Development and Analysis of a Stoichiometric Model of Candidate Bacterium for Bioethanol Production Clostridium Thermocellum

Adel Maghsoudpour
Department of Mechanical and Aerospace Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

Ali Ghaffari
Department of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran

Mohammad Teshnehlab
Department of Electrical Engineering, K. N. Toosi University of Technology, Tehran, Iran

ABSTRACT
In this study we developed a computational model of core metabolism of *Clostridium thermocellum* which allows for *in silico* analysis of central metabolic fluxes in this strain. The model was used to predict a number of experimentally observed metabolic phenotypes in *C. thermocellum* growth cultures including cell growth rate and by-product secretion. Results of the model analysis show a good agreement between experimental data and model predictions. All the reactions included in the model are based on experimental evidence on their actual activity in the metabolism of *C. thermocellum*. Using flux balance analysis the distribution of flux over central metabolism and fermentative pathways are predicted. The results of model prediction can be further validated when experimental data on actual distribution carbon flux in *C. thermocellum* becomes available. The developed stoichiometric model can be applied for predicting the consequences of introducing pathways manipulation in this organism with the aim of improving ethanol and hydrogen production yields.

Keywords:
Stoichiometric Modeling, Central Metabolism, Consolidated Bioprocessing, *Clostridium thermocellum*.

1. INTRODUCTION
Ethanol production from plant biomass in the new paradigm of energy supply to currently fossil fuel based energy production. While several technology frameworks have been proposed for realizing ethanol production bioprocess, the large-scale industrial production of bioethanol has not reached its potential situation because of some economic considerations including low production yield and productivity. Conversion of lignocellulosic material to ethanol is carried out in nature through four biological steps including production of cellulase enzymes, hydrolysis of cellulose and hemicellulose to soluble sugares, fermentation of hexose sugares (resulting from cellulose hydrolysis) and fermentation of pentose sugares (resulting from hemicelluose hydrolysis). An economically viable process design for industrial conversion of plant biomass to ethanol has to harness all these biological steps within the process configuration. However, a potential approach to both lowering investment cost and increasing production yield of such a process is to combine the mentioned biological steps within more integrated process unites. This approach has been followed by several lines of research and has led to conceptual design of bioprocesses that integrate the second and third steps (simultaneous saccharification and fermentation; SSF), second, third and forth steps (simultaneous saccharification and co-fermentation; SSCF). The SSF bioprocess has been successfully implemented and the SSCF process is viewed as an accessible goal in middle term [1]. The logical end point to evolution of ethanol bioprocess design is manifested in a process configuration enabling combination of all four biologic steps in a single processing unit. Such a process in called consolidated bioprocessing (CBP) and its realization is awaiting for achievement of several technical milestones [2]. One major requisites of industrial realization of a CBP is to develop microbial production systems with the capabilities of releasing cellulase activities, efficiently hydrolyzing of cellulose and hemicellulose, and high yield fermentation of the resultant sugars to ethanol. A promising subset of microorganisms for development of CBD-enabling biocatalysts is thermophilic cellulytic anaerobes. Among this class of organisms, *C. thermocellum* is known as the most suitable candidate to be applied in a CBP configuration regarding its native characteristics such as very high cellulytic activity and its ability to efficiently convert microcrystalline cellulose to soluble sugars. Several studies have been investigated different biological features of this strain [3-11] and the efforts are increasingly growing to uncover the unknown aspects of the organism with respect to its application in CBP. In this study we follow a mathematical approach to analysis of core metabolism of *C. thermocellum* with the aim of providing a predictive view of flux distribution over the central and fermentative metabolic pathways. We develop a stoichiometric model of the central metabolism and analyze the model using a number of computational methods. The results of our study provide useful insight into the nature of flux distribution over core metabolic pathways of *C. thermocellum* which can be used toward achieving efficient strategies for redesign of the strain for be application in a consolidated bioprocess.

2. METHOD
2.1. Construction of a stoichiometric model
In an overall view, biochemical evidences confirm the activities of Embden-Meyerhof-Parnas pathway, the non-oxidative branch of pentose phosphate pathway and the routes from pyruvate toward fermentative products including ethanol, acetate and lactate when *C. thermocellum* is grown on cellobiose and glucose [12,13]. By referring to genome-annotation data in KEGG database [14], a list of central metabolic reactions of *C. thermocellum* and the corresponding genes was provided. Moreover transport reactions responsible for import of carbon and nitrogen sources were incorporated to

the network based on both bioinformatic and experimental evidences [15,16,19]. When the pathways of central metabolic reactions was completed and refined the network was converted to a mathematical model. An S matrix of stoichiometric coefficients was constructed with rows representing metabolites and columns representing reactions. Following a constraint based modeling approach [20], flux through the import direction of all transport reactions but those responsible for import of carbon and nitrogen sources were constrained to zero. By choosing biomass production as an objective function a flux balance analysis was performed resulting in prediction of *in silico* growth rate as well as by-ptoduct secretion rates of ethanol, acetate and formate.

2.2. Biomass composition

To enable analysis of constructed model using LP-based approaches a biomass objective function was formulated and incorporated to the model. There is no complete analysis of biomass of *C. thermocellum* therefore the composition of biomass objective function was determined according to the consensus cell constituents of bacteria, specifically the relative strains such as *C. acetobutylicum* [21]. Wherever *C. thermocellum* specific data for composition of macromolecules was available this data was used. The overall macromolecular composition of biomass objective function was specified based on that of the gram positive bacterium of the same phylum *B. subtillis* [22]. Amino acid composition of proteins was adopted from the measure amino acid composition data of *C. acetobutylicum* [21]. The fatty acid composition of lipid macromolecules was also taken from the same reference. The nucleotide composition of DNA and RNA was calculated based on *C. thermocellum* genomic data [23].

Table 1. The estimated values of biomass constituents of *C. thermocellum* utilized in formulation of biomass objective function in the stoichiometric model

Metabolites	mmolgDW-1	Metabolites	mmolgDW-1	Metabolites	mmolgDW-1
Protein (54.5 %)	0.8091	RNA (6.75 %)		Small molecule (1.9 %)	
Alanine	0.4922	ATP	0.0495	Menaquinone 7	0.0029
Arginine	0.564	CTP	0.0555	10-Formyltetrahydrofolate	0.004
Asparginine	0.1853	GTP	0.0467	NAD	0.0029
Aspartate	0.1853	UTP	0.0416	NADH	0.0028
Cysteine	0.2262			NADP	0.0026
Glutamine	0.1952	DNA (4 %)		NADPH	0.0025
Glutamate	0.064	dATP	0.0241	Coenzyme A	0.0025
Glycine	0.2479	dCTP	0.0166	FMN	0.0041
Histidine	0.0129	dGTP	0.0248	FAD	0.0024
Isoleucine	0.5668	dTTP	0.0147		
Leucine	0.2939			Ion (3.24)	
Lysine	0.1292	Lipid (7.83 %)		K	0.7126
Methionen	0.0422	Phosphatidylethanolamine	0.0006159	Mg	0.1027
Phenylalanine	0.0545	Phophstidylglycerol	0.0003053	Fe(+3)	0.0035
Proline	0.068	Cardiolipin	0.0000843	Ca	0.0032
Serine	0.0506			Phosphate	0.0145
Threonine	0.068	Cell Wall (23.1)		Diphospate	0.0009
Tryptophan	0.0506	Peptidoglycan	0.1018		
Tyrosine	0.2249	Carbohydrate	0.0806		
Valine		Teichoic acid	0.0178		

2.3. Growth medium

The composition of *in silico* medium used for simulation of growth is specified according to the chemically defined minimal medium for growth of *C. thermocellum* [24] composed of limited substrate (cellobiose or glucose), p-aminobenzoate and the ions of sodium, potassium, ferrous, calcium, magnesium, ammonium, phosphate and diphosphate. However, during model simulations not all of medium components were uptaken by because there was no demand reaction for their utilization in the model.

3. RESULTS AND DISCUSSION

3.1. Model description

A stoichiometric biochemical model of core metabolism of *C. thermocellum* including glycolysis, non-oxidative branch of pentose phosphate pathway, fermentative routes as well as artificial anabolic reactions toward biomass production was constructed. All catabolic reactions included in the model are based on experimental evidence on their activity during organism growth on cellulosic substrates [12-18]. The anabolic processes were considered as lumped reactions converting growth precursors to the building block of biomass macromolecules. Our model captures the major features of central metabolism in *C. thermocellum* and allows for *in silico* study of the effects of flux distribution on growth and by-prodcut secretion rates.

3.2. Model analysis

Using flux balance model of central metabolism of *C. thermocellum* the growth and distribution of fluxes over metabolic pathways was simulated. The results show a good agreement with experimental data obtained from growth cultures of *C. thermocellum*. Our model was able to predict the growth rate of the organism in the range of that observed in *C. thermocellum* cultures on cellobiose [13,25]. Moreover, consistent to the experimental data [7,17,26-27], the stoichiometric model predicted secretion of ethanol, acetate and formate as by-products of growth. However lactate secretion which is reported in several cultures of the strains was not predicted by FBA. This observation is consistent with the finding that lactate production is a phenotype related to over flow metabolism [17] which occurs when cell operates under sub-optimal conditions. As FBA is based on assumption that organism optimally allocates its resource toward maximizing the objective function (growth rate) it is not surprising to see that lactate is not produced during FBA simulation.

A theoretical molar growth yield of 20 and 34 was obtained for simulation of growth on glucose and cellobiose which is in very good agreement with experimental data obtained from growth cultures of this strain on these substrates [13].

3.3. Evaluating Central Metabolic Capabilities

C. thermocellum is a strictly anaerobic organism and the substrate level phosphorylation is the sole source of ATP generation required for cell growth and maintenance activities. This low level redundancy to obtain required energy, results in low flexibility of central metabolic pathways in distribution of flux over reactions. In order to evaluate maximal capacity of central metabolism we simulated maximum theoretical yields of ethanol, acetate, lactate and formate for growth on cellobiose (Figure 1). It was observed that the model can produce these organic compounds in amounts close to their maximum theoretical yields.

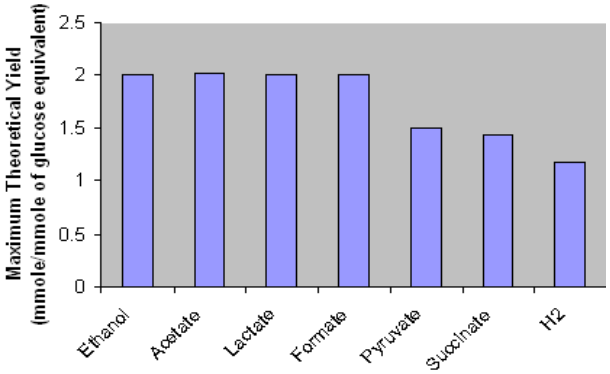

Figure 1. Predicted maximum production yield of organic compounds in *C. thermocellum* using the stoichiometric model of central metabolism

3.4. Relation between ethanol production and organic acid formation rates

The competitive relation between ethanol synthesis pathway and the routes toward other fermentation end products can be deduced from the structure and the stoichiometry of reactions in catabolism of pyruvate. However, the degree to which ethanol production flux is negatively affected by the activities of theses competing fluxes at various growth rates is difficult to be evaluated intuitively. The constraint based analysis allows for two-variable sensitivity analysis of metabolic fluxes which can be applied to gain a better understanding of the interactions between fluxes through the competing reactions and the growth rate. By applying this technique, we obtained a trade-off between ethanol production rate, growth rate and formation rates of acetate and lactate (Figure 2). As seen and already expected the growth rate has the most negative effect on the production rate of ethanol while the formation rate of lactate and acetate stand at later order of negatively affecting ethanol production.

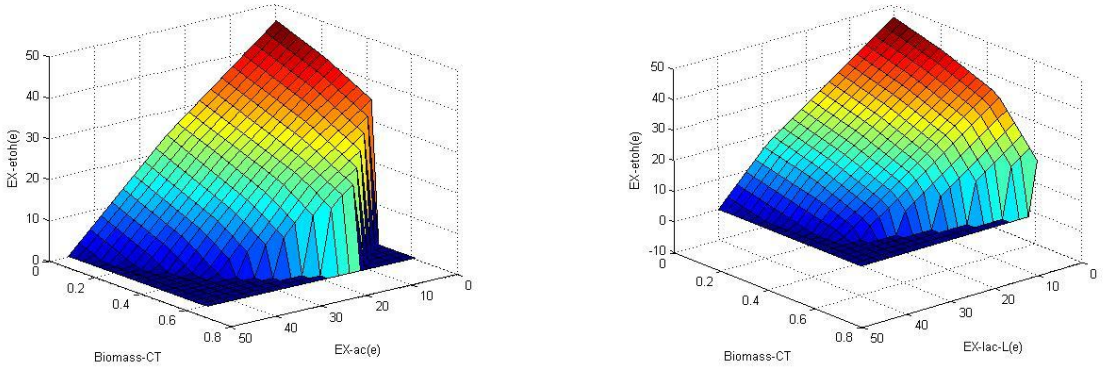

Figure 2. The relationship between ethanol production rate, growth rate and acetate (A) and lactate (B) secretion rates.

4. CONCLUSION

A stoichiometric model of central metabolism of *C. thermocellum* was developed. The model was used to predict a number of major metabolic phenotypes in growth cultures of the strain including growth rate and by-product secretion. The agreement between modeling results and experimental data show model capability to reproduce the metabolic behavior of the organism. Such a model capability can be used to *in silico* sturdy of core metabolism of *C. thermocellum* with the aim of finding efficient strategies for improving hydrogen and ethanol production rates in this organism. We applied our model to investigate the relationship between fluxes toward two major economically important by-products of *C. thermocellum* growth on cellulosic substrates including ethanol and hydrogen. Our model predicts a competitive relation between production rates of these two products. Also consistent to experimental data, the sensitivity analysis of central metabolic pathways show a competitive relationship between flux through ethanol and the fluxes toward other fermentative products such as acetate, lactate and succinate.

5. ACKNOWLEDGMENTS
The author would like to express his appreciation to the Department of Mechanical and Aerospace Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran for financial support.

6. REFERENCES
[1] Wooley, R., M. Ruth, D. Glassner, and J. Sheehan. 1999. Process design and costing of bioethanol technology: a tool for determining the status and direction of research and development. Biotechnol. Prog. 15:794–803.

[2] Lee R. Lynd, Paul J. Weimer, Willem H. van Zyl, and Isak S. Pretorius: Microbial Cellulose Utilization: Fundamentals and Biotechnology. Microbiol Mol Biol Rev. 2002 December; 66(4): 739.

[3] Demain AL, Newcomb M, Wu JH. Cellulase, clostridia, and ethanol. Microbiol Mol Biol Rev. 2005 Mar;69(1):124-54.

[4] Yi-Heng Percival Zhang* and Lee R. Lynd: Cellulose utilization by *Clostridium thermocellum*: Bioenergetics and hydrolysis product assimilation. Proc Natl Acad Sci U S A. 2005 May 17; 102(20): 7321–7325.

[5] Lu Y, Zhang YH, Lynd LR : Enzyme-microbe synergy during cellulose hydrolysis by *Clostridium thermocellum*. Proc Natl Acad Sci U S A. 2006 Oct 31;103(44):16165-9. Epub 2006 Oct 23.

[6] Islam R, Cicek N, Sparling R, and Levin DB. 2006. Effect of substrate loading on hydrogen production during anaerobic fermentation by *Clostridium thermocellum* 27405. Appl. Microbiol. Biotechnol. (accepted, Jan. 2006)

[7] McBee RH: The Characteristics of *Clostridium thermocellum*. J Bacteriol 1954, 67(4):505-506.

[8] Eric A. Johnson, Fredereque Bouchot and Arnold L. Demain: Regulation of Cellulase Formation in *Clostridium thermocellum*. Journal of General Microbiology 131 (1985), 2303-2308.

[9] Yi-Heng Percival Zhang and Lee R. Lynd, Kinetics and Relative Importance of Phosphorolytic and Hydrolytic Cleavage of Cellodextrins and Cellobiose in Cell Extracts of *Clostridium thermocellum*. Applied and Environmental Microbiology, March 2004, p. 1563-1569, Vol. 70, No. 3.

[10] Yi-Heng Percival Zhang and Lee R. Lynd, Regulation of Cellulase Synthesis in Batch and Continuous Cultures of *Clostridium thermocellum*. Journal of Bacteriology, January 2005, p. 99-106, Vol. 187, No. 1.

[11] Gold ND, Martin VJ: Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. J Bacteriol. 2007 Oct;189(19):6787-95. Epub 2007 Jul 20.

[12] Lamed R, Zeikus JG: Ethanol production by thermophilic bacteria: relationship between fermentation product yields of and catabolic enzyme activities in *Clostridium thermocellum* and *Thermoanaerobium brockii*. J Bacteriol. 1980 Nov;144(2):569-78.

[13] Ng TK, Zeikus JG: Differential metabolism of cellobiose and glucose by *Clostridium thermocellum* and *Clostridium thermohydrosulfuricum*. J Bacteriol. 1982 Jun;150(3):1391-9.

[14] Kanehisa M, Goto S: KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucl Acids Res 2000, 28(1):27-30.

[15] Patni NJ, Alexander JK: Utilization of glucose by *Clostridium thermocellum*: presence of glucokinase and other glycolytic enzymes in cell extracts. J Bacteriol. 1971 Jan;105(1):220-5.

[16] Patni NJ, Alexander JK: Catabolism of fructose and mannitol in *Clostridium thermocellum*: presence of phosphoenolpyruvate: fructose phosphotransferase, fructose 1-phosphate kinase, phosphoenolpyruvate: mannitol phosphotransferase, and mannitol 1-phosphate dehydrogenase in cell extracts.

[17] Carlo R. Carere, Vipin Kalia, Richard Sparling, Nazim Cicek, David B. Levin: Pyruvate catabolism and hydrogen synthesis pathway genes of *Clostridium thermocellum* ATCC 27405.

[18] Stevenson DM, Weimer PJ: Expression of 17 genes in *Clostridium thermocellum* ATCC 27405 during fermentation of cellulose or cellobiose in continuous culture. Appl Environ Microbiol. 2005 Aug;71(8):4672-8.

[19] Yakir Nataf, Sima Yaron, Frank Stahl, Raphael Lamed, Edward A. Bayer, Thomas-Helmut Scheper, Abraham L. Sonenshein, and Yuval Shoham, Cellodextrin and Laminaribiose ABC Transporters in *Clostridium thermocellum.* J Bacteriol. 2009 January; 191(1): 203–209.

[20] Nathan D. Price, Jennifer L. Reed[1] & Bernhard Ø. Palsson, Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature Reviews Microbiology* 2, 886-897.

[21] Lee J, Yun H, Feist AM, Palsson BØ, Lee SY. Genome-scale reconstruction and in silico analysis of the *Clostridium acetobutylicum* ATCC 824 metabolic network. Appl Microbiol Biotechnol. 2008 Oct;80(5):849-62. Epub 2008 Aug 29.

[22] Oh YK, Palsson BO, Park SM, Schilling CH, Mahadevan R. Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data. J Biol Chem. 2007 Sep 28;282(39):28791-9. Epub 2007 Jun 15.

[23] Clostridium thermocellum ATCC 27405, complete genome, http://www.ncbi.nlm.nih.gov/nuccore/CP000568

[24] Johnson EA, Madia A, Demain AL: Chemically Defined Minimal Medium for Growth of the Anaerobic Cellulolytic Thermophile *Clostridium thermocellum*. Appl Environ Microbiol 1981, 41(4):1060-1062.

[25] Lee Rybeck Lynd, Hans E. Grethlein, and Richard H. Wolkin, Fermentation of Cellulosic Substrates in Batch and Continuous Culture by *Clostridium thermocellum*. *Appl* Environ Microbiol. 1989 December; 55(12): 3131–3139.

[26] Rydzak T, Levin DB, Cicek N, Sparling R: Growth phase-dependant enzyme profile of pyruvate catabolism and end-product formation in *Clostridium thermocellum* ATCC 27405. J Biotechnol. 2009 Mar 25;140(3-4):169-75. Epub 2009 Feb 7.

[27] Sparling R, Islam R, Cicek N, Carere C, Chow H, Levin DB: Formate synthesis by *Clostridium thermocellum* during anaerobic fermentation. Can J Microbiol. 2006 Jul;52(7):681-8.