R., Ransom, H.W., Wang, A., Xuan, J., Liu, M.C., Gehan, E.A., Wang, Y. (2008) The properties of high-dimensional data spaces: implications for exploring gene and protein expression data. *Nature reviews. Cancer*, Vol. 8(1), pp. 37–49.
28. Cline, M.S. et al. (2007) Integration of biological networks and gene expression data using Cytoscape. *Nature protocols*, Vol. 2, pp. 2366–2382.
29. Colijn, C. et al. (2009) Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production. *PLoS computational biology*, Vol. 5(8), p. e1000489.
30. Cotten, C., Reed, J.L. (2013) Mechanistic analysis of multi-omics datasets to generate kinetic parameters for constraint-based metabolic models. *BMC bioinformatics*, Vol. 14, p. 32.
31. Craigon, D.J., James, N., Okyere, J., Higgins, J., Jotham, J., May, S. (2004) NASCArrays: a repository for microarray data generated by NASC’s transcriptomics service. *Nucleic acids research*, Vol. 32(Database issue), p. D575–7.
32. Croft, D. et al. (2014) The Reactome pathway knowledgebase. *Nucleic Acids Research*, Vol. 42.
33. D A Fell and J R Small (1986) Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochemical Journal*, Vol. 238, pp. 781–786.
34. Dandekar, T., Schuster, S., Snel, B., Huynen, M., Bork, P. (1999) Pathway alignment: application to the comparative analysis of glycolytic enzymes. *The Biochemical journal*, 343 Pt 1, p. 115–24.
35. Das, I., Dennis, J.E. (1998) Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems. *SIAM Journal on Optimization*, Vol. 8(3), pp. 631–657.
36. Dirks, R.M., Bois, J.S., Schaeffer, J.M., Winfree, E., Pierce, N.A. (2007) Thermodynamic Analysis of Interacting Nucleic Acid Strands. *SIAM Review*, 49, pp.65–88.
37. Droste, P., Miebach, S., Niedenfür, S., Wiechert, W., Nöh, K. (2011) Visualizing multi-omics data in metabolic networks with the software Omix-A case study. *BioSystems*, Vol. 105, pp. 154–161.

38. Duarte, N.C., Becker, S.A., Jamshidi, N., Thiele, I., Mo, M.L., Vo, T.D., Srivas, R., Palsson, B.Ø. (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 104(6), p. 1777–82.
39. Durot, M., Bourguignon, P.-Y., Schachter, V. (2009) Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS microbiology reviews*, Vol. 33(1), p. 164–90.
40. Ebrahim, A., Lerman, J.A., Palsson, B.O., Hyduke, D.R. (2013) COBRApy: COntstraints-Based Reconstruction and Analysis for Python. *BMC systems biology*, Vol. 7, p. 74.
41. Edwards, J.S., Covert, M., Palsson, B. (2002) Metabolic modelling of microbes: the flux-balance approach. *Environmental Microbiology*, Vol. 4(3), pp. 133–140.
42. Edwards, J.S., Ibarra, R.U., Palsson, B.O. (2001) In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. *Nature biotechnology*, Vol. 19(2), p. 125–30.
43. Edwards, J.S., Palsson, B.O. (1999) Systems Properties of the Haemophilus influenzae Rd Metabolic Genotype. *Journal of Biological Chemistry*, Vol. 274(25), pp. 17410–17416.
44. Egelhofer, V., Schomburg, I., Schomburg, D. (2010) Automatic assignment of EC numbers. *PLoS computational biology*, Vol. 6(1), p. e1000661.
45. Entezari, M.H., Pétrier, C. (2004) A combination of ultrasound and oxidative enzyme: sono-biodegradation of phenol. *Applied Catalysis B: Environmental*, Vol. 53(4), pp. 257–263.
46. Farasat, I., Kushwaha, M., Collens, J., Easterbrook, M., Guido, M., Salis, H.M. (2014) Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. *Molecular systems biology*, Vol. 10, p. 731.
47. Feist, A.M. et al. (2007) A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular systems biology*, Vol. 3, p. 121.
48. Feist, A.M., Herrgård, M.J., Thiele, I., Reed, J.L., Palsson, B.Ø. (2009) Reconstruction of biochemical networks in microorganisms. *Nature reviews. Microbiology*, Vol. 7(2), p. 129–43.
49. Fell, D.A. (1992) Metabolic control analysis: a survey of its theoretical and experimental development. *Biochemical Journal*, Vol. 286, pp. 313–330.
50. Fell, D.A. (1993) *Modern Trends in Biothermokinetics*. Boston, MA: Springer US.
51. Fell, D.A. (1997) *Understanding the Control of Metabolism*. 1st ed. London: Portland Press Ltd.
52. Ferrannini, E. (1988) The theoretical bases of indirect calorimetry: A review. *Metabolism*, Vol. 37(3), pp. 287–301.
53. Fischer, E., Zamboni, N., Sauer, U. (2004) High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived <sup>13</sup>C constraints. *Analytical biochemistry*, Vol. 325(2), p. 308–16.
54. Flamholz, A., Noor, E., Bar-Even, A., Milo, R. (2012) eQuilibrator--the biochemical thermodynamics calculator. *Nucleic acids research*, Vol. 40(Database issue), p. D770–5.
55. Fleischmann, A. et al. (2004) IntEnz, the integrated relational enzyme database. *Nucleic acids research*, Vol. 32(Database issue), p. D434–7.
56. Fleischmann, R.D. et al. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science (New York, N.Y.)*, Vol. 269(5223), pp. 496–512.

57. Foretz, M. et al. (2010) Metformin inhibits hepatic gluconeogenesis in mice independently of the LKB1/AMPK pathway via a decrease in hepatic energy state. *The Journal of clinical investigation*, Vol. 120(7), p. 2355–69.
58. Funahashi, A., Matsuoka, Y., Jouraku, A., Morohashi, M., Kikuchi, N., Kitano, H. (2008) CellDesigner 3.5: A versatile modeling tool for biochemical networks. *Proceedings of the IEEE*, Vol. 96, pp. 1254–1265.
59. Funahashi, A., Matsuoka, Y., Jouraku, A., Morohashi, M., Kikuchi, N., Kitano, H. (2008) CellDesigner 3.5: A Versatile Modeling Tool for Biochemical Networks. *Proceedings of the IEEE*, Vol. 96(8), pp. 1254–1265.
60. Garcia-Albornoz, M., Thankaswamy-Kosalai, S., Nilsson, A., Våremo, L., Nookaew, I., Nielsen, J. (2014) BioMet Toolbox 2.0: Genome-wide analysis of metabolism and omics data. *Nucleic Acids Research*, Vol. 42.
61. Garcia-Viloca, M., Gao, J., Karplus, M., Truhlar, D.G. (2004) How enzymes work: analysis by modern rate theory and computer simulations. *Science (New York, N.Y.)*, Vol. 303(5655), p. 186–95.
62. Gasteiger, E. (2003) ExpASY: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, Vol. 31(13), pp. 3784–3788.
63. Ge, H., Liu, Z., Church, G.M., Vidal, M. (2001) Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nature genetics*, Vol. 29(4), p. 482–6.
64. Gebauer, J., Schuster, S., de Figueiredo, L.F., Kaleta, C. (2012) Detecting and investigating substrate cycles in a genome-scale human metabolic network. *The FEBS journal*, Vol. 279(17), p. 3192–202.
65. Genee, H.J., Bonde, M.T., Bagger, F.O., Jespersen, J.B., Sommer, M.O. a, Wernersson, R., Olsen, L.R. (2014) Software-supported USER cloning strategies for site-directed mutagenesis and DNA assembly. *ACS synthetic biology*.
66. Georgi, T., Rittmann, D., Wendisch, V.F. (2005) Lysine and glutamate production by *Corynebacterium glutamicum* on glucose, fructose and sucrose: roles of malic enzyme and fructose-1,6-bisphosphatase. *Metabolic engineering*, Vol. 7(4), pp. 291–301.
67. Ghosh, S., Matsuoka, Y., Asai, Y., Hsin, K.-Y., Kitano, H. (2011) Software for systems biology: from tools to integrated platforms. *Nature reviews. Genetics*, Vol. 12(12), p. 821–32.
68. Golightly, A., Gillespie, C.S. (2013) Simulation of stochastic kinetic models. *Methods in molecular biology (Clifton, N.J.)*, Vol. 1021, p. 169–87.
69. Gombert, A.K., Moreira dos Santos, M., Christensen, B., Nielsen, J. (2001) Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *Journal of bacteriology*, Vol. 183(4), p. 1441–51.
70. De Graaf, A.A., Striegel, K., Wittig, R.M., Laufer, B., Schmitz, G., Wiechert, W., Sprenger, G.A., Sahm, H. (1999) Metabolic state of *Zymomonas mobilis* in glucose-, fructose-, and xylose-fed continuous cultures as analysed by <sup>13</sup>C- and <sup>31</sup>P-NMR spectroscopy. *Archives of microbiology*, Vol. 171(6), p. 371–85.
71. Gunawardena, J. (2012) Some lessons about models from Michaelis and Menten. *Molecular biology of the cell*, Vol. 23(4), p. 517–9.
72. Gustafsson, C. (2009) For anyone who ever said there's no such thing as a poetic gene. *Nature*, 458, pp.703–703.
73. Haki, G. (2003) Developments in industrially important thermostable enzymes: a review. *Bioresource Technology*, Vol. 89(1), pp. 17–34.

74. Hastings, J. et al. (2013) The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic acids research*, Vol. 41(Database issue), p. D456–63.
75. Hastay, J., McMillen, D., Isaacs, F., Collins, J.J. (2001) Computational studies of gene regulatory networks: in numero molecular biology. *Nature reviews. Genetics*, Vol. 2(4), p. 268–79.
76. Heller, S., McNaught, A., Stein, S., Tchekhovskoi, D., Pletnev, I. (2013) InChI - the worldwide chemical structure identifier standard. *Journal of cheminformatics*, Vol. 5(1), p. 7.
77. Hoover, D.M., Lubkowski, J. (2002) DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic acids research*, Vol. 30, p. e43.
78. Hoppe, A., Hoffmann, S., Gerasch, A., Gille, C., Holzhütter, H.-G. (2011) FASIMU: flexible software for flux-balance computation series in large metabolic networks. *BMC bioinformatics*, Vol. 12, p. 28.
79. Hosseini, N., Ebrahimi, S., Salehi, P., Asghari, B., Ahmadi, M., Salhi Arjmand, H. (2008) Simultaneous determination of atropine and scopolamine in Hyoscyamus arachnoideus by HPLC. *Planta Medica*, Vol. 74(09).
80. Howe, G.A., Jander, G. (2008) Plant immunity to insect herbivores. *Annual review of plant biology*, Vol. 59, pp. 41–66.
81. Hsiung, H.M., Cantrell, A., Luirink, J., Oudega, B., Veros, A.J., Becker, G.W. (1989) Use of Bacteriocin Release Protein in E. Coli for Excretion of Human Growth Hormone into the Culture Medium. *Bio/Technology*, Vol. 7(3), pp. 267–271.
82. Hung, J.-H. (2013) Gene Set/Pathway enrichment analysis. *Methods in molecular biology (Clifton, N.J.)*, Vol. 939, p. 201–13.
83. Huntley, R.P., Sawford, T., Mutowo-Meullenet, P., Shypitsyna, A., Bonilla, C., Martin, M.J., O'Donovan, C. (2015) The GOA database: Gene Ontology annotation updates for 2015. *Nucleic Acids Research*.
84. Hyduke, D.R., Lewis, N.E., Pálsson, B.Ø. (2013) Analysis of omics data with genome-scale models of metabolism. *Molecular bioSystems*, Vol. 9(2), p. 167–74.
85. Ibarra, R.U., Edwards, J.S., Pálsson, B.O. (2002) Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*, Vol. 420(6912), p. 186–9.
86. Ideker, T., Krogan, N.J. (2012) Differential network biology. *Molecular systems biology*, Vol. 8, p. 565.
87. Iyer, P. V., Ananthanarayan, L. (2008) Enzyme stability and stabilization—Aqueous and non-aqueous environment. *Process Biochemistry*, Vol. 43(10), pp. 1019–1032.
88. James, D. (2010) *Crafting Digital Media*. Berkeley, CA: Apress.
89. Jamshidi, N., Pálsson, B.Ø. (2008) Formulating genome-scale kinetic models in the post-genome era. *Molecular systems biology*, Vol. 4, p. 171.
90. Jensen, P.A., Papin, J.A. (2014) MetDraw: automated visualization of genome-scale metabolic network reconstructions and high-throughput data. *Bioinformatics (Oxford, England)*, Vol. 30(9), p. 1327–8.
91. Jensen, P.R., Williams, P.G., Oh, D.-C., Zeigler, L., Fenical, W. (2007) Species-specific secondary metabolite production in marine actinomycetes of the genus Salinispora. *Applied and environmental microbiology*, Vol. 73(4), p. 1146–52.
92. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabási, A.L. (2000) The large-scale organization of metabolic networks. *Nature*, Vol. 407(6804), p. 651–4.

93. Johnson-Kerner, B.L. et al. (2014) Intermediate filament protein accumulation in motor neurons derived from giant axonal neuropathy iPSCs rescued by restoration of gigaxonin. *Human molecular genetics*.
94. De Jong, H. (2002) Modeling and simulation of genetic regulatory systems: a literature review. *Journal of computational biology: a journal of computational molecular cell biology*, Vol. 9(1), pp. 67–103.
95. De Jongh, W.A., Nielsen, J. (2008) Enhanced citrate production through gene insertion in *Aspergillus niger*. *Metabolic engineering*, Vol. 10(2), pp. 87–96.
96. Joyce, A.R., Palsson, B.Ø. (2006) The model organism as a system: integrating “omics” data sets. *Nature reviews. Molecular cell biology*, Vol. 7(3), pp. 198–210.
97. Kakkar Tarundeep, Boxenbaum Harold, M.M. (1998) ESTIMATION OF K I IN A COMPETITIVE ENZYME-INHIBITION MODEL: COMPARISONS AMONG THREE METHODS OF DATA ANALYSIS. *DRUG METABOLISM AND DISPOSITION*, Vol. 27(6), pp. 756–762.
98. Kalnenieks, U., Pentjuss, A., Rutkis, R., Stalidzans, E., Fell, D.A. (2014) Modeling of *Zymomonas mobilis* central metabolism for novel metabolic engineering strategies. *Frontiers in microbiology*, Vol. 5, p. 42.
99. Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M. (2014) Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Research*, Vol. 42.
100. Keating, K.A., Quinn, J.F. (1998) Estimating Species Richness: The Michaelis-Menten Model Revisited. *Oikos*, Vol. 81(2), p. 411.
101. Kell, D.B., Mendes, P. (1999) SNAPSHOTS OF SYSTEMS - METABOLIC CONTROL ANALYSIS AND BIOTECHNOLOGY IN THE POST-GENOMIC ERA.
102. Kerkhoven, E.J., Lahtvee, P.-J., Nielsen, J. (2014) Applications of computational modeling in metabolic engineering of yeast. *FEMS Yeast Research*, p. n/a–n/a.
103. Klamt, S. (2006) Generalized concept of minimal cut sets in biochemical networks. *Bio Systems*, Vol. 83(2-3), p. 233–47.
104. Klamt, S., von Kamp, A. (2011) An application programming interface for CellNetAnalyzer. *Bio Systems*, Vol. 105(2), p. 162–8.
105. Klamt, S., Saez-Rodriguez, J., Gilles, E.D. (2007) Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC systems biology*, Vol. 1, p. 2.
106. Klitgord, N., Segrè, D. (2010) Environments that induce synthetic microbial ecosystems. *PLoS computational biology*, Vol. 6(11), p. e1001002.
107. Klukas, C., Schreiber, F. (2010) Integration of -omics data and networks for biomedical research with VANTED. *Journal of integrative bioinformatics*, Vol. 7, p. 112.
108. Kumar, A., Suthers, P.F., Maranas, C.D. (2012) MetRxn: a knowledgebase of metabolites and reactions spanning metabolic models and databases. *BMC Bioinformatics*, 13, p.6.
109. Kumar, V.S., Maranas, C.D. (2009) GrowMatch: An automated method for reconciling in silico/in vivo growth predictions. *PLoS Computational Biology*, Vol. 5.
110. Leflaive, J., Ten-hage, L. (2007) Algal and cyanobacterial secondary metabolites in freshwaters: a comparison of allelopathic compounds and toxins. *Freshwater Biology*, Vol. 52(2), pp. 199–214.
111. Leiser, J., Blum, J.J. (1987) On the analysis of substrate cycles in large metabolic systems. *Cell Biophysics*, Vol. 11(1), pp. 123–138.
112. Lewis, N.E. et al. (2010) Large-scale in silico modeling of metabolic interactions between cell types in the human brain. *Nature biotechnology*, Vol. 28(12), p. 1279–85.



113. Lewis, N.E., Nagarajan, H., Palsson, B.O. (2012) Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nature reviews. Microbiology*, Vol. 10(4), pp. 291–305.
114. Li, F., Thiele, I., Jamshidi, N., Palsson, B.Ø. (2009) Identification of potential pathway mediation targets in Toll-like receptor signaling. *PLoS computational biology*, Vol. 5(2), p. e1000292.
115. Liao, J.C., Hou, S.Y., Chao, Y.P. (1996) Pathway analysis, engineering, and physiological considerations for redirecting central metabolism. *Biotechnology and bioengineering*, Vol. 52(1), p. 129–40.
116. Lieber, D.J., Catlett, J., Madayiputhiya, N., Nandakumar, R., Lopez, M.M., Metcalf, W.W., Buan, N.R. (2014) A multienzyme complex channels substrates and electrons through acetyl-CoA and methane biosynthesis pathways in *Methanosarcina*. *PLoS one*, Vol. 9(9), p. e107563.
117. Liebermeister, W., Klipp, E. (2006) Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theoretical biology & medical modelling*, Vol. 3, p. 41.
118. Link, H., Christodoulou, D., Sauer, U. (2014) Advancing metabolic models with kinetic information. *Current opinion in biotechnology*, Vol. 29C, pp. 8–14.
119. Liu, J.K., O'Brien, E.J., Lerman, J.A., Zengler, K., Palsson, B.O., Feist, A.M. (2014) Reconstruction and modeling protein translocation and compartmentalization in *Escherichia coli* at the genome-scale. *BMC Systems Biology*, Vol. 8(1), p. 110.
120. Longabaugh, W.J.R., Davidson, E.H., Bolouri, H. (2009) Visualization, documentation, analysis, and communication of large-scale gene regulatory networks. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, Vol. 1789, pp. 363–374.
121. Low, A.G. (2007) The activity of pepsin, chymotrypsin and trypsin during 24 h periods in the small intestine of growing pigs. *British Journal of Nutrition*, Vol. 48(01), p. 147.
122. Ludäscher, B., Raschid, L. eds. (2005) *Data Integration in the Life Sciences*. Berlin, Heidelberg: Springer Berlin Heidelberg.
123. Machado, D., Herrgård, M. (2014) Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS computational biology*, Vol. 10(4), p. e1003580.
124. Mahadevan, R., Schilling, C.H. (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic Engineering*, Vol. 5(4), pp. 264–276.
125. Majewski, R.A., Domach, M.M. (1990) Simple constrained-optimization view of acetate overflow in *E. coli*. *Biotechnology and bioengineering*, Vol. 35(7), p. 732–8.
126. Mendes, P., Hoops, S., Sahle, S., Gauges, R., Dada, J., Kummer, U. (2009a) Computational modeling of biochemical networks using COPASI. *Methods in Molecular Biology*, Vol. 500, pp. 17–59.
127. Mendes, P., Hoops, S., Sahle, S., Gauges, R., Dada, J., Kummer, U. (2009b) Computational modeling of biochemical networks using COPASI. *Methods in molecular biology (Clifton, N.J.)*, Vol. 500, pp. 17–59.
128. Mo, M.L., Palsson, B.O., Herrgård, M.J. (2009) Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC systems biology*, Vol. 3, p. 37.
129. Moodie, S., Moodie, S., Le Novere, N., Demir, E., Mi, H., Villeger, A. (2011) Systems Biology Graphical Notation: Process Description language Level 1. *Nature Precedings*.
130. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., Kanehisa, M. (2007) KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, Vol. 35.

131. Nagasaki, M., Saito, A., Jeong, E., Li, C., Kojima, K., Ikeda, E., Miyano, S. (2010a) Cell illustrator 4.0: A computational platform for systems biology. *In Silico Biology*, Vol. 10, pp. 5–26.
132. Nagasaki, M., Saito, A., Jeong, E., Li, C., Kojima, K., Ikeda, E., Miyano, S. (2010b) Cell Illustrator 4.0: a computational platform for systems biology. *In silico biology*, Vol. 10(1), pp. 5–26.
133. Newsholme, E.A., Challiss, R.A.J., Crabtree, B. (1984) Substrate cycles: their role in improving sensitivity in metabolic control. *Trends in Biochemical Sciences*, Vol. 9(6), pp. 277–280.
134. Nielsen, J. (2003) It Is All about Metabolic Fluxes. *Journal of Bacteriology*, Vol. 185(24), pp. 7031–7035.
135. Nielsen, J., Jewett, M.C. (2008) Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. *FEMS yeast research*, Vol. 8(1), p. 122–31.
136. Le Novère, N. et al. (2009) The Systems Biology Graphical Notation. *Nature biotechnology*, Vol. 27(8), p. 735–41.
137. Olivier, B.G., Rohwer, J.M., Hofmeyr, J.H.S. (2005) Modelling cellular systems with PySCeS. *Bioinformatics*, Vol. 21, pp. 560–561.
138. Olson, J.A. (1997) An Analysis of the Algebraic Method for Balancing Chemical Reactions. *Journal of Chemical Education*, Vol. 74(5), p. 538.
139. Overbeek, R. et al. (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Research*, Vol. 33, pp. 5691–5702.
140. Overbeek, R. et al. (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic acids research*, Vol. 42(Database issue), p. D206–14.
141. Pál, C., Papp, B., Lercher, M.J. (2005) Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nature genetics*, Vol. 37(12), p. 1372–5.
142. Palsson, B.Ø. (2011) *Systems Biology: Simulation of Dynamic Network States*. New York: Cambridge University Press.
143. Papin, J.A., Palsson, B.O. (2004) The JAK-STAT signaling network in the human B-cell: an extreme signaling pathway analysis. *Biophysical journal*, Vol. 87(1), pp. 37–46.
144. Papin, J.A., Price, N.D., Wiback, S.J., Fell, D.A., Palsson, B.O. (2003) Metabolic pathways in the post-genome era. *Trends in biochemical sciences*, Vol. 28(5), p. 250–8.
145. Pentjuss, A., Kalnenieks, U. (2013) Assessment of *Zymomonas mobilis* biotechnological potential in ethanol production by flux variability analysis. *Biosystems and Information technology*, Vol. 3(1), pp. 1–5.
146. Pentjuss, A., Odzina, I., Kostromins, A., Fell, D.A., Stalidzans, E., Kalnenieks, U. (2013) Biotechnological potential of respiring *Zymomonas mobilis*: a stoichiometric analysis of its central metabolism. *Journal of biotechnology*, Vol. 165(1), pp. 1–10.
147. Pentjuss, A., Rubenis, O., Bauze, D., Aprupe, L., Lace, B. (2013) Flux variability analysis approach of autism related metabolism in stoichiometric model of mitochondria. *Biosystems and Information technology*, Vol. 2(2), pp. 37–42.
148. Pentjuss, A., Odzina, I., Gailums, A. (2012) Multi agent genome-scale metabolic reconstruction modeling software schema. In: *11th International Scientific Conference on Engineering for Rural Development*, Jelgava, pp. 189–194.
149. Pfeiffer, T., Sánchez-Valdenebro, I., Nuño, J.C., Montero, F., Schuster, S. (1999) METATOOL: for studying metabolic networks. *Bioinformatics (Oxford, England)*, Vol. 15(3), p. 251–7.

150. Pharkya, P., Burgard, A.P., Maranas, C.D. (2004) OptStrain: A computational framework for redesign of microbial production systems. *Genome Research*, Vol. 14, pp. 2367–2376.
151. Poolman, M.G. (2006) ScrumPy: metabolic modelling with Python. *Systems biology*, Vol. 153(5), p. 375–8.
152. Poolman, M.G., Sebu, C., Pidcock, M.K., Fell, D.A. (2007) Modular decomposition of metabolic systems via null-space analysis. *Journal of theoretical biology*, Vol. 249(4), pp. 691–705.
153. Pöpkén, T., Götze, L., Gmehling, J. (2000) Reaction Kinetics and Chemical Equilibrium of Homogeneously and Heterogeneously Catalyzed Acetic Acid Esterification with Methanol and Methyl Acetate Hydrolysis. *Industrial & Engineering Chemistry Research*, Vol. 39(7), pp. 2601–2611.
154. Price, N.D., Papin, J.A., Schilling, C.H., Palsson, B.O. (2003) Genome-scale microbial in silico models: the constraints-based approach. *Trends in biotechnology*, Vol. 21(4), p. 162–9.
155. Price, T.D., Qvarnström, A., Irwin, D.E. (2003) The role of phenotypic plasticity in driving genetic evolution. *Proceedings. Biological sciences / The Royal Society*, Vol. 270(1523), p. 1433–40.
156. Pruitt, K.D., Tatusova, T., Maglott, D.R. (2007) NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, Vol. 35.
157. Quek, L.-E., Nielsen, L.K. (2014) A depth-first search algorithm to compute elementary flux modes by linear programming. *BMC systems biology*, Vol. 8, p. 94.
158. Quek, L.-E., Wittmann, C., Nielsen, L.K., Krömer, J.O. (2009) OpenFLUX: efficient modelling software for <sup>13</sup>C-based metabolic flux analysis. *Microbial cell factories*, Vol. 8, p. 25.
159. Racine, J. (2006) gnuplot 4.0: a portable interactive plotting utility. *Journal of Applied Econometrics*, Vol. 21(1), pp. 133–141.
160. Rahman, S.A., Cuesta, S.M., Furnham, N., Holliday, G.L., Thornton, J.M. (2014) EC-BLAST: a tool to automatically search and compare enzyme reactions. *Nature methods*, Vol. 11(2), p. 171–4.
161. Raza, S. et al. (2010) Construction of a large scale integrated map of macrophage pathogen recognition and effector systems. *BMC systems biology*, Vol. 4, p. 63.
162. Reder, C. (1988) Metabolic control theory: A structural approach. *Journal of Theoretical Biology*, Vol. 135(2), pp. 175–201.
163. Ren, Q., Paulsen, I.T. (2007) Large-scale comparative genomic analyses of cytoplasmic membrane transport systems in prokaryotes. In: *Journal of Molecular Microbiology and Biotechnology*, pp. 165–179.
164. Reznik, E., Segre, D., Sherwood, W.E. (2011) The Quasi-Steady State Assumption in an Enzymatically Open System., p. 28.
165. Ricquier, D. (2007) Respiration uncoupling and metabolism in the control of energy expenditure. *Proceedings of the Nutrition Society*, Vol. 64(01), pp. 47–52.
166. Rocha, I. et al. (2010) OptFlux: an open-source software platform for in silico metabolic engineering. *BMC systems biology*, Vol. 4, p. 45.
167. Sauer, U. (2006) Metabolic networks in motion: <sup>13</sup>C-based flux analysis. *Molecular systems biology*, Vol. 2, p. 62.
168. Sauro, H.M., Kholodenko, B.N. (2004) Quantitative analysis of signaling networks. *Progress in biophysics and molecular biology*, Vol. 86(1), pp. 5–43.

169. Schellenberger, J. et al. (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature protocols*, Vol. 6(9), p. 1290–307.
170. Schellenberger, J., Park, J.O., Conrad, T.M., Palsson, B.Ø. (2010) BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. *BMC bioinformatics*, Vol. 11, p. 213.
171. Schilling, C.H., Letscher, D., Palsson, B.O. (2000) Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *Journal of theoretical biology*, Vol. 203(3), p. 229–48.
172. Schilling, C.H., Schuster, S., Palsson, B.O., Heinrich, R. (1999) Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. *Biotechnology progress*, Vol. 15(3), pp. 296–303.
173. Schomburg, I. et al. (2013) BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. *Nucleic acids research*, Vol. 41(Database issue), p. D764–72.
174. Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M., Sauer, U. (2012) Multidimensional optimality of microbial metabolism. *Science (New York, N.Y.)*, Vol. 336(6081), p. 601–4.
175. Schuster, S. (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends in Biotechnology*, Vol. 17(2), pp. 53–60.
176. Schuster, S., Dandekar, T., Fell, D.A. (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends in biotechnology*, Vol. 17(2), pp. 53–60.
177. Schuster, S., Fell, D.A., Dandekar, T. (2000) A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature biotechnology*, Vol. 18(3), p. 326–32.
178. Schuster, S., Hilgetag, C. (1994) ON ELEMENTARY FLUX MODES IN BIOCHEMICAL REACTION SYSTEMS AT STEADY STATE. *Journal of Biological Systems*, Vol. 02(02), pp. 165–182.
179. Segrè, D., Vitkup, D., Church, G.M. (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 99(23), p. 15112–7.
180. Selkov, E. et al. (1997) The metabolic pathway collection: An update. *Nucleic Acids Research*, Vol. 25, pp. 37–38.
181. Seo, J.-S. et al. (2005) The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nature biotechnology*, Vol. 23(1), p. 63–8.
182. Shih, P.M. et al. (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 110(3), p. 1053–8.
183. Shlomi, T., Cabili, M.N., Herrgård, M.J., Palsson, B.Ø., Ruppin, E. (2008) Network-based prediction of human tissue-specific metabolism. *Nature biotechnology*, Vol. 26(9), p. 1003–10.
184. Smith, S.R. (2009) A critical review of the bioavailability and impacts of heavy metals in municipal solid waste composts compared to sewage sludge. *Environment international*, Vol. 35(1), p. 142–56.
185. Srivastava, R., You, L., Summers, J., Yin, J. (2002) Stochastic vs. Deterministic Modeling of Intracellular Viral Kinetics. *Journal of Theoretical Biology*, Vol. 218(3), pp. 309–321.

186. Stanzak, R., Matsushima, P., Baltz, R.H., Rao, R.N. (1986) Cloning and Expression in *Streptomyces lividans* of Clustered Erythromycin Biosynthesis Genes from *Streptomyces erythreus*. *Bio/Technology*, Vol. 4(3), pp. 229–232.
187. Stearns Stephen C (1989) The Evolutionary Significance of Phenotypic Plasticity. *BioScience*, Vol. 39(7), pp. 436–445.
188. Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., Gilles, E.D. (2002) Metabolic network structure determines key aspects of functionality and regulation. *Nature*, Vol. 420(6912), p. 190–3.
189. Sungman Cha (1968) A Simple Method for Derivation of Rate Equations for Enzyme-catalyzed Reactions under the Rapid Equilibrium Assumption or Combined Assumptions of Equilibrium and Steady State. *The journal of biological chemistry*, Vol. 243, pp. 820–825.
190. Suthers, P.F., Zomorodi, A., Maranas, C.D. (2009) Genome-scale gene/reaction essentiality and synthetic lethality analysis. *Molecular systems biology*, Vol. 5, p. 301.
191. Terzer, M., Stelling, J. (2008) Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics (Oxford, England)*, Vol. 24(19), p. 2229–35.
192. Thiele, I. et al. (2013) A community-driven global reconstruction of human metabolism. *Nature biotechnology*, Vol. 31(5), p. 419–25.
193. Thiele, I., Jamshidi, N., Fleming, R.M.T., Palsson, B.Ø. (2009) Genome-scale reconstruction of *Escherichia coli*'s transcriptional and translational machinery: a knowledge base, its mathematical formulation, and its functional characterization. *PLoS computational biology*, Vol. 5(3), p. e1000312.
194. Thiele, I., Palsson, B.Ø. (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature protocols*, Vol. 5(1), pp. 93–121.
195. Thiele, I., Price, N.D., Vo, T.D., Palsson, B.Ø. (2005) Candidate metabolic network states in human mitochondria. Impact of diabetes, ischemia, and diet. *The Journal of biological chemistry*, Vol. 280(12), p. 11683–95.
196. Thomas Simon, Mooney J. F. Peter, Burrell M. Michael, F.A.D. (1997) Metabolic Control Analysis of glycolysis in tuber tissue of potato (*Solanum tuberosum*): explanation for the low control coefficient of phosphofructokinase over respiratory flu. *Biochemical Journal*, Vol. 322, pp. 111–117.
197. Trinh, C.T., Sreenc, F. (2009) Metabolic engineering of *Escherichia coli* for efficient conversion of glycerol to ethanol. *Applied and Environmental Microbiology*, Vol. 75(21), pp. 6696–6705.
198. Trinh, C.T., Unrean, P., Sreenc, F. (2008) Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. *Applied and environmental microbiology*, Vol. 74(12), p. 3634–43.
199. Unrean, P., Trinh, C.T., Sreenc, F. (2010) Rational design and construction of an efficient *E. coli* for production of diapolycopendioic acid. *Metabolic engineering*, Vol. 12(2), p. 112–22.
200. Vīgants, A. (2008) *Starpaugstskolu akadēmiskā maģistra studiju programma "UZTURZINĀTNE."* Rīga.
201. Villéger, A.C., Pettifer, S.R., Kell, D.B. (2010) Arcadia: A visualization tool for metabolic pathways. *Bioinformatics*, Vol. 26, pp. 1470–1471.
202. Villéger, A.C., Pettifer, S.R., Kell, D.B. (2010) Arcadia: a visualization tool for metabolic pathways. *Bioinformatics (Oxford, England)*, Vol. 26(11), p. 1470–1.
203. Welch, G.R. (1985) Some problems in the usage of gibbs free energy in biochemistry. *Journal of Theoretical Biology*, Vol. 114(3), pp. 433–446.

204. Wheeler, D.L. et al. (2004) Database resources of the National Center for Biotechnology Information: update. *Nucleic acids research*, Vol. 32(Database issue), p. D35–D40.
205. Wiechert, W. (2007) The thermodynamic meaning of metabolic exchange fluxes. *Biophysical journal*, Vol. 93(6), p. 2255–64.
206. Wilkinson, D.J. (2007) Bayesian methods in bioinformatics and computational systems biology. *Briefings in bioinformatics*, Vol. 8(2), p. 109–16.
207. Wink, M. (2003) Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry*, Vol. 64(1), pp. 3–19.
208. Wright, J., Wagner, A. (2008) The Systems Biology Research Tool: evolvable open-source software. *BMC systems biology*, Vol. 2, p. 55.
209. Xia, J., Mandal, R., Sinelnikov, I. V., Broadhurst, D., Wishart, D.S. (2012) MetaboAnalyst 2.0--a comprehensive server for metabolomic data analysis. *Nucleic acids research*, Vol. 40(Web Server issue), p. W127–33.
210. Yeung, M., Thiele, I., Palsson, B.O. (2007) Estimation of the number of extreme pathways for metabolic networks. *BMC bioinformatics*, Vol. 8(1), p. 363.
211. Yousofshahi, M., Ullah, E., Stern, R., Hassoun, S. (2013) MC3: a steady-state model and constraint consistency checker for biochemical networks. *BMC systems biology*, Vol. 7, p. 129.
212. Zamboni, N., Fendt, S.-M., Rühl, M., Sauer, U. (2009) (13)C-based metabolic flux analysis. *Nature protocols*, Vol. 4, pp. 878–892.
213. Zettlmeissl, G., Rudolph, R., Jaenicke, R. (1979) Reconstitution of lactic dehydrogenase. Noncovalent aggregation vs. reactivation. 1. Physical properties and kinetics of aggregation. *Biochemistry*, Vol. 18(25), pp. 5567–5571.
214. Zhao, J., Shimizu, K. (2003) Metabolic flux analysis of *Escherichia coli* K12 grown on <sup>13</sup>C-labeled acetate and glucose using GC-MS and powerful flux calculation method. *Journal of biotechnology*, Vol. 101(2), p. 101–17.
215. Zrenner, R., Stitt, M., Sonnwald, U., Boldt, R. (2006) Pyrimidine and purine biosynthesis and degradation in plants. *Annual review of plant biology*, Vol. 57, pp. 805–836.









