MULTI AGENT GENOME-SCALE METABOLIC RECONSTRUCTION MODELING SOFTWARE SCHEMA

Agris Pentjuss, Ilona Odzina, Aleksandrs Gailums
Latvia University of Agriculture
agris.pentjuss@gmail.com

Abstract. Systems biology is a new science, where the research field as constraint based genome – scale reconstruction modeling and analysis past 10 years has been rapidly advanced and created new biological and computational advance for computer modeling the metabolism of the living organisms. The example could be bacteria *Z. Mobilis* for bioethanol production, cyanobacteria *Synechococcus Elongatus* for biobutanol production, cyanobacteria *Cyanothece 51142* for biohydrogen production. This research field is used to find solutions for improving the organism metabolism to produce more fuel substance than in a conventional way. Constraint based genome – scale reconstruction modeling and analysis methods limiting feature is that the modeling process is simulated only within a fixed pH level and temperature, where the cells do not take into account the changing thermodynamic properties contrary to the nature of ongoing. In the paper an insight is described into the multi agents standards-based genome-scale reconstruction modeling and analyze software scheme, where each bacterium is viewed as a separate agent with own properties in dynamic environment and analyzed as a multi agent system as whole. As the model reconstruction analysis tool Cobra Toolbox 2.0 is used and for thermodynamic properties analyze von Bertalanffy extension is used, which provides reactions direction adjustment. The authors' research result is the schema, which describes agents collaboration, deliberation, goal based behavior, communication and planning processes, interaction interface with user and multi agent genome – scale reconstruction modeling software.

Keywords: genome-scale analysis, multi agent systems, thermodynamic properties analysis.

Introduction

The genome scale reconstruction (GSR) approach is to make GSR analyses in one organism and at specific conditions [1]. Some starting researches in bioprocess control algorithm are made [2], which show in theory that there should be possibilities to combine GSR models with other impacting parameters in fermentation environment to model results. In nature environment could impact differently the environment and biological processes in real life than the GSR analyses shows. In precision agriculture the fermentation opportunities of bacteria for biogass, biohydrogen and bioethanol and even biobutanol for real life conditions have not even been modeled because of expensive experiments. The main goal of the paper is to create the Multi Agent Genome-Scale reconstruction Modeling Software (MAGSMS) scheme, which describes how to divide environment in environmental agents, bacterial agents, how the pH level, temperature, osmotic and hydrostatic pressure, food and water existence influence on whole modeling environment should be calculated. GSR analysis tool from COBRA toolbox will be used in the MAGSMS scheme [3]. For thermodynamically feasible GSR Von Bertalanffy toolbox extension should be used [4], MAGSMS modeling environment should be made by object oriented language Python [5].

Materials and methods

Microorganisms that are specialized to convert sugars into ethanol in the most efficient way are expected to be cells with minimal functionality, equipped only with the specialized catalytic capability for the formation of the desired product and for the replication and renewal of this catalytic function. The creation of a minimal cell that is able to self-assemble and to self-replicate based on a limited number of genes is an intriguing goal in biology. Several approaches have been proposed to realize this goal, using comparative genomic, genetic, or biochemical tools [6 – 10].

Growth medium

The growth medium for microorganisms could be anything starting with biogass production [11] different types of digester, Ethanol production from wide spectrum of sugars or even sugar mix [12], even making biobutanol from carbon dioxide for cyanobacteria Synechococcus Elongatus [13] and also biohydrogen production necessary growth medium for cyanobacteria Cyanothece 51142 [14]. Each organism needs its own growth medium. Making MAGMMS there is a need to include medium definition parameters with unlimited amount of substance for microorganism growth. Also there

should be an opportunity to add some ions, chemical compounds and even different microorganisms with different growth medium requirements. The growth medium should not include physical things like digester, because it is too difficult to split into basic chemical elements, which means the MAGMMS user should choose the chemical compound from the software offered ones rather than add by himself. This means that the user will need some basic knowledge in constraint based modeling approach [15].

Growing environment

In the growing environment four basic components (food, moisture, warmth and time) should be calculated which taken together cause bacterial growth and should be included in modeling MAGSMS [16, 17]. As food or necessary nutrients (as nitrogen, sulfur, phosphorus and others) for each type of bacteria is different, for example, to produce bioethanol [18] or even biobutanol [13], we need sugars which can consume – a carbon source to transform sugar to ethanol. Some bacteria to produce biogas use digester as food. Some nutrients catalases bacteria growth some do not do any impact on the growth rate, but it all should be added as the parameters in MAGSMS.

Bacteria cannot grow without water. Bacteria need some degree of moister to start consume food. Each bacteria has its own moister degree level to start consume food. Many bacteria are quickly killed buy dry conditions, but some can survive some time without water and later added these bacteria could start consume it. MAGSMS will need to have an ability get the data about water activity in the fermentation modeling process [16].

Difference of temperature is one of the main factors which influence bacterial growth. Theoretically, bacteria have an ability to grow at all temperatures between the freezing point of water and the temperature at which protein coagulates. If temperature reaches below the minimal growth temperature, bacteria stop the growth, but if temperature rises above the maximal growth temperature bacteria should soon be killed [16]. Temperature dependence should be added in MAGSMS as mathematical equation.

Time is one of the constraints in bacteria growth. At ideal conditions bacteria will grow and multiply by dividing into two pieces every 20 minutes. After 6 hours in ideal conditions one bacterium should divide in 131 072 bacteria and this method should be included in MAGSMS as the changing value by environment conditions [16].

There are also some factors what affect the growth rate of bacteria – Ph level, by oxygen dependency for growth and competition conditions between different numbers of bacteria in environment for food [16]. These competition conditions should be described between different bacteria as dependency between agents interaction in MAGSMS.

Ph level in environment as the acidity factor should be calculated also and in different places in the environment independently, because moisture or water Ph level cannot be equal everywhere in the environment. Bacteria sensitivity to ph acidity level in environment is described in two terms – ph maximum and ph optimum [16].

Osmotic pressure is known as diffusion of water across cell membranes in response to solute concentrations, if the osmotic pressure is higher in bacteria than in environment then it will prevent bacteria consumption of food through the membrane. This parameter should be gained from the user of MAGSMS to calculate the diffusion possibility of bacteria food (substrate) [16].

Hydrostatic pressure [19] is a physical parameter which changes in ion fluxes across the membrane, which drives the influx or efflux of water and leads to cell volume recovery. The changes of these parameters should be included as mathematical equations in MAGSMS as additional conditions for genome – scale reconstruction model reaction directionality [16].

Dynamically changing environment

For modeling it is very important to define the environment parameters and changing speed of environment. As in precision agriculture there are already defined agent based software environment properties [20] then MAGSMS modeling environment is dynamical, stochastic, discrete and accessible environment. Methodology of modeling should include the ability to calculate each bacterium parameter, GSR fluxes of all even different types of bacteria. The multi agent system approach [21]

could make the most necessary calculations of MAGSMS, include in modeling constraints of the growth medium parameters and count them for each bacterium in MAGSMS environment. The agent type could be a practical reasoning agent or deductive reasoning agent with their abilities. At simplifier MAGSMS modeling with several parameters which make modeling more theoretical then realistic to the real life dynamic environment it is better to use deductive reasoning agents, but in case if MAGSMS modeling environment parameters will be more closer to the real dynamical changing environment, then it would be better to use practical reasoning agents, because the deliberation process of the deductive reasoning agent can take much more calculation time in this case and it means – end reasoning should be faster in calculations on deliberation, planning processes.

Analytical techniques

The main goal of MAGSMS is to model the necessary product creation from some food for bacteria taking into account also the physical constraints of environment and bacteria count. At the first MAGSMS version the temperature, pH, pressure and other parameters will change linearly, it means that there will not be helping tools which could improve the speed of these parameter changings in the appropriate time. The genome – scale reconstruction reactions directionality will change according to Von Bertalanffy 1.0 toolbox extension methods [4] using Gibbs energy, metabolite concentrations, algorithm bounds.

Yield and flux calculation and carbon balance

GSRs calculate the flux of objective function, for example, Biomass reaction flux, product, and substrate. This flux shows the ability to produce products from substrate. These calculations are made for one bacterium and to calculate the product yield we need to calculate each bacterium flux taking into account all constraints of environment and growth media, calculate each bacterium product flux under different MAGSMS environment conditions and get average flux value from all bacteria if it is possible.

Also we need to calculate bacteria biomass reaction flux which means increase of the bacteria biomass yield which will impact the substrate consumption rate. This should be calculated for each bacterium separately and converteded as time in which bacterium will divide. Calculation of full dividing time for each bacterium will depend of flux equation of bacterium multiply with the time coefficient.

Calculation programming languages

For GSR calculations Matlab [22] programming language or Python programming language in Windows operating system should be used. GSR analyses tool COBRA toolbox (available at http://opencobra.sourceforge.net) is developed in Matlab environment and has access to already made scripts. These scripts are rewritten in programming language Python code, but directionality calculation COBRA toolbox extension Von Bertelanffy 1.0 has not yet been transformed to Python programming language, that means Genome – scale reconstruction analysis should be done using Matlab based COBRA toolbox scripts. Also there is a need to make agent based programmed environment for the modeling approach. For that we would suggest to use Python programming language as it supports the agent based object oriented programming and also has the ability in future to be used for genome – scale reconstruction analyses.

Results and discussion

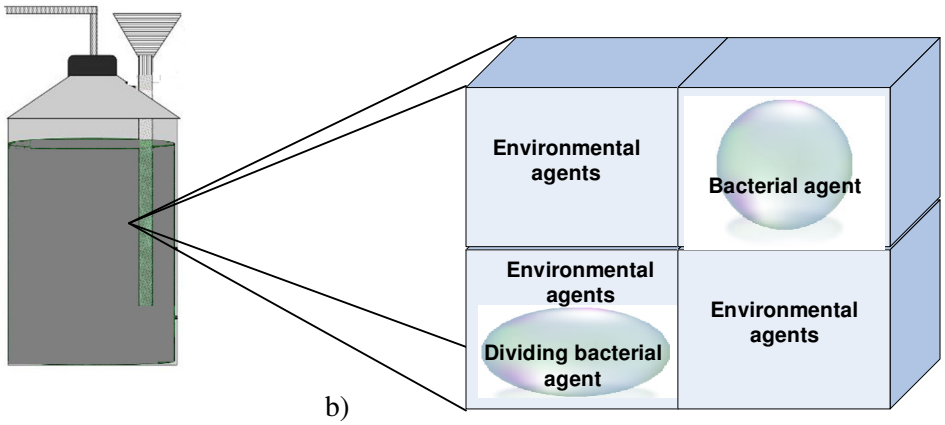
As a result of this paper the authors purpose MAGSMS schema as software for the systems biology approach modeling using genome – scale reconstruction analysis and combining this approach with the Python agent object – oriented programming language approach for growth media and growth environment constraint modeling. MAGSMS architecture should consist of 2 general types of agents:

Environmental agent – environment is divided in many cubes and one cube should be able to carry out two bacteria in environment Fig. 1 in b section. They are responsible for the growth medium and environment medium properties like food, moisture or water, warmth, pH level, oxygen existence, temperature, osmotic pressure, hydrostatic pressure. This kind of agent should have the following functions.

- Mobility – these agents are mobile. Their properties are calculated using the GSR calculation results of substrate consumption flux, product flux, biomass flux.
- Collaboration – bacterial agent collaborates in two methods:
  - environmental medium collaboration – they should calculate global difference in temperature, pH, hydrostatic and osmatic pressure and also count of bacterial agents in this environment.
  - bacterial collaboration – gives information about environmental agent local temperature, pH, hydrostatic and osmatic pressure to bacterial agent for the recalculation process of new genome – scale reconstruction constraints in each bacteria.
- Agent properties – it includes temperature, pH, hydrostatic and osmatic pressure value of environment, stores substrate and product yield. All agents should have the same dimensional size.
- Communication information – communicating with environmental agents they should transfer the recalculated global information about temperature, pH, hydrostatic and osmatic pressure value. Communicating with bacterial agents they should transfer local temperature, pH, hydrostatic and osmatic pressure value for genome – scale reconstruction constraint recalculations.

Bacterial agent – bacterial agent could be much in count and type where changing information could be GSRs between the bacterial type, one type bacteria genome – scale reconstruction constraints, dividing time calculation constraints, thermodynamically feasible GSR model calculations and results in reaction directionality. This kind of agent should have the following functions (Fig. 1).

- Mobility – these agents are not mobile they are stationary. Their properties are calculated using growth and environment medium to model bacteria in MAGSMS closer to real life conditions.
- Collaboration – bacterial agent collaboration is divided into two methods:
  - environmental medium collaboration – bacterial agent sends information about calculated genome – scale reconstruction reaction fluxes, it receives information about environmental agent calculated local temperature, pH, hydrostatic and osmatic pressure for thermodynamically balanced genome – scale reconstruction. Information is sent about the produced product and consumed substrate yield.
  - bacterial collaboration – sends information about movement direction to nearest bacterial agent in one environmental agent radius.
- Agent properties – it includes GSR reaction flux calculations, biomass flux calculation, dividing time calculation, calculates substrate consumption and product production yield.
- Communication information – communicating with bacterial agent they should transfer movement and collision information, with environmental agent they transfer local temperature, pH level, hydrostatic and osmatic pressure. Also information about the produced product and consumed substrate yield is sent.



a) b)

Fig. 1. **Graphical interpretation of environmental agent and bacterial agent in modelling MAGSMS environment:** a – modelling environment; b – structure modelling environment

Conclusions

1. This modeling approach gives an opportunity to simulate different bacteria collaboration, with even different genome-scale reconstructions for product production. This approach modeling also is good for precision agriculture science, because it allows modeling the fermentation process starting with feasible different agricultural land climatic conditions and ending with bioreactor strict conditions.
2. This modeling weakness is calculation resources, because if we use many count bacteria and each bacterium has own large genome – scale reconstruction to run then one personal computer calculation resources will be not enough. There should be also tested possibilities to make MAGSMS for cluster type calculation resources distribution.
3. In future there is a need to create agents and their environment constraints, define their properties as close as it could be made relaying to real dynamical environment, to create agents collaboration and information transferring processes as close as real conditions in nature, to take research of the physical properties for bacteria in that kind of environment. Make real experiments and compare the obtained results with the modeling data.

Acknowledgements

This study is supported by ESF project 'Latvian Interdisciplinary Interuniversity Scientific Group of Systems'. Biology project number 2009/0207/1DP/1.1.1.2.0/09/APIA/VIAA/128.

References

1. Notebaart RA, van Enckevort FH, Francke C, Siezen RJ, Teusink B., Accelerating the reconstruction of genome-scale metabolic networks 7, 2006, 296 p.
2. Viļums - S., Towards application of cellular metabolic model of Saccharomyces cerevisiae in bioprocess control algorithm" 10th International Scientific Conference Engineering for Rural development 2011, Vol 10, May 26-27, 2011, Jelgava, pp. 97 - 102.
3. Schellenberger J., Que R., Fleming R. M. T., Thiele I., Orth J. D., Adam M st, Zielinski D. C., Bordbar A. , Lewis N. E., Rahmanian S. , Kang J., Hyduke D. R., Palsson B. Ø., Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0, Nature Protocols 6, 2011, pp. 1290–1307.
4. Fleming R. M. T., Thiele I. , von Bertalanffy 1.0: a COBRA toolbox extension to thermodynamically constrain metabolic models, 2011. [online] [11.03.2012]. Available at http://opencobra.sourceforge.net.
5. Kuchling A. M., Functional Programming HOWTO, Python v2.7.2 documentation, Python Software Foundation. [online] [2012-02-09]. Available at : http://docs.python.org/howto/functional.html.
6. Forster A. C., and G. M. Church, Towards synthesis of a minimal cell.Mol. Syst. Biol. 2:45, 2006, pp. 153 - 164.
7. Fraser C. M., Gocayne J. D., White O., Adams M. D., Clayton R. A., Fleischmann R. D., Bult C. J., Kerlavage A. R., Sutton G., Kelley J. M., Fritchman R. D., Weidman J. F., Small K. V., Sandusky M., Fuhrmann J., Nguyen D., Utterback T. R, Saudek D. M., Phillips C. A., Merrick J. M., Tomb J. F., Dougherty B. A., Bott K. F., Hu P. C., Lucier T. S., Peterson S. N., Smith H. O., Hutchison III C. A., Venter J. C., The minimal gene complement of Mycoplasma genitalium. Science 270, 1995, pp. 397–403.
8. Glass, J. I., Assad-Garcia N., Alperovich N., Yooseph S., Lewis M. R., Maruf M., Hutchison III C. A., Smith H. O., and Venter J. C., Essential genes of a minimal bacterium. Proc. Natl. Acad. Sci. USA, 2006, pp. 425–430.
9. Hutchison, C. A., Peterson S. N., Gill S. R., Cline R. T., White O., Fraser C. M., Smith H. O., Venter J. C., Global transposon mutagenesis and a minimal Mycoplasma genome. Science 286, 1999, pp. 2165–2169.
10. Mushegian, A. R., and Koonin E. V., A minimal gene set for cellular life derived by comparison of complete bacterial genomes, Proc. Natl. Acad. Sci. USA 93, 1996, pp. 10268–10273.

11. Kostenberga D., Marchaim U., Anaerobic digestion and horticultural value of solid waste from manufacture of instant coffee, Environmental Technology, Volume 14, Issue 10, 1993, pp. 973-980.
12. Dien. B. S. , Cotta, M. A., Jeffries T. W. , Bacteria engineered for fuel ethanol production: current status, Appl Microbiol Biotechnol, 2003, pp. 258 -266.
13. Ethan I., Lan , James C., Liao, Metabolic engineering of cyanobacteria for 1-Butanol production from carbon dioxide, Metabolic Engineering,Volume 13, Issue 4, 2011, pp. 353–36.
14. Min H., Sherman L. A. , Hydrogen Production by the Unicellular, Diazotrophic Cyanobacterium Cyanothece sp. Strain ATCC 51142 under Conditions of Continuous Light, applied and environmental microbiology, vol. 76, no. 13, 2010, pp. 4293–4301.
15. COBRA toolbox modeling approach. [online] [05.03.2012]. Available at : http://opencobra.sourceforge.net/openCOBRA/Welcome.html, 17.58, 05.03.2012.
16. O'Mahony F., Rural diary technology experiences in Ethiopia, Designed and printed at ILCA, Typeset on Linotype CRTronic, 1988, pp. 15 – 17.
17. Pommerville A., Jeffrey C., Alcamo's Fundamentals of Microbiology. Eighth Edition, Grube M., Rutkis R., Gavare M., Lasa Z., Strazdina I., Galinina N. and Kalnenieks U. Aerobic energy metabolism in Zymomonas mobilis respiratory knock-out mutants, FEMS, Geneva, 2012.
18. Thom S. R. and Marquis R. E., Microbial Growth Modification by Compressed Gases and Hydrostatic Pressure, Appl Environ Microbiol., 47(4), 2008, pp. 780–787.
19. Pentjušs A., Zacepins A., Gailums A. : Improving precision agriculture methods with multi agent systems in latvian agricultural field, 10. International scientific conference: „Engineering for Rural Development", LLU, TF 26 – 27.05.2011, pp. 109 - 114.
20. Wooldrige M., An Introduction to multi agent systems: second edition. United Kingdom: West Sussex, 2009, 461 p.
21. Mathlab mathematical calculation software. [online] [05.03.2012]. Available at: www.mathworks.com/products/matlab/, 17.58, 05.03.2012.