

Latvijas Lauksaimniecības universitāte
Informācijas tehnoloģiju fakultāte
Datoru sistēmu katedra

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

Jurijs Meitalovs

**STEHIOMETRISKO MODELŪ PIELIETOJUMS
METABOLISKĀS INŽENIERIJAS RISINĀJUMU
TELPAS AUTOMATIZĒTĀ ANALĪZĒ**

**THE USE OF STOICHIOMETRIC MODELS FOR
METABOLIC ENGINEERING SOLUTION SPACE
AUTOMATED ANALYSIS**

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

Jurijs Meitalovs

Paraksts/Signature

Jelgava 2015

INFORMĀCIJA

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

Promocijas darba zinātniskais vadītājs: Dr.sc.ing., prof. Egils Stalidzāns.

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

Oficiālie recenzenti:

1. Latvijas Universitātes asoc. profesors, Dr. sc.comp. Juris Viķsna;
2. Rīgas Tehniskās Universitātes profesors, Dr.habil.sc.ing. Jurijs Merkurjevs;
3. Oxford Brookes universitātes galvenais pētnieks, Dr. Mark Poolman.

Promocijas darba aizstāvēšana notiks LLU Informācijas tehnoloģiju nozares promocijas padomes atklātā sēdē 2015. gada ___. _____, Jelgavā, Lielajā ielā 2, Informācijas tehnoloģiju fakultātes ___. auditorijā, plkst. ____.

Ar promocijas darbu var iepazīties LLU Fundamentālajā bibliotēkā, Lielajā ielā 2, Jelgavā un tiešsaistē http://llufb.llu.lv/promoc_darbi.html

Atsauksmes sūtīt Promocijas padomes sekretārei – Lielajā ielā 2, Jelgava, LV-3001; tālrunis 63005621; e-pasts tatjana.tabunova@llu.lv. Atsauksmes vēlam sūtīt skenētā veidā ar parakstu.

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

SATURS

Promocijas darba aprobācija	4
Ievads.....	6
Tēmas aktualitāte	6
Promocijas darba mērķis un uzdevumi	8
Pētījumu metodes	8
Zinātniskais jauninājums un praktiskā vērtība	8
Pētījuma tēzes	9
Promocijas darba struktūra un apjoms	9
1. Metaboliskā inženierija	9
2. Metabolisko tīklu datoranalīze	11
3. Automatizēta bioķimisko tīklu konstruēšanas procedūra	12
4. Programmatūras rīks sAnalyzer	14
5. Automatizēta metaboliskās inženierijas risinājumu telpas datoranalīze...	16
5.1.Risinājuma telpas analīze atkarībā no ierobežojumiem	16
5.2.Baktērijas <i>Z.mobilis</i> metaboliskās inženierijas analīze glicerīna utilizācijai ..	18
5.3.Baktērijas <i>E.coli</i> metaboliskas inženierijas analīze α -Pinene ražošanai	19
5.4.Baktērijas <i>E.coli</i> metaboliskās inženierijas analīze <i>1-butanol</i> ražošanai	20
Secinājumi.....	22
Galvenie darba rezultāti	22
Secinājumi un attīstības perspektīvas	23
Literatūra.....	47

PROMOCIJAS DARBA APROBĀCIJA

Pētījumu rezultāti ir atspoguļoti sekojošās publikācijās:

1. **Meitalovs, J.** (2012a) Process of genetic modification of microorganisms from the point of view of the software engineering. In: *Proceedings of the 5th International Scientific Conference on Applied Information and Communication Technologies*, pp. 348–351 (Indeksēts „AGRIS” un „EBSCO” datubāzēs).
2. **Meitalovs, J.** (2012b) The concept of metabolic engineering software tool. In: *Proceedings of the 5th International Scientific Conference on Applied Information and Communication Technologies*, pp. 352–355 (Indeksēts „AGRIS” un „EBSCO” datubāzēs).
3. **Meitalovs, J.**, Bulipopa, N., Kovalonoka, O. (2012) Modeling and design tools for synthetic biology. In: *Proceedings of the 5th International Scientific Conference on Applied Information and Communication Technologies*, pp. 341–347 (Indeksēts „AGRIS” un „EBSCO” datubāzēs).
4. **Meitalovs, J.** (2012c) Software tool for probabilistic metabolic pathways construction. In: *Proceedings of the IEEE 13th International Symposium on Computational Intelligence and Informatics (CINTI)*, pp. 405–408 (Indeksēts „Scopus”, „Web of knowledge”, un „IEEE explore” daubāzēs).
5. **Meitalovs, J.**, Stalidzans, E. (2013b) Connectivity analysis of metabolites in synthetic metabolic pathways. In: *Proceedings of the 12th International Scientific Conference ‘Engineering for Rural Development’*, pp. 435–440 (Indeksēts „Scopus” un „AGRIS” datubāzēs).
6. **Meitalovs, J.**, Stalidzans, E. (2013a) Analysis of synthetic metabolic pathways solution space. In: *Proceedings of the 2013 International Conference on System Science and Engineering (ICSSE)*, pp. 183–187 (Indeksēts „Scopus”, „Web of knowledge”, un „IEEE explore” daubāzēs).
7. **Meitalovs, J.**, Stalidzans, E. (2015) Impact of thermodynamic constraint to the solution space of metabolic pathway design using *sAnalyzer* tool. *Baltic Journal of Modern Computing*, 3, pp. 164–178 (Indeksēts „ProQuest” un „EBSCO” datubāzēs).

Pētījumos iegūtie rezultāti tika prezentēti sekojošās konferencēs:

1. Rutkis R., **Meitalovs J.**, Odzina I., Kalnenieks U., In silico metabolic engineering of Zymomonas mobilis for glycerol consumption, *SB5.0: the fifth International meeting on synthetic biology*, June 15-17, 2011, Stanford, USA.

2. **Meitalovs J.** Process of genetic modification of microorganisms from the point of view of the software engineering. *Applied Information and Communication Technologies*, April 26-27, 2012, Jelgava, Latvia.
3. **Meitalovs J.** The concept of metabolic engineering software tool. *Applied Information and Communication Technologies*, April 26-27, 2012, Jelgava, Latvia.
4. **Meitalovs J.**, Bulipopa N., Kovalonoka O. Modeling and design tools for synthetic biology. *Applied Information and Communication Technologies*, April 26-27, 2012, Jelgava, Latvia.
5. **Meitalovs J.**, Software tool for probabilistic metabolic pathways construction, *13th IEEE International Symposium on Computational Intelligence and Informatics*, November 20-22, 2012, Budapest, Hungary.
6. **Meitalovs J.**, Stalidzans E., Connectivity analysis of metabolites in synthetic metabolic pathways, *12th International Scientific Conference ‘Engineering for rural development’*, May 23-24, 2013, Jelgava, Latvia.
7. **Meitalovs J.** , Stalidzans E., Analysis of synthetic metabolic pathways solution space, *ICSSE 2013 IEEE International Conference on System Science and Engineering*, July 4-6, 2013, Budapest, Hungary.
8. **Meitalovs J.**, Probabilistic metabolic pathways construction and analysis tool, *SB6.0: The Sixth International Meeting on Synthetic Biology*, July 9-11, 2013, London, UK.
9. **Meitalovs J.**, Stalidzans E., Applications of *sAnalyzer* – tool for metabolic network construction, *2nd Congress of baltic Microbiologists*, October 16-19, 2014, Tartu, Estonia.

Pētījumos iegūtie rezultāti tika prezentēti sekojošos semināros:

Meitalovs J., *sAnalyzer* rīka pielietojumi, *110. Biosistēmu grupas seminārs*, 2014. gada 23. maijā, Jelgava, Latvija.

Dalība ar promocijas darbu saistītajos projektos:

„Latvijas starpaugstskolu zinātniskās grupas izveide sistēmbioloģijā” Nr.2009/0207/1DP/1.1.1.2.0/09/APIA/VIAA/128, (2009. – 2012., pētnieks).

IEVADS

Tēmas aktualitāte

Pēcgenoma ērā zinātnieki arvien vairāk izmanto dabā sastopamo reakciju izmantošanu tām neraksturīgos organismos biotehnoloģisku mērķu sasniegšanai. Piemēram, 1928. gadā *A. Fleming* atklāja penicilīnu – pirmo antibiotiku, un 1941. gadā tika izstrādāta tehnoloģija, lai ražotu antibiotiku ar fermentācijas procesu palīdzību. 1978. gadā insulīna kodējošo gēnu izdevās ievietot *E.coli* baktērijā un ļaut dzīvai sistēmai risināt insulīna sintēzes uzdevumu ("Herbert W. Boyer"). Šāds bioķīmijas pielietojums tiek saukts par biotehnoloģiju, un, iespējams, tā ir viena no visstraujāk augošajām zinātnes nozarēm.

Biotehnoloģija ir industrija, kura darbojas ar produktu iegūšanu cilvēka vajadzībām, izmantojot mikroorganismus. Biotehnoloģiskā procesa produkti tiek plaši izmantoti vairākās industrijās, piemēram, pārtikas, medicīnas, ķīmijas, biodegvielu ražošanā, lauksaimniecībā (Adrio, Demain, 2010; Atsumi et al., 2008; Carothers et al., 2009; Serrano, 2007; Trinh, Srienc, 2009; Yang et al., 2013; Yu et al., 2010; "Contained use of genetically ...").

Mazāk nekā 200 gados zinātnieki pārgāja no vienas reakcijas starp minerāliem uz tūkstošiem savstarpēji savienotām un vadītām reakcijām, kurās ļauj baktērijai dzīvot un ražot cilvēkiem nepieciešamās vielas. Šādu tūkstošiem reakciju veidotu sistēmu matemātiskais apraksts un informācijas tehnoloģijas to analīzei ir klūvušas par priekšnosacījumu pieejamās informācijas daudzveidības sistematizēšanai un izmantošanai (Kuprianov et al., 2013; Lehrach et al., 2011; Zazzu et al., 2013). Mūsdienās biotehnoloģija ir multidisciplināra nozare - tā ietver vairākas zinātnes, tādās kā bioloģija, ķīmija, metaboliskā inženierija, informācijas tehnoloģijas, procesu vadība un optimizācija.

Par metabolismu jeb vielmaiņas procesu pieņemts saukt dažādu savstarpēji saistītu bioķīmisko reakciju kopumu šūnās vai organismā, kas nepieciešams, lai nodrošinātu to dzīvotspēju (Vīgants, 2008).

Viens no galvenajiem metaboliskās inženierijas mērķiem ir palielināt organismu produktivitāti ķīmisko vielu ražošanā, izmantojot ģenētiskas modifikācijas. Inženierijas termina piesaukšana norāda, ka arvien vairāk darbībās ar bioloģiskajām sistēmām tiek pielietota inženierijai raksturīga pieeja: sistēma tiek būvēta vai modificēta tikai pēc pietiekoši detalizētas sistēmas izpratnes un tās datormodeļu analīzes (Tomlin and Axelrod, 2005; Tyo et al., 2007; Vital-Lopez et al., 2006; Mendes and Kell, 1998).

Metaboliskais ceļš attēlo funkcionālo vienību tīklā. Bioķīmija definē metabolisko ceļu kā ķīmisko reakciju secību, ar ko sākotnējais metabolīts tiek pakāpeniski pārveidots par kādu citu molekulu vai molekulām (Berg et al., 2002). Ir iespējams veikt ģenētiskas manipulācijas organismā, lai mainītu vielmaiņas tīklu un panāktu nepieciešamās vielas ražošanu.

Viens no svarīgiem organismu izmaiņu soļiem ir mērķa organisma projektēšana ar datormodelēšanas palīdzību. To var realizēt ar modelēšanu un simulācijām: projektēt modificēto organismu atbilstoši mērķa funkcijai un pārbaudīt tā sagaidāmo potenciālu (Macdonald et al., 2011). Bioloģisko sistēmu modeļu un datoreksperimentu izmantošana bioloģisko eksperimentu vietā ir lētāka un neprasī iekārtu un ķimikāliju resursus (Gendrault et al., 2011, Wiechert, 2002). Šūnu sistēmu modelēšana nozīmē molekulāro komponenšu apvienošanu matemātiski un bioloģiski pareizā veidā, lai vesela sistēma pastāvētu *in silico* kā viens matemātiskais objekts. Šādu rekonstrukciju veikšanai ir nepieciešams izmantot specifisku bioloģisko informāciju.

Metabolisko ceļu datormodelēšanai var izmantot bioķīmisko tīklu stehiometriskās matricas analīzi. Šādas analīzes galvenā problēma ir skaitļošanas apjoms, ko izraisa iespējamo ceļu skaita kombinatorisks sprādziens genoma līmeņa metaboliskajos tīklos (Blum, 2009). Parasti jauno ieviešamo metabolisko ceļu ievietošana šās jās organismu modelī un analīze notiek manuāli, atbilstoši biotehnoloģijas speciālistu idejai. Lielas iespējamās modifikācijas risinājuma telpas pārmeklēšanai ir nepieciešams daudz manuāla darba. Līdz ar to, izvirzās arī prasība pēc liela apjoma analīzes automatizācijas, kas ir vēl viens informācijas theoloģiju potenciālais pielietojums.

Esošo rīku klāsts ļauj risināt dažādas problēmas, bet jauno, dabā neeksistējošo metabolisko ceļu meklēšanas problēmas pašlaik nav risināmas ar specializētās programmatūras palīdzību. Šobrīd neeksistē datorizēti rīki, kuri ļauj pārbaudīt plašu iespējamo risinājumu telpu un sniegt ieteikumus par iespējamas modifikācijas rezultātu. Lai arī šādas problēmas risināšanai ir bijuši mēģinājumi izstrādāt visu iespējamo transformācijas ceļu meklēšanas algoritmus (Blum, Kohlbacher, 2008; Ranganathan, Maranas, 2010; Rodrigo et al. 2008; Wang, Hatzimanikatis, 2006; Arita, 2000; Croes et al., 2006), problēma vēl pastāv, jo pieejamie algoritmi nav realizēti datorizētos rīkos un netiek plaši izmantoti, jo daudzi minētie algoritmi nav realizēti datorizētos rīkos, vai realizētu rīku pieejamās versijas netiek uzturētas funkcionālā stāvoklī, līdz ar to tās nav darboties spējīgas. Kā arī, to izmantoti dati ir novecojušie, jo netiek izmantota tiešsaistes datu ieguve.

Daudzi programmatūras produkti ir izstrādāti vienas noteiktās problēmas risināšanai. Turklat tajos nav izmantotas vairākas analīzes metodes, kuras ļauj atbildēt uz jautājumiem, vai izvēlētais ceļš spēj nodrošināt gan nepieciešamās vielas ražošanu, gan organisma augšanu, netiek pielietoti vairāki ierobežojumi, piemēram, reakciju termodinamikas dati, kā arī netiek ņemtas vērā šās jās organismu metabolisma īpašības.

Minētā situācija norāda uz datortehnoloģiju pielietojuma aktualitāti metaboliskajā inženierijā un nosaka promocijas darba mērķi un uzdevumus.

Promocijas darba mērķis un uzdevumi

Promocijas darba mērķis ir automatizēt bioķīmisko ceļu konstruēšanu un analīzi šasijas organismos metaboliskās inženierijas uzdevumu risināšanai.

Pētījuma mērķa sasniegšanai tika izvirzīti šādi uzdevumi:

1. pastāvošo metaboliskās inženierijas un datorizētās bioķīmisko ceļu konstruēšanas metožu analīze;
2. bioķīmisko tīklu automātiskas konstruēšanas un analīzes procesā iekļaujamo kritēriju un parametru analīze;
3. bioķīmisko tīklu automatizētas konstruēšanas procedūras izstrāde ar iespēju iekļaut esošo šasijas organisma metaboliskā tīkla modeli;
4. automātiska bioķīmisko ceļu konstruēšanas un analīzes rīka izstrāde;
5. iespējamās risinājumu telpas analīze atkarībā no izvirzītajiem ierobežojumiem;
6. metaboliskās inženierijas uzdevumu risināšanas procedūras demonstrēšana.

Pētījumu metodes

1. Automatizētas bioķīmisko ceļu konstruēšanas un analīzes rīks, izstrādāts *Matlab* vidē.
2. Metabolītu transformācijas bioķīmiska tīkla konstruēšanai ir izmantoti dati no tiešsaistes brīvpieejas datu bāzes KEGG.
3. Bioķīmisko ceļu meklēšanai konstruētā metabolītu transformācijas bioķīmiskajā tīklā ir izmantots modificēts visu ceļu meklēšanas dzīlumā algoritms (DFS). Tika izstrādāta metode bioķīmisko ceļu automātiskai vizualizācijai, kura izmanto *Matlab* rīkā iebūvētas iespējas grafu vizualizācijai.
4. Bioķīmisko modeļu analīzē izmantots plūsmas bilances analīzes rīks *COBRA*, kas balstīts uz lineāras algebras metodēm.
5. Bioķīmisko procesu modelēšanā izmantotas stehiometriskās modelēšanas metodes, pielietots SBML standarts modeļa definēšanai.
6. Konstruēto bioķīmisko ceļu analīzei tiek izmantots izstrādāts daudzkriteriāls analīzes algoritms. Ceļu ranžēšanai ir izmantoti metaboliskos ceļus raksturojoši parametri ar maināmiem svara koeficientiem.

Zinātniskais jauninājums un praktiskā vērtība

1. Izstrādāta bioķīmisko tīklu automātiskās konstruēšanas procedūra.
2. Izstrādāts automātiski realizējams bioķīmisko ceļu analīzes algoritms, kas ņem vērā termodinamiku, šasijas organisma metabolītus un plūsmu bilances analīzes rezultātus.
3. Izstrādāta konstruēto bioķīmisko ceļu daudzparametru svērtās ranžēšanas metode.

- Izstrādāts rīks, kas automātiski pārlasa iespējamo ģenētiskas modifikācijas risinājumu telpu un konstruē bioķīmiskos tīklus atbilstoši metaboliskās inženierijas uzdevumam.
- Konstatēts Ričarda līknes tipa telpas pieaugums atkarībā no risinājuma telpā pieļaujamo reakciju skaita.

Pētījuma tēzes

- Palielinot pieļaujamo reakciju skaitu ceļā, iespējamo modifikāciju risinājumu telpa sākotnēji strauji pieaug, bet turpinājumā sasniedz ierobežota reaģētu skaita noteiktu piesātinājumu.
- Ir iespējams automātiski konstruēt un ranžēt bioķīmiskos ceļus, nemot vērā to termodinamiskos ierobežojumus, šasijas organisma datormodelus un plūsmu bilances analīzes rezultātus.
- Risinājumu telpas pilna automātiska skenēšana piedāvā literatūrā publicētajiem risinājumiem konkurētspējīgu risinājumu kopu.

Promocijas darba struktūra un apjoms

Promocijas darbs ir uzrakstīts latviešu valodā, satur anotāciju, ievadu, 5 nodaļas, secinājumus, literatūras sarakstu un 6 pielikumus, tajā skaitā 5 tabulas, 58 attēlus, kopā 128 lappuses. Darbā izmantoti 133 literatūras avoti.

1. METABOLISKĀ INŽENIERIJA

Šūnu metabolisms sastāv no sarežģīta ķīmisko reakciju tīkla, kurš transformē molekulās un kurš darbojas, lai pārveidotu barojošās vielas enerģijā un biomasā. Šādā tīklā metaboliskie ceļi attēlo funkcionālās vienības, kuras ir atbildīgas par specifiskiem metaboliskiem procesiem, piemēram, oglekļa avotu degradāciju vai aminoskābju sintēzi. Tas ir, bioķīmija definē metabolisko ceļu kā ķīmisko reakciju sekvenci, ar ko sākotnējais metabolīts ir pakāpeniski pārveidots par kādu citu molekulu vai molekulām (Berg et al., 2002). Piemēram, glikolīzes metaboliskais ceļš pārveido glikozi par piruvātu. Termodinamikas informācija var tikt izmantota, lai aprēķinātu ķīmisko reakciju līdzsvaru un prognozētu, vai konkrētā reakcija var noritēt pie dotajos apstākļos.

Šobrīd zināmie metaboliskie ceļi tika atklāti, pateicoties cītīgam darbam ar specifiskiem organismu modeļiem. Tomēr realitātē metaboliskie tīkli ir vairāk mainīgi un vairāk savstarpēji saistīti nekā (pārsvara lineārie) modeļu metaboliskie ceļi (Cordwell, 1999; Vīgants, 2008). Izmantojot gēnu-enzīmu (gēnu-proteīnu), enzīmu-reakciju, kā arī reakciju-metabolītu attiecības, var tikt rekonstruēti organismu specifiskie šūnu metaboliskie tīkli, balstīti uz datiem no bioloģisko objektu datu bāzēm, un var tikt veikta šādu tīklu *in silico* analīze.

Analizējot šādus vielmaiņas tīklus, var noteikt, kādi produkti tiek ražoti mikroorganismu augšanas procesā, kā arī ir iespējams modelēt vielmaiņas ceļu

izmaiņas, lai panāktu biotehnoloģiskajā procesā nepieciešamo vielu ražošanas palielināšanu, vai sākt ražot jaunu vielu, kura nav sastopama savvaļas mikroorganismu sugā.

Alteraņtīvo metabolisko ceļu eksistence ir organismu adaptācijas rezultāts apkārtējai videi, tāpēc, zināšanas par visiem iespējamajiem ceļiem, kuri transformē avota metabolītu par mērķa metabolītu var palīdzēt labāk saprast metabolismu. Tomēr, metabolisko ceļu eksperimentāla noteikšana ir laika un resursu ietilpīgs process (Blum, 2009).

Metaboliskā inženierija ir ģenētisko un regulācijas procesu optimizācija šūnās, lai uzlaboti noteikta produkta ražošanu. Dabiski metabolisko reakciju tūkli var tikt uzlaboti ar modelēšanas un optimizācijas metodēm. Mikroorganismu uzlabošana, izmantojot metaboliskas inženierijas metodes, paaugstina organisma produktivitāti, samazina izmaksas un apkārtējo piesārņojumu, un vispārēji uzlabo rezultātus vairākās industrijās un nozarēs. Organismu metabolismus var transformēt izejvielas par plašu produktu klāstu, vairāki no šiem materiāliem var tikt izmantoti industriāliem un komerciāliem mērķiem. Visplašāk sintētisko bioloģiju un sistēmbiotehnoloģiju izmanto medicīnā un apkārtēja vidē, t.i., medikamentu sintēzē, terapijā, biosensoros, biodegvielas ražošanā, farmācijā, biomateriālu un ķīmikāliju ražošanā, dzīvības izpratnes uzlabošanā (Gendrault et al., 2011; Khalil, Collins, 2010; Tepper, Shlomi, 2010).

Biotehnoloģija ir zinātnē, kura ļauj izmantot esošo reakciju kopumu mikroorganismos, lai ražotu cilvēkiem nepieciešamas vielas, izmantojot metabolisko inženieriju, un ar sistēmbioloģijas un sintētiskas bioloģijas palīdzību modificētus mikroorganismus. Sistēmbioloģija kā zinātnē nav jauna, bet tā ir strauji attīstījusies, pateicoties jaunu tehnoloģiju attīstībai. Sistēmbioloģija tiek definēta kā globāla mēroga bioloģisku procesu pētīšana šūnā, organismā vai arī pat kopienas līmenī, to molekulāro komponentu un funkciju mijiedarbību ("What is systems biology?"). Terms „sintētiskā bioloģija” parādījās nesen un ir saistīts ar pilnīgi atšķirīgiem pētījumu virzieniem. Vairāki no tiem ir fokusēti uz inženieriju, kurā dabiskās molekulas tiek apvienotas definētos kompleksos un ievietotas šūnās, piešķirot šādi uzbūvētām šūnām jaunas īpašības un funkcionalitāti, ģenētiski tos modificējot (Khalil, Collins, 2010; Mannelli et al., 2003). Sintētiskā bioloģija ir definēta kā inženierijas principu pielietojums bioloģisko sistēmu izstrādē ("What's in a name?").

Metaboliskajā inženierijā ir nepieciešamas vairākas iterācijas, lai pārliecinātos par atsevišķu soļu veiksmīgu integrāciju. Viens no svarīgiem organismu izmaiņu procedūras soļiem ir procesa projektēšana. Viena no metodēm, kas ļauj to izpildīt, ir datormodelēšana un simulācijas (*in silico* metodes) – metodes, ar kuru palīdzību ir iespējams projektēt modificēto organismu atbilstoši mērķa funkcijai un pārbaudīt tas potenciālu (Macdonald et al., 2011). Šobrīd ir vairāki programmatūras rīki, ar kuru palīdzību ir iespējams atvieglot un paātrināt projektēšanu. Šādus rīkus var sadalīt vairākās kategorijās pēc to pielietojuma.

2. METABOLISKO TĪKLU DATORANALĪZE

Šūnas satur daudzus metaboliskos ceļus, lai nodrošinātu to dabiskos izdzīvošanas procesus. Lai saprastu, kā šūnā tiek sadalītas tās funkcijas, var tikt izmantota metabolisko tīklu analīze. Iespējamo metabolisko ceļu liela skaita un sarežģītības dēļ nav vienmēr intuitīvi saprotams, kuri ceļi tiek izmantoti, lai sasnietu šūnas mērķa funkcijas. Metaboliskos tīklus var attēlot grafa veidā un pielietot grafa teorijas metodes un ceļu meklēšanas konceptus, lai saskaitītu lineārās biotransformācijas ceļus starp doto avota metabolītu un mērķa metabolītu grafā, kurš attēlo metabolisko tīklu. Ceļu meklēšanas algoritmi var tikt izmantoti, lai aprēķinātu vienu vai vairākus ceļus, kuri savieno divas atšķirīgas virsotnes grafā. Sie algoritmi atšķiras pēc to potenciālā pielietojuma un to sarežģītības.

Metabolisko tīklu analīzi var veikt, izmantojot organismu stehiometrisko modelēšanu. Tā ir metode, kura ir balstīta uz ierobežojumiem. Metabolisko tīklu stehiometrija apraksta vielmaiņas procesā notiekošo ķīmisko reakciju reāgentu un produktu attiecības. Izmantojot šādu modelēšanas pieeju, var noskaidrot metabolītu plūsmas organismā modeļā un secināt par organismā notiekošajiem vielmaiņas procesiem. Viens no plašāk izmantojamajiem rīkiem, ar kuru palīdzību var analizēt visus iespējamos metaboliskos ceļus stehiometriskos modeļos, ir plūsmu bilances analīze (angl., *FBA, Flux Balance Analysis*). Tīklu vizuālā attēlošana ir svarīga to saprašanai. Lai gan vairāk informācijas var iegūt no matemātiskas attēlošanas, ir iespējams attēlot metabolisko tīklu matemātiski, izveidojot stehiometrisko (savienojumu) matricu, kura attēlo tīklu. Reakciju un metabolītu identifikatori tiek izmantoti, kā kolonnu un rindu nosaukumi. Biotehnoloģiskajā ražošanā, meklējot risinājumu, kā var saražot noteikto vielu, izmantojot mikroorganismus, rodas problēma - kādas bioķīmiskas reakcijas izmantotas, vienu vielu transformējot citā. Papildus ir nepieciešams noskaidrot, kādā veidā ir iespējams ražot noteiktu vielu, ieviešot pēc iespējas mazāk modifikāciju.

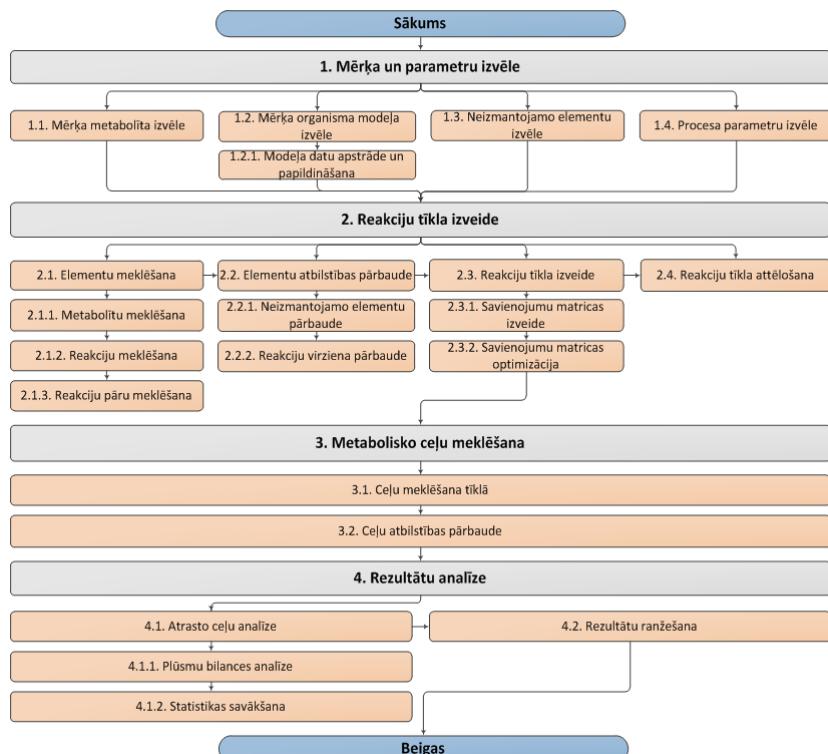
Automātiski izveidoti metaboliskie ceļi var tikt izmantoti, lai pārbaudītu noteiktas teorijas par metabolītu konvertāciju, iespējamo produktu biosintēzi vai substrātu biodegradāciju. Pārsvarā pieejamās metodes ļauj izveidot metaboliskus ceļus, bet lietotājam manuāli ir jāpārbauda izveidotā ceļa precizitāte, un arī, ja ir nepieciešams, jāimplementē to esošajā modeļā, lai to testētu. Šāds risinājums der neliela izmēra modeļiem, bet lielākos metaboliskos tīklus to manuāli izdarīt ir darbietilpīgi. Cita biotehnoloģiskas ražošanas problēma varētu būt, kad ir zināms substrāts un produkts, bet nav zināms šāsījas organisms – organisms, kurš tiks izmantots ģenētisko modifikāciju veikšanai, lai pārveidotu substrātu par produktu.

Šajā gadījumā nav neviena rīka, kurš palīdzētu izvēlēties piemērotus risinājumus. Pašreiz organisma izvēle un metabolisko ceļu meklēšana pārsvarā notiek eksperimentāli vai/un balstoties uz zinātnieku zināšanām un pieredzi. Bet bieži vien tas nesniedz optimālo rezultātu, nav pārbaudīti visi iespējamie varianti. Līdz ar to var tikai cerēt, ka šāda izvēlētā modifikācija izdosies un būs konkurenčspējīga ar ciemtiem biotehnoloģiskajiem risinājumiem.

Šajā darbā tiek piedāvāta metodoloģija, to realizējošs rīks un praktiski piemēri, kas parāda automatizētu metabolisko ceļu izveides algoritmu un to implementāciju. Piedāvātais risinājums apvieno esošās metodes un uzlabo tās izmantošanu industriālajos pielietojumos, nodrošina uzlaboto metabolisko ceļu validācijas shēmu.

3. AUTOMATIZĒTA BIOĶIMISKO TĪKLU KONSTRUĒŠANAS PROCEDŪRA

Sistēmbioloģijā un sintētiskajā bioloģijā bieži nāas noteiktā secībā cikliski pielietot dažādus datoranalīzes algoritmus. Šo algoritmu pielietošanas secību bieži sauc par procedūru (Balsa-Canto et al., 2010; Mozga, 2012; Rubina, 2013; Vital-Lopez et al., 2006). Lai novērstu iepriekš minētos trūkumus metabolisko tīklu meklēšanā, darba ietvaros tika izstrādāta automatizēta datormodelēšanā un optimizācijā balstīta bioķimisko tīklu konstruēšanas un analīzes procedūra. Tā ir sadalīta četros galvenajos soļos (sk. 1. att.). Katra soļa rezultāts tiek izmantots nākamajā solī.



1. att. Bioķimisko tīklu automatizētas konstruēšanas procedūras soļi

Pirmais solis tiek sadalīts četros apakšsolos, kas ir savstarpēji saistīti. Pirmais solis sākas ar kritēriju un modeļa izvēli, kas ir paredzēts procesa vispārējai iizzināšanai – kāds ir biotehnoloģiskās ražošanas uzdevums, kāds ir bioķīmiskais process, kādi kritēriji ir izvirzīti. Tieki definēts mērķa metabolīts un mērķa šasijas organisms, kurš tieks izmantots biotehnoloģiskajā procesā. Šim šasijas organismam tiek izvēlēts metaboliskais modelis, atbilstoši kuram tiek izveidots saraksts ar reakcijām un metabolītiem, kuri tiek izslēgti no tālākās analīzes, kā arī definēti papildu parametri.

Mērķa metabolīts, elementi, kuri nepiedalās ceļu konstruēšanā, un papildu parametri tiek izmantoti otrajā solī reakciju tīkla izveidē. Papildu parametri tiek izmantoti arī mērķa organisma modeļa izvēles procesā, kas savukārt tiek izmantots trešajā solī – metabolisko ceļu meklēšanā.

Nākamajā solī notiek bioķīmisko reakciju meklēšana un savienošana reakciju tīklā. Šajā solī tiek pārbaudīti iespējamie metabolītu transformācijas ceļi. Darba autors piedāvā izmantot visu iespējamo metabolītu savienojumu pārlasi, t.i., pārbaudot visas iespējamās ķīmiskās reakcijas, kuras var pārveidot mērķa metabolītu. Katram metabolītam KEGG datu bāzē tiek atrastas visas reakcijas, kurās tas piedalās. Katrai reakcijai tiek sameklēti reakciju pāri. Katram reakciju pārim tiek sameklēti metabolīti, kuri piedalās noteiktā reakcijā. No atrastajiem metabolītiem un reakcijām tiek izveidota savienojumu matrica, kura attēlo elementu saistības. Tieki izveidots nākamā meklēšanas līmena metabolītu saraksts. Savukārt, katram metabolītu līmenim netiek ietverti tie metabolīti, kuri tika definēti kā apkārtējie metabolīti un nepiedalās atomu pārveides kēdē starp substrātu un produktu. Beigās tiek atgriezta tīkla savienojumu matrica, kura attēlo izveidotu reakcijas tīklu.

Trešajā solī notiek metabolisko ceļu meklēšana izveidotā reakciju tīkla risinājuma telpā. Šajā solī ir nepieciešams sameklēt visus iespējamos metabolītu transformācijas ceļus, sākot ar mērķa metabolītu līdz kādam metabolītam mērķa organisma modeļi, tas ir, jāatrod visi ceļi, ar kuriem mērķa metabolīts var tikt savienots ar modeļi.

Tādēļ tiek piedāvāts izmantot automatizētu pieeju visu iespējamo metabolītu transformācijas ceļu meklēšanai. Tā kā iepriekšējā solī tika izveidots metabolītus savienojošs reakciju tīkls, kurš tiek attēlots grafa veidā, tiek piedāvāts izmantot algoritmu visu ceļu meklēšanai grafā. Eksistē vairāki algoritmi, kuri piedāvā šādu funkcionalitāti, līdz ar to ir nepieciešams izvēlēties piemērotāko un realizēt to. Šim algoritmam jāatgriež saraksts ar visiem metaboliskie ceļiem, kuri savieno divus metabolītus.

Pēdējais solis paredz iepriekšējā solī iegūto metabolisko ceļu analīzi un ranžēšanu. Šis solis ir sadalīts četros posmos, pirmais no kuriem ir iepriekšējā solī atrasto ceļu analīze. Analīzei ir nepieciešams atrast visus papildus metabolītus, kuri piedalās katrā no ceļa reakcijām, balstoties uz pirmajā solī izvēlētajiem parametriem. Ir nepieciešams pievienot šo ceļu ar visiem metabolītiem modeļi un pārbaudīt, vai šis ceļš spēj nodrošināt atomu plūsmu caur to. Šim mērķim var izmantot plūsmu

bilances analīzes (FBA) algoritmu, kurš nosaka, cik liela atomu plūsma ir iespējama caur noteiktu metabolisko ceļu. Ja FBA analīzes rezultātā ir redzams, ka šis ceļš nespēj nodrošināt pietiekamu plūsmu, kā arī, ja biomasa augšanas plūsma ir zema, šo ceļu izslēdz no turpmākās analīzes.

Darba autors piedāvā izmantot sekojošu metabolisko ceļu novērtēšanas formu:

- ņemot vērā ceļa kopējo metabolītu skaitu;
- ņemot vērā modeļa un ceļa kopējo eksistējošo metabolītu skaitu;
- ņemot vērā modelī pievienoto metabolītu skaitu, kuri eksistē ceļā;
- ņemot vērā ceļa reakciju skaitu;
- reakciju skaitu ar zināmo termodinamiku, ja analīze tiek izmantoti termodinamikas dati;
- modelī un ceļā kopējo reakciju skaitu;
- nelīdzsvaroto reakciju skaitu;
- metaboliskā ceļa plūsmas bilances analīzes rezultāti;
- modeļa biomasa plūsmas bilances analīzes rezultātus ar pievienotu ceļu.

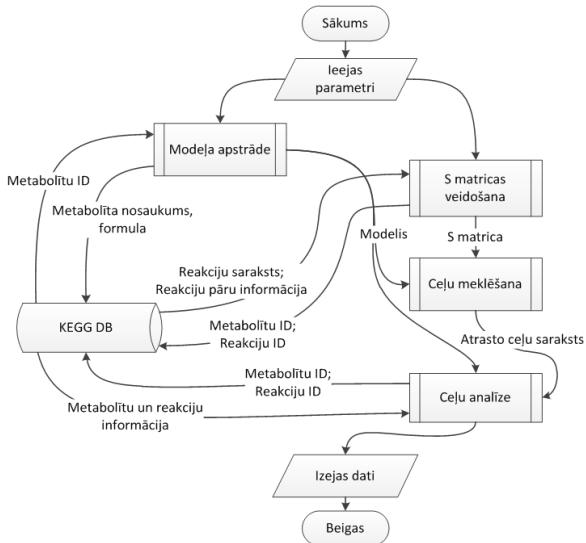
Tika izstrādāta konstruēto bioķīmisko tīklu daudzparametru svērtās ranžēšanas metode, kurā ņem vērā 20 atsevišķos parametrus un svarīguma koeficientus. Šādu metodi ir iespējams pielāgot, izvēloties tikai tos parametrus, kuriem ir lielāka priekšrocība biotehnoloģiskā uzdevuma risināšanā.

4. PROGRAMMATŪRAS RĪKS SANALYZER

Balstoties uz iepriekš aprakstītās procedūras soļiem, tika izstrādāts rīks *sAnalyzer*, kurš secīgi implementē visus procedūras soļus un pierāda, ka ar šādas procedūras palīdzību ir iespējams veikt jauno, dabā neeksistējošo metabolisko ceļu izstrādi. *sAnalyzer* rīks ļauj pārbaudīt visas iespējamās kīmisko reakciju kombinācijas, kuras savieno nepieciešamo metabolītu ar šāsijas modeli.

Šis rīks ir izstrādāts *Matlab* vidē. Programmēšana *Matlab* vidē notiek skriptu izpildāmos failos, kur katrs fails var saturēt vienu vai vairākas funkcijas. *Matlab* vide tika izvēlēta, jo tā ļauj izmantot *Matlab* iebūvēto funkcionalitāti, kā arī tam ir pieejami vairāki paplašinājumi.

Viens no tiem, „*COBRA toolbox*” (Becker et al., 2007), tika izmantots, lai veiktu metabolisko ceļu analīzi ar plūsmu bilances analīzes algoritma palīdzību. *sAnalyzer* ietver četrus galvenos funkcionālos moduļus, kuri ļauj manipulēt ar modeli un iegūtajiem datiem (sk. 2. att.).



2. att. *sAnalyzer* programmas vienkāršots darbības algoritms

Šim rīkam piemīt vairākas inovatīvas īpašības, kuras to atšķir no jau esošajiem rīkiem. Kā svarīgākās īpašības var pieminēt:

1. *sAnalyzer* ļauj savienot mērķa metabolītu ar doto šāsijas modeli;
2. netiek definēta ieejas/izejas vieta, programma meklē tās automātiski;
3. *sAnalyzer* meklē reakcijas no dažādiem organismiem tiešsaistes datu bāzē;
4. tiek pārbaudīta plaša modifikāciju risinājuma telpa, ņemot vērā šāsijas organizma īpašības.

Pēc ieejas datu ievades tiek izsauktas divas apakšprogrammas – izejvielas transformācijas ceļu savienojuma matricas izveides apakšprogramma un šāsijas organizma modeļa apstrādes apakšprogramma. Modeļa apstrādes apakšprogramma pārbauda vai ieejas modelī eksistē metabolītu KEGG datu bāzes identifikatori. Ja šādas informācijas modelī nav, tā tiek automātiski sameklēta KEGG datu bāzē un pievienota modelim. Dati tiek nolasīti no KEGG datu bāzes, izmantojot API pieprasījumus. Šī informācija tiek izmantota, lai atrastu metabolītu transformācijas ceļus starp ieejas metabolītu un modeļi, izmantojot izveidotu otrā apakšprogrammā savienojumu matricu.

Savienojumu matricas veidošanas apakšprogramma izveido sākotnēja metabolīta transformācijas tīklu, kurš var tikt attēlots kā grafs, sākot ar ieejas metabolītu un ar noteikto reakciju skaitu katrā ceļā. Šī apakšprogramma atgriež ieejas metabolīta transformācijas stehiometrisko matricu, kura apraksta elementu savienojumu attiecības.

Šī matrica tiek izmantota visu metabolisko ceļu meklēšanai starp diviem metabolītiem, izmantojot modificēto meklēšanas dziļumā algoritmu (angl., *Depth-first search, DFS*) ("Graph Traversal"). Modificēts DFS algoritms ir implementēts ceļu meklēšanas apakšprogrammā. Iegūtie rezultāti tiek nodoti ceļu analīzes apakšprogrammai, kura veic katru metaboliska ceļa analīzi.

Katrs analizētais metaboliskais ceļš tiek pievienots modelim, izmantojot *COBRA* rīka funkcionalitāti, veicot tā FBA analīzi, pārbaudot iespējamo plūsmu caur pievienoto ceļu, un FBA analīzi caur pievienoto ceļu un biomasa augšanas reakciju. Katram analīzes parametram tiek piešķirts noteikts svara koeficients, pēc kura notiek ceļu atmešana un ranžēšana.

Papildu vienkāršam biotehnoloģiskam uzdevumam – substrāta utilizācija vai produkta ražošanai, *sAnalyzer* piedāvā risināt šādu apvienotu uzdevumu – substrāta utilizācija un produkta ražošana, izmantojot šāsijas organisma modeli.

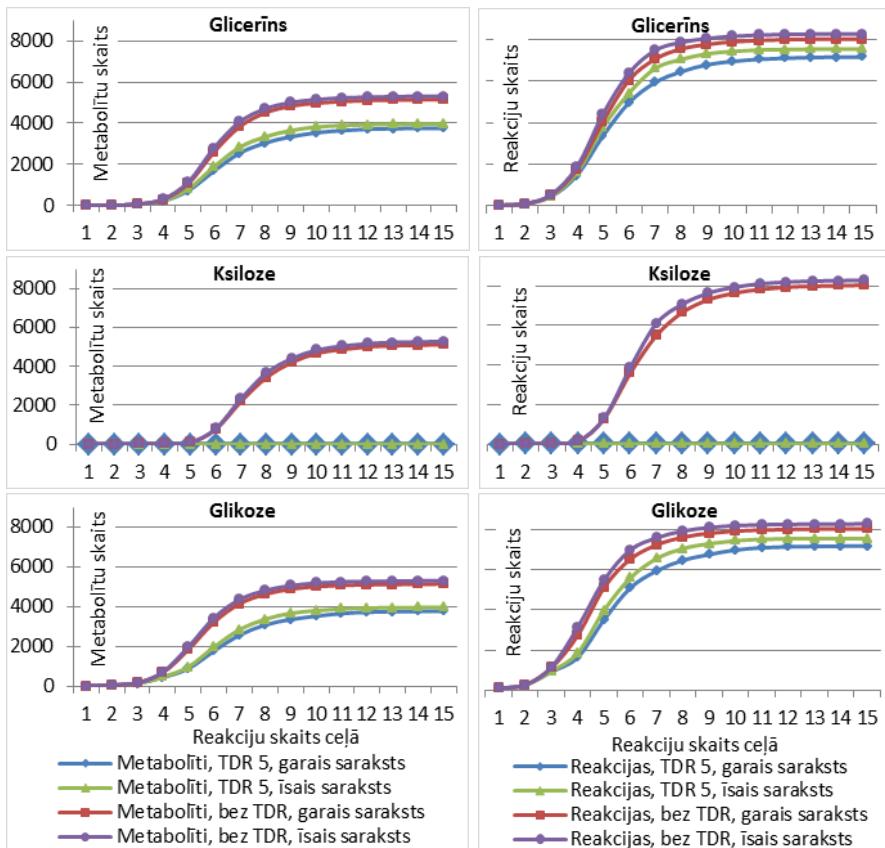
5. AUTOMATIZĒTA METABOLISKĀS INŽENIERIJAS RISINĀJUMU TELPAS DATORANALĪZE

Nodaļa veltīta izstrādātās metaboliskās inženierijas risinājumu telpas datoranalīzes pieejas un rīka darbības analīzei.

5.1. Risinājuma telpas analīze atkarībā no ierobežojumiem

Tika veikti 12 eksperimenti ar 3 metabolītiem, lai noteiktu citu metabolītu sasniedzamību ar KEGG datu bāzē pieejamo reakciju palīdzību. Tika izveidoti katrā metabolīta transformācijas tīkli. Eksperimentam tika izvēlēti metabolīti ar atšķirīgu reakciju skaitu, kurās tie ir tieši iesaistīti.

Šādu reakciju skaits norāda uz metabolītu sākotnējo savienojamību tīklā. Izvēlētas trīs biotehnoloģijā plaši pielietotas izejvielas, kā sākotnējie metabolīti: glikoze (tieši iesaistīta 111 reakcijās), ksiloze (tieši iesaistīta 8 reakcijās) un glicerīns (tieši iesaistīts 25 reakcijās) (sk. 3. att.).



3. att. Metabolītu transformācijas tīkłā izmantoto metabolītu (pa kreisi) un reakciju (pa labi) skaits atkarībā no reakciju skaita ceļā

Eksperimenti notika ar reakciju termodinamikas robežvērtību (TDR) 5kJ/mol un neizmantojot TDR, kā arī izmantojot divus sarakstus ar elementiem, kuri nepiedalās ceļu konstruēšanā, t.i., definēti kā apkārtējie metabolīti. Garais saraksts tika ņemts no publicētajiem datiem un satur 159 metabolītu nosaukumus (Ranganathan, Maranas, 2010), kuri nevar piedalīties ceļu konstruēšanā. Īsais saraksts ir autora veidots ar 18 plaši izplatītiem apkārtējiem metabolītiem.

Līkņu raksturs atbilst Ričarda līknei (angl., *Richards' curve*) (Richards, 1959) – ar zemām argumenta vērtībām pieaugums ir ar kavējumu, bet ar lielām argumenta vērtībām tas kļūst asymptotiski piesātināts.

Neieviešot termodinamikas parametrus, sasniedzamo metabolītu skaits atkarībā no reakciju skaita ceļā norāda, ka ar 15 reakciju garu ceļu no katra pētītā metabolīta ir

iespējams sasniegt ap 90% citu KEGG datu bāzē minēto reakcijās iesaistīto metabolītu.

Glicerīna un glikozes iesaistīto elementu skaits ir līdzīgs visā reakciju skaitā diapazonā. Atšķirības ir nebūtiskas un ir skaidrojamas ar sākotnējo reakciju skaitu, kurās šīs vielas ir tieši iesaistītas. Tajā pašā laikā ir arī novērojams sākotnējās savienojamības efekts: ksilozei ir vērojama savienojamības līmena atpalīšana no glicerīna un glikozes par aptuveni divām reakcijām ceļā. Šī relatīvi stiprākā izolācija no citiem metabolītiem netieši norāda uz iemesliem, kas rada ksilozes kā substrāta izmantošanas problēmas koksnes biomassas pārstrādes procesā (Jeffries, 2006). Bet garākā reakciju tīklā kopējais iesaistīto elementu skaits sasniedz tādu pašu līmeni kā glicerīnam un glikozei. Tas ir, ar visiem 3 metabolītiem ir sasniedzama tā pati metabolītu kopa.

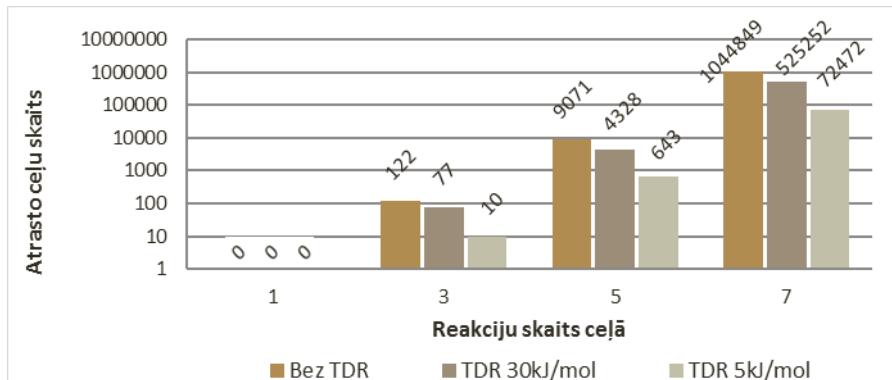
Glicerīnam un glikozei termodinamikas robežvērtībai (TDR) 5kJ/mol ir līdzīga ietekme – iesaistīto metabolītu skaits ir mazāks par aptuveni 1400, un reakciju skaits par 700 reakcijām ar maksimālo reakciju skaitu tīklā 15 reakcijas. Būtiska TDR ietekme ir redzama, sākot ar 5 reakcijām ceļā. Izmantojot TDR ksilozei, jauno iesaistīto elementu skaita augšana apstājas pie reakciju skaita 2 reakcijas, iesaistot 8 reakcijas pirmajā līmenī un 14 reakcijas otrajā līmenī. Tas nozīmē, ka neviens no 14 reakcijām termodinamiski nevar notikt ksilozes pārstrādes virzienā.

5.2. Baktērijas *Z. mobilis* metaboliskās inženierijas analīze glicerīna utilizācijai

Šajā darbā tika izmantots *Z.mobilis* centrālā metabolisma modelis (Pentjuss et al., 2013), kurš satur 96 reakcijas un 81 metabolītus, modeļa pamatā ir *Entner-Dudoroff* ceļš etanola ražošanai.

Tika veikti 12 datoreksperimenti, lai novērtētu glicerīna kā izejvielas izmantošanas iespējas ar atšķirīgu reakciju Gibbs enerģijas robežvērtībām un maksimālo reakciju skaitu ceļā 1, 3, 5 un 7 reakcijas. Tika izmantots īsais saraksts ar ceļu konstruēšanā neizmantotajiem metabolītiem. Tika izveidotas 12 savienojumu matricas, kuras attēlo glicerīna transformācijas ceļus.

Nākamajā solī tika meklēti visi ceļi izveidotā savienojumu matricā, kuri ļauj savienot glicerīnu un *Z.mobilis* datormodeli, izmantojot KEGGID modeļa laukus (sk. 4. att.) (Y ass – logaritmiskā mērogā).



4. att. Iespējamo bioķīmisko ceļu skaits starp glicerīnu un *Z.mobilis* modeli

Pēdējā solī tiek veikta sameklēto ceļu analīze. Analīzes laikā tiek noraidīti ceļi, kas neizpilda minimālās iestatītās prasības, piemēram – nespēj saražot biomasu vai nenodrošina pietiekami lielu mērķa plūsmu biotehnoloģiski nepieciešamajā ceļā stacionārajā stāvoklī. Ceļi, kas izpilda minimālās prasības, tiek saranžēti pēc izvēlētās kritēriju secības vai pēc kopējā svērtā kritērija (KSK) izmantojot daudzparametru svērtās ranžēšanas metodi, atkarībā no uzdevuma iestatījumiem. Pēc analīzes tika noraidīta puse no ceļiem pie ceļa garuma 5 reakcijas, un ar maksimālo ceļu garuma 7 reakcijas analizēto ceļu skaits pārsniedz robežu 5000 ceļi.

Minimālais reakciju skaits, kurā var savienot glicerīnu ar *Z.mobilis* modeli, ir 5 reakcijas. Var secināt, ka *Zymomonas mobilis* piedāvātās modifikācijas glicerīna pārstrādei (Pentjuss et al., 2013) varētu nebūt darboties spējīgas termodinamisko parametru dēļ (tieki piedāvātas 2 reakcijas).

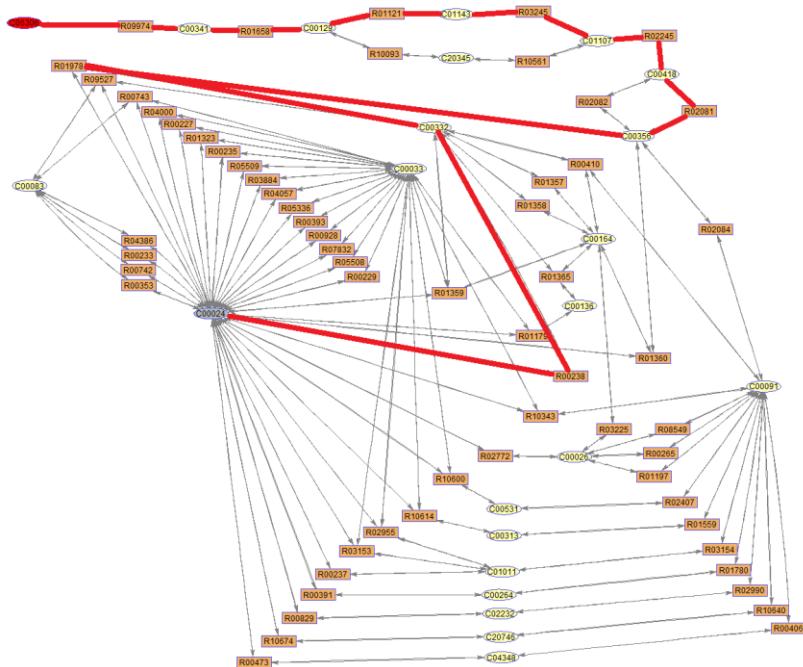
5.3. Bakterijas *E.coli* metaboliskas inženierijas analīze α -Pinene ražošanai

α -Pinene ir svarīgs dabisks produkts, kurš plaši tiek izmantots atšķirīgās nozarēs. Šobrīd α -Pinene, kurš tiek izmantots ražošanā, pārsvārā ir iegūts no kokiem vai koksnes pārstrādes procesā. Yang pētījumā (Yang et al., 2013) ir parādīts, ka *E.coli* mikroorganismā tika ievietots metaboliskais ceļš, kurš sastāv no 9 gēniem, ekspresētiem no citiem mikroorganismiem.

Šie dati tika pārbaudīti ar *sAnalyzer* palīdzību - tika konstruēta α -Pinene vielas transformācijas ceļu savienojuma matrica. Tika pārlasītas 7046 reakcijas un 3168 metabolīti no KEGG datu bāzes. Izveidotajā metaboliskajā tīklā ir 6176 virsotnes, kur katra virsotne ir metabolīts vai reakcija, un 9808 šķautnes, kuras ir savienojumi starp metabolītu-reakciju-metabolītu.

Izmantojot daudzparametru svērtās ranžēšanas metodi, *sAnalyzer* atgrīzeza 252 iespējamos metaboliskos ceļus, kuri var savienot sākotnējo metabolītu ar kādu no modelī esošajiem metabolītiem (sk. 5. att.). Ar sarkano šajā attēlā ir parādīts ceļš,

par kuru ir rakstīts izvēlētajā piemērā (Yang et al., 2013). Iegūtie ceļi ir atšķirīgi ar reakciju skaitu. Tika atrasti 8 ceļi ar 7 reakcijām un 16 ar 8 reakcijām. Pārējos ceļus veido 9 reakcijas. Pirmajiem 12 labākajiem atrastajiem ceļiem konstruēta ceļa FBA vērtība ir 4.05 mmol/(gDW·h). Yang pētījumā minētais ceļš atrodas 45. vietā saskaņā ar ranžēšanas datiem ar ceļa FBA vērtību 4.05 mmol/(gDW·h). Pretstatā publikācijā (Yang et al., 2013) piedāvātajam 9 reakciju metaboliskās inženierijas risinājumam *sAnalyzer* piedāvāja virkni ceļu ar lielāku *a-Pinene* ražošanas plūsmu ar TDR 45kJ/mol.



5. att. *a-Pinene* iespējamo 252 ražošanas ceļu vizualizācija ar TDR 45 kJ/mol

5.4. Baktērijas *E.coli* metaboliskās inženierijas analīze *1-butanol* ražošanai

1-butanol ir garo kēžu alkohols, kurš var tikt izmantots vairākos veidos. Viens no tiem ir biodegvielas ražošana. Salīdzinot ar etanolu, *1-butanol* vielai ir vairākas priekšrocības tā izmantošanai kā degvielai augstākas enerģijas saturā un augstākas hidrofobitātes dēļ. *1-butanol* var tikt ražots gan ķīmiski, gan mikroorganismu fermentācijas procesā.

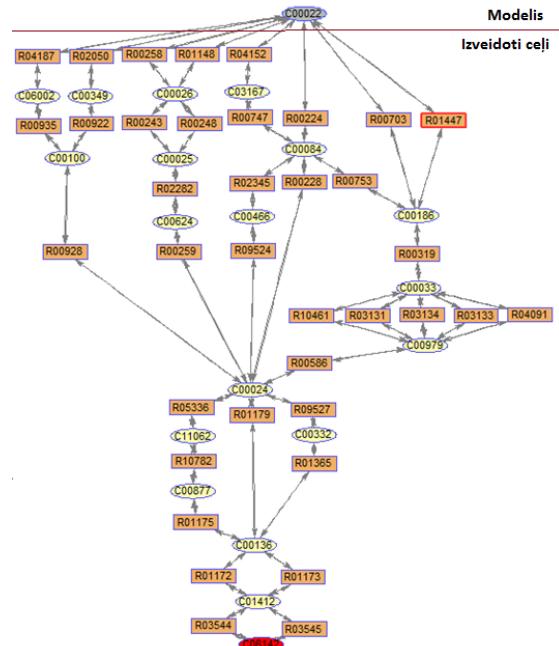
Iz vairāki pētījumi par *E.coli* organizma izmantošanu *1-butanol* ražošanā (Atsumi et al., 2008; Lee et al., 2008; Ranganathan, Maranas, 2010). Viens no iespējamajiem

E.coli modifikācijas veidiem ir *C.acetobutylicum* strādājošā *1-butanol* ražošanas ceļa ievietošana *E. coli* (Atsumi et al., 2008).

Citā pētījumā papildu šādam dabiskam ceļam tika meklēti dabā neeksistējušie *1-butanol* ražošanas ceļi (Ranganathan et al., 2010). Tika izmantota uz grafiem bāzēta jauno ceļu meklēšanas procedūra, kas izmanto datus no KEGG un BRENDU datu bāzes, jo KEGG datu bāzē nav pilno datu par garo ķēžu alkohola ražošanas ceļiem, tāpēc šie dati tika iegūti no BRENDU datu bāzes (Ranganathan et al., 2010).

Šie dati tika pārbaudīti, izmantojot *sAnalyzer* rīku. Kā modeļa savienojuma metabolīts tika izvēlēts *pyruvate* (C00022), maksimālais ceļa garums 7 reakcijas, minimālais biomasas augšanas ātrums 0.01% no nemodificēta modeļa biomasas reakcijas plūsmas un reakciju termodinamikas robežvērtība 35kJ/mol. Tika izmantots garais saraksts ar ceļu konstruēšanā neizmantotajiem metabolītiem.

sAnalyzer analīzē tika atrasti 23 iespējamie ceļi (sk. 6. att.), kuri savieno *E.coli* mikroorganisma modeli ar *1-butanol* vielu. Visi publikācijās minētie ceļi tika noraidīti. Iesējams noraidīšanas iemesls - augstāk minētie ceļi nenodrošina organisma iepriekš minimālā līmenī (0.01 mmol/(gDW·h)) vai kāda no ceļa reakcijām, saskaņā ar termodinamikas datiem, var notikt tikai pretējā virzienā.



6. att. *sAnalyzer* piedāvātie *1-butanol* ražošanas ceļi ar TDR 35 kJ/mol

Minētie praktiskā pielietojuma piemēri ir pārbaudīti tikai modelēšanas līmenī. Pirms to ieviešanas ir nepieciešama bioloģiskā analīze. Tomēr piedāvātais plašais risinājumu klāsts ir konkurētspējīgs ar analizētajās publikācijās izmantotajiem risinājumiem.

SECINĀJUMI

Galvenie darba rezultāti

Tika automatizēta bioķīmisko ceļu konstruēšana un analīze šasijas organismos metaboliskās inženierijas uzdevumu risināšanai.

- Ir analizētas pastāvošas metaboliskās inženierijas un datorizētās bioķīmisko ceļu konstruēšanas metodes.*

Eksistē vairāku tipu rīki metaboliskās inženierijas uzdevumu risināšanai un datu analīzei. Ir publicēti rīki, kuri ļauj meklēt un konstruēt bioķīmiskos ceļus, tomēr lietošanai pieejami un darbojas tikai daži no tiem. Šādu rīku trūkums ir tas, ka esošie rīki neļauj analizēt visu risinājumu telpu, bet gan aprobežojas ar zināmo organismu metaboliskiem ceļiem. Līdz ar to – netiek analizēti iespējamie jaunie metabolītu transformācijas ceļi. Turklat pieejamie rīki neļauj analizēt konstruētos ceļus ar šasijas mikroorganisma modeļa palīdzību.

- Ir analizēti bioķīmisko ceļu automātiskās konstruēšanas un analīzes procesā iekļaujamie kritēriji un parametri.*

Par svarīgākajiem parametriem ceļu analīzei tika izvēlēti analizētā ceļa un šasijas organismu biomasa reakcijas plūsmu bilances rezultāts, minimālās plūsmas lielums caur biomasa augšanas reakciju, reakciju skaits ceļā, no analīzes izslēgtas reakcijas un metabolīti, reakciju termodinamikas dati reakciju virziena noteikšanai, nelīdzsvarotu reakciju skaits ceļā, ceļā ietverto un organismā esošo metabolītu daudzums.

- Ir izstrādāta bioķīmisko tīklu automātiskās konstruēšanas procedūra ar iespēju iekļaut esošo šasijas organismu metaboliskā tīkla modeli.*

Izstrādātā procedūra sastāv no 4 soļiem: mērķa un parametru izvēles, reakciju tīkla izveides, metabolisko ceļu meklēšanas un rezultātu analīzes. Izstrādātā procedūra ļauj pārmeklēt visu modifikāciju risinājuma telpu un analizēt atbilstošus metaboliskos ceļus ar šasijas organismu modeļa palīdzību.

Bioķīmisko ceļu konstruēšana (reakciju tīkla izveide) tiek veikta, balstoties uz definēto mērķi un sākotnējiem parametriem, izmantojot datus no tiešsaistes bioloģisko objektu datu bāzes. Tieki iegūti visi ceļi, kuri var savienot mērķa metabolītu ar doto šasijas organismu metaboliskā tīkla modeli. Ceļu analīze tiek veikta, balstoties uz iekļautajiem kritērijiem un parametriem. Ir iespējami vairāki analīzes veidi.

4. Ir izstrādāts automatizētas bioķīmisko ceļu konstruēšanas un analīzes rīks.

Uz izstrādātās procedūras pamata tika izstrādāts brīvpieejas rīks *sAnalyzer*, ar kura palīdzību ir iespējams veikt metabolisko ceļu risinājuma telpas pārmeklēšanu un ceļu analīzi. Šis rīks implementē visus piedāvātās procedūras soļus. Dažus soļus ir iespējams atkārtot, t.i., ir iespējams veikt ceļu analīzi ar atšķirīgiem analīzes parametriem, neatkarojot vēlreiz iepriekšējos procedūras soļus. *sAnalyzer* atgriež sarakstu ar ceļiem un to ranžēšanas kritēriju vērtībām, kā arī šasijas organisma modeli ar pievienošanai ieteicamajiem metaboliskajiem ceļiem.

5. Ir analizēta iespējamā risinājumu telpa atkarībā no izvirzītajiem ierobežojumiem.

Tika pārbaudīta iespējamā risinājuma telpa, veidojot lielus metabolītu transformāciju bioķīmiskos tīklus. Tika konstatēts Ričarda līknes tipa analīzē iekļauto reakciju un metabolītu skaita pieaugums, un kopējais risinājuma telpas pieaugums atkarībā no pielaujamo reakciju skaita. Izmantojot maksimālo reakciju skaitu ceļā un termodynamikas ierobežojumus, ir iespējams samazināt kopējo risinājuma telpu, iekļaujot tikai reakcijas, kuras var notikt tikai vienā virzienā ar konkrētiem procesa parametriem.

6. Demonstrēta metaboliskās inženierijas uzdevumu risināšanas procedūra.

Tika veikti eksperimenti, lai demonstrētu procedūras darbību metaboliskās inženierijas uzdevumu risināšanā, izmantojot *sAnalyzer* rīku.

Pārbaudītas *Z.mobilis* baktērijas iespējas nodrošināt glicerīna patēriņu un organisma augšanu pie atšķirīgiem termodinamikas parametriem. *Z.mobilis* eksperimentā tika parādīts, ka glicerīna pārveidei par etanolu nepieciešams ievietot 5-9 reakcijas. Tika veikta *E.coli* baktērijas izmantošanas analīze jauno produktu ražošanai, salīdzinot rezultātus ar iepriekš publicētajiem datiem. *a-Pinene* eksperimentā atrasti ceļi, kuru *a-Pinene* ražošanas plūsmas ātrums ir 40 reizes lielāks par publikācijā piedāvātā risinājuma plūsmas ātrumu. *1-butanol* eksperimentā, pretstatā publikācijā minētajiem ceļiem, kuru izmantošana saskaņā ar *sAnalyzer* datiem neļauj ražot biomasu, tika atrasti ceļi, kas ļauj ražot gan *1-butanol* vielu pieckārtīgā apjomā, salīdzinot ar publikācijā minētajiem, gan arī biomasu.

sAnalyzer spēja atrast gan publikācijās nosauktos ceļus, gan arī piedāvāja ceļus, kuriem iespējamais vielas ražošanas potenciāls ir lielāks par piedāvāto ceļu. Tika konstatēta trūkstošās vai nepilnīgās informācijas ietekme uz ceļu analīzes rezultātiem.

Secinājumi un attīstības perspektīvas

Darba gaitā tika gūti vairāki secinājumi, būtiskākie no tiem ir minēti šajā apakšnodalā.

1. Ir iespējams automātiski atlasīt šasijas organismā ievietojamos ceļus, kuri nodrošina produkta ražošanu un/vai substrāta patēriņu, kā arī organisma augšanu.

2. Modeļu un darba autora izstrādātās procedūras pielietošana modifikāciju projektēšanas stadijā ļauj samazināt manuālo darbu, nepieciešamo resursu daudzumu un laiku, pateicoties automātiskajai risinājumu analīzei un ranžēšanai.
3. Bioloģisko risku formalizācija un to ietveršana datorizētā risinājumu telpas pārmeklēšanas procesā novērš riskantu risinājumu pārbaudi daudz dārgākajos bioloģiskajos eksperimentos.
4. Konstatēts Ričarda līknes tipa risinājumā telpā iesaistīto metabolītu un reakciju skaita pieaugums atkarībā no pieļaujamo reakciju skaita ceļā.
5. Reakciju termodynamikas robežvērtību izmantošana ievērojami samazina risinājuma telpu un analīzei patēriņto laiku.
6. Metaboliskās inženierijas risinājumu meklēšanu ir lietderīgi veikt, sākot ar mazu reakciju skaitu ceļā, reakciju skaitu palielinot, ja piemērots risinājums nav atrasts.

Kā darba attīstības perspektīvas var iezīmēt vairākus virzienus.

1. *sAnalyzer* darbībai ir nepieciešams *Matlab* rīks, kas ir maksas produkts, ar *COBRA* paplašinājumu. Ir iespējams implementēt *COBRA* atbalstu atvērta koda rīkos un aizvietot maksas *Matlab* rīku ar atvērta koda risinājumiem, piemēram, *Octave*, *Scilab*.
2. Izstrādāt *sAnalyzer* rīka versiju ar *Python* programmēšanas valodu, jo *COBRA* rīks ir pielietojams, arī izmantojot *Python* programmēšanas valodu. Šādā gadījumā *Matlab* rīks nav nepieciešams, kas ļauj *sAnalyzer* rīku brīvi izmanot.
3. Pievienot papildus reakciju un metabolītu datu avotus (datu bāzes) *sAnalyzer* rīkā.
4. Izstrādāt iespējamo gēnu izvēles un analīzes algoritmu, kurš piedāvās iespējamo gēnu kombinācijas, balstoties uz reakciju enzīmu datiem. Papildus, var tikt izmantoti enzīmu kinētiskie dati, lai prognozētu reakciju iespējamību šasijas organismā.

PARTICULARS

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

Scientific Adviser of the Doctoral Thesis: Dr.sc.ing., Professor Egils Stalidzans.

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, 2012. Minutes No. 4.

Official Reviewers:

1. Associate Professor of University of Latvia, Dr.sc.comp. Juris Viksna;
2. Professor of Riga Technical University, Dr.habil.sc.ing. Jurijs Merkurjevs;
3. Senior Research Fellow 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 ____ on _____, 2015, Room ___, Faculty of Information Technologies. Liela 2, Jelgava.

The thesis can be accessed at the LUA Fundamental Library, Liela 2, Jelgava, and online at http://llufb.llu.lv/promoc_darbi.html

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

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

TABLE OF CONTENTS

Approbation of PhD thesis.....	27
Introduction	29
Theme topicality	29
The aim and the tasks of the PhD thesis	30
Research methods.....	31
Scientific novelty.....	31
Theses.....	32
PhD thesis structure and volume	32
1. Metabolic engineering	32
2. Metabolic networks analysis	33
3. Automated procedure for biochemical networks construction.....	35
4. Software tool sAnalyzer	37
5. Automated metabolic engineering solution space analysis.....	38
5.1.Constraint depending analysis of solution space	39
5.2. <i>Z.mobilis</i> bacteria metabolic engineering analysis for glycerol utilisation	40
5.3. <i>E.coli</i> bacteria metabolic engineering analysis for <i>a-Pinene</i> production.....	41
5.4. <i>E.coli</i> bacteria metabolic engineering analysis for <i>l-butanol</i> production.....	42
Conclusions	44
The main PhD thesis results	44
Conclusions and development prospects	46
Bibliography	47

APPROBATION OF PHD THESIS

The research results are presented in the following publications:

1. **Meitalovs, J.** (2012a) Process of genetic modification of microorganisms from the point of view of the software engineering. In: Proceedings of the *5th International Scientific Conference on Applied Information and Communication Technologies*, pp. 348–351 (Indexed in the AGRIS and EBSCO databases).
2. **Meitalovs, J.** (2012b) The concept of metabolic engineering software tool. In: Proceedings of the *5th International Scientific Conference on Applied Information and Communication Technologies*, pp. 352–355 (Indexed in the AGRIS and EBSCO databases).
3. **Meitalovs, J.**, Bulipopa, N., Kovalonoka, O. (2012) Modeling and design tools for synthetic biology. In: Proceedings of the *5th International Scientific Conference on Applied Information and Communication Technologies*, pp. 341–347 (Indexed in the AGRIS and EBSCO databases).
4. **Meitalovs, J.** (2012c) Software tool for probabilistic metabolic pathways construction. In: Proceedings of the *IEEE 13th International Symposium on Computational Intelligence and Informatics (CINTI)*, pp. 405–408 (Indeksēts „Scopus”, „Web of knowledge”, un „IEEE explore” daubāzēs).
5. **Meitalovs, J.**, Stalidzans, E. (2013a) Connectivity analysis of metabolites in synthetic metabolic pathways. In: Proceedings of the *12th International Scientific Conference ‘Engineering for Rural Development’*, pp. 435–440 (Indexed in the Scopus and AGRIS databases).
6. **Meitalovs, J.**, Stalidzans, E. (2013b) Analysis of synthetic metabolic pathways solution space. In: Proceedings of the *2013 International Conference on System Science and Engineering (ICSSE)*, pp.183–187 (Indexed in the Scopus, Web of Knowledge, and IEEE Explore databases).
7. **Meitalovs, J.**, Stalidzans, E. (2015) Impact of thermodynamic constraint to the solution space of metabolic pathway design using *sAnalyzer* tool. *Baltic Journal of Modern Computing*, 3, pp. 164-178 (Indexed in the ProQuest and EBSCO databases).

The research results were presented at the following conferences:

1. Rutkis R., **Meitalovs J.**, Odzina I., Kalnenieks U., In silico metabolic engineering of Zymomonas mobilis for glycerol consumption, *SB5.0: the Fifth International Meeting on Synthetic Biology*, July 15-17, 2011, Stanford, USA.

2. **Meitalovs J.** Process of genetic modification of microorganisms from the point of view of the software engineering. *Applied Information and Communication Technologies*, April 26-27, 2012, Jelgava, Latvia.
3. **Meitalovs J.** The concept of metabolic engineering software tool. *Applied Information and Communication Technologies*, April 26-27, 2012, Jelgava, Latvia.
4. **Meitalovs J.**, Bulipopa N., Kovalonoka O. Modeling and design tools for synthetic biology. *Applied Information and Communication Technologies*, April 26-27, 2012, Jelgava, Latvia.
5. **Meitalovs J.**, Software tool for probabilistic metabolic pathways construction, *13th IEEE International Symposium on Computational Intelligence and Informatics*, November 20-22, 2012, Budapest, Hungary.
6. **Meitalovs J.**, Stalidzans E., Connectivity analysis of metabolites in synthetic metabolic pathways, *12th International Scientific Conference “Engineering for rural development”*, April 26-27, 2012, Jelgava, Latvia.
7. **Meitalovs J.**, Stalidzans E., Analysis of synthetic metabolic pathways solution space, *ICSSE 2013 IEEE International Conference on System Science and Engineering*, July 4-6, 2013, Budapest, Hungary.
8. **Meitalovs J.**, Probabilistic metabolic pathways construction and analysis tool, *SB6.0: The Sixth International Meeting on Synthetic Biology*, July 9-11, 2013, London, UK.
9. **Meitalovs J.**, Stalidzans E., Applications of sAnalyzer – tool for metabolic network construction, *2nd Congress of Baltic Microbiologists*, October 16-19, 2014, Tartu, Estonia.

The research results were presented at the following seminars:

Meitalovs J., Applications of the *sAnalyzer* tool, *110th Biosystems Group Seminar*, May 23, 2014, Jelgava, Latvia.

Participation in the projects related to the PhD thesis development:

Establishment of Latvian Interdisciplinary Interuniversity Scientific Group of Systems Biology No.2009/0207/1DP/1.1.2.0/09/APIA/VIAA/128, (2009 – 2012, researcher).

INTRODUCTION

Theme topicality

In post-genome era, scientists are increasingly using naturally occurring reactions in not typical organisms for achieving biotechnological objectives. For example, in 1928, *A. Fleming* discovered the penicillin – the first antibiotic, and in 1941 the technology was developed for production of antibiotics using fermentation process. In 1978, the insulin coding gene was successfully implemented in the *E.coli* bacteria to solve the task of insulin synthesis ("Herbert W. Boyer"). Such biochemical applications are called biotechnology, and perhaps it is one of the fastest growing scientific disciplines.

Biotechnology is the industry dealing with the product development for human needs using microorganisms. Biotechnological products are widely used in many industries, e.g. food, medicine, chemistry, biofuel, agriculture, etc. (Adrio, Demain, 2010; Atsumi et al., 2008; Carothers et al., 2009; Serrano, 2007; Trinh, Srienc, 2009; Yang et al., 2013; Yu et al., 2010; "Contained use of genetically ...").

In less than 200 years scientists have moved from one reaction between minerals to thousands of interconnected and controlled reactions which allow the bacteria to live and produce important compounds. Mathematical description and information technologies for analysis of such systems built from thousands reactions has become a prerequisite for available information systematization and exploitation (Kuprijanov et al., 2013; Lehrach et al., 2011; Zazzu et al., 2013). Nowadays, biotechnology is a multidisciplinary field – it includes a number of different sciences, e.g. biology, chemistry, metabolic engineering, information technologies, process control, and optimisation.

Metabolism, or metabolic process, is defined as a set of various interrelated biochemical reactions in cells or in an organism, which is necessary to ensure the ability to live (Vīgants, 2008).

One of the most important objectives of metabolic engineering is increasing production of biochemical compounds using genetic modifications. In increasing number of activities with biological systems engineering approach is used: the system is built or modified only after a sufficient detailed understanding of the system and its computer model analysis (Tomlin and Axelrod, 2005; Tyo et al., 2007; Vital-Lopez et al., 2006; Mendes and Kell, 1998).

Metabolic pathway is a functional unit of the biochemical network. In biochemistry, metabolic pathway is defined as a sequence of reactions, which transform the initial metabolite into another molecule or molecules (Berg et al., 2002). It is possible to make genetic manipulations in the organism to change metabolic network and reach a production of the necessary substance.

One of the most important steps of the metabolic engineering is the organism design. It can be done using modeling techniques and simulation: to make the

design of the modified organism according to the objective function and to verify it's expected potential (Macdonald et al., 2011).

The use of computer experiments and models of biological systems instead of biological experiments is cheaper and does not require equipment and chemical resources (Gendrault et al., 2011, Wiechert, 2002). Modeling of cell systems means merging of molecular components in mathematically and biologically accurate way to create *in silico* mathematical objects. For such reconstructions it is required to use specific biological information.

For computational modeling the biochemical network's stoichiometric matrix analysis can be used. The main problem for such analysis is the computational amount caused by combinatorial explosion of possible number of pathways in genome-level metabolic networks (Blum, 2009). Usually, the insertion of the new metabolic pathways into chassis organism model and its analysis is performed manually, based on the idea of biologist. The scanning of the wide solution space of possible modifications needs a lot of manual work. Hence, requirements for large-scale analysis automation rise, which is another information technologies potential application.

The range of existing software tools allow solving a variety of problems, but the problem of new non-naturally occurring metabolic pathway searching cannot be currently solved with the help of specialised software. At present, there are no computerised tools which allow testing a wide range of possible solution space and making recommendations on possible modification results. Although for such problem solving there have been attempts to develop algorithms for searching all possible transformation pathways (Blum, Kohlbacher, 2008; Ranganathan, Maranas, 2010; Rodrigo et al. 2008; Wang, Hatzimanikatis, 2006; Arita, 2000; Croes et al., 2006), the problem still exists, as the available algorithms are not implemented in computational tools, or available versions of existing tools are not maintained in functional state, so they are not fully operational. It used data is obsolete because of online data capture is not used. Many of software products are designed for one specific problem solving. In addition, they are not using a number of analysis methods which allow to answer the question – is the chosen pathway able to provide both needed substance production and the organism growth? Many constraints, e.g. reaction's thermodynamic data and metabolic characteristics of chassis organism model, are not used either.

The situation indicates topicality of the use of computer technology in metabolic engineering and determines the aim and tasks of the PhD thesis.

The aim and the tasks of the PhD thesis

The aim of the PhD thesis is to automate the design and analysis of biochemical networks in chassis organisms for solving metabolic engineering tasks. In order to achieve the aim of the PhD thesis several tasks were solved:

1. to analyse existing metabolic engineering and computational methods for biochemical networks design;
2. to analyse criteria and parameters for automatic design and analysis of biochemical networks;
3. to develop a procedure for automatic biochemical networks design with an option to include metabolic model of chassis organism;
4. to develop a tool for automated metabolic networks design and analysis;
5. to analyse possible solution space of metabolic engineering task depending on given restrictions;
6. to demonstrate operation of metabolic engineering tasks solving procedure.

Research methods

1. *Matlab*-based automated tool for biochemical pathways design and analysis.
2. Data from online KEGG database is used for constructing metabolites transformation network.
3. For biochemical pathways search in constructed metabolites transformation network the modified Depth-First search algorithm is used. A method for automatic visualisation of biochemical pathways was developed, which uses *Matlab* built-in functions for graph visualisation.
4. For analysis of biochemical models the flux balance analysis tool COBRA is used, which is based on linear algebra methods.
5. For modeling biochemical processes stoichiometric modeling methods are used, SBML standard is used for model definition.
6. For analysis of constructed biochemical pathways the developed multi-criteria analysis algorithm is used. Metabolic pathways characterising parameters with changeable weighting coefficients are used for ranking pathways.

Scientific novelty

1. Automatic biochemical networks design and construction procedure is developed.
2. Automatic biochemical pathways analysis algorithm is developed, which takes into account thermodynamics, chassis organism's metabolites and flux balance analysis results.
3. Multiparameter weighted ranking method of constructed biochemical pathways was developed.
4. The tool was developed for automatic scanning of possible genetic modifications solution space and constructing of biochemical networks according to the metabolic engineering task.
5. Richard curve type of the growth of solution space was found depending on the allowable number of reactions in the pathway.

Theses

1. Increasing the allowable number of reactions in the pathway, the possible modification solution space is growing rapidly, but then reaches a saturation limited by number of certain metabolites.
2. It is possible to automatically construct and rank biochemical pathways taking into account thermodynamic constraints, chassis organism models and flux balance analysis data.
3. The automatic scanning of the solution space offers a set of solutions competitive to the published data.

PhD thesis structure and volume

The PhD thesis is written in Latvian containing abstract, introduction, 5 chapters, conclusions, bibliography, and 6 annexes, including 5 tables, 58 figures, 128 pages in total. 133 literature sources were used.

1. METABOLIC ENGINEERING

Cell metabolism consists of a complex chemical reactions network, which transforms molecules and acts to transform nutrients into energy and biomass. Metabolic pathways represent functional units of such network; they are responsible for specific metabolic processes, e.g. carbon source degradation or amino acids synthesis. Biochemistry defines metabolic pathway as a sequence of chemical reactions, where the initial metabolite is gradually transformed into another molecule or molecules (Berg et al., 2002).

For example, glycolysis metabolic pathway converts glucose to pyruvate. Thermodynamic data can be used to calculate reactions balance and to predict whether a particular reaction can take place at the given conditions.

The presently known metabolic pathways were discovered thanks to meticulous work with specific organism models. However, in reality, metabolic networks are more variable and more interrelated than the mostly linear model metabolic pathways (Cordwell, 1999; Vīgants, 2008).

Using gene-enzyme (protein), enzyme-reaction and reaction-metabolite relations, organism specific cellular metabolic networks can be reconstructed based on data from biological object databases. For such networks *in silico* analysis can be performed.

By analysing these metabolic networks, it can be determined which products are produced during the growth of microorganisms. It is possible to simulate changes in metabolic pathways in order to increase a yield of biotechnological process or to start new substance production, which is not present in wild type microorganism species.

The existence of alternative metabolic pathways is the result of organism adaptation to environmental conditions. Therefore, the knowledge about all possible pathways that can transform source metabolite into target metabolite may help to understand metabolism. However, the determination of metabolic pathways in experimental way is time and resources consuming process (Blum, 2009).

Metabolic engineering is optimisation of genetic and regulatory processes in cells to increase production of certain product metabolites. Natural metabolic networks can be improved both with modeling and optimisation methods, and systems and synthetic biology methods.

Metabolism of organisms can transform source substances to variety of products, many of them can be used for industrial and commercial purposes. Most widely systems and synthetic biology methods are used in medicine and environment, e.g. drugs synthesis, therapy, biosensors, biofuels, pharmacy, biomaterials and biochemical production, etc. (Gendrault et al., 2011; Khalil, Collins, 2010; Tepper, Shlomi, 2010).

Biotechnology is the science that allows the use of existing set of reactions in microorganisms to produce substances needed to humans using organisms modified with systems and synthetic biology methods. Systems biology as a science is not new but it is growing rapidly due to development of new technologies. Systems biology is defined as a global research of biological processes in the cell, organism, or even community level, its molecular interactions of components and functions ("What is systems biology?"). The term *synthetic biology* appeared recently, and it is related to completely different research directions. Several of them are focused on engineering, where natural molecules are combined into specific complexes and inserted into cells, giving them new features and functionality genetically modifying them (Khalil, Collins, 2010; Mannelli et al., 2003). Synthetic biology is defined as use of engineering principles in development of biological systems ("What's in a name?").

In metabolic engineering, several iterations are needed to ensure the successful integration of the individual steps. One of the most important steps of organism modification procedure is process design. One of the methods that allow performing it is computational modeling and simulation. Those are methods to design modified organisms based on the objective function and test its potential (Macdonald et al., 2011). Currently, there are several software tools that allow facilitating and accelerating the design step. Such tools can be divided into several categories according to their application.

2. METABOLIC NETWORKS ANALYSIS

Cells contain many metabolic pathways to ensure the natural survival processes. To understand how functions are divided in a cell the metabolic networks analysis can be applied. Due to a large number of possible metabolic pathways and their

complexity, it is not always intuitively understandable which pathways are used to reach the target cell functions.

Metabolic pathways can be represented as graphs, and graph theory methods can be applied to calculate biotransformation pathways between source and target metabolites. The pathway searching algorithms can be used to find one or many pathways that connect two graph nodes. These algorithms vary both in their potential application and their complexity.

Metabolic networks analysis can be done using organism stoichiometric modeling. It is a constraint-based modeling method. Metabolic network stoichiometry describes relations of substrates and products in chemical reactions of the cell metabolism. By using such modeling technique the metabolic fluxes in organism model can be determined and conclusions about metabolic processes in the organism can be made. One of the frequently used tools for metabolic pathways stoichiometric analysis is flux balance analysis (FBA). Visual representation of the network is important for its understanding; however, more can be obtained from its mathematical representation. It is possible to represent metabolic network using stoichiometric matrix where reactions and metabolites are used as columns and rows.

In the biotechnological production, one of the questions is as follows: by using which biochemical reactions one substance can be transformed into another one? Additionally, it is necessary to find the minimum set of reactions that allow producing the necessary metabolite.

Automatically created metabolic pathways can be used to test certain theories about conversion of metabolites, possible product biosynthesis or substrates biodegradation. Generally available methods allow creating metabolic pathways, but in order to test them the user should manually check the accuracy of the pathway and implement it to the chassis model. This solution is suitable for small models but for larger metabolic networks it requires a lot of manual work. Another problem of biotechnological production is when necessary substrate and product are known but the chassis organism which can be used for genetic modifications is not known.

In such case, there is no tool which can help to choose appropriate solution. Currently, the selection of organism and search of metabolic pathways happens based on experiments and/or based on scientific knowledge and experience. But often it does not provide the optimal result, because not all possible solutions are tested. Therefore, biologist can only hope that such selected modification will succeed and will be competitive with other biotechnological solutions.

This work offers a methodology, a software tool implementing it and practical examples that show an automated metabolic pathways construction algorithm and its implementation. The proposed solution combines existing methods and improves their use in industrial application, as well as provides enhanced metabolic pathways validation scheme.

3. AUTOMATED PROCEDURE FOR BIOCHEMICAL NETWORKS CONSTRUCTION

To avoid the above mentioned disadvantages in metabolic network modeling, automated procedure for biochemical networks construction and analysis was developed. It is divided into four main steps (Figure 1). The result of each step is used in the next procedure steps.

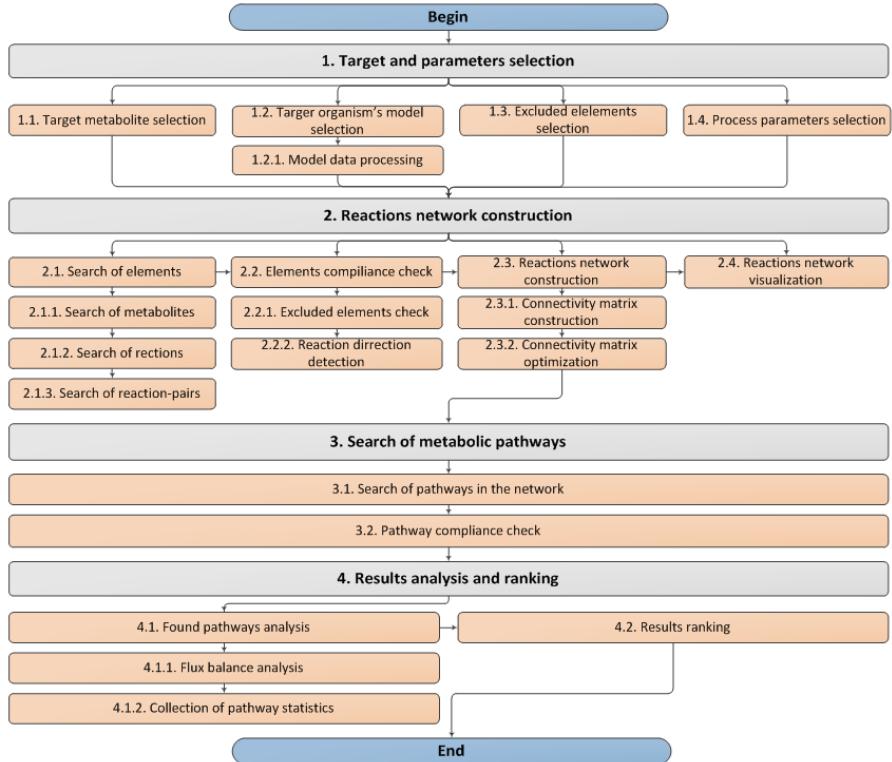


Figure 1. Steps of automated procedure for biochemical networks construction

The first step is divided into four interlinked sub-steps. The first step starts with criteria and model selection to define the biotechnological process. During this step the source metabolite and chassis organism model are selected, the list of metabolites and reactions that cannot be used in network constructions is defined and additional parameters are chosen, e.g. minimum biomass reaction flux, connection metabolites, thermodynamic data, maximum number of reactions in pathway. Such process parameters are used in the next procedure steps.

At the next step biochemical reactions and metabolites are searched and connected in the transformation network of the source metabolite. The author offers to use all possible connections of metabolites, e.g. examine all possible reactions that can transform the source metabolite. For each metabolite all reactions where the particular metabolite is participating are selected from the KEGG database.

For each reaction metabolites reactions-pairs are found. The found metabolites and reactions are connected into connection matrix which represents biochemical network. For each metabolite all reactions are found. It continues until the maximum number of reactions in pathway is reached. As the result a connection matrix is returned.

At the third step all metabolic pathways should be found in the constructed biochemical network that can connect the source metabolite with any other metabolite in the chassis organism model. Therefore, it is proposed to use an automated approach for searching every possible metabolite transformation pathway.

Since the matrix created in the previous step can be represented as a graph, it is proposed to use the algorithm for searching all paths in the graph. Many algorithms are offering the following functionality, thus it is necessary to choose the most appropriate and implement it. This algorithm should return a list of all paths that connect two nodes of the graph, i.e., metabolic pathways in biochemical network.

The last step provides analysis and ranking of the metabolic pathways found in the previous step. First of all, pathways should be analysed one by one. In this step, all metabolites from all reactions in the pathway should be found, based on parameters defined in the first step. Next, the analysed pathway should be added to the model and tested if it is capable of providing atom flow. Flux balance analysis algorithm can be used to determine the flux in the constructed pathway. If the fluxes of the analysed pathway and biomass reaction are very low or zero, this pathway can be excluded from the future analysis. Other pathways can be ranked based on calculated pathway parameters.

The author proposes to use the following metabolic pathways evaluation form:

- taking into account total number of metabolites;
- taking into account the number of common metabolites of model and pathway;
- taking into account the number of metabolites added to the model;
- taking into account the number of reactions in the pathway;
- taking into account the number of reactions with known thermodynamics;
- taking into account the number of common reactions of model and pathway;
- taking into account the number of unbalanced reactions in the pathway;
- flux balance analysis results of constructed pathway;

- flux balance analysis results of biomass reaction in model with added constructed pathway.

Multiparameter weighted ranking method of constructed biochemical pathways was developed, which takes into account 20 separate parameters and coefficients. This method can be adapted, choosing only those parameters that have higher priority in biotechnological task solving.

4. SOFTWARE TOOL SANALYZER

Based on the above described steps of the procedure a software tool *sAnalyzer* was developed. *sAnalyzer* implements all steps of the procedure and proves that by using such procedure it is possible to create new, non-naturally occurring metabolic pathways. *sAnalyzer* allows to test all possible combinations of biochemical reactions that connect the source metabolite with the model.

The tool is developed in *Matlab* environment. Programming in *Matlab* is made in executable script files where each file may contain one or more functions. *Matlab* environment was chosen, because it allows using both *Matlab* built-in functionality and many add-in toolboxes. One of them *COBRA toolbox* (Becker et al., 2007) was used for metabolic pathways flux balance analysis.

sAnalyzer includes four main modules that allow to manipulate with the model and obtained data (Figure 2).

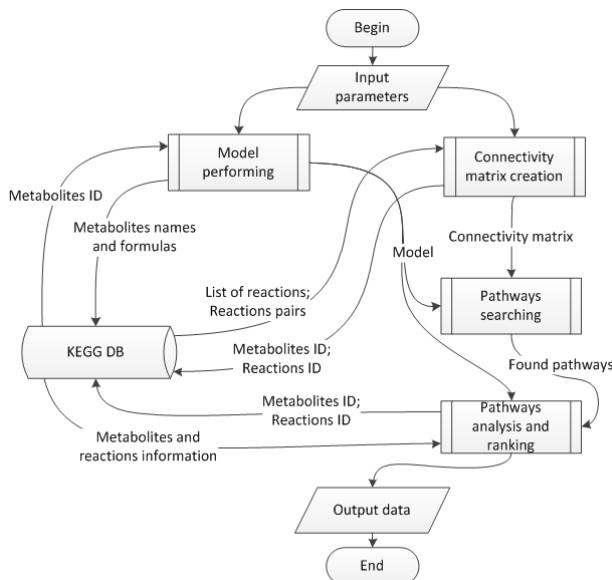


Figure 2. *sAnalyzer* program simplified operation algorithm

This tool has a number of innovative features that differentiate it from the existing tools. As the most important characteristics may be mentioned:

1. *sAnalyzer* allows to automatically connect a given metabolite to a chassis model;
2. input/output metabolites are not defined, *sAnalyzer* can find them automatically;
3. *sAnalyzer* searches for reactions from various organisms in the online database;
4. a wide solution space is tested taking into account the characteristics of the chassis organism

After the data input two subprograms are called – chassis organism model processing program and pathways connectivity matrix construction program.

Model processing program checks for KEGG identifiers existence in the model. If the model does not contain KEGG data, they are automatically searched in the online KEGG database and added to the model. Data are read from the KEGG database using API requests. Such information is used to find all transformation pathways between the source metabolite and any metabolite in the model using connectivity matrix created by the other subprogram.

Connectivity matrix construction program creates a source metabolite transformation network that can be represented as a graph tree starting with the source metabolite and defined number of reactions in each pathway. In the end, stoichiometric matrix of constructed network is returned.

This matrix is used to find all possible pathways between two metabolites using modified Depth-first algorithm ("Graph Traversal"). The found pathways are passed to the pathways analysis and ranking subprogram.

Each metabolic pathway is connected to the model using COBRA tool functionality and is tested for flux balance in this pathway and the biomass reaction. For pathways ranking a multiparameter weighted ranking method is used.

Additionally to the simple biotechnological task of substrate degradation or product synthesis, *sAnalyzer* allows solving a combined task of utilising the substrate and producing the product while using chassis organism model.

5. AUTOMATED METABOLIC ENGINEERING SOLUTION SPACE ANALYSIS

This Chapter is devoted to the developed metabolic engineering solution space analysis approach and operations analysis of the developed tool.

5.1. Constraint depending analysis of solution space

Twelve experiments were performed with three different metabolites to identify reachability of other metabolites using reactions from KEGG database. The transformation network of each metabolite was done. For experimenting three metabolites were chosen with different number of reactions where that metabolite is participating. Such number of reactions indicates the initial connectivity of the metabolite. The following source metabolites widely used in biotechnology were chosen: glucose (participating in 111 initial reactions), xylose (participating in 8 initial reactions), and glycerol (participating in 25 initial reactions) (Figure 3).

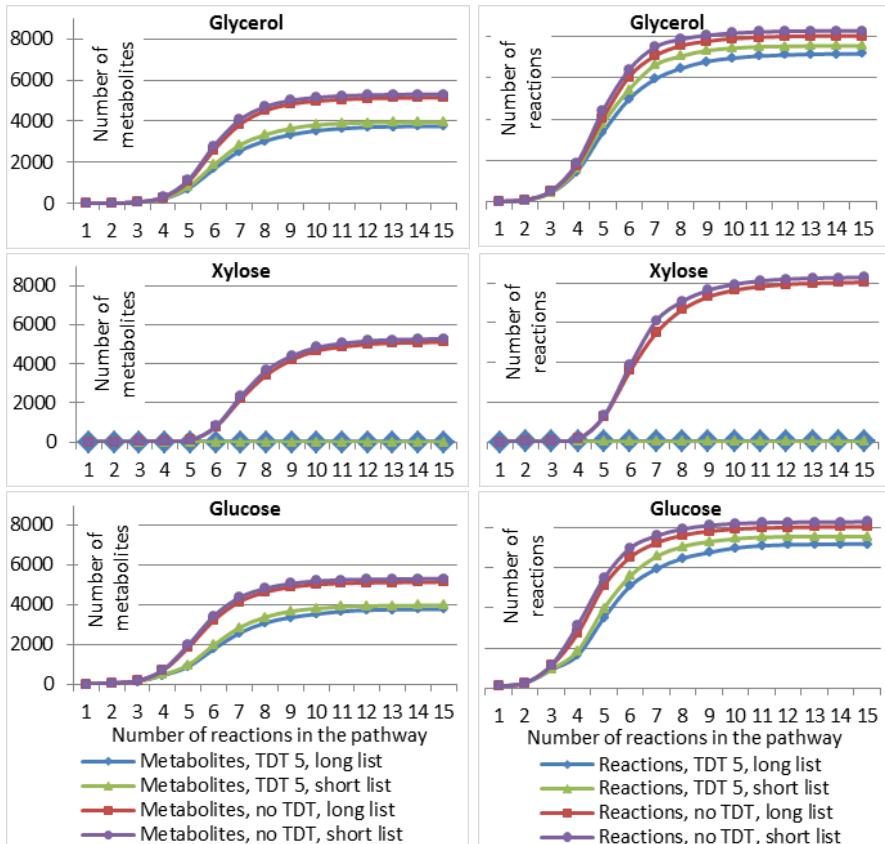


Figure 3. Number of used metabolites (left) and reactions (right) in transformation network construction

Experiments were done using Gibbs energy value thermodynamic threshold (TDT) 5 kJ/mol and without TDT. Also, two different lists of metabolites which are excluded from metabolic pathways construction were used. The long list with 159

metabolites was taken from the published data (Ranganathan, Maranas, 2010). The short list was developed by the author consisting of 18 widely known pool metabolites.

The character of curves corresponds to Richard's curve type (Richards, 1959): at low argument values the increase of growth starts with delay, but at high values of the argument it becomes asymptotically saturated.

Without introducing thermodynamic threshold, the achievable number of metabolites depending on the number of reactions in the pathway shows that with 15 reactions from each metabolite about 90% of other metabolites, involved in reactions, can be reached in the KEGG database.

The total number of included elements in glycerol and glucose transformation network construction is similar at the whole range of number of reactions in the pathway. The differences are insignificant and can be explained by the initial number of reactions in which this metabolite is directly involved. At the same time, the initial connectivity effect is observed – in the xylose case the connectivity level lagging from glycerol and glucose by about two reactions can be seen. This relatively stronger isolation from other metabolites indicates the reasons which lead to problems of xylose as a substrate use in wood biomass recycling process (Jeffries, 2006). But with longer reaction networks the number of total involved elements reaches the same level as glycerol and glucose. That is, all three metabolites can reach the same set of metabolites.

In glycerol and glucose cases thermodynamic threshold (TDT) value 5kJ/mol have similar effect – the number of involved metabolites is smaller by about 1,400 metabolites, and the number of reactions by about 700 reactions with the maximum 15 reactions in the pathway. A significant TDT influence is visible starting from 5 reactions in the pathway. By using TDT in the xylose case, the growth of the number of involved elements stops at two reactions in the pathway. It involves only 8 new reactions at the first reaction in the pathway and 14 reactions at the second reaction in the pathway. That happens because there are no reactions which can thermodynamically happen in xylose degradation direction.

5.2. *Z.mobilis* bacteria metabolic engineering analysis for glycerol utilisation

In this example, *Z.mobilis* central metabolism model (Pentjuss et al., 2013) is used, which consists of 96 reactions and 81 metabolites, the base of this model is *Entner-Dudoroff* pathway for ethanol production.

Twelve experiments were performed to evaluate the use of glycerol as a possible substrate at different Gibbs energy value thermodynamic thresholds (TDT) 30kJ/mol, 5kJ/mol, without TDT, and at different maximum number of reactions in the pathway – 1, 3, 5 or 7 reactions. The long list of excluded metabolites was used. In this example, a total of 12 connection matrixes were created, which represent glycerol transformation networks using different parameters.

In the next step, all pathways were searched in created connection matrixes, which allow connecting glycerol and *Z.mobilis* model (Figure 4.) Y axe is given in a logarithmical scale.

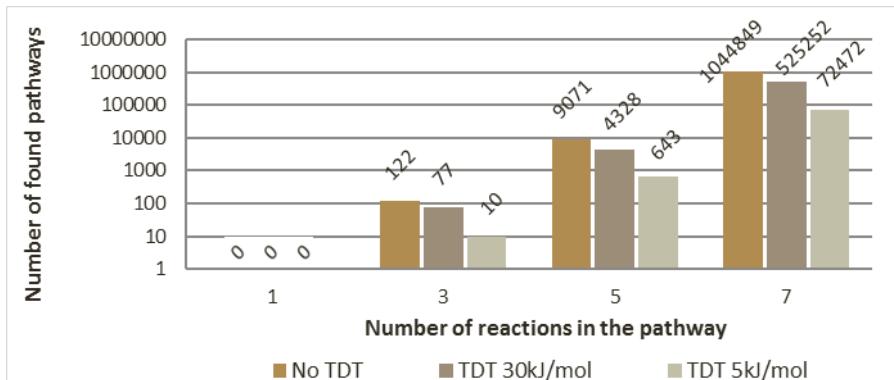


Figure 4. Number of possible biochemical pathways between glycerol and *Z.mobilis* model

In the last step, the found pathways are analysed, where pathways not fulfilling the minimum flux requirements are discarded. Other pathways were ranked based on selected parameters sequence or multiparameter weighted ranking method. During analysis, a half of the found pathways were discarded with 5 reactions in the pathway, but with 7 reactions in the pathway the number of analysed pathways reached the limit of 5,000 pathways.

The minimum number of reactions to connect glycerol to the *Z.mobilis* model is 5 reactions. It can be concluded that the proposed two reactions of *Z.mobilis* modification (Pentjuss et al., 2013) for glycerol recycling might not be operational due to the thermodynamic parameters.

5.3. *E.coli* bacteria metabolic engineering analysis for α -Pinene production

α -Pinene is important natural product which is widely used in different fields. At present α -Pinene, which is used in manufacturing, mostly is derived from trees or in wood recycling process (Yang et al., 2013). Yang research shows an example of implementation of 9 genes into the *E.coli* bacteria for α -Pinene production.

This data was tested using *sAnalyzer* tool. α -Pinene transformation network was constructed with maximum 9 reactions in the pathway. In total 7,046 reactions and 3,136 metabolites were used.

By using the multiparameter weighted ranking method, in total 252 were found and analysed. All found pathways are represented in Figure 5.

With red in this figure is marked the pathway, which is described in the Yang's publication (Yang et al., 2013). The found pathways differ by the number of reactions in the pathway. 8 pathways with 7 reactions and 16 pathways with 8

reactions in the pathway were found. Other pathways have 9 reactions. The pathway mentioned in Yang publication was at the 45th place after ranking with flux on it 4.05 mmol/(gDW^h). In contrast to the 9 reaction pathway proposed in the publication (Yang et al., 2013) , *sAnalyzer* offered a range of pathways with larger *a*-Pinene producing flux using TDT 45kJ/mol.

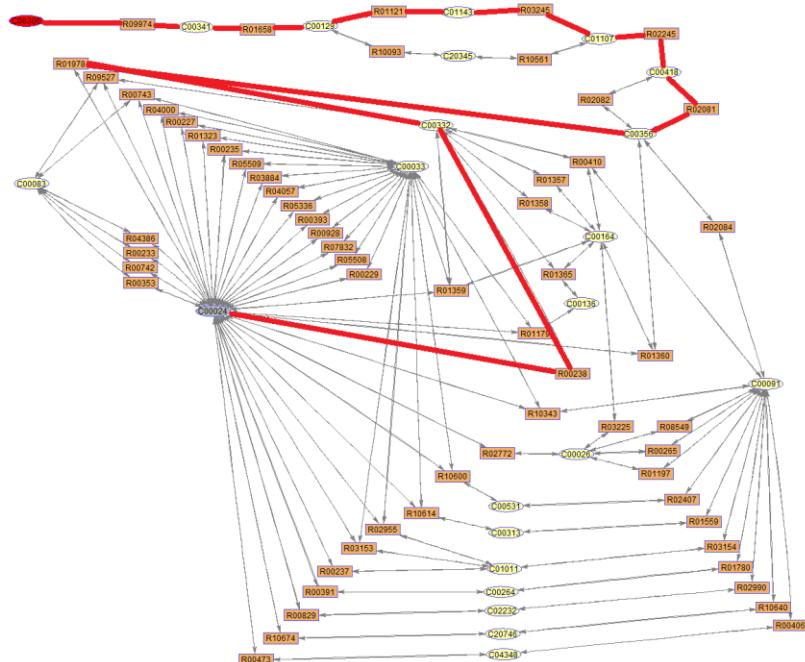


Figure 5. *a*-Pinene producing metabolic pathways visualisation

5.4. *E.coli* bacteria metabolic engineering analysis for 1-butanol production

1-butanol is long chain alcohol, which can be used in several ways. One of them is biofuel production. In comparison with an ethanol, *1-butanol* has many advantages of its use as a fuel, due to higher energy amount and higher hydrophobicity. *1-butanol* can be produced both in chemical way and in fermentation process of microorganisms.

There are several studies on *E.coli* use for *1-butanol* production (Atsumi et al., 2008; Lee et al., 2008; Ranganathan, Maranas, 2010). One of possible *E.coli* modifications is implementing of *C.acetobutylicum* genes of *1-butanol* production (Atsumi et al., 2008). In another research, additionally to existing pathway, non-existing pathways were searched (Ranganathan et al., 2010). A graph-based new pathways search algorithm was used, which uses data from KEGG and BRENDA

databases. BRENDA database was used, because KEGG does not contain data for long chain alcohols transformation.

The found pathways were tested using *sAnalyzer* tool. As a model connection metabolite *pyruvate* (C00022) was chosen – the same as in the publication (Ranganathan et al., 2010). Other parameters of *sAnalyzer* were selected based on the tasks setting the limit of maximum 7 reactions in the pathways, minimum biomass growth reaction flux 0.1% from the wild-type model biomass growth, TDT 35kJ/mol and the long list of excluded elements.

In the result of *sAnalyzer* analysis 23 possible pathways were found that can connect *pyruvate* in the *E.coli* metabolism with the *1-butanol* (Figure 6). All pathways mentioned in the publication were discarded. Possible reasons could be that pathways do not provide the minimum requirements of atom flux in the pathway and biomass reaction or some of reactions cannot happen in *1-butanol* producing direction using the chosen TDT.

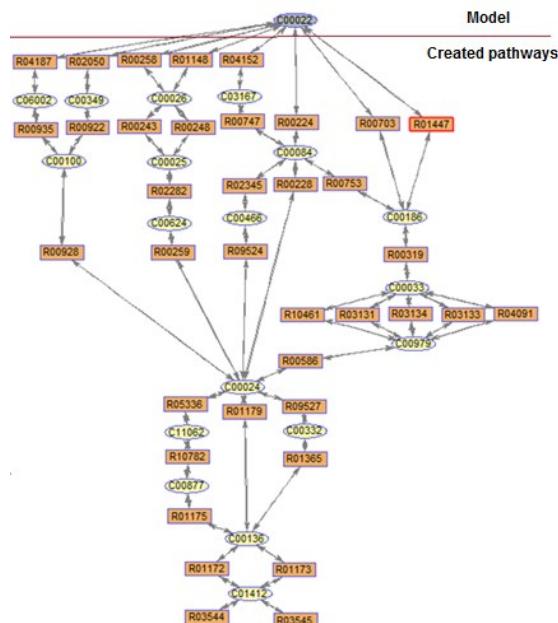


Figure 6. *sAnalyzer* proposed *1-butanol* producing pathways using TDT 35 kJ/mol

The mentioned practical application examples are tested only in modeling level. Before their implementation, biological analysis is needed. However, a wide set of proposed solutions is competitive with the solutions used in publications.

CONCLUSIONS

The main PhD thesis results

Biochemical pathways construction and analysis in chassis organisms was automated for solving metabolic engineering tasks.

1. *The existing metabolic engineering and computational tools for biochemical pathways construction were analysed.*

There are several types of tools for solving metabolic engineering tasks and data analysis: gene system modeling tools, gap filling tools in metabolic networks, flux balance analysis tools, metabolic pathways optimisation tools, dynamic analysis tools, and biological object databases.

Many tools are published that allows searching and constructing biochemical pathways; however, only some of them are available for use. The lack of such tools is the use of out-dated data, because they do not use online biological object databases and the returned result cannot be automatically used in other analysis tools. Existing tools do not allow analysing the whole solution space, but are limited to known metabolic pathways. Not all new possible metabolites transformation pathways are analysed. In addition, the available tools do not allow analysing constructed pathways using chassis organism model.

2. *Criteria and parameters used in automatic construction and analysis of biochemical pathways were analysed.*

Possible criteria and parameters were evaluated, which can be used in automatic biochemical pathways construction and analysis. As most important parameters in pathway analysis were selected: flux balance analysis results of constructed pathway and biomass reaction, biomass growth minimal level, the number of reactions in the pathway, a set of reactions and metabolites excluded from pathway construction, thermodynamic data of reactions, the number of unbalanced reactions in the pathway, a ratio of metabolites in constructed pathway and existing metabolites in chassis model. It is important to have the possibility to exclude any of these criteria during analysis depending on specific analysis task, or to arrange these parameters according to their importance in analysis process.

3. *Automatic biochemical networks construction procedure with the possibility of including existing chassis organism metabolic network model was developed.*

The developed procedure consists of 4 steps: selection of target metabolite and parameters, reaction network construction, search of metabolic pathways, and results analysis and ranking. The procedure allows to scan the whole possible modifications solution space and to analyse relevant metabolic pathways with chassis organism model. Construction of biochemical network is based on the defined target metabolite and initial parameters, using data from online biological object database. For search of metabolic pathways in the constructed network the

graph-based modified Depth-first search algorithm is used. The pathways analysis is performed based on included criteria and parameters. There are different types of analysis available.

4. *Automated biochemical pathways construction and analysis tool was developed.*

On the basis of the developed procedure, open access tool *sAnalyzer* was developed, which can be used to scan the whole possible modifications solution space and to analyse possible metabolic pathways. *sAnalyzer* implements all the steps of the proposed procedure. Some of the steps can be repeated, e.g. pathways analysis can be done with different analysis parameters without repeating previous steps. *sAnalyzer* returns the list of pathways, its criteria ranking scores and chassis organism model with added pathways.

5. *Possible solution space was analysed depending on specified constraints.*

Possible solution space was tested, creating large metabolites transformation biochemical networks.

A Richard curve type increase in the number of metabolites and reactions, used in biochemical network construction, was found depending on the allowable number of reactions in the pathway. Using constraints of maximum number of reactions in the pathway and thermodynamics, it is possible to reduce the overall solution space including only reactions that can occur in one direction at the given process parameters. Using minimum biomass reaction flow constraint, it is possible to control the growth of the organism and discard pathways which do not allow for the growth of organism biomass.

6. *Metabolic engineering task solving procedure is demonstrated.*

Experiments were done to demonstrate the procedure for solving tasks of metabolic engineering using *sAnalyzer* tool.

The ability of *Z.mobilis* bacteria to provide both consumption of glycerol and the growth of the organism at different thermodynamic parameters was tested. *Z.mobilis* case showed that for glycerol conversion to ethanol it is needed to insert 5-9 additional reactions into organism (higher productivity is achieved with 9 reactions) contrary to the published two reactions long pathway which remains experimentally untested. The analysis of *E.coli* bacteria usage options for new substances production was performed, comparing results with the previously published data. In *a-Pinene* case, the pathways were found which can produce *a-Pinene* 40 times greater than solution flow proposed in publication using the same number of inserted reactions. In *l-butanol* case, contrary to the published pathways which cannot produce biomass based on *sAnalyzer* data, some pathways were found which allow producing both biomass and five times bigger amount of *l-butanol*.

sAnalyzer was able to find the published pathways, and it offered pathways which might have greater production potential than proposed pathways. The impact of missing or incomplete information on the results of pathway analysis was found.

Conclusions and development prospects

A number of conclusions were obtained, the most important of them are listed in this section.

1. It is possible to automatically select in the chassis organism insertable pathways, which provide both production and/or consumption of substances and the organism growth.
2. The use of chassis organism models and proposed procedure during modification design allows reducing the required amount of manual work, resources and time due to automatic solution space scanning and pathway analysis and ranking.
3. Biological risk formalisation and inclusion into computational solution space scanning process eliminates risky solutions from more expensive biological experiments.
4. A Richard curve type increase in the number of metabolites and reactions used in analysis and depending on the allowable number of reactions in the pathway was found.
5. The use of reaction thermodynamic threshold significantly reduces the solution space and amount of timespent for analysis.
6. The search of metabolic engineering solutions is appropriate to start from a small number of reactions in the pathway, raising it if a suitable solution was not found.

Several directions can be highlighted as future development prospects.

1. For *sAnalyzer* both *Matlab* tool, which is a product to be paid for, and *COBRA* toolbox is required. It is possible to implement *COBRA* functionality in open source solutions and replace *Matlab* tool with open source tools, e.g. *Octave* or *Scilab*.
2. To develop *sAnalyzer* tool version using *Python* language, since *COBRA* tool can be used in *Python* application. In this case, *Matlab* tool is not needed and it makes *sAnalyzer* fully available free of charge.
3. To improve *sAnalyzer* tool with additional reactions and metabolites data sources (online databases support).
4. To develop possible gene searching and analysis algorithm which can provide possible gene combinations of the selected pathway based on gene-protein-reaction relations. Additionally, enzyme kinetic data can be used to predict reaction likelihood in the chassis organism.

LITERATŪRA

BIBLIOGRAPHY

1. Adrio, J. L., Demain, A. L. (2010) Recombinant organisms for production of industrial products. *Bioengineered Bugs*, 1(2), pp. 116–131.
2. Akutsu, T. (2004) Efficient extraction of mapping rules of atoms from enzymatic reaction data. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, 11(2-3), pp. 449–462.
3. Arita, M. (2000) Metabolic reconstruction using shortest paths. *Simulation Practice and Theory*, 8(1-2), pp. 109–125.
4. Arita, M. (2003) In silico atomic tracing by substrate-product relationships in Escherichia coli intermediary metabolism. *Genome Research*, 13(11), pp. 2455–2466.
5. Atsumi, S., Cann, A. F., Connor, M. R., Shen, C. R., Smith, K. M. et al. (2008) Metabolic engineering of Escherichia coli for 1-butanol production. *Metabolic Engineering*, 10(6), pp. 305–311.
6. Bajpai, P. K., Margaritis, A. (1986) Effect of temperature and pH on immobilized Zymomonas mobilis for continuous production of ethanol. *Biotechnology and Bioengineering*, 28(6), pp. 824–828.
7. Balsa-Canto, E., Banga, J. R. (2011) AMIGO, a toolbox for advanced model identification in systems biology using global optimization. *Bioinformatics (Oxford, England)*, 27(16), pp. 2311–2313.
8. Balsa-Canto, E., Alonso, A.A., Banga, J.R. (2010). An iterative identification procedure for dynamic modeling of biochemical networks. *BMC Systems Biology*, 4(11), p. 18.
9. Barrett, C. L., Kim, T. Y., Kim, H. U., Palsson, B. Ø., Lee, S. Y. (2006) Systems biology as a foundation for genome-scale synthetic biology. *Current Opinion in Biotechnology*, pp. 488–492.
10. Becker, S., Feist, A. M., Mo, M. L., Hannum, G., Palsson, B. Ø. et al. (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nature Protocols*, 2(3), pp. 727–738.
11. Berg, J., M., Tymoczko, J., L., Stryer, L. (2002) Biochemistry, 5th edition. New York: W.H.Freeman. p.1050.
12. Blum, T. (2009) Computational Approaches for Analyzing Metabolic Pathways. PhD thesis. Eberhard-Karls-Universitat Tubingen, p. 148.
13. Blum, T., Kohlbacher, O. (2008) MetaRoute: fast search for relevant metabolic routes for interactive network navigation and visualization. *Bioinformatics (Oxford, England)*, 24(18), pp. 2108–2109.

14. Boele, J., Olivier, B. G., Teusink, B. (2012) FAME, the Flux Analysis and Modeling Environment. *BMC Systems Biology*, 6(1), p. 8.
15. Brohee, S., Faust, K., Lima-Mendez, G., Sand, O., Janky, R. et al. (2008) NeAT: a toolbox for the analysis of biological networks, clusters, classes and pathways. *Nucleic Acids Research*, 36(Web Server issue), pp. 444–451.
16. 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*, 84(6), pp. 647–657.
17. Carothers, J. M., Goler, J., Keasling, J. D. (2009) Chemical synthesis using synthetic biology. *Current Opinion in Biotechnology*, 20(4), pp. 498–503.
18. Caspi, R., Foerster, H., Fulcher, C., Kaipa, P., Krummenacker, M. et al. (2008) The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Research*, 36(Database issue), pp. 623–631.
19. Chandran, D., Bergmann, F. T., Sauro, H. M. (2009) TinkerCell: modular CAD tool for synthetic biology. *Journal of Biological Engineering*, 3, p. 19.
20. Clancy, K., Voigt, C. (2010) Programming cells: towards an automated “Genetic Compiler”. *Current Opinion in Biotechnology*, 21(4), pp. 572–581.
21. Contained use of genetically modified micro-organisms (GMMs) [tiešsaiste] [skatīts 2014. g. 13. dec.]. Pieejams: http://europa.eu/legislation_summaries/agriculture/food/sa0015_en.htm.
22. Corby-Harris, V., Drexler, A., Watkins de Jong, L., Antonova, Y., Pakpour, N. et al. (2010) Activation of Akt signaling reduces the prevalence and intensity of malaria parasite infection and lifespan in Anopheles stephensi mosquitoes. *PLoS Pathogens*, 6(7), p. 10.
23. Cordwell, S. J. (1999) Microbial genomes and “missing” enzymes: redefining biochemical pathways. *Archives of Microbiology*, 172(5), pp. 269–279.
24. Croes, D., Couche, F., Wodak, S. J., van Helden, J. (2006) Inferring meaningful pathways in weighted metabolic networks. *Journal of Molecular Biology*, 356(1), pp. 222–36.
25. Czar, M. J., Cai, Y., Peccoud, J. (2009) Writing DNA with GenoCAD. *Nucleic Acids Research*, 37(Web Server issue), pp. 40–47.
26. Densmore, D., Devender, A. Van. (2009) A platform-based design environment for synthetic biological systems. *The Fifth Richard Tapia*, pp. 24–29.
27. Endy, D. (2005) Foundations for engineering biology. *Nature*, 438(7067), pp. 449–453.
28. Feist, A.M., Henry, C.S., Reed, J.L., Krummenacker, M., Joyce, A.R., 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*, 3(121), p. 18.
- 29. eQuilibrator - biochemical thermodynamics calculator [tiešsaiste] [skatīts 2014. g. 14. aug.]. Pieejams: <http://equilibrator.weizmann.ac.il/>
 - 30. Fernández-Suárez, X. M., Galperin, M. Y. (2013) The 2013 Nucleic Acids Research Database Issue and the online molecular biology database collection. *Nucleic Acids Research*, 41(Database issue), pp. 1–7.
 - 31. Flamholz, A., Noor, E., Bar-Even, A., Milo, R. (2012) eQuilibrator - The biochemical thermodynamics calculator. *Nucleic Acids Research*, 40, pp. 770–775.
 - 32. Fleming, R.M.T., Thiele, I., Provan, G., Nasheuer, H.P. (2010). Integrated stoichiometric, thermodynamic and kinetic modelling of steady state metabolism. *Journal of Theoretical Biology*, 264(3), pp. 683–692.
 - 33. Forth, T. (2012) Metabolic Systems Biology of the Malaria Parasite. PhD thesis, University of Leeds. p. 240.
 - 34. Fotadar, U., Zaveloff, P., Terracio, L. (2005) Growth of Escherichia coli at elevated temperatures. *Journal of Basic Microbiology*, 45(5), pp. 403–404.
 - 35. Freilich, S., Zarecki, R., Eilam, O., Segal, E.S., Henry, C.S., et al. (2011). Competitive and cooperative metabolic interactions in bacterial communities. *Nature Communications*, 2(589).
 - 36. Fu, P., Panke, S. (2009) Systems biology and synthetic biology. Hoboken, NJ, USA: John Wiley & Sons. p. 657.
 - 37. Funahashi, A., Morohashi, M., Kitano, H., Tanimura, N. (2003) CellDesigner : a process diagram editor for gene-regulatory and. *Biosilico*, 1(5), pp. 159–162.
 - 38. Gendrault, Y., Madec, M., Lallement, C., Pecheux, F., Haiech, J. (2011) Synthetic biology methodology and model refinement based on microelectronic modeling tools and languages. *Biotechnology Journal*, 6(7), pp. 796–806.
 - 39. Gevorgyan, A., Poolman, M.G., Fell, D.A. (2008). Detection of stoichiometric inconsistencies in biomolecular models. *Bioinformatics*, 24(19), pp. 2245–2251.
 - 40. Genetically modified animals. [tiešsaiste] [skatīts 2014. g. 10. dec.]. Pieejams: <http://www.efsa.europa.eu/en/topics/topic/gmanimals.htm>.
 - 41. GMO Biography. [tiešsaiste] [skatīts 2014. g. 20. nov.]. Pieejams: Retrieved from <http://www.artistopia.com/gmo/biography>.
 - 42. Graph Traversal. [tiešsaiste] [skatīts 2014. g. 10. dec.]. Pieejams: <http://web.cse.ohio-state.edu/~gurari/course/cis680/cis680Ch14.html>.
 - 43. Hallinan, J. S., Misirli, G., Wipat, A. (2010) Evolutionary computation for the design of a stochastic switch for synthetic genetic circuits. In: Proceedings of

- the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 768–774.
- 44. Hastings, J., de Matos, P., Dekker, A., Ennis, M., Harsha, B. et al. (2013) The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic Acids Research*, 41(Database issue), pp. 456–463.
 - 45. Hatzimanikatis, V., Li, C., Ionita, J., Henry, C. S., Jankowski, M. D. et al. (2005) Exploring the diversity of complex metabolic networks. *Bioinformatics (Oxford, England)*, 21(8), pp. 1603–1609.
 - 46. Henry, C. S., DeJongh, M., Best, A., Frybarger, P. M., Lindsay, B. et al. (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotechnology*, 28(9), pp. 977–982.
 - 47. Herbert W. Boyer. [tiešsaiste] [skatīts 2014. g. 10. dec.]. Pieejams: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3160001/>.
 - 48. Herrgård, M. J., Fong, S. S., Palsson, B. Ø. (2006) Identification of genome-scale metabolic network models using experimentally measured flux profiles. *PLoS Computational Biology*, 2(7), pp. 676–686.
 - 49. Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J. et al. (2006) COPASI-a COmplex PAthway SImlator. *Bioinformatics (Oxford, England)*, 22(24), pp. 3067–3074.
 - 50. 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*, 12(28), p. 7.
 - 51. Jeffries, T.W. (2006) Engineering yeasts for xylose metabolism. *Current Opinion in Biotechnology*, 17, pp. 320–326.
 - 52. Jing, L.S., Shah, F.F.M., Mohamad, M.S., Hamran, N.L., Salleh, A.H.M., et al. (2014). Database and tools for metabolic network analysis. *Biotechnology and Bioprocess Engineering*, 19(4), pp. 568–585.
 - 53. Jung, S.-K., McDonald, K. (2011) Visual gene developer: a fully programmable bioinformatics software for synthetic gene optimization. *BMC Bioinformatics*, 12(340), p. 13.
 - 54. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M. (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, 40(Database issue), pp. 109–114.
 - 55. Karp, P. D., Paley, S., Romero, P. (2002) The Pathway Tools software. *Bioinformatics (Oxford, England)*, 18, pp. 225–232.
 - 56. Khalil, A. S., Collins, J. J. (2010) Synthetic Biology : Applications Come of Age, *Nature*, 11(5), pp. 367–379.
 - 57. Kharchenko, P., Chen, L., Freund, Y., Vitkup, D., Church, G. M. (2006) Identifying metabolic enzymes with multiple types of association evidence. *BMC Bioinformatics*, 7(177), p. 16.

58. Kim, H. U., Kim, T. Y., Lee, S. Y. (2008) Metabolic flux analysis and metabolic engineering of microorganisms. *Molecular bioSystems*, 4(2), pp. 113–120.
59. Kim, J., Reed, J. L. (2010) OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Systems Biology*, 4(53), p. 19.
60. Kitney, R., Freemont, P. (2012) Synthetic biology - the state of play. *FEBS Letters*, 586(15), pp. 2029–2036.
61. Klamt, S., Stelling, J. (2002) Combinatorial complexity of pathway analysis in metabolic networks. *Molecular Biology Reports*, 29(1-2), pp.233–236.
62. Klamt, S., Stelling, J., Ginkel, M., Gilles, E. D. (2003) FluxAnalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps. *Bioinformatics*, 19(2), pp. 261-269.
63. Kostromins, A. (2012). Altfluxes: COBRA Toolbox extension for flux variability analysis of stoichiometric models of metabolism. In: Proceedings of the 5th International Scientific Conference on Applied Information and Communication Technologies. April 26-27, 2012, Jelgava, Latvia, pp. 294–299.
64. Kouvelis, V. N., Saunders, E., Brettin, T. S., Bruce, D., Detter, C. et al. (2009) Complete genome sequence of the ethanol producer Zymomonas mobilis NCIMB 11163. *Journal of Bacteriology*, 191(22), pp. 7140–7141.
65. Kumar, V. S., Maranas, C. D. (2009) GrowMatch: an automated method for reconciling in silico/in vivo growth predictions. *PLoS Computational Biology*, 5(3), p. 13.
66. Kuprijanov, A., Schaepe, S., Simutis, R., Lübbert, A. (2013). Model predictive control made accessible to professional automation systems in fermentation technology. *Biosystems and Information Technology*, 2(2), pp. 26–31.
67. Lee, K. Y., Park, J. M., Kim, T. Y., Yun, H., Lee, S. Y. (2010) The genome-scale metabolic network analysis of Zymomonas mobilis ZM4 explains physiological features and suggests ethanol and succinic acid production strategies. *Microbial Cell Factories*, 9(94), p. 12.
68. Lee, S. K., Chou, H., Ham, T. S., Lee, T. S., Keasling, J. D. (2008) Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Current Opinion in Biotechnology*, 19(6), pp. 556–563.
69. Lehrach, H., Subrak, R., Boyle, P., Pasterk, M., Zatloukal, K., et al. (2011). ITfOM – The IT Future of Medicine. *Procedia Computer Science*, 7, pp. 26–29.
70. Li, S.-D., Huang, L. (2006) Gene therapy progress and prospects: non-viral gene therapy by systemic delivery. *Gene Therapy*, 13(18), pp. 1313–1319.

71. Macdonald, J. T., Barnes, C., Kitney, R. I., Freemont, S., Stan, G. V. (2011) Integrative Biology Computational design approaches and tools for synthetic biology. *Society*, 3, pp. 97–108.
72. Mannelli, I., Minunni, M., Tombelli, S., Mascini, M. (2003) Quartz crystal microbalance (QCM) affinity biosensor for genetically modified organisms (GMOs) detection. *Biosensors & Bioelectronics*, 18(2-3), pp. 129–140.
73. Matlab Overview [tiešsaiste] [skatīts 2015. g. 10. jūl.]. Pieejams: <http://se.mathworks.com/products/matlab/>.
74. McShan, D. C., Rao, S., Shah, I. (2003) PathMiner: predicting metabolic pathways by heuristic search. *Bioinformatics*, 19(13), pp. 1692–1698.
75. Meitalovs, J. (2011) Fermentācijas procesa hemostata nodrošināšanas sistēmas izstrāde. Magistra darbs. Jelgava: LLU. 64 lpp.
76. Meitalovs, J. (2012a) Process of genetic modification of microorganisms from the point of view of the software engineering. In: Proceedings of the 5th International Scientific Conference on Applied Information and Communication Technologies. April 26-27, 2012, Jelgava, Latvia, pp. 348–351.
77. Meitalovs, J. (2012b) Software tool for probabilistic metabolic pathways construction. In: Proceedings of the 2012 IEEE 13th International Symposium on Computational Intelligence and Informatics (CINTI). November 20-22, 2012, Budapest, Hungary, pp. 405–408.
78. Meitalovs, J. (2012c) The concept of metabolic engineering software tool. In: Proceedings of the 5th International Scientific Conference on Applied Information and Communication Technologies. April 26-27, 2012, Jelgava, Latvia, pp. 352–355.
79. Meitalovs, J., Bulipopa, N., Kovalonoka, O. (2012) Modeling and design tools for synthetic biology. In: Proceedings of the 5th International Scientific Conference on Applied Information and Communication Technologies. April 26-27, 2012, Jelgava, Latvia, pp. 341–347
80. Meitalovs, J., Stalidzans, E. (2013a) Analysis of synthetic metabolic pathways solution space. In: Proceedings of the 2013 International Conference on System Science and Engineering (ICSSE). July 4-6, 2013, Budapest, Hungary, pp. 183–187.
81. Meitalovs, J., Stalidzans, E. (2013b) Connectivity analysis of metabolites in synthetic metabolic pathways. In: Proceedings of the 12th International Scientific Conference “Engineering for Rural Development”. May 23-24, 2013, Jelgava, Latvia, pp. 435–440.
82. Meitalovs, J., Stalidzans, E. (2015) Impact of thermodynamic constraint to the solution space of metabolic pathway design using *sAnalyzer* tool. *Baltic Journal of Modern Computing*, 3, pp. 164-178.

83. Mendes, P., Kell, D. (1998). Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics*, 14, pp. 869–883.
84. Mirschel, S., Steinmetz, K., Rempel, M., Ginkel, M., Gilles, E. D. (2009) PROMOT: modular modeling for systems biology. *Bioinformatics*, 25(5), pp. 687–689.
85. Mozga, I. (2012). Biokīmisko tīklu stacionāro stāvokļu optimizācijas procedura. Promocijas darbs. Jelgava:LLU. 158 lpp.
86. Nielsen, J. (1998). Metabolic engineering: techniques for analysis of targets for genetic manipulations. *Biotechnology and Bioengineering*, 58(2-3), pp. 125–132.
87. Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H. et al. (1999) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 27(1), pp. 29–34.
88. Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H. et al. (2011) A comprehensive genome-scale reconstruction of Escherichia coli metabolism. *Molecular Systems Biology*, 7(535), p. 9.
89. Orth, J. D., Palsson, B. Ø. (2010) Systematizing the generation of missing metabolic knowledge. *Biotechnology and Bioengineering*, 107(3), pp. 403–412.
90. Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H.-Y., et al. (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Research*, 33(17), pp. 5691–5702.
91. Park, J. H., Lee, S. Y., Kim, T. Y., Kim, H. U. (2008) Application of systems biology for bioprocess development. *Trends in Biotechnology*, 26(8), pp. 404–412.
92. Pentjuss, A., Odzina, I., Kostromins, A., Fell, D. A., Stalidzans, E. et al. (2013) Biotechnological potential of respiring Zymomonas mobilis: a stoichiometric analysis of its central metabolism. *Journal of Biotechnology*, 165(1), pp. 1–10.
93. Pharkya, P., Burgard, A. P., Maranas, C. D. (2004) OptStrain: a computational framework for redesign of microbial production systems. *Genome Research*, 14(11), pp. 2367–2376.
94. Rahman, S. A., Advani, P., Schunk, R., Schrader, R., Schomburg, D. (2005) Metabolic pathway analysis web service (Pathway Hunter Tool at CUBIC). *Bioinformatics (Oxford, England)*, 21(7), pp. 1189–1193.
95. Ranganathan, S., Maranas, C. D. (2010) Microbial 1-butanol production: Identification of non-native production routes and in silico engineering interventions. *Biotechnology Journal*, 5(7), pp. 716–725.

96. Ranganathan, S., Suthers, P. F., Maranas, C. D. (2010) OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. *PLoS Computational Biology*, 6(4), p.11.
97. Reed, J. L., Patel, T. R., Chen, K. H., Joyce, A. R., Applebee, M. K. et al. (2006) Systems approach to refining genome annotation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(46), pp. 17480–17484.
98. Richards, F.J. (1959) A flexible growth function for empirical use. *Journal of Experimental Botany*, 10, pp. 290–301.
99. Richardson, S. M., Nunley, P. W., Yarrington, R. M., Boeke, J. D., Bader, J. S. (2010) GeneDesign 3.0 is an updated synthetic biology toolkit. *Nucleic Acids Research*, 38(8), pp. 2603–2606.
100. Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L. et al. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 440(7086), pp. 940–943.
101. Rocha, I., Maia, P., Evangelista, P., Vilaça, P., Soares, S. et al. (2010) OptFlux: an open-source software platform for in silico metabolic engineering. *BMC Systems Biology*, 4(45), p. 12.
102. Rodrigo, G., Carrera, J., Jaramillo, A. (2011) Computational design of synthetic regulatory networks from a genetic library to characterize the designability of dynamical behaviors. *Nucleic Acids Research*, 39(20), p. 12.
103. Rodrigo, G., Carrera, J., Prather, K. J., Jaramillo, A. (2008) DESHARKY: Automatic design of metabolic pathways for optimal cell growth. *Bioinformatics*, 24(21), pp. 2554–2556.
104. Rubina, T. (2013). The procedure of evolution modelling of biochemical networks structure. *Biosystems and Information Technology*, 2(2), pp. 19–25.
105. Salvadó, Z., Arroyo-López, F. N., Guillamón, J. M., Salazar, G., Querol, A. et al. (2011) Temperature adaptation markedly determines evolution within the genus *Saccharomyces*. *Applied and Environmental Microbiology*, 77(7), pp. 2292–2302.
106. 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*, 203(3), pp. 229–248.
107. Schomburg, I., Chang, A., Schomburg, D. (2002) BRENDA, enzyme data and metabolic information. *Nucleic Acids Research*, 30(1), pp. 47–49.
108. 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*, 17(2), pp. 53–60.
109. Seo, J.-S., Chong, H., Park, H. S., Yoon, K.-O., Jung, C. et al. (2005) The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. *Nature Biotechnology*, 23(1), pp. 63–68.

110. Serrano, L. (2007) Synthetic biology: promises and challenges. *Molecular Systems Biology*, 3(158), p. 5.
111. Shetty, R. P., Endy, D., Knight, T. F. (2008) Engineering BioBrick vectors from BioBrick parts. *Journal of Biological Engineering*, 2(5), p. 12.
112. Soh, K. C., Hatzimanikatis, V. (2010) DREAMS of metabolism. *Trends in Biotechnology*, 28(10), pp. 501–508.
113. Spiller, D. G., Wood, C. D., Rand, D. A., White, M. R. H. (2010) Measurement of single-cell dynamics. *Nature*, 465(7299), pp. 736–745.
114. Symbiopedia. [tiešsaiste] [skatīts 2015. g. 30. mar.]. Pieejams: http://benheavner.com/systemsbio/index.php?title=File:Sc_iND750_GlcMM.xml
115. Tepper, N., Shlomi, T. (2010) Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways. *Bioinformatics (Oxford, England)*, 26(4), pp. 536–543.
116. The Nobel Prize in Chemistry 2008. [tiešsaiste] [skatīts 2014. g. 10. dec.]. Pieejams: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2008/press.html
117. Thiele, I., Palsson, B. Ø. (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, 5, pp. 93–121.
118. Tomlin, C.J., Axelrod, J.D. (2005). Understanding biology by reverse engineering the control. *Proceedings of the National Academy of Sciences*, 102, pp. 4219–4220.
119. Trinh, C. T., Srienc, F. (2009) Metabolic engineering of Escherichia coli for efficient conversion of glycerol to ethanol. *Applied and Environmental Microbiology*, 75(21), pp. 6696–6705.
120. Tsantili, I. C., Karim, M. N., Klapa, M. I. (2007) Quantifying the metabolic capabilities of engineered Zymomonas mobilis using linear programming analysis. *Microbial Cell Factories*, 6(8), p. 23.
121. Tyo, K.E., Alper, H.S., Stephanopoulos, G. (2007). Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends in Biotechnology*, 25, pp. 132–137.
122. Villalobos, A., Ness, J. E., Gustafsson, C., Minshull, J., Govindarajan, S. (2006) Gene Designer: a synthetic biology tool for constructing artificial DNA segments. *BMC Bioinformatics*, 7(285), p. 8.
123. Vīgants, A. (2008). Cilvēka bioķīmija un molekulārā ķīmija. Rīga: LU. 123 lpp.
124. Vital-Lopez, F.G., Armaou, A., Nikolaev, E. V, Maranas, C.D. (2006). A computational procedure for optimal engineering interventions using kinetic models of metabolism. *Biotechnology Progress*, 22(6), pp. 1507–1517.

125. Wang, L., Hatzimanikatis, V. (2006) Metabolic engineering under uncertainty. I: framework development. *Metabolic Engineering*, 8(2), pp. 133–141.
126. What is systems biology? [tiešsaiste] [skatīts 2014. g. 10. dec.]. Pieejams: http://blogs.nature.com/sevenstones/2007/07/what_is_systems_biology3.html.
127. What's in a name? (2009) *Nature Biotechnology*, 27(12), pp. 1071–1073.
128. Widiastuti, H., Kim, J. Y., Selvarasu, S., Karimi, I., Kim, H. et al. (2011) Genome-scale modeling and in silico analysis of ethanologenic bacteria *Zymomonas mobilis*. *Biotechnology and Bioengineering*, 108(3), pp. 655–665.
129. Wiechert, W. (2002). Modeling and simulation: tools for metabolic engineering. *Journal of Biotechnology*, 94(1), pp. 37–63.
130. Yang, J., Nie, Q., Ren, M., Feng, H., Jiang, X. et al. (2013) Metabolic engineering of *Escherichia coli* for the biosynthesis of alpha-pinene. *Biotechnology for Biofuels*, 6(60), p. 10.
131. Young, E., Alper, H. (2010) Synthetic biology: tools to design, build, and optimize cellular processes. *Journal of Biomedicine & Biotechnology*, 2010, p. 12.
132. Yu, K. O., Kim, S. W., Han, S. O. (2010) Engineering of glycerol utilization pathway for ethanol production by *Saccharomyces cerevisiae*. *Bioresource Technology*, 101(11), pp. 4157–4161.
133. Zazzu, V., Regierer, B., Kühn, A., Sudbrak, R., Lehrach, H. (2013). IT Future of Medicine: from molecular analysis to clinical diagnosis and improved treatment. *New Biotechnology*, 30, pp. 362–365.