RAFAEL DAVID DE OLIVEIRA

**Desenvolvimento de uma ferramenta
computacional para a análise de fluxos
metabólicos empregando carbono marcado**

São Paulo

2017

RAFAEL DAVID DE OLIVEIRA

**Desenvolvimento de uma ferramenta computacional para
a análise de fluxos metabólicos empregando carbono
marcado**

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do título de Mestre em Ciências.

São Paulo
2017

RAFAEL DAVID DE OLIVEIRA

**Desenvolvimento de uma ferramenta computacional para
a análise de fluxos metabólicos empregando carbono
marcado**

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do título de Mestre em Ciências.

Área de concentração: Engenharia Química

Orientador: Prof. Dr. Galo A. C. Le Roux

Coorientador: Prof. Dr. José Gregório Cabrera
Gomez

São Paulo

2017

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catálogo-na-publicação

Oliveira, Rafael David

Desenvolvimento de uma ferramenta computacional para a análise de fluxos metabólicos empregando carbono marcado / R. D. Oliveira -- versão corr. -- São Paulo, 2017.

87 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Química.

1. Análise de fluxos metabólicos 2. Engenharia metabólica 3. Estimção de parâmetros I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Química II. t.

À pessoa mais sábia que conheci em toda minha vida, minha mãe Maria Francisca.

Agradecimentos

Aos meus orientadores Prof. Dr. Galo A. C. Le Roux e Prof. Dr. José Gregório Cabrera Gomez pelo incentivo, confiança e pela orientação. À Caroline Satye pelo apoio em diversos aspectos desse trabalho e pela amizade. À minha família por tornar possível a realização desse mestrado. À Maíra pela motivação e pela parceria. Aos meus caros amigos de casa: Leonardo, Gustavo, Vaca e Gledson pelo ganho de potência de agir diário. Aos meus caros amigos que fiz nesses dois anos em São Paulo: Irena, Priscila, Natália, Zé Otávio, Zé Eduardo, María, Matheus, Fernanda, Fred, Layane e Denise pela amizade e pelo café de todo dia. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa concedida.

“Exatamente da mesma maneira que a luz revela a si própria e as trevas, assim também a verdade é norma de si própria e do falso.”
(Baruch Spinoza)

Resumo

OLIVEIRA, Rafael David. **Desenvolvimento de uma ferramenta computacional para a análise de fluxos metabólicos empregando carbono marcado** 2017. 87 f. Dissertação (Mestrado em Engenharia Química) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2017.

A ^{13}C -Análise de Fluxos Metabólicos (^{13}C -MFA) tornou-se uma técnica de alta precisão para estimar fluxos metabólicos e obter informações importantes sobre o metabolismo. Este método consiste em procedimentos experimentais, técnicas de medição e em cálculos para análise de dados. Neste contexto, os grupos de pesquisa de engenharia metabólica necessitam de ferramentas computacionais precisas e adequadas aos seus objetos de estudo. No presente trabalho, foi construída uma ferramenta computacional na plataforma MATLAB que executa cálculos de ^{13}C -MFA, com balanços de metabólitos e cumômeros. Além disso, um módulo para estimar os fluxos metabólicos e um módulo para quantificar as incertezas das estimativas também foram implementados. O programa foi validado com dados presentes na literatura e aplicado a estudos de caso. Na estimação de fluxos de *Pseudomonas sp.* LFM046, identificou-se que esse micro-organismo possivelmente utiliza a Via das Pentoses em conjunto com a Via Entner-Doudoroff para a biossíntese de Polihidroxialcanoato (PHA). No design ótimo de experimentos para uma rede genérica de *Pseudomonas*, identificou-se a glicose marcada no átomo cinco como um substrato que permitirá determinar o fluxo na Via das Pentoses com menor incerteza.

Palavras-chave: ^{13}C -análise de fluxos metabólicos, engenharia metabólica, modelagem, estimação de parâmetros, *Pseudomonas*, PHA

Abstract

OLIVEIRA, Rafael David. **Development of a computational tool for metabolic flux analysis with labeled carbon**. 2017. 87 p. Dissertation (Master of Science) – Polytechnic School, University of São Paulo, São Paulo, 2017.

^{13}C -Metabolic Flux Analysis (^{13}C -MFA) has become a high-precision technique to estimate metabolic fluxes and get insights into metabolism. This method consists of experimental procedures, measurement techniques and data analysis calculations. In this context, metabolic engineering research groups demand accurate and suitable computational tools to perform the calculations. A computational tool was implemented in MATLAB platform that performs ^{13}C -MFA calculation, using metabolite and cumomer balances, as well as a module to estimate the fluxes and a module to quantify their uncertainty. The program was validated with some classical cases from literature. From the flux estimates of *Pseudomonas sp.* LFM046, it was identified that the microorganism possibly uses the Pentose Phosphate Pathway along with the Entner-Doudoroff Pathway for Polyhydroxyalkanoate (PHA) biosynthesis. From the optimal experimental design for a generic *Pseudomonas* network, it was possible to conclude that glucose labeled at atom five is the best option to determine the flux in the Pentose Phosphate Pathway with smaller uncertainty.

Key-words: ^{13}C -metabolic flux analysis, metabolic engineering, modeling, parameter estimation, *Pseudomonas*, PHA

Lista de figuras

Figura 1 – Processo cíclico da engenharia metabólica.	19
Figura 2 – Limitações do MFA.	22
Figura 3 – Princípios do ^{13}C -MFA	23
Figura 4 – Representação dos isotopômeros de uma molécula com três átomos.	24
Figura 5 – Rede metabólica simples para representar o metabolismo do AcCoA.	25
Figura 6 – Formação do α – <i>cetoglutarato</i>	29
Figura 7 – Representação da cascata para resolução dos balanços de cumômeros	31
Figura 8 – Representação do processo de estimação de fluxos metabólicos	33
Figura 9 – Fluxograma representativo da ferramenta desenvolvida.	44
Figura 10 – Modelo estrutural simplificado do ciclo TCA.	45
Figura 11 – Modelo estrutural simplificado da <i>E.coli</i>	48
Figura 12 – Metabolismo central da <i>Pseudomonas sp.</i> LFM046	55
Figura 13 – Fragmentos característicos de propil-ésteres de 3HA obtidos por espectroscopia de massas	56
Figura 14 – Espectro de massa dos propil-ésteres dos monômeros 3HO e 3HD.	57
Figura 15 – Rede metabólica simplificada da <i>Pseudomonas sp.</i> LFM046 considerando todo o fluxo de carbono pela via ED	58
Figura 16 – Rede metabólica simplificada da <i>Pseudomonas sp.</i> LFM046 considerando um reciclo na via ED	59
Figura 17 – Variação dos isotopômeros de massa do PHA com a mudança no reciclo.	60
Figura 18 – Rede metabólica simplificada da <i>Pseudomonas sp.</i> LFM046 considerando a via das pentoses em equilíbrio.	61
Figura 19 – Rede metabólica simplificada de <i>Pseudomonas putida</i> . Abreviações dos metabólitos: G6P, glicose-6-fosfato; F6P, frutose-6-fosfato; G3P, gliceraldeído-3-fosfato; PYR, piruvato; P5P, <i>pool</i> agrupado de ribose-5-fosfato e xilose-5-fosfato; S7P, sedoheptulose-7-fosfato; E4P, eritrose-4-fosfato; AcCoA, acetilcoenzima A. Os fluxos têm unidades arbitrárias.	64
Figura 20 – Variância do parâmetro θ para diferentes proporções de $[\text{U-}^{13}\text{C}]$	65
Figura 21 – Variação dos isotopômeros de AcCoA com o fluxo na via das pentoses.	66
Figura 22 – Variação dos isotopômeros de massa de PHA com o fluxo na via das pentoses.	67

Figura 23 – Variância do parâmetro θ para diferentes proporções de $[5-^{13}\text{C}]$	68
Figura 24 – Exemplo de uma rede metabólica.	77

Lista de tabelas

Tabela 1 – Sumário dos softwares disponíveis para ^{13}C -MFA	35
Tabela 2 – Comparação da simulação dos isotopômeros de massa do glutamato no ciclo TCA.	46
Tabela 3 – Dados de marcação de entrada (GLC) e dos intermediários.	49
Tabela 4 – Diferentes valores iniciais para o processo de estimação	50
Tabela 5 – Comparação entre os fluxos estimados para E.coli.	51
Tabela 6 – Comparação do processo de estimação com e sem o gradiente analítico	51
Tabela 7 – Valores estimados para os fluxos livres e seus respectivos intervalos de confiança.	52
Tabela 8 – Valores estimados para os fluxos na coordenada aplicada e seus respectivos intervalos de confiança.	53
Tabela 9 – Medidas de Isotopômeros de massa e seus respectivos desvios padrão.	56
Tabela 10 – Comparação entre os Isotopômeros de massa simulados e os medidos	58
Tabela 11 –	62
Tabela 12 – Comparação entre os Isotopômeros de massa estimados e os medidos	63
Tabela 13 – Comparação entre a variância de θ para diferentes marcações de glicose.	65
Tabela 14 – Estequiometria e transições da rede apresentada na Figura 24. As letras minúsculas representam átomos sucessivos em cada metabólito.	77
Tabela 15 – Estequiometria e transições do ciclo TCA	85
Tabela 16 – Estequiometria e transições da rede simplificada de E.coli.	85
Tabela 17 – Rede metabólica simplificada da <i>Pseudomonas sp.</i> LFM046 considerando somente a via ED cíclica.	86
Tabela 18 – Rede metabólica simplificada da <i>Pseudomonas sp.</i> LFM046 considerando a via das pentoses em equilíbrio.	86
Tabela 19 – Rede metabólica genérica para <i>Pseudomonas</i>	87

Lista de abreviaturas e siglas

AMM	Matriz de transição atômica (<i>Atom Mapping Matrices</i>)
IMM	Matriz de transição de isotopômeros (<i>Isotopomer Mapping Matrices</i>)
EMU	Unidade metabólica elementar (<i>Elementary Metabolite Unit</i>)
MFA	Análise de Fluxos Metabólicos (<i>Metabolic Flux Analysis</i>)
¹³ C-MFA	¹³ C-Análise de Fluxos Metabólicos (<i>¹³C-Metabolic Flux Analysis</i>)
FBA	Análise de balanço de fluxos (<i>Flux Balance analysis</i>)
RMN	Ressonância magnética nuclear (<i>Resonance magnetic nuclear</i>)
EMP	Embden-Meyerhof Parnas
ED	Entner-Doudoroff
VP	Via das pentoses
SQP	Programação Quadrática Sequencial (<i>Sequential quadratic programming</i>)
<i>E.coli</i>	<i>Escherichia coli</i>
GC-MS	cromatografia gasosa acoplada à espectroscopia de massas (<i>Gas chromatography–mass spectrometry</i>)
PHA	Polihidroxialcanoato

Lista de símbolos

v	Fluxo metabólico (unidade arbitrária)
θ	Fluxos livres (unidade arbitrária)
$\bar{x}$	Fração isotopomérica (adimensional)
x	Fração de marcação de cumômero (adimensional)
Φ	Função objetivo (unidade arbitrária)

Sumário

1	Introdução	16
1.1	Objetivo	17
2	Fundamentos Teóricos	18
2.1	Engenharia Metabólica	18
2.2	Análise de Fluxos Metabólicos	21
2.3	Análise de Fluxos Metabólicos com carbono marcado	23
3	Revisão bibliográfica	25
3.1	Modelos matemáticos	25
3.1.1	Modelos com balanços de átomos marcados	25
3.1.2	Modelos com balanços de isotopômeros	28
3.2	Métodos de otimização	32
3.3	Softwares para ¹³C-MFA	34
4	Metodologia	36
4.1	Modelagem	36
4.1.1	Parametrização	36
4.1.2	Balanco de cumômeros	39
4.2	Estimação	40
4.3	Quantificação da incerteza	41
4.3.1	Sensibilidade analítica	41
4.3.2	Ajuste	42
4.3.3	Intervalo de confiança marginal	42
4.4	Ferramenta computacional desenvolvida	43
5	Resultados e discussão	45
5.1	Validação do Programa	45
5.1.1	Simulação do ciclo TCA	45
5.1.2	Estimação de fluxos em <i>E.coli</i>	46
5.2	Estudo de Caso: Estimação de fluxos em <i>Pseudomonas</i> <i>sp.</i> LFM046	53

5.2.1	Construção do modelo metabólico	54
5.2.2	Dados experimentais	54
5.2.3	Estimação de fluxos	57
5.3	Estudo de Caso:Design ótimo de experimentos em <i>Pseudomonas</i>	63
6	Conclusões	69
6.0.1	Sugestões para trabalhos futuros	69
	Referências ¹	71
	Apêndice A – Detalhes do Modelo matemático	77
A.1	Equações de restrição	77
A.2	Balanço de cumômeros	81
	Apêndice B – Destino dos átomos nas redes metabólicas	85

¹ De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

1 Introdução

O conhecimento dos genes e proteínas e suas interconexões não é suficiente para entender o metabolismo de um microrganismo, uma vez que a relação entre genoma e fenótipo é não linear e complexa (KITANO, 2002). Dessa forma, a biologia de sistemas se tornou uma poderosa ferramenta para o desenvolvimento de processos biotecnológicos, pois é possível adquirir informações importantes sobre o funcionamento de células (OTERO; NIELSEN, 2010). A engenharia metabólica utiliza a biologia de sistemas como um guia de mudanças genéticas para aumentar a produção em células de um produto de interesse. Ela funciona de forma cíclica em três estágios: modificações genéticas são realizadas em um microrganismo de interesse através de técnicas de modificação genética; experimentos de bioprocessos são realizados para coletar dados e realizar análises; os resultados são interpretados para se definir as próximas modificações. Este processo mantém-se até o microrganismo atingir um rendimento e produtividade aceitáveis (STEPHANOPOULOS, 2002).

Na etapa de análise no ciclo de engenharia metabólica é necessária a utilização de um modelo matemático, pois muitas reações químicas estão envolvidas em uma célula. Esses modelos são bastante utilizados para se realizar a análise fluxômica da célula, porque os fluxos metabólicos são de grande valor para se estudar as vias metabólicas utilizadas pelos micro-organismos. Portanto, muitos tipos de modelos foram criados com este objetivo, variando em complexidade e informação requerida. Dentre eles se destacam os baseados na estequiometria da reação, como a Análise de Fluxo Metabólico (MFA), que consiste em balanços em torno dos metabólitos e medidas de fluxos extracelulares para se calcular fluxos intracelulares no estado estacionário. No entanto, este modelo está limitado a redes de metabólitos sem ciclos e fluxos bidirecionais (STEPHANOPOULOS et al., 1998), o que exclui casos relevantes.

Para superar esta limitação as experiências com substratos marcados foram propostas, em geral a marcação utilizada é o isótopo 13 do carbono ^{13}C . Nestas experiências, os substratos marcados são alimentados em um sistema e esta marcação dos carbonos se espalha para todos os metabólitos de acordo com as transições de carbono nas reações. Em seguida, o padrão de marcação de carbono nos metabólitos internos é analisado com alguma técnica analítica, geralmente Ressonância Magnética Nuclear (RMN) e Espectrometria de Massa dos aminoácidos que são formados a partir de seus precursores. Quando as informações de marcação de carbono estão disponíveis, uma análise mais complexa é

possível, ^{13}C - Análise de Fluxo Metabólico (^{13}C -MFA). O novo modelo inclui os balanços de carbonos marcados ou de Isotopômeros, que são isômeros com átomos isotópicos. Nos balanços de carbono, um balanço de marcação para cada átomo de cada metabólito interno é realizado. Nos balanços de isotopômeros, um balanço para cada isotopômero possível é criado, para se calcular as frações isotopoméricas. Este novo modelo permite a estimação de uma quantidade maior de fluxos metabólicos.

Tendo em vista a relevância do tema, qualquer grupo de pesquisa que pretenda realizar uma pesquisa de excelência na área de engenharia metabólica, utilizando a ^{13}C -MFA, deve possuir uma ferramenta computacional que seja precisa e se adapte bem aos seus casos de interesse. E ainda, que possa ser alterada com facilidade para incluir modificações que se julguem necessárias, e que permita, por exemplo, quantificar a incerteza dos parâmetros, com o objetivo de possibilitar o projeto experimental.

1.1 Objetivo

O presente trabalho teve como objetivo desenvolver uma ferramenta computacional na plataforma *MATLAB*® v8.1 que realize cálculos de ^{13}C -MFA, com balanços de metabólitos e isotopômeros, e que utilize um módulo para estimar os fluxos e um módulo para a quantificação da incerteza das estimativas.

2 Fundamentos Teóricos

2.1 Engenharia Metabólica

Por volta dos anos 80, teve início a manipulação de genes em células, e começou a se pensar nelas como minifábricas que poderiam gerar produtos de interesse comercial (TYO et al., 2010). Nesse contexto, emerge um novo campo de pesquisa chamado Engenharia Metabólica que começa a investigar aspectos fundamentais no projeto de células (TETHNOL et al., 1991; BAILEY, 1991). O que diferencia esse novo campo da já existente Engenharia Genética, é a forma de olhar para a rede metabólica. Enquanto a Engenharia Genética olha cada reação bioquímica de forma individual, a Engenharia Metabólica olha de forma sistêmica, tentando entender como as modificações genéticas influenciariam no metabolismo como um todo (STEPHANOPOULOS, 1994). Assim o objetivo principal deste campo de estudo é a construção de cepas com alto rendimento e produtividade passando por um processo cíclico (STEPHANOPOULOS, 1994), o que inclui a criação de novas vias; avaliação da viabilidade termodinâmica de vias; localização de pontos chaves nas redes metabólicas.

Notoriamente o avanço da área só foi possível graças a tecnologias de modificação genética como o DNA recombinante, e mais recentemente da utilização da CRISPR–Cas9 (LI et al., 2015). No entanto, embora essas técnicas permitam modificações específicas nas redes metabólicas, não trouxeram avanço na capacidade de se localizar reações críticas que necessitam ser manipuladas, sendo esse processo feito de maneira *ad hoc*. Com isso, a Engenharia Metabólica tem um papel crucial na análise sistêmica da rede, sendo utilizadas ferramentas matemáticas para esse fim.

O processo então ocorre de maneira cíclica, como ilustrado na Figura 1. Pode-se separá-lo em três partes: a síntese que realiza modificações genéticas nas células; a análise que caracteriza os resultados obtidos; o design que interpreta os resultados e projeta novas modificações. Esse ciclo se mantém até que as células tenham as características esperadas.

Figura 1 – Processo cíclico da engenharia metabólica.

Fonte: Adaptado de (NIELSEN, 2001)

Dentre as aplicações possíveis da engenharia metabólica em processos biotecnológicos pode-se citar a utilização de diferentes substratos por uma célula; produção de novos produtos químicos; melhoramento do rendimento e da seletividade; aumento da robustez da célula, ou seja, tolerância a compostos tóxicos (NIELSEN et al., 2014). A engenharia metabólica tem uma grande interface com o setor industrial, e diversos casos de sucesso já foram relatados, destacando-se entre eles:

- A DuPont empregou as técnicas de engenharia metabólica para modificar uma *E.coli* e possibilitar a utilização de um substrato de baixo custo na produção de 1,3-propanodiol, produto químico chave na produção de alguns polímeros (NAKAMURA; WHITED, 2003).
- A Gevo and Butamal, uma joint venture da DuPont e BP, desenvolveu um processo para a produção de isobutanol que pode ser utilizado tanto em biocombustíveis, como em fármacos (HONG; NIELSEN, 2012).
- A Amyris aumentou a produção de artemisinina de 1.6 g/L para 25 g/L aplicando técnicas de engenharia metabólica (PADDON et al., 2013). A Artemisina é um grupo de fármacos utilizado na cura da malária.

A etapa de análise é crítica no processo mostrado na Figura 1, pois milhares de reações químicas estão envolvidas em um micro-organismo. Sendo assim, faz-se necessária a utilização de modelos matemáticos para essa análise. Diversos modelos foram desen-

volvidos até o presente momento, com diferentes graus de detalhamento, para o estado estacionário ou dinâmico. Destacam-se dentre os modelos para o estado estacionário:

- Modelos baseados na estequiometria das reações bioquímicas, que são os mais simples, mas podem ser extremamente valiosos nas análises. Dentre esses se destacam: Análise de Fluxos Metabólicos (MFA) ([STEPHANOPOULOS et al., 1998](#)), o mais antigo, que emprega os balanços dos metabólitos internos junto a dados dos fluxos extracelulares; Análise de balanço de fluxos (FBA) ([ORTH et al., 2010](#)), que não necessita de dados experimentais e utiliza um processo de otimização para encontrar fluxos plausíveis; Modos elementares ([SCHUSTER; HLGETAG, 1994](#)), que também não necessita de dados experimentais, e encontra as vias metabólicas elementares que podem manter a célula em um estado estacionário.
- Modelos baseados em fluxo de carbono são mais detalhados, e para além da estequiometria das reações utilizam a informação da transição dos átomos de carbono dentro da célula. Necessitam da realização de experimentos com carbono marcado e de análises que descrevam como esse carbono se distribui na rede metabólica. A ferramenta mais utilizada é a Análise de Fluxos Metabólicos com carbono marcado (^{13}C -MFA) ([WIECHERT, 2001](#)), que emprega os balanços dos metabólitos internos, bem como balanços de carbonos marcados nesses metabólitos.
- Modelos baseados no mecanismo da reação ([WIECHERT; NOACK, 2011](#)), modelos cinéticos. Em comparação aos modelos estequiométricos incluem a regulação metabólica que leva a uma determinada distribuição de fluxo, ou seja, explica como um determinado padrão emerge, e tem um poder preditivo maior. No entanto, esse modelo exige um conhecimento *a priori* do mecanismo de reação das enzimas, e uma quantidade grande de dados de intermediários das reações, informações que em muitos casos não estão disponíveis.
- Modelos baseados em regulação gênica ([WIECHERT, 2002](#)), incluem ainda mais informações e são os modelos mais detalhados existentes. Eles incluem a regulação gênica que é responsável pela quantidade de enzimas ativas. Esses modelos são complexos pois demandam um conhecimento dos mecanismo de regulação que são fenômenos inerentemente estocásticos, o que dificulta a modelagem.

2.2 Análise de Fluxos Metabólicos

A MFA é a mais antiga ferramenta na Engenharia Metabólica (AIBA; MATSUOKA, 1979), e consiste no cálculo de fluxos intracelulares através de um modelo estequiométrico que contém balanços em torno dos metabólitos internos. Para isso, o modelo precisa de medidas de fluxos externos, que normalmente são taxas de consumo de substrato e taxas de excreção de metabólitos. O cálculo desses fluxos é essencial para se compreender o papel das diferentes vias nos processos metabólicos, tendo importantes aplicações como: identificação de vias alternativas; cálculo de fluxos extracelulares não medidos; cálculo do rendimento máximo teórico; verificação da flexibilidade de pontos de ramificação na rede metabólica (STEPHANOPOULOS et al., 1998).

Os balanços para os metabólitos internos podem ser escritos da seguinte forma (STEPHANOPOULOS et al., 1998):

$$\frac{dX_{met}}{dt} = r_{met} - \mu X_{met} \quad (2.1)$$

Em que X_{met} é o vetor que contém as concentrações dos metabólitos internos, r_{met} o vetor contendo a taxa de formação de cada metabólito e μ a taxa específica de crescimento. O termo à esquerda na equação se refere ao acúmulo nos *pools* metabólicos. Como a concentração desses *pools* se ajusta rapidamente a novos níveis, mesmo com grandes perturbações no ambiente (STEPHANOPOULOS et al., 1998), é razoável se assumir um estado pseudo-estacionário dos metabólitos, ou seja, não há acúmulo. Portanto esse termo pode ser desprezado:

$$0 = r_{met} - \mu X_{met} \quad (2.2)$$

O segundo termo à direita é o termo de diluição do *pool* decorrente do crescimento celular. No entanto, esse efeito é na maioria dos casos desprezível (STEPHANOPOULOS et al., 1998), pois a concentração nos *pools* é muito pequena se comparada aos fluxos de entrada e saída. Dessa forma, esse termo também pode ser eliminado da equação:

$$r_{met} = G^T v = 0 \quad (2.3)$$

O termo restante é o que contabiliza a soma dos fluxos de entrada e saída do *pool* metabólico, o qual pode ser separado em uma forma simples e linear que é a multiplicação

da matriz estequiométrica das reações bioquímicas G^T e o vetor contendo os fluxos de cada reação v . Pode-se ainda separar a matriz estequiométrica e o vetor dos fluxos em variáveis medidas (m) e variáveis calculadas (c):

$$0 = G^T v = G_m^T v_m + G_c^T v_c \quad (2.4)$$

Sendo assim, o número de incógnitas do sistema linear de equações acima é igual ao tamanho do vetor v_c , e o número de balanços é igual ao número de linhas da matriz G^T (número de metabólitos internos). Caso haja mais equações do que incógnitas, o sistema é sobredeterminado e um estimador de mínimos quadrados pode ser utilizado para encontrar uma solução (STEPHANOPOULOS; VALLINO, 1993). Já no caso de se ter o mesmo número de equações e balanços, a matriz G_c^T é quadrada, e se ela tiver uma inversa, o sistema é determinado, com a seguinte solução:

$$v_c = -(G_c^T)^{-1} G_m^T v_m \quad (2.5)$$

Portanto, essa solução pode ser usada sempre que a matriz G_c^T possuir inversa, o que não acontece em algumas situações. Essa singularidade emerge frequentemente de questões físicas da célula, pois ela pode utilizar mais de uma via metabólica para realizar a mesma função. Isso então acarreta em uma dependência linear entre reações na matriz estequiométrica, o que gera a singularidade. Desse fato surgem as limitações do modelo MFA, e a Figura 2 ilustra como essas limitações se dão. Elas podem ocorrer em: a) vias paralelas; b) vias cíclicas; c) reações bidirecionais; d) fluxos que se dividem contendo cofatores não balanceados.

Figura 2 – Limitações do MFA.

Fonte: (STEPHANOPOULOS et al., 1998)

2.3 Análise de Fluxos Metabólicos com carbono marcado

Como visto, o MFA falha em descrever alguns tipos de vias importantes do metabolismo celular. Além disso, em redes muito grandes, as medições dos fluxos externos não são suficientes para estimar todos os fluxos internos. Dado esse cenário, faz-se necessário a introdução de mais dados experimentais. No começo da década de 80 (WIECHERT, 2001), iniciou-se a realização de experimentos com substratos marcados, dentre eles sendo o mais utilizado o isótopo ^{13}C (WIECHERT, 2001). Uma ilustração desses experimentos se encontra na Figura 3, onde um substrato com carbono marcado é alimentado no meio celular. Essa marcação se distribui pelos metabólitos internos a depender das transições de carbono nas reações bioquímicas no interior da célula. Uma vez atingido o estado estacionário isotópico, essas marcações se mantêm constantes e pode-se medi-las. Dessa maneira, têm-se além das medidas dos fluxos extracelulares, medidas das marcações internas dos átomos de carbono. As técnicas de medições dessas marcações são usualmente Ressonância Magnética Nuclear (RMN) e Espectroscopia de Massa (SZYPERSKI, 1998), e a análise normalmente é feita a partir de aminoácidos constituintes da biomassa celular, as quais vêm de precursores internos do metabolismo central (SZYPERSKI, 1995).

Figura 3 – Princípios do ^{13}C -MFA .

Fonte: (WIECHERT, 2001)

Para analisar as informações de marcação, são necessários novos balanços além dos balanços de metabólitos, são estes os balanços de átomos marcados e balanço de isotopômeros. Os balanços de átomos marcados consistem em balanços para cada átomo de cada metabólito interno. Com isso é possível saber qual a fração de marcação de cada átomo na célula, como ilustrado na Figura 3 . Esse tipo de modelagem já é aplicado desde

a década de 80, e foi utilizado para modelar importantes vias como gliconeogênese(KATZ, 1985), ciclo de Krebs(KELLEHER, 1985), dentre outras.

Os modelos baseados em balanços de átomos marcados são simples e possuem solução analítica quando o vetor de fluxo é fixado, o que explica a grande utilização deles em diferentes trabalhos (CHRISTENSEN; NIELSEN, 2000; MARX et al., 1996; PORTAIS et al., 1993). No entanto, embora esse modelo, a depender do sistema, seja suficiente para a estimativa de todos os fluxos (WINDEN et al., 2001a), ele não fornece toda a informação de marcação nos metabólitos.

Por vezes se faz necessário utilizar o conceito de Isotopômeros, moléculas isômeras que possuem átomos isótopos. Para ilustrar esse conceito, a Figura 4 exibe uma representação dos isotopômeros possíveis para uma molécula com três átomos.

Figura 4 – Representação dos isotopômeros de uma molécula com três átomos. Os números em cima de cada isotopômero representam sua fração, enquanto os números à direita representam as frações de marcação atômica. Os valores foram escolhidos para simplificar a exemplificação, e o fato de certas frações se repetirem em nada representa alguma regra física.

Primeiramente, é oportuno atentar para a quantidade de isotopômeros na Figura 4; são 8 para uma molécula com 3 átomos. Como regra geral, o número de isotopômeros para um metabólito é igual a 2^n onde n é o número de átomos. Desta forma, a informação sobre o sistema cresce exponencialmente.

Para se compreender melhor como os isotopômeros fornecem um cenário completo do padrão de marcação na célula, é necessário olhar para as frações com mais atenção e notar que caso as frações isotopoméricas de isotopômeros que possuem dois átomos marcados fossem distribuídas igualmente entre os isotopômeros com apenas um átomo marcado, as frações de marcação de átomos não se alterariam. Portanto, há duas distribuições diferentes de marcações em isotopômeros que geram a mesma distribuição atômica.

3 Revisão bibliográfica

3.1 Modelos matemáticos

Os modelos propostos para representar a distribuição dos carbonos marcados nas vias metabólicas são, via de regra, mais complexos que os modelos para o MFA. A razão para isso vem de dois fatores: o aumento do número de equações que são utilizadas para descrever as marcações, e o surgimento de termos bilineares em alguns modelos. Esses modelos podem ser classificados em duas classes, os que utilizam balanços para átomos de carbono marcados e os que utilizam balanços para isotopômeros, sendo o último, o que fornecerá a melhor descrição do sistema. Pretende-se assim apresentar aqui as principais contribuições de modelagem em estado estacionário na literatura para os dois tipos de balanços, bem como as características principais de cada uma delas.

3.1.1 Modelos com balanços de átomos marcados

Os modelos com balanços de átomos marcados foram os primeiros modelos criados, sendo utilizados desde a década de 80 na área de Engenharia Metabólica. Para ilustrar como são construídos os modelos nesses trabalhos, doravante chamado de modelo clássico, um exemplo apresentado por [Blum e Stein \(1982\)](#) será utilizado. Considere um sistema formado por dois *pools* de Acetilcoenzima A (AcCoA), apresentado na Figura 5. Nesse exemplo, Piruvato e Hexanoato são os metabólitos que podem ser alimentados ao sistema com seus carbonos marcados. A distribuição isotópica nos *pools* de AcCoA será portanto, função dos fluxos de carbono e das marcações nos substratos.

Figura 5 – Rede metabólica simples para representar o metabolismo do AcCoA. $V_1 - V_6$ representam os fluxos de AcCoA. Piruvato e Hexanoato são os metabólitos passíveis de marcação.

Fonte: (BLUM; STEIN, 1982)

O número de balanços possíveis para cada metabólito é igual ao seu número de átomos. Nesse exemplo, é possível construir quatro balanços, um para cada átomo dos dois AcCoA, que são os metabólitos internos nessa rede. Por questão de simplicidade, apresentar-se-á somente as equações referentes aos átomos do segundo *pool* de AcCoA, são elas:

$$(V_3 + V_6)AcCoA_{II}(1) = V_2AcCoA_I(1) + \left(\frac{V_5}{3}\right)(Hex(1) + Hex(3) + Hex(5)) \quad (3.1)$$

$$(V_3 + V_6)AcCoA_{II}(2) = V_2AcCoA_I(2) + \left(\frac{V_5}{3}\right)(Hex(2) + Hex(4) + Hex(6)) \quad (3.2)$$

Em que $AcCoA(i)$ é a fração de marcação do átomo i do metabólito AcCoA, e $Hex(i)$ é a fração de marcação do átomo i do metabólito hexanoato. V_n é o fluxo metabólico na reação n . Tem-se então no lado esquerdo das equações os termos de saída de átomos marcados, e no lado direito os termos de entrada. Os números entre parêntesis indicam o átomo do metabólito representado. As saídas do *pool* de $AcCoA_{II}$ são os fluxos V_3 e V_6 , que aparecem multiplicados pelas frações de marcação do próprio *pool* $AcCoA_{II}$. As entradas aos fluxos V_2 e V_5 , que vêm do $AcCoA_I$ e Hexanoato respectivamente. Nesse momento é crucial que se saiba quais as transições atômicas envolvidas nessas reações, pois os termos que vão multiplicar esses fluxos são exatamente as frações de marcação de origem desses átomos. Assim, no balanço do átomo $AcCoA_{II}(1)$, o fluxo V_2 é multiplicado pelo termo $AcCoA_I(1)$ e o fluxo V_5 é multiplicado por $(Hex(1) + Hex(3) + Hex(5))/3$, que são os átomos que originam $AcCoA_I(1)$ em cada fluxo. Ao se olhar mais atentamente as equações 3.1 e 3.2, principalmente no que diz respeito às suas estruturas, percebe-se que as equações possuem termos bilineares dos fluxos em relação às marcações. No entanto, essas equações são resolvidas geralmente tendo como entrada os fluxos de carbono, com isso elas se tornam lineares e de fácil solução analítica. Para a solução do sistema composto por todas as equações de balanço, um procedimento possível é organizar as equações de estado estacionário em forma de matriz e resolver o sistema por uma inversão de matriz, da seguinte forma:

$$As = b \quad \rightarrow \quad s = A^{-1}b \quad (3.3)$$

Onde A é a matriz cujos elementos são os coeficientes das equações de balanço, s é o vetor contendo as frações de marcação desconhecidas e b o vetor com as frações

de marcação conhecidas. Para o exemplo dado na Figura 5, o sistema fica com a seguinte forma:

$$\begin{bmatrix} \frac{V_2+V_4}{V_1} & 0 & -\frac{V_3}{V_1} & 0 \\ 0 & \frac{V_2+V_4}{V_1} & 0 & -\frac{V_3}{V_1} \\ -\frac{3V_2}{V_5} & 0 & \frac{3(V_3+V_6)}{V_5} & 0 \\ 0 & -\frac{3V_2}{V_5} & 0 & \frac{3(V_3+V_6)}{V_5} \end{bmatrix} x \begin{pmatrix} AcCoA_I(1) \\ AcCoA_I(2) \\ AcCoA_{II}(1) \\ AcCoA_{II}(2) \end{pmatrix} = \begin{pmatrix} Pyr(2) \\ Pyr(3) \\ Hex(1) + Hex(3) + Hex(5) \\ Hex(2) + Hex(4) + Hex(6) \end{pmatrix}$$

Olhando essa equação fica evidente quais as informações necessárias para se determinar as frações de marcação dos dois *pools* de AcCoA, são elas, a marcação do segundo e terceiro átomos do Piruvato, e a marcação dos seis átomos do Hexanoato.

Embora esse método de resolução possa ser bem simples dada sua operação algébrica, em casos de redes metabólicas maiores, a matriz A possui grandes dimensões e a solução desse sistema pode tomar bastante tempo computacional. Além disso, qualquer alteração que se queira fazer na rede metabólica, pode ser onerosa. Para reduzir esses problemas, Zupke e Stephanopoulos (1994) propuseram uma nova maneira de representar essas equações, eles introduziram as Atom Mapping Matrices (AMM), que descrevem as transições atômicas entre dois átomos.

Novamente, para o exemplo da Figura 5, a transição entre os *pools* de AcCoA pode ser descrita pela seguinte AMM:

$$[AcCoA_I > AcCoA_{II}]_{v_2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Nesta AMM lê-se que o primeiro átomo do $AcCoA_I$ vai para o primeiro átomo do $AcCoA_{II}$, e o segundo átomo do $AcCoA_I$ vai para o segundo átomo do $AcCoA_{II}$ na reação V_2 , o que gera, como esperado, uma matriz identidade. Já para a transição do Pyr para o $AcCoA_I$ temos:

$$[Pyr > AcCoA_I]_{v_1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

O que informa que o segundo e terceiro átomos do Piruvato vão para o $AcCoA_I$. Assim temos uma matriz AMM para cada reação. Para fins ilustrativos, coloca-se então as equações 3.1 e 3.2 para os átomos de $AcCoA_{II}$ com a nova notação:

$$(V_3 + V_6)AcCoA_{II} = V_5[Hex > AcCoA_{II}]Hex + V_2[AcCoA_I > AcCoA_{II}]AcCoA_I \quad (3.4)$$

É notório observar então que em relação ao modelo clássico, essa nova abordagem organiza as equações de balanço atômico em uma equação matricial para cada metabólito, o que aumenta o número de equações se comparado ao método em 3.3. No entanto, com a vantagem de se poder resolver sistemas menores em vez de um grande sistema. Além disso, essa nova notação permite a alteração mais fácil do modelo estrutural. A solução do sistema de equações pode ser encontrada iterativamente, sem a necessidade de inversão de matrizes.

Em paralelo ao modelo anterior, outra abordagem semelhante foi proposta por Wiechert e Graaf (1997) no primeiro de uma série de quatro artigos que estudavam reações bidirecionais em redes metabólicas. Foi apresentado um modelo para balanços atômicos que permite a determinação de todas as marcações atômicas através de uma única expressão analítica:

$$x = -\left(\sum_i \vec{v}_i \cdot \vec{P}_i + \sum_i \overleftarrow{v}_i \cdot \overleftarrow{P}_i\right)^{-1} \cdot \left(\sum_i \vec{v}_i \cdot P_i^{inp}\right) \cdot x^{inp} \quad (3.5)$$

Onde x é vetor contendo as frações de marcação desconhecidas; x^{inp} o vetor contendo as frações de marcação dos substratos; $\vec{v}_i$ e $\overleftarrow{v}_i$ os fluxos nas direções direta e inversa respectivamente; $\vec{P}_i$, $\overleftarrow{P}_i$ e P_i^{inp} são as Atom Transitions Matrices, análogas as AMM's, das reações direta, inversa e de entrada (com o substrato) respectivamente.

3.1.2 Modelos com balanços de isotopômeros

Um dos primeiros modelos propostos para representar a distribuição de isotopômeros foi o de Schmidt et al. (1997), que propuseram as "Isotopomer Mapping Matrices"(IMM), análogas às AMM's. Cada matriz descreve a transição entre os isotopômeros dos reagentes para os isotopômeros dos produtos. Para ilustrar esse conceito, a Figura 6 traz um exemplo da formação do α – *cetoglutarato*(*akg*) cujo balanço se encontra na equação 3.6.

Figura 6 – Formação do α – *cetoglutarato*

Fonte: (SCHMIDT et al., 1997)

$$r_1 \cdot (IMM_{aca>akg} \cdot I_{aca}) \otimes (IMM_{oaa>akg} \cdot I_{oaa}) - r_2 \cdot I_{akg} = 0 \quad (3.6)$$

Onde I é o vetor contendo as frações isotopoméricas do metabólito, r os fluxos metabólicos e o operador $\otimes$ significa 'multiplicação ponto a ponto'. Essa notação é importante já que ela concatena os 2^5 balanços para o α –*cetoglutarato* em uma única equação matricial. Além do grande número de equações geradas, uma outra dificuldade desse novo modelo é o surgimento de termos bilineares entre frações de marcação em reações bimoleculares ou de ordem maior. Como vemos na equação 3.6, na multiplicação de um termo do oxaloacetato (oaa) pelo AcCoA (aca). Dessa maneira, a equação não possui solução analítica mesmo que seja dada a distribuição de fluxos metabólicos.

Para resolver a questão da falta de solução analítica, Wurzel e Graaf (1999) propõem, no terceiro artigo da série sobre reações bidirecionais, uma transformação de variáveis que permite que os balanços de isotopômeros sejam resolvidos como uma cascata de equações lineares. Para isso, foi criado um novo conceito chamado cumômero (cumulated isotopomer fraction), que consiste em variáveis que são somatórios das frações isotopoméricas. A transformação do espaço de isotopômeros para o espaço de cumômeros é unívoca, ou seja, temos exatamente o mesmo número de cumômeros (d) e isotopômeros (I), como é ilustrado abaixo.

$$(I_{000}, I_{001}, I_{010}, I_{011}, I_{100}, I_{101}, I_{110}, I_{111}) \Leftrightarrow (d_{xxx}, d_{xx1}, d_{x1x}, d_{x11}, d_{1xx}, d_{1x1}, d_{11x}, d_{111})$$

A transformação do espaço de isotopômeros para o espaço de cumômeros é exemplificada abaixo para uma molécula de três átomos de carbono.

Frações de Cumômeros-peso 0

$$d_{xxx} = \sum_{i,j,k=0}^1 I_{ijk} = 1$$

Frações de Cumômeros- peso 1

$$d_{1xx} = \sum_{j,k=0}^1 I_{1jk}$$

$$d_{x1x} = \sum_{i,k=0}^1 I_{i1k}$$

$$d_{xx1} = \sum_{i,j=0}^1 I_{ij1}$$

Frações de Cumômeros- peso 2

$$d_{11x} = \sum_{k=0}^1 I_{11k}$$

$$d_{1x1} = \sum_{j=0}^1 I_{1j1}$$

$$d_{x11} = \sum_{i=0}^1 I_{i11}$$

Frações de Cumômeros- peso 3

$$d_{111} = I_{111}$$

Nas equações acima, I_{ijk} representa os isotopômeros onde seus índices i,j e k podem assumir os valores 1 ou 0, que representam átomos marcados ou não respectivamente; d_{xxx} representa o cumômero que é igual ao somatório de todos os isotopômeros que podem ter seus átomos marcados ou não, o que claramente é igual a 1, pois é a soma de todas as frações isotopoméricas. As frações de cumômeros podem então ser divididas de acordo com o seu "peso", que significa quantos dos seus átomos têm a sua marcação fixada. o cumômero d_{xxx} tem peso zero, pois todos os seus átomos podem estar marcados ou não. Quanto aos átomos de peso 1, temos três possibilidades, o somatório de todos os isotopômeros com o primeiro átomo marcado d_{1xx} , com o segundo d_{x1x} ou com o terceiro d_{xx1} . É essencial notar aqui, que os cumômeros de peso 1 são exatamente as frações de marcação atômicas, com isso temos sempre n cumômeros desse tipo. Os cumômeros de peso 2 são análogos aos anteriores com duas marcações fixadas, e por fim o de peso 3 que é na verdade um só cumômero que é igual ao isotopômero com todos os átomos marcados, sendo assim, temos sempre como peso máximo de cumômeros o número de átomos de carbono no metabólito.

A Figura 7 apresenta uma representação de como os balanços são resolvidos em cascata; Em que o nível 0 contém os balanços de peso 0, que como visto, são iguais a

1, assim esse nível é igual ao MFA clássico, onde só fluxos metabólicos estão envolvidos; O nível 1 representa os balanços atômicos que vão fornecer os cumômeros de peso 1 para o nível 2. Note então que as duas primeiras etapas na resolução desse sistema são a resolução de um MFA e de um balanço de átomos marcados. No nível 2, todos os termos bilineares são fixados com a utilização dos cumômeros de peso 1 calculados no nível 1, o que permite o cálculo dos cumômeros de peso 2, que por sua vez são alimentados ao nível 3 que fixa os seus termos bilineares e calcula os cumômeros de peso 3. Esse processo então prossegue até o último nível existente, o que permite a transformação das equações bilineares em equações lineares, exatamente na mesma quantidade.

Figura 7 – Representação da cascata para resolução dos balanços de cumômeros

Fonte: (WURZEL; GRAAF, 1999)

O modelo dos cumômeros representou um grande avanço no cálculo das frações isotopoméricas, principalmente devido à simplicidade de sua resolução. No entanto, quando aplicado a redes grandes, o número de equações pode crescer rapidamente, pois a transformação preserva o número de variáveis, 2^n para cada metabólito. Essa característica pode se tornar um empecilho principalmente em experimentos que utilizam mais de um traçador, no caso da gliconeogênese por exemplo, utilizando três traçadores (^2H , ^{13}C , ^{18}O), o número de isotopômeros supera os 2 milhões (ANTONIEWICZ et al., 2007).

Para solucionar essa questão, surgiram dois novos modelos que têm a finalidade de reduzir o número de equações geradas e assim diminuir o esforço computacional necessário para simulação. O primeiro deles foi desenvolvido por [Windén et al. \(2002\)](#) que apresentou o novo conceito chamado de Bondomer que é similar ao conceito de isotopômeros, porém descreve as ligações C-C e não os átomos de carbono propriamente. Os Bondomers podem ser transformado em Cumulative Bondomers assim como os Isotopômeros podem ser transformados em cumômeros. Essa transformação permite também uma solução analítica para o conjunto de balanços gerados. A redução do número de equações é feita eliminando-se as variáveis que não estão relacionadas a fragmentos de metabólitos que poderão ser analisadas por 2D [^{13}C , ^1H] RMN. O que impõe uma limitação ao modelo, pois restringe-se o método analítico a ser utilizado nas medições. Uma outra limitação desse modelo é que ele só pode ser aplicado a experimentos onde todos os substratos são uniformemente marcados.

O outro modelo que visa diminuir o número de balanços gerados, e não possui as desvantagens do Bondomers, foi proposto por [Antoniewicz et al. \(2007\)](#), que consiste em transformação de variáveis chamadas Elementary Metabolite Unit (EMU), que são subconjuntos de átomos dos metabólitos. Essa transformação permite, como as anteriores, uma solução analítica em cascata e tem ainda a peculiaridade de utilizar o mínimo de informação necessária para simular os metabólitos de interesse. Com isso essa abordagem diminui substancialmente o número de equações e pode ser aplicado a qualquer substrato. Por fim, [Srour et al. \(2011\)](#) introduziram uma outra transformação de variáveis chamada Fluxomers, nessa abordagem as variáveis de estado de fluxo e de marcação de isotopômeros são combinadas. Os balanços dessas novas variáveis são resolvidos analiticamente com um esforço computacional menor que as propostas anteriores e sem eliminar nenhuma informação do sistema, ou seja, todos os fluxos e marcações são calculados.

3.2 Métodos de otimização

Como visto na seção anterior, o balanço de cumômeros permite o cálculo das frações isotopoméricas através de uma cascata de equações lineares, tendo como entrada os fluxos metabólicos. Contudo, o problema inverso é o que possui mais aplicações práticas ([WIECHERT, 2001](#)), ou seja, determinar a distribuição de fluxos metabólicos dentro de uma célula, a partir das marcações e fluxos medidos. Esse processo de estimação dos fluxos

está representado na Figura 8, onde um processo cíclico ocorre, e os dados simulados pelo modelo são comparados aos dados experimentais por algum critério, geralmente a soma dos erros quadrados. O método de estimação (otimização) avalia esse critério e propõe uma nova distribuição de fluxos, até que se atinja um mínimo.

A otimização em ^{13}C -MFA possui características que a torna em um problema bastante complexo, como uma função objetivo quadrática, restrições lineares (equação estequiométrica) e não lineares (balanços de isotopômeros); espaço de grande dimensão de procura (WIECHERT, 2001); função objetivo não diferenciável; e a presença de múltiplos ótimos. É importante salientar que embora o balanço de isotopômero possa ser resolvido de forma linear para obter as frações de marcação, ele continua sendo bilinear em relação aos fluxos.

Figura 8 – Representação do processo de estimação de fluxos metabólicos

Fonte: adaptado de (WIECHERT, 2001)

Diversos métodos foram aplicados para o processo de estimação, dentre eles os métodos baseados em gradientes foram utilizados devido à sua rápida convergência, embora os resultados dependam bastante do ponto inicial (ZOMORRODI et al., 2012). Os métodos híbridos foram o que obtiveram os melhores resultados, como em Wiechert e Graaf (1997), onde se utilizou o método de Programação Quadrática Sequencial (SQP) junto ao método de Levenberg-Maquardt. E em Yang et al. (2008), que utilizou a SQP junto ao método de *Trust Region*.

Métodos estocásticos também começaram a ser aplicados a problemas de ^{13}C -MFA, como os Algoritmos Evolucionários (EA), que possuem bom desempenho em espaços de

busca de alta dimensão, apesar de sofrerem com esforço computacional e convergência lenta para ótimos locais (CHEN et al., 2007; YANG et al., 2007). A maioria dos trabalhos recentes que apresentaram novas propostas de otimização propuseram abordagens estocásticas, embora esses tenham se mostrado ineficientes em encontrar o mínimo global (ZOMORRODI et al., 2012). Como exemplo, tem-se os algoritmos de enxame de partículas e enxame de partículas com comportamento quântico, ambos utilizando uma função de penalidade que deixa o problema irrestrito (LONG et al., 2009; LONG et al., 2014).

Por fim, há as abordagens de otimização global. Zhao e Shimizu (2003) aplicaram um algoritmo híbrido, onde o método estocástico EA é utilizado para busca global, e os melhores resultados são levados para uma busca local usando o método de Levenberg-Maquardt. Riascos et al. (2005) desenvolveram um método *Branch and Bound* baseado em uma relaxação convexa do problema. E ainda, outra abordagem determinística utilizando o pacote BARON (branch-and-reduce optimization navigator) também foi aplicada a problemas pequenos (ZOMORRODI et al., 2012).

3.3 Softwares para ^{13}C -MFA

Diversos softwares para ^{13}C -MFA foram desenvolvidos pelos principais grupos de pesquisa da área em todo o mundo. Esse fato permitiu um grande aumento de publicações na área, o que não necessariamente foi acompanhado por um aumento da qualidade dos trabalhos (CROWN; ANTONIEWICZ, 2013). Isso se deve principalmente ao fato de que embora as diversas ferramentas disponíveis facilitem os cálculos, elas ainda necessitam de usuários que entendam bem o problema, e as características de cada software utilizado. Na Tabela 1 são exibidos os softwares existentes, bem como os tipos de modelo, *solver*, linguagem de programação e qual o grupo desenvolvedor.

Mais detalhes sobre os diferentes softwares, bem como os diferentes *solvers* utilizados se encontram em Guo et al. (2015). Note-se que os principais modelos utilizados são EMU e cumômero, o primeiro pela sua aplicação a sistemas de grande porte, e o segundo devido à sua fácil implementação e performance. A principal linguagem utilizada foi a MATLAB, provavelmente devido à sua facilidade de lidar com matrizes esparsas. E como consequência o solver fmincon, foi o mais utilizado, pois é o único presente no MATLAB que é aplicável ao problema.

Tabela 1 – Sumário dos softwares disponíveis para ^{13}C -MFA

Software	Modelo	Solver	Linguagem	Referência
13CFLUX2	Cumômero/EMU	IPOPT	C++	(WEITZEL et al., 2013)
Metran	EMU	fmincon	MATLAB	(ANTONIEWICZ et al., 2007)
FIA	Fluxomer	SNOPT	Python	(SROUR et al., 2011)
influx_s	Cumômero	NLSIC	C++	(SOKOL et al., 2012)
C13	SFL	fmincon	MATLAB	(CVIJOVIC et al., 2010)
OpenFlux2	EMU	fmincon	MATLAB	(SHUPLETSOV et al., 2014)
FiatFlux	MDV	fmincon	MATLAB	(ZAMBONI et al., 2005)
INCA	EMU	Próprio	MATLAB	(YOUNG, 2014)
OpenMebius	EMU	Próprio	MATLAB	(KAJIHATA et al., 2014)

Fonte: adaptado de (GUO et al., 2015)

4 Metodologia

4.1 Modelagem

O modelo utilizado no presente trabalho consiste na utilização da equação estequiométrica e nos balanços de isotopômeros, ambos formulados para o estado pseudo-estacionário metabólico e estado estacionário isotópico. O modelo possui duas variáveis de estado, a variável de fluxos metabólicos (v) e variável de fração de marcação de cumômeros (x). As hipóteses sobre o sistema biológico necessárias são simples se comparadas a modelos dinâmicos e cinéticos. Nenhuma hipótese sobre os mecanismos de reações enzimáticas ou transporte de metabólitos precisa ser feita, bem como o balanço energético não precisa ser fornecido. As principais hipóteses feitas são:

- O sistema pode ser representado por compartimentos com concentração de isotopômeros homogênea (*pools*).
- O sistema está em estado pseudo-estacionário metabólico.
- O sistema está em estado estacionário isotópico.
- A rede metabólica considerada representa adequadamente as reações *in vivo*.
- Todos os destinos dos átomos na rede metabólica são conhecidos.
- Não há efeitos mensuráveis nos fluxos metabólicos devido às marcações nas moléculas.

Neste capítulo, primeiro se apresentará a parametrização da variável de fluxo na seção 4.1.1; Posteriormente serão apresentadas as equações de balanços de isotopômeros e cumômeros na seção 4.1.2; o problema de estimação será apresentado na seção 4.2, onde os fluxos são tratados como parâmetros; na seção 4.3 os métodos para a quantificação da incerteza serão apresentados; e por fim, a estrutura do software construído pode ser encontrada na seção 4.4.

4.1.1 Parametrização

Tradicionalmente os fluxos metabólicos em reações bidirecionais são representados por uma variável de fluxo para cada direção da reação, assim como é apresentado na Figura 5. Essa formulação tem uma interpretação física muito fácil, no entanto, problemas

numéricos e de identificabilidade causados por altos fluxos bidirecionais podem surgir (STEPHANOPOULOS et al., 1998). Sendo assim, duas transformações de variáveis são aplicadas (WIECHERT; GRAAF, 1997), a primeira transforma os fluxos de cada reação nas duas direções, variáveis naturais ($\vec{v}$, $\overleftarrow{v}$), em dois outros fluxos; O fluxo *net* (v^{net}), a diferença entre os fluxos nas duas direções, e o fluxo *exchange* (v^{xch}) que é definido como o menor fluxo entre os dois. Essas variáveis são chamadas doravante de aplicadas. A transformação é unívoca e pode ser definida matematicamente como:

$$v^{net} = \vec{v} - \overleftarrow{v} \quad (4.1)$$

$$v^{xch} = \min(\vec{v}, \overleftarrow{v}) \quad (4.2)$$

A segunda transformação, também unívoca, mantém a variável v^{net} e redefine o fluxo *exchange* entre [0,1] através de uma transformação hiperbólica, sendo β uma constante da ordem de v^{net} . Essas variáveis são chamadas doravante de numéricas:

$$v^{xch[0,1]} = \frac{v^{xch}}{\beta + v^{xch}} \quad (4.3)$$

A variável $v^{xch[0,1]}$ é restrita a um intervalo de valores [0,1], a variável v^{net} tem que ser uma solução da equação estequiométrica ($Sv = 0$). Para satisfazer essas e outras restrições, equações de restrição de igualdade e desigualdade são formuladas. Abaixo essas equações são apresentadas, mais detalhes da construção delas encontram-se no Apêndice A. As restrições lineares de igualdade dos fluxos são dadas pelas seguintes equações:

$$N^{net} \cdot v^{net} = n^{net} \quad (4.4)$$

$$N^{xch[0,1]} \cdot v^{xch[0,1]} = n^{xch[0,1]} \quad (4.5)$$

A equação 4.4 restringe os fluxos *net*, onde N^{net} é a matriz de restrição de fluxo *net*, e n^{net} o vetor de restrição de fluxo *net*. A equação 4.5 restringe os fluxos *exchange* [0,1], onde $N^{xch[0,1]}$ é a matriz de restrição de fluxo *exchange* [0,1], e $n^{xch[0,1]}$ o vetor de restrição de fluxo *exchange* [0,1]. A equação 4.4 inclui a restrição de estado estacionário, representado pela equação estequiométrica. A equação 4.5 inclui as restrições de unidirecionalidade de uma reação ($v_i^{xch[0,1]} = 0$), e de equilíbrio rápido ($v_i^{xch[0,1]} = 1$). O espaço de fluxo é

restringindo com as equações acima, para fechar os graus de liberdade restantes são definidos os chamados fluxos livres (θ). Assim, teremos variáveis de fluxos dependentes, e outras variáveis de fluxos que serão escolhidas como independentes, os fluxos livres. Com isso, agora todos os fluxos podem ser representados por um conjunto menor de variáveis, diminuindo assim o espaço de fluxos e reduzindo o esforço computacional no processo de estimação. Os fluxos livres são definidos da seguinte forma:

$$N^{free} \cdot \begin{bmatrix} v^{net} \\ v^{xch[0,1]} \end{bmatrix} = \theta \quad (4.6)$$

Onde N^{free} é a matriz que determina quais são os fluxos livres e θ fornece o valor para cada um deles. Pode-se então concatenar essas três equações lineares e formar uma única equação de restrição linear 4.9, que é função somente dos fluxos livres:

$$N = \begin{bmatrix} N^{net} & 0 \\ 0 & N^{xch[0,1]} \\ & N^{free} \end{bmatrix} \quad (4.7)$$

$$n = \begin{bmatrix} n^{net} \\ n^{xch[0,1]} \\ \theta \end{bmatrix} \quad (4.8)$$

$$N \cdot \begin{bmatrix} v^{net} \\ v^{xch[0,1]} \end{bmatrix} = n \quad (4.9)$$

A escolha dos fluxos livres é restrita aos casos que a matriz N gerada seja quadrada e possua inversa. Por fim, para se restringir valores de fluxos sem sentido fisiológico, como fluxos de carbono negativos, definem-se as restrições de desigualdade lineares, aplicada também as variáveis numéricas.

$$U \cdot \begin{bmatrix} v^{net} \\ v^{xch[0,1]} \end{bmatrix} \geq u \quad (4.10)$$

Sendo U e u, a matriz e o vetor de restrição de desigualdade, respectivamente. Os fluxos *exchange* são sempre restritos no intervalo $0 \geq v^{xch[0,1]} \geq 1$. E a direção de uma reação unidirecional pode ser fixada com $v_i^{net} \geq 0$ ou $v_i^{net} \leq 0$. Para fluxos extracelulares essa restrição é obrigatória.

4.1.2 Balanço de cumômeros

Como descrito na seção 3.1.2, os balanços de isotopômeros são equações com termos bilineares, que surgem devido a reações bimoleculares. A forma geral dessas equações pode ser descrita como (WURZEL; GRAAF, 1999):

$$\bar{f}(v, \bar{x}, \bar{x}^{inp}) = \frac{1}{2} \bar{x}^T \cdot \left(\sum_i \vec{v} \cdot \vec{Q}_i + \overleftarrow{v} \cdot \overleftarrow{Q}_i \right) \cdot \bar{x} + \left(\sum_i \vec{v}_i \cdot \vec{P}_i + \sum_i \overleftarrow{v}_i \cdot \overleftarrow{P}_i \right) \cdot \bar{x} + \left(\sum_i \vec{v}_i \cdot \vec{P}_i^{inp} \right) \cdot \bar{x}^{inp} = 0 \quad (4.11)$$

Onde $\bar{x}$ e $\bar{x}^{inp}$ são as frações isotopoméricas nos metabólitos internos e no substrato, respectivamente. $\vec{P}_i$ e $\vec{P}_i^{inp}$ são as matrizes de transição de isotopômeros unimolecular para os fluxos internos e para o fluxo de entrada respectivamente. $\vec{Q}_i$ são as matrizes de transição de isotopômeros bimolecular para os fluxos internos. Detalhes sobre a construção das matrizes P e Q são fornecidos no Apêndice A.

A equação 4.11 é algébrica não linear, e métodos numéricos para esses casos são em geral aplicáveis, no entanto, a convergência pode ser muito lenta, e é preterido quando possível (BEERS, 2007). Como a equação é não linear em relação à variável v e quadrática em relação a $\bar{x}$, é possível através de uma transformação de variáveis encontrar uma solução analítica para $\bar{x}$ em função de v . Essa nova coordenada é chamada de cumômero, e foi apresentada na seção 3.1.2, ela se mostra bem menos interligada do que os isotopômeros, como consequência disso, podemos separar o vetor de cumômeros x em pesos, assim, o vetor ${}^n x$ contém todas as frações de cumômeros de peso n . Como demonstrado por Wurzel e Graaf (1999), pode-se construir uma cascata de equações lineares, uma para cada peso, da seguinte forma:

$$\begin{aligned} 1 &= {}^0 x \\ 0 &= {}^1 A(v) \cdot {}^1 x + {}^1 b(v) \\ 0 &= {}^2 A(v) \cdot {}^2 x + {}^2 b(v, {}^1 x) \\ 0 &= {}^3 A(v) \cdot {}^3 x + {}^3 b(v, {}^1 x, {}^2 x) \\ &\vdots \end{aligned} \quad (4.12)$$

Os termos ${}^n A(v)$ e ${}^n b(v)$ são construídos a partir das matrizes de transição de isotopômeros, detalhes sobre a construção dessas matrizes são fornecidos no Apêndice A. Esse sistema mostrou-se ser uma extensão não linear da teoria de sistemas compartimentados,

e dessa forma, as matrizes ${}^nA(v)$ possuem inversa sempre que o vetor v estiver na região viável (WIECHERT; WURZEL, 2001).

4.2 Estimação

A maneira como o problema de estimação é formulado influencia bastante no desempenho do método da estimação, e reparametrizações e modificações nas restrições são sempre bem-vindas. No problema de estimação de fluxos metabólicos, os parâmetros a serem estimados são os fluxos livres (θ), e há duas restrições lineares, uma de igualdade (4.9), e a outra de desigualdade (4.10). Além disso, temos a restrição da equação de balanço de isotopômeros (4.11), que é não linear em relação aos fluxos. Todas as três restrições podem então ser expressas em função dos fluxos livres. E ainda, as restrições 4.9 e 4.11 podem ser incorporadas à função objetivo.

Além disso, o vetor de fluxos medidos e o vetor de marcações medidas foram agrupados em um único vetor $w_m = [x_m, v_m]$, assim como os calculados $w_c = [x_c, v_c]$. Pesos foram adicionados nessas variáveis, uma vez que as medidas de fluxo e marcação podem ter dimensões bem distintas e algumas medidas podem ser menos confiáveis que outras. Sabe-se que os pesos que levam a um estimador com variância mínima para um modelo linear e com erros normalmente distribuídos são formados pelo inverso da matriz de covariância dos erros (BARD, 1974). Embora essa propriedade não possa ser provada fora dessas especificações, é razoável esperar resultados aproximadamente ótimos em problemas não lineares (BARD, 1974). O estimador de mínimos quadrados com pesos nas variáveis, com a matriz de covariância dos erros V é formulado da seguinte forma:

$$\min_{\theta} \Phi = (w_c(\theta) - w_m) \cdot V^{-1} \cdot (w_c(\theta) - w_m)^T \quad (4.13)$$

$$\text{Sujeito a: } U' \cdot \theta \geq u'$$

A restrição do problema é a equação 4.10 formulada em termos dos fluxos livres. Sendo assim, o espaço de busca do problema é convexo, o que facilita a utilização de alguns métodos de otimização. Além disso, como a não linearidade do modelo foi introduzida na função objetivo, há um ganho de eficiência quando se utilizam métodos que aproximam a função objetivo por uma função quadrática, como o método SQP.

4.3 Quantificação da incerteza

Na seção 4.2 foi apresentado o método de estimativa pontual dos fluxos metabólicos. Embora bastante informativo, a estimativa pontual não traz toda a informação sobre o sistema em estudo. Para uma análise mais detalhada da fluxômica da célula é preciso entender o quão confiáveis e precisas são essas estimativas, ou seja, qual o grau de incerteza dos parâmetros estimados. Além disso, é necessário também se compreender o quão confiável é o modelo, o que pode ser verificado através do ajuste que ele obtiver dos dados experimentais.

4.3.1 Sensibilidade analítica

Uma das ferramentas para quantificar as incertezas no seu modelo é o cálculo da sua sensibilidade, que indica como a resposta do modelo se modifica se os parâmetros forem alterados. No caso do ^{13}C -MFA, o quanto as frações de marcação de cumômero variam quando os fluxos metabólicos são alterados. O cálculo da sensibilidade pode ser feito numericamente por diferenças finitas ou por uma equação analítica, o que deixa o processo mais eficiente e preciso. Para se obter a equação analítica da sensibilidade pode-se diferenciar implicitamente a equação 4.11. Embora computacionalmente fácil de ser implementado, o cálculo é custoso computacionalmente porque é necessário se inverter uma matriz da dimensão de x . Uma maneira mais eficiente é a diferenciação de toda a cascata de cumômeros (equação 4.12). Diferenciando-se a equação de peso n , obtém-se:

$$0 = \frac{\partial^n A(v)}{\partial v_i} \cdot x + {}^n A(v) \cdot \frac{\partial({}^n x)}{\partial v_i} + \frac{\partial({}^n b)}{\partial v} \cdot \frac{\partial v}{\partial v_i} + \sum_{i=1}^{n-1} \frac{\partial({}^n b)}{\partial({}^i x)} \cdot \frac{\partial({}^i x)}{\partial v_i} \quad (4.14)$$

O único termo desconhecido da equação 4.14 é a própria sensibilidade dos cumômeros de peso n em relação ao fluxo i ($\frac{\partial({}^n x)}{\partial v_i}$). Como a fatoração da matriz ${}^n A(v)$ já foi realizada no cálculo dos cumômeros de peso n , ela pode ser aproveitada aqui. A sensibilidade do modelo pode ser aplicada no design de experimentos, para se verificar se a marcação de um determinado metabólito varia quando se modifica a distribuição de fluxos na célula, quanto maior essa resposta melhor é a sua estimativa. A sensibilidade também é utilizada para o cálculo do gradiente do processo de estimação dos fluxos. Em um algoritmo de otimização numérica, uma das etapas de maior custo computacional é o cálculo da estimativa do

gradiente por diferenças finitas, já que a função deve ser acessada diversas vezes. Para se calcular como a função objetivo (Φ) varia com os fluxos livres tem-se que realizar uma derivação em cadeia, uma vez que o modelo realiza transformações de variáveis:

$$\frac{\partial \Phi}{\partial \theta} = \frac{\partial \Phi}{\partial^n x} \cdot \frac{\partial^n x}{\partial v_{nat}} \cdot \frac{\partial v_{nat}}{\partial v_{app}} \cdot \frac{\partial v_{app}}{\partial v_{num}} \cdot \frac{\partial v_{num}}{\partial \theta} \quad (4.15)$$

Onde $\frac{\partial^n x}{\partial v_{nat}}$ é a sensibilidade do modelo calculada em 4.14. Os outros termos podem ser calculados a partir das equações 4.13, 4.1, 4.2 e 4.3, respectivamente.

4.3.2 Ajuste

Uma vez realizada a estimação dos fluxos é possível se questionar se o modelo foi bem ajustado aos dados experimentais. Se o resíduo final puder ser explicado por erros aleatórios nas medições, então podemos dizer que o modelo se ajustou bem aos dados. Um teste bastante utilizado é o teste qui-quadrado, onde um teste de hipótese é realizado para se verificar a consistência estatística do ajuste (CROWN; ANTONIEWICZ, 2013). O resíduo calculado na equação 4.13 é uma variável aleatória que possui uma distribuição aproximadamente qui-quadrado com graus de liberdade (p) igual ao número de medidas independentes menos o número de parâmetros estimados. O teste de hipótese pode então ser formulado, onde a hipótese nula é de que o resíduo é igual a zero. A hipótese nula é rejeitada se:

$$\int_0^{\Phi^*} \chi_p^2(\Phi) d\Phi \geq 1 - \alpha \quad (4.16)$$

Onde $1 - \alpha$ é o grau de confiabilidade e Φ^* é o resíduo minimizado. É importante ressaltar que esse teste só deve ser aplicado caso se conheça a variância das medidas realizadas, uma vez isso influencia no limite superior da integral.

4.3.3 Intervalo de confiança marginal

O cálculo dos intervalos de confiança marginal para os fluxos metabólicos é crucial para interpretação do estado fisiológico da célula, uma vez que fornece uma forma prática de analisar as incertezas de cada estimativa. Neste trabalho foi implementado o método conhecido como *linearized statistics*, que se baseia nas hipóteses de que o estimador tem

uma distribuição aproximadamente normal e o modelo tem um comportamento aproximadamente linear próximo ao mínimo (WIECHERT; GRAAF, 1997). Nesse método, calcula-se um intervalo para o parâmetro estimado em que se espera que o valor real esteja contido com uma probabilidade de $1-\alpha$. O intervalo de confiança para um parâmetro θ_i definido pelos limites superior (L_{sup}) e inferior (L_{inf}) pode ser definido como:

$$P[L_{inf} \leq \theta_i \leq L_{sup}] = 1 - \alpha \quad (4.17)$$

Os limites superior e inferior podem ser calculados a partir do parâmetro estimado ($\hat{\theta}_i$) como se segue:

$$L_{sup} = \hat{\theta}_i + z_{1-\frac{\alpha}{2}} \cdot \sqrt{V_{\theta_i}} \quad (4.18)$$

$$L_{inf} = \hat{\theta}_i - z_{1-\frac{\alpha}{2}} \cdot \sqrt{V_{\theta_i}} \quad (4.19)$$

Onde $z_{1-\frac{\alpha}{2}}$ é o quantil $1 - \alpha$ da distribuição normal padrão e V_{θ_i} é variância do parâmetro θ_i . Considerando, novamente, que a distribuição dos parâmetros é aproximadamente normal, com a variância das medições conhecida e igual a V , a matriz de covariância dos parâmetros pode ser aproximada como (BARD, 1974):

$$V_{\theta} \approx \frac{\partial w_c}{\partial \theta} \cdot V^{-1} \cdot \frac{\partial w_c}{\partial \theta} \quad (4.20)$$

$\frac{\partial w_c}{\partial \theta}$ representa o quanto as variáveis simuladas variam com os parâmetros e pode ser calculada a partir da sensibilidade do modelo (equação 4.14).

4.4 Ferramenta computacional desenvolvida

O modelo matemático e o estimador apresentados nas seções anteriores foram implementados na plataforma *MATLAB*® v8.1. O modelo é estruturado e pode realizar cálculos de qualquer rede metabólica, contanto que se forneça os dados necessários sobre ela. Foi utilizada a função "sparse" para representar as matrizes de transição, a fim de se diminuir a memória utilizada para armazená-las. Na estimação, utilizou-se o método SQP.

A Figura 9 apresenta um fluxograma representativo do funcionamento da ferramenta desenvolvida. O programa pode funcionar em três modos de operação:

- No modo simulação, é preciso fornecer o valor dos fluxos livres (θ), na coordenada de fluxo numérico. Além disso, é preciso fornecer a rede metabólica e as transições dos

átomos nas reações. Como resultado o programa fornece as frações isotopoméricas para todos os metabólitos internos.

- No modo estimação, é preciso fornecer os valores iniciais para os fluxos livres (θ_0) e os dados medidos, de fluxo e/ou marcação, bem como os respectivos desvios padrão. Além disso, é preciso fornecer a rede metabólica e as transições dos átomos nas reações. Como resultado o programa fornece os fluxos estimados, o resultado do teste de ajuste e os intervalos de confiança marginais.
- No modo design, é preciso fornecer o valor dos fluxos livres (θ), na coordenada de fluxo numérico. Além disso, é preciso fornecer a rede metabólica e as transições dos átomos nas reações. Como resultado o programa fornece a marcação do substrato que permitirá estimar os fluxos livres com maior precisão.

No fornecimento da rede metabólica pelo usuário, só são aceitas reações com até dois reagentes e/ou produtos no máximo, reações de ordem maior devem ser decompostas.

Figura 9 – Fluxograma representativo da ferramenta desenvolvida.

5 Resultados e discussão

5.1 Validação do Programa

5.1.1 Simulação do ciclo TCA

Para validar o modelo que simula as frações isotopoméricas, foram comparados os resultados da simulação do Ciclo dos Ácidos Tricarboxílicos (Ciclo TCA) com os presentes em [Antoniewicz et al. \(2007\)](#). A rede metabólica para o ciclo TCA, Figura 10, é composta por 9 metabólitos (5 internos e 4 externos) e 8 fluxos, tendo somente uma reação reversível. As estequiometrias das reações, bem como as transições atômicas se encontram no Apêndice B na Tabela 15.

Figura 10 – Modelo estrutural simplificado do ciclo TCA. Abreviações dos metabólitos: OAC, oxaloacetato; Asp, aspartato; AcCoA, acetilcoenzima A; Cit, citrato; AKG, -cetoglutarato; Glu, glutamato; Fum, fumarato; Suc, succinato. Os fluxos têm unidades arbitrárias.

Fonte: ([ANTONIEWICZ et al., 2007](#))

Foi considerada marcação de entrada somente no metabólito AcCoA, com uma mistura de 25% $[2-^{13}C]$, 25% $[1,2-^{13}C]$ e 50% sem marcação. A marcação natural de carbono não foi considerada. Foram avaliados os isotopômeros de massa do glutamato em fração molar, $M+N$, onde N representa a quantidade de átomos marcados na molécula. Essas entidades são exatamente iguais ao somatório dos isotopômeros de peso N do Glutamato. Como exemplo, o termo $M+4$ é exibido na equação 5.1.

$$M + 4 = \#01111 + \#10111 + \#11011 + \#11101 + \#11110 \quad (5.1)$$

Os resultados encontrados pelo programa desenvolvido neste trabalho e os encontrados por [Antoniewicz et al. \(2007\)](#) se encontram na Tabela 2. [Antoniewicz et al. \(2007\)](#) fizeram simulações com dois tipos de modelos, um em que as equações de balanços de isotopômeros são resolvidas numericamente, e outro com o modelo EMU. Os três modelos apresentaram o mesmo resultado para a marcação do glutamato. Esse padrão de marcação do glutamato é totalmente compatível com as características da rede metabólica, uma vez que os átomos do AcCoA vão para os dois últimos átomos da molécula de glutamato. Sendo assim, é de se esperar uma fração razoável de isotopômeros de massa M+1 e M+2. Além disso, observa-se que embora a porcentagem de moléculas não marcadas na entrada seja de 50%, os isotopômeros M+0 tem uma porcentagem menor. Isso porque o fluxo v4 leva uma parte da marcação do AcCoA para o restante do ciclo, e permite que isotopômeros de ordem maior apareçam.

Tabela 2 – Comparação da simulação dos isotopômeros de massa do glutamato no ciclo TCA.

Isotopômeros de massa do glutamato simulados	Isotopômero (ANTONIEWICZ et al., 2007)	EMU (ANTONIEWICZ et al., 2007)	Próprio autor
M+0	0.3464	0.3464	0.3464
M+1	0.2695	0.2695	0.2695
M+2	0.2708	0.2708	0.2708
M+3	0.0807	0.0807	0.0807
M+4	0.0286	0.0286	0.0286
M+5	0.0039	0.0039	0.0039

5.1.2 Estimação de fluxos em *E.coli*

Para avaliar o desempenho do programa para a estimação de fluxos, foi avaliada uma rede metabólica simplificada de *E.coli*, e os resultados comparados com os presentes em [Srouf et al. \(2011\)](#). Uma ilustração da rede utilizada se encontra na Figura 11. O sistema é composto pela via Embden-Meyerhof-Parnas (EMP) e pela Via das Pentoses, contendo assim 18 fluxos (incluindo os bidirecionais) e 11 metabólitos internos. Dessa forma tem-se 7 graus de liberdade, fixando o fluxo de entrada upt em 1.02 (unidade arbitrária), passa-se a ter 6 graus de liberdade. Como mencionado na seção 4.1.1, o espaço de fluxos é parametrizado em função dos fluxos livres (θ). Os fluxos escolhidos como livres, na coordenada numérica, foram:

- $\theta(1) = \text{ppp1}_{net}$
- $\theta(2) = \text{ppp2}_{xch[0,1]}$
- $\theta(3) = \text{ppp3}_{xch[0,1]}$
- $\theta(4) = \text{ppp4}_{xch[0,1]}$
- $\theta(5) = \text{ppp5}_{xch[0,1]}$
- $\theta(6) = \text{ppp6}_{xch[0,1]}$

A escolha desses fluxos foi feita de forma que a equação 4.9 tenha solução única, o que explica a escolha dos fluxos bidirecionais, uma vez que os fluxos nos dois sentidos são linearmente dependentes. A estequiometria das reações e as transições entre os átomos nas reações estão no Apêndice B na Tabela 16. As reações inversas da via das pentoses possuem tanto a estequiometria como as transições no sentido contrário ao apresentado na tabela. Na Tabela 3 estão presentes a marcação do substrato para cada isotopômero, e também as medidas de cumômeros de alguns intermediários, bem como os respectivos desvios padrão.

Uma análise simples de identificabilidade *a priori* pode ser feita na rede, como em Winden et al. (2001a), onde se mostrou que para cada metabólito que tem sua marcação medida, o número de restrições independentes adicionadas ao problema é igual ao menor valor entre o número de balanços de isotopômeros, e o número de fluxos que geram o metabólito. Ou seja, $\min\{m-1, 2^n-1\}$, onde m é a quantidade de fluxos de entrada no *pool* do metabólito, e n o número de átomos. Neste problema, tem-se medidas de quatro metabólitos: PEP, Ru5P, Ery4P e GA3P. Dessa forma, a medida de PEP não adiciona nenhuma restrição ao problema, pois só tem um fluxo de entrada. E as medidas de Ru5P, Ery4P e GA3P adicionam 2, 1 e 3 restrições independentes cada, respectivamente. Somando então 6 restrições independentes para 6 graus de liberdade, ou seja, *a priori* os fluxos são identificáveis. Obviamente essa é uma condição necessária, mas não suficiente para a identificabilidade real dos fluxos.

Figura 11 – Modelo estrutural simplificado da *E.coli*. Abreviações dos metabólitos: GLC, glicose; Glc6P, glicose-6-fosfato; Fru6P, frutose-6-fosfato; FruBP, frutose-1,6-bisfosfato; GA3P, gliceraldeído-3-fosfato; PGA, fosfoglicerato; PEP, fosfoenolpiruvato; PYR, piruvato; Ru15P, ribulose-5-fosfato; Rib5P, ribose-5-fosfato; Xul5P, xilose-5-fosfato; Sed7P, sedoheptulose-7-fosfato; Ery4P, eritrose-4-fosfato.

Fonte: (SROUR et al., 2011)

Como a convexidade do problema não pôde ser demonstrada, diferentes valores iniciais (θ_0) foram fornecidos ao programa, a fim de que se possa verificar a presença de múltiplos mínimos. É aconselhável que esses valores se encontrem dentro da região viável do problema. Embora o método SQP possa obter a solução mesmo para valores iniciais fora dessa região, todos os seus passos subsequentes obedecem as restrições. Como o primeiro parâmetro ($\theta_0(1)$) é um fluxo net de uma reação unidirecional, seu valor deve ser maior que zero. Além disso, o valor máximo dele é limitado ao quanto entra de substrato, que nesse trabalho foi fixado em 1.02. Já os outros parâmetros, são $\text{exchange}[0,1]$ e devem estar no intervalo entre 0 e 1. No entanto, percebeu-se que inicializações para esses parâmetros iguais a 1, causavam descontinuidades nas matrizes de transição, tornando inviável o cálculo. Dessa forma o limite máximo para esses parâmetros foi fixado em 0.99.

Tabela 3 – Dados de marcação de entrada (GLC) e dos intermediários. Dados provenientes de simulações realizadas por (SROUR et al., 2011).

Fluxo	índice	Valor	Desvio padrão
GLC	#000000	0.445	-
	#100000	0.500	-
	#010000	0.011	-
	#001000	0.011	-
	#000100	0.011	-
	#000010	0.011	-
	#000001	0.011	-
Ru5P	#1xxxx	0.1979	0.002
	#x1xxx	0.0153	0.002
	#xx1xx	0.0284	0.002
	#xxx1x	0.0122	0.002
	#xxxx1	0.0976	0.002
Ery4P	#1xxx	0.0568	0.002
	#x1xx	0.0229	0.002
	#xx1x	0.0118	0.002
	#xxx1	0.0704	0.002
GA3P	#1xx	0.0330	0.002
	#x1x	0.0126	0.002
	#xx1	0.1207	0.002
PEP	#1xx	0.0330	0.002
	#x1x	0.0126	0.002
	#xx1	0.1207	0.002

Fonte: (SROUR et al., 2011)

As inicializações fornecidas ao programa estão na Tabela 4. Os dois primeiros valores iniciais partem de pontos presentes em Srou et al. (2011). Os outros quatro pontos são próximos das fronteiras da região viável do problema. Como todos os cálculos convergiram para o mesmo ponto, assumiu-se que ele é o mínimo global. O valor dos parâmetros no ponto ótimo se encontra na última linha da tabela 4. Percebe-se que esse ponto encontrado é bem próximo dos dois primeiros valores iniciais, e distante das fronteiras do espaço viável, portanto nenhuma restrição está ativa.

Para uma melhor visualização da solução encontrada, calcularam-se os fluxos nas coordenadas naturais, o que facilita a interpretação física do problema. Esses fluxos se encontram na quarta coluna da Tabela 5. Analisando-se a rede metabólica da Figura 11, nota-se que alguns fluxos são cruciais para se entender a distribuição de marcação pelos metabólitos internos, nomeadamente ppp1 e emp1. O fluxo de entrada, upt, é dividido nesses dois fluxos, e a marcação dos metabólitos GA3P e PEP varia bastante com a razão entre eles. Quanto maior o fluxo ppp1, menos marcação se tem no sistema, uma vez que o

átomo com maior marcação na entrada, o primeiro, sai na forma de CO_2 . Já quanto maior o fluxo emp1, maior será a marcação no sistema, principalmente nos terceiros átomos dos carbonos de GA3P e PEP. Como a Tabela 3 indica uma porcentagem em torno de 12% no átomo três desses metabólitos, é natural esperar uma razão próxima de 1:1 entres os fluxos ppp1 e emp1, como de fato aconteceu. Outra análise qualitativa que se pode fazer quanto à plausibilidade do resultado obtido, se dá na marcação no primeiro átomo da espécie Ru5P. Que só pode ter esse padrão, caso haja fluxos inversos na via das pentoses, pois como já mencionado, o fluxo ppp1 remove o átomo com maior marcação do substrato.

Tabela 4 – Diferentes valores iniciais para o processo de estimação

	$\theta_0(1)$	$\theta_0(2)$	$\theta_0(3)$	$\theta_0(4)$	$\theta_0(5)$	$\theta_0(6)$
1	0.5100	0.8033	0.8113	0.8030	0.2024	0.2032
2	0.5101	0.7995	0.7197	0.8067	0.2114	0.2038
3	0.1	0	0	0	0	0
4	1.02	0	0	0	0	0
5	0	0.99	0.99	0.99	0.99	0.99
6	1.02	0.99	0.99	0.99	0.99	0.99
θ^*	0.5103	0.8015	0.7093	0.8041	0.2131	0.2025

Fonte: Próprio autor.

Por fim, a Tabela 5 também fornece os fluxos calculados por outros dois programas, presentes em (SROUR et al., 2011). Mais detalhes sobre esses programas são fornecidos na seção 3.3. As três distribuições de fluxos são distintas, e a quarta coluna apresenta a distribuição que gera a menor função objetivo (Φ) na ferramenta aqui desenvolvida. Apesar disso, os valores para os fluxos unidirecionais são similares, e a diferença se dá nos fluxos bidirecionais, que geralmente possuem problemas de identificabilidade.

A sensibilidade analítica (equação 4.15) e o gradiente (equação 4.15) foram validados com a rede metabólica da *E.coli* utilizando derivadas numéricas. Uma comparação na estimação dos fluxos da *E.coli* é feita na Tabela 6, onde a vantagem da utilização do gradiente analítico no desempenho da otimização é demonstrado. Como pode ser observado, o número de iterações nos dois cálculos é similar, no entanto o número de acessos a função cai significativamente, e o tempo de cálculo cai para um terço com a utilização do gradiente analítico.

Tabela 5 – Comparação entre os fluxos estimados para *E.coli*, e a função objetivo (Φ) nos três pontos. O índice r se refere às reações no sentido inverso.

Fluxo	FIA	13CFLUX	Próprio autor
upt	1.0200	1.0200	1.0200
emp1	0.5100	0.5099	0.5097
emp2	0.8500	0.8500	0.8499
emp3	0.8500	0.8500	0.8495
emp4	1.8700	1.8700	1.8699
emp5	1.8700	1.8700	1.8699
emp6	1.8700	1.8700	1.8699
ppp1	0.5100	0.5001	0.5003
ppp2	4.4234	4.3281	4.3770
ppp2r	4.0834	3.9880	4.0368
ppp3	4.4689	2.7370	2.6100
ppp3r	4.2989	2.5670	2.4399
ppp4	4.2468	4.3440	4.2759
ppp4r	4.0768	4.1740	4.1059
ppp5	0.4238	0.4381	0.4408
ppp5r	0.2538	0.2680	0.2708
ppp6	0.4250	0.4260	0.4240
ppp6r	0.2550	0.2560	0.2539
Φ	7.95×10^{-3}	6.98×10^{-3}	2.26×10^{-3}

Fonte: (SROUR et al., 2011) /Próprio autor.

Tabela 6 – Comparação do processo de estimação com e sem o gradiente analítico

	iterações	número de acessos a função	tempo (min)
Gradiente numérico	47	340	192
Gradiente analítico	50	53	59

Fonte: Próprio autor.

O teste de ajuste da seção 4.3.2 pode ser aplicado no caso da *E.coli*. Tem-se 15 medidas independentes e 6 parâmetros independentes, dessa forma, assume-se que o resíduo tem uma distribuição aproximadamente qui-quadrado com 9 graus de liberdade. Realizando o teste com um grau de confiabilidade de 95%, 9 graus de liberdade e resíduo igual a 0.0026, obtemos:

$$\int_0^{0.0026} \chi_9^2(\Phi) d\Phi = 1.96 \times 10^{-15} \leq 0.95$$

Como a integral da nossa função cumulativa foi menor que 0.95, não devemos rejeitar a hipótese nula. Como qualquer teste de ajuste para um modelo, ele se limita a descartar modelos que tenham se ajustado mal aos dados e não em validá-los. Caso o teste tivesse

rejeitado a hipótese nula, teríamos um grande indício de que a nossa rede metabólica para o sistema está incorreta, ou os dados experimentais contêm erros grosseiros.

Por fim, pode-se determinar o intervalo de confiança para os fluxos estimados como demonstrado na seção 4.3.3. O resultado do cálculo do intervalo de confiança com 95% de confiabilidade para os fluxos livres se encontra na Tabela 7. Como pode ser observado, os parâmetros $\theta(1)$ e $\theta(6)$ parecem estar bem definidos, enquanto os outros tem um intervalo muito grande, uma vez que, excluindo-se o primeiro, os parâmetros variam no intervalo de $[0,1]$. Se voltarmos para a tabela 4, veremos que os parâmetros com grande intervalo são exatamente os parâmetros cujos valores encontrados nesse trabalho mais se diferenciaram dos obtidos com os outros dois programas (pontos 1 e 2). Provavelmente há um problema de identificabilidade prática desses parâmetros, sendo impossível assim determinar os seus valores.

Tabela 7 – Valores estimados para os fluxos livres e seus respectivos intervalos de confiança.

	Valores estimados	Intervalo de confiança 95%
$\theta(1)$	0.510	[0.505,0.515]
$\theta(2)$	0.802	[0.693,0.910]
$\theta(3)$	0.709	[0.000,1.000]
$\theta(4)$	0.804	[0.698,0.910]
$\theta(5)$	0.213	[0.000,0.794]
$\theta(6)$	0.203	[0.177,0.228]

Fonte: Próprio autor.

A Tabela 8 apresenta os intervalos de confiança para todos os fluxos na coordenada aplicada. Embora a coordenada natural seja a mais fácil de se interpretar fisicamente, ela não daria os intervalos para os fluxos bidirecionais, uma vez que estes são funções dos fluxos *exchange*, que como visto não podem ser determinados. Como nota-se na Tabela 8, todos os fluxos do tipo *net* tem intervalos bem definidos, enquanto os do tipo *exchange* são indeterminados, excetuando o fluxo ppp6. Este caso demonstra claramente a vantagem da utilização das transformadas de variáveis aplicadas na seção 4.1.1, pois consegue-se estimar com eficiência o fluxo líquido nas reações bidirecionais, apesar da indeterminação na ordem de grandeza.

Tabela 8 – Valores estimados para os fluxos na coordenada numérica e seus respectivos intervalos de confiança. net corresponde aos fluxos do tipo *net* e xch aos fluxos do tipo *exchange*

Fluxo	Valores estimados	Intervalo de confiança 95%
upt(net)	1.020	[1.020,1.020]
emp1(net)	0.510	[0.505,0.515]
emp2(net)	0.850	[0.847,0.853]
emp3(net)	0.850	[0.847,0.853]
emp4(net)	1.870	[1.867,1.873]
emp5(net)	1.870	[1.867,1.873]
emp6(net)	1.870	[1.867,1.873]
ppp1(net)	0.510	[0.505,0.515]
ppp2(net)	0.340	[0.336,0.344]
ppp2(xch)	4.038	[2.255,10.133]
ppp3(net)	0.170	[0.167,0.173]
ppp3(xch)	2.440	[0.000,∞]
ppp4(net)	0.170	[0.167,0.173]
ppp4(xch)	4.105	[2.313,10.111]
ppp5(net)	0.170	[0.167,0.173]
ppp5(xch)	0.271	[0.000,3.851]
ppp6(net)	0.170	[0.167,0.173]
ppp6(xch)	0.254	[0.215,0.295]

Fonte: Próprio autor.

5.2 Estudo de Caso: Estimação de fluxos em *Pseudomonas sp.* LFM046

Um estudo de caso que visa exemplificar a aplicabilidade da ferramenta desenvolvida foi realizado em *Pseudomonas sp.* LFM046. *Pseudomonas sp.* LFM046 é uma bactéria que produz Polihidroxialcanoatos (PHAs), que são poliésteres acumulados sob a forma de grânulos intracelulares, podendo representar até 80% de sua massa seca celular (KAWAI, 2013). PHA é um material de reserva de carbono, energia ou equivalentes redutores, dessa forma, ele é principalmente produzido em condições de excesso de carbono e limitação de um nutriente essencial para o crescimento, como nitrogênio ou fósforo. Esses polímeros despertam grande interesse comercial devido às suas propriedades termoplásticas, e seu potencial para substituir plásticos de origem petroquímica (KAWAI, 2013). Riascos et al. (2013) estudaram as principais vias na produção de PHA em *Pseudomonas sp.* LFM046 por ¹³C-MFA, no entanto, após o genoma ter sido sequenciado (CARDINALI-REZENDE et al., 2015), concluiu-se que a via das pentoses não possuía a parte oxidativa, e portanto o modelo

metabólico estava incorreto. Esse é um problema comum em ^{13}C -MFA (WINDEN et al., 2001b). Dessa forma, esse estudo de caso tem como objetivo encontrar uma distribuição de fluxos que descreva bem os dados de carbono marcado obtidos, e possa indicar quais as principais vias na produção de PHA, utilizando uma rede metabólica derivada do sequenciamento dos genes.

5.2.1 Construção do modelo metabólico

O modelo metabólico aqui utilizado é baseado no sequenciamento do genoma da *Pseudomonas sp.* LFM046 feito por Cardinali-Rezende et al. (2015). O metabolismo central da *Pseudomonas sp.* LFM046 é apresentado na Figura 12, a rede possui duas entradas de substrato, glicose e frutose e uma saída de PHA; A rede inclui a Via das Pentoses, a via ED, que pode atuar como cíclica, e o Ciclo de Krebs. Ao todo a rede contém 35 reações, incluindo as bidirecionais, e 19 metabólitos internos, sendo assim, o espaço de fluxo desse problema tem dimensão 16.

5.2.2 Dados experimentais

Os dados experimentais utilizados para a estimação de fluxos neste trabalho são provenientes de Kawai (2013), que utilizou *Pseudomonas sp.* LFM046 isolada de amostras de solo de canavial da Usina São José, localizada no estado de Pernambuco. Todos os ensaios foram feitos em frascos agitados com fonte de glicose marcada na proporção de 20% marcada em todos os átomos [$\text{U-}^{13}\text{C}$] e 80% sem marcação. Foram feitos experimentos com limitação de nitrogênio no meio, e experimentos com limitação de fósforo no meio. Os monômeros de PHA acumulado nos dois tipos de experimentos foram analisados por cromatografia gasosa acoplada à espectroscopia de massas (GC-MS) com o objetivo de medir a marcação dos isotopômeros de massa do polímero.

De acordo com Kawai (2013), dois tipos principais de PHAs são produzidos por bactérias a partir de carboidratos: poli-3-hidroxibutirato (P3HB) produzido por diversas bactérias, e PHA com monômeros de cadeia média PHA_{MCL} produzido principalmente por *Pseudomonas sp.*. Quando carboidratos são fornecidos como fonte de carbono, P3HB e PHA_{MCL} são sintetizados a partir de AcCoA. P3HB é produzido em três etapas: primeiro, duas moléculas de AcCoA são condensadas, formando o acetoacetil-CoA, que é reduzido

a 3-hidroxibutiril-CoA, e finalmente polimerizado pela PHA sintase a P3HB. Os monômeros de cadeia média 3-hidroxiálcanoatos (3HAs) são gerados a partir de intermediários da via de biossíntese de ácido graxo *de novo* e consiste principalmente de 3-Hidroxihexanoato (3HHx), 3-Hidroxiocetanoato (3HO), 3-Hidroxiodecanoato (3HD), 3-Hidroxiodecanoato (3HDd) e 3-Hidroxiobutirato (3HB). A via de biossíntese de ácido graxo *de novo* é uma via cíclica, onde ácidos graxos são aumentados em dois carbonos (AcCoA) a cada ciclo. Monômeros 3HAs intermediários da via de biossíntese de ácidos graxos são levados à biossíntese de PHA_{MCL} .

Figura 12 – Metabolismo central da *Pseudomonas sp.* LFM046. Abreviações dos metabólitos: G6P, glicose-6-fosfato; F6P, frutose-6-fosfato; F16P, frutose-1,6-bisfosfato; 6PG, 6-fosfogluconato; G3P, gliceraldeído-3-fosfato; PEP, fosfoenolpiruvato; PYR, piruvato; RIB5P, ribose-5-fosfato; Xul5P, xilose-5-fosfato; S7P, sedoheptulose-7-fosfato; E4P, eritrose-4-fosfato; OXA, oxaloacetato; AcCoA, acetilcoenzima A; AKG, α -cetoglutarato; Fum, fumarato; Suc, succinato; , succinil-Coenzima A; GLX, glioxilato. Os fluxos têm unidades arbitrárias.

Fonte: Próprio autor.

Para se analisar a marcação no PHA, foram analisados por GC-MS os propil-ésteres dos monômeros 3HAs obtidos por propanólise. Na Figura 13 são apresentados os três

fragmentos típicos dos propil-ésteres de 3HA. O fragmento $m/z=[M-59]$ é específico de cada monômero 3HA, e contém carbonos provenientes de todas as moléculas de AcCoA usadas na biossíntese de 3HA, e nenhum do propanol utilizado na propanólise. O fragmento $m/z=131$ é proveniente da clivagem α do grupo funcional hidroxil e inclui três carbonos derivados da biossíntese e três carbonos derivados do propanol. O fragmento $m/z=89$ é gerado pelo rearranjo dos três primeiros carbonos do 3HA.

Figura 13 – Fragmentos característicos de propil-ésteres de 3HA obtidos por espectroscopia de massas

Fonte: (KAWAI, 2013)

Os isotopômeros de massa foram calculados a partir dos propil-ésteres dos monômeros 3HO e 3HD que estavam em maior quantidade. E analisaram-se os picos do fragmento $m/z=89$ que é comum a todos os monômeros e contém somente carbonos provenientes de 3HA. Esses carbonos vêm de uma molécula e meia de AcCoA. Um espectro ilustrativo dessas medidas é apresentado na Figura 14 e os isotopômeros de massa medidos e seus respectivos desvios padrão para os casos limitados em nitrogênio e fósforo estão apresentados na Tabela 9.

Tabela 9 – Medidas de Isotopômeros de massa e seus respectivos desvios padrão.

	N limitado	P limitado
M+0	0.58 ± 0.02	0.60 ± 0.02
M+1	0.22 ± 0.01	0.22 ± 0.01
M+2	0.15 ± 0.01	0.14 ± 0.01
M+3	0.04 ± 0.01	0.04 ± 0.01

Fonte: (KAWAI, 2013)

Figura 14 – Espectro de massa dos propil-ésteres dos monômeros 3HO e 3HD.

Fonte: (KAWAI, 2013)

5.2.3 Estimação de fluxos

A rede metabólica apresentada na seção 5.2.1 possui dimensão do espaço de fluxo igual a 16. Como as únicas medidas de marcação disponíveis fora da condição de crescimento são as do PHA, hipóteses simplificadoras precisam ser feitas. Sendo assim, considerou-se que não há fluxo no Ciclo de Krebs pois como há limitação de nutrientes nos experimentos, não há crescimento celular, e todo o fluxo de carbono vai para a produção do polímero. Considerou-se ainda a glicose como única fonte de carbono, e reações sequenciais foram agrupadas. Todos os casos considerados utilizam uma unidade arbitrária de fluxo, com a entrada de glicose fixada em 1.00.

Três vias metabólicas podem produzir PHA a partir de glicose, a via ED, a via ED cíclica e a Via das Pentoses. Considerando que todo o fluxo de carbono vai pela via ED, tem-se a situação ilustrada na Figura 15. Uma simulação foi feita com a rede e distribuição de fluxos presentes na Figura 15 e marcação de entrada idêntica aos utilizados por Kawai (2013) em seus experimentos. Os destinos dos átomos nas reações são apresentados no Apêndice B, Tabela 17. Os Isotopômeros de massa calculados se encontram na terceira coluna da Tabela 10. Como pode-se observar, o padrão de isotopômeros de massa gerados difere do experimental e portanto a via ED não deve ser a única via de síntese de PHA.

Figura 15 – Rede metabólica simplificada da *Pseudomonas sp.* LFM046 considerando todo o fluxo de carbono pela via ED. Abreviações dos metabólitos: G6P, glicose-6-fosfato; F6P, frutose-6-fosfato; F16P, frutose-1,6-bisfosfato; G3P, gliceraldeído-3-fosfato; PYR, piruvato; AcCoA, acetilcoenzima A. Os fluxos têm unidades arbitrárias.

Fonte: Próprio autor.

Tabela 10 – Comparação entre os Isotopômeros de massa simulados e os medidos

	N limitado	P limitado	simulado: ED
M+0	0.58	0.60	0.64
M+1	0.22	0.22	0.16
M+2	0.15	0.14	0.16
M+3	0.04	0.04	0.04

Fonte: (KAWAI, 2013)

Se considerarmos agora a via ED operando de forma cíclica, podemos ter uma distribuição de fluxos igual a ilustrada na Figura 16, nela temos um fluxo de reciclo igual a 0.5, esse fluxo pode ter qualquer valor entre 0 e 1, ou seja, entre nenhum reciclo e um reciclo total. Foram feitas simulações variando-se o fluxo do reciclo e calculando-se os isotopômeros de massa do PHA, os destinos dos átomos e as reações consideradas

Figura 17 – Variação dos isotopômeros de massa do PHA com a mudança no reciclo.

Fonte: Próprio autor

Uma outra possibilidade para tentar descrever os dados experimentais é utilizar uma rede metabólica que acople a Via das Pentoses (VP) e a via ED, como ilustrado na Figura 18. Para satisfazer todos os balanços de metabólitos, a Via das Pentoses no modelo aqui utilizado deve ter fluxo líquido nulo, usando a notação da seção 4.1.1, os fluxos *net* devem ser nulos, no entanto os fluxos *exchange* podem não ser nulos. Dessa forma, propõe-se aqui uma situação hipotética onde a VP opere no equilíbrio, assim, apesar de não afetar a distribuição de fluxos, a VP influencia na marcação dos metabólitos.

Um problema de estimação de fluxos foi formulado utilizando-se a rede metabólica e a distribuição de fluxos presentes na Figura 18, mais detalhes dos destinos dos átomos nas reações são fornecidos no Apêndice B, Tabela 18. Os parâmetros θ presentes na Figura 18 correspondem aos fluxos *exchange* das respectivas reações. Para o processo de estimação foram utilizados os dados de isotopômeros medidos por Kawai (2013) para o caso em que foi feita a limitação de Nitrogênio no meio.

Figura 18 – Rede metabólica simplificada da *Pseudomonas sp.* LFM046 considerando a via das pentoses em equilíbrio. Abreviações dos metabólitos: G6P, glicose-6-fosfato; F6P, frutose-6-fosfato; G3P, gliceraldeído-3-fosfato; PYR, piruvato; RIB5P, ribose-5-fosfato; Xul5P, xilose-5-fosfato; S7P, sedoheptulose-7-fosfato; E4P, eritrose-4-fosfato; AcCoA, acetilcoenzima A. Os fluxos têm unidades arbitrárias.

Fonte: Próprio autor

Os resultados da estimação dos fluxos *exchange* da VP para diferentes valores iniciais se encontram na Tabela 11. Todas as estimativas levaram a mesma função objetivo, $\Phi^* = 2.56$, no entanto apresentaram valores distintos dos parâmetros. Provavelmente há um problema de identificabilidade desses fluxos, ou seja, os fluxos não podem ser estimados a partir do modelo proposto e dos dados medidos. Problemas de identificabilidades envolvendo fluxos *exchange* são comuns na literatura (WURZEL; GRAAF, 1999; WINDEN et al., 2001a; KAPPELMANN et al., 2016). Uma análise de identificabilidade local *a posteriori* pode ser feita, como em Isermann e Wiechert (2003). Escolheu-se realizar a análise no ponto correspondente à terceira coluna da Tabela 11. Foi realizada a decomposição em valores singulares da matriz $\frac{\partial w_c}{\partial \theta}^T \cdot \frac{\partial w_c}{\partial \theta}$, sendo a matriz $\frac{\partial w_c}{\partial \theta}$ a variação dos isotopômeros calculados com os parâmetros. Os valores singulares calculados se encontram em 5.2 e os autovetores estão em 5.3. Como só há um valor singular não nulo, pode-se concluir que as medidas

de isotopômero só fixam um grau de liberdade do problema, e portanto tem-se infinitas soluções nesse ponto. A partir dos vetores do espaço nulo do problema, quatro últimas colunas de V , percebe-se que pelo menos localmente, a marcação dos isotopômeros é insensível ao parâmetro $\theta(5)$. Essa insensibilidade deve vir do fato de que a reação de isomerização de *RIB5P* a *Xul5P* não provoca embaralhamento. Portanto, esse parâmetro deve ser fixado no processo de estimação.

Tabela 11 – Diferentes parâmetros estimados a partir de pontos iniciais diferentes. Sendo Φ^* a função objetivo no ponto de mínimo.

Valor inicial θ_0	$\theta(1)$	$\theta(2)$	$\theta(3)$	$\theta(4)$	$\theta(5)$	Φ^*
(0.1,0.1,0.1,0.1,0.1)	0.9755	0.4376	0.4891	0.4735	0.1000	2.56
(0.3,0.3,0.3,0.3,0.3)	0.6673	0.6677	0.6678	0.6678	0.3000	2.56
(0.5,0.5,0.5,0.5,0.5)	0.6675	0.6674	0.6675	0.6675	0.5000	2.56
(0.7,0.7,0.7,0.7,0.7)	0.6675	0.6675	0.6675	0.6675	0.7000	2.56
(0.9,0.9,0.9,0.9,0.9)	0.6671	0.6690	0.6690	0.6660	0.9000	2.56

Fonte: Próprio autor.

$$S = \begin{pmatrix} 0.0048 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (5.2)$$

$$V = \begin{pmatrix} 0.9045 & -0.4239 & -0.0460 & -0.0052 & 0 \\ 0.1005 & 0.2172 & 0.0833 & -0.9673 & 0 \\ 0.1005 & 0.1066 & 0.9820 & 0.1189 & 0 \\ 0.4020 & 0.8728 & -0.1630 & 0.2237 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.3)$$

Apesar desses fluxos não poderem ser determinados, um resultado interessante surge dessa estimação; Na Tabela 12 é apresentada uma comparação entre os dados experimentais utilizados e os estimados, que é o mesmo para todos os parâmetros da Tabela 11. Embora não se possa afirmar quais são os fluxos em equilíbrio na VP, um padrão de marcação próximo ao experimental pode ser gerado, indicando que a *Pseudomonas sp.* LFM046 pode utilizar as VP e ED em conjunto para sintetizar PHA a partir de glicose.

Tabela 12 – Comparação entre os Isotopômeros de massa estimados e os medidos

	N limitado	Estimado: VP
M+0	0.58 ± 0.02	0.61
M+1	0.22 ± 0.01	0.22
M+2	0.15 ± 0.01	0.14
M+3	0.04 ± 0.01	0.03

Fonte: (KAWAI, 2013)

5.3 Estudo de Caso: Design ótimo de experimentos em *Pseudomonas*

Um segundo estudo de caso realizado neste trabalho exemplifica a aplicação da ferramenta desenvolvida no Design ótimo de experimentos. O design ótimo de experimentos se tornou um estudo imprescindível em ^{13}C -MFA para se diminuir os custos experimentais e aumentar a qualidade das estimativas (ANTONIEWICZ, 2013). O design ótimo de experimentos pode ter diferentes objetivos, os mais comuns são (BARD, 1974): otimizar a precisão da estimativa; otimizar as predições do seu modelo; possibilitar a seleção de modelos. No presente trabalho decidiu-se realizar o design visando otimizar a precisão da estimativa dos fluxos metabólicos. Um critério para alcançar esse fim é escolher os experimentos que fornecerão a menor variância dos parâmetros estimados, para isso, se minimiza o determinante da matriz de covariância dos parâmetros, que equivale a minimizar a região de confiança desses parâmetros (BARD, 1974). As variáveis que influenciam no design de experimentos com carbono marcado são o padrão de marcação dos substratos e quais medições de marcação de metabólito serão feitas.

Foi realizado o design ótimo para *Pseudomonas*, onde a única medição possível é a da marcação do polímero PHA. Sendo assim, o problema de design se resume a determinar qual a melhor marcação de glicose para se otimizar a qualidade da estimação. Como este é um design de experimento *a priori*, ou seja, antes de qualquer experimento, é necessário se assumir uma distribuição de fluxos, e assumir que o experimento ótimo será o mesmo independente da distribuição de fluxos, essa é uma hipótese razoável para ^{13}C -MFA (ANTONIEWICZ, 2013). As medidas de isotopômeros de massa utilizadas foram as do caso de nitrogênio limitado presentes na Tabela 9, bem como os seus respectivos desvios padrão. A rede metabólica utilizada e a distribuição de fluxo assumida para esse problema se encontram na Figura 19, os destinos dos átomos nas reações se encontram no Apêndice B, Tabela 19. A rede metabólica assumida tem como via de síntese do PHA a Via

das Pentose e a via ED, dessa forma o objetivo do design aqui escolhido foi determinar qual experimento permite estimar o fluxo que vai pela Via das Pentoses com a menor variância. O fluxo na VP será designado por θ e sua variância por $var(\theta)$ cuja estimação foi feita como na seção 4.3.3.

Figura 19 – Rede metabólica simplificada de *Pseudomonas putida*. Abreviações dos metabólitos: G6P, glicose-6-fosfato; F6P, frutose-6-fosfato; G3P, gliceraldeído-3-fosfato; PYR, piruvato; P5P, *pool* agrupado de ribose-5-fosfato e xilose-5-fosfato; S7P, sedoheptulose-7-fosfato; E4P, eritrose-4-fosfato; AcCoA, acetilcoenzima A. Os fluxos têm unidades arbitrárias.

Fonte: Próprio autor

Como as possibilidades de padrões de marcação da glicose são muitas, optou-se por estudar os tipos de marcação mais comuns. Primeiro, estudou-se o caso em que somente moléculas totalmente marcadas [$U\text{-}^{13}\text{C}$] e moléculas sem nenhuma marcação estão disponíveis. A Figura 20 apresenta como a variância do parâmetro θ varia conforme se aumenta a proporção de [$U\text{-}^{13}\text{C}$] utilizada. O melhor experimento nesse caso é o que se utiliza 59% da glicose [$U\text{-}^{13}\text{C}$]. Uma maior proporção de [$U\text{-}^{13}\text{C}$] fornece a menor variância porque se assumiu que as medidas do isotopômero de massa $M+0$ são as que possuem o maior erro experimental (Tabela 9).

Figura 20 – Variância do parâmetro θ para diferentes proporções de $[U-^{13}C]$

Fonte: Próprio autor

Outro tipo de marcação bastante utilizada é a molécula de glicose com apenas um de seus átomos marcados. A Tabela 13 fornece o valor da variância de θ para experimentos com 100% de glicose marcada em um de seus átomos. Nota-se primeiramente que as variâncias dos casos com marcações nos átomos 1 e 4 são infinitas, o que significa que não é possível se estimar o fluxo na via das pentoses com esses tipos de marcação. Isso porque esses átomos sempre saem como dióxido de carbono, e assim suas marcações não podem ser analisadas a partir do PHA. Excetuando-se esses dois casos, todos os outros experimentos apresentam uma qualidade na estimativa bem superior se comparado ao caso demonstrado na figura 20.

Tabela 13 – Comparação entre a variância de θ para diferentes marcações de glicose.

	[1- ^{13}C]	[2- ^{13}C]	[3- ^{13}C]	[4- ^{13}C]	[5- ^{13}C]	[6- ^{13}C]
Var(θ)	∞	0.0190	0.0022	∞	0.0012	0.0021

Fonte: Próprio autor.

Para compreender o porquê do átomo cinco fornecer a melhor estimativa para o fluxo na VP, devemos entender como os padrões de marcação de isotopômeros emergem nesse caso. Para isso, as Figuras 21 e 22 apresentam a variação dos isotopômeros do *AcCoA* e dos isotopômeros de massa de *PHA* em relação ao fluxo na via das pentoses,

respectivamente. Quando o fluxo na VP é nulo, todo o fluxo é direcionado pela via ED, e o *pool* de *AcCoA* tem metade de suas moléculas sem nenhuma marcação e a outra metade com marcação no primeiro átomo. No outro extremo, onde todo o fluxo é direcionado pela VP, todas as moléculas de *AcCoA* têm o primeiro átomo marcado. Dessa forma, para cada relação de fluxo entre as duas vias, há um padrão de marcação diferente e assim é possível se estimar a relação entre eles. Se atentarmos para a variação dos isotopômeros de massa de *PHA*, o isotopômero sem nenhuma marcação $M + 0$ vai diminuindo o seu pico conforme o fluxo na via das pentose aumenta. A marcação no átomo cinco é a que gera o menor pico de $M + 0$, sendo essa medida justamente a que fornece um erro maior, por essa razão a glicose marcada no carbono cinco é a melhor escolha para esse experimento.

Figura 21 – Variação dos isotopômeros de AcCoA com o fluxo na via das pentoses. Em que '00' representa o isotopômero onde dois átomos de AcCoA não estão marcados; '10' representa o isotopômero onde somente o primeiro átomo de AcCoA está marcado; Os outros isotopômeros do AcCoA têm marcação nula para qualquer valor do fluxo na VP.

Fonte: Próprio autor

Figura 22 – Variação dos isotopômeros de massa de PHA com o fluxo na via das pentoses. O pico M+3 é nulo para qualquer valor do fluxo na VP.

Fonte: Próprio autor

Um outro fator a ser considerado no design de experimentos é a questão econômica (ANTONIEWICZ, 2013). Embora experimentos com apenas um átomo marcado sejam mais informativos, o custo desses substratos é bem maior. Como exemplo, o preço de 100 mg de glicose [$U-^{13}C$] está em torno de R\$253.00 (SIGMA-ALDRICH, 2017b) e o preço de 100 mg de glicose [$5-^{13}C$] está em torno de R\$1,718.00 (SIGMA-ALDRICH, 2017a). A Figura 23 traz um gráfico importante na escolha da compra entre esses dois substratos, nele temos como a variância de θ varia para diferentes proporções de [$5-^{13}C$] e glicose sem marcação. Mesmo que se utilize somente 10% de glicose [$5-^{13}C$], a variância ainda é bem menor que o experimento com 60% de glicose [$U-^{13}C$]. Portanto a escolha da glicose [$5-^{13}C$] implica em uma melhora significativa na estimação, sem aumentar demasiadamente os custos experimentais.

Figura 23 – Variância do parâmetro θ para diferentes proporções de $[5-^{13}\text{C}]$

Fonte: Próprio autor

6 Conclusões

Como conclusões do trabalho, pode-se afirmar que:

- A ferramenta de simulação desenvolvida exibiu resultados consistentes e comparáveis aos obtidos com outros programas na literatura para a simulação de experimentos com carbono marcado.
- A ferramenta de estimação desenvolvida exibiu resultados consistentes e comparáveis aos obtidos com outros programas na literatura para a estimação de fluxos metabólicos com base em medições de marcação de carbono em metabólitos intermediários.
- O módulo de quantificação de incerteza nas estimativas pode ser utilizado para uma avaliação dos fluxos estimados.
- Do estudo de caso para *Pseudomonas sp.* LFM046 pode se concluir que esse micro-organismo deve utilizar a Via Entner-Doudoroff e a Via das Pentoses em conjunto na biossíntese de PHA.
- Do estudo de caso para uma rede metabólica genérica de *Pseudomonas sp.*, pode se concluir que a utilização de uma mistura com glicose marcada no carbono cinco e glicose não marcada é preferível à utilização de uma mistura de glicose totalmente marcada e glicose sem marcação, mesmo que se leve em conta fatores econômicos.

6.0.1 Sugestões para trabalhos futuros

Como para qualquer ferramenta computacional desenvolvida, a robustez só é alcançada com o tempo, quando aplicada a problemas cada vez mais complexos, as suas eventuais falhas são corrigidas. Dessa forma, propõe-se aqui que a ferramenta seja aplicada a micro-organismos com rotas distintas das estudadas aqui.

Uma outra sugestão é a inclusão de um módulo para realizar uma análise de identificabilidade na rede metabólica (KAPPELMANN et al., 2016), uma vez que, como visto nesse trabalho, fluxos não identificáveis são um problema comum em ^{13}C -MFA.

Um outro módulo que seria de grande utilidade para a ferramenta desenvolvida, seria a possibilidade de realizar cálculos com dados de experimentos com outros tipos de marcação, como ^2H e ^{18}O .

Por fim, como proposta mais desafiadora, pode-se incluir métodos para cálculos fora do estado estacionário isotópico conhecido como ^{13}C -NMFA ([ANTONIEWICZ, 2015](#)) e desenvolver métodos para cálculos fora do estado estacionário metabólico conhecido como ^{13}C -DMFA ([ANTONIEWICZ, 2015](#)). Essas técnicas permitem teoricamente a estimação de uma quantidade muito maior de fluxos com gastos experimentais muito menores.

Referências¹

- 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. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/bit.260210806/full>>.
- ANTONIEWICZ, M. R. ¹³C metabolic flux analysis: optimal design of isotopic labeling experiments. *Current Opinion in Biotechnology*, v. 24, n. 6, p. 1116–1121, dez. 2013. ISSN 09581669. Disponível em: <[http://linkinghub.elsevier.com/retrieve/pii-S0958166913000177](http://linkinghub.elsevier.com/retrieve/pii/S0958166913000177)>.
- ANTONIEWICZ, M. R. Methods and advances in metabolic flux analysis: a mini-review. *Journal of Industrial Microbiology & Biotechnology*, v. 42, n. 3, p. 317–325, mar. 2015. ISSN 1367-5435, 1476-5535. Disponível em: <<http://link.springer.com/10.1007/s10295-015-1585-x>>.
- ANTONIEWICZ, M. R.; KELLEHER, J. K.; STEPHANOPOULOS, G. Elementary metabolite units (EMU): A novel framework for modeling isotopic distributions. *Metabolic Engineering*, v. 9, n. 1, p. 68–86, jan. 2007. ISSN 10967176. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S109671760600084X>>.
- BAILEY, J. E. Toward a science of metabolic engineering. *Science*, v. 252, n. 5013, p. 1668–1675, jun. 1991. ISSN 0036-8075.
- BARD, Y. *Nonlinear Parameter Estimation*. [S.l.: s.n.], 1974.
- BEERS, K. J. *Numerical methods for chemical engineering: applications in Matlab*. Cambridge; New York: Cambridge University Press, 2007. OCLC: 708245930. ISBN 978-0-511-25538-0 978-0-511-64881-6 978-0-511-25483-3 978-0-511-25593-9 978-0-511-25650-9 978-0-511-81219-4. Disponível em: <<http://www.books24x7.com/marc.asp?bookid=17219>>.
- BLUM, J. J.; STEIN, R. B. On the analysis of metabolic networks. In: _____. *Biological Regulation and Development: Hormone Action*. Boston, MA: Springer US, 1982. p. 99–125. ISBN 978-1-4684-1125-6. Disponível em: <http://dx.doi.org/10.1007/978-1-4684-1125-6_3>.
- CARDINALI-REZENDE, J. et al. Draft Genome Sequence of *Pseudomonas* sp. Strain LFM046, a Producer of Medium-Chain-Length Polyhydroxyalkanoate. *Genome Announc.*, v. 3, n. 4, p. e00966–15, ago. 2015. ISSN , 2169-8287. Disponível em: <<http://genomea.asm.org/content/3/4/e00966-15>>.
- CHEN, J. et al. Improving metabolic flux estimation via evolutionary optimization for convex solution space. *Bioinformatics*, v. 23, n. 9, p. 1115–1123, maio 2007. ISSN 1367-4803, 1460-2059. Disponível em: <<https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btm050>>.
- CHRISTENSEN, B.; NIELSEN, J. Metabolic network analysis of *penicillium chrysogenum* using ¹³C-labeled glucose. *Biotechnol. Bioeng.*, v. 68, n. 6, p. 652–659, jun. 2000. ISSN 1097-0290. Disponível em: <[http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-0290\(20000620\)68:6<652::AID-BIT8>3.0.CO;2-J/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0290(20000620)68:6<652::AID-BIT8>3.0.CO;2-J/abstract)>.

¹ De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

CROWN, S. B.; ANTONIEWICZ, M. R. Publishing ¹³C metabolic flux analysis studies: A review and future perspectives. *Metabolic Engineering*, v. 20, p. 42–48, nov. 2013. ISSN 10967176. Disponível em: <[http://linkinghub.elsevier.com/retrieve/pii-S1096717613000839](http://linkinghub.elsevier.com/retrieve/pii/S1096717613000839)>.

CVIJOVIC, M. et al. BioMet Toolbox: genome-wide analysis of metabolism. *Nucl. Acids Res.*, v. 38, n. suppl 2, p. W144–W149, jul. 2010. ISSN 0305-1048, 1362-4962. Disponível em: <http://nar.oxfordjournals.org/content/38/suppl_2/W144>.

GUO, W.; SHENG, J.; FENG, X. ¹³C-Metabolic Flux Analysis: An Accurate Approach to Demystify Microbial Metabolism for Biochemical Production. *Bioengineering*, v. 3, n. 1, p. 3, dez. 2015. ISSN 2306-5354. Disponível em: <<http://www.mdpi.com/2306-5354/3/1/3>>.

HONG, K.-K.; NIELSEN, J. Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cell. Mol. Life Sci.*, v. 69, n. 16, p. 2671–2690, ago. 2012. ISSN 1420-9071.

ISERMANN, N.; WIECHERT, W. Metabolic isotopomer labeling systems. Part II: structural flux identifiability analysis. *Mathematical Biosciences*, v. 183, n. 2, p. 175–214, jun. 2003. ISSN 00255564. Disponível em: <[http://linkinghub.elsevier.com/retrieve/pii-S0025556402002225](http://linkinghub.elsevier.com/retrieve/pii/S0025556402002225)>.

KAJIHATA, S. et al. Research article, *OpenMebius: An Open Source Software for Isotopically Nonstationary ¹³C-Based Metabolic Flux Analysis*. 2014. DOI: 10.1155/2014/627014. Disponível em: <<https://www.hindawi.com/journals/bmri/2014/627014/>>.

KAPPELMANN, J.; WIECHERT, W.; NOACK, S. Cutting the Gordian Knot: Identifiability of anaplerotic reactions in *Corynebacterium glutamicum* by means of ¹³C-metabolic flux analysis. *Biotechnol. Bioeng.*, v. 113, n. 3, p. 661–674, mar. 2016. ISSN 1097-0290. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/bit.25833/abstract>>.

KATZ, J. Determination of gluconeogenesis in vivo with ¹⁴C-labeled substrates. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, v. 248, n. 4, p. R391–R399, abr. 1985. ISSN 0363-6119, 1522-1490. Disponível em: <<http://ajpregu.physiology.org/content/248/4/R391>>.

KAWAI, L. A. *Melhoramento da eficiência de produção de polihidroxialcanoatos por Pseudomonas sp. através da análise molecular e modificação genética*. Tese (Doutorado) — USP, Instituto de Ciências Biomédicas, abr. 2013.

KELLEHER, J. K. Analysis of tricarboxylic acid cycle using [¹⁴C]citrate specific activity ratios. *American Journal of Physiology - Endocrinology and Metabolism*, v. 248, n. 2, p. E252–E260, fev. 1985. ISSN 0193-1849, 1522-1555. Disponível em: <<http://ajpendo.physiology.org/content/248/2/E252>>.

KITANO, H. Systems biology: a brief overview. *Science*, v. 295, n. 5560, p. 1662–1664, 2002. Disponível em: <<http://science.sciencemag.org/content/295/5560/1662.short>>.

LI, Y. et al. Metabolic engineering of *Escherichia coli* using CRISPR–Cas9 mediated genome editing. *Metabolic Engineering*, v. 31, p. 13–21, set. 2015. ISSN 1096-7176. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1096717615000750>>.

LONG, H.; WU, S.; FU, H. Improving Metabolic Flux Estimation of Metabolic Networks by QPSO with Penalty Function. In: *2014 Tenth International Conference on Computational Intelligence and Security*. [S.l.: s.n.], 2014. p. 52–57.

LONG, H.-X.; XU, W.-B.; SUN, J. Metabolic flux estimation using particle swarm optimization with penalty function. In: *Biology Forum/Rivista di Biologia*. [S.l.: s.n.], 2009. v. 102, n. 2.

MARX, A. et al. Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol. Bioeng.*, v. 49, n. 2, p. 111–129, jan. 1996. ISSN 1097-0290. Disponível em: <[http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-0290\(19960120\)49:2<111::AID-BIT1>3.0.CO;2-T/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0290(19960120)49:2<111::AID-BIT1>3.0.CO;2-T/abstract)>.

NAKAMURA, C. E.; WHITED, G. M. Metabolic engineering for the microbial production of 1,3-propanediol. *Current Opinion in Biotechnology*, v. 14, n. 5, p. 454–459, out. 2003. ISSN 0958-1669. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0958166903001265>>.

NIELSEN, J. Metabolic engineering. *Applied Microbiology and Biotechnology*, v. 55, n. 3, p. 263–283, abr. 2001. ISSN 0175-7598, 1432-0614. Disponível em: <<http://link.springer.com/10.1007/s002530000511>>.

NIELSEN, J. et al. Engineering synergy in biotechnology. *Nat Chem Biol*, v. 10, n. 5, p. 319–322, maio 2014. ISSN 1552-4450. Disponível em: <<http://www.nature.com/nchembio/journal/v10/n5/full/nchembio.1519.html>>.

ORTH, J. D.; THIELE, I.; PALSSON, B. What is flux balance analysis? *Nature Biotechnology*, v. 28, n. 3, p. 245–248, mar. 2010. ISSN 1087-0156, 1546-1696. Disponível em: <<http://www.nature.com/doifinder/10.1038/nbt.1614>>.

OTERO, J. M.; NIELSEN, J. Industrial systems biology. *Biotechnology and Bioengineering*, v. 105, n. 3, p. 439–460, fev. 2010. ISSN 00063592, 10970290. Disponível em: <<http://doi.wiley.com/10.1002/bit.22592>>.

PADDON, C. J. et al. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*, v. 496, n. 7446, p. 528–532, abr. 2013. ISSN 1476-4687.

PORTAIS, J.-C. et al. Metabolic flux determination in C6 glioma cells using carbon-13 distribution upon [1-13c]glucose incubation. *European Journal of Biochemistry*, v. 217, n. 1, p. 457–468, out. 1993. ISSN 1432-1033. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1993.tb18265.x/abstract>>.

RIASCOS, C. A.; GOMBERT, A. K.; PINTO, J. M. A global optimization approach for metabolic flux analysis based on labeling balances. *Computers & Chemical Engineering*, v. 29, n. 3, p. 447–458, fev. 2005. ISSN 00981354. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0098135404002352>>.

RIASCOS, C. A. M. et al. Metabolic pathways analysis in PHAs production by *Pseudomonas* with 13c-labeling experiments. *Computer Aided Chemical Engineering*, v. 32, p. 121–126, jan. 2013. ISSN 1570-7946. Disponível em: <<http://www.sciencedirect.com/science/article/pii/B978044463234050021X>>.

SCHMIDT, K. et al. Modeling isotopomer distributions in biochemical networks using isotopomer mapping matrices. *Biotechnology and Bioengineering*, v. 55, n. 6, p. 831–840, set. 1997. ISSN 1097-0290. Disponível em: <[http://onlinelibrary.wiley.com/doi/10.1002-\(SICI\)1097-0290\(19970920\)55:6<831::AID-BIT2>3.0.CO;2-H/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0290(19970920)55:6<831::AID-BIT2>3.0.CO;2-H/abstract)>.

SCHUSTER, S.; HLGETAG, C. ON ELEMENTARY FLUX MODES IN BIOCHEMICAL REACTION SYSTEMS AT STEADY STATE. *Journal of Biological Systems*, v. 2, n. 2, p. 165–182, jan. 1994.

SHUPLETSOV, M. S. et al. OpenFLUX2: ¹³C-MFA modeling software package adjusted for the comprehensive analysis of single and parallel labeling experiments. *Microbial Cell Factories*, v. 13, p. 152, nov. 2014. ISSN 1475-2859. Disponível em: <<https://doi.org/10.1186/s12934-014-0152-x>>.

SIGMA-ALDRICH. *D-Glucose-5*. 2017. Disponível em: <<http://www.sigmaaldrich.com/catalog/product/aldrich/310808>>.

SIGMA-ALDRICH. *D-Glucose-U*. 2017. Disponível em: <<http://www.sigmaaldrich.com/catalog/product/aldrich/389374>>.

SOKOL, S.; MILLARD, P.; PORTAIS, J.-C. influx_s: increasing numerical stability and precision for metabolic flux analysis in isotope labelling experiments. *Bioinformatics*, v. 28, n. 5, p. 687–693, mar. 2012. ISSN 1367-4811.

SROUR, O.; YOUNG, J. D.; ELDAR, Y. C. Fluxomers: a new approach for ¹³C metabolic flux analysis. *BMC systems biology*, v. 5, n. 1, p. 1, 2011. Disponível em: <<http://bmcsystbiol.biomedcentral.com/articles/10.1186/1752-0509-5-129>>.

STEPHANOPOULOS, G. Metabolic engineering. *Current Opinion in Biotechnology*, v. 5, n. 2, p. 196–200, 1994. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0958166905800369>>.

STEPHANOPOULOS, G. Metabolic engineering: perspective of a chemical engineer. *AIChE journal*, v. 48, n. 5, p. 920–926, 2002. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/aic.690480502/full>>.

STEPHANOPOULOS, G.; ARISTIDOU, A. A.; NIELSEN, J. *Metabolic Engineering Principles and Methodologies*. [S.l.]: Academic Press, 1998.

STEPHANOPOULOS, G.; VALLINO, J. Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnology and Bioengineering*, v. 41, p. 633–646, mar. 1993.

SZYPERSKI, T. Biosynthetically Directed Fractional ¹³C-labeling of Proteinogenic Amino Acids. *European Journal of Biochemistry*, v. 232, n. 2, p. 433–448, set. 1995. ISSN 1432-1033. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1111/j.1432-1033.1995.433zz.x/abstract>>.

SZYPERSKI, T. ¹³C-NMR, MS and metabolic flux balancing in biotechnology research. *Quarterly reviews of biophysics*, v. 31, n. 01, p. 41–106, 1998. Disponível em: <http://journals.cambridge.org/abstract_S0033583598003412>.

TETHNOL, M.; RAPOPORT, M.; STEPHANOPOULOS, G. Network rigidity and metabolic engineering in metabolite overproduction. *Biol. Chem*, v. 259, p. 9646, 1991. Disponível em: <<http://search.proquest.com/openview/98c7d4bc8fa3881b9757a6e627c79597/1.pdf?pq-origsite=gscholar&cbl=1256>>.

TYO, K. E.; KOCHARIN, K.; NIELSEN, J. Toward design-based engineering of industrial microbes. *Current Opinion in Microbiology*, v. 13, n. 3, p. 255–262, jun. 2010. ISSN 13695274. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1369527410000226>>.

WEITZEL, M. et al. 13cflux2—high-performance software suite for 13c-metabolic flux analysis. *Bioinformatics*, v. 29, n. 1, p. 143–145, jan. 2013. ISSN 1367-4803, 1460-2059. Disponível em: <<http://bioinformatics.oxfordjournals.org/content/29/1/143>>.

WIECHERT, W. 13c Metabolic Flux Analysis. *Metabolic Engineering*, mar. 2001.

WIECHERT, W. Modeling and simulation: tools for metabolic engineering. *Journal of biotechnology*, v. 94, n. 1, p. 37–63, 2002. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168165601004187>>.

WIECHERT, W. The Thermodynamic Meaning of Metabolic Exchange Fluxes. *Biophysical Journal*, v. 93, n. 6, p. 2255–2264, set. 2007. ISSN 0006-3495. Disponível em: <[http://www.cell.com/biophysj/abstract/S0006-3495\(07\)71480-7](http://www.cell.com/biophysj/abstract/S0006-3495(07)71480-7)>.

WIECHERT, W.; GRAAF, A. A. de. Bidirectional reaction steps in metabolic networks: I. Modeling and simulation of carbon isotope labeling experiments. *Biotechnology and Bioengineering*, v. 55, n. 1, p. 101–117, jul. 1997. ISSN 1097-0290. Disponível em: <[http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-0290\(19970705\)55:1<101::AID-BIT12>3.0.CO;2-P/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0290(19970705)55:1<101::AID-BIT12>3.0.CO;2-P/abstract)>.

WIECHERT, W.; NOACK, S. Mechanistic pathway modeling for industrial biotechnology: challenging but worthwhile. *Current Opinion in Biotechnology*, v. 22, n. 5, p. 604–610, out. 2011. ISSN 09581669. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0958166911000152>>.

WIECHERT, W.; WURZEL, M. Metabolic isotopomer labeling systems: Part I: global dynamic behavior. *Mathematical biosciences*, v. 169, n. 2, p. 173–205, 2001. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0025556400000596>>.

WINDEN, W. A. van; HEIJNEN, J. J.; VERHEIJEN, P. J. Cumulative bondomers: A new concept in flux analysis from 2d [13c,1h] COSY NMR data. *Biotechnology and Bioengineering*, v. 80, n. 7, p. 731–745, dez. 2002. ISSN 0006-3592, 1097-0290. Disponível em: <<http://doi.wiley.com/10.1002/bit.10429>>.

WINDEN, W. A. van et al. A priori analysis of metabolic flux identifiability from 13c-labeling data. *Biotechnol. Bioeng.*, v. 74, n. 6, p. 505–516, set. 2001. ISSN 1097-0290. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/bit.1142/abstract>>.

WINDEN, W. van; VERHEIJEN, P.; HEIJNEN, S. Possible Pitfalls of Flux Calculations Based on 13c-Labeling. *Metabolic Engineering*, v. 3, n. 2, p. 151–162, abr. 2001. ISSN 1096-7176. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1096717600901745>>.

WURZEL, M.; GRAAF, A. A. de. Bidirectional Reaction Steps in Metabolic Networks: III. Explicit Solution and Analysis of Isotopomer Labeling Systems. *BIOTECHNOLOGY AND BIOENGINEERING*, v. 66, n. 2, 1999. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.467.6917&rep=rep1&type=pdf>>.

YANG, J. et al. Metabolic flux estimation—a self-adaptive evolutionary algorithm with singular value decomposition. *IEEE/ACM Trans Comput Biol Bioinform*, v. 4, n. 1, p. 126–138, mar. 2007. ISSN 1545-5963.

YANG, T. H.; FRICK, O.; HEINZLE, E. Hybrid optimization for ^{13}C metabolic flux analysis using systems parametrized by compactification. *BMC Systems Biology*, v. 2, p. 29, 2008. ISSN 1752-0509. Disponível em: <<http://dx.doi.org/10.1186/1752-0509-2-29>>.

YOUNG, J. D. INCA: a computational platform for isotopically non-stationary metabolic flux analysis. *Bioinformatics*, v. 30, n. 9, p. 1333–1335, maio 2014. ISSN 1367-4803. Disponível em: <<https://academic.oup.com/bioinformatics/article/30/9/1333/236911/INCA-a-computational-platform-for-isotopically-non>>.

ZAMBONI, N.; FISCHER, E.; SAUER, U. FiatFlux—a software for metabolic flux analysis from ^{13}C -glucose experiments. 2005. Disponível em: <<http://e-collection.library.ethz.ch/view/eth:28082>>.

ZHAO, J.; SHIMIZU, K. Metabolic flux analysis of *Escherichia coli* K12 grown on ^{13}C -labeled acetate and glucose using GC-MS and powerful flux calculation method. *Journal of Biotechnology*, v. 101, n. 2, p. 101–117, mar. 2003. ISSN 0168-1656. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168165602003164>>.

ZOMORRODI, A. R. et al. Mathematical optimization applications in metabolic networks. *Metabolic Engineering*, v. 14, n. 6, p. 672–686, nov. 2012. ISSN 1096-7176. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1096717612001000>>.

ZUPKE, C.; STEPHANOPOULOS, G. Modeling of isotope distributions and intracellular fluxes in metabolic networks using atom mapping matrices. *Biotechnology Progress*, v. 10, n. 5, p. 489–498, 1994. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1021/bp00029a006/full>>.

Apêndice A – Detalhes do Modelo matemático

A.1 Equações de restrição

Na seção 4.1.1 foram apresentadas as transformações de variáveis realizadas no vetor de fluxo, bem como as equações que restringem o espaço viável de fluxo. Nesta seção serão apresentados os detalhes da construção dessas equações de restrição por intermédio de um exemplo ilustrativo. A Figura 24 apresenta uma rede metabólica simples, onde o metabólito A é convertido no metabólito B , que por sua vez se transforma em C por dois fluxos deferentes, sendo um deles bidirecional. Por fim o metabólito C é convertido em D . Dessa forma, o sistema possui 5 variáveis de fluxo e dois balanços de metabólitos. A Tabela 14 apresenta os destinos dos átomos em cada reação. Todas as moléculas possuem dois átomos e a única reação que promove o embaralhamento entre os átomos é a reação v_3 .

Figura 24 – Exemplo de uma rede metabólica.

Fonte: Próprio autor

Tabela 14 – Estequiometria e transições da rede apresentada na Figura 24. As letras minúsculas representam átomos sucessivos em cada metabólito.

Reação	Estequiometria	Transição entre os átomos
v_1	$A \rightarrow B$	$ab \rightarrow ab$
v_2	$B \rightarrow C$	$ab \rightarrow ab$
v_3	$B \rightarrow C$	$ab \rightarrow ba$
v_4	$C \rightarrow D$	$ab \rightarrow ab$

Fonte: Próprio autor

Todas as restrições serão aplicadas na coordenada numérica, como descrito na seção 4.1.1. O conjunto de coordenadas numéricas inclui as variáveis v^{net} e $v^{xch[0,1]}$, restrições

lineares são aplicáveis a ambas. A equação de restrição linear na coordenada *net* (equação 4.4) inclui as seguintes restrições:

- Balanços dos metabólitos internos. Esses balanços são incluídos através da matriz estequiométrica, que sempre é uma sub matriz da matriz de restrição linear do fluxo *net* (N^{net}).

$$B : v_1^{net} = v_2^{net} + v_3^{net}$$

$$C : v_2^{net} + v_3^{net} = v_4^{net}$$

- Balanços de energia para ATP, NADH ou NADPH podem ser expressos como restrições no fluxo *net* (WIECHERT; GRAAF, 1997).
- Algum fluxo v_i^{net} pode ser fixado em um determinado valor n_i^{net} ; ou seja, $v_i^{net} = n_i^{net}$.

No exemplo da Figura 24, além da restrição da matriz estequiométrica, o fluxo de entrada pode ser fixado $v_1^{net} = 1$. Dessa forma a equação de restrição linear para o fluxo *net* fica:

$$N^{net} \cdot v^{net} = n^{net}$$

$$N^{net} = \begin{pmatrix} 1 & -1 & -1 & \cdot \\ \cdot & 1 & 1 & -1 \\ 1 & \cdot & \cdot & \cdot \end{pmatrix}$$

$$n^{net} = \begin{pmatrix} \cdot \\ \cdot \\ 1 \end{pmatrix}$$

$$v^{net} = \begin{pmatrix} v_1^{net} \\ v_2^{net} \\ v_3^{net} \\ v_4^{net} \end{pmatrix}$$

A equação de restrição linear para os fluxos *exchange* (equação 4.5) também pode ser formulada para esse problema, e inclui as seguintes restrições:

- Unidirecionalidade da reação *i* (WIECHERT, 2007); ou seja, $v_i^{xch[0,1]} = 0$.
- Equilíbrio rápido da reação *i* (WIECHERT, 2007); ou seja, $v_i^{xch[0,1]} = 1$.
- Fixar um fluxo *exchange i* em um determinado valor $n_i^{xch[0,1]}$; ou seja, $v_i^{xch[0,1]} = n_i^{xch[0,1]}$.
- Igualar dois fluxos *exchange*; ou seja, $v_i^{xch[0,1]} = v_j^{xch[0,1]}$.

No exemplo da Figura 24, os fluxos v_1 , v_3 e v_4 são unidirecionais. Desta forma a equação de restrição linear para o fluxo *exchange* fica:

$$N^{xch[0,1]} \cdot v^{xch[0,1]} = n^{xch[0,1]}$$

$$N^{xch[0,1]} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}$$

$$n^{xch[0,1]} = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

$$v^{xch[0,1]} = \begin{pmatrix} v_1^{xch[0,1]} \\ v_2^{xch[0,1]} \\ v_3^{xch[0,1]} \\ v_4^{xch[0,1]} \end{pmatrix}$$

As relações de restrição linear para o vetor de fluxos apresentadas geralmente não são suficientes para a completa determinação de todos os fluxos metabólicos. Sendo assim, alguns fluxos *net* e/ou *xch* devem ser fixados para que se satisfaçam os graus de liberdade restantes. Para definir esses fluxos, o conceito de fluxos livres (θ) foi introduzido na seção 4.1.1, equação 4.6. Os fluxos livres podem ser fixados para uma simulação, ou podem variar no processo de estimação. No exemplo da Figura 24, temos 8 fluxos, quatro *net* e 4 *exchange*, e as restrições lineares fixam 6 deles. Assim, os fluxos v_2^{net} e $v_2^{xch[0,1]}$ são escolhidos para serem os fluxos livres.

$$N^{free} \cdot \begin{pmatrix} v^{net} \\ v^{xch[0,1]} \end{pmatrix} = \theta$$

$$N^{free} = \begin{pmatrix} \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \end{pmatrix}$$

Todas essas restrições lineares podem ser concatenadas em uma única equação:

$$N = \begin{pmatrix} N^{net} & 0 \\ 0 & N^{xch[0,1]} \\ & N^{free} \end{pmatrix}$$

$$n = \begin{pmatrix} n^{net} \\ n^{xch[0,1]} \\ \theta \end{pmatrix}$$

$$N \cdot \begin{pmatrix} v^{net} \\ v^{xch[0,1]} \end{pmatrix} = n$$

Se os fluxos livres forem escolhidos corretamente, essa equação irá de forma única determinar todos os fluxos. Para se evitar situações onde as restrições são insuficientes ou redundantes, exige-se que a matriz N seja quadrada e possua inversa. Para se excluir fluxos sem significado fisiológico como fluxos negativos de carbono, uma equação de restrição linear de desigualdade é construída. Nela se inclui as seguintes restrições:

- Fluxos *exchange* são restritos ao intervalo $0 \leq v^{xch[0,1]} \leq 1$.
- A direção das reações unidirecionais pode ser avaliada. $v_i^{net} \geq 0$ ou $v_i^{net} \leq 0$.
- Algum fluxo *net* ou *exchange* pode ter um limite superior ou inferior que venha de outros experimentos, ou que evite algum problema numérico.

No exemplo da Figura 24, os fluxos v_1^{net} , v_3^{net} e v_4^{net} devem ser positivos, para que o fluxo seja no sentido indicado na rede. Essas restrições estão nas três primeiras linhas da matriz U , as outras linhas garantem que os fluxos *exchange* estejam no intervalo entre $[0,1]$.

$$U \cdot \begin{pmatrix} v^{net} \\ v^{xch[0,1]} \end{pmatrix} \geq u$$

$$U = \begin{pmatrix} 1 & . & . & . & . & . & . & . & . \\ . & . & 1 & . & . & . & . & . & . \\ . & . & . & 1 & . & . & . & . & . \\ . & . & . & . & 1 & . & . & . & . \\ . & . & . & . & -1 & . & . & . & . \\ . & . & . & . & . & 1 & . & . & . \\ . & . & . & . & . & -1 & . & . & . \\ . & . & . & . & . & . & 1 & . & . \\ . & . & . & . & . & . & -1 & . & . \\ . & . & . & . & . & . & . & 1 & . \\ . & . & . & . & . & . & . & . & -1 \end{pmatrix}$$

$$u = \begin{pmatrix} . \\ . \\ . \\ . \\ -1 \\ . \\ -1 \\ . \\ -1 \\ . \\ -1 \end{pmatrix}$$

A.2 Balanço de cumômeros

Os balanços de isotopômeros, equação 4.11, foram formulados por Wurzel e Graaf (1999) da seguinte forma:

$$\bar{f}(v, \bar{x}, \bar{x}^{inp}) = \frac{1}{2} \bar{x}^T \cdot \left(\sum_i \vec{v} \cdot \vec{Q}_i + \overleftarrow{v} \cdot \overleftarrow{Q}_i \right) \cdot \bar{x} + \left(\sum_i \vec{v}_i \cdot \vec{P}_i + \sum_i \overleftarrow{v}_i \cdot \overleftarrow{P}_i \right) \cdot \bar{x} + \left(\sum_i \vec{v}_i \cdot \vec{P}_i^{inp} \right) \cdot \bar{x}^{inp} = 0$$

Onde $\bar{x}$ e $\bar{x}^{inp}$ são as frações isotopoméricas nos metabólitos internos e no substrato, respectivamente. $\vec{P}_i$ e $\overleftarrow{P}_i^{inp}$ são as matrizes de transição unimoleculares de isotopômeros para os fluxos internos e para o fluxo de entrada respectivamente. $\vec{Q}_i$ são as matrizes de transição bimoleculares de isotopômeros para os fluxos internos. Os vetores de fração isotopomérica são ordenados primeiro em metabólitos e depois em ordem binária. No exemplo da Figura 24, os vetores são construídos da seguinte forma:

$$\bar{x}^{inp} = \begin{pmatrix} a_{00} & a_{01} & a_{10} & a_{11} \end{pmatrix}^T$$

$$\bar{x} = \begin{pmatrix} b_{00} & b_{01} & b_{10} & b_{11} & c_{00} & c_{01} & c_{10} & c_{11} \end{pmatrix}^T$$

Na notação adotada, 0 representa os átomos não marcados e 1 os átomos marcados, como apresentado na seção 3.1.2. Na equação de balanço de isotopômero há três termos, o primeiro contabiliza as transições de marcações de reações bimoleculares; o segundo contabiliza as transições em reações unimoleculares; e o último as transições nas reações de entrada, que também são unimoleculares. Qualquer reação que tenha uma ordem maior que 2 deve ser fragmentada em reações de ordem menor. As matrizes de transição de isotopômeros unimolecular $\vec{P}_i$ são matrizes quadradas com a dimensão do número de isotopômeros total. Na construção dessa matriz, as seguintes regras são seguidas:

$$(\vec{P}_i)_{j,k} = \begin{cases} l & \text{se a reação } i \text{ no sentido direto alimenta o isotopômero } j \\ & \text{com } l \text{ átomos de carbono do } pool \text{ do isotopômero } k \\ -l & \text{se } j=k \text{ e a reação } i \text{ no sentido direto retira} \\ & l \text{ átomos de carbono do } pool \text{ do isotopômero } j \\ 0 & \text{caso contrário} \end{cases}$$

As mesmas regras prevalecem para $\overleftarrow{P}_i$ e $\overleftarrow{P}_i^{inp}$. No exemplo da Figura 24, a matriz para o fluxo v_3 é construídos da seguinte forma:

$$\vec{P}_3 = \begin{pmatrix} -1 & . & . & . & . & . & . & . \\ . & -1 & . & . & . & . & . & . \\ . & . & -1 & . & . & . & . & . \\ . & . & . & -1 & . & . & . & . \\ 1 & . & . & . & . & . & . & . \\ . & . & 1 & . & . & . & . & . \\ . & 1 & . & . & . & . & . & . \\ . & . & . & 1 & . & . & . & . \end{pmatrix}$$

As matrizes de transição de isotopômeros bimolecular são mais complexas, pois descrevem transições de marcação de dois reagentes para um ou mais produtos. Dessa forma, essa matriz precisa ter três dimensões e a matriz $\vec{Q}_i$ que aparece nos balanços é na verdade um vetor de matrizes:

$$\vec{Q}_i = \begin{pmatrix} \vec{Q}_{i,1} \\ \vdots \\ \vec{Q}_{i,j} \\ \vdots \\ \vec{Q}_{i,dim\bar{x}} \end{pmatrix}$$

Cada posição nesse vetor descreve as transições de marcação do fluxo i para o isotopômero j . As regras para a construção dessas matrizes são:

$$(\vec{Q}_{i,j})_{k,s} = \begin{cases} 1 & \text{se a reação } i \text{ no sentido direto gera o isotopômero } j \\ & \text{combinando os isotopômeros } k \text{ e } s \\ 0 & \text{caso contrário} \end{cases}$$

As mesmas regras prevalecem para $\vec{Q}_i^{\leftarrow}$. É notório observar que as matrizes Q_i são simétricas, e assim é necessário o termo $\frac{1}{2}$ na equação de balanço para não se contabilizar as transições em dobro. Se usarmos agora a transformação de variáveis apresentada na seção 3.1.2, podemos criar uma equação para o balanço de cumômeros (x) que possui a mesma estrutura:

$$f(v, x, x^{inp}) = \frac{1}{2} x^T \cdot \left(\sum_i \vec{v} \cdot \vec{Q}_i + \vec{v} \cdot \vec{Q}_i^{\leftarrow} \right) \cdot x + \left(\sum_i \vec{v}_i \cdot \vec{P}_i + \sum_i \vec{v}_i^{\leftarrow} \cdot \vec{P}_i^{\leftarrow} \right) \cdot x + \left(\sum_i \vec{v}_i \cdot P_i^{inp} \right) \cdot x^{inp} = 0 \quad (\text{A.1})$$

As matrizes de transição dos cumômeros podem ser construídas a partir das matrizes de transição de isotopômeros, utilizando a propriedade da conservação dos pesos dos cumômeros, como demonstrado por Wurzel e Graaf (1999). A definição dos pesos dos cumômeros foi feita na seção 3.1.2.

$$(\vec{P}_i)_{j,k} = \begin{cases} (\vec{P}_i)_{j,k} & \text{se peso(k) = peso(j)} \\ 0 & \text{caso contrário} \end{cases}$$

$$(\vec{Q}_{i,j})_{k,s} = \begin{cases} (\vec{Q}_{i,j})_{k,s} & \text{se peso(k) + peso(s) = peso(j)} \\ 0 & \text{caso contrário} \end{cases}$$

Como já mencionado neste texto, os cumômeros são variáveis menos acopladas do que os isotopômeros, e essa propriedade facilita os cálculos. Com isso, a equação A.1 pode ser rearranjada de forma que facilite a sua solução. O primeiro passo para alcançar esse objetivo é ordenar os vetores de marcação de cumômero em uma ordem diferente, primeiro se ordena por pesos e depois por metabólitos. Como exemplo, os vetores para a rede da Figura 24 são:

$$x^{inp} = \begin{pmatrix} a_{xx} & a_{x1} & a_{1x} & a_{11} \end{pmatrix}^T$$

$$x = \begin{pmatrix} b_{xx} & c_{xx} & b_{x1} & c_{x1} & b_{1x} & c_{1x} & b_{11} & c_{11} \end{pmatrix}^T$$

Uma notação mais concisa para esses vetores pode ser feita utilizando ${}^n x^{inp}$ e ${}^n x$ que são frações de marcação de peso n para o *input* e para os metabólitos internos, respectivamente. Os vetores de marcação de cumômeros podem então ser representados da seguinte forma:

$$x^{inp} = \begin{pmatrix} {}^0 x^{inp} & {}^1 x^{inp} & {}^2 x^{inp} \end{pmatrix}^T$$

$$x = \begin{pmatrix} {}^0 x & {}^1 x & {}^2 x \end{pmatrix}^T$$

As matrizes de transição também podem ser particionadas em pesos. Uma vez que há a conservação de pesos nas transições entre cumômeros, as matrizes de transição só operam nos vetores de cumômeros de peso correspondente, consequentemente:

$$\vec{P}_i \cdot x = \begin{pmatrix} {}^0 \vec{P}_i \cdot {}^0 x \\ {}^1 \vec{P}_i \cdot {}^1 x \\ \vdots \end{pmatrix}$$

$$x^T \cdot \vec{Q}_i \cdot x = \begin{pmatrix} {}^0 x^T \cdot {}^{0,0} \vec{Q}_i \cdot {}^0 x \\ {}^0 x^T \cdot {}^{0,1} \vec{Q}_i \cdot {}^1 x + {}^1 x^T \cdot {}^{1,0} \vec{Q}_i \cdot {}^0 x \\ {}^0 x^T \cdot {}^{0,2} \vec{Q}_i \cdot {}^2 x + {}^1 x^T \cdot {}^{1,1} \vec{Q}_i \cdot {}^1 x + {}^2 x^T \cdot {}^{2,0} \vec{Q}_i \cdot {}^0 x \\ \vdots \end{pmatrix}$$

A equação A.1 pode ser particionada em uma equação para cada peso, para a equação de peso n temos:

$$\frac{1}{2} \sum_{k+l=n} [{}^k x^T \cdot (\sum_i \vec{v}_i \cdot {}^{k,l} \vec{Q}_i + \overleftarrow{v} \cdot {}^{k,l} \overleftarrow{Q}_i) \cdot {}^l x] + (\sum_i \vec{v}_i \cdot {}^n \vec{P}_i + \sum_i \overleftarrow{v}_i \cdot {}^n \overleftarrow{P}_i) \cdot {}^n x + (\sum_i \vec{v}_i \cdot {}^n P_i^{inp}) \cdot {}^n x^{inp} = 0 \quad (A.2)$$

Utilizando o fato de que os cumômeros de peso zero são todos iguais a um (${}^0 x = 1$), como demonstrado na seção 3.1.2, e a seguinte propriedade de simetria das matrizes Q vale:

$${}^0 x^T \cdot {}^{0,n} \vec{Q}_i \cdot {}^n x = 1 \cdot {}^{0,n} \vec{Q}_i \cdot {}^n x$$

$${}^n x^T \cdot {}^{n,0} \vec{Q}_i \cdot {}^0 x = 1 \cdot {}^{n,0} \vec{Q}_i \cdot {}^n x$$

Pode-se construir uma equação onde todos os termos se tornam lineares em relação a ${}^n x$:

$$\begin{aligned} & [\vec{v}_i \cdot 1 \cdot ({}^{0,n} \vec{Q}_i + {}^{n,0} \vec{Q}_i) + \overleftarrow{v}_i \cdot 1 \cdot ({}^{0,n} \overleftarrow{Q}_i + {}^{n,0} \overleftarrow{Q}_i) + \sum_i \vec{v}_i \cdot {}^n \vec{P}_i + \sum_i \overleftarrow{v}_i \cdot {}^n \overleftarrow{P}_i] \cdot {}^n x \\ & + \left[\frac{1}{2} \sum_{k+l=n} [{}^k x^T \cdot (\sum_i \vec{v}_i \cdot {}^{k,l} \vec{Q}_i + \overleftarrow{v} \cdot {}^{k,l} \overleftarrow{Q}_i) \cdot {}^l x] + (\sum_i \vec{v}_i \cdot {}^n P_i^{inp}) \cdot {}^n x^{inp} \right] = 0 \end{aligned}$$

Para facilitar a visualização da equação, o termo à esquerda de ${}^n x$ é substituído por ${}^n A$, e o termo à direita por ${}^n b$:

$$0 = {}^n A(v) \cdot {}^n x + {}^n b(v, {}^1 x, \dots, {}^{n-1} x)$$

Observa-se das equações acima que o termo ${}^n A$ contém as marcações que chegam e saem de apenas um reagente, e portanto, todos os termos são lineares. O termo ${}^n b$ contém as marcações que vêm diretamente do substrato e as marcações que vêm simultaneamente de dois reagentes. Este último termo, contém termos bilineares, no entanto, esses termos só dependem dos fluxos e das frações de marcação de pesos menores que n . Sendo assim, essas equações podem ser resolvidas sequencialmente da seguinte forma:

$$\begin{aligned} 1 &= {}^0 x \\ 0 &= {}^1 A(v) \cdot {}^1 x + {}^1 b(v) \\ 0 &= {}^2 A(v) \cdot {}^2 x + {}^2 b(v, {}^1 x) \\ 0 &= {}^3 A(v) \cdot {}^3 x + {}^3 b(v, {}^1 x, {}^2 x) \\ &\vdots \end{aligned}$$

Apêndice B – Destino dos átomos nas redes metabólicas

Tabela 15 – Estequiometria e transições do ciclo TCA. As letras minúsculas representam átomos sucessivos em cada metabólito. As abreviações são as mesma utilizadas na Figura 10.

Reação	Estequiometria	transição atômica
1	OAC + AcCoA → Cit	abcd + ef → dcbfea
2	Cit → AKG + CO ₂	abcdef → abcde + f
3	AKG → Glu	abcde → abcde
4	AKG → Suc + CO ₂	abcde → bcde + a
5	Suc → Fum	$\frac{1}{2}$ abcd + $\frac{1}{2}$ dcba → $\frac{1}{2}$ abcd + $\frac{1}{2}$ dcba
6	Fum → OAC	$\frac{1}{2}$ abcd + $\frac{1}{2}$ dcba → abcd
7	OAC → Fum	abcd → $\frac{1}{2}$ abcd + $\frac{1}{2}$ dcba
8	Asp → OAC	abcd → abcd

Fonte: (ANTONIEWICZ et al., 2007)

Tabela 16 – Estequiometria e transições da rede simplificada de E.coli. As letras minúsculas representam átomos sucessivos em cada metabólito. As abreviações são as mesma utilizadas na Figura 11.

Reação	Estequiometria	transição atômica
upt	GLC → Glc6P	abcdef → abcdef
emp1	Glc6P → Fru6P	abcdef → abcdef
emp2	Fru6P → FruBP	abcdef → abcdef
emp3	FruBP → GA3P	abcdef → cba + def
emp4	GA3P → PGA	abc → abc
emp5	PGA → PEP	abc → abc
emp6	PEP → PYR	abc → abc
ppp1	Glc6P → Ru5P + CO ₂	abcdef → bcdef + a
ppp2	Ru5P → Xu5P	abcde → abcde
ppp3	Ru5P → Rib5P	abcde → abcde
ppp4	Xu5P + Ery4P → GA3P + Fru6P	abcde + fghi → cde + abfghi
ppp5	Xu5P + Rib5P → Sed7P + GA3P	abcde + fghij → abfghij + cde
ppp6	GA3P + Sed7P → Ery4P + Fru6P	abc + defghij → ghij + defabc

Fonte: (SROUR et al., 2011)

Tabela 17 – Rede metabólica simplificada da *Pseudomonas sp.* LFM046 considerando somente a via ED cíclica. As letras minúsculas representam átomos sucessivos em cada metabólito. As abreviações são as mesma utilizadas na Figura 15.

Reação	Estequiometria	transição atômica
1	Glicose → G6P	abcdef → abcdef
2	G6P → F6P	abcdef → abcdef
3	F16P → F6P	abcdef → abcdef
4	G3P → F16P	abc + def → cbadef
5	G3P → PYR	abc → abc
6	G6P → G3P + PYR	abcdef → def + abc
7	PYR → AcCoA	abc → bc
8	AcCoA + AcCoA → PHA	ab + cd → abc

Fonte: (RIASCOS et al., 2013)

Tabela 18 – [Rede metabólica simplificada da *Pseudomonas sp.* LFM046 considerando a via das pentoses em equilíbrio. As letras minúsculas representam átomos sucessivos em cada metabólito. As abreviações são as mesma utilizadas na Figura 17.

Reação	Estequiometria	transição atômica
1	Glicose → G6P	abcdef → abcdef
2	G6P → F6P	abcdef → abcdef
3	G3P → PYR	abc → abc
4	G6P → G3P + PYR	abcdef → def + abc
5	PYR → AcCoA	abc → bc
6	AcCoA + AcCoA → PHA	ab + cd → abc
7	RIB5P → XuI5P	abcde → abcde
8	XuI5P + E4P → G3P + F6P	abcde + fghi → cde + abfghi
9	XuI5P + RIB5P → S7P + G3P	abcde + fghij → abfghij + cde
10	G3P + S7P → E4P + F6P	abc + defghij → ghij + defabc

Fonte: (RIASCOS et al., 2013)

Tabela 19 – Rede metabólica simplificada da *Pseudomonas putida*. As letras minúsculas representam átomos sucessivos em cada metabólito. As abreviações são as mesma utilizadas na Figura 19.

Reação	Estequiometria	transição atômica
1	Glucose $\rightarrow$ G6P	abcdef $\rightarrow$ abcdef
2	G6P $\rightarrow$ F6P	abcdef $\rightarrow$ abcdef
3	G3P $\rightarrow$ PYR	abc $\rightarrow$ abc
4	G6P $\rightarrow$ G3P + PYR	abcdef $\rightarrow$ def + abc
5	PYR $\rightarrow$ AcCoA	abc $\rightarrow$ bc
6	AcCoA + AcCoA $\rightarrow$ PHA	ab + cd $\rightarrow$ abc
7	P5P + E4P $\rightarrow$ G3P + F6P	abcde + fghi $\rightarrow$ cde + abfghi
8	P5P + P5P $\rightarrow$ S7P + G3P	abcde + fghij $\rightarrow$ abfghij + cde
9	G3P + S7P $\rightarrow$ E4P + F6P	abc + defghij $\rightarrow$ ghij + defabc
10	G6P $\rightarrow$ P5P	abcdef $\rightarrow$ bcdef

Fonte: (RIASCOS et al., 2013)