



Universidade Federal de Santa Catarina
Centro Tecnológico - CTC
Departamento de Engenharia Química e Engenharia de Alimentos
Programa de Pós-Graduação em Engenharia Química

Samara Silva de Souza

**BIOLOGIA SISTÊMICA DA PRODUÇÃO DE CELULOSE
BACTERIANA ATRAVÉS DA RECONSTRUÇÃO
METABÓLICA DA *Gluconacetobacter hansenii***

Florianópolis
2014

Samara Silva de Souza

**BIOLOGIA SISTÊMICA DA PRODUÇÃO DE CELULOSE
BACTERIANA ATRAVÉS DA RECONSTRUÇÃO
METABÓLICA DA *Gluconacetobacter hansenii***

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Química da Universidade Federal de Santa Catarina como requisito parcial para obtenção do grau de Mestre em Engenharia Química.

Orientador: Prof. Dr. Luismar Marques Porto

Florianópolis
2014

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Souza, Samara Silva de

Biologia sistêmica da produção de celulose bacteriana
através da Reconstrução Metabólica da *Gluconacetobacter
hansenii* / Samara Silva de Souza ; orientador, Luismar
Marques Porto - Florianópolis, SC, 2014.

144 p.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-Graduação em
Engenharia Química.

Inclui referências

1. Engenharia Química. 2. *Gluconacetobacter hansenii*
ATCC 23769. 3. Reconstrução da rede metabólica. 4. Análise
baseada em restrições. 5. Celulose. I. Porto, Luismar
Marques. II. Universidade Federal de Santa Catarina.
Programa de Pós-Graduação em Engenharia Química. III. Título.

Samara Silva de Souza

**BIOLOGIA SISTÊMICA DA PRODUÇÃO DE CELULOSE
BACTERIANA ATRAVÉS DA RECONSTRUÇÃO
METABÓLICA DA *Gluconacetobacter hansenii***

Esta dissertação foi julgada adequada para a obtenção do Título de Mestre em Engenharia Química e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Química da Universidade Federal de Santa Catarina

Florianópolis, 15 de Abril de 2014

Prof. Dr. Ricardo Antonio Francisco Machado
Coordenador do Programa de
Pós-Graduação em Engenharia Química

Banca examinadora:

Prof. Dr. Luismar Marques Porto
Presidente
UFSC

Prof. Dr. Everton Skoronski
Membro Externo
UDESC/Lajes

Prof. Dr. Eugênio Simão
Membro Externo
UFSC/Araranguá

Dr. Leandro Antunes Berti
Membro Externo
UFSC/CERTI

*“Alguns homens vêem as coisas como são, e dizem Por quê?
Eu sonho com as coisas que nunca
foram e digo Por que não?”
George Bernard Shaw*

AGRADECIMENTOS

Aos meus pais, Ronaldo e Adriana, pelo amor e apoio em todos os momentos da minha vida. Vocês são minhas referências e a base do que sou hoje.

Ao meu namorado Felipe pela compreensão, carinho e conselhos dedicados durante esta jornada.

Ao Professor Dr. Luismar Marques Porto, pela oportunidade que me foi dada de fazer parte da equipe do InteLab e pelos ensinamentos.

A todos os colegas do laboratório, pela companhia, trocas de experiência e risadas. Trabalhar com vocês foi muito gratificante.

Aos professores da banca de defesa pela disponibilidade e pelas valiosas contribuições.

Ao Programa de Pós-Graduação de Engenharia Química da UFSC, por possibilitar um grande enriquecimento intelectual permitindo o desenvolvimento deste trabalho.

Aos docentes do programa, pela dedicação ao ensino de qualidade.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudo.

Enfim, a tudo e a todos que acreditaram em mim e no meu potencial.

Muito obrigada!

RESUMO

Nas últimas décadas, reconstruções e aplicações de modelos metabólicos em escala genômica influenciaram fortemente o campo da biologia de sistemas, com base em uma combinação de informações sobre a sequência genômica e informações bioquímicas, proporcionando uma plataforma sobre a qual a análise computacional de uma rede metabólica pode ser realizada. A bactéria Gram-negativa *Gluconacetobacter* tem sido extensivamente usada para a síntese de celulose. Recentemente a sequência do genoma da bactéria *Gluconacetobacter hansenii* ATCC 23769 (GenBank: CM000920 e taxonomia ID: 714995), se tornou disponível em bancos de dados, permitindo os estudos deste organismo e sua produção de celulose bacteriana. Dentre os campos de aplicações da celulose bacteriana, está o da medicina como curativo de feridas, material de pele artificial entre outras aplicações. Para uma visão global das capacidades da *G. hansenii* é fundamental a construção de um modelo metabólico. O conhecimento sobre a via metabólica de um organismo permite a compreensão da fisiologia celular e regulação de seu metabolismo, sendo a quantificação de fluxos metabólicos um importante objetivo, principalmente no que diz respeito à obtenção de produtos úteis comercialmente ou cientificamente. A abordagem da reconstrução e análise baseada em restrições tem sido utilizada para analisar e construir *in silico* reconstruções da rede metabólica. Neste trabalho a rede foi reconstruída a partir da sequência genômica anotada deste organismo usando o software Pathway Tools para gerar uma rede preliminar. Além disso, um recurso para a geração e análise de modelos metabólicos em escala genômica, chamado Modelo SEED foi usado criando um banco de dados do organismo. O passo seguinte foi a etapa de cura manual baseada numa combinação de informações genômicas, bioquímicas e fisiológicas em diversos bancos de dados e literatura. O objetivo foi reconstruir a rede com as principais vias e reações para a produção da celulose. Foi utilizado o software MATLAB® utilizando as funções disponíveis no COBRA *toolbox*. Além disso, um conjunto de ferramentas de biologia de sistemas computacionais, desenvolvido pelo nosso grupo chamado GENSys (*Genomic Engineering System*), foi aplicado à rede. O GENSys compreende vários módulos que permitem a análise e simulação de redes bioquímicas, tais como ABF (análise de

balanço de fluxo) e AFM (análise de fluxo metabólico). Usando o modelo matemático em conjunto com os métodos baseados em restrição, foram simulados os fluxos metabólicos, mimetizando três diferentes condições ambientais para compreender os efeitos de produção de celulose, crescimento bacteriano e a distribuição dos fluxos na rede variando as fontes de carbonos: glicose, glicerol e manitol. As modificações no protocolo desenvolvido para reconstruções metabólicas permitiram a construção do primeiro modelo *core* da *G. hansenii*.

Palavras Chave: *Gluconacetobacter hansenii* ATCC 23769; reconstrução da rede; análise baseada em restrições; celulose

ABSTRACT

Systems Biology of the Bacterial Cellulose Production by the Metabolic Reconstruction of *Gluconacetobacter hansenii*

In the last decade, reconstruction and applications of genome-scale metabolic network have strongly influenced the field of systems biology, based on a combination of genome sequence information and detailed biochemical information, providing a platform on which high throughput computational analysis of metabolic network can be accomplished. The Gram-negative bacterium *Gluconacetobacter* has been extensively used for cellulose synthesis. Recently *Gluconacetobacter hansenii* ATCC 23769 genome sequence has become available (GenBank number: CM000920 and taxonomy ID: 714995), allowing studies of this organism and its production of bacterial cellulose. Among the application fields of bacterial cellulose is the medicine, used as wound dressing, artificial skin material and other applications. For a global understanding of the capabilities of *G. hansenii* it is essential the construction of a metabolic model. The knowledge about the metabolic pathway of an organism allows the understanding of cellular physiology and regulation of metabolism, which the quantification of metabolic fluxes is an important objective, especially with regard to obtaining commercially or scientifically useful products. The constraint-based reconstruction and analysis approach has been used to analyze and build *in silico* metabolic network reconstructions. In this work, the network was reconstructed from the annotated genomic sequence of this organism using Pathway Tools software to generate a preliminary network. Also a resource for generation and analysis of genome-scale metabolic models, called Model SEED was used by creating a database of the organism. The next step was the manual curation based on a combination of genomic, biochemical and physiological information in several databases and literature. The goal was to rebuild the network with the main pathways and reactions for the production of cellulose. MATLAB® using the functions available in the COBRA toolbox was also used. Moreover, a set of computational systems biology tools, written by our group, called GEnSys, (Genomic Engineering System), was applied to the network. GEnSys comprises several modules that allow analysis and simulation of biochemical reaction networks, such as FBA (flux balance analysis)

and MFA (metabolic flux analysis). Using the model in conjunction with the constraint-based methods, the metabolic fluxes were simulated mimicking three different environmental conditions to understand the effects of cellulose production, bacterial growth and flux distribution in the network by varying the carbon sources: glucose, mannitol and glycerol. The modifications in the protocol developed for metabolic reconstructions allowed the construction of the first core model of *G. hansenii*.

Keywords: *Gluconacetobacter hansenii* ATCC 23769; network reconstruction; constraint based analysis; cellulose

ÍNDICE DE FIGURAS

Figura 1 - Vias metabólicas da <i>Escherichia coli</i>	40
Figura 2 - Número acumulativo de reconstruções metabólicas. Lista desenvolvida e atualizada pelo grupo <i>System Biology Research</i>	42
Figura 3 - Interface do GENsYS 1.0 no ambiente MATLAB®	51
Figura 4 - Representação das ligações de hidrogênio inter- e intramolecular da celulose.	53
Figura 5 - Série de enzimas reguladoras na biossíntese de celulose bacteriana.	55
Figura 6 - Enzima Celulose Sintase:	55
Figura 7 - Subunidades da Celulose Sintase: <i>operon acs</i>	56
Figura 8 - Estruturas de celulose bacteriana obtidas variando-se as condições de cultura bacteriana.....	59
Figura 9 - Esquema geral da metodologia aplicada. A partir da bactéria <i>G. hansenii</i> foi mapeado seu metabolismo e utilizado ferramentas computacionais para analisar sua rede metabólica.	61
Figura 10 - Etapas do processo de reconstrução metabólica	62
Figura 11 - Fluxograma representativo das etapas de construção da rede metabólica.	63
Figura 12 - Trabalho submetido ao sistema do RAST.	64
Figura 13 - Interface dos bancos de dados: KEGG e Brenda.	66
Figura 14 - Definição das condições de contorno.	68

Figura 15 - A matriz estequiométrica como centro do processo de reconstrução metabólica.....	69
Figura 16 - <i>Toolboxes</i> utilizados no ambiente MATLAB®: COBRA E GENSYS.....	70
Figura 17 - Interface rBioNet mostrando o editor de reações e metabólitos e o criador de reconstruções	71
Figura 18 - Genoma anotado no Banco de Dados do NCBI	78
Figura 19 - Genoma da bactéria e formatos disponíveis para exportar a sequência genômica	79
Figura 20 - Formatos FASTA e GenBank da <i>G. hansenii</i>	79
Figura 21 - PGDB da <i>Gluconacetobacter hansenii</i> criado no Pathway Tools a partir do genoma anotado	80
Figura 22 - Visualizações das vias no Mapa celular (<i>Cellular overview</i>)	81
Figura 23 - Componentes do mapa celular destacados nas vias.....	82
Figura 24 - Visualizações do Mapa genômico (<i>Genome Overview</i>).....	82
Figura 25 - Visualização do arquivo SBML do modelo gerado.....	83
Figura 26 - Planilha gerada com as informações de metabólitos e reações.....	84
Figura 27 - Visualização da página principal dos resultados obtidos com o RAST Server	85
Figura 28 - Busca em bancos de dados do organismo - IMG	86
Figura 29 - Subsistema glicólise e gliconeogênese da <i>G. hansenii</i>	88

Figura 30 - Via das pentoses fosfato da <i>G. hansenii</i>	90
Figura 31 - Ciclo tricarboxílico da <i>G. hansenii</i>	90
Figura 32 - Via Entner-Doudoroff (ED) da <i>G. hansenii</i>	91
Figura 33 – Cadeia de transporte de elétrons e fosforilação oxidativa. .	92
Figura 34 - Modelo simplificado da rota metabólica para a biossíntese de celulose a partir de glicose, manitol ou glicerol por <i>G. hansenii</i>	94
Figura 35 – Representação da rede metabólica simplificada da <i>G. hansenii</i>	95
Figura 36 – Componentes da estrutura do Modelo no MATLAB®	96
Figura 37 - Validação do arquivo SBML	97
Figura 38 - Layout gerado pelo CellDesigner	98
Figura 39 - Estrutura da matriz esparsa gerada no COBRA.	99
Figura 40 - AFM do modelo com entrada de glicose	101
Figura 41 - AFM do modelo com entrada de manitol	101
Figura 42 – AFM do modelo com entrada de glicerol	102
Figura 43 – Análise de fluxo metabólico para as três fontes de carbono.	103
Figura 44 – Variações de pH ao longo de 96h na produção de CB.	105
Figura 45 – Análise de Balanço de Fluxo para as três fontes de carbono com o crescimento celular como função objetivo para maximizar.	110

ÍNDICE DE TABELAS

Tabela 1 - Componentes e concentração do meio Yamanaka (g/L).....	73
Tabela 2 - Recursos <i>online</i> para a reconstrução da rede metabólica e os respectivos <i>links</i> de acesso	74
Tabela 3 - Dados e identificações da bactéria	78
Tabela 4 - Demanda dos precursores para <i>G. xylinus</i>	93
Tabela 5 - Valores dos fluxos de ácido glucônico para cada fonte de carbono.....	104
Tabela 6 - Parâmetros e dados para calcular as taxas de entrada dos componentes do meio mínimo	107
Tabela 7 - Taxas de consumo dos componentes do meio mínimo em mmol/gDW/h	108
Tabela 8 - Composição de meio mínimo com as condições de contorno	108
Tabela 9 - Taxas de crescimento obtidos pela ABF para os três substratos ao maximizar o crescimento celular.....	111
Tabela 10 - Resultados das simulações com ABF para maximização da celulose com entrada do meio mínimo de crescimento.....	112

LISTA DE ABREVIATURAS

13DPG	1-3-Difosfoglicerato
2DDG6P	2-Ceto-3-desoxi-6-fosfoglucanato
2PG	2-Fosfoglicerato
3PG	3-Fosfoglicerato
6PGC	6-Fosfoglucanato
6PGL	Glicose-1,5-lactona-6-fosfato
ABF	Análise de Balanço de Fluxo
ACCOA	Acetil-CoA
ACON-C	Aconitato
ADP	Adenosina difosfato
AFM	Análise de Fluxo Metabólico
AKG	2-Oxoglutarato
AMP	Adenosina monofosfato
ATP	Adenosina trifosfato
BIOMASS	Biomassa
CELL	Celulose ou 1,4- β -D-glucana
CIT	Citrato
CO2	Dióxido de carbono
COA	Coenzima A
DHAP	Diidroxiacetona fosfato
E4P	D-Eritrose 4-fosfato
F6P	D-Frutose 6-fosfato
FAD	Flavina adenina dinucleotídeo oxidada
FADH2	Flavina adenina dinucleotídeo reduzida
FDP	Frutose 1,6-difosfato
FERRICYTOCHROME C	Citocromo c oxidado
FERROCYTOCHROME C	Citocromo c reduzido
FRU	D-Frutose
FUM	Fumarato
G1P	D-Glicose 1-fosfato

G3P	Gliceraldeído 3-fosfato
G6P	D-Glicose 6-fosfato
GLC-D	D-Glicose
GLCN	Ácido glucônico
GLN-L	L-Glutamina
GLU-L	L-Glutamato
GLYC	Glicerol
GLYC3P	Glicerol -3-fosfato
H	H+ /Próton
H2O	Água
IMG	Integrated Microbial Genomes
ICIT	Isocitrato
KEGG	Kyoto Encyclopedia of Genes and Genomes
MAL	(S)-Malate
MANN	D-Manitol
NAD	Nicotinamida adenina dinucleotídeo
NADH	Nicotinamida adenina dinucleotídeo – reduzida
NADP	Nicotinamida adenina dinucleotídeo fosfato
NADPH	Nicotinamida adenina dinucleotídeo fosfato – reduzida
NCBI	National Center for Biotechnology Information
NGAM	<i>Non-growth associated maintenance</i>
NH4	Amônia
O2	Oxigênio
OAA	Oxaloacetato
PEP	Fosfoenolpiruvato
PGDB	<i>Pathway Genome Database</i>
PGL	D-Glicona-1,5-lactona
PI	Fosfato
PPI	Difosfato
PYR	Piruvato
Q8	Ubiquinona

Q8H2	Ubiquinol
R5P	D-Ribose 5-fosfato
RU5P	D-Ribulose 5-fosfato
S7P	Sedoepulose 7-fosfato
SUCC	Succinato
SUCCOA	Succinil-CoA
UDP	Uridina difosfato
UDPG	UDP-glicose
UTP	Uridina trifosfato
XU5P	D-Xilulose 5-fosfato

SUMÁRIO

CAPÍTULO 1 - INTRODUÇÃO, OBJETIVOS, JUSTIFICATIVA E MOTIVAÇÃO	27
1.1 Introdução	27
1.2 Objetivos.....	32
1.2.1 Objetivo Geral	32
1.2.2 Objetivos Específicos.....	32
1.3 Justificativa e motivação.....	33
1.4 Estrutura do documento.....	34
CAPÍTULO 2 - FUNDAMENTAÇÃO TEÓRICA	35
2.1 Engenharia metabólica e biologia de sistemas.....	35
2.1.1 Abordagens na engenharia metabólica para reconstrução de redes.....	36
2.2 Vias metabólicas.....	38
2.3 Reconstrução metabólica em escala genômica	40
2.3.1 Análise de Fluxo Metabólico (AFM).....	42
2.3.2 Análise de Balanço de Fluxos (ABF).....	43
2.3.3 Fluxos metabólicos e matriz estequiométrica.....	44
2.4 Ferramentas computacionais.....	46
2.4.1 Bancos de dados e ferramentas genômicas.....	47
<i>2.4.1.1 NCBI.....</i>	<i>47</i>

2.4.1.2 <i>Integrated Microbial Genomes (IMG)</i>	47
2.4.1.3 <i>Kyoto Encyclopedia of Genes and Genomes(KEGG)</i> ...	48
2.4.2 Pathway Tools	48
2.4.3 Model SEED e RAST Server	49
2.4.4 System Biology Markup Language (SBML)	49
2.4.5 Implementações em MATLAB®	50
2.4.5.1 <i>COBRA</i> toolbox	50
2.4.5.1.1 <i>rBioNet</i>	51
2.4.5.2 <i>GEnSys</i>	51
2.5 <i>Gluconacetobacter hansenii</i> ATCC 23769	52
2.5.1 Celulose Bacteriana	53
2.5.1.1 <i>Biossíntese de celulose bacteriana</i>	54
2.5.1.2 <i>Efeito das fontes de carbono na produção de celulose bacteriana</i>	56
2.5.1.3 <i>Aplicações da celulose bacteriana</i>	57
CAPÍTULO 3 - METODOLOGIA	61
3.1 <i>Reconstrução da rede metabólica em escala genômica</i>	62
3.1.1 Etapa 1 - Reconstrução preliminar – o <i>draft</i> do genoma	63
3.1.2 Etapa 2 - Cura manual	64
3.1.3 Etapa 3 - Conversão da construção metabólica para o formato computacional	69
3.1.3.1 <i>AFM e ABF no GEnSys</i>	72
3.1.3.2 <i>Funções objetivo</i>	72
3.1.3.3 <i>Meio mínimo de crescimento</i>	73

3.1.3.4 Ferramentas de bioinformática.....	73
3.1.4 Etapa 4 - Validação do modelo matemático	75
CAPÍTULO 4 - RESULTADOS E DISCUSSÃO.....	77
4.1 Construção do <i>draft</i> da rede metabólica da <i>G. hanseni</i>	77
4.2 Cura Manual do <i>draft</i> gerado.....	86
4.2.1 Montagem da rede metabólica <i>core</i>	87
4.2.2 Vias do metabolismo <i>core</i> da <i>G. hanseni</i>	86
4.3 Conversão da construção metabólica para o formato computacional	96
4.3.1 Análise de Fluxo Metabólico no GEnSys	100
4.3.2 Análise de Balanço de Fluxo no GEnSys	106
4.3.2.1 Meio mínimo e funções objetivo	107
4.3.2.2 Caso I: Maximizando a produção de biomassa.....	109
4.3.2.3 Caso II: Maximizando a produção de celulose	111
4.3.2.4 Caso III: Maximizando metabólitos externos e avaliando limitação de nutrientes	112
4.3.2.5 Caso IV: Outras simulações: reações de demanda....	113
4.4 Validação da rede metabólica	113
CAPÍTULO 5 - CONCLUSÕES, CONSIDERAÇÕES FINAIS E SUGESTÕES.....	115
REFERÊNCIAS BIBLIOGRÁFICAS	119
APÊNDICE A - MFA no GEnSys.....	137
APÊNDICE B - FBA no GEnSys	138
APÊNDICE C - REAÇÕES DO MODELO	139

CAPÍTULO 1

INTRODUÇÃO, OBJETIVOS, JUSTIFICATIVA E MOTIVAÇÃO

1.1 Introdução

Os avanços no sequenciamento do genoma têm progredido rapidamente, acompanhado do alto rendimento das análises e custos cada vez menores. Tais tecnologias estão rotineiramente sendo aplicadas em distintas áreas da biologia e da medicina, permitindo que os pesquisadores abordem questões biológicas relevantes, inacessíveis anteriormente (METZKER, 2005, 2010; ZHAO e GRANT, 2011).

Com o sequenciamento do genoma humano, em 1990, o campo da biologia entrou na era pós-genômica, e com isto veio a promessa de que o comportamento fenotípico de um organismo poderia ser previsto por análises computacionais do genoma. A era pós-genômica tem como fundamento relacionar a sequência do genoma anotado com as funções fisiológicas de uma célula, possibilitando reconstituir vias metabólicas completas (DUROT *et al.*, 2009; EDWARDS *et al.*, 2001; FONG e PALSSON, 2004). Aliado a estes avanços surgiu a oportunidade de prever a nível computacional o diagnóstico e o tratamento de doenças, através da especificidade de medicamentos para cada paciente e avanços na bioengenharia a nível industrial (BARABASI e OLTVAI, 2004; DUROT *et al.*, 2009; NG e KIRKNESS, 2010; SOON *et al.*, 2013; SWAINSTON, 2011).

Neste sentido, ocorreu uma mudança na situação da pesquisa biológica através do uso de computadores e a Web. Nos últimos anos, a quantidade de novos dados biológicos produzidos pelos esforços dos pesquisadores vem mudando o padrão no qual as pesquisas que costumavam ser iniciadas em laboratório, começam agora no computador, com a busca de informações que possam sugerir novas hipóteses (BAXEVANIS, 2001; COVERT *et al.*, 2001; HENRY *et al.*, 2010).

Assim, surgiu o campo da bioinformática e da biologia computacional que consiste na aplicação de técnicas de computação intensivas para a compreensão dos processos biológicos. A geração de quantidades de dados biológicos cada vez maiores levou ao desenvolvimento dessa área, que permitiu a manipulação confiável e

rápida de grande quantidade de dados e que viabiliza e amplia o potencial das pesquisas. A tarefa da bioinformática é organizar e analisar dados complexos resultantes de modernas técnicas de biologia molecular e bioquímica e compreender como as informações contidas nos genes são refletidas em características fisiológicas (BORK e SERRANO, 2005; COVERT *et al.*, 2001; FERNALD *et al.*, 2011).

Nos últimos anos, outra área tem se destacado, a biologia de sistemas, ou biologia sistêmica, que tem focado em reconstruções metabólicas e na análise de dados baseada em restrições, que fornecem representações matemáticas e computacionais das capacidades metabólicas conhecidas de um dado organismo (BLANK e EBERT, 2013; BORDBAR e PALSSON, 2012). Em outras palavras, o objetivo da biologia de sistemas é entender as interações entre os vários componentes celulares, para eventualmente, prever o comportamento da rede, célula e do organismo em questão (BORK e SERRANO, 2005; FEIST *et al.*, 2009).

Cada vez mais as bactérias são utilizadas em processos industriais, para produção de produtos químicos, alimentos, medicamentos, entre outros (LEE *et al.*, 2012). Através das ferramentas de engenharia metabólica, as principais vias bioquímicas de bactérias podem ser manipuladas e otimizadas para produzir de forma mais eficiente os compostos industriais de interesse em diversas áreas (MCCLOSKEY *et al.*, 2013).

Várias linhagens de bactérias já têm genomas sequenciados e anotados, os quais têm sido utilizados em associação com dados bioquímicos e fisiológicos para reconstruir redes metabólicas em escala genômica (BECKER e PALSSON, 2005b; TERZER *et al.*, 2009). Recentemente modelos genômicos foram reconstruídos para algumas espécies bacterianas com o objetivo de ampliar a quantidade e qualidade das informações que haviam sido anotadas na literatura ou nos bancos de dados previamente (DUROT *et al.*, 2009). A rede preliminar gerada a partir do genoma anotado, ainda exige uma cura manual significativa apenas para definir o conjunto abrangente de capacidades metabólicas que o organismo possui (FEIST *et al.*, 2009; MCCLOSKEY *et al.*, 2013). A reconstrução metabólica mais detalhada e completa até o momento é referente à bactéria *Escherichia coli* (ALMAAS *et al.*, 2004; EDWARDS *et al.*, 2001; EDWARDS e PALSSON, 2000; KESELER *et al.*, 2005; ORTH *et al.*, 2011).

O processo de reconstrução metabólica geralmente é trabalhoso e pode levar anos para ser bem detalhado e estudado quando envolve um

genoma bacteriano, e envolvendo diversos pesquisadores (DUARTE *et al.*, 2007). Muitas vezes, o processo de reconstrução é interativo e contínuo, como por exemplo, a rede metabólica da *Escherichia coli*, cuja reconstrução foi expandida e refinada ao longo dos últimos 20 anos (FEIST e PALSSON, 2008).

Como o número de organismos reconstruídos aumenta a cada dia, a necessidade de encontrar maneiras automatizadas, ou pelo menos semi-automatizada, para reconstruir redes metabólicas a partir da anotação do genoma está crescendo. Apesar da crescente experiência e conhecimento, ainda não se é capaz de reconstruir completamente de maneira automática as redes metabólicas de alta qualidade que podem ser usadas como modelos preditivos. Estudos recentes destacam os problemas atuais com anotações do genoma e bancos de dados, que fazem reconstruções automáticas e assim, necessitam de uma avaliação manual (FEIST *et al.*, 2009; REED *et al.*, 2006). Um número limitado de ferramentas de software e pacotes estão disponíveis (livres e comerciais), que visam a ajudar e facilitar o processo de reconstrução. O protocolo disponível atualmente com 96 etapas é um processo bem estabelecido para montagem, cura e validação da reconstrução metabólica (THIELE e PALSSON, 2010). Esse protocolo é combinado com as ferramentas computacionais, como o pacote de software de visualização e cálculo numérico, MATLAB[®]. O MATLAB[®] também é usado para codificar COBRA *toolbox*, que é um conjunto de funções utilizadas para facilitar a reconstrução, depuração e processo de cura manual (BECKER *et al.*, 2007).

Nessas redes metabólicas reconstruídas a modelagem baseada em restrição é preditiva na identificação de reações críticas, metabólitos e genes. Isto é extremamente útil para a determinação das necessidades metabólicas celulares e simula o comportamento fenotípico sob diferentes condições genéticas e ambientes fisiológicos (EDWARDS *et al.*, 2001; FONG e PALSSON, 2004; MAHADEVAN *et al.*, 2011; TRAWICK e SCHILLING, 2006).

Há uma extensa diversidade de redes metabólicas inexploradas nos genomas de vários microrganismos e ainda existe uma grande lacuna na compreensão da interligação entre o código genético e o fenótipo correspondente (BLANK e EBERT, 2013; MAHADEVAN *et al.*, 2011). O metabolismo celular é complexo e engloba as interações moleculares mais importantes da biologia de sistemas. O conhecimento sobre a via metabólica de um organismo permite a compreensão da fisiologia celular em função da regulação de seu metabolismo. A

quantificação e modulação de fluxos metabólicos é muito importante em escala científica e comercial, pois permite que a síntese de um determinado produto de interesse seja realizada de forma mais eficiente (BORDBAR e PALSSON, 2012; DUARTE *et al.*, 2007).

Gluconacetobacter hansenii (anteriormente denominado *Acetobacter xylinus*) (LISDIYANTI *et al.*, 2006) foi extensivamente caracterizada e tem sido utilizada para a síntese de celulose (BENZIMAN *et al.*, 1980; DEINEMA e ZEVENHUIZEN, 1971; ROSS *et al.*, 1991; SAXENA *et al.*, 1994). A *G. hansenii* produz celulose extracelular livre de lignina ou hemicelulose, tornando-o uma excelente fonte de celulose pura. A celulose bacteriana (CB) representa uma alternativa potencial para substituir a celulose de origem vegetal, em diversas aplicações industriais, assim como na indústria biomédica (ROSS *et al.*, 1991; SAXENA *et al.*, 1994). Esta bactéria foi sequenciada, mas seu genoma anotado ainda não está completo, havendo *gaps* na sequência (IYER *et al.*, 2010). Um grande número de estudos vem sendo realizados com o objetivo de melhorar o rendimento de celulose bacteriana por esta espécie para aplicação industrial (ROSS *et al.*, 1991).

A via de biossíntese da celulose envolve algumas enzimas individuais e um complexo de proteínas, onde a celulose sintase apresenta-se como a principal enzima desse complexo. Em diversas bactérias observou-se que os genes que codificam as proteínas desse complexo estão organizados em um operon *acs*. O processo da biossíntese da celulose inclui a síntese de uridina difosfato glicose (UDP-glicose), a qual é substrato da celulose sintase, seguida de sua polimerização com formação de cadeia β -1,4 glicana e a liberação de uridina difosfato (UDP). As associações de cadeias nascentes irão formar as fibras de celulose (BROWN *et al.*, 1976; WONG *et al.*, 1990; SAXENA *et al.*, 1994; SAXENA *et al.*, 1990;1991).

Ferramentas computacionais para prever fluxos em redes bioquímicas são aplicadas nas áreas que integram a biologia de sistemas, bioinformática e genômica (BEARD *et al.*, 2002; DUARTE *et al.*, 2007). Através da aplicação de técnica de análise de balanço de fluxo (ABF), a rede metabólica pode ser convertida para um modelo em escala genômica, permitindo uma avaliação qualitativa da relação entre os comportamentos genotípicos e fenotípicos, e uma estimativa global da distribuição de fluxo dentro do metabolismo de um organismo, que não pode ser medido usando um delineamento experimental padrão

(HENRY *et al.*, 2010; PRICE *et al.*, 2004; TERZER *et al.*, 2009; THIELE *et al.*, 2013; TRAWICK e SCHILLING, 2006).

Atualmente, uma ferramenta popular para investigar modelos metabólicos estequiométricos é a reconstrução e análise baseada em restrições (COBRA), uma caixa de ferramentas no ambiente MATLAB[®]. Esta conversão metabólica de uma reconstrução de um organismo em um modelo requer a imposição de restrições físico-químicas e ambientais para definir os limites dos sistemas. A conversão também inclui a transformação de uma lista de reações e metabólitos para um formato de matriz estequiométrica (BECKER *et al.*, 2007; SCHELLENBERGER *et al.*, 2011).

De uma maneira geral a engenharia metabólica é um campo multidisciplinar, onde a bioquímica prove a base das rotas metabólicas e informações sobre os mecanismos das reações bioquímicas, além da estequiometria, cinética e regulação (STEPHANOPOULOS, 1999; STEPHANOPOULOS *et al.*, 1998). A genética e a biologia molecular fornecem as ferramentas aplicadas na construção da base genética, um importante passo no estudo de controle de fluxo (BRINDLE, 1996). A fisiologia celular promove uma maior integração na visão das funções celulares metabólicas, através da definição da plataforma para estudo de taxas de reações e representação dos estados fisiológicos (ALBERTS *et al.*, 2002; BARABASI e OLTVAI, 2004).

E finalmente, a engenharia química é o mais apropriado caminho para aplicação de uma visão de engenharia no estudo de sistemas biológicos. Esta visão introduz os conceitos de integração, quantificação e relevância no estudo de sistemas biológicos. Mais especificamente, fornece as ferramentas e experiências para análise dos sistemas onde as condições de processos são limitadas, um campo no qual a engenharia química tem contribuído fortemente e frequentemente se superado (STEPHANOPOULOS e STAFFORD, 2002).

Esta dissertação apresenta o primeiro modelo metabólico em escala genômica da *Gluconacetobacter hansenii* ATCC 23769. Para a reconstrução metabólica partiu-se das informações que estão presentes no seu genoma e na literatura. A construção deste modelo metabólico envolve diversos passos, tais como, anotação funcional do genoma, a identificação das reações e determinação da sua estequiometria, construção da rede simplificada, adição de reações ausentes e biomassa e por fim a definição das restrições do modelo. Utilizando-se ferramentas computacionais os fluxos metabólicos do modelo foram simulados, para três condições ambientais diferentes baseadas em fonte

de carbono. Para compreenderem-se os efeitos de produção de celulose, crescimento bacteriano e a distribuição dos fluxos na rede foram utilizados como fontes de carbonos: glicose, glicerol e manitol.

1.2 Objetivos

1.2.1 Objetivo Geral

Analisar a biossíntese de celulose bacteriana a partir de diferentes fontes de carbono através da reconstrução metabólica da *Gluconacetobacter hansenii* ATCC 23769.

1.2.2 Objetivos Específicos

- Avaliar o protocolo desenvolvido para reconstruções metabólicas e propor modificações;
- reconstruir a rede metabólica simplificada do organismo para elucidar a via metabólica envolvida na biossíntese de celulose;
- converter a reconstrução em um modelo computacional, para utilizá-lo como uma ferramenta de análise das principais vias metabólicas envolvidas;
- analisar *in silico* a reconstrução metabólica simulando o modelo metabólico com três fontes de carbono distintas.

1.3 Justificativa e Motivação

A biologia de sistemas tenta entender sistemas biológicos através da geração de modelos preditivos que permitam que o comportamento do sistema possa ser modelado *in silico* (BORDBAR e PALSSON, 2012). Reconstruções metabólicas em escala genômica têm sido reconhecidas como uma ferramenta muito importante para uma variedade de aplicações que vão desde a engenharia metabólica a estudos evolutivos (BORDBAR e PALSSON, 2012; DE OLIVEIRA DAL'MOLIN e NIELSEN, 2013; DUROT *et al.*, 2009; FEIST e PALSSON, 2008; HENRY *et al.*, 2010).

Esta pesquisa se insere em uma linha estratégica de estudos do Grupo de Engenharia Genômica e Engenharia Tecidual da Universidade Federal de Santa Catarina, que pretende contribuir com os avanços na área de bioinformática e biologia de sistemas. Aplicando os princípios das ciências de informação e abordagens matemáticas e computacionais é possível tratar questões biológicas de uma maneira mais compreensível e útil.

Ao longo dos últimos anos, com o desenvolvimento da bioinformática, e das novas ferramentas para melhorar a extração da informação gênica a partir de grandes conjuntos de dados biológicos, tais como os perfis de transcrição, sequências e os fluxos metabólicos, permitiu-se realizar experimentos mais significativos, e extrair mais informações sobre o sistema.

A bactéria *Gluconacetobacter hansenii* ATCC 23769 é capaz de produzir biofilmes de celulose bacteriana, um dos produtos de grande interesse para a indústria, principalmente na área médica. As análises e ferramentas utilizadas e desenvolvidas neste trabalho permitiram a primeira reconstrução metabólica deste organismo.

1.4 Estrutura do documento

Esta dissertação está organizada da seguinte forma:

O Capítulo 1 apresenta a introdução sobre quais foram os fatores que motivaram a realização deste trabalho e suas justificativas, assim como os objetivos que levaram à execução deste trabalho;

O Capítulo 2 refere-se à fundamentação teórica e introduz os conceitos fundamentais, desde engenharia metabólica e biologia de sistemas, o uso da bioinformática para a reconstrução em escala genômica, as vias metabólicas da bactéria *Gluconacetobacter hansenii* com foco na biossíntese de celulose bacteriana;

O Capítulo 3 apresenta a metodologia utilizada com as principais ferramentas e recursos computacionais abordados para reconstruir uma rede metabólica;

O Capítulo 4 apresenta os resultados e discussões do trabalho.

O Capítulo 5, finalmente apresenta as conclusões e perspectivas futuras para a pesquisa.

CAPÍTULO 2

FUNDAMENTAÇÃO TEÓRICA

2.1 Engenharia metabólica e biologia de sistemas

A biologia de sistemas utiliza tecnologias de ponta para obter o máximo de informações possíveis sobre um sistema biológico e, em seguida, usa essas informações para construir modelos matemáticos que fornecem uma base sólida para a concepção da engenharia metabólica utilizando técnicas computacionais (DUROT *et al.*, 2009; LEE *et al.*, 2005).

No início da década de 90, o rápido avanço das técnicas de biologia molecular e, principalmente, da tecnologia do DNA recombinante, aumentou muito as aplicações desta tecnologia em vários setores da biotecnologia como, por exemplo, no melhoramento de organismos utilizados na obtenção de produtos de interesse. Nesta época, o termo engenharia metabólica foi introduzido por Bailey (1984) que o definiu da seguinte forma: “engenharia metabólica é o melhoramento das atividades celulares através da manipulação de funções enzimáticas, de transporte e reguladoras das células, fazendo uso da tecnologia do DNA recombinante”.

Stephanopoulos e colaboradores (1998) propuseram a seguinte definição em seu livro texto: “engenharia metabólica é o melhoramento dirigido da formação de produto ou de propriedades celulares através da modificação de reações bioquímicas específicas ou da introdução de novas reações fazendo uso da tecnologia do DNA recombinante”. Observa-se assim que a engenharia metabólica envolve a análise e modificação de vias metabólicas, e que o resultado final é sempre o metabolismo modificado com um determinado propósito.

Neste ponto, observa-se que há uma interface entre engenharia metabólica e bioinformática, à medida que as análises do genoma, transcriptoma, proteoma e metaboloma fornecerão informações importantes para a análise do metabolismo (EDWARDS e PALSSON, 1998; BASSETT *et al.*, 1999; PANDEY e MANN, 2000). Assim, a engenharia metabólica serve como uma abordagem integrada, que combina trabalho computacional e experimental. A combinação de métodos analíticos para quantificar fluxos e seus controles com técnicas de biologia molecular na implementação de modificações genéticas é a

essência da engenharia metabólica (PATIL *et al.*, 2004; STEPHANOPOULOS, 1999).

Segundo Stephanopoulos e colaboradores (1998), uma das mais importantes contribuições da engenharia metabólica é, provavelmente, a ênfase empregada no fluxo metabólico e o seu controle sob condições *in vivo*. O campo da engenharia metabólica está emergindo porque seu foco não está em reações isoladas, mas na inter-relação das reações em redes (KOFFAS *et al.*, 1999).

2.1.1 Abordagens na engenharia metabólica para reconstrução de redes

Os primeiros trabalhos que fizeram uso de um modelo estequiométrico para a interpretação de dados experimentais datam de décadas passadas (AIBA e MATSUOKA, 1979; PAPOUTSAKIS e MEYER, 1985). No entanto, a técnica de Análise de Balanço de Fluxo foi introduzida como uma técnica de análise do metabolismo e sistematizada através do uso de álgebra linear e cálculo matricial anos mais tarde (GOMBERT, 2001; VALLINO e STEPHANOPOULOS, 1990).

Nissen e colaboradores (1997) desenvolveram um modelo estequiométrico para descrever o metabolismo anaeróbico de *Saccharomyces cerevisiae* durante o crescimento em meio definido. O modelo foi usado para calcular os fluxos intracelulares baseados no consumo de substrato do meio, na excreção de produtos da célula e na velocidade de formação de biomassa. O modelo foi usado para calcular as velocidades de produção do malato e fumarato.

Takaç e colaboradores (1998) propuseram uma rede metabólica para uma bactéria ácida glutâmica a qual foi usada como base estequiométrica no modelo de balanço de fluxos para produção de L-glutamato. As rotas metabólicas teóricas ligadas à otimização da produção de glutamato foram determinadas para várias velocidades específicas de crescimento celular onde foram obtidas as variações de fluxos.

Naeimpoor e Mavituna (2000) utilizaram análise de fluxos metabólicos para estudar o metabolismo de *Streptomyces coelicolor*. Nesse trabalho, uma rede de reações bioquímicas foi construída baseada em dados da literatura. Um modelo metabólico foi proposto e resolvido via programação linear por definição de uma função objetivo, onde os

fluxos internos e os fluxos externos, não medidos experimentalmente, foram calculados. Foi investigado o efeito de vários parâmetros do modelo na distribuição dos fluxos.

Rossa e colaboradores (2002) analisaram o metabolismo do carbono no crescimento de *Streptomyces lividans* e na produção de antibióticos, sob limitações de fosfato em cultivos contínuos. Aplicaram análise de fluxos, com base num modelo estequiométrico, para analisar a relação entre a produção do antibiótico, o acúmulo de biomassa, e a distribuição do fluxo de carbono através da maioria das vias metabólicas.

Sanchez e colaboradores (2006) analisaram extensivamente os fluxos metabólicos de várias linhagens mutantes de *Escherichia coli* expressando a piruvato carboxilase de *Lactococcus lactis* para a produção de succinato. Nesse estudo, uma análise de redistribuição de fluxos metabólicos para uma biblioteca de mutantes foi realizado e as diferenças entre os fluxos dessas linhagens foram comparadas, incluindo as bifurcações dos principais nós, com o objetivo de elucidar a importância da fermentação e da via glicoxilato na produção de succinato. Um modelo estequiométrico foi implementado com sucesso para estimar os fluxos intracelulares em várias linhagens produtoras.

Oliveira (2008) construiu a partir de dados da literatura, uma rede metabólica da *Clostridium acetobutylicum* envolvendo todas as reações usadas para a produção de ácidos, solventes e hidrogênio molecular a partir de dois substratos distintos: glicose e glicerol. No total, a rede possui 25 metabólitos e 20 reações. Este trabalho reúne ferramentas computacionais de biologia de sistemas num aplicativo chamado GEnSys (*Genomic Engineering System*), desenvolvido no Laboratório de Tecnologias Integradas (InteLab/UFSC), usando o MATLAB® (Mathworks, Natick, MA, USA). O GenSys é formado por um conjunto de módulos interdependentes que permite a análise e simulação de redes de reações bioquímicas. Foram efetuadas duas abordagens diferentes para o estudo da rede metabólica: análise de balanço de fluxo e a determinação dos modos elementares, para elucidar as possíveis vias de produção de H₂ por esta rede. Através da ABF foram realizadas diferentes simulações que mostraram que este modelo matemático pode ser usado para estudar *in silico* possíveis formas de melhorar a produção de H₂, ou até mesmo de solventes.

Castro (2010) reportou o potencial de produção de hidrogênio por *C. acetobutylicum* ATCC 824, aproveitando um resíduo industrial da produção de biodiesel, o glicerol. Ferramentas de engenharia

metabólica foram utilizadas para explorar a capacidade da *C. acetobutylicum* de crescer e produzir hidrogênio a partir de glicerol como única fonte de carbono. Foi utilizado um modelo metabólico simplificado de 20 reações e 25 metabólitos. Com o resultado da Análise de Balanço de Fluxo (ABF), foi possível prever crescimento e produção de biohidrogênio.

Klanchui e colaboradores (2012) apresentaram a análise da rede metabólica em escala genômica de *Spirulina platensis* C1, iAK692, e suas propriedades topológicas, capacidades e funções metabólicas. A rede foi reconstruída a partir do genoma anotado da *S. platensis* C1 usando software Pathway Tools (SRI International, Menlo Park, CA, USA). A cura manual foi realizada com base em uma combinação de informações genômica, bioquímica e fisiológica. Foi simulado o modelo matemático em condições de crescimento autotrófico, heterotrófico e mixotrófico usando COBRA *toolbox* (UC San Diego).

McCloskey e colaboradores (2013) elucidaram o modelo em escala genômica do metabolismo da bactéria *Escherichia coli* K-12 que está em desenvolvimento há mais de uma década e agora está em ampla utilização. Os estudos da rede em escala genômica da *E. coli* tem sido focado em diferentes aplicações: (1) a engenharia metabólica, (2) predição de fenótipos celulares, (3) estudos de análise das propriedades biológicas da rede, (4) processos evolutivos, e (5) modelos de interações entre espécies.

2.2 Vias metabólicas

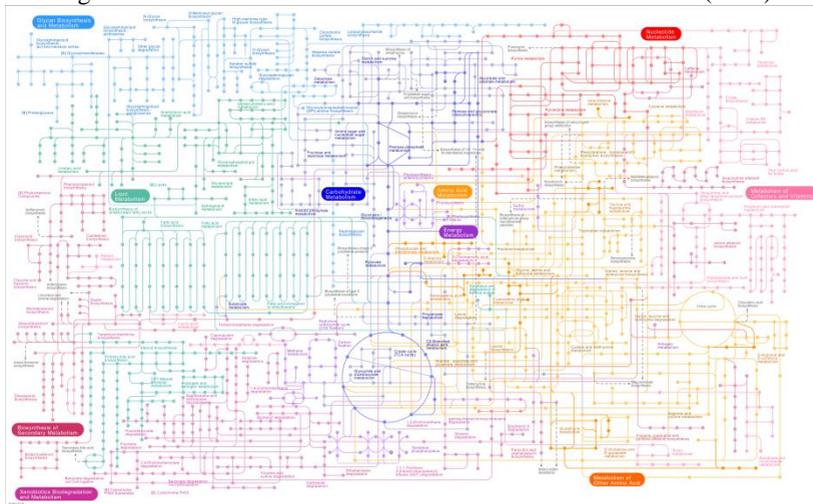
Os caminhos pelos quais a célula converte compostos externos que entram pela membrana celular (proteínas, ácidos nucleicos e carboidratos) em compostos celulares que serão de seu interesse, são chamados de vias metabólicas do organismo. Estes processos são usados inclusive para a obtenção da energia necessária para a sobrevivência e duplicação da célula (KIM e GADD, 2008).

As vias metabólicas são processos bioquímicos nos quais as enzimas executam o papel de alterar a velocidade dos processos, sem, no entanto, ser consumida ou produzida no processo. O mecanismo central nos processos metabólicos é controlado pela regulação da atividade enzimática e expressão gênica. A maquinaria celular executa, de forma coordenada, a complexa série de interações bioquímicas, que darão às células condições de se adaptarem a variações nutricionais e ambientais (ALBERTS *et al.*, 2002). O metabolismo compreende dois

tipos de transformações químicas: vias catabólicas que quebram vários substratos em metabólitos comuns e vias anabólicas que sintetizam coletivamente aminoácidos, ácidos graxos, ácidos nucléicos e outros componentes. Durante esses processos, uma troca complexa de vários grupos químicos potenciais de redução e oxidação ocorre através de um conjunto de moléculas portadoras (por exemplo, ATP e NADH) (PALSSON, 2006).

As vias metabólicas colocam os genes num contexto biológico amplo, e são primordiais para o entendimento dos processos complexos codificados pelo genoma. O desafio é caracterizar as múltiplas funções bioquímicas num contexto genômico. Através da análise dessas vias, também se pode observar vias que, apesar de serem amplamente conhecidas e fazerem parte das funções básicas do organismo, possuem muitos passos desconhecidos; isto pode ser devido a erros de anotação ou mesmo refletir uma variação na via deste organismo (KARP *et al.*, 1999).

Segundo Lange e Ghassemian (2005), a representação visual de vias ajuda a entender a complexa relação entre os compostos das redes metabólicas e disponibilizam um recurso para a integração dos conjuntos de dados dos transcriptomas, proteomas e metabolomas. Para compreender as células, é necessário o uso da matemática, da computação e da informação quantitativa. Por meio de tais métodos, utilizando o conhecimento das sequências genômicas, pode-se listar os genes e as proteínas em uma célula, e começar a descrever a rede de interações entre eles (ALBERTS *et al.*, 2002; BARABASI e OLTVAI, 2004). Uma rede de vias metabólicas pode ser muito complexa, conforme Figura 1, que apresenta as vias metabólicas da bactéria *E.coli*.

Figura 1 – Vias metabólicas da *Escherichia coli*. Fonte: KEGG (2013)

2.3 Reconstrução metabólica em escala genômica

A reconstrução metabólica foi introduzida há mais de 18 anos como um método de interpretação da sequência do genoma (PALSSON, 2006). Ela usa uma variedade de abordagens de bioinformática e integra informações com os dados de vias metabólicas disponíveis do organismo, para estabelecer quais vias são susceptíveis de estarem presentes no organismo de interesse (PALSSON, 2006).

Embora existam excelentes fontes de dados, métodos e ferramentas desenvolvidas para a reconstrução metabólica, o processo ainda depende fortemente da cura manual, devido tanto à complexidade e à diversidade do metabolismo microbiano, quanto à inconsistência nos bancos de dados, levando a um desempenho inadequado das ferramentas automatizadas. No entanto, as vantagens de ter uma rede metabólica totalmente reconstruída superam os problemas encontrados durante o processo da reconstrução. A reconstrução metabólica conduz frequentemente à descoberta de novas funções, vias e traços fenotípicos, enquanto que a rede reconstruída pode ser usada para a modelagem matemática do metabolismo microbiano e previsões quantitativas para a engenharia genética (COVERT *et al.*, 2001; FEIST *et al.*, 2009; STEPHANOPOULOS *et al.*, 1998; HENRY *et al.*, 2010).

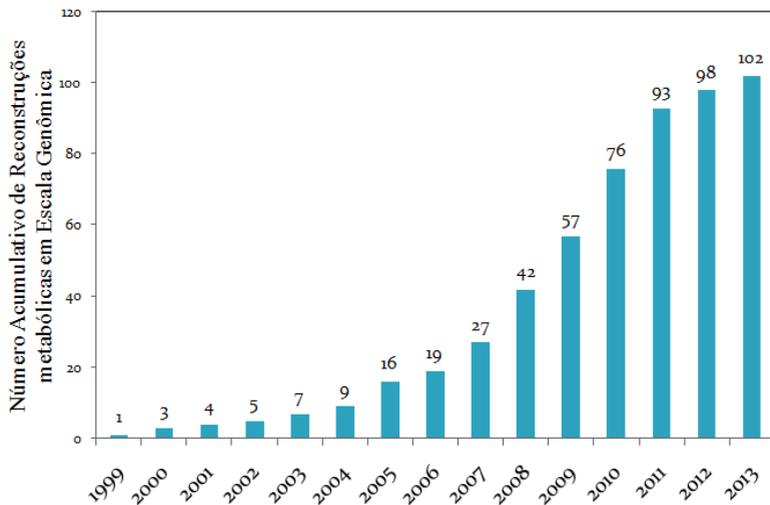
A construção de uma rede metabólica em escala genômica é um processo longo e consiste em quatro passos principais, que requer o uso de diferentes tipos de dados biológicos (THIELE e PALSSON, 2010). Um protocolo de reconstrução metabólica em escala genômica bem estabelecido consiste em 96 etapas, desde o genoma anotado até a validação da rede (THIELE e PALSSON, 2010). Em geral, quanto mais informações disponíveis sobre a fisiologia, bioquímica e genética do organismo, melhor a capacidade de previsão dos modelos reconstruídos (ORTH *et al.*, 2011; THORLEIFSSON e THIELE, 2011).

As funções das redes reconstruídas são definidas pelas interconexões de suas partes. Uma vez que estas ligações envolvem reações químicas, elas podem ser descritas por uma relação estequiométrica. O terceiro passo da reconstrução, portanto, é converter essas reações para um modelo matemático com o intuito de se realizar as simulações necessárias. A matriz estequiométrica que contém todas essas relações em uma rede é, portanto, uma representação matemática concisa das redes reconstruídas. O último passo é integrar dados experimentais ao modelo, permitindo possíveis descobertas biológicas e o seu refinamento (PALSSON, 2006; PATIL *et al.*, 2004; THIELE e PALSSON, 2010).

Uma reconstrução metabólica fornece uma plataforma matemática estruturada para se compreender a biologia de sistemas de vias metabólicas de um organismo (THIELE e PALSSON, 2010). Existem diferentes tipos de reconstrução representando vários tipos de redes biológicas (metabólica, reguladora, transcrição/translação), sendo que a reconstrução metabólica tem sido extensivamente utilizada (ORTH *et al.*, 2010). As reconstruções são frequentemente atualizadas com outros genes, proteínas, metabólitos e reações (DUARTE *et al.*, 2007; EDWARDS e PALSSON, 2000). Com a crescente influência dessas reconstruções em pesquisa biomédica e biológica, é essencial compreender sua utilização e o futuro potencial de análises de sistemas baseados no metabolismo (MCCLOSKEY *et al.*, 2013; OBERHARDT *et al.*, 2009).

De acordo com a Figura 2 é possível visualizar o aumento do número de reconstruções metabólicas ao longo dos anos. Esse crescimento reflete os avanços alcançados por meio de novas ferramentas computacionais disponíveis e genomas publicados.

Figura 2 – Número acumulativo de reconstruções metabólicas. Lista desenvolvida e atualizada pelo grupo *System Biology Research*
Disponível em: <http://gcrg.ucsd.edu/Home>



Assim, o processo de reconstrução pode ser visto como um processo iterativo em que as informações obtidas a partir de várias fontes são combinadas para construir um conjunto preliminar de reações e restrições (PRICE *et al.*, 2004). As reações e conjuntos de restrições podem ser integrados num modelo estequiométrico que pode ser usado para a simulação utilizando diversas abordagens de métodos baseados em restrição, dentre elas, a Análise de Fluxo Metabólico e Análise de Balanço de Fluxo (BEARD *et al.*, 2002; BLANK e EBERT, 2013; ORTH *et al.*, 2010; PALSSON, 2006).

2.3.1 Análise de Fluxo Metabólico (AFM)

A Análise de Fluxo Metabólico (AFM), em inglês, *Metabolic Flux Analysis (MFA)*, usa dados medidos experimentalmente para estimar o estado metabólico das células em condições específicas. Tem sido utilizada para estudar a fase de crescimento exponencial e estados estacionários em processos de fermentação contínua. Existe também interesse na utilização de AFM para monitorar variáveis no tempo,

particularmente fluxos em ambientes industriais (HERWIG e VON STOCKAR, 2002; TAKIGUCHI *et al.*, 1997).

De um modo geral, a Análise de Fluxo Metabólico (AFM) combina um conjunto de fluxos medidos (muitas vezes os extracelulares), com um modelo baseado em restrição para obter uma estimativa de todos os fluxos (velocidade de reações bioquímicas). Isto resulta num fluxo metabólico que representa o estado estacionário em que cada reação ocorre na rede. Este padrão de fluxo informa sobre a contribuição de cada reação para os processos metabólicos gerais de utilização de substrato e formação de produto. O sistema precisa ser determinado e redundante. Caso contrário, a estimativa fornecida será uma das muitas possíveis (LLANERAS, 2010).

2.3.2 Análise de Balanço de Fluxo (ABF)

Análise de Balanço de Fluxo (ABF), em inglês, *Flux Balance Analysis (FBA)*, é uma poderosa ferramenta para identificar as distribuições dos fluxos no regime estacionário e capacidades metabólicas de redes bioquímicas (ORTH *et al.*, 2010). Usando ABF, relações quantitativas globais podem ser previstas entre as taxas de entrada de nutrientes, as taxas de subprodutos de saída, e a taxa de crescimento (EDWARDS *et al.*, 2001; PRICE *et al.*, 2004; STEPHANOPOULOS *et al.*, 1998).

Modelos de balanço de fluxos do metabolismo usam estequiometria das vias metabólicas, as demandas metabólicas de crescimento e prevê a distribuição de fluxo metabólico e crescimento celular em condições ambientais específicas (VARMA e PALSSON, 1993). Em outras palavras, o método de ABF calcula o rendimento do crescimento máximo alcançável no modelo metabólico, maximizando o fluxo da reação de biomassa (que representa a taxa de crescimento), dado um conjunto de taxas de entrada delimitadas para substratos externos (PATIL *et al.*, 2004; VARMA e PALSSON, 1994).

Cortassa e colaboradores (2002) sugerem dois tipos de função objetivo: (i) a maximização do crescimento de biomassa, metodologia amplamente utilizada quando o modelo metabólico é complexo demais para a obtenção dos dados experimentais necessários a um sistema determinado, e (ii) a minimização do gasto energético, que pode ser obtida pela minimização das taxas das reações em que ocorre produção de ATP.

Uma vez que os fluxos intracelulares são difíceis de medir, eles são muitas vezes obtidos por métodos computacionais. Durante o cálculo dos fluxos, alguns dados experimentais reais, tais como a absorção de substrato e taxas de excreção do produto, são muitas vezes fornecidos como restrições para tornar os fluxos calculados mais realistas (LEE *et al.*, 2005; SCHILLING *et al.*, 2000).

As etapas críticas da ABF incluem a reconstrução de uma rede metabólica, seguida do balanço de massa, a imposição de restrições, a escolha de uma função objetivo adequada (biologicamente relevante) e otimização (linear) (ORTH *et al.*, 2010; RAMAN e CHANDRA, 2009). A ABF tem sido aplicada para: (1) calcular o estado ótimo para o crescimento sob diferentes condições de cultura, (2) aumentar a produção de ATP ou NADH e (3) maximizar a produção de metabólitos alvo (RAMAKRISHNA *et al.*, 2001; VARMA e PALSSON, 1994; LEE *et al.*, 2006; RAMAN e CHANDRA, 2009).

A *Escherichia coli* tem sido o microrganismo mais bem estudado, devido à quantidade considerável de dados experimentais (EDWARDS *et al.*, 2001; REED *et al.*, 2003; FEIST *et al.*, 2007). Outros modelos de FBA foram desenvolvidos, por exemplo, para *Haemophilus influenzae* (EDWARDS e PALSSON, 1999), *Helicobacter pylori* (SCHILLING *et al.*, 2002), *Saccharomyces cerevisiae* (FORSTER *et al.*, 2003), *Methano sarcina barkeri* (FEIST *et al.*, 2006) e *Synechocystis sp.* (MONTAGUD *et al.*, 2010).

Como resultado desses esforços, muitas aplicações de análise de balanço de fluxo têm sido investigadas. Nos últimos anos, as primeiras aplicações médicas foram realizadas. Thiele e colaboradores (2005) utilizaram ABF para investigar a rede metabólica da mitocôndria humana e para avaliar o efeito de potenciais tratamentos contra doenças. A ABF tem sido igualmente aplicada para otimizar o metabolismo de hepatócitos em cultura utilizados em dispositivos de fígado bioartificial (NOLAN *et al.*, 2006; SHARMA *et al.*, 2005).

2.3.3 Fluxos metabólicos e matriz estequiométrica

A teoria que se aplica para a determinação dos fluxos metabólicos baseia-se na simples conservação de massa em um sistema fechado. No caso uma célula ou um compartimento celular fechado com metabólitos especifica-se a reversibilidade das reações e quais metabólitos são considerados como internos e externos. A Equação 1

descreve a conservação de massa de metabólitos em um sistema de volume definido:

$$\frac{dC}{dt} = S \cdot v - \mu C \quad (1)$$

Na equação acima, C (mol/L) é um vetor da concentração de m metabólitos internos; v ((mol/L) /h) é o vetor de todos os fluxos de reação na rede, de n reações que convertem metabólitos; S é a matriz estequiométrica de dimensões $m \times n$; e μ (1/h) é o grau específico de diluição associado à mudança no volume de um sistema, o que é muito importante considerar no modelo, pois o grau de diluição afeta diretamente as velocidades de reação.

Como em uma célula o grau de diluição é muito baixo quando comparado com os graus de reação, as mudanças de volume no sistema são consideradas negligenciáveis. Temos, portanto, a Equação 2 mais simplificada:

$$\frac{dC}{dt} = S \cdot v \quad (2)$$

Em regime estacionário, que é o que se considera na análise de fluxo metabólico, não há acúmulo de metabólitos, e, portanto suas concentrações, bem como a população bacteriana, tornam-se constantes, fazendo com que $dC/dt = 0$ assim, temos a Equação 3:

$$S \cdot v = 0 \quad (3)$$

Uma reconstrução pode ser convertida para um formato matemático, representada pela matriz estequiométrica (PALSSON, 2006). A matriz estequiométrica, $S_{m \times n}$, consiste de m linhas de metabólitos e n colunas de reações.

Utilizando programação linear para resolver o modelo estequiométrico, o vetor de fluxo pode ser encontrado através da especificação de uma função objetivo (Z) que pode ser minimizada ou maximizada. As distribuições do fluxo metabólico da rede são estimadas sob dadas condições.

Geralmente, o objetivo é maximizar a velocidade de crescimento ou otimizar (max/min) alguns dos fluxos da rede metabólica. Esse sistema de equações, na grande maioria das vezes, é indeterminado e possui várias soluções, o que implica na formação de um espaço de soluções. Podem-se adicionar mais restrições a este espaço especificando-se valores máximos ou mínimos para os fluxos da rede:

$$v_{\min j} \leq v_j \leq v_{\max j}$$

As restrições para os limites inferiores (*lower bound*) e limites superiores (*upper bound*) das reações reversíveis e irreversíveis, ficam definidas como:

$$\begin{aligned} -\infty \leq v_j \leq \infty & \text{ Reversíveis} \\ 0 \leq v_j \leq \infty & \text{ Irreversíveis} \end{aligned}$$

Para reações irreversíveis, o limite inferior é definido como zero e para reações reversíveis ambos os valores são tipicamente definidos para valores arbitrariamente grandes. Dessa forma a direção das reações (restrição termodinâmica) é respeitada.

Assim, um modelo metabólico em escala genômica é basicamente uma representação estequiométrica de todas as possíveis reações metabólicas na célula. A estequiometria e a reversibilidade de cada reação, juntamente com a suposição de regime estacionário para os metabólitos internos, permitem definir uma região de distribuições de fluxo desejada (KLANCHUI *et al.*, 2012; LIU *et al.*, 2010; STEPHANOPOULOS *et al.*, 1998).

2.4 Ferramentas computacionais

Em engenharia metabólica são utilizadas diversas ferramentas computacionais para a construção de modelos *in silico* (LEE *et al.*, 2005). Algumas das principais ferramentas *online* e/ou disponíveis publicamente são descritas a seguir, além do *toolbox* GEnSys, desenvolvido pelo nosso grupo.

2.4.1 Bancos de dados e ferramentas genômicas

Devido à imensa quantidade de dados gerados em inúmeros laboratórios de todo o mundo, faz-se necessário organizá-los de maneira acessível, de modo a evitar redundância na pesquisa científica e possibilitar a análise por um maior número possível de cientistas. A construção de bancos de dados para armazenamento de informações de sequências de DNA e genomas inteiros, proteínas e suas estruturas tridimensionais, bem como vários outros produtos da era genômica, têm sido um grande desafio, mas simultaneamente são extremamente importantes para o desenvolvimento da nova biotecnologia.

2.4.1.1 NCBI

O NCBI, ou Centro Nacional de Informação Biotecnológica dos Estados Unidos, é considerado a base de dados central sobre informações genômicas, talvez a mais importante do mundo. Vários outros bancos de dados similares estão distribuídos por países da Europa e Japão, mas todos trocam dados em um intervalo de 24 horas com o NCBI. O *GenBank* é o principal banco de dados do NCBI e armazenam todas as sequências disponíveis publicamente de DNA (de sequências pequenas a genomas inteiros), RNA e proteínas. Além do *GenBank*, que coleta todas as entradas de sequências, outros bancos do NCBI apresentam as informações organizadas de diferentes maneiras. Por exemplo, o UniGene agrupa todas as sequências parciais do transcriptoma de um organismo em aglomerados ou *clusters*, onde cada aglomerado representa a sequência concisa de um gene (NCBI, 2013).

2.4.1.2 *Integrated Microbial Genomes* (IMG)

O sistema integrado de genomas microbianos, *Integrated Microbial Genomes* (IMG) é um sistema de navegação de genomas e uma plataforma de anotação desenvolvida pelo Departamento de Energia dos EUA (DOE) - *Joint Genome Institute*. O IMG contém todos os *drafts* e genomas microbianos completos sequenciados pelo DOE-JGI integrados com outros genomas disponíveis publicamente (incluindo arqueas, bactérias, eucariotos, vírus e plasmídeos) (IMG, 2013; MARKOWITZ *et al.*, 2012). O IMG fornece aos usuários um conjunto de ferramentas para a análise comparativa de genomas

microbianos ao longo de três dimensões: genes, genomas e funções. Os usuários podem selecioná-los e transferi-los para diferentes análises comparativas com base em uma variedade de critérios. O IMG inclui ainda uma conexão da anotação do genoma que integra a informação a partir de várias ferramentas, incluindo KEGG (KEGG, 2013), Pfam (PFAM, 2013), InterPro (INTERPRO, 2013) e GeneOntology (GENEONTOLOGY, 2013), entre outros. Neste trabalho, o IMG foi incluído como uma importante alteração do protocolo de reconstrução metabólica.

2.4.1.3 Kyoto Encyclopedia of Genes and Genomes (KEGG)

O aumento da quantidade de informação genômica e molecular é a base para a compreensão dos sistemas biológicos mais complexos, incluindo célula e organismos, e suas interações com o meio ambiente, bem como para aplicações práticas médicas, industriais e outras. O recurso KEGG (KEGG, 2013) proporciona uma base de conhecimento de referência para a ligação de genomas aos sistemas biológicos. O KEGG (*Kyoto Encyclopedia of Genes and Genomes*) é um recurso de base de dados integrado, composto por 16 bases de dados principais; em geral, são categorizadas em sistemas de informação, informação genômica e informação química (KANEHISA *et al.*, 2006, 2010).

2.4.2 Pathway Tools

Pathway Tools é um software para gestão, análise e visualização de coleções integradas de genoma, vias e dados reguladores. Ele suporta o armazenamento e manipulação de informações de regulação através de uma variedade de estratégias. O software está em desenvolvimento contínuo dentro do Grupo de Pesquisa em Bioinformática na SRI International (Menlo Park, CA, USA) desde o início da década de 1990 (Karp *et al.*, 2002; Paley *et al.*, 2012). O Pathway Tools é um software que suporta a criação de novos PGDBs (*Pathway/Genome Database*), utilizando o componente *PathoLogic*. Um PGDB como o EcoCyc integra a evolução do nosso entendimento dos genes, proteínas da rede metabólica, e rede genética de um organismo (KARP *et al.*, 2002; KESELER *et al.*, 2005).

Os quatro principais componentes do Pathway Tools são: 1) O componente *PathoLogic* suporta a criação de novos PGDBs do genoma anotado de um organismo, contendo genes, proteínas, reações

bioquímicas, e prevê caminhos metabólicos do organismo; 2) O *Pathway/Genome Navigator* fornece consulta, visualização e serviços Web para PGDBs; 3) O *Pathway/Genome Editors* suporta atualização interativa dos PGDBs. Finalmente, 4) O *Pathway Tools Ontology* define o esquema de PGDBs (KARP *et al.*, 2002, 2010). Neste trabalho, utilizamos a versão 16.5 do Pathway Tools.

2.4.3 Model SEED e RAST Server

O Model SEED é um recurso para a geração, otimização, cura e análise de modelos metabólicos em escala genômica. É uma ferramenta livre e está disponível online (http://www.theseed.org/wiki/Main_Page). O primeiro passo da reconstrução é a partir do genoma anotado (formato FASTA ou GenBank) no RAST (*Rapid Annotation using Subsystem Technology*) Server criando um banco do organismo com diversas informações, desde reações, vias, metabólitos ou genes. Uma das maneiras de exportar seu conteúdo é o formato SBML que serve de entrada para o COBRA *toolbox*.

2.4.4 System Biology Markup Language (SBML)

Existem comunidades de especialistas de várias áreas que se reúnem para discutir, planejar, executar e curar a reconstrução de uma rede metabólica. Esse processo tem o potencial de agregar o conhecimento de muitas pessoas, e assim permitir a criação de uma rede metabólica de alta qualidade e disponibilizar os dados para a comunidade, se possível em um formato padrão. A criação do padrão *Systems Biology Markup Language* (SBML) foi um avanço nessa direção, pois muitas redes metabólicas são disponibilizadas nesse formato e existe uma série de ferramentas capaz de processá-lo.

Um modelo SBML é um arquivo XML estruturado com três componentes básicos:

- 1) definição de compartimentos de reação (nome, id);
- 2) definição de espécie de metabólitos (nome, id, compartimento);
- 3) definição de reações (nome, id, reagentes, produtos, limite superior e inferior).

2.4.5 Implementações em MATLAB®

O software MATLAB® é uma ferramenta comercial muito utilizada em ciências físicas e engenharias e está disponível em muitos centros de pesquisa. A principal vantagem do MATLAB® sobre outras linguagens de programação, como C++, Java e outras, é que este possui uma linguagem de *script* e ambiente de desenvolvimento interativo, o que permite ao usuário editar, facilmente, as funções escritas nesta linguagem de acordo com as suas necessidades (ULLAH *et al.*, 2006). O MATLAB® tem outras funcionalidades, como: acesso rápido às ferramentas de análise de dados e otimização; funções para visualização de dados (amplo recurso para plotagem de gráficos) e recursos para se criar aplicativos contendo uma interface gráfica. O MATLAB® possui ainda vários *toolboxes* específicos para modelagem e simulação em várias áreas das engenharias e ciências físicas. Além disso, o MATLAB® contém funções para a representação e manipulação de vetores e matrizes multidimensionais. Há versões do MATLAB® para vários sistemas operacionais, como MS-Windows, Linux, OSX entre outros. Com estas características é fácil escrever *scripts* relativamente complexos de forma rápida e compacta.

2.4.5.1 COBRA *toolbox*

A abordagem COBRA (em inglês: *constraint-based reconstruction and analysis*) tem sido amplamente utilizada para analisar as reconstruções de redes metabólicas (Palsson, 2006; Price *et al.*, 2004). O COBRA é um pacote de software em execução no ambiente MATLAB®, que permite a previsão quantitativa do comportamento celular, usando uma abordagem baseada em restrições e a Análise de Balanço de Fluxos (ABF). O problema de programação linear da análise de balanço de fluxo é resolvido usando GLPKMEX, uma interface MEX do pacote glpk (GNU *linear programming kit*) (BECKER *et al.*, 2007; SCHELLENBERGER *et al.*, 2011).

Especificamente, este software permite cálculos preditivos do comportamento de crescimento tanto em regime estacionário quanto e em dinâmico, transiente. O conhecimento das vias metabólicas de um organismo é um constante desafio para os cientistas devido à complexidade nos processos celulares (BECKER *et al.*, 2007; SCHELLENBERGER *et al.*, 2011).

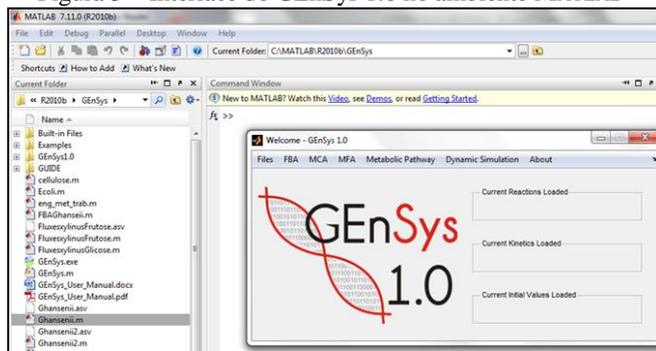
2.4.5.1.1 rBioNet

A extensão rBioNet exige o COBRA *toolbox* como requisito e está inserida no ambiente de programação MATLAB® (BECKER *et al.*, 2007; THORLEIFSSON e THIELE, 2011). Esta ferramenta ajuda a carregar as reações e metabólitos do modelo. Reconstruções bioquímicas, montadas com rBioNet, podem ser convertidas em modelos matemáticos, e podem então ser investigadas utilizando métodos baseados em restrição (THORLEIFSSON e THIELE, 2011; Palsson, 2006). A extensão possui uma interface intuitiva, em três partes: i) um criador de metabólitos associado ao banco de dados; ii) um criador de reação, e iii) um criador de reconstrução (THORLEIFSSON e THIELE, 2011; THIELE e PALSSON, 2010).

2.4.5.2 GEnSys

Com o crescente avanço da informação genômica, perfis metabólicos e projetos de proteômica, engenheiros químicos se deparam com centenas de vias bioquímicas do genoma e redes de sinalização. Os cálculos clássicos de engenharia metabólica foram ampliados nos últimos anos para incluir redes de regulação e análises de vias complexas, bem como para descrever a dinâmica do sistema. O *Genomic Engineering Systems* (GEnSys) foi implementado como um conjunto de ferramentas do MATLAB®, criado pelo grupo Intelab/UFSC e compreende vários módulos interdependentes que permitem a análise e simulação de redes em um ambiente de fácil utilização.

Figura 3 – Interface do GEnSys 1.0 no ambiente MATLAB®



Várias abordagens têm sido propostas para estudar redes metabólicas, tais como Análise de Fluxo Metabólico (AFM) e Análise de Balanço de Fluxo (ABF). Infelizmente, esses métodos para análise *in silico* destas redes não estão disponíveis em fácil entendimento para os pesquisadores, particularmente aqueles que não têm experiência em ferramentas analíticas e computacionais (OLIVEIRA, 2008).

Existem diversos módulos do GENSys para a modelagem e simulação de redes metabólicas e reguladoras. A principal característica desta organização em módulos é que eles podem ser herdados por outros módulos e estes podem usar todos os dados e métodos do módulo base.

Cada módulo pode usar os métodos e variáveis do módulo no qual está ligado, dentre eles, o módulo para a análise de balanço de fluxo e o módulo para a análise de fluxo metabólico. Existe também módulos que são ferramentas (*toolboxes*) auxiliares, como o COBRA (*constraint-based reconstruction and analysis*) que faz análise de balanço de fluxo estático e dinâmico (OLIVEIRA, 2008).

2.5 *Gluconacetobacter hansenii* ATCC 23769

A *Gluconacetobacter hansenii* (anteriormente denominada *Acetobacter xylinus*) (LISDIYANTI *et al.*, 2006), tem sido extensivamente utilizada para síntese de celulose (BENZIMAN *et al.*, 1980; DEINEMA e ZEVENHUIZEN, 1971; ROSS *et al.*, 1991; SAXENA *et al.*, 1994). Embora Brown tenha identificado a bactéria *Gluconacetobacter xylinus* (*G. xylinus*), já em 1886, como uma bactéria produtora de celulose (BROWN, 1886), até hoje esta espécie é reconhecida como a única capaz de produzir celulose em escala comercial (BROWN *et al.*, 1976; KLEMM *et al.*, 2006). Esta bactéria foi reclassificada de *Gluconacetobacter xylinus* ATCC 23769 para *Gluconacetobacter hansenii* ATCC 23769 (YAMADA *et al.*, 1997).

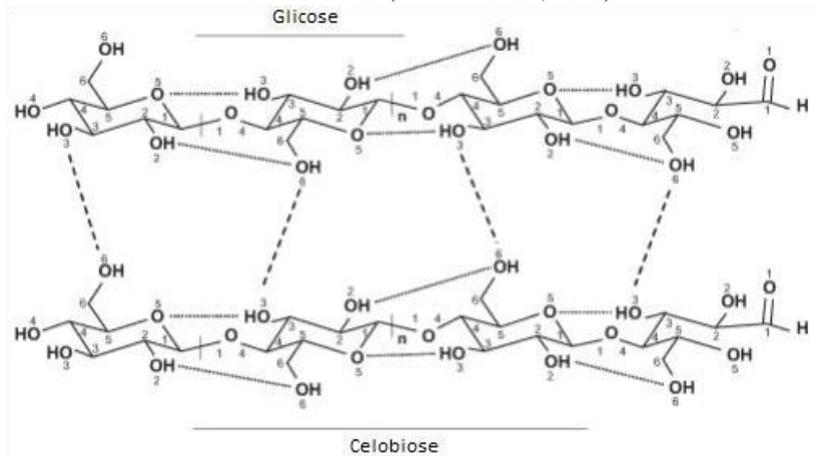
A *G. hansenii* é uma bactéria Gram-negativa, não-patogênica, em forma de bastonete, estritamente aeróbica que produz uma membrana de celulose na interface líquido-ar do meio de cultivo. *G. hansenii* produz celulose extracelular livre de lignina ou hemicelulose, tornando-a uma excelente fonte de celulose pura (BACKDAHL *et al.*, 2008; BROWN, 2004; BROWN *et al.*, 1976; IGUCHI *et al.*, 2000; WATANABE *et al.*, 1998). Ela é capaz de converter várias fontes de carbono em celulose, mas o seu mecanismo de síntese ainda não foi completamente elucidado; presume-se, no entanto, que seja semelhante ao da *Gluconacetobacter xylinus*, que começa sempre com o precursor

de UDP-glicose e termina formando uma cadeia polissacarídica (BODIN *et al.*, 2007).

2.5.1 Celulose Bacteriana (CB)

Quimicamente a celulose é um homopolissacarídeo linear, cuja unidade estrutural, a celobiose, é formada pela união de duas moléculas de D-glicose unidas por ligações glicosídicas β -(1 \rightarrow 4). Essas ligações dão um alinhamento linear para a molécula, o que permite a formação de duas ligações hidrogênio intramoleculares dentro de cada resíduo de glicose: uma ligação liga o O(6) com o O(2)H do resíduo seguinte e a outra liga o O(3)H com o O(5). A ligação intermolecular de hidrogênio liga as diferentes cadeias de celulose, a partir de O(3) para O(6)H (Figura 4). A cadeia glicana pode ter de 2000 a mais de 5000 resíduos de glicose. A hidrólise total da celulose produz moléculas de glicose, enquanto que a hidrólise parcial produz moléculas de celobiose (RAVEN *et al.*, 2010).

Figura 4 - Representação das ligações de hidrogênio inter- e intramolecular da celulose. Linhas tracejadas mostram as ligações de hidrogênio intermoleculares, enquanto as linhas pontilhadas mostram as ligações de hidrogênio intramoleculares (RAVEN *et al.*, 2010).



Grupos hidroxilas estabelecem interações tipo ponte de hidrogênio intra e intermoleculares que são responsáveis pela rigidez da cadeia, pela formação de nanofibras e estáveis e fazem com que a

celulose seja insolúvel em água e na maioria dos solventes orgânicos (KENNEDY *et al.*, 1982). Na maior parte a CB consiste de água (98-99%), mas as propriedades mecânicas são resultados do conteúdo da matriz das fibras de celulose, em que as células estão incorporadas (BACKDAHL *et al.*, 2006; KLEMM *et al.*, 2005).

O oxigênio é um fator importante para a produção de celulose. Em culturas líquidas a bactéria, mesmo sendo estritamente aeróbia, frequentemente cresce estaticamente e leva à formação de uma camada na forma de membrana, também chamada película, cobrindo a superfície inteira. A membrana de CB formada na interface entre o meio de cultivo e o ar em condições estáticas apresenta duas superfícies com microestruturas distintas. A membrana de CB formada na interface com o ar possui uma superfície formada por um arranjo altamente denso de nanofibras de CB (lado denso) e a outra superfície possui um arranjo anisotrópico e poroso de nanofibras de CB (lado poroso) (BACKDAHL *et al.*, 2006; BERTI, 2012; CHAVEZ-PACHECO *et al.*, 2005; CHENG *et al.*, 2009; HORNUNG *et al.*, 2006; JONAS e FARAH, 1998).

2.5.1.1 A biossíntese de celulose bacteriana

Segundo estudos, a bactéria *G. hansenii* produz celulose por duas razões: como subproduto do seu metabolismo e como mecanismo de defesa do ambiente. Esse mecanismo permite que a bactéria flutue na interface líquido-ar tendo acesso a oxigênio e nutrientes. A capacidade desse microorganismo de se agregar e crescer em superfícies expostas forma um biofilme que é caracterizado por células, as quais se encontram incorporadas na matriz, composta por substâncias poliméricas extracelulares (em inglês, *exopolysaccharides*, ou EPSs), produzidas por elas próprias (FREITAS *et al.*, 2011). A celulose bacteriana é caracterizada por ter uma estrutura em forma de rede tridimensional, onde as cadeias de celulose se agrupam por ligações ou pontes de hidrogênio formando nano e microfibrilas.

O polímero (CB) é secretado para o ambiente extracelular da bactéria e as cadeias são agrupadas em paralelo. A organização das fibras de celulose é consequência de seu mecanismo de biossíntese, uma vez que as cadeias são formadas por polimerização em uma direção específica seguida da agregação espontânea, enquanto ainda estão ligadas à superfície celular (MORGAN *et al.*, 2013).

A biossíntese de celulose envolve várias enzimas reguladoras, sendo que a mais importante desta via é a celulose sintase, que participa na transição a partir de UDP-glicose para celulose. A enzima celulose sintase é ativada alostericamente pela molécula diguanosina monofosfato cíclico (c-di-GMP), formado a partir de duas moléculas de guanosina trifosfato (GTP). A enzima é ativada por essa ligação, transformando UDP-glicose em UDP enquanto são formadas as cadeias β -1,4-glicanas, que são então utilizados para sintetizar a celulose (ROSS *et al.*, 1991; STRAP *et al.*, 2011).

Figura 5 – Série de enzimas reguladoras na biossíntese de celulose bacteriana.

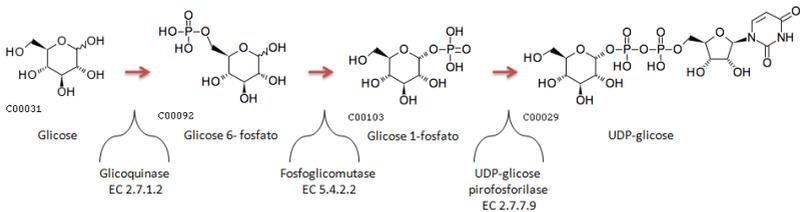
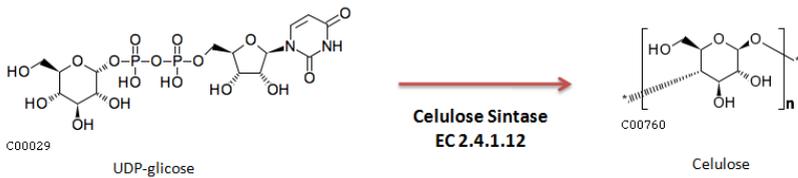


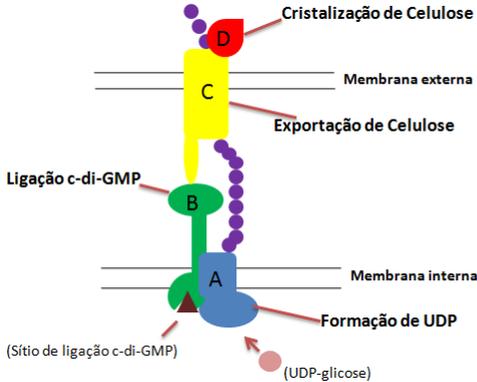
Figura 6 – Enzima Celulose Sintase:
Reação a partir de UDP-glicose para celulose



A polimerização da glicose em cadeias envolve um complexo presente na membrana celular, cujos genes estão organizados em um *operon* (conjunto de genes que se encontram funcionalmente relacionados e controlados coordenadamente), conforme mostra a Figura 7. O cluster de genes que desempenham o papel de produção da celulose é o operon *acs*. Ele contém três genes principais, *acsAB*, *acsC* e *acsD* e o produto final é a celulose sintase. O *acsAB* catalisa a formação de celulose bacteriana; *acsA* é a subunidade catalítica que utiliza UDP-glicose para formar a unidade básica de celulose, enquanto que *acsB* fornece uma subunidade reguladora que tem um domínio de ligação de di-GMP cíclico. O gene *acsC* é o principal componente para a formação do complexo de celulose sintase proposto para ser envolvido na

exportação do polímero através da parede celular. O gene *acsD* está envolvido na cristalização da celulose madura por clivagem entre as cadeias de celulose e se propõe a ter um papel na liberação do crescimento do polímero a partir da célula (IYER *et al.*, 2010).

Figura 7 – Subunidades da Celulose Sintase: *operon acs*



As perspectivas mais interessantes para o aumento de produção das EPSs bacterianas são através de engenharia metabólica ou manipulação de genes que codificam as enzimas que catalisam as reações nas vias, ou alterando as vias reguladoras que afetam a expressão gênica e a atividade enzimática (FREITAS *et al.*, 2011).

2.5.1.2 Efeito das fontes de carbono na produção de celulose bacteriana

A produção de celulose depende de vários fatores, tais como o meio de crescimento, condições ambientais, e a formação de subprodutos. Em meados da década de 1900, ficou comprovada a necessidade de glicose e oxigênio na síntese de celulose bacteriana (SCHRAMM e HESTRIN, 1954). O meio contém carbono, nitrogênio e outros nutrientes necessários para o crescimento de bactérias (RAMANA *et al.*, 2000).

A bactéria *G. hansenii* tem a capacidade de utilizar uma variedade de substratos carbônicos para a biossíntese de celulose. Os meios de cultivo bacterianos mais utilizados são constituídos basicamente de uma fonte de carboidrato (sacarose, glicose, frutose,

lactose ou manitol entre outros), além da utilização do extrato de levedura como fonte de nitrogênio (YAMADA *et al.*, 1997). Tradicionalmente a celulose bacteriana é produzida em culturas estáticas utilizando a bactéria *G. hansenii*, que produz uma membrana de celulose na superfície do meio de cultivo que pode ser cultivada até a formação de uma membrana de vários milímetros de espessura.

Em geral, a glicose tem sido usada como uma fonte de carbono para a produção de celulose. Porém estudos relatam o uso de diferentes fontes de carbono, simplesmente substituindo a glicose por glicerol, manitol, frutose, galactose ou sacarose (MASAOKA *et al.*, 1993; MIKKELSEN *et al.*, 2009; ROSS *et al.*, 1991; SUWANPOSRI *et al.*, 2013).

O manitol é conhecido por ser convertido em frutose e, em seguida, metabolizado por este organismo para a produção de celulose, enquanto que a glicose e frutose são transportadas através da membrana da célula e integradas na via biossintética de celulose (OIKAWA *et al.*, 1995; ROSS *et al.*, 1991). Oikawa e colaboradores comprovaram que sob as mesmas condições de cultura a quantidade de celulose a partir de manitol foi três vezes superior do que a partir de glicose para a bactéria *Acetobacter xylinum* (OIKAWA *et al.*, 1995)

2.5.1.3 Aplicações da celulose bacteriana

Nas últimas décadas, inúmeros avanços têm sido realizados na área de biomateriais, o que foi favorecido pelo crescimento conjunto da biologia celular e molecular. Estes biomateriais são novos materiais projetados para substituir partes do corpo e permitir a recuperação de funções biológicas afetadas por doenças ou acidentes (Palsson e Bhatia, 2004). O desafio neste caso é a concepção de *scaffolds* que interagem com sistemas biológicos e promovam a adesão e função celular (FINK, 2009).

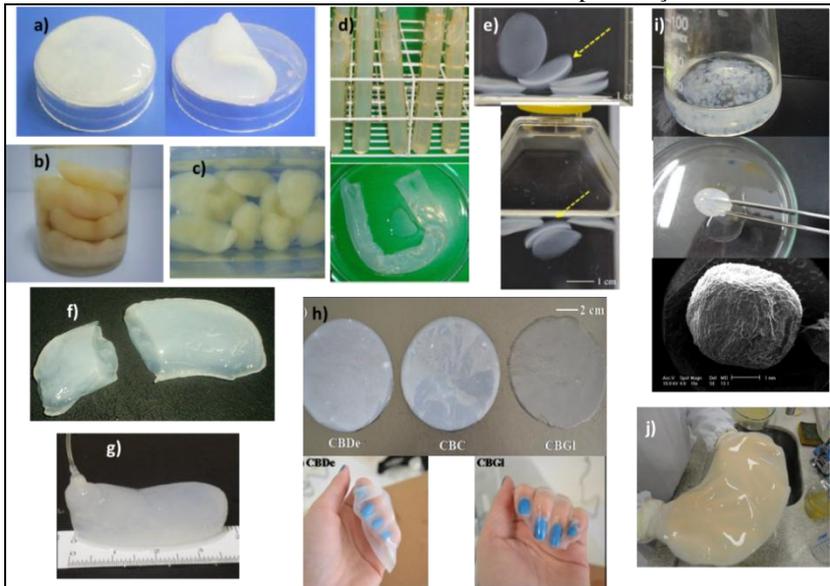
Diversos biopolímeros têm sido investigados como *scaffolds*, dentre os quais se destaca a celulose bacteriana, CB (Svensson *et al.*, 2005). O *scaffold* deve mimetizar a matriz extracelular e guiar as células no processo de formação do tecido (GODBNEY e ATALA, 2002). No entanto, a funcionalidade dos *scaffolds* pode ser alcançada para diferentes tecidos, através de modificações das propriedades de superfície, porosidade, e da própria estrutura tridimensional. Entretanto, o requisito mais importante é a sua biocompatibilidade esta pode ser

definida, de acordo com Williams como "a capacidade que um material tem de realizar uma resposta apropriada do hospedeiro objetivando aplicações específicas" (WILLIAMS, 1987). As aplicações biomédicas da CB incluem, entre outros, substitutos temporários da pele para recuperação de ferimentos e queimaduras; sistema para liberação de fármacos; produção de vasos sanguíneos, entre outros (CHENG *et al.*, 2009; CZAJA *et al.*, 2006; CZAJA *et al.*, 2007; FU *et al.*, 2013; KLEMM *et al.*, 2001).

A CB pode ser moldada em estruturas 3D *in vitro* durante a cultura. Variações estruturais, tais como a espessura e a densidade da rede podem ser controladas por alterações das condições de cultura. (BACKDAHL *et al.*, 2006; KLEMM *et al.*, 2001). A Figura 8 apresenta algumas das estruturas de celulose bacteriana obtidas variando-se condições de cultura no Laboratório de Tecnologias Integradas - Intelab, UFSC.

Figura 8 – Estruturas de celulose bacteriana obtidas variando-se as condições de cultura bacteriana.

(a) Membrana de CB obtida em cultura estática (RECOUVREUX *et al.*, 2008), (b) Hidrogel de celulose bacteriana-3D de formato variado, produzido por *G. hansenii* em agitador orbital sem tratamento de purificação (RECOUVREUX *et al.*, 2008), (c) a mesma estrutura gelatinosa que em (b) após tratamento de purificação (RECOUVREUX *et al.*, 2008), (d) Vasos de celulose bacteriana em seu aparato e depois de pronto (OLIVEIRA, 2011), (e) Membranas de celulose bacteriana e suas camadas assimétricas percebe-se uma diferença de densidade de fibras da camada superior, formada em contato com o ar (mais esbranquiçada), e com a inferior em contato com o meio (mais transparente e porosa) (BERTI, 2012), (f) Estrutura do hidrogel 3D – cocoon (Recouvreux *et al.*, 2011), (g) Macroestrutura de CB com fio de Nylon no interior (BERTI, 2012), (h) Membranas de celulose bacteriana sintetizadas em placas de Petri, resultantes da adição de dextrina (CBDe) e glicose (CBGI) ao meio de cultura (SILVA, 2012), (i) Esferas de celulose bacteriana obtidas em cultura agitada, (j) Membrana de CB em cultura estática sem tratamento de purificação.

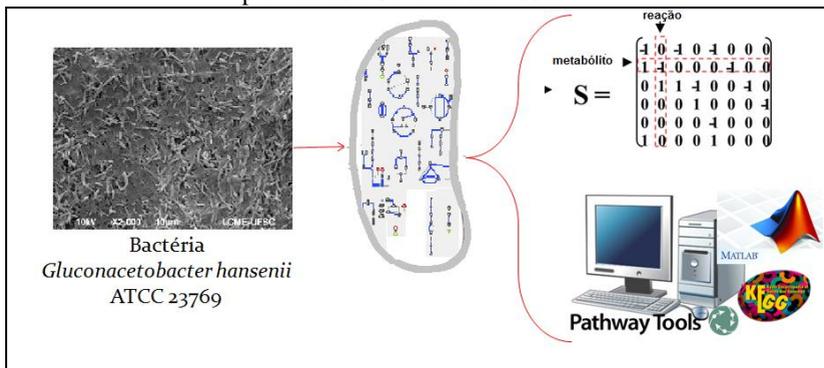


CAPÍTULO 3

METODOLOGIA

Neste capítulo estão descritas as etapas do processo de reconstrução metabólica em escala genômica da bactéria *Gluconacetobacter hansenii*, que consistiu primeiramente na utilização do genoma anotado do organismo para construção de uma rede preliminar. Segue-se o processo de cura manual através do estudo do metabolismo e vias bioquímicas, com foco na biossíntese de celulose bacteriana. A reconstrução é então convertida para um modelo matemático que pode ser usado para diversos propósitos. Utilizando-se ferramentas computacionais foram realizadas diferentes simulações com o modelo metabólico construído até a avaliação da rede e validação do modelo teórico. A Figura 9 resume a metodologia empregada, ou seja, partindo-se do organismo alvo, a bactéria *G.hansenii*, mapear o seu metabolismo e aplicar abordagens matemáticas e ferramentas computacionais que permitam realizar as análises propostas.

Figura 9 – Esquema geral da metodologia aplicada. A partir da bactéria *G. hansenii* foi mapeado seu metabolismo e utilizado ferramentas computacionais para analisar sua rede metabólica.

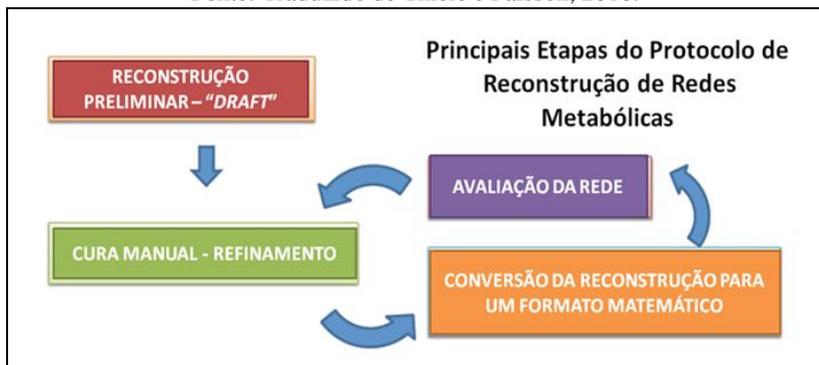


3.1 Reconstrução da rede metabólica em escala genômica

O processo de reconstrução de um modelo metabólico em escala genômica envolveu quatro etapas principais: (1) a criação de um modelo inicial (*draft*), (2) a reconstrução desse modelo de maneira detalhada, através da cura manual, (3), a conversão para um formato matemático e (4) avaliação da rede.

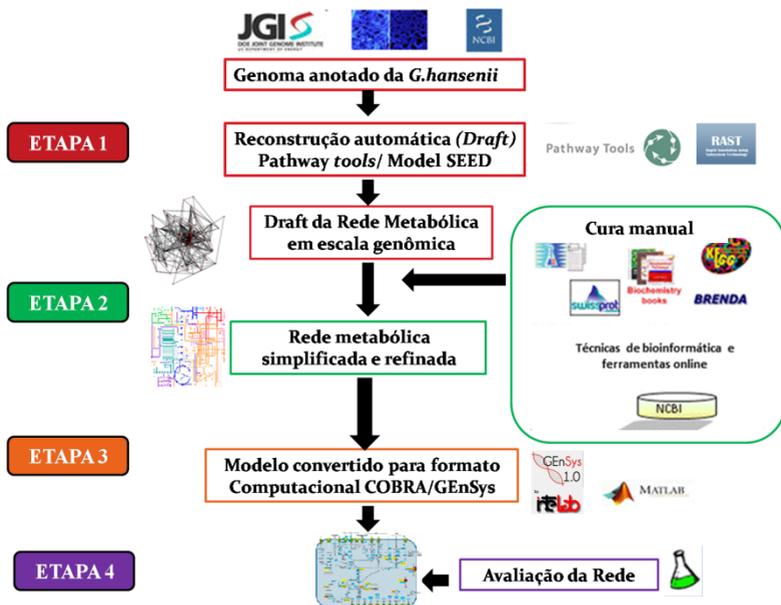
O protocolo desenvolvido e proposto por Pálsson e seu grupo para reconstruções de redes metabólicas (THIELE e PALSSON, 2010) consiste nestas quatro etapas principais. Ele é um protocolo bem estruturado para organismos bem conhecidos e com abundância de informações experimentais. A bactéria *G. hansenii*, de nosso interesse, por outro lado, tem fisiologia relativamente pouco conhecida. Para a execução dos passos do protocolo foram realizadas algumas alterações para realizar a reconstrução desta bactéria. Tais modificações serão discutidas ao longo das etapas. A Figura 10 apresenta as etapas do processo de reconstrução.

Figura 10 - Etapas do processo de reconstrução metabólica
Fonte: Traduzido de Thiele e Pálsson, 2010.



O fluxograma com as etapas pré-definidas do processo e as principais ferramentas utilizadas encontra-se na Figura 11.

Figura 11 - Fluxograma representativo das etapas de construção da rede metabólica.



3.1.1 Etapa 1 – Reconstrução preliminar – o draft do genoma

A primeira etapa de qualquer reconstrução em escala genômica é o uso do genoma anotado do organismo em estudo. Assim esta etapa consiste dos seguintes passos:

1 – Obtenção do genoma anotado: Os dados anotados da sequência genômica de *G. hansenii* (NCBI ID 714995) foram obtidos do banco de dados GenBank do NCBI. A versão mais recente disponível (ACCESSION NZ CM000920, 3636659 bp, DNA circular, de 07/03/2013) foi baixada para garantir que as atualizações e correções desde a publicação original do genoma fossem contabilizadas. Assim, a qualidade do genoma anotado é crucial para a qualidade da reconstrução.

2 – Montagem do *draft* - Esta etapa foi feita de maneira automática, utilizando duas ferramentas: Pathway Tools versão 16.5 e o recurso Model SEED, que apesar de não conter em seu banco de dados a bactéria *G. hansanii*, através do RAST Server o genoma pode ser submetido a um trabalho totalmente automático, criando um banco do organismo com diversas informações, desde reações, vias, metabólitos e genes. A inserção desta ferramenta Model SEED foi uma modificação ao protocolo. Conforme a Figura 12 foi realizada a submissão do genoma anotado ao RAST Server.

Figura 12 – Trabalho submetido ao sistema do RAST.

Job Details #95837

Jobs you have access to :

Job ▲▼	Owner ▲▼	ID ▲▼	Name ▲▼	Num contigs ▲▼	Size (bp) ▲▼	Creation Date	Annotation Progress	Status
95837	SILVA DE SOUZA, SAMARA	714995.5	Gluconacetobacter hansanii ATCC 23769	73	3547097	2013-06-30 10:16:51	<div style="width: 100%; height: 10px; background-color: green;"></div>	complete

3.1.2 Etapa 2 - Cura Manual

A próxima etapa do processo foi a cura manual, que consistiu em corrigir as eventuais inconsistências, uma vez que a extração das reações metabólicas foi realizada de forma automática e não permite estabelecer certas características do organismo alvo. Logo, a cura manual foi necessária para adicionar e corrigir informações que o processo automático deixou escapar ou errou na reconstrução inicial da rede. Esse processo requereu uma pesquisa extensiva em bancos de dados online, literatura e publicações disponíveis sobre o organismo. Quanto mais informações disponíveis do organismo alvo mais realista a reconstrução metabólica.

Seguindo o protocolo já desenvolvido os seguintes passos foram realizados nesta etapa de cura manual:

1 - Refinamento do *draft* e montagem do modelo simplificado: Com as informações geradas na etapa 1, avaliaram-se as reações em cada via e organizou-se as informações em subsistemas para montar o modelo simplificado. Com o intuito de compreender a produção de celulose

bacteriana foram identificadas as principais vias metabólicas envolvidas na sua biossíntese (metabolismo *core*). A escolha das vias baseou-se na obtenção de dados da literatura através da consulta a livros-texto de bioquímica e bancos de dados online relacionados com o gênero *Gluconacetobacter* como, por exemplo, KEGG, IMG e os *drafts* gerados no PathoLogic e RAST Server. Isso incluiu a análise de diversas vias como: glicólise, gliconeogênese, ciclo dos ácidos tricarboxílicos (TCA), via das pentoses-fosfato, metabolismo do nitrogênio, transporte de elétrons e fosforilação oxidativa, biossíntese da celulose, além das reações de transporte e biomassa. Segundo Noor e colaboradores (2010), excelentes modelos simplificados podem corresponder a importantes aspectos biológicos, sendo o metabolismo central o caminho mínimo entre os precursores para a produção de biomassa e energia. A escolha por um modelo simplificado foi uma alteração ao protocolo.

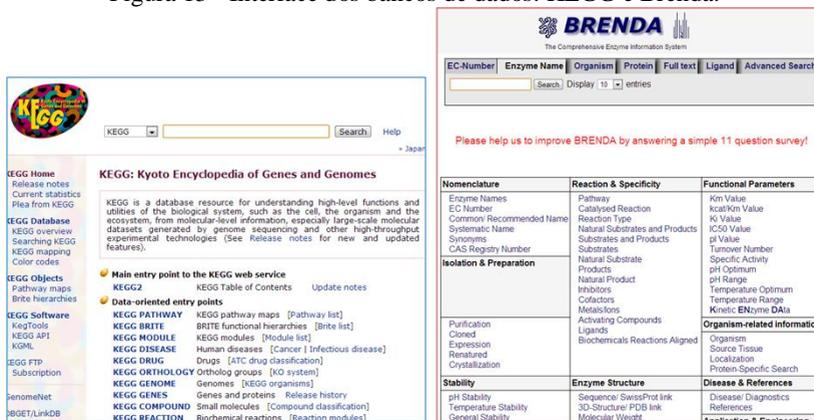
2 - Determinação e verificação do uso de substrato e cofatores:

Através da literatura determinaram-se os substratos e especificidade dos cofatores das enzimas no organismo. Recursos como o KEGG foram utilizados para visualizar a associação das enzimas com as reações. Informações de organismos com filogenia próxima, como *Gluconacetobacter xylinus* e *Gluconobacter oxydans* foram utilizados.

3 - Determinação da fórmula neutra e carregada para cada metabólito na reação e seu identificador:

A fórmula neutra e carregada foi obtida a partir de vários recursos, incluindo, KEGG, Brenda, PubChem, ChEBI e livros-textos (Figura 13). Cada metabólito teve pelo menos um dos seguintes identificadores: ChEBI, KEGG e PubChem identificados e anotados. Muitas reconstruções disponíveis utilizam esses identificadores para escrever as reações.

Figura 13 - Interface dos bancos de dados: KEGG e Brenda.



KEGG: Kyoto Encyclopedia of Genes and Genomes

KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies (See Release notes for new and updated features).

- Main entry point to the KEGG web service
- KEGG2 KEGG Table of contents Update notes
- Data-oriented entry points
- KEGG PATHWAY KEGG pathway maps [Pathway list]
- KEGG BRITE BRITE functional hierarchies [Brite list]
- KEGG MODULE KEGG modules [Module list]
- KEGG DISEASE Human diseases [Cancer | Infectious disease]
- KEGG DRUG Drugs [ATC drug classification]
- KEGG ORTHOLOGY Ortholog groups [KO system]
- KEGG GENOME Genomes [KEGG organisms]
- KEGG GENES Genes and proteins [Release history]
- KEGG COMPOUND Small molecules [Compound classification]
- KEGG REACTION Biochemical reactions [Reaction modules]

BRENDA
The Comprehensive Enzyme Information System

EC-Number | Enzyme Name | Organism | Protein | Full text | Ligand | Advanced Search

Search Display 10 entries

Please help us to improve BRENDA by answering a simple 11 question survey!

Nomenclature	Reaction & Specificity	Functional Parameters		
Enzyme Names	Pathway	Km Value		
EC Number	Catalysed Reaction	Acid/Km Value		
Common/Recommended Name	Reaction Type	Ki Value		
Systematic Name	Natural Substrates and Products	IC50 Value		
Synonyms	Substrates and Products	pH Value		
CAS Registry Number	Substrates	Turnover Number		
Isolation & Preparation	Natural Substrate	Specific Activity		
	Products	pH Optimum		
	Natural Product	pH Range		
	Inhibitors	Temperature Optimum		
Cofactors	Temperature Range	Kinetic Enzyme Data		
Metallofens				
Purification	Activating Compounds	Organism-related information		
	Ligands		Organism	
	Biochemicals Reactions Aligned		Source Tissue	
Cleaved		Localization		
Expression		Protein-Specific Search		
Renatured				
Crystallization				
Stability	Enzyme Structure	Disease & References		
			pH Stability	Sequence: SwissProt link
			Temperature Stability	3D-Structure: PDB link
General Stability	Molecular Weight	Disease: Diagnostics		
		References		
		Annotation & Enzyme.....		

4 - Verificação da estequiometria e reversibilidade das reações:

Neste passo todas as reações do modelo simplificado foram corrigidas estequiometricamente. Prótons e água necessitaram ser adicionados às reações. Os dados bioquímicos para o organismo são muito importantes para determinar a reversibilidade da reação, mas podem não estar disponíveis. De acordo com o protocolo novas abordagens estão disponíveis, tais como a estimativa da energia livre de Gibbs padrão de formação (ΔG_f°) e de reação (ΔG_r°) num sistema bioquímico. Para isso foram utilizados bancos de dados e a ferramenta eEquilibrator (FLAMHOLZ *et al.*, 2012) para analisar a reversibilidade das reações. Com isto as restrições termodinâmicas foram respeitadas.

5 - Adição das informações do gene e localização da reação.

Alguns genes foram identificados e anotados para futuras análises. Utilizou-se o *Genome Region* do NCBI para identificar o gene das reações (enzimas) a partir da localização no genoma, obtida pelo *draft* gerado pelo Model SEED.

6 - Adição das informações de subsistemas para as reações:

Para auxiliar na depuração e avaliação da rede a atribuição dos subsistemas foi feita com base em livros de bioquímica ou com os mapas do KEGG. Algumas vezes uma reação ou uma enzima podem aparecer em vários mapas e, portanto, o subsistema deve refletir sua principal função.

7 - Verificação das associações gene-proteína-reação (GPR): Indica qual a função de cada gene. Utilizando bases de dados específicos do organismo e literatura as GPR foram construídas, mas não estão completas por falta de informações para este organismo. A maioria das reações do modelo simplificado teve apenas seus genes identificados.

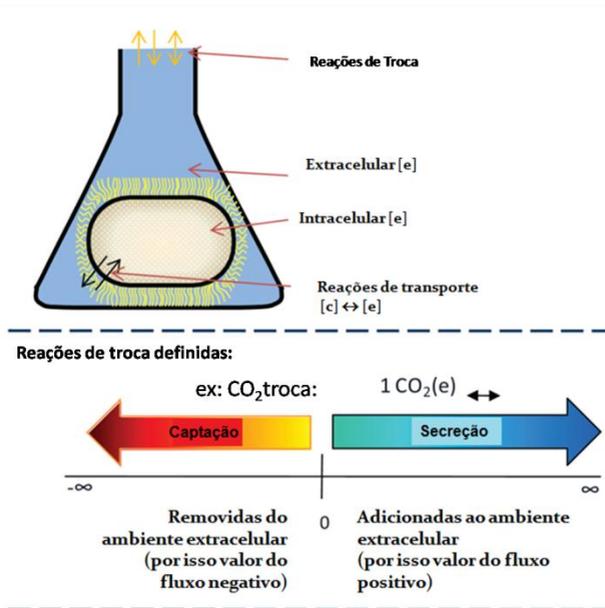
8 - Definição de restrições e adição de reações: Nesta etapa de cura manual algumas restrições (*constraints*) são definidas no modelo, dentre elas, restrições estequiométricas (impostas pelo balanço de massa) e as restrições termodinâmicas (através da reversibilidade e irreversibilidade dos fluxos).

De acordo com o protocolo, durante a cura manual um passo importante na reconstrução da rede metabólica é a adição de reações que não são inferidas da anotação genômica e que definem as condições de contorno do sistema, conforme Figura 14, como:

- **Reações espontâneas** - Utilizou-se a literatura e bases de dados (KEGG e Brenda) para identificar as possíveis reações espontâneas e incluí-las na matriz. Foram incluídas apenas aquelas reações que têm pelo menos um metabólito presente na reconstrução assim minimiza-se o número de metabólitos sem saída (*dead-ends*).
- **Reações de troca (*Exchange reaction*)** - Reações de troca foram adicionadas para todos os metabólitos extracelulares. As reações de troca representaram os limites do sistema.
- **Reações de transporte extracelular** - Uma reação de transporte necessita ser incluída (a partir do espaço entre a matriz extracelular e o citoplasma) para cada metabólito que entra ou deixa a célula. Também se incluem reações de transporte para metabólitos que podem difundir-se através das membranas celulares.
- **Reações de transporte intracelular** - Não foram adicionadas ao modelo, pois apenas reconstruções com múltiplos compartimentos incluem essas reações.
- **Reações de demanda** - São reações não balanceadas que permitem o acúmulo de um composto. Este tipo de reação não é permitido em modelos que envolvem o regime estacionário. Essas reações são adicionadas durante as simulações para avaliar a consistência da rede.

Figura 14 - Definição das condições de contorno.

Fonte: Traduzido de THIELE, 2010



9 - Adição da reação da biomassa: Uma das reações que precisa ser adicionada ao modelo matemático é a reação da produção de biomassa (células). Em modelos metabólicos, a reação de biomassa é incluída para permitir a simulação de crescimento e divisão celular através da produção simultânea de todos os blocos de construção de moléculas pequenas de biomassa (por exemplo, aminoácidos, lipídios, nucleotídeos e cofatores). A reação da biomassa representa assim a combinação ponderada dos componentes, formando a massa seca da célula e a quantidade de ATP hidrolisado necessário para a produção de energia durante o crescimento e a manutenção celular.

Para escrever essa reação utilizou-se dados experimentais da bactéria *Gluconacetobacter xylinus* (ZHONG *et al.*, 2013). A partir da demanda de todos os precursores a reação foi escrita e adicionada ao modelo.

Outra reação adicionada ao modelo foi referente a manutenção de ATP associada ao não-crescimento (NGAM) que é definida como uma demanda de manutenção de ATP equivalente, sendo a unidade

usual mmol ATP/ gDW·h] (gDW: grama de massa seca). Para isso, a seguinte reação foi adicionada à lista de reações: $1 \text{ ATP} + 1 \text{ H}_2\text{O} \rightarrow 1 \text{ ADP} + 1 \text{ Pi} + 1 \text{ H}$.

Todas as informações obtidas durante a cura manual foram registradas em planilhas de Excel® (Microsoft Corporation, Redmond, WA, USA) utilizadas no processo de reconstrução.

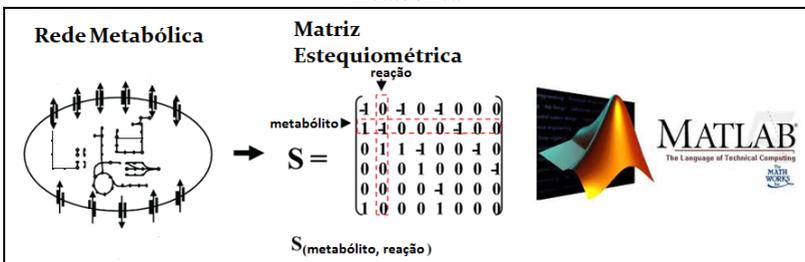
3.1.3 Etapa 3 - Conversão da construção metabólica para o formato computacional

O modelo metabólico construído e refinado após a etapa de cura manual foi convertido em uma representação matemática (matriz estequiométrica) que permitiu implementar uma grande variedade de ferramentas computacionais para analisar as propriedades da rede.

Nesta etapa foram adicionadas as restrições de capacidade que são os limites superiores e inferiores, que definem o fluxo máximo e mínimo permitido para as reações. Neste caso, estas restrições foram utilizadas para definir a quantidade disponível de substrato fixando um fluxo máximo de uma das três fontes de carbono e a ausência das outras duas na alimentação do sistema. Além disso, foram determinadas as taxas de consumo das fontes de nitrogênio e fosfato de acordo com o meio mínimo definido.

As entradas da matriz estequiométrica são os coeficientes dos metabólitos nas reações, com valores negativos para metabólitos consumidos (substratos) e valores positivos quando os metabólitos produzidos ou secretados (produtos) (Figura 15).

Figura 15 - A matriz estequiométrica como centro do processo de reconstrução metabólica



As propriedades da matriz estequiométrica foram investigadas através de dois métodos de análise baseados em restrições: a Análise de Balanço de Fluxo (ABF) e Análise de Fluxo Metabólico (AFM).

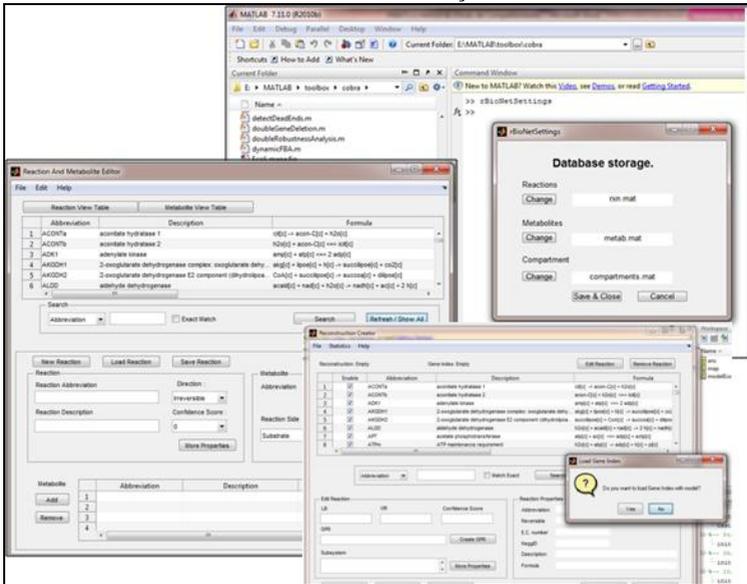
Para a etapa computacional foram utilizados o software MATLAB® e os *toolboxes* COBRA e GEnSys. O GEnSys foi desenvolvido pelo grupo Intelab e foi adicionado ao protocolo para realizar as simulações (Figura 16).

Figura 16 - *Toolboxes* utilizados no ambiente MATLAB®: COBRA e GEnSys.



O COBRA *toolbox* apresenta uma interface rBioNet (Figura 17) onde foi possível adicionar reações e metabólitos, estequiometria, compartimentos (citoplasma e meio extracelular) adequados e restrições. O rBioNet é uma ferramenta útil para a verificação da consistência da estequiometria das reações, pois uma vez adicionados os metabólitos, as reações eram automaticamente editadas com a estequiometria balanceada. Outras funções do COBRA foram utilizadas, dentre elas, a visualização da matriz esparsa (*sparse matrix*).

Figura 17 - Interface rBioNet mostrando o editor de reações e metabólitos e o criador de reconstruções



O modelo construído no rBioNet foi convertido para SBML e validado utilizando a ferramenta SBML validator. Este formato é uma linguagem de marcação em reconstruções metabólicas, e uma série de ferramentas é capaz de processá-la.

Utilizou-se o CellDesigner para a visualização da rede bioquímica. Foi possível construir diferentes *layouts* e visualizar todos os metabólitos da rede e suas ligações.

A modificação no protocolo nesta etapa computacional foi a inserção do *toolbox* GEnSys. Utilizando programação linear para resolver o modelo estequiométrico, o vetor de fluxo pode ser encontrado através da especificação de uma função objetivo que foi maximizada. As simulações e visualização gráfica utilizando as abordagens de AFM e ABF permitiram avaliar a consistência do modelo.

3.1.3.1 AFM e ABF no GEnSys

As redes de reações metabólicas são difíceis de modelar, considerando todos os mecanismos de funcionamento do metabolismo a ser detalhado em função das previsões quantitativas sobre a dinâmica celular. Além disso, os mecanismos das reações intracelulares são complexos e ainda pouco compreendidos cientificamente; aliada a isto, a falta de dados quantitativos referentes à dinâmica intracelular, torna difícil a construção de modelos estruturados complexos. Por isso os modelos estequiométricos ignoram as dinâmicas das reações intracelulares e assumem que os metabólitos internos encontram-se no regime estacionário. Esta hipótese é suportada pela observação de que a dinâmica intracelular é muito mais rápida do que a dinâmica extracelular. Portanto, é sensato ignorar o seu comportamento transitório e considerar que o fluxo rapidamente alcança o regime estacionário.

Para resolver o sistema de equações resultantes do modelo foram utilizadas as técnicas de AFM e ABF para o crescimento de *G. hansenii* em glicose, manitol e glicerol em condições aeróbicas (oxigênio foi ajustado para não haver limitação). O modelo foi simulado fixando as velocidades de consumo de glicose e manitol de $10 \text{ mmol/gDW}^{-1} \text{ h}^{-1}$, e para o glicerol de $20 \text{ mmol/gDW}^{-1} \text{ h}^{-1}$, ou seja, em equivalência de carbomol, uma vez que glicose e manitol apresentam seis carbonos e o glicerol com três carbonos. Para a ABF foram definidas as funções objetivo que mimetizam diferentes situações de maximização/minimização que o organismo necessita em determinada condição.

3.1.3.2 - Funções Objetivo

Para cada condição existe uma função objetivo mais relevante. Funções objetivo cuidadosamente escolhidas alcançam intrinsecamente boa previsão não só em função da fisiologia de crescimento, mas também dos fluxos intracelulares. A escolha dessas funções objetivo deve ser biologicamente consistente e relevante para o problema que está sendo estudado.

Existem diversas possibilidades de funções objetivo para sistemas biológicos. Foram avaliados diferentes cenários considerados relevantes para o presente estudo. Um caso geral (maximização da biomassa) que analisa o caso universal da maximização do crescimento, onde o objetivo é a produção de biomassa, ou seja, os compostos

metabólicos são preferencialmente convertidos em componentes da biomassa. O segundo caso foi a maximização da celulose, o produto de maior interesse derivado da bactéria *G. hansenii*. O terceiro caso avaliou a consistência do balanço através da maximização dos metabólitos externos do modelo. Além disso, o modelo foi simulado limitando alguns nutrientes do metabolismo da bactéria, como fonte de fosfato, nitrogênio e oxigênio. No quarto e último caso, foi incluída uma reação de demanda para cada metabólito da reação da biomassa e otimizada para esse fluxo. A capacidade para sintetizar os precursores de biomassa foi analisada através da adição dessas reações. Normalmente as reações são representadas da seguinte forma: DM_pep. Neste caso, DM_pep refere-se à reação de demanda de fosfoenolpiruvato, e assim sucessivamente para os outros precursores. Se algum metabólito for identificado e não for produzido, é necessário percorrer as vias metabólicas e determinar o motivo pelo qual ele não está sendo produzido.

3.1.3.3 - Meio mínimo de crescimento

Foi adicionado ao modelo matemático o meio de crescimento mínimo da bactéria. Informações sobre o meio utilizado foi essencial para esta etapa computacional.

Foi necessário analisar alguns estudos experimentais para avaliar a melhor composição e definição do meio mínimo. Dentre os meios de cultura citados na literatura para esta bactéria, foi escolhido o meio de crescimento Yamanaka, cuja composição está descrita na Tabela 1 (RUKA *et al.*, 2012; YAMANAKA *et al.*, 1989). A escolha pelo meio foi devido à complexidade da sua composição, o meio Yamanaka apresenta uma composição definida.

Tabela 1 - Componentes e concentração do meio Yamanaka (g/L).

Fonte: Adaptado de RUKA, 2012.

Fonte de carbono	50
(NH ₄) ₂ SO ₄	5
KH ₂ PO ₄	3

Foi prevista uma composição de meio mínimo capaz de suportar o crescimento de *G. hansenii*. Utilizando dados disponíveis na literatura, diferentes combinações de metabólitos foram adicionadas à

rede metabólica até o meio mínimo *in silico* ser determinado, ou seja, a produção de biomassa diferente de zero. Para a determinação do meio mínimo, cada fonte de carbono entrava no modelo, uma por uma, adicionando-se reações de troca caso não houvesse uma correspondente para simular o crescimento do modelo sob diferentes condições ambientais.

Para a simulação do crescimento aeróbico em meio mínimo, os seguintes metabólitos externos foram permitidos a entrar e sair da rede livremente: NH_4^+ , O_2 , H^+ , PO_4^{3-} , CO_2 e H_2O . Todos os outros metabólitos externos, exceto os substratos testados foram autorizados a deixar o sistema. Com o meio definido foram calculadas as taxas de consumo das fontes de carbono, nitrogênio e fosfato para serem utilizadas durante as simulações.

3.1.3.4 Ferramentas de bioinformática

As plataformas *online* para pesquisas biológicas, os bancos de dados biológicos e os softwares utilizados neste trabalho estão listados na Tabela 2 com seus respectivos *links* de acesso.

Tabela 2- Recursos *online* para a reconstrução da rede metabólica e os respectivos *links* de acesso.

RECURSO	URL
Sequência Genômica NCBI	http://www.ncbi.nlm.nih.gov/genome/?term=gluconacetobacter%20hansenii
IMG – Integrated Microbial Genomes	http://img.jgi.doe.gov/cgi-bin/w/main.cgi
Kyoto Encyclopedia of Genes and Genomes (KEGG)	http://www.genome.jp/kegg/pathway.html
ExPASy Biochemical Pathways	http://www.expasy.ch/cgi-bin/search-biochem-index
BRENDA	http://www.brenda-enzymes.org
Uniprot	http://www.uniprot.org/uniprot/?query=gluconacetobacter%20hansenii&sort=score
SEED	http://pubseed.theseed.org/

Genome Region NCBI	http://www.ncbi.nlm.nih.gov/genome/?term=GLUCONACETOBACTER%20HANSENI
eQuilibrator	http://equilibrator.weizmann.ac.il/
SBML validator	http://sbml.org/validator/
MATLAB®	http://www.mathworks.com
Pathway Tools version 16.5	http://bioinformatics.ai.sri.com/ptools/
COBRA (Constraint-based reconstruction and analysis) toolbox	http://opencobra.sourceforge.net/openCOBRA/Welcom.html
GNU linear programming toolk it (GLPK)	http://glpkmex.sourceforge.net
SBMLToolbox version 4.0.1	http://www.sbml.org
libSBML programming library 4.0.1	http://sbml.org/Software/libSBML
rBioNet	<a href="http://sourceforge.net/projects/opencobra/files/cobra/fou
ndry/rBioNet/">http://sourceforge.net/projects/opencobra/files/cobra/fou ndry/rBioNet/
GenSys	http://www.intelab.ufsc.br/

3.1.4 Etapa 4 – Validação do modelo matemático

Com o modelo computacional buscou-se investigar as capacidades metabólicas e avaliar as propriedades sistêmicas que elas podem executar sob as restrições conferidas à célula. Sendo assim, a última etapa no processo de reconstrução consistiu na verificação, avaliação e validação da rede.

O modelo metabólico foi testado, e realizaram-se diversas simulações e análises para avaliar a sua consistência. Por se tratar de uma rede simplificada, grande parte dos *gaps* já havia sido preenchida durante a cura manual.

Uma questão importante é decidir quando parar o processo e considerar a reconstrução como finalizada. Esta decisão baseia-se normalmente no propósito da reconstrução. Observa-se que o modelo metabólico mais completo atualmente é o da bactéria *E.coli* que vem sendo investigado e aprimorado há mais de 10 anos de pesquisa (ALMAAS *et al.*, 2004; EDWARDS *et al.*, 2001; EDWARDS e PALSSON, 2000; KESELER *et al.*, 2005; ORTH *et al.*, 2011; FEIST e PALSSON, 2008).

CAPÍTULO 4

RESULTADOS E DISCUSSÃO

A perspectiva de compreender a relação entre o genoma e a fisiologia do organismo foi um incentivo importante para reconstruir redes metabólicas. O protocolo desenvolvido por Thiele e Palsson (2010) foi muito bem estruturado para redes metabólicas em grande escala e organismos bem estudados e com detalhados dados experimentais. Para organismos com poucas informações relatadas sobre suas capacidades metabólicas, como a *G. hanseni*, as adaptações no protocolo foram necessárias para a reconstrução da rede metabólica.

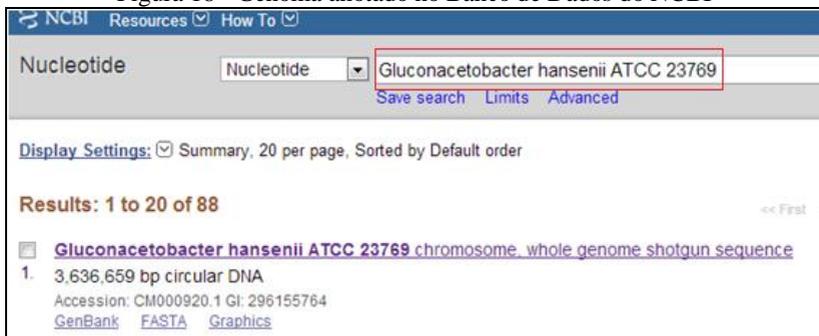
A modelagem experimental *in silico* envolveu a combinação de resultados biológicos da literatura associada a modelos de sistemas biológicos baseados em computadores. Este tipo de estudo serve como precursor para a pesquisa laboratorial experimental, permitindo aos pesquisadores refinar suas metodologias experimentais com o objetivo de aumentar a eficiência da pesquisa.

A primeira rede metabólica *core* da *G. hanseni* seguindo os passos definidos no protocolo revelou-se uma grande ferramenta. Os primeiros passos do processo foram automatizados, no entanto, a cura manual foi, como de praxe, trabalhosa, pois exigiu uma busca intensiva das informações disponíveis.

4.1 Construção do *draft* da rede metabólica da *G. hanseni*

Seguindo as etapas do processo de reconstrução, primeiramente foi construído o *draft* do modelo. Partiu-se do genoma anotado da bactéria *G. hanseni* no banco de dados GenBank do NCBI, conforme apresentado na Figura 18.

Figura 18 - Genoma anotado no Banco de Dados do NCBI

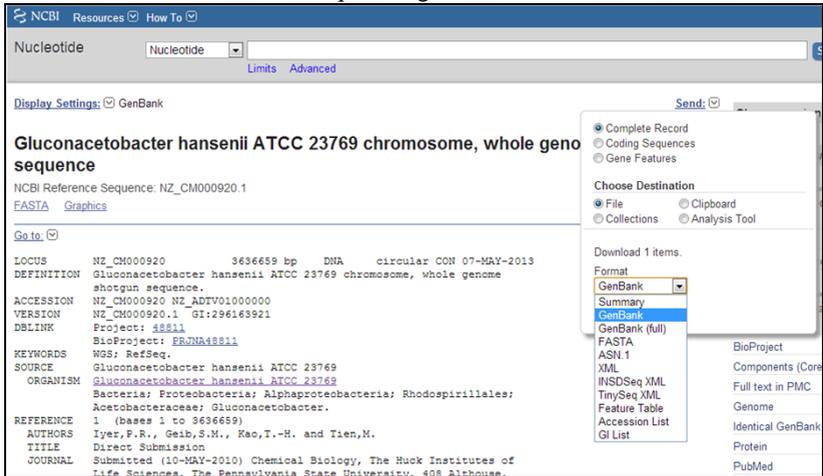


O NCBI forneceu inúmeras informações em relação ao organismo conforme a Tabela 3. A Figura 19 refere-se ao genoma anotado da bactéria e os formatos disponíveis para exportar o arquivo do banco de dados.

Tabela 3 – Dados e identificações da bactéria

Nome organismo	<i>G. hansenii</i> ATCC 23769
Genoma ID	647000257
NCBI Taxonomia ID	714995
NCBI Projeto ID	43711
GenBank	CM000920
GOLD ID	Gi05942
Tamanho Genoma	3.547.112 bp
Total de genes	3.351
Genes codificadores de proteínas (CDS)	3.308 (84% do genoma)
Genes de RNA	45
• Genes de tRNA	43
• Genes de rRNA	2

Figura 19 - Genoma da bactéria e formatos disponíveis para exportar a sequência genômica



Os dois arquivos gerados foram salvos para utilização posteriormente. Os arquivos apresentam estruturas diferentes conforme apresentadas na Figura 20.

Figura 20 - Formatos FASTA e GenBank da *G. hanseni*

FASTA

```

>gi|296163921|ref|NZ_CM000920.1| Gluconacetobacter
hanseni ATCC 23769 chromosome, whole genome
shotgun sequence
TCATCACAGACTACGCATAAGGGGCTATCCACCTCT
ATGGCCGGAATTCATTCCTGTTCTGCTTCTCTGGAAC
AATACAGGGCGTGTCCGGTTCGCTCGCCACTACTAG
CGGAATCTCAAITGATGTCTTCTCCGGGTACTTAG
ATGTTTCAGTTCCTCCGGGTTCGCCTATGCCCTATGT
ATTCAGAACATGATACCCATCCGCTGGTGGGTTGCCCC
ATTCAAGATATCCAGGATCAAAGCTGTGCTCGGGCTCC
CATGGCTTTTCGACAGCTGGCCACGTCCTTTCATCGCT
CTGTGTGCAAGGATCCACCCGAATGCCCTTATCGCGC
TCATTACCAACATGCAAGGAGGACATCCACCTGGG
GCAGACATGCCGACATAGAAAGTGAGTTATCTTC
ACGACCACCTATTCTTTACGTCGTCATGTTCGCTT
ACGCCACATCGCACAGAGCCGAATACCCCGGTTCCCGG
GACCATCGATGGTTCAGACCAATCCCGGACAGCGGA
CGCACACAGAGCCCAACCTTATCCACTGCAGAAAGA
ACAAATTCACAGACCAAAAGCGCGGCCGATCGC
GACTCGCGACAGATGTCGATCTCTTTCACAAAACA
ACAATCATAACGCTGTGATATCACAGACCCATCG
CATCCGGATCCGGGCAACCTGTGTGATAATGACACAG
AAAGGCCACCTGGTGAAGGCGGAGGGATCGAACCGA
CGACCCCTGCTGCAAGACAGAGTCTCCACAGCTGA
GCTACGCCCCGATGCTGATCCGATGGTGGCCAGGG
AGGACTTGAACCTCCGACCCACGCTTATCAAGCGTGT
... (continua)
    
```

GenBank

```

LOCUS NZ_CM000920 3636659 bp DNA circular CON 07-
MAY-2013 DEFINITION Gluconacetobacter hanseni ATCC
23769 chromosome, whole genome shotgun sequence.
ACCESSION NZ_CM000920 NZ_ADVT01000000
VERSION NZ_CM000920.1 GI:296163921 DBLINK Project: 4881
BioProject: PRNA4881
KEYWORDS WGS; RefSeq.
SOURCE Gluconacetobacter hanseni ATCC 23769
ORGANISM Gluconacetobacter hanseni ATCC 23769 Bacteria;
Proteobacteria; Alphaproteobacteria; Rhodospirillales;
Acetobacteraceae; Gluconacetobacter.
REFERENCE 1 (bases 1 to 3636659)
AUTHORS Iyer,P.R., Geib,S.M., Kao,T.-H. and Tien,M.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2010) Chemical Biology, The
Huck Institutes of Life Sciences, The Pennsylvania State
University, 408 Altohouse, University Park, PA 16802, USA
COMMENT PROVISIONAL REFSEQ: This record has not yet
been subject to final NCBI review. The reference sequence is
identical to CM000920. Annotation was added by the NCBI
Prokaryotic Genomes Automatic Annotation Pipeline Group.
Information about the Pipeline can be found here:
http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html.
Please be aware that the annotation is done automatically with
little or no manual curation. ##Genome-Assembly-Data-
START## Genome Coverage :: 22x Sequencing Technology ::
454 ##Genome-Assembly-Data-END##
FEATURES             Location/Qualifiers             source             1.3636659
/organism="Gluconacetobacter hanseni ATCC 23769"
/mol_type="genomic DNA" /strain="ATCC 23769"
/culture_collection="ATCC:23769" /db_xref="taxon:74995"
... (continua)
    
```

Para o desenvolvimento do primeiro *draft* foi utilizado o componente PathoLogic do Pathway Tools e criou-se um novo PGDB (*Pathway/Genome Database*), aqui denominado de GhaCyc contendo as vias metabólicas da *G. hanseni*, utilizando como entrada os arquivos GenBank/FASTA, previamente obtidos. Após esse processo, o PGDB da *G. hanseni* foi criado, conforme mostrado na Figura 21.

Figura 21 - PGDB da *Gluconacetobacter hanseni* criado no Pathway Tools a partir do genoma anotado

The figure displays two screenshots of the Pathway Tools software interface. The top screenshot shows the 'Available Databases' selection screen, and the bottom screenshot shows the main database information page for *Gluconacetobacter hanseni*.

Pathway Tools -- Available Databases

Organisms	Pathways	Genes (ORF %)	Genome Size (bp)	Citations	Download
<i>Escherichia coli</i> K-12 substr. MG1655	300	4623 (18.3%)	4,639,675	23908	
<i>Gluconacetobacter hanseni</i> ATCC 23769	226	3351 (32.7%)	3,636,659	629	
<i>Gluconacetobacter xylinus</i> NBRC 3288	238	2885 (24.5%)	609		
MetaCyc	1928	8319 (0.3%)	0	33647	

Gluconacetobacter hanseni

Strain: ATCC 23769 GhaCyc version: 1.0

Authors: Samara Silva de Souza, UFSC-Brazil

Replicon	Total Genes	Protein Genes	RNA Genes	Pseudogenes	Size (bp)	NCBI Link
GEN-EL203304	3351	3308	43	0	3,636,659	

Summary Statistics:

- Pathways: 226
- Enzymatic Reactions: 1255
- Transport Reactions: 4
- Polypeptides: 3308
- Protein Complexes: 0
- Enzymes: 777
- Transporters: 23
- Compounds: 955
- Transcription Units: 0
- RNAs: 43

Taxonomic lineage: cellular organisms, Bacteria, Proteobacteria, Alphaproteobacteria, Rhodospirillales, Acetobacteraceae, Gluconacetobacter, Gluconacetobacter hanseni group, Gluconacetobacter hanseni, Gluconacetobacter hanseni ATCC 23769

Unification Links: NCBI-Taxonomy:714995

Genetic Code Number:

Com a construção do *draft* no Pathway Tools foi possível visualizar as vias, reações, metabólitos e genes. As vias importadas são baseadas no número de reações únicas e não-exclusivas presentes no genoma. Quando o PathoLogic terminou a sua execução, no entanto, algumas destas reações permaneceram sem associação com sequências genéticas do genoma utilizado. Essas reações são susceptíveis de existir

no organismo para completar a rede metabólica, mas as evidências experimentais permanecem indisponíveis; assim estas vias permanecem identificadas como *gaps* no banco de dados do organismo. As reações de transporte intracelular e extracelular, de formação de biomassa e de transporte precisaram ser adicionadas manualmente.

As Figuras 22 e 23 se referem às visualizações da rede metabólica da *G. hansenii* do projeto GhaCyc, construído pelo autor. Os retângulos representam os compartimentos celulares e as linhas azuis são reações de conexão com compostos químicos. Como pode ser visto, existem algumas reações não agrupadas nas vias e, inclusive, algumas reações aparecem desconectadas do resto da rede. Além disso, é possível acessar vias, genes, reações, metabólitos, reações em múltiplas vias, reações sem o número EC, *dead-ends*, entre outras visualizações, para facilitar o entendimento da rede.

Figura 22 - Visualizações das vias no Mapa celular (*Cellular overview*)

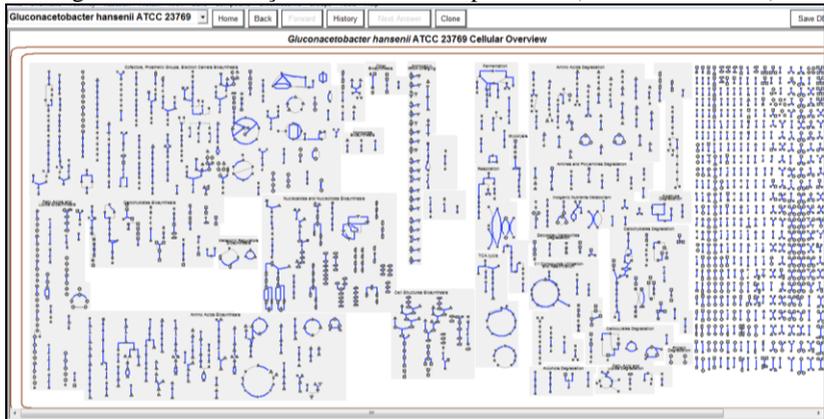
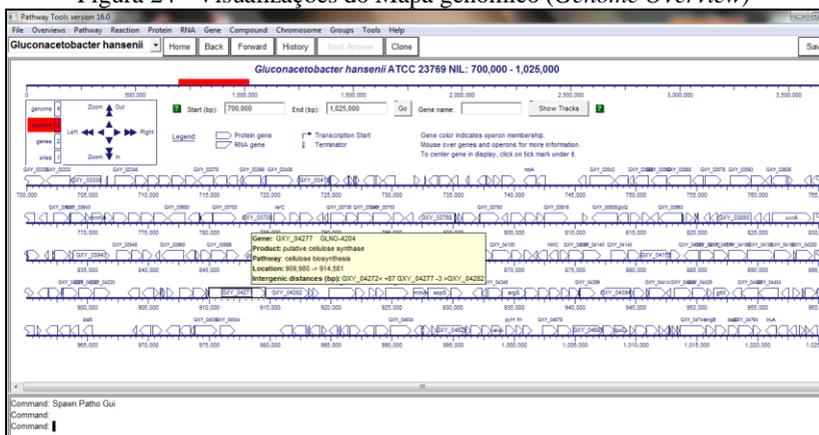


Figura 23 - Componentes do mapa celular destacados nas vias



Além do mapa celular, os dados de GhaCyc foram visualizados em um mapa genômico, conforme mostrado na Figura 24. Esse tipo de informação será muito útil futuramente para os estudos dos mecanismos reguladores. A inclusão de informações de interações reguladoras que torna a rede mais complexa é um passo importante para o futuro desenvolvimento do modelo. Neste mapa os genes podem ser identificados quanto à sua localização no genoma.

Figura 24 - Visualizações do Mapa genômico (*Genome Overview*)

A plataforma Pathway Tools permitiu exportar o *draft* do modelo em SBML usando um utilitário de exportação de arquivos. O utilitário consultou todas as reações e metabólitos dentro do PGDB, e

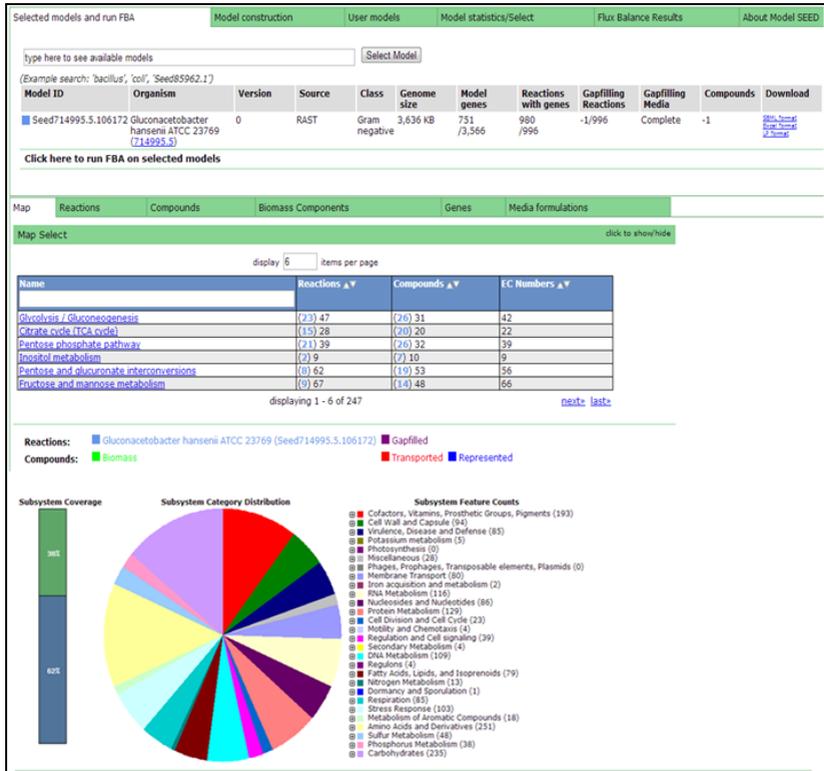
Figura 26 - Planilha gerada com as informações de metabólitos e reações

1	A	B	C
2	Rxn_name	Rxn description	Formula
3	RXN_45_2043	RXN-2043	CELLULOSE + WATER ↔ Celloedextrins
4	MALTODEG_45_RXN	MALTODEG-RXN	MALTOHEXAOSE + MALTOSE ↔ CPD0_45_1133 + GLC
5	2_46_7_46_7_46_8_45_RXN	2.7.7.8-RXN	Pi ↔ RNA_45_Holder + Nucleoside_45_Diphosphates
6	RXNO_45_6522	RXNO-6522	9S_45_RNA ↔ 5S_45_RNAs + 2 sRNAs
7	RXNO_45_6523	RXNO-6523	RNASE_45_G_45_DEGRADATION_45_SUBSTRATE_45_MRNA ↔
8	3_46_1_46_20_46_12_45_RXN	3.1.26.12-RXN	RNASE_45_E_45_DEGRADATION_45_SUBSTRATE_45_MRNA ↔
9	RXNO_45_6524	RXNO-6524	RNASE_45_H_45_POLY_45_A_45_SUBSTRATE_45_MRNA ↔ R
10	RXNO_45_6521	RXNO-6521	RNASE_45_E_45_MRNA_45_PROCESSING_45_SUBSTRATE ↔ R
11	DNA_45_DIRECTED_45_RNA_45_POLYMERASE_45_RXN	DNA-DIRECTED-RNA-POLYMERASE-RXN	Nucleoside_45_Triphosphates ↔ 99 + RNA_45_Holder
12	GLUTNAREDUCT_45_RXN	GLUTNAREDUCT-RXN	Charged_45_GUT_45_RNAs + NADPH → GLITAMATE_45_1_45_4
13	2_46_9_46_1_46_1_45_RXN	2.9.1.1-RXN	SEPO3 + L_45_seryl_45_SEC_45_RNAs → Charged_45_SEC_45
14	RXN_45_12104	RXN-12104	tRNAs_45_containing_45_epoxy_45_quenone → tRNAs_45_w
15	TRNA_45_GUANINE_N7_45_45_METHYLTRANSFERASE_45_RXN	TRNA-GUANINE-N7-METHYLTRANSFERAS	Guanine46_45_in_45_RNA + S_45_ADEOSYLMETHIONINE ↔ TR
16	RXNO_45_1342	RXNO-1342	TRNA_45_with_45_7_45_aminomethyl_45_7_45_deazaguanin
17	METHIONYL_45_TRNA_45_FORMYLTRANSFERASE_45_RXN	METHIONYL-TRNA-FORMYLTRANSFERASE	10_45_FORMYL_45_THF + L_45_methylomyl_45_RNAmet ↔ TH
18	6_46_3_46_5_46_6_45_RXN	6.3.5.6-RXN	GLN + L_45_aspartyl_45_tRNAAsn + ATP + WATER → GLT + Charged_45
19	6_46_3_46_5_46_7_45_RXN	6.3.5.7-RXN	L_45_glutamyl_45_tRNAgin + GLN + ATP + WATER → Charged_45_
20	TRNA_45_NUCLEOTIDYLTRANSFERASE_45_RXN	TRNA-NUCLEOTIDYLTRANSFERASE-RXN	Pi ↔ RNA_45_Holder + Nucleoside_45_Diphosphates
21	RXNO_45_6482	RXNO-6482	CPD0_45_2353 + P → tRNAs + Ribonucleoside_45_Diphosphates
22	TRNA_45_CYTIDYLTRANSFERASE_45_RXN	TRNA-CYTIDYLTRANSFERASE-RXN	tRNA_45_precursors + 2 CTP + ATP ↔ tRNAs_45_with_45_CCA + 4
23	LEUCYLTRANSFERASE_45_RXN	LEUCYLTRANSFERASE-RXN	General_45_Proteins_45_Substrates + Charged_45_LEU_45_tRNA
24	3_46_1_46_13_46_5_45_RXN	3.1.13.5-RXN	tRNA_45_precursors ↔ Nucleoside_45_Monophosphates + tRNA_
25	RXNO_45_6484	RXNO-6484	CPD0_45_2353 → Nucleoside_45_Monophosphates + tRNAs
26	RXN_45_1961	RXN-1961	TRNA_45_1962 LYS + ATP ↔ lysidine_45_RNA_45_1962 + 2 PROT
27	RXNO_45_6480	RXNO-6480	CPD_45_2354 → tRNAs + ssRNAs
28	RXNO_45_6274	RXNO-6274	CPD_45_4211 + tRNA_45_Holder ↔ tRNA_45_Containing_45_6
29	QUEUOSINE-TRNA-RIBOSYLTRANSFERASE_45_RXN	QUEUOSINE-TRNA-RIBOSYLTRANSFERASE	Guanine34_45_in_45_RNAs + QUEUINE + PROTON ↔ GUANINE + P
30	RXNO_45_1321	RXNO-1321	7_45_AMINOMETHYL_45_7_45_DEAZAGUANINE + Guanine34_45_
31	AMINOACYL_45_TRNA_45_HYDROLASE_45_RXN	AMINOACYL-TRNA-HYDROLASE-RXN	N_45_Substituted_45_Aminomethyl_45_RNA + WATER ↔ N_45_1
32	RXN_45_12460	RXN-12460	Charged_45_45_RNAs + WATER → ASN + ASN_45_RNAs + 2
33	TRNA_45_PSEUDOURIDINE_45_SYNTHASE_45_1_45_RXN	TRNA-PSEUDOURIDINE-SYNTHASE-I-RXN	TRNA_45_uridine_45_38_45_40 ↔ tRNA_45_pseudouridine_45

Além da plataforma Pathway Tools, utilizou-se a ferramenta Model SEED, que foi adicionada à primeira etapa do processo e complementou as informações obtidas pelo Pathway Tools. Como a *G. hansenii* não se encontrava nos bancos de dados do Model SEED, o genoma desta bactéria foi submetido ao RAST Server, um dos ambientes do SEED, para gerar o banco de dados correspondente. O sistema forneceu informações úteis nesta etapa de construção do *draft* e também contribuiu para a etapa de cura manual. Os arquivos gerados por ambas as plataformas serviram para gerenciar os dados de cura manual.

A Figura 27 apresenta as funcionalidades do sistema, sendo possível também exportar o modelo nos formatos SBML e Excel. É gerado um banco de informações com reações, metabólitos e genes.

Figura 27 - Visualização da página principal dos resultados obtidos com o RAST Server



Com as informações obtidas pelo Model SEED foi possível visualizar os subsistemas e gerar fluxogramas das vias do organismo. O interesse nesta etapa foi identificar algumas peculiaridades do organismo e auxiliar posteriormente na etapa de cura manual.

4.2 Cura manual do *draft* gerado

Essa etapa do processo foi a mais extensa, pois houve a necessidade de incluir diversos detalhes à análise. Após a escolha das vias que definiram o modelo simplificado da bactéria, a cura manual foi realizada seguindo todas as etapas definidas pelo protocolo.

Foi realizada uma extensa revisão do organismo em bancos de dados, sendo que a plataforma IMG foi muito utilizada por conter informações referentes a esta bactéria (Figura 28). O uso dessa plataforma foi uma modificação do protocolo. O IMG é integrado com diversos bancos de dados e forneceu informações úteis do organismo. Por exemplo, a identificação do ID da taxonomia foi muito importante, pois existem diversas bactérias do gênero *Gluconacetobacter* e o número é um identificador da linhagem em estudo.

Figura 28 - Busca em bancos de dados do organismo - IMG

The screenshot shows the IMG website interface. At the top, there is a search bar with 'ALL Genomes' and a 'GO' button. Below the search bar is the IMG logo and the text 'INTEGRATED MICROBIAL GENOMES'. The main navigation bar includes links for 'IMG Home', 'Find Genomes', 'Find Genes', 'Find Functions', 'Compare Genomes', 'Analysis Cart', and 'My IMG'. There is also a 'Companion Systems' and 'Using IMG' link. The main content area is titled 'Genome Browser' and includes a search filter for 'Proposal Name' with the text 'gluconacetobacter' entered. Below the search bar is a table with columns: 'Select', 'Domain', 'Status', 'Proposal Name', 'Genome Name / Sample Name', 'Sequencing Center', 'Genome Size', and 'Gene Count'. The table contains four rows of data, with the last row highlighted in red.

Select	Domain	Status	Proposal Name	Genome Name / Sample Name	Sequencing Center	Genome Size	Gene Count
<input type="checkbox"/>	B	F	Gluconacetobacter diazotrophicus PAI 5, DSM 5601	Gluconacetobacter diazotrophicus PAI 5, DSM 5601	AGROBIOLOGIA, UNCC, UERJ	3999591	3997
<input type="checkbox"/>	B	F	Gluconacetobacter xylinus NBRC 3288	Gluconacetobacter xylinus NBRC 3288	National Institute of Technology and Evaluation (NITE)	3513191	3269
<input type="checkbox"/>	B	D	Gluconacetobacter sp. SVCC-1	Gluconacetobacter sp. SVCC-1	Tianjin University	4233336	4967
<input type="checkbox"/>	B	D	Gluconacetobacter hansenii ATCC 23769	Gluconacetobacter hansenii ATCC 23769	Penn State University	3547122	3353

Uma rede metabólica pode apresentar um nível variado de detalhes, como uma grande quantidade de reações, informações de regulação, informação espacial das reações, como compartimento celular, entre outros dados. Uma reconstrução completa de um organismo como a *G. hansenii* que possui poucas informações experimentais relacionadas ao seu metabolismo pode levar anos para ser finalizada, exigindo assim um grande esforço por parte dos pesquisadores.

4.2.1 Montagem da rede metabólica core

Com as informações obtidas das plataformas Pathway Tools e Model SEED, iniciou-se o processo de cura manual. As reações apresentavam estequiometria incorreta ou desequilibrada em função da ausência de reações e erros provenientes do genoma anotado, uma vez que o genoma disponível ainda continha *gaps*.

As reações dos subsistemas definidos para o modelo simplificado foram anotadas e refinadas. Os metabólitos foram organizados em dois compartimentos (citoplasma e extracelular) com base na localização das enzimas associadas a cada via. Os cofatores e a reversibilidade das reações foram compilados a partir de informações da literatura. Para cada metabólito, a carga, a fórmula e informações de identificação foram anotadas. O número de classificação das enzimas (EC number), i.e., o número identificador da reação bioquímica, foi anotado e os genes foram identificados. Todas estas informações foram compiladas em planilhas Excel.

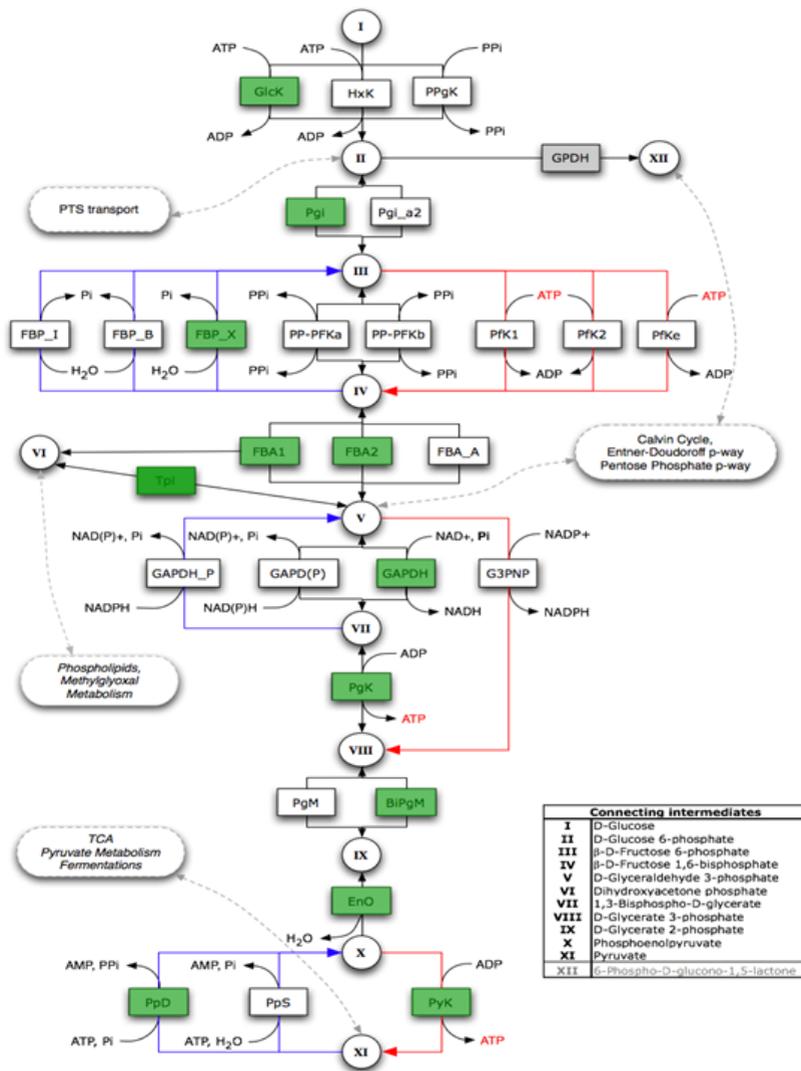
Com a ferramenta Model SEED, os subsistemas definidos para construir o modelo metabólico simplificado foram gerados.

Glicólise e Gliconeogênese

A Figura 29 representa o subsistema glicólise e gliconeogênese, responsável pela conversão de glicose (I) à piruvato (XI). Os blocos verdes referem-se as vias realizadas pela *Gluconacetobacter hansenii*; as linhas vermelhas representam as enzimas glicolíticas irreversíveis; e as linhas azuis são enzimas gliconeogênicas.

Figura 29 - Subsistema glicólise e gliconeogênese da *G. hanseni*.

Fonte: Model SEED



Analisando as informações obtidas nos fluxogramas, a Figura 29 mostra uma indicação de irreversibilidade da reação do composto III (frutose 6-fosfato) para composto IV (frutose 1,6-difosfato). A fosfofrutoquinase é uma das enzimas chave na via Embden-Meyerhof-

Parnas (EMP) e é responsável por catalisar a formação de frutose-1,6-difosfato a partir de frutose-6-fosfato na glicólise, mas ela se encontra ausente ou apresenta baixa atividade na *Gluconacetobacter*.

Segundo Ross (1991), a *G. hansenii* é incapaz de metabolizar a glicose anaerobicamente devido à carência da fosfofrutoquinase. Desta forma a síntese da celulose resulta de um *pool* metabólico de hexose fosfato que é sustentado diretamente pela fosforilação de hexoses exógenas e indiretamente pela via das pentoses e gliconeogênese.

As três maiores vias metabólicas da célula são: as rotas Embden-Meyerhof-Parnas (EMP ou glicólise), via das pentoses fosfato (PPP) e Entner-Doudoroff (ED) que são responsáveis pela conversão de glicose à piruvato. A primeira etapa destas rotas metabólicas, referente à conversão de glicose a fosfogliceraldeído, ocorre através de diferentes vias. Num segundo estágio, o fosfogliceraldeído é convertido a piruvato. O piruvato também pode ser convertido a acetil-CoA, passando pelo ciclo dos ácidos tricarbóxicos onde é oxidado em dióxido de carbono nos organismos aeróbicos como a *G. hansenii*.

A síntese da celulose, a partir de fontes endógenas, se dá a partir da gliconeogênese que, em *G. hansenii*, ocorre a partir do oxalacetato, via piruvato, devido a um mecanismo de regulação não comum das enzimas oxalacetato descarboxilase, e a piruvato fosfatodiquinase, que converte o piruvato a fosfoenolpiruvato. A reação de síntese de CB constitui um processo que sobrecarrega a célula, consumindo parte do ATP gerado no metabolismo bacteriano. Assim, a energia empregada para a síntese da CB é proveniente do metabolismo aeróbio.

Via das pentoses e Ciclo dos ácidos tricarbóxicos

A biossíntese da celulose bacteriana depende de duas vias anfóbicas importantes (que envolvem catabolismo e anabolismo), a via das pentoses (Figura 30) e a via do ciclo dos ácidos tricarbóxicos que é fonte de importantes precursores metabólicos (Figura 31).

Figura 30 - Via das pentoses fosfato da *G. hansenii*. Fonte: Model SEED

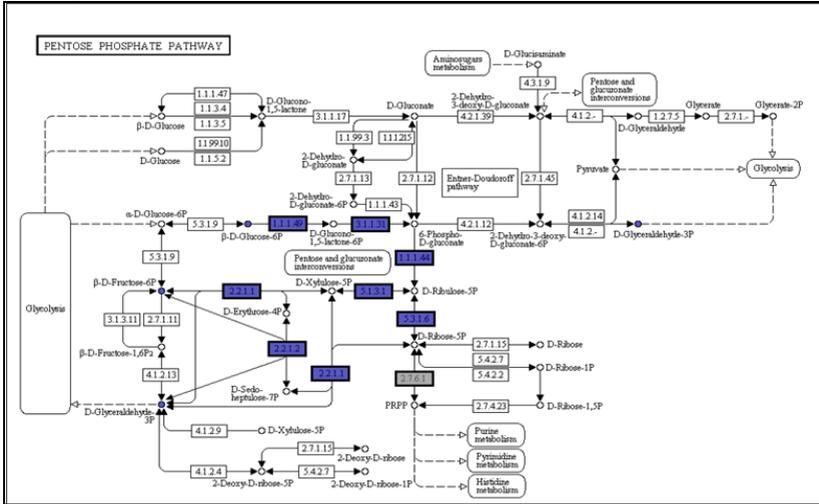
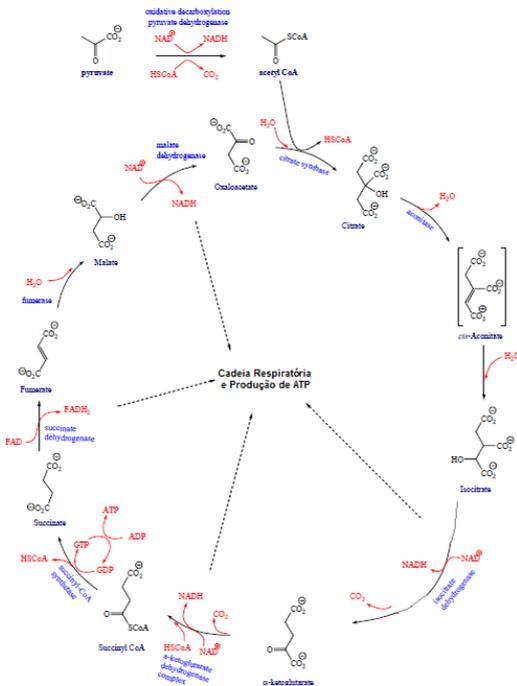


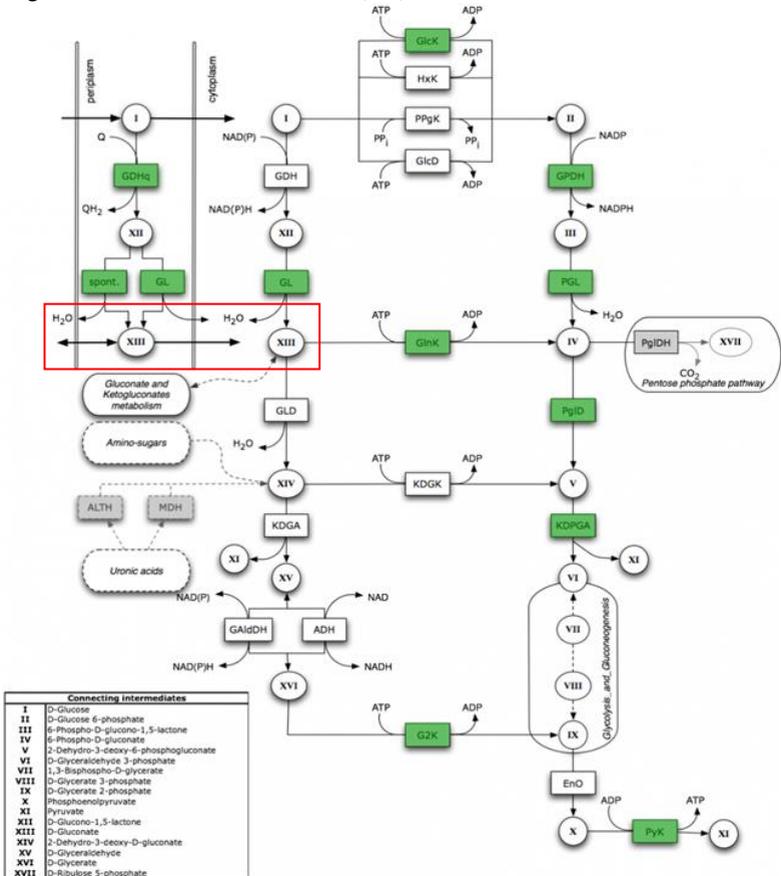
Figura 31 - Ciclo tricarboxílico da *G. hansenii*. Fonte: Model SEED



Via Entner-Doudoroff

A Figura 32 apresenta a via Entner-Doudoroff da *G. hansenii* que utiliza um conjunto de enzimas diferente da glicólise para catabolizar glicose (I) a piruvato (XI): enzimas fosfogluconato desidrogenase e 2-ceto-3-desoxi-6-fosfogluconato aldolase. Analisando esta via evidencia-se que esta bactéria pode oxidar glicose diretamente na forma de gluconato (XIII). Dependendo das condições ambientais, o ácido glucônico se acumula no sistema, o que foi comprovado através da análise de fluxo metabólico durante as simulações, mostradas adiante.

Figura 32 - Via Entner-Doudoroff (ED) da *G. hansenii*. Fonte: Model SEED

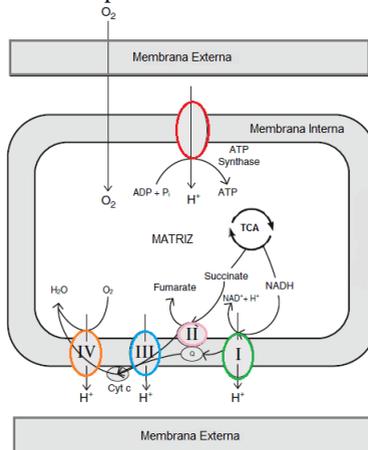


Transporte de elétrons e fosforilação oxidativa

A celulose bacteriana é formada na interface entre o meio de cultivo e o ar, assim entender a respiração deste organismo e as reações envolvidas é de interesse fundamental. Foram identificadas na *G. hansenii* as enzimas responsáveis pela transferência de elétrons do NADH para ubiquinona (ou coenzima Q), de ubiquinol para o citocromo c, e do citocromo c para o oxigênio.

A Figura 33 mostra a cadeia de transporte de elétrons, onde ocorre a fosforilação oxidativa. Todas as bactérias têm uma membrana plasmática onde ocorre a oxidação fosforilativa, uma vez que bactérias não possuem mitocôndrias. O NADH e succinato gerados no ciclo dos ácidos tricarboxílicos são oxidados liberando energia para a ATP sintase. O complexo I (NADH-desidrogenase) catalisa a transferência de elétrons do NADH para ubiquinona. O complexo II (Succ-desidrogenase) catalisa a transferência de elétrons do succinato para ubiquinona. O complexo III catalisa a transferência de elétrons da ubiquinona para citocromo c. E o complexo IV (citocromo oxidase) catalisa a transferência de elétrons do citocromo c para o oxigênio. Os elétrons são transferidos entre diferentes carreadores e a cada transferência de elétrons, um número de prótons é bombeado para fora, e esses retornam para o interior da membrana interna através de poros de prótons associados com a enzima ATP sintase. Esse movimento de prótons pelos poros fornece a energia necessária para a síntese de ATP.

Figura 33 – Cadeia de transporte de elétrons e fosforilação oxidativa.



Reação da biomassa

A reação de proliferação celular, isto é, de produção de biomassa foi escrita de acordo com dados experimentais obtidos para a bactéria *Gluconacetobacter xylinus* (ZHONG *et al.*, 2013), conforme parâmetros apresentados na Tabela 4.

Tabela 4 – Demanda dos precursores para *G. xylinus* (Zhong *et al.*, 2013)

Precursor	G6P	F6P	R5P	E4P	G3P	3PG	PEP
mmol/g	0,205	0,052	0,870	0,233	0,269	1,253	0,509
Precursor	PYR	ACCOA	AKG	OAA	NADH	NADPH	CO2
mmol/g	3,113	2,407	1,415	1,726	4,669	15,779	2,811

Uma função importante da célula é a produção dos chamados compostos precursores, ou *buildingblocks*. São esses precursores que podem gerar tudo dentro da célula desde seu DNA até as membranas celulares. Isso significa que a grande maioria das reações da *G. hansenii* em algum momento forma e/ou consome essas substâncias em suas etapas de reação. Portanto, um modelo que inclui a formação e consumo desses precursores é de grande interesse fisiológico.

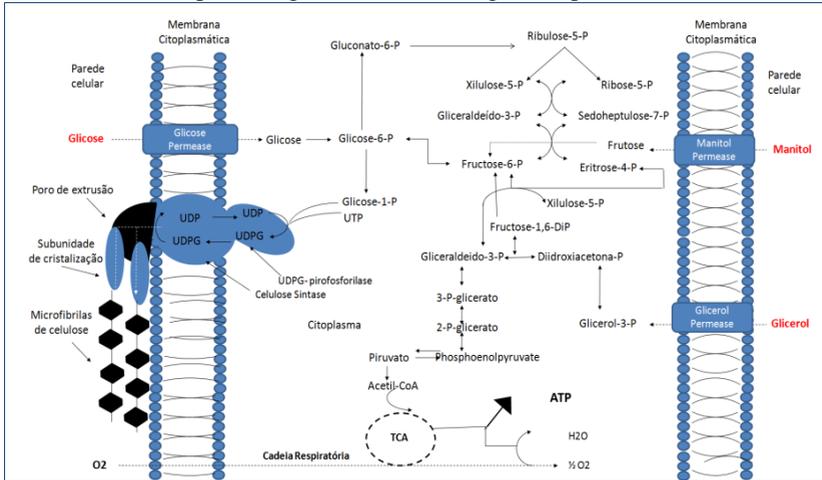
Em nosso estudo, a reação da biomassa foi modificada para incluir prótons internos e água. A quantidade de água necessária é igual à quantidade de ATP hidrolisado para satisfazer o requisito do crescimento associado à necessidade de ATP. A hidrólise do ATP resulta na produção de um próton, enquanto a utilização de NADPH e NADH consome um próton, o que resulta na produção de prótons da reação a biomassa. Ou seja, todos os precursores somados às moléculas para se realizar oxidações (NAD), reduções (NADPH) e fornecer energia (ATP), resultam em 1 mol de biomassa, que é a quantidade de biomassa produzida com esses compostos.

Experimentos com a bactéria em questão, dirigidos às necessidades metabólicas para a síntese de biomassa e manutenção associados e não-associados ao crescimento, são de grande interesse prático, mas atualmente inexistentes. A grande dificuldade de se descrever a reação da biomassa deve-se a que todos os precursores da reação precisavam estar sendo produzidos pelo sistema do contrário o modelo não é consistente. Durante as simulações com o GEnSys foram adicionadas reações de demanda que avaliam se os precursores estão devidamente produzidos e/ou consumidos pela célula.

4.2.2 Vias do metabolismo *core* da *G. hansenii*

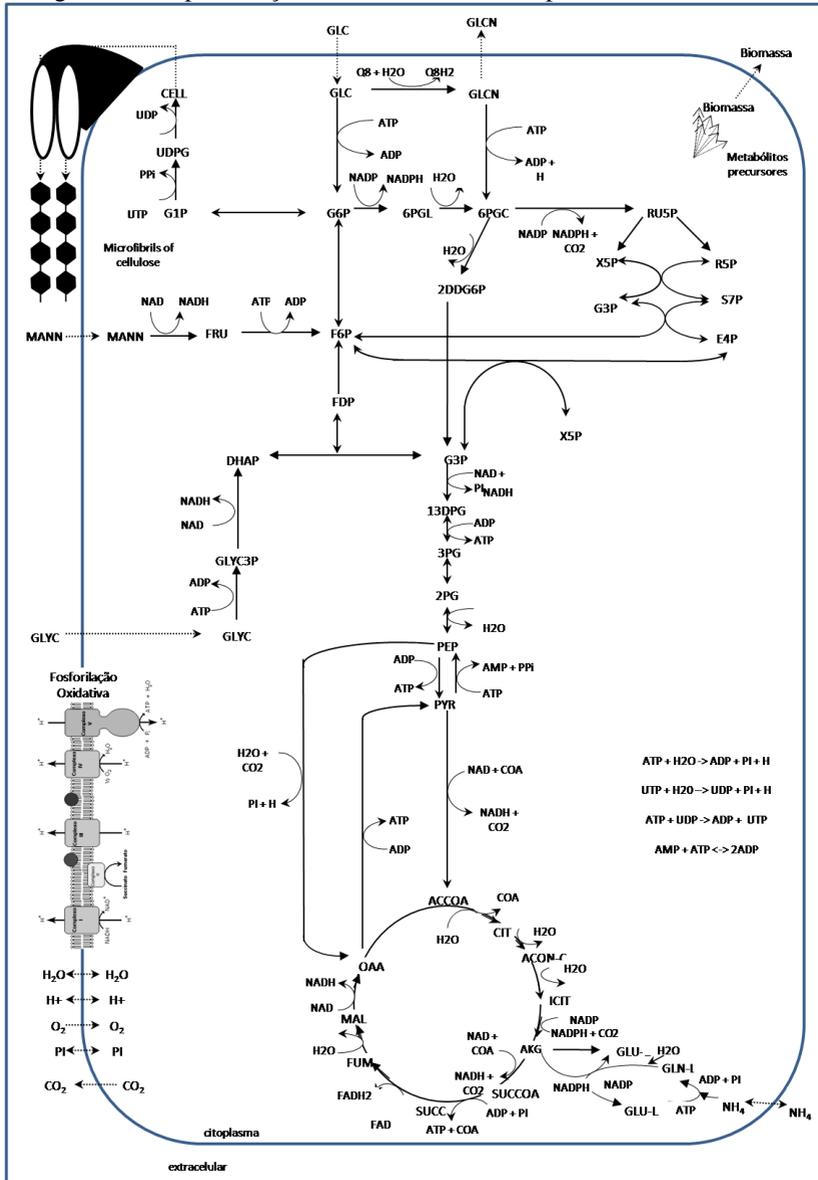
De modo simplificado a rota metabólica da bactéria *G. hansenii* para a biossíntese de celulose bacteriana a partir de glicose, manitol e glicerol pode ser vista na Figura 34.

Figura 34 - Modelo simplificado da rota metabólica para a biossíntese de celulose a partir de glicose, manitol ou glicerol por *G. hansenii*.



Tendo identificado as principais reações necessárias para mimetizar o metabolismo da *G. hansenii*, foi construída a rede com as vias para o metabolismo *core*, conforme Figura 35. As reações correspondentes ao modelo apresentado encontram-se no Apêndice C.

Figura 35 – Representação da rede metabólica simplificada da *G. hansenii*.



4.3 Conversão da construção metabólica para o formato computacional

Através das etapas anteriores foram incorporadas as restrições com relação à estequiometria (balanço químico), à termodinâmica (a questão da reversibilidade das reações) e à fisiologia (com respeito aos cofatores utilizados e à composição da biomassa). O produto dessa investigação gera um sistema de equações que descreve o metabolismo da célula envolvendo os metabólitos de interesse.

O modelo proposto consiste de uma rede representada matematicamente, através da matriz estequiométrica, considerando-se as condições de contorno e restrições nas reações envolvidas. O modelo final proposto como *core* consiste em 75 metabólitos e 74 reações. A hipótese de regime estacionário pode ser aplicada para 60 metabólitos, resultando em uma matriz estequiométrica de 60×74 .

A tarefa confiada a programas de análise de vias metabólicas como o COBRA e GENsYS foi a de encontrar possíveis soluções para esse sistema que digam quais são os possíveis fluxos de todas as reações envolvidas.

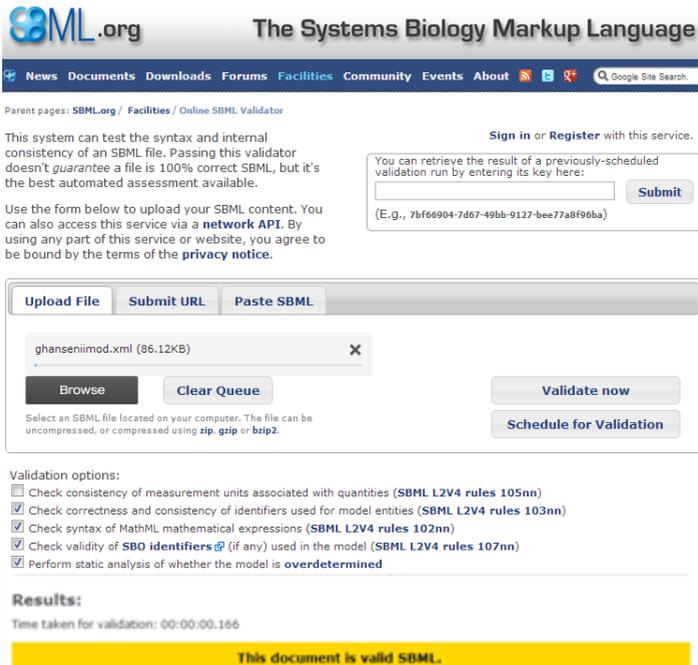
No COBRA *toolbox*, a partir da interface rBioNet, o modelo construído foi salvo como um modelo de reconstrução em uma estrutura de matriz estequiométrica. A Figura 36 ilustra os principais campos contidos na estrutura. As informações ficaram armazenadas em submatrizes.

Figura 36 – Componentes da estrutura do Modelo no MATLAB®

Field	Value	Min	Max
metS	<76x1 cell>		
metNames	<76x1 cell>		
metFormulas	<76x1 cell>		
rns	<73x1 cell>		
rznNames	<73x1 cell>		
subSystems	<73x1 cell>		
lb	<73x1 double>	-1000	0
ub	<73x1 double>	1000	1000
rev	<73x1 double>	0	1
c	<73x1 double>	0	0
b	<76x1 double>	0	0
S	<76x73 double>		
rznGeneMat	[]		
rules	<73x1 cell>		
grRules	<73x1 cell>		
genes	<0x1 cell>		
comments	<73x1 cell>		
citations	<73x1 cell>		
confidenceScores	<73x1 cell>		
ecNumbers	<73x1 cell>		
onkEggID	<73x1 cell>		
metCharge	<76x1 double>	NaN	NaN
metCHEBBID	<76x1 cell>		

O modelo construído no rBioNet foi convertido para SBML uma linguagem muito utilizada em modelos computacionais de processos biológicos. Para isso foi utilizada uma função específica do COBRA. O arquivo SBML foi validado ao utilizar a ferramenta SBML validator, de acordo com Figura 37.

Figura 37 - Validação do arquivo SBML



SBML.org The Systems Biology Markup Language

News Documents Downloads Forums Facilities Community Events About

Parent pages: SBML.org / Facilities / Online SBML Validator

This system can test the syntax and internal consistency of an SBML file. Passing this validator doesn't guarantee a file is 100% correct SBML, but it's the best automated assessment available.

Use the form below to upload your SBML content. You can also access this service via a **network API**. By using any part of this service or website, you agree to be bound by the terms of the **privacy notice**.

Sign in or Register with this service.

You can retrieve the result of a previously-scheduled validation run by entering its key here:

(E.g., 7bf66904-7d67-49bb-9127-bee77a8f96ba)

Upload File Submit URL Paste SBML

ghansenimod.xml (86.12KB)

Browse Clear Queue Validate now Schedule for Validation

Select an SBML file located on your computer. The file can be uncompressed, or compressed using **zip**, **gzip** or **bzip2**.

Validation options:

- Check consistency of measurement units associated with quantities (**SBML L2V4 rules 105nn**)
- Check correctness and consistency of identifiers used for model entities (**SBML L2V4 rules 103nn**)
- Check syntax of MathML mathematical expressions (**SBML L2V4 rules 102nn**)
- Check validity of **SBO identifiers** (if any) used in the model (**SBML L2V4 rules 107nn**)
- Perform static analysis of whether the model is **overdetermined**

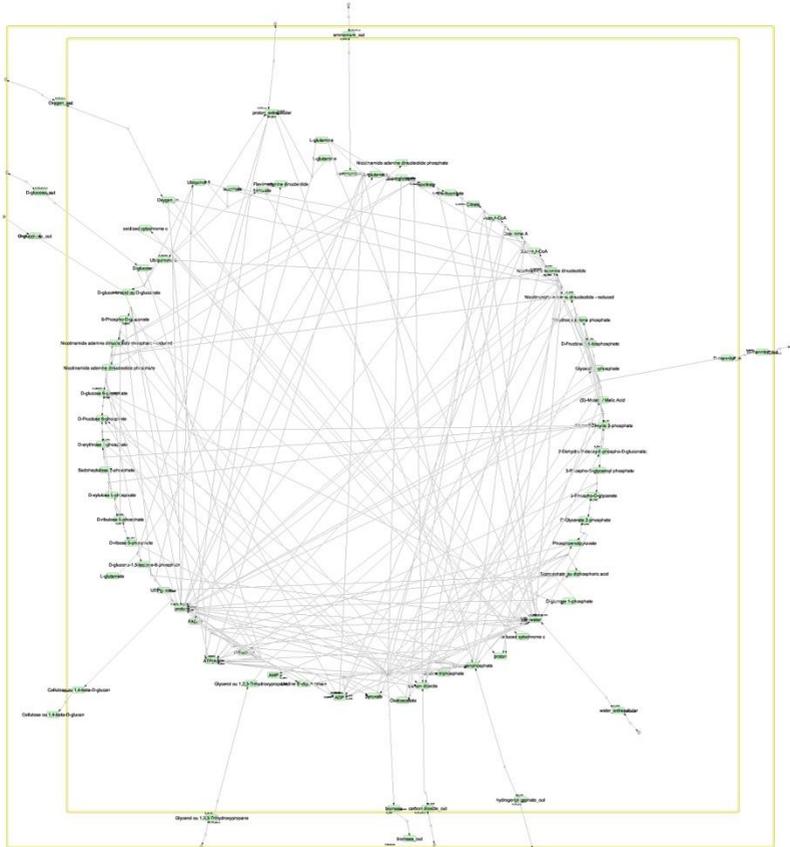
Results:

Time taken for validation: 00:00:00.166

This document is valid SBML.

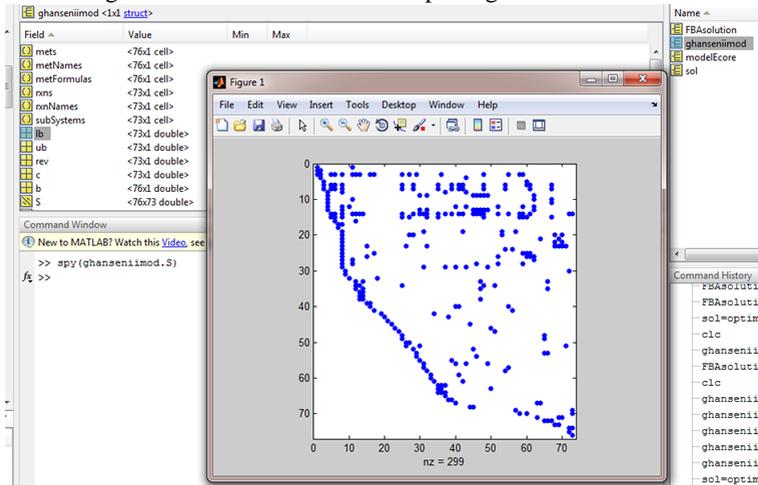
A partir do arquivo SBML diversas ferramentas computacionais puderam ser utilizadas. Optou-se pelo CellDesigner para importar este arquivo e visualizar o modelo. No CellDesigner foi possível visualizar as interconexões entre os metabólitos da rede. Esta é uma plataforma que consegue integrar todas as informações do arquivo em SBML e pode ser usada futuramente para editar o modelo e melhorá-lo com novas informações. Um dos *layouts* de visualização da rede pode ser visto na Figura 38.

Figura 38 - Layout gerado pelo CellDesigner



Nesta etapa do processo, diversas rotinas disponíveis no COBRA *toolbox* foram utilizadas. Uma dessas funções permitiu visualizar a matriz em uma representação diferente da matriz estequiométrica, na qual todos os pontos representam as entradas diferentes de zero (*non-zero*).

Figura 39 - Estrutura da matriz esparsa gerada no COBRA



Essa matriz conseguiu indicar a participação dos metabólitos em cada uma das reações. Existem alguns metabólitos que participam em muitas reações e com isso têm várias conexões, tais como ATP, NADPH, NADH, que são metabólitos-chave no metabolismo central. Estes metabólitos mais interconectados desempenham papéis importantes na rede metabólica: por exemplo, a estabilidade no transporte de tais metabólitos altamente ligados dentro ou fora da rede pode afetar o fenótipo metabólico *in silico* e, possivelmente, também *in vitro*. Além disso, a organização dos mecanismos de regulação pode ser significativamente afetada por tais metabólitos altamente conectados. Assim, a matriz esparsa consegue indicar a participação dos metabólitos em cada reação e, futuramente, em um modelo ampliado com regulação, pode ser muito útil para o entendimento das vias de interesse.

4.3.1 Análise de Fluxo Metabólico no GEnSys

Ao realizar a AFM do modelo simplificado para as condições determinadas no sistema com cada uma das fontes de carbono, primeiramente a função *MFA* do GEnSys permitiu visualizar as equações do balanço para cada metabólito, descrevendo sua dependência em função do tempo, conforme descrito abaixo:

```

ODE system:
d(GLC)/dt = v1_gli - v12 - v29
d(O2)/dt = v2 - v59 - v60 - 0.5v67 - 0.5v68
d(CO2)/dt = v3 + v22 + v24 - v25 + v33 + v48 + v49 + 2.811v69
d(BIOMASS)/dt = -v4 + v69
d(CELL)/dt = -v5_cell + v28
d(MANN)/dt = v6_mann - v41
d(NH3)/dt = v7 - v54 + v55
d(GLYC)/dt = v8_glyc - v43
d(H2O)/dt = v10 - v14 + v20 - v25 - v29 - v32 + v34 - v45 + v46 - v47 - v52 - v55 + 2v59 + 2v60 - v63 + v64 - v65 +
3.5v67 + 2.5v68 - 41.3v69
d(GLCN)/dt = -v11 + v29 - v30
d(ADP)/dt = v12 - v18 - v21 - v24 + v30 + v42 + v43 - v50 + v54 + 2v62 + v63 - v64 - 2.5v67 - 1.5v68 + 41.3v69
d(G6P)/dt = v12 - v13 - v26 - v31 - 0.205v69
d(F6P)/dt = v13 + v14 + v38 + v40 + v42 - 0.052v69
d(P1)/dt = v14 - v17 - v18 - v23 - v24 + v25 - v50 + v54 + v63 - v64 + 2v65 - 2.5v67 - 1.5v68 + 41.3v69
d(DHAP)/dt = v15 - v16 + v44
d(G3P)/dt = v15 + v16 - v17 + v35 - v38 + v39 + v40 - 0.269v69
d(NAD)/dt = -v17 - v22 - v41 - v44 - v49 - v53 + v57 - v66 + v67
d(13-DPG)/dt = v17 - v18
d(H)/dt = v17 - v21 + v25 - v27 + v41 + v42 + v44 + v45 + v48 + v53 + v54 - v56 - 5v57 - 4v59 - 4v60 - 2v61 + v63 +
4v64 + v65 - 3.5v67 - 2.5v68 + 41.3v69
d(3-PG)/dt = v18 - v19 - 1.253v69
d(2-PG)/dt = v19 - v20
d(PEP)/dt = v20 - v21 + v23 - v25 - 0.509v69
d(PYR)/dt = v21 - v22 - v23 + v24 + v35 - 3.113v69
d(COA)/dt = -v22 + v45 - v49 + v50
d(AMP)/dt = v23 - v62
d(PPI)/dt = v23 + v27 - v65
d(OAA)/dt = -v24 + v25 - v45 + v53 - 1.726v69
d(G1P)/dt = v26 - v27
d(UDPG)/dt = v27 - v28
d(6-PGC)/dt = v30 + v32 - v33 - v34
d(6-PGL)/dt = v31 - v32
d(NADPH)/dt = v31 + v33 + v48 - v56 - v66 - 15.779v69
d(RUSP)/dt = v33 - v36 - v37
d(2-DDG6P)/dt = v34 - v35
d(XUSP)/dt = v36 - v39 - v40
d(R5P)/dt = v37 - v39 - 0.87v69
d(57P)/dt = -v38 + v39
d(E4P)/dt = v38 - v40 - 0.233v69
d(FRU)/dt = v41 - v42
d(GLYC3P)/dt = v43 - v44
d(CIT)/dt = v45 - v46
d(ACON-C)/dt = v46 - v47
d(ICIT)/dt = v47 - v48
d(AKG)/dt = v48 - v49 - v56 - 1.415v69
d(SUCCOA)/dt = v49 - v50
d(SUCC)/dt = v50 - v51 - v58
d(FADH2)/dt = v51 - v68
d(FUM)/dt = v51 - v52 + v58
d(MAL-L)/dt = v52 - v53
d(GLN-L)/dt = v54 - v55 - v56
d(QSH2)/dt = v57 + v58 - 2v59 - v61
d(FERRICYTOCHROME)/dt = 4v60 - 2v61

```

Este sistema de equações diferenciais ordinárias foi resolvido pela hipótese de estado estacionário. O modelo estequiométrico ignora as dinâmicas das reações intracelulares e assume que os metabólitos internos encontram-se no regime estacionário.

A AFM forneceu uma primeira distribuição dos fluxos do modelo simplificado para cada fonte de carbono e algumas análises puderam ser feitas. Os gráficos gerados pelo GEnSys estão apresentados abaixo (Figuras 40 a 42).

Figura 40 - AFM do modelo com entrada de glicose

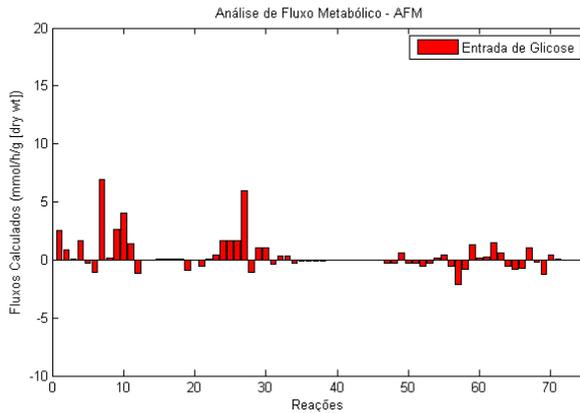


Figura 41 - AFM do modelo com entrada de manitol

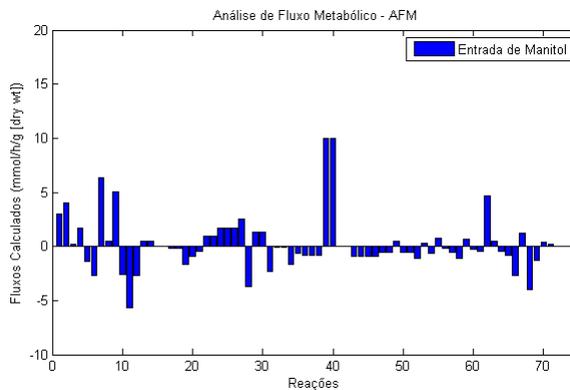
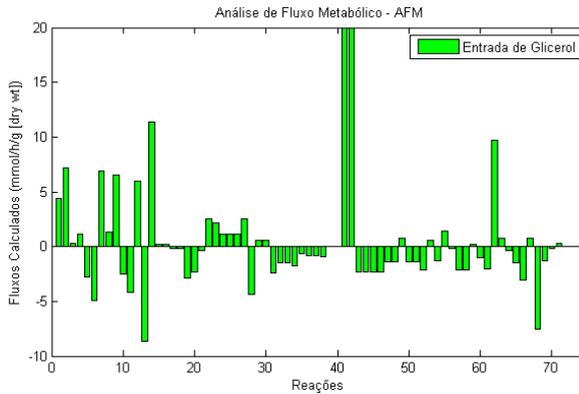
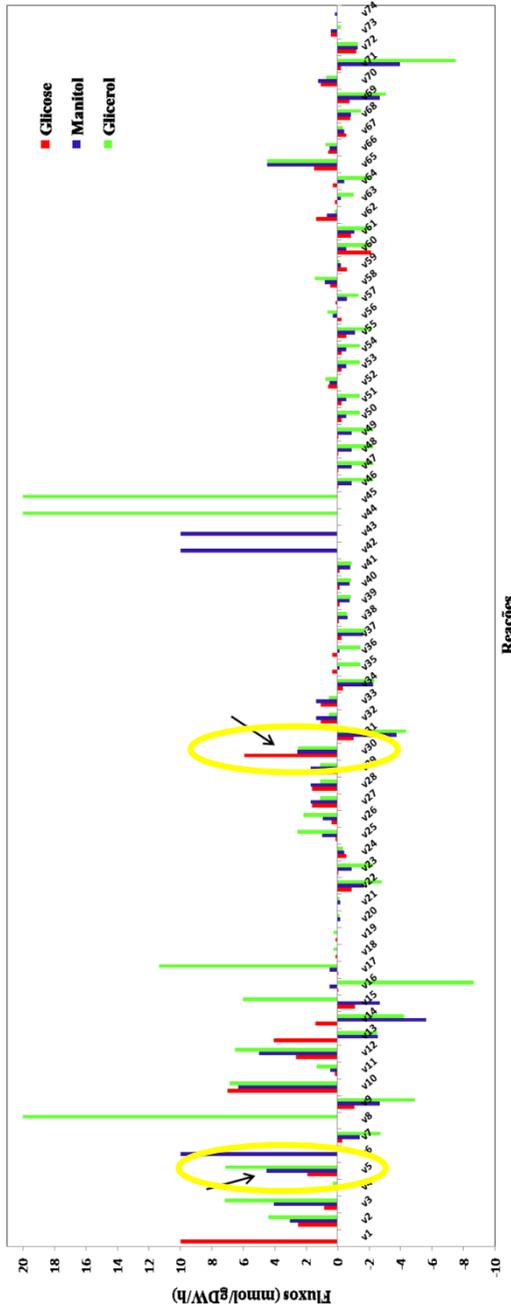


Figura 42 – AFM do modelo com entrada de glicerol



O resultado da distribuição dos fluxos para as três fontes de carbono, obtidos no GenSys, foram plotados em um mesmo eixo fornecendo um gráfico comparativo para analisar a variação nos fluxos, conforme Figura 43.

Figura 43 – Análise de fluxo metabólico para as três fontes de carbono. Os fluxos indicados referem-se às reações v5 e v30, referentes a produção de celulose e ácido gluconico, respectivamente.



O resultado das simulações indica a viabilidade da produção de celulose por *G. hansenii* a partir das três fontes de carbono avaliadas. Analisando o fluxo de produção de celulose (v5) de acordo com a Figura 43, verificou-se que 19%, 45,65% e 35,69% de glicose, manitol e glicerol foram usados para a biossíntese de CB, respectivamente.

Existem diversos estudos experimentais que avaliam a melhor fonte de carbono para a produção de celulose. Alguns estudos apontam o glicerol como melhor fonte de carbono, outros para o manitol como a melhor escolha. A escolha da fonte de carbono irá depender do meio mínimo de crescimento escolhido (Ruka *et al.*, 2012; Yamanaka *et al.*, 1989; Keshk e Sameshima, 2005; Schramm e Hestrin, 1954; Zhong *et al.*, 2013).

É importante ressaltar que o modelo metabólico utilizado neste trabalho é simplificado e que uma análise utilizando um modelo mais completo poderia apresentar um resultado mais confiável.

Ao analisar o fluxo de produção de celulose com glicose como fonte de carbono foi possível verificar um decréscimo neste fluxo comparado às outras fontes de carbono. Com isso foram analisados outros fluxos, o que levou a um resultado importante em relação a produção de ácido glucônico (v30). Ao utilizar glicose, este fluxo foi duas vezes superior ao obtido com a utilização de manitol e de glicerol.

A Tabela 5 apresenta os valores dos fluxos de ácido glucônico em $\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$ (gDW: grama de massa seca) do modelo simplificado.

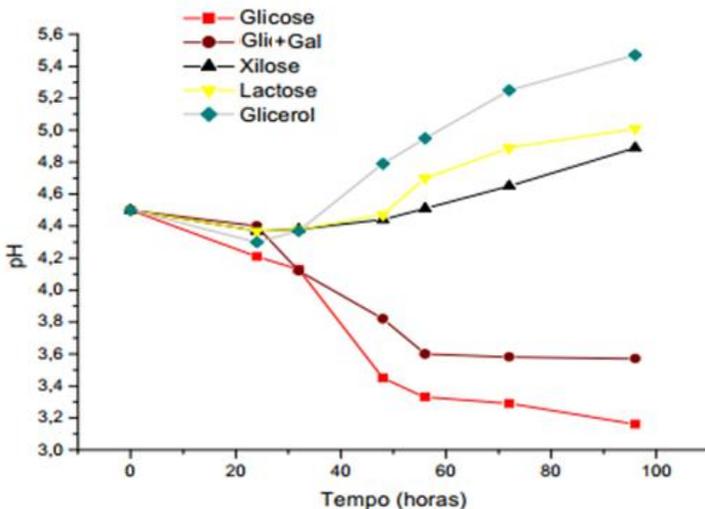
Tabela 5 - Valores dos fluxos de ácido glucônico para cada fonte de carbono através da AFM do modelo simplificado

Fonte de Carbono	Fluxo Ácido Glucônico ($\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$)
Glicose	5,9366
Manitol	2,5605
Glicerol	2,5281

A AFM conseguiu avaliar que utilizando como substrato a glicose, a produção de celulose é inferior devido ao desvio do fluxo metabólico para a produção do ácido glucônico como subproduto nesta via. De acordo com a literatura, esta bactéria oxida glicose diretamente na forma de ácido glucônico, o qual se acumula no sistema levando à redução do pH. Segundo Chawla (2009) e Carreira (2010), o pH dos meios de cultura contendo glicose como fonte de carbono decresce ao

longo do tempo. Isto é resultado das bactérias do gênero *Gluconacetobacter* durante a produção de celulose, que inclui formar um subproduto ácido no meio, o ácido glucônico. Cabe salientar que estes dados experimentais são para operação em batelada, e serviram como indicação para mostrar a redução do pH utilizando-se glicose (em vermelho) como fonte de carbono, conforme Figura 44.

Figura 44 – Variações de pH ao longo de 96h na produção de CB.
Fonte: CARREIRA (2010)



O objetivo desta primeira análise foi obter a distribuição de fluxos e avaliar se o modelo estava representando a estimativa a partir do conjunto de fluxos fornecidos. A AFM forneceu informações sobre a contribuição de cada reação para os processos metabólicos de utilização de substrato e formação de produto.

Quanto mais dados experimentais fornecidos ao modelo mais realística a AFM. Pode-se fornecer medições de fluxos extracelulares, ou até mesmo de alguns fluxos intracelulares, através da técnica ^{13}C MFA, que utiliza traçadores isotópicos, uma técnica onde os átomos de carbono marcados são distribuídos por toda a rede metabólica, até que o enriquecimento isotópico possa ser medido por técnicas de ressonância magnética e espectrometria de massa. Esses dados fornecem

informações para quantificar os fluxos intracelulares (WIECHERT, 2001; WIECHERT e NÖH, 2013).

4.3.2 Análise de Balanço de Fluxos no GENsYS

Na ABF (análise de balanço de fluxo), as restrições e a função objetivo foram colocadas no arquivo do modelo para serem processadas e a distribuição de fluxo determinada de forma a otimizar a função objetivo.

A função FBA retornou a estrutura *FBAmodel* contendo os seguintes campos:

- **intModel**: estrutura que contém o modelo interno (apenas os metabólitos internos). Um dos campos é a matriz estequiométrica (m metabólitos por n reações).
- **beq**: vetor $m \times 1$ com todos elementos iguais a zero. Reservado para uso futuro, como recurso para adicionar fluxos com valores conhecidos;
- **c**: vetor $n \times 1$ contendo os n coeficientes da função objetivo. É determinado pela soma das linhas correspondentes aos metabólitos a serem otimizados;
- **lb**: vetor $n \times 1$ onde são armazenados os limites inferiores das velocidades dos fluxos (velocidade mínima / *lower bound*);
- **ub**: vetor $n \times 1$ onde são armazenados os limites superiores das velocidades dos fluxos (velocidade máxima / *upper bound*);
- **minMax**: tipo do objetivo: -1 para maximizar e 1 para minimizar.

O modelo da *G. hansenii* foi gerado e apresentou a seguinte estrutura:

m =

```
intModel: [1x1 struct]
beq: [60x1 double]
c: [1x74 double]
lb: [74x1 double]
ub: [74x1 double]
minMax: -1
```

A função retornou em uma matriz estequiométrica de 60 metabólitos e 74 reações, com o propósito de maximizar uma função objetivo.

4.3.2.1 Meio Mínimo e Funções Objetivo

A partir do meio mínimo definido, cada fonte de carbono entrava no modelo, uma por uma, adicionando reações de troca (caso não houvesse uma correspondente) para simular o crescimento do modelo sob diferentes condições ambientais. Para determinar as taxas de consumo das fontes de carbono, nitrogênio e fosfato de acordo com a composição do meio mínimo, foi necessário utilizar dados experimentais em cultura contínua, pois a ABF assume regime estacionário e gera previsões que são consistentes com a cultura contínua.

Existem poucas informações disponíveis sobre este organismo, a *G. hansenii*, no que se refere a experimentos em cultura contínua, ou seja, que avaliaram o seu crescimento e produção de celulose, taxas de consumo de substrato, taxa de produção de metabólitos externos e concentração de biomassa. Por isso, alguns dados foram baseados em experimentos de cultura contínua com a bactéria *Gluconobacter oxydans* (HANKE *et al.*, 2013; OLIJVE e KOK, 1979), utilizando-se alguns parâmetros da mesma, como a concentração da biomassa e a taxa de diluição. Os parâmetros e dados necessários para calcular as taxas foram definidos conforme Tabela 6.

Tabela 6 – Parâmetros e dados para calcular as taxas de entrada dos componentes do meio mínimo

X: Concentração de biomassa (g/L)	1,5 ^a
D: Taxa de Diluição (h ⁻¹)	0,05 ^b
PM: Massa Molar Glicose/Manitol (g/mol)	180
PM: Massa Molar Glicerol (g/mol)	92
PM: Massa Molar (NH ₄) ₂ SO ₄ (g/mol)	132
PM: Massa Molar KH ₂ PO ₄ (g/mol)	136
C: Concentração fonte de carbono (g/L) - Meio Yamanaka	50 ^c
C: Concentração fonte de nitrogênio (g/L) - Meio Yamanaka	5 ^c
C: Concentração fonte de fosfato (g/L) - Meio Yamanaka	3 ^c

^a(HANKE *et al.*, 2013); ^b(OLIJVE E KOK, 1979); ^c(RUKA *et al.*, 2012)

Com os dados da Tabela 6, foram calculadas as taxas de consumo, que correspondem às condições de contorno utilizadas como dados de entrada no modelo computacional. As taxas foram obtidas dividindo-se a concentração ($\text{g}\cdot\text{L}^{-1}$) de cada componente pela sua massa molar ($\text{g}\cdot\text{mol}^{-1}$), depois multiplicando-se pela vazão específica de alimentação (h^{-1}) e dividindo-se pela quantidade de biomassa ($\text{gDW}\cdot\text{L}^{-1}$), conforme Equação 4:

$$r\left(\frac{\text{mmol}}{\text{gDW}\cdot\text{h}}\right) = \frac{C\left(\frac{\text{g}}{\text{L}}\right)}{PM\left(\frac{\text{g}}{\text{mol}}\right)} \times \frac{D(\text{h}^{-1})}{X\left(\frac{\text{gDW}}{\text{L}}\right)} \times \frac{1000\text{mmol}}{1\text{mol}} \quad (4)$$

Tabela 7 – Taxas de consumo dos componentes do meio mínimo em $\text{mmol}\cdot\text{gDW}^{-1}\cdot\text{h}^{-1}$

Taxa de consumo Glicose/Manitol	10
Taxa de consumo Glicerol	20
Taxa de consumo $(\text{NH}_4)_2\text{SO}_4$	1,3
Taxa de consumo KH_2PO_4	0,75

A Tabela 8 apresenta a composição do meio definido, sendo apenas alterada a fonte de carbono glicose, por manitol ou glicerol. As taxas de consumo calculadas referem-se aos limites das reações.

Tabela 8 – Composição de meio mínimo com as condições de contorno

Descrição Reação	Equação	LB *	UB *
Glicose	[e]: glc-D ->	0	10
Manitol	[e]: mann ->	0	10
Glicerol	[e]: glic ->	0	20
O ₂ troca	[e]: o ₂ <->	-1000	1000
NH ₄ troca (fonte de nitrogênio)	[e]: nh ₄ <->	-1,3	1000
PO ₄ troca (fonte de fosfato)	[e]: pi <->	-0,78	1000
H ₂ O troca	[e]: h ₂ o <->	-1000	1000
CO ₂ troca	[e]: co ₂ <->	-1000	1000
Próton troca	[e]: h <->	-1000	1000

*LB: *lower bound* (limite inferior); UB: *upper bound* (limite superior), cuja unidade é ($\text{mmol}/\text{gDW}/\text{h}$).

Os outros metabólitos externos definidos no meio mínimo, como O_2 , H^+ , CO_2 e H_2O mantiveram as condições de contorno com entrada ilimitada.

Com todas as restrições adicionadas ao modelo metabólico, desde a estequiometria e reversibilidade durante a etapa de cura manual, e os limites superiores e inferiores das reações durante a etapa computacional, aplicou-se programação linear (otimização) definindo uma das reações do modelo como função objetivo. A escolha da função objetivo ideal foi baseada em argumentos biológicos relevantes para o problema em questão.

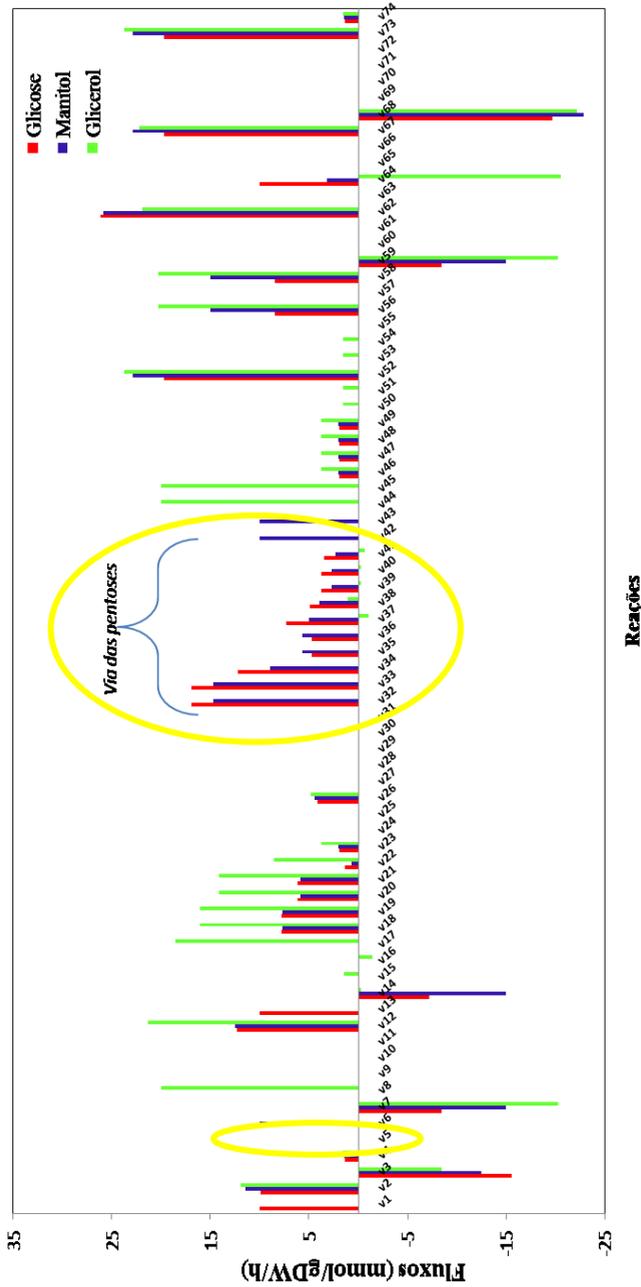
Cabe salientar, como (ORTH *et al.*, 2010) adverte, que a Análise de Balanço de Fluxo tem limitações, uma vez que não levam em conta os parâmetros cinéticos ou efeitos reguladores, logo, não pode oferecer previsões exatas de concentrações metabólicas.

A seguir são apresentados os estudos de caso (quatro, no total) que foram analisados neste trabalho:

4.3.2.2 Caso I: Maximizando a produção de biomassa

Foi obtida a distribuição de fluxos mimetizando uma situação de maximização da biomassa, ou seja, a taxa de compostos metabólicos foi convertida em componentes da biomassa. A distribuição dos fluxos foi analisada comparando as fontes de carbono utilizadas, conforme Figura 45.

Figura 45 – Análise de Balanço de Fluxo para as três fontes de carbono com o crescimento celular como função objetivo para maximizar. Os fluxos indicados referem-se a produção de celulose (v5) e a via das pentoses identificada nas reações v31 a v41.



O resultado das simulações, ao maximizar-se a produção de biomassa indicou que o fluxo de carbonos foi utilizado para o crescimento da bactéria, sem haver produção de celulose. O fluxo referente à produção de celulose (v5) foi nulo. Isso é um caso possível, uma vez que as bactérias não necessitam de celulose para crescer, é uma estratégia de sobrevivência e serve para as bactérias manterem a sua posição perto dos nutrientes, para se proteger contra os inimigos e radiação ultravioleta, e migrar para superfícies de alcance do oxigênio (IGUCHI *et al.*, 2000).

As taxas de crescimento para as condições usadas neste estudo utilizando cada uma das fontes de carbono estão descritas na Tabela 9.

Tabela 9 - Taxas de crescimento obtidos pela ABF para os três substratos ao maximizar o crescimento celular.

Fonte de Carbono	Taxa de crescimento (h^{-1})
Glicose	1,3303
Manitol	1,4209
Glicerol	1,5420

Também é importante ressaltar os fluxos referentes via das pentoses (PPP) (fluxos v31–v41), apresentados na Figura 45. A ABF encontrou uma indicação de que a utilização de glicerol como fonte de carbono, pela via das pentoses, não foi favorecida teoricamente. Além disso, os valores das taxas de crescimento mostraram uma dependência com essa fração de carbono direcionada para a via das pentoses. Se a bactéria tem como objetivo o seu crescimento, a ABF mostrou que a melhor função objetivo foi a maximização da biomassa utilizando glicerol como substrato.

4.3.2.3 Caso II: Maximizando a produção de celulose

O segundo caso estudado foi a simulação de maximização da celulose para as mesmas condições de contorno do meio mínimo, de acordo com a Tabela 8. Determinou-se a distribuição de fluxos com as restrições impostas ao modelo, para as três fontes de carbono em estudo. A função *solveFBA*, do GenSys, retornou o valor da função objetivo calculada (f), utilizando o *solver* GLPK. O campo x é o vetor de tamanho n (74 neste caso) com os fluxos calculados. O *status* do *solver*

(*stat*) quando igual a 1 significa solução ideal. Os resultados estão descritos na Tabela 10.

Tabela 10 – Resultados das simulações com ABF para maximização da celulose com entrada do meio mínimo de crescimento

Glicose	Manitol	Glicerol
s =	s =	s =
f: 10	f: 10	f: 10
x: [74x1 double]	x: [74x1 double]	x: [74x1 double]
stat: 1	stat: 1	stat: 1
solver: 'glpk'	solver: 'glpk'	solver: 'glpk'
time: 0.0034	time: 0.0031	time: 0.0035

Os resultados das simulações mostraram que a bactéria tem a capacidade de utilizar cada uma das fontes de carbono para estas condições de crescimento. Se o objetivo for produzir celulose a maximização da celulose é uma boa função objetivo e prediz o rendimento teórico máximo de produção de celulose, uma vez que para estas condições a taxa de crescimento foi nula. Assim, os fluxos foram direcionados para a produção de celulose e não houve exportação de carbono para produzir biomassa. Isso também sugere que o modelo é consistente, apresentando os resultados que seriam teoricamente esperados.

Os rendimentos teóricos máximos para produção de celulose foram calculados: 0,95g/g glicose; 0,95g/g manitol e 0,93g/g glicerol. Esses limites teóricos podem ser utilizados para verificar se os rendimentos de crescimento experimentais não excedem os máximos teóricos dados pelo modelo, e verificar se o modelo mostra concordância com os dados experimentais, ao mesmo tempo em que rejeita dados inválidos gerados.

4.3.2.4 Caso III: Maximizando metabólitos externos e avaliando limitação de nutrientes

Para avaliar a consistência do modelo, analisou-se o balanço de massa, realizando algumas simulações maximizando os metabólitos externos. Por exemplo, ao maximizar gás carbônico entrando com 10 mmol/gDW/h de glicose no modelo, a resposta da análise de fluxo foi 60 mmol·gDW⁻¹·h⁻¹ de CO₂. O resultado é coerente, uma vez que a glicose tem seis carbonos e a molécula de gás carbônico tem apenas um

único carbono. Ao maximizar outros metabólitos externos que não apresentam carbono a análise de fluxo apresentou resultados nulos.

O crescimento microbiano e a formação de produtos metabólicos são condicionados à disponibilidade de nutrientes no meio de cultivo. As fontes de nitrogênio, fosfato e entrada de oxigênio também foram avaliadas no modelo. Definindo as taxas de entrada de NH_4 , PO_4 e O_2 como $0 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{h}^{-1}$, a bactéria não consegue crescer, obtendo-se um distribuição de fluxos nula. Porém quando teve limitação de nitrogênio ou de fosfato apenas, isso não foi suficiente para impedir o seu crescimento. De acordo com Ross (1991), em *Acetobacter xylinum* células com limitação de uma fonte de nitrogênio, a produção de celulose se manteve quando alimentada com uma quantidade adequada de fonte de carbono. Sendo assim, a presença de excesso de substrato disponível e limitações de outros nutrientes podem promover a produção de celulose.

A taxa de oxigênio foi definida como ilimitada, visto que não há dados disponíveis na literatura ou experimentais.

4.3.2.5 Caso IV: Outras simulações: Reações de Demanda

O fluxo de biomassa refere-se a um agrupamento do resultado de vias e distribuição de fluxos de carbono referente aos precursores. A consistência dos fluxos implica que cada um dos metabólitos precursores seja produzido pelo microrganismo. As simulações mostraram que o modelo produziu cada um dos constituintes de biomassa, com todas as três fontes de carbono, de forma consistente com o que deveria ser esperado.

4.4 Validação da Rede Metabólica

A consistência do modelo teórico proposto foi possível através das simulações e observações realizadas. De acordo com as vias previamente estabelecidas, as condições de contorno e o meio de crescimento, foi possível realizar as análises do modelo simplificado e comprovar a sua consistência.

Nesta etapa os erros mais comuns que surgem nas reconstruções metabólicas são: restrições em reações incorretas; falta de reações de

transporte; cofatores podem não ser produzidos ou consumidos; precursores não estão sendo produzidos; matriz estequiométrica com solução indeterminada, impedindo que o modelo mimetize o crescimento da bactéria e resultando em uma velocidade de crescimento nula.

Segundo LLANERAS (2010), os modelos baseados em restrição são frequentemente construídos para redes grandes ou em escala genômica de organismos conhecidos, mas também existem redes mais simples com apenas alguns metabólitos-chave. Desta forma, os dados disponíveis são explorados para construir modelos simplificados mais confiáveis. O procedimento pode ser estendido para detectar limitações do modelo e guiar a sua melhoria.

Esse processo de reconstrução é cíclico, e o modelo pode ser sempre estendido e seus parâmetros ajustados para melhorar a consistência com as situações investigadas. Os requisitos de manutenção podem ser experimentalmente determinados e adicionados como restrições do modelo. É possível combinar o conhecimento dos mecanismos metabólicos e reguladores. A rede reguladora de genes é fundamental para se entender a complexidade da interdependência entre os genes. Os métodos *in silico* representam um caminho promissor no objetivo de extrair redes reguladoras de genes de conhecimento biológico prévios, com os dados genômicos disponibilizados. Para o caso em questão, isto é, para produção de celulose por *G. hansenii*, os principais genes da rede metabólica foram identificados, conforme consta do Apêndice C.

CAPÍTULO 5

CONCLUSÕES, CONSIDERAÇÕES FINAIS E SUGESTÕES

A construção da rede metabólica e as ferramentas de engenharia metabólica foram utilizadas para entender a biossíntese de celulose bacteriana a partir de diferentes fontes de carbono da *Gluconacetobacter hansenii* ATCC 23769.

As modificações no protocolo proposto pelo grupo de Palsson, desenvolvido para reconstruções metabólicas de alcance genômico, permitiram a construção do primeiro modelo *core* da *G. hansenii*. A criação de novos bancos de dados como o PGDB GhaCyc, e ferramentas correspondentes, a seleção de vias específicas para compor a rede simplificada e o uso do *toolbox* GEnSys desenvolvido pelo grupo, formam algumas das principais alterações adicionadas ao protocolo.

Utilizando ferramentas computacionais como o Pathway Tools e Model SEED, muitas informações sobre o organismo foram obtidas a partir do genoma anotado. A etapa de cura manual, através da análise conjunta de diversos bancos de dados disponíveis, permitiu integrar diversos dados e construir o modelo *core* da rede metabólica. A metodologia utilizada nesse trabalho pode servir para compilação de muitos dados disponíveis na literatura com intuito de propor rotas metabólicas alternativas e assim poder fazer um estudo mais aprofundado do metabolismo a ser estudado.

A determinação das funções objetivo que foram consideradas relevantes foi essencial para os estudos *in silico*. Além disso, o meio mínimo de crescimento da bactéria definido pelas simulações foi fundamental para compor um modelo metabólico consistente com a realidade.

A rede metabólica simplificada foi representativa e consistente sob todos os limites e condições de contorno definidos. A partir da conversão desta reconstrução para um formato matemático-computacional, os métodos de AFM e ABF foram utilizados para investigar diversas possibilidades. A aplicação de modelos baseados em restrição mostrou que muitas informações valiosas podem ser extraídas a partir deles, mesmo se a cinética intracelular é desconhecida. Através de análise *in silico* realizada por meio da técnica de Análise de Balanço de Fluxo, foi possível prever crescimento e produção de celulose, em condições ótimas para casos limites. A previsão realizada pela análise

computacional foi ponderada levando em conta seu caráter qualitativo e suas simplificações inerentes.

A AFM forneceu uma primeira distribuição de fluxos variando as fontes de carbono: glicose, manitol e glicerol. Ao entrar no sistema com glicose, o fluxo de produção de ácido glucônico é superior, e com isso o fluxo de produção de celulose é inferior, conforme esperado teoricamente. O modelo conseguiu avaliar o que já havia sido evidenciado experimentalmente, tal como a redução do pH do meio, a formação de um bioproduto, o ácido glucônico e o rendimento de celulose inferior, quando a glicose é utilizada como única fonte de carbono. Analisando o fluxo de produção de celulose, indicou que 19,0%, 45,6% e 35,7% de glicose, manitol e glicerol foram usados para biossíntese de CB, respectivamente.

Através da ABF foi avaliado se o modelo era capaz de prever o fenótipo que as células vão mostrar em determinadas condições. Para isso, foram realizadas simulações em diferentes condições de otimização (maximização do crescimento celular e maximização da produção de celulose) e de restrições (com e sem restrição das velocidades específicas de consumo das fontes de carbono, nitrogênio e fosfato). As restrições impostas para estas velocidades específicas se basearam na composição do meio mínimo Yamanaka. A descrição desse meio de cultura foi de extrema importância para a obtenção de um modelo preditivo.

Ao analisar a distribuição de fluxos maximizando a produção de celulose, os resultados mostraram que a bactéria tem a capacidade de utilizar cada uma das fontes de carbono para as condições de crescimento impostas. Se o objetivo for produzir celulose a maximização da celulose é uma boa função objetivo e prediz o rendimento teórico máximo de produção de celulose, uma vez que para estas condições a taxa de crescimento foi nula. Assim, os fluxos foram direcionados para a produção de celulose, e não houve exportação de carbono para produzir biomassa.

Ao analisar a distribuição de fluxos maximizando o crescimento celular o fluxo de carbonos foi utilizado para o crescimento da bactéria, sem haver produção de celulose. É um caso possível, uma vez que as bactérias não necessitam de celulose para crescer; de outra forma é uma estratégia de sobrevivência e serve para as bactérias manterem a sua posição perto dos nutrientes, para se protegerem contra os inimigos e radiação ultravioleta, e flutuar sobre superfícies de alcance do oxigênio.

As taxas de crescimento específicos calculadas por hora foram: em glicerol, $\mu = 1,5420$; em manitol, $\mu = 1,4209$; e em glicose, $\mu = 1,3303$, sob as condições de crescimento usadas neste estudo. Analisando o vetor de distribuição de fluxo para estas restrições, mostrou-se que utilizando glicerol como fonte de carbono, os fluxos na via das pentoses não foram favorecidos assim, a taxa de crescimento foi dependente da fração de carbono direcionada por esta via.

Outras simulações investigaram se o modelo era consistente, através da análise do balanço de massa, maximizando-se alguns metabólitos externos. Além disso, foram feitas simulações com limitação de nutrientes. Ao definir as taxas de entrada de NH_4^+ , PO_4^{3-} e O_2 para 0 mmol/gDW/h , a bactéria não conseguiu crescer, obtendo-se um distribuição de fluxos nula. Porém, com limitação de nitrogênio ou de fosfato apenas, isso não foi suficiente para impedir o seu crescimento. Segundo a literatura, se houver limitação de algum nutriente, este organismo continua promovendo a produção de celulose se houver excesso de substrato disponível.

Outro resultado importante foi através da adição de reações de demanda. A consistência dos fluxos implicou que cada um dos metabólitos precursores precisa ser produzido pelo organismo. Ao incluir uma reação de demanda para cada metabólito da reação da biomassa e otimizar para este fluxo, os resultados mostraram que o modelo produz cada um dos constituintes de biomassa, com todas as três fontes de carbono.

SUGESTÕES FUTURAS

Uma reconstrução metabólica sempre pode ser melhorada. Como sugestões futuras propõem-se a melhoria da rede através da inserção de mais informações ao modelo e a cura manual de um número maior de reações. Um modelo mais complexo e com mais informações pode, em princípio, prever melhor o comportamento do microorganismo. A adição de mecanismos reguladores à rede e entender como o organismo utiliza essas estratégias de regulação para o controle de expressão gênica.

Algumas análises experimentais podem fornecer informações adicionais para essas simulações computacionais. Além disso, a cada dia novas ferramentas computacionais surgem, combinando novas técnicas e gerando mais informações.

Por fim, a utilização de modelos metabólicos é de extrema importância quando se pretende apontar estratégias e propor modificações no metabolismo. Este conhecimento pode potencialmente levar a uma intervenção no interior das células, sugerindo metabólitos alvo ou bifurcação bioquímica e a otimização do processo através da manipulação de variáveis externas. Assim as técnicas de engenharia metabólica podem ser extremamente úteis para previsão de crescimento e produção de metabólitos de interesse, e devem ser utilizadas sempre com apoio de procedimentos experimentais em respeito às suas limitações intrínsecas.

REFERÊNCIAS BIBLIOGRÁFICAS

AIBA, S.; MATSUOKA, M. Identification of metabolic model: Citrate production from glucose by *Candida lipolytica*. **Biotechnology and Bioengineering**, v. 21, n. 8, p. 1373-1386, 1979. ISSN 1097-0290.

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Molecular Biology of the Cell**. 4. New York: 2002.

ALMAAS, E.; KOVACS, B.; VICSEK, T.; OLTVAI, Z.; BARABASI, A. Global organization of metabolic fluxes in the bacterium *Escherichia coli*. **Nature**, v. 427, n. 6977, p. 839 - 843, 2004.

BACKDAHL, H.; ESGUERRA, M.; DELBRO, D.; RISBERG, B.; GATENHOLM, P. Engineering microporosity in bacterial cellulose scaffolds. **J Tissue Eng Regen Med**, v. 2, n. 6, p. 320-30, Aug 2008.

BACKDAHL, H.; HELENIUS, G.; BODIN, A.; NANNMARK, U.; JOHANSSON, B. R.; RISBERG, B.; GATENHOLM, P. Mechanical properties of bacterial cellulose and interactions with smooth muscle cells. **Biomaterials**, v. 27, n. 9, p. 2141-9, Mar 2006.

BAILEY, R. B.; WOODWORD, A. Isolation and characterization of a pleiotropic glucose repression resistant mutant of *Saccharomyces cerevisiae*. **Mol Gen Genet**, v. 193, n. 3, p. 507-12, 1984.

BARABASI, A.; OLTVAI, Z. Network biology: understanding the cell's functional organization. **Nature reviews**, v. 5, n. 2, p. 101 - 113, 2004.

BASSETT, D. E., JR.; EISEN, M. B.; BOGUSKI, M. S. Gene expression informatics--it's all in your mine. **Nat Genet**, v. 21, n. 1 Suppl, p. 51-5, Jan 1999.

BAXEVANIS, A. D. Information retrieval from biological databases. **Methods Biochem Anal**, v. 43, p. 155-85, 2001.

BEARD, D. A.; LIANG, S. D.; QIAN, H. Energy balance for analysis of complex metabolic networks. **Biophys J**, v. 83, n. 1, p. 79-86, Jul 2002.

BECKER, S.; FEIST, A.; MO, M.; HANNUM, G.; PALSSON, B.; HERRGARD, M. Quantitative Prediction of Cellular Metabolism with Constraint-based Models: The COBRA Toolbox. **Nat Protoc**, v. 2, n. 3, p. 727 - 738, 2007.

BECKER, S.; PALSSON, B. Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: an initial draft to the two-dimensional annotation. **BMC Microbiol**, v. 5, p. 8, 2005a.

BECKER, S.; PALSSON, B. Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: an initial draft to the two-dimensional annotation. **BMC Microbiol**, v. 5, n. 1, p. 8, 2005b.

BENZIMAN, M.; HAIGLER, C. H.; BROWN, R. M.; WHITE, A. R.; COOPER, K. M. Cellulose biogenesis: Polymerization and crystallization are coupled processes in *Acetobacter xylinum*. **Proc Natl Acad Sci U S A**, v. 77, n. 11, p. 6678-82, Nov 1980.

BERTI, F. V. **Desenvolvimento de Estruturas Vasculares Endotelizadas em Scaffolds de Celulose Bacteriana**. 2012. 147 (Doutorado). Programa de Pós Graduação em Engenharia Química, Universidade Federal de Santa Catarina, Florianópolis.

BLANK, L. M.; EBERT, B. E. From measurement to implementation of metabolic fluxes. **Current Opinion in Biotechnology**, v. 24, n. 1, p. 13-21, 2013.

BODIN, A.; BACKDAHL, H.; FINK, H.; GUSTAFSSON, L.; RISBERG, B.; GATENHOLM, P. Influence of cultivation conditions on mechanical and morphological properties of bacterial cellulose tubes. **Biotechnol Bioeng**, v. 97, n. 2, p. 425-34, Jun 1 2007.

BORDBAR, A.; PALSSON, B. O. Using the reconstructed genome-scale human metabolic network to study physiology and pathology. **J Intern Med**, v. 271, n. 2, p. 131-41, Feb 2012.

BORK, P.; SERRANO, L. Towards cellular systems in 4D. **Cell**, v. 121, n. 4, p. 507-9, May 20 2005.

- BRINDLE, K. M. Analysis of metabolic control in vivo using molecular genetics. **Cell Biochem Funct**, v. 14, n. 4, p. 269-76, Dec 1996.
- BROWN, A. J. XIX.-The chemical action of pure cultivations of bacterium aceti. **Journal of the Chemical Society, Transactions**, v. 49, p. 172-187, 1886.
- BROWN, R. M. Bacterial cellulose: Its potential for new products of commerce. **Abstracts of Papers of the American Chemical Society**, v. 227, 2004.
- BROWN, R. M., JR.; WILLISON, J. H.; RICHARDSON, C. L. Cellulose biosynthesis in *Acetobacter xylinum*: visualization of the site of synthesis and direct measurement of the in vivo process. **Proc Natl Acad Sci U S A**, v. 73, n. 12, p. 4565-9, Dec 1976.
- CARREIRA, P. M. D. C. **Produção de celulose bacteriana a partir de resíduos industriais** 2010. (Mestrado). Departamento de Química, Universidade de Aveiro, Portugal.
- CASTRO, J. V. **Produção biológica de hidrogênio a partir de glicerol por *Clostridium acetobutylicum* ATCC 824**. 2010. (Mestrado). Engenharia Química, UFSC, Florianópolis.
- CHAVEZ-PACHECO, J. L.; MARTINEZ-YEE, S.; CONTRERAS, M. L.; GOMEZ-MANZO, S.; MEMBRILLO-HERNANDEZ, J.; ESCAMILLA, J. E. Partial bioenergetic characterization of *Gluconacetobacter xylinum* cells released from cellulose pellicles by a novel methodology. **J Appl Microbiol**, v. 99, n. 5, p. 1130-40, 2005.
- CHAWLA, P. R.; BAJAJ, I. B.; SURVASE, S. A.; SINGHAL, R. S. Microbial Cellulose: Fermentative Production and Applications. . **Food Technology and Biotechnology and Bioengineering**, v. (47), n. 2, p. 107-124, 2009.
- CHENG, K. C.; CATCHMARK, J. M.; DEMIRCI, A. Enhanced production of bacterial cellulose by using a biofilm reactor and its material property analysis. **J Biol Eng**, v. 3, p. 12, 2009.

CORTASSA, S.; AON, M. A.; IGLESIAS, A. A.; LLOYD, D. **An Introduction to Metabolic and Cellular Engineering**. 2002.

COVERT, M. W.; SCHILLING, C. H.; FAMILI, I.; EDWARDS, J. S.; GORYANIN, II; SELKOV, E.; PALSSON, B. O. Metabolic modeling of microbial strains in silico. **Trends Biochem Sci**, v. 26, n. 3, p. 179-86, Mar 2001.

CZAJA, W.; KRYSZYNOWICZ, A.; BIELECKI, S.; BROWN, R. M., JR. Microbial cellulose--the natural power to heal wounds. **Biomaterials**, v. 27, n. 2, p. 145-51, Jan 2006.

CZAJA, W. K.; YOUNG, D. J.; KAWECKI, M.; BROWN, R. M., JR. The future prospects of microbial cellulose in biomedical applications. **Biomacromolecules**, v. 8, n. 1, p. 1-12, Jan 2007.

DE OLIVEIRA DAL'MOLIN, C. G., NIELSEN, L. K. Plant genome-scale metabolic reconstruction and modelling. **Curr Opin Biotechnol**, v. 24, n. 2, p. 271-7, Apr 2013.

DEINEMA, M. H.; ZEVENHUIZEN, L. P. Formation of cellulose fibrils by gram-negative bacteria and their role in bacterial flocculation. **Arch Mikrobiol**, v. 78, n. 1, p. 42-51, 1971.

DUARTE, N.; BECKER, S.; JAMSHIDI, N.; THIELE, I.; MO, M.; VO, T.; SRIVAS, R.; PALSSON, B. Global reconstruction of the human metabolic network based on genomic and bibliomic data. **Proceedings of the National Academy of Sciences**, v. 104, n. 6, p. 1777 - 1782, 2007.

DUROT, M.; BOURGUIGNON, P. Y.; SCHACHTER, V. Genome-scale models of bacterial metabolism: reconstruction and applications. **FEMS Microbiol Rev**, v. 33, n. 1, p. 164-90, Jan 2009.

EDWARDS, J.; IBARRA, R.; PALSSON, B. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. **Nat Biotechnol**, v. 19, n. 2, p. 125 - 130, 2001.

EDWARDS, J. S.; PALSSON, B. O. How will bioinformatics influence metabolic engineering? **Biotechnol Bioeng**, v. 58, n. 2-3, p. 162-9, Apr 20-May 5 1998.

EDWARDS, J. S.; PALSSON, B. O. Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. **J Biol Chem**, v. 274, n. 25, p. 17410-6, Jun 18 1999.

EDWARDS, J. S.; PALSSON, B. O. The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. **Proc Natl Acad Sci U S A**, v. 97, n. 10, p. 5528-33, May 9 2000.

FEIST, A.; HENRY, C.; REED, J.; KRUMMENACKER, M.; JOYCE, A.; KARP, P.; BROADBELT, L.; HATZIMANIKATIS, V.; PALSSON, B. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. **Mol Syst Biol**, v. 3, 2007.

FEIST, A.; PALSSON, B. O. Metabolic Flux Balancing: Basic concepts, Scientific and Practical Use - 13 Years Later. **Nat Biotechnol**, v. 26, n. 6, p. 659 - 667, 2008.

FEIST, A. M.; HERRGARD, M. J.; THIELE, I.; REED, J. L.; PALSSON, B. O. Reconstruction of biochemical networks in microorganisms. **Nat Rev Microbiol**, v. 7, n. 2, p. 129-43, Feb 2009.

FEIST, A. M.; SCHOLTEN, J. C. M.; PALSSON, B. O.; BROCKMAN, F. J.; IDEKER, T. Modeling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*. **Molecular Systems Biology**, v. 2, n. 1, 2006.

FERNALD, G. H.; CAPRIOTTI, E.; DANESHJOU, R.; KARCZEWSKI, K. J.; ALTMAN, R. B. Bioinformatics challenges for personalized medicine. **Bioinformatics**, v. 27, n. 13, p. 1741-8, Jul 1 2011.

FINK, H. **Artificial blood vessels: Studies on endothelial cell and blood interactions with bacterial cellulose**. 2009. (Doctoral). Vascular Engineering Centre at the Department of Surgery, Sahlgrenska Academy at the University of Gothenburg, Göteborg, Suécia.

FLAMHOLZ, A.; NOOR, E.; BAR-EVEN, A.; MILO, R. eQuilibrator--the biochemical thermodynamics calculator. **Nucleic Acids Res**, v. 40, n. Database issue, p. D770-5, Jan 2012.

FONG, S.; PALSSON, B. O. Metabolic gene deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. **Nature Genetics**, v. 36, n. 10, p. 1056 - 1058, 2004.

FORSTER, J.; FAMILI, I.; FU, P.; PALSSON, B. O.; NIELSEN, J. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. **Genome Res**, v. 13, p. 244 - 253, 2003.

FREITAS, F.; ALVES, V. D.; REIS, M. A. M. Advances in bacterial exopolysaccharides: from production to biotechnological applications. **Trends in Biotechnology**, v. 29, n. 8, p. 388-398, 2011.

FU, L.; ZHANG, J.; YANG, G. Present status and applications of bacterial cellulose-based materials for skin tissue repair. **Carbohydrate Polymers**, v. 92, n. 2, p. 1432-1442, 2013.

GENEONTOLOGY. Gene Ontology Database. 2013. Disponível em: < <http://www.geneontology.org/> >.

GODBEY, W. T.; ATALA, A. In vitro systems for tissue engineering. **Ann N Y Acad Sci**, v. 961, p. 10-26, Jun 2002.

GOMBERT, A. K. **Análise de redes metabólicas em *Saccharomyces cerevisiae***. 2001. (Doutorado). Escola Politécnica, Universidade de São Paulo, São Paulo.

HANKE, T.; NOH, K.; NOACK, S.; POLEN, T.; BRINGER, S.; SAHM, H.; WIECHERT, W.; BOTT, M. Combined fluxomics and transcriptomics analysis of glucose catabolism via a partially cyclic pentose phosphate pathway in *Gluconobacter oxydans* 621H. **Appl Environ Microbiol**, v. 79, n. 7, p. 2336-48, Apr 2013.

HENRY, C. S.; DEJONGH, M.; BEST, A. A.; FRYBARGER, P. M.; LINSAY, B.; STEVENS, R. L. High-throughput generation, optimization and analysis of genome-scale metabolic models. **Nat Biotechnol**, v. 28, n. 9, p. 977-82, Sep 2010.

HERWIG, C.; VON STOCKAR, U. A small metabolic flux model to identify transient metabolic regulations in *Saccharomyces cerevisiae*. **Bioprocess and Biosystems Engineering**, v. 24, n. 6, p. 395-403, 2002.

HORNUNG, M.; LUDWIG, M.; GERRARD, A. M.; SCHMAUDER, H. P. Optimizing the production of bacterial cellulose in surface culture: Evaluation of substrate mass transfer influences on the bioreaction (Part 1). **Engineering in Life Sciences**, v. 6, n. 6, p. 537-545, Dec 2006.

IGUCHI, M.; YAMANAKA, S.; BUDHIONO, A. Bacterial cellulose—a masterpiece of nature's arts. **Journal of Materials Science**, v. 35, n. 2, p. 261-270, 2000.

IMG. Integrated Microbial Genomes. 2013. Disponível em: < <http://img.jgi.doe.gov/cgi-bin/w/main.cgi> >.

INTERPRO. Protein sequence analysis & classification. 2013. Disponível em: < <http://www.ebi.ac.uk/interpro/> >.

IYER, P. R.; GEIB, S. M.; CATCHMARK, J.; KAO, T. H.; TIEN, M. Genome sequence of a cellulose-producing bacterium, *Gluconacetobacter hansenii* ATCC 23769. **J Bacteriol**, v. 192, n. 16, p. 4256-7, Aug 2010.

JONAS, R.; FARAH, L. F. Production and application of microbial cellulose. **Polymer Degradation and Stability**, v. 59, n. 1-3, p. 101-106, 1998.

KANEHISA, M.; GOTO, S.; FURUMICHI, M.; TANABE, M.; HIRAKAWA, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. **Nucleic Acids Res**, v. 38, n. Database issue, p. D355-60, Jan 2010.

KANEHISA, M.; GOTO, S.; HATTORI, M.; AOKI-KINOSHITA, K.; ITOH, M.; KAWASHIMA, S.; KATAYAMA, T.; ARAKI, M.; HIRAKAWA, M. From genomics to chemical genomics: new developments in KEGG. **Nucl Acids Res**, v. 34, n. suppl_1, p. D354 - 357, 2006.

KARP, P. D.; KRUMMENACKER, M.; PALEY, S.; WAGG, J. Integrated pathway-genome databases and their role in drug discovery. **Trends Biotechnol**, v. 17, n. 7, p. 275-81, Jul 1999.

KARP, P. D.; PALEY, S.; ROMERO, P. The Pathway Tools software. **Bioinformatics**, v. 18 Suppl 1, p. S225-32, 2002.

KARP, P. D.; PALEY, S. M.; KRUMMENACKER, M.; LATENDRESSE, M.; DALE, J. M.; LEE, T. J.; KAIPA, P.; GILHAM, F.; SPAULDING, A.; POPESCU, L.; ALTMAN, T.; PAULSEN, I.; KESELER, I. M.; CASPI, R. Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. **Brief Bioinform**, v. 11, n. 1, p. 40-79, Jan 2010.

KEGG. Kyoto Encyclopedia of Genes and Genomes. 2013. Disponível em: < <http://www.genome.jp/kegg/>>.

KENNEDY, J. F.; KEEP, P. A.; CATTY, D. The use of cellulose carbonate-based immunoadsorbents in the isolation of minor allotypic components of rabbit immunoglobulin populations. **J Immunol Methods**, v. 50, n. 1, p. 57-75, 1982.

KESELER, I.; COLLADO-VIDES, J.; GAMA-CASTRO, S.; INGRAHAM, J.; PALEY, S.; PAULSEN, I.; PERALTA-GIL, M.; KARP, P. EcoCyc: a comprehensive database resource for *Escherichia coli*. **Nucleic Acids Res**, v. 33, p. D334 - 7, 2005.

KESHK, S. M. A. S.; SAMESHIMA, K. Evaluation of different carbon sources for bacterial cellulose production. **African Journal of Biotechnology**, v. 4, n. 6, p. 478-482, 2005.

KIM, B. H.; GADD, G. M. **Bacterial Physiology and Metabolism**. Cambridge University Press: 2008.

KLANCHUL, A.; KHANNAPHO, C.; PHODEE, A.; CHEEVADHANARAK, S.; MEECHAI, A. iAK692: a genome-scale metabolic model of *Spirulina platensis* C1. **BMC Syst Biol**, v. 6, p. 71, 2012.

KLEMM, D.; HEUBLEIN, B.; FINK, H. P.; BOHN, A. Cellulose: fascinating biopolymer and sustainable raw material. **Angew Chem Int Ed Engl**, v. 44, n. 22, p. 3358-93, May 30 2005.

KLEMM, D.; SCHUMANN, D.; KRAMER, F.; HEßLER, N.; HORNUNG, M.; SCHMAUDER, H.P.; MARSCH, S. Nanocelluloses as Innovative Polymers in Research and Application. In: KLEMM, D. (Ed.). **Polysaccharides II**: Springer Berlin Heidelberg, v.205, 2006. cap. 97, p.49-96. (Advances in Polymer Science).

KLEMM, D.; SCHUMANN, D.; UDHARDT, U.; MARSCH, S. Bacterial synthesized cellulose — artificial blood vessels for microsurgery. **Progress in Polymer Science**, v. 26, n. 9, p. 1561-1603, 2001.

KOFFAS, M.; ROBERGE, C.; LEE, K.; STEPHANOPOULOS, G. Metabolic engineering. **Annu Rev Biomed Eng**, v. 1, p. 535-57, 1999.

LANGE, B. M.; GHASSEMIAN, M. Comprehensive post-genomic data analysis approaches integrating biochemical pathway maps. **Phytochemistry**, v. 66, n. 4, p. 413-451, 2005.

LEE, J. M.; GIANCHANDANI, E. P.; PAPIN, J. A. Flux balance analysis in the era of metabolomics. **Brief Bioinform**, v. 7, n. 2, p. 140-50, Jun 2006.

LEE, J. W.; NA, D.; PARK, J. M.; LEE, J.; CHOI, S.; LEE, S. Y. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. **Nat Chem Biol**, v. 8, n. 6, p. 536-46, Jun 2012.

LEE, S.; LEE, D.; KIM, T. Systems biotechnology for strain improvement. **Trends Biotechnol**, v. 23, n. 7, p. 349 - 358, 2005.

LISDIYANTI, P.; NAVARRO, R. R.; UCHIMURA, T.; KOMAGATA, K. Reclassification of *Gluconacetobacter hansenii* strains and proposals of *Gluconacetobacter saccharivorans* sp. nov. and *Gluconacetobacter nataicola* sp. nov. **Int J Syst Evol Microbiol**, v. 56, n. Pt 9, p. 2101-11, Sep 2006.

LIU, L.; AGREN, R.; BORDEL, S.; NIELSEN, J. Use of genome-scale metabolic models for understanding microbial physiology. **FEBS Lett**, v. 584, n. 12, p. 2556-64, Jun 18 2010.

LLANERAS, F. **Interval and Possibilistic Methods for Constraint-Based Metabolic Models**. 2010. (PhD). Instituto Universitario de Automática e Informática Industrial, Universidad Politécnica de Valencia

MAHADEVAN, R.; PALSSON, B. O.; LOVLEY, D. R. In situ to in silico and back: elucidating the physiology and ecology of *Geobacter* spp. using genome-scale modelling. **Nat Rev Microbiol**, v. 9, n. 1, p. 39-50, Jan 2011.

MARKOWITZ, V. M.; CHEN, I. M.; PALANIAPPAN, K.; CHU, K.; SZETO, E.; GRECHKIN, Y.; RATNER, A.; JACOB, B.; HUANG, J.; WILLIAMS, P.; HUNTEMANN, M.; ANDERSON, I.; MAVROMATIS, K.; IVANOVA, N. N.; KYRPIDES, N. C. IMG: the Integrated Microbial Genomes database and comparative analysis system. **Nucleic Acids Res**, v. 40, n. Database issue, p. D115-22, Jan 2012.

MASAOKA, S.; OHE, T.; SAKOTA, N. Production of cellulose from glucose by *Acetobacter xylinum*. **Journal of Fermentation and Bioengineering**, v. 75, n. 1, p. 18-22, 1993.

MCCLOSKEY, D.; PALSSON, B. O.; FEIST, A. M. Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. **Mol Syst Biol**, v. 9, p. 661, 2013.

METZKER, M. L. Emerging technologies in DNA sequencing. **Genome Res**, v. 15, n. 12, p. 1767-76, Dec 2005.

METZKER, M. L. Sequencing technologies - the next generation. **Nat Rev Genet**, v. 11, n. 1, p. 31-46, Jan 2010.

MIKKELSEN, D.; FLANAGAN, B. M.; DYKES, G. A.; GIDLEY, M. J. Influence of different carbon sources on bacterial cellulose production by *Gluconacetobacter xylinus* strain ATCC 53524. **J Appl Microbiol**, v. 107, n. 2, p. 576-83, Aug 2009.

MONTAGUD, A.; NAVARRO, E.; FERNANDEZ DE CORDOBA, P.; URCHUEGUIA, J.; PATIL, K. Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. **BMC Systems Biology**, v. 4, n. 1, p. 156, 2010.

MORGAN, J. L.; STRUMILLO, J.; ZIMMER, J. Crystallographic snapshot of cellulose synthesis and membrane translocation. **Nature**, v. 493, n. 7431, p. 181-6, Jan 10 2013.

NAEIMPOOR, F.; MAVITUNA, F. Metabolic Flux Analysis in *Streptomyces coelicolor* under Various Nutrient Limitations. **Metabolic Engineering**, v. 2, n. 2, p. 140-148, 2000.

NCBI. National Center for Biotechnology Information. 2013. Disponível em: < <http://www.ncbi.nlm.nih.gov/>>.

NG, P.; KIRKNESS, E. Whole Genome Sequencing. In: BARNES, M. R. e BREEN, G. (Ed.). **Genetic Variation**: Humana Press, v.628, 2010. cap. 12, p.215-226. (Methods in Molecular Biology).

NISSEN, T. L.; SCHULZE, U.; NIELSEN, J.; VILLADSEN, J. Flux distributions in anaerobic, glucose-limited continuous cultures of *Saccharomyces cerevisiae*. **Microbiology**, v. 143 (Pt 1), p. 203-18, Jan 1997.

NOLAN, R. P.; FENLEY, A. P.; LEE, K. Identification of distributed metabolic objectives in the hypermetabolic liver by flux and energy balance analysis. **Metab Eng**, v. 8, n. 1, p. 30-45, Jan 2006.

NOOR, E.; EDEN, E.; MILO, R.; ALON, U. Central carbon metabolism as a minimal biochemical walk between precursors for biomass and energy. **Mol Cell**, v. 39, n. 5, p. 809-20, Sep 10 2010.

OBERHARDT, M. A.; PALSSON, B. O.; PAPIN, J. A. Applications of genome-scale metabolic reconstructions. **Mol Syst Biol**, v. 5, p. 320, 2009.

OIKAWA, T.; OHTORI, T.; AMEYAMA, M. Production of Cellulose from d-Mannitol by *Acetobacter xylinum* KU-1. **Bio/Technology**, v. 59, p. 331-332, 1995.

OLIJVE, W.; KOK, J. J. An analysis of the growth of *Gluconobacter oxydans* in chemostat cultures. **Archives of Microbiology**, v. 121, n. 3, p. 291-297, 1979/06/01 1979.

OLIVEIRA, I. L. **Análise Computacional de Redes Metabólicas com regulação**. 2008. (Doutorado). Engenharia Química, UFSC, Florianópolis.

OLIVEIRA, V. A. **Produção e Degradação in vitro de Vasos Sanguíneos Artificiais à Base de Celulose Bacteriana**. 2011. (Mestrado). Departamento de Engenharia Química, Universidade Federal de Santa Catarina Florianópolis.

ORTH, J. D.; CONRAD, T. M.; NA, J.; LERMAN, J. A.; NAM, H.; FEIST, A. M.; PALSSON, B. O. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism--2011. **Mol Syst Biol**, v. 7, p. 535, 2011.

ORTH, J. D.; THIELE, I.; PALSSON, B. O. What is flux balance analysis? **Nat Biotechnol**, v. 28, n. 3, p. 245-8, Mar 2010.

PALEY, S. M.; LATENDRESSE, M.; KARP, P. D. Regulatory network operations in the Pathway Tools software. **BMC Bioinformatics**, v. 13, p. 243, 2012.

PALSSON, B. O.; BHATIA, S. **Tissue engineering**. Pearson Prentice Hall, 2004. 407p

PALSSON, B. O. **Systems Biology: Properties of Reconstructed Networks**. New York: 2006.

PANDEY, A.; MANN, M. Proteomics to study genes and genomes. **Nature**, v. 405, n. 6788, p. 837-46, Jun 15 2000.

PAPOUTSAKIS, E. T.; MEYER, C. L. Equations and calculations of product yields and preferred pathways for butanediol and mixed-acid fermentations. **Biotechnol Bioeng**, v. 27, n. 1, p. 50-66, Jan 1985.

PATIL, K. R.; AKESSON, M.; NIELSEN, J. Use of genome-scale microbial models for metabolic engineering. **Curr Opin Biotechnol**, v. 15, n. 1, p. 64-9, Feb 2004.

PFAM. Protein Families Database. 2013. Disponível em: <<http://pfam.sanger.ac.uk/>>.

PRICE, N.; REED, J.; PALSSON, B. O. Genome-scale models of microbial cells: evaluating the consequences of constraints. **Nat Rev Micro**, v. 2, n. 11, p. 886 - 897, 2004.

RAMAKRISHNA, R.; EDWARDS, J.; MCCULLOCH, A.; PALSSON, B.O. Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. **American journal of physiology**, v. 280, n. 3, p. R695 - 704, 2001.

RAMAN, K.; CHANDRA, N. Flux balance analysis of biological systems: applications and challenges. **Brief Bioinform**, v. 10, n. 4, p. 435-49, Jul 2009.

RAMANA, K. V.; TOMAR, A.; SINGH, L.. Effect of various carbon and nitrogen sources on cellulose synthesis by *Acetobacter xylinum*. **World Journal of Microbiology and Biotechnology**, v. 16, n. 3, p. 245-248, 2000/04/01 2000.

RAVEN, P. H.; EVERT, R. F.; EICHHORN, S. E. **Biologia vegetal**. 7. Rio de Janeiro: 2010. 830.

RECOUVREUX, D. O.; CARMINATTI, C. A.; PITLOVANCIV, A. K.; RAMBO, C. R.; PORTO, L. M.; ANTONIO, R. V. Cellulose biosynthesis by the beta-proteobacterium, *Chromobacterium violaceum*. **Curr Microbiol**, v. 57, n. 5, p. 469-76, Nov 2008.

RECOUVREUX, D. O. S.; RAMBO, C. R.; BERTI, F. V.; CARMINATTI, C. A.; ANTÔNIO, R. V.; PORTO, L. M. Novel three-dimensional cocoon-like hydrogels for soft tissue regeneration. **Materials Science and Engineering: C**, v. 31, n. 2, p. 151-157, 2011.

REED, J.; FAMILI, I.; THIELE, I.; PALSSON, B.O. Towards multidimensional genome annotation. **Nat Rev Genet**, v. 7, n. 2, p. 130 - 141, 2006.

REED, J.; VO, T.; SCHILLING, C.; PALSSON, B.O. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). **Genome Biology**, v. 4, n. 9, p. R54.51 - R54.52, 2003.

ROSS, P.; MAYER, R.; BENZIMAN, M. Cellulose biosynthesis and function in bacteria. **Microbiol Rev**, v. 55, n. 1, p. 35-58, Mar 1991.

ROSSA, C. A.; WHITE, J.; KUIPER, A.; POSTMA, P. W.; BIBB, M.; TEIXEIRA DE MATTOS, M. J. Carbon Flux Distribution in Antibiotic-Producing Chemostat Cultures of *Streptomyces lividans*. **Metabolic Engineering**, v. 4, n. 2, p. 138-150, 2002.

RUKA, D. R.; SIMON, G. P.; DEAN, K. M. Altering the growth conditions of *Gluconacetobacter xylinus* to maximize the yield of bacterial cellulose. **Carbohydrate Polymers**, v. 89, n. 2, p. 613-622, 2012.

SANCHEZ, A. M.; BENNETT, G. N.; SAN, K. Y. Batch culture characterization and metabolic flux analysis of succinate-producing *Escherichia coli* strains. **Metab Eng**, v. 8, n. 3, p. 209-26, May 2006.

SAXENA, I. M.; KUDLICKA, K.; OKUDA, K.; BROWN, R. M., JR. Characterization of genes in the cellulose-synthesizing operon (acs operon) of *Acetobacter xylinum*: implications for cellulose crystallization. **J Bacteriol**, v. 176, n. 18, p. 5735-52, Sep 1994.

SAXENA, I. M.; LIN, F. C.; BROWN, R. M. Cloning and sequencing of the cellulose synthase catalytic subunit gene of *Acetobacter xylinum*. **Plant Mol Biol**, v. 15, n. 5, p. 673-83, Nov 1990.

SAXENA, I. M.; LIN, F. C.; BROWN, R. M. Identification of a new gene in an operon for cellulose biosynthesis in *Acetobacter xylinum*. **Plant Mol Biol**, v. 16, n. 6, p. 947-54, Jun 1991.

SCHELLENBERGER, J.; QUE, R.; FLEMING, R. M.; THIELE, I.; ORTH, J. D.; FEIST, A. M.; ZIELINSKI, D. C.; BORDBAR, A. A.;

- LEWIS, N. E.; RAHMANIAN, S.; KANG, J.; HYDUKE, D. R.; PALSSON, B. O. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. **Nat Protoc**, v. 6, n. 9, p. 1290-307, Sep 2011.
- SCHILLING, C. H.; EDWARDS, J. S.; LETSCHER, D.; PALSSON, B. O. Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems. **Biotechnol Bioeng**, v. 71, n. 4, p. 286 - 306, 2000.
- SCHILLING, C. H.; COVERT, M. W.; FAMILI, I.; CHURCH, G. M.; EDWARDS, J. S.; PALSSON, B. O. Genome-scale metabolic model of *Helicobacter pylori* 26695. **J Bacteriol**, v. 184, n. 16, p. 4582-93, Aug 2002.
- SCHRAMM, M.; HESTRIN, S. Factors affecting production of cellulose at the air/liquid interface of a culture of *Acetobacter xylinum*. **J Gen Microbiol**, v. 11, n. 1, p. 123-9, Aug 1954.
- SHARMA, N. S.; IERAPETRITOU, M. G.; YARMUSH, M. L. Novel quantitative tools for engineering analysis of hepatocyte cultures in bioartificial liver systems. **Biotechnol Bioeng**, v. 92, n. 3, p. 321-35, Nov 5 2005.
- SILVA, T. R. S. **Desenvolvimento de hidrogéis de celulose bacteriana para cultura de células e permeação de biomoléculas**. 2012. (Mestrado). Programa de Pós-Graduação em Engenharia Química, Universidade Federal de Santa Catarina, Florianópolis.
- SOON, W. W.; HARIHARAN, M.; SNYDER, M. P. High-throughput sequencing for biology and medicine. **Mol Syst Biol**, v. 9, p. 640, 2013.
- STEPHANOPOULOS, G. Metabolic fluxes and metabolic engineering. **Metab Eng**, v. 1, n. 1, p. 1-11, Jan 1999.
- STEPHANOPOULOS, G.; ARISTIDOU, A. A.; NIENSEN, J. **Metabolic Engineering: Principles and Methodologies**. San Diego, CA: 1998.

STEPHANOPOULOS, G.; STAFFORD, D. E. Metabolic engineering: a new frontier of chemical reaction engineering. **Chemical Engineering Science**, v. 57, n. 14, p. 2595-2602, 2002.

STRAP, J. L.; LATOS, A.; SHIM, I.; BONETTA, D. T. Characterization of pellicle inhibition in *Gluconacetobacter xylinus* 53582 by a small molecule, pellicin, identified by a chemical genetics screen. **PLoS One**, v. 6, n. 12, p. e28015, 2011.

SUWANPOSRI, A.; YUKPHAN, P.; YAMADA, Y.; OCHAIKUL, D. Identification and biocellulose production of *Gluconacetobacter* strains isolated from tropical fruits in Thailand. **Maejo International Journal of Science and Technology**, v. 7, p. 70-82, 2013.

SVENSSON, A.; NICKLASSON, E.; HARRAH, T.; PANILAITIS, B.; KAPLAN, D. L.; BRITTBURG, M.; GATENHOLM, P. Bacterial cellulose as a potential scaffold for tissue engineering of cartilage. **Biomaterials**, v. 26, n. 4, p. 419-31, Feb 2005.

SYSTEM BIOLOGY RESEARCH GROUP. Disponível em: < <http://gerg.ucsd.edu/Home>>

SWAINSTON, N. **Systems biology informatics for the development and use of genome scale metabolic models**. 2011. (PhD thesis). School of Computer Science, University of Manchester

TAKAÇ, S.; ÇALIK, G.; MAVITUNA, F.; DERVAKOS, G. Metabolic flux distribution for the optimized production of l-glutamate. **Enzyme and Microbial Technology**, v. 23, n. 5, p. 286-300, 1998.

TAKIGUCHI, N.; SHIMIZU, H.; SHIOYA, S. An on-line physiological state recognition system for the lysine fermentation process based on a metabolic reaction model. **Biotechnol Bioeng**, v. 55, n. 1, p. 170-81, Jul 5 1997.

TERZER, M.; MAYNARD, N. D.; COVERT, M. W.; STELLING, J. Genome-scale metabolic networks. **Wiley Interdiscip Rev Syst Biol Med**, v. 1, n. 3, p. 285-97, Nov-Dec 2009.

THIELE, I.; HEINKEN, A.; FLEMING, R. M. T. A systems biology approach to studying the role of microbes in human health. **Current Opinion in Biotechnology**, v. 24, n. 1, p. 4-12, 2013.

THIELE, I.; PALSSON, B. O. A protocol for generating a high-quality genome-scale metabolic reconstruction. **Nat Protoc**, v. 5, n. 1, p. 93-121, Jan 2010.

THIELE, I.; VO, T.; PRICE, N.; PALSSON, B.O. An Expanded Metabolic Reconstruction of *Helicobacter pylori* (iIT341 GSM/GPR): An in silico genome-scale characterization of single and double deletion mutants. **J Bacteriol**, v. 187, n. 16, p. 5818 - 5830, 2005.

THORLEIFSSON, S. G.; THIELE, I. rBioNet: A COBRA toolbox extension for reconstructing high-quality biochemical networks. **Bioinformatics**, v. 27, n. 14, p. 2009-10, Jul 15 2011.

TRAWICK, J.; SCHILLING, C. Use of constraint-based modeling for the prediction and validation of antimicrobial targets. **Biochem Pharmacol**, v. 71, p. 1026 - 1035, 2006.

ULLAH, M.; SCHMIDT, H.; CHO, K. H.; WOLKENHAUER, O. Deterministic modelling and stochastic simulation of biochemical pathways using MATLAB. **Syst Biol (Stevenage)**, v. 153, n. 2, p. 53-60, Mar 2006.

VALLINO, J. J.; STEPHANOPOULOS, G. Flux determination in cellular bioreaction networks: applications to lysine fermentations. In: SIKDAR, S. K.; BIER, M., *et al* (Ed.). **Frontiers in Bioprocessing**, cap. 18, p.205-219. 1990.

VARMA, A.; PALSSON, B. O. Metabolic flux balancing: basic concepts, scientific, and practical use. **Bio/Technology**, v. 12, n. 10, 1994.

VARMA, A.; PALSSON, B. O. Metabolic capabilities of *Escherichia coli*: I. synthesis of biosynthetic precursors and cofactors. **J Theor Biol**, v. 165, n. 4, p. 477-502, Dec 21 1993.

WATANABE, K.; TABUCHI, M.; MORINAGA, Y.; YOSHINAGA, F. Structural Features and Properties of Bacterial Cellulose Produced in Agitated Culture. **Cellulose**, v. 5, n. 3, p. 187-200, 1998.

WIECHERT, W. 13C metabolic flux analysis. **Metab Eng**, v. 3, n. 3, p. 195-206, Jul 2001.

WIECHERT, W.; NÖH, K. Isotopically non-stationary metabolic flux analysis: complex yet highly informative. **Current Opinion in Biotechnology**, v. 24, n. 6, p. 979-986, 2013.

WILLIAMS, D. F. Definitions in biomaterials. In: WILLIAMS, D. F. (Ed.). **Proceedings of a consensus conference of the European Society for Biomaterials**. Chester, England, 1987.

WONG, H. C.; FEAR, A. L.; CALHOON, R. D.; EICHINGER, G. H.; MAYER, R.; AMIKAM, D.; BENZIMAN, M.; GELFAND, D. H.; MEADE, J. H.; EMERICK, A. W. Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. **Proc Natl Acad Sci U S A**, v. 87, n. 20, p. 8130-4, Oct 1990.

YAMADA, Y.; HOSHINO, K.; ISHIKAWA, T. The phylogeny of acetic acid bacteria based on the partial sequences of 16S ribosomal RNA: the elevation of the subgenus *Gluconoacetobacter* to the generic level. **Biosci Biotechnol Biochem**, v. 61, n. 8, p. 1244-51, Aug 1997.

YAMANAKA, S.; WATANABE, K.; KITAMURA, N.; IGUCHI, M.; MITSUHASHI, S.; NISHI, Y.; URYU, M. The structure and mechanical properties of sheets prepared from bacterial cellulose. **Journal of Materials Science**, v. 24, n. 9, p. 3141-3145, 1989/09/01 1989.

ZHAO, J.; GRANT, S. F. Advances in whole genome sequencing technology. **Curr Pharm Biotechnol**, v. 12, n. 2, p. 293-305, Feb 1 2011.

ZHONG, C.; ZHANG, G. C.; LIU, M.; ZHENG, X. T.; HAN, P. P.; JIA, S. R. Metabolic flux analysis of *Gluconacetobacter xylinus* for bacterial cellulose production. **Appl Microbiol Biotechnol**, v. 97, n. 14, p. 6189-99, Jul 2013.

APÊNDICE A – MFA no GEnSys

```
%Genome-scale reconstruction and in silico  
analysis of the Gluconacetobacter hansenii ATCC  
23769 metabolic network.
```

```
%Metabolic Flux analysis (MFA) to calculate  
internal fluxes on the network
```

```
clc;
```

```
%usado apenas para visualizar as equações do  
balanço de cada metabólito
```

```
m = DetSim('GhanseniiMFA.m');  
printODE(m);
```

```
%calcula os fluxos
```

```
m = mfa('GhanseniiMFA.m');  
m = solvemfa(m);
```

```
figure, bar(m.vc, 'g')
```

```
title('Análise de Fluxo Metabólico - AFM')
```

```
ylabel('Fluxos (mmol/h/g [dry wt])')
```

```
xlabel('Reações')
```

```
% legend('Unmeasured rates')
```

```
for i=1:length(m.iCalcFlux)
```

```
    str = num2str(m.vc(i));
```

```
    str = [m.lRxn{m.iCalcFlux(i)} ' = ' str];
```

```
    fprintf('%s\n', str);
```

```
end
```

APÊNDICE B – FBA no GEnSys

```
%Genome-scale reconstruction and in silico  
analysis of the Gluconacetobacter hansenii ATCC  
23769 metabolic network.
```

```
%Flux balance analysis (FBA) to predict growth  
rates on three different carbon sources:  
glucose, glycerol and mannitol.
```

```
clc  
m = fba('Ghansenii.m')  
  
s = solveFBA(m)  
  
printFluxVector(m.intModel.lRxn, s.x);  
  
fprintf('\nValor fun objetivo:%s  
\n', num2str(s.f))  
  
figure, bar(s.x, 'g')  
title('Análise de Balanço de Fluxo - ABF')  
ylabel('Fluxos (mmol/h/g [dry wt])')  
xlabel('Reações')  
legend('Glicose como Substrato')  
%legend('Manitol como Substrato')  
%legend('Glicerol como Substrato')
```

APÊNDICE C - REAÇÕES DO MODELO

Abb	Reação	Descrição	Equação	EC	GENE	Subsistema
v1_gli		transporte glicose/glicose permease	GLC_x -> GLC			transporte
v2_o2		transporte oxigênio	O2_x -> O2			transporte
v3_co2		transporte co2	CO2_x <-> CO2			transporte
v4_bio		transporte biomassa	BIOMASS -> BIOMASS_x			transporte
v5_cell		transporte cellulose	CELL -> CELL_x			transporte
v6_man		transporte manitol	MANN_x -> MANN			transporte
v7_nh4		fonte de nitrogênio	NH4_x <-> NH4			transporte
v8_glyc		transporte glicerol	GLYC_x -> GLYC			transporte
v9_h2o		Transporte água	H2O_x <-> H2O			transporte
v10_glc_n		transporte gluconato	GLCN -> GLCN_x			transporte
v11_h		transporte próton	H_x <-> H			transporte
v12_pi		fonte de fosfato	PI_x <-> PI			transporte
v13	R00299	glicoquinase	GLC + ATP -> ADP + G6P	2.7.1.2	GXY_05501, GXY_13683(pu tative)	Glicólise/gliconeogênese
v14	R00741	Fosfoglicose isomerase	G6P <-> F6P	5.3.1.9	GXY_02166	Glicólise/gliconeogênese
v15	R00762	frutose difosfatase	FDP + H2O -> F6P + PI	3.1.3.11	GXY_08300 (gfpX)	Glicólise/gliconeogênese
v16	R01068	frutose-bisfosfate aldolase	FDP <-> DHAP + G3P	4.1.2.13	GXY_08305, GXY_09124	Glicólise/gliconeogênese
v17	R01015	Triosefosfato isomerase	DHAP <-> G3P	5.3.1.1	GXY_10284	Glicólise/gliconeogênese
v18	R01061	gliceraldeído 3-fosfato desidrogenase	G3P + NAD + PI -> 13-DPG + NADH_x + H	1.2.1.12	GXY_04003	Glicólise/gliconeogênese

v19	R01512	fosfoglicerato quinase	13-DPG + ADP + PI <-> ATP + 3-PG	2.7.2.3	GXY_03998 (pgk)	Glicólise/gli- coneogênese
v20	R01518	fosfoglicerato mutase	3-PG <-> 2-PG	5.4.2.1	GXY_02671,GX Y_12768	Glicólise/gli- coneogênese
v21	R00658	Fosfopiruvato hidratase/ enolase	2-PG <-> PEP + H2O	4.2.1.11	GXY_10254 (eno)	Glicólise/gli- coneogênese
v22	R00200	Piruvato quinase	PEP + ADP + H -> ATP + PYR	2.7.1.40	GXY_00359	Glicólise/gli- coneogênese
v23	R00209	Piruvato deidrogenase complexo	PYR + NAD + COA -> ACCOA_x + NADH_x + CO2	(2.3.1.12 and 1.8.1.4 and 1.2.4.1)	(GXY_16242 OR GXY_10329 OR GXY_07680) - (GXY_03931 OR(GXY_03931 AND GXY_03943) - (GXY_15912 OR (ilvH AND GXY_00049) OR (GXY_10324 AND GXY_10319) OR (GXY_13548 OR GXY_15937)	Glicólise/gli- coneogênese
v24	R00206	Piruvato fosfato diquinase	PYR + ATP + PI -> AMP + PEP + PPI	2.7.9.1	GXY_08205	Glicólise/gli- coneogênese
v25	R00217	oxaloacetato decarboxilase	OAA + ADP + PI -> PYR + ATP + CO2	4.1.1.3		Glicólise/gli- coneogênese
v26	R00345	Fosfoenolpiru- vato carboxilase	PEP + H2O + CO2 -> OAA + PI + H	4.1.1.31	GXY_12143	Glicólise/gli- coneogênese
v27	R08639	Fosfoglico- mutase	G6P -> G1P	5.4.2.2	GXY_09809	Glicólise/gli- coneogênese
v28	R00289	uridina glicose pirofosforilase	G1P + UTP + H -> UDPG + PPI	2.7.7.9	GXY_10109	Metabolis- mo amido e sacarose
v29	R02889	celulose - UDP	UDPG -> UDP + CELL	2.4.1.12	GXY_04277, GXY_08869	Metabolis- mo amido e sacarose

v30	R06620 / R01519	Glicose deidrogenase quinoproteína/ gliconolactona- nase	GLC + Q8 + H2O -> GLCN + Q8H2	1.1.5.2 / 3.1.1.17	GXY_11284; GXY_13868	Ciclo das pentoses
v31	R01737	gliconoquinase	GLCN + ATP -> ADP + 6-PGC + H	2.7.1.12	GXY_02201,GX Y_12403	ciclo das pentoses
v32	R00835	glicose-6- fosfato deidrogenase	G6P + NADP_x <-> 6-PGL + NADPH	1.1.1.49	GXY_01616,GX Y_02176,GXY_ 11509	ciclo das pentoses
v33	R02035	6-fosfoglicono- lactonase	6-PGL + H2O - > 6-PGC	3.1.1.31	GXY_02191,GX Y_10154	ciclo das pentoses
v34	R01528	6- fosfogliconato deidrogenase	6-PGC + NADP_x -> RU5P + CO2 + NADPH	1.1.1.44	GXY_04594	ciclo das pentoses
v35	R02036	fosfogliconato deidratase	6-PGC -> 2- DDG6P + H2O	4.2.1.12	GXY_03863	ciclo das pentoses
v36	R05605	2-deidro-3- deoxifosfo gluconato aldolase	2-DDG6P -> G3P + PYR	4.1.2.14	GXY_03858	ciclo das pentoses
v37	R01529	epimerase	RU5P <-> XU5P	5.1.3.1		ciclo das pentoses
v38	R01056	isomerase	RU5P <-> R5P	5.3.1.6	GXY_02196	ciclo das pentoses
v39	R01827	transaldolase	S7P + G3P <-> E4P + F6P	2.2.1.2	GXY_02166	ciclo das pentoses
v40	R01641	transketolase	R5P + XU5P <- > S7P + G3P	2.2.1.1	GXY_02161,GX Y_04008	ciclo das pentoses
v41	R01067	transketolase	E4P + XU5P <-> F6P + G3P	2.2.1.1	GXY_02161,GX Y_04008	ciclo das pentoses
v42	R00868	manitol 2- deidrogenase	MANN + NAD <-> FRU + NADH_x + H	1.1.1.67	GXY_02161	manitol
v43	R00760	frutoquinase	FRU + ATP -> ADP + F6P + H	2.7.1.4	GXY_10569	
v44	R00847	glicerol quinase	GLYC + ATP -> ADP + GLYC3P	2.7.1.30	GXY_08295	Glycerophos pholipid metabolism
v45	R00842	glicerol-3- fosfato deidrogenase	GLYC3P + NAD <-> DHAP + NADH_x + H	1.1.1.94	GXY_04966	Glycerophos pholipid metabolism
v46	R00351	citrato sintase	ACCOA_x + H2O + OAA -> CIT + COA + H	2.3.3.1	GXY_10922	ciclo de krebs
v47	R01325	aconitate hidratase 1	CIT -> ACON-C + H2O	4.2.1.3	GXY_01403	ciclo de krebs

v48	R01900	aconitase hidratase 2	ACON-C + H2O <-> ICIT	4.2.1.3	GXY_01403	ciclo de krebs
v49	R00267	Isocitrato deidrogenase	ICIT + NADP_x <-> AKG + CO2 + NADPH + H	1.1.1.42	GXY_08180	ciclo de krebs
v50	R01197	2-oxoglutarato sintase	AKG + NAD + COA -> SUCCOA + NADH_x + CO2	1.2.7.3		ciclo de krebs
v51	R00405	sucinyl-CoA sintetase	SUCCOA + ADP + PI <-> ATP + SUCC + COA	6.2.1.5	GXY_05758, GXY_05763	ciclo de krebs
v52	R00408	Sucinato deidrogenase	SUCC + FAD <- > FADH2 + FUM	1.3.99.1	GXY_01598 - sdhB	ciclo de krebs
v53	R01082	fumarase	FUM + H2O <- > MAL-L	4.2.1.2	GXY_13863, GXY_02031	ciclo de krebs
v54	R00342	Malato deidrogenase	MAL-L + NAD <-> NADH_x + OAA + H	1.1.1.37		
v55	R00253	Glutamina sintetase	ATP + GLU-L + NH4 -> ADP + PI + GLN-L + H	6.3.1.2	GXY_02336	
v56	R00256	glutaminase	GLN-L + H2O -> GLU-L + NH4	3.5.1.2	GXY_12733 ; carB	
v57	R00114	glutamato sintase	GLN-L + AKG + NADPH + H -> 2GLU-L + NADP_x	1.4.1.13	GXY_04844, gxy_04839	
v58			GLN-L_x + H_x <-> GLN-L + H			transporte
v59			GLU-L_x + H_x <-> GLU-L + H			transporte
v60		adenilato quinase	AMP + ATP <-> 2ADP	2.7.4.3	GXY_12943	purina metabolism
v61		ATP manutenção	ATP + H2O -> ADP + PI + H			NGAM
v62		ATP sintase - Complexo V	ADP + PI + 4H_x -> ATP + H2O + 4H	3.6.3.14	GXY_00619, GXY_00624, GXY_00629, GXY_00634, GXY_15649, GXY_15672, GXY_15677, GXY_15682	fosforilação oxidativa
v63		Difosfatase	PPI + H2O -> 2PI + H	3.6.1.1	GXY_01896	

v64	NAD trans hidrogenase	NAD + NADPH -> NADH_x + NADP_x	1.6.1.2		
v65	Nucleosídeo trifosfatase (UTP)	UTP + H2O -> UDP + H + PI	3.6.1.5		
v66	Nucleosídeo difosfatase quinase (ATP:UDP)	ATP + UDP -> ADP + UTP	2.7.4.6	ndk	
v67	Complexo I (NADH deidrogenase)	Q8 + NADH_x + 5H -> Q8H2 + NAD + 4H_x	1.6.5.3	GXY_08325,GX Y_11983,GXY_ 11988,GXY_11 993,GXY_1258 3,GXY_15579	fosforilação oxidativa
v68	Complexo II (sucinato desidrogenase)	Q8 + SUCC -> Q8H2 + FUM	1.3.5.1		fosforilação oxidativa
v69	Ubiquinol Oxidase (citocromo bd oxidase)	2Q8H2 + 4H + O2 -> 2Q8 + 2H2O + 4H_x	1.10.3.10	GXY_05121,GX Y_05126,GXY_ 05131,GXY_05 136	fosforilação oxidativa
v70	Complexo IV (citocromo c oxidase)	O2 + 4FERROCYTOC HROME + 4H - > 4FERRICYTOCH ROME + 2H2O + 4H_x	1.9.3.1	GXY_04894,GX Y_07135	fosforilação oxidativa
v71	Complexo III (citocromo bc1)	Q8H2 + 2FERRICYTOCH ROME + 2H -> Q8 + 2FERROCYTOC HROME + 4H_x	1.10.2.2	GXY_00569,GX Y_00574,GXY_ 16474	fosforilação oxidativa
v72		NADH_x + 0.5O2 + 2.5ADP + 2.5PI + 3.5H -> 3.5H2O + NAD + 2.5ATP			fosforilação oxidativa
v73		FADH2 + 0.5O2 + 1.5ADP + 1.5PI + 2.5H -> 2.5H2O + FAD + 1.5ATP			fosforilação oxidativa

v74	41.3ATP + 0.205G6P + 0.052F6P + 0.87R5P + 0.233E4P + 0.269G3P + 1.253*3-PG + 0.509PEP + 3.113PYR + 2.407ACCOA_x + 1.415AKG + 1.726OAA + 15.779NADPH + 41.3H2O -> 41.3ADP + 41.3PI + 2.811CO2 + 4.669NADH_x + 41.3H + BIOMASS	biomassa
-----	---	----------
