



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA
DEPARTAMENTO DE QUÍMICA ANALÍTICA

Tese de Doutorado

**Avaliação comparativa dos perfis proteômicos e
metaloproteômicos de sementes de soja (*Glycine max* (L.)
Merrill) modificadas geneticamente**

Herbert de Sousa Barbosa
Orientador: Prof. Dr. Marco Aurélio Zezzi Arruda

Campinas – São Paulo
2011

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO INSTITUTO
DE QUÍMICA DA UNICAMP**

B234a Barbosa, Herbert de Sousa.
Avaliação comparativa dos perfis proteômicos e metaloproteômicos de semente de soja (*Glycine max* (L.) Merrill) modificadas geneticamente / Herbert de Sousa Barbosa. -- Campinas, SP: [s.n], 2011.

Orientador: Prof. Dr. Marco Aurelio Zezzi Arruda.

Doutorado - Universidade Estadual de Campinas,
Instituto de Química.

1. Soja. 2. Eletroforese em gel. 3. Proteômica.
4. Metaloproteômica. I. Arruda, Marco Aurelio Zezzi.
II. Universidade Estadual de Campinas. Instituto de
Química. III. Título.

Título em inglês: Comparative evaluation of proteomics and metalloproteomics profiles of soybean seeds (*Glycine max* (L.) Merrill) genetically modified

Palavras-chaves em inglês: Soybean, Gel electrophoresis, Proteomics, Metalloproteomics

Área de concentração: Química Analítica

Titulação: Doutor em Ciências

Banca examinadora: Prof. Dr. Marco Aurelio Zezzi Arruda (orientador), Prof. Dr. Pedro de Magalhães Padilha (DQB-UNESP), Prof. Dr. Ricardo Erthal Santelli (IQ-UFRJ), Prof. Dr. Fábio Cesar Gozzo (IQ-UNICAMP), Prof. Dr. Dosil Pereira de Jesus (IQ-UNICAMP)

Data de defesa: 21/03/2011

Dedico esta tese:

aos meus pais Hilton e Lenice,
aos meus irmãos Hilton e Helton,
a minha esposa Ivana e aos meus filhos Herbert Filho e Francisco,
pelo amor, carinho, incentivo, paciência e apoio incondicional.

**“O Senhor é meu pastor, nada me faltará.
Em verdes prados ele me faz repousar.
Conduz-me junto às águas refrescantes,
restaura as forças de minha alma.
Pelos caminhos retos ele me leva,
por amor do seu nome.
Ainda que eu atravessasse o vale escuro,
nada temerei, pois estais comigo.
Vosso bordão e vosso báculo são o meu amparo”.**

(Salmo 22, 1-4)

AGRADECIMENTOS

À Universidade Estadual de Campinas, em específico ao Instituto de Química, pela possibilidade da realização desta Tese de Doutorado.

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

Ao Prof. Dr. Marco Aurélio Zezzi Arruda por ter me recebido em seu laboratório, pela orientação e ensinamentos e por todas as oportunidades recebidas durante o desenvolvimento deste doutorado.

À Profa. Dra. Siu Mui Tsai, do Centro de Energia Nuclear na Agricultura (CENA-USP), pelo fornecimento das sementes de soja utilizadas neste trabalho de pesquisa.

A Rede Proteoma do Estado de São Paulo, o qual este projeto está inserido.

Ao Prof. Fábio Cezar Gozzo, por ter me recebido no Laboratório Nacional de Luz Síncrotron (LNLS) para os estudos de caracterização e identificação das proteínas por espectrometria de massas junto a Rede Proteoma do Estado de São Paulo.

A Pesquisadora Adriana Franco Paes Leme e a Bianca Pauletti, do Laboratório de Espectrometria de Massas (LNLS) pelas medidas por ESI-QTOF, junto a Rede Proteoma do Estado de São Paulo.

Aos amigos do Grupo de Espectrometria, Preparo de Amostras e Mecanização (GEPAM), inclusive aqueles que não se encontram mais no grupo: Adilson, Alessandra, Alessandra Melo, Alessandra Sussulini, Aline Klassen, Aline Lopes, Ana Cristi, Cicero, Cristiane, Daiane Torres, Daiane Leticia, Eduardo, Elidiane, Eraldo, Fabiane, Geraldo, Gustavo, Jerusa, Lidiane, Marcelo, Márcia, Renata e Sara pelo apoio e ótimo convívio durante o desenvolvimento desta Tese de Doutorado.

Aos técnicos do Instituto de Química, Ivanilce e Danilo, por todo o suporte na realização deste trabalho de pesquisa.

Aos funcionários do Instituto de Química e, em especial, a Bel da CPG, pela colaboração sempre que necessário.

Ao meu grande amigo e “irmão” Edgar por todo convívio e companheirismo durante o curso de graduação e até hoje.

Aos meus pais e irmãos, pelo carinho, apoio e ensinamentos ao longo da vida, os quais foram fundamentais para que conseguisse alcançar os objetivos traçados até o presente momento da minha vida.

Por fim, mas não menos importante, a minha esposa Ivana e aos meus filhos Herbert Filho e Francisco, os quais se constituem como os alicerces da minha vida, pelo fundamental apoio e por acreditarem e confiarem em minha força de vontade, determinação e em meu trabalho.

Enfim, gostaria de agradecer a todos, que de alguma forma, contribuíram para a realização desta Tese de Doutorado.

SÚMULA CURRICULAR

DADOS PESSOAIS

Herbert de Sousa Barbosa
Data de nascimento: 04 de abril de 1981
Naturalidade: Teresina/PI
E-mail: herbertquimica@yahoo.com.br

FORMAÇÃO ACADÊMICA

2007-2011 **Doutorado em Ciências**

Universidade Estadual de Campinas

2005-2007 **Mestrado em Química Analítica**

Universidade Federal do Piauí

2005-2006 **Licenciatura em Química**

Universidade Federal do Piauí

2000-2004 **Bacharelado em Química com Atribuições Tecnológicas**

Universidade Federal do Piauí

HISTÓRICO PROFISSIONAL

a) Universidade Estadual de Campinas

- Monitoria junto ao projeto “*Análise de Proteínas em Sementes de Soja por Eletroforese Unidimensional em Gel de Poliacrilamida*”, sob orientação do Prof. Dr. Marco Aurélio Zezzi Arruda, como parte da sexta edição do programa *Ciência & Arte nas Férias*, promovido de 04 de janeiro a 01 de fevereiro de 2008. Universidade Estadual de Campinas/Instituto de Química - Departamento de Química Analítica;
- Membro da Comissão Organizadora do 1º Encontro Brasileiro sobre Especiação Química (12/2008) e 2º Encontro Brasileiro sobre Especiação Química (12/2010);
- Estagiário docente nas disciplinas: Química Analítica III (2º semestre de 2008), Química Analítica Instrumental I (1º semestre de 2009), Química Analítica Instrumental II (2º semestre de 2009) e Química Analítica IV (1º semestre de 2010).

b) Universidade Federal do Piauí

- Estagiário docente na disciplina Química Analítica Quantitativa Experimental I (1º semestre de 2004).

- Iniciação Científica: Desenvolvimento de Método Espectrofotométrico para Determinação de Nitrito e Nitrato envolvendo o sistema Ácido Nitroso/Iodeto, bolsa Pibic-CNPq (01/08/2002 a 31/07/2003).

PRODUÇÃO CIENTÍFICA

Artigo publicado

- Brandão, A.R.; **Barbosa, H.S.**; Arruda, M.A.Z. *Image analysis of two-dimensional gel electrophoresis for comparative proteomics of transgenic and non-transgenic soybean seeds*. **J. Proteomics**, 73 (2010) 1433-1440.

Principais trabalhos apresentados em congressos

- Arruda, M.A.Z.; Sussulini, A.; Tasic, L.; Banzato, C.E.M.; **Barbosa, H S.**; Brandão, A.R. *Biomarkers and comparative "omics"*. 1st International Congress on Analytical Proteomics, 30 de setembro a 03 de outubro de 2009, Costa da Caparica, Lisboa, Portugal. Forma de Apresentação – oral. Anais divulgados em CD-Rom.
- **Barbosa, H.S.**; Brandão, A.R.; Arruda, M.A.Z. *Comparative proteomic study involving transgenic and non-transgenic soybean seeds*. 1st International Congress on Analytical Proteomics, 30 de setembro a 03 de outubro de 2009, Costa da Caparica, Lisboa, Portugal. Forma de Apresentação – painel. Anais divulgados em CD-Rom.
- Magalhães, C.S.; **Barbosa, H.S.**; Arruda, M.A.Z. *Protein Map and Metal Determination in Proteins of Horse Chestnut Seeds*. Tenth Rio Symposium on Atomic Spectrometry, 7 a 12 de Setembro de 2008, Salvador/BA. Forma de apresentação – painel. Anais divulgados em Livro de Resumos.
- **Barbosa, H.S.**; Silva, M.A.O.; Brandão, A.R.; Sussulini, A.; Gozzo, F.C.; Arruda, M.A.Z. *Comparative proteome analysis in transgenic and non-transgenic soybeans*. 2º Congresso Brasileiro de Espectrometria de Massas, 9 a 11 de Dezembro de 2007, Campinas, SP. Forma de Apresentação – painel. Anais divulgados em CD-Rom.
- Sussulini, A.; Souza, G.H.M.F.; Eberlin, M.N.; Gozzo, F.C.; Silva, M.A.O.; **Barbosa, H.S.**; Arruda, M.A.Z. *Comparative metallomics: a study involving natural and transgenic soybeans*. 35th Colloquium Spectroscopicum Internationale, de 23 a 27 de setembro de 2007, Xiamen, China. Forma de Apresentação – painel. Anais divulgados em Livro de Resumos.

Prêmio

- Prêmio “*CSI Excellent Poster Award*”, Colloquium Spectroscopicum Internationale XXXV, 2007

RESUMO

AVALIAÇÃO COMPARATIVA DOS PERFIS PROTEÔMICOS E METALOPROTEÔMICOS DE SEMENTES DE SOJA (*GLYCINE MAX* (L.) MERRILL) MODIFICADAS GENETICAMENTE

Autor: Herbert de Sousa Barbosa

Orientador: Prof. Dr. Marco Aurélio Zezzi Arruda

Este trabalho de pesquisa consiste na comparação entre dois tipos de sementes de soja [*Glycine max* (L.) Merrill], transgênica e não-transgênica, em termos proteômicos e metaloproteômicos, de modo a avaliar possíveis diferenças existentes entre as amostras (Ex. diferença de expressão protéica, concentração das espécies metálicas e não-metálicas ligadas às biomoléculas). Para isso, inicialmente, as biomoléculas foram extraídas usando um tampão extrator específico e, após o tratamento adequado da amostra, separadas por meio da técnica de eletroforese bidimensional em gel de poliacrilamida (2D-PAGE) e caracterizadas por meio de técnicas de espectrometria de massas (MALDI-QTOF e ESI-QTOF) e identificadas por meio de buscas em banco de dados relacionados utilizando o programa *Mascot*, a partir dos dados gerados pelo espectrômetro de massas. Sendo assim, foram identificadas um total de 192 proteínas, sendo 13 proteínas na faixa de pH de 3 a 10 e 179 na faixa de pH de 4 a 7, sendo que esta última faixa de pH mostrou melhor resolução dinâmica na separação das proteínas. A distribuição funcional das proteínas identificadas mostrou que a maior parte (49%) destas estão envolvidas em armazenamento de nutrientes e atividade proteolítica, citando como exemplo deste grupo as proteínas glicinina e β -conglitinina e suas subunidades. Para uma avaliação proteômica comparativa, com o intuito de se avaliar possíveis diferenças quanto à expressão das proteínas separadas entre as amostras avaliadas, foi utilizada a técnica de eletroforese em gel diferencial bidimensional (2D-DIGE), a partir de dados como intensidade e volume dos *spots* protéicos. Sendo assim, foram selecionadas um total de 04

proteínas diferenciais entre as amostras, sendo que todas foram identificadas via espectrometria de massas. Para uma avaliação metaloproteômica comparativa, um estudo para identificação de íons metálicos e semimetálicos livres ou ligados a proteínas foi realizado. Para isso, foi feita uma varredura semi-quantitativa de espécies metálicas e semimetálicas presentes nos extratos protéicos das sementes de soja transgênica e não-transgênica usando a técnica de ICP-MS, sendo que os íons Na^+ , Mg^+ , Si^+ , K^+ , Ti^+ , Mn^+ , Zn^+ e Rb^+ foram mais detectados na soja não-transgênica. Já o íon Cu^+ foi mais detectado nas sementes de soja transgênica. Estes elementos diferenciais entre as amostras foram quantificados via ICP-MS. Além disso, os *spots* diferenciais obtidos com a técnica de 2D-DIGE foram avaliados quanto a presença de cobre, já que este elemento mostrou diferenças em concentrações em proteínas de soja, conforme trabalhos recentemente publicados na literatura pelo nosso grupo de pesquisa.

ABSTRACT

COMPARATIVE EVALUATION OF PROTEOMICS AND METALLOPROTEOMICS PROFILES OF SOYBEAN SEEDS (GLYCINE MAX (L.) MERRILL) GENETICALLY MODIFIED

Author: Herbert de Sousa Barbosa

Adviser: Prof. Dr. Marco Aurélio Zezzi Arruda

This research compares two types of soybean [Glycine max (L.) Merrill], transgenic and non-transgenic, in terms of proteomics and metalloproteomics in order to evaluate possible differences among the samples (E.g. difference in protein expression, concentration of metal species and non-metal bounded to biomolecules). Therefore, initially, the biomolecules were extracted using a specific extraction buffer and, after proper treatment of the sample, separated by the technique of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and characterized by mass spectrometry techniques (MALDI-QTOF and ESI-QTOF) and identified through searches related database using the Mascot program, from the data generated by mass spectrometer. Thus, we identified a total of 192 proteins, 13 proteins in the pH range 3 to 10 and 179 at pH 4-7, and this latest range of pH showed a better dynamic resolution in the separation of proteins. The functional distribution of identified proteins showed that most (49%) of these are involved in nutrient storage and proteolytic activity, citing as an example of this group glycinin and β -conglycinin and their subunits. For an evaluation of proteomics in order to evaluate possible differences in protein expression between the samples, the differential gel electrophoresis (2D-DIGE) technique was used based on data such as intensity and volume of proteins. Therefore, we selected a total of 04 differential proteins between samples, all of them were identified by mass spectrometry. For a comparative assessment metalloproteomics, a study for identification of metal ions and semimetal free or bound to proteins was performed. In this way, a semi-quantitative sweeping of semi-metal and metal species present

in protein extracts of transgenic and non-transgenic soybean seeds was carried out using ICP-MS, and the ions Na^+ , Mg^+ , Si^+ , K^+ , Ti^+ , Mn^+ , Zn^+ and Rb^+ were detected in seeds of non-transgenic soybean. The ion Cu^+ was detected in transgenic soybeans seeds. These differential elements between the samples were quantified via ICP-MS. Moreover, the differential spots obtained with the 2D-DIGE technique were evaluated according to the presence of copper, since this element showed differences in concentrations in soybean proteins, as recently published in the literature by our research group.

ÍNDICE

LISTA DE ACRÔNIMOS E ABREVIATURAS	xxi
LISTA DE TABELAS	xxv
LISTA DE FIGURAS	xxvii
 INTRODUÇÃO	 03
 CAPÍTULO 1: <i>OBTENÇÃO DE UM MAPA PROTEÔMICO DE SEMENTES DE SOJA TRANSGÊNICA: APLICAÇÃO DAS TÉCNICAS DE 2D-PAGE E ESPECTROMETRIA DE MASSAS</i>	 07
1. OBJETIVOS	09
2. REVISÃO BIBLIOGRÁFICA	09
2.1 Soja: aspectos proteômicos	09
2.2 Soja: modificação genética	11
2.3 Eletroforese bidimensional em gel de poliacrilamida (2D-PAGE)	13
2.4 Espectrometria de massas em proteômica: MALDI-QTOF MS e ESI-QTOF MS	15
3. PARTE EXPERIMENTAL	18
3.1 Equipamentos utilizados	18
3.2 Reagentes e soluções	19
3.3. Procedimento Experimental.....	21
3.3.1 Amostras de soja	21
3.3.2 Extração das proteínas de soja	21
3.3.3 Separação das proteínas por 2D-PAGE	22

3.3.4 Digestão proteolítica e identificação das proteínas por Espectrometria de massas	24
3.3.5 Identificação da enzima <i>CP4 EPSPS</i> por SDS-PAGE e espectrometria de massas	26
4. RESULTADOS E DISCUSSÃO	28
4.1 Identificação das proteínas por MALDI-QTOF-MS e ESI-QTOF-MS .	28
4.2 Avaliação funcional das proteínas identificadas	31
4.2.1 Armazenamento e atividade proteolítica	31
4.2.2 Crescimento/Divisão Celular	33
4.2.3 Metabolismo/Energia	34
4.2.4 Transporte	35
4.2.5 Doença/Defesa	35
4.2.6 Transdução de sinais / Metabolismo secundário	36
4.2.7 Não-classificada / Classificação desconhecida	36
4.3 Identificação da enzima <i>CP4 EPSPS</i>	37
5. CONCLUSÕES PARCIAIS	45
 CAPÍTULO 2: <i>PROTEÔMICA E METALOPROTEÔMICA COMPARATIVAS EM SEMENTES DE SOJA TRANSGÊNICA E NÃO TRANSGÊNICA</i>	 47
1. OBJETIVOS	49
2. REVISÃO BIBLIOGRÁFICA	49
2.1 Eletroforese em gel bidimensional (2D-PAGE) para estudos proteômicos comparativos	50
2.2 Eletroforese diferencial em gel bidimensional (2-D DIGE)	51
2.3 Metais associados a proteínas	54

2.4 Espectrometria de massas em metalômica: ICP-MS	55
3. PARTE EXPERIMENTAL	58
3.1 Equipamentos utilizados	58
3.2 Reagentes e soluções	59
3.3. Procedimento Experimental.....	59
3.3.1 Amostras de soja	59
3.3.2 Extração das proteínas de soja	59
3.3.3 Preparo dos extratos de soja para análise por 2D-DIGE	60
3.3.4 Separação das proteínas por 2D-DIGE	60
3.3.5 Digestão trípica e identificação das proteínas por Espectrometria de massas	61
3.3.6 Decomposição do extrato protéico das sementes de soja por radiação microonda	62
3.3.7 Identificação e quantificação dos íons livres no extrato protéico de soja por ICP-MS	63
3.3.8 Decomposição dos <i>spots</i> protéicos diferenciados obtidos pela técnica de 2D-DIGE por radiação microonda	64
3.3.9 Quantificação de cobre nos <i>spots</i> protéicos diferenciados por ICP-MS	65
4. RESULTADOS E DISCUSSÃO	66
4.1 Separação por 2D-DIGE	66
4.2 Identificação dos <i>spots</i> diferenciais por espectrometria de massas ..	71
4.3 Avaliação do perfil ionômico nos extratos protéicos por ICP-MS	73
4.4 Quantificação de cobre nos <i>spots</i> protéicos diferenciados por ICP-MS.	77
5. CONCLUSÕES PARCIAIS	80

REFERÊNCIAS BIBLIOGRÁFICAS	81
CONSIDERAÇÕES FINAIS	91
ANEXOS	95

LISTA DE ACRÔNIMOS E ABREVIATURAS

2-D PAGE	Eletroforese bidimensional em gel de poliacrilamida, do inglês, <i>two-dimensional polyacrylamide gel electrophoresis</i>
2-D DIGE	Eletroforese diferencial em gel bidimensional, do inglês, <i>two-dimensional differential in-gel electrophoresis</i>
ACN	Acetonitrila
CBB	Azul de Coomassie, do inglês, <i>Coomassie Brilliant Blue</i>
CHAPS	Sulfonato de 3-[(3-cloroamidopropil)-dimetilamônio]-1-propano, do inglês, <i>3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate</i>
CHCA	Ácido α -ciano-4-hidroxicinâmico, do inglês, <i>α-cyano-4-hydroxycinnamic acid</i>
CID	Dissociação induzida por colisão, do inglês, <i>collision-induced dissociation</i>
Da	Dalton ($1 \text{ Da} = 1,661 \cdot 10^{-24} \text{ g}$)
DC	Corrente contínua, do inglês <i>Direct current</i>
DNA	Ácido desoxirribonucléico, do inglês, <i>deoxyribonucleic acid</i>
DTT	1,4-ditiotreitol
ELISA	Do inglês <i>Enzyme-linked immunosorbent assay</i>
EPSPS	Enolpiruvil shikimato-3-fosfato sintase, do inglês <i>5-enolpyruvylshikimate-3-phosphate synthase</i>
ESI	Ionização por <i>electrospray</i> , do inglês, <i>electrospray ionization</i>
ICP-MS	Espectrometria de massas com fonte de plasma indutivamente acoplado, do inglês, <i>inductively coupled plasma mass spectrometry</i>
IEF	Focalização isoelétrica, do inglês, <i>isoelectric focusing</i>
IPG	Gradiente de pH imobilizado, do inglês, <i>immobilized pH</i>

gradient

LA	Ablação a laser, do inglês, <i>laser ablation</i>
LC-MS	Cromatografia líquida de alta eficiência acoplada à espectrometria de massas, do inglês, <i>liquid chromatography mass spectrometry</i>
MALDI	Ionização por desorção a laser assistida por matriz, do inglês, <i>matrix-assisted laser desorption ionization</i>
MM	Massa molar
MS	Espectrometria de massas, do inglês, <i>mass spectrometry</i>
MS/MS	Espectrometria de massas em sequência
MSDB	Do inglês <i>Mass spectrometry data bank</i>
NCBI	Do inglês <i>National Center for Biotechnology Information</i>
PCR	Reação em cadeia da polimerase, do inglês <i>Polymerase Chain Reaction</i>
pI	Ponto isoelétrico
PMF	Determinação da impressão digital pela massa de peptídeos, do inglês <i>peptide mass fingerprinting</i>
PMSF	Do inglês <i>Phenylmethylsulfonyl fluoride</i>
Q	Quadrupolo, do inglês <i>quadrupole</i>
RF	Radio frequência
RR	Do inglês <i>Roundap Ready</i>
SF	Do inglês, <i>sector field</i>
SDS	Dodecil sulfato de sódio, do inglês, <i>sodium dodecyl sulfate</i>
SDS-PAGE	Eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio, do inglês, <i>sodium dodecyl sulfate polyacrilamide gel electrophoresis</i>
TCA	Ácido tricloroacético

TEMED	<i>N,N',N,N'</i> -tetrametiletenodiamina
TFA	Ácido trifluoroacético
TOF	Tempo de vôo, do inglês, <i>time of flight</i>
UV	Ultravioleta

LISTA DE TABELAS

- TABELA 1.1.** Sequência de peptídeos, obtidos pelo *Mascot*, usados na identificação da enzima *CP4 EPSPS* por ESI-QTOF MS (Banco NCBI - *Glycine max*, enzima tripsina) p.39
- TABELA 1.2.** Identificação da enzima *CP4 EPSPS* na semente de soja transgênica usando SDS-PAGE e espectrometria de massas, tipo ESI-QTOF. p.40
- TABELA 2.1.** Programa aplicado para a decomposição dos extratos protéicos das sementes de soja por radiação microonda, p.57
- TABELA 2.2.** Condições de operação do equipamento de ICP-MS para análise semi-quantitativa, p.58
- TABELA 2.3.** Programa aplicado para a decomposição dos *spots* protéicos diferenciais por radiação microonda, p.60
- TABELA 2.4.** Níveis de diferenciação entre os *spots* com diferenças de expressão, p.63
- TABELA 2.5.** Resultados obtidos para a identificação das proteínas com diferenças de expressão entre as sementes de soja transgênica e não-transgênica, p. 66
- TABELA 2.6.** Parametros das curvas analíticas de calibração dos elementos diferenciais obtidos por ICP-MS, p.69
- TABELA 2.7.** Concentrações dos elementos Na, K e Mg nos extratos protéicos das sementes de soja transgênica e não-transgênica (média $\pm$ desvio padrão), p.69
- TABELA 2.8.** Concentrações dos elementos Mn, Cu, Rb e Zn nos extratos protéicos das sementes de soja transgênica e não-transgênica (média $\pm$ desvio padrão) p.70
- TABELA 2.9.** Parametros da curva analítica de calibração e algumas figuras de mérito para a quantificação de cobre nos *spots* protéicos diferenciais, p.72
- TABELA 2.10.** Concentrações de cobre para os *spots* protéicos diferenciais obtidos com a técnica de 2D-DIGE, p.73

LISTA DE FIGURAS

- FIGURA 1.1.** Inibição da enzima Epsps pelo glifosato. Figura adaptada de Ocaña *et al* [19], p.08
- FIGURA 1.2.** Proteínas identificadas para sementes de soja transgênica: na faixa de pl de 3 a 10 (A), correspondendo a um total de 13 proteínas identificadas e na faixa de pl de 4 a 7 (B), correspondendo a um total de 179 proteínas identificadas, p. 25
- FIGURA 1.3.** Distribuição funcional das proteínas de soja transgênica identificadas, p. 27
- FIGURA 1.4.** Estrutura tridimensional da proteína: (a) glicinina sub-unidade A₃B₄ (G4); (b) sub-unidade α da β-conglicinina [79], p. 28
- FIGURA 1.5.** Estrutura tridimensional da proteína: (a) *lipxygenase-3*; (b) *Mn-Superoxide dismutase* [79], p.31
- FIGURA 1.6.** Possível localização do *spot* no gel 2D-PAGE, (pl 4-7) correspondente a enzima *CP4 EPSPS* de acordo o gel obtido neste trabalho, p. 34
- FIGURA 1.7.** Perfil eletroforético obtido para a semente de soja transgênica usando diferentes quantidades em massa de proteínas. Em destaque (vermelho) a banda protéica cortada e submetida a digestão enzimática e posterior análise por ESI-QTOF MS, p.35
- FIGURA 1.8.** Seqüência de aminoácidos da enzima *CP4 EPSPS* usando a enzima tripsina na etapa de digestão enzimática. Em vermelho, são destacados os peptideos identificados por ESI-QTOF MS (Banco NCBI - *Glycine max*). p. 36
- FIGURA 1.9.** Seqüência de aminoácidos da enzima *CP4 EPSPS* usando a enzima quimiotripsina na etapa de digestão enzimática. Em vermelho, são destacados os peptideos identificados por ESI-QTOF MS (Banco NCBI - *Glycine max*). p. 36
- FIGURA 1.10.** Espectros de MS/MS relativos aos peptideos da enzima *CP4 EPSPS*, obtidos no modo “*top three*” (três picos mais intensos) por ESI-QTOF MS/MS, p. 38
- FIGURA 2.1.** Estrutura química dos corantes Cy2, Cy3 e Cy5, usados na técnica de 2-D DIGE. Figura Adaptada de Tonge *et al.* 2001 [62], p. 46

FIGURA 2.2. Análise por 2-D DIGE (faixa de pH entre 4 e 7) para sementes de soja transgênica (7575 RR) e não-transgênica (7501). Em azul estão marcadas as proteínas sub-expressas e, em vermelho, as sobre-expressas. Os números são relacionados às identificações das proteínas (Ver tabela 2.4), p. 62

FIGURA 2.3. *Spots* com variação de expressão entre as amostras de soja transgênica (T) e não-transgênica (NT). A área delimitada pelo software com coloração amarela, mostra o spot com variação de expressão. No canto superior, à direita, está o número do *spot* localizado no gel 2-D DIGE (Ver Figura 2.2),p. 64

1 INTRODUÇÃO

INTRODUÇÃO

Atualmente, a soja é considerada umas das principais fontes de proteínas vegetais, pois, aproximadamente, 40% (massa seca) de todos os componentes da semente são proteínas [1]. Esta quantidade expressiva de proteínas, e também um percentual considerável de óleo (aproximadamente 21%, também em massa seca), tornam a soja um produto de grande importância para o setor industrial, pois várias aplicações podem ser propícias, como, por exemplo, na fabricação de farinha protéica, que pode ser empregada nos mais diversos setores alimentícios. Além destas aplicações, a soja é também apontada como uma fonte alternativa de combustível, ou seja, o biodiesel [2].

Nos últimos anos, muitas variedades de soja têm sido desenvolvidas para adaptação em diferentes regiões geográficas, com o intuito de gerar plantas com maiores produtividade e qualidade. As plantações de soja geneticamente modificadas já ocupam, aproximadamente, 60 % da área mundial cultivada [3]. Neste contexto, a soja transgênica denominada *Roundup Ready* (RR) desenvolvida pela empresa Monsanto (EUA), é um dos organismos transgênicos mais amplamente comercializados [2].

Em anos recentes, vários trabalhos científicos, versando sobre temas como genômica e proteômica, foram publicados na literatura, implicando em grandes avanços nas mais diferentes áreas do conhecimento, incluindo a química, bioquímica, biologia, dentre outras. Tais estudos estabeleceram condições favoráveis para consolidação de um amplo campo de estudo envolvendo biomoléculas [4].

O proteoma de um organismo, objeto principal de estudos da proteômica, se define como a totalidade de proteínas que são expressas por um genoma em um dado momento e sob determinadas condições. Cada célula de um mesmo organismo contém o mesmo genoma, entretanto, os diferentes tipos celulares expressam centenas de proteínas diferentes, e cada uma delas pode experimentar numerosas modificações em resposta a um microambiente determinado [5].

A introdução, eliminação e/ou substituição de genes pode modificar o mecanismo de produção de proteínas, acarretando em algumas consequências: síntese de novas substâncias, ausência de proteínas que eram sintetizadas antes da modificação genética e a síntese de uma grande quantidade de substâncias já presentes no organismo [6]. Tais efeitos são conhecidos por gerar significativas mudanças no proteoma do organismo [7].

Uma das classes mais importantes de marcadores biológicos, verificados em diferenciações fisiológicas ou em alterações de células ou tecidos biológicos, envolve íons metálicos ou semimetálicos ligados a proteínas [8]. Dentro deste contexto, ressalta-se que um número significativo de proteínas e enzimas (ca. 30 %) contém espécies metálicas ou metalóides em suas estruturas [9,10] sendo estas denominadas de metaloproteínas, ou proteínas ligantes de metal. As metaloproteínas são definidas como proteínas as quais suas funções são conferidas por sua ligação com o metal, como por exemplo: atividade catalítica, implicação na reação de transferência de elétrons ou estabilização da estrutura terciária ou quaternária das proteínas. Já as proteínas ligantes de metal apenas complexam íons metálicos em condições termodinamicamente favoráveis e, neste caso, o metal não possui uma função específica [4,8,11].

Como, até o presente momento, existem poucos estudos na literatura envolvendo a busca de biomarcadores protéicos e/ou metaloprotéicos para a modificação genética em sementes de soja, a hipótese deste trabalho é que a busca destes marcadores biológicos possa gerar indícios que auxiliem na identificação de sementes de soja com alterações genéticas a partir da inserção do gene hexógeno oriundo do microorganismo *Agrobacterium* sp. Cepa CP4. Além disso, outra hipótese deste trabalho é que as modificações genéticas em sementes de soja podem gerar alterações não somente em nível proteômico, mas, também, em nível metaloproteômico.

A identificação e quantificação de biomarcadores protéicos e/ou metaloprotéicos exigem a aplicação de diferentes técnicas analíticas eficientes, robustas e sensíveis. Neste sentido, o presente trabalho de Tese utilizou a combinação de técnicas de separação como a eletroforese bidimensional em gel

de poliacrilamida (2D-PAGE) e eletroforese diferencial em gel de poliacrilamida (2D-DIGE), com técnicas de espectrometria de massas em proteômica (MALDI-QTOF e ESI-QTOF) e metaloproteômica (ICP-MS) para a identificação de possíveis biomarcadores correlacionados com o processo de modificação genética em sementes de soja.

Neste trabalho de Tese, os resultados obtidos foram divididos em dois capítulos, para melhor distribuição e compreensão dos dados gerados. No Capítulo 1, foi feito estudo para obtenção de um mapa proteômico de referência para sementes de soja transgênica, usando eletroforese bidimensional em gel de poliacrilamida (2D-PAGE) e espectrometria de massas. Além disso, com o intuito de comprovarmos a modificação genética do organismo em estudo, foi feita a identificação específica da enzima *CP4 EPSPS* na semente de soja transgênica, usando SDS-PAGE e espectrometria de massas.

No capítulo 2, foi abordado um estudo proteômico comparativo entre as sementes de soja transgênica e não-transgênica usando a técnica de eletroforese diferencial bidimensional em gel de poliacrilamida (2D-DIGE). Os spots com diferenças de expressão entre as amostras avaliadas foram submetidos a identificação por espectrometria de massas, tipo ESI-QTOF. Além disso, um estudo metaloproteômico comparativo foi feito entre as amostras avaliadas. Para isso, foi feita uma varredura semi-quantitativa de espécies metálicas e semimetálicas presentes nos extratos protéicos das sementes de soja transgênica e não-transgênica usando a técnica de ICP-MS. Aquelas espécies metálicas que mostraram diferenças de intensidades entre as amostras foram quantificadas. Além disso, com relação aos *spots* diferenciais obtidos com a técnica de 2D-DIGE entre as amostras, foi feita a determinação de cobre nestes *spots* levando em consideração que este elemento já ter sido encontrado como diferencialmente expresso em trabalhos já reportados na literatura pelo nosso grupo de pesquisa.

CAPÍTULO 1:

OBTENÇÃO DE UM MAPA PROTEÔMICO DE SEMENTES DE

SOJA TRANSGÊNICA: APLICAÇÃO DAS TÉCNICAS DE 2D-

PAGE E ESPECTROMETRIA DE MASSAS

1. OBJETIVOS

Os objetivos deste capítulo consistem na elaboração de um mapa proteômico de referência para sementes de soja transgênica, como também avaliar a distribuição funcional das proteínas identificadas. Para isso, foram empregados a eletroforese bidimensional em gel de poliacrilamida (2D-PAGE) e espectrometria de massas, dos tipos MALDI-QTOF e ESI-QTOF. Além disso, foi feita a identificação da enzima *CP4 EPSPS* presente na semente de soja transgênica, sintetizada a partir do gene hexógeno introduzido no genoma da semente, usando as técnicas de SDS-PAGE e espectrometria de massas, tipo ESI-QTOF para identificação, com o intuito de melhor evidenciarmos o processo de modificação genética na semente de soja.

2. REVISÃO BIBLIOGRÁFICA

A revisão bibliográfica referente a este capítulo aborda uma descrição sobre a amostra em estudo (sementes de soja), bem como aspectos com relação ao processo de modificação genética. Em seguida, serão feitas considerações sobre a separação de proteínas empregando eletroforese bidimensional em gel de poliacrilamida (2D-PAGE) e sua identificação empregando espectrometria de massas (MALDI-QTOF e ESI-QTOF), técnicas estas bastante empregadas em estudos proteômicos.

2.1 Soja: aspectos proteômicos

Por muitos anos, a soja [*Glycine max* (L.) Merrill] tem constituído a base das dietas humana e animal em muitos países orientais, sendo uma importante fonte de ácidos graxos e proteínas. Nos últimos anos, seu uso tem aumentado em todo

o mundo devido as suas qualidades nutricionais, propriedades funcionais, abundância, baixo custo e também devido aos desenvolvimentos tecnológicos recentes que visam incluir proteínas de soja em uma grande variedade de produtos comestíveis [12]. Tendo em vista a grande importância do cultivo de soja na economia mundial, aumentou-se, significativamente, os estudos científicos por parte de institutos de pesquisa e indústrias de biotecnologia de modo a aperfeiçoar a qualidade dos seus nutrientes bem como aumentar sua resistência a vários tipos de pestes [13].

Do total de proteínas presentes na semente de soja, glicinina e β -conglucina e suas sub-unidades são componentes predominantes nas sementes de soja. Estas proteínas exercem a função de armazenamento e correspondem à cerca de 70-80% do conteúdo total de proteínas, sendo grandemente responsáveis pelas propriedades nutricionais e físico-químicas da soja [14]. A β -conglucina, uma globulina 7S, é uma glicoproteína trimérica consistindo de três subunidades, α , α' e β . A glicinina, uma globulina 11S, consiste de cinco subunidades, G1, G2, G3, G4 e G5 [14].

Na literatura, são encontrados vários estudos proteômicos utilizando a soja [12-15]. Boa parte destes estudos estão focados no estudo das duas maiores proteínas de reserva da soja, glicinina e β -conglucina. Além disso, a comparação proteômica entre diferentes genótipos de soja também esta sendo bastante avaliado. Alguns trabalhos envolvendo estes estudos, reportados na literatura, serão discutidos a seguir.

Natarajan *et al* [14] estudaram e compararam a composição protéica das duas maiores proteínas de armazenamento presente em soja, β -conglucina e glicina, em genótipos de soja nativa (*Glycine soja*) e cultivada (*Glycine max*). Para isso, aplicaram eletroforese bidimensional em gel de poliacrilamida (2D-PAGE) para separação das proteínas em três diferentes faixas de pontos isoeletricos (3-10, 4-7 e 6-11) em combinação com espectrometria de massas e cromatografia líquida (MALDI-TOF MS e LC-MS/MS) para identificação das proteínas avaliadas. Os resultados obtidos neste trabalho demonstraram um elevado grau de diferença

entre os genótipos avaliados, com relação às diferentes subunidades das proteínas β -conglucina e glicina.

Xu *et al* [15] realizaram a separação e identificação de proteínas presentes em folhas de soja empregando 2D-PAGE e espectrometria de massas (MALDI-TOF MS e LC-MS/MS). Os autores usaram o programa *Mascot* para identificação das proteínas a partir de três bancos de dados: *NCBI*, *SwissProt* e *EST*. Com este trabalho, os autores conseguiram identificar um total de 116 proteínas, sendo que, deste total, mais de 50% correspondem a proteínas envolvidas no metabolismo energético da planta.

Estes trabalhos mostram a importância da realização de estudos com a soja, especialmente em nível proteômico. Ressalta-se, ainda, que um processo amplamente empregado na produção da soja e que acarreta em mudanças no proteoma deste vegetal é a sua modificação genética. A seguir, este processo será discutido.

2.2 Soja: modificação genética

Atualmente, a soja transgênica mais conhecida e comercializada é aquela denominada *Roundup Ready* (RR), a qual foi desenvolvida pela empresa Monsanto (EUA). Em dados recentes, a soja *Roundup Ready* ocupou 77% do total de 90 milhões de hectares plantados no ano de 2009 [16]. A escolha da soja *Roundup Ready* é atribuída à simplicidade do uso, eficiência de custo para o controle de ervas daninhas e compatibilidade com a conservação da lavoura [17].

Neste caso, o processo de modificação genética na semente de soja consiste na inserção do gene *CP4 EPSPS* (5-enolpiruvil shikimato-3-fosfato sintase) da bactéria do solo *Agrobacterium* sp. tipo CP4, o que torna a planta resistente a herbicidas que possuem o glifosato (N-fosfonometil-glicina) como princípio ativo [7]. Este herbicida é considerado como não-seletivo, ou seja, além de eliminar as ervas daninhas, ele também pode atacar as plantas de soja [18].

Entretanto, o glifosato é considerado um herbicida ambientalmente seguro, pois apresenta baixa toxicidade, pequena ou ausente mobilidade do solo para a água e persistência limitada.

A enzima Epsps está envolvida na biosíntese de aminoácidos aromáticos e, em uma das etapas desta síntese, catalisa a reação do shikimato-3-fosfato (S3P) e fosfoenolpiruvato para 5-enolpiruvil shikimato-3-fosfato (Figura 1.1). No entanto, o glifosato inativa este processo pela formação de um complexo com a enzima, Epsps-S3P-glifosato [7,19]. Diferentemente, a proteína *CP4 EPSPS*, presente na soja transgênica devido à inserção do gene bacteriano, é funcionalmente similar a Epsps produzida pelas plantas não-transgênicas, porém, esta possui uma menor afinidade pelo glifosato [20].

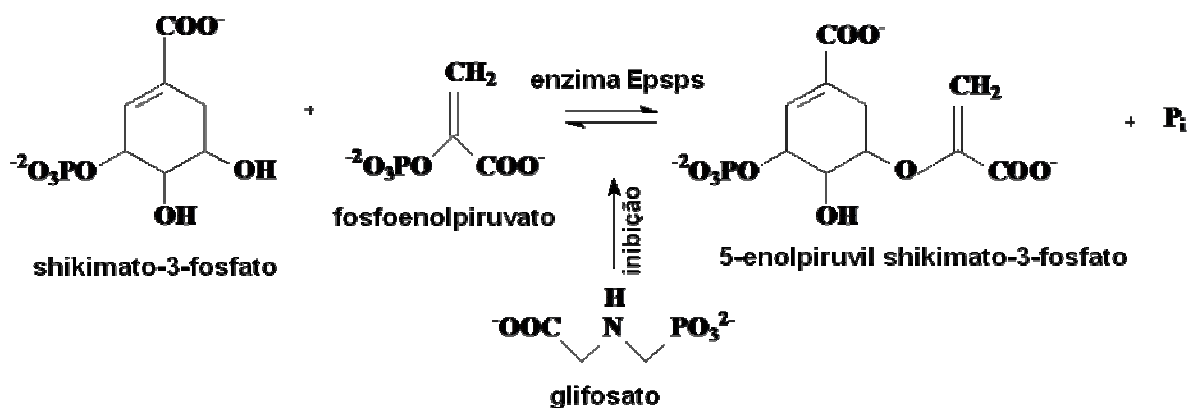


Figura 1.1. Inibição da enzima Epsps pelo glifosato. Figura adaptada de Ocaña *et al* [19].

Entre os procedimentos mais comumente empregados para a detecção de organismos geneticamente modificados, destacam-se os métodos PCR (*Polymerase Chain Reaction*) e ELISA (*Enzyme-linked immunosorbent assay*), os quais são baseados na análise de DNA e proteína, respectivamente. Estes métodos são bastante sensíveis e com baixos limites de detecção. No entanto, o método PCR é relativamente caro e depende da integridade do DNA, o qual pode ser degradado por aquecimento, atividade de nucleases e baixo valor de pH. O método ELISA requer um anticorpo específico para cada OGM e não pode ser

aplicado para produtos de soja processados devido à desnaturação e degradação das proteínas, após o processamento da soja [7,19].

Sendo assim, uma grande alternativa a estes métodos seria a aplicação de métodos de eletroforese em gel juntamente como a técnica de espectrometria de massas. Neste contexto, Ocaña *et al* [19] utilizaram a técnica de SDS-PAGE juntamente com espectrometria de massas para identificar a enzima *CP4 EPSPS* em sementes de soja e milho modificadas geneticamente. Para isso, os autores, inicialmente, fizeram um pré-fracionamento dos extratos protéicos usando cromatografia por exclusão de tamanho e por troca iônica. As frações foram, então, separadas por SDS-PAGE, e a banda correspondente a massa da enzima, bem como as frações cromatográficas, foram digeridas e caracterizadas por MALDI-TOF MS e nanoLC/nanoESI-QTOF MS. A identificação da enzima foi feita usando o programa *Mascot* a partir de busca no banco de dados do *NCBI*.

Sendo assim, métodos baseados em eletroforese em gel (SDS-PAGE e 2D-PAGE) tem sido bastante aplicados em estudos proteômicos, seja na elaboração de mapas proteômicos bem como na identificação de proteínas específicas. Estes métodos serão descritos a seguir.

2.3 Eletroforese bidimensional em gel de poliacrilamida (2D-PAGE)

A eletroforese em gel é uma técnica de separação baseada na migração de espécies carregadas em um determinado meio, sob a influência de um campo elétrico contínuo. Na eletroforese, as propriedades mais importantes das proteínas utilizadas são a sua massa molar, a sua carga e a sua conformação [21]. Neste contexto, cabe enfatizar que a aplicação da eletroforese bidimensional em gel de poliacrilamida neste trabalho de tese é bastante conveniente, pois a mesma possui elevada resolução e ampla faixa dinâmica, permitindo uma separação eficiente das proteínas do organismo em estudo [22]

A poliacrilamida é uma matriz de gel muito comum devido as suas características intrínsecas como, por exemplo, ser um material quimicamente inerte, transparente e estável em extensas faixas de pH, temperatura e força iônica. A poliacrilamida é formada pelo processo de copolimerização entre a acrilamida e a bis-acrilamida (N,N'-metilenobisacrilamida) [21].

A eletroforese bidimensional em gel de poliacrilamida é amplamente empregada para separação de proteínas, presentes em diferentes sistemas biológicos, pois combina duas técnicas distintas no processo de separação: a focalização isoelétrica e a eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE) [23]. Na primeira dimensão, a focalização isoelétrica tem como objetivo separar as proteínas de acordo com os seus pontos isoelétricos (pI), sendo este o pH no qual uma molécula apresenta uma carga líquida igual a zero. Além disso, a proteína com um determinado pI irá migrar pelo gel com gradiente de pH até que sua carga líquida seja nula; neste ponto, o pI da proteína é igual ao pH do gradiente [21,23,24]

Na SDS-PAGE, é utilizado um surfactante para desnaturar as proteínas antes da aplicação do campo elétrico. Seu principal objetivo é utilizar o SDS para transferir carga negativa para as proteínas, mascarando suas cargas intrínsecas, de modo que elas possuam formas similares e razões massa/carga parecidas e, desta maneira, seja capaz de separar as proteínas de acordo com as suas massas molares [22,24,25]. Portanto, a técnica 2D-PAGE separa as proteínas pela combinação de duas propriedades diferentes: na primeira dimensão, a IEF separa as proteínas de acordo com o seu ponto isoelétrico e, na segunda dimensão, a SDS-PAGE separa as proteínas de acordo com suas massas molares [23,24].

Para que o processo de separação empregando a técnica 2D-PAGE tenha resultados satisfatórios, o preparo de amostras constitui-se como etapa determinante para a obtenção de uma boa resolução na separação das proteínas. Para isso, as proteínas presentes no sistema biológico em estudo devem ser desnaturadas, desagregadas, reduzidas e solubilizadas [25]. Devido a grande

diversidade de proteínas em uma amostra, um procedimento deve ser otimizado em laboratório, de modo a garantir boa resolução na separação protéica [24].

Em estudos proteômicos, após a separação das proteínas aplicando a técnica de 2D-PAGE, os spots protéicos obtidos são submetidos a caracterização com o uso da técnicas de espectrometria de massas, o qual será discutida a seguir.

2.4 Espectrometria de massas em proteômica: MALDI-QTOF MS e ESI-QTOF MS

Atualmente, a espectrometria de massas, juntamente com a bioinformática, tornaram-se essenciais em estudos proteômicos, devido não apenas a sua sensibilidade, mas, também, ao conteúdo total de informações que podem ser obtidas. A técnica de MALDI-QTOF-MS apresenta-se como uma das mais aplicadas para a caracterização de proteínas após realizar a separação das mesmas pela técnica de eletroforese bidimensional em gel de poliacrilamida (2D-PAGE), uma vez que esta técnica exige que as proteínas sejam previamente isoladas. Basicamente, a técnica utiliza o feixe de um laser para fornecer energia às moléculas, proporcionando sua dessorção, levando-as para a fase gasosa em sua forma iônica [26,27].

No procedimento de análise, a amostra é misturada com excesso de uma matriz orgânica ácida que absorva a energia do laser incidente (geralmente utiliza-se um laser de N₂ altamente energético). A absorção da energia do laser causa a evaporação da matriz, sendo que o analito, resultante na fase gasosa, torna-se ionizado via mecanismos de transferência de prótons em fase gasosa. Os íons formados são, então, conduzidos até o detector por meio da aplicação de uma ddp [26-28].

Além da fonte de ionização do tipo MALDI, outra técnica utilizada para identificação de proteínas é a ionização por *eletrospray* (ESI). Esta técnica tem como princípio transferir os íons existentes em uma solução para a fase gasosa.

Esta versatilidade tem aumentado significativamente a gama de substâncias capazes de serem determinadas, possibilitando, por exemplo, que polipeptídeos possam ser analisados sem que haja sua decomposição. A ionização por *electrospray* envolve a formação de um *spray* eletrolítico, que gera pequenas gotas carregadas, e, destas, são liberados os íons. É necessária uma fonte de alta tensão (1,0 a 7,0 kV) que esteja em contato com a solução contendo os eletrólitos. Esta solução é bombeada através de um microcapilar (d.i. 50 a 100 μm) com uma velocidade de fluxo da ordem de 1 a 20 $\mu\text{L}/\text{min}$ ou menores. Quando um potencial positivo é aplicado na solução, os íons positivos tendem a se afastar para uma região menos positiva, isto é, em direção ao contra-eletrodo. Conforme a densidade de carga aumenta na gota, o campo elétrico formado entre o capilar e o contra eletrodo aumenta provocando uma deformação na gota que está presa na ponta do capilar. A gota ganha forma de um cone e permanece “presa” ao capilar até o momento em que a densidade de carga na superfície da gota e o aumento da repulsão entre os íons vençam a tensão superficial, ocorrendo à liberação de pequenas gotas com alta densidade de carga [29].

No espectrômetro de massas, a determinação das relações massa/carga pode ser feita utilizando tanto um analisador por tempo de voo (TOF), bem como um analisador tipo quadrupolo, possibilitando que íons específicos sejam selecionados e sujeitos à fragmentação para posterior obtenção de um espectro de íons precursores. Em um quadrupolo, os íons são selecionados aplicando, nas barras do quadrupolo, voltagens RF/DC, possibilitando que ocorra a inversão de polaridade entre as barras. Com isso, o quadrupolo passa a atuar como um filtro de massas, permitindo apenas que os íons que apresentam estabilidade neste campo possam ser selecionados. Em analisadores tipo TOF, os íons são separados e analisados ao longo de um tubo de tamanho específico de acordo com suas diferentes velocidades. No modo Q-TOF, os íons são inicialmente selecionados no quadrupolo, fragmentados em uma célula de colisão e as massas dos fragmentos dos íons são determinados pelo segundo analisador TOF [26-28].

Anterior a análise no espectrômetro de massas, realiza-se, primeiramente, a digestão proteolítica das proteínas (geralmente usa-se tripsina), sendo que os

peptídeos formados são, então, misturados com uma matriz, em análises usando ionização do tipo MALDI. Assim, obtém-se um espectro característico para a proteína em questão, numa espécie de “impressão digital” da macromolécula.

Após a obtenção dos espectros dos fragmentos das proteínas, é feita a identificação das mesmas a partir de buscas em bancos de dados utilizando algoritmos específicos para identificação de proteínas [30]. Este processo exige conhecimento prévio dos tratamentos aplicados ao longo do preparo da amostra, seja na etapa de separação eletroforética ou durante a digestão proteolítica das proteínas, incluindo, por exemplo, a enzima utilizada para realizar a digestão, bem como os agentes de redução e alquilação aplicados.

3. PARTE EXPERIMENTAL

3.1 Equipamentos utilizados

Os equipamentos utilizados no desenvolvimento dos experimentos descritos neste capítulo são listados a seguir. Eles se encontram disponíveis no Grupo de Espectrometria, Preparo de Amostras e Mecanização (GEPAM), localizado no Instituto de Química, Unicamp. Os Espectrômetros de massas, para identificação das proteínas, encontram-se disponíveis no Laboratório Dalton de Espectrometria de Massas, no Instituto de Química, Unicamp, bem como no Laboratório Nacional de Biociências (LNBio), localizado no Laboratório Nacional de Luz Síncrotron (LNLS), Campinas, SP. Estes equipamentos foram utilizados dentro da Rede Proteoma do Estado de São Paulo.

- Agitador de tubos (Phoenix, BR)
- Balança analítica, modelo 260 (Mettler, EUA);
- Espectrofotômetro UV/Visível, modelo B582 (Micronal, Brasil);
- Espectrômetro de massas do tipo MALDI-QTOF (Waters Q-Tof Premier, Inglaterra) e ESI-QTOF (Waters Q-Tof Ultima, Inglaterra);
- Estufa (Quimis, Brasil);
- Mesa agitadora, modelo Q225M (Quimis, Brasil);
- Misturador vortex, modelo M-37600 (Thermolyne, EUA);
- *Scanner*, modelo ImageScannerTM II (GE Healthcare, Suécia);
- Sistema de purificação Milli-Q, modelo QuantumTM cartridge (Millipore, França);
- Sistema para eletroforese 2-D PAGE, modelo EttanTM Daltsix (GE Healthcare, Suécia);
- Ultracentrífuga, modelo Bio-Spin-R (BioAgency, Brasil);
- Speed Vac modelo SPD1010-115 (Thermo Electron Corporation).

- Vidrarias de uso rotineiro em laboratório.

3.2 Reagentes e soluções

Todas as soluções foram preparadas utilizando água deionizada ($\geq 18,2$ M Ω), a partir de sistema de purificação Milli-Q. Os reagentes utilizados são pureza analítica, provenientes das empresas Amersham Biosciences (Suécia), BioAgency (Brasil), J.T. Baker (EUA) e Merck (Alemanha), e serão listados a seguir. As vidrarias usadas são de uso rotineiro na área de bioanalítica.

- 3-[(3-Colamidopropil)-dimetilamonio]-1-propanosulfonato (CHAPS) – $C_{32}H_{58}N_2O_7S$ - 614,88 g/mol (USB);
- Acetado de amônio, $C_2H_7NO_2$ - 77,08 g/mol (Mallinckrodt);
- Acetona, C_3H_6O - 58,08 g/mol (Synth);
- Acetonitrila, grau HPLC, C_2H_3N – 58,1 g/mol (Synth);
- Ácido acético glacial, $C_2H_4O_2$ - 60,05 g/mol (J. T. Baker);
- Ácido clorídrico, HCl - 36,5 g/mol (Merck);
- Ácido trifluoroacético, $C_2HF_3O_2$ – 114,03 g/mol (Merck);
- Ácido fosfórico, H_3PO_4 - 98,00 g/mol (Synth);
- Ácido α -ciano-4-hidroxicinâmico (CHCA), $C_{10}H_7NO_3$ – 189,17 g/mol (Sigma-Aldrich);
- Acrilamida, C_3H_5NO - 71,08 g/mol (USB);
- Agarose (BioAgency);
- Albumina de soro bovino (Merck);
- Anfólitos de pH 3-10 (Amersham Biosciences);
- Azul de Bromofenol, $C_{19}H_9Br_4NaO_5S$ - 691,94 g/mol (BioAgency);
- Azul de coomassie G-250, $C_{47}H_{50}N_3NaO_7S_2$ - 854,03 g/mol (J. T. Baker);
- Bicarbonato de amônio, $(NH_4)HCO_3$ – 79,06 g/mol (Mallinckrodt);

- Cloreto de potássio, KCl - 74,55 g/mol (Merck);
- Ditioneitol (DTT), $C_4H_{10}O_2S_2$ - 154,24 g/mol (Amersham Biosciences);
- Dodecil sulfato de sódio (SDS), $C_{12}H_{25}NaO_4S$ - 288,28 g/mol (Synth);
- Etanol, C_2H_6O - 46,07 g/mol (J. T. Baker);
- Éter de petróleo, p.e. 35-60 °C (Mallinckrodt);
- Fluoreto de fenilmetanosulfonila (PMSF), $C_7H_7FO_2S$ - 174,19 g/mol (Sigma-Aldrich);
- Glicerol, $C_3H_8O_3$ - 92,09 g/mol (Amersham Biosciences);
- Glicina, $C_2H_5O_2N$ - 75,07 (Amersham Biosciences);
- Iodoacetamida, C_2H_4INO - 184,96 g/mol (Amersham Biosciences);
- Metanol, CH_4O - 32,04 g/mol (J. T. Baker);
- N,N'-metilenobisacrilamida (bisacrilamida), $C_7H_{10}N_2H_2$ - 154,17 g/mol (Amersham Biosciences);
- N,N',N,N'-tetrametiletenodiamina (TEMED), $C_6H_{16}N_2$ - 116,20 g/mol (J. T. Baker);
- Óleo mineral (Amersham Biosciences);
- Padrão protéico de massa molar (Amersham Bioscience);
- Persulfato de amônio, $N_2H_8S_2O_8$ - 228,20 g/mol (Mallinckrodt)
- Sulfato de amônio, $N_2H_8SO_4$ - 132,14 g/mol (J. T. Baker)
- Tiouréia, CH_4N_2S - 76,12 g/mol (Amersham Bioscience);
- Tripsina (Sigma-Aldrich);
- Tris-hidroximetilaminometano, $C_4H_{11}O_3N$ - 121,14 g/mol (Amersham Biosciences);
- Uréia, CH_4N_2O - 60,06 g/mol (Amersham Bioscience);

3.3 Procedimento Experimental

3.3.1 Amostras de soja

As amostras de sementes de soja transgênica (MSOY 7575 RR) e não-transgênica (MSOY 7501) foram gentilmente cedidas pela Prof^a. Dra. Siu Mui Tsai (CENA/USP – Piracicaba, SP), sendo estas previamente adquiridas da empresa Monsanto – Brasil [31]. A soja transgênica utilizada neste trabalho é aquela denominada *Roundup Ready*, cujo processo de modificação genética consiste na inserção do gene *CP4 EPSPS* da bactéria *Agrobacterium* sp. tipo CP4, o qual confere resistência da planta a herbicidas que contenham glifosato como princípio ativo.

Como neste trabalho as amostras em estudo incluem organismos geneticamente modificados é informado o número de registro do Certificado de Qualidade em Biossegurança: 240/2007, publicado em 24 de julho de 2007.

3.3.2 Extração das proteínas de soja

O procedimento de extração das proteínas foi feito de acordo com o método proposto por Sussulini *et al* [32]. As sementes de soja foram congeladas em nitrogênio líquido e maceradas com o auxílio de um almofariz e um pistilo. Para cada 100 mg de soja moída, adicionou-se 1 mL de éter de petróleo, e, então, o solvente foi deixado em contato com a amostra durante 15 minutos, sob agitação. O solvente foi removido com o auxílio de uma pipeta e a amostra foi novamente macerada com nitrogênio líquido. Este processo foi repetido por mais duas vezes. A seguir, as proteínas foram extraídas da amostra pela maceração desta com 1 mL de uma solução contendo Tris-HCl 50 mmol L⁻¹ (pH 8,8), cloreto de potássio

1,5 mmol L⁻¹, DTT 10 mmol L⁻¹, PMSF 1,0 mmol L⁻¹ e SDS 0,1 % (m/v). O extrato protéico obtido foi, então, agitado por 10 min em banho de gelo, e os materiais insolúveis foram removidos por centrifugação durante 5 min, a 5000 *g* e a 4 °C. O sobrenadante, contendo as proteínas, foi coletado em tubos do tipo Eppendorf e armazenado a -20 °C.

3.3.3 Separação das proteínas por 2D-PAGE

Antes da obtenção do gel de eletroforese bidimensional, com o intuito de remover possíveis interferentes no processo eletroforético (como sais e surfactantes aniônicos), o extrato protéico foi submetido a uma etapa de precipitação das proteínas, empregando solução de acetato de amônio 0,1 mol.L⁻¹ em metanol durante 12 h, a -20 °C. O precipitado protéico foi coletado por centrifugação a 5000 *g* (4 °C) por 10 min e, em seguida, lavado duas vezes com solução gelada de acetato de amônio 0,1 mol L⁻¹ em metanol, duas vezes com acetona gelada, na concentração de 80 % (v/v) e uma vez com etanol a 70 % (v/v), também gelado.

Para a obtenção dos géis de eletroforese bidimensional, o precipitado protéico foi ressolubilizado em tampão que consistiu de uma solução de uréia 7 mol L⁻¹, tiouréia 2 mol L⁻¹, CHAPS 2 % (m/v), anfólitos 0,5 % (v/v) em pH variando de 3 a 10 e azul de bromofenol 0,002 % (m/v). Em seguida, foi feita a centrifugação da solução resultante por 2 min a 5000 *g*, de modo a remover qualquer resíduo sólido que ainda permanecesse na amostra. O precipitado protéico foi quantificado usando o método de Bradford [33].

A primeira dimensão do processo eletroforético foi executada empregando-se uma fita de 13 cm, sendo que esta fita continha um gel de poliacrilamida onde o gradiente de pH foi formado pela presença de anfólitos. Neste trabalho, foram utilizadas duas faixas lineares de pH: 3 a 10 e 4 a 7. Antes da focalização isoeletrica, foram aplicados 250 µL da amostra ressolubilizada, contendo 300 µg

de proteínas para fitas de 3 a 10 e 500 µg de proteínas para fitas de 4 a 7, em um aparato sobre o qual a fita foi colocada. Esta foi reidratada a temperatura ambiente por 12 h. Após este período, a fita foi levada ao sistema focalizador e o programa com quatro etapas foi aplicado: (1) 500 V até 500 Vh, (2) 1000 V até 800 Vh, (3) 10000 V até 11300 Vh e (4) 10000 V até 3000 Vh, para fitas de 3 a 10, e, em fitas de 4 a 7, o programa adotado foi o mesmo, com exceção da última etapa, onde foi aplicado 10000 V até 2000 Vh.

Ao término da focalização, a fita foi equilibrada em duas etapas. Na primeira delas, foram usados 10 mL de solução contendo uréia 6 mol L⁻¹, SDS 2 % (m/v), glicerol 30 % (v/v), Tris-HCl 50 mmol L⁻¹ (pH 8,8), azul de bromofenol 0,002 % (m/v) e DTT 2 % (m/v). Esta etapa é denominada de redução e tem como finalidade manter as proteínas em suas formas reduzidas [23]. Na segunda etapa, foi utilizada uma solução de composição similar, porém, o DTT foi substituído por iodoacetamida 2,5 % (m/v). Esta etapa é denominada de alquilação e tem como finalidade alquilar os grupos tióis das proteínas, prevenindo que elas sejam reoxidadas durante o processo eletroforético [23]. Cada uma das etapas durou 15 min e foi feita sob leve agitação.

Após o equilíbrio da fita, foi feita a segunda dimensão do processo eletroforético (SDS-PAGE). Para isso, a fita foi aplicada em um gel de poliacrilamida 12,5 % (m/v) previamente preparado em placa de 180 x 160 x 1,5 mm. Foi colocado sobre o gel, ao lado da fita, um pedaço de papel de filtro contendo 8 µL de um padrão de massa molar contendo as proteínas β-fosforilase (97,0 kDa), albumina (66,0 kDa), ovalbumina (45,0 kDa), anidrase carbônica (30,0 kDa), inibidor de tripsina (20,1 kDa) e α-lactoalbumina (14,4 kDa). A fita e o papel de filtro foram vedados com uma solução quente de agarose 0,5 % (m/v), de modo a garantir o contato destes com o gel de poliacrilamida. A corrida eletroforética foi executada em duas etapas: (1) 90 V, 25 mA gel⁻¹ e 100 W, durante 30 min e (2) 250 V, 25 mA gel⁻¹ e 100 W, durante, aproximadamente, 5 h. Após o término desta etapa, as proteínas foram fixadas no gel, durante 1 h, empregando-se uma solução contendo ácido acético 10 % (v/v) e etanol 40 % (v/v). Em seguida, as proteínas foram reveladas empregando-se o corante Coomassie coloidal, que

consistiu em uma solução de sulfato de amônio 8 % (m/v), ácido fosfórico 1,6 % (v/v), CBB G-250 0,08 % (m/v) e metanol 25 % (v/v). O corante ficou em contato com o gel durante 48 h e, depois, foi removido por lavagens com água desionizada.

O gel obtido foi escaneado e sua imagem foi analisada pelo programa de tratamento de imagens *ImageMaster 2D Platinum*, versão 6.0, que permite obter uma estimativa dos pontos isoeletricos e das massas molares das proteínas separadas, bem como calcular o número de *spots* obtidos no gel de eletroforese.

3.3.4 Digestão proteolítica e identificação das proteínas por Espectrometria de massas

Inicialmente, os *spots* de proteínas a serem caracterizados por espectrometria de massas foram recortados do gel de poliácridamida e cada um deles foi colocado em um pocinho de uma placa, denominada Zip Plate (Millipore), na qual foi feita a digestão trípica. A digestão trípica foi feita com o uso de um kit específico (*In-Gel Digest_{ZP} Kit*) que, além de digerir a proteína, purifica os peptídeos obtidos.

Para as análises por MALDI, as amostras resultantes da digestão trípica foram aplicadas em uma microplaca de MALDI, e esta foi deixada à temperatura ambiente por alguns minutos, de modo que o volume das gotas fosse reduzido por evaporação (método *dried droplet*). A matriz CHCA (ácido α -ciano-4-hidróxicinâmico) foi preparada adicionando 10 mg de ácido em uma mistura de água/acetonitrila 1:1 (v/v) contendo TFA 0,1 % (v/v) e, então, 2 μ L desta solução foi adicionada a 2 μ L amostra. A amostra foi aplicada na placa de MALDI e deixou-se secar a temperatura ambiente.

As caracterizações das proteínas foram feitas empregando dois modos de obtenção de dados: em um único estágio, usando MALDI-MS e em modo seqüencial, usando MALDI-MS/MS e ESI-MS/MS. Para as análises por MALDI, os

espectros foram adquiridos no equipamento Q-Tof Premier (Waters Co.). As medidas foram feitas no modo íon positivo, com fonte de íons nitrogênio, no modo “V”, empregando os seguintes parâmetros: frequência do laser – 200 Hz, faixa de massa 800-3000 Da, com o mínimo de 40 contagens para a obtenção dos espectros MS/MS, e estes com o mínimo de 10 contagens para registro.

Para as análises por ESI, os peptídeos obtidos foram secos, e posteriormente, ressuspensos em solução de ácido fórmico 0,1 % (v/v). Os espectros foram adquiridos em um equipamento Q-Tof Ultima (Waters Co.). Uma alíquota de 5,0 µL dos peptídeos resultantes foram, inicialmente, separados por um RP-nanoUPLC (nanoAcquity, Waters) em uma coluna C18 de fase reversa (75 µm x 100 mm) acoplado no espectrômetro de massas com uma fonte de nano-electrospray a um fluxo de 0,6 µL.min⁻¹. Os peptídeos foram eluídos em um gradiente de 2-90% de acetonitrila em ácido fórmico 0,1% (v/v) por 20 min. Os espectros de massas foram adquiridos no modo “*top three*”, no qual um espectro no modo MS é adquirido seguido por espectros de massas no modo MS/MS dos três picos mais abundantes detectados.

A identificação das proteínas foi feita utilizando os dados gerados pelo espectrômetro de massas juntamente com o programa *Mascot*. Os espectros obtidos no formato “pkl” foram processados por meio do programa *Mascot Distiller*, os quais foram convertidos no formato “mgf”, compatível para o processo de identificação. As buscas foram feitas em bancos de dados não-redundantes, usando os seguintes parâmetros: bancos de dados MSDB e NCBI, os quais contem proteínas gerais e com seqüências não-idênticas, e *Glycine Max*, o qual é específico para proteínas de soja; um sítio de clivagem; digestão com tripsina; carbamidometilação como modificação fixa; oxidação em metionina como modificação variável e erro, em massa, de 0,1 Da.

3.3.5 Identificação da enzima *CP4 EPSPS* por SDS-PAGE e espectrometria de massas

Antes da obtenção do gel de eletroforese unidimensional (SDS-PAGE) e com o intuito de remover possíveis interferentes no processo eletroforético (como sais e surfactantes aniônicos), o extrato protéico foi submetido a uma etapa de precipitação das proteínas, empregando solução de acetato de amônio $0,1 \text{ mol.L}^{-1}$ em metanol durante 12 h, a -20°C . O precipitado protéico foi coletado por centrifugação a 5000 g (4°C) por 10 min, e, em seguida, lavado duas vezes com solução gelada de acetato de amônio $0,1 \text{ mol L}^{-1}$ em metanol, duas vezes com acetona gelada, na concentração de 80 % (v/v) e uma vez com etanol a 70 % (v/v), também gelado.

Para a obtenção do gel de eletroforese unidimensional, o precipitado protéico foi ressolubilizado em tampão dissociante composto por Tris-HCl pH 6,8, 10 % (m/v) de glicerol, 20 % (v/v) de SDS, 5 % (v/v) de β -mercaptoetanol e 0,001 % (m/v) de azul de bromofenol. Como padrão protéico, utilizou-se uma mistura, disponível comercialmente, composta de 7 proteínas purificadas: β -galactose de *E. Coli* (116,0 kDa); albumina bovina (66,2 kDa); ovalbumina (45,0 kDa); lactato desidrogenase (35,0 kDa); enzima de restrição endonuclease Bsp98I de *E. coli* (25,0 kDa); β -lactoglobulina (18,4 kDa) e lisozima (14,4 kDa). Para aplicação no gel, o padrão foi submetido a aquecimento em banho com água a 95°C por 5 min, para promover a desnaturação das proteínas.

O gel de eletroforese unidimensional, para a separação das proteínas, consistia de um gel concentrador de 1,5 cm com concentração de 3,5 % (m/m) de acrilamida e tampão Tris-HCl a $1,0 \text{ mol L}^{-1}$ pH 6,8 e outro gel separador a 10 % (m/m) de acrilamida e tampão Tris-HCl a $1,5 \text{ mol L}^{-1}$ pH 8,8. Após polimerização completa, diferentes valores em massa de proteínas totais (25, 50, 75 e $100 \mu\text{g}$) foram aplicadas na região superior do gel concentrador. A voltagem aplicada na corrida eletroforética bem como o modo de revelação das proteínas no gel foi o mesmo como descrito anteriormente.

Após a conclusão da revelação das bandas, o gel obtido foi digitalizado e as bandas obtidas analisadas pelo software Gel-Pro Analyser 3.0 para estimativa da massa molar de cada banda. A banda correspondente a massa de 47 kDa (enzima *CP4 EPSPS*) foi recortada, reduzida, alquilada e submetida a digestão enzimática usando duas enzimas de clivagem: tripsina e quimiotripsina.

Os peptídeos obtidos foram secos, e posteriormente ressuspensos em solução de ácido fórmico 0,1 % (v/v). A obtenção dos espectros de MS/MS bem como o modo de identificação das proteínas foram semelhantes a aqueles descritos no item anterior (3.3.4), com exceção do tempo de corrida na obtenção dos espectros, os quais foi de 40 minutos.

4. RESULTADOS E DISCUSSÃO

4.1 Identificação das proteínas por MALDI-QTOF-MS e ESI-QTOF-MS

Em um trabalho recentemente publicado por Brandão et al, 2010 [34], foram avaliados os perfis eletroforéticos entre amostras de soja transgênica e não transgênica, usando a técnica de 2D-PAGE e análise das imagens dos géis obtidos. Os resultados obtidos neste trabalho mostraram que o número de *spots* protéicos presentes na faixa linear de 4 a 7 foi maior (369 ± 46) em comparação com a faixa linear de 3 a 10 (217 ± 41) para sementes de soja transgênica, concluindo que esta faixa mostrou melhor resolução na separação das proteínas. Neste sentido, para este capítulo, um maior número de *spots* para a faixa de 4 a 7 foram selecionados e submetidos ao processo de identificação em relação a faixa de 3 a 10, onde, nesta faixa, em específico, foram selecionados *spots*, preferencialmente localizados na faixa de pH de 7 a 10, para o processo de identificação por espectrometria de massas.

No presente trabalho de Tese, empregando a técnica de 2D-PAGE como componente de separação e espectrometria de massas, do tipo MALDI-QTOF e ESI-QTOF como componente estrutural, foi possível identificar um total de 192 proteínas de sementes de soja transgênica, sendo que, deste total, 179 proteínas foram identificadas para a faixa de pH de 4 a 7 e 13 proteínas para a faixa de 3 a 10. A tabela com todas as proteínas identificadas esta localizada nos anexos desta tese. Estes resultados evidenciam que o perfil de distribuição das proteínas das sementes de soja está diretamente mais localizado em faixa de pH de 4 a 7 [34]. Os mapas proteômicos nas duas faixas de pH podem ser visualizados na Figuras 1.2.

Relacionado ao processo de identificação usando o programa *Mascot*, a busca em banco de dados de vários *spots* gerou a identificação de, pelo menos, mais de uma proteína. Sendo assim, o critério de corte utilizado para a seleção da possível proteína relacionada ao *spot* foi aquela que mostrou maior pontuação (*score*).

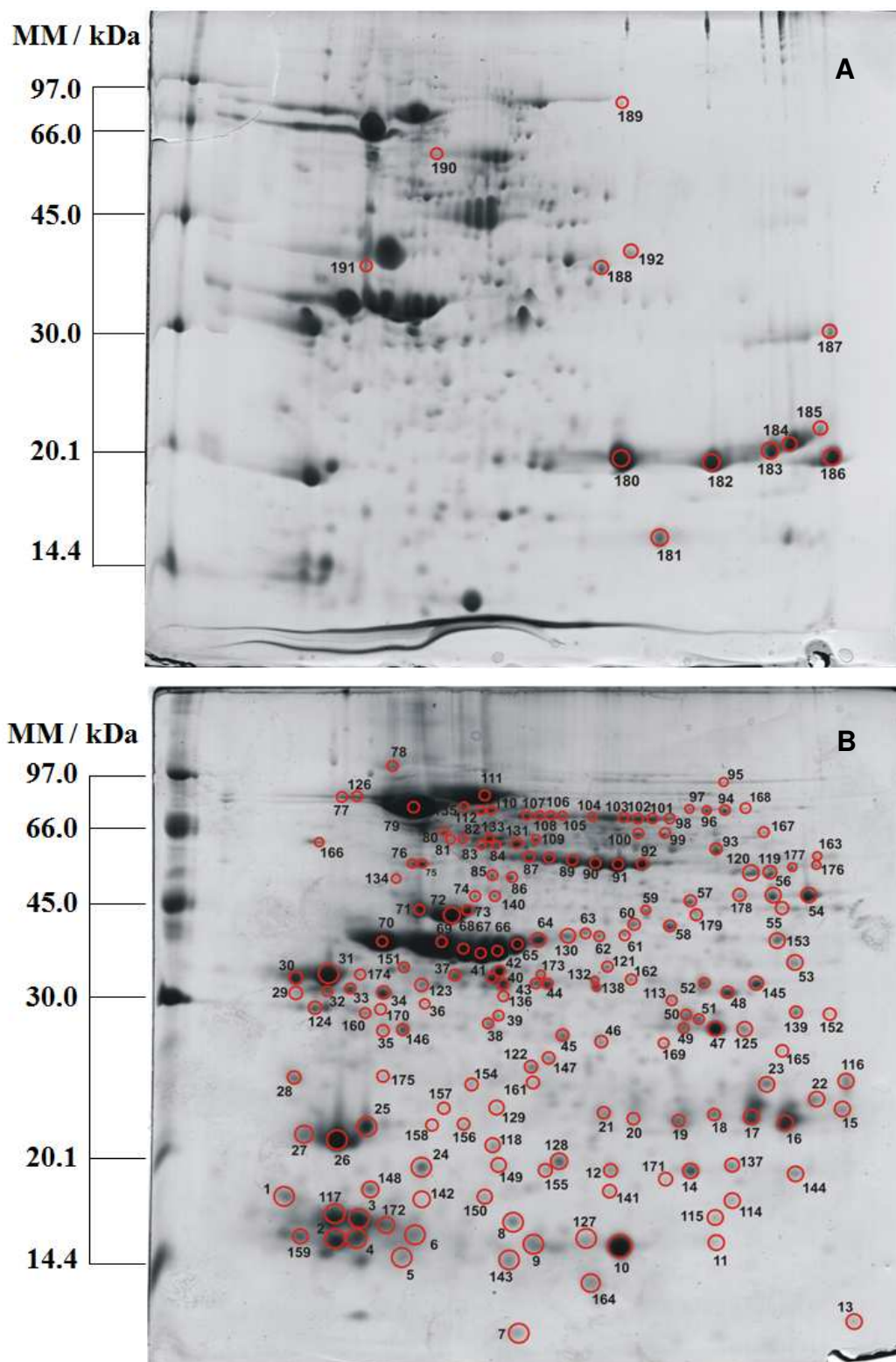


Figura 1.2. Proteínas identificadas para sementes de soja transgênica: na faixa de pl de 3 a 10 (A), correspondendo a um total de 13 proteínas identificadas e na faixa de pl de 4 a 7 (B), correspondendo a um total de 179 proteínas identificadas.

Um ponto importante a se considerar quando o objetivo é a elaboração de um mapa proteômico relaciona-se com o percentual de eficiência obtido no processo de identificação. Neste trabalho, foi obtido um percentual de 49 % relacionando-se o número de *spots* detectados no gel de eletroforese bidimensional e *spots* identificados em banco de dados. Este valor percentual obtido neste trabalho apresenta uma boa eficiência quando comparado com estudos similares, envolvendo soja, publicados na literatura. No trabalho de Xu *et al* [15], foi obtida uma eficiência de 46 % na identificação de proteínas de folhas de soja usando eletroforese bidimensional em gel de poliacrilamida (2D-PAGE) e espectrometria de massas (MALDI-TOF-MS e LC-MS/MS). Melhores resultados foram obtidos no trabalho de Brechenmacher *et al* [35] no estabelecimento de um mapa proteômico de referência para células presentes nas raízes capilares de plantas de soja, obtendo-se uma eficiência de 57 % usando 2D-PAGE e espectrometria de massas. Entretanto, ressalta-se que, neste trabalho, o número de *spots* detectados no gel foi de 1165, o que pode acarretar em maiores possibilidades no processo de identificação.

Os resultados de identificação obtidos neste trabalho mostram a grande eficiência na utilização conjunta das técnicas de 2D-PAGE como componente de separação e espectrometria de massas como componente estrutural no estabelecimento de mapas proteômicos de referência. Entretanto, o número de proteínas identificadas poderia ser maior, levando em consideração algumas limitações associadas a técnica de 2D-PAGE, tais como restrição de faixa de massa molecular na separação (proteínas de alta e baixa massa molecular), proteínas muito ácidas ou básicas, proteínas hidrofóbicas e com baixa abundância [21,36,37].

Sendo assim, a utilização da técnica de 2D-PAGE em conjunto com metodologias baseadas em cromatografia multidimensional associadas a espectrometria de massas (LC-MS/MS) podem fornecer resultados bastante interessantes no estabelecimento de mapas proteômicos de referência em amostras biológicas complexas [38].

De forma a melhor avaliarmos o conjunto de proteínas identificadas neste trabalho, foi feita uma classificação funcional das proteínas de sementes de soja transgênica, a qual será abordada no tópico a seguir.

4.2 Avaliação funcional das proteínas identificadas.

As proteínas identificadas foram classificadas de acordo com suas atividade biológicas em categorias, de acordo com o trabalho de Bevan *et al* [39], e são mostradas na Figura 1.3. Em seguida, serão discutidas as principais classes de proteínas identificadas neste trabalho.

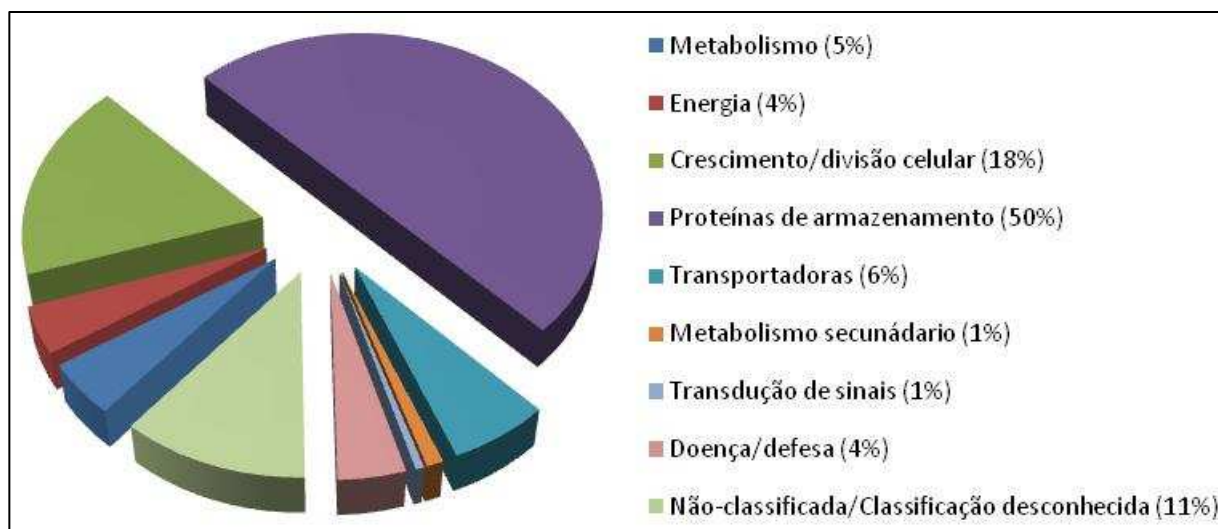


Figura 1.3. Distribuição funcional das proteínas de soja transgênica identificadas.

4.2.1 Armazenamento e atividade proteolítica

Neste estudo, 50 % (96 *spots*) das proteínas identificadas estão relacionadas com função de armazenamento e atividade proteolítica. Nesta classe, incluem-se as proteínas majoritárias glicinina e β -conglucina e suas sub-

unidades, as quais são componentes predominantes nas sementes de soja e representam as principais proteínas de reserva da soja [14].

A proteína 11S globulina, comumente conhecida como glicinina, é a proteína de reserva predominante entre as globulinas da soja [14]. A glicinina nativa é um hexâmero de massa molar estimada entre 320 e 360 kDa, composto de diferentes combinações de 5 subunidades: G1 a G5, sendo que cada uma das subunidades, as quais apresentam massa molar entre 58-69 kDa, pode ser dissociada, dentro de condições redutoras, em cadeias polipeptídicas ácidas (A: 31-45 kDa) e básicas (B: 18-20 kDa), os quais são ligadas por meio de pontes dissulfeto internas. As cinco subunidades identificadas e seus constituintes polipeptídicos caracterizados são: $A_{1a}B_2$ (G1), $A_{1b}B_{1b}$ (G2), A_2B_{1a} (G3), A_3B_4 (G4) (Figura 1.4) e $A_5A_4B_3$ (G5) [40,41]. Neste trabalho de tese, todas as cinco variedades foram identificadas, sendo um total de 20, 03, 09, 06 e 11 *spots* para as subunidades G1, G2, G3, G4 e G5 respectivamente.

A β -conglícinina, uma globulina 7S, é uma glicoproteína de massa molar que varia entre 140 e 170 kDa. É um trímero composto por três subunidades: α (57-72 kDa) (Figura 1.4), α' (57-68 kDa) e β (42-52 kDa) [40,42]. Todas estas subunidades foram identificadas neste trabalho, sendo um total de 16, 01 e 09 *spots* para as subunidades α , α' e β respectivamente.

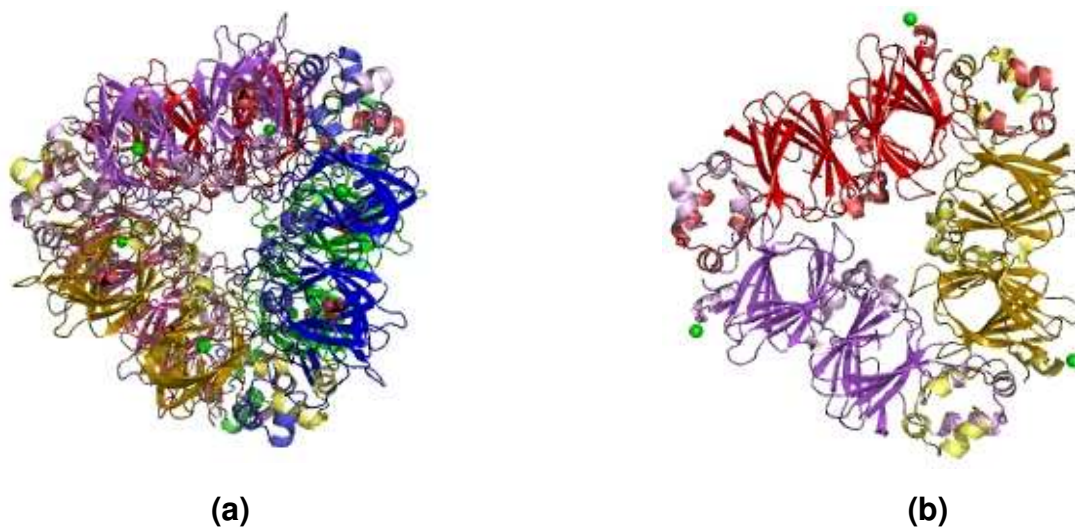


Figura 1.4. Estrutura tridimensional da proteína: (a) glicinina sub-unidade A_3B_4 (G4); (b) sub-unidade α da β -conglícinina [79].

Entretanto, neste ponto, destaca-se que, devido ao elevado percentual para esta classe de proteínas, pode haver certa limitação na resolução dinâmica de separação das proteínas presentes, já que, possivelmente, podem haver perdas na detecção de proteínas pouco abundantes presentes nas sementes de soja transgênica [43].

Outras proteínas pertencentes a esta classificação funcional pertencem a duas classes de inibidores da protease: Kunitz e Bowman-Birk. Com relação a inibidor tipo Kunitz (Kti), foi possível identificar um total de 05 *spots* (1, 25-27 e 158). Esta proteína possui massa molar de aproximadamente 21 kDa, e atua como inibidor da atividade da enzima tripsina [44]. A proteína do tipo Bowman-Birk está inserida dentro de uma classe de inibidores de proteases o qual compreendem várias proteínas relacionadas, sendo que estas são específicas para tripsina, quimiotripsina e elastase [44]. Desta proteína, foram identificados 05 *spots* (4-6, 117 e 164).

4.2.2 Crescimento/Divisão Celular

Após as proteínas de armazenamento, a segunda classe funcional de proteínas mais abundantes são as proteínas envolvidas no processo de crescimento/divisão celular, o qual corresponde a 18 % (34 *spots*) do total de proteínas identificadas. Nesta classe, a principal proteína é a *late embryogenic abundant protein* (LEA), sendo esta bastante acumulada nos últimos estágios de embriogenese da semente [45]. Além disso, esta proteína está relacionada a processos de resposta ao estresse osmótico e na deficiência de água resultante dos estágios de dessecação da semente [45,46].

As proteínas do tipo LEA estão inseridas dentro de um grupo de famílias com diversas estruturas e funções [46]. Sendo assim, duas proteínas e suas subclasses foram identificadas neste grupo: *late embryonic abundant protein* (11 *spots*: 08, 48, 52, 94, 96, 97, 115, 118, 129, 167 e 168) e *seed maturation protein*

(14 *spots*). Desta proteína, foram identificadas 09 variedades, sendo elas: PM 22 (*spot* 24), PM24 (*spot* 33), PM25 (*spot* 34), PM31 (*spots* 12 e 14), PM 32 (*spots* 154 e 157), PM 34 (*spot* 53), PM36 (*spot* 38), 35 KDa *seed maturation protein* (*spots* 58, 60 e 138) e 51 kDa *seed maturation protein* (*spot* 163).

Outra proteína pertencente ao grupo de proteínas tipo LEA é *dehydrin* (*spots* 46, 47, 49, 50, 113, 125, 128 e 161). Esta proteína está envolvida em mecanismos de resposta ao stress abiótico, como desidratação e baixas temperaturas [46,47]. Ainda dentro desta categoria funcional, outra proteína identificada é a *Class III acidic endochitinase precursor* (*spot* 124). Esta proteína pertence à família das *Chitinases* e é especificamente expressa durante o estágio de desenvolvimento da semente de soja [48].

4.2.3 Metabolismo/Energia

Cerca de 17 proteínas identificadas (9%) estão envolvidas em processos metabólicos/energéticos. De forma geral, estas proteínas participam no metabolismo de carboidratos e degradação de polissacarídeos [46]. Como exemplos, pode-se citar: *Agglutinin* (*spots* 37, 41, 42, 44 e 173), *Triosephosphate isomerase* (*spot* 51), *Glyceraldehyde-3-phosphate dehydrogenase* (*spots* 55 e 118), *Cytosolic malate dehydrogenase* (*spot* 59), *Cytosolic glutamine synthetase* (*spot* 74), *Beta-amylase* (*spots* 84 e 131), *Enolase* (*spot* 108) e *Formate dehydrogenase* (*spots* 119 e 120).

Outra proteína pertencente a esta categoria funcional é a *lipoxygenase* (*spot* 95) (Figura 1.5), uma enzima ligante de ferro que catalisa a hidroperoxidação de lipídios [46]. Esta enzima está envolvida em vários aspectos da fisiologia da planta como crescimento e desenvolvimento, resistência a pragas e resposta a ferimentos na planta [46].

4.2.4 Transporte

Aproximadamente 6 % (12 *spots*) se relacionam com proteínas de transporte. Nesta categoria funcional, apenas uma enzima foi relacionada, sendo ela a *sucrose-binding protein* (*spots* 93, 98, 101-109, 133 e 190). Esta enzima é responsável pelo transporte de sacarose, durante o estágio de desenvolvimento da semente [46]. Além disso, esta enzima também pode atuar como reserva de nutrientes para a planta [46].

4.2.5 Doença/Defesa

Cerca de 4% (08 *spots*) correspondem a proteínas envolvidas dentro desta categorial funcional, sendo que a maior parte dos *spots* detectados pertencem a enzima *Alcohol dehydrogenase* (*spots* 54, 56, 57 e 132). Esta enzima é ligante de zinco e está envolvida na atividade catalítica de reações de oxido-redução [46]. Outra enzima pertencente a esta classe é a *Mn-Superoxide dismutase* (*spot* 165) (Figura 1.5), sendo esta uma metaloenzima antioxidante atuando em mecanismos de defesa da planta contra espécies reativas de oxigênio (ROS) [49].

Outras proteínas pertencentes a esta classe são: *In 2-1 protein* (*spot* 39), *18.5 kDa class I heat shock protein* (*spot* 137), e *glutathione S-transferase* (*spot* 169).

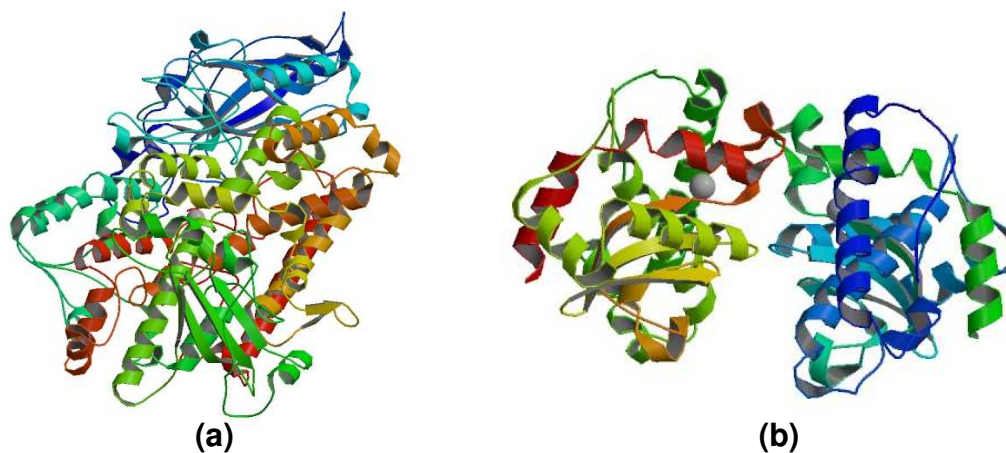


Figura 1.5. Estrutura tridimensional da proteína: (a) *lipoxigenase-3*; (b) *Mn-Superoxide dismutase* [79].

4.2.6 Transdução de sinais / Metabolismo secundário

Apenas uma proteína foi relacionada a categoria funcional de transdução de sinais, a enzima *Nucleoside diphosphatase kinase* (spot 114). Esta enzima, ligante de ATP e magnésio, está envolvida em reações bioquímicas de nucleotídeos, principalmente na transferência de fosfato, usualmente de moléculas de ATP, para substratos secundários [46]. A enzima *Isoflavone reductase homolog 2* (spots 61 e 62) foi classificada na categoria de metabolismo secundário, e está envolvida na biosíntese de isoflavonoides, sendo que estes compostos possuem importante papel na planta, durante os estágios de crescimento, desenvolvimento, nodulação e na defesa contra microorganismos e estresse abiótico [50].

4.2.7 Não-classificada / Classificação desconhecida

Dentro da categoria de proteínas não-classificadas, foram identificadas um total de 21 proteínas. Destas, é dado destaque às proteínas *Gly m Bd 30K allergen* (spots 30 e 31) e *Allergen Gly m Bd 28K* (spots 45 e 116). Estas proteínas são conhecidas por serem as mais alérgicas em indivíduos que mostram elevada sensibilidade ao consumo de derivados de soja [51]. Além destas, foram identificadas as proteínas *Hypothetical protein related* (spots 121 e 139), *PreproMP73 related* (spot 122) e *unknown* (spots 140, 144, 145, 147-150, 152, 153, 155, 156, 166, 171 e 178). A proteína *glycine-rich RNA-binding protein* (spots 142 e 150) foi inserida na categoria classificação desconhecida. Esta proteína é assim denominada por exibir uma elevada concentração de glicina no C-terminal da proteína [52].

4.3 Identificação da enzima *CP4 EPSPS*

A partir do mapa proteômico estabelecido para a semente de soja, esperava-se que a enzima *CP4 EPSPS* pudesse ter sido identificada, mostrando assim que o organismo em questão passou pelo processo de modificação genética. Na literatura, foi encontrado apenas um trabalho relacionado a utilização da técnica de 2D-PAGE e espectrometria de massas para identificação da enzima *CP4 EPSPS* em sementes de soja, tipo *Roudap Ready*. No trabalho de Kim *et al* [7], o perfil eletroforético do gel 2D-PAGE obtido para sementes de soja transgênica mostrou-se bastante diferente em relação ao gel obtido neste trabalho, em termos de procedimento de extração, massa de proteínas aplicada no gel e, principalmente, em termos de número de spots protéicos identificados no gel, impossibilitando uma comparação com os resultados obtidos nesta tese.

Para a identificação da enzima, os autores fizeram um gel 2D-PAGE da enzima *CP4 EPSPS* purificada do microrganismo *Escherichia coli* clonada como controle. Levando em consideração o gel obtido neste trabalho e com os valores do pI e a massa moléculas da enzima, que são 5,1 e 47,0 kDa, supõe-se que sua possível posição no gel, (Figura 1.6) seria correspondente a região onde estão localizadas *spots* bastante intensos de sub-unidades da glicinina (*spots* 71,72 e 73). Cabe ressaltar que possíveis modificações pós-traducionais podem ter contribuído na diferença de posição do *spot* em ambos os géis, já que estas modificações podem alterar os valores de pI e/ou massa moléculas em relação aos valores teóricos .

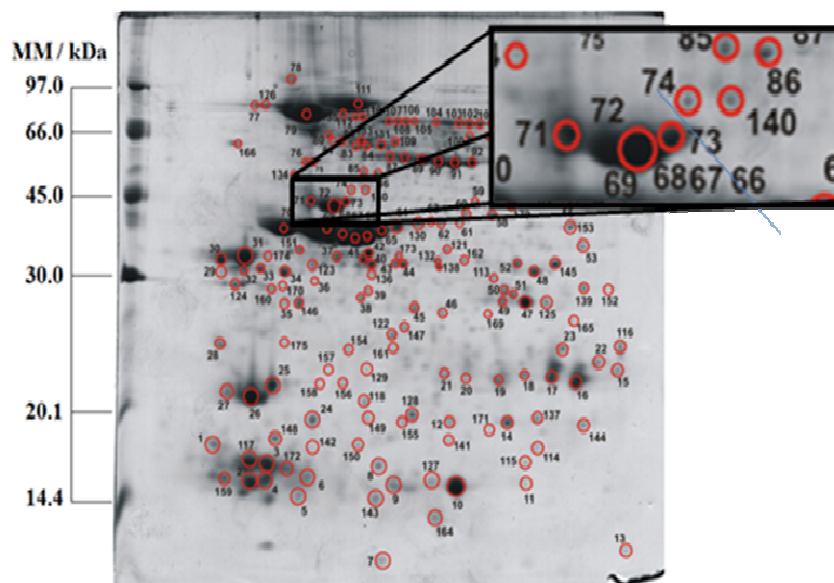


Figura 1.6. Possível localização do *spot* no gel 2D-PAGE, (pl 4-7) correspondente a enzima *CP4 EPSPS*, de acordo o gel obtido neste trabalho.

Sendo assim, supõe-se que a elevada concentração destas sub-unidades protéicas possa ter sido fator limitante para a identificação da enzima usando a técnica de 2D-PAGE e MALDI-QTOF MS.

Para a identificação da enzima *CP4 EPSPS*, foi feito um gel de eletroforese unidimensional (SDS-PAGE) aplicando-se diferentes massas de proteínas de sementes de soja solubilizadas em tampão dissociante. O perfil eletroforético do gel obtido pode ser visualizado na Figura 1.7. A partir deste gel, foi selecionada uma banda de *ca.* 47 kDa e, posteriormente submetida a digestão proteolítica usando duas enzimas de clivagem: a tripsina, com especificidade de clivagem em ligações peptídicas contendo os aminoácidos lisina e arginina, e a quimiotripsina, com especificidade de clivagem em ligações peptídicas dos aminoácidos fenilalanina, tirosina, triptofano e leucina. A opção pela escolha das duas enzimas foi para aumentar a probabilidade de identificação da enzima usando a técnica de ESI-QTOF MS.

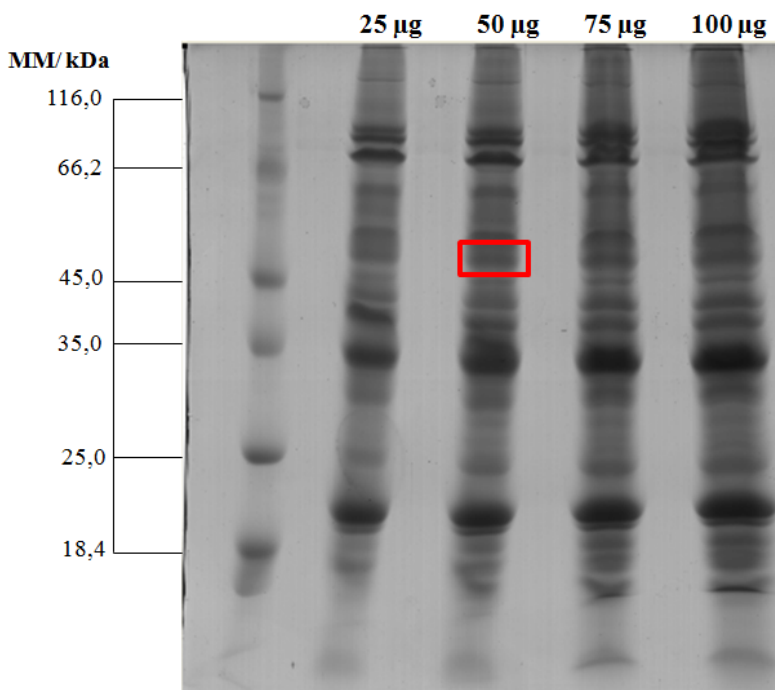


Figura 1.7. Perfil eletroforético obtido para a semente de soja transgênica usando diferentes quantidades em massa de proteínas. Em destaque (vermelho) a banda protéica cortada e submetida a digestão enzimática e posterior análise por ESI-QTOF MS.

Com este procedimento foi possível identificar a enzima *CP4 EPSPS*, utilizando as duas enzimas avaliadas, ou seja, a tripsina e a quimiotripsina. Dentre as duas enzimas de clivagem, a tripsina mostrou os melhores resultados, em termos de pontuação, bem como no percentual de cobertura. Ressalta-se que esta metodologia empregada mostra-se simples e bastante eficiente, já que, em trabalho semelhante reportado na literatura [19], houve a necessidade de pre-fracionamento da amostra usando duas colunas cromatograficas (exclusão por tamanho e troca iônica). Neste trabalho, o gel de SDS-PAGE foi feito diretamente com o extrato de proteínas e a banda correspondente a massa da enzima foi digerida diretamente, sem necessidade de pre-fracionamento para obtenção do gel. A sequência de aminoácidos para a enzima *CP4 EPSPS*, identificada usando a enzima tripsina, está mostrada na Figura 1.8, onde foram destacados, em vermelho, os peptídeos identificados.

```

1 MLHGASSRPA TARKSSGLSG TVRIPGDKSI SHRSFMFGGL ASGETRITGL
51 LEGEDVINTG KAMQAMGARI RKEGDTWIID GVGNGGLLAP EAPLDFGNAA
101 TGCRLTMGLV GVDYDFDSTFI GDASLTKRPM GRVLNPLREM GVQVKSEDGD
151 RLPVTLRGPK TPTPITYRVP MASAQVKSAV LLAGLNTPGI TTVIEPIMTR
201 DHTEKMLQGF GANLTVETDA DGVRTIRLEG RGKLTGQVID VPGDPSSTAF
251 PLVAALLVPG SDVTILNVLM NPTRTGLILT LQEMGADIEV INPRLAGGED
301 VADLRVRSST LKGVTVPEDR APSMIDEYPI LAVAAFAEG ATVMNGLEEL
351 RVKESDRLSA VANGLKLNGV DCDEGETSLV VRGRPDGKGL GNASGAAVAT
401 HLDHRIAMSF LVMGLVSENP VTVDDATMIA TSFPEFMDLM AGLGAKIELS
451 DTKAA

```

Figura 1.8. Seqüência de aminoácidos da enzima *CP4 EPSPS* usando a enzima tripsina na etapa de digestão enzimática. Em vermelho, são destacados os peptídeos identificados por ESI-QTOF MS (Banco NCBI - *Glycine max*).

Como pode ser visto na Figura 1.8, a clivagem da proteína se dá nas ligações peptídicas contendo os aminoácido lisina (K) e arginina (R). A Figura 1.9 mostra a sequência de aminoácidos da enzima *CP4 EPSPS* utilizando quimiotripsina como enzima de digestão. Como pode ser visto na Figura, a clivagem das ligações peptídicas ocorrem nos aminoácidos fenilalanina (F), tirosina (Y) e, principalmente, em leucina (L). Além disso, poucas sequências peptídicas foram identificadas usando esta enzima de clivagem.

```

1 MLHGASSRPA TARKSSGLSG TVRIPGDKSI SHRSFMFGGL ASGETRITGL
51 LEGEDVINTG KAMQAMGARI RKEGDTWIID GVGNGGLLAP EAPLDFGNAA
101 TGCRLTMGLV GVDYDFDSTFI GDASLTKRPM GRVLNPLREM GVQVKSEDGD
151 RLPVTLRGPK TPTPITYRVP MASAQVKSAV LLAGLNTPGI TTVIEPIMTR
201 DHTEKMLQGF GANLTVETDA DGVRTIRLEG RGKLTGQVID VPGDPSSTAF
251 PLVAALLVPG SDVTILNVLM NPTRTGLILT LQEMGADIEV INPRLAGGED
301 VADLRVRSST LKGVTVPEDR APSMIDEYPI LAVAAFAEG ATVMNGLEEL
351 RVKESDRLSA VANGLKLNGV DCDEGETSLV VRGRPDGKGL GNASGAAVAT
401 HLDHRIAMSF LVMGLVSENP VTVDDATMIA TSFPEFMDLM AGLGAKIELS
451 DTKAA

```

Figura 1.9. Seqüência de aminoácidos da enzima *CP4 EPSPS* usando a enzima quimiotripsina na etapa de digestão enzimática. Em vermelho, são destacados os peptídeos identificados por ESI-QTOF MS (Banco NCBI - *Glycine max*).

A Figura 1.10 mostra os espectros de MS/MS dos peptídeos relativos à enzima *CP4 EPSPS*. Além disso, a Tabela 1.1 mostra dados, como razão m/z observados nos espectros de MS/MS, e, a partir destas, as massas experimentais e calculadas pelo programa *Mascot* relativos aos peptídeos que foram utilizados para a identificação da enzima *CP4 EPSPS*. A somatória dos *scores* individuais de cada peptídeo corresponde ao *score* total para identificação da enzima junto ao banco de dados.

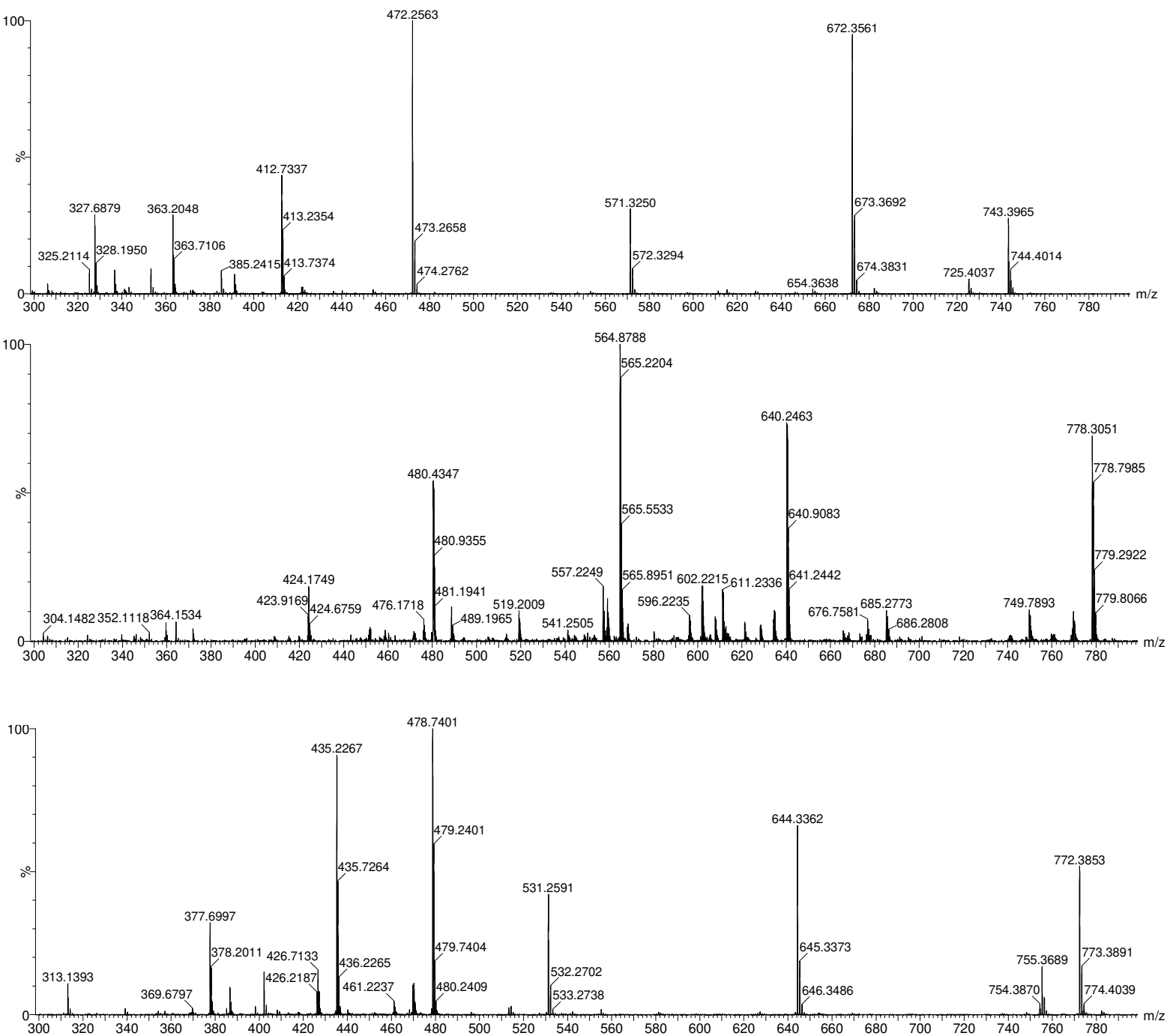


Figura 1.10. Espectros de MS/MS relativos aos peptídeos da enzima *CP4 EPSPS*, obtidos no modo “*top three*” (três picos mais intensos) por ESI-QTOF MS/MS.

Tabela 1.1. Sequência de peptídeos, obtidos pelo *Mascot*, usados na identificação da enzima *CP4 EPSPS* por ESI-QTOF MS (Banco NCBI - *Glycine max*, enzima tripsina).

m/z observada	Massa experimental	Massa calculada	Score	Sequência
356,2377	710,4608	710,4439	22	VLNPLR
395,7168	789,4191	789,4055	32	EMGVQVK
403,2268	804,4390	804,4229	17	IELSDTK
418,2110	834,4073	834,3840	10	AMQAMGAR
432,2354	862,4562	862,4509	25	SSGLSGTVR
436,7231	871,4316	871,4400	35	GVTVPEDR
436,7611	871,5076	871,5127	46	LSAVANGLK
465,7571	929,4997	929,5004	41	VPMASAQVK
474,7486	947,4827	947,5076	50	TPTPITYR
558,2818	1114,5490	1114,5618	86	LAGGEDVADLR
453,2421	1356,7044	1356,6997	44	SEDGDRLPVTLR
680,3288	1358,6430	1358,6289	90	SFMFGGLASGETR
779,9091	1557,8036	1557,8250	117	ITGLLEGEDVINTGK
549,6180	1645,8322	1645,8285	72	GLGNASGA AVATHLDHR
728,4029	2182,1868	2182,1667	77	TGLILTLQEMGADIEVINPR
789,7845	2366,3317	2366,3243	105	SAVLLAGLNTPGITTVIEPIMTR

De modo a confirmar a origem do gene que foi utilizado no processo de modificação genética para síntese da enzima *CP4 EPSPS* na semente de soja transgênica, também foi feita busca em um banco de dados geral (NCBI), contendo sequências do organismo *Agrobacterium* sp. Os resultados mostram a identificação da enzima neste organismo, corroborando o fato do gene ter sido originado da bactéria do solo *Agrobacterium* sp. Os resultados de identificação, para ambos os bancos de dados, podem ser vistos na Tabela 1.2.

Tabela 1.2. Identificação da enzima *CP4 EPSPS* na semente de soja transgênica usando SDS-PAGE e espectrometria de massas, tipo ESI-QTOF.

Nome da Proteína	Banco	Score	Acesso NCBI	MM (kDa)/pI	Cobertura (%)
<i>CP4 EPSPS</i>	NCBI (<i>Glycine max</i>)	868 (27 [*])	gi 18266432	47699 / 5,13	40
<i>5-enolpyruvylshikimate-3-phosphate synthase</i>	NCBI (<i>Agrobacterium</i> sp)	573 (53 [*])	gi 8469107	47673 / 5,13	34

^{*} Score mínimo para identificação

5. CONCLUSÕES PARCIAIS

Com os resultados apresentados neste capítulo, foi possível caracterizar e identificar, usando eletroforese bidimensional em gel de poliacrilamida (2D-PAGE), espectrometria de massas (MALDI-QTOF e ESI-QTOF) e banco de dados relacionados (MSDB, NCBI e *Glycine Max*), um total de 192 proteínas de soja, sendo que 179 proteínas foram identificadas para a faixa de pH de 4 a 7. Esta faixa mostrou melhor resolução na separação das proteínas. Foram identificadas 13 proteínas na faixa de 3 a 10. No total, foi obtida uma eficiência de 49% de proteínas identificadas em relação ao número de *spots* detectados no gel de eletroforese, dado este com boa eficiência em comparação a estudos semelhantes publicados na literatura.

Estes resultados permitiram estabelecer um mapa proteômico de referência para sementes de soja transgênica, tipo *Roundap Ready*, constatando a grande eficiência da técnica de 2D-PAGE como componente de separação e da técnica de espectrometria de massas como componente estrutural em estudos proteômicos.

De acordo com a avaliação funcional das proteínas de soja, observou-se, como esperado, que a maioria das proteínas identificadas estão relacionadas com a categoria de função de armazenamento e atividade protéica (96 proteínas), seguido de crescimento/divisão celular (34 proteínas), metabolismo/energia (17 proteínas), transporte (12 proteínas), doença/defesa (08 proteínas), metabolismo secundário (02 proteínas) e transdução de sinais (01 proteína). Ressalta-se um considerável número de proteínas relacionadas com a categoria não-classificadas/classificação desconhecida, totalizando um total de 22 proteínas.

Ressalta-se, ainda, a identificação da enzima *CP4 EPSPS* presente na semente de soja transgênica utilizando as técnicas de SDS-PAGE e espectrometria de massas. Com este resultado obtido, é possível afirmar, de maneira inequívoca, que o organismo em questão passou pelo processo de modificação genética. Além disso, destaca-se a simplicidade do método aplicado,

onde não houve a necessidade de pré-fracionamento da amostra, utilizando procedimentos cromatográficos, no qual a identificação foi possível usando duas enzimas de clivagem, a tripsina e a quimiotripsina.

CAPÍTULO 2:

PROTEÔMICA E METALOPROTEÔMICA COMPARATIVAS EM

SEMENTES DE SOJA TRANSGÊNICA E NÃO-TRANSGÊNICA

1. OBJETIVOS

Os objetivos deste capítulo consistem em uma avaliação proteômica e metaloproteômica comparativas entre as amostras de soja transgênica e não-transgênica. Para isso, foi utilizada a técnica de eletroforese em gel diferencial bidimensional (2D-DIGE), no intuito de se obter possíveis *spots* com diferenças quanto à expressão das proteínas entre as amostras. Após identificação, as proteínas que apresentaram alterações na sua expressão, foram caracterizadas por espectrometria de massas e identificadas a partir de buscas em bancos de dados relacionados com o auxílio do programa *Mascot*. Além disso, foi feito um estudo para identificação de íons metálicos e semimetálicos livres ou ligados a proteínas. Para isso, foi feita uma varredura semi-quantitativa de espécies metálicas e semimetálicas presentes nos extratos protéicos das sementes de soja transgênica e não-transgênica usando a técnica de ICP-MS. Além disso, naqueles *spots* diferenciais obtidos com a técnica de 2D-DIGE, foi feita a quantificação de cobre nestes *spots*, considerando que este elemento já foi encontrado como diferencialmente expresso em trabalhos já reportados na literatura pelo nosso grupo de pesquisa.

2. REVISÃO BIBLIOGRÁFICA

A revisão bibliográfica referente a este capítulo aborda uma descrição sobre proteômica comparativa usando eletroforese em gel bidimensional (2D-PAGE), justificando a escolha pela técnica de eletroforese diferencial bidimensional em gel de poliacrilamida (2D-DIGE) para este trabalho de tese. Em seguida, aspectos relacionados à interação metal-proteína serão discutidos. Por fim, será comentada a aplicação da técnica de ICP-MS em estudos metaloproteômicos comparativos.

2.1 Eletroforese em gel bidimensional (2D-PAGE) para estudos proteômicos comparativos

Os estudos das alterações protéicas em organismos biológicos, quando estas são expressas em géis de poliacrilamida, são feitos por meio de avaliações minuciosas das imagens destes géis. O resultado deste trabalho pode fornecer informações relevantes para estudos proteômicos comparativos [53]. Neste contexto, uma das principais aplicações destes estudos está relacionada com a obtenção de possíveis biomarcadores em sistemas biológicos [54].

Dentro do cenário de estudos proteômicos comparativos, a eletroforese bidimensional em gel de poliacrilamida (2D-PAGE) é a plataforma de separação mais aplicada, pois sua alta resolução possibilita uma elevada eficiência na separação protéica, permitindo, assim, a identificação de possíveis *spots* protéicos com diferenças de expressão nos géis avaliados [54,55]. No entanto, a comparação entre géis, o qual é feita pela comparação das imagens digitalizadas destes géis, mostra-se como uma atividade bastante árdua, que exige elevada atenção por parte do analista [53].

Além disso, outra questão envolvida nestes estudos proteômicos usando análise comparativa de imagens de géis 2D-PAGE, situa-se na alta variação que existe entre os géis obtidos, o que acarreta na reprodutibilidade insuficiente dos géis, como também na dificuldade em fazer quantificações exatas das proteínas resolvidas nos géis [56]. Estas variações podem ser atribuídas ao método de preparo da amostra empregado, variações naturais quando se considera amostras biológicas, variações relacionadas ao sistema eletroforético empregado, etc. [53,57,58].

Neste contexto, uma alternativa promissora a estas limitações consiste na utilização da técnica de eletroforese diferencial em gel bidimensional (2D-DIGE). Esta técnica baseia-se no uso de corantes fluorescentes, os quais possibilitam comparações proteômicas quantitativas exatas entre duas amostras, os quais são resolvidas no mesmo gel, minimizando, assim, os problemas acima mencionados

[59]. Além disso, destaca-se a grande sensibilidade destes corantes, o que possibilita a detecção de proteínas pouco abundantes, quando comparado a outros corantes usados na detecção de *spots* protéicos [60].

A seguir, será feita uma breve discussão sobre a técnica de eletroforese diferencial em gel bidimensional.

2.2 Eletroforese diferencial em gel bidimensional (2-D DIGE)

Na técnica 2-D DIGE, misturas complexas de proteínas são marcadas com corantes fluorescentes antes da separação eletroforética, a partir da ligação covalente destes com as proteínas, permitindo a detecção e quantificação de diferenças na abundância de proteínas em um único gel, o que minimiza os problemas de reprodutibilidade. Nesta técnica, a análise quantitativa das proteínas é feita pela excitação dos diferentes corantes (Cy2, Cy3 e Cy5) em comprimentos de onda específicos e a intensidade de fluorescência pode ser comparada, permitindo a quantificação de cada proteína presente no gel [61, 62]. A Figura 2.1 mostra a estrutura química destes corantes.

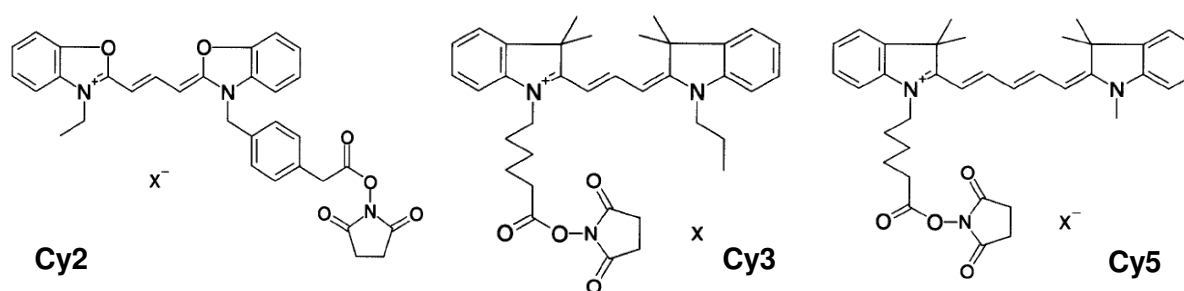


Figura 2.1. Estrutura química dos corantes Cy2, Cy3 e Cy5, usados na técnica de 2-D DIGE. Figura Adaptada de Tonge et al. 2001 [62].

Os corantes fluorescentes sofrem reação de substituição nucleofílica com o grupo ϵ -amino dos resíduos de lisina para formar uma amida. Os corantes são positivamente carregados para compensar a carga da lisina, que é perdida

durante a reação de marcação. Além disso, aproximadamente de 1-3% das proteínas são marcadas pelos corantes, sendo que cada proteína é marcada por uma única molécula de corante, onde as demais proteínas permanecem desmarcadas. Esta técnica fluorescente possui uma ampla faixa dinâmica, de quatro a cinco ordens de magnitude, permitindo que todas as proteínas visualizadas sejam exatamente quantificadas [62, 63].

Para a análise quantitativa das proteínas, é utilizado um padrão interno, que é marcado por um corante (em geral, o Cy2) e aplicado ao gel de eletroforese, juntamente com duas amostras a serem analisadas (marcadas com Cy3 e Cy5). O padrão interno consiste em uma mistura de quantidades iguais de proteínas das duas amostras e, portanto, ele contém todas as proteínas que estão presentes em cada uma das amostras avaliadas [59, 63]. Neste sentido, cada proteína possui um sinal único no padrão interno, que é utilizado para a comparação quantitativa direta dentro de cada gel e para normalizar os valores de abundância para cada proteína ao comparar géis diferentes, removendo, assim, possíveis variações experimentais gel-a-gel.

Em um estudo recente, Alvarez *et al.* [64] utilizaram a técnica de 2-D DIGE para detectar proteínas de raízes de *Brassica juncea* expressas diferencialmente, em resposta a tratamento com cádmio, já que esta planta exibe um grande potencial na remediação de solos contaminados com metais potencialmente tóxicos. Para isso, as plantas foram germinadas em câmara de crescimento e, após 13 dias de germinação, foi adicionado solução de CdCl_2 250 $\mu\text{mol/L}$. As proteínas extraídas das raízes das plantas controle foram marcadas com o corante Cy5 e as plantas tratadas com cádmio foram marcadas com o corante Cy3. O padrão interno foi marcado com o corante Cy2, consistindo de uma mistura equivalente das amostras controle e tratada com cádmio. Em seguida, as amostras foram misturadas para a etapa de separação por eletroforese bidimensional. Como resultados, foram identificados 69 *spots* diferencialmente expressos em resposta ao tratamento com cádmio, onde 43 *spots* foram sub-expressos e 26 *spots* foram sobre-expressos. Estes *spots* protéicos foram, posteriormente, identificados usando a técnica de nano-LC-MS/MS.

O processo de modificação genética em organismos biológicos vem sendo cada vez mais discutido na literatura, pois a introdução de genes exógenos, possivelmente, pode alterar a expressão de genes endógenos e/ou a produção de metabólicos. Neste sentido, Carli *et al* [65] investigaram possíveis alterações no proteoma de folhas de plantas transgênicas a partir da inserção de genes exógenos que expressam anticorpos recombinantes direcionados a agir contra proteínas virais, usando a técnica de 2D-DIGE. Como referência, as plantas não-transgênicas também foram utilizadas no trabalho. Os autores sugerem, como hipótese, que a modificação genética nas plantas não causa efeitos pleiotrópicos nas mesmas, validando a proteção mediada por anticorpos como ferramenta para gerar organismos resistentes ao ataque de vírus. Para isso, duas amostras de plantas foram utilizadas neste trabalho: tomate (*Lycopersicon esculentum* cv. Micro-Tom), expressando o anticorpo scFv (G4) que age contra uma proteína do vírus do mosaico do pepino (*Cucumber mosaic virus*, CMV) , e a planta *Nicotiana benthamiana*, o qual expressa o anticorpo scFv (B9) que atua contra uma glicoproteína presente no vírus do tomateiro. Os genes, os quais codificaram os anticorpos, foram clonados a partir do vetor pBI usando o organismo *Agrobacterium* como mediador da transformação. Após extração das proteínas, as amostras foram marcadas com os corantes Cy3 e Cy5, onde o corante Cy2 foi utilizado na marcação de todas as amostras. Seis replicatas biológicas foram feitas para cada amostra de planta.

Como resultados, 10 proteínas da planta de tomate e 08 proteínas da planta *Nicotiana benthamiana* foram diferencialmente expressas, em termos de volume, onde o fator de regulação foi de 1,5 a um nível de confiança de 95% ($p \leq 0,05$, de acordo com o teste *t* de Student). Todos os *spots* foram identificados por MALDI-TOF MS e μ LC-ESI-IT-MS/MS. Os autores sugerem que este número limitado de proteínas expressas diferencialmente pode estar relacionado aos eventos de transformação.

2.3 Metais associados a proteínas

Íons metálicos são componentes vitais para a química da vida [66]. Um significativo número de proteínas e enzimas (ca. 30%) contem espécies metálicas ou semimetálicas em suas estruturas, sendo que 40% destes elementos desempenham atividades cruciais nas atividades biológicas de proteínas, principalmente atuando como cofatores nestas proteínas [11]. Isto mostra que estudos envolvendo metais e proteínas tem sido extensivamente desenvolvidos para elucidar a essencialidade biológica e toxicidade destas espécies metálicas em nível molecular, sendo estas informações bastante necessárias em diversos campos de pesquisa, tais como a bioquímica, biologia, medicina, farmácia, nutrição, agricultura e ciências ambientais [11,67].

Levando em consideração as definições recentemente propostas na literatura, o metaloma pode ser entendido como a entidade de espécies metálicas e semimetálicas, na forma livre ou complexada, presentes em um sistema biológico, onde deve ser definida a sua identidade e/ou concentração. Sendo assim, a metalômica corresponde ao estudo do metaloma, como também as interações entre íons metálicos e outras espécies metálicas com genes, proteínas, metabolitos, e outras biomoléculas em sistemas biológicos [66].

Dentro de um estudo metalômico, quando se busca a composição elementar de um sistema biológico, é possível avaliar a presença de espécies metálicas e semimetálicas livres, sendo este estudo denominado de ionômica, como também a determinação destas espécies complexadas a proteínas, ou seja, a metaloproteômica [66]. Estas duas ultimas abordagens em estudos metalômicos foram avaliadas neste trabalho de Tese.

O estudo metalômico pode fornecer importantes informações sobre um determinado sistema biológico, tais como: (i) como o elemento (metal ou semimetal) esta distribuído dentro dos compartimentos celulares de um dado tipo de célula; (ii) seu ambiente de coordenação, ou seja, em qual biomolécula está incorporado ou em qual bioligante está complexado, e (iii) a concentração das

espécies metálicas individuais [68]. O monitoramento de mudanças do metaloma em função do tempo e/ou a exposição a um estímulo externo pode ser referido como metalômica comparativa [68].

Assim, para compreensão dos fenômenos que envolvem as interações entre metais e proteínas, deve-se fazer uso de uma gama diversificada de técnicas analíticas. Desta forma, a etapa inicial de um estudo metalômico consiste na separação das proteínas, o qual podem ser utilizadas técnicas como eletroforese em gel, eletroforese capilar ou cromatografia líquida de alta eficiência. Em seguida, é feita a caracterização estrutural das proteínas aplicando técnicas de espectrometria de massas, tais como a ionização/dessorção de matriz induzida por laser (MALDI), bem como a ionização por electrospray (ESI). E, por fim, técnicas destinadas à identificação e quantificação de íons metálicos e semimetálicos ligados as proteínas, que incluem a espectrometria de massas com plasma indutivamente acoplado (ICP-MS), espectrometria de absorção atômica com atomização eletrotérmica (ETAAS), fluorescência de raios-X com radiação Sincrotron, dentre outras [66-68].

2.4 Espectrometria de massas em metalômica: ICP-MS

Atualmente, a técnica de ICP-MS tem se consolidado como uma das mais empregadas na área de espectrometria atômica para determinações de metais e/ou semimetais ligados a biomoléculas. Esta técnica constitui-se como uma ferramenta complementar aos estudos proteômicos quando o objetivo é identificar e quantificar diferentes elementos ligados em metaloproteínas e proteínas ligantes de metal [11].

Dentre as principais vantagens desta técnica, pode-se citar sua ampla faixa linear dinâmica (até oito ordens de magnitude), baixos limites de detecção para muitos elementos sendo esta característica fundamental quando se pretende avaliar espécies inorgânicas ligadas a proteínas, mínimo efeito de matriz, bem

como a possibilidade de se fazer análises multielementares, essencial neste tipo de estudo, uma vez que as etapas de preparo das amostras são complexas e demandam um tempo relativamente longo [69]. Levando em consideração esta última característica, em uma única varredura, é possível avaliar de forma qualitativa a maior parte dos elementos presentes na tabela periódica. No modo quantitativo, a precisão e exatidão obtidas são comparáveis às obtidas com as técnicas monoelementares [70].

Após a separação das proteínas por 2D-PAGE, os spots protéicos podem ser analisados por ICP-MS, a partir de dois caminhos: (i) em solução, onde os *spots* protéicos são digeridos aplicando a técnica de decomposição assistida por radiação microondas, e a solução resultante é introduzida no equipamento via sistema nebulizador; e (ii) em *in situ*, onde o *spot* protéico é diretamente analisado pela técnica de LA-ICP-MS [11]. Neste último caso, um pulso de laser é incidido sobre a superfície do *spot* com posterior ejeção de material na forma de vapor e matéria particulada, que serão, então, conduzidos para o ICP com auxílio do fluxo de gás carregador [71,72].

Em trabalho recentemente publicado, Mataveli *et al* [73] avaliaram os efeitos da modificação genética na composição elementar em sementes de soja transgênicas *Roundap Ready* e não-transgênicas, em termos de nível elementar total nas sementes, fracionamento por extração seqüencial e em digestão gastrointestinal simulada. As amostras foram analisadas usando a técnica de ICP-SF-MS e cromatografia por exclusão de tamanho acoplado ao ICP-MS. A análise das sementes de soja digeridas por radiação microondas mostrou diferenças estatisticamente significativas no conteúdo de Cu, Fe e Sr entre as amostras de soja, sendo que o nível destes elementos foi maior na soja transgênica. Além disso, as frações bioacessíveis dos elementos Cu, Fe, Mn, S e Zn foram maiores nas sementes de soja transgênica.

Sussulini *et al* (2010) [74] aplicaram estratégias metalômicas para avaliar diferenças no perfil metalômico do soro sanguíneo entre indivíduos saudáveis (grupo controle) e pacientes com transtorno afetivo bipolar (tratados com lítio ou com outras drogas com exceção do lítio). Para isso, foram utilizadas as técnicas

de eletroforese bidimensional em gel de poliacrilamida (2D-PAGE) para separação das proteínas, MALDI-TOF MS/MS para identificação das biomoléculas e LA-ICP-MS para a identificação das espécies metálicas ligadas a estas proteínas presentes nos géis. Como principais resultados, os autores detectaram diferenças entre os grupos avaliados, em termos de metais ligados às proteínas. Foram detectados Na, Mg, Zn, Ca e Fe como metais ligados às proteínas em todos os grupos. Mn foi encontrado somente no grupo controle. Co foi encontrado nos grupos controle e pacientes bipolares tratados com lítio. K e Ti foram encontrados somente nos grupos de pacientes bipolares e P foi encontrado nos grupos controle e pacientes bipolares não tratados com lítio. Este trabalho demonstra que estratégias metalômicas empregando a combinação de LA-ICP-MS com MALDI-TOF-MS/MS apresentam dados complementares para observar possíveis diferenças em termos de metais ligados a proteínas de soro, entre indivíduos controle e pacientes com transtorno bipolar tratados ou não com lítio.

3. PARTE EXPERIMENTAL

3.1 Material e equipamentos

Os equipamentos utilizados no desenvolvimento dos experimentos descritos neste capítulo são listados a seguir. Eles se encontram disponíveis no Grupo de Espectrometria, Preparo de Amostras e Mecanização (GEPAM), localizado no Instituto de Química, Unicamp. Os Espectrômetros de massas moleculares, para identificação das proteínas, encontram-se disponíveis no Laboratório Dalton de Espectrometria de Massas, no Instituto de Química, Unicamp, bem como no Laboratório Nacional de Biociências, localizado no Laboratório Nacional de Luz Síncrotron (LNLS), Campinas, SP. Estes equipamentos foram utilizados dentro da Rede Proteoma do Estado de São Paulo. Levando em consideração que muitos dos equipamentos utilizados neste capítulo já foram listados no item 3.1 (Capítulo 1), aqui serão somente listados os equipamentos adicionais.

- *Scanner*, modelo EttanTM DIGE Imager (GE Healthcare, Suécia);
- Espectrômetro de massas com fonte de plasma indutivamente acoplado, ICP-MS (Perkin Elmer, EUA);
- Forno de Microondas tipo cavidade, modelo DGT 100 Plus (Provecto Analítica, Brasil);
- Placa aquecedora (Quimis, Brasil);

3.2 Reagentes e soluções

Todas as soluções foram preparadas utilizando água deionizada ($\geq 18,2$ M Ω), a partir de sistema de purificação Milli-Q. Os reagentes utilizados foram de pureza analítica, provenientes das empresas Amersham Biosciences (Suécia), BioAgency (Brasil), J.T. Baker (EUA) e Merck (Alemanha), e serão listados a seguir. As vidrarias usadas são de uso rotineiro na área de bioanalítica. Levando em consideração que muitos dos reagentes utilizados neste capítulo já foram listados no item 3.2 (Capítulo 1), aqui serão somente listados os reagentes adicionais.

- Ácido Nítrico, HNO₃ – 65,0 g/mol (Merck);
- Corantes fluorescentes CyDye™ DIGE Fluors (GE Healthcare);
- Lisina, C₆H₁₄N₂O₂ – (146,19 g/mol) (Sigma);
- *N, N*-dimetilformamida (DMF), C₇H₇NO – 73,09 g/mol (Sigma);
- Péroxido de Hidrogênio, H₂O₂ - 30 % (m/m) (Merck);
- Solução padrão multielementar (Merck);
- Solução padrão de zinco (1000 mg L⁻¹), grau ICP (Merck);
- Solução padrão ródio (1000 mg L⁻¹), grau ICP (EMD™).

3.3 Procedimento experimental

3.3.1 Amostras de soja

As amostras utilizadas são as mesmas já descritas no capítulo 1, item 3.3.1.

3.3.2 Extração das proteínas de soja

O procedimento de extração das proteínas foi o mesmo já descrito no capítulo 1, item 3.3.2.

3.3.3 Preparo dos extratos de soja para análise por 2D-DIGE

Para a análise por 2-D DIGE, realizada na faixa de pH entre 4 e 7, com o intuito de remover possíveis interferentes no processo eletroforético (como sais e surfactantes aniônicos), o extrato protéico foi submetido a uma etapa de precipitação das proteínas, empregando solução de acetato de amônio $0,1 \text{ mol.L}^{-1}$ em metanol durante 12 h, a -20°C . O precipitado protéico foi separado por centrifugação a 5000 g (4°C) por 10 min e, em seguida, lavado duas vezes com solução gelada de acetato de amônio $0,1 \text{ mol L}^{-1}$ em metanol, duas vezes com acetona gelada, na concentração de 80% (v/v) e uma vez com etanol a 70% (v/v), também gelado [3].

Após esta etapa, o precipitado foi ressolubilizado em um tampão que consistiu em uma solução de uréia 7 mol L^{-1} , tiouréia 2 mol L^{-1} , CHAPS 2% (m/v) e Tris-HCl 40 mmol L^{-1} (pH 8,8). Em seguida, realizou-se a centrifugação da solução resultante por 2 min a 5000 g , de modo a remover qualquer resíduo sólido que ainda permanecesse na amostra.

3.3.4 Separação das proteínas por 2D-DIGE

Para a separação das proteínas por 2D-DIGE, o precipitado protéico foi ressolubilizado em um tampão que consiste em uma solução de uréia 7 mol L^{-1} , tiouréia 2 mol L^{-1} , CHAPS 2 % (m/v) e Tris-HCl 40 mmol L^{-1} (pH 8,8). Em seguida, após quantificação, um volume correspondente a $50 \text{ }\mu\text{g}$ de proteínas de cada amostra e $50 \text{ }\mu\text{g}$ do padrão interno, composto por $25 \text{ }\mu\text{g}$ de cada amostra, foram marcados com os corantes CyDye™ DIGE Fluors. A reconstituição dos corantes foi feita seguindo as instruções do fabricante. Os corantes Cy2, Cy3 e Cy5 foram, então, misturados com as proteínas, de modo que cada $50 \text{ }\mu\text{g}$ de proteínas de cada amostra fossem marcados com 400 pmol do corante selecionado. A reação

foi feita (em gelo) a 4 °C por 30 minutos. Foi adicionado 1 µL de lisina a 10 mmol L⁻¹ por 50 µg de proteína para interromper a reação de marcação com incubação a 4 °C, em gelo, por 10 min no escuro. As amostras (padrão interno marcado com Cy2 e duas amostras marcadas com Cy3 e Cy5) foram misturadas, e, em seguida, o volume foi completado a 250 µL com tampão de rehidratação, o qual é composto por uréia 7 mol L⁻¹, tiouréia 2 mol L⁻¹, CHAPS 2% (m/v), anfólitos 0,5% (v/v) em pH variando de 3 a 10, e azul de bromofenol 0,002% (m/v). O procedimento para separação eletroforética bidimensional em gel foi o mesmo como mencionado anteriormente. As imagens dos géis foram obtidas por um *scanner* específico para detectar a fluorescência das amostras e padrão interno, e a análise das imagens foi feita com o programa DeCyder™ 2-D Differential Analysis Software v.7.0 (GE Healthcare), que permite a detecção dos possíveis *spots* com diferença de expressão entre as amostras avaliadas. Com este procedimento foram feitas três replicatas para as amostras de sementes de soja.

3.3.5 Digestão trípica e identificação das proteínas por Espectrometria de massas

Para identificação das proteínas diferenciais, um gel 2D-PAGE preparativo foi feito usando o mesmo procedimento descrito no capítulo 1. Da mesma forma, o procedimento de digestão enzimática, análise por espectrometria de massas e identificação das proteínas também foram seguidos. Ressalta-se que a identificação das proteínas foi feita em espectrômetro de massas do tipo ESI-QTOF. Para isso, os peptídeos foram secos, e posteriormente ressolubilizados em solução de ácido fórmico 0,1% (v/v).

Os espectros de massas contendo a lista com os picos dos peptídeos foram processados usando os seguintes parâmetros: bancos de dados *NCBI* (*Glycine Max*) e *NCBI* geral, um sítio de clivagem; digestão com tripsina;

carbamidometilação como modificação fixa; oxidação em metionina como modificação variável e erro, em massa, de 0,1 Da.

3.3.6 Decomposição do extrato protéico das sementes de soja por radiação microonda

Para a decomposição assistida por radiação microonda dos extratos protéicos das sementes de soja transgênica e não-transgênica, adicionou-se 5 mL de HNO_3 sub-destilado concentrado e 1 mL de H_2O_2 30% (m/m) a cada frasco de teflon. Foi dado um intervalo de tempo de espera de 20 min para pré-reação, antes do início da decomposição por radiação microondas. O programa de decomposição aplicado está descrito na Tabela 2.1. Foram decompostas três replicatas para o branco de decomposição bem como para as amostras. Após a decomposição, os frascos foram removidos do forno microondas e esperou-se o resfriamento até a temperatura ambiente (pelo menos, 2 h).

Tabela 2.1 Programa aplicado para a decomposição dos extratos protéicos das sementes de soja por radiação microonda.

Etapa	Tempo (min)	Potência (W)
1	2	330
2	5	330
3	5	460
4	4	590

3.3.7 Identificação e quantificação dos íons livres no extrato protéico de soja por ICP-MS

Para as análises semi-quantitativas (*TotalQuantity*) por ICP-MS, os extratos decompostos foram submetidos a aquecimento, em banho de areia, para diminuição do volume até quase secura. Em seguida, as amostras foram filtradas e os volumes finais ajustados para 10 mL, usando solução de HNO₃ sub-destilado 1% (v/v). Para as medidas no equipamento, foi feita uma diluição de 50x nos extratos das sementes, bem como no branco. A análise semi-quantitativa foi feita em triplicata e as seguintes regiões de razão massa/carga foram avaliadas: 6-15, 19-39, 42-210 e 230-240, com voltagem no quadrupolo (RPq) igual 0,4 V. O sistema de introdução de amostras empregado consiste em um nebulizador Meinhard® acoplado a uma câmara de nebulização ciclônica. As condições de operação do equipamento são descritas na Tabela 2.2. O instrumento foi otimizado a cada dia de trabalho (*daily performance*) para máxima sensibilidade para os íons M⁺ e mínimos sinais para M²⁺ e MO⁺ usando uma solução de 1 µg L⁻¹ contendo Mg, In, Ba, Ce, U e Ce e monitoramento das contagens de ²⁴Mg⁺, ¹¹⁵In⁺, ¹³⁸Ba⁺, ¹³⁸Ba⁺⁺, ¹⁴⁰Ce⁺, ¹⁵⁶CeO⁺ e ²³⁸U⁺.

Tabela 2.2 Condições de operação do equipamento de ICP-MS para análise semi-quantitativa.

Parâmetro	Valor
Fluxo de gás nebulizador	0,92 L/min
Fluxo de gás auxiliar	1,1 L/min
Fluxo de gás do plasma	15 L/min
Voltagem das lentes	6,25 V
Potência da radiofrequência	1280 kW
Fluxo de injeção da amostra	1,5 mL/min

Aquelas espécies metálicas que mostraram-se como diferenciais, em termos de contagens, entre as amostras foram quantificadas usando curvas de calibração cujos padrões variaram de 0,1 a 50 $\mu\text{g L}^{-1}$ dos padrões individuais dos elementos quantificados. Sendo assim, foram quantificados os elementos $^{24}\text{Mg}^+$, $^{39}\text{K}^+$, $^{23}\text{Na}^+$, $^{55}\text{Mn}^+$, $^{63}\text{Cu}^+$, $^{60}\text{Ni}^+$, $^{85}\text{Rb}^+$ e $^{66}\text{Zn}^+$. Além disso, foi utilizado um padrão interno de ródio ($^{103}\text{Rh}^+$) para a normalização dos resultados. Os dados de quantificação obtidos para todos os elementos foram transferidos para uma planilha do programa *Excel* para avaliação estatística.

3.3.8 Decomposição dos *spots* protéicos diferenciais obtidos pela técnica de 2D-DIGE por radiação microonda.

Para a decomposição em microonda dos *spots* protéicos diferenciais, foram feitas quatro réplicas de géis para cada amostra, sendo que, para os *spots* avaliados, foi feito um “*pool*” dos materiais digeridos de modo a aumentarmos a eficiência na quantificação por ICP-MS. Para esta análise, foi utilizado o procedimento proposto por Verbi *et al* [75]. Sendo assim, adicionou-se 4 mL de HNO_3 sub-destilado concentrado e 1 mL de H_2O_2 30% (v/v) a cada frasco de teflon. Foi dado um intervalo de tempo de espera de 20 min para pré-reação, antes do início da decomposição por radiação microonda. O programa de decomposição aplicado é mostrado na Tabela 2.3. Foram decompostas três replicatas para o branco de decomposição bem como para as amostras. Após a decomposição, os frascos foram removidos do forno microondas e esperou-se o resfriamento até a temperatura ambiente (pelo menos 2 h).

Tabela 2.3 Programa aplicado para a decomposição dos *spots* protéicos diferenciais por radiação microonda.

Etapa	Tempo (min)	Potência (W)
1	3	400
2	6	790
3	3	000

3.3.9 Quantificação de cobre nos *spots* protéicos diferenciais por ICP-MS.

Para a quantificação por ICP-MS, os *spots* decompostos foram submetidos a aquecimento, em banho de areia, para diminuição do volume até quase *secura*. Em seguida, as amostras foram filtradas e os volumes finais ajustados para 5 mL, usando solução de HNO₃ sub-destilado 1% (v/v). Para as leituras no equipamento, realizou-se uma diluição de 10x nos extratos das sementes, bem como no branco. Para a quantificação de $^{63}\text{Cu}^+$ foi utilizada uma curva de calibração cujos padrões variaram de 0,1 a 3 $\mu\text{g L}^{-1}$. Além disso, foi utilizado um padrão interno de ródio ($^{103}\text{Rh}^+$) para normalização dos resultados. Os dados obtidos foram transferidos para uma planilha do programa *Excel* para avaliação estatística. As condições de operação do equipamento são as mesmas com descritas no item 3.3.7 e Tabela 2.2.

4. RESULTADOS E DISCUSSÃO

4.1 Separação por 2D-DIGE

A técnica de 2D-DIGE, associada a ESI-QTOF MS/MS foi utilizada para visualizar a expressão dos perfis protéicos em amostras de soja transgênica e não-transgênica. A Figura 2.2 mostra o gel de eletroforese diferencial em gel bidimensional (DIGE) de sementes de soja transgênica e não-transgênica. Em azul, estão marcados os *spots* correspondentes às proteínas sub-expressas, ou seja, mais expressas na soja não-transgênica, e, em vermelho, estão mostrados os *spots* correspondentes às proteínas sobre-expressas, ou seja, mais expressas na soja transgênica. O fator de regulação foi considerado como sendo $\geq 1,5$ (50 % de variação), conforme determinado pelo programa de análise de imagens Decyder 2D software, e considerando diferenças estatisticamente significantes na expressão ($p \leq 0,05$, de acordo com o teste *t* de Student).

De acordo com os dados obtidos pelo software, um total de 4 possíveis *spots* protéicos foram diferencialmente expressos, levando em consideração uma variação de expressão ao nível de 50 % (1,5). Deste total, 2 foram sobre-expressos, ou seja, foram mais expressos na soja transgênica e 2 foram sub-expressos no gel, ou seja, menos expressos na soja transgênica. Sendo assim, estes 4 *spots* foram selecionados, retirados do gel, digeridos com tripsina e analisados por ESI-QTOF MS/MS. A Tabela 2.4 mostra os resultados obtidos para a identificação das proteínas. Ressalta-se que, em todas as proteínas identificadas, os valores de *score* foram estatisticamente significantes ($p < 0,05$).

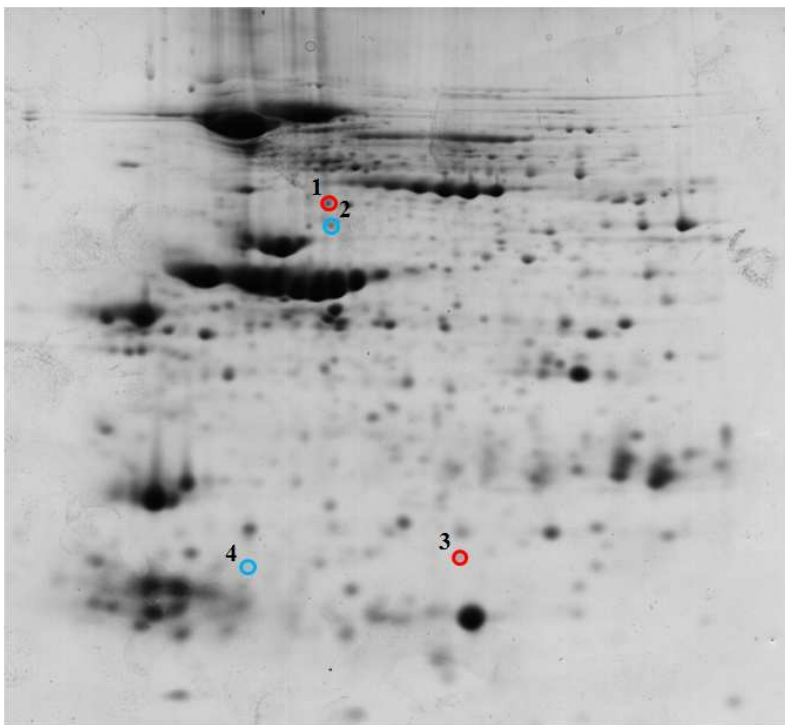


Figura 2.2. Análise por 2-D DIGE (faixa de pH entre 4 e 7) para sementes de soja transgênica (7575 RR) e não-transgênica (7501). Em azul estão marcadas as proteínas sub-expressas e, em vermelho, as sobre-expressas. Os números são relacionados às identificações das proteínas (Ver tabela 2.4).

Um ponto interessante a se destacar nesta técnica seria a quantidade de *spots* que podem ser detectados. Neste Capítulo, empregando 2-D DIGE, foi possível identificar, aproximadamente, 780 ± 83 *spots* no gel obtido. Entretanto, levando em consideração que existe um único gel para ambas as amostras, ressalta-se que o Decyder 2D software faz a seleção automática para ambas as amostras, ou seja, a mesma quantidade de *spots* de uma amostra estará presente na outra amostra, não permitindo uma seleção manual mais acurada, como realizado no trabalho de Brandão *et al.* [34]. Sendo assim, na técnica de 2D-DIGE, a remoção manual de um *spot* em uma amostra removerá este mesmo *spot* na outra amostra. Como consequência disso, a possibilidade de *spots* falsos positivos estarem presentes nas amostras é significativa.

A Figura 2.3 mostra os *spots* protéicos obtidos com diferença de expressão neste trabalho. Nesta figura são mostradas as imagens dos *spots*, as quais mostraram diferenças em sua expressão, levando em consideração as três

replicadas avaliadas entre as amostras. A área dos *spots* está delimitada em amarelo. Os níveis de diferenciação de expressão entre os *spots* detectados como diferenças de expressão podem ser visualizados na tabela 2.4.

Tabela 2.4. Níveis de diferenciação entre os *spots* com diferenças de expressão.

Spot	Nível
1	1,89±0,14
2	1,63±0,39
3	3,34±1,79
4	1,51±0,06

Em comparação ao trabalho publicado por Brandão *et al* [34], foram detectados um total de 07 proteínas diferenciais considerando a faixa de pI de 4 a 7 em sementes de soja transgênica em relação a soja não-transgênica. A diferença obtida deste resultado com o resultado obtido neste trabalho de tese pode estar relacionado com o modo de seleção dos *spots* nas duas técnicas eletroforéticas.

Na técnica de 2D-PAGE, o programa de análise de imagens utilizado pelos autores, após a detecção automática dos *spots*, permite a aplicação da edição manual, a fim de eliminar *spots* falso positivos, inserir *spots* falso negativos e efetuar correções quanto a delimitação das regiões *spots*, sendo que estas devem ser, aproximadamente, circulares ou elípticas [58]. Consequentemente, a edição manual acarretará em modificações nos valores das intensidades e volumes dos *spots* quando comparado com valores obtidos com o modo de detecção automático.

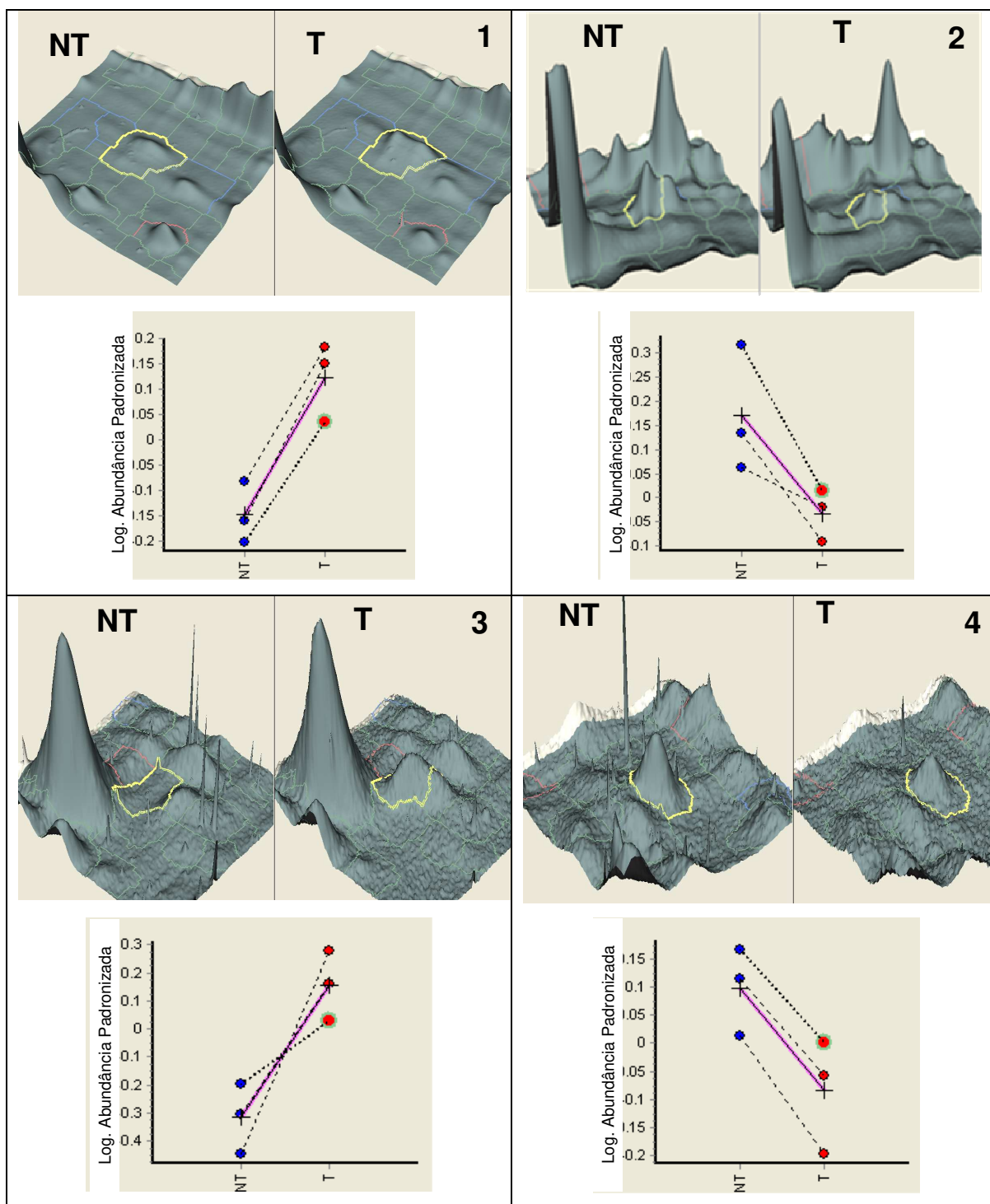


Figura 2.3. Spots com variação de expressão entre as amostras de soja transgênica (T) e não-transgênica (NT). A área delimitada pelo software com coloração amarela, mostra o spot com variação de expressão. No canto superior, à direita, está o número do spot localizado no gel 2-D DIGE (Ver Figura 2.2).

Neste trabalho de Tese, o programa DeCyder™ 2-D Differential Analysis v.7.0 (GE Healthcare) foi utilizado para as análises comparativas das imagens dos géis referentes aos corantes fluoróforos misturados às amostras de sementes de soja transgênica e não-transgênica. Este software utiliza um algoritmo de co-deteção que detecta simultaneamente *spots* protéicos marcados em diferentes imagens presentes em um mesmo gel. Esta co-deteção aumenta significativamente a precisão no estudo proteômico comparativo, ou seja, áreas equivalentes de cada imagem podem ser comparadas de forma mais exata [76,77]. Sendo assim, a detecção dos *spots* a partir do programa utilizado permite apenas a seleção automática dos *spots* não sendo possível nenhuma edição manual, seja na adição de *spots* ou quanto à região delimitada nos mesmos [76].

A vantagem deste modo baseia-se no fato de que, como duas amostras são corridas no mesmo gel, a delimitação das áreas dos *spots* correlacionados entre as amostras é a mesma. Entretanto, uma possível limitação deste modo de seleção é que alguns *spots* podem ser delimitados de forma incorreta na imagem do gel. Isto pode ser evidenciado observando-se o *spot* 1, onde a área delimitada para o *spot* na soja não-transgênica seria menor quando comparado a área que seria delimitada manualmente. Além disso, no *spot* 3, a área delimitada para a soja transgênica poderia ser maior, já que não há uma distinção clara entre os dois *spots*, como mostrado pela figura.

Desta forma, é possível dizer que o número de *spots* diferenciais entre as amostras analisadas poderia ser diferente (tanto para mais quanto para menos) levando em consideração a possível edição manual dos *spots* detectados pelo programa utilizado, seja na inserção ou remoção de *spots* bem como na correção das áreas delimitadas pelo software utilizado.

4.2 Identificação dos *spots* diferenciais por espectrometria de massas

A Tabela 2.5 mostra os resultados para a caracterização dos *spots* que apresentaram variação de expressão (ca. 50 %), usando ESI-QTOF MS/MS. Como resultados obtidos, todos os 4 *spots* foram identificados. Informações como sequência dos peptídeos identificados, ponto isoelétrico e massa molar experimental (estimado dos *spots* do gel de eletroforese), ponto isoelétrico e massa molar teórico (baseado na proteína identificada), e sequência de cobertura dos peptídeos identificados são mostrados.

Tabela 2.5. Resultados obtidos para a identificação das proteínas com diferenças de expressão entre as sementes de soja transgênica e não-transgênica.

Spot	Identidade da proteína	Acesso	MM (kDa) pI	Score	Cobertura	Peptídeos sequenciados
1	Actin (Fragment).- [Glycine max]	Q96445_SOYBN (MSDB)	37236/5,57	65 (35 [*])	9	K.SYELPDGQVITIGDER.F K.LSYIALDYEQELETAR.T
2	cytosolic glutamine synthetase (Cucumis melo)	gi 113171384 (NCBI)	39338/5,23	198 (54 [*])	13	K.GYFEDR.R R.DIVDAHVK.A R.EEGGYEVIK.K K.EHIAAYGEGNER.R K.SMREEGGYEVIVK.K
3	glycinin subunit G1 [Glycine max]	gi 18635 (NCBI)	56299/5,89	91 (54 [*])	4	R.VFDGELQEGR.V R.SQSDNFEYVSFK.T
4	glycine-rich RNA-binding protein [Glycine max]	gi 5726567 (NCBI)	15894/6,58	30 (27 [*])	6	R.NITVNEAQRS.G

^{*} Score mínimo para identificação

A proteína do *spot* 1 foi identificada como a proteína *Actin* (Fragmento). Esta proteína já havia sido identificada no mapa proteômico estabelecido no capítulo 1 e foi classificada na categoria de armazenamento. Esta proteína está envolvida em vários tipos de motilidade celular e está amplamente expressa em

todas as células eucariotas. Além disso, esta proteína também se liga a ATP e a outras proteínas [46]. Ressalta-se, ainda, que esta proteína também foi detectada como diferencialmente expressa no trabalho de Brandão *et al*, 2010 [34], usando a técnica de 2D-PAGE e análise de imagens.

A proteína do *spot 2* não foi identificada usando o banco de dados de soja. Sendo assim, em busca feita no banco de dados do *NCBI* geral, esta proteína foi identificada como *cytosolic glutamine synthetase (Cucumis melo)* e foi classificada na categoria de metabolismo. É considerada com uma enzima ligante, sendo bastante expressa em vários tipos de raízes, principalmente na parte capilar. Esta enzima se liga a moléculas de ATP, participa na fixação de nitrogênio e na biosíntese da glutamina (ácido 2-amino-4-carbamoilbutanóico) [46].

A proteína do *spot 3* refere-se a *glycinin subunit G1* [Glycine max]. Esta proteína, classificada na categoria de armazenamento, é uma sub-unidades de uma das mais abundantes proteínas de armazenamento de soja, a *glycinin*. Esta proteína é responsável pelas propriedades nutricionais, físico-químicas e fisiológicas das sementes de soja [78].

A proteína do *spot 4* foi identificada como *glycine-rich RNA-binding protein* [Glycine max]. Esta proteína pertence a uma classe de proteínas ligantes de RNA em plantas e está envolvida na resposta celular ao ambiente e as condições de desenvolvimento. Entretanto, sua relação com o metabolismo do RNA ainda permanece desconhecida [46,52]. Sendo assim, esta proteína possui classificação desconhecida.

Até o presente momento não foi reportado na literatura nenhuma relação destas proteínas diferenciais, identificadas neste trabalho, com processos de modificação genética em plantas.

4.3 Avaliação do perfil ionômico nos extratos protéicos por ICP-MS

A análise ionômica foi feita a partir da semi-quantificação dos extratos protéicos das sementes de soja transgênica e não-transgênica por ICP-MS. As quantificações feitas seguiram o padrão das contagens das análises semi-quantitativas.

De acordo com os resultados obtidos para a análise semiquantitativa, verifica-se que os íons Mg^+ , Si^+ , K^+ , Ti^+ , Mn^+ , Zn^+ e Rb^+ foram mais detectados na soja não-transgênica. Já o íon Cu^+ foi mais detectado nas sementes de soja transgênica. Além disso, cabe ressaltar que os elementos Na, Mg e K estão presentes em maiores concentrações, o que corrobora o fato destes elementos serem considerados macroelementos. Já os elementos Si, Ti, Mn, Cu, Zn e Rb estão presentes em concentrações menores em comparação aos macroelementos citados. Vale salientar aqui que o branco utilizado neste estudo consiste da solução de extração de proteínas utilizada sem a amostra para que qualquer contaminação vinda do sistema pudesse ser verificada e descontada dos sinais analíticos dos elementos avaliados.

A partir dos elementos identificados como diferenciais entre as amostras, foi feita a quantificação destes de modo a corroborar com os resultados observados previamente nas análises semiquantitativas. Ressalta-se que os elementos Si e Ti não foram quantificados pela indisponibilidade de padrões no laboratório. As curvas de calibração construídas para os elementos diferenciais apresentaram ótima correlação linear, sendo que os parâmetros obtidos das são mostrados na Tabela 2.6.

Tabela 2.6. Parametros das curvas analíticas de calibração dos elementos diferenciais obtidos por ICP-MS.

Metal	a (s[*])	b (s[*])	R
Mg	0,0064 (0,0038)	0,0311 (0,0001)	0,9999
K	- 0,0011 (0,0391)	0,074 (0,0013)	0,9995
Na	- 0,0021 (0,0077)	0,0476 (0,0002)	0,9999
Mn	-0,0119 (0,0111)	0,8388 (0,0024)	0,9999
Cu	0,0137 (0,0132)	0,3598 (0,0028)	0,9998
Rb	-0,0155 (0,0035)	0,8423 (0,0007)	1
Zn	-0,2972 (0,0128)	0,2105 (0,0013)	0,9999

^{*}s é o erro da medida

Devido à diferença de concentração obtida entre os elementos diferenciais, os resultados foram divididos em duas tabelas. A Tabela 2.7 mostra a concentração dos elementos Na, K e Mg e a Tabela 2.8 mostra a concentração dos elementos Mn, Cu, Rb e Zn.

Tabela 2.7. Concentrações dos elementos Na, K e Mg nos extratos protéicos das sementes de soja transgênica e não-transgênica (média $\pm$ desvio padrão)

Concentrações / mg L⁻¹				
Metal	Limite de detecção	Limite de quantificação	Soja Transgênica	Soja Não-Transgênica
Na	0,096	0,320	3,58 $\pm$ 0,31	8,24 $\pm$ 1,50
K	1,388	4,625	38,24 $\pm$ 0,34	46,83 $\pm$ 0,84
Mg	0,041	0,136	4,89 $\pm$ 0,06	5,88 $\pm$ 0,32

Como pode ser visto na Tabela 2.7, os elementos Na, K e Mg estiveram mais presentes na soja não-transgênica, corroborando assim os dados de semi-quantificação. Levando em consideração proteínas catalogadas em banco de dados que se ligam a metais [46], o elemento Mg mostrou maiores resultados. O Mg é o cátion divalente mais abundante no citosol da planta e é essencial para a função de varias enzimas, incluindo RNA polimerases, ATPases, fosfatases e carboxilases [80]. Além disso foi identificada neste trabalho a enzima *Nucleoside diphosphatase kinase (spot 114)*, o qual é ligante de magnésio [46]. O estudo de Duke *et al* [81] mostrou que a presença de glifosato nas raízes e folhas de plantas de soja pode reduzir a absorção e o translocamento de Mg.

O elemento K é considerado um macronutriente essencial para o crescimento e desenvolvimento da planta, participando na síntese de polipeptídeos no citosol celular, além de estar presente no fluxo osmótico celular [82]. Dos elementos avaliados quantitativamente, foi o que mostrou maior concentração no extrato protéico das sementes. Já o elemento Na é considerado como um elemento benéfico quando presente na planta, já que sua ausência não prejudica o crescimento e desenvolvimento da planta. Em alguns processos celulares, este elemento pode substituir outros metais alcalinos monovalentes, como K, principalmente no fluxo osmótico celular [82].

Tabela 2.8. Concentrações dos elementos Mn, Cu, Rb e Zn nos extratos protéicos das sementes de soja transgênica e não-transgênica (média $\pm$ desvio padrão)

Concentrações / $\mu\text{g L}^{-1}$				
Metal	Limite de deteccção	Limite de quantificação	Soja Transgênica	Soja Não-Transgênica
Mn	0,004	0,013	48,9 $\pm$ 1,5	66,9 $\pm$ 8,8
Cu	0,037	0,124	73,2 $\pm$ 4,6	45,3 $\pm$ 3,3
Rb	0,006	0,018	177,5 $\pm$ 4,6	221,3 $\pm$ 0,5
Zn	0,591	1,968	186,4 $\pm$ 11,2	254,8 $\pm$ 37,9

Como pode ser visto na Tabela 2.8, três elementos estiveram mais presentes na soja não-transgênica: Mn, Rb e Zn. Somente o elemento Cu mostrou maior concentração no extrato protéico da semente de soja transgênica.

O cobre é considerado um elemento essencial, que desempenha papel estrutural e funcional em todos os organismos aeróbicos. Existe um grande numero de enzimas que contem cobre ou proteínas que ligam a cobre e que participam de diversas funções biológicas [82]. Em culturas de soja, a deficiência de íons cobre causa necrose das folhas e, por outro lado, o excesso causa clorose das mesmas [83]. Além disso, foi identificada neste trabalho, uma enzima ligante de cobre, a *glutathione S-transferase* (*spot* 169).

O Mn é necessário em vários processos essenciais as plantas, onde, como exemplo, cita-se sua participação na desintoxicação de radicais livres de oxigênio (Mn-superóxido dismutase), sendo que esta enzima foi identificada neste trabalho (*spot* 165) [82].

O Zn é um componente catalítico essencial de mais de 300 enzimas de plantas, e, por isso, é o elemento cuja pesquisa em bancos de dados de soja gera o maior numero de resultados de proteínas ligadas a ele [82]. Neste trabalho, não foi possível identificar alguma metaloproteína ou proteína ligante de zinco. Como exemplos destas em banco de dados de soja, pode-se citar a metaloenzima *Metalloendoproteinase 1*, uma enzima que utiliza zinco como cofator, ligando-se a um íon de Zn por subunidade de proteína [82].

Dos elementos quantificados e mostrados na Tabela 2.8, apenas o Rb não mostrou resultados em busca de banco de dados de proteínas que se ligam a este metal. Sua atividade na planta ainda é desconhecida, entretanto sua absorção pelas raízes da planta está ligada com a deficiência de boro [82].

Com os resultados mostrados acima, é importante ressaltar que a concentração dos elementos avaliados nas sementes de soja transgênica e não-transgênica é dependente de vários fatores, sendo que os principais incluem características do solo e composição das fontes hídricas utilizadas. Estes fatores podem influenciar no crescimento e desenvolvimento da planta [84]. Levando em consideração que estes fatores são de considerável controle em plantações de

soja, seria possível esperar que as diferenças encontradas nas concentrações de Mg, K, Na, Mn, Cu, Rb e Zn possam estar relacionadas ao processo de modificação genética.

4.4 Quantificação de cobre nos *spots* protéicos diferenciados por ICP-MS.

A escolha do elemento cobre para a quantificação por ICP-MS foi feita levando em consideração que este elemento já foi encontrado com concentrações diferentes nas amostras de soja transgênica e não-transgênica em trabalhos publicados pelo nosso grupo de pesquisa usando diferentes técnicas de separação de proteínas [3,73]. Além disso, ressalta-se a sua presença como ligante em várias proteínas e enzimas [82].

Para a quantificação de cobre nos *spots* protéicos diferenciais obtidos com a técnica de 2D-DIGE, foi feito um “*pool*” dos *spots* obtidos em quatro replicatas de géis 2D-PAGE para ambas as amostras de soja, de modo a aumentarmos a eficiência no processo de quantificação, levando em consideração a sensibilidade da técnica empregada para este fim. A Tabela 2.9 mostra os parâmetros obtidos para a curva analítica de calibração por ICP-MS.

Tabela 2.9. Parâmetros da curva analítica de calibração e algumas figuras de mérito para a quantificação de cobre nos *spots* protéicos diferenciais.

Metal	a (s*)	b (s*)	R	LD ($\mu\text{g L}^{-1}$)	LQ ($\mu\text{g L}^{-1}$)
Cu	-0,0083 (0,0042)	0,3692 (0,0025)	0,9999	0,0128	0,0427

s é o erro da medida

Os resultados obtidos para a quantificação de cobre nos quatro *spots* protéicos diferenciais são mostrados na Tabela 2.10. Ressalta-se que foi feita a subtração das concentrações de cobre dos *spots* protéicos e brancos dos géis com a concentração obtida no branco de decomposição (HNO_3 e H_2O_2). Além disso, também foi feita a subtração da concentração de cobre dos *spots* protéicos diferenciais com a concentração de cobre nos *spots* sem presença de proteínas (branco dos géis), selecionados no gel 2D-PAGE, para cada amostra. Este procedimento foi realizado para descontar a concentração de cobre oriunda do corante usado no processo de revelação dos géis, já que a tendência do metal presente no corante é se distribuir por toda a região do gel.

Como pode ser visto na Tabela 2.10, não foi possível quantificar os *spots* 1 e 2 para nos géis de soja transgênica, já que os valores obtidos estavam abaixo do valor obtido para a concentração de cobre nos brancos de ambos os géis. Situação semelhante foi obtida para os *spots* 3 e 4 para os géis de soja não-transgênica. Para avaliarmos a influência da solução corante no valor de concentração do branco dos géis, foi feita a quantificação desta solução e o valor obtido foi de $16,23 \pm 3,61 \mu\text{g L}^{-1}$. Sendo assim, pode-se verificar a grande influência do teor de metal desta solução na análise dos *spots*, já que este valor obtido é, em média, 18 vezes superior a concentração encontrada nos *spots*.

Tabela 2.10. Concentrações de cobre para os *spots* protéicos diferenciais obtidos com a técnica de 2D-DIGE.

<i>Spot</i> ¹	Concentrações / $\mu\text{g L}^{-1}$	
	Soja	Soja
	Transgênica	Não-Transgênica
1	< C.B ²	$0,749 \pm 0,115$
2	< C.B ²	$0,686 \pm 0,177$
3	$0,867 \pm 0,075$	< C.B ²
4	$0,433 \pm 0,036$	< C.B ²

¹vide figura 2.2; ²Concentração do branco dos géis.

Sendo assim, de acordo com a Tabela 2.10, é possível supor que a concentração de cobre foi maior na soja não-transgênica para os *spots* 1 e 2, enquanto que a concentração de cobre foi maior na soja transgênica para os *spots* 3 e 4. Este resultado mostrou certa tendência, de modo aleatório, com os resultados obtidos com a técnica de 2D-DIGE, relativo as diferenças de expressão destas proteínas, já que dois *spots* (1 e 3) foram mais expressos na soja transgênica e dois *spots* (2 e 4) foram mais expressos na soja não-transgênica.

Com relação à presença de cobre nas proteínas diferenciais, a busca em banco de dados mostrou que nenhuma destas proteínas identificadas se ligam a cobre ou a outro metal qualquer [82]. Sendo assim, sua presença pode ser justificada por interações iônicas destas proteínas com o metal.

5. CONCLUSÕES PARCIAIS

Estudos envolvendo proteômica comparativa em organismos biológicos usando a técnica de 2-D DIGE vêm sendo cada vez mais abordados na literatura, com o intuito de se obter possíveis biomarcadores que respondam a algum estímulo externo. Neste capítulo, a técnica de 2-D DIGE foi utilizada para um estudo proteômico comparativo entre as sementes de soja transgênica e não-transgênica.

Como resultados, foi possível identificar 4 proteínas com diferenças de expressão entre as amostras avaliadas, levando em consideração um nível de, pelo menos, 50 % de variação (1,5). Deste total, 2 foram sobre-expressas (*spots* 1 e 3), ou seja, foram mais expressas na soja transgênica e 2 foram sub-expressas (*spots* 2 e 4) no gel, ou seja, menos expressas na soja transgênica. Ressalta-se que as 4 proteínas com variação de expressão foram identificadas. Com isso, é possível constatar a grande eficiência na utilização desta técnica para estudos proteômicos comparativos.

Os resultados obtidos para a avaliação semi-quantitativa mostrou que os íons Na^+ , Mg^+ , Si^+ , K^+ , Ti^+ , Mn^+ , Zn^+ e Rb^+ foram mais detectados na soja não-transgênica. Já o íon Cu^+ foi mais detectado nas sementes de soja transgênica. Destes elementos, Na, Mg, K, Mn, Zn, Rb e Cu foram selecionados para quantificação e os dados obtidos mostraram concordância com os dados obtidos para a análise semi-quantitativa.

Em relação aos *spots* protéicos diferenciados obtidos com a técnica de 2D-DIGE, foi feita a quantificação de cobre, levando em consideração que este elemento já mostrou resultados diferenciais em amostras de soja em trabalhos reportados na literatura pelo grupo de pesquisa GEPAM. Como resultados obtidos, foi possível quantificar cobre em dois *spots* protéicos (3 e 4) para a soja transgênica e dois *spots* protéicos (1 e 2) para a soja não-transgênica.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] H.A El-Shemy, M.M Khalafalla, K. Fujita, M. Ishimoto. *Improvement of protein quality in transgenic soybean plants*. **Biologia Plantarum** 51 (2007) 277-284.
- [2] Empresa Brasileira de Pesquisa Agropecuária (Embrapa), disponível em <http://www.cnpso.embrapa.br>, acessado em 19 de fevereiro de 2009.
- [3] A. Sussulini, G.H.M.F. Souza, M.N. Eberlin, M.A.Z Arruda, *Comparative metallomics for transgenic and non-transgenic soybeans*. **Journal of Analytical Atomic Spectrometry** 22 (2007) 1501-1506.
- [4] J.S. Garcia, C.S. Magalhães, M.A.Z. Arruda, *Trends in metal-binding and metalloprotein analysis*, **Talanta** 69 (2006) 1-15.
- [5] M. R. Wilkins, C. Pasquali, R.D. Appel. *From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acids analysis*. **Biotechnology** 14 (1996) 61-65.
- [6] S.G. Uzogara, *The impact of genetic modification of human foods in the 21st century: A review*. **Biotechnology Advances** 18 (2000) 179-206.
- [7] V.H. Kim, S.J Choi, H.A. Lee, T.W. Moon, *Quantitation of CP4 5-enolpyruvylshikimate-3-phosphate synthase in soybean by two-dimensional gel electrophoresis*. **Journal of Microbiology and Biotechnology** 16 (2006) 25-31.
- [8] S. Mounicou, J. Szpunar, R. Lobinski, *Metallomics: the concept and methodology*, **Chemical Society Reviews**, 38 (2009) 1119-1138.
- [9] Y.X. Gao, C.Y. Chen, Z.F. Chai, *Advanced nuclear analytical techniques for metalloproteomics*, **Journal of Analytical Atomic Spectrometry**, 22 (2007) 856-866.
- [10] L. Banci, *Molecular dynamics simulations of metalloproteins*, **Current Opinion in Chemical Biology**, 7 (2003) 143-149.

- [11] M.A.O. da Silva, A. Sussulini, M.A.Z. Arruda, *Metalloproteomics as an interdisciplinary area involving proteins and metals*, **Expert Reviews of Proteomics**, 7 (2010) 387-400.
- [12] J. M. Saz, M. L. Marina, *High performance liquid chromatography and capillary electrophoresis in the analysis of soybean proteins and peptides in foodstuffs*. **Journal of Separation Science** 30 (2007) 431-451.
- [13] S. S. Natarajan, C. Xu, P. Cregan, T. J. Caperna, W. M. Garrett, D. Luthria, *Utility of proteomics techniques for assessing protein expression*. **Regulatory Toxicology and Pharmacology**, 54 (2009) 32-36.
- [14] S.S. Natarajan, C. Xu, H. Bae, T.J. Caperna, W. Garret, *Characterization of storage proteins in wild (*Glycine soja*) and cultivated (*Glycine max*) soybean seeds using proteomics analysis.*, **Journal of Agricultural and Food Chemistry**, 54 (2006) 3114-3120.
- [15] C. Xu, W.M. Garrett, J. Sullivan, T.Caperna, S.S. Natarajan, *Separation and identification of soybean leaf proteins by two-dimensional gel electrophoresis and mass spectrometry*. **Phytochemistry** 67 (2006) 2431-2440.
- [16] C. James, Situação Global das Culturas Biotecnológicas/GM Comercializadas: 2009. Em ISAAA Brief 41, ISAAA: Ithaca, New York, 2009; p. 3.
- [17] D.R. Lundry, W.P. Ridley, J.J. Meyer, S.G. Riordan, M.A. Nemeth, W.A. Trujillo, M.L. Breeze, R. Sorbet. *Composition of Grain, Forage, and Processed Fractions from Second-Generation Glyphosate-Tolerant Soybean, MON 89788, Is Equivalent to That of Conventional Soybean (*Glycine max* L.)*. **Journal of Agricultural and Food Chemistry**. 56 (2008) 4611-4622.
- [18] C.A. Moldes, L.O. Medici, O.S. Abrahão, S.M. Tsai, R.A. Azevedo, *Biochemical responses of glyphosate resistant and susceptible soybean plants exposed to glyphosate*, **Acta Physiology Plant** 30 (2008) 469-479.

- [19] M.F. Ocana, P.D. Fraser, R.K.P. Patel, J.M. Halket, P.M. Bramley, *Mass spectrometric detection of CP4 EPSPS in genetically modified soya and maize. Rapid Communications in Mass Spectrometry*, 21 (2007) 319-328.
- [20] M.C. McCann, K. Liu, W.A. Trujillo, R.C. Dobert, *Glyphosate-tolerant soybeans remain compositionally equivalent to conventional soybeans (Glycine max L.) during three years of field testing. Journal of Agricultural and Food Chemistry*. 53 (2005) 5331-5335.
- [21] J. G. Silva Jr., *Eletroforese de Proteínas: Guia Teórico-Prático*, ed. **Interciência**, Rio de Janeiro (2001).
- [22] M. Eravci, S. Fuxius, O. Broedel, S.E. Weist, U. Mansmann, H. Schluter, J. Tiemann, A. Baumgartner, *Improved Comparative Proteome Analysis Based on Two-Dimensional Gel Electrophoresis, Proteomics* 7 (2007) 513-523.
- [23] T. Berklman, e T. Stenstedt, *2-D Electrophoresis Using Immobilized pH Gradients: Principles and Methods, Amershan Biosciences*, Uppsala (1998).
- [24] A. Sussulini, *Avaliação das Alterações Protéicas e Metaloproteicas em Soja Após o Processo de Modificação Genética*, Dissertação de Mestrado, **Universidade Estadual de Campinas**, Campinas (2007).
- [25] C.S. Magalhães, J.S. Garcia, A.S. Lopes, E.C. Figueiredo, M.A.Z. Arruda (editor). *Strategies for sample preparation focusing biomolecules determination/characterization em: Trends in sample preparation*, Nova Science Publishers, New York, 2006.
- [26] B. Domon, R. Aebersold, *Mass Spectrometry and Protein Analysis. Science* 312 (2006) 212-217.
- [27] R. Aebersold, M. Mann, *Mass spectrometry-based proteomics, Nature* 422 (2003) 198-207.

- [28] M. Kinter, N.E. Sherman, **Protein sequencing and identification using tandem mass spectrometry**, John Wiley & Sons, New York (2000).
- [29] R.B. Cole, **Electrospray Ionization Mass Spectrometry: fundamentals, instrumentation, and applications**, John Wiley & Sons, New York, (1997).
- [30] Matrix Science – Mascot Search, disponível em <http://www.matrixscience.com/>.
- [31] O.S. Abrahão, Rastreabilidade de soja Roundup Ready® em produtos agrícolas e derivados: produção de materiais de referência e uso de marcadores ALFP, Tese de Doutorado, **Centro de Energia Nuclear na Agricultura (CENA/USP)**, Piracicaba – SP (2008).
- [32] A. Sussulini, J.S. Garcia, M.F. Mesko, D.P. Moraes, E.M.M. Flores, C.A. Perez, M.A.Z. Arruda, *Evaluation of soybean seed protein extraction focusing metalloprotein analysis*, **Microchimica Acta** 158 (2007) 173–180.
- [33] M.M. Bradford, *Rapid and sensitive method for quantitation of microgram quantities of proteins utilizing principle of protein-dye binding*, **Analytical Biochemistry** 72 (1976) 248-254.
- [34] A.R. Brandão, H.S. Barbosa, M.A.Z. Arruda, *Image analysis of two-dimensional gel electrophoresis for comparative proteomics of transgenic and non-transgenic soybean seeds*. **Journal of Proteomics**, 73 (2010) 1433-1440.
- [35] L. Brechenmacher, J. Lee, S. Sachdev, Z. Song, T.H.N. Nguyen, T. Joshi, N. Oehrle, M. Libault, B. Mooney, D. Xu, B. Cooper, G. Stacey, *Establishment of a Protein Reference Map for Soybean Root Hair Cells*, **Plant Physiology**, 149 (2009) 670-682.
- [36] T. Rabillouda, M. Chevalletb, S. Lucheb, C. Lelongb, *Two-dimensional gel electrophoresis in proteomics: Past, present and future*, **Journal of Proteomics** 73 (2010) 2064-2077.

- [37] B. Wittmann-Liebold, H.R. Graack, T. Pohl, *Two-dimensional gel electrophoresis as tool for proteomics studies in combination with protein identification by mass spectrometry*, **Proteomics** 6 (2006) 4688-4703.
- [38] P. Horvatovich, B. Hoekman, N. Govorukhina, R. Bischoff, *Multidimensional chromatography coupled to mass spectrometry in analysing complex proteomics samples*, **Journal of Separation Science** 33 (2010) 1421-1437.
- [39] M. Bevan, I. Bancroft, E. Bent, K. Love, H. Goodman, C. Dean, R. Bergkamp, W. Dirkse, M. Van Staveren, W. Stiekema, *et al. Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana*, **Nature** 391 (1998) 485-488.
- [40] T. Holzhauser, O. Wackermann, B.K. Ballmer-Weber, C. Bindslev-Jensen, J. Scibilia, L. Perono-Garoffo, S. Utsumi, L.K. Poulsen, S. Vieths, *Soybean (Glycine max) allergy in Europe: Gly m 5 (β -conglycinin) and Gly m 6 (glycinin) are potential diagnostic markers for severe allergic reactions to soy*, **Journal of Allergy and Clinical Immunology** 123 (2009) 452-458.
- [41] S. Natarajan, C. Xu, H. Bae, B.A. Bailey, P. Cregan, T.J. Caperna, W.M. Garrett, D. Luthria, *Proteomic and genetic analysis of glycinin subunits of sixteen soybean genotypes*, **Plant Physiology and Biochemistry** 45 (2007) 436-444.
- [42] H. Zhang, L. Li, G.S. Mittal, *Effects of high pressure processing on soybean beta-conglycin*, **Journal of Food Process Engineering** 33 (2010) 568-583.
- [43] H.B. Krishnan, N.W. Oehrle, S.S. Natarajan. *A rapid and simple procedure for the depletion of abundant storage proteins from legume seeds to advance proteome analysis: A case study using Glycine max*, **Proteomics** 9 (2009) 3174-3188.
- [44] H.B. Krishnan, *Characterization of a soybean [Glycine max (L.) Merr.] mutant with reduced levels of Kunitz trypsin inhibitor*, **Plant Science** 160 (2001) 979-986.

- [45] W. Wang, B. Vinocur, A. Altman, *Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance*, **Planta** 218 (2003) 1-14.
- [46] *Expasy Proteomics Server*, disponível em <http://www.expasy.org>.
- [47] S.A. Campbell, T. Close, *Dehydrins: genes, proteins and associations with phenotypic traits*, **New Phytologist** 137 (1997) 61-74.
- [48] N.A. Yeboah, M. Arahira, V.H. Nong, D. Zhang, K. Kadokura, A. Watanabe, C. Fukazawa, *A class III acidic endochitinase is specifically expressed in the developing seeds of soybean (*Glycine max* [L.] Merr.)*, **Plant Molecular Biology** 36 (1998) 407-415.
- [49] P.L. Gratão, A. Polle, P.J. Lea, R.A. Azevedo, *Making the life of heavy metal-stressed plants a little easier*, **Functional Plant Biology** 32 (2005) 481-494.
- [50] I. Alam, D.G. Lee, K.H. Kim, C.H. Park, S.A. Sharmin, H. Lee, K.W. Oh, B.W. Yun, B.H. Lee, *Proteome analysis of soybean roots under waterlogging stress at an early vegetative stage*, **Journal of Biosciences** 35 (2010) 49-62.
- [51] M. Hiemoria, H. Itoa, M. Kimotoa, H. Yamashitaa, K. Nishizawab, N. Maruyamab, S. Utsumib, H. Tsuji, *Identification of the 23-kDa peptide derived from the precursor of Gly m Bd 28K, a major soybean allergen, as a new allergen*, **Biochimica et Biophysica Acta** 1675 (2004) 174-183.
- [52] J.Y. Kim, W.Y. Kim, K.J. Kwak, S.H. Oh, Y.S. Han, H. Kang, *Glycine-rich RNA-binding proteins are functionally conserved in *Arabidopsis thaliana* and *Oryza sativa* during cold adaptation process*, **Journal of Experimental Botany** 61 (2010) 2317-2325.
- [53] M. Berth, F.M. Moser, M. Kolbe, J. Bernhardt, *The state of the art in the analysis of two-dimensional gel electrophoresis images*, **Applied Microbiology and Biotechnology** 76 (2007) 1223-1243.

- [54] J.E. Bandow, J.D. Baker, M. Berth, C. Painter, O.J. Sepulveda, K.A. Clark, I. Kilty, R.A. VanBogelen, *Improved image analysis workflow for 2-D gels enables large-scale 2-D gel-based proteomics studies – COPD biomarker discovery study*, **Proteomics** 8 (2008) 3030-3041.
- [55] A.W. Dowsey, M.J. Dunn, G.Z. Yang, *The role of bioinformatics in two-dimensional gel electrophoreses*, **Proteomics** 3 (2003) 1567-1596.
- [56] J. Heinemeyer, B. Scheibe, U.K. Schmitz, H.P. Braun, *Blue native DIGE as a tool comparative analyses of protein complexes*, **Journal of Proteomics** 72 (2009) 539 – 544.
- [57] M. Eravci, S. Fuxius, O. Broedel, S.E. Weist, U. Mansmann, H. Schluter, J. Tiemann, A. Baumgartner, *Improved comparative proteome analysis based on two-dimensional gel electrophoresis*, **Proteomics** 7 (2007) 513-523.
- [58] P. Cutler, G. Heald, I.R. White, J. Ruan, *A novel aproach to spot detection for two-dimensional gel electrophoresis images using pixel value collection*, **Proteomics** 3 (2003) 392-401.
- [59] J.F. Timms, R. Cramer, *Difference gel electrophoresis*, **Proteomics** 8 (2008) 4886-4897.
- [60] R. Marouga, S. David, E. Hawkins, *The development of the DIGE system: 2D fluorescence difference gel analysis technology*, **Analytical and Bioanalytical Chemistry** 382 (2005) 669 – 678.
- [61] M. Unlü, M.E. Morgan, J.S. Minden, *Difference gel electrophoresis: a single gel method for detecting changes in protein extracts*. **Electrophoresis** 18 (1997) 2071-2077.
- [62] R. Tonge, J. Shaw, B. Middleton, R. Rowlinson, S. Rayner, J. Young, F. Pognan, E. Hawkins, I. Currie, M. Davison, *Validation and development of*

fluorescence two-dimensional differential gel electrophoresis proteomics technology, **Proteomics** 1 (2001) 377-396.

[63] A. Alban, S.O. David, L. Bjorkesten, C. Andersson, E. Sloge, S. Lewis, I. Currie, *A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard*, **Proteomics** 3 (2003) 36-44.

[64] S. Alvarez, B.M. Berla, J. Sheffield, R.E. Cahoon, J.M. Jez, L.M. Hicks, *Comprehensive analysis of the Brassica juncea root proteome in response to cadmium exposure by complementary proteomic approaches*, **Proteomics** 9 (2009) 2419-2431.

[65] M.D. Carli, M.E. Villani, G. Renzone, L. Nardi, A. Pasquo, R. Franconi, A. Scaloni, E. Benvenuto, A. Desiderio, *Leaf Proteome Analysis of Transgenic Plants Expressing Antiviral Antibodies*. **Journal of Proteome Research** 8 (2009) 838-848.

[66] R. Lobinski, J.S. Becker, H. Haraguchi, B. Sarkar, *Metallomics: Guidelines for terminology and critical evaluation of analytical chemistry approaches (IUPAC Technical Report)*, **Pure and Applied Chemistry** 82 (2010) 493-504.

[67] S. Mounicou, R. Lobinski, *Challenges to metallomics and analytical chemistry solutions*, **Pure and Applied Chemistry** 80 (2008) 2565-2575.

[68] J. Szpunar, *Metallomics: a new frontier in analytical chemistry*. **Analytical and Bioanalytical Chemistry** 378 (2004) 54-56.

[69] R. Lobinski, D. Schaumlöffel, J. Szpunar, *Mass spectrometry in bioinorganic analytical chemistry*, **Mass Spectrometry Reviews** 25 (2006) 255-289.

[70] A.P. OLIVEIRA, *Determinacao de Cu, Zn e Se em soro humano por ICPMS*, **Revista Analytica** 25 (2006) 76-83.

- [71] D. Günther, B. Hattendorf, *Solid sample analysis using laser ablation inductively coupled plasma mass spectrometry*, **Trends in Analytical Chemistry** 24 (2005) 255-265.
- [72] R.E. Russo, X. Mao, H. Liu, J. Gonzalez, S.S. Mao, *Laser ablation in analytical chemistry – a review*, **Talanta** 57 (2002) 425-451.
- [73] L.R.V. Mataveli, P. Pohl, S. Mounicou, M.A.Z. Arruda, J. Szpunar, *A comparative study of element concentrations and binding in transgenic and non-transgenic soybean seeds*, **Metallomics** 2 (2010) 800-805.
- [74] A. Sussulini, H. Kratzin, O. Jahn, C.E.M. Banzato, M.A.Z. Arruda, J.S. Becker, *Metallomics Studies of Human Blood Serum from Treated Bipolar Disorder Patients*, **Analytical Chemistry** 82 (2010) 5859-5864.
- [75] F.M. Verbi, S.C.C. Arruda, A.P.M. Rodriguez, C.A. Pérez, M.A.Z. Arruda. *Metal-binding proteins scanning and determination by combining gel electrophoresis, synchrotron radiation X-ray fluorescence and atomic spectrometry*, **Journal of Biochemical and biophysical methods** 62 (2005) 97-109.
- [76] N.A. Karp, D.P. Kreil, K.S. Lilley, *Determining a significant change in protein expression with DeCyderTM during a pair-wise comparison using two-dimensional difference gel electrophoresis*, **Proteomics** 4 (2004) 1421-1432.
- [77] I.K. Fodor, D.O. Nelson, M. Alegria-Hartman, K. Robbins, R.G. Langlois, K.W. Turteltaub, T.H. Corzett, S.L. McCutchen-Maloney, *Statistical challenges in the analysis of two-dimensional difference gel electrophoresis experiments using DeCyderTM*, **Bioinformatics** 21 (2005) 3733-3740.
- [78] V.H. Thanh, K. Shibasaki, *Major proteins of soybean seeds – subunit structure of beta-conglycinin*. **Journal of Agricultural and Food Chemistry** 26 (1978) 692-695.

- [79] *RCSB Protein Data Bank*, disponível em <http://www.pdb.org>, acessado em 30 de janeiro (2011).
- [80] T.C. FOX, M.L. GUERINOT, *Molecular biology of cation transport in plants. Annual Review of Plant Physiology and Plant Molecular Biology*, 49 (1998) 669-696.
- [81] S.O. Duke, K.C. Vaughn, R.D. Wauchope, *Effects of glyphosate on uptake, translocation, and intracellular localization of metal cations in soybean (Glycine max) seedlings*, **Pesticide Biochemistry and Physiology** 24 (1985) 384-394.
- [82] A.V. BARKER, D.J. PILBEAM, **Handbook of plant nutrition**. CRCpress, 1^a edição (2006), Boca Raton, Estados Unidos, 613 p.
- [83] C.F.B. Coutinho, L.H. Mazo, *Complexos metálicos com o herbicida glifosato: revisão*, **Química Nova** 28 (2005) 1038-1045.
- [84] J. Naozuka, P.V. Oliveira, *Cu, Fe, Mn and Zn Distribution in Protein Fractions of Brazil-Nut, Cupuassu Seed and Coconut Pulp by Solid-Liquid Extraction and Electrothermal Atomic Absorption Spectrometry*, **Journal of the Brazilian Chemical Society** 18 (2007) 1547-1553.

CONSIDERAÇÕES FINAIS

Os estudos apresentados nesta Tese de Doutorado permitiram corroborar a hipótese formulada, a qual foi possível constatar que o processo de modificação genética em sementes de soja pode ocasionar alterações em nível protéico e metaloprotéico, aplicando-se as técnicas de 2D-DIGE no estudo proteômico comparativo e ICP-MS no estudo metaloproteômico comparativo. Além disso, foi possível identificar, com um método simples, a enzima alvo (*CP4 EPSPS*) do processo de modificação genética.

No Capítulo 1, foi possível caracterizar e identificar, usando eletroforese bidimensional em gel de poliacrilamida (2D-PAGE), espectrometria de massas (MALDI-QTOF e ESI-QTOF) e banco de dados relacionados (MSDB, NCBI e *Glycine Max*), um total de 192 proteínas de soja, sendo que 179 proteínas foram identificadas para a faixa de pH de 4 a 7 e 13 proteínas para a faixa de 3 a 10, obtendo-se uma eficiência de 49% de proteínas identificadas em relação ao número de *spots* detectados no gel de eletroforese. Estes resultados permitiram estabelecer um mapa proteômico de referência para sementes de soja transgênica, tipo *Roundap Ready*, constatando a grande eficiência da técnica de 2D-PAGE como componente de separação e da técnica de espectrometria de massas como componente estrutural em estudos proteômicos.

De acordo com a avaliação funcional das proteínas de soja, observou-se, como esperado, que a maioria das proteínas identificadas estão relacionadas com a categoria de função de armazenamento e atividade protéica (96 proteínas), seguido de crescimento/divisão celular (34 proteínas), metabolismo/energia (17 proteínas), transporte (12 proteínas), doença/defesa (08 proteínas), metabolismo secundário (02 proteínas) e transdução de sinais (01 proteína). Além disso, foram identificadas um total de 22 proteínas relacionadas com a categoria não-classificadas/classificação desconhecida.

Levando em consideração que a enzima *CP4 EPSPS* não havia sido identificada no mapa proteômico estabelecido para a semente de soja transgênica, utilizou-se as técnicas de SDS-PAGE e espectrometria de massas para sua

identificação. Com este resultado obtido, é possível afirmar, de maneira inequívoca, que o organismo em questão passou pelo processo de modificação genética. Além disso, destaca-se a simplicidade do método aplicado, onde não houve a necessidade de pré-fracionamento da amostra, utilizando procedimentos cromatográficos, onde a identificação foi possível usando duas enzimas de clivagem, a tripsina e a quimiotripsina.

No Capítulo 2, foi feito um estudo proteômico comparativo entre as sementes de soja transgênica e não-transgênica aplicando-se, a técnica de 2-D DIGE. Como resultados obtidos, foi possível identificar 4 proteínas com diferenças de expressão entre as amostras avaliadas, levando em consideração um nível de, pelo menos, 50 % de variação (1,5). Deste total, 2 foram sobre-expressas (*spots* 1 e 3), ou seja, foram mais expressas na soja transgênica e 2 foram sub-expressas (*spots* 2 e 4) no gel, ou seja, menos expressas na soja transgênica. Ressalta-se que as 4 proteínas com variação de expressão foram identificadas.

Para o estudo metaloproteômico comparativo, foram obtidos resultados para a avaliação semi-quantitativa, onde observou-se que os íons Na^+ , Mg^+ , Si^+ , K^+ , Ti^+ , Mn^+ , Zn^+ e Rb^+ foram mais detectados na soja não-transgênica. Já o íon Cu^+ foi mais detectado nas sementes de soja transgênica. Destes elementos, Na, Mg, K, Mn, Zn, Rb e Cu foram selecionados para quantificação e os dados obtidos mostraram concordância com os dados obtidos para a análise semi-quantitativa.

Em relação aos *spots* protéicos diferenciados obtidos com a técnica de 2D-DIGE, foi feita a quantificação de cobre, levando em consideração que este elemento já mostrou dados de concentração diferenciais entre amostras de soja em trabalhos respotados na literatura pelo grupo de pesquisa GEPAM. Como resultados obtidos, foi possível quantificar cobre em dois *spots* protéicos (3 e 4) para a soja transgênica e dois *spots* protéicos (1 e 2) para a soja não-transgênica.

Finalizando, este trabalho de tese evidencia que uma abordagem multidisciplinar é extremamente necessária aos estudos proteômicos e metaloproteômicos comparativos em organismos biológicos, os quais necessitam da aplicação de diversas ferramentas analíticas para a obtenção dos resultados, bem como uma conexão entre diferentes áreas científicas, como a química

analítica, a bioquímica, fisiologia de plantas e a biotecnologia, para que os resultados obtidos possam ser melhores interpretados e aplicados.

Anexo

Proteínas identificadas em sementes de soja transgênica usando 2D-PAGE e espectrometria de massas.

Spot Gel	Proteína	Acesso	Score	pI/MM spot	Cobertura (%)	Peptídeos seqüenciados
1	trypsin inhibitor KTi1 (Kunitz) - soybean	JQ1091	85	4,20 / 17622	13	R.DTVDGWFNIER.V K.EICPLTVVQSPNELDK.G
2	Seed maturation protein PM41.- Glycine max (Soybean)	Q9SWB2_SOYBN	54	4,46 / 15103	17	K.TATTYESVEGGENR.T
3	2S albumin precursor - soybean	T05710	119	4,58 / 16351	20	K.CCTEMSEL.R.S K.ELINLATMCR.F + Oxidação (M) R.FGPMIQCDLSSDD.-
4	proteinase inhibitor (Bowman-Birk) D-II precursor - soybean	TISYD2	146	4,58 / 15344	48	R.SMPPQCSCEDIR.L + Oxidação (M) R.CLDTNDFCYKPKCK.S K.SDQSSSYDDDEYSKPCCDLCMCTR.S
5	proteinase inhibitor (Bowman-Birk) D-II precursor - soybean	TISYD2	48	4,80 / 14400	24	R.SMPPQCSCEDIR.L + Oxidação (M) R.CLDTNDFCYKPKCK.S
6	Bowman-Birk type proteinase inhibitor D-II precursor related cluster	SOYCIPIB	48	4,86 / 15449	34	R.SMPPQCSCEDIR.L R.CLDTNDFCYKPKCK.S K.SDQSSSYDDDEYSKPCCDLCMCTR.S
7	Napin-type 2S albumin 1 precursor.- Glycine max (Soybean)	Q9ZNZ4_SOYBN	59	5,37 / 11296	14	K.WQQHQQESCR.E K.GINLNPCEHIMEK.I
8	embryonic abundant protein - soybean	T07087	147	5,35 / 16166	24	K.SLEAQQHLEGR.S K.EQLGTEGYQEMGR.K R.KEQLGTEGYQEMGR.K
9	Glycinin A5A4B3.- Glycine max (Soybean).	Q9SB11_SOYBN	63	5,41 / 14342	5	R.NGLHLPSYSPYPR.M K.LNECQLNNLNALEPDHR.V
10	Glycinin A5A4B3.- Glycine max (Soybean).	Q9SB11_SOYBN	102	5,86 / 14342	8	R.NGLHLPSYSPYPR.M K.LNECQLNNLNALEPDHR.V R.VESEGGLIQTWNSQHPELK.C
11	Glycinin A5A4B3.- Glycine max (Soybean)	Q9SB11_SOYBN	63	6,37 / 15034	5	R.NGLHLPSYSPYPR.M K.LNECQLNNLNALEPDHR.V
12	Seed maturation protein PM31.- Glycine max (Soybean)	Q9XET1_SOYBN,	87	5,83 / 17710	40	-.MDWIGAYR.G R.VGDADDITSSLAHAHVDWR.E R.ETDKAHIFR.A K.ILQISGER.V K.EKEDQNDK.W K.EDQNDKWHR.V R.QCGSFLR.R
13	ubiquitin precursor - soybean (fragment)	UQSY	63	7,07 / 11974	25	K.ESTLHLVLR.L K.IQDKEGIPPDQQR.L
14	Seed maturation protein PM31.- Glycine max (Soybean)	Q9XET1_SOYBN	105	6,26 / 18949	24	R.DWCDPSSPFTDLWDPR.R R.VGDADDITSSLAHAHVDWR.E R.SRDWCDPSSPFTDLWDPR.R

15	glycinin chain A2B1a precursor - soybean	FWSYG1	90	7,05 / 20497	9	K.LSAQYGSLR.K K.NNNPFSFLVPPQESQR.R K.NAMFVPHYTLNANSIIYALNGR.A + Oxidação (M)
16	glycinin G1 precursor - soybean	S10851	106	6,76 / 20297	8	R.LSAEFGSLR.K R.VFDGELQEGR.V R.VLIVPQNFVVAAR.S R.SQSDNFEYVSFK.T
17	glycinin chain A2B1a precursor - soybean	FWSYG1	71	6,56 / 20040	7	K.LSAQYGSLR.K K.SQSDNFEYVSFK.T K.NNNPFSFLVPPQESQR.R
18	glycinin chain A1aBx precursor - soybean	FWSYG2	60	6,35 / 20497	7	R.VFDGELQEGR.V R.VLIVPQNFVVAAR.S R.SQSDNFEYVSFK.T
19	glycinin G1 precursor - soybean	S10851	107	6,18 / 20577	8	R.LSAEFGSLR.K R.VFDGELQEGR.V K.FLVPPQESQK.R R.VLIVPQNFVVAAR.S
20	glycinin chain A2B1a precursor - soybean	FWSYG1	53	5,94 / 20417	7	K.LSAQYGSLR.K K.SQSDNFEYVSFK.T K.NNNPFSFLVPPQESQR.R
21	glycinin G1 precursor - soybean	S10851	52	5,77 / 20417	3	R.LSAEFGSLR.K R.VFDGELQEGR.V
22	Glycine G2 A2B1a subunit - Glycine max	AAA33963	89	6,88 / 23624	37	K.LSAQYGSLR.K K.LSAQYGSLRK.N K.NAMFVPHYTLNANSIIYALNGR.A R.VFDGELQEGGVLPVQNFVAAK.S K.SQSDNFEYVSFK.T K.NNNPFSFLVPPQESQR.R
23	glycinin chain A2B1a precursor - soybean	FWSYG1	72	6,64 / 24874	7	K.NNNPFSFLVPPQESQR.R K.NAMFVPHYTLNANSIIYALNGR.A + Oxidação (M)
24	Seed maturation protein PM22.- Glycine max (Soybean)	Q9XER5_SOYBN	118	4,90 / 19165	30	K.VPYSILMSLAK.D K.VNDMAKPEASVTDVDFKR.V R.NPYSTSIPICEINYSFK.S
25	trypsin inhibitor KTi1 (Kunitz) - soybean	JQ1091	145	4,61 / 21938	18	K.NKPLVVQFQK.F R.DTVDGWFNIER.V K.EICPLTVVQSPNELDK.G
26	trypsin inhibitor A (Kunitz) precursor - soybean	TISY	167	4,41 / 20218	13	K.DAMDGWFR.L + Oxidação (M) K.GIGTISSPYR.I K.IGENKDAMDGWFR.L R.NELDKGIGTISSPYR.I
27	trypsin inhibitor A (Kunitz) precursor - soybean	TISY	189	4,28 / 20417	22	R.CPLTVVQSR.N K.GIGTISSPYR.I R.FIAEGHPLSLK.F K.IGENKDAMDGWFR.L R.NELDKGIGTISSPYR.I

28	beta-conglycinin alpha chain precursor - soybean	FWSYBA	123	4,13 / 24199	6	R.QFPFPRPPHQK.E K.GSEEEDEDEDEEQDER.Q R.QEEHEQREEQEWP.R.K
29	Glycinin A5A4B3.- Glycine max (Soybean).	Q9SB11_SOYBN	119	4,18 / 30828	9	K.HFLAQSFNTNEDIAEK.L R.EQDQDQDEDEDEDEDQPR.K R.EQDEDEDEDEDKPRPSRPSQ.GK.R
30	Gly m Bd 30K allergen - Glycine max (Soybean)	O64458_SOYBN	81	4,25 / 34900	8	K.EQYSCDHPPASWDWR.K K.KEQYSCDHPPASWDWR.K K.NSWGEDWGEDGYIWIQR.N
31	Gly m Bd 30K allergen - Glycine max (Soybean)	O64458_SOYBN	50	4,33 / 34480	8	K.EQYSCDHPPASWDWR.K K.NSWGEDWGEDGYIWIQR.N
32	Glycinin A5A4B3.- Glycine max (Soybean).	Q9SB11_SOYBN	143	4,32 / 31968	13	K.HFLAQSFNTNEDIAEK.L R.EQDQDQDEDEDEDEDQPR.K K.QGQHQQEEEEEGGSVLGFSK.H R.EQDEDEDEDEDKPRPSRPSQ.GK.R
33	Seed maturation protein PM24.- Glycine max (Soybean)	Q9SEL0_SOYBN	66	4,46 / 31488	14	K.AATLQDAEGVASAEAR.N R.NNPDATAATPGGVAASVAAAAR.L
34	Seed maturation protein PM25.- Glycine max (Soybean).	Q9SEK9_SOYBN	226	4,65 / 31015	50	R.NEQAGLVGHR.D K.AVDQSDASAIQAAAEVR.A K.AATLQDAEGVACAEVR.N R.NNPDATAATPGGVAASVAAAAR.L R.VLGQTQPGGAASVMQSAATR.N + Oxidação (M) K.ENAITIGEALATAQTVGQK.A R.IITEAVGGQVVEQYVEATPVEAGR.S
35	Glycinin A3B4 subunit.- Glycine max (Soybean).	Q7GC77_SOYBN	56	4,64 / 28738	9	R.SQQQLQDSHQK.I K.HFLAQSFNTNEDIAEK.L K.QGQHQQEEEEEGGSVLGFSK.H R.KQGQHQQEEEEEGGSVLGFSK.H
36	glycinin chain A2B1a precursor - soybean	FWSYG1	47	4,84 / 29650	2	R.RFYLAGNQEQEFLK.Y
37	soybean agglutinin - soybean	1SBF	70	5,03 / 33554	13	K.TSLPEWVR.I -.AETVSFSWNK.F K.VDENGTPKPSSLGR.A R.NSWDPPNPHIGINVNSIR.S
38	Seed maturation protein PM36 - Glycine max	AAD51624	89	5,18 / 27963	16	R.WGNEAFGK.Y R.HPLIISIR.D K.YCQSLQNIANR.C K.TWLAQDYLFVR.A
39	In2-1 protein.- Glycine max (Soybean)	Q9FQ95_SOYBN	130	5,27 / 30991	54	M.ASLGVRPVLPPPLTSISDPPPLFDGTTR.L R.LYISYSCPYAQR.V K.INLVPINLQDRPAWYK.E K.VYPENKVPSLEHNGK.V K.GDAVQQASPAFEYLENALGK.F R.FQIVFAEVFK.H K.LATWFEELNK.L K.LNAYTETR.V R.VDPQEIVDLFK.K

40	glycinin G1 precursor - soybean	S10851	48 54	5,31 / 34255	10	R.RFYLAGNQEQEFLK.Y R.RPSYTNGPQEIYIQQKG.G K.GIFGMIYPGCPSTFEEPQQPQQR.G
41	soybean agglutinin, fragment 2 - soybean	1SBD2	58	5,20 / 32357	22	K.TSLPEWVR.I R.NSWDPPNPHIGINVNSIR.S
42	soybean agglutinin - soybean	1SBF	52	5,23 / 33758	16	K.TSLPEWVR.I R.ALYSTPIHIWDK.E R.NSWDPPNPHIGINVNSIR.S
43	glycinin G1 precursor - soybean	S10851	91	5,47 / 34355	18	R.EQPQQNECQIQK.L K.LNALKPDNR.I R.RPSYTNGPQEIYIQQKG.G K.GIFGMIYPGCPSTFEEPQQPQQR.G R.RFYLAGNQEQEFLK.Y R.FYLAGNQEQEFLK.Y K.NLQGENEGEDKGAIIVTK.G
44	soybean agglutinin - soybean	1SBF	50	5,49 / 31679	24	K.TSLPEWVR.I K.VDENGTPKPSSLGR.A K.QPNMILQGDAIVTSSGK.L + Oxidação (M) R.NSWDPPNPHIGINVNSIR.S
45	Allergen Gly m Bd 28K (Fragment).- Glycine max (Soybean)	Q9AVK8_SOYBN	69	5,57 / 26013	2	K.LGFIYDDELAER.R
46	Dehydrin-like protein.- Glycine max (Soybean).	Q39805_SOYBN	79	5,73 / 24389	42	K.QTDEYGNPVHAASVTYVATR.T R.TAAGGYSDINKQHDTTNAYGVDTR.Q R.HHGTGTVYGIDTDR.Q R.QQHGTGGYAGDTGR.Q R.QHGNIGGPYYGTNTADTGTGPR.S
47	Dehydrin.- Glycine max (Soybean).	Q70EM0_SOYBN	161	6,41 / 26578	31	R.QQHGTGGYAGDTGR.Q K.HHGITGGYNDDTNR.H K.QTDEYGNPVHAASVTYVATR.T R.QHGNIGGPYYGTNTADTGTGPR.S
48	Lea protein.- Glycine max (Soybean).	Q7XAW0_SOYBN	129	6,47 / 29305	22	R.KVDEYGNVER.Q R.DHGTTGVYGLNTDR.H K.QHDATGVYPGIDIGR.D K.QHDITGVYPEKDTGR.H
49	Dehydrin.- Glycine max (Soybean)	Q70EM0_SOYBN	70	6,21 / 29631	24	K.HHGITGGYNDDTNR.H K.QTDEYGNPVHAASVTYVATR.T R.QHGNIGGPYYGTNTADTGTGPR.S
50	Dehydrin.- Glycine max (Soybean)	Q70EL8_SOYBN	57	6,23 / 30991	36	R.HHGTGTVYGIDTDR.Q K.QTDEYGNPVHAASVTYVATR.T R.QHGNIGGPYYGTNTADTGTGPR.S R.TAAGGYSDINKQHDTTNAYGVDTR.Q
51	Triosephosphate isomerase (EC 2.3.1.16).- Glycine max (Soybean)	Q6GW08_SOYBN	79	6,29 / 30508	20	K.FFVGGNWK.C K.VIACIGETLEQR.E K.WVHDNVSAEVAASVR.I K.SLLRPDFHVSAQNCWVR.K

52	Lea protein.- Glycine max (Soybean).	Q7XAW0_SOYBN	47	6,35 / 29941	19	K.EAVVGKPHHEEVYAK.E R.TNERDEGYDLHDEANFK.Q
53	Seed maturation protein PM34.- Glycine max (Soybean)	Q9LLQ6_SOYBN	78	6,77 / 36651	28	K.FPPQQQQTQPGK.E K.EHAMTPVPQFTSPDYKPSNK.L K.EHAMTPVPQFTSPDYKPSNK.L Oxição (M) K.IALVTGGDSGIGR.A R.AVCNLFALLEGATVAFTYVK.G R.TNIFYFFMAR.H K.GAIVAYTR.G
54	Alcohol dehydrogenase 1 (Fragment).- Glycine max (Soybean)	Q8LJR2_SOYBN	55	6,86 / 48245	7	R.ILFNSLCR.S K.GTFYGHYRPR.T K.SSESNLCELLR.I
55	Putative glyceraldehyde-3-phosphate dehydrogenase	Q9FX54_ARATH	56	6,73 / 43003	14	R.VPTVDVSVVDLTVR.L K.LVSWYDNEWGYSSR.V K.GILGYTEDDVSTDFVGDNR.S
56	Alcohol dehydrogenase 1 (Fragment).- Glycine max (Soybean)	Q8LJR2_SOYBN	96	6,66 / 45813	10	R.ILFNSLCR.S K.GTFYGHYRPR.T K.DHNKPVQEVIAEMTNGGVDR.A
57	Alcohol-dehydrogenase (Fragment).- Glycine max (Soybean)	Q9ZT38_SOYBN	61	6,16 / 47002	8	K.GQTPLFPR.I K.FGVNEFVNPK.D K.THPVNFLNER.T
58	35 kDa seed maturation protein.- Glycine max (Soybean).	Q9ZTY1_SOYBN	86	6.10 / 39991	20	K.EAVVGKPHHEEVYAK.E R.AMDGDMEEEGVHLHVER.R + 2 Oxição (M) R.TNERDEGYDLHDEANFK.Q K.EQHQHDENKPGVIGSMFR.A + Oxição (M)
59	Cytosolic malate dehydrogenase (EC 2.3.1.16).- Glycine max (Soybean)	Q6RIB6_SOYBN	84	6,02 / 43527	28	R.VLVTGAAGQIGYALVPMIAR.G Oxição (M) K.MELVDAAFPLK.G Oxição (M) K.VLVVANPANTNALILK.E R.ALQGISER.L R.ELIADDAWLNGEFITTQQR.G R.KLSSALSAASAACDHIR.D K.LSSALSAASAACDHIR.D
60	35 kDa seed maturation protein.- Glycine max (Soybean)	Q9ZTY1_SOYBN	56	5,97 / 41667	15	K.EAVVGKPHHEEVYAK.E R.TNERDEGYDLHDEANFK.Q K.EQHQHDENKPGVIGSMFR.A
61	Isoflavone reductase homolog 2.- Glycine max (Soybean)	Q9SDZ0_SOYBN	101	5,91 / 40329	32	K.ILFIGGTGYIGK.F K.AGHPTFLLVR.E K.KFYPSEFGNDVDR.T K.FYPSEFGNDVDR.T K.AVFNKEEDIGTYTINSVDDPR.T R.IYVPEEQLLK.Q K.QIEESAPPVNVILSINHSSYVK.G K.YITVDEYLNQFV.-

62	Isoflavone reductase homolog 2.- Glycine max (Soybean).	Q9SDZ0_SOYBN	73	5,70 / 39391	14	K.AGHPTFLLVR.E K.FYPSEFGNDVDR.T K.KFYPSEFGNDVDR.T K.AVFNKEEDIGTYTINSVDDPR.T
63	Glycinin G1 precursor [Contains: Glycinin A1a subunit; Glycinin Bx subunit] related cluster	TA439_3847	69	5,55 / 39391	8	R.RFYLAGNQEQEFLK.Y K.NLQGENEGEDKGAIVTVK.G
64	glycinin chain A1aBx precursor - soybean	FWSYG2	54	5,47 / 38799	9	R.RFYLAGNQEQEFLK.Y K.NLQGENEGEDKGAIVTVK.G R.RPSYTNGPQEIYIQQGK.G
65	glycinin chain A1aBx precursor - soybean	FWSYG2	64	5,36 / 37757	6	R.RFYLAGNQEQEFLK.Y R.RPSYTNGPQEIYIQQGK.G
66	glycinin G1 precursor - soybean	S10851	105	5,26 / 36631	10	R.FYLAGNQEQEFLK.Y R.RFYLAGNQEQEFLK.Y R.RPSYTNGPQEIYIQQGK.G K.GIFGMIYPGCPSTFEPPQQPQQR.G
67	glycinin chain A1aBx precursor - soybean	FWSYG2	62	5,15 / 35864	6	R.RFYLAGNQEQEFLK.Y R.RPSYTNGPQEIYIQQGK.G
68	glycinin G1 precursor - soybean	S10851	55	5,06 / 36631	10	R.FYLAGNQEQEFLK.Y R.RFYLAGNQEQEFLK.Y R.RPSYTNGPQEIYIQQGK.G K.GIFGMIYPGCPSTFEPPQQPQQR.G
69	glycinin chain A2B1a precursor - soybean	FWSYG1	53	4,97 / 38216	5	K.EAFGVNMQIVR.N R.RFYLAGNQEQEFLK.Y
70	Glycinin A5A4B3.- Glycine max (Soybean).	Q9SB11_SOYBN	156	4,66 / 38565	13	K.HFLAQSFNTNEDIAEK.L R.EQDQDQDEDEDEDQPR.K K.QGQHQQEEEEEGGSLSGFSK.H R.EQDEDEDEDEDKPRPSRPSQKG.R
71	Glycinin A3B4 subunit.- Glycine max (Soybean).	Q7GC77_SOYBN	82	4,86 / 44593	14	K.HFLAQSFNTNEDTAEK.L K.FNECQLNNLNALEPDHR.V K.QGQHQQEEEEEGGSLSGFSK.H R.KQGQHQQEEEEEGGSLSGFSK.H K.GAIGFAFPGCPETFEKPQQQSSR.R
72	glycinin G5 precursor - soybean	FWSYG3	83	5,00 / 43527	10	K.HFLAQSFNTNEDTAEK.L K.FNECQLNNLNALEPDHR.V K.GAIGFAFPGCPETFEKPQQQSSR.R
73	Glycinin A3B4 subunit.- Glycine max (Soybean).	Q7GC77_SOYBN	74	5,08 / 44459	19	K.HFLAQSFNTNEDTAEK.L K.FNECQLNNLNALEPDHR.V K.QGQHQQEEEEEGGSLSGFSK.H K.GAIGFAFPGCPETFEKPQQQSSR.R R.VFYLAGNPDIHPETMQQQQQK.S + Oxidação(M)
74	Cytosolic glutamine synthetase alpha (Fragment).- Glycine max (Soybean)	Q94KI3_SOYBN	60	5,15 / 45486	37	R.HKEHIAAYGEGNER.R R.HETADINTFSWGVANR.G

75	glycinin G4 precursor - soybean	S11004	79	4,86 / 54498	16	K.LNECQLNNLNALEPDHR.V R.NGLHLPSYSPYPR.M K.GALGVAIPGCPETFEEPQEQRNR.R K.GALGVAIPGCPETFEEPQEQRNR.R K.QGQHQQEEEEEGSVLSGFSK.H K.HFLAQSFNTNEDIAEK.L
76	Glycinin A5A4B3.- Glycine max (Soybean).	Q9SB11_SOYBN	103	4,81 / 54498	12	R.NGLHLPSYSPYPR.M K.HFLAQSFNTNEDIAEK.L K.LNECQLNNLNALEPDHR.V K.GALGVAIPGCPETFEEPQEQRNR.R
77	Beta-conglycinin alpha prime subunit.- Glycine max (Soybean).	Q4LER6_SOYBN	131	4,15 / 84770	14	K.LFEITPEKNPQLR.D R.VPAGTTYVVNPDNDENLR.M R.QPHQEEHEQKEEHEWHR.K
78	beta-conglycinin alpha chain precursor - soybean	FWSYBA	79	4,70 / 104626	10	R.SPQLQNL.R.D K.NPFLFGSNR.F K.TISSEDKPFNL.R.S K.EQQQEQQQEEQPLEVR.K R.VPSGTTYVVNPDNNENLR.L
79	beta-conglycinin alpha chain precursor - soybean	FWSYBA	303	4,84 / 76620	16	R.QFPFPRPPHQK.E K.FFEITPEKNPQLR.D R.QFPFPRPPHQKEER.N K.EQQQEQQQEEQPLEVR.K R.QEEEEHQREEQEWPR.K R.VPSGTTYVVNPDNNENLR.L R.FESFLLSSTEAAQSYLQGFSR.N
80	beta-conglycinin alpha chain precursor - soybean	FWSYBA	215	5,04 / 59974	15	K.NPFLFGSNR.F R.QFPFPRPPHQK.E K.FFEITPEKNPQLR.D R.QFPFPRPPHQKEER.N K.EQQQEQQQEEQPLEVR.K R.VPSGTTYVVNPDNNENLR.L
81	beta-conglycinin alpha chain precursor - soybean	FWSYBA	139	5,04 / 59974	13	K.NPFLFGSNR.F R.LITLAIPVKNPGR.F R.KTISSEDKPFNL.R.S K.FFEITPEKNPQLR.D K.EQQQEQQQEEQPLEVR.K R.VPSGTTYVVNPDNNENLR.L
82	beta-conglycinin alpha chain precursor - soybean	FWSYBA	197	5,11 / 59800	13	K.NPFLFGSNR.F K.NKNPFLFGSNR.F R.LITLAIPVKNPGR.F K.TISSEDKPFNL.R.S K.FFEITPEKNPQLR.D K.EQQQEQQQEEQPLEVR.K R.VPSGTTYVVNPDNNENLR.L
83	beta-conglycinin alpha chain precursor - soybean	FWSYBA	151	5,16 / 59627	11	R.LITLAIPVKNPGR.F K.FFEITPEKNPQLR.D K.EQQQEQQQEEQPLEVR.K R.VPSGTTYVVNPDNNENLR.L R.FESFLLSSTEAAQSYLQGFSR.N

84	Beta-amylase - Glycine max (Soybean).	AMYB_SOYBN	86	5,36 / 59110	10	R.HHAILNFTCLEMR.D + Oxidação (M) K.EYLTVGVDNEPIFHGR.T K.VENHAAELTAGYYNLNDR.D
85	Actin (Fragment).- Glycine max (Soybean)	Q96445_SOYBN	65	5,25 / 51006	9	K.SYELPDGQVITIGDER.F K.LSYIALDYEQELETAR.T
86	Actin (Fragment).- Glycine max (Soybean)	Q96447_SOYBN	103	5,34 / 50281	16	R.AVFPSIVGRPR.H K.IWHHTFYNELR.V K.SYELPDGQVITIGSER.F K.YPIEHGIVSNWDDMEK.I
87	Beta-conglycinin beta subunit (Fragment).- Glycine max (Soybean).	Q50JD8_SOYBN	246	5,42 / 55616	22	R.NFLAGEKDNVVR.Q R.SRNPIYSNNFGK.F K.QKQEEEEPLEVQR.Y R.KTISSEDEPFNL.R.S K.FFEITPEKNPQLR.D R.SSNSFQTLFENQNGR.I R.QVQELAFPGSAQDVER.L
88	Beta-conglycinin beta subunit (Fragment).- Glycine max (Soybean).	Q50JD8_SOYBN	210	5,51 / 55134	22	R.NFLAGEKDNVVR.Q R.SRNPIYSNNFGK.F K.QKQEEEEPLEVQR.Y R.KTISSEDEPFNL.R.S K.FFEITPEKNPQLR.D R.SSNSFQTLFENQNGR.I R.QVQELAFPGSAQDVER.L
89	Beta subunit of beta conglycinin (Fragment).- Glycine max (Soybean).	O22121_SOYBN	110	5,63 / 55134	16	R.EDENNPFFYFR.S R.SRNPIYSNNFGK.F K.VREDENNPFFYFR.S K.FFEITPEKNPQLR.D R.SSNSFQTLFENQNGR.I R.QVQELAFPGSAQDVER.L
90	Beta subunit of beta conglycinin (Fragment).- Glycine max (Soybean).	O22121_SOYBN	175	5,73 / 54974	16	R.EDENNPFFYFR.S R.SRNPIYSNNFGK.F K.VREDENNPFFYFR.S K.FFEITPEKNPQLR.D R.SSNSFQTLFENQNGR.I R.QVQELAFPGSAQDVER.L
91	Beta subunit of beta conglycinin (Fragment).- Glycine max (Soybean).	O22121_SOYBN	194	5,87 / 54656	16	R.EDENNPFFYFR.S R.SRNPIYSNNFGK.F K.VREDENNPFFYFR.S K.FFEITPEKNPQLR.D R.SSNSFQTLFENQNGR.I R.QVQELAFPGSAQDVER.L
92	Beta subunit of beta conglycinin (Fragment).- Glycine max (Soybean).	O22121_SOYBN	207	5,99 / 55134	16	R.EDENNPFFYFR.S R.SRNPIYSNNFGK.F K.VREDENNPFFYFR.S K.FFEITPEKNPQLR.D R.SSNSFQTLFENQNGR.I R.QVQELAFPGSAQDVER.L

93	Sucrose binding protein homolog S-64.- Glycine max (Soybean).	Q9SP11_SOYBN	53	6,34 / 57922	4	K.LLQGIENFR.L K.FEEFFGPGGR.D
94	seed biotin-containing protein LEA - soybean	T07064	113	6,41 / 74082	24	K.MATHFEHLAEQAK.E Oxidação (M) K.ESDITAGKDT PQGSIEALQAGER.V K.DTPQGSIEALQAGER.V K.GGGTQKPIREEER.E K.QAASETLNNTTTQTAQEKSAAQAK.N K.ATAVGWAAAHFSAEK.T K.AAAHVVEGAAGYAGHK.A K.KPSQPQEAEEERPSEGIGETVR.Q R.QYAQKPKPSEGNPQK.E K.EGTGSIVFTAIGETVSSAGEK.V
95	lipoxygenase-3 - soybean (fragments)	1BYT	48	6,37 / 93336	4	K.LLHPHYR.D R.HASDEVYLGGER.D R.EFDSFDEVHGLYSGGIK.L
96	seed biotin-containing protein LEA- soybean	T07064	81	6,30 / 74082	10	K.GGGTQKPIREEER.E K.AAAHVVEGAAGYAGHK.A K.DTPQGSIEALQAGER.V K.EGTGSIVFTAIGETVSSAGEK.V
97	seed biotin-containing protein LEA - soybean	T07064	109	6,21 / 74799	30	-.MASEQLAR.R Oxidação (M) K.MATHFEHLAEQAK.E Oxidação (M) K.ESDITAGKDT PQGSIEALQAGER.V K.DTPQGSIEALQAGER.V R.ETHELGAHFESLADK.V R.DVGKFEMR.T Oxidação (M) R.EVIGRTEK.E K.GGGTQKPIREEER.E R.SAWEQISNYSQATQGVK.E K.QAASETLNNTTTQTAQEKSAAQAK.N R.DKATAVGWAAAHFSAEK.T K.ATAVGWAAAHFSAEK.T K.AAAHVVEGAAGYAGHK.A K.KPSQPQEAEEERPSEGIGETVR.Q R.QYAQKPKPSEGNPQK.E
98	62K sucrose-binding protein precursor - soybean	JQ1730	94	6,16 / 69924	8	R.AHTFVSPR.H K.LLQGIENFR.L K.RPTISNGYGR.L K.HQCQQQQYTEGDKR.V
99	glycinin G1 precursor - soybean	S10851	81	6,10 / 63006	17	R.RPSYTNGPQEIIYIQGK.G R.RFYLAGNQEQEFLK.Y R.LSAEFGSLR.K R.ALIQVVNCNGER.V R.VFDGELQEGR.V R.VLIVPQNFFVVAAR.S R.SQSDNFEYVSFK.T

100	Beta subunit of beta conglycinin (Fragment).- Glycine max (Soybean).	O22121_SOYBN	139	5,97 / 63189	10	K.VREDENNPFFYFR.S R.EDENNPFFYFR.S R.SSNSFQTLFENQNGR.I R.IPAGTTYLVNPHDHQNLK.I R.KTISSEDEPFNL.R.S R.SRNPIYSNNFGK.F K.FFEITPEKNPQLR.D K.QKQEEEEPLEVQR.Y R.NFLAGEKDNVVR.Q R.QVQELAFPGSAQDVER.L
101	62K sucrose-binding protein precursor - soybean	JQ1730	152	6,07 / 70601	11	R.AHTFVSPR.H K.LLQGIENFR.L K.RPTISNGYGR.L R.LTEVGPDDDEKSWLQR.L K.HQCQQQQQYTEGDKR.V
102	62K sucrose-binding protein precursor - soybean	JQ1731	126	6,00 / 69924	8	K.LLQGIENFR.L K.RPTISNGYGR.L K.ESFFFPFELPR.E K.HQCQQQQQYTEGDKR.V
103	62K sucrose-binding protein precursor - soybean	JQ1732	144	5,93 / 70601	10	R.AHTFVSPR.H K.LLQGIENFR.L K.RPTISNGYGR.L K.ESFFFPFELPR.E K.HQCQQQQQYTEGDKR.V
104	62K sucrose-binding protein precursor - soybean	JQ1733	62	5,85 / 70601	6	R.AHTFVSPR.H K.ESFFFPFELPR.E R.LTEVGPDDDEKSWLQR.L
105	62K sucrose-binding protein precursor - soybean	JQ1734	111	5,70 / 70626	7	K.ESFFFPFELPR.E R.HFDSEVVFFNIK.G R.LTEVGPDDDEKSWLQR.L
106	62K sucrose-binding protein precursor - soybean	JQ1735	56	5,68 / 70262	6	R.AHTFVSPR.H K.RPTISNGYGR.L R.LTEVGPDDDEKSWLQR.L
107	62K sucrose-binding protein precursor - soybean	JQ1736	67	5,55 / 69589	6	R.AHTFVSPR.H K.RPTISNGYGR.L R.LTEVGPDDDEKSWLQR.L
108	Enolase (EC 2.3.1.16).- Glycine max (Soybean).	Q6RIB7_SOYBN	103	5,47 / 60850	18	K.VNQIGSVTESIEAVR.M R.AAVPSGASTGIYEALR.D R.IEEELGAEAVYAGANFR.T K.TYDLNFKEDNNDGSQR.I R.GNPTVEVDLTCSGTFAR.A
109	62K sucrose-binding protein precursor - soybean	JQ1736	93 75	5,45 / 69924	10	R.AHTFVSPR.H K.LLQGIENFR.L K.RPTISNGYGR.L K.ESFFFPFELPR.E R.LTEVGPDDDEKSWLQR.L

110	Beta-conglycinin alpha prime subunit.- Glycine max (Soybean).	Q4LER6_SOYBN	77	5,21 / 73726	8	R.QQQEEQPLEVR.K K.TISSEDKPFNLR.S K.LFEITPEKNPQLR.D R.VPAGTTYVVNPDNDENLR.M
111	Beta-conglycinin alpha prime subunit.- Glycine max (Soybean).	Q4LER6_SOYBN	166	5,17 / 81568	12	R.QQQEEQPLEVR.K K.LFEITPEKNPQLR.D K.VLFGREEGQQQGEER.L R.VPAGTTYVVNPDNDENLR.M R.QPHQEEEEHQKEEHEWHR.K
112	Beta-conglycinin alpha prime subunit.- Glycine max (Soybean).	Q4LER6_SOYBN	97	5,16 / 73020	5	K.LFEITPEKNPQLR.D R.VPAGTTYVVNPDNDENLR.M
113	Dehydrin.- Glycine max (Soybean)	Q70EL8_SOYBN	73	6,07 / 29728	31	R.QQHGTGGYAGDTGR.Q R.HHGTGTVYIDTD.R.Q K.QTDEYGNPVHAASVTYVATR.T R.QHGNIGGPYYGTNTADTGTGPR.S
114	Nucleoside diphosphatase kinase (EC 2.7.4.6) Glycine max (Soybean)	Q8GV24_SOYBN	151	6,37 / 17295	18	K.LVTVDRPFAEK.H K.IIGATNPAQSEPGTIR.G R.KIIGATNPAQSEPGTIR.G
115	embryonic abundant protein - soybean	JQ2273	79	6,29 / 16309	11	K.SLEAQQHLEGR.S
116	Allergen Gly m Bd 28K (Fragment) - Glycine max (Soybean).	Q9AVK8_SOYBN	146	6,93 / 24289	7	K.QGDVFVVR.Y K.NKPQFLAGAASLLR.T R.DGPLEFFGFSTSAR.K
117	Bowman-Birk type proteinase inhibitor C-II precursor related cluster	TA5957_3847	35	4,42 / 16270	7	R.CLDTTDFCYKPCK.S
118	Late embryogenic abundant protein related cluster	TA9647_3847	40	5,21 / 20432	6	K.SLEAQEHLEGR.S
119	Formate dehydrogenase related cluster	TA1261_3847	75	6,56 / 51810	12	K.EGQLASQYR.- K.GEWNVAGIAHR.A R.NFLPGYHQAVK.G K.GEDFPEQNYIVK.E
120	Formate dehydrogenase related cluster	TA1261_3847	55	6,47 / 52374	12	K.EGQLASQYR.- K.GEWNVAGIAHR.A R.NFLPGYHQAVK.G K.GEDFPEQNYIVK.E
121	Hypothetical protein related cluster	AW307413	43	5,77 / 34307	9	R.FLHVYR.V K.DPDGYAFELIQR.S
122	PreproMP73 related cluster	TA28641_3847	118	5,40 / 25098	13	K.ELPYHLR.F K.LTWANDDGTSTIR.L K.EGSVFIQSNLEAER.R
123	Glycinin G1 precursor	TA439_3847	37	4,86 / 31500	5	R.FYLAGNQEQFLK.Y R.RFYLAGNQEQFLK.Y R.RPSYTNQGPQEIIQQGK.G
124	Class III acidic endochitinase precursor related cluster	TA2420_3847	40	4,33 / 29139	3	R.YHDVLNYHSDQIK.D

125	Putative dehydrin (Fragment).- Glycine max (Soybean)	Q5K3Q9_SOYBN	49	6,43 / 27691	21	R.QQHGGTTGGYAGDTGR.Q R.HHGGTTGVYGIDTDR.Q K.QHDTTNAYGVDTGR.Q
126	Beta-conglycinin alpha prime subunit.- Glycine max (Soybean)	Q4LER6_SOYBN	153	4,53 / 82304	7	K.NQYGHVR.V R.SQQLQNL.R.D R.QQQEEQPLEVR.K R.EHPRPHQPHQK.E K.TISSEDKPFNLR.S
127	glycinin chain A1aBx precursor - soybean	FWSYG2	81	5,65 / 15022	4	R.VFDGELQEGR.V R.VLIVPQNFVVAAR.S
128	Dehydrin related cluster	TA266_3847	49	5,53 / 19496	11	R.QQHGGTTGGYAGDTGR.Q R.HHGGTTGVYGIDTDR.Q K.QHDTTNAYGVDTGR.Q
129	LEA protein (Fragment)	Q4U4M1_ARAHY	56	5,22 / 22666	11	K.SLEAQEHLAEGR.S
130	glycinin chain A1aBx precursor - soybean	FWSYG2	49	5,57 / 38756	6	R.FYLAGNQEQEFLK.Y R.RFYLAGNQEQEFLK.Y R.RPSYTNGPQEIIYQQGK.G
131	Beta-amylase - Glycine max (Soybean).	AMYB_SOYBN	103	5,32 / 61399	5	R.DGYRPIAR.M K.VSGIHWYK.V K.LSMFGVTYLR.L
132	Alcohol dehydrogenase 1 related cluster	TA1574_3847	56	5,69 / 32277	3	K.GQTPLFPR.I K.ILFNSLCR.T
133	Sucrose-binding protein 2.- Glycine max (Soybean).	Q84V19_SOYBN	109	5,19 / 62068	5	R.AHTFVSPR.H K.LLQGIENFR.L K.FEEFFGPGGR.D
134	beta-conglycinin alpha chain precursor - soybean	S20007	175	4,73 / 52185	7	R.SPQLQNL.R.D R.EEQEWPR.K R.KQEEDEDEEQQR.E K.GSEEEDEDEEQDER.Q
135	beta-conglycinin alpha chain precursor - soybean	FWSYBA	260	5,07 / 75813	9	R.SPQLQNL.R.D K.NPFLFGSNR.F R.QFPFPRPPHQB.E K.TISSEDKPFNLR.S K.EQQQEQQQEQPLEVR.K
136	Beta-conglycinin alpha-subunit Glycine max (Soybean).	Q948X9_SOYBN	131	5,27 / 30183	7	K.FFEITPEK.N R.ESYFVDAQPK.K K.TISSEDKPFNLR.S K.EQQQEQQQEQPLEVR.K
137	18.5 kDa class I heat shock protein related cluster	TA4152_3847	54	6,37 / 19268	6	R.VLQISGER.N R.FRLPENAK.V
138	35 kDa seed maturation protein Glycine max (Soybean).	Q9ZTY1_SOYBN	90	5,71 / 31308	9	K.LWHSWAYEIK.S K.FWCTWQVDR.G R.QIETHHYCSR.L

139	Hypothetical protein P0456A01.21 related cluster	TA3824_3847	75	6,68 / 28927	14	R.MGIDYEEER.K R.MGIDYEEER.K + Oxidação (M) K.VFHFWQVDK.G R.LWHSPLYEVK.S R.DGQIYDHLQSSAER.M
140	cytosolic glutamine synthetase (Cucumis melo)	gi 113171384	198	5,31 / 49752	13	K.GYFEDR.R R.DIVDAHVK.A R.EEGGYEVK.K K.EHIAAYGEGNER.R K.SMREEGGYEVK.K
141	Glycinin A1aBx	gi 225651	91	6,00 / 17928	4	R.VFDGELQEGR.V R.SQSDNFEYVSFK.T
142	glycine-rich RNA-binding protein [Glycine max]	gi 18643	242	5,08 / 17569	6	R.NITVNEAQS.R.G
143	napin-type 2S albumin 1 precursor [Glycine max]	gi 4097894	367	5,43 / 14128	40	K.SPICQCK.A R.ELMNLAIR.C K.CCSEMSELK.S + Oxidação (M) R.LGPMIGCDLSSDD.- + Oxidação (M) K.IMDNQSEQLEGK.E + Oxidação(M) K.EGKEEEEEGHMQK.C + Oxidação (M)
144	unnamed protein product [Glycine max]	gi 296510449	314	6,90 / 18893	26	K.VEVEDGR.V R.VLQISGEK.T R.FRLPENAK.M K.EDVKVEVEDGR.V K.AAMENGVLTVTPK.E + Oxidação (M)
145	unnamed protein product [Glycine max]	gi 218319206	475	6,69 / 34675		K.MLLEAK.T + Oxidação (M) K.FPAQSQK.T K.GAIVAFTR.G K.QVIDLVVK.E R.GLSQQLASR.G R.AVCLCFAK.E K.EGATVAFTYVK.G K.VALVTGGDSGIGR.A K.EHVMNPLPQATNPDKH.A + Oxidação (M)
146	RecName: Full=Beta-conglycinin, alpha' chain; Flags: Precursor	gi 121286	40	4,88 / 27934	1	R.NFLAGSK.D
147	unknown [Glycine max]	gi 255630413	40	5,60 / 25654	8	K.AAMENGVLTVTPK.E + Oxidação (M)
148	unknown [Glycine max]	gi 255633070	602	4,68 / 18078	60	K.ITFECK.L K.AKSDALFK.A K.ITFLEDGETK.F K.GDAQPNQDDLK.I K.ALVTADNVIPK.A K.KITFLEDGETK.F R.SVENVEGNGGPGTIK.K K.AVEAYLLAHPDYN. R.SVENVEGNGGPGTIKK.I M.GVFTFEDETTSPVAPATLYK.A

149	unnamed protein product [Glycine max]	gi 227247706	166	5,34 / 19279	19	K.TFPENK.H K.FLADGAAK.Y K.HVPGFIER.A R.FALLVEDLK.V
150	glycine-rich RNA-binding protein [Glycine max]	gi 5726567	171	5,31 / 17506	20	R.GGDGGDGGW.R.N R.NITVNEAQSR.G R.AFSQYGEIVETK.I
151	RecName: Full=P34 probable thiol protease; Flags: Precursor	gi 129353	173	4,84 / 36411	4	K.GVITQVK.Y K.KGVITQVK.Y K.EESETLVSAR.V
152	unknown [Glycine max]	gi 255628279	365	7,05 / 30493	34	R.LREVELK.H K.HDPSPPSVK.A R.MGIDYEEER.K + Oxidação (M) K.NGEPTTGTYLR.D R.QCLIYDSPEK.K R.DTATSAIQNFAPINK.I R.EYMTGPVHGIHPLANGGK.G + Oxidação (M)
153	unnamed protein product [Glycine max]	gi 253775905	380	6,79 / 43557	30	K.GAIVAYTR.G R.DTLEMIK.R + Oxidação (M) R.GLALQLVSK.G K.LLDYTSTK.G R.DTLEMIK.R.A + Oxidação (M) K.IALVTGGDSGIGR.A K.KFPPQQQTQPGK.E K.EEETAQFGAQVPMK.R + Oxidação (M) K.EGSSIINTTSVNAYK.G
154	maturation protein pPM32 [Glycine max]	gi 5733686	312	5,23 / 24108	37	K.EYAYDTK.E R.EADEMSESAK.E K.TKEYAYDTK.E K.ETMNEGYDR.A + Oxidação (M) K.QGVQGAWETAK.D K.TEEVAASAGEALK.N R.TNWTYDSAEAR.R K.EKTEEVAASAGEALK.N
155	unnamed protein product [Glycine max]	gi 227247706	205	5,60 / 18893	19	K.TFPENK.H K.FLADGAAK.Y K.HVPGFIER.A R.FALLVEDLK.V
156	unknown [Glycine max]	gi 255627799	29	5,18 / 21537	19	K.LLVSDEAR.R K.FVDTVTPQYKPK.F K.LSTMTADEYFEK.H + Oxidação (M)
157	maturation protein pPM32 [Glycine max]	gi 5733686	199	5,12 / 22812	31	K.EYAYDTK.E K.TKEYAYDTK.E K.ETMNEGYDR.A + Oxidação (M) K.QGVQGAWETAK.D K.TEEVAASAGEALK.N R.TNWTYDSAEAR.R K.EKTEEVAASAGEALK.N

158	trypsin inhibitor [Glycine max]	gi 9367042	57	5,04 / 21786	6	K.WVAFVDNEIQK.A
159	2S albumin pre-propeptide [Glycine max]	gi 2305020	121	4,35 / 15689	19	K.CCTEMSEL.R.S + Oxidação (M) K.ELINLATMCR.F + Oxidação (M) K.IMENQSEEELEK.Q + Oxidação (M)
160	glycinin [Glycine max]	gi 18641	341	4,67 / 28617	14	K.LQSPDDER.K K.LQSPDDER.K.Q K.QIVTVEGGLSVISPK.W K.HFLAQSFNTNEDIAEK.L K.QGQHQQEEEEEGGSVLSGFSK.H R.VFYLAGNPDIIEYPETMQQQQQK.S + Oxidação (M)
161	dehydrin [Glycine max]	gi 2270990	135	5,46 / 24108	21	K.LPGVGGGNNNK.E K.EKLPGVGGGNNNK.E R.DQHGNPVPLTDQYGNPVILTDER.G
162	glycinin subunit G2 [Glycine max]	gi 18637	261	6,00 / 34487	14	R.VTAPAMR.K + Oxidação (M) K.LNALKPDNR.I K.EAFGVNMQIVR.N + Oxidação (M) R.EQAQQNECQIQK.L R.FYLAGNQEQEFLK.Y R.RFYLAGNQEQEFLK.Y R.NLQGENEEEDSGAIVTVK.G
163	51 kDa seed maturation protein [Glycine max]	gi 414977	1056	7,00 / 63520	45	K.LSEGLGFK.H K.GYVGDAQK.T K.DYAGDAAQK.S K.DYASDAAQK.T K.EYVGDAQK.T K.DYVGDAQQR.S K.DYASDAVQK.R K.DYASDTAQR.T K.DYASDATDAK.K K.EYAGDVALNAK.D K.DYAGDAAQNVK.D K.SKDYAGDAAQK.S K.EASDYASETAK.K K.DYASDTAQTSK.D K.TKDYASDAAQK.T K.VSDYATDTAQK.S K.GAAEYASDAAQR.T K.TKDYATDAAQK.T K.TKDYVGDAQQR.S K.DYATDTAQKSK.D K.TKDYASDTAQR.T K.TKDYASDATDAK.K K.SKEYAGDVALNAK.D K.SKDYAGDAAQNVK.D K.TKDYASDTAQTSK.D K.DYASEASDVAQNTK.D K.LQDIASEAGQYSAEK.A K.TKDYASEASDVAQNTK.D

164	Bowman-Birk Proteinase Inhibitor	gi 157833497	80	5,70 / 13359	44	R.CLDTNDFCYKPK.S -.DEYSKPCCDLCMCTR.S + Oxidação (M)
165	MnSOD [Glycine max]	gi 147945633	217	6,84 / 26677	17	K.YASEVYEK.E K.NVRPDYLK.N K.ALEQLQDAIAK.K R.LVVETTANQDPLVTK.G
166	unnamed protein product [Glycine max]	gi 296524578	180	4,40 / 67840	25	K.VFFEER.F K.IKNPNYK.G K.ELPDPEAK.K K.TAFEEAEK.K K.LLSGDVDQK.K K.LDCGGGYMK.L + Oxidação (M) K.GIQTSEDYR.F K.LLSGDVDQKK.F K.APLIDNPDFK.D K.KPEGYDDIPK.E R.FYAISAQYPEFSNK.D K.TEAGEDTKEEGVHDEL.-
167	late embryogenesis abundant protein [Glycine max]	gi 170010	920	6,68 / 71808	41	R.AVGYLSGK.K R.GVIGSMFK.A + Oxidação (M) K.ELEQVNR.E K.TAETAEVTK.N K.RAVGYLSGK.K K.RGVIGSMFK.A + Oxidação (M) K.ENVSGAGETAR.R K.AVQDTYENAK.E K.VGEYADYASQK.A K.AGEYTDYASQK.A R.FDEEGVVNVER.R K.LGELKDTASDAK.R K.MGEYKDYAAEK.A + Oxidação (M) K.EMASETAEATANK.A K.VEESRPGAIAETLK.A K.ENTKENVSGAGETAR.R K.AADQIAGQTFNDVGR.F K.THDIYDSATDNNNNK.T R.DRDFGVVAEQQQHHQEDQQK.R

168	LEA protein [Glycine max]	gi 1389897	529	6,61 / 80306	36	K.SLTSGIEK.L R.QSVVLETR.V K.GLAASAGETAK.E R.DVGKFEMR.T + Oxidação (M) R.ESGGQVVAEK.G R.VMVSGEDER.V + Oxidação (M) K.DYTLQAAEK.A R.GRESGGQVVAEK.G K.AAELASMSAGAVK.G + Oxidação (M) K.SAGGTTASYVGEK.A R.TTAVITCTLEK.G K.AKDYTLQAAEK.A K.VTDHAAANVVGNK.E K.AAAHVVEGAAGYAGHK.A K.MATHFEHLAEQAK.E + Oxidação (M) K.DTPQGSIEALQAGER.V R.ETHELGAHFESLADK.V K.QAASETLNTTTQTAQEK.S K.EGTGSIVFTAIGETVSSAGEK.V K.KPSQPQEAERPSEGIGETVR.Q
169	glutathione S-transferase GST 14 [Glycine max]	gi 11385443	317	6,31 / 28063	28	K.IEEEEIK.G K.QYPLLPR.D R.TALSSTFQG.- K.MLVYFHSR.R + Oxidação (M) R.FWANFAEQK.L K.LLSFFASPFQK.R K.MLVYFHSRR.T + Oxidação (M) K.TSLLLQLNPVHK.K
170	glycinin subunit G2 [Glycine max]	gi 18637	191	4,79 / 30714	7	R.VTAPAMR.K + Oxidação (M) K.EAFGVNMQIVR.N + Oxidação (M) R.NLQGENEEEDSGAIVTVK.G
171	unknown [Glycine max]	gi 255637485	206	6,24 / 18580	18	K.FSYPQQCER.K R.NGFIYYVLPAILR.G K.DTCPITVVQSPNPNSK.G
172	2S albumin pre-propeptide [Glycine max]	gi 2305020	221	4,77 / 16010	19	K.CCTEMSEL.R.S + Oxidação (M) K.ELINLATMCR.F + Oxidação (M) K.IMENQSEEELEEK.Q + Oxidação (M)
173	Chain A, Soybean Agglutinin Complexed With 2,6-Pentasaccharide	gi 6729836	240	5,53 / 35372	32	K.TSLPEWVR.I K.TTSWDLANNK.V -.AETVSFSWNK.F R.TSNILSDVVDLK.T R.ALYSTPIHIWDK.E K.VDENGTPKPSSLGR.A K.QPNMILQGDAIVTSSGK.L + Oxidação (M)

174	glycinin [Glycine max]	gi 18641	191	4,63 / 35758	12	K.LQSPDDER.K K.LQSPDDER.Q K.HFLAQSFNTNEDIAEK.L K.QGQHQQEEEEEGGSVLSGFSK.H R.KQGQHQQEEEEEGGSVLSGFSK.H R.VFYLAGNPDIETMPPQQQQQK.S + Oxidação (M)
175	glycinin [Glycine max]	gi 18642	222	4,78 / 24584	12	K.LQSPDDER.K K.LQSPDDER.Q K.HFLAQSFNTNEDIAEK.L K.QGQHQQEEEEEGGSVLSGFSK.H R.VFYLAGNPDIETMPPQQQQQK.S + Oxidação (M)
176	RecName: Full=Beta-conglycinin, beta chain; Flags: Precursor	gi 121282	42	6,96 / 56823	1	R.SPQLENLR.D
177	RecName: Full=Beta-conglycinin, beta chain; Flags: Precursor	gi 121282	50	6,85 / 55455	1	R.SPQLENLR.D
178	unnamed protein product [Glycine max]	gi 218336143	337	6,60 / 51098	31	R.GV <u>M</u> LSDGK.T + Oxidação (M) K.GQTPLFPR.I K.INPAAPLDK.V R.TDIPGVVEK.Y K.ILFNLCR.T R.IIGVDLLTNR.F K.QFGVTDFVNPK.D K.GTFYGHYRPR.T K.SSESNLCELLR.I R.ILGHEASGIVESVGK.G
179	malate dehydrogenase [Glycine max]	gi 5929964	454	6,39 / 45950	54	K.TLCTAIK.Y K.ANLDDDDVIK.A R.LFGVTTLDVVR.A R.TQDGGTEVVEAK.A R.DDLFNINAGIVK.T R.SEVVGYYGGDEELGK.A K.ALEGADVVIIPAGVPR.K K.VAVLGAAGGIGQPLSLLMK.L + Oxidação (M) R.KVAVLGAAGGIGQPLSLLMK.L + Oxidação (M) K.YCPHALVNMISNPVNSTVPIAAEVFK.K + Oxidação (M) K.LNPLVSSLSLYDIAGTPGVAADISHINTR.S K.ANPVAGVNVPPVGGHAGITILPLFSQATPK.A
180	glycinin chain A2B1a precursor - soybean	FWSYG1	165	7,45 / 19314	7	K.NNNPFSFLVPPQESQR.R K.NAMFVPHYTLNANSIIYALNGR.A
181	Basic 7S globulin - Glycine max	BAA03681	125	7,88 / 15343	7	R.QLEENLVVFDLAR.S K.INAYPSVDLVMDKPNPQVWR.I
182	Glycinin A1bB2-784.- Glycine max (Soybean).	Q852U4_SOYBN	151	8,49 / 19039	7	K.NNNPFSFLVPPQESQR.R K.NAMFVPHYTLNANSIIYALNGR.A
183	glycinin chain A5A4B3 precursor - soybean	FWSYG5	272	9,17 / 19693	11	R.ISTLNSLTLPALR.Q R.AIPSEVLAHSYNLR.Q K.YEGNWGPLVNPESQGGSPR.V K.NGIYSPHWNLNANSVIYVTR.G

184	glycinin chain A5A4B3 precursor - soybean	FWSYG5	285	9,38 / 20100	11	R.ISTLNSLTLPALR.Q R.AIPSEVLAHSYNLR.Q K.YEGNWGPLVNPESQQGSPR.V K.NGIYSPHWNLNANSVIYVTR.G
185	glycinin chain A5A4B3 precursor - soybean	FWSYG6	121	9,73 / 21122	8	R.ISTLNSLTLPALR.Q R.AIPSEVLAHSYNLR.Q K.YEGNWGPLVNPESQQGSPR.V
186	Glycinin A3B4 subunit.- Glycine max (Soybean)	Q7GC77_SOYBN	318	9,88 / 19314	14	K.LHENIARPSR.A R.ISTLNSLTLPALR.Q R.QFGLSAQYVVLRYR.N R.AIPSEVLSNSYNLGQSQVR.Q R.NGIYSPHWNLNANSVIYVTR.G
187	basic 7S globulin precursor - soybean	S06750	74	9,84 / 29411	9	K.GAIIFGDAPNNMR.Q K.GAIIFGDAPNNMR.Q + Oxidação (M) R.NTQGVAGLGHAPISLPNQLASHFGLQR.Q
188	Glyceraldehyde-3-phosphate dehydrogenase.- Glycine max (Soybean).	Q2I0H4_SOYBN	92	7,24 / 36815	8	K.TLLFGDKPVTIFGHR.N K.LVSWYDNEWGYSSR.V
189	glycinin chain A2B1a precursor - soybean	FWSYG1	113	7,44 / 78920	7	K.LSAQYGSLR.K K.SQSDNFEYVSFK.T K.NNNPFSFLVPPQESQR.R
190	Sucrose-binding protein precursor related cluster	TA1389_3847	41	5,35 / 57373	2	R.AHTFVSPR.H K.FEEFFGPGGR.D
191	Glycinin precursor [Contains: Glycinin A3 subunit; Glycinin B4 subunit]	TA389_3847	32	4,55 / 37255	2	K.HFLAQSFNTNEDTAEK.L
192	glycinin G1 precursor - soybean	S10851	126	7,56 / 38864	3	R.LSAEFGSLR.K R.VFDGELQEGR.V