CARLOS ARTURO MARTINEZ RIASCOS

**TÉCNICAS DE PROGRAMAÇÃO MATEMÁTICA PARA A
ANÁLISE E PROJETO DE SISTEMAS BIOTECNOLÓGICOS**

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção do
título de Doutor em Engenharia.

São Paulo

2005

CARLOS ARTURO MARTINEZ RIASCOS

**TÉCNICAS DE PROGRAMAÇÃO MATEMÁTICA PARA A
ANÁLISE E PROJETO DE SISTEMAS BIOTECNOLÓGICOS**

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção do
título de Doutor em Engenharia.

Área de concentração:
Engenharia Química

Orientador:
Prof. Dr. José Maurício Pinto

São Paulo

2005

RIASCOS, Carlos Arturo Martinez

Técnicas de programação matemática para a análise e projeto de sistemas biotecnológicos – São Paulo, 2005.

167 p.

Tese (Doutorado) – Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Química.

1. Programação matemática, 2. Otimização global, 3. Otimização mista-inteira, 4. Quantificação de fluxos metabólicos, 5. síntese e projeto de plantas multiproduto.

I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Química. II t.

Dedico este trabalho:

A minha mãe, que apesar da distância está
sempre no meu coração, e a Bianca pelo
apoio, amor e compreensão.

AGRADECIMENTOS

Um agradecimento especial ao meu orientador Prof. Dr. José Maurício Pinto, pela confiança, incentivo e paciência, pela liberdade que me conferiu para elaborar este trabalho e, sobretudo, por sua amizade e exemplo profissional.

Também, ao Prof. Dr. Andreas K. Gombert, cujo aporte foi fundamental no presente trabalho.

À CAPES pela bolsa e ao CNPq pelo financiamento do estágio em Nova Iorque.

A Dona Vani, a Sr. Antônio e a toda família Bonani Sacchitelli pelo acolhimento sincero no seio da família. A Bianca, meu presente divino.

A Luis Alberto, meu irmão e amigo, e ao Jairo companheiros nas baladas.

A Rubens, Adriano, Alessandro, Denis, Kike, Jorge e Galo colegas no estudo, no futebol e no Rei, e amigos de coração.

Aos companheiros do Laboratório de Simulação e Controle de Processos: Maria Cristina, Francisco, Elsa, Sérgio, Maurício, Juliana, sempre alegres.

A Maria Lúcia e Fátima pelo trabalho alegre e eficiente na biblioteca. A Elisete, Graça e Alexandre, sempre atenciosos na secretaria.

A Carminha pelo trabalho incessante para o funcionamento do LSCP. A Silvia pelo esforço para o funcionamento da rede. E a todo o pessoal do Dep. de Engenharia Química que direta ou indiretamente colaboraram para o desenvolvimento e conclusão deste trabalho.

RESUMO

A complexidade de alguns sistemas biotecnológicos impossibilita seu estudo sem o uso de técnicas de programação matemática avançadas. A quantificação de fluxos metabólicos e a síntese e projeto ótimos de plantas multiproduto são problemas com esta característica, abordados na presente tese.

A quantificação de fluxos metabólicos empregando balanços de marcações é representada como um problema de otimização não-linear, o qual se resolve através da minimização da diferença entre as medidas experimentais e as predições do modelo da rede metabólica. Este problema surge da necessidade de se caracterizar o metabolismo mediante a estimação das velocidades das reações bioquímicas. O modelo matemático para problemas deste tipo é composto basicamente por balanços de metabólitos e de isótopos; os primeiros são lineares, enquanto os segundos introduzem não-linearidades ao problema e, neste trabalho, são modelados mediante uma modificação da técnica de matrizes de mapeamento de átomos.

Para quantificar os fluxos metabólicos considerando a existência de ótimos locais, desenvolveu-se um algoritmo *branch & bound* espacial, no qual a busca global é feita mediante a divisão da região de busca (*branching*) e a geração de seqüências de limites (*bounding*) que convergem para a solução global. Como estudo de caso, estimaram-se os fluxos no metabolismo central de *Saccharomyces cerevisiae*. Os resultados confirmam a existência de soluções locais e a necessidade de desenvolver uma estratégia de busca global; a solução global obtida apresenta semelhanças, nos fluxos centrais, com a melhor solução obtida por um algoritmo evolucionário.

As abordagens mais empregadas para resolver problemas de síntese e projeto de sistemas biotecnológicos multiproduto são a definição e dimensionamento seqüencial das operações unitárias, e a fixação dos parâmetros de dimensionamento e de estimação do tempo de operação (com valores obtidos em laboratório ou planta piloto); porém ambas abordagens fornecem soluções subótimas. Por outro lado, a solução simultânea da síntese e projeto de sistemas biotecnológicos multiproduto gera modelos misto-inteiros não-lineares (MINLP) de grande porte, devido à combinação das decisões, ligadas à existência de alternativas no processo, com as restrições não-lineares geradas dos modelos das operações.

Como estudo de caso considera-se uma planta para produção de insulina, vacina para hepatite B, ativador de plasminogênio tecidual (*tissue plasminogen activator*) e superóxido dismutase, mediante três hospedeiros diferentes: levedura (*S. cerevisiae*) com expressão extra ou intracelular, *Escherichia coli* e células de mamíferos. O projeto deve satisfazer a meta de produção para cada produto, minimizando os custos de capital e selecionando os hospedeiros, as operações e o arranjo dos equipamentos em cada estágio. Os resultados obtidos mostram que a formulação das decisões por abordagem *big-M* permite resolver o modelo MINLP gerado e que a consideração de múltiplos produtos com seqüências e condições de processamento diferentes gera grande ociosidade nos equipamentos e aumenta o custo total do projeto. Para o estudo de caso observou-se que a alocação de tanques intermediários tem um efeito limitado na diminuição do custo do projeto, porém a implementação simultânea da flexibilização do *scheduling*, do projeto de equipamentos auxiliares e tanques intermediários permite obter projetos satisfatórios.

ABSTRACT

The complexity of biotechnological systems does not allow their study without the use of advanced mathematical programming techniques. Metabolic flux quantification and optimal synthesis and design of multiproduct plants are problems with this characteristic, and are addressed in this thesis.

The metabolic flux quantification employing labeling balances is formulated as a nonlinear optimization problem that is solved by the minimization of the difference between experimental measurements and predictions of the metabolic network model. This problem is generated by the necessity of estimating the rates of biochemical reactions that characterize the metabolism. The mathematical model for this class of problems is composed by balances of metabolites and isotopes; the former are linear whereas the latter are nonlinear and, in this work, are modeled by a modification of the atom mapping matrix technique.

A spatial branch & bound algorithm was developed to quantify the metabolic fluxes, that considers the existence of local optima; in this algorithm, the global search is developed by the division of the searching region (branching) and the generation of sequences of bounds (bounding) that converge to the global solution. As a case study, fluxes in central metabolism of *Saccharomyces cerevisiae* were estimated. The results confirm the existence of local solutions and the necessity of develop a global search strategy; the central fluxes in the obtained global solution are similar to those ones obtained by an evolutionary algorithm.

To solve problems of synthesis and design of multiproduct biotechnological systems, the most employed approaches are the sequential selection and sizing of the unit operations, and the fixing of sizing and time parameters (employing values from laboratory or pilot plants); nevertheless, both approaches generate suboptimal solutions. On the other hand, the simultaneous solution of the synthesis and design of multiproduct biotechnological systems generates large size mixed-integer nonlinear models (MINLP), due to the combination of options into the processing with nonlinear constraints from the operation models.

As case study, a plant for production of insulin, hepatitis B vaccine, tissue plasminogen activator and superoxide dismutase was considered, by three hosts: yeast (*S. cerevisiae*) with extra or intracellular expression, *Escherichia coli* and mammalian cells. The design must satisfy the production target for each product, minimizing the capital cost and considering the selection of hosts, the operations and the number of parallel units in each stage. The obtained results show that the formulation of decisions by the big-M approach allows the solution of the generated MINLP model and that consideration of several products with different processing sequences and conditions generates large idleness at the equipment and increases the total cost of the design. In the case study it was observed that the allocation of storage tanks has a limited effect on cost reduction, but the simultaneous implementation of flexible scheduling, design of auxiliary equipments and intermediate storage tanks allow the generation of satisfactory designs.

SUMÁRIO

LISTA DE FIGURAS	vi
LISTA DE TABELAS	viii
CAPÍTULO 1 INTRODUÇÃO	1
1.1 MOTIVAÇÃO	1
1.1.1 O problema da quantificação de fluxos metabólicos	1
1.1.2 O problema de síntese e projeto ótimo de bioprocessos	3
1.2 OBJETIVOS	5
1.3 ESTRUTURA DA TESE	6
CAPÍTULO 2 QUANTIFICAÇÃO DE FLUXOS METABÓLICOS	7
2.1 ENGENHARIA METABÓLICA	7
2.2 ANÁLISE DE FLUXOS METABÓLICOS	10
2.2.1 Análise de fluxos metabólicos baseada em balanços de metabólitos	11
2.2.2 Análise de fluxos metabólicos em experimentos com substrato marcado	15
2.2.3 Cuidados especiais em experimentos de marcação	17
2.2.4 Informação obtida dos experimentos de marcação	18
2.2.5 Técnicas para medição dos padrões de marcação	20
2.2.6 Quantificação dos fluxos metabólicos em experimentos de marcação	23
2.3 ESTIMAÇÃO DE PARÂMETROS POR ABORDAGEM DE ERRO-NAS - VARIÁVEIS	24
2.4 OTIMIZAÇÃO DE PROBLEMAS NÃO-CONVEXOS	26
CAPÍTULO 3 MODELAGEM E OTIMIZAÇÃO DO PROBLEMA DE MFA	27
3.1 MODELAGEM MATEMÁTICA	27
3.1.1 Fluxos bidirecionais	27
3.1.2 Balanços de metabólitos	29
3.1.3 Balanços de marcações	30
3.1.4 Estimação das marcações fracionárias somadas (SFLs)	33
3.1.5 Função objetivo	33
3.2 ABORDAGEM DE OTIMIZAÇÃO GLOBAL PARA QUANTIFICAÇÃO DE FLUXOS METABÓLICOS	34
3.2.1 Aproximação convexa do problema	34
3.2.2 Estimação dos limites das variáveis	35
3.2.3 Algoritmo de busca	36
3.3 ESTUDO DE CASO	39
3.4 ANÁLISE DE RESULTADOS	40
3.4.1 Identificação de soluções locais	42
3.4.2 Otimização global	44
3.5 CONCLUSÕES	50
3.6 NOMENCLATURA NA QUANTIFICAÇÃO DE FLUXOS METABÓLICOS	50
CAPÍTULO 4 SÍNTESE E PROJETO DE BIOPROCESSOS	53
4.1 BREVE RESUMO DA EVOLUÇÃO DOS PROCESSOS DE BIOSÍNTESE	53

4.2 PURIFICAÇÃO DE BIOPRODUTOS	54
4.2.1 Remoção/coleta de células, restos celulares ou corpos de inclusão.....	55
4.2.2 Rompimento de células.....	57
4.2.3 Dissolução de corpos de inclusão e renaturação de proteínas.....	57
4.2.4 Concentração e purificação inicial.....	58
4.2.5 Purificação de alta resolução por cromatografia.....	59
4.2.6 Integração dos estágios de processamento.....	64
4.3 CARACTERÍSTICAS DE ALGUNS BIOPRODUTOS	66
4.3.1 Insulina humana	66
4.3.2 Vacina para hepatite B	67
4.3.3 Ativador de plasminogênio tecidual (tPA).....	67
4.3.4 Superóxido dismutase	68
4.4 AUMENTO DE ESCALA DOS BIOPROCESSOS	69
4.5 ABORDAGENS PARA SÍNTESE E PROJETO DE PROCESSOS	69
4.5.1 A otimização isolada das tarefas.....	70
4.5.2 Otimização da superestrutura do processo.....	71
4.6 SÍNTESE E PROJETO ÓTIMO DE PLANTAS BIOQUÍMICAS DESCONTÍNUAS	73
4.6.1 Modelos de desempenho	74
4.6.2 Projeto com equipamentos em paralelo e tanques intermediários	74
4.7 EXEMPLOS DE SÍNTESE E PROJETO ÓTIMO DE PLANTAS MULTIPRODUTO	75

CAPÍTULO 5 SOLUÇÃO DO PROBLEMA DE SÍNTESE E PROJETO DE BIOPROCESSOS MULTIPRODUTO

5.1 CARACTERÍSTICAS GERAIS DO PROBLEMA	78
5.2 MODELAGEM MATEMÁTICA	79
5.2.1 Arranjos de equipamentos em paralelo.....	80
5.2.2 Modelagem das decisões na síntese do processo.....	81
5.2.3 Modelagem do efeito das decisões no desempenho das unidades	83
5.2.4 Modelagem de estágios descontínuos	84
5.2.5 Modelagem de estágios semi-contínuos	85
5.2.6 Modelagem de estágios cromatográficos	87
5.2.7 Restrições para garantir o cumprimento das metas de produção.....	87
5.2.8 Função objetivo	87
5.2.9 Formulação do problema de síntese e projeto ótimo de bioprocessos por GDP.....	88
5.2.10 Redução das não-convexidades do problema	89
5.2.11 Alocação de tanques intermediários em posições predefinidas.....	90
5.2.12 Alocação de tanques intermediários em posições otimizadas	91
5.3 ESTUDO DE CASO.....	92
5.4 REFORMULAÇÃO DO PROBLEMA POR ABORDAGEM <i>BIG-M</i>	95
5.4.1 Modelo básico para a planta multiproduto.....	95
5.4.2 Modelagem da alocação de tanques intermediários em posições predefinidas	100
5.4.3 Modelagem da alocação de tanques intermediários em posições otimizadas	100
5.4.4 Considerações adicionais para a implementação do modelo.....	101
5.5 RESULTADOS	103

5.5.1 Resultados para o modelo de plantas monoproduto (cenário 1).....	103
5.5.2 Resultados para o modelo multiproduto sem tanques intermediários (cenário 2)	105
5.5.3 Resultados para o modelo multiproduto com tanques intermediários em posições predefinidas ..	108
5.5.4 Resultados para o modelo multiproduto com tanques intermediários em posições otimizadas	111
5.5.5 Esforço computacional	114
5.6 CONCLUSÕES	116
5.7 NOMENCLATURA DOS MODELOS DE SÍNTESE E PROJETO DE PROCESSOS	116
CAPÍTULO 6 SÍNTESE E PROJETO COM EQUIPAMENTOS AUXILIARES E CONSIDERAÇÕES NO <i>SCHEDULING</i>	119
6.1 CONSIDERAÇÕES ESPECIAIS NO <i>SCHEDULING</i>	120
6.2 PROJETO DE EQUIPAMENTOS AUXILIARES	122
6.3 RESULTADOS	124
6.3.1 Resultados para o modelo multiproduto com <i>scheduling</i> flexível (cenário 5)	125
6.3.2 Resultados para o modelo multiproduto com <i>scheduling</i> flexível e tanques intermediários ..	126
6.3.3 Resultados para o modelo multiproduto com biorreator auxiliar (cenário 7).....	127
6.3.4 Resultados para o modelo multiproduto com biorreator auxiliar e tanques intermediários	129
6.4 CONCLUSÕES	131
CAPÍTULO 7 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS	132
7.1 ASPECTOS GERAIS	132
7.2 SUMÁRIO DO TRABALHO	132
7.2.1 Análise de fluxos metabólicos	133
7.2.2 Síntese e projeto de bioprocessos.....	133
7.3 CONTRIBUIÇÕES DESTE TRABALHO	135
7.4 RECOMENDAÇÕES PARA TRABALHOS FUTUROS	136
APÊNDICE A MODELO PARA ANÁLISE DE FLUXOS METABÓLICOS	138
APÊNDICE B AVALIAÇÃO DO PARÂMETRO DE INCERTEZA	140
APÊNDICE C REPRESENTAÇÃO DE DECISÕES POR PROGRAMAÇÃO DISJUNTIVA GENERALIZADA	141
APÊNDICE D FORMULAÇÃO DE DISJUNÇÕES POR ABORDAGEM <i>BIG-M</i>	142
APÊNDICE E PARÂMETROS PARA O ESTUDO DE CASO DE SÍNTESE E PROJETO ÓTIMO	143
APÊNDICE F PROJETOS PARA O MODELO MONOPRODUTO	151
APÊNDICE G RESULTADOS COMPLEMENTARES PARA MODELOS MULTIPRODUTO ..	154
REFERÊNCIAS BIBLIOGRÁFICAS	161

LISTA DE FIGURAS

Figura 2.1	O ciclo da engenharia metabólica (Nielsen, 2001).....	10
Figura 2.2	Situações que impedem a aplicação de MFA estequiométrica.	14
Figura 2.3	Caracterização por isotopômeros de posição e massa.	19
Figura 2.4	Espectro de ^{13}C NMR característico para um composto C_3	21
Figura 2.5	Espectro de massa para glicina.	22
Figura 2.6	Princípio da MFA em experimentos de marcação.	24
Figura 3.1	Modelagem de fluxos bidirecionais.	28
Figura 3.2	Rede metabólica para ilustração da modelagem.	28
Figura 3.3	Fluxograma do algoritmo de busca global e ilustração da convergência.	38
Figura 3.4	Rede para o metabolismo central de <i>S. cerevisiae</i>	39
Figura 3.5	Comparação qualitativa dos resultados.	43
Figura 3.6	Distribuição de fluxos para o estudo de caso.	45
Figura 3.7	Comparação entre os fluxos calculados e os limites pre-estabelecido.	46
Figura 3.8	Árvore da busca <i>Branch & Bound</i>	48
Figura 4.1	Locais e número aproximado de proteínas em <i>E. coli</i>	55
Figura 4.2	Esquema da microfiltração e de um cartucho com várias membranas	56
Figura 4.3	Princípio do fracionamento de proteínas por filtração em gel.	61
Figura 4.4	Princípio da retenção e eluição na cromatografia de troca iônica	62
Figura 4.5	Princípio da cromatografia por afinidade.	64
Figura 4.6	Seqüências de estágios para purificação de proteínas.	65
Figura 4.7	Esquema do processo multiproduto analisado por Pinto <i>et al.</i> (2001).	76
Figura 5.1	Efeito do projeto de equipamentos em paralelo na produtividade.	80
Figura 5.2	Exemplo de seqüência de processamento com opções.	82

Figura 5.3	Configuração de um estágio semi-contínuo com unidades em paralelo.	86
Figura 5.4	Configuração geral da planta para o cenário 2.	106
Figura 5.5	Configuração geral da planta para o cenário 3.	109
Figura 5.6	Configuração geral da planta para o cenário 4.	113
Figura 6.1	Gráfico de Gantt para ilustração da disponibilidade dos equipamentos.....	119
Figura 6.2	Gráfico de Gantt para ilustração da flexibilização do <i>scheduling</i>	121
Figura 6.3	Gráfico de Gantt para ilustração da alocação de um biorreator auxiliar. ...	123
Figura 6.4	Configuração geral da planta para o cenário 5.	126
Figura 6.5	Configuração geral da planta para o cenário 6.	127
Figura 6.6	Configuração geral da planta para o cenário 7.	129
Figura 6.7	Configuração geral da planta para o cenário 8.	130

LISTA DE TABELAS

Tabela 3.1	Variáveis medidas na quantificação de fluxos metabólicos.	41
Tabela 3.2	Resultados obtidos com NLP local e algoritmo evolucionário.	42
Tabela 3.3	<i>SFLs</i> medidas e calculadas por algoritmo evolucionário e global.	49
Tabela 3.4	Fluxos medidos e calculados por algoritmo evolucionário e global.	49
Tabela 5.1	Definição das rotas de produção no exemplo da Figura 5.2.	83
Tabela 5.2	Produtos e hospedeiros para o estudo de caso.	93
Tabela 5.3	Estágios necessários para a produção, dependendo do hospedeiro.	94
Tabela 5.4	Características dos cenários resolvidos para o estudo de caso.	95
Tabela 5.5	Resultados gerais da síntese e projeto para o modelo monoproduto.	104
Tabela 5.6	Resultados gerais por produto para o cenários 1 e 2.	105
Tabela 5.7	Projeto para planta multiproduto sem tanques intermediários.	106
Tabela 5.8	Custos do projeto (US\$) para o cenário 2.	107
Tabela 5.9	Resultados gerais por produto para o cenário 3.	110
Tabela 5.10	Projeto para o cenário 3 e comparação da ociosidade.	110
Tabela 5.11	Custos do projeto (US\$) para o cenário 3.	111
Tabela 5.12	Projeto para o cenário 4 e comparação da ociosidade.	112
Tabela 5.13	Custos do projeto (US\$) para o cenário 4.	114
Tabela 5.14	Desempenho computacional para problemas de síntese e projeto.	115
Tabela 6.1	Características dos cenários resolvidos para o problema de síntese e projeto com considerações especiais no <i>scheduling</i> e unidades auxiliares.	125
Tabela 6.2	Resultados gerais por produto para os cenários 2 e 5.	125
Tabela 6.3	Resultados gerais por produto para os cenário 2, com tPA/CHO, e 7.	127
Tabela 6.4	Projeto e custos para o cenário 7.	128

Tabela 6.5	Projeto e custos para o cenário 8.	130
Tabela A.1	Metabólitos na rede para o metabolismo central de <i>S. cerevisiae</i>	138
Tabela A.2	Reações na rede para o metabolismo central de <i>S. cerevisiae</i>	139
Tabela B.1	Análise da eliminação de variáveis na MFA.	140
Tabela E.1	Rendimentos das operações.	143
Tabela E.2	Constantes para os parâmetros de dimensionamento e tempo.	144
Tabela E.3	Funções para o tempo de processamento em cada estágio (h).	147
Tabela E.4	Custo dos equipamentos.	149
Tabela E.5	Limites na dimensão dos equipamentos	150
Tabela F.1	Projetos para modelo monoproduto para produção de insulina (INS).	151
Tabela F.2	Projetos para modelo monoproduto para produção de vacina (HBV).	152
Tabela F.3	Projetos para modelo monoproduto para produção de tPA.	152
Tabela F.4	Projetos para modelo monoproduto para produção de SOD.	153
Tabela G.1	Resultados gerais por produto para o cenário 3 considerando $F = 5$	154
Tabela G.2	Projeto e custos para o cenário 3 considerando $F = 5$	155
Tabela G.3	Variáveis de dimensionamento para o cenário 4.	156
Tabela G.4	Projeto e custos para o cenário 5.	157
Tabela G.5	Projeto e custos para o cenário 6.	158
Tabela G.6	Variáveis de dimensionamento para o cenário 6.	159
Tabela G.7	Projeto e custos para o cenário 2 com expressão de tPA por CHO.	159
Tabela G.8	Variáveis de dimensionamento para o cenário 8.	160

CAPÍTULO 1 INTRODUÇÃO

1.1 Motivação

O desenvolvimento atual das ciências biológicas criou a necessidade de se empregar abordagens de programação matemática para resolver satisfatoriamente alguns dos seus problemas. A quantificação das velocidades das reações bioquímicas numa rede metabólica, e a síntese e projeto ótimos de processos biotecnológicos são exemplos de problemas que requerem a ação integrada das ciências biológicas e da engenharia de processos. A ação da engenharia de processos justifica-se pela capacidade, desenvolvida durante décadas, de lidar com as técnicas matemáticas e computacionais para a solução de problemas semelhantes na área dos processos químicos.

1.1.1 O problema da quantificação de fluxos metabólicos

A partir dos anos 70, o desenvolvimento da engenharia genética gerou uma revolução na biotecnologia; porém, as mudanças genéticas direcionadas não são possíveis sem análise prévia do microorganismo em estudo. Essas análises caracterizam o código genético (genoma); identificam as proteínas presentes (proteoma), e definem e quantificam as reações bioquímicas que acontecem na célula (fluxoma). O conhecimento de tais reações permite modelar os processos de transformação dos substratos em produtos, dentro da célula, e o seu conjunto é denominado rede metabólica.

A quantificação das velocidades das reações que definem uma rede, conhecida como análise de fluxos metabólicos (*Metabolic Flux Analysis* - MFA), é uma ferramenta importante na caracterização do metabolismo celular. Na MFA são evidenciadas quatro etapas: cultivo, medição, modelagem e quantificação dos fluxos. As condições de cultura dependem do sistema a ser estudado, assim como dos objetivos da análise e das técnicas de medição disponíveis; as medições incluem a concentração de biomassa e de

metabólitos extracelulares, a composição da biomassa, e o padrão de marcação gerado em metabólitos extra e intracelulares quando se emprega um substrato marcado¹. As medidas de marcação possibilitam a análise de fluxos em redes metabólicas complexas, mas aumentam a sofisticação dos experimentos e das técnicas matemáticas para a quantificação. As implicações do uso de medidas de marcação são abordadas na seção 2.2.

Quando a MFA inclui medidas de marcação, a modelagem matemática, que se baseia no conhecimento das reações bioquímicas, requer balanços de metabólitos e de isótopos. Os de metabólitos consideram as transformações que acontecem nas reações e geram equações lineares em função dos fluxos, enquanto os balanços de isótopos consideram a migração dos átomos de carbono na rede e introduzem equações não-lineares em função dos fluxos e dos padrões de marcação, o que impede a solução analítica do problema de quantificação; neste caso, os fluxos metabólicos devem ser estimados com estratégias de otimização, que minimizam a diferença entre as medições e as predições do modelo.

Alguns métodos de otimização empregados na quantificação de fluxos metabólicos são: algoritmos evolucionários (Gombert *et al.*, 2001), algoritmos genéticos (Zhao e Shimizu, 2003) ou programação quadrática sequencial (Wiechert *et al.*, 2001). Apesar de gerarem soluções satisfatórias, estes métodos não consideram a possibilidade de existirem múltiplas soluções para a rede metabólica², logo não garantem a obtenção da melhor solução possível (o ótimo global).

A abordagem *Branch & Bound* (B&B) espacial vem sendo empregada para a otimização global de problemas não-lineares (Quesada e Grossmann, 1995; Esposito e Floudas, 1998; Sahinidis e Tawarmalani, 2000; Alle e Pinto, 2004), mostrando a importância de se identificar a solução global quando se abordam problemas não-convexos.

¹ Que possui isótopos de carbono (¹³C ou ¹⁴C) em abundância diferente à natural e em posições definidas.

² A multiplicidade de soluções é devida às equações não-lineares do modelo.

No presente trabalho, desenvolveu-se um algoritmo de otimização global baseado na abordagem B&B espacial, para resolver o problema de quantificação de fluxos metabólicos. Como estudo de caso, abordou-se o metabolismo central de *Saccharomyces cerevisiae* com marcações medidas mediante cromatografia gasosa acoplada a espectrometria de massa (GC-MS). Os balanços de isótopos foram formulados na forma de balanços de marcações fracionárias, empregando-se uma generalização da técnica de matrizes de mapeamento de átomos (Zupke e Stephanopoulos, 1994).

1.1.2 O problema de síntese e projeto ótimo de bioprocessos

O problema de síntese do processo envolve a definição das operações necessárias para a transformação desejada, enquanto no projeto dimensionam-se os equipamentos e define-se, entre outras características, a alocação de arranjos com múltiplos equipamentos em paralelo e de tanques intermediários; em ambos problemas visa-se minimizar o investimento de capital, satisfazendo as demandas do mercado e os requerimentos de qualidade e pureza para os produtos.

Nas últimas décadas, a evolução das ferramentas computacionais permitiu aos engenheiros de processos considerar as plantas de produção como sistemas integrados e não mais como a soma de unidades independentes; por outro lado, a engenharia de processos biotecnológicos está começando a aplicar uma abordagem holística³ em seus processos produtivos. Neste sentido, Steffens *et al.* (2000) observam que os bioprocessos são projetados de forma sequencial, iniciando no biorreator até obter o produto com as especificação desejadas, e que embora esta abordagem gera processos economicamente viáveis, outras alternativas não consideradas podem ser mais lucrativas. Cabe ressaltar que as características específicas dos sistemas biotecnológicos (como diluição e impureza dos produtos, e a existência de incertezas geradas pela simplificação e desconhecimento das complexidades metabólicas) adicionam dificuldades tanto à síntese e projeto de plantas quanto à determinação da operação. Assim, é necessário o direcionamento de

³ A abordagem holística de processos supõe que cada estágio (operação unitária) afeta o desempenho das demais e conseqüentemente de todo o sistema.

esforços para o desenvolvimento de metodologias que abordem a síntese e projeto de processos bioquímicos, considerando o processo integrado em lugar de unidades individuais.

Bogle *et al.* (1996) consideram os problemas e desafios que surgem quando os processos bioquímicos são considerados no conjunto e avaliam as possibilidades de desenvolver abordagens gerais que permitam melhorar a eficiência dos processos, empregando as técnicas desenvolvidas na engenharia de processos. Quanto ao projeto de processos bioquímicos descontínuos, os autores argumentam que existem vários softwares com diferentes abordagens, mas a obtenção do projeto e das estratégias de operação ótimas para plantas multiproduto e multipropósito, especialmente quando se devem selecionar equipamentos entre várias opções existentes, é uma tarefa computacionalmente difícil pois existem complexidades que tornam insolúveis alguns problemas. Os autores argumentam que, embora o projeto dos processos baseie-se freqüentemente em modelos simplificados para cada etapa, é necessário empregar modelos mais complexos, já que existem parâmetros que não podem ser fixados previamente.

Esses trabalhos permitem inferir que a otimização do projeto para plantas biotecnológicas deve gerar ganhos quanto ao investimento de capital, especialmente para as plantas multiproduto. Além disso, acredita-se que a consideração dos processos biotecnológicos como sistemas e a implementação de modelos detalhados para cada operação permite a obtenção de melhores resultados, apesar de aumentar a complexidade e dificultar a resolução do problema de otimização.

No presente trabalho, para resolver simultaneamente o problema de síntese e projeto ótimo de plantas bioquímicas multiproduto, desenvolveram-se modelos de Programação Mista-Inteira Não-Linear (MINLP) nos quais as operações são modeladas com modelos de desempenho, cujas escolhas são formuladas na forma mista-inteira pela abordagem *big-M*. Como estudo de caso aborda-se a síntese e projeto de uma planta biotecnológica para produção de insulina, vacina para hepatite B, ativador de plasminogênio tecidual ou *tissue plasminogen activator* (tPA) e superóxido dismutase

(SOD), mediante três hospedeiros diferentes: levedura (*S. cerevisiae*), *Escherichia coli* e células de mamíferos: *chinese hamster ovary cells* (CHO). O processo requer até 15 estágios, dependendo dos hospedeiros escolhidos e em algumas etapas é necessário escolher entre várias opções existentes.

1.2 Objetivos

O objetivo central deste trabalho é desenvolver metodologias de programação matemática para a análise e projeto de sistemas biotecnológicos, especificamente os problemas de análise de fluxos metabólicos (MFA) e de síntese e projeto ótimo de processos.

No problema de quantificação de fluxos metabólicos, os objetivos específicos são:

- Estudar a modelagem matemática destes problemas, considerando as implicações das medidas de marcação.
- Avaliar a existência de soluções locais e as características dessas soluções.
- Analisar os efeitos de incertezas das medições na solução do problema.
- Desenvolver um algoritmo de otimização global para quantificação dos fluxos.
- Identificar as dificuldades na resolução global do problema e avaliar alternativas para contorná-las.
- Comparar a otimização global com outras abordagens empregadas na solução destes problemas, identificando as vantagens de cada uma, no intuito de estabelecer as potencialidades da metodologia desenvolvida.

No problema de síntese e projeto ótimo de bioprocessos, os objetivos específicos são:

- Avaliar os conceitos empregados na seleção das operações unitárias em cada etapa e de suas políticas de operação (contínua, semi-contínua ou descontínua) durante a síntese e projeto de processos biotecnológicos.
- Estudar a implementação de características especiais no projeto, como a definição de arranjos com múltiplos equipamentos que operam em-fase ou fora-de-fase e a

alocação de tanques intermediários, identificando suas vantagens, desvantagens e os efeitos no custo.

- Analisar a relação entre as decisões tomadas durante a síntese e o projeto de plantas biotecnológicas.
- Investigar e estudar as técnicas empregadas para a síntese e projeto de plantas químicas e bioquímicas, visando desenvolver uma metodologia eficiente para a otimização estrutural e operacional de bioprocessos.
- Aplicar a metodologia desenvolvida a processos biotecnológicos de interesse industrial, e analisar os resultados obtidos para determinar suas potencialidades e limitações.

1.3 Estrutura da tese

A presente tese está conformada por sete capítulos: o presente faz a introdução, estabelecendo a relevância dos problemas abordados e os objetivos do trabalho; o segundo contém uma revisão dos fundamentos da análise de fluxos metabólicos e das metodologias empregadas para quantificar os fluxos; o terceiro apresenta a modelagem deste problema, a metodologia de otimização global desenvolvida, a rede metabólica para *S. cerevisiae* considerada como estudo de caso, os resultados da quantificação de fluxos por otimização global e a comparação desta abordagem com um algoritmo evolucionário.

O quarto capítulo apresenta uma revisão na área de síntese e projeto de bioprocessos, caracterizando os estágios de processamento e as abordagens empregadas para resolver problemas deste tipo; no quinto capítulo define-se a metodologia empregada para resolver simultaneamente a síntese e o projeto de bioprocessos considerando uma planta multiproduto com linha única de produção, e apresentam-se os resultados obtidos. A partir da análise dos resultados do capítulo quinto, desenvolveram-se modificações na modelagem para considerar a alocação de linhas auxiliares de produção e a flexibilização do *scheduling* que permitam melhorar o projeto da planta; estas modificações e os resultados, com elas, obtidos são apresentados no sexto capítulo. Finalmente, as conclusões e sugestões para trabalhos futuros são apresentadas no sétimo capítulo.

CAPÍTULO 2 QUANTIFICAÇÃO DE FLUXOS METABÓLICOS

Este capítulo caracteriza o problema de quantificação de fluxos metabólicos com balanços de marcações, estabelecendo a sua relevância dentro da engenharia metabólica, e apresentando os fundamentos para a modelagem, algumas técnicas empregadas para a medição das marcações e a informação obtida com cada técnica.

2.1 Engenharia metabólica

Em 1973, a produção, pela primeira vez, de um gene clonado que permitiu a bem sucedida engenharia genética da *Escherichia coli* (Cohen *et al.*, 1973) abriu um novo caminho para melhorar os processos biotecnológicos existentes e para o desenvolvimento de novos processos: a engenharia metabólica, que emprega a engenharia genética (mudanças genéticas direcionadas, não aleatórias, através de técnicas de recombinação do DNA). Stephanopoulos *et al.* (1998) definem a engenharia metabólica como: “O melhoramento dirigido da geração de produtos ou das propriedades celulares, através da modificação de reações bioquímicas específicas ou da introdução de novas reações, com uso da tecnologia do DNA recombinante”.

Nielsen (2001) classifica as aplicações de engenharia metabólica em sete grupos:

- 1) Produção de proteínas heterólogas⁴; estas aplicações permitem a produção de proteínas para uso farmacêutico e enzimas, e são desenvolvidas mediante a inserção dos genes heterólogos no hospedeiro (microorganismo ou célula produtora) e a manipulação da via metabólica⁵ de produção para melhorar a produtividade ou permitir a secreção das proteínas.

⁴ Proteínas estranhas ao hospedeiro, *i.e.* que não são sintetizadas pelas cepas selvagens.

⁵ Seqüência de reações que ligam um grupo de substratos a um grupo de produtos.

- 2) Diversificação de substratos, aumentando a eficiência no uso da matéria prima, mediante a inserção da enzima ou via metabólica necessária para a utilização do novo substrato; em alguns casos é necessário manipular várias vias para obter uma taxa de metabolização do substrato razoável e prevenir a formação de produtos secundários indesejados.
- 3) Direcionamento de vias metabólicas para novos produtos, gerando hospedeiros que produzem várias substâncias diferentes; estas aplicações requerem o prolongamento de vias existentes e frequentemente são necessárias manipulações adicionais para melhorar a taxa de produção e prevenir a formação de subprodutos.
- 4) Degradação de xenobióticos⁶; na natureza existem muitos organismos que degradam estas substâncias, mas com a inserção de vias metabólicas ou mediante manipulação das existentes podem ser gerados organismos que degradem simultaneamente várias dessas substâncias.
- 5) Modificações na fisiologia celular para melhorar os processos produtivos, tais como aumentar a tolerância a baixas concentrações de oxigênio, reduzir a sensibilidade a altas concentrações de glicose ou melhorar a morfologia.
- 6) Eliminação ou redução da formação de subprodutos, visando prevenir a perda do carbono metabolizado, o efeito tóxico e a interferência destes subprodutos na purificação do produto desejado; às vezes a eliminação do gene envolvido é viável, mas se o subproduto for vital para o hospedeiro, é necessário analisar a rede metabólica completa.
- 7) Aumento do rendimento ou da produtividade, especialmente na produção de substâncias com baixo valor agregado; estes objetivos são atingidos adicionando cópias dos genes envolvidos ou com estratégias mais complexas, que surgem do estudo da rede metabólica.

⁶ Do grego “*xenos*” que significa “estranho”, são compostos orgânicos artificiais, que por sua estabilidade e solubilidade em compostos graxos, acumulam-se na cadeia alimentar.

Stephanopoulos *et al.* (1998) apresentam uma compilação de exemplos de aplicação de engenharia metabólica para diversos propósitos: aumento da produtividade na geração de etanol, aminoácidos e solventes; diversificação de substratos; obtenção de novos produtos como antibióticos, vitaminas, biopolímeros, pigmentos, hidrogênio, etc.; melhoramento das características celulares como metabolismo do nitrogênio, utilização do oxigênio, assimilação do substrato, estabilidade genética, etc.; e degradação de xenobióticos.

Uma característica comum para qualquer aplicação de engenharia metabólica é que a etapa de engenharia genética é precedida por uma análise que determina as modificações genéticas necessárias e seguida de uma nova análise que freqüentemente estabelece modificações genéticas adicionais. Deste modo, a engenharia metabólica de um microorganismo é a junção da análise do seu metabolismo e da engenharia genética do mesmo. Nielsen (2001) expõe que a engenharia metabólica é o aperfeiçoamento contínuo das características celulares, mediante várias aplicações de engenharia genética, e representa este processo como um ciclo (ver Figura 2.1) formado por três passos:

- 1) Análise: caracterização do metabolismo do microorganismo original; ou comparação entre o metabolismo do microorganismo modificado e original, para verificar os efeitos das mudanças genéticas.
- 2) Projeto: interpretação da análise para estabelecer as modificações genéticas necessárias.
- 3) Síntese: implementação das mudanças genéticas para melhorar as características do microorganismo, empregando biologia molecular e técnicas de recombinação de DNA.

Figura 2.1 O ciclo da engenharia metabólica (Nielsen, 2001).

2.2 Análise de fluxos metabólicos

Stephanopoulos *et al.* (1998) argumentam que os mapas metabólicos, obtidos da análise do genoma e do proteoma, apesar da sua complexidade, só apresentam as rotas possíveis para conversão de nutrientes em produtos e energia, mas não possuem informação a respeito dos fluxos reais de carbono, nitrogênio e energia através das vias metabólicas, nas condições particularmente estudadas. Como citado anteriormente, a quantificação das velocidades para as reações bioquímicas em uma rede metabólica, conhecida como análise de fluxos metabólicos (MFA), caracteriza o metabolismo celular; tal análise possibilita a comparação entre condições de cultura, e entre cepas. O resultado final da análise de fluxos metabólicos é o mapa quantificado de fluxos, que apresenta a rede das reações bioquímicas e a velocidade estimada para cada reação. Os autores apresentam diversas aplicações da análise de fluxos:

- 1) Identificação de nós de controle '*branch point control*': a rigidez ou flexibilidade dos nós das vias metabólicas é identificada com a avaliação dos efeitos das condições ambientais e das mutações genéticas. A identificação destes nós é fundamental na definição das mudanças a serem implementadas.
- 2) Identificação de vias alternativas: em alguns microorganismos, devido à existência de vias alternativas, a identificação do caminho pelo qual os produtos estão sendo gerados pode não ser trivial. A MFA permite identificar vias possíveis, aquelas que

satisfazem os balanços de material, e descartar as não viáveis, aquelas que não os satisfariam.

- 3) Estimação de fluxos não medidos: fluxos de interesse, mas difíceis de medir, como as taxas de produção de subprodutos, podem ser determinados empregando o modelo da rede metabólica e as medidas de outros fluxos.
- 4) Estimativa de rendimentos teóricos máximos: estes rendimentos estabelecem limites para o processo real e quantificam atividades metabólicas sob as condições de rendimento máximo, gerando parâmetros para estratégias de controle ótimo.

Independentemente dos objetivos visados, a análise de fluxos metabólicos pode ser baseada em balanços de metabólitos intracelulares, balanços de co-fatores⁷, experimentos de marcação isotópica ou combinações de várias metodologias.

2.2.1 Análise de fluxos metabólicos baseada em balanços de metabólitos

Para desenvolver a MFA estequiométrica, que emprega exclusivamente o conhecimento da estequiometria da rede metabólica e medidas da taxa de produção de metabólitos, deve-se assumir que o sistema encontra-se em estado estacionário ou pseudo-estacionário. Sob esta condição a concentração celular dos metabólitos intracelulares é constante. Dessa forma, para cada metabólito, a somatória dos fluxos de produção iguala a dos fluxos de consumo, gerando uma equação linear. Se alguns fluxos podem ser diretamente medidos, de modo a tornar o sistema observável⁸, a rede metabólica pode ser totalmente quantificada. Esse procedimento permitiu a primeira quantificação de uma rede metabólica central completa (Vallino e Stephanopoulos, 1993).

⁷ Co-fatores ou co-enzimas são espécies de baixo peso molecular (tais como: ATP, NADH, NADPH, etc.) cuja função de receber ou dar elétrons ou grupos funcionais as torna indispensáveis em muitas reações enzimáticas, especialmente as de geração de energia.

⁸ O número mínimo de fluxos independentes a serem medidos é definido pelo número de fluxos intracelulares (incógnitas) menos o número de balanços independentes (equações).

A seguir é apresentada a formulação do balanço de metabólitos exposta em Stephanopoulos *et al.* (1998). Considerando uma rede metabólica com N substratos, M produtos metabólicos, Q compostos que constituem a biomassa, K metabólitos intracelulares, a estequiometria para cada uma das $j = 1, \dots, J$ reações será:

$$\sum_{i=1}^N \mathbf{a}_{ji} S_i + \sum_{i=1}^M \mathbf{b}_{ji} P_i + \sum_{i=1}^Q \mathbf{g}_{ji} Z_{macro i} + \sum_{i=1}^K \mathbf{g}_{ji} Z_{met i} = 0 \quad j = 1, \dots, J \quad (2.1)$$

Em cada termo da equação (2.1), o fator da esquerda é o coeficiente estequiométrico para o composto i na reação j , negativo se o composto é consumido, positivo caso contrário, e o da direita representa um mol da substância (substrato, produto, constituinte da biomassa ou metabólito intracelular); este balanço permite prever a quantidade consumida ou gerada de uma substância (em moles), conhecendo o consumo ou geração de outra. O modelo estequiométrico da rede é formado por uma equação na forma de (2.1) para cada reação e pode-se escrever, de forma compacta, empregando notação matricial:

$$\underline{\underline{A}} \times \underline{\underline{S}} + \underline{\underline{B}} \times \underline{\underline{P}} + \underline{\underline{G}} \times \underline{\underline{Z}}_{macro} + \underline{\underline{G}} \times \underline{\underline{Z}}_{met} = \underline{\underline{0}} \quad (2.2)$$

onde $\underline{\underline{A}}$, $\underline{\underline{B}}$, $\underline{\underline{G}}$ e $\underline{\underline{G}}$ são as matrizes com os coeficientes estequiométricos dos substratos, produtos, constituintes da biomassa e metabólitos intracelulares, respectivamente, em cada uma das reações; $\underline{\underline{S}}$, $\underline{\underline{P}}$, $\underline{\underline{Z}}_{macro}$ e $\underline{\underline{Z}}_{met}$ são os vetores das substâncias.

A taxa de produção (ou consumo) de uma substância é estimada em função da velocidade da reação (v_j). Então, se o coeficiente estequiométrico do composto i na reação j é $\mathbf{b}_{ji}$, a sua taxa de produção é $\mathbf{b}_{ji} \cdot v_j$. As taxas líquidas de consumo de substratos, geração de produtos, macromoléculas e metabólitos intracelulares, considerando todas as reações, são:

$$r_{s,i} = \sum_{j=1}^J \mathbf{a}_{ji} \mathbf{n}_j \quad (2.3)$$

$$r_{p,i} = \sum_{j=1}^J \mathbf{b}_{ji} \mathbf{n}_j \quad (2.4)$$

$$r_{macra,i} = \sum_{j=1}^J g_{ji} \mathbf{n}_j \quad (2.5)$$

$$r_{met,i} = \sum_{j=1}^J g_{ji} \mathbf{n}_j \quad (2.6)$$

O balanço molar para compostos extracelulares (substratos e produtos) deve considerar a taxa líquida de produção ou consumo, assim como as taxas de alimentação e remoção do composto, se o cultivo não é em batelada. Para os constituintes da biomassa (macromoléculas) e para os metabólitos intracelulares deve-se considerar a taxa de produção e o efeito de diluição ocasionado pelo crescimento da biomassa. Na forma vetorial, o balanço para os metabólitos intracelulares é:

$$\frac{d\mathbf{X}_{met}}{dt} = \mathbf{r}_{met} - \mathbf{m}\mathbf{X}_{met} \quad (2.7)$$

onde $\mathbf{X}_{met}$ é o vetor de concentração dos metabólitos intracelulares, $\mathbf{r}_{met}$ é o vetor das taxas líquidas de produção e $\mathbf{m}$ é a taxa específica de crescimento de biomassa. Assumindo estado pseudo-estacionário e desconsiderando o termo de diluição, já que a concentração dos metabólitos intracelulares é muito pequena, obtém-se o seguinte balanço simplificado:

$$\mathbf{0} = \mathbf{r}_{met} = \mathbf{G}^T \mathbf{v} \quad (2.8)$$

onde $\mathbf{G}$ é a matriz estequiométrica para os metabólitos intracelulares, e $\mathbf{v}$ é o vetor das velocidades de reação. A equação vetorial (2.8) é o balanço de metabólitos, base para a análise de fluxos, que no item (3.1.2) é exemplificado para uma rede metabólica simples.

Apesar das múltiplas aplicações de MFA baseada em balanços de metabólitos intracelulares, esta abordagem tem limitações e não pode ser usada em qualquer problema. Wiechert (2001) aponta quatro situações que impedem a aplicação satisfatória de MFA estequiométrica:

- 1) Vias metabólicas paralelas, em que nenhum dos caminhos está associado com alguma variável que possa ser diretamente medida (Figura 2.2a).
- 2) Ciclos metabólicos não associados a fluxos metabólicos mensuráveis (Figura 2.2b).

- 3) Reações reversíveis, que são um caso especial de ciclos metabólicos (Figura 2.2c); esta situação é freqüente, já que só as reações fortemente irreversíveis, por motivos termodinâmicos, podem ser consideradas com direção única.
- 4) Caminhos associados sem balanços completos de co-fatores (Figura 2.2d). Para desenvolver os balanços de co-fatores, deve-se assumir que todas as reações que produzem ou consomem energia e todas as reações de transformação entre co-fatores são totalmente conhecidas. Estudos baseados em balanços de marcações concluem que tanto os balanços de co-fatores quanto as reações de transformação entre eles não são totalmente conhecidos (Marx *et al.*, 1996; Sauer e Bailey, 1999).

Christensen e Nielsen (2000) também expõem as limitações deste tipo de análise argumentando que se são empregados exclusivamente balanços de metabólitos intracelulares, a análise de fluxos metabólicos deve resolver um sistema de equações lineares, tornando este método matematicamente atraente, porém a informação contida nesses balanços não é suficiente para quantificar os fluxos em redes complexas, as quais incluem reações reversíveis, ou vias metabólicas paralelas ou cíclicas. Então, a análise é freqüentemente desenvolvida adicionando-se balanços de co-fatores, que requerem hipóteses que devem ser verificadas, como a demanda de co-fatores em cada reação.

Figura 2.2 Situações que impedem a aplicação de MFA estequiométrica.

2.2.2 Análise de fluxos metabólicos em experimentos com substrato marcado

Para este tipo de análise, o sistema biológico que se deseja analisar é alimentado com um substrato marcado⁹; devido à metabolização do substrato, os átomos de carbono marcados incorporam-se nos metabólitos intracelulares e às macromoléculas que constituem a biomassa. A incorporação desses átomos não é homogênea¹⁰ e depende da atividade das diferentes reações que compõem a rede metabólica; por isso, os padrões de marcação gerados são úteis para definir e quantificar a rede.

Os primeiros experimentos com substrato marcado para quantificação de fluxos metabólicos foram desenvolvidos com ¹⁴C, o qual é radiativo, (Katz e Wood 1963; Hanson e Rhodes, 1983; Katz, 1985). No final da década de 1980, começou-se a empregar ¹³C (Malloy *et al.*, 1987), e atualmente existem várias metodologias e inúmeras aplicações de ¹³C MFA.

Se o objetivo da análise é estimar a atividade relativa de algumas vias metabólicas, pode-se basear exclusivamente em balanços de marcações, que é conhecida como análise de razões de fluxos metabólicos “*metabolic flux ratio analysis*” (Sauer *et al.*, 1999). Finalmente, a combinação de balanços de metabólitos e de marcações permite a quantificação completa da rede metabólica; análises deste tipo para *Escherichia coli*, *Saccharomyces cerevisiae*, *Penicillium chrysogenum*, *Synechocystis* sp. e células de mamíferos são encontradas respectivamente em Zhao e Shimizu (2003), Christensen *et al.* (2002), Christensen e Nielsen (2000), Yang *et al.* (2002) e Jeffrey *et al.* (2002).

Na MFA são evidenciadas quatro etapas: cultura, medição, modelagem e estimação dos fluxos. Algumas características de cada etapa são:

- 1) Geralmente, os experimentos de marcação são desenvolvidos em culturas contínuas com glicose marcada como a única fonte de carbono, mas alguns estudos de MFA têm sido desenvolvidos em condições diferentes. Gombert *et al.* (2001) quantificaram os

⁹ Substrato contendo isótopos do carbono em proporções diferentes das naturais e em posições definidas.

¹⁰ A proporção de átomos marcados varia entre as posições atômicas em um metabólito e entre os diversos metabólitos.

fluxos metabólicos para *S. cerevisiae* sob condições de baixa e alta concentração de glicose; a primeira condição foi obtida em cultura contínua com taxa específica de crescimento de $0,1 \text{ h}^{-1}$, enquanto a segunda condição foi obtida em cultura descontínua com taxa específica de crescimento de $0,37 \text{ h}^{-1}$. Zhao e Shimizu (2003) e Zhang *et al.* (2003A) quantificaram os fluxos metabólicos em *E. coli* e *S. cerevisiae* para diferentes fontes de carbono (glicose, glicerol e acetato); nos resultados destes trabalhos observa-se que tanto a concentração quanto a natureza da fonte de carbono têm efeito importante na atividade das vias metabólicas. Jeffrey *et al.* (2002) empregam $[3-^{13}\text{C}]$ piruvato e uma mistura de $[2-^{13}\text{C}]$ acetato e $[1-^{13}\text{C}]$ glicose para a análise metabólica de células de mamíferos. Adicionalmente, as condições de cultura devem considerar as técnicas de medição a serem empregadas.

- 2) As medições incluem a concentração e composição da biomassa, a concentração de metabólitos extracelulares, e a distribuição do ^{13}C (marcação) em metabólitos intracelulares. Na medição de marcação podem ser empregadas técnicas de ressonância magnética nuclear (NMR) (Jeffrey *et al.*, 2002; Schmidt *et al.*, 1999A; Yang *et al.*, 2002) ou espectrometria de massa (MS) (Christensen e Nielsen, 2000; Christiansen *et al.*, 2002; Thykaer *et al.*, 2002), porém devido à baixa concentração dos metabólitos intracelulares, a medição de marcação tem sido feita indiretamente, medindo-a em aminoácidos e macromoléculas, e relacionando-a com a marcação nos metabólitos centrais. Só recentemente desenvolveu-se uma abordagem para medir a marcação diretamente nos metabólitos centrais (van Winden *et al.*, 2005); algumas vantagens que os autores atribuem a esta abordagem residem no fato de que permite estimar a marcação para metabólitos que não participam da síntese de macromoléculas e fornece informação sobre a reversibilidade de algumas reações.
- 3) A modelagem matemática baseia-se no conhecimento da estequiometria das reações bioquímicas para formular os balanços de metabólitos e de isótopos. Os primeiros consideram a transformação dos metabólitos nas reações, e os últimos consideram a migração dos átomos de carbono marcados na rede. A formulação matemática para a modelagem depende da técnica empregada na medição das marcações, e portanto só é introduzida no item 3.1.3.

- 4) Os balanços de isótopos introduzem não-linearidades no modelo matemático, já que a contribuição de cada reação na produção/consumo de cada isótopo depende da taxa de reação (fluxo metabólico) e da abundância dos isótopos envolvidos. Essas não-linearidades impedem a solução analítica da rede metabólica, mas como as medidas podem ser expressas em função dos fluxos, estes podem ser quantificados com um programa de otimização, minimizando-se a diferença entre as medições e as predições do modelo (ver item 2.2.6).

2.2.3 Cuidados especiais em experimentos de marcação

Wiechert (2001) enumera algumas características especiais deste tipo de análise:

- 1) Assim como na análise baseada em balanços de metabólitos, deve-se garantir que o sistema esteja em estado estacionário quanto ao acúmulo e consumo de metabólitos intracelulares, denominado pelo autor como estado estacionário metabólico.
- 2) A escolha do substrato marcado é fundamental para o sucesso da análise. No caso da fonte de carbono ser glicose, os substratos mais empregados são [1-¹³C]glicose e glicose totalmente marcada, com custos aproximados de US\$ 100 por grama; glicose com outra marcação pode ser empregada, mas seus custos são ainda maiores. Adicionalmente, misturas de [1-¹³C]glicose, glicose totalmente marcada e glicose não marcada podem ser empregadas. Outras fontes de carbono são empregadas em Zhao e Shimizu (2003), Zhang *et al.* (2003A) e Jeffrey *et al.* (2002)
- 3) O biorreator deve ter o menor tamanho possível, para reduzir o custo do experimento, mas é importante considerar que a diminuição do tamanho aumenta a dificuldade no controle das condições desejadas; volumes usuais estão entre 300 e 1000 mL. Christiansen *et al.* (2002) e Thykaer *et al.* (2002) quantificaram os fluxos metabólicos para *Bacillus clausii* e *Penicillium chrysogenum*, empregando volumes de trabalho de 220 mL e 180 mL, respectivamente. Recentemente, sistemas de *microtiter plates* têm sido empregados para culturas descontínuas, estes envolvem volumes de aproximadamente 1 mL e os resultados com eles obtidos reproduzem os gerados em biorreatores de 500 mL.

- 4) Além do estado estacionário metabólico, o sistema deve alcançar o estado estacionário isotópico; neste estado, o substrato marcado está totalmente incorporado aos metabólitos intracelulares e aos aminoácidos e macromoléculas empregados na medição das marcações. Wiechert (2001) argumenta que o tempo para a incorporação da marcação nos metabólitos intracelulares é da ordem do tempo de residência do biorreator, e que o tempo total para a incorporação nas macromoléculas é entre 2 e 4 vezes o tempo de residência. Gombert *et al.* (2001) avaliam a incorporação da marcação em *S. cerevisiae* em cultura contínua, concluindo que este fenômeno pode ser modelado com uma cinética de primeira ordem e que depois de quatro tempos de residência a marcação atinge 98,2% do seu valor no estado estacionário isotópico.
- 5) Depois do experimento, durante os processos de acondicionamento das amostras para medição da marcação (coleta da biomassa, ruptura das células, isolamento das macromoléculas, hidrólise das mesmas, etc.), deve-se ter cuidado especial para não destruir ou alterar as substâncias.

2.2.4 Informação obtida dos experimentos de marcação

MFA baseada em experimentos de marcação ou ^{13}C MFA (Wiechert, 2001) está ligada ao conceito de isotopômero (*isotopomer*). Este termo origina-se da combinação de isótopo e isômero, e emprega-se para designar cada um dos possíveis estados de marcação que um composto pode apresentar. A distribuição isotopomérica de um metabólito é definida pela percentagem de cada isotopômero.

A caracterização por isotopômeros de posição define o número de átomos de ^{13}C e as suas posições na molécula, determinando o padrão de marcação completo, enquanto a caracterização por isotopômeros de massa define o número de átomos de ^{13}C na molécula sem considerar as suas posições. Para uma molécula com n átomos de carbono (molécula C_n) existem 2^n isotopômeros de posição e $n+1$ de massa. A Figura 2.3 ilustra as duas caracterizações para uma molécula C_3 .

Os isotopômeros de posição podem ser denominados por uma sequência binária, na qual o *zero* significa que existe um ^{12}C na posição correspondente, e o *um*, que existe um

^{13}C ; ou podem ser enumerados transformando-se a sequência binária em um número inteiro, usando-se a seguinte relação:

$$p = \sum_{k=1}^n 2^{k-1} c_k \quad (2.9)$$

Figura 2.3 Caracterização por isotopômeros de posição e massa para uma molécula C_3 .

Em (2.9), c_k é um algarismo binário ($c_k = 1$ se há um ^{13}C na posição k). Denotando A_p a abundância do isotopômero de posição p , e a_q a abundância do isotopômero de massa q , podemos relacionar as duas caracterizações: $a_0 = A_0$, $a_1 = A_1 + A_2 + A_4$, $a_2 = A_3 + A_5 + A_6$, $a_3 = A_7$.

Os resultados de experimentos de marcação também podem ser interpretados na forma de marcações fracionárias (*fractional labelings*), como apresentado em Christensen e Nielsen (2000). A marcação fracionária para o metabólito i na posição k , l_{ik} , é definida como a fração de átomos de ^{13}C naquela posição¹¹. Para o exemplo da Figura 2.3, a

¹¹ Christensen e Nielsen (2000) denotam a marcação fracionária para metabólito x na posição i como $x(i)$.

relação entre marcações fracionárias e frações isotopoméricas é a seguinte¹²:

$$\begin{aligned} I_1 &= A_1 + A_3 + A_5 + A_7 \\ I_2 &= A_2 + A_3 + A_6 + A_7 \\ I_3 &= A_4 + A_5 + A_6 + A_7 \end{aligned} \quad (2.10)$$

Na equação 2.10, tanto as marcações fracionárias quanto as frações isotopoméricas possuem um único índice, que identifica a posição ou o número característico do isotopômero; para a formulação dos balanços de marcações é ainda necessário adicionar um outro índice para identificar o metabólito.

2.2.5 Técnicas para medição dos padrões de marcação

A ressonância magnética nuclear (NMR) e a espectrometria de massa (MS) são as bases das técnicas empregadas para medir marcações. ¹H ou próton NMR e ¹³C NMR requerem o isolamento dos diferentes compostos a serem analisados. Com ¹H NMR, cada átomo de carbono do composto gera um sinal diferente; então, comparando-se a intensidade dos sinais, pode-se obter a percentagem de moléculas marcadas em cada posição (marcação fracionária). Por outro lado, com ¹³C NMR o sinal gerado depende também do estado de marcação dos átomos de carbono adjacentes (ver Figura 2.4).

Partindo do espectro gerado para a mistura de isotopômeros de um metabólito, pode-se obter informação da distribuição de marcações. Uma combinação de ¹H e ¹³C NMR denominada espectrometria de ressonância magnética nuclear em duas dimensões (*two-dimensional NMR*) permite medir as marcações sem isolamento dos compostos, já que gera um espectro em duas dimensões para cada composto que pode ser analisado sem interferência dos outros (Schmidt *et al.*, 1999A).

¹² Diferente das frações isotopoméricas de posição e massa (A_p e a_q), a soma das marcações fracionárias não é obrigatoriamente igual à unidade.

Figura 2.4 Espectro de ^{13}C NMR característico para um composto C_3 (Schmidt *et al.*, 1999A). O carbono central (C_2) é o núcleo observado, o isotopômero com marcação em C_2 gera um sinal simples (*singlet* – S), os isotopômeros com marcação em C_1 e C_2 ou C_2 e C_3 geram sinais duplos (*doublets* – D1, D2) que as vezes não podem ser diferenciados, e o isotopômero totalmente marcado gera um duplo sinal duplo ou um sinal triplo (*doublet of doublets* ou *triplet* – DD, em ambos os casos). Uma mistura de isotopômeros gera um espectro de padrão múltiplo (P), e deste espectro pode se obter informação da distribuição de marcações, pois a intensidade relativa de cada sinal depende de abundância dos isotopômeros que a geram.

Diferente da NMR, a espectrometria de massa caracteriza as espécies existentes em uma amostra de acordo com seu peso molecular, gerando a distribuição dos isotopômeros de massa, e trabalha acoplada a uma técnica de separação, como cromatografia gasosa (GC) ou líquida (LC). Quando se emprega GC-MS, os compostos que eluem da coluna cromatográfica devem ser ionizados, o que causa a fragmentação das moléculas; conseqüentemente, o espectro gerado permite a caracterização do composto e dos seus fragmentos, os quais contêm diferentes partes da cadeia de átomos de carbono do composto original; assim, a distribuição dos seus isotopômeros de massa contém informação adicional.

No caso da glicina (ver Figura 2.5) geram-se dois íons: um com o carbono C_2 e outro com $\text{C}_1\text{-C}_2$; o íon molecular contendo os dois átomos de carbono da glicina (na direita do espectro) pode ter massa M, M+1 e M+2, enquanto o íon do fragmento contendo C_2 (na esquerda do espectro) pode ter massa M' e M'+1. Devido à abundância

natural dos isótopos dos átomos presentes nos fragmentos (^{15}N , ^{17}O , ^{18}O , ^{13}C , ^2H , etc.) é necessário corrigir os resultados do espectro de massa (Lee *et al.*, 1991 e 1992).

Figura 2.5 Espectro de massa para glicina com correções por marcações naturais (Christensen e Nielsen, 2000).

A intensidade de cada sinal, dividida pela soma das intensidades de todos os sinais do mesmo íon, define a abundância do isotopômero de massa correspondente. A relação entre os sinais, as frações isotopoméricas de massa dos íons e de posição da glicina é:

Íon do fragmento	Íon molecular	
$M': a_0^2 = A_{00} + A_{10}$	$M: a_0^{1-2} = A_{00}$	(2.11)
$M'+1: a_1^2 = A_{01} + A_{11}$	$M+1: a_1^{1-2} = A_{01} + A_{10}$	
	$M+2: a_2^{1-2} = A_{11}$	

onde o superscrito na fração isotopomérica de massa indica os átomos de carbono presentes no íon e os isotopômeros de posição são identificados com a notação binária (item 2.2.4). No exemplo da glicina, as duas marcações fracionárias podem ser estimadas; a soma das marcações fracionárias (denominada como Gly(1,2) em Christensen e Nielsen, 2000) é calculada a partir do íon molecular, e a marcação fracionária para C_2 ($I_{\text{Gly},2}$), a partir do íon do fragmento; a relação entre as marcações fracionárias e as frações isotopoméricas de massa dos íons é:

$$l_{\text{Gly},1} + l_{\text{Gly},2} = \frac{0 \cdot a_0^{1-2} + a_1^{1-2} + 2 \cdot a_2^{1-2}}{a_0^{1-2} + a_1^{1-2} + a_2^{1-2}} \quad (2.12)$$

$$l_{\text{Gly},2} = \frac{0 \cdot a_0^2 + a_1^2}{a_0^2 + a_1^2} \quad (2.13)$$

As frações isotopoméricas de massa obtidas por GC-MS permitem empregar uma formulação matemática baseada em marcações fracionárias em lugar de isotopômeros (Schmidt *et al.*, 1997), o que simplifica a modelagem e reduz o esforço matemático na quantificação dos fluxos metabólicos. Esta formulação é apresentada no item 3.1.3. Frequentemente, a fragmentação das espécies não é suficiente para determinar as marcações fracionárias individuais; então, a formulação deve empregar marcações fracionárias somadas, como apresentado no item 3.1.4.

2.2.6 Quantificação dos fluxos metabólicos em experimentos de marcação

O conhecimento completo da rede metabólica implica na identificação das vias metabólicas ativas e na quantificação dos fluxos nas reações que as constituem. Com este conhecimento, pode-se prever o padrão de marcação (na forma de isotopômeros de posição, de massa ou marcações fracionárias) dos metabólitos intracelulares e em consequência dos aminoácidos por eles gerados. Porém o objetivo da MFA é quantificar os fluxos na rede metabólica a partir das medidas de marcação e dos fluxos extracelulares, já que isto permite a caracterização do metabolismo, como exposto anteriormente. A relação entre as medidas de marcação e a distribuição de fluxos metabólicos é complexa e não-linear, tornando praticamente impossível a estimação analítica das velocidades das reações intracelulares em função das medidas, com exceção de redes metabólicas simples e sem interesse prático (Schmidt *et al.*, 1999A).

Devido a essa dificuldade, a quantificação dos fluxos metabólicos é resolvida minimizando-se a diferença entre os valores medidos e os preditos pelo modelo que simula o metabolismo celular (ver Figura 2.6). Esta abordagem é conhecida como formulação de erro-nas-variáveis, e é apresentada na seção seguinte.

Figura 2.6 Princípio da quantificação de fluxos metabólicos em experimentos de marcação.

Diferentes algoritmos têm sido empregados na quantificação de fluxos metabólicos, os quais podem ser agrupados em métodos estocásticos, como os algoritmos evolucionários (Gombert *et al.*, 2001) e genéticos (Zhao e Shimizu, 2003), e métodos determinísticos, como programação quadrática sequencial (Wiechert *et al.*, 2001). Apesar de gerar soluções satisfatórias, esses métodos não consideram a possível existência de múltiplas soluções para a rede metabólica, conseqüentemente, não garantem a obtenção da melhor solução possível (o ótimo global).

2.3 Estimação de parâmetros por abordagem de erro-nas-variáveis

Modelos matemáticos para simular fenômenos físicos são repetidamente usados na engenharia química para o projeto, controle e otimização de processos químicos. Frequentemente, devido à complexidade dos processos, os modelos aplicados são não-lineares aumentando a dificuldade na estimação dos parâmetros a partir dos dados experimentais disponíveis; uma das abordagens mais empregadas para estimá-los é a minimização do erro-nas-variáveis (*error-in-variables*). Esta formulação baseia-se no princípio da máxima verossimilhança (*maximum likelihood*) e nas seguintes hipóteses (Esposito e Floudas, 1998):

- O erro gerado na medição, diferença entre o valor real e medido, pode ser descrito por uma distribuição normal com média zero e covariância conhecida.
- O erro em cada experimento é independente.

Considerando que o processo em questão seja modelado mediante um sistema de equações algébricas não-lineares:

$$\underline{f}(\underline{q}, \underline{x}) = 0 \quad (2.14)$$

onde $\underline{q}$ é o vetor dos parâmetros desconhecidos, e $\underline{x}$ é o vetor das variáveis medidas experimentalmente. Simplificando a abordagem para um único experimento, e considerando que os valores medidos ($\bar{x}$) e preditos pelo modelo ($\hat{x}$), os quais devem aproximar os valores reais, são relacionados pelo erro ($\bar{x} = \hat{x} + \underline{e}$), a abordagem gera o seguinte problema de otimização:

$$\begin{aligned} \min_{\hat{x}, \underline{q}} \sum_{i=1}^m \frac{(\hat{x}_i - \bar{x}_i)^2}{\underline{s}_i^2} \\ \text{sujeito a} \quad \underline{f}(\hat{x}, \underline{q}) = 0 \end{aligned} \quad (2.15)$$

onde $\underline{s}_i$ é o desvio padrão da variável i . Algumas propriedades matemáticas desta formulação são:

- A função objetivo é convexa, já os valores medidos permanecem fixos e os preditos são variáveis de otimização.
- O modelo do processo é incorporado como um conjunto de restrições e portanto modelos não-lineares geram problemas de otimização não-convexos.
- A otimização é desenvolvida em função dos parâmetros e dos valores ajustados para as variáveis; logo, modelos que são lineares em abordagens de simulação, nas quais os valores dos parâmetros permanecem fixos, podem gerar problemas não-lineares de estimação dos parâmetros.

Estas propriedades devem ser consideradas no desenvolvimento de uma estratégia para a solução do problema, particularmente no caso em que procura-se a solução global do mesmo.

2.4 Otimização de problemas não-convexos

Em geral, na solução de problemas de otimização não-convexos há duas linhas de pesquisa. A primeira agrupa os trabalhos baseados em metaheurísticas (métodos de busca estocástica), os quais procuram uma solução viável que atinja um critério de qualidade desejado e possa ser obtida em tempo razoável. A segunda linha de pesquisa é formada por trabalhos baseados em abordagens determinísticas, que empregam programação matemática e métodos de otimização. Estas abordagens procuram uma solução que além de ser viável seja a melhor dentro de uma vizinhança da região de busca, no caso da otimização local, ou na região inteira, no caso da otimização global.

Uma compilação de técnicas e aplicações de otimização global por abordagens determinísticas é apresentada em Grossmann (1996). Essas técnicas podem ser classificadas em:

- Métodos primais-duais (Floudas e Visweswaran, 1990), os quais empregam a decomposição de Benders generalizada (*Generalized Benders Decomposition* - GBD).
- Métodos de intervalo (Ratschek e Rokne, 1991), que aplicam técnicas matemáticas de intervalo para gerar os limites na função objetivo e restringir a região viável.
- Métodos *Branch and Bound* (B&B) (Quesada e Grossmann, 1995; Esposito e Floudas, 1998; Sahinidis e Tawarmalani, 2000; Alle e Pinto, 2004), que empregam aproximações convexas do problema original para gerar o limite inferior do ótimo global enquanto o limite superior é a melhor solução viável conhecida. Estes métodos são conhecidos como *branch and bound* espacial porque a convergência dos limites para o ótimo global é propiciada pela divisão do espaço de busca em sub-regiões. Esta é a abordagem escolhida para resolver o problema de quantificação de fluxos metabólicos, estando melhor detalhada na seção 3.2.

CAPÍTULO 3 MODELAGEM E OTIMIZAÇÃO DO PROBLEMA DE MFA

Neste capítulo apresenta-se de forma detalhada a modelagem e a solução do problema de análise de fluxos metabólicos por otimização global. O estudo de caso considera medidas de marcação por GC-MS e balanços de marcações fracionárias, o que gera um problema de otimização não-linear.

A seção 3.1 apresenta a modelagem e inclui a formulação dos balanços de metabólitos e marcações, das restrições para relacionar as marcações medidas com as modeladas e da função objetivo. As seções 3.2 e 3.3 apresentam a abordagem de otimização global empregada e o estudo de caso considerado. As seções 3.4 e 3.5 apresentam os resultados e as conclusões. Finalmente, a seção 3.6 apresenta a nomenclatura empregada na formulação e solução deste problema.

3.1 Modelagem matemática

A modelagem matemática do problema de quantificação de fluxos metabólicos gera um problema de otimização não-linear (NLP). O qual é formado pela função objetivo, os balanços de metabólitos (restrições lineares), os balanços de marcações na forma de marcações fracionárias (restrições bi-lineares), as funções para estimação das marcações fracionárias somadas (SFLs) a partir das marcações individuais (restrições lineares) e os limites das variáveis e dos parâmetros.

3.1.1 Fluxos bidirecionais

Para os fluxos bidirecionais emprega-se a modelagem apresentada em Wiechert e de Graaf (1997), onde cada fluxo é considerado como a associação do fluxo líquido (*net flux*), que descreve a conversão de metabólitos, e o fluxo de troca de átomos (*exchange flux*), que modela o intercâmbio de átomos entre os metabólitos envolvidos na reação, mas não contribui para a conversão líquida dos metabólitos. A Figura 3.1 ilustra a modelagem de uma reação bidirecional simples.

Figura 3.1 Modelagem de fluxos bidirecionais. As letras minúsculas representam os átomos de carbono. As magnitudes dos fluxos são: $v_{\text{net}} = v^{\rightarrow} - v^{\leftarrow}$; $v_{\text{exch}} = \min\{v^{\rightarrow}, -v^{\leftarrow}\}$.

Para ilustrar a formulação dos balanços, emprega-se a rede metabólica apresentada na Figura 3.2. Esta rede contém metabólitos com até 3 átomos de carbono e apresenta vias paralelas de geração dos produtos, o que impossibilita o cálculo dos fluxos metabólicos com balanços só de metabólitos. Porém, a inversão na ordem dos átomos de carbono na reação v_3 (Figura 3.2b) permite a quantificação completa da rede, empregando-se juntamente os balanços de marcações e de metabólitos.

Figura 3.2 Rede metabólica para ilustração da modelagem. O fluxo reversível (v_3) é modelado como a associação dos fluxos líquido e de troca.

3.1.2 Balanços de metabólitos

Os balanços de metabólitos consideram a contribuição de cada reação da rede metabólica para o acúmulo de cada metabólito intracelular, o qual é igual a zero na situação de estado estacionário metabólico (ver item 2.2.1). Para a rede da Figura 3.2, os balanços de metabólitos são:

$$\begin{aligned}
 \text{Met. A: } v_1 - v_2 - v_{3\text{net}} &= 0 \\
 \text{Met. B: } v_{3\text{net}} - v_4 - v_5 &= 0 \\
 \text{Met. C: } v_2 + v_5 - v_7 &= 0 \\
 \text{Met. D: } v_2 + v_5 - v_6 &= 0
 \end{aligned} \tag{3.1}$$

onde v_j são os fluxos metabólicos (velocidades de reação) em unidades de mol/g cel. h. A forma matricial dos balanços de metabólitos (Stephanopoulos *et al.*, 1998) é a seguinte (ver item 2.2.1):

$$\underline{\underline{G}}^T \cdot \underline{v} = 0 \tag{3.2}$$

onde $\underline{\underline{G}}$ é a matriz estequiométrica e $\underline{v}$ é o vetor dos fluxos. A matriz $\underline{\underline{G}}$ contém os coeficientes estequiométricos dos metabólitos intracelulares nas reações da rede. Para a rede metabólica da Figura 3.2, $\underline{\underline{G}}$ e $\underline{v}$ são:

$$\underline{\underline{G}} = \begin{matrix} & \begin{matrix} \text{A} & \text{B} & \text{C} & \text{D} \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix} & \begin{matrix} v_1 \\ v_2 \\ v_{3\text{net}} \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{matrix} \end{matrix} \quad \underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_{3\text{net}} \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} \tag{3.3}$$

O balanço de metabólitos para este exemplo de ilustração gera um sistema de quatro equações (3.1) com três incógnitas (v_2 , $v_{3\text{net}}$ e v_5), já que os outros fluxos são mensuráveis. Porém dos balanços de C e D observamos que $v_6 = v_7$, então a medição de ambos fluxos é redundante; com a manipulação matemática dos balanços, pode-se chegar

a $v_1 - v_7 = v_4$, o que nos indica que uma destas três medidas também é redundante; como só duas medidas são independentes (v_1, v_4 ou v_1, v_6 ou v_1, v_7 ou v_4, v_6 ou v_4, v_7) os fluxos não podem ser estimados apenas com balanços de metabólitos.

3.1.3 Balanços de marcações

Quando a modelagem é baseada em balanços de isotopômeros, deve-se formular um balanço para cada um dos 2^n isotopômeros possíveis para cada metabólito (onde n é o número de átomos de carbono); a contribuição de cada reação depende da sua velocidade (fluxo) e da abundância dos isotopômeros envolvidos. Para ilustração, apresentam-se os balanços de isotopômeros para o metabólito B da Figura 3.2.

$$\begin{aligned}
B_0 = B_{000}: v_{3net} \cdot A_{A,0} + v_{3exch} \cdot A_{A,0} - (v_4 + v_{3exch} + v_5) \cdot A_{B,0} &= 0 \\
B_1 = B_{100}: v_{3net} \cdot A_{A,4} + v_{3exch} \cdot A_{A,4} - (v_4 + v_{3exch} + v_5) \cdot A_{B,1} &= 0 \\
B_2 = B_{010}: v_{3net} \cdot A_{A,2} + v_{3exch} \cdot A_{A,2} - (v_4 + v_{3exch} + v_5) \cdot A_{B,2} &= 0 \\
B_3 = B_{110}: v_{3net} \cdot A_{A,6} + v_{3exch} \cdot A_{A,6} - (v_4 + v_{3exch} + v_5) \cdot A_{B,3} &= 0 \\
B_4 = B_{001}: v_{3net} \cdot A_{A,1} + v_{3exch} \cdot A_{A,1} - (v_4 + v_{3exch} + v_5) \cdot A_{B,4} &= 0 \\
B_5 = B_{101}: v_{3net} \cdot A_{A,5} + v_{3exch} \cdot A_{A,5} - (v_4 + v_{3exch} + v_5) \cdot A_{B,5} &= 0 \\
B_6 = B_{011}: v_{3net} \cdot A_{A,3} + v_{3exch} \cdot A_{A,3} - (v_4 + v_{3exch} + v_5) \cdot A_{B,6} &= 0 \\
B_7 = B_{111}: v_{3net} \cdot A_{A,7} + v_{3exch} \cdot A_{A,7} - (v_4 + v_{3exch} + v_5) \cdot A_{B,7} &= 0
\end{aligned} \tag{3.4}$$

onde $A_{i,p}$ é a abundância para o isotopômero de posição p do metabólito i . No caso de metabólitos sintetizados pela condensação de moléculas, a contribuição de uma reação é o produto do fluxo vezes a abundância de cada um dos isotopômeros envolvidos.

Diferente da abordagem anterior, quando a modelagem é baseada em marcações fracionárias é necessário formular um balanço de átomos marcados para cada uma das n posições em cada metabólito, o que reduz drasticamente o número de restrições no problema de quantificação de fluxos. A contribuição de cada reação é o produto do fluxo vezes a marcação fracionária do átomo envolvido; dessa forma, todos os termos das restrições serão bi-lineares. Para a rede metabólica da Figura 3.2 os balanços na forma de marcações fracionárias são:

$$\begin{aligned}
\text{Átomo A,1: } v_1 \cdot l_{S,1} + v_{3\text{exch}} \cdot l_{B,3} - (v_2 + v_{3\text{net}} + v_{3\text{exch}}) \cdot l_{A,1} &= 0 \\
\text{Átomo A,2: } v_1 \cdot l_{S,2} + v_{3\text{exch}} \cdot l_{B,2} - (v_2 + v_{3\text{net}} + v_{3\text{exch}}) \cdot l_{A,2} &= 0 \\
\text{Átomo A,3: } v_1 \cdot l_{S,3} + v_{3\text{exch}} \cdot l_{B,1} - (v_2 + v_{3\text{net}} + v_{3\text{exch}}) \cdot l_{A,3} &= 0 \\
\text{Átomo B,1: } v_{3\text{net}} \cdot l_{A,3} + v_{3\text{exch}} \cdot l_{A,3} - (v_4 + v_{3\text{exch}} + v_5) \cdot l_{B,1} &= 0 \\
\text{Átomo B,2: } v_{3\text{net}} \cdot l_{A,2} + v_{3\text{exch}} \cdot l_{A,2} - (v_4 + v_{3\text{exch}} + v_5) \cdot l_{B,2} &= 0 \\
\text{Átomo B,3: } v_{3\text{net}} \cdot l_{A,1} + v_{3\text{exch}} \cdot l_{A,1} - (v_4 + v_{3\text{exch}} + v_5) \cdot l_{B,3} &= 0 \\
\text{Átomo C,1: } v_2 \cdot l_{A,1} + v_5 \cdot l_{B,1} - v_7 \cdot l_{C,1} &= 0 \\
\text{Átomo C,2: } v_2 \cdot l_{A,2} + v_5 \cdot l_{B,2} - v_7 \cdot l_{C,2} &= 0 \\
\text{Átomo D,1: } v_2 \cdot l_{A,3} + v_5 \cdot l_{B,3} - v_6 \cdot l_{D,1} &= 0
\end{aligned} \tag{3.5}$$

onde $l_{i,k}$ é a marcação fracionária do átomo de carbono k no metabólito i .

Empregando-se a formulação com matrizes de mapeamento de átomos (AMM), desenvolvida em Zupke e Stephanopoulos (1994), os balanços de marcações da equação (3.5) são formulados como:

$$\begin{aligned}
[S>A]_1 \cdot \underline{S} \cdot v_1 + [B>A]_{3\text{exch}} \cdot \underline{B} \cdot v_{3\text{exch}} + (g_{2,A} \cdot v_2 + g_{3\text{net},A} \cdot v_{3\text{net}} + g_{3\text{exch},A} \cdot v_{3\text{exch}}) \cdot \underline{A} &= 0 \\
[A>B]_{3\text{net}} \cdot \underline{A} \cdot v_2 + [A>B]_{3\text{exch}} \cdot \underline{A} \cdot v_{3\text{exch}} + (g_{4,B} \cdot v_4 + g_{3\text{exch},B} \cdot v_{3\text{exch}} + g_{5,B} \cdot v_5) \cdot \underline{B} &= 0 \\
[A>C]_2 \cdot \underline{A} \cdot v_2 + [B>C]_5 \cdot \underline{B} \cdot v_5 + g_{7,C} \cdot \underline{C} \cdot v_7 &= 0 \\
[A>D]_2 \cdot \underline{A} \cdot v_2 + [B>D]_5 \cdot \underline{B} \cdot v_5 + g_{6,D} \cdot \underline{D} \cdot v_6 &= 0
\end{aligned} \tag{3.6}$$

onde $[i>ip]_j$ é a matriz que descreve a transferência dos átomos do metabólito i para o metabólito ip na reação j ; $\underline{S}$, $\underline{A}$, $\underline{B}$, $\underline{C}$ e $\underline{D}$ são os vetores de marcação nos metabólitos. $g_{j,i}$ é o coeficiente estequiométrico do metabólito i na reação j . Cabe observar que todos os coeficientes considerados são negativos, já que fazem parte dos termos que consideram as reações de consumo dos metabólitos. Para a rede da Figura 3.2, as matrizes de mapeamento de átomos são:

$$\begin{aligned}
& \begin{matrix} & S_1 & S_2 & S_3 \\ [S>A]_1 = & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} \end{matrix} \\
& [A>B]_{3\text{net}} = [A>B]_{3\text{exch}} = [B>A]_{3\text{exch}} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\
& [A>C]_2 = [B>C]_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\
& [A>D]_2 = [B>D]_5 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}
\end{aligned} \tag{3.7}$$

A matriz $[S>A]_1$ indica que na reação v_1 o carbono 1 do substrato (S_1) gera o carbono 1 do metabólito A (A_1), S_2 gera A_2 e S_3 gera A_3 . Todas as matrizes relacionam os átomos dos metabólitos consumidos e produzidos de forma análoga. Os vetores de marcação para a rede metabólica estão formados pelas marcações fracionárias:

$$\underline{S} = \begin{bmatrix} l_{S,1} \\ l_{S,2} \\ l_{S,3} \end{bmatrix} \quad \underline{A} = \begin{bmatrix} l_{A,1} \\ l_{A,2} \\ l_{A,3} \end{bmatrix} \quad \underline{B} = \begin{bmatrix} l_{B,1} \\ l_{B,2} \\ l_{B,3} \end{bmatrix} \quad \underline{C} = \begin{bmatrix} l_{C,1} \\ l_{C,2} \end{bmatrix} \quad \underline{D} = [l_{D,1}] \tag{3.8}$$

Agrupando as matrizes de mapeamento em um arranjo de matrizes de mapeamento de átomos (AMMA), os balanços de marcações podem ser escritos como:

$$\sum_{j \in J_{ip}} \sum_{i \in IS_j} AMMA_{j,i,p,i} \cdot MMM_i \cdot v_j + \sum_{j \in J_{ip}^*} g_{j,ip} \cdot MMM_{ip} \cdot v_j = 0 \quad \forall ip, k \in K_{ip} \tag{3.9}$$

Na equação (3.9), o balanço para o carbono k do metabólito ip considera as contribuições de todos os metabólitos $i \in IS_j$ em cada uma das reações $j \in J_{ip}$, onde J_{ip} é o conjunto de reações que geram ip , IS_j é o conjunto de metabólitos que participam como substratos na reação j , e J_{ip}^* é o conjunto de reações que consomem ip . No arranjo, quando se definem a reação (j) e os metabólitos (ip e i), caracteriza-se uma matriz de mapeamento ($AMMA_{j,ip,i} = [i>ip]_j$). MMM é a matriz de marcação nos metabólitos onde MMM_{ip} caracteriza o vetor de marcação para o metabólito ip .

3.1.4 Estimação das marcações fracionárias somadas (SFLs)

Quando se emprega GC-MS, obtêm-se as marcações fracionárias somadas (*SFLs*) para alguns metabólitos e para os fragmentos deles gerados, enquanto os balanços de marcação empregam as marcações fracionárias individuais ($l_{i,k}$). A *SFL* para um fragmento é igual à soma das marcações fracionárias dos átomos de carbono nele presentes (Gombert *et al.*, 2001). Assim, as marcações fracionárias somadas e individuais são relacionadas por:

$$SFL_f = \sum_{i \in I_f} \sum_{k \in K_{i,f}} l_{i,k} \quad \forall f \in F \quad (3.10)$$

onde F é o conjunto de fragmentos medidos, I_f é o conjunto de metabólitos que geram o fragmento f , e $K_{i,f}$ é o conjunto de átomos do metabólito i , que são incorporados no fragmento f .

3.1.5 Função objetivo

Para quantificar os fluxos metabólicos em experimentos com marcação, minimiza-se o erro-nas-variáveis (ver item 2.3), onde as variáveis são os fluxos metabólicos medidos e as marcações calculadas a partir dos espectros de NMR ou MS. Para ponderar a contribuição individual de cada variável na função objetivo, Zhao e Shimizu (2003) e Schmidt *et al.* (1999A) empregam uma estimativa do erro absoluto nos valores medidos, obtida com a comparação de medidas redundantes ou fazendo múltiplas medições; enquanto Gombert *et al.* (2001) empregam o desvio padrão estimado a partir de medidas independentes de no mínimo seis amostras diferentes.

No presente trabalho, os fluxos medidos experimentalmente e as marcações fracionárias somadas são empregados como variáveis, e a ponderação do erro é feita com o desvio padrão (s); assim, a função objetivo é:

$$\min \sum_{j \in JM} \left(\frac{\hat{v}_j - \bar{v}_j}{s_j} \right)^2 + \sum_{f \in F} \left(\frac{\hat{SFL}_f - \bar{SFL}_f}{s_f} \right)^2 \quad (3.11)$$

onde $\hat{v}_j$, $\hat{SFL}_f$ e $\bar{v}_j$, $\bar{SFL}_f$ são os valores estimados e medidos para os fluxos e as *SFLs* respectivamente, e JM , F são os conjuntos das variáveis medidas.

3.2 Abordagem de otimização global para quantificação de fluxos metabólicos

Na otimização global do problema de quantificação de fluxos metabólicos empregou-se uma abordagem *Branch and Bound* espacial (Esposito e Floudas, 1998). Nessa abordagem, a busca global é desenvolvida mediante a geração de seqüências convergentes de limites inferiores, ou *Lower Bounds (LB)*, e superiores, ou *Upper Bounds (UB)*, da solução global. Para obter os limites da solução é necessário formular uma relaxação convexa do problema, e a convergência destes limites é promovida pela divisão sucessiva da região de busca (*branching*) que permite melhorar os limites da solução (*bounding*). Visto que a MFA é modelada como um problema de minimização, o *UB* é a menor das soluções para o problema original não-convexo, e o *LB* é a menor das soluções para a aproximação convexa do problema:

$$UB = \min_s(ub_s) \quad (3.12)$$

$$LB = \min_s(lb_s) \quad (3.13)$$

Em (3.13) e (3.12) lb_s e ub_s são as soluções para o problema convexo e não-convexo em cada sub-região s .

3.2.1 Aproximação convexa do problema

No modelo apresentado, a função objetivo é não-linear, porém convexa. Então, só é necessário linearizar os balanços de marcações para obter a aproximação convexa do problema; as não-linearidades existentes são os termos bi-lineares gerados pelo produto de fluxos e marcações ($v_j \cdot l_{i,k}$). McCormick (1976) sugere substituir cada termo bi-linear por uma variável auxiliar, $vl_{j,i,k}$ para o presente problema, e adicionar as seguintes restrições:

$$\begin{aligned}
v_j^L \cdot l_{i,k} + l_{i,k}^L \cdot v_j - v_j^L \cdot l_{i,k}^L - v l_{j,i,k} &\leq 0 \\
v_j^U \cdot l_{i,k} + l_{i,k}^U \cdot v_j - v_j^U \cdot l_{i,k}^U - v l_{j,i,k} &\leq 0 \\
-v_j^U \cdot l_{i,k}^L - l_{i,k}^L \cdot v_j + v_j^U \cdot l_{i,k}^L + v l_{j,i,k} &\leq 0 \\
-v_j^L \cdot l_{i,k}^U - l_{i,k}^U \cdot v_j + v_j^L \cdot l_{i,k}^U + v l_{j,i,k} &\leq 0
\end{aligned} \tag{3.14}$$

onde os sobrescritos L e U designam os limites inferiores e superiores para as variáveis.

3.2.2 Estimação dos limites das variáveis

Devido à estrutura dos estimadores de McCormick (Restrições (3.14)), os limites para as variáveis (medidas) e para os parâmetros (não conhecidos) são fundamentais na convergência ao ótimo global. Para o problema de quantificação de fluxos metabólicos, os limites das variáveis são obtidos dos seus valores medidos $(\bar{v}_j, \bar{SFL}_m)$ e de um parâmetro ($\mathbf{d}$) que traduz a incerteza das medições:

$$v_j^L = (1 - \mathbf{d}) \cdot \bar{v}_j \quad v_j^U = (1 + \mathbf{d}) \cdot \bar{v}_j \quad \forall j \in JM \tag{3.15}$$

$$SFL_f^L = (1 - \mathbf{d}) \cdot \bar{SFL}_f \quad SFL_f^U = (1 + \mathbf{d}) \cdot \bar{SFL}_f \quad \forall f \in F \tag{3.16}$$

Para os parâmetros ($\mathbf{q}$), que são os fluxos não medidos ($v_j, j \notin JM$) e as marcações fracionárias ($l_{i,k}$), o conhecimento do problema permite gerar um conjunto de limites iniciais que, durante a otimização global, devem ser ajustados. Os limites iniciais das marcações são a marcação natural do carbono 1,1% e a marcação máxima teórica 100%, e para o fluxos estimam-se valores iniciais a partir do conhecimento da rede metabólica.

O ajuste dos limites é feito através de uma série de problemas min/max ($\mathbf{q}^L = \min \mathbf{q}, \mathbf{q}^U = \max \mathbf{q}$), considerando as restrições do problema convexo. Estes problemas estimam os limites como os valores extremos que satisfazem o modelo convexo. Para melhorar o ajuste dos limites adicionaram-se as seguintes restrições:

$$\sum_{j \in JM} \left(\frac{\hat{v}_j - \bar{v}_j}{s_j} \right)^2 - \mathbf{I}_v \cdot UB \leq 0 \tag{3.17}$$

$$\sum_{f \in F} \left(\frac{S\hat{F}L_f - S\bar{F}L_f}{s_f} \right)^2 - I_{SFL} \cdot UB \leq 0 \quad (3.18)$$

onde $?_v$, $?_{SFL}$ são as parcelas máximas que cada grupo de variáveis pode ter na função objetivo; assim, as restrições (3.17) e (3.18) limitam a contribuição de cada grupo de variáveis.

3.2.3 Algoritmo de busca

Em cada iteração da busca global, a sub-região geradora do LB é escolhida para o seguinte *branching*, que é feito por bissecção. Existem várias alternativas para selecionar a variável de divisão (Esposito e Floudas, 1998; Smith e Partelides, 1999; Quesada e Grossmann, 1995). Neste trabalho opta-se pela variável com a maior diferença entre seus valores nas soluções convexa e não-convexa. O algoritmo de busca global é descrito a seguir e é apresentado na forma de fluxograma na Figura 3.3; adicionalmente, esta figura ilustra uma iteração da busca $B\&B$, para um problema em que a função objetivo (FO) depende de uma única variável (x).

Passo 1 Definição de limites iniciais e de parâmetros do algoritmo.

- 1.1 Definir a tolerância relativa na convergência global, $?$.
- 1.2 Definir as contribuições máximas na função objetivo, $?_v$ e $?_{SFL}$.
- 1.3 Definir limites ajustados para as variáveis (v_j medidos e $SFLs$).
- 1.4 Definir limites iniciais para os parâmetros (v_j não-medidos e $l_{j,k}$).

Passo 2 Ajuste dos limites iniciais.

- 2.1 Resolver o problema original não-convexo a partir de diferentes pontos iniciais e escolher a melhor solução como limite superior inicial da solução global (UB^1). Este valor é empregado na geração das restrições adicionais para o ajuste dos limites (equações 3.17 e 3.18).
- 2.2 Resolver o problema convexo a partir da solução escolhida no passo 2.1.
- 2.3 Ajustar os limites dos parâmetros, *i.e.* solucionar o problema min/max para cada parâmetro, atualizando cada valor já ajustado.

Passo 3 Determinação do limite inferior para a solução global.

- 3.1 Resolver o problema convexo empregando os limites determinados no passo 2, para gerar o limite inferior da solução global (LB^1).

Passo 4 Verificação da convergência.

- 4.1 Se $(1 - \epsilon) \cdot UB = LB$, terminar a busca; *i.e.* o ótimo global (OG) é o atual UB .
Caso contrário, continue a busca global; *i.e.* vá para o passo 5.

Passo 5 *Branching*; caso seja a primeira divisão da região de busca, a mesma será a definida pelos limites estabelecidos no Passo 2; caso contrário, será a selecionada no passo 7.

- 5.1 Escolher a variável para dividir a região de busca, variável com a maior diferença entre seus valores para as soluções convexas e não-convexas.
5.2 Dividir a região de busca por bisseção na variável selecionada.

Passo 6 *Bounding*; avaliação das sub-regiões geradas (ambas devem ser avaliadas).

- 6.1 Iniciar o contador de sub-regiões avaliadas, $CSA = 0$.
6.2 Resolver o problema convexo para a sub-região atual, estimando lb .
6.3 Se a sub-região é um nó aberto, *i.e.* $(1 - \epsilon) \cdot UB > lb$, deve ser avaliada.
Caso contrário, vá para o passo 6.7.
6.4 Ajustar os limites dos parâmetros e das variáveis.
6.5 Resolver o problema convexo com os limites ajustados e gravar a solução na lista de nós abertos.
6.6 Resolver o problema não-convexo a partir da solução convexa. Se a solução obtida (ub) é menor que o atual UB , atualiza-se o UB . Vá para o passo 6.8.
6.7 Se existe solução viável para o problema convexo e $(1 - \epsilon) \cdot UB = lb$, grave a solução na lista de nós fechados.
De outro modo, descarte a sub-região.
6.8 Incrementar o contador de sub-regiões avaliadas, $CSA = CSA + 1$.
6.9 Se as duas sub-regiões novas foram avaliadas, *i.e.* $CSA = 2$, vá para o passo 7.
Caso contrário, considere a segunda sub-região e vá para o passo 6.2.

Passo 7 Seleção da região para *branching*.

- 7.1 Selecionar a região com a menor solução convexa ($\min lb$).
7.2 Apagar a região da lista de nós abertos e atualizar o limite inferior ($LB = \min lb$).
7.3 Voltar para o passo 4.

É importante ressaltar que o procedimento de *branch & bound* só termina quando são eliminados todos os *nós abertos*; uma sub-região do espaço de busca é considerada um *nó aberto* se a sua solução convexa (lb) é menor que o valor de convergência, $lb < (1 - \xi) UB$, onde ξ é a tolerância percentual para convergência global.

Figura 3.3 Fluxograma do algoritmo de busca global e ilustração da convergência.

3.3 Estudo de caso

A modelagem apresentada na seção 3.1 e a abordagem de otimização global da seção 3.2 são empregadas para quantificar os fluxos no metabolismo central de carbono em *Saccharomyces cerevisiae*, em cultivo contínuo em aerobiose, com velocidade específica de crescimento de $0,1 \text{ h}^{-1}$.

Baseando-se na rede descrita por Gombert *et al.* (2001), o metabolismo central de carbono em *S. cerevisiae* inclui: a) as vias metabólicas centrais (Embden-Meyerhoff-Parnas, EMP; pentoses-fosfato, PP, e o ciclo dos ácidos tricarboxílicos, TCA); b) algumas reações para formação de aminoácidos; c) compartimentação do piruvato, do acetil-CoA e do oxaloacetato. A Figura 3.4 apresenta a rede metabólica e os valores para os fluxos medidos, enquanto a descrição completa do estudo de caso, incluindo a migração dos átomos de carbono entre os metabólitos em cada reação, é apresentada no Apêndice A.

Figura 3.4 Rede para o metabolismo central de *S. cerevisiae*, em cultivo contínuo aeróbio, com $D = 0,1 \text{ h}^{-1}$. Os fluxos medidos referem-se a 100 unidades arbitrárias de glicose consumidas como única fonte de carbono.

Gombert *et al.* (2001) justificam a compartimentação dos metabólitos com base no proteoma de *S. cerevisiae*, já que as enzimas ativas identificadas no citosol e na mitocôndria são diferentes, assim como os aminoácidos gerados a partir dos precursores presentes nesses dois compartimentos celulares. Por exemplo, acredita-se que a valina é gerada do piruvato mitocondrial, enquanto a origem da alanina é incerta, pois as enzimas necessárias para sua formação encontram-se tanto no citosol quanto na mitocôndria. Adicionalmente, tendo avaliado a rede metabólica com e sem compartimentação para o piruvato, os autores verificaram que se obtém melhor ajuste quando essa característica é considerada. Os resultados apresentados nesse trabalho de Gombert *et al.* (2001) consideram a compartimentação de piruvato, acetil-CoA e oxaloacetato, ainda que o modelo apresenta o piruvato como um único metabólito.

Para este problema, Christensen *et al.* (2002) calcularam desvios padrão nos fluxos metabólicos centrais (vias PP e EMP, e ciclo TCA) inferiores a 3%, a partir dos quais os autores concluem que a estimação destes fluxos é bastante consistente. Os autores também estabelecem limites para todos os fluxos da rede, baseados em soluções obtidas por um algoritmo evolucionário para vários grupos independentes de medidas, e argumentam que a análise de fluxos não é capaz de resolver o problema da formação de acetil-CoA mitocondrial, *i.e.* definir a fração formada através da reação catalisada pela enzima piruvato desidrogenase (pdh) e através do desvio da pdh (as reações que formam a rede metabólica são apresentadas na Tabela A1, do Apêndice A). Isto se deve ao fato de que em nenhuma das vias há reordenamento dos átomos de carbono, o que é um requisito para a quantificação de fluxos paralelos mediante experimentos de marcação (conforme ilustrado na modelagem, Figura 3.2).

3.4 Análise de resultados

A modelagem matemática e o algoritmo de otimização foram implementados na plataforma GAMS (Brooke *et al.*, 2000), empregando o *solver* CONOPT2 (Drud, 1996) baseado no método do gradiente reduzido generalizado. Os resultados apresentados nesta seção foram obtidos em um PC Pentium 4, 2 GHz e 256 MB de memória RAM.

Os fluxos obtidos estão relacionados a uma incorporação de glicose de 100 (unidades arbitrárias). Na função objetivo foram empregadas 24 marcações (*SFLs*) e 18 fluxos metabólicos correspondentes à incorporação de intermediários nas macromoléculas e à excreção de produtos (ver Tabela A1). Para os limites iniciais das variáveis não medidas, empregam-se os seguintes valores $0 < v_{exch} < 1500$, $0 < v_{net} < 150$, $1,1 < l_{i,k} < 100$. A Tabela 3.1 apresenta o valor medido e o desvio padrão para cada variável; adicionalmente, para os fragmentos apresenta-se a origem dos átomos, o que permite relacionar as *SFLs* com as marcações fracionárias individuais (ver item 3.1.4).

Tabela 3.1 Variáveis medidas na quantificação de fluxos metabólicos.

Marcações fracionárias somadas				Fluxos metabólicos		
Fragmento	Origem	Valor medido	Desvio	Fluxo	Valor medido	Desvio
Glc331	G6P(1-6)	91,0	1,0	Incorporação de intermediários		
Gly175	GLY(1,2)	5,9	1,0	G6P →	22,20	2,2
Gly144	GLY(1,2)	6,2	1,0	P5P →	2,86	0,3
Gly85	GLY(2)	3,7	1,0	E4P →	2,50	0,3
Ser175	SER(1,2)	3,1	1,0	G3P →	0,82	0,1
Ser132	SER(2,3)	33,8	1,0	PEP →	4,99	0,5
Ala116	PYRcit(2,3)	36,0	1,0	PYRmit →	12,40	1,3
Ala99	PYRcit(2,3)	36,0	1,0	PYRcit →	4,07	0,5
Ala158	PYRcit(1-3)	38,5	1,0	OAAcit →	5,3	0,4
Leu158	PYRmit(2,2,3,3) + AcCoAmit(2)	106,1	1,0	AKG →	9,62	1,0
				AcCoAcit →	22,0	2,2
Val144	PYRmit(2,2,3,3)	73,3	1,0	AcCoAmit →	2,8	0,3
Val143	PYRmit(1,2)	6,9	1,0	SER →	1,81	0,2
Val127	PYRmit(2,2,3,3)	73,1	1,0	GLY →	3,66	0,4
Val186	PYRmit(1,2,2,3,3)	74,9	1,0	C-1 →	2,94	0,5
Asp188	OAAcit(2,3,4)	57,1	1,0	THR →	3,63	0,4
Asp115	OAAcit(2)	12,0	1,0	Excreção de produtos		
Asp216	OAAcit(1-4)	64,4	1,0	ETH →	0,0	0,01
Ile158	OAAcit(2-4) + PYRmit(2,3)	93,7	1,0	ACE →	0,0	0,01
Glu143	AKG(1,2)	39,9	1,0	GLYC →	0,0	0,01
Glu230	AKG(1-5)	100,1	1,0			
Pro142	AKG(2-5)	85,5	1,0			
Lys156	AKG(2-5) + AcCoAcit(2)	117,2	1,0			
Phe192	E4P(1-4) + PEP(2,2,3,3)	90,6	3,0			
Phe143	PEP(1,2)	3,0	1,0			

3.4.1 Identificação de soluções locais

Inicialmente, para confirmar a existência de soluções locais, o problema original (não-convexo) foi resolvido localmente, com o *solver* CONOPT2, a partir de diferentes pontos iniciais gerados randomicamente. Para os limites dos parâmetros empregaram-se os valores iniciais, sem ajuste, e para verificar o efeito dos limites das variáveis consideraram-se vários valores do parâmetro de incerteza (d).

Os resultados da busca local são apresentados na Tabela 3.2 junto com a melhor solução viável obtida por um algoritmo evolucionário. Para $d=0,1$ só uma solução local foi obtida; porém, uma variável adotou um dos seus valores-limite, sugerindo que soluções melhores (com menor valor para a função objetivo) poderiam ser obtidas se a região de busca fosse ampliada. Outra alternativa avaliada foi desconsiderar a(s) variável(is) que adotam valores-limite na solução, porém esta alternativa foi descartada, conforme mostrado no Apêndice B. Considerando $d=0,2$ obtiveram-se dois ótimos locais (OL), os quais correspondem a soluções diferentes para a rede metabólica e não apresentam variáveis nos seus valores-limite. Estes resultados confirmam a existência de soluções locais e sustentam a necessidade de desenvolver uma estratégia de busca global e de empregar um parâmetro de incerteza adequado.

Tabela 3.2 Resultados obtidos com NLP local e algoritmo evolucionário.

	Algoritmo NLP local			Algoritmo evolucionário (Gombert <i>et al.</i>)
	$d = 0,1$	$d = 0,2$		
Função Objetivo	23,06	22,74	201,85	26,31
Soma do erro quadrático nos fluxos	0,91	0,86	1,02	1,70
Soma do erro quadrático nas marcações	22,15	21,88	200,83	24,61
Fluxos metabólicos centrais				
Fluxo EMP	34,5	34,4	47,2	33,6
Fluxo PP	43,8	43,7	29,9	44,2
Fluxo TCA	62,4	62,1	64,2	59,9
Via pdh	65,2	65,0	67,1	39,9
Desvio da pdh	22,3	22,3	22,4	45,8
AcCoAcit ? AcCoAmit	0,0	0,0	0,0	22,8

A Figura 3.5 apresenta uma comparação qualitativa das soluções locais obtidas com programação não-linear, para $d=0,2$, e com o algoritmo evolucionário (Gombert *et al.*, 2001). A rede da Figura 3.5a corresponde à melhor solução obtida através do NLP (FO = 22,74), a Figura 3.5b é a solução com maior valor para a função objetivo (FO = 201,85) e a 3.5c é a solução obtida com abordagem evolucionária.

Figura 3.5 Comparação qualitativa dos resultados obtidos na quantificação de fluxos metabólicos: a) FO = 22,74, b) FO = 201,85, c) solução por algoritmo evolucionário. As linhas pontilhadas representam a compartimentação entre o citosol (acima) e a mitocôndria (abaixo). A largura das setas é proporcional à magnitude dos fluxos. As setas sombreadas são os fluxos com diferenças significativas entre as soluções.

Os fluxos metabólicos centrais (EMP, PP e ciclo TCA) para a melhor solução local e para a solução obtida por algoritmo evolucionário são semelhantes, ainda que pequenas diferenças são observadas. Por outro lado, a solução com maior função objetivo apresenta diferenças marcantes nos fluxos centrais, caracterizando uma distribuição realmente diferente para o carbono na rede metabólica. Nas vias para produção de acetil-CoA citosólico e mitocondrial são observadas diferenças importantes: as soluções obtidas com NLP local sugerem taxa de produção no citosol, desvio da pdh, inferior à da solução obtida com algoritmo evolucionário, e transferência de AcCoA para a mitocôndria nula. Estudos anteriores (Chistensen *et al.*, 2002) concluíram que há bastante incerteza na estimação dos fluxos nas vias para produção de AcCoA, este aspecto será novamente abordado na análise dos resultados da otimização global.

3.4.2 Otimização global

Para implementar o algoritmo de otimização global devem ser definidos os valores dos parâmetros de otimização e outras características específicas do algoritmo; para isso foram necessárias várias tentativas que não são apresentadas individualmente, apesar da importância no aprimoramento da metodologia. As características do algoritmo e os valores definitivos para os parâmetros são apresentados a seguir:

- O parâmetro de incerteza para estimar os limites nas variáveis medidas foi $d=0,2$. Valores menores para este parâmetro foram avaliados, porém, os resultados não foram considerados satisfatórios (ver item 3.4.1).
- Nas restrições adicionais para ajuste dos limites das variáveis (equações 3.16 e 3.17), as contribuições máximas, dos fluxos metabólicos e das marcações, na função objetivo foram $\gamma_v=0,2$ e $\gamma_{SL}=1,0$. Estes valores melhoram o procedimento de ajuste dos limites e eliminam soluções consideradas *inconsistentes* devido às grandes diferenças entre os fluxos medidos e estimados.
- Na convergência global, a tolerância foi de 10% ($\epsilon=0,1$). Este valor garante que não existe nenhuma solução com função objetivo menor do que $0,9 \cdot OG$, onde OG é o valor da função objetivo na solução considerada como o ótimo global.
- As variáveis escolhidas para a divisão da região de busca (*branching*) são os fluxos metabólicos líquidos não medidos e as marcações fracionárias individuais ($l_{j,k}$). Esta escolha baseia-se em três observações: 1) os limites nas variáveis não medidas (fluxos não medidos e marcações individuais) são os maiores geradores de divergência entre o problema convexo e o não-convexo; 2) a magnitude dos fluxos de troca tem efeito pequeno na quantificação da rede, e 3) após várias tentativas, verificou-se que este grupo de variáveis é o que melhor conduz à convergência.

Os valores para os parâmetros são particulares para o estudo de caso abordado, porém servem como estimativa inicial em outros problemas; enquanto, acredita-se que algumas características do algoritmo, como a escolha das variáveis para a divisão da região de busca, podem ser mantidas em outros estudos de caso. Para a convergência global foram necessárias múltiplas divisões da região de busca, as quais sempre incluíram

o procedimento de ajuste dos limites das variáveis para cada nova sub-região, como apresentado no algoritmo (item 3.2.3). No total, foram necessárias 132 divisões e 10,4 horas de tempo computacional para atingir a convergência.

A solução global obtida corresponde à melhor solução local da Tabela 3.2 e é apresentada na Figura 3.6, junto à solução obtida através da abordagem evolucionária. Para cada fluxo, o primeiro valor (de cima para baixo) é o gerado pelo algoritmo de busca global, o segundo corresponde à solução evolucionária, enquanto que o terceiro (quando disponível) é o valor medido; para os fluxos reversíveis, os valores dos fluxos de troca encontram-se entre parênteses.

Figura 3.6 Distribuição de fluxos para o estudo de caso. O valor de cima é o gerado com algoritmo de busca global, o segundo corresponde à solução evolucionária, enquanto o terceiro (quando disponível) é o valor medido. Para os fluxos reversíveis, o valor dos fluxos de troca encontra-se entre parênteses.

Os fluxos metabólicos centrais são semelhantes para as duas abordagens, embora o erro quadrático seja menor na solução obtida mediante otimização global (ver Tabela 3.2), enquanto os fluxos estimados para o transporte de oxaloacetato através da membrana mitocondrial e para as vias de geração do acetil-CoA são visivelmente diferentes. Os valores estimados, nesta pesquisa, para alguns fluxos estão fora dos limites anteriormente estabelecidos por Christensen *et al.* (2002), como pode ser observado na Figura 3.7.

Figura 3.7 Comparação entre os fluxos calculados com otimização global e os valores mínimos e máximos estabelecidos em Christensen *et al.* (2002). Todos os valores apresentados são para fluxos líquidos.

Na quantificação dos fluxos para formação de acetil-CoA (via enzima pdh e desvio da pdh), os átomos de carbono do acetil-CoA têm a mesma origem no piruvato citosólico,

o que não contribui para a quantificação. Porém a mudança no padrão de marcação do piruvato mitocondrial devido à reação catalisada pela enzima málica¹³ permite, ainda que com moderada certeza, estimar os fluxos nessas vias. A diferença observada nos padrões de marcação do piruvato citosólico e mitocondrial é pequena. A partir dos dados da Tabela 3.1 temos: $PYR_{cit}(2,3) = 36,0$ (de Ala116 e Ala99), e $PYR_{mit}(2,3) = 36,6$ (de Val144 e Val127); $PYR_{cit}(1) = 2,5$ (de Ala158 e Ala99), e $PYR_{mit}(1) = 1,7$ (de Val186, Val144 e Val127). Essas diferenças indicam que o fluxo na reação catalisada pela enzima málica também é pequeno, como observado na solução obtida neste e em trabalhos anteriores (ver Figuras 3.6 e 3.7).

A Figura 3.8 apresenta algumas seções da árvore gerada pela busca *Branch & Bound*, incluindo a evolução da solução para os modelos convexo e não-convexo (*lb* e *ub*) e dos limites na solução global (*LB* e *UB*). É importante observar que o número de iterações e o esforço computacional podem ser reduzidos empregando-se limites iniciais mais estreitos, os quais são obtidos com um valor menor para o parâmetro de incerteza (*d*); porém, o resultado final pode ser diferente, já que se corre o risco de excluir a solução global da região de busca, como foi observado na identificação de soluções locais (item 3.4.1). Nesta figura observa-se que a maior redução na diferença entre *LB* e *UB*, conhecida como *gap*, acontece no ajuste dos limites; assim, apesar destes ajustes consumirem 58% do esforço computacional (tempo de CPU), esse passo é indispensável para a convergência global.

A Tabela 3.3 apresenta os valores calculados e medidos para as *SFLs*, enquanto a Tabela 3.4 mostra os valores para os fluxos. Pode-se observar que os valores obtidos por otimização global são similares aos valores medidos e aos estimados em trabalhos anteriores (Gombert *et al.*, 2001), porém os erros quadráticos obtidos por abordagem de otimização global são menores para ambos grupos de variáveis.

¹³ Esta reação transforma o malato em piruvato e dióxido de carbono, porém no modelo da rede metabólica (Apêndice A) malato e oxaloacetato não são diferenciados, e a reação é apresentada como: $OOA_{mit} \rightarrow PYR_{mit} + CO_2$.

Figura 3.8 Árvore da busca *Branch & Bound*. lb e ub são os valores da função objetivo para os problemas convexos e não-convexos em cada sub-região avaliada. UB^n e LB^n são os limites para a solução após a divisão n .

Tabela 3.3 Valores medidos, e calculados por algoritmo evolucionário e global para as *SFLs*.

Fragmento	Gombert <i>et al.</i> (2001)		abordagem global
	medido	calculado	calculado
Glc331	91,0	91,07	91,14
Gly175	5,9	6,24	6,27
Gly144	6,2	6,24	6,27
Gly85	3,7	3,19	3,27
Ser175	3,1	3,48	3,33
Ser132	33,8	34,43	34,57
Ala116	36,0	34,43	34,57
Ala99	36,0	34,43	34,57
Ala158	38,5	36,56	36,59
Leu158	106,1	105,08	105,04
Val144	73,3	72,67	72,76
Val143	6,9	7,69	7,52
Val127	73,1	72,67	72,76
Val186	74,9	76,19	76,19
Asp188	57,1	58,90	58,77
Asp115	12,0	10,31	10,37
Asp216	64,4	65,42	65,36
Ile158	93,7	95,24	95,15
Glu143	39,9	40,19	40,01
Glu230	100,1	99,56	99,66
Pro142	85,5	85,21	84,52
Lys156	117,2	118,05	117,92
Phe192	90,6	91,83	91,55
Phe143	3,0	3,48	3,32
Erro quadrático somado		24,61	21,88

Tabela 3.4 Valores medidos, e calculados por algoritmo evolucionário e global para os fluxos.

Reação	Gombert <i>et al.</i> (2001)		abordagem global
	medido	calculado	calculado
G6P → G6Puot	22,20	22,20	21,89
P5P → P5Pout	2,86	2,86	2,82
E4P → E4Pout	2,50	2,50	2,46
G3P → G3Pout	0,82	0,82	0,82
PEP → PEPout	4,99	5,00	4,96
PYRmit → PYRmitout	12,40	12,46	12,34
PYRcit → PYRcitout	4,07	4,04	4,04
OAAcit → OAAcitout	5,3	5,36	5,34
AKG → AKGout	9,62	10,83	10,49
AcCoAcit → AcCoAcitout	22,0	22,91	22,34
AcCoAmit → AcCoAmitout	2,8	2,82	2,82
SER → SERout	1,81	1,81	1,80
GLY → GLYout	3,66	3,66	3,68
C-1 → C-1out	2,94	2,91	2,91
THR → THRout	3,63	3,63	3,66
ETH → ETHout	0,0	0,10	0,0
ACE → ACEout	0,0	0,10	0,0
GLYC → GLYCout	0,0	0,10	0,0
Erro quadrático somado		1,70	0,86

3.5 Conclusões

Os resultados obtidos com NLP local permitem verificar que problemas de quantificação de fluxos metabólicos com medidas de marcação têm soluções locais, as quais correspondem a distribuições diferentes para a rede metabólica, o que expõe a fragilidade das abordagens locais na solução destes problemas. Apesar disso, o número de ótimos locais identificados no estudo de caso e o elevado esforço computacional necessário para a convergência global não permitem concluir contundentemente quanto à vantagem de implementar metodologias rigorosas de busca global.

A abordagem *Branch & Bound* espacial é uma alternativa viável para resolver globalmente os problemas de quantificação de fluxos metabólicos com medidas de marcação, e os estimadores de McCormick foram eficientes na relaxação convexa do problema original, permitindo gerar os limites inferiores para a solução global.

Quanto à convergência para o ótimo global, observou-se que as *SFLs* e os fluxos líquidos não-medidos são as variáveis mais vantajosas para a divisão da região de busca; que são necessárias múltiplas divisões em variáveis de ambos grupos, e que a maior redução no *gap* deve-se ao procedimento de ajuste dos limites, o que torna este “passo” indispensável na resolução do problema.

3.6 Nomenclatura na quantificação de fluxos metabólicos

Esta seção apresenta a notação utilizada para a formulação do problema de quantificação de fluxos metabólicos.

<i>A</i>	abundância para isotopômeros de posição.
<i>a</i>	abundância para isotopômeros de massa.
<i>CSA</i>	contador de sub-regiões avaliadas.
<i>g</i>	coeficiente estequiométrico para metabólitos intracelulares.
<i>l</i>	marcação fracionária individual.
<i>lb</i>	solução do problema convexo em uma sub-região.
<i>LB</i>	limite inferior da solução global.

P	produto metabólico.
r	taxa líquida de reação, considerando todas as reações.
S	substrato do metabolismo.
SFL	marcação fracionária somada.
t	tempo.
ub	solução do problema não-convexo em uma sub-região.
UB	limite superior da solução global.
v	fluxo metabólico (velocidade de reação).
x	variável medida experimentalmente na abordagem de erro-nas-variáveis.
$\hat{x}, \bar{x}$	valores estimados e medidos para as variáveis.
X_{met}	concentração dos metabólitos intracelulares.
Z_{macro}	composto constituinte da biomassa.
Z_{met}	metabólito intracelular.

Letras Gregas

a	coeficiente estequiométrico do substrato.
β	coeficiente estequiométrico do produto.
$?$	coeficiente estequiométrico do constituinte da biomassa.
$?$	contribuição máxima para cada grupo de variáveis na função objetivo.
d	parâmetro que introduz a incerteza das medições na estimação dos limites.
$?$	parâmetro desconhecido na abordagem de erro-nas-variáveis.
μ	velocidade específica de crescimento de biomassa.
$?$	tolerância relativa na convergência global.
s	desvio padrão.

Conjuntos

F	fragmentos com medida de marcação.
I_f	metabólitos que geram o fragmento f .
IS_j	metabólitos que participam como substratos na reação j .
J	reações que compõem a rede metabólica.
J_{ip}	reações que geram o metabólito ip .
J_{ip}^*	reações que consomem o metabólito ip .
JM	reações com fluxo medido experimentalmente.

$K_{i,f}$ átomos do metabólito i incorporados no fragmento f .

Arranjos

AMMA arranjo de matrizes de mapeamento de átomos.

$\underline{\underline{A}}$ matriz estequiométrica para os substratos.

$\underline{\underline{B}}$ matriz estequiométrica para os produtos.

$\underline{\underline{G}}$ matriz estequiométrica para os constituintes da biomassa.

$\underline{\underline{G}}$ matriz estequiométrica para os metabólitos intracelulares.

MMM matriz de marcação nos metabólitos.

$[i > ip]_j$ matriz de mapeamento de átomos na transformação de i para ip na reação j .

Subscritos

exch fluxo de troca de átomos nas reações bidirecionais.

f fragmento.

i metabólito intracelular ou composto em geral (só nas equações 2.1 a 2.6).

j reação.

k posição do átomo de carbono.

net fluxo líquido nas reações bidirecionais.

p número que caracteriza o isotopômero de posição.

q número que caracteriza o isotopômero de massa.

s sub-região criada na divisão da região de busca.

Superscritos

L limite inferior para uma variável ou parâmetro.

U limite superior para uma variável ou parâmetro.

CAPÍTULO 4 SÍNTESE E PROJETO DE BIOPROCESSOS

Este capítulo introduz o problema de síntese e projeto de bioprocessos, apresentando brevemente a evolução histórica dos bioprocessos e caracterizando os estágios de purificação; também define os produtos considerados no estudo de caso e os hospedeiros usados para sua expressão, estabelece a importância de otimizar a síntese e projeto do processo, e define as estratégias que vêm sendo empregadas para resolver problemas tanto de projeto quanto de síntese de processos.

4.1 Breve resumo da evolução dos processos de biossíntese

Vários processos biotecnológicos, como produção de cervejas, vinhos e queijos, têm sido empregados e aperfeiçoados durante séculos; estes processos envolvem a ação de enzimas ou microorganismos para a conversão de substratos em produtos. Alguns processos que envolvem proteínas, como a inclusão de enzimas proteolíticas em detergentes, foram desenvolvidos no começo do Século XX mas só foram industrialmente aplicados no final dos anos 60; outros processos industriais que dependem da ação catalítica das enzimas foram desenvolvidos entre os anos 40 e 60 (Walsh e Headon, 1994). Porém, foi após o desenvolvimento das técnicas de engenharia genética nos anos 70 que a biotecnologia tornou-se uma área de rápida evolução. Hoje em dia, o gene para codificar qualquer proteína pode ser em princípio isolado e inserido em um hospedeiro apropriado, para permitir a sua expressão e produção em escala comercial.

As fontes disponíveis para obtenção de bioprodutos também apresentam constante expansão. Nos anos 20, a diabetes começou a ser tratada com insulina extraída e purificada de células do pâncreas de animais, mas a disponibilidade do tecido usado como fonte e seu baixo teor de insulina impediram a produção em grande escala. A biossíntese de insulina a partir de *Escherichia coli* recombinante foi desenvolvida e comercializada em 1982 (Johnson, 1983), sendo o primeiro produto terapêutico obtido com tecnologia de recombinação do DNA. Atualmente, os produtos da indústria biotecnológica são altamente diversificados (ácidos orgânicos, antibióticos,

polissacarídeos, hormônios, peptídeos e proteínas), bem como os hospedeiros empregados para sua produção e o local de acúmulo com relação à célula. Os hospedeiros mais empregados para processos de biossíntese são microorganismos geneticamente modificados, como bactérias, leveduras e fungos, e células animais, vegetais e de insetos também modificadas.

4.2 Purificação de bioprodutos

A maioria dos bioprodutos comercialmente disponíveis podem ser agrupados em duas categorias: os produzidos em grandes quantidades sem necessidade de pureza elevada, como ácidos orgânicos e enzimas com aplicações nas indústrias de alimentos e bebidas, e os destinados a aplicações terapêuticas que são produzidos em quantidades menores e com alto grau de pureza. As proteínas e substâncias protéicas, como enzimas e hormônios, de uso terapêutico são produzidas, quase exclusivamente, em culturas de microorganismos ou células animais, onde podem ser expressas intra ou extracelularmente.

Quando a proteína é intracelular, o processo de purificação costuma ser mais complexo já que a ruptura das células gera uma mistura complexa de substâncias; porém se o produto de interesse acumula-se na forma de corpos de inclusão¹⁴, o isolamento destes aglomerados é relativamente simples e permite um grau de pureza considerável, já que eles são constituídos em grande parte pela proteína alvo (Singh e Panda, 2005), mas a dissolução dos corpos de inclusão e a renaturação¹⁵ da proteína diminuem consideravelmente a produtividade do processo. Também existe a possibilidade de que o produto se acumule no espaço periplásmico; neste caso, a ruptura da parede celular permite ter o produto dentro de um meio de baixa complexidade simplificando os estágios posteriores. A Figura 4.1 ilustra os locais de acúmulo das proteínas expressas em *E. coli*.

¹⁴ Aglomerados de proteína desnaturada devido ao dobramento e enovelamento incorreto das cadeias.

¹⁵ Alguns autores preferem o termo *naturação* (*naturation*) porque entendem que está se levando o produto a seu estado nativo, o qual nunca teve (Singh e Panda, 2005).

Figura 4.1 Locais e número aproximado de proteínas em *E. coli* (Janson e Rydén, 1989).

Apesar de não envolver o rompimento das células para a liberação do produto, a purificação das proteínas que são secretadas no meio extracelular tem suas complexidades, especialmente em culturas de células animais nas quais substâncias protéicas são adicionadas no meio de cultura para promover o crescimento celular (Scopes, 1994). A continuação apresentam-se os estágios para recuperação e purificação de bioprodutos, tanto extracelulares como intracelulares, considerando as operações empregadas industrialmente para desenvolver cada tarefa e algumas considerações para a integração dos estágios de processamento.

4.2.1 Remoção/coleta de células, restos celulares ou corpos de inclusão

No caso de produtos extracelulares, o primeiro passo para a recuperação do produto é a clarificação do meio de cultura que consiste na remoção das células. Se o produto é intracelular, as células devem ser colhidas para posterior rompimento, e um segundo estágio de clarificação deve ser projetado para remover os restos celulares e as células remanescentes ou para colher os corpos de inclusão. As técnicas mais empregadas para desempenhar estas tarefas, em processos de escala industrial, são centrifugação e microfiltração.

A centrifugação consiste na aceleração da sedimentação por ação de um campo gravitacional centrífugo. As centrífugas de escala industrial, diferentemente dos equipamentos para laboratório, operam de forma contínua; porém seu desempenho

depende dos mesmos parâmetros, como o tamanho de partícula, a diferença de densidade e a viscosidade do meio. Assim, a sedimentação de células de leveduras é menos exigente que a sedimentação de células bacterianas ou de restos celulares, devido ao tamanho; bem como os corpos de inclusão sedimentam mais facilmente que os restos celulares, pela maior densidade. Uma centrífuga de discos de escala industrial consegue separar 99% dos sólidos, gerando um sedimento com 80-90% de sólidos, logo torna-se alternativa tanto na remoção de restos celulares (Harris e Angal, 1995) como na coleta de células.

A microfiltração ou filtração de membrana separa as partículas pelo tamanho, empregando uma película fina como elemento filtrante, geralmente fabricada a partir de polímeros como acetato de celulose, nitrato de celulose ou náilon. A filtração tangencial, na qual o fluxo é paralelo ao elemento filtrante (ver Figura 4.2) é a forma de operação preferida na purificação de bioprodutos. Um dos principais problemas dos sistemas de filtração por membrana é o entupimento (*fouling*) que diminui o fluxo e aumenta a pressão do sistema, o que pode danificar a membrana. Uma alternativa para evitar esse problema é empregar um sistema com várias membranas com tamanho de poro decrescente em um único cartucho (ver Figura 4.2). Schmidell *et al.* (2001) sugerem que o entupimento pode ser minimizado com o controle da velocidade de escoamento e da pressão de transmembrana¹⁶.

Figura 4.2 Esquema da microfiltração e de um cartucho com várias membranas

¹⁶ Diferença entre a média das pressões de alimentação e do retido, e a pressão do filtrado.

4.2.2 Rompimento de células

O rompimento de células pode ser realizado por métodos químicos (tratamento com detergentes, solventes orgânicos), enzimáticos (tratamento com lisoenzimas), físicos (homogeneização, ultra-som) ou mecânicos (moinhos de bolas). O desempenho de cada técnica depende do tipo de célula e das condições de cultura. As principais desvantagens dos sistemas químicos são a possibilidade de desnaturação do produto e os efeitos das substâncias adicionadas nos estágios posteriores; já os tratamentos com enzimas e ultra-som apresentam custos elevados, o que limita sua aplicação em escala industrial.

As células animais, diferentemente das de microorganismos, carecem de parede celular, o que as torna mais frágeis e vulneráveis ao rompimento. Os principais métodos para rompimento de células, tanto de microorganismos como animais, em escala comercial, são a homogeneização e a moagem com bolas.

Na homogeneização, a suspensão celular é obrigada a passar sob elevada pressão através de um pequeno orifício, o que gera grandes tensões de cisalhamento e produz o rompimento celular. Os equipamentos modernos possuem sistemas de resfriamento que dissipam o calor gerado, reduzindo a desnaturação das proteínas. A aplicação de pressão elevada proporciona alto desempenho em uma única etapa; no entanto, pode-se operar com múltiplas etapas para aumentar o rendimento do processo.

Nos moinhos de bolas, as células são agitadas vigorosamente na presença de esferas de vidro de 0,2-0,3 mm de diâmetro e acabam rompendo-se devido à colisão com as esferas. Assim como na homogeneização, o calor gerado é considerável e deve ser retirado para evitar a degradação do produto.

4.2.3 Dissolução de corpos de inclusão e renaturação de proteínas

A maioria das proteínas de uso terapêutico expressas em *E. coli* apresentam-se na forma de corpos de inclusão, que são aglomerados insolúveis e podem conter até 40% do total de proteína celular (Misawa e Kumagai, 1999). Os métodos tradicionais para recuperar proteínas dos corpos de inclusão requerem quatro estágios: separação dos

corpos de inclusão, dissolução, renaturação e purificação da proteína (Singh e Panda, 2005). A dissolução emprega agentes desnaturantes como uréia e hidrócloro de guanidina em alta concentração (6-8 M) ou detergentes, como dodecil sulfato de sódio; em alguns casos é necessária a adição de tióis¹⁷ que agem como agentes redutores.

A renaturação das proteínas dissolvidas começa com a remoção dos agentes desnaturantes por diluição ou diafiltração; a eficiência da renaturação depende da competição entre o *folding*¹⁸ correto e a agregação de moléculas, e portanto este estágio é desenvolvido em baixa concentração de proteína. A renaturação de proteínas com múltiplas ligações dissulfeto é feita na presença de agentes oxidantes e redutores que desestabilizam as ligações erradas. Protocolos específicos para dissolução e renaturação de várias proteínas expressas em *E. coli* são apresentados em Misawa e Kumagai (1999), Singh e Panda (2005) e Harris e Angal (1995).

4.2.4 Concentração e purificação inicial

Muitas vezes, depois dos estágios de recuperação, o produto encontra-se diluído em baixa concentração. Para reduzir o tempo e o custo dos estágios de purificação é necessário concentrar a solução. Em escala industrial, a concentração de soluções é desenvolvida por precipitação, cromatografia de troca iônica ou ultrafiltração.

A precipitação de proteínas pode ser induzida mediante a adição de sais neutros, solventes orgânicos ou polímeros de peso molecular elevado, ou por mudanças de pH. O sulfato de amônio é, possivelmente, o agente de precipitação mais empregado (Walsh e Headon, 1994); sua popularidade deve-se à alta solubilidade, ao baixo custo, à característica de não desnaturar a maioria das proteínas e ao efeito estabilizador em muitas delas. O aumento da concentração do sal desestabiliza as proteínas na solução e induz a sua precipitação, o que é conhecido como *salting out*. Este processo depende da

¹⁷ Ou tioálcool (R-SH), composto no qual o oxigênio da função álcool é substituído pelo enxofre.

¹⁸ Dobramento e enovelamento da cadeia, que definem a estrutura terciária da molécula, mantidos por interações eletrostáticas, pontes de hidrogênio, ligações dissulfeto, interações entre cadeias laterais hidrofóbicas, etc.

existência de zonas hidrofóbicas na superfície da proteína; quando se adiciona o sal, os íons gerados capturam moléculas de água, permitindo a interação entre as regiões hidrofóbicas das moléculas de proteína, o que induz sua agregação e precipitação.

A cromatografia de troca iônica, além de concentrar a solução, envolve uma purificação considerável do produto e portanto é considerada nos procedimentos de purificação de alta resolução (item 4.2.5).

A ultrafiltração também é amplamente usada na concentração de soluções. Enquanto a microfiltração só retém células inteiras e seus restos, e emprega elementos filtrantes com poros de 0,1 μm a 10 μm , a ultrafiltração retém proteínas de baixo peso molecular e emprega membranas com poros de 1 nm a 20 nm. Comercialmente, as membranas de ultrafiltração são denominadas pelo peso molecular nominal de corte (*cut-off*), que é definido como o mínimo peso que uma molécula globular requer para ser retida. O crescente uso de ultrafiltração na concentração de soluções de proteínas deve-se a múltiplos fatores como o baixo efeito sobre a atividade do produto, as elevadas taxas de recuperação do mesmo e a elevada velocidade de processamento.

A diafiltração permite remover moléculas de baixo peso molecular e mudar as condições da solução entre dois estágios de purificação. Este procedimento é semelhante à ultrafiltração, porém nova solução é adicionada continuamente ao reservatório. As moléculas a serem eliminadas podem ser sais, etanol e outros solventes, aminoácidos, peptídeos, estabilizadores de proteínas adicionados, etc. Já a mudança no *buffer* é feita quando, por exemplo, o pH e a força iônica devem ser alterados entre dois estágios de purificação.

4.2.5 Purificação de alta resolução por cromatografia

Como exposto anteriormente, os bioprodutos destinados a aplicações terapêuticas requerem elevado grau de pureza, o que envolve a aplicação de um ou vários estágios cromatográficos. A cromatografia pode ser definida como a separação dos componentes de uma amostra que se movimenta através de uma fase sólida estacionária, empacotada numa coluna. Os componentes que têm mais afinidade com a fase sólida atravessam a

coluna mais lentamente ou ficam nela retidos. A eficiência da purificação pode ser avaliada em função do grau de separação das moléculas, denominado resolução. Outras características que devem ser consideradas para a seleção de um processo cromatográfico são: a capacidade, que é a quantidade de proteína que pode ser adsorvida por unidade de volume ou de massa da fase sólida, e a seletividade, que é a capacidade de reter a proteína e de rejeitar os contaminantes.

Cada técnica cromatográfica permite a separação de proteínas com base em alguma característica, como tamanho e forma, carga total, existência de grupos hidrofóbicos na superfície e capacidade de ligação com outras espécies, entre outras. Quando são comparadas as características de várias proteínas, observam-se algumas semelhanças, porém a combinação de características é única para cada proteína, o que permite obter produtos com elevado grau de pureza mediante a combinação de técnicas cromatográficas.

A cromatografia por exclusão ou filtração em gel (*gel filtration*) separa as proteínas pelo seu tamanho e forma. Nesta técnica, o fracionamento¹⁹ é feito quando a solução passa através da coluna empacotada com partículas esféricas de um gel poroso. As proteínas maiores não conseguem entrar nos poros do gel e atravessam os espaços existentes entre as partículas, então percorrem a coluna rapidamente e eluem primeiro; por outro lado, as proteínas menores entram nos poros do gel e são retidas na coluna por mais tempo (ver Figura 4.3).

Em muitos casos, o gel de filtração é formado por polímeros com elevado número de ligações cruzadas, o qual define o tamanho médio dos poros. O tamanho de poro a ser empregado depende do peso molecular da proteína de interesse e dos contaminantes. O tamanho das partículas do gel também tem efeito na separação; quando as partículas são menores a resolução é melhor, porém a vazão se torna menor. Este tipo de cromatografia geralmente é usada no final do processo de purificação, pois trabalha com volumes pequenos e a coluna é facilmente entupida por impurezas. Depois da aplicação

¹⁹ Separação dos componentes de uma solução por divisão da mesma em frações de composição diferente.

da solução, as frações são eluídas pela adição de uma solução apropriada; a solução obtida geralmente é mais diluída que a solução inicial, porém mais pura. A filtração em gel é um dos métodos mais empregados na purificação de proteínas e geralmente complementa a cromatografia de troca iônica

Figura 4.3 Princípio do fracionamento de proteínas por filtração em gel.

A cromatografia de troca iônica (*ion-exchange*) emprega a carga característica da molécula para induzir sua retenção na matriz, e a eluição é feita com uma solução de elevada força iônica (ver Figura 4.4). A carga total de uma proteína depende dos aminoácidos²⁰ e grupos funcionais que a conformam, e do pH da solução. O pH no qual a molécula apresenta carga total zero é denominado ponto isoelétrico (pI); uma proteína apresenta carga total negativa quando o pH da solução é maior do que o seu pI, e carga total positiva na condição oposta.

²⁰ No pH 7.0, ácidos aspártico e glutâmico apresentam carga negativa, enquanto lisina, arginina e histidina têm carga positiva.

Figura 4.4 Princípio da retenção e eluição na cromatografia de troca iônica

A maioria dos protocolos de purificação de proteínas empregam pelo menos um estágio de troca iônica; o amplo uso desta técnica deve-se ao seu bom nível de resolução, à simplicidade no aumento de escala, e à facilidade de uso e regeneração da coluna. Trocadores iônicos derivados da celulose são muito empregados como matriz pela eficiência e pelo baixo custo relativo, pois podem ser regenerados e reutilizados, e ainda permitem a eliminação de impurezas como partículas que não foram removidas, algumas moléculas de lipídeos e carboidratos, proteínas parcialmente desnaturadas ou agregadas, e proteínas com carga diferente.

A cromatografia de interação hidrofóbica (*hydrophobic interaction*) baseia-se no grau de hidrofobicidade da superfície protéica. Dos aminoácidos que constituem as proteínas oito são considerados hidrófobos pela natureza apolar dos seus grupos terminais. A estrutura terciária das proteínas permite que a maioria destes aminoácidos permaneça no interior da estrutura, podendo interagir entre eles e não com o meio aquoso, o que estabiliza a estrutura, enquanto os aminoácidos presentes na superfície da molécula são expostos ao meio e definem a interação da proteína com outras moléculas.

O grau de hidrofobicidade de uma molécula depende do número e do tipo de aminoácidos hidrófobos presentes na superfície; a separação das proteínas, mediante esta técnica, é baseada nas interações hidrofóbicas entre a superfície das proteínas e os grupos

hidrófobos ligados à matriz da coluna cromatográfica. O incremento da força iônica da solução, pela adição de sais como sulfato de amônio, aumenta a hidrofobicidade das proteínas²¹ e favorece sua ligação na matriz. Quanto mais hidrofóbica uma proteína, mais forte a ligação; a eluição emprega uma solução de baixa força iônica e a regeneração da coluna requer lavagem intensa para remover as moléculas fortemente ligadas à matriz, as quais diminuem sua capacidade de retenção. A cromatografia de interação hidrofóbica geralmente é desenvolvida depois de um estágio de precipitação por adição de sal ou de cromatografia por troca iônica, já que a solução obtida por estas técnicas tem elevada força iônica.

A cromatografia em fase reversa (*reverse-phase*) também separa as proteínas com base na hidrofobicidade, porém nesta técnica emprega-se uma coluna altamente hidrofóbica, que retém a maioria das proteínas, e a eluição é feita com uma solução de baixa polaridade. A forte condição de eluição ocasiona a desnaturação de muitas proteínas (Walsh e Headon, 1994), o que torna mais restrito o uso desta técnica. Schmidell *et al.* (2001) expõem que a desnaturação das proteínas é causada pela interação do solvente apolar com as regiões hidrófobas internas.

A cromatografia de afinidade (*affinity chromatography*) é considerada a técnica mais poderosa e de melhor seletividade (Walsh e Headon, 1994). Esta técnica baseia-se na capacidade que as proteínas têm de se ligar a substâncias específicas, denominadas ligantes (ver Figura 4.5), e na reversibilidade desta ligação. Os ligantes estão fixos por ligações covalentes à matriz inerte; quando a solução atravessa a coluna, as moléculas da proteína alvo são retidas na coluna, enquanto os contaminantes não conseguem se ligar e são lavados com a solução; posteriormente, aplica-se um eluente contendo um composto que compete pela ligação, e permite a liberação e eluição da proteína.

Uma grande variedade de ligantes pode ser empregada para purificar proteínas por afinidade, mas o uso em escala industrial de moléculas produzidas biologicamente como

²¹ Assim como no processo de precipitação (item 4.2.4), a hidratação das moléculas do sal expõe as regiões hidrofóbicas das proteínas.

adsorventes só é viável quando o produto final tem alto valor agregado. Quando anticorpos são usados para purificar o antígeno que estimula sua produção, a técnica é conhecida como cromatografia de imuno-afinidade (*immunoaffinity*).

Figura 4.5 Princípio da cromatografia por afinidade.

4.2.6 Integração dos estágios de processamento

A escolha dos estágios para a purificação de bioprodutos considera múltiplas características do sistema. Os estágios iniciais visam a recuperação dos produtos a partir do meio de cultura, e dependem especialmente do hospedeiro e das condições do meio de cultura, enquanto os estágios de purificação de alta resolução dependem das características da molécula alvo e dos contaminantes. Além destas características, a definição da sequência de operações, conhecida como síntese do processo, deve considerar as características inerentes de cada operação e a compatibilidade com as demais. Alguns critérios válidos são:

- A filtração em gel normalmente é desenvolvida no final do processo, pois envolve diluição do produto e aumento do volume da solução, o que incrementa os custos dos estágios seguintes (Schmidell *et al.*, 2001).
- A cromatografia de interação hidrofóbica, geralmente, é desenvolvida depois de um estágio de troca iônica ou precipitação por *salting-out*, já que a solução obtida nestes estágios apresenta elevada força iônica. Assim, a mudança das condições da solução é menor.

Figura 4.6 Sequências de estágios para purificação de proteínas extra e intracelulares (Walsh e Headon, 1994), e em corpos de inclusão (Misawa e Kumagai, 1999).

Diversos protocolos para purificação de bioprodutos são encontrados na literatura. Walsh e Headon (1994) sugerem sequências de operações para a biossíntese de produtos de uso terapêutico com expressão extra e intracelular, enquanto Misawa e Kumagai

(1999) sugerem uma sequência para o processamento de proteínas expressas em corpos de inclusão. A Figura 4.6 apresenta estas sequências com alguns elementos complementares, como os estágios de diafiltração para a modificação da solução entre estágios cromatográficos.

4.3 Características de alguns bioprodutos

Nesta seção caracterizam-se os bioprodutos envolvidos no estudo de caso, identificando a sua importância e função, e apresentando brevemente a evolução das técnicas e fontes de produção.

4.3.1 Insulina humana

A insulina é um hormônio essencial para a regulação metabólica da glicose. Como citado na seção 4.1, a produção de insulina para uso terapêutico começou a partir de tecido pancreático como fonte, porém a produção em grande escala só foi possível com o desenvolvimento da expressão em *E. coli* recombinante. A insulina humana é um polipeptídeo formado por duas cadeias (A e B) com 21 e 30 aminoácidos respectivamente. Nas células beta do pâncreas, a insulina é sintetizada como pró-insulina numa única cadeia, para posteriormente ser transformada na molécula de insulina com duas cadeias.

A produção de insulina recombinante em *E. coli* segue dois caminhos (Schmidt *et al.*, 1999B); um é a produção de pró-insulina e sua posterior transformação em insulina, o outro é a produção e purificação de forma separada das cadeias A e B para sua posterior combinação. Enquanto a produção em *S. cerevisiae* é feita na forma do precursor de insulina (IP, sigla do inglês), que contém 29 aminoácidos da cadeia B ligados aos 21 da cadeia A, o sistema de secreção em leveduras é complexo e ainda não bem entendido (Kjeldsen *et al.*, 1999), e requer que a molécula esteja corretamente dobrada (*folded*), o que garante alto nível de atividade do produto final, e ligada a um composto guia (*leader*), o qual permite seu trânsito dentro da célula e sua secreção.

4.3.2 Vacina para hepatite B

Inicialmente, as infecções por vírus da hepatite B eram prevenidas pela vacinação com partículas da superfície do vírus, com propriedades de antígeno, *surface-antigen* (HBsAg), as quais eram purificadas do plasma de portadores assintomáticos do vírus. Apesar das vacinas derivadas do plasma serem seguras e efetivas apresentam algumas desvantagens (Hardy *et al.*, 2000):

1. A produção envolve procedimentos complexos para a inativação dos vírus da hepatite B e de outros patógenos que possam estar presentes no plasma.
2. O custo de produção é relativamente elevado, inviabilizando os programas de imunização de grande escala.
3. A disponibilidade de plasma humano é limitada.
4. São necessários longos testes de inoculação em animais.

Em 1986, a vacinação contra hepatite B com HBsAg produzido em *S. cerevisiae* recombinante foi aprovada (Liljeqvist e Ståhl, 1999), e atualmente existem várias alternativas para biossíntese de HBsAg. Zhang *et al.* (2003B) reportam que na China 90% do HBsAg é produzido a partir de *S. cerevisiae* recombinante, e o restante de células de ovário de Hamster Chinês (CHO, sigla do inglês). Outras pesquisas têm estudado a produção e purificação de HBsAg a partir de *Pichia pastoris* (Hardy *et al.*, 2000), mostrando que o produto obtido satisfaz os parâmetros de segurança da Organização Mundial da Saúde.

4.3.3 Ativador de plasminogênio tecidual (tPA)

Os ativadores de plasminogênio são um grupo de enzimas proteolíticas empregadas como agentes trombolíticos no tratamento de obstruções vasculares, já que transformam o plasminogênio da proteína do sangue em plasmina, que dissolve a fibrina que forma os coágulos.

Outros ativadores de plasminogênio conhecidos, além do tPA, são o originado da uroquinase (uPA) e o de origem bacteriana (estreptoquinase). A estreptoquinase tem custo

relativamente baixo, mas devido à origem não-humana gera respostas imunológicas indesejáveis; a uroquinase, assim como a estreptoquinase, ativa o plasminogênio em todo o sistema circulatório, aumentando o risco de hemorragias, enquanto o tPA ativa preferencialmente o plasminogênio preso nos coágulos (Rouf *et al.*, 1996). A expressão de tPA tem sido estudada em múltiplos sistemas, como células de mamíferos originais e recombinantes (Takagi *et al.*, 2001), e outras culturas de células recombinantes como *E. coli* (Misawa e Kumagai, 1999) e *S. cerevisiae*. A revisão das técnicas de produção apresentada em Rouf *et al.* (1996) permite verificar que as células originais de mamíferos mais eficientes na produção são as malignas, porém a produção comercial só foi possível a partir de células de mamíferos modificadas geneticamente.

Kohnert *et al.* (1996) avaliaram o produto obtido a partir de *E. coli*, observando que apresenta nível de atividade semelhante ao de CHO, a especificidade sobre a fibrina, formadora dos coágulos, também foi observada. Misawa e Kumagai (1999) reportaram produtividade em *E. coli* de aproximadamente 10% do total de proteína celular, e sugerem uma sequência com quinze estágios de produção (ver figura 4.6).

4.3.4 Superóxido dismutase

A enzima superóxido dismutase (SOD) catalisa a dismutação do íon superóxido (O_2^-), altamente reativo, a peróxido de hidrogênio (H_2O_2), que é posteriormente convertido em oxigênio e água sob a ação da enzima catalase, segundo as seguintes reações:

Em mamíferos têm sido observadas três isoenzimas de SOD; a primeira é Cu-, Zn-SOD que é encontrada no citoplasma e no núcleo das células, a segunda é a Mn-SOD mitocondrial e a terceira é SOD extracelular (EC-SOD) que contém Cu/Zn (He *et al.*, 2002). A superóxido dismutase purificada a partir das células do fígado ou de eritrócitos de bovinos tem sido usada como agente anti-inflamatório (Walsh e Headon, 1994). A produção de Cu/Zn-SOD em *E. coli* recombinante foi reportada pela primeira vez em

1985 e posteriormente em leveduras (Ahl *et al.*, 2004). Em ambos os casos, a expressão é intracelular, formando corpos de inclusão no caso da *E. coli*. Produtividades de 30 a 50% do total de proteína são observadas em *E. coli*.

4.4 Aumento de escala dos bioprocessos

Shimizu (1996) analisou os desenvolvimentos na engenharia de bioprocessos e concluiu que, durante os últimos anos, a indústria biotecnológica evoluiu rapidamente, incorporando grande parte dos seus produtos no mercado e gerando a necessidade de projetar criteriosamente os processos desenvolvidos nos laboratórios a dimensões que permitam sua comercialização.

Um dos principais problemas do aumento de escala (*scale-up*) de bioprocessos, segundo o autor, é a necessidade de incrementar a produtividade volumétrica e a concentração dos produtos para tornar a produção mais eficiente e econômica, o que pode-se conseguir com a otimização do projeto e da operação dos biorreatores. Pinto *et al.* (2001) sugerem que, além de otimizar a fase de produção, devem-se determinar a configuração da planta e as dimensões dos equipamentos que minimizam o capital a investir (projeto e otimização estrutural da planta).

Outro aspecto importante quando se procura melhorar o desempenho de processos de biossíntese é a escolha do hospedeiro, o que compõe a síntese do processo. A forma de expressão do produto (extracelular, intracelular livre ou em corpos de inclusão), os subprodutos gerados, suas concentrações e em consequência a produtividade do processo dependem do hospedeiro, e influem fortemente no projeto do biorreator e das unidades posteriores (*downstream*). Assim, acredita-se que a integração da síntese e projeto traz benefícios econômicos para estes empreendimentos.

4.5 Abordagens para síntese e projeto de processos

A síntese de processos envolve a definição da sequência de operações necessárias para gerar o produto desejado e portanto implica na escolha entre alternativas existentes

para a produção e processamento. A resolução destes problemas é abordada por metodologias diferentes: heurísticas, otimização das tarefas (*task targeting optimization*) e otimização da superestrutura. As heurísticas (Lin e Miller, 2004; Steffens *et al.*, 2000; Hostrup *et al.*, 1999) empregam a experiência e criatividade do projetista na tomada de decisões. As duas últimas são baseadas em programação matemática e o projetista deve propor objetivos, restrições e superestruturas alternativas, formular modelos com diferentes níveis de complexidade, porém consistentes e adequadamente conexos, e avaliar e verificar os resultados. Kravanja e Grossmann (1997) argumentam que a vantagem da abordagem de otimização das tarefas é a capacidade de considerar o efeito de múltiplos fenômenos no desempenho de cada estágio do processo, enquanto a otimização da superestrutura emprega modelos simplificados para as operações, porém a principal vantagem da abordagem de otimização da superestrutura é a de analisar as operações de forma integrada, gerando uma representação mais realista do processo.

4.5.1 A otimização isolada das tarefas

Steffens *et al.* (2000) observaram que a abordagem mais empregada para a síntese de bioprocessos é a definição sequencial das operações com otimização de cada unidade, visando melhorar o desempenho do processo, e que embora esta abordagem gere processos rentáveis, algumas alternativas não consideradas podem ser mais lucrativas. Quanto à otimização da operação de processos biotecnológicos, Groep *et al.* (2000) observam que a abordagem mais usada é a otimização individual de cada operação unitária, apesar do consenso de que as mudanças na operação de cada unidade influem no desempenho das seguintes. Tal observação permite concluir que esta abordagem gera um desempenho sub-ótimo da planta. Os autores modelaram a produção da enzima intracelular álcool desidrogenase (*alcohol dehydrogenase*) em culturas descontínuas alimentadas e contínuas. O processo é constituído, além do biorreator, por uma centrífuga para coletar as células, um homogeneizador de alta pressão para romper as mesmas e liberar a proteína, uma centrífuga para remover os restos sólidos, enquanto na purificação emprega-se um sistema de precipitação e dupla separação por centrífuga.

Mediante múltiplas simulações, os autores avaliaram os efeitos das variações da taxa de diluição, que controla implicitamente a taxa específica de crescimento, e do número de passos pelo homogeneizador sobre a produtividade e o rendimento do processo, o qual considera o valor do produto, e os custos de matéria-prima e de operação. Para o processo contínuo, a otimização tradicional (que visa maior produtividade sem considerar a inter-relação das operações unitárias) indica que a taxa de diluição ótima é de $0,135 \text{ h}^{-1}$, enquanto no homogeneizador sugere-se entre 5 e 7 passos. Com estas condições de operação, o rendimento global do processo é de £5000 em cada ciclo de operação. Por outro lado, com a otimização integrada das operações, o rendimento global é de £37000, o qual é obtido operando com uma taxa de diluição de $0,105 \text{ h}^{-1}$ e executando só 3 passos no homogeneizador.

Com isto, os autores verificam que a otimização individual das condições de operação para cada etapa pode conduzir a um desempenho sub-ótimo da planta. Groep *et al.* (2000) também argumentam que, de fato, existe um grande número de processos bioquímicos nos quais pode se obter benefício econômico quando se desenvolve a otimização integrada que considera a inter-relação entre as operações.

4.5.2 Otimização da superestrutura do processo

Quando os problemas de síntese e projeto ótimo de plantas descontínuas envolvem a escolha de alternativas, nos equipamentos a serem empregados ou na sua alocação, geram-se problemas de otimização mista-inteira que podem ser lineares, não-lineares ou dinâmicos, dependendo dos modelos empregados para a operação das unidades. Kravanja e Grossmann (1997) consideram que uma das principais dificuldades da otimização da superestrutura é a complexidade dos problemas, que pode superar a capacidade dos algoritmos de otimização MINLP existentes; as abordagens empregadas para reduzir a complexidade destes problemas implicam na simplificação dos modelos das operações unitárias, na divisão do problema em partes tratáveis e no uso de modelos de agregação ou projeções preliminares. Devido à complexidade da fase de purificação dos produtos em processos bioquímicos, alguns trabalhos focam no desenvolvimento de metodologias

para sua síntese, sem considerar as fases iniciais do processo (Vásquez-Alvarez *et al.*, 2001; Vásquez-Alvarez e Pinto, 2004)

Segundo Samsatli e Shah (1996), o problema do uso de modelos simplificados é a perda de precisão que pode permitir a geração de resultados inviáveis, enquanto a divisão do problema mantém a complexidade dos modelos, mas gera a possibilidade de não se obter o ótimo global. Para contornar a dificuldade gerada pela complexidade do problema, Kravanja e Grossmann (1997) propõem uma abordagem com vários níveis de hierarquia para resolver o problema MINLP de síntese de processos. A abordagem resolve problemas em três níveis: 1) síntese da rede de reatores, que define as variáveis binárias para a seleção da superestrutura, empregando modelos detalhados dos reatores e simplificados das unidades de separação; 2) síntese da etapa de separação e da rede de reatores, que emprega modelos detalhados dos reatores e das unidades de separação, considerando fixa a superestrutura definida no problema anterior, e 3) síntese da rede de trocadores de calor, que emprega a estrutura definida para os reatores e separadores. Os autores argumentam que esta abordagem permite a otimização do projeto do processo através de um caminho mais sistemático e confiável, permitindo a integração, no projeto, de diferentes fenômenos como reação, separação e transferência de calor, e evitando lidar com um problema extremamente complexo como é a síntese e projeto simultâneo com modelos detalhados de todos os equipamentos e decisões em todos os níveis.

Outras metodologias que resolvem o problema em múltiplos níveis de complexidade são apresentadas em Allgor e Barton (1997) e Galan e Barton (1998); nesses trabalhos, são estudados problemas de otimização mista-inteira dinâmica (MIDO). Por outro lado, a metodologia aplicada por Pahor e Kravanja (1995) para síntese de um processo produtivo com uma seqüência de separações multi-componente, modelado como um problema MINLP dinâmico, implica na discretização do sistema de equações diferenciais por colocação ortogonal em elementos finitos. Este procedimento gera um conjunto de equações não-lineares que substituem as diferenciais na formulação do MINLP; os resultados obtidos foram satisfatórios, evidenciando as potencialidades da metodologia. Riascos e Pinto (2004) estudaram a viabilidade de resolver problemas de otimização dinâmicos por colocação ortogonal em elementos finitos; o estudo de caso

abordado foi a biossíntese de poli-hidroxialcanoatos, empregando um modelo cibernético estruturado. Os autores concluem que essa metodologia gera uma boa aproximação para sistemas complexos de equações diferenciais e que o erro gerado pela aproximação pode ser reduzido com algumas restrições adicionais, as quais não aumentam significativamente a carga computacional na solução do problema.

4.6 Síntese e projeto ótimo de plantas bioquímicas descontínuas

Há várias décadas, a engenharia de processos tem direcionado esforços na procura de metodologias para a síntese e projeto ótimo de diversos sistemas, tais como: redes de trocadores de calor (Cheung e Hui, 2001), seqüências de sistemas de destilação (Novak *et al.*, 1996), redes de trocadores de massa (Foo *et al.*, 2004), plantas multipropósito (Lin e Floudas, 2001). Além disso, segundo Montagna *et al.* (1994), a possibilidade de obter diversos produtos, empregando seções da planta ou mudando o esquema de processamento, faz dos processos descontínuos uma boa alternativa para manufaturar produtos de alto valor e de demanda limitada.

Desafortunadamente, como observado em Samsatli e Shah (1996), as técnicas desenvolvidas na engenharia de processos não podem ser diretamente aplicadas na engenharia bioquímica, devido às grandes diferenças existentes entre os processos químicos e bioquímicos, às quais são geradas pela natureza peculiar dos últimos (o uso de biocatalisadores: microorganismos, células animais, vegetais ou enzimas), já que como consequência direta, os produtos desejados encontram-se em baixas concentrações dentro de uma mistura de substâncias semelhantes. Assim, o projeto destes processos apresenta dificuldades particulares:

- A modelagem matemática das operações de separação é complexa e os parâmetros são difíceis de medir com exatidão.
- Estes processos agrupam um conjunto de operações descontínuas, descontínuas alimentadas e semi-contínuas, e as vezes várias operações podem ser executadas em um mesmo equipamento (multipropósito) ou deve-se escolher entre múltiplos equipamentos que desempenham a mesma finalidade; assim, para definir a

programação da produção (*scheduling*) é necessário verificar a disponibilidade dos equipamentos.

4.6.1 Modelos de desempenho

Montagna *et al.* (1994) observam que no projeto de plantas descontínuas multiproduto empregam-se tempos de processamento e parâmetros de dimensionamento para cada produto em cada equipamento no lugar de modelos gerais para as operações; estes parâmetros podem ser fixados a partir de valores obtidos em laboratório ou planta piloto ou definidos em função das variáveis de operação do processo, empregando modelos de desempenho (*process performance models*). Os autores argumentam que quando estes parâmetros são fixados, como em Ravemark e Rippin (1998), impõe-se uma forte restrição, impedindo a análise de soluções que poderiam ser interessantes.

Como os modelos de desempenho definem os parâmetros para dimensionamento e estimação do tempo de operação em função das variáveis de processo selecionadas para otimização da planta geram uma representação mais detalhada do processo e conseqüentemente um projeto mais exato. Estes modelos têm sido empregados em trabalhos que visam o desenvolvimento de metodologias para o projeto ótimo de bioprocessos (Montagna *et al.*, 1994; Asenjo *et al.*, 2000; Pinto *et al.*, 2001). Nesses trabalhos observa-se que os projetos com modelos de desempenho apresentam redução considerável nos custos da planta (capital e operação); além disso, permitem diminuir a ociosidade dos equipamentos.

4.6.2 Projeto com equipamentos em paralelo e tanques intermediários

Durante o projeto do processo, podem ser empregadas algumas estratégias para reduzir os custos de capital, tais como o projeto de múltiplos equipamentos em paralelo, operando em-fase ou fora-de-fase, e a alocação de tanques intermediários (Ravemark e Rippin, 1998; Pinto *et al.*, 2001). O projeto de equipamentos em paralelo permite flexibilizar as restrições das condições de operação da planta, que originalmente estão definidas pelas capacidades e tempos de operação dos equipamentos; assim, unidades operando em-fase são projetadas quando a dimensão máxima de uma unidade não

permite o processamento da batelada, e equipamentos fora-de-fase quando o tempo de operação do estágio restringe o tempo de ciclo do processo e gera tempos ociosos em outros equipamentos (ver item 5.2.1). A alocação de tanques intermediários é uma alternativa para reduzir os custos gerados pela duplicação de equipamentos (ver item 5.2.11).

4.7 Exemplos de síntese e projeto ótimo de plantas multiproduto

Ravemark e Rippin (1998) desenvolveram o projeto ótimo de plantas descontínuas multiproduto, considerando a alocação de tanques intermediários e de unidades paralelas que operam em-fase ou fora-de-fase. O modelo original é submetido a uma transformação exponencial, gerando um problema MINLP convexo que é resolvido por programação matemática. O estudo de caso avaliado envolve uma planta para geração de 15 produtos com 3 estágios descontínuos e 7 semi-contínuos, e uma única possibilidade para alocação de tanques intermediários. Os autores obtêm uma solução de qualidade superior à previamente reportada e concluem que a formulação apresentada tem a vantagem de gerar um problema convexo, garantindo a obtenção da melhor solução possível.

Pinto *et al.* (2001) estudaram o projeto e otimização estrutural de plantas biotecnológicas, observando que a estratégia mais empregada é fixar os parâmetros de dimensionamento e de estimação do tempo de operação dos equipamentos, empregando valores obtidos em laboratório ou plantas piloto. Para cada equipamento, os parâmetros definem uma dimensão específica, por unidade de produto final: o volume (em reatores e tanques), a taxa de processamento (em homogeneizadores e moinhos) ou a área (em filtros e colunas cromatográficas); além do mais, o fator de dimensionamento para um equipamento depende da eficiência dos posteriores.

Para abordar tal problema, os autores empregam modelos de desempenho das operações unitárias, os quais estimam os fatores de dimensionamento e de tempo de processamento. O processo avaliado foi a produção de quatro proteínas diferentes por *Saccharomyces cerevisiae*, das quais duas têm aplicação medicinal (insulina e vacina para hepatite B), a terceira deve ter grau alimentício (quimosina) e a outra é uma enzima

detergente (protease crio-fílica). A diversidade dos produtos faz com que os processos de purificação sejam diferentes; assim, a vacina e a protease, por serem intracelulares, precisam passar por um homogeneizador que rompe as células e libera as proteínas, e por um filtro para remover os restos das células; além disso, os produtos medicinais requerem vários passos cromatográficos para obter a pureza necessária. Na Figura 4.7 é apresentado o fluxograma simplificado para a planta multiproduto do problema analisado.

Figura 4.7 Esquema do processo multiproduto analisado por Pinto *et al.* (2001).

No projeto da planta considera-se a possibilidade de ter equipamentos em paralelo operados em-fase ou fora-de-fase, e comparam-se os resultados obtidos quando se empregam modelos de desempenho ou parâmetros fixos, e quando se alocam ou não tanques intermediários; assim, o projeto é desenvolvido para quatro cenários diferentes: 1) sem tanques e usando modelos de desempenho, 2) sem tanques e parâmetros constantes, 3) com tanques e modelos de desempenho, e 4) com tanques e parâmetros constantes. Os problemas de otimização gerados são modelos MINLP não-convexos, nos quais as variáveis binárias são as decisões para o projeto estrutural da planta.

Os autores observam que o uso de modelos de desempenho gera economias de 15% quando não se consideram tanques intermediários (cenário 1: US\$ 1.505.326 contra cenário 2: US\$ 1.770.418) e 13% quando consideram-se os tanques (cenário 3: US\$ 800.138 contra cenário 4: US\$ 920.790). Adicionalmente, a aplicação dos modelos gerou redução significativa no tempo ocioso dos equipamentos. Assim, os autores concluem que o uso de modelos de desempenho em lugar de parâmetros fixos permite a obtenção de melhores soluções, devido aos graus de liberdade que são adicionados no problema.

Esses trabalhos confirmam que a síntese e projeto de plantas biotecnológicas é um problema de extrema relevância, pois as técnicas existentes estão longe de ser infalíveis e abrangentes, e sua solução por técnicas de programação matemática deve melhorar a competitividade dos processos, os quais se encontram em contínua evolução, devido aos avanços tanto na expressão como na purificação dos produtos. Este problema é abordado nos capítulos seguintes, mediante a formulação de modelos de programação mista-inteira não-linear.

CAPÍTULO 5 SOLUÇÃO DO PROBLEMA DE SÍNTESE E PROJETO DE BIOPROCESSOS MULTIPRODUTO

Este capítulo aborda a solução do problema de síntese e projeto ótimo de plantas bioquímicas multiproduto com linha de produção única. O problema considera tanto a escolha entre alternativas para expressar e processar o produto como o dimensionamento dos equipamentos escolhidos, considerando que vários produtos não podem estar simultaneamente no mesmo estágio de produção.

As seções 5.1 e 5.2 apresentam as características gerais do problema e a sua formulação, que inclui a modelagem dos arranjos de equipamentos em paralelo e das decisões da síntese, os modelos gerais para os estágios, as restrições para garantir as metas de produção, a função objetivo, o tratamento matemático do modelo para permitir a sua solução e a alocação de tanques intermediários. As seções 5.3 e 5.4 apresentam, respectivamente, o estudo de caso e a reformulação do problema por abordagem *big-M*. As seções 5.5 e 5.6 apresentam os resultados obtidos e as conclusões. Finalmente, a seção 5.7 apresenta a nomenclatura empregada ao longo deste capítulo e do seguinte.

5.1 Características gerais do problema

As características gerais do problema analisado (síntese e projeto de uma planta para biossíntese de múltiplas proteínas recombinantes) são:

- A planta deve ser projetada para satisfazer a meta de produção Q_i para cada um dos produtos i ($i = 1, \dots, P$) com o menor custo possível.
- Para gerar cada produto existem vários hospedeiros h possíveis, e a escolha da melhor opção faz parte do problema de síntese, introduzindo variáveis binárias no modelo matemático.
- A planta é composta por uma sequência de estágios j , necessários para a geração de cada produto. Esta sequência, assim como alguns parâmetros do modelo, depende do hospedeiro escolhido.

- Em alguns estágios existem várias operações capazes de realizar a tarefa. Por exemplo, para separação sólido/líquido, pode-se empregar centrifugação ou filtração; a escolha da operação também é introduzida no modelo através de variáveis binárias.
- Os estágios podem ser descontínuos ou semi-contínuos; os estágios descontínuos possuem só o equipamento em batelada enquanto os estágios semi-contínuos possuem um tanque para alimentação, o equipamento contínuo e eventualmente um tanque para o produto.
- As unidades são representadas através de modelos de desempenho, os quais empregam fatores de dimensionamento e de trabalho (*size factors* e *duty factors*) para os estágios descontínuos e semi-contínuos, respectivamente.
- Em cada estágio pode-se projetar um arranjo de equipamentos em paralelo, operando em-fase ou fora-de-fase. O arranjo em-fase visa a divisão da batelada, sendo útil quando o tamanho da batelada supera a capacidade máxima do equipamento; por outro lado, o arranjo fora-de-fase visa a diminuição do tempo ocioso nos equipamentos, já que diminui o tempo de operação nos estágios que restringem o tempo de ciclo.
- Tanques intermediários podem ser projetados entre estágios consecutivos, desvinculando a operação em seções com diferentes tempos de ciclo e tamanhos de batelada.

5.2 Modelagem matemática

No presente trabalho, o problema de síntese e projeto de processos bioquímicos é modelado como um problema MINLP no qual as variáveis binárias correspondem às decisões de alternativas: os hospedeiros para a produção, as operações (nos estágios com alternativas) e o número de equipamentos em paralelo. Os parâmetros de dimensionamento das unidades são estimados com modelos de desempenho, os quais permitem quantificar os efeitos que as decisões tomadas no projeto de outros estágios têm sobre o desempenho de cada unidade.

5.2.1 Arranjos de equipamentos em paralelo

As restrições que a capacidade dos equipamentos e seu tempo de operação impõem ao processamento geram a necessidade de projetar arranjos de múltiplos equipamentos, em lugar de equipamentos individuais para alguns estágios do processo produtivo. Quando a dimensão máxima do equipamento não permite o processamento da batelada, equipamentos operando em-fase devem ser projetados; estes equipamentos processam frações iguais da batelada e, como o nome indica, trabalham simultaneamente. Quando o tempo de operação de um estágio restringe o tempo de ciclo do processo, e gera tempos ociosos em outros equipamentos, o projeto de unidades trabalhando fora-de-fase pode aumentar a produtividade da planta e reduzir a ociosidade.

Para ilustrar a modelagem e os benefícios dos arranjos de múltiplos equipamentos, apresenta-se um processo hipotético com dois estágios para a produção de um único produto, ilustrado na Figura 5.1.

Figura 5.1 Efeito do projeto de equipamentos em paralelo na produtividade: a) projeto com equipamentos individuais, b) projeto com unidades em paralelo.

No projeto com equipamentos individuais, Figura 5.1a, o tempo de operação do primeiro estágio (T_1) define o tempo de ciclo (TL) e gera um grande tempo ocioso no

segundo estágio, enquanto o volume máximo do equipamento no segundo estágio (V_2^U) restringe o tamanho máximo da batelada (B^U). Se fossem projetadas duas unidades iguais fora-de-fase no primeiro estágio ($M_1 = 2$) e três em-fase no segundo ($G_2 = 3$), Figura 5.1b, o tamanho da batelada poderia ser aumentado, os tempos de ciclo e ociosos seriam diminuídos e a produtividade máxima aumentaria em cinco vezes.

5.2.2 Modelagem das decisões na síntese do processo

Considerando a existência de opções para expressar cada produto, definem-se variáveis binárias y_{ih} para a escolha dos hospedeiros; se o hospedeiro h é escolhido para gerar o produto i , então $y_{ih} = 1$. Para garantir que um único hospedeiro seja escolhido para cada produto, emprega-se a seguinte restrição:

$$\sum_{h \in H_i} y_{ih} = 1 \quad \forall i \quad (5.1)$$

onde H_i denota o conjunto de hospedeiros que podem expressar i . Estudos anteriores (Montagna *et al.*, 2004) modelam a seleção de alternativas em cada estágio de forma independente, considerando um grupo de variáveis binárias z_{ihjd} , em que $z_{ihjd} = 1$, se a opção d é escolhida para o processamento no estágio j do produto i expresso pelo hospedeiro h ; porém esta abordagem dificulta a formulação das restrições para os equipamentos nos quais a operação depende das escolhas feitas em outros estágios. A alternativa empregada para contornar este obstáculo é definir os parâmetros de processamento em função da rota de produção e não só da escolha no estágio.

Uma rota de produção é caracterizada por uma seqüência de operações que permitem o processamento; dessa forma, ela envolve a seleção entre as opções existentes em cada um dos estágios necessários. Como a seqüência de estágios é definida pelo hospedeiro (ver características gerais do problema, na seção 5.1), o número de rotas é estimado por:

$$\prod_{j \in J_h} ND_j \quad \forall h \quad (5.2)$$

onde J_h é o conjunto de estágios para a produção quando se emprega o hospedeiro h e ND_j é o número de opções no estágio j . Para ilustrar a definição das rotas de produção em função das escolhas nos estágios, considera-se uma sequência com cinco estágios ($j = 1, \dots, 5$) na qual há duas opções de processamento em cada um dos estágios 2 a 4, de acordo com a Figura 5.2.

Figura 5.2 Exemplo de sequência de processamento com opções.

A seleção das rotas de produção emprega variáveis binárias (yr_{ihr}) que definem a rota r escolhida para a produção de i com o hospedeiro h . Para garantir que uma única rota seja escolhida para cada produto, empregam-se as seguintes restrições:

$$\sum_{h \in H_i} \sum_{r \in R_h} yr_{ihr} = 1 \quad \forall i \quad (5.3)$$

$$\sum_{r \in R_h} yr_{ihr} - y_{ih} = 0 \quad \forall i, h \in H_i \quad (5.4)$$

onde H_i é o conjunto de hospedeiros que podem expressar o produto i , e R_h é o conjunto de rotas possíveis para o hospedeiro h . Por definição, a rota de produção caracteriza as escolhas em todos os estágios; dessa forma, a opção d no estágio j para o produto i expresso por h é escolhida (*i.e.* $z_{ihjd}=1$ na modelagem de Montagna *et al.*, 2004) se uma das rotas que consideram esta escolha é selecionada, *i.e.* se $\sum_{r \in R_{jd}} yr_{ihr} = 1$ onde R_{jd} é o grupo de rotas possíveis quando se escolhe a opção d no estágio j . As rotas possíveis para o exemplo da Figura 5.2 são enumeradas na Tabela 5.1. De acordo com a Tabela 5.1, os conjuntos de rotas possíveis são: $R_{1,1} = \{1, \dots, 8\}$; $R_{2,1} = \{1, \dots, 4\}$; $R_{2,2} = \{5, \dots, 8\}$; $R_{3,1} = \{1, 2, 5, 6\}$; $R_{3,2} = \{3, 4, 7, 8\}$; $R_{4,1} = \{1, 3, 5, 7\}$; $R_{4,2} = \{2, 4, 6, 8\}$, e $R_{5,1} = \{1, \dots, 8\}$.

Tabela 5.1 Definição das rotas de produção no exemplo da Figura 5.2.

Rota de produção	Estágio/opção de operação				
	Estágio 1	Estágio 2	Estágio 3	Estágio 4	Estágio 5
1	1	1	1	1	1
2	1	1	1	2	1
3	1	1	2	1	1
4	1	1	2	2	1
5	1	2	1	1	1
6	1	2	1	2	1
7	1	2	2	1	1
8	1	2	2	2	1

5.2.3 Modelagem do efeito das decisões no desempenho das unidades

Visto que a modelagem das unidades emprega modelos de desempenho de processo, a estimativa da dimensão dos equipamentos e do tempo de processamento requer a definição de parâmetros de dimensionamento e tempo (S_{ihjr} e T_{ihjr}^1). No presente trabalho, estes parâmetros são estimados de forma simples, o que permite, porém, considerar efeitos das decisões tomadas em outros estágios. Tais parâmetros consideram uma constante, que depende da procedência do fluxo alimentado, e os rendimentos dos estágios posteriores, que dependem das escolhas nos mesmos:

$$S_{ihjr} = \frac{CS_{ihjr}}{\prod_{\substack{k \geq j \\ k \in J_h}} h_{ihkr}} \quad \forall i, h \in H_i, j \in J_h, r \in R_i \quad (5.5)$$

$$T_{ihjr}^1 = \frac{CT_{ihjr}^1}{\prod_{\substack{k \geq j \\ k \in J_h}} h_{ihkr}} \quad \forall i, h \in H_i, j \in J_h, r \in R_i \quad (5.6)$$

onde CS_{ihjr} , CT_{ihjr}^1 e h_{ihkr} são as constantes para os parâmetros de dimensionamento e tempo, e o rendimento para o produto i , hospedeiro h no estágio k ou j , considerando a rota de produção r . Durante a otimização, a escolha dos parâmetros de operação requer a definição do seguinte grupo de variáveis:

$$S_{ihjd}^* = \sum_{r \in R_{jd}} y_{ihr} S_{ihjr} \quad \forall i, h \in H_i, j \in J_h, d \in D_j \quad (5.7)$$

$$T_{ihjd}^{1*} = \sum_{r \in R_{jd}} y_{r_{ih}} T_{ihjr}^1 \quad \forall i, h \in H_i, j \in J_h, d \in D_j \quad (5.8)$$

Devido ao efeito das decisões nos parâmetros, e visando eliminar restrições e variáveis desnecessárias, os cálculos são desenvolvidos na seguinte sequência:

Previamente à etapa de otimização:

1. Definem-se todas as rotas de produção possíveis para cada hospedeiro.
2. Para os estágios nos quais a operação depende das decisões em outros estágios, estimam-se os parâmetros em função das rotas (S_{ihjr} e T_{ihjr}^1) pelas equações (5.5) e (5.6).
3. Para os estágios nos quais a operação não depende das decisões em outros estágios estimam-se os parâmetros em função da opções no estágio (S_{ihjd} e T_{ihjd}^1):

$$S_{ihjd} = \frac{CS_{ihjd}}{\prod_{\substack{k \geq j \\ k \in J_h}} h_{ihkd}} \quad \forall i, h \in H_i, j \in J_h, d \in D_j \quad (5.9)$$

$$T_{ihjd}^1 = \frac{CT_{ihjd}^1}{\prod_{\substack{k \geq j \\ k \in J_h}} h_{ihkd}} \quad \forall i, h \in H_i, j \in J_h, d \in D_j \quad (5.10)$$

Decisões de otimização

4. Estimam-se as variáveis de operação nos estágios dependentes da síntese (equações 5.7 e 5.8).
5. As restrições para operação dos equipamentos são construídas com as variáveis estimadas nas equações (5.7) e (5.8) ou com os parâmetros de (5.9) e (5.10), dependendo do estágio.

5.2.4 Modelagem de estágios descontínuos

Em estágios descontínuos deve-se estimar o volume do equipamento para a opção d no estágio j (V_{jd}) e o tempo para processar cada produto i gerado pelo hospedeiro h (T_{ihjd}). O volume deve permitir o processamento da batelada de qualquer produto i , sendo que o volume necessário para processar uma batelada de i é proporcional ao tamanho da

batelada total (B_i) dividida pelo número de unidades operando em-fase (G_{jd}), e a constante de proporcionalidade é o parâmetro de dimensionamento (S_{ihjd} ou S_{ihjd}^* , dependendo do estágio). Esse fator define o volume necessário para processar uma unidade de massa de produto. O tempo de processamento é estimado pela soma de dois termos: o primeiro considera o tempo fixo T_{ihjd}^0 e o segundo é proporcional ao volume da batelada dividido pelo número de unidades em-fase, com o parâmetro de tempo (T_{ihjd}^1 ou T_{ihjd}^{1*} , dependendo do estágio). Assim, os estágios descontínuos são modelados por:

$$V_{jd} \geq \frac{S_{ihjd} B_i}{G_{jd}} \quad \forall i, h \in H_i, j \in (J_b \cap J_h), d \in D_j \quad (5.11)$$

$$T_{ihjd} = T_{ihjd}^0 + T_{ihjd}^1 \frac{B_i}{G_{jd}} \quad \forall i, h \in H_i, j \in (J_b \cap J_h), d \in D_j \quad (5.12)$$

onde H_i é o conjunto de hospedeiros possíveis para o produto i , J_b é o conjunto de estágios descontínuos, J_h é o conjunto de estágios necessários quando se escolhe o hospedeiro h , e D_j é o conjunto de opções para o estágio j .

5.2.5 Modelagem de estágios semi-contínuos

Estágios semi-contínuos possuem um tanque de alimentação, a unidade semi-contínua e eventualmente um tanque para o produto. No intuito de reduzir custos, pode-se projetar um tanque de alimentação e um de produto (se necessário) para cada grupo de unidades que operam em-fase. Assim, os volumes dos tanques ($V1_{jd}$ e $V2_{jd}$) são estimados como o produto do fator de dimensionamento pelo volume da batelada total.

$$V1_{jd} \geq S1_{ihjd} B_i \quad V2_{jd} \geq S2_{ihjd} B_i \quad \forall i, h \in H_i, j \in (J_s \cap J_h), d \in D_j \quad (5.13)$$

onde J_s é o conjunto dos estágios semi-contínuos. A Figura 5.3 exemplifica o projeto para um estágio semi-contínuo com dois grupos de unidades operando fora-de-fase ($M = 2$), cada um com três unidades operando em-fase ($G = 3$).

Figura 5.3 Configuração de um estágio semi-contínuo com unidades em paralelo.

Na unidade semi-contínua, deve-se estimar a dimensão característica (R_{jd}) que corresponde à taxa de processamento (m^3/h) para um homogeneizador ou moinho, à potência (kW) para uma centrífuga ou à área (m^2) para unidades de filtração; em qualquer desses casos, a dimensão é proporcional à batelada processada em cada unidade, dividida pelo tempo de processamento (q_{ihjd}); a proporcionalidade é dada pelo fator de trabalho (D_{ihjd}).

$$R_{jd} \geq \frac{D_{ihjd} B_i}{q_{ihjd} G_{jd}} \quad \forall i, h \in H_i, j \in (J_s \cap J_h), d \in D_j \quad (5.14)$$

Salomone *et al.* (1994) propõem que nos estágios com unidades semi-contínuas alimentadas por descontínuas, estime-se o volume da unidade descontínua (equação 5.13) e o tempo de operação, que inclui o tempo fixo (T_{ihjd}^0) e o tempo de processamento (q_{ihjd}). Assim:

$$T_{ihjd} = T_{ihjd}^0 + T_{ihjd}^1 \frac{B_i}{G_{jd} R_{jd}} \quad \forall i, h \in H_i, j \in (J_s \cap J_h), d \in D_j \quad (5.15)$$

onde T_{ihjd}^1 corresponde a D_{ihjd} .

5.2.6 Modelagem de estgios cromatogrficos

Estgios cromatogrficos operam em forma descontnua, mas possuem tanques de alimentao e produto, portanto as colunas so dimensionadas de forma anloga aos reatores:

$$R_{jd} \geq \frac{S_{ihjd} B_i}{G_{jd}} \quad \forall i, h \in H_i, j \in (J_{cr} \cap J_h), d \in D_j \quad (5.16)$$

Para os tanques emprega-se a restrio (5.13), enquanto que o tempo de processamento  estimado com a equao (5.15), pois depende da dimenso da coluna.

5.2.7 Restries para garantir o cumprimento das metas de produo

Para garantir o cumprimento da meta de produo Q_i de cada produto i no horizonte de tempo HT , emprega-se a seguinte restrio:

$$\sum_i \frac{Q_i TL_i}{B_i} \leq HT \quad (5.17)$$

onde TL_i  o tempo de ciclo para o produto i , que  estimado como o maior tempo de processamento nos estgios envolvidos:

$$TL_i \geq \frac{T_{ihjd}}{M_{jd}} \quad \forall i, h \in H_i, j \in J_h, d \in D_j \quad (5.18)$$

onde T_{ihjd}  obtido com a equao (5.12) ou (5.15), dependendo do tipo de estgio, e M_{jd}  o nmero de unidades fora-de-fase para a opo d no estgio j .

5.2.8 Funo objetivo

O objetivo no problema  minimizar o custo de capital, que  dado pela soma dos custos de instalao dos equipamentos em todos os estgios. Os estgios descontnuos (J_b) so constitudos pela unidade de processamento com volume V_{jd} , enquanto os semi-contnuos e cromatogrficos (J_s, J_{cr}) incluem a unidade de processamento com dimenso R_{jd} , e os tanques de alimentao e produto:

$$\begin{aligned} \min \text{custo} = & \sum_{j \in J_b} \sum_{d \in D_j} \mathbf{a}_{jd} M_{jd} G_{jd} V_{jd}^{b_{jd}} + \\ & \sum_{j \in J_s \cup J_{cr}} \sum_{d \in D_j} (\mathbf{a}_{jd} M_{jd} V I_{jd}^{b_{jd}} + \mathbf{l}_{jd} M_{jd} G_{jd} R_{jd}^{w_{jd}} + \mathbf{a}_{jd} M_{jd} V 2_{jd}^{b_{jd}}) \end{aligned} \quad (5.19)$$

onde $(\mathbf{a}_{jd}, \mathbf{b}_{jd})$ e $(\mathbf{l}_{jd}, \mathbf{w}_{jd})$ são coeficientes de custo para as unidades descontínuas e semi-contínuas, respectivamente.

5.2.9 Formulação do problema de síntese e projeto ótimo de bioprocessos por programação disjuntiva generalizada (GDP)

A representação baseada em Programação Disjuntiva Generalizada-GDP (Apêndice C) pode ser aplicada à modelagem do problema de síntese e projeto, gerando a seguinte formulação:

$$\begin{aligned} \min \text{custo} \quad & (5.19) \\ \text{s.a} \quad & \sum_i \frac{Q_i T L_i}{B_i} \leq HT \\ & \left[\begin{array}{c} \bigvee_{h \in H_i} \left[\begin{array}{c} \bigvee_{r \in R_h} \left[\begin{array}{c} YR_{ihr} \\ V_{jd} \geq \frac{S_{ihjr} B_i}{G_{jd}} \quad \forall j \in J_b \cap J_h \\ TL_i \geq \frac{T_{ihjd}^0 + T_{ihjr}^1 \frac{B_i}{G_{jd}}}{M_{jd}} \quad \forall j \in J_b \cap J_h \\ R_{jd} \geq \frac{S_{ihjr} B_i}{G_{jd}} \quad \forall j \in J_{cr} \cap J_h \\ VI_{jd} \geq S I_{ihjr} B_i \quad \forall j \in (J_s \cup J_{cr}) \cap J_h \\ V2_{jd} \geq S2_{ihjr} B_i \quad \forall j \in (J_s \cup J_{cr}) \cap J_h \\ TL_i \geq \frac{T_{ihjd}^0 + T_{ihjr}^1 \frac{B_i}{G_{jd} R_{jd}}}{M_{jd}} \quad \forall j \in (J_s \cup J_{cr}) \cap J_h \end{array} \right] \end{array} \right] \quad \forall i \end{array} \right] \quad (5.20) \\ & Y_{ih} \Leftrightarrow \sum_{r \in R_h} YR_{ihr} \quad \forall i, h \in H_i \end{aligned}$$

Nesta formulação, observa-se que a restrição para garantir o cumprimento das metas de produção independe das escolhas, porém a dimensão dos equipamentos, o

tempo de operação para cada estágio, e em consequência o tempo de ciclo para cada produto (TL_i), dependem do hospedeiro e da rota de produção escolhidos; o que é formulado como uma dupla disjunção que depende das variáveis Booleanas. A proposição lógica indica que, quando o hospedeiro h é escolhido para o produto i (*i.e.* $Y_{ih} = \text{verdadeiro}$), uma rota de produção dentro das possíveis (R_h) deve ser escolhida. De outro modo, se o hospedeiro h não é escolhido, nenhuma rota deve ser considerada.

5.2.10 Redução das não-convexidades do problema

As não-convexidades dificultam, ou até impossibilitam, a solução de um problema de otimização. Devido à simplicidade das funções que definem os parâmetros de operação (equações 5.5 e 5.6 ou 5.9 e 5.10, dependendo do estágio), as não-linearidades nas restrições do problema limitam-se a produtos/quocientes entre variáveis, as quais, na maioria dos casos podem ser convexificadas com transformações logarítmicas:

$$v_{jd} = \ln V_{jd} \quad \forall j \in (J_b \cap J_h), d \in D_j \quad (5.21)$$

$$v1_{jd} = \ln V1_{jd} \quad v2_{jd} = \ln V2_{jd} \quad \forall j \in ((J_s \cup J_{cr}) \cap J_h), d \in D_j \quad (5.22)$$

$$r_{jd} = \ln R_{jd} \quad \forall j \in ((J_s \cup J_{cr}) \cap J_h), d \in D_j \quad (5.23)$$

$$b_i = \ln B_i \quad \forall i \quad (5.24)$$

$$tl_i = \ln TL_i \quad \forall i \quad (5.25)$$

$$m_{jd} = \ln M_{jd} \quad \forall j, d \in D_j \quad (5.26)$$

$$g_{jd} = \ln G_{jd} \quad \forall j, d \in D_j \quad (5.27)$$

Com estas transformações, os produtos/quocientes entre variáveis são substituídos por somas/diferenças entre variáveis transformadas, como mostrado a seguir:

$$\mathbf{a}_{jd} M_{jd} G_{jd} V_{jd}^{b_{jd}} \Rightarrow \mathbf{a}_{jd} \exp(m_{jd} + g_{jd} + \mathbf{b}_{jd} v_{jd}) \quad (5.28)$$

$$V_{jd} \geq \frac{S_{ihjd} B_i}{G_{jd}} \xRightarrow{\ln} v_{jd} \geq \ln(S_{ihjd}) + b_i - g_{jd} \quad (5.29)$$

No entanto, nas restrições de estimação de tempo de processamento (equações 5.12 e 5.15), quando os dois termos que definem o tempo devem ser considerados, ainda restam termos não-convexos. Como M_{jd} e G_{jd} são variáveis inteiras, é necessário desenvolver transformações adicionais para tratá-las como variáveis binárias:

$$m_{jd} = \sum_n c_n ym_{jdn} \quad \forall j, d \in D_j \quad (5.30)$$

$$g_{jd} = \sum_n c_n yg_{jdn} \quad \forall j, d \in D_j \quad (5.31)$$

onde ym_{jdn} e yg_{jdn} são variáveis binárias e $c_n = \ln n$. Assim, $ym_{jdn} = 1$ se $M_{jd} = n$, i.e. n unidades operando fora-de-fase são projetadas para a opção d do estágio j ; o raciocínio é análogo para yg_{jdn} . Para garantir que só uma configuração seja escolhida em cada estágio, devem ser adicionadas as seguintes restrições:

$$\sum_n ym_{jdn} = 1 \quad \text{e} \quad \sum_n yg_{jdn} = 1 \quad \forall j, d \in D_j \quad (5.32)$$

5.2.11 Alocação de tanques intermediários em posições predefinidas

Quando um tanque intermediário ou tanque pulmão é alocado entre os estágios k e $k+1$, o processo é desacoplado em duas seções, as quais devem ter taxas médias de produção iguais, pois o tanque não tem função de estocar produto, mas o tempo de ciclo e o tamanho da batelada podem ser diferentes. Assim, a restrição do tempo de ciclo (equação 5.18) deve ser independente para cada sub-processo:

$$TL_{t,1} \geq \frac{T_{ihjd}}{M_{jd}} \quad \forall i, h \in H_i, j \in J_h, j \leq k, d \in D_j \quad (5.33)$$

$$TL_{t,2} \geq \frac{T_{ihjd}}{M_{jd}} \quad \forall i, h \in H_i, j \in J_h, j > k, d \in D_j \quad (5.34)$$

onde $TL_{i,1}$ e $TL_{i,2}$ são os tempos de ciclo para o sub-processo anterior e posterior ao tanque intermediário, respectivamente. A restrição das metas de produção também deve ser duplicada:

$$\sum_i \frac{Q_i TL_{i,1}}{B_{i,1}} \leq HT \quad \text{e} \quad \sum_i \frac{Q_i TL_{i,2}}{B_{i,2}} \leq HT \quad (5.35)$$

onde $B_{i,1}$ e $B_{i,2}$ são os tamanhos da batelada para cada seção. A alocação de tanques pulmão visa obter um projeto mais eficiente aumentando a utilização dos equipamentos. Como a função destes tanques não é estocagem, as taxas de produção nas duas seções do processo devem ser idênticas:

$$\frac{B_{i,1}}{TL_{i,1}} = \frac{B_{i,2}}{TL_{i,2}} \quad \forall i \quad (5.36)$$

Uma abordagem simples para o dimensionamento destes tanques, apresentada em Ravemark e Rippin (1998), sugere que o volume seja o dobro da maior batelada, o qual é modelado como:

$$VT_k \geq 2 ST_{ih} B_{i,1} \quad \forall i, h \in H_i \quad (5.37)$$

$$VT_k \geq 2 ST_{ih} B_{i,2} \quad \forall i, h \in H_i \quad (5.38)$$

onde VT_k é o volume do tanque na posição k e ST_{ih} é o fator de dimensionamento. Adicionalmente, a razão dos tamanhos de batelada das seções contíguas deve ser restrita para evitar diferenças exageradas entre os tamanhos de batelada. Para isto empregam-se as seguintes restrições:

$$\frac{1}{F} \leq \frac{B_{i,1}}{B_{i,2}} \leq F \quad \forall i \quad (5.39)$$

Com a alocação de tanques intermediários, um termo para estimar o custo de cada tanque, $a_k VT_k^{b_k}$, deve ser adicionado à função objetivo; o qual deve sofrer a mesma transformação exponencial das variáveis, item 5.2.10, conforme mostrado em 5.4.2

5.2.12 Alocação de tanques intermediários em posições otimizadas

Considera-se a possibilidade de alocar tanques-pulmão em cada uma das possíveis posições, e deseja-se otimizar a decisão de alocação. Assim, deve-se incluir um grupo de

variáveis binárias ytk_j , em que $ytk_j = 1$ se um tanque é alocado na posição j (*i.e.* entre os estágios j e $j+1$); com esta consideração, os tempos de ciclo e os tamanhos de batelada devem ser definidos para cada estágio ($TL_{i,j}$ e $B_{i,j}$). Como na alocação de tanques em posições predefinidas, deve-se garantir que a taxa de produção seja constante em todos os sub-processos. Montagna *et al.*, (2000) sugerem uma forma simples para expressar a restrição da taxa de produção entre os sub-processos:

$$E_i = \frac{TL_{i,j}}{B_{i,j}} \quad \forall i, j \quad (5.40)$$

onde E_i é o inverso da taxa de produção para o produto i . Com isto as restrições para o horizonte de tempo podem ser unificadas:

$$\sum_i Q_i E_i \leq HT \quad (5.41)$$

A restrição para a relação entre os tamanhos de batelada deve considerar que, se na posição j o tanque não é alocado (*i.e.* $ytk_j = 0$), as bateladas dos estágios j e $j+1$ devem ser idênticas. Assim, a restrição torna-se:

$$1 + \left(\frac{1}{F} - 1 \right) ytk_j \leq \frac{B_{i,j}}{B_{i,j+1}} \leq 1 + (F - 1) ytk_j \quad (5.42)$$

5.3 Estudo de caso

Como estudo de caso aborda-se a síntese e projeto de uma planta biotecnológica para produção de insulina, vacina para hepatite B, ativador de plasminogênio tecidual ou *tissue plaminogen activator* (tPA) e superóxido dismutase (SOD), mediante três hospedeiros diferentes: levedura (*S. cerevisiae*) com expressão extra ou intracelular, *Escherichia coli* e células de mamíferos ou *chinese hamster ovary cells* (CHO). Este estudo de caso foi desenvolvido e foi resolvido em Montagna *et al.* (2004) com reformulação das disjunções por abordagem *big-M* e *convex hull*. A Tabela 5.2 apresenta os produtos com sua demanda anual (meta de produção) e os hospedeiros que podem produzi-los.

Tabela 5.2 Produtos e hospedeiros para o estudo de caso.

Produto	Sigla	meta de produção (kg/ano)	hospedeiro		
			<i>S. cerevisiae</i>		<i>E. coli</i> CHO
			extracel.	intracel.	
Insulina	INS	1500	x		x
Vacina para hepatite B	HBV	1000		x	x
Ativador de plasminogênio tecidual	tPA	10			x
Superóxido dismutase	SOD	200		x	x

Os estágios requeridos para o processamento de cada produto dependem do hospedeiro escolhido, e em algumas etapas é necessário escolher entre várias opções existentes, como apresentado na Tabela 5.3. Para qualquer produto, o processo possui três fases: a geração do produto (estágio 1), a recuperação (estágios 2 a 9) e a purificação (estágios 10 a 15); no entanto, dependendo do hospedeiro escolhido, alguns estágios na fase de recuperação do produto são desnecessários. As seqüências de estágios para este estudo de caso são semelhantes às sugeridas por outros autores, apresentadas no item 4.2.6.

Tanto no estágio 2 (remoção/coleta de células) como no 4 (remoção de restos celulares ou coleta dos corpos de inclusão), as opções consideradas são centrifugação ou microfiltração, enquanto no estágio 3 (rompimento de células) as opções são homogeneização e moinho de bolas. Os estágios 2, 3, 4, 6, 9, 11, 13 e 15 possuem uma unidade semi-contínua que processa o material procedente de um tanque (*holding vessel*). Adicionalmente, para os estágios 2, 4 e 15 devem-se projetar tanques de produto; porém se os estágios 2 e 4 são desenvolvidos por microfiltração e o produto encontra-se no fluxo de retido (ver item 4.2.1) o tanque de produto é desnecessário. Os demais estágios (1, 5, 7, 8, 10, 12, 14) operam de forma descontínua, porém os estágios cromatográficos (10, 12, 14) possuem tanques de alimentação e produto.

Tabela 5.3 Estágios necessários para a produção, dependendo do hospedeiro.

Estágio	Operação	Hospedeiro			
		<i>S. cerevisiae</i>		<i>E. coli</i>	CHO
		extracel.	intracel.		
1	Fermentação	X	X	X	X
2	Separação de células	X	X	X	X
	A. Centrifugação				
	B. Microfiltração				
3	Rompimento de células		X	X	
	A. Homogeneização				
	B. Moinho de bolas				
4	Separação sólido/líquido		X	X	
	A. Centrifugação				
	B. Microfiltração				
5	Dissolução de IB			X	
6	Diafiltração			X	
7	Sulfonação			X	
8	<i>Refolding</i>			X	
9	Ultrafiltração			X	X
10	Cromatografia	X	X	X	X
11	Ultra-diafiltração	X	X	X	X
12	Cromatografia	X	X	X	X
13	Ultra-diafiltração	X	X	X	X
14	Gel filtração	X	X	X	X
15	Filtração estéril	X	X	X	X

As características de cada equipamento (rendimento, parâmetros de dimensionamento, funções para estimar o tempo de processamento e o custo, e os limites na dimensão característica) são apresentadas no Apêndice E.

Para o estudo de caso foram avaliados diferentes modelos. O primeiro considera a síntese e projeto de plantas monoproducto, e foi desenvolvido para servir como base de comparação dos seguintes modelos; o segundo modelo considera a síntese e projeto de uma planta multiproducto. O modelo multiproducto foi resolvido com e sem alocação de tanques intermediários. Os cenários gerados são apresentados na Tabela 5.4.

Tabela 5.4 Características dos cenários resolvidos para o estudo de caso.

Modelo	Cenário
I Monoproducto	1. Sem tanques intermediários
	2. Sem tanques intermediários
II Multiproducto	3. Tanques em posições prefixadas
	4. Tanques em posições otimizadas

Tanto os cenários acima enumerados como os do capítulo 6 foram resolvidos para um horizonte de tempo (*HT*) de 6000 horas.

5.4 Reformulação do problema por abordagem *big-M*

Como apresentado no item 5.2.9, a existência de opções para o processamento envolve a escolha entre diversos grupos de equações para o dimensionamento das unidades; formulada mediante a abordagem *big-M* (Apêndice D). A seguir, descreve-se o problema de otimização MINLP gerado do estudo de caso da seção anterior. Inicialmente considera-se o modelo básico para a planta multiproducto, posteriormente, nos itens 5.4.2, 5.4.3 e 5.4.4, descrevem-se as modificações no modelo para a alocação de tanques intermediários e as considerações adicionais para a implementação do modelo.

5.4.1 Modelo básico para a planta multiproducto

À continuação apresenta-se a modelagem básica para a síntese e projeto da planta multiproducto, aproveitando as características do modelo de cada estágio para reduzir a complexidade do problema. Com a transformação das variáveis, a função objetivo (5.19) toma-se:

$$\min \text{custo} = \sum_{j \in J_b} \mathbf{a}_{j,1} \exp(m_{j,1} + g_{j,1} + \mathbf{b}_{j,1} v_{j,1}) + \sum_{j \in \{J_s \cup J_{cr}\}} \sum_{d \in D_j} [\mathbf{a}_{jd} \exp(m_{jd} + \mathbf{b}_{jd} v_{jd}) + \mathbf{l}_{jd} \exp(m_{jd} + g_{jd} + \mathbf{w}_{jd} r_{jd}) + \mathbf{a}_{jd} \exp(m_{jd} + \mathbf{b}_{jd} v_{jd}^2)] \quad (5.43)$$

A primeira somatória inclui os estágios descontínuos $J_b = \{1, 5, 7, 8\}$, os quais não requerem tanques; como nestes estágios não há alternativas, não é necessário desenvolver a somatória em d , e este índice torna-se fixo ($d = 1$). A segunda somatória considera os

estágios que possuem tanques $J_s \cup J_{cr} = \{2, 3, 4, 6, 9, 10, 11, 12, 13, 14, 15\}$; em alguns destes estágios há alternativas, o que torna necessária a somatória em d . O primeiro e o terceiro elemento desta somatória referem-se aos custos dos tanques de alimentação e produto, pelo qual não consideram o número de unidades que operam em fase (g_{jd}), como estabelecido no item 5.2.5.

Para formular as restrições, os estágios são agrupados segundo as suas características, o que permite reduzir o número de variáveis e de restrições e, conseqüentemente, a complexidade numérica do problema. O biorreator é o único estágio descontínuo em que o parâmetro de dimensionamento depende da rota de produção. Então, a restrição para o volume torna-se:

$$v_{jd} \geq \ln(S_{ihjd}^*) + b_i - g_{jd} - \text{Ml}_{ihjd}(1 - y_{ih}) \quad \forall i, h \in H_i, j = 1, d \in D_j \quad (5.44)$$

onde a variável S_{ihjd}^* é estimada com a equação (5.7):

$$S_{ihjd}^* = \sum_{r \in R_{jd}} y r_{ihjr} S_{ihjr} \quad \forall i, h \in H_i, j = 1, d \in D_j \quad (5.7)$$

Para os outros equipamentos descontínuos (estágios 5, 7, 8) o volume é estimado como:

$$v_{jd} \geq \ln(S_{ihjd}) + b_i - g_{jd} - \text{Ml}_{ihjd}(1 - y_{ih}) \quad \forall i, h \in H_i, j \in (J_h \cap \{5, 7, 8\}), d \in D_j \quad (5.45)$$

onde o parâmetro S_{ihjd} é estimado com a equação (5.9). Para os estágios cromatográficos (10, 12, 14), a área da coluna (restrição 5.16) é estimada de forma semelhante:

$$r_{jd} \geq \ln(S_{ihjd}) + b_i - g_{jd} - \text{Ml}_{ihjd}(1 - y_{ih}) \quad \forall i, h \in H_i, j \in (J_h \cap J_{cr}), d \in D_j \quad (5.46)$$

Nos estágios semi-contínuos ou cromatográficos com opções ou nos quais o volume dos tanques, dado pela restrição (5.13), depende da rota de produção (estágios 2, 3, 4, 9 e 10 para o tanque de alimentação e estágios 2 e 4 para o tanque de produto), tem-se:

$$v1_{jd} \geq \ln(S1_{ihjd}^*) + b_i - M2_{ihjd} (1 - \sum_{r \in R_{jd}} yr_{ihr}) \quad \forall i, h \in H_i, j \in (J_h \cap \{2,3,4,9,10\}), d \in D_j \quad (5.47)$$

$$v2_{jd} \geq \ln(S2_{ihjd}^*) + b_i - M3_{ihjd} (1 - \sum_{r \in R_{jd}} yr_{ihr}) \quad \forall i, h \in H_i, j \in (J_h \cap \{2,4\}), d \in D_j \quad (5.48)$$

onde os parâmetros de dimensionamento tornam-se variáveis e são estimados como:

$$S1_{ihjd}^* = \sum_{r \in R_{jd}} yr_{ihr} S1_{ihjr} \quad \forall i, h \in H_i, j \in (J_h \cap \{2,3,4,9,10\}), d \in D_j \quad (5.49)$$

$$S2_{ihjd}^* = \sum_{r \in R_{jd}} yr_{ihr} S2_{ihjr} \quad \forall i, h \in H_i, j \in (J_h \cap \{2,4\}), d \in D_j \quad (5.50)$$

Nos estágios semi-contínuos ou cromatográficos nos quais o volume dos tanques independe da rota de produção (estágios 6, 11, 12, 13, 14 e 15 para o tanque de alimentação e estágios 10, 12, 14 e 15 para o de produto), tem-se:

$$v1_{jd} \geq \ln(S1_{ihjd}) + b_i - M2_{ihjd} (1 - y_{ih}) \quad \forall i, h \in H_i, j \in (J_h \cap \{6,11,12,13,14,15\}), d \in D_j \quad (5.51)$$

$$v2_{jd} \geq \ln(S2_{ihjd}) + b_i - M3_{ihjd} (1 - y_{ih}) \quad \forall i, h \in H_i, j \in (J_h \cap \{10,12,14,15\}), d \in D_j \quad (5.52)$$

onde $S1_{ihjd}$ e $S2_{ihjd}$ são parâmetros estimados de forma análoga à equação (5.9).

Para formular as restrições do tempo de ciclo, além da existência de opções, deve-se considerar as características dos estágios. Nos estágios 1, 5, 7, 8 e 14 não há alternativas, e o tempo de processamento é constante, então as restrições do tempo de ciclo, equações (5.12) e (5.15), podem ser formuladas como:

$$tl_i \geq \ln(T_{ihjd}^0) - m_{jd} - M4_{ihjd} (1 - y_{ih}) \quad \forall i, h \in H_i, j \in (J_h \cap \{1,5,7,8,14\}), d \in D_j \quad (5.53)$$

Nos estágios semi-contínuos, o tempo é estimado só com o termo dependente do tamanho da batelada. Logo nos estágios 6, 11, 13 e 15, nos quais não há efeito da rota de produção, a restrição do tempo de ciclo torna-se:

$$tl_i \geq \ln(T_{ihjd}^1) + b_i - r_{jd} - g_{jd} - m_{jd} - M5_{ihjd} (1 - y_{ih}) \quad \forall i, h \in H_i, j \in (J_h \cap \{6,11,13,15\}), d \in D_j \quad (5.54)$$

Por outro lado, nos estágios 2, 3, 4 e 9, devido ao efeito da rota de produção a restrição deve ser formulada como:

$$tl_i \geq \ln(T_{ihjd}^{1*}) + b_i - r_{jd} - g_{jd} - m_{jd} - M5_{ihjd} (1 - \sum_{r \in R_{jd}} yr_{ir}) \quad \forall i, h \in H_i, j \in (J_h \cap \{2,3,4,9\}), d \in D_j \quad (5.55)$$

onde o parâmetro do tempo de processamento torna-se variável e deve ser estimado com a equação (5.8).

$$T_{ihjd}^{1*} = \sum_{r \in R_{jd}} yr_{ir} T_{ihjr}^1 \quad \forall i, h \in H_i, j \in (J_h \cap \{2,3,4,9\}), d \in D_j \quad (5.8)$$

Nos estágios 10 e 12 o tempo é estimado com a função completa (o termo constante mais o termo dependente do tamanho da batelada). Assim, a restrição para o tempo de ciclo é:

$$\exp(tl_i + m_{jd}) \geq T_{ihjd}^0 + T_{ihjd}^1 \exp(b_i - r_{jd} - g_{jd}) - M6_{ihjd} (1 - y_{ih}) \quad \forall i, h \in H_i, j \in (J_h \cap \{10,12\}), d \in D_j \quad (5.56)$$

Finalmente, a restrição do horizonte de tempo (5.17), em função das variáveis transformadas, torna-se:

$$\sum_i Q_i \exp(tl_i - b_i) \leq HT \quad (5.57)$$

Como os valores das constantes M_i influem no desempenho dos *solvers* de otimização, seus valores não devem ser arbitrariamente elevados, mas estimados (ver Apêndice D). Para o volume do biorreator tem-se:

$$M1_{ihjd} = \ln(\max_{r \in R_{jd}} S_{ihjr}) - v_{jd}^{lo} + b_i^{up} - g_{jd}^{lo} \quad \forall i, h \in H_i, j = 1, d \in D_j \quad (5.58)$$

Para as unidades descontínuas nos estágios 5, 7 e 8, M1 é definido como:

$$M1_{ihjd} = \ln(S_{ihjd}) - v_{jd}^{lo} + b_i^{up} - g_{jd}^{lo} \quad \forall i, h \in H_i, j \in (J_h \cap \{5,7,8\}), d \in D_j \quad (5.59)$$

Analogamente, para a área nas colunas cromatográficas (estágios 10, 12 e 14), tem-se:

$$M1_{ihjd} = \ln(S_{ihjd}) - r_{jd}^{lo} + b_i^{up} - g_{jd}^{lo} \quad \forall i, h \in H_i, j \in (J_h \cap J_{cr}), d \in D_j \quad (5.60)$$

Para o volume dos tanques nos estágios com opções ou dependentes da rota (2, 3, 4, 9 e 10 para o tanque de alimentação e 2 e 4 para o de produto) deve-se fazer:

$$M2_{ihjd} = \ln(\max_{r \in R_{jd}} S1_{ihjr}) - v1_{jd}^{lo} + b_i^{up} \quad \forall i, h \in H_i, j \in (J_h \cap \{2,3,4,9,10\}), d \in D_j \quad (5.61)$$

$$M3_{ihjd} = \ln(\max_{r \in R_{jd}} S2_{ihjr}) - v2_{jd}^{lo} + b_i^{up} \quad \forall i, h \in H_i, j \in (J_h \cap \{2,4\}), d \in D_j \quad (5.62)$$

Nos estágios sem opções e independentes da rota (6, 11, 12, 13, 14 e 15 para o tanque de alimentação e 10, 12, 14 e 15 para o de produto) são definidos:

$$M2_{ihjd} = \ln(S1_{ihjd}) - v1_{jd}^{lo} + b_i^{up} \quad \forall i, h \in H_i, j \in (J_h \cap \{6,11,12,13,14,15\}), d \in D_j \quad (5.63)$$

$$M3_{ihjd} = \ln(S2_{ihjd}) - v2_{jd}^{lo} + b_i^{up} \quad \forall i, h \in H_i, j \in (J_h \cap \{10,12,14,15\}), d \in D_j \quad (5.64)$$

Para o tempo de ciclo nos estágios com tempo de processamento constante (1, 5, 7, 8 e 14) tem-se:

$$M4_{ihjd} = \ln(T_{ihjd}^0) - tl_i^{lo} - m_{jd}^{lo} \quad \forall i, h \in H_i, j \in (J_h \cap \{1,5,7,8,14\}), d \in D_j \quad (5.65)$$

Nos estágios com tempo em função da batelada e independente da rota (6, 11, 13 e 15):

$$M5_{ihjd} = \ln(T_{ihjd}^1) - tl_i^{lo} + b_i^{up} - r_{jd}^{lo} - g_{jd}^{lo} - m_{jd}^{lo} \quad \forall i, h \in H_i, j \in (J_h \cap \{6,11,13,15\}), d \in D_j \quad (5.66)$$

enquanto nos estágios 2, 3, 4 e 9 que dependem da rota:

$$M5_{ihjd} = \ln(\max_{r \in R_{jd}} T_{ihjr}^1) - tl_i^{lo} + b_i^{up} - r_{jd}^{lo} - g_{jd}^{lo} - m_{jd}^{lo} \quad \forall i, h \in H_i, j \in (J_h \cap \{2,3,4,9\}), d \in D_j \quad (5.67)$$

Nos estágios 10 e 12, os quais não dependem da rota:

$$M6_{ihjd} = T_{ihjd}^0 \exp(-tl_i^{lo} - m_{jd}^{lo}) + T_{ihjd}^1 \exp(b_i^{up} - r_{jd}^{lo} - g_{jd}^{lo} - tl_i^{lo} - m_{jd}^{lo}) - 1 \quad \forall i, h \in H_i, j \in (J_h \cap \{10,12\}), d \in D_j \quad (5.68)$$

5.4.2 Modelagem da alocação de tanques intermediários em posições predefinidas

Quando se consideram tanques intermediários, as restrições decorrentes da operação dos equipamentos são mantidas, porém devem-se definir variáveis independentes para o tempo de ciclo e para o tamanho de batelada em cada sub-processo. Assim, o tempo de operação de um estágio deve ser considerado na estimativa do tempo de ciclo do sub-processo ao qual pertence, como apresentado no item 5.2.11. Com a transformação das variáveis, as restrições para garantir o cumprimento das metas de produção e para igualar as taxas de produção são:

$$\sum_i Q_i \exp(tl_{i,1} - b_{i,1}) \leq HT \quad \text{e} \quad \sum_i Q_i \exp(tl_{i,2} - b_{i,2}) \leq HT \quad (5.69)$$

$$b_{i,1} - tl_{i,1} = b_{i,2} - tl_{i,2} \quad \forall i \quad (5.70)$$

As restrições para estimar o volume de um único tanque são:

$$vt_k \geq \ln(2 ST_{ih}) + b_{i,1} \quad \forall i, h \in H_i \quad (5.71)$$

$$vt_k \geq \ln(2 ST_{ih}) + b_{i,2} \quad \forall i, h \in H_i \quad (5.72)$$

onde vt é a variável transformada para o volume do tanque (VT_k). As restrições para a razão dos tamanhos de batelada nas variáveis transformadas são formuladas como:

$$-\ln(F) \leq b_{i,1} - b_{i,2} \leq \ln(F) \quad \forall i \quad (5.73)$$

Finalmente, o custo do tanque intermediário deve ser incorporado na função objetivo através do termo $a_k \exp(\beta_k vt_k)$.

5.4.3 Modelagem da alocação de tanques intermediários em posições otimizadas

Como indicado no item 5.2.12, a alocação de tanques intermediários em posições otimizadas requer a definição de um grupo de variáveis binárias yt_{kj} . A restrição para igualar as taxas de produção (5.40) deve ser formulada nas variáveis transformadas, enquanto a que garante as metas de produção (5.41) não muda:

$$E_i = \exp(tl_{i,j} - b_{i,j}) \quad \forall i, j \quad (5.74)$$

$$\sum_i Q_i E_i \leq HT \quad (5.41)$$

As restrições para estimar o volume do tanque j dependem do tamanho da batelada nos estágios j e k , onde k é o próximo estágio na sequência de produção, que no estudo de caso não é necessariamente o estágio $j+1$; para considerar a possibilidade de não alocar o tanque, deve-se adicionar um termo $big-M$ e as restrições tornam-se:

$$vt_j \geq \ln(2 ST_{ihj}) + b_{i,j} - M7_{ihj}(1 - ytk_j) \quad \forall i, h \in H_i, j \in J_h \quad (5.75)$$

$$vt_j \geq \ln(2 ST_{ihj}) + b_{i,k} - M8_{ihj}(1 - ytk_j) \quad \forall i, h \in H_i, j \in J_h, k = j_h^+ \quad (5.76)$$

onde j_h^+ identifica o estágio posterior a j na sequência definida para o hospedeiro h . As constantes M são:

$$M7_{ihj} = \ln(2 ST_{ihj}) - vt_j^{lo} + b_{i,j}^{up} \quad \forall i, h \in H_i, j \in J_h \quad (5.77)$$

$$M8_{ihj} = \ln(2 ST_{ihj}) - vt_j^{lo} + b_{i,k}^{up} \quad \forall i, h \in H_i, j \in J_h, k = j_h^+ \quad (5.78)$$

5.4.4 Considerações adicionais para a implementação do modelo

Um grande obstáculo que enfrentam os algoritmos para solução de problemas não-lineares é a obtenção da primeira solução viável; no caso concreto do problema de síntese e projeto ótimo, as restrições do horizonte de tempo (5.57), (5.69) ou (5.41) são as mais difíceis de serem satisfeitas. Para facilitar a obtenção de soluções viáveis, adicionaram-se variáveis (HT^* ou HT_j^*) que relaxam tais restrições, e caracterizam o tempo adicional necessário para cumprir as metas de produção; logo, na solução final, devem ter valor nulo. Para o modelo básico da planta multiproduto a restrição (5.57) torna-se:

$$\sum_i Q_i \exp(tl_i - b_i) \leq HT + HT^* \quad (5.79)$$

Para forçar a variável HT^* a tomar um valor nulo, adiciona-se uma penalidade na função objetivo ($t HT^*$), em que t é uma ponderação apropriada.

A transformação logarítmica das variáveis de projeto impede que as dimensões possam adotar valores nulos. Assim, caso a alocação de um equipamento não seja necessária existe um custo remanescente, igual ao custo mínimo do equipamento (custo do equipamento com dimensão mínima). Para eliminar esses custos devem-se adicionar à função objetivo termos de compensação que empregam variáveis binárias, as quais caracterizam a necessidade de projetar o equipamento. Para as operações opcionais dos estágios 2, 3 e 4, essas variáveis binárias são definidas por:

$$yuso_{jd} \geq \sum_{r \in R_{jd}} yr_{thr} \quad \forall i, h \in H_i, j \in (J_h \cap J_a), d \in D_j \quad (5.80)$$

onde $yuso_{jd} = 1$ se uma rota de produção envolvendo a opção d no estágio j é escolhida para algum produto, e J_a é o conjunto de estágios com alternativas (2, 3 e 4). Os termos que eliminam os custos remanescentes das opções não escolhidas nestes estágios são:

$$-\sum_{j \in J_a} \sum_{d \in D_j} \left[\mathbf{a}_{jd} \exp(m_{jd}^{lo} + \mathbf{b}_{jd} v 1_{jd}^{lo}) + \mathbf{l}_{jd} \exp(m_{jd}^{lo} + g_{jd}^{lo} + \mathbf{w}_{jd} r_{jd}^{lo}) + \mathbf{a}_{jd} \exp(m_{jd}^{lo} + \mathbf{b}_{jd} v 2_{jd}^{lo}) \right] (1 - yuso_{jd}) \quad (5.81)$$

Observando as opções de hospedeiro para cada produto, e a sequência de estágios para cada hospedeiro (Tabelas 5.2 e 5.3) verifica-se que os estágios 5, 6, 7 e 8 tornam-se desnecessários se para nenhum dos produtos seleciona-se a expressão por *E. coli*. A variável que caracteriza a necessidade de projetar esta sequência de estágios é:

$$yseq58 \geq y_{ih} \quad \forall i, h = coli \quad (5.82)$$

Dentro deste grupo de estágios, 5, 7 e 8 são descontínuos e então o termo de compensação é:

$$-\sum_{j \in \{5,7,8\}} \mathbf{a}_{jd} m_{jd}^{lo} g_{jd}^{lo} \exp(\mathbf{b}_{jd} v 1_{jd}^{lo}) \cdot (1 - yseq58) \quad (5.83)$$

enquanto o estágio 6 é semi-contínuo e não requer tanque de produto, então o termo correspondente se torna:

$$-[\mathbf{a}_{jd} \exp(m_{jd}^{lo} + \mathbf{b}_{jd} v 1_{jd}^{lo}) + \mathbf{l}_{jd} \exp(m_{jd}^{lo} + g_{jd}^{lo} + \mathbf{w}_{jd} r_{jd}^{lo})] \cdot (1 - yseq58) \quad j = 6, d \in D_j \quad (5.84)$$

5.5 Resultados

Os modelos matemáticos MINLP para o problema de síntese e projeto ótimo foram implementados na plataforma GAMS (Brooke *et al.*, 2000), empregando o *solver* DICOPT (Viswanathan e Grossmann, 1990), que usa o método OA/AP/ER (*outer approximation/ augmented penalty/ equality relaxation*). Este método resolve alternadamente problemas mestres MILP e subproblemas NLP. Primeiramente, resolve-se uma relaxação para o MINLP (as variáveis binárias são transformadas em contínuas), que corresponde a um modelo NLP; em seguida, resolve-se um modelo MILP mestre formado por uma aproximação linear da solução obtida no NLP; o método continua pela resolução do NLP fixando-se as variáveis binárias nos valores obtidos pelo MILP mestre; um novo MILP mestre é construído e o algoritmo procede até atingir convergência. Os subproblemas NLP foram resolvidos com o *solver* CONOPT2 (Drud, 1996) e os problemas mestres MILP com o OSL (IBM, 1991). Os resultados obtidos são apresentados nos itens 5.5.1 a 5.5.4; inicialmente resolveu-se o modelo monoproduto que serve como base de comparação para os modelos multiproduto e, posteriormente, o modelo multiproduto para três cenários que diferem no projeto de tanques intermediários (ver Tabela 5.4, p. 95). O item 5.5.5 apresenta o esforço computacional para a solução.

5.5.1 Resultados para o modelo de plantas monoproduto (cenário 1)

Este modelo foi desenvolvido para estabelecer parâmetros que permitam a comparação do custo e a análise das decisões na síntese dos empreendimentos sugeridos pelos modelos multiproduto.

Quanto à síntese, observa-se que a escolha ótima dos hospedeiros requer a consideração de múltiplos fatores que incidem no custo total da planta, entre os quais podem ser identificados o número de estágios requeridos para a recuperação do produto e os tempos de processamento. O efeito do número de estágios no custo da planta deve-se tanto ao custo dos equipamentos adicionados, como à redução do rendimento geral que induz o aumento das dimensões em outros estágios. Por outro lado, o efeito dos tempos de processamento deve-se à necessidade de alocar múltiplas unidades em paralelo ou

aumentar o tamanho da batelada para cumprir a meta de produção. Os hospedeiros escolhidos para cada produto e o custo de capital para cada empreendimento são apresentados na Tabela 5.5, que contém também os custos obtidos dos empreendimentos quando se escolhem hospedeiros diferentes dos ótimos.

Nos processos para produção de insulina e SOD, o fator decisivo na escolha do hospedeiro é o número de estágios, já que os tempos de processamento apresentam diferenças pequenas, como pode ser observado no Apêndice E. Para a produção da vacina, o fator decisivo na escolha do hospedeiro é o tempo de processamento, especialmente o tempo de fermentação, já que a melhor solução quando se considera a expressão por células de CHO projeta dez fermentadores operando fora-de-fase, e uma batelada maior que o dobro daquela projetada para a produção por *S. cerevisiae* (ver Tabela 5.5), o que incrementa o custo nos estágios cromatográficos. Na produção de tPA, o tempo de fermentação novamente tem peso maior do que o número de estágios, já que a produção com CHO requer a alocação de cinco fermentadores e o custo das colunas cromatográficas também é maior devido ao tamanho da batelada. A comparação dos projetos para o modelo monoproduto é apresentada no Apêndice F.

Tabela 5.5 Resultados gerais da síntese e projeto para o modelo monoproduto.

Produto	Hospedeiro	Escolha nos estágios			$B_i(\text{kg})$	$TL_i(\text{h})$	Custo total (US\$)
		$j = 2$	$j = 3$	$j = 4$			
INS	<i>S. cerevisiae</i> extracel.*	microfilt.			1,25	5,00	1.845.825*
	<i>E. coli</i> **	microfilt.	moinho	microfilt.	1,25	5,00	2.279.853**
HBV	<i>S. cerevisiae</i> intracel.*	microfilt.	moinho	microfilt.	1,29	7,71	2.223.386*
	CHO**	microfilt.			2,80	16,80	4.689.961**
tPA	<i>E. coli</i> *	microfilt.	homogen.	microfilt.	0,030	18,00	324.817*
	CHO**	microfilt.			0,056	33,60	540.347**
SOD	<i>S. cerevisiae</i> intracel.*	microfilt.	homogen.	microfilt.	0,51	15,43	761.994*
	<i>E. coli</i> **	microfilt.	homogen.	microfilt.	0,50	15,00	973.898**
Custo total* (US\$)							5.156.018

* hospedeiros ótimos.

** hospedeiros fixados.

Quanto à escolha da operação nos estágios 2, 3 e 4 observou-se que para as operações de separação/coleta de células, restos celulares ou corpos de inclusão houve

preferência pela microfiltração sobre a centrifugação. Para a ruptura das células, a escolha mudou entre os produtos (ver Tabela 5.5).

5.5.2 Resultados para o modelo multiproduto sem tanques intermediários (cenário 2)

Quanto à síntese observou-se que, para a planta multiproduto, selecionaram-se os mesmos hospedeiros que para as plantas monoproduto, e que a microfiltração foi a operação selecionada nos estágios 2 e 4; para a ruptura de células selecionou-se o moinho de bolas; a escolha de uma única operação para cada estágio era esperada, já que o projeto de ambas opções acarretaria custos adicionais. Os resultados gerais por produto são apresentados na Tabela 5.6.

Tabela 5.6 Resultados gerais por produto para o cenários 1 e 2.

Produto	Hospedeiro	monoproduto (cenário 1)		multiproduto (cenário 2)	
		B_i (kg)	TL_i (h)	B_i (kg)	TL_i (h)
INS	<i>S. cerevisiae</i> extracel.	1,25	5,00	3,86	5,00
HBV	<i>S. cerevisiae</i> intracel.	1,29	7,71	3,19	7,72
tPA	<i>E. coli</i>	0,03	18,00	0,18	9,00
SOD	<i>S. cerevisiae</i> intracel.	0,51	15,43	2,71	15,43

A Tabela 5.7 apresenta o projeto para a planta multiproduto, e a Figura 5.4 a configuração geral da planta, enquanto a Tabela 5.8 apresenta os custos do projeto e a comparação com os custos somados para as plantas monoproduto (cenário 1). O custo da planta multiproduto sem tanques pulmão (US\$ 5.925.056) é 14,9% maior do que a soma dos custos das plantas monoproduto (US\$ 5.158.207).

Na Tabela 5.8 pode-se observar que os estágios 5 a 9 e os estágios cromatográficos são os principais responsáveis pelo aumento no custo para o cenário 2. O aumento das dimensões dos equipamentos e do número de unidades em paralelo deve-se ao fato de que a geração de todos os produtos numa única linha de produção leva à diminuição do tempo disponível para a produção de cada produto, o qual se torna uma fração do horizonte de tempo em lugar do horizonte completo. Logo, é necessário tanto o dimensionamento de equipamentos maiores e a duplicação de unidades em-fase para

reduzir o número total de bateladas, como a duplicação de unidades fora-de-fase para diminuir os tempos de ciclo.

Figura 5.4 Configuração geral da planta multiproduto sem tanques intermediários (cenário 2); as unidades superpostas operam em-fase.

Tabela 5.7 Projeto para planta multiproduto sem tanques intermediários (cenário 2).

Estágio	unidades em-fase / fora-de-fase / dimensão			OC _j (%)	
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 2	Cenário 1 *
1		1/ 5/ 0,929		33,5	20,5
2B	1/ 1/ 0,929	1/ 1/ 0,840	1/ 1/ 0,840	44,3	37,1
3B	1/ 1/ 0,464	1/ 1/ 0,155		84,8	64,6
4B	1/ 1/ 0,464	1/ 1/ 0,645	1/ 1/ 0,383	58,4	9,8
5		1/ 2/ 0,743		94,5	8,3
6	1/ 1/ 0,743	1/ 1/ 5,780		91,8	0,0
7		1/ 2/ 0,892		91,8	0,0
8		1/ 4/ 9,364		91,8	0,0
9	1/ 1/ 9,364	2/ 1/ 26,010		91,8	0,0
10	1/ 1/ 0,840	5/ 1/ 0,750	1/ 1/ 7,969	24,4	29,7
11	1/ 1/ 7,969	1/ 1/ 25,810		69,5	22,1
12	1/ 2/ 0,268	5/ 2/ 0,750	1/ 2/ 7,969	15,1	6,8
13	1/ 1/ 7,969	2/ 1/ 25,810		20,9	0,0
14	1/ 1/ 0,241	3/ 1/ 0,643	1/ 1/ 0,386	30,7	43,8
15	1/ 1/ 0,386	1/ 1/ 0,771	1/ 1/ 0,386	41,4	41,3

Dimensão dos equipamentos: m³ para unidades descontínuas (tanques e reatores), m² para filtros e colunas cromatográficas, m³/h para homogeneizador ou moinho de bolas e kW para centrífugas.

* No cenário 1 estima-se como a ociosidade média nas quatro plantas.

Na comparação dos modelos mono e multiproduto (Tabela 5.6), o aumento do tamanho de batelada é observado em todos os produtos, enquanto a duplicação de unidades fora-de-fase é observada no projeto dos estágios 5, 7 e 8, que são exclusivos para a produção de tPA, a qual teve o tempo de ciclo reduzido. No biorreator e no estágio cromatográfico 14, o número de unidades fora-de-fase foi igual para o projeto monoproduto da insulina (Apêndice F) e para o multiproduto sem tanques (Figura 5.4). No caso dos estágios 5 a 9, a ociosidade torna-se extremamente grande porque estes estágios são exclusivos para produção de tPA (ver Tabela 5.7). A ociosidade (OC_j) do equipamento j é estimada como a percentagem do horizonte de tempo em que o equipamento permanece inoperante:

$$OC_j = 100\% \left(1 - \sum_i \frac{T_{ij} Q_i}{B_j HT} \right) \quad (5.85)$$

onde T_{ij} é o tempo de processamento no estágio j para o produto i , estimado com a equação (5.12) ou (5.15).

Tabela 5.8 Custos do projeto (US\$) para o cenário 2 e comparação com o cenário 1.

Estágio	Custo do arranjo de equipamentos			Custo total dos estágios	
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 2	Cenário 1
1		303.285		303.285	342.627
2B	5.501	2.501	5.180	13.182	14.403
3B	3.630	10.872		14.502	17.048
4B	3.630	1.998	3.233	8.861	14.693
5		53.448		53.448	13.864
6	4.812	12.883		17.695	3.733
7		58.549		58.549	13.864
8		379.441		379.441	77.020
9	22.005	92.533		114.540	17.453
10	5.180	1.323.169	19.976	1.348.325	1.242.708
11	19.975	45.963		65.940	33.631
12	5.217	2.646.338	39.951	2.691.506	2.562.692
13	19.976	91.927		111.903	106.817
14	2.249	729.366	3.247	735.063	677.797
15	3.247	2.325	3.247	8.820	19.857
Custo total				5.925.056	5.158.207

Os resultados para o cenário 2 permitem concluir que o projeto de plantas multiproduto pode aumentar os custos de capital se não for aplicada alguma estratégia para melhorar o aproveitamento dos equipamentos, *i.e.* diminuir a ociosidade. Na tentativa de diminuir a ociosidade e reduzir o custo da planta foram avaliadas duas alternativas, a alocação de tanques pulmão, e o projeto de linhas de produção auxiliares com flexibilização do *scheduling*. Os resultados obtidos com a alocação de tanques-pulmão são apresentados nos itens 5.5.3 e 5.5.4, enquanto a consideração de equipamentos auxiliares é apresentada no Capítulo 6.

5.5.3 Resultados para o modelo multiproduto com tanques intermediários em posições predefinidas (cenário 3)

Na Tabela 5.7 observa-se que para o cenário 2 os maiores tempos ociosos ocorrem nos estágios 3 a 9, os quais não processam todos os produtos: a produção de insulina, que é expressa de modo extracelular, não requer nenhum destes estágios, enquanto a produção de HBV e SOD não envolve os estágios 5 a 9. A partir da observação anterior, acredita-se que a alocação de dois tanques-pulmão, um entre os estágios 2 e 3 e outro entre 9 e 10 permite diminuir a ociosidade, melhorando o aproveitamento dos equipamentos, e reduzindo o custo da planta.

Inicialmente, avaliou-se o efeito da razão máxima entre as bateladas dos subprocessos. Para uma razão $F = 2$ a melhor solução obtida tem um custo de US\$ 5.703.418; por outro lado, para $F = 5$ obtém-se um custo de US\$ 5.627.280. A economia de 1,3% não justifica o aumento da razão entre as bateladas, cujos valores são apresentados no Apêndice G, já que a operação com grandes diferenças nos tamanhos de batelada requer tempos elevados de permanência nos tanques, o que na prática acarreta maiores custos operacionais e pode ser prejudicial para os produtos.

Quanto à síntese, os resultados para ambos valores de F foram idênticos, porém o projeto apresenta diferenças para os estágios 1 a 9; especificamente na fermentação e no *refolding* (estágio 8), a redução no número de unidades que operam fora-de-fase permite a diminuição do custo de capital (ver Apêndice G).

Os resultados apresentados a seguir e no próximo capítulo foram obtidos com uma razão máxima entre as bateladas $F = 2$. Como mencionado anteriormente, a melhor solução obtida para este cenário tem um custo de US\$ 5.703.418, o que representa uma economia de 3,7 % em relação ao projeto para a planta sem tanques intermediários, porém mantém-se um sobrecusto de 10,6 % em relação ao projeto das plantas monoproduto.

Em relação ao cenário 2, esta solução apresenta as mesmas escolhas, tanto para os hospedeiros como para as operações; por outro lado, no projeto foram necessárias menos unidades em paralelo nos estágios 1, 5, 7 e 8 (ver Figura 5.5), o que gerou a diminuição de custo previamente mencionada (a comparação de custo para cada estágio é apresentada na Tabela 5.11).

Figura 5.5 Configuração geral da planta para o cenário 3 (tanques intermediários em posições predefinidas), as unidades superpostas operam em-fase.

Na comparação das variáveis de dimensionamento das unidades para os cenários 2 e 3 (Tabelas 5.6 e 5.9) observa-se que os tamanhos de batelada e os tempos de ciclo para a planta sem tanques são semelhantes aos da seção de purificação ($10 \leq j$) da planta com tanques, o que permite verificar que os estágios cromatográficos de purificação, devido ao custo, têm peso maior na otimização do projeto. Logo, acredita-se que, desvinculando

a operação de cada um destes estágios através da alocação de tanques intermediários, pode-se reduzir o custo total do empreendimento.

Tabela 5.9 Resultados gerais por produto para o cenário 3.

Produto	Hospedeiro	B_{ij} (kg)			TL_{ij} (h)		
		$j \leq 2$	$3 \leq j \leq 9$	$10 \leq j$	$j \leq 2$	$3 \leq j \leq 9$	$10 \leq j$
INS	<i>S. cerevisiae</i> extracel.	6,17	---	3,86	8,00	---	5,00
HBV	<i>S. cerevisiae</i> intracel.	3,33	1,66	3,19	8,06	4,03	7,72
tPA	<i>E. coli</i>	0,18	0,36	0,72	9,00	18,00	35,44
SOD	<i>S. cerevisiae</i> intracel.	1,40	1,35	2,71	8,00	7,72	15,44

Tabela 5.10 Projeto para o cenário 3 e comparação da ociosidade.

Estágio	unidades em-fase / fora-de-fase / dimensão			OC_j (%)		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 3	Cenário 2	Cenário 1 *
1		1/ 3/ 0,929		1,2	33,5	20,5
2B	1/ 1/ 0,929	1/ 1/ 0,840	1/ 1/ 1,344	44,3	44,3	37,1
	Tanque pulmão 2	2,123				
3B	1/ 1/ 0,929	1/ 1/ 0,155		84,8	84,8	64,6
4B	1/ 1/ 0,929	1/ 1/ 0,645	1/ 1/ 0,200	58,4	58,4	9,8
5		1/ 1/ 1,486		94,5	94,5	8,3
6	1/ 1/ 1,486	1/ 1/ 5,780		91,8	91,8	0,0
7		1/ 1/ 1,783		91,8	91,8	0,0
8		1/ 2/ 18,727		91,8	91,8	0,0
9	1/ 1/ 18,727	2/ 1/ 26,010		91,8	91,8	0,0
	Tanque pulmão 9	5,714				
10	1/ 1/ 0,840	5/ 1/ 0,750	1/ 1/ 7,969	26,9	24,4	29,7
11	1/ 1/ 7,969	1/ 1/ 25,810		69,5	69,5	22,1
12	1/ 2/ 0,268	5/ 2/ 0,750	1/ 2/ 7,969	20,2	15,1	6,8
13	1/ 1/ 7,969	2/ 1/ 25,810		20,9	20,9	0,0
14	1/ 1/ 0,241	3/ 1/ 0,643	1/ 1/ 0,386	34,1	30,7	43,8
15	1/ 1/ 0,386	1/ 1/ 0,771	1/ 1/ 0,386	41,4	41,4	41,3

Dimensão dos equipamentos: m^3 para unidades descontínuas (tanques e reatores), m^2 para filtros e colunas cromatográficas, m^3/h para homogeneizador ou moinho de bolas e kW para centrífugas.

* No cenário 1 estima-se como a ociosidade média nas quatro plantas.

Os resultados obtidos na síntese e projeto com tanques intermediários em posições predefinidas não são considerados satisfatórios. A flexibilização nas condições de

operação para os sub-processos permite a diminuição no número de unidades em paralelo nos estágios 1, 5, 7 e 8, e é nestes estágios que o projeto envolve redução no custo, comparado com o cenário 2; porém o custo continua sendo maior que o das plantas monoproduto (ver Tabela 5.11) e a ociosidade dos equipamentos é da mesma ordem de magnitude que para o cenário 2 (Tabela 5.10), exceto para o biorreator, no qual se consegue a maior redução no custo.

Tabela 5.11 Custos do projeto (US\$) para o cenário 3 e comparação com os cenários 1 e 2.

Estágio	Custo do arranjo de equipamentos			Custo total dos estágios		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 3	Cenário 2	Cenário 1
1		181.971		181.971	303.285	342.627
2B	5.501	2.501	6.868	14.870	13.182	14.403
Tanque-pulmão 2		9.034		9.034		
3B	5.501	10.872		16.373	14.502	17.048
4B	5.501	1.998	2.189	9.688	8.861	14.693
5		37.793		37.793	53.448	13.864
6	7.293	12.884		20.177	17.695	3.733
7		41.400		41.400	58.549	13.864
8		268.306		268.306	379.441	77.020
9	33.354	92.534		125.888	114.540	17.453
Tanque-pulmão 9		16.362		16.362		
10	5.180	1.323.169	19.976	1.348.325	1.348.325	1.242.708
11	19.976	45.964		65.940	65.940	33.631
12	5.217	2.646.338	39.951	2.691.506	2.691.506	2.562.692
13	19.976	91.927		111.903	111.903	106.817
14	2.449	729.367	3.247	735.063	735.063	677.797
15	3.247	2.326	3.247	8.820	8.820	19.857
Custo total				5.703.418	5.925.056	5.158.207

5.5.4 Resultados para o modelo multiproduto com tanques intermediários em posições otimizadas (cenário 4)

Quanto à síntese, a melhor solução obtida para este cenário considera as mesmas escolhas do cenário anterior, que correspondem a *S. cerevisiae* para expressão de INS, HBV e SOD, e *E. coli* para tPA; para o processamento definiu-se microfiltração para os

estágios 2 e 4, e moinho de bolas para o 3. Quanto ao projeto, foram alocados 6 tanques intermediários (ver Tabela 5.12 e Figura 5.6) e o número de unidades em paralelo diminuiu para os estágios 1, 8 e 9; porém a redução do custo total (US\$ 5.635.063) é pequena comparada à planta com tanques em posições prefixadas e sem tanques (1,2 e 4,9%, respectivamente). As variáveis de dimensionamento para cada estágio são apresentadas no Apêndice G.

Tabela 5.12 Projeto para o cenário 4 e comparação da ociosidade.

Estágio	unidades em-fase / fora-de-fase / dimensão			OC_j (%)		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 4	Cenário 2	Cenário 1 *
1		1/ 2/ 1,142		10,6	33,5	20,5
	Tanque-pulmão 1	11,214				
2B	1/ 1/ 1,669	1/ 1/ 0,925	1/ 1/ 1,110	49,4	44,3	37,1
	Tanque-pulmão 2	20,186				
3B	1/ 1/ 0,556	1/ 1/ 0,139		83,1	84,8	64,6
4B	1/ 1/ 0,556	1/ 1/ 0,580	1/ 1/ 0,200	53,7	58,4	9,8
5		1/ 1/ 0,890		90,8	94,5	8,3
	Tanque-pulmão 5	12,215				
6	1/ 1/ 1,335	1/ 1/ 5,193		90,8	91,8	0,0
	Tanque-pulmão 6	14,658				
7		1/ 1/ 3,204		95,4	91,8	0,0
8		1/ 1/ 33,643		90,8	91,8	0,0
9	1/ 1/ 33,643	1/ 1/ 46,749		90,8	91,8	0,0
10	1/ 1/ 0,869	5/ 1/ 0,739	1/ 1/ 7,849	10,3	24,4	29,7
11	1/ 1/ 7,849	1/ 1/ 25,422		69,0	69,5	22,1
	Tanque-pulmão 11	0,554				
12	1/ 2/ 0,257	5/ 2/ 0,719	1/ 2/ 7,640	5,6	15,1	6,8
13	1/ 1/ 7,640	2/ 1/ 26,524		23,0	20,9	0,0
	Tanque-pulmão 13	0,530				
14	1/ 1/ 0,265	3/ 1/ 0,707	1/ 1/ 0,424	6,7	30,7	43,8
15	1/ 1/ 0,424	1/ 1/ 0,849	1/ 1/ 0,424	46,8	41,4	41,3

Dimensão dos equipamentos: m³ para unidades descontínuas (tanques e reatores), m² para filtros e colunas cromatográficas, m³/h para homogeneizador ou moinho de bolas e kW para centrífugas.

* No cenário 1 estima-se como a ociosidade média nas quatro plantas.

A ociosidade obtida para este cenário (Tabela 5.12) é semelhante aos valores do cenário 2 (planta multiproduto sem tanques intermediários) exceto para a fermentação e os estágios cromatográficos; porém pode-se observar que o projeto e os custos nos

estágios cromatográficos são semelhantes aos dos cenários 3 e 2; logo, a redução na ociosidade destes estágios deveu-se a mudanças nas variáveis para dimensionamento (tempos de ciclo e tamanhos de batelada) que permitem a diminuição dos custos em outros estágios.

A redução no custo para este cenário deve-se principalmente às mudanças no projeto dos estágios 1, 5 e 7. Isto confirma que o peso dos estágios cromatográficos no custo da planta faz com que, para os demais estágios, superdimensionem-se os equipamentos, sugerindo que se devem procurar abordagens para melhorar o aproveitamento dos equipamentos, especialmente em estágios diferentes dos cromatográficos.

Figura 5.6 Configuração geral da planta para o cenário 4 (tanques intermediários em posições otimizadas), as unidades superpostas operam em-fase.

Nos resultados obtidos para este cenário pode-se observar que a alocação de tanques intermediários, ainda que em posições otimizadas, não consegue melhorar satisfatoriamente o aproveitamento dos equipamentos, *i.e.* reduzir a ociosidade. Logo, outras estratégias como a alocação de unidades auxiliares e a flexibilização das restrições que garantem a disponibilidade dos equipamentos serão abordadas no capítulo seguinte.

Tabela 5.13 Custos do projeto (US\$) para o cenário 4 e comparação com os cenários 1 e 2.

Estágio	Custo do arranjo de equipamentos			Custo total dos estágios		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 4	Cenário 2	Cenário 1
1		137.293		137.293	303.285	342.627
	Tanque-pulmão 1	24.521		24.521		
2B	7.818	2.713	6.120	16.651	13.182	14.403
	Tanque-pulmão 2	34.890		34.890		
3B	4.045	10.306		14.351	14.502	17.048
4B	4.045	1.825	2.189	8.059	8.861	14.693
5		29.253		29.253	53.448	13.864
	Tanque-pulmão 5	25.811		25.811		
6	6.839	11.763		18.602	17.695	3.733
	Tanque-pulmão 6	28.794		28.794		
7		55.490		55.490	58.549	13.864
8		179.809		179.809	379.441	77.020
9	47.403	76.156		123.559	114.540	17.453
10	5.286	1.312.210	19.795	1.337.291	1.348.325	1.242.708
11	19.795	45.377		65.172	65.940	33.631
	Tanque-pulmão 11	4.034		4.034		
12	5.087	2.585.798	38.955	2.629.840	2.691.506	2.562.692
13	19.478	94.085		113.563	111.903	106.817
	Tanque-pulmão 13	3.931		3.931		
14	2.593	768.719	3.438	774.750	735.063	677.797
15	3.438	2.523	3.438	9.399	8.820	19.857
Custo total				5.635.063	5.925.056	5.158.207

5.5.5 Esforço computacional

A seguir apresenta-se a complexidade dos modelos resolvidos, quanto ao número de variáveis (contínuas e binárias), de restrições e de decisões, e seu efeito sobre o esforço computacional para solução, o qual é medido em termos de tempo de CPU utilizando um PC Pentium 4, 2 GHz e 512 MB de memória RAM. A Tabela 5.14 apresenta a dimensão dos problemas resolvidos e seu respectivo esforço computacional. Na modelagem empregada, um grande número de variáveis binárias é gerado pela formulação das variáveis inteiras, que definem o número de unidades em paralelo como variáveis binárias, ver item 5.2.10; porém para cada equipamento, as n variáveis y_{gjn}

estão condicionadas por uma única decisão, o mesmo acontece com as n variáveis ym_{jdn} e com as variáveis yr_{ihj} que definem a rota de produção para cada combinação produto/hospedeiro (a estimação do número de rotas de produção está no item 5.2.2).

Tabela 5.14 Desempenho computacional para problemas de síntese e projeto de plantas mono e multiproduto.

Modelo	Cenário	Restrições	Variáveis binárias	Decisões discretas	Variáveis contínuas	Tempo de CPU (s)
I	1					
	INS	287	235	38	107	7
	HBV	273	186	28	94	8
	tPA	313	235	38	108	13
	SOD	313	241	38	114	10
II	2	947	295	44	207	619
	3	1047	295	44	228	583
	4	1127	309	58	366	1500

No cenário 1 resolveram-se os quatro problemas monoproducto. A descrição dos modelos e cenários é apresentada na Tabela 5.4.

Outro parâmetro que afeta o esforço computacional é o critério de parada do *solver*. Os problemas misto-inteiros requerem a definição de um *gap* mínimo para considerar a solução satisfatória, e a redução desse parâmetro induz a obtenção de soluções melhores, porém aumenta o esforço computacional. Todos os resultados apresentados neste capítulo consideram critérios de parada entre 0,5 e 1,0%.

Na tabela 5.14 observa-se que o maior aumento proporcional no esforço computacional é gerado pela consideração de múltiplos produtos, já que, apesar do número de variáveis binárias e de decisões discretas aumentarem em menor proporção, as relações entre decisões discretas e restrições tornam-se muito mais complexas. Cabe lembrar que, no projeto multiproduto o dimensionamento de cada equipamento deve considerar as restrições impostas para cada um dos produtos. A otimização da alocação de tanques intermediários também aumenta consideravelmente o esforço computacional, o que se atribui ao aumento no número de configurações possíveis para a planta.

5.6 Conclusões

Os resultados obtidos mostram que a formulação das decisões por abordagem *big-M* permite resolver simultaneamente os problemas de síntese e projeto, com esforço computacional razoável. Na escolha de hospedeiros conclui-se que a expressão extracelular não é obrigatoriamente a mais econômica, já que as condições de operação (tempo de processamento em cada estágio, tamanho de batelada, etc.) podem ter um efeito preponderante, reafirmando que os problemas de síntese e projeto devem ser resolvidos simultaneamente.

A escolha de opções diferentes para a ruptura de células (estágio 3), nos modelos monoproduto e multiproduto, mostra que a escolha de opções também deve ser incluída no problema de síntese e projeto, já que não depende só dos produtos como também das variáveis de dimensionamento dos equipamentos (tamanhos de batelada e tempos de ciclo), as quais devem conciliar a geração dos múltiplos produtos nas quantidades requeridas.

Observa-se ainda que a conciliação do processamento de vários produtos dentro do horizonte de tempo disponível torna necessária a duplicação de unidades e gera ociosidade elevada, o que aumenta o custo total do projeto. Os resultados obtidos com alocação de tanques intermediários mostram que esta abordagem tem um efeito limitado na diminuição dos tempos de ociosidade e do custo do projeto. Assim, mudanças na modelagem para flexibilizar o *scheduling* tornam-se necessárias, na tentativa de melhorar o aproveitamento dos equipamentos. Também observou-se que a otimização da localização de tanques intermediários aumenta consideravelmente o esforço computacional para solução do problema, devido ao incremento no número de variáveis binárias.

5.7 Nomenclatura dos modelos de síntese e projeto de processos

Esta seção apresenta a notação utilizada neste capítulo, em referências bibliográficas empregadas, bem como no capítulo seguinte.

Variáveis

B, b	tamanho de batelada e seu logaritmo.
E	inverso da taxa de produção.
G, g	número de unidades em-fase e seu logaritmo.
HT^*	tempo adicional necessário para cumprir as metas de produção.
M, m	número de unidades fora-de-fase e seu logaritmo.
N	número de bateladas de cada produto no super-ciclo (capítulo 6).
OC	ociosidade, percentagem do tempo que o equipamento permanece inoperante.
R, r	dimensão de unidades contínuas ou colunas cromatográficas e seu logaritmo.
S^*	parâmetro de dimensionamento escolhido.
$S1^*, S2^*$	parâmetros de dimensionamento escolhidos, para tanques de alimento e produto.
T	tempo de operação.
T^{1*}	parâmetro escolhido para estimação do tempo de operação.
TL, tl	tempo de ciclo e seu logaritmo.
V, v	volume da unidade descontínua e seu logaritmo.
VT, vt	volume de tanque intermediário e seu logaritmo.
$VI, v1$	volume do tanque de alimentação e seu logaritmo.
$V2, v2$	volume do tanque de produto e seu logaritmo.
y	variável binária para escolha de hospedeiro.
yg	variável binária para definir o número de unidades em-fase.
ym	variável binária para definir o número de unidade fora-de-fase.
yr	variável binária para escolha da rota de produção.
$yuso$	variável binária que caracteriza a necessidade de projetar cada opção.
$yseq58$	variável binária que caracteriza a necessidade de projetar os estágios 5 a 8.
ytk	variável binária para alocação de tanque intermediário.
z	variável binária para escolha de opção de operação (Montagna <i>et al.</i> , 2004).

Parâmetros

CS, CT^1	constantes para estimação dos parâmetros de dimensionamento e tempo.
c	coeficiente para definir como variáveis binárias, o número de unidades em paralelo.
D	fator de trabalho em estágios descontínuos.
HT	horizonte de tempo.
$M1, M8$	parâmetros para formulação <i>big-M</i> da disjunção.
ND	número de opções num estágio.

Q	meta de produção.
S	parâmetro de dimensionamento.
$S1, S2$	parâmetros de dimensionamento para tanques de alimento e produto.
T^0, T^1	parâmetros para estimação do tempo de operação.
a, b	parâmetros de custo para unidades descontínuas e tanques.
I, w	parâmetros de custo para unidades contínuas e cromatográficas.
h	coeficiente de rendimento da operação.
q	tempo de processamento em estágios descontínuos.
F	relação entre os tamanhos de batelada para os sub-processos.
t	peso do tempo adicional (HT^*) na função objetivo.

Conjuntos

D_j	opções de operação para o estágio j .
H_i	hospedeiros disponíveis para o produto i .
J_a	estágios com alternativas de operação.
J_b	estágios descontínuos.
J_{cr}	estágios cromatográficos.
J_h	estágios necessários para o processamento quando emprega-se o hospedeiro h .
J_s	estágios semi-contínuos.
R_h	rotas de produção possíveis para o hospedeiro h .
R_{jd}	rotas de produção possíveis para a opção d no estágio j .

Subscritos

d	opção de operação.
h	hospedeiro.
i	produto.
j, k	estágio.
j_h^+	estágio posterior a j na sequência definida para o hospedeiro h .
n	número de unidades em paralelo.
r	rota de produção.

Superscritos

lo, up	limite inferior e superior para a variável.
aux	equipamento auxiliar (capítulo 6)

CAPÍTULO 6 SÍNTESE E PROJETO COM EQUIPAMENTOS AUXILIARES E CONSIDERAÇÕES NO *SCHEDULING*

Quando se considera o processamento em plantas monoproduto com equipamentos multipropósito, o *scheduling* da operação requer a verificação da disponibilidade dos equipamentos, como observado por Samsatli e Shah (1996). De forma semelhante, quando se analisam plantas multiproduto ainda que sem equipamentos multipropósito, a disponibilidade dos equipamentos deve ser considerada. A modelagem apresentada no capítulo anterior considera que todos os produtos seguem a mesma linha de processamento, o que implica que vários produtos não podem estar simultaneamente no mesmo estágio de processamento, e define tempos de ciclo para o processamento completo de cada produto ou para os sub-processos, quando há alocação de tanques intermediários; o projeto da planta mediante esta abordagem verifica a disponibilidade dos equipamentos já que aloca um período de tempo para cada produto em cada estágio, igual ao tempo de ciclo, e evita lidar diretamente com o problema da disponibilidade no *scheduling*, como ilustrado no gráfico de Gantt para um processo com dois produtos e três estágios (Figura 6.1).

Figura 6.1 Gráfico de Gantt para ilustração da disponibilidade dos equipamentos. Os tempos de ciclo são $TL_A = 3$ h, $TL_B = 2$ h.

No exemplo da Figura 6.1, observa-se que a ordem de processamento é idêntica para todos os estágios e que apesar do produto B não requerer processamento no estágio 2, existe um período de tempo alocado para este propósito. Como mencionado anteriormente, estas considerações garantem a disponibilidade dos equipamentos; porém inflexibilizam o *scheduling* e podem prejudicar o objetivo de minimizar o custo de capital pois impõem condições muito restritivas ao projeto dos equipamentos, gerando ociosidade elevada. A comparação dos resultados obtidos para plantas mono e multiproduto confirma que a conciliação do processamento de vários produtos dentro do horizonte de tempo disponível, não só obriga a duplicação de unidades como também gera ociosidade elevada em alguns estágios, o que aumenta o custo total da planta.

Os modelos apresentados neste capítulo visam a obtenção de melhores soluções para o problema de síntese e projeto ótimo apresentado na seção 5.3, através de duas abordagens. A primeira busca melhorar o aproveitamento dos equipamentos nos estágios 5 a 9 (operações para dissolução de corpos de inclusão e renaturação) para reduzir os custos da planta quando se considera a expressão de tPA por *E. coli*; este hospedeiro é a melhor opção tanto nos cenários 1, 2, 3 e 4. O segundo caminho para obter melhores soluções do problema é projetar um biorreator auxiliar para a expressão de tPA por CHO e flexibilizar a operação do estágio 9, que se tornaria exclusivo para o processamento de tPA; apesar da expressão de tPA por CHO não ser a melhor opção nos cenários resolvidos anteriormente, na prática é um caminho intensamente empregado e acredita-se que, contornando a dificuldade gerada pelo elevado tempo de fermentação, podem-se obter soluções interessantes. Ambas abordagens foram avaliadas sem e com alocação de tanques intermediários.

6.1 Considerações especiais no *scheduling*

Considere a planta multiproduto com a expressão de tPA por *E. coli* e dos outros produtos por *S. cerevisiae*, como nos resultados para os cenários 2, 3 e 4. Uma parcela importante do sobrecusto (em relação ao projeto de plantas monoproduto) deve-se ao projeto dos estágios 5 a 9 que são exclusivos para produção de tPA e conseqüentemente apresentam ociosidade muito elevada. Visando melhorar o aproveitamento destes

equipamentos, desenvolvem-se modificações na modelagem, as quais consideram um *scheduling* flexível. Neste, o tempo de processamento para tPA nos estágios exclusivos (5 a 9) desvincula-se do tempo nos estágios compartilhados com outros produtos (1 a 4 e 10 a 15). Com isto, o *schedule* da operação na planta multiproduto torna-se semelhante ao apresentado na Figura 6.2.

Figura 6.2 Gráfico de Gantt para ilustração da operação da planta multiproduto com expressão de tPA por *E. coli* e flexibilização do *scheduling*. Os tempos de ciclo são $TL_{INS} = TL_{HBV} = TL_{SOD} = TL_{tPA} = 12$ h; o biorreator consiste em um arranjo com duas unidades fora-de-fase. As linhas pontilhadas indicam a seqüência de processamento para cada produto.

Com a desvinculação do tempo de operação dos estágios 5 a 9, a restrição para o tempo de ciclo de tPA torna-se:

$$TL_{tPA} \geq \frac{t_{tPA,j}}{M_{j,d}} \quad \forall j \in (J_{coli} - \{5, \dots, 9\}), d \in D_j \quad (6.1)$$

Logo estes estgios devem ser excludos das restries (5.53, 5.54 e 5.55). A dimenso dos equipamentos descontnuos, contidos nesta seo do processo, continua sendo determinada pelo tamanho de batelada; ento as restries (5.45, 5.47 e 5.51) no so modificadas. Por outro lado, o tempo de processamento nos estgios semi-contnuos (6 e 9) depende do tamanho de batelada e da dimenso das unidades; se for considerado que estes tempos devem ser da ordem de magnitude dos tempos nos estgios descontnuos, pode-se empregar a seguinte restrio:

$$t_{\text{PA},8} \geq \frac{t_{\text{PA},j}}{M_{jd}} \quad \forall j \in \{6,9\}, d \in D_j \quad (6.2)$$

6.2 Projeto de equipamentos auxiliares.

O projeto de equipamentos auxiliares visa aumentar a disponibilidade dos equipamentos durante a operao e viabilizar algumas opes de processamento que na abordagem do captulo anterior so descartadas devido ao elevado tempo de operao que envolvem. No caso de estudo considerado, a expresso de HBV e tPA por clulas de CHO  inviabilizada pelo tempo de fermentao (168 h); porm esta expresso pode ser atraente, especialmente para o tPA, j que tornaria desnecessrios os estgios 5 a 8 (dissoluo de corpos de incluso e renaturao). Para viabilizar a produo de tPA por clulas de CHO projeta-se um arranjo de biorreatores exclusivo para este produto, o qual operaria de forma simultnea ao arranjo que gera os produtos restantes. Adicionalmente, se for considerada expresso de tPA por CHO, mantendo a escolha de hospedeiro para os outros produtos (*S. cerevisiae*), o estgio 9 (ultrafiltrao) torna-se exclusivo para produo de tPA, o que gera um elevado tempo de ociosidade e um sobre-dimensionamento do mesmo. Para reduzir este efeito, desvincula-se o tempo de processamento para este estgio do tempo de ciclo para o tPA. O grfico de Gantt da Figura 6.3 ilustra uma possibilidade de operao para a planta multiproduto com as consideraes apresentadas nesta seo; o tempo de fermentao para o tPA  168 h e para os outros produtos  24 h; os tempos de processamento em qualquer outro estgio so inferiores a 12 horas, com exceo do tempo para tPA na ultrafiltrao.

Figura 6.3 Gráfico de Gantt para ilustração da operação da planta multiproduto com biorreator auxiliar para a produção de tPA por CHO, os tempos de ciclo são $TL_{INS} = TL_{HBV} = TL_{SOD} = TL_{tPA} = 12$ h; o biorreator principal consiste em um arranjo com duas unidades fora-de-fase, enquanto o auxiliar consiste de uma unidade simples. As linhas pontilhadas indicam a sequência de processamento para cada produto.

Na Figura 6.3 observa-se que as bateladas são agrupadas em super-ciclos, pois para cada batelada de tPA devem ser processadas seis bateladas de INS, seis de HBV e duas de SOD. Também observa-se que os tempos de fermentação e ultrafiltração para tPA tornam-se restritos pelo tempo do super-ciclo.

O projeto de um biorreator auxiliar para tPA requer a definição de variáveis para seu projeto, tais como volume (V_1^{aux}), número de unidades em paralelo (M_1^{aux} e G_1^{aux}), tamanho de batelada (B_{tPA}^{aux}) e tempo de ciclo (TL_{tPA}^{aux}). O tempo de ciclo para tPA na linha de processamento principal deixa de ser restrito pelos tempos de fermentação e ultrafiltração (estágio 9):

$$TL_{tPA} \geq \frac{t_{tPA,j}}{M_{j,d}} \quad \forall j \in (J_{CHO} - \{1,9\}), d \in D_j \quad (6.3)$$

Por outro lado, na linha de processamento auxiliar tem-se:

$$TL_{tPA}^{aux} \geq \frac{t_{tPA,j}}{M_{jd}} \quad \forall j = 1, j = 9, d \in D_j \quad (6.4)$$

Como mencionado anteriormente, o *scheduling* agrupa as bateladas em super-ciclos. Desta forma, o tempo de processamento para tPA na linha auxiliar é restrito pelo tempo do super-ciclo e o número de bateladas dos outros produtos deve permitir a conformação dos super-ciclos, para garantir a meta de produção de tPA. Estas considerações são modeladas como:

$$TL_{tPA}^{aux} \leq \sum_i N_i TL_i \quad (6.5)$$

$$\frac{Q_{tPA}}{B_{tPA}} \leq \frac{Q_i}{B_i N_i} \quad \forall i \neq tPA \quad (6.6)$$

onde N_i é o número de bateladas do produto i no super-ciclo ($N_{tPA} = 1$), Q_i/B_i caracteriza o número total de bateladas para atingir a meta de produção de i , e portanto $Q_i/B_i N_i$ é o número máximo de super-ciclos que cada produto permite. O volume do reator auxiliar é definido de forma análoga à dos equipamentos principais:

$$V_1^{aux} \geq \frac{S_{tPA,CHO,1}^* B_{tPA}^{aux}}{G_1^{aux}} \quad (6.7)$$

Se não são considerados tanques intermediários, os tamanhos de batelada para tPA em ambas linhas de processamento devem ser iguais. Estas restrições devem ser convexificadas com transformação logarítmica das variáveis, conforme item 5.2.10.

6.3 Resultados

Os resultados computacionais apresentados nesta seção foram obtidos com as mesmas estratégias e recursos computacionais da seção 5.5. Inicialmente, os modelos com considerações especiais no *scheduling* e com projeto do fermentador auxiliar foram resolvidos sem tanques intermediários, para verificar o efeito isolado destas modificações; posteriormente analisou-se a alocação de tanques intermediários em posições otimizadas. A Tabela 6.1 apresenta os cenários resolvidos.

Tabela 6.1 Características dos cenários resolvidos para o problema de síntese e projeto com considerações especiais no *scheduling* e unidades auxiliares.

Modelo	Cenário
III Multiproduto com <i>scheduling</i> flexível, tPA por <i>E.coli</i> .	5. Sem tanques intermediários 6. Tanques em posições otimizadas
IV Multiproduto com unidades auxiliares, tPA por CHO.	7. Sem tanques intermediários 8. Tanques em posições otimizadas

6.3.1 Resultados para o modelo multiproduto com *scheduling* flexível (cenário 5)

A Tabela 6.2 apresenta os resultados principais; vale citar que o cenário 2 corresponde à planta multiproduto sem alocação de tanques. Pode-se observar que a desvinculação do tempo de operação nos estágios 5 a 9 não modificou drasticamente as variáveis de projeto das unidades (tamanhos de batelada e tempos de ciclo) em relação às do problema multiproduto básico (cenário 2), porém gerou uma economia de custo de 8,0% (US\$ 5.453.555 contra US\$ 5.925.056), a qual é resultado da diminuição no número de unidades em paralelo e da diminuição do volume dos biorreatores. A configuração geral da planta é apresentada na Figura 6.4 e a descrição completa do projeto no Apêndice G.

Tabela 6.2 Resultados gerais por produto para os cenários 2 e 5.

Produto	Hospedeiro	cenário 2		cenário 5	
		B_i (kg)	TL_i (h)	B_i (kg)	TL_i (h)
INS	<i>S. cerevisiae</i> extracel.	3,86	5,00	3,86	5,00
HBV	<i>S. cerevisiae</i> intracel.	3,19	7,72	3,19	7,72
tPA	<i>E. coli</i>	0,18	9,00	0,17	7,51
SOD	<i>S. cerevisiae</i> intracel.	2,71	15,43	2,61	15,44

O tempo de ciclo para tPA no cenário 5 corresponde aos estágios 1 a 4 e 10 a 15, enquanto os tempos de processamento para os estágios desvinculados são: $t_{PA,5} = 12$, $t_{PA,6} = 36$, $t_{PA,7} = 18$, $t_{PA,8} = 36$ e $t_{PA,9} = 36$.

O número de unidades em paralelo definidas para o biorreator e os valores estabelecidos para as variáveis de processamento sugerem que o projeto de tanques intermediários pode reduzir o custo do projeto, o qual será avaliado no seguinte item.

Figura 6.4 Configuração geral da planta multiproduto com desvinculação do tempo de processamento nos estágios 5 a 9 e sem tanques intermediários (cenário 5); as unidades superpostas operam em-fase.

6.3.2 Resultados para o modelo multiproduto com *scheduling* flexível e tanques intermediários (cenário 6)

Quando se consideram tanques intermediários, o processamento requer que seqüências com várias bateladas do mesmo produto sejam programadas; porém a diferença nos tempos de processamento para tPA, devido à flexibilização do *scheduling*, geraria tempos de espera entre os estágios. Para evitar este problema restringiram-se os tamanhos de batelada para este produto, tornando-os iguais, o que implica que o processamento de tPA não se beneficia da alocação de tanques intermediários.

O melhor resultado obtido para este cenário apresenta um custo de US\$ 4.976.480, o que representa uma economia de 16,0 % em comparação com o problema básico (cenário 2). As escolhas nas opções são as mesmas dos cenários anteriores (microfiltração e moinho de bolas), e apenas dois tanques intermediários foram projetados. A configuração geral da planta é apresentada na Figura 6.5, e o projeto completo é mostrado no Apêndice G. Em comparação ao cenário 5, a redução do custo é obtida na fermentação e nos estágios cromatográficos. Cabe lembrar que só com a alocação de tanques intermediários (cenários 3 e 4) não se obteve redução de custo nos estágios cromatográficos.

Figura 6.5 Configuração geral da planta multiproduto com desvinculação do tempo de processamento nos estágios 5 a 9 e tanques intermediários (cenário 6); as unidades superpostas operam em-fase.

6.3.3 Resultados para o modelo multiproduto com biorreator auxiliar (cenário 7)

Os resultados apresentados neste item e no próximo visam estudar a viabilização econômica da planta multiproduto com expressão de tPA por CHO através do projeto de um biorreator auxiliar para este produto. Para verificar os benefícios econômicos desta abordagem, resolveu-se o problema de síntese e projeto fixando a produção de tPA por CHO, cujo melhor resultado apresenta um custo de US\$ 6.491.336 e cujo projeto completo é apresentado no Apêndice G. As características principais da operação são comparadas na Tabela 6.3.

Tabela 6.3 Resultados gerais por produto para os cenário 2 com tPA/CHO e cenário 7.

Produto	Hospedeiro	Cenário 2 (tPA-CHO)		Cenário 7	
		B_i (kg)	TL_i (h)	B_i (kg)	TL_i (h)
INS	<i>S. cerevisiae</i> extracel.	3,86	5,00	3,86	5,00
HBV	<i>S. cerevisiae</i> intracel.	3,19	7,72	3,19	7,72
tPA	CHO	0,67	33,60	0,18	9,00
SOD	<i>S. cerevisiae</i> intracel.	2,71	15,43	2,71	15,43

Comparando os resultados para o cenário 2 com expressão de tPA por *E. coli* (item 5.5.2) e por CHO (Apêndice G), pode-se observar que o incremento no custo da planta deve-se principalmente ao aumento na dimensão do biorreator. Este aumento é necessário para processar a batelada de tPA, que se tornou 272% maior (0,18 contra 0,67 kg), enquanto as condições de operação para a produção de INS, HBV e SOD foram mantidas. Com o projeto de um biorreator auxiliar, a melhor solução obtida incorre num custo de US\$ 5.452.713, o que representa uma economia de 16,0% em relação ao projeto de planta multiproduto básica com expressão de tPA por CHO, e de 8,0% comparado ao projeto com expressão de tPA por *E. coli*. A economia de custo com a implementação do biorreator auxiliar é conseguida principalmente no estágio de fermentação se comparado ao cenário 2 com tPA por CHO e nos estágios 5 a 9 se comparado ao cenário 2 com tPA por *E. coli*, como pode ser observado na Tabela 6.4.

Tabela 6.4 Projeto e custos para planta multiproduto com biorreator auxiliar, sem tanques intermediários (cenário 7).

Estágio	unidades em-fase / fora-de-fase / dimensão			Custo total do estágio (US\$)		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 7	Cenário 2 (tPA- <i>E. coli</i>)	Cenário 2 (tPA-CHO)
1 auxiliar		1/ 1/ 8,576		230.183		
1		1/ 5/ 0,742		265.013	303.285	1.383.286
2B	1/ 1/ 8,576	1/ 1/ 2,503	1/ 1/ 12,007	52.748	13.182	59.948
3B	1/ 1/ 0,200	1/ 1/ 0,050		8.367	14.502	4.649
4B	1/ 1/ 0,200	1/ 1/ 0,481	1/ 1/ 0,371	6.916	8.861	7.562
5					53.448	
6					17.695	
7					58.549	
8					379.441	
9	1/ 1/ 12,007	1/ 1/ 3,574		34.107	114.540	74.336
10	1/ 1/ 0,814	5/ 1/ 0,718	1/ 1/ 7,623	1.315.854	1.348.325	1.348.325
11	1/ 1/ 7,623	1/ 1/ 24,691		63.717	65.940	65.939
12	1/ 2/ 0,259	5/ 2/ 0,726	1/ 2/ 7,719	2.644.662	2.691.506	2.691.506
13	1/ 1/ 7,719	1/ 1/ 50,000		100.232	111.903	111.903
14	1/ 1/ 0,234	3/ 1/ 0,623	1/ 1/ 0,374	722.278	735.063	735.063
15	1/ 1/ 0,374	1/ 1/ 0,747	1/ 1/ 0,374	8.634	8.820	8.820
Custo total				5.452.712	5.925.056	6.491.336

Dimensão dos equipamentos: m³ para unidades descontínuas (tanques e reatores), m² para filtros e colunas cromatográficas, m³/h para homogeneizador ou moinho de bolas e kW para centrífugas.

A Figura 6.6 apresenta a configuração geral da planta para o presente cenário, na qual pode-se observar uma redução importante no número de unidades e estágios, em relação aos projetos obtidos para os cenários anteriores. No seguinte item serão avaliados os benefícios adicionais que podem se obter com a alocação de tanques intermediários.

Figura 6.6 Configuração geral da planta multiproduto com reator auxiliar sem tanques intermediários (cenário 7); as unidades superpostas operam em-fase.

6.3.4 Resultados para o modelo multiproduto com biorreator auxiliar e tanques intermediários (cenário 8)

Assim como no cenário 6 (item 6.3.2), para a síntese e projeto com biorreator auxiliar e tanques intermediários em posições otimizadas, o tamanho de batelada de tPA foi igualado em todos os estágios. Logo, o processamento deste produto não se beneficia da alocação dos tanques intermediários. A melhor solução obtida tem um custo de US\$ 5.405.607, que representa uma economia de 0,9 % em relação ao cenário 7 e considera alocação de três tanques intermediários, como pode ser observado na Figura 6.4. Na Tabela 6.5 observa-se que em alguns equipamentos, especialmente no biorreator principal, obtém-se economia de capital considerável, porém os custos adicionais gerados pela alocação de tanques intermediários e o aumento do custo em outros estágios reduzem o ganho obtido. As variáveis de dimensionamento para cada estágio são apresentados no Apêndice G.

Tabela 6.5 Projeto e custos para planta multiproduto com biorreator auxiliar e tanques intermediários (cenário 8).

Estágio	unidades em-fase / fora-de-fase / dimensão			Custo total do estágio (US\$)		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 8	Cenário 7	Cenário 2 (tPA -CHO)
1 auxiliar		1/ 1/ 7,101		205.542	230.183	
1		1/ 2/ 1,191		140.804	265.013	1.383.286
	Tanque-pulmão 1	14,203		28.255		
2B	1/ 1/ 7,102	1/ 1/ 2,074	1/ 1/ 9,942	46.844	52.748	59.948
	Tanque-pulmão 2	19,884		34.576		
3B	1/ 1/ 0,200	1/ 1/ 0,050		8.367	8.367	4.649
4B	1/ 1/ 0,200	1/ 1/ 0,496	1/ 1/ 0,201	5.982	6.916	7.562
9	1/ 1/ 9,942	1/ 1/ 2,959		30.104	34.107	74.336
10	1/ 1/ 0,789	5/ 1/ 0,673	1/ 1/ 7,149	1.270.202	1.315.854	1.348.325
	Tanque-pulmão 10	18,750		33.379		
11	1/ 1/ 3,575	1/ 1/ 23,155		54.261	63.717	65.939
12	1/ 2/ 0,268	5/ 2/ 0,750	1/ 2/ 7,969	2.691.506	2.644.662	2.691.506
13	1/ 1/ 7,969	2/ 1/ 25,810		111.903	100.232	111.903
14	1/ 1/ 0,241	3/ 1/ 0,643	1/ 1/ 0,386	735.063	722.278	735.063
15	1/ 1/ 0,386	1/ 1/ 0,771	1/ 1/ 0,386	8.820	8.634	8.820
Custo total				5.405.607	5.452.712	6.491.336

Dimensão dos equipamentos: m³ para unidades descontínuas (tanques e reatores), m² para filtros e colunas cromatográficas, m³/h para homogeneizador ou moinho de bolas e kW para centrífugas.

Figura 6.7 Configuração geral da planta multiproduto com reator auxiliar e tanques intermediários (cenário 8); as unidades superpostas operam em-fase.

6.4 Conclusões

Os resultados obtidos com as considerações de flexibilização do *scheduling* e projeto de equipamentos auxiliares mostraram que estas modificações na modelagem são alternativas com grande potencial na otimização da síntese e projeto de plantas bioquímicas descontínuas. A combinação destas considerações com a alocação de tanques intermediários gerou projetos com economia de custo considerável e melhor aproveitamento dos equipamentos.

Outra observação importante é que a expressão de tPA por células de CHO torna-se economicamente viável, dentro da planta multiproduto estudada, quando se considera o projeto de um biorreator auxiliar exclusivo para este produto, pois esta consideração contorna a principal desvantagem deste tipo de produção, que corresponde ao longo tempo de fermentação.

A diversidade dos projetos definidos para os cenários resolvidos neste capítulo e no anterior permite concluir que a tomada de decisões na síntese e projeto de processos é uma tarefa complexa, devido à quantidade dos parâmetros e variáveis envolvidos, o que torna necessária a continuação no estudo de técnicas de programação matemática para a solução destes problemas.

Da mesma forma, a definição das posições para alocação de tanques intermediários nos cenários 4, 6 e 8 mostram que estas decisões não são simples e devem ser integradas no modelo de síntese e projeto.

CAPÍTULO 7 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

7.1 Aspectos gerais

A complexidade de alguns sistemas biotecnológicos requer o uso de técnicas de programação matemática avançadas para seu estudo quantitativo rigoroso. A quantificação de fluxos metabólicos e a síntese e projeto ótimos de plantas multiproduto são dois problemas que apresentam esta dificuldade; o primeiro gera um problema de otimização não-linear com a possibilidade, confirmada neste trabalho, de existirem soluções locais; por outro lado, o segundo é formulado como um problema misto-inteiro não-linear devido à introdução de variáveis binárias para representar as decisões na síntese e aos modelos não-lineares para a operação das unidades.

Em virtude disso, o presente trabalho estudou técnicas de programação matemática para a análise e síntese de tais sistemas biotecnológicos. Para o problema de análise de fluxos metabólicos verificou-se a existência de soluções locais e desenvolveu-se uma abordagem de otimização global, baseada na técnica de *Branch & Bound* espacial (Esposito e Floudas, 1998). O problema de otimização simultânea da síntese e projeto de processos bioquímicos foi formulado por GDP, empregando a abordagem *big-M* para relaxar as disjunções. Para reduzir o custo do projeto avaliou-se: 1) alocação de tanques intermediários; 2) considerações especiais no *scheduling* para flexibilizar as restrições da operação em alguns equipamentos e evitar o seu sobre-dimensionamento, e 3) projeto de equipamentos auxiliares em estágios com tempo de processamento crítico.

7.2 Sumário do trabalho

Este trabalho propôs técnicas de programação matemática para abordar problemas de análise e síntese de sistemas biotecnológicos. O sumário do trabalho para os dois problemas considerados é apresentado a seguir.

7.2.1 Análise de fluxos metabólicos

Os experimentos para quantificação de fluxos metabólicos que envolvem medições de marcação viabilizam a análise de redes metabólicas complexas, porém as não-linearidades, que os balanços de marcações ou isotopômeros adicionam ao modelo, impedem a solução direta do problema, tornando necessária a aplicação de técnicas de otimização. Neste trabalho, empregaram-se medidas de marcação por GC-MS que permitem a modelagem na forma de marcações fracionárias e desenvolveu-se uma generalização das técnicas de matrizes de mapeamento de átomos. Os resultados obtidos para o metabolismo central do carbono na *S. cerevisiae* permitem concluir que:

- Problemas deste tipo têm soluções locais, as quais correspondem a distribuições diferentes para a rede metabólica; devido a isso, é necessário desenvolver abordagens que garantam a otimalidade global da solução.
- A abordagem *Branch & Bound* espacial é uma alternativa viável para resolver globalmente estes problemas, e os estimadores de McCormick são eficientes na relaxação convexa do problema original, permitindo gerar limites inferiores para a solução global.
- As marcações fracionárias somadas e os fluxos líquidos não-medidos são as variáveis mais vantajosas para a divisão da região de busca, sendo necessárias múltiplas divisões em variáveis de ambos grupos. O maior avanço na convergência é devido ao procedimento de ajuste dos limites, o que torna esta etapa indispensável na resolução do problema.

7.2.2 Síntese e projeto de bioprocessos

A solução simultânea dos problemas de síntese e projeto permite reduzir os custos de capital, mas gera problemas de alta complexidade devido à combinação de decisões (escolha de hospedeiro para produção e de operações em alguns estágios, e definição de arranjos de equipamentos em paralelo) com modelos não-lineares para o projeto das unidades, os quais dependem das decisões tomadas para todos os estágios da planta. Os

resultados obtidos considerando uma planta descontínua para produção de insulina, vacina para hepatite B (HBV), ativador de plasminogênio tecidual (tPA) e superóxido dismutase mediante *S. cerevisiae* com expressão extra ou intracelular, *Escherichia coli* e células de mamíferos permitem concluir que:

- A formulação das decisões por abordagem *big-M* permite resolver simultaneamente os problemas de síntese e projeto com esforço computacional razoável.
- A complexidade do problema não permite que as decisões sejam tomadas com base unicamente na experiência. Exemplo disto é que para o problema multiproduto básico (cenário 2) escolheu-se expressão intracelular para HBV e tPA, apesar de envolver mais estágios na recuperação do produto; nesta escolha outras variáveis como o tempo de processamento tiveram efeito preponderante.
- Assim como a seleção dos hospedeiros, a escolha entre as operações disponíveis não depende só das características dos produtos mas também das variáveis de operação, reafirmando que os problemas de síntese e projeto estão fortemente associados e devem ser resolvidos simultaneamente.
- O projeto de plantas multiproduto requer a conciliação do processamento de vários produtos dentro do horizonte de tempo disponível, tornando necessária a duplicação de unidades e gerando tempos ociosos elevados, o que pode aumentar o custo total do projeto se não for aplicada alguma estratégia para melhorar o aproveitamento dos equipamentos.
- A alocação de tanques intermediários, ainda que em posições otimizadas, tem um efeito limitado na diminuição da ociosidade e do custo da planta. Por outro lado, a flexibilização do *scheduling* e o projeto de unidades auxiliares associados com a alocação de tanques mostrou um grande potencial na diminuição dos custos no projeto.

7.3 Contribuições deste trabalho

Como principais contribuições deste trabalho podem-se citar:

- Generalização da técnica de matrizes de mapeamento de átomos para a modelagem dos balanços de marcações.
- Confirmação da superioridade das abordagens determinísticas (programação matemática) sobre as metaheurísticas (algoritmo evolucionário) na solução de problemas não-lineares.
- Confirmação da existência de soluções locais nos problemas de quantificação de fluxos metabólicos com balanços de marcações.
- Desenvolvimento e avaliação de uma abordagem de otimização global para a quantificação de fluxos metabólicos, incluindo a análise dos parâmetros envolvidos na convergência da solução.
- Implementação de um modelo misto-inteiro não-linear (MINLP) para otimizar simultaneamente a síntese e projeto de plantas descontínuas multiproduto.
- Estudo da alocação de unidades em paralelo e tanques intermediários, considerando as modificações na modelagem e os efeitos na síntese e projeto da planta.
- Desenvolvimento de modificações na modelagem para reduzir o sobre-dimensionamento e melhorar o aproveitamento de unidades que não participam do processo para todos os produtos. Estas modificações requerem flexibilização do *scheduling*, porém garantem a disponibilidade dos equipamentos.
- Consideração do projeto de unidades auxiliares em estágios com tempo de operação elevado. Estas unidades viabilizam opções de processamento que envolvem tempos de processamento crítico.
- Formulação e solução de problemas relevantes em sistemas biotecnológicos através de programação matemática.

7.4 Recomendações para trabalhos futuros

A análise de fluxos metabólicos permite descrever o funcionamento das células, o que é interessante e necessário; porém, mais interessante ainda é poder prever o metabolismo em novas condições ou após modificações genéticas. O anterior implica na modelagem ainda mais completa do metabolismo celular, com testes *in vivo* que permitam verificar e ajustar os modelos; assim, um trabalho desses requer a combinação de dados do genoma (genes), transcriptoma (mRNA), proteoma (proteínas), metaboloma (metabólitos), fluxoma (fluxos metabólicos) com técnicas de programação matemática para modelagem e otimização. Outros estudos que podem ser desenvolvidos, dentro da análise de fluxos metabólicos são:

- Estudo de alternativas para a redução da carga computacional durante a convergência na busca global.
- Novos estudos de caso para verificar a existência ou não de soluções locais e gerar conclusões contundentes a respeito da necessidade de aplicar abordagens de otimização global para a estimação de fluxos metabólicos.
- Aplicação de técnicas de programação matemática na quantificação de fluxos metabólicos empregando medidas de ressonância magnética nuclear (NMR).
- Comparação da modelagem incluindo balanços de marcações fracionárias, empregada na presente tese, com as que incluem balanços de isotopômeros e cumômeros.

Quanto à síntese e projeto de processos, é interessante a implementação de modelos mais complexos que permitam descrever o problema com maior precisão. No entanto, o aumento de complexidade pode inviabilizar a solução do problema com a metodologia proposta; assim, uma alternativa é o desenvolvimento de modelos em vários níveis de complexidade, para resolver o problema com abordagens hierárquicas. Alguns desenvolvimentos que podem ser considerados são:

- Implementar modelos de desempenho mais complexos, como os empregados em Pinto *et al.* (2001), os quais descrevem as operações com maior precisão, porém

aumentam a complexidade do problema devido à introdução de novas variáveis contínuas para o dimensionamento das unidades.

- Generalizar a modelagem da flexibilização do *scheduling* e da alocação de equipamentos auxiliares, permitindo a otimização da escolha dos estágios para a implementação destas considerações.
- Considerar valores discretos para as dimensões dos equipamentos, o que reproduz com maior fidelidade o projeto de plantas de processamento.
- Adicionar custos operacionais na função objetivo, os quais, acredita-se que têm peso na escolha dos hospedeiros e das operações.

APÊNDICE A MODELO PARA ANÁLISE DE FLUXOS METABÓLICOS

O modelo para o metabolismo central de *S. cerevisiae* é baseado no modelo descrito por Gombert *et al.* (2001). A Tabela A.1 apresenta os metabólitos envolvidos com as abreviaturas do inglês para facilitar a relação com publicações internacionais.

Tabela A.1 Metabólitos na rede para o metabolismo central de *S. cerevisiae*.

Metabólitos intracelulares		Outros metabólitos	
ACA	Acetaldeído	ACEout	Acetato excretado ao meio
ACE	Acetato	AcCoAcitout	Acetil-CoA citosólico usado para formação de biomassa
AcCoAcit	Acetil-CoA citosólico	AcCoAmitout	Acetil-CoA mitocondrial usado para formação de biomassa
AcCoAmit	Acetil-CoA mitocondrial	AKGout	α -Cetoglutarato usado para formação de biomassa
AKG	α -Cetoglutarato	COout	Dióxido de carbono excretado
CO	Dióxido de carbono	E4Pout	Eritrose 4-fosfato usado para formação de biomassa
E4P	Eritrose 4-fosfato	ETHout	Etanol excretado ao meio
ETH	Etanol	GLC	Glicose
F6P	Frutose 6-fosfato	GLYout	Glicina usada para formação de biomassa
FUM	Fumarato	GLYCout	Glicerol usado para formação de biomassa
GLY	Glicina	G3Pout	Gliceraldeído 3-fosfato usado para formação de biomassa
GLYC	Glicerol	G6Pout	Glicose 6-fosfato usado para formação de biomassa
G3P	Gliceraldeído 3-fosfato	OAAcitout	Oxaloacetato citosólico usado para formação de biomassa
G6P	Glicose 6-fosfato	PEPout	fosfoenolpiruvato usado para formação de biomassa
ICIT	Isocitrato	P5Pout	Pentose 5-fosfato usado para formação de biomassa
OAAcit	Oxaloacetato citosólico	PYRcitout	Piruvato citosólico usado para formação de biomassa
OAAmit	Oxalacetato mitocondrial	PYRmitout	Piruvato mitocondrial usado para formação de biomassa
PEP	fosfoenolpiruvato	SERout	Serina usada para formação de biomassa
P5P	Pentose 5-fosfato	THRout	Treonina usada para formação de biomassa
PYRcit	Piruvato citosólico		
PYRmit	Piruvato mitocondrial		
SER	Serina		
S7P	Sedoheptulose 7-fosfato		
THR	Treonina		

A Tabela A.2 apresenta as reações consideradas na rede, incluindo a migração dos átomos de carbono dentro de cada uma.

Tabela A.2 Reações na rede para o metabolismo central de *S. cerevisiae*.

Incorporação de glicose		AcCoAcit AcCoAmit	
GLC	G6P	ab	ab
abcde	abcdef	PYRcit	PYRmit
		abc	abc
Via EMP		Metabolismo da treonina, serina e glicina	
G6P	F6P	(assume-se que todas as enzimas estão no citosol)	
abcdef	abcdef	G3P	SER
F6P + G6P	G6P + F6P	abc	abc
abcdef ghijkl	abcdef ghijkl	SER	GLY + C-1
F6P	G3P + G3P	abc	ab c
abcdef	cba def	SER + GLY + C-1	GLY + C-1 + SER
G3P	PEP	abc de f	ab c def
abc	abc	OAAcit	THR
PEP	PYRcyt	abcd	abcd
abc	abc	THR	GLY + ACA
Via PP		abcd	ab cd
G6P	CO + P5P	GLY + ACA + THR	THR + GLY + ACA
abcdef	a bcdef	ab cd efgh	abcd ef gh
P5P + P5P	S7P + G3P	Enzima málica (descarboxilação do malato mitocond.)	
abcde fghij	fgabcde hij	OAAmit	PRYmit + CO
S7P + G3P + P5P + P5P	P5P + P5P + S7P + G3P	abcd	abc d
fgabcde hij klmno pqrst	abcde fghij klpqrst mno	Incorporação de intermediários nas macro-moléculas	
S7P + G3P	F6P + E4P	G6P	G6Pout
abcdefg hij	abchij defg	abcdef	abcdef
F6P + E4P + S7P + G3P	S7P + G3P + F6P + E4P	P5P	P5Pout
abchij defg klmnopq rst	abcdefg hij klmrst nopq	abcde	abcde
P5P + E4P	F6P + G3P	E4P	E4Pout
abcde fghi	abfghi cde	abcd	abcd
F6P + G3P + P5P + E4P	P5P + E4P + F6P + G3P	G3P	G3Pout
abfghi cde jklmn opqr	abcde fghi jkopqr lmn	abc	abc
Desvio da pdh		PEP	PEPout
PYRcit	ACA + CO	abc	abc
abc	ab c	PYRmit	PYRmitout
ACA	ACE	abc	abc
ab	ab	PYRcit	PYRcitout
ACE	AcCoAcit	abc	abc
ab	ab	OAAcit	OAAcitout
Formação de glicerol e etanol		abcd	abcd
G3P	GLYC	AKG	AKGout
abc	abc	abcde	abcde
ACA	ETH	AcCoAcit	AcCoAcitout
ab abc	ab	ab	ab
Reação anaplerótica (no citosol)		AcCoAmit	AcCoAmitout
PYRcit + CO	OAAcit	ab	ab
abc d	abcd	SER	SERout
Ciclo TCA (considerando embaralhamento no FUM)		abc	abc
pdh: PYRmit	AcCoAmit + CO	GLY	GLYout
abc	ab c	ab	ab
OAAmit + AcCoAmit	ICIT	C-1	C-1out
abcd ef	dcbafe	a	a
ICIT	AKG + CO	THR	THRout
abcdef	abcef d	abcd	abcd
AKG	FUM + CO	Excreção de produtos	
abcde	bcde a	ETH	ETHout
FUM + FUM	OAAmit + OAAmit	ab	ab
abcd efgh	abcd hgfe	ACE	ACEout
OAAmit	FUM	ab	ab
abcd	abcd	GLYC	GLYCout
Fluxos de transporte		abc	abc
OAAmit	OAAcit	Evolução do CO₂	
abcd	abcd	CO	Coout
OAAcit	OAAmit	a	a
abcd	abcd		

APÊNDICE B AVALIAÇÃO DO PARÂMETRO DE INCERTEZA

Durante a identificação de soluções locais, item 3.4.1, para um valor pequeno do parâmetro de incerteza ($d=0,1$), o qual define a região de busca para as variáveis medidas, foi observado que a marcação fracionária somada (*SFL*) para Gly85 adota o seu valor-limite inferior (ver Tabela B.1), o que sugere que a solução pode ser melhorada. Duas alternativas foram consideradas: retirar a variável da função objetivo ou expandir a região de busca mediante o aumento do valor para o parâmetro de incerteza.

Tabela B.1 Análise da eliminação de variáveis na MFA.

Função objetivo	Variável problemática	Valor-limite inferior	Valor na solução	Valor-limite superior
Incluindo todas as variáveis				
23,06	Gly85	3,30	3,30	4,07
Eliminando Gly85				
22,83	Phe143	2,70	3,30	3,30
Eliminando Gly85 e Phe143				
22,68	Ser175	2,79	3,41	3,41
	Val143	6,21	7,59	7,59
	Asp115	10,80	10,80	13,20

Uma desvantagem da eliminação da variável é o desconhecimento da qualidade do seu valor medido em relação aos das outras variáveis. A eliminação da *SFL* para Gly85 induziu uma leve melhoria na função objetivo, porém o problema (adoção de valores-limite) apresentou-se na *SFL* para Phe143, e com a eliminação da *SFL* para Phe143 o problema apresentou-se nas *SFL* para Ser175, Val143 e Asp115 (ver Tabela B1). Isto sugere que os valores medidos para estas variáveis não têm qualidade inferior aos de outras variáveis, e que a causa do problema é o tamanho da região viável.

Por outro lado, para $d=0,2$, os valores das variáveis nas soluções locais não adotam seus valores-limite, e para $d=0,3$ não identificou-se nenhuma solução local diferente às apresentadas no item 3.4.1; isto sugere que o valor $d=0,2$ é consistente com os dados experimentais empregados.

APÊNDICE C REPRESENTAÇÃO DE DECISÕES POR PROGRAMAÇÃO DISJUNTIVA GENERALIZADA

Para a formulação de problemas com seleção de alternativas foi desenvolvida recentemente a programação disjuntiva generalizada ou *generalized disjunctive programming* (GDP), formalizada em Lee e Grossmann (2000) e Vecchietti *et al.* (2003). Lee e Grossmann (2000) propõem que o problema seja representado como:

$$\begin{aligned}
 \min \quad & Z = \sum_{k \in K} c_k + f(x) \\
 \text{s.a} \quad & r(x) \leq 0 \\
 & \bigvee_{j \in D_k} \begin{bmatrix} Y_{jk} \\ g_{jk}(x) \leq 0 \\ c_k = \underline{g}_{jk} \end{bmatrix} \quad k \in K \\
 & \Omega(Y) = \text{verdadeiro} \\
 & x \geq 0, \quad c_k \geq 0, \quad Y_{jk} \in \{\text{verdadeiro}, \text{falso}\}
 \end{aligned} \tag{C.1}$$

onde $x \in R^n$ é o vetor de variáveis contínuas, Y_{jk} são variáveis Booleanas, $c_k \in R^1$ são variáveis contínuas, $\underline{g}_k$ são os parâmetros fixos, $f: R^n \rightarrow R^1$ é o termo contendo as variáveis contínuas x na função objetivo, $r: R^n \rightarrow R^q$ é o grupo de restrições que não dependem das decisões discretas, enquanto $g_{jk}(x)$ é o conjunto de restrições não-lineares que depende das decisões, $g_{jk}^i(x): R^n \rightarrow R^1$, $i \in I_{jk}$, $k \in K$ onde I_{jk} é um grupo de restrições indexado, $g_{jk}(x)$, $f(x)$ e $r(x)$ são funções convexas.

A disjunção é composta por um operador OU ($\vee$) com vários termos; cada um formado pelas variáveis Booleanas Y_{jk} , um grupo de restrições $g_{jk}(x)$ e uma variável de custo c_k ; se Y_{jk} é verdadeira, então as restrições $g_{jk}(x) \leq 0$ e $c_k = \underline{g}_k$ são consideradas; do contrário, são ignoradas. A proposição lógica $\Omega(Y)$ contém a relação entre as variáveis Booleanas, definindo as condições possíveis. Cada condição atribui o valor *verdadeiro* para um conjunto de variáveis Booleanas.

APÊNDICE D FORMULAÇÃO DE DISJUNÇÕES POR ABORDAGEM *BIG-M*

Vecchietti *et al.* (2003) consideraram o seguinte grupo de disjunções não-lineares:

$$F = \bigvee_{i \in D} [h_i(x) \leq 0] \quad x \in R^n \quad (D.1)$$

onde $h_i(x)$ é uma função contínua convexa, F é uma expressão lógica que ativa um único grupo de restrições e D é o conjunto de opções possíveis. A relaxação *big-M* para essa disjunção é dada por:

$$\begin{aligned} h_i(x) &\leq M1_i(1 - y_i) & i \in D \\ \sum_{i \in D} y_i &= 1 \end{aligned} \quad (D.2)$$

Em (D.2), $M1_i$ é um parâmetro grande o suficiente para que se $y_i = 0$ a restrição torne-se redundante; por outro lado, se $y_i = 1$ a restrição deve ser satisfeita. A qualidade da relaxação, e conseqüentemente o esforço na solução do problema de otimização gerado, depende dos valores definidos para as constantes $M1_i$. O melhor valor para $M1_i$ é definido por:

$$M1_i = \max \left[h_i(x), \quad x^{lo} \leq x \leq x^{up} \right] \quad i \in D \quad (D.3)$$

APÊNDICE E PARÂMETROS PARA O ESTUDO DE CASO DE SÍNTESE E PROJETO ÓTIMO

Este estudo de caso é baseado no problema apresentado por Montagna *et al.* (2004). O desempenho de cada estágio é modelado com o tempo de operação e o parâmetro de dimensionamento ou trabalho, este último transformado em parâmetro de tempo (ver item 5.2.5). Como mencionado no item 5.2.3, os parâmetros dependem do rendimento dos estágios posteriores e da procedência do fluxo alimentado. A Tabela E.1 apresenta o rendimento para cada equipamento nas possíveis combinações produto-hospedeiro.

Tabela E.1 Rendimentos das operações.

Operação		Produto/hospedeiro							
		INS		HBV		tPA		SOD	
		<i>Sacch.</i>	<i>E. coli</i>	<i>Sacch.</i>	CHO	<i>E. coli</i>	CHO	<i>Sacch.</i>	<i>E. coli</i>
1	Fermentação	1	1	1	1	1	1	1	1
2A	Centrifugação	0,75	1	1	0,8	1	0,8	1	1
2B	Microfiltração	0,85	1	1	0,9	1	0,9	1	1
3A	Homogeneização		0,7	0,75		0,7		0,75	0,7
3B	Moinho de bolas		0,8	0,85		0,8		0,85	0,8
4A	Centrifugação		1	0,8		1		0,8	1
4B	Microfiltração		1	0,9		1		0,9	1
5	Solubilização de IB		0,7			0,7			0,7
6	Diafiltração		1			1			1
7	Sulfonação		0,9			0,9			0,9
8	Refolding		0,6			0,6			0,6
9	Ultrafiltração		1		1	1	1		1
10	Cromatografia	0,75	0,75	0,8	0,8	0,9	0,9	0,85	0,85
11	Ultra-diafiltração	1	1	1	1	1	1	1	1
12	Cromatografia	0,9	0,9	0,85	0,85	0,9	0,9	0,85	0,85
13	Ultra-diafiltração	1	1	1	1	1	1	1	1
14	Gel filtração	0,8	0,8	0,8	0,8	0,8	0,8	0,8	0,8
15	Filtração estéril	1	1	1	1	1	1	1	1

A Tabela E.2 apresenta as constantes CS_{ihjd} e CT_{ihjd}^l ou CS_{ihjr} e CT_{ihjr}^l e os estágios considerados na produtória dos rendimentos, para a estimação dos parâmetros de dimensionamento e tempo (equações 5.5 e 5.6 ou 5.9 e 5.10). Além disso, a Tabela E.3 apresenta a função para estimação do tempo de processamento.

Tabela E.2 Constantes para estimação dos parâmetros de dimensionamento e tempo.

Estágio, equipamento	Produto/hospedeiro							
	INS <i>Sacch.</i>	INS <i>E.coli</i>	HBV <i>Sacch.</i>	HBV CHO	tPA <i>E.coli</i>	tPA CHO	SOD <i>Sacch.</i>	SOD <i>E.coli</i>
1. Fermentação	0,05 <i>j: {2-15}</i>	0,03 <i>j: {2-15}</i>	0,1 <i>j: {2-15}</i>	0,5 <i>j: {2-15}</i>	1 <i>j: {2-15}</i>	10 <i>j: {2-15}</i>	0,04 <i>j: {2-15}</i>	0,05 <i>j: {2-15}</i>
2.A Centrifugação								
Tanque alimento	0,05 <i>j: {2-15}</i>	0,03 <i>j: {2-15}</i>	0,1 <i>j: {2-15}</i>	0,5 <i>j: {2-15}</i>	1 <i>j: {2-15}</i>	10 <i>j: {2-15}</i>	0,04 <i>j: {2-15}</i>	0,05 <i>j: {2-15}</i>
Centrífuga D_{ih}	0,0025 <i>j: {2-15}</i>	0,15 <i>j: {2-15}</i>	0,005 <i>j: {2-15}</i>	0,0025 <i>j: {2-15}</i>	5 <i>j: {2-15}</i>	0,05 <i>j: {2-15}</i>	0,002 <i>j: {2-15}</i>	0,125 <i>j: {2-15}</i>
Tanque produto	0,05 <i>j: {2-15}</i>	0,015 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>	0,5 <i>j: {2-15}</i>	0,5 <i>j: {2-15}</i>	10 <i>j: {2-15}</i>	0,01 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>
2.B Microfiltração								
Tanque de retido	0,05 <i>j: {2-15}</i>	0,03 <i>j: {2-15}</i>	0,1 <i>j: {2-15}</i>	0,5 <i>j: {2-15}</i>	1 <i>j: {2-15}</i>	10 <i>j: {2-15}</i>	0,04 <i>j: {2-15}</i>	0,05 <i>j: {2-15}</i>
Microfiltro D_{ih}	0,5 <i>j: {2-15}</i>	0,12 <i>j: {2-15}</i>	0,375 <i>j: {2-15}</i>	2,2 <i>j: {2-15}</i>	4 <i>j: {2-15}</i>	44 <i>j: {2-15}</i>	0,15 <i>j: {2-15}</i>	0,2 <i>j: {2-15}</i>
Tanque de filtrado	0,1 <i>j: {2-15}</i>	Não	Não	0,7 <i>j: {2-15}</i>	Não	14 <i>j: {2-15}</i>	Não	Não
3.A Homogeneização								
Tanque de <i>holding</i>	-	0,015 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>	-	0,5 <i>j: {2-15}</i>	-	0,01 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>
Homogeneizador	-	0,045 <i>j: {2-15}</i>	0,1 <i>j: {2-15}</i>	-	1,5 <i>j: {2-15}</i>	-	0,04 <i>j: {2-15}</i>	0,075 <i>j: {2-15}</i>
3.B Moinho de bolas								
Tanque de <i>holding</i>	-	0,015 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>	-	0,5 <i>j: {2-15}</i>	-	0,01 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>
Moinho bolas D_{ih}	-	0,045 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>	-	1,5 <i>j: {2-15}</i>	-	0,01 <i>j: {2-15}</i>	0,075 <i>j: {2-15}</i>
4.A Centrifugação								
Tanque alimento	-	0,015 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>	-	0,5 <i>j: {2-15}</i>	-	0,01 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>
Centrífuga D_{ih}	-	0,75 <i>j: {2-15}</i>	1,25 <i>j: {2-15}</i>	-	25 <i>j: {2-15}</i>	-	0,5 <i>j: {2-15}</i>	1,25 <i>j: {2-15}</i>
Tanque produto	-	0,00375 <i>j: {2-15}</i>	0,02 <i>j: {2-15}</i>	-	0,125 <i>j: {2-15}</i>	-	0,008 <i>j: {2-15}</i>	0,00625 <i>j: {2-15}</i>
4.B Microfiltração								
Tanque de retido	-	0,015 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>	-	0,5 <i>j: {2-15}</i>	-	0,01 <i>j: {2-15}</i>	0,025 <i>j: {2-15}</i>
Microfiltro D_{ih}	-	0,1875 <i>j: {2-15}</i>	0,5 <i>j: {2-15}</i>	-	6,25 <i>j: {2-15}</i>	-	0,2 <i>j: {2-15}</i>	0,3125 <i>j: {2-15}</i>

Tabela E.2 (continuação) Constantes para estimação dos parâmetros.

Estágio, equipamento	Produto/hospedeiro							
	INS <i>Sacch.</i>	INS <i>E.coli</i>	HBV <i>Sacch.</i>	HBV CHO	tPA <i>E.coli</i>	tPA CHO	SOD <i>Sacch.</i>	SOD <i>E.coli</i>
Tanque de filtrado	-	Não	0,05 <i>j</i> : {2-15}	-	Não	-	0,02 <i>j</i> : {2-15}	Não
5. IB solubilização	-	0,1 <i>j</i> : {5-15}	-	-	1 <i>j</i> : {5-15}	-	-	0,5 <i>j</i> : {5-15}
6. Diafiltração								
Tanque de <i>holding</i>	-	0,1 <i>j</i> : {5-15}	-	-	1 <i>j</i> : {5-15}	-	-	0,5 <i>j</i> : {5-15}
Diafiltro D_{ih}	-	7 <i>j</i> : {5-15}	-	-	70 <i>j</i> : {5-15}	-	-	35 <i>j</i> : {5-15}
7. Sulfonação	-	0,12 <i>j</i> : {5-15}	-	-	1,2 <i>j</i> : {5-15}	-	-	0,6 <i>j</i> : {5-15}
8. Refolding	-	1 <i>j</i> : {8-15}	-	-	20 <i>j</i> : {8-15}	-	-	2 <i>j</i> : {8-15}
9. Ultrafiltração								
Tanque de <i>holding</i>	-	1 <i>j</i> : {8-15}	-	de 2.A 0,5 <i>j</i> : {3-15} de 2.B 0,7 <i>j</i> : {2-15}	20 <i>j</i> : {8-15}	de 2.A 10 <i>j</i> : {3-15} de 2.B 14 <i>j</i> : {2-15}	-	2 <i>j</i> : {8-15}
Ultrafilter D_{ih}	-	50 <i>j</i> : {8-15}	-	de 2.A 25 <i>j</i> : {3-15} de 2.B 35 <i>j</i> : {2-15}	1000 <i>j</i> : {8-15}	de 2.A 500 <i>j</i> : {3-15} de 2.B 700 <i>j</i> : {2-15}	-	100 <i>j</i> : {8-15}
10. Cromatografia								
Tanque alimento	de 2.A 0,05 <i>j</i> {3-15} de 2.B 0,1 <i>j</i> {2-15}	0,4 <i>j</i> {10-15}	de 4.A 0,02 <i>j</i> {2-15} de 4.B 0,05 <i>j</i> {2-15}	0,4 <i>j</i> {10-15}	0,4 <i>j</i> {10-15}	0,4 <i>j</i> {10-15}	de 4.A 0,008 <i>j</i> {2-15} de 4.B 0,02 <i>j</i> {2-15}	0,4 <i>j</i> {10-15}
Coluna Cromatográfica	0,5 <i>j</i> {10-15}	0,5 <i>j</i> {10-15}	0,5 <i>j</i> {10-15}	0,8 <i>j</i> {10-15}	0,5 <i>j</i> {10-15}	0,8 <i>j</i> {10-15}	0,8 <i>j</i> {10-15}	0,8 <i>j</i> {10-15}
Tanque produto	0,1 <i>j</i> {11-15}	0,1 <i>j</i> {11-15}	0,1 <i>j</i> {11-15}	2 <i>j</i> {11-15}	0,1 <i>j</i> {11-15}	2 <i>j</i> {11-15}	2 <i>j</i> {11-15}	2 <i>j</i> {11-15}
11. Ultra-diafiltração								
Tanque de <i>holding</i>	0,1 <i>j</i> {11-15}	0,1 <i>j</i> {11-15}	0,1 <i>j</i> {11-15}	2 <i>j</i> {11-15}	0,1 <i>j</i> {11-15}	2 <i>j</i> {11-15}	2 <i>j</i> {11-15}	2 <i>j</i> {11-15}
Ultrafiltro D_{ih}	5 <i>j</i> {11-15}	5 <i>j</i> {11-15}	5 <i>j</i> {11-15}	100 <i>j</i> {11-15}	5 <i>j</i> {11-15}	100 <i>j</i> {11-15}	100 <i>j</i> {11-15}	100 <i>j</i> {11-15}

Tabela E.2 (continuação) Constantes para estimação dos parâmetros.

Estágio, equipamento	Produto/hospedeiro							
	INS <i>Sacch.</i>	INS <i>E.coli</i>	HBV <i>Sacch.</i>	HBV CHO	tPA <i>E.coli</i>	tPA CHO	SOD <i>Sacch.</i>	SOD <i>E.coli</i>
12. Cromatografia								
Tanque alimento	0,05 $j\{12-15\}$	0,05 $j\{12-15\}$	0,05 $j\{12-15\}$	0,05 $j\{12-15\}$	0,05 $j\{12-15\}$	0,05 $j\{12-15\}$	0,05 $j\{12-15\}$	0,05 $j\{12-15\}$
Coluna Cromatográfica	0,7 $j\{12-15\}$	0,7 $j\{12-15\}$	0,8 $j\{12-15\}$	0,8 $j\{12-15\}$	0,8 $j\{12-15\}$	0,8 $j\{12-15\}$	0,8 $j\{12-15\}$	0,8 $j\{12-15\}$
Tanque produto	1 $j\{13-15\}$	1 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$
13. Ultra-diafiltração								
Tanque de <i>holding</i>	1 $j\{13-15\}$	1 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$	2 $j\{13-15\}$
Ultrafiltro D_{ih}	50 $j\{13-15\}$	50 $j\{13-15\}$	100 $j\{13-15\}$	100 $j\{13-15\}$	100 $j\{13-15\}$	100 $j\{13-15\}$	100 $j\{13-15\}$	100 $j\{13-15\}$
14. Gel filtração								
Tanque alimento	0,05 $j\{14-15\}$	0,05 $j\{14-15\}$	0,05 $j\{14-15\}$	0,05 $j\{14-15\}$	0,05 $j\{14-15\}$	0,05 $j\{14-15\}$	0,05 $j\{14-15\}$	0,05 $j\{14-15\}$
Coluna	0,4 $j\{14-15\}$	0,4 $j\{14-15\}$	0,4 $j\{14-15\}$	0,4 $j\{14-15\}$	0,4 $j\{14-15\}$	0,4 $j\{14-15\}$	0,4 $j\{14-15\}$	0,4 $j\{14-15\}$
Tanque produto	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$
15. Filtração estéril								
Tanque alimento	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$
Filtro D_{ih}	1	1	1	1	1	1	1	1
Tanque produto	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$	0,1 $j\{15\}$

Tabela E.3 Funções para o tempo de processamento em cada estágio (h).

Estágio, operação	Produto/hospedeiro			
	INS/ <i>Sacch.</i>	INS/ <i>E. Coli</i>	HBV/ <i>Sacch.</i>	HBV/CHO
1. Fermentação	24	24	24	168
2.A Centrifugação	$\frac{0,0025 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{0,15 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{0,005 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{0,0025 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$
2.B Microfiltração	$\frac{0,5 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{0,12 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{0,375 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{2,2 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$
3.A Homogeneização	-	$\frac{0,045 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{0,1 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	-
3.B Moinho de bolas	-	$\frac{0,045 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{0,025 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	-
4.A Centrifugação	-	$\frac{0,75 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{1,25 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	-
4.B Microfiltração	-	$\frac{0,1875 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	$\frac{0,5 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$	-
5. IB solubilização	-	8	-	-
6. Diafiltração	-	$\frac{7 B_i}{G_{jd} R_{jd} \prod_{j=5}^{15} h_{ihj}}$	-	-
7. Sulfonação	-	12	-	-
8. Refolding	-	12	-	-
9. Ultrafiltração		$\frac{50 B_i}{G_{jd} R_{jd} \prod_{j=8}^{15} h_{ihj}}$	de 2.A $\frac{25 B_i}{G_{jd} R_{jd} \prod_{j=3}^{15} h_{ihj}}$	de 2.B $\frac{35 B_i}{G_{jd} R_{jd} \prod_{j=2}^{15} h_{ihj}}$
10. Cromatografia	$3,5 + \frac{0,6 B_i}{G_{jd} R_{jd} \prod_{j=10}^{15} h_{ihj}}$	$3,5 + \frac{0,6 B_i}{G_{jd} R_{jd} \prod_{j=10}^{15} h_{ihj}}$	$3,5 + \frac{0,6 B_i}{G_{jd} R_{jd} \prod_{j=10}^{15} h_{ihj}}$	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{j=10}^{15} h_{ihj}}$
11. Ultra-diafiltração	$\frac{5 B_i}{G_{jd} R_{jd} \prod_{j=11}^{15} h_{ihj}}$	$\frac{5 B_i}{G_{jd} R_{jd} \prod_{j=11}^{15} h_{ihj}}$	$\frac{5 B_i}{G_{jd} R_{jd} \prod_{j=11}^{15} h_{ihj}}$	$\frac{100 B_i}{G_{jd} R_{jd} \prod_{j=11}^{15} h_{ihj}}$
12. Cromatografia	$8 + \frac{0,5 B_i}{G_{jd} R_{jd} \prod_{j=12}^{15} h_{ihj}}$	$8 + \frac{0,5 B_i}{G_{jd} R_{jd} \prod_{j=12}^{15} h_{ihj}}$	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{j=12}^{15} h_{ihj}}$	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{j=12}^{15} h_{ihj}}$
13. Ultra-diafiltração	$\frac{50 B_i}{G_{jd} R_{jd} \prod_{j=13}^{15} h_{ihj}}$	$\frac{50 B_i}{G_{jd} R_{jd} \prod_{j=13}^{15} h_{ihj}}$	$\frac{100 B_i}{G_{jd} R_{jd} \prod_{j=13}^{15} h_{ihj}}$	$\frac{100 B_i}{G_{jd} R_{jd} \prod_{j=13}^{15} h_{ihj}}$
14. Gel filtração	5	5	5	5
15. Filtração estéril	$\frac{1 B_i}{G_{jd} R_{jd}}$	$\frac{1 B_i}{G_{jd} R_{jd}}$	$\frac{1 B_i}{G_{jd} R_{jd}}$	$\frac{1 B_i}{G_{jd} R_{jd}}$

Tabela E.3 (continuação) Funções para o tempo de processamento em cada estágio (h).

Estágio, operação	Produto/hospedeiro			
	tPA/ <i>E. Coli</i>	tPA/CHO	SOD/ <i>Sacch.</i>	SOD/ <i>E. Coli</i>
1. Fermentação	24	168	24	24
2.A Centrifugação	$\frac{5B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{0,05B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{0,002 B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{0,125 B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$
2.B Microfiltração	$\frac{4B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{44B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{0,15 B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{0,2B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$
3.A Homogeneização	$\frac{1,5B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	-	$\frac{0,04 B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{0,075B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$
3.B Moinho de bolas	$\frac{1,5B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	-	$\frac{0,01 B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{0,075B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$
4.A Centrifugação	$\frac{25B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	-	$\frac{0,5B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{1,25B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$
4.B Microfiltração	$\frac{6,25B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	-	$\frac{0,2B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$	$\frac{0,3125B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$
5. IB solubilização	12	-	-	10
6. Diafiltração	$\frac{70B_i}{G_{jd} R_{jd} \prod_{i=5}^{15} h_{ihj}}$	-	-	$\frac{35B_i}{G_{jd} R_{jd} \prod_{i=5}^{15} h_{ihj}}$
7. Sulfonação	18	-	-	15
8. Refolding	36	-	-	24
9. Ultrafiltração	de 2.A $\frac{1000B_i}{G_{jd} R_{jd} \prod_{i=8}^{15} h_{ihj}}$	$\frac{500B_i}{G_{jd} R_{jd} \prod_{i=3}^{15} h_{ihj}}$ de 2.B $\frac{700B_i}{G_{jd} R_{jd} \prod_{i=2}^{15} h_{ihj}}$		$\frac{100B_i}{G_{jd} R_{jd} \prod_{i=8}^{15} h_{ihj}}$
10. Cromatografia	$3,5 + \frac{0,6B_i}{G_{jd} R_{jd} \prod_{i=10}^{15} h_{ihj}}$	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{i=10}^{15} h_{ihj}}$	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{i=10}^{15} h_{ihj}}$	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{i=10}^{15} h_{ihj}}$
11. Ultra-diafiltração	$\frac{5B_i}{G_{jd} R_{jd} \prod_{i=11}^{15} h_{ihj}}$	$\frac{100B_i}{G_{jd} R_{jd} \prod_{i=11}^{15} h_{ihj}}$	$\frac{100B_i}{G_{jd} R_{jd} \prod_{i=11}^{15} h_{ihj}}$	$\frac{100B_i}{G_{jd} R_{jd} \prod_{i=11}^{15} h_{ihj}}$
12. Cromatografia	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{i=12}^{15} h_{ihj}}$	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{i=12}^{15} h_{ihj}}$	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{i=12}^{15} h_{ihj}}$	$15 + \frac{0,35 B_i}{G_{jd} R_{jd} \prod_{i=12}^{15} h_{ihj}}$
13. Ultra-diafiltração	$\frac{100B_i}{G_{jd} R_{jd} \prod_{i=13}^{15} h_{ihj}}$	$\frac{100B_i}{G_{jd} R_{jd} \prod_{i=13}^{15} h_{ihj}}$	$\frac{100B_i}{G_{jd} R_{jd} \prod_{i=13}^{15} h_{ihj}}$	$\frac{100B_i}{G_{jd} R_{jd} \prod_{i=13}^{15} h_{ihj}}$
14. Gel filtração	5	5	5	5
15. Filtração estéril	$\frac{1B_i}{G_{jd} R_{jd}}$	$\frac{1B_i}{G_{jd} R_{jd}}$	$\frac{1B_i}{G_{jd} R_{jd}}$	$\frac{1B_i}{G_{jd} R_{jd}}$

Os custos para os equipamentos, em cada uma das operações são apresentados na Tabela E.4, e as dimensões limites são mostradas na Tabela E.5.

Tabela E.4 Custo dos equipamentos.

Estágio, operação		Equipamento/função de custo	
1. Fermentação		fermentador $63400 V^{0.6}$	
2.A Centrifugação	tanque de alimentação $5750 V^{0.6}$	centrífuga $28600 R^{0.7}$	tanque de produto $5750 V^{0.6}$
2.B Microfiltração	tanque de alimentação $5750 V^{0.6}$	microfiltro $2900 R^{0.85}$	tanque de filtrado $5750 V^{0.6}$
3.A Homogeneização	tanque de <i>holding</i> $5750 V^{0.6}$	homogeneizador $12100 R^{0.75}$	
3.B Moinho de bolas	tanque de <i>holding</i> $5750 V^{0.6}$	moinho de bolas $27630 R^{0.5}$	
4.A Centrifugação	tanque de alimentação $5750 V^{0.6}$	centrífuga $28600 R^{0.7}$	tanque de produto $5750 V^{0.6}$
4.B Microfiltração	tanque de alimentação $5750 V^{0.6}$	microfiltro $2900 R^{0.85}$	tanque de filtrado $5750 V^{0.6}$
5. IB solubilização		reator de solubilização $31000 V^{0.5}$	
6. Diafiltração	tanque de <i>holding</i> $5750 V^{0.6}$	diafiltro $2900 R^{0.85}$	
7. Sulfonação		reator de sulfonação $31000 V^{0.5}$	
8. Refolding		reator de refolding $31000 V^{0.5}$	
9. Ultrafiltração	tanque de <i>holding</i> $5750 V^{0.6}$	ultrafiltro $2900 R^{0.85}$	
10. Cromatografia	tanque de <i>holding</i> $5750 V^{0.6}$	coluna cromatográfica $310\,000 A^{0.55}$	tanque de produto $5750 V^{0.6}$
11. Ultra-diafiltração	tanque de <i>holding</i> $5750 V^{0.6}$	ultrafiltro $2900 R^{0.85}$	
12. Cromatografia	tanque de alimentação $5750 V^{0.6}$	coluna cromatográfica $310\,000 A^{0.55}$	tanque de produto $5750 V^{0.6}$
13. Ultra-diafiltração	tanque de <i>holding</i> $5750 V^{0.6}$	ultrafiltro $2900 R^{0.85}$	
14. Gel filtração	tanque de alimentação $5750 V^{0.6}$	coluna $310\,000 A^{0.55}$	tanque de produto $5750 V^{0.6}$
15. Filtração estéril	tanque de alimentação $5750 V^{0.6}$	filtro estéril $2900 R^{0.85}$	tanque de filtrado $5750 V^{0.6}$

Tabela E.5 Limites na dimensão dos equipamentos.

Equipamento/dimensão	Limite inferior	Limite superior
Fermentador/volume	0,2 m ³	100 m ³
Filtros/área	0,1 m ²	50 m ²
Homogeneizador/taxa de processamento	0,1 m ³ /h	20 m ³ /h
Moinho de bolas/taxa de processamento	0,05 m ³ /h	10 m ³ /h
Centrífuga/potência	0,1 kW	20 kW
Reatores/volume	0,2 m ³	100 m ³
Colunas cromatográficas/área	0,0001 m ²	0,75 m ²
Tanques/volume	0,2 m ³	100 m ³

APÊNDICE F PROJETOS PARA O MODELO MONOPRODUTO

Neste apêndice são apresentados os projetos ótimos e sub-ótimos (com hospedeiro fixo) obtidos com o modelo de plantas monoproduto. Nas Tabelas F.1 a F.4, as unidades das dimensões dos equipamentos são m^3 para unidades descontínuas (tanques e reatores), m^2 para filtros e colunas cromatográficas, m^3/h para homogeneizador e moinho de bolas e kW para centrífugas.

Tabela F.1 Projetos para modelo monoproduto para produção de insulina (INS).

Estágio	unidades em-fase / fora-de-fase / dimensão					
	projeto ótimo (<i>S. cerevisiae</i>)			projeto sub-ótimo (<i>E. coli</i>)		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Tanque de alimentação	Unidade de processamento	Tanque de produto
1		1/ 5/ 0,200			1/ 5/ 0,230	
2B	1/ 1/ 0,200	1/ 1/ 0,272	1/ 1/ 0,272	1/ 1/ 0,230	1/ 1/ 0,184	
3B				1/ 1/ 0,200	1/ 1/ 0,069	
4B				1/ 1/ 0,200	1/ 1/ 0,287	
5					1/ 2/ 0,612	
6				1/ 1/ 0,612	1/ 1/ 8,573	
7					1/ 3/ 0,735	
8					1/ 3/ 3,858	
9				1/ 1/ 3,858	1/ 1/ 38,580	
10	1/ 1/ 0,272	2/ 1/ 0,579	1/ 1/ 0,200	1/ 1/ 0,926	2/ 1/ 0,579	1/ 1/ 0,200
11	1/ 1/ 0,200	1/ 1/ 1,736		1/ 1/ 0,200	1/ 1/ 1,736	
12	1/ 2/ 0,200	2/ 2/ 0,608	1/ 2/ 1,563	1/ 2/ 0,200	2/ 2/ 0,608	1/ 2/ 1,563
13	1/ 1/ 1,563	1/ 1/ 15,625		1/ 1/ 1,563	1/ 1/ 15,625	
14	1/ 1/ 0,200	1/ 1/ 0,625	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,625	1/ 1/ 0,200
15	1/ 1/ 0,200	1/ 1/ 0,250	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,250	1/ 1/ 0,200

Tabela F.2 Projetos para modelo monoproduto para produção de vacina (HBV).

Estágio	unidades em-fase / fora-de-fase / dimensão					
	projeto ótimo (<i>S. cervisiae</i>)			projeto sub-ótimo (CHO)		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Tanque de alimentação	Unidade de processamento	Tanque de produto
1		1/ 4/ 0,309			1/ 10/ 2,859	
2B	1/ 1/ 0,309	1/ 1/ 0,150		1/ 1/ 2,859	1/ 1/ 0,749	1/ 1/ 4,003
3B	1/ 1/ 0,200	1/ 1/ 0,050				
4B	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,200			
9				1/ 1/ 4,003	1/ 1/ 11,914	
10	1/ 1/ 0,200	2/ 1/ 0,591	1/ 1/ 0,200	1/ 1/ 2,059	6/ 1/ 0,686	1/ 1/ 8,235
11	1/ 1/ 0,200	1/ 1/ 1,225		1/ 1/ 8,235	1/ 1/ 24,510	
12	1/ 2/ 0,200	3/ 2/ 0,504	1/ 2/ 3,216	1/ 1/ 0,206	5/ 1/ 0,659	1/ 1/ 7,000
13	1/ 1/ 3,216	1/ 1/ 20,833		1/ 1/ 7,000	1/ 1/ 20,833	
14	1/ 1/ 0,200	1/ 1/ 0,643	1/ 1/ 0,200	1/ 1/ 0,200	2/ 1/ 0,700	1/ 1/ 0,280
15	1/ 1/ 0,200	1/ 1/ 0,167	1/ 1/ 0,200	1/ 1/ 0,280	1/ 1/ 0,167	1/ 1/ 0,280

Tabela F.3 Projetos para modelo monoproduto para produção de tPA.

Estágio	unidades em-fase / fora-de-fase / dimensão					
	projeto ótimo (<i>E. coli</i>)			projeto sub-ótimo (CHO)		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Tanque de alimentação	Unidade de processamento	Tanque de produto
1		1/ 2/ 0,200			1/ 5/ 0,960	
2B	1/ 1/ 0,200	1/ 1/ 0,100		1/ 1/ 0,960	1/ 1/ 0,126	1/ 1/ 1,344
3A	1/ 1/ 0,200	1/ 1/ 0,100				
4B	1/ 1/ 0,200	1/ 1/ 0,100				
5		1/ 1/ 0,200				
6	1/ 1/ 1,144	1/ 1/ 0,476				
7		1/ 1/ 0,200				
8		1/ 1/ 1,543				
9	1/ 1/ 2,883	1/ 1/ 4,287		1/ 1/ 1,344	1/ 1/ 2,000	
10	1/ 1/ 0,346	1/ 1/ 0,023	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,069	1/ 1/ 0,200
11	1/ 1/ 1,471	1/ 1/ 0,100		1/ 1/ 0,200	1/ 1/ 0,231	
12	1/ 1/ 0,200	1/ 1/ 0,033	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,062	1/ 1/ 0,200
13	1/ 1/ 1,250	1/ 1/ 0,208		1/ 1/ 0,200	1/ 1/ 0,208	
14	1/ 1/ 0,200	1/ 1/ 0,015	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,028	1/ 1/ 0,200
15	1/ 1/ 0,200	1/ 1/ 0,100	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,100	1/ 1/ 0,200

Tabela F.4 Projetos para modelo monoproduto para produção de SOD.

Estágio	unidades em-fase / fora-de-fase / dimensão					
	projeto ótimo (<i>S. cerevisiae</i>)			projeto sub-ótimo (<i>E. coli</i>)		
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Tanque de alimentação	Unidade de processamento	Tanque de produto
1		1/ 2/ 0,200			1/ 2/ 0,200	
2B	1/ 1/ 0,200	1/ 1/ 0,100		1/ 1/ 0,200	1/ 1/ 0,100	
3A	1/ 1/ 0,200	1/ 1/ 0,100		1/ 1/ 0,200	1/ 1/ 0,100	
4B	1/ 1/ 0,200	1/ 1/ 0,100	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,100	
5					1/ 1/ 1,144	
6				1/ 1/ 1,144	1/ 1/ 5,340	
7					1/ 1/ 1,373	
8					1/ 2/ 2,883	
9				1/ 1/ 2,884	1/ 1/ 9,611	
10	1/ 1/ 0,200	1/ 1/ 0,712	1/ 1/ 1,513	1/ 1/ 0,346	1/ 1/ 0,692	1/ 1/ 1,471
11	1/ 1/ 1,513	1/ 1/ 4,902		1/ 1/ 1,471	1/ 1/ 4,902	
12	1/ 1/ 0,200	1/ 1/ 0,605	1/ 1/ 1,286	1/ 1/ 0,200	1/ 1/ 0,588	1/ 1/ 1,250
13	1/ 1/ 1,286	1/ 1/ 4,167		1/ 1/ 1,250	1/ 1/ 4,167	
14	1/ 1/ 0,200	1/ 1/ 0,257	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,250	1/ 1/ 0,200
15	1/ 1/ 0,200	1/ 1/ 0,100	1/ 1/ 0,200	1/ 1/ 0,200	1/ 1/ 0,100	1/ 1/ 0,200

APÊNDICE G RESULTADOS COMPLEMENTARES PARA MODELOS MULTIPRODUTO

Este apêndice apresenta dados complementares dos resultados obtidos nos capítulos 5 e 6. Nas Tabelas G.1 e G.2 tem-se os resultados para a planta multiproduto com alocação de tanques intermediários em posições definidas considerando uma relação entre os tamanhos de batelada para os sub-processos $F = 5$. A Tabela G.3 apresenta os valores obtidos para as variáveis de dimensionamento dos equipamentos na planta multiproduto com tanques intermediários em posições otimizadas (cenário 4).

A Tabela G.4 apresenta o projeto e os custos para cenário 5, que considera o modelo multiproduto com flexibilização do *scheduling* (desvinculação do tempo de processamento para os estágios 5 a 9), e as G.5 e G.6 apresentam o projeto, os custos e as variáveis de dimensionamento para o cenário 6, que adiciona a alocação de tanques intermediários em posições otimizadas. Os hospedeiros escolhidos para ambos cenários foram os mesmos dos cenários 2, 3 e 4, o que permite a comparação do projeto.

A Tabela G.7 apresenta o projeto e os custos para o cenário 2 fixando a expressão de tPA por CHO, estes resultados servem como parâmetro de comparação para o resultados dos cenários 7 e 8. Finalmente, a Tabela G.8 apresenta as variáveis de dimensionamento para o cenário 8. As dimensões dos equipamentos estão em unidades de m^3 para equipamentos descontínuos (tanques e reatores), m^2 para filtros e colunas cromatográficas, m^3/h para homogeneizador e moinho de bolas e kW para centrífugas.

Tabela G.1 Resultados gerais por produto para o cenário 3 considerando $F = 5$.

Produto	Hospedeiro	B_{ij} (kg)			TL_{ij} (h)		
		$j \leq 2$	$3 \leq j \leq 9$	$10 \leq j$	$j \leq 2$	$3 \leq j \leq 9$	$10 \leq j$
INS	<i>S. cerevisiae</i> extracel.	18,51	--	3,86	24,00	--	5,00
HBV	<i>S. cerevisiae</i> intracel.	9,91	1,98	3,19	24,00	4,80	7,72
tPA	<i>E. coli</i>	0,49	0,73	0,15	24,00	36,00	7,51
SOD	<i>S. cerevisiae</i> intracel.	4,21	1,00	2,71	24,00	5,70	15,44

Tabela G.2 Projeto e custos para o cenário 3 considerando $F = 5$.

Estágio	unidades em-fase / fora-de-fase / dimensão			Custo total dos estágios (US\$)	
	Tanque de alimentação	Unidade de processamento	Tanque de produto	$F = 5$	$F = 2$
1		1/ 1/ 2,477		109.260	181.971
2B	1/ 1/ 2,477	1/ 1/ 0,840	1/ 1/ 4,033	25.686	14.870
	Tanque-pulmão 2	4,082		13.693	9.034
3B	1/ 1/ 1,858	1/ 1/ 0,155		19.210	16.373
4B	1/ 1/ 1,858	1/ 1/ 0,645	1/ 1/ 0,238	12.767	9.688
5		1/ 1/ 2,973		53.448	37.793
6	1/ 1/ 2,973	1/ 1/ 5,780		23.939	20.177
7		1/ 1/ 3,567		58.549	41.400
8		1/ 1/ 37,455		189.721	268.306
9	1/ 1/ 37,455	2/ 1/ 26,010		143.090	125.888
	Tanque-pulmão 9	5,714		16.362	16.362
10	1/ 1/ 0,840	5/ 1/ 0,750	1/ 1/ 7,969	1.348.325	1.348.325
11	1/ 1/ 7,969	1/ 1/ 25,810		65.940	65.940
12	1/ 2/ 0,268	5/ 2/ 0,750	1/ 2/ 7,969	2.691.506	2.691.506
13	1/ 1/ 7,969	2/ 1/ 25,810		111.903	111.903
14	1/ 1/ 0,241	3/ 1/ 0,643	1/ 1/ 0,386	735.063	735.063
15	1/ 1/ 0,386	1/ 1/ 0,771	1/ 1/ 0,386	8.820	8.820
Custo total				5.627.280	5.703.418

Tabela G.3 Variáveis de dimensionamento para o cenário 4.

Estágio	INS		HBV		tPA		SOD	
	B_{ij} (kg)	TL_{ij} (h)	B_{ij} (kg)	TL_{ij} (h)	B_{ij} (kg)	TL_{ij} (h)	B_{ij} (kg)	TL_{ij} (h)
1	10,19	12,00	4,75	12,00	0,22	12,31	4,49	24,00
Tanque-pulmão 1								
2B	5,09	6,00	3,33	8,41	0,33	17,99	2,24	12,98
Tanque-pulmão 2								
3B			1,66	4,20	0,22	12,00	2,67	15,44
4B			1,66	4,20	0,22	12,00	2,67	15,44
5					0,22	12,00		
Tanque-pulmão 5								
6					0,33	17,99		
Tanque-pulmão 6								
7					0,65	36,00		
8					0,65	36,00		
9					0,65	36,00		
10	3,99	4,70	1,66	4,20	0,65	36,00	2,67	15,44
11	3,99	4,70	1,66	4,20	0,65	36,00	2,67	15,44
Tanque-pulmão 11								
12	3,70	4,36	3,06	7,72	0,33	17,99	1,33	7,72
13	3,70	4,36	3,06	7,72	0,33	17,99	1,33	7,72
Tanque-pulmão 13								
14	4,24	5,00	1,98	5,00	0,16	9,00	1,00	5,78
15	4,24	5,00	1,98	5,00	0,16	9,00	1,00	5,78

Tabela G.4 Projeto e custos para o cenário 5.

Estágio	unidades em-fase / fora-de-fase / dimensão			Custo total dos estágios (US\$)	
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 5	Cenário 2 (tPA- <i>E. coli</i>)
1		1/ 5/ 0,845		286.645	303.285
2B	1/ 1/ 0,845	1/ 1/ 0,840	1/ 1/ 0,840	12.880	13.182
3B	1/ 1/ 0,423	1/ 1/ 0,169		14.784	14.502
4B	1/ 1/ 0,423	1/ 1/ 0,704	1/ 1/ 0,383	8.814	8.861
5		1/ 1/ 0,676		25.496	53.448
6	1/ 1/ 0,676	1/ 1/ 1,315		8.209	17.695
7		1/ 1/ 0,812		27.930	58.549
8		1/ 1/ 8,523		90.503	379.441
9	1/ 1/ 8,523	1/ 1/ 11,838		44.494	114.540
10	1/ 1/ 0,840	5/ 1/ 0,724	1/ 1/ 7,691	1.322.353	1.348.325
11	1/ 1/ 7,691	1/ 1/ 24,911		64.154	65.940
12	1/ 2/ 0,268	5/ 2/ 0,75	1/ 2/ 7,969	2.691.506	2.691.506
13	1/ 1/ 7,969	2/ 1/ 25,810		111.903	111.903
14	1/ 1/ 0,241	3/ 1/ 0,643	1/ 1/ 0,386	735.063	735.063
15	1/ 1/ 0,387	1/ 1/ 0,771	1/ 1/ 0,386	8.820	8.820
Custo total				5.453.555	5.925.056

Tabela G.5 Projeto e custos para o cenário 6.

Estágio	unidades em-fase / fora-de-fase / dimensão			Custo total dos estágios (US\$)	
	Tanque de alimentação	Unidade de processamento	Tanque de produto	Cenário 6	Cenário 5
1		1/ 3/ 1,042		194.938	286.645
	Tanque-pulmão 1	8,856		21.282	
2B	1/ 1/ 1,042	1/ 1/ 0,714	1/ 1/ 0,572	12.183	12.880
3B	1/ 1/ 0,521	1/ 1/ 0,195		16.100	14.784
4B	1/ 1/ 0,521	1/ 1/ 0,814	1/ 1/ 0,260	8.888	8.814
5		1/ 1/ 0,833		28.301	25.496
6	1/ 1/ 0,833	1/ 1/ 1,621		9.527	8.209
7		1/ 1/ 1,000		31.003	27.930
8		1/ 1/ 10,502		100.461	90.503
9	1/ 1/ 10,502	1/ 1/ 14,589		51.876	44.494
10	1/ 1/ 0,572	5/ 1/ 0,579	1/ 1/ 6,153	1.168.952	1.322.353
11	1/ 1/ 6,153	1/ 1/ 19,928		53.998	64.154
	Tanque-pulmão 11	0,455		3.587	
12	1/ 2/ 0,228	5/ 2/ 0,638	1/ 2/ 6,775	2.461.282	2.691.506
13	1/ 1/ 6,775	2/ 1/ 33,873		133.949	111.903
14	1/ 1/ 0,205	3/ 1/ 0,547	1/ 1/ 0,328	672.236	735.063
15	1/ 1/ 0,328	1/ 1/ 0,656	1/ 1/ 0,328	7.916	8.820
Custo total				4.976.480	5.453.555

Tabela G.6 Variáveis de dimensionamento para o cenário 6.

Estágio	INS		HBV		tPA		SOD	
	B_{ij} (kg)	TL_{ij} (h)	B_{ij} (kg)	TL_{ij} (h)	B_{ij} (kg)	TL_{ij} (h)	B_{ij} (kg)	TL_{ij} (h)
1	5,25	8,00	4,34	8,00	0,20	8,00	1,22	8,00
Tanque-pulmão 1								
2B	2,62	4,00	2,17	4,00	0,20	8,00	2,09	15,44
3B			2,17	4,00	0,20	8,00	2,09	15,44
4B			2,17	4,00	0,20	8,00	2,09	15,44
5					0,20	12,00		
6					0,20	36,00		
7					0,20	18,00		
8					0,20	36,00		
9					0,20	36,00		
10	2,62	4,00	2,17	4,00	0,20	8,00	2,09	15,44
11	2,62	4,00	2,17	4,00	0,20	8,00	2,09	15,44
Tanque-pulmão 11								
12	3,28	5,00	2,71	5,00	0,20	8,00	1,18	8,72
13	3,28	5,00	2,71	5,00	0,20	8,00	1,18	8,72
14	3,28	5,00	2,71	5,00	0,20	8,00	1,18	8,78
15	3,28	5,00	2,71	5,00	0,20	8,00	1,18	8,78

Tabela G.7 Projeto e custos para o cenário 2 considerando expressão de tPA por CHO.

Estágio	unidades em-fase / fora-de-fase / dimensão			Custo total dos estágios
	Tanque de alimentação	Unidade de processamento	Tanque de produto	
1		1/ 5/ 11,653		1.383.286
2B	1/ 1/ 11,653	1/ 1/ 1,526	1/ 1/ 16,314	59.948
3A	1/ 1/ 0,217	1/ 1/ 0,112		4.649
4B	1/ 1/ 0,217	1/ 1/ 0,562	1/ 1/ 0,434	7.562
9	1/ 1/ 16,314	1/ 1/ 26,276		74.336
10	1/ 1/ 0,840	5/ 1/ 0,750	1/ 1/ 7,969	1.348.325
11	1/ 1/ 7,969	1/ 1/ 25,810		65.939
12	1/ 2/ 0,268	5/ 2/ 0,750	1/ 2/ 7,969	2.691.506
13	1/ 1/ 7,969	2/ 1/ 25,810		111.903
14	1/ 1/ 0,241	3/ 1/ 0,643	1/ 1/ 0,386	735.063
15	1/ 1/ 0,386	1/ 1/ 0,771	1/ 1/ 0,386	8.820
Custo total				6.491.336

Tabela G.8 Variáveis de dimensionamento para o cenário 8.

Estágio	INS		HBV		tPA		SOD	
	B_{ij} (kg)	TL_{ij} (h)	B_{ij} (kg)	TL_{ij} (h)	B_{ij} (kg)	TL_{ij} (h)	B_{ij} (kg)	TL_{ij} (h)
1	10,93	12,00	4,96	12,00	0,41	15,07	3,59	12,00
Tanque-pulmão 1								
2B	5,53	7,17	3,34	8,09	0,41	15,07	1,80	11,41
Tanque-pulmão 2								
3B			1,67	4,05			2,43	15,44
4B			1,67	4,05			2,43	15,44
9					0,41	15,07		
10	3,62	4,70	1,67	4,05	0,41	15,07	2,43	15,44
Tanque-pulmão 10								
11	3,86	5,00	3,19	7,72	0,41	15,07	2,43	7,72
12	3,86	5,00	3,19	7,72	0,41	15,07	2,43	7,72
13	3,86	5,00	3,19	7,72	0,41	15,07	1,22	7,72
14	3,86	5,00	3,19	7,72	0,41	15,07	1,22	7,72
15	3,86	5,00	3,19	7,72	0,41	15,07	1,22	7,72

REFERÊNCIAS BIBLIOGRÁFICAS.

- Ahl, I. M.; Lindberg, M. e Tibell, L. (2004) Coexpresion of yeast cooper chaperon (yCCS) and CuZn-superoxide dismutases in *Escherichia coli* yields protein with high copper contents. *Protein Expression and Purification* **37**, pp. 311-319.
- Alle, A. e Pinto, J. M. (2004) Global optimization for the cyclic scheduling and operation of multistage continuous plants. *Industrial & Engineering Chemistry Research* **43**, pp. 1485-1498.
- Asenjo, J. A.; Montagna, J. M.; Vecchiotti, A. R.; Iribarren, O. A. e Pinto, J. M. (2000) Strategies for the simultaneous optimization of the structure and the process variables of a protein production plant. *Computers and Chemical Engineering* **24**, pp. 2277-2290.
- Allgor, R. J. e Barton, P. I. (1997) Mixed-integer dynamic optimization. *Computers and Chemical Engineering* **21** (suppl.), pp. S451-S456.
- Bogle, I. D. L.; Conckshott, A. R.; Bulmer, M.; Thornhill, N.; Gregory, M. e Dehghani, M. (1996) A process systems engineering view of biochemical process operations. *Computers and Chemical Engineering* **20**, pp. 943-949.
- Brooke, A.; Kendrick, D.; Meeraus, A.; Raman, R. e Rosenthal, R. E. (1998) *GAMS a user's guide*. Washington: GAMS Development Corporation.
- Cheung, K. e Hui, C. (2001) Heat exchanger network optimization with discontinuous exchanger cost function. *Applied Thermal Engineering* **21**, pp. 1397-1405.
- Christensen, B. e Nielsen, J. (2000) Metabolic network analysis of *Penicillium chrysogenum* using ^{13}C -labeled glucose. *Biotechnology and Bioengineering* **68** (6), pp. 652-659.
- Christensen, B.; Gombert, A. K. e Nielsen, J. (2002) Analysis of flux estimates based on ^{13}C -labeling experiments. *European Journal of Biochemistry* **269**, pp. 2795-2800.
- Christiansen, T; Christensen, B. e Nielsen, J. (2002) Metabolic network analysis of *Bacillus clausii* on minimal and semirich medium using ^{13}C -labeled glucose. *Metabolic Engineering* **4**, pp. 159-169.
- Cohen, S. N.; Chang, A. C. W.; Boyer, H. e Helling, R. (1973) Construction of biologically functional bacterial plasmids *in vitro*. *Proceedings of the National Academy of Sciences of USA* **70**, pp. 3240-3244.
- Drud, A. (1996) CONOPT: A system for large scale nonlinear optimization. Reference manual. Bagsvaerd: ARKI Consulting and Development A/S.

- Esposito, W. E. e Floudas, C. A. (1998) Parameter estimation in nonlinear algebraic models via global optimization. *Computers and Chemical Engineering* **22** (suppl.), pp. S213-S220.
- Fischer, E.; Zamboni, N. e Sauer, U. (2004) High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived ^{13}C constraints. *Analytical Biochemistry* **325**, pp. 308-316.
- Floudas, C. A. e Visweswaran, V. (1990) A global optimization algorithm (GOP) for certain classes of nonconvex NLPs – I. Theory. *Computers and Chemical Engineering* **14** (12), pp. 1397-1417.
- Foo, C. Y.; Manan, Z. A.; Yunus, R. M. e Aziz, R. A. (2004) Synthesis of mass exchange network for batch processes-Part I: Utility targeting. *Chemical Engineering Science* **59**, pp. 1009-1026.
- Galan, S. e Barton, P. I. (1998) Dynamic optimization of hybrid systems. *Computers and Chemical Engineering* **22** suppl., pp. S183-S190.
- Gombert, A. K.; dos Santos, M. M.; Christensen, B. e Nielsen, J. (2001) Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *Journal of Bacteriology* **183** (4), pp. 1441-1451.
- Groep, M. E.; Gregory, M. E.; Kershenbaum, L. S. e Bogle I. D. L. (2000) Performance modeling and simulation of biochemical process sequences with interacting unit operations. *Biotechnology and Bioengineering* **67** (3), pp. 301-310.
- Grossmann, I. E. (1996) *Nonconvex Optimization and Its Applications vol. 9. Global Optimization in Engineering Design*. Netherlands: Kluwer Academic Publishers.
- Hanson, A. D. e Rhodes, D. (1983) ^{14}C tracer evidence for synthesis of choline and betaine via phosphoryl base intermediates in salinized sugarbeet leaves. *Plant Physiol.* **71**, pp. 692-700
- Hardy, E.; Martínez, E.; Diago, D.; Díaz, R.; González, D. e Herrera, L. (2000) Large-scale production of recombinant hepatitis B surface antigen from *Pichia pastoris*. *Journal of Biotechnology* **77**, pp. 157-167.
- Harris, E. L. V. e Angal, S. (1994) *Protein purification methods: a practical approach*. England: IRL Press.
- Harris, E. L. V. e Angal, S. (1995) *Protein purification applications: a practical approach*. England: IRL Press.
- He, H. J.; Yuan, Q. S.; Yang, G. Z. e Wu, X. F. (2002) High-level expression of human extracellular superoxide dismutase in *Escherichia coli* and insect cells. *Protein Expression and Purification* **24**, pp. 13-17.

- Hostrup, M; Harper, P. M. e Gani, R. (1999) Design of environmentally benign processes: integration of solvent design and separation process synthesis. *Computers and Chemical Engineering* **23**, pp. 1395-1414.
- IBM (1991) *IBM-OSL guide and reference* (release 2). Kingston.
- Janson, J. C. e Rydén, L. (1989) *Protein purification: principles, high resolution methods and applications*. New York: VCH Publishers Inc.
- Jeffrey, F. M. H; Roach, J. S.; Storey, C. J.; Sherry, A. D. e Malloy, C. R. (2002) 13C isotopomer analysis of glutamate by Tandem Mass Spectrometry. *Analytical Biochemistry* **300**, pp. 192-205.
- Johnson, I. S. (1983) Human insulin from recombinant DNA technology. *Science* **219**, pp. 632-637.
- Katz, J. (1985) Determination of gluconeogenesis in vivo with ^{14}C -labeled substrates *Am. J. Physiol.-Regulatory, Integrative and Comparative Physiology* **248**, pp. R391-R399.
- Katz, J. e Wood, H. G.(1963) The use of C^{14}O_2 yields from glucose-1- and -6- C^{14} for the evaluation of the pathways of glucose metabolism. *Journal of Biological Chemistry* **238**, pp. 517-523.
- Kjeldsen, T.; Pettersson, A. F. e Hach, M. (1999) The role of leaders in intracellular transport and secretion of the insulin precursor in the yeast *Saccharomyces cerevisiae*. *Journal of Biotechnology* **75**, pp. 195-208.
- Kravanja, Z. e Grossmann, I. E. (1997) Multilevel-hierarchical MINLP synthesis of process flowsheets. *Computers and Chemical Engineering* **21** (suppl.), pp. S421-S426.
- Lee, W. N. P.; Bergner, E. A. e Guo, Z. K. (1992) Mass isotopomer pattern and precursor-product relationship. *Biological Mass Spectrometry* **21**, pp. 114-122.
- Lee, W. N. P.; Byerley, L. O.; Bergner, E. A. e Edmond, J. (1991) Mass isotopomer analysis: theoretical and practical considerations. *Biological Mass Spectrometry* **20**, pp. 451-458.
- Lee, S. e Grossmann, I. E. (2000) New algorithms for nonlinear generalized disjunctive programming. *Computers and Chemical Engineering* **24**, pp. 2125-2142.
- Liljeqvist, S. e Ståhl, S. (1999) Production of recombinant subunit vaccine: protein immunogens, live delivery systems and nucleic acid vaccines: Review. *J. Biotechnology* **73**, pp. 1-33.
- Lin, B. e Miller, D. C. (2004) Tabu search algorithm for chemical process optimization. *Computers and Chemical Engineering* **28**, pp. 2287-2306.

- Lin, X. e Floudas, C. A. (2001) Design, synthesis and scheduling of multipurpose batch plants via an effective continuous-time formulation. *Computers and Chemical Engineering* **25**, pp. 665-674.
- Malloy C. R.; Sherry A. D. e Jeffrey F. M. H. (1987) Carbon flux through citric acid cycle pathways in perfused heart by ^{13}C NMR spectroscopy. *FEBS Letters* **212**, pp. 58-62.
- Marx, A.; de Graaf, A. A.; Wiechert W.; Eggeling, L. e Sahm, H. (1996) Determination of the fluxes in central metabolism of *Corynebacterium glutamicum* by NMR spectroscopy combined with metabolite balancing. *Biotechnology and Bioengineering* **49**, pp. 111-129.
- McCormick, G. P. (1976) Computability of global solutions to factorable nonconvex programs: Part I- convex underestimations problems. *Mathematical Programming* **10**, pp. 147-175.
- Misawa, S. e Kumagai, I. (1999) Refolding of therapeutic proteins produced in *Escherichia coli* as inclusion bodies. *Biopolymers (Peptide Science)* **51**, pp. 297-307.
- Montagna, J. M.; Iribarren O. A. e Galiano, F. C. (1994) The design of multiproduct batch plants with process performance models. *Trans IcemE*, **72** (A), pp. 783-791.
- Montagna, J. M.; Iribarren, O. A. e Vecchiotti, A. R. (2004) Synthesis of biotechnological processes using generalized disjunctive programming. *Industrial & Engineering Chemistry Research* **43**, pp. 4220-4232.
- Nielsen, J. (2001) Metabolic engineering. *Applied Microbiology Biotechnology* **55**, pp. 263-283.
- Novak, Z.; Kravanja, Z. e Grossmann, I. E. (1996) Simultaneous synthesis of distillation sequences in overall process schemes using an improved MINLP approach. *Computers and Chemical Engineering* **20**, pp. 1425-1440.
- Pahor, B. e Kravanja, Z. (1995) Simultaneous solution and MINLP synthesis of DAE process problems: PFR networks in overall processes. *Computers and Chemical Engineering* **19** (suppl.), pp. S181-S188.
- Pinto, J. M.; Montagna, J. M.; Vecchiotti, A. R.; Iribarren, O. A. e Asenjo, J. A. (2001) Process performance models in the optimization of multiproduct protein production plants. *Biotechnology and Bioengineering*, **74** (6), pp. 451-465.
- Quesada, I. e Grossmann, I. E. (1995) Global optimization of bilinear process networks with multicomponent flows. *Computers and Chemical Engineering* **19**, pp. 1219-42.
- Ratschek, H. e Rokne, J. (1991) Interval tools for global optimization. *Computers & Mathematics with Applications* **21** (6-7), pp.41-50.

- Ravemark, D. E. e Rippin, D. W. (1998) Optimal design of a multi-product batch plant. *Computers and Chemical Engineering* **22**, pp. 177-183.
- Riascos, C. A. M e Pinto, J. M. (2004) Optimal control of bioreactors: a simultaneous approach for complex systems. *Chemical Engineering Journal* **99**, pp. 23-34.
- Rouf, S. A.; Moo-Young, M. e Chisti, Y. (1996) Tissue-type plasminogen activator: characteristics, applications and production technology. *Biotechnology Advances* **14** (3), pp. 239-266.
- Sahinidis, N. V. e Tawarmalani, M. (2000) Applications of global optimization to process and molecular design. *Computers and Chemical Engineering* **24**, 2157-69.
- Salomone, H. E.; Montagna, J. M. e Iribarren, O. A. (1994) Dynamic simulations in the design of batch processes. *Computers and Chemical Engineering* **18**, pp. 191-204.
- Samsatli N. J. e Shah N. (1996) Optimal integrated design of biochemical processes. *Computers and Chemical Engineering* **20** (suppl.), pp. S315-S320.
- Sauer, U. e Bailey, J. E. (1999) Estimation of P-to-O ratio in *Bacillus subtilis* and its influence on maximum riboflavin yield. *Biotechnology and Bioengineering* **64**, pp. 750-754.
- Sauer, U.; Lasko, D. R.; Fiaux, J.; Hochuli, M.; Glaser, R.; Szyperski, T.; Wüthrich, K. e Bailey, J. E. (1999) Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism. *Journal of Bacteriology* **181**, pp. 6679-6688.
- Schmidell, W.; Lima, U. A.; Arquarone, E. e Borzani, W. (2001) *Biotechnologia Industrial: Vol. 2 engenharia bioquímica*. Brasil: Edgard Blücher Ltda.
- Schmidt, K.; Carlsen, M.; Nielsen, J. e Villadsen, J. (1997) Modeling isotopomer distributions in biochemical networks using isotopomer mapping matrices. *Biotechnology and Bioengineering* **55**, pp. 831-840.
- Schmidt, K.; Nielsen, J. e Villadsen, J. (1999A) Quantitative analysis of metabolic fluxes in *Escherichia coli*, using two-dimensional NMR spectroscopy and complete isotopomer models. *Journal of Biotechnology* **71**, pp. 175-190.
- Schmidt, M.; Babu, K. R.; Khanna, N.; Marten, S. e Rinas, S. (1999B) Temperature-induced production of recombinant human insulin in high-cell density of recombinant *Escherichia coli*. *Journal of Biotechnology* **68**, pp. 71-83.
- Scopes, R. K. (1994) *Protein purification: principles and practice* (3^a ed.). New York: Springer-Verlag.
- Shimizu K. (1996) A tutorial review on bioprocess systems engineering. *Computers and Chemical Engineering* **20** (6/7), pp. 915-941.

- Singh, S. M. e Panda, A. K. (2005) Solubilization and refolding of bacterial inclusion body proteins: Review. *Journal of Bioscience and Bioengineering* **99** (4), pp. 303-310.
- Smith, E. M. B. e Pantelides, C. C. (1999) A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Computers and Chemical Engineering* **23**, pp. 457-478.
- Steffens, M. A.; Fraga, E. S. e Bogle, I. D. L. (2000) Synthesis of bioprocesses using physical properties data. *Biotechnology and Bioengineering*, **68** (2), pp. 218-230.
- Stephanopoulos, G. N.; Aristidou, A. A. e Nielsen, J. (1998) *Metabolic Engineering Principles and Methodologies*. San Diego: Academic Press.
- Takagi, M.; Moriyama, T. e Yoshida, T. (2001) Effects of Shifts up and down in osmotic pressure on production of tissue plasminogen activator by chinese hamster ovary cells in suspension. *Journal of Bioscience and Bioengineering* **91** (5), 509-514.
- Thykaer, J.; Christensen, B. e Nielsen, J. (2002) Metabolic network analysis of an adipoyl-7-ADCA-producing strain of *Penicillium chrysogenum*: elucidatio of adipate degradation. *Metabolic Engineering* **4**, pp. 151-158.
- Vallino, J. J. e Stephanopoulos, G. (1993) Metabolic flux distribution in *Corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnology and Bioengineering* **41**, pp. 633-646.
- van Winden, W. A.; van Dam, J. C.; Ras, C.; Kleijn, R. J.; Vinke, J. L. van Gulik, W. M. e Heijnen, J. J. (2005) Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of ¹³C-labeled primary metabolites. *FEMS Yeast Research* **5**, pp. 559-568.
- Vásquez-Alvarez, E.; Lienqueo, M. E. e Pinto, J. M. (2001) Optimal synthesis of protein purification processes. *Biotechnology Progress* **17**, pp. 685-696.
- Vásquez-Alvarez, E. e Pinto, J. M. (2004) Efficient MILP formulations for the optimal synthesis of chromatographic protein purification processes. *Journal of Biotechnology* **110**, pp. 295-311.
- Vecchiotti, A.; Lee, S. e Grossmann, I. E. (2003) Modeling of discrete/continuous optimization problems: characterization and formulation of disjunctions and their relaxations. *Computers and Chemical Engineering* **27** pp. 433-448.
- Viswanathan, J. e Grossmann, I. E. (1990) A combined penalty function and outer approximation method for MINLP optimization. *Computers and Chemical Engineering* **14**, pp. 769-782.
- Walsh, G. e Headon, D. (1994) *Protein Biotechnology*. England: John Wiley & Sons.

- Wiechert, W. (2001) ^{13}C Metabolic flux analysis: Minireview. *Metabolic Engineering* **3**, pp. 195-206.
- Wiechert, W. e de Graaf, A. A. (1997) Bidirectional reaction steps in metabolic networks, I modeling and simulation of carbon isotope labeling experiments. *Biotechnology and Bioengineering* **55**, pp. 101-117.
- Wiechert, W.; Möllney, M.; Petersen, S. e de Graaf, A. A. (2001) A universal framework for ^{13}C metabolic flux analysis. *Metabolic Engineering* **3**, pp. 256-281.
- Yang, C.; Hua, Q. e Shimizu, K. (2002) Metabolic analysis in *Synechocystis* using isotope distribution from ^{13}C -labeled glucose. *Metabolic Engineering* **4**, pp. 202-216.
- Zhang, H.; Shimizu, K. e Yao, S. (2003A) Metabolic flux analysis of *Saccharomyces cerevisiae* grown on glucose, glycerol or acetate by ^{13}C -labeling experiments. *Biochemical Engineering Journal* **16**, pp. 211-220.
- Zhang, Y.; Dai, L. e Sun, X. (2003B) Kinetic aspect of hepatitis B surface antigen production in recombinant *Saccharomyces cerevisiae* fermentation. *Process Biochemistry* **38**, pp. 1593-1598.
- Zhao, J. e Shimizu, K. (2003) 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* **101**, pp. 101-117.
- Zupke, C. e Stephanopoulos, G. (1994) Modeling of isotope distribution and intracellular fluxes in metabolic networks using atom mapping matrices. *Biotechnology Progress* **10** (5), pp. 489-498.