

---

**Universidade Estadual de Campinas – UNICAMP**

**Instituto de Química**



**Amadeu Hoshi Iglesias**

**Utilização de Espectrometria de Massas,  
Ligação Cruzada (Cross-Linking) e Footprinting  
no Estudo de Interações Proteína-Proteína**

Tese apresentada no Instituto de Química da  
Universidade Estadual de Campinas para a obtenção do  
título de Doutor em Química (Área de concentração:  
Química Orgânica).

**Orientador: Prof. Dr. Fábio Cesar Gozzo**

**Campinas, SP  
Abril/2009**

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO  
INSTITUTO DE QUÍMICA DA UNICAMP

Ig4u Iglesias, Amadeu Hoshi.  
Utilização de espectrometria de massas, ligação cruzada (*cross-linking*) e *footprinting* no estudo de interações proteína-proteína / Amadeu Hoshi Iglesias.  
-- Campinas, SP: [s.n], 2009.

Orientador: Fábio César Gozzo.

Tese - Universidade Estadual de Campinas,  
Instituto de Química.

1. Espectrometria de massas. 2. *Cross-Linking*.  
3. *Footprinting*. 4. Interações proteína-proteína.  
I. Gozzo, Fábio Cesar. II. Universidade Estadual de  
Campinas. Instituto de Química. III. Título.

**Título em inglês:** Use of mass spectrometry, cross-linking and footprinting in the study of protein-protein interactions

**Palavras-chaves em inglês:** Mass spectrometry, Cross-linking, Footprinting, Protein-protein interactions

**Área de concentração:** Química Orgânica

**Titulação:** Doutor em Ciências

**Banca examinadora:** Prof. Dr. Fábio Cesar Gozzo (orientador), Prof. Dr. Carlos Henrique Inácio Ramos (IQ-UNICAMP), Prof. Dr. Munir Salomão Skaf (IQ-UNICAMP), Prof. Dr. Luiz Alberto Beraldo de Moraes (FFCLRP-USP), Prof. Dr. Emer Suavinho Ferro (ICB-USP)

**Data de defesa:** 03/04/2009

Dedico esta dissertação  
às pessoas mais importantes da minha vida:  
meu pais Antonio Carlos e Norica,  
minha irmã Patricia  
e minha namorada Daniela,  
pelo amor, carinho, paciência  
e apoio incondicional.

## Agradecimentos

Ao meu orientador e amigo, Prof. Fábio Cesar Gozzo, pela orientação e paciência.

Ao meu pai Antonio Carlos, minha mãe Norica e minha irmã Pat por representarem tudo que um ser humano necessita para ser feliz: amor, amizade, carinho, confiança, generosidade e paciência.

À minha querida namorada Dani, pela paciência nos momentos difíceis, por representar a esperança quando tudo parecia dar errado, pelo carinho incondicional; enfim, por me ensinar a definição da palavra *amor*.

Aos meus parentes, em especial às minhas avós Ba-tian e Silvina, padrinhos China e Márcia, tio Hikaru e a nova geração de primos Diogo e Douglas, por torcerem, rezarem e acreditarem em mim, mesmo estando eu muitas vezes longe e ausente.

Aos amigos do grupo MAS de todas as “gerações”, pela ajuda no desenvolvimento desse trabalho, pelo ótimo ambiente e, principalmente, por partilharem muitos e bons momentos de alegria. Não posso deixar de destacar algumas pessoas que foram e continuarão sendo muito especiais em minha vida: Lu, Márrrcia, Jane, Pablito, Luizinho e Pilau.

Aos amigos do LNLS, em especial Alex e Ju, pelos momentos de descontração e pelos ensinamentos de biologia.

A todos os técnicos e funcionários do LNLS pelo carinho dedicado em me ajudar em todas as ocasiões, em especial a Tereza, Eugênia, Gi, Maurício, Evandro, Liu, James, Margareth, Adriana e Andréia.

Aos meus amigos do Rio, Diogo, Rapha, Bi e Guibosa, pelas férias inesquecíveis e pela amizade eterna.

Aos meus amigos de Campinas, Rubão, Andreas, Raphield, Daniel Bozo, Heitor, Danilo, Odair, Leila, Garsa, Marcão, Bujas, Carla Pilau, Carla Veríssimo, Almir, dentre outros, pelos ótimos momentos compartilhados e porque graças a eles sempre me senti em casa mesmo estando longe de casa.

Ao Prof. Marcos Eberlin, pelo bom humor e disponibilidade em ceder seu tempo e laboratório.

Ao Prof. Carlos Roque, por ceder tempo e espaço de laboratório para as tentativas de síntese.

Ao Dr. Nilson Zanchin e a Celisa Caldana, por terem me mostrados os meios para trabalhar com as Tifs.

Ao Laboratório Nacional de Luz Síncrotron e Instituto de Química da UNICAMP, pela infra-estrutura disponibilizada para realização desse projeto.

À FAPESP e à CAPES, pelo auxílio financeiro.

---

# Curriculum Vitae

## **Dados Pessoais:**

Nome: Amadeu Hoshi Iglesias

Nascimento: 01/12/1982

## **Formação acadêmica:**

- **Bacharelado em Química**

Local: Universidade Estadual de Campinas - UNICAMP - Campinas - SP.

Período: março/2001 - dezembro/2004.

## **Participações em Congressos Científicos:**

- Iglesias A. H.; Ferreira O. P.; Alves O. L. Síntese e comportamento térmico de hidróxidos duplos lamelares contendo  $\text{Co}^{2+}$ ,  $\text{Cu}^{2+}$  e  $\text{Fe}^{3+}$ . XXVI Congresso Latinoamericano de Química/27ª Reunião Anual da Sociedade Brasileira de Química, 2004, Salvador.
- Iglesias A. H.; Eberlin M. N.; Gozzo F. C. New Mass Spectrometric Methods for Protein Structure Elucidation: Chemical Cross-Linking and X-Ray Footprinting. 1º Congresso da Sociedade Brasileira de Espectrometria de Massas, 2005, Campinas.
- Iglesias A. H.; Gozzo F. C. Desenvolvimento de Novas Estratégias Utilizando Espectrometria de Massas para a Determinação de Estrutura de Proteínas. 16ª Reunião Anual de Usuários do LNLS, 2006, Campinas.
- Iglesias A. H.; Gozzo F. C. Chemical cross-linking and X-Ray footprinting coupled to mass spectrometry as a tool for protein structure elucidation. XXXV Reunião Anual da SBBq, 2006, Águas de Lindóia.
- Gozzo F. C.; Iglesias A. H. Chemical cross-linking and X-Ray footprinting coupled to mass spectrometry as a tool for protein structure elucidation. 54<sup>th</sup> ASMS Conference on Mass Spectrometry, 2006, Seattle, Estados Unidos.
- Iglesias A. H.; Gozzo F. C. Utilização de Técnicas de Espectrometria de Massas para o Estudo de Interação das Proteínas Tif34p/HisTif35N. 17ª Reunião Anual de Usuários do LNLS, 2007, Campinas.
- Gozzo F. C.; Iglesias A. H. Footprinting coupled to UPLC/MS/MS to map the interaction surface between Tif34 and HisTif35 proteins. 55<sup>th</sup> ASMS Conference on Mass Spectrometry, 2007, Indianapolis, Estados Unidos.

- Iglesias A. H.; Gozzo F. C. Cross-Linking and UPLC/MS/MS analysis of the Tif34/Tif35 Protein Complex. 55<sup>th</sup> ASMS Conference on Mass Spectrometry, 2007, Indianapolis, Estados Unidos.
- Santos L. F. A.; Iglesias A. H.; Gozzo F. C. Analysis of Cross-Linked Peptides by ESI- and MALDI-MS/MS. 2<sup>o</sup> Congresso da Sociedade Brasileira de Espectrometria de Massas, 2007, Campinas.
- Iglesias A. H.; Gozzo F. C. Mapping Tif34/Tif35 Interaction Domain Using Mass Spectrometry. 2<sup>o</sup> Congresso da Sociedade Brasileira de Espectrometria de Massas, 2007, Campinas.
- Iglesias A. H.; Santos, L. F. A.; Gozzo F. C. Fragmentation Pattern of Intermolecular Cross-Linked Tryptic Peptides. 56<sup>th</sup> ASMS Conference on Mass Spectrometry, 2008, Denver, Estados Unidos.
- Santos, L. F. A.; Iglesias A. H.; Gozzo F. C. Study of Fragmentation Pattern of Intramolecular Cross-Linked Peptides by ESI and MALDI-MS/MS. 56<sup>th</sup> ASMS Conference on Mass Spectrometry, 2008, Denver, Estados Unidos.
- Santos L. F. A.; Iglesias A. H.; Gozzo F. C. Analysis of Cross-Linked Peptides by ESI- and MALDI-MS/MS. 18<sup>a</sup> Reunião Anual de Usuários do LNLS, 2008, Campinas.
- Iglesias A. H.; Gozzo F. C. Mapping Tif34/Tif35 Interaction Domain Using Mass Spectrometry. 18<sup>a</sup> Reunião Anual de Usuários do LNLS, 2008, Campinas.

### **Produção Bibliográfica:**

- Iglesias A. H.; Ferreira O. P.; Gouveia D. X.; Souza Filho A. G.; Paiva J. A. C.; Mendes Filho J.; Alves O. L. Structural and thermal properties of Co-Cu-Fe hydrotalcite-like compounds. *Journal of Solid State Chemistry*, 178, 142-152, 2005.
- Iglesias A. H.; Santos L. F. A.; Gozzo F. C. Collision Induced Dissociation of Lys-Lys Intramolecular Cross-Linked Peptides. *Journal of the American Society for Mass Spectrometry*, 20, 557-566, 2009.
- Santos L. F. A.; Iglesias A. H.; Gozzo F. C. Fragmentation of Intermolecular Lys-Lys Cross-Linked Peptides generated by MALDI- and ESI-MS/MS for Structural Proteomics. Submetido para publicação no *Journal of Proteome Research*.

## Resumo

### UTILIZAÇÃO DE ESPECTROMETRIA DE MASSAS, LIGAÇÃO CRUZADA (*CROSS-LINKING*) E *FOOTPRINTING* NO ESTUDO DE INTERAÇÕES PROTEÍNA-PROTEÍNA

A Espectrometria de Massas (MS) é hoje a principal técnica de caracterização de estrutura primária de proteínas devido às vantagens intrínsecas da técnica. As técnicas de *cross-linking* e *footprinting* visam desfrutar de tais vantagens para obtenção de informações estruturais de complexos protéicos. Nesse projeto foram realizados estudos de fragmentação de peptídeos contendo como *cross-linker* o DSS, reagente mais utilizado para esse propósito. Os mecanismos de fragmentação foram propostos baseados na dissociação de peptídeos modelos sintetizados para gerar espécies que mimetizassem experimentos com proteínas. Esses estudos permitiram a identificação de íons marcadores, que foram posteriormente utilizados em experimentos de Varredura de Íons Precursores para auxiliar na detecção de peptídeos modificados. Para realização dos experimentos de *footprinting*, foi desenvolvida uma linha de luz no LNLS. Posteriormente, foi proposto um novo método de quantificação de cinética de oxidação baseado nos dados de LC-MS, levando em consideração todos os produtos de oxidação formados. Esses métodos foram utilizados no estudo de interação das proteínas Tif34 e Tif35 do fator de iniciação de tradução de *Saccharomyces cerevisiae*. Os resultados obtidos indicam que Tif34 apresenta dois possíveis sítios de interação para a proteína Tif35.

## Abstract

### USE OF MASS SPECTROMETRY, CROSS-LINKING AND FOOTPRINTING IN THE STUDY OF PROTEIN-PROTEIN INTERACTIONS

Mass Spectrometry (MS) is the most important tool for analyses related to protein primary structure. Cross-Linking and footprinting aim to bring those advantages to gain insights in the spatial structure of protein complexes. In this work the fragmentation of peptides containing DSS, the most used cross-linker, was studied. Fragmentation mechanisms were proposed based on the dissociation of model peptides which were synthesized in order to generate species that would resemble species found in experiments with proteins. Diagnostic ions were identified which allowed the use of Precursor Ion Scan to detect those species. In order to perform footprinting experiments, a new beamline was developed at the LNLS. A new method for quantitation of oxidation kinetics was developed based on LC-MS analysis, which considers every single oxidation product that can be generated. Those methods were thereafter used in the study of the interaction between Tif34 and Tif35, proteins which compose a translation initiation factor of *Saccharomyces cerevisiae*. The results indicate that Tif34 presents two different sites for the interaction with Tif35.

# ÍNDICE

Lista de Figuras .....	xvii
Abreviaturas, siglas e símbolos .....	xxiii
<b>I INTRODUÇÃO .....</b>	<b>1</b>
I.1 PROTEÍNAS E COMPLEXOS PROTÉICOS.....	1
I.2 MÉTODOS DE CARACTERIZAÇÃO ESTRUTURAL DE PROTEÍNAS E SUAS INTERAÇÕES .....	4
I.3 MÉTODOS BASEADOS EM MS PARA CARACTERIZAÇÃO ESTRUTURAL DE PROTEÍNAS E SUAS INTERAÇÕES.....	9
I.3.1 <i>Cross-Linking</i> .....	15
I.3.2 <i>Footprinting</i> .....	21
I.4 COMPLEXO TIF34/TIF35 DE <i>SACCHAROMYCES CEREVISIAE</i> .....	27
<b>II OBJETIVOS .....</b>	<b>32</b>
II.1 OBJETIVO GERAL .....	32
II.2 OBJETIVOS ESPECÍFICOS .....	32
<b>III PARTE EXPERIMENTAL .....</b>	<b>34</b>
III.1 <i>CROSS-LINKING</i> .....	34
III.1.1 <i>Estudo de Fragmentação de Peptídeos Contendo DSS</i> .....	34
III.1.1.1 Reação de Cross-Linking .....	34
III.1.1.2 Análise dos Produtos de <i>Cross-Linking</i> por MS .....	35
III.1.2 <i>Varredura de Íons Precursores (PIS)</i> .....	36
III.1.2.1 Preparação das Amostras .....	36
III.1.2.2 Análise por LC-MS/MS .....	37
III.2 <i>FOOTPRINTING</i> .....	38
III.2.1 <i>Desenvolvimento da Linha D06B</i> .....	38
III.2.2 <i>Quantificação Label-Free da Cinética de Oxidação</i> .....	39
III.3 ESTUDO DE INTERAÇÃO DO COMPLEXO TIF34/TIF35 .....	40
III.3.1 <i>Expressão e Purificação de HisTif34p, HisTif35N e Tif34p/ HisTif35N</i> .....	40
III.3.1.1 Expressão das Proteínas em <i>Escherichia coli</i> .....	40
III.3.1.2 Cromatografia de Afinidade.....	41
III.3.1.3 Cromatografia de Troca Iônica .....	42
III.3.1.4 Cromatografia de Gel Filtração.....	44

---

III.3.2	<i>Cross-Linking</i> .....	45
III.3.3	<i>Footprinting</i> .....	46
<b>IV</b>	<b>RESULTADOS E DISCUSSÃO</b> .....	<b>47</b>
IV.1	<i>CROSS-LINKING</i> .....	47
IV.1.1	<i>Introdução</i> .....	47
IV.1.2	<i>Desenho dos Peptídeos Modelos para Estudos de Fragmentação</i> .....	55
IV.1.3	<i>Determinação dos Mecanismos de Fragmentação de Peptídeos Intermoleculares Contendo DSS</i> .....	63
IV.1.4	<i>Identificação de Espécies Contendo Cross-Linkers por Varredura de Íon Precursor</i> .....	80
IV.1.5	<i>Conclusões</i> .....	96
IV.2	<i>FOOTPRINTING</i> .....	98
IV.2.1	<i>Introdução</i> .....	98
IV.2.2	<i>Experimentos de Footprinting na Linha de Luz DFX do LNLS</i> .....	104
IV.2.3	<i>Experimentos de Footprinting na Linha de Luz D06B do LNLS</i> .....	108
IV.2.4	<i>Quantificação de Cinética de Oxidação pelo Método Label-Free</i> .....	113
IV.2.5	<i>Conclusões</i> .....	117
IV.3	<i>ESTUDO DE INTERAÇÃO ENTRE AS PROTEÍNAS TIF34 E TIF35 DE SACCHAROMYCES CEREVISIAE</i> .....	119
IV.3.1	<i>Introdução</i> .....	119
IV.3.2	<i>Cross-Linking do Complexo Tif34/Tif35</i> .....	123
IV.3.3	<i>Footprinting do Complexo Tif34/Tif35</i> .....	127
IV.3.4	<i>Conclusões</i> .....	136
<b>V</b>	<b>CONCLUSÕES GERAIS</b> .....	<b>138</b>
<b>VI</b>	<b>BIBLIOGRAFIA</b> .....	<b>141</b>

## LISTA DE FIGURAS

Figura 1 - (A) Gráfico de correlação entre o número de proteínas e o tamanho estimado do interactoma (referência 25). (B), (C) e (D) Representação esquemática dos interactomas de <i>C. elegans</i> , <i>D. melanogaster</i> e <i>H. sapiens</i> respectivamente. ....	4
Figura 2 – Classificação dos diversos métodos de caracterização de complexos protéicos (referência 28).....	6
Figura 3 – Estatísticas do NYSGRG: das aproximadamente 8000 proteínas selecionadas para o estudo, cerca de apenas 5% apresentam estruturas depositadas no PDB.....	8
Figura 4 – (A) Determinação do modo de enovelamento das unidades $\alpha$ e $\beta$ da urease de <i>H. pilory</i> em função do tempo (a – 15 minutos, b – 120 minutos) (referência 68); (B) Espectro de MS do capsídeo do vírus da hepatite B; os espectros em amarelo referem-se ao arranjo quaternário da proteína do capsídeo na forma truncada, demonstrando a presença de dois arranjos diferentes; o espectro em violeta apresenta o arranjo quaternário da proteína inteira, mostrando a formação de um arranjo espacial preferencial (referência 65); (C) Espectro de $V_1$ -ATPase de levedura, onde é possível observar a presença de outros dois subcomplexos devido à perda de espécies mais lábeis. (referência 70).....	12
Figura 5 – Série de <i>cross-linkers</i> homólogos derivados de NHS. ....	17
Figura 6 – Esquema experimental da utilização de <i>cross-linking</i> associado a MS visando a obtenção de dados estruturais de proteínas. (Adaptado da referência 102).....	20
Figura 7 – Experimento de <i>footprinting</i> de DNA visando a identificação das bases que interagem com a proteína. ....	22
Figura 8 – Esquema experimental da técnica de <i>footprinting</i> acoplada a MS para determinação da superfície de interação proteína-ligante. (Adaptada da referência ). ....	26
Figura 9 – (A) Modelo de iniciação de síntese de proteínas em eucariotos. Os números na Figura indicam os diferentes fatores de iniciação de tradução (eIF). M7G indica 7-metilguanossina. (referência 157) (B) Modelo de interação entre as unidades do eIF3 (referência 158). TC indica o complexo ternário eIF2/GTP/tRNA <sub>i</sub> , ctd indica o C-terminal, hld refere-se ao domínio tipo HCR1 e rrm significa sítio de reconhecimento de RNA.....	31
Figura 10 - SDS-PAGE da purificação por afinidade do complexo Tif34p/HisTif35N. E: extrato total; FT: fração não retida na resina; 1-3: tampão sem imidazol; 4-5: 5mM imidazol; 6-8: 20 mM imidazol; 9-11: 50 mM imidazol; 12: 100 mM imidazol; M: Marcador de peso molecular. As frações 6-12 foram combinadas para o próximo passo da purificação. ....	42
Figura 11 - Cromatograma do complexo fracionado utilizando-se a coluna HiTrap Q Sepharose 5 mL (troca iônica).....	43
Figura 12 - Exemplo de gel de SDS-PAGE da purificação por troca iônica do complexo Tif34p/HisTif35N. inj: amostra injetada; C5-C7 e D6-D14: frações coletadas, de acordo com o cromatograma da figura anterior; M: Marcador de peso molecular. As frações D6-D13 foram combinadas para o próximo passo da purificação. ....	43

---

Figura 13 - Cromatograma do complexo fracionado utilizando-se a coluna Superdex 75 HR 10/30 (gel filtração). .....	44
Figura 14 – Exemplo de gel de SDS-PAGE da purificação por gel filtração do complexo Tif34p/HisTif35N. A15 e B1-B6: frações coletadas, de acordo com o cromatograma da figura anterior; M: Marcador de peso molecular. As frações B1-B6 foram combinadas. ....	45
Figura 15 – (A) Estrutura do DSS. (B) Possíveis produtos de reação entre um peptídeo/proteína e um <i>cross-linker</i> bifuncional. (C) Exemplo de reação do DSS com a cadeia lateral de duas lisinas: para que esse produto seja formado, a distância entre esses resíduos deve ser compatível com o tamanho da cadeia espaçadora do DSS. ....	49
Figura 16 – Exemplos de <i>cross-linkers</i> descritos na literatura especificamente desenhados para auxiliar na detecção de peptídeos modificados: reagentes marcados isotopicamente (A) (referência 163), cliváveis com TCEP (B) (referência 171) e DTT (C) (referência 170) e com biotina como grupo de afinidade (D) (a biotina está destacada com o quadrado vermelho) (referência 169). ....	51
Figura 17 - (A) Esquema geral da reação dos peptídeos com <i>cross-linking</i> (DSS nesse caso) para a formação de espécies intra e intermoleculares. (B) Sequências dos peptídeos utilizadas nesse estudo. ....	56
Figura 18 e Esquema 1 - Espectros de MALDI-MS do peptídeo P1 intacto (A), após a reação com DSS (B) e digestão com tripsina por 1 hora (C) e 3 horas (D). Os íons assinalados com uma seta correspondem às espécies de interesse. ....	58
Figura 19 e Esquema 2 - Espectros de MALDI-MS do peptídeo P2 intacto (A), após a reação com DSS (B) e digestão com tripsina por 1 hora (C) e 3 horas (D). Os íons assinalados com uma seta correspondem às espécies de interesse. ....	59
Figura 20 e Esquema 3 - Espectros de MALDI-MS do peptídeo P3 intacto (A), após a reação com DSS (B) e digestão com tripsina por 1 hora (C) e 3 horas (D). Os íons assinalados com uma seta correspondem às espécies de interesse. ....	60
Figura 21 e Esquema 4 - Espectros de MALDI-MS do peptídeo PX intacto (A), após a reação com DSS (B) e digestão com tripsina por 1 hora (C) e 3 horas (D). Os íons assinalados com uma seta correspondem às espécies de interesse. ....	61
Figura 22 – Exemplo de possível coincidência de massas para o peptídeo P1 entre as espécies intermolecular e <i>dead-end</i> . ....	62
Figura 23 – Espectros de ESI-MS dos produtos de reação dos peptídeos P1 (A), P3 (B) e PX (C) com DSS seguido da digestão com tripsina. Os íons indicados com as setas correspondem às espécies intermoleculares. No caso de PX, as setas contínuas indicam os íons referentes à PX1 e as setas pontilhadas apontam aos íons referentes à PX2. ....	65
Figura 24 - Espectros de ESI-MS/MS dos peptídeos intermoleculares de P1 (A), P3 (B), PX1 2+ (C), PX1 3+ (D), PX2 2+ (E) e PX2 3+ (F). ....	69
Figura 25 – Nomenclatura utilizada para a estrutura de fragmentos observados na dissociação de peptídeos em fase gasosa. ....	70

---

Figura 26 - Espectros de MALDI-MS/MS dos peptídeos intermoleculares de P1 (A), P3 (B), PX1 (C) e PX2 (D).....	78
Figura 27 – Representação esquemática dos diferentes tipos possíveis de experimentos de MS/MS.....	82
Figura 28 - Cromatograma da amostra PX com baixa energia de colisão, indicando a presença dos íons de interesse.....	84
Figura 29 - Cromatograma de íons selecionados da amostra PX com alta energia de colisão, indicando a presença dos íons de $m/z$ 222.1494 (linha roxa), 239.1759 (linha verde) e 305.2229 (linha vermelha) (Erro de 0.02 Da).....	85
Figura 30 - Cromatograma de íons selecionados com alta energia de colisão (linhas em vermelho e verde correspondem aos íons de $m/z$ 222.1 e 239.1.....	87
Figura 31 - Cromatogramas com baixa energia de colisão para a hemoglobina digerida com (linha verde) e sem PX (linha vermelha). ....	88
Figura 32 - Cromatogramas de íons selecionados com alta energia de colisão para as amostras contendo hemoglobina + PX (linha verde) e somente hemoglobina (linha vermelha), indicando a presença dos íons marcadores de $m/z$ 222.1 (A), 239.1 (B) e 305.2 (C). ....	89
Figura 33 - Cromatograma com alta energia de colisão para o controle (vermelho) e experimento (verde), indicando a presença dos íons marcadores de $m/z$ 222.1 (A), 239.1 (B) e 305.2 (C).....	91
Figura 34 – (A) Cromatogramas de alta energia de colisão de íons selecionados para as $m/z$ 222.1494 e 239.1759 ( $\pm$ 0.02 Da), em vermelho e verde respectivamente. O retângulo destaca o pico cromatográfico referente ao espectro de MS (baixa energia de colisão) apresentado em (B). (C) Espectro anotado de MS/MS do íon de $m/z$ 381.7, peptídeo N-terminal da lisozima modificado com DSS ( <i>dead-end</i> ). Em (D), ampliação do espectro (C), destacando a presença dos íons marcadores.....	92
Figura 35 – Fluxograma de interpretação de dados obtidos de um experimento de PIS para detecção de peptídeos contendo <i>cross-linkers</i> derivados de diésteres de NHS (CE indica a massa da cadeia espaçadora do <i>cross-linker</i> utilizado no experimento). ....	94
Figura 36 – Mecanismo e principais produtos de oxidação de resíduos alifáticos (A), aromáticos (B), básicos (C), ácidos (D), contendo enxofre (E) e álcoois (F) em experimentos de <i>footprinting</i> .....	100
Figura 37 – (A) Área superficial exposta ao solvente para os 20 aminoácidos em proteínas depositadas no PDB. (B) Cinética de reação dos aminoácidos frente ao radical hidroxila, frequência de cada aminoácido em proteínas, fração na qual esses aminoácidos apresentam área superficial exposta (SEA) maior que 30 Å <sup>2</sup> e o produto entre a frequência dos aminoácidos em proteínas e a SEA. ....	102
Figura 38 – Cálculo da cinética de oxidação de peptídeos em experimentos de <i>footprinting</i> : para cada intervalo de tempo de exposição, a fração de peptídeo não modificado é calculada pela divisão entre peptídeo não modificado e peptídeo total (modificado + não modificado). ....	103
Figura 39 – (A) Gráfico da formação de produtos de oxidação da mioglobina em função do tempo de exposição e da presença de peróxido. (B), (C) e (D) apresentam os espectros de MS deconvolvidos para a mioglobina para 1 s sem peróxido, 5 s sem peróxido e 1 s com peróxido, respectivamente. ....	106

---

Figura 40 - (A) Espectros de MS da amostra com diferentes intervalos de tempo de exposição. Os íons de $m/z$ 908.2, 916.2, 924.2 e 932.2 indicam respectivamente o precursor e as espécies mono, di e trihidroxilada. (B) Cinética de formação dos produtos apresentados no espectro anterior. ....	107
Figura 41 – Estação experimental da linha de luz XRL, onde foram realizados os experimentos de <i>footprinting</i> . Os índices I e II indicam respectivamente o porta-amostras (capacidade para 6 amostras) e o <i>shutter</i> rápido. ....	109
Figura 42 – Estruturas do Alexa Flúor 488 (A) e Diclorofluoresceína (B). ....	110
Figura 43 – Intensidade da fluorescência da amostra irradiada nas diferentes posições do porta-amostras ( $n = 3$ ). ....	111
Figura 44 – (A) Intensidade da fluorescência de DCF em função do tempo de exposição. (B) Linearização da curva (A) de acordo com as equações anteriores. ....	113
Figura 45 – Cromatogramas correspondentes a digestos de mioglobina após 5 intervalos de tempo de exposição diferentes e corrente iônica extraída para 3 peptídeos diferentes da mioglobina. ....	116
Figura 46 – Espectro de MS deconvolvido para as amostras controle (linha vermelha) e irradiada por 800 ms (linha verde). A ampliação mostra a formação dos produtos de oxidação com o aumento do tempo de exposição. ....	117
Figura 47 – (A) Seqüência característica de proteínas da família <i>WD-repeat</i> . (B) Regiões da seqüência que definem as características estruturais das proteínas dessa família; os resíduos em destaque referem-se às regiões conservadas nessa estrutura. (C) Representação esquemática da estrutura tipo $\beta$ - <i>propeller</i> , sendo as diferentes cores as regiões por onde as proteínas dessa família podem interagir. ....	120
Figura 48 – Sobreposição dos modelos obtidos por SAXS e homologia para a proteína Tif34 (A) e o complexo Tif34/Tif35 (B). ....	122
Figura 49 – SDS-PAGE da reação de <i>cross-linking</i> entre Tif34, Tif35 e DSS. ....	123
Figura 50 – Representação esquemática das ligações obtidas pela análise de MS do complexo Tif34/Tif35 ligado por DSS. Os retângulos vermelhos representam as posições dos resíduos de lisinas ao longo da seqüência das proteínas. ....	125
Figura 51 – Modelo de homologia da proteína Tif34p, onde as lisinas marcadas em vermelho, azul e amarelo representam respectivamente resíduos com maior tendência a formação de peptídeos com <i>cross-linking</i> intra ou intermolecular; resíduos onde ocorre formação preferencial de <i>dead-end</i> ; e um resíduo que formou tanto <i>cross-linking</i> intermolecular quanto <i>dead-end</i> . ....	126
Figura 52 – Sobreposição de cromatogramas da proteína Tif34 exposta à radiação síncrotron por diferentes intervalos de tempo (400-800 ms). ....	128
Figura 53 – SDS-PAGE da proteína Tif34 irradiada por diferentes intervalos de tempo. ....	129
Figura 54 – Perfil cromatográfico do íon de $m/z$ 512.5 (3+) em função do tempo de exposição para a proteína Tif34 isolada (A) e complexada com Tif35 (B). (C) Gráfico da intensidade do íon em função do tempo de exposição. (D) Linearização do gráfico anterior, de onde é possível obter as diferentes taxas de oxidação. ....	131

---

Figura 55 - Perfil cromatográfico do íon de $m/z$ 624.0 (3+) em função do tempo de exposição para a proteína Tif34 isolada (A) e complexada com Tif35 (B). (C) Gráfico da intensidade do íon em função do tempo de exposição. (D) Linearização do gráfico anterior, de onde é possível obter as diferentes taxas de oxidação.....	132
Figura 56 – Comparação da cinética de oxidação dos peptídeos de Tif34 na presença e ausência de Tif35. As regiões com maior diferença de taxas de oxidação na presença e ausência do ligante estão marcadas com o quadrado verde. ....	133
Figura 57 – Modelo de homologia de Tif34, onde os sítios protegidos de oxidação pela ligação com Tif35 estão mostrados em amarelo. ....	134
Figura 58 - Regiões de Tif34 oxidadas nos experimentos de <i>footprinting</i> (em vermelho) e modificadas com DSS nos experimentos de <i>cross-linking</i> (em vermelho).....	135
Figura 59 - Modelo de interação entre as proteínas Tif34 (modelo de homologia) e Tif35 (em laranja), baseado nos dados provenientes de experimentos de <i>cross-linking</i> e <i>footprinting</i> . ....	136

## Abreviaturas, siglas e símbolos

ACN	Acetonitrila
CE	Cadeia Espaçadora (do <i>cross-linker</i> )
CHCA	Ácido $\alpha$ -ciano-4-hidróxi-cinâmico
CID	Dissociação Induzida por Colisão
CRP	Cristalografia de Proteínas
Da	Dalton
DCF	Diclorofluoresceína
DDA	Análise Dependente de Dados
DMF	Dimetilformamida
DSS	Dissuccinimidil Suberato
DTT	Ditiotreitol
ECD	Dissociação por Captura de Elétrons
eIF3	Fator de Iniciação de Tradução 3 de Eucariotos
ESI	Ionização por <i>Electrospray</i>
FT-ICR	Ressonância Ciclotrônica de Íons com Transformada de Fourier
HDX	Troca Hidrogênio-Deutério
IAA	Iodoacetamida
IMS	Espectrometria de Mobilidade Iônica
IPTG	Isopropil- $\beta$ -D-1-tiogalactopiranosídeo
IRMPD	Dissociação Multifotônica por Infravermelho
LC (HPLC/UPLC)	Cromatografia Líquida (Alta/Ultra Performance)
MALDI	Ionização por Dessorção a Laser Auxiliada por Matriz
MS	Espectrometria de Massas
MS3D	Espectrometria de Massas Aplicada a Estrutura de Proteínas

NHS .....	N-hidroxisuccinimida
OD <sub>600</sub> .....	Densidade Óptica a 600 nm
PIS .....	Varredura de Íons Precursores
PMSF .....	Fluoreto de Fenilmetilsulfona
PTM .....	Modificações Pós-Traducionais
Q-Tof .....	Quadrupolo-Tempo de Vôo
RMN .....	Ressonância Magnética Nuclear
SAXS .....	Espalhamento de Raios-X a Baixos Ângulos
SDS-PAGE ...	Eletroforese em Gel de Poliacrilamida-Dodecilsulfato de Sódio
TFA .....	Ácido Trifluoracético
THP .....	Tetrahidropiridina

# I INTRODUÇÃO

## I.1 PROTEÍNAS E COMPLEXOS PROTÉICOS

Proteínas constituem um dos tipos de macromoléculas biológicas mais abundantes, versáteis e complexas. Sua complexidade se deve ao fato de que estas podem ser constituídas por diferentes seqüências de aminoácidos que podem ou não estar modificados, além de apresentarem uma ampla faixa de tamanho (variando de alguns poucos kDa a alguns MDa) e estruturas. Em decorrência dessa diversidade, são atribuídas a essas macromoléculas as mais diversas funções, tais como catálise (enzimas em geral), transporte (desde moléculas simples como O<sub>2</sub> até outras proteínas), estrutura, defesa (anticorpos), dentre outras.<sup>1</sup>

Com o fim do seqüenciamento de genomas de inúmeros organismos ao longo das últimas duas décadas, foi possível obter seqüências de um grande número de proteínas, muitas das quais eram desconhecidas até o momento. A informação contida na seqüência desses genes é fundamental na caracterização evolucionária e funcional de genes correlatos assim como na identificação de produtos genéticos envolvidos em doenças. Por outro lado, o conhecimento somente da seqüência primária das proteínas não é suficiente

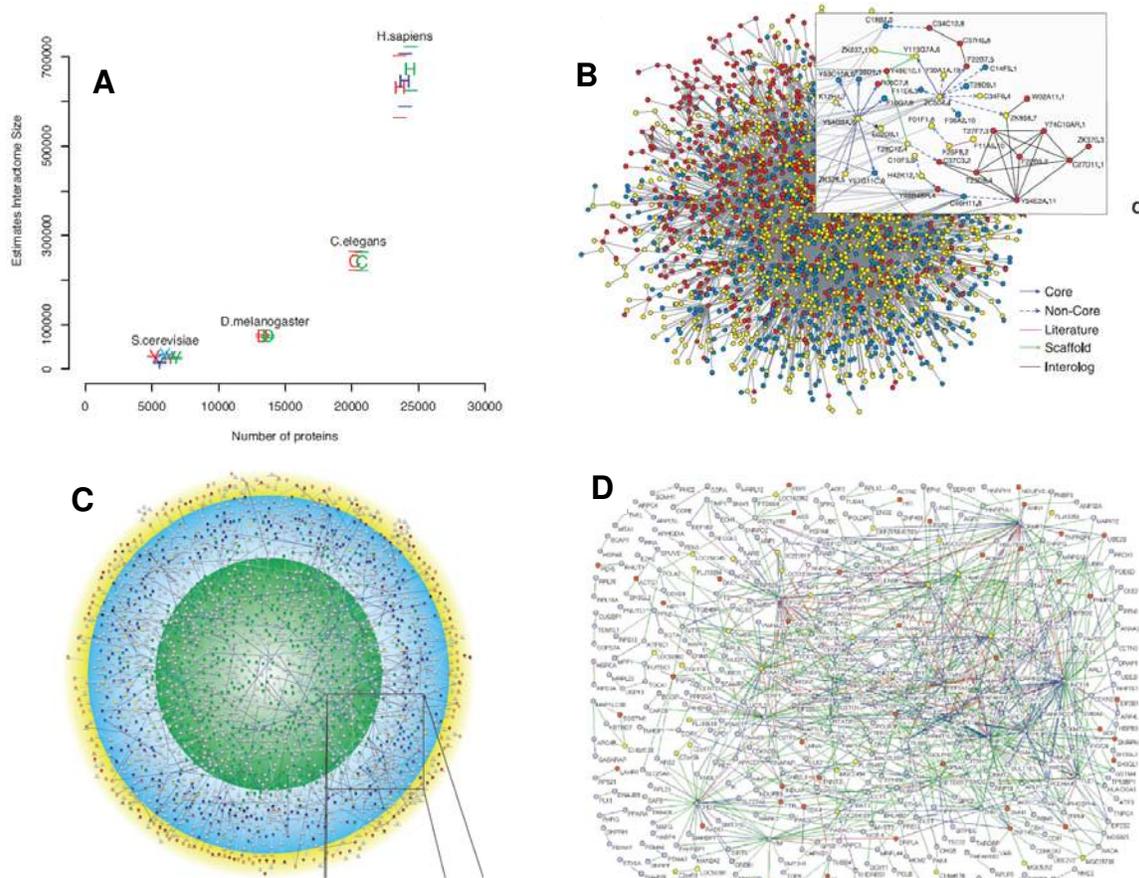
para a caracterização dos mecanismos bioquímicos fundamentais à sobrevivência de todas as células, tanto do ponto de vista das funções que uma proteína pode exercer como do modo que elas a executam.<sup>2,3</sup> Com isso, interpretação dos mecanismos de ação de proteínas, avaliação de padrões de evolução e o desenho de novas drogas são dependentes do conhecimento da estrutura espacial dessas macromoléculas.<sup>4,5</sup>

Uma das principais características das proteínas é que muitas de suas funções celulares tais como replicação de DNA, transcrição, tradução, *splicing*, secreção e controle de ciclo celular não são realizadas individualmente, mas na forma de complexos protéicos, que podem variar desde pequenos oligômeros (dímeros, trímeros) a complexos formados por dezenas de cadeias polipeptídicas e com a presença de ácidos nucléicos e cofatores.<sup>6,7,8,9,10,11,12,13</sup> Isso pode ser resultado tanto de uma determinada mudança conformacional induzida pelo ligante quanto pelo fato de a proteína no estado ligado adquirir uma conformação adequada para realização de sua função.<sup>14</sup> Desse modo, interações entre proteínas, descritas como *sociologia molecular* da célula, são vitais para a manutenção de qualquer processo biológico e, conseqüentemente, da vida de determinado organismo.<sup>9</sup>

A determinação do modo de interação entre proteínas (composição, cinética, estequiometria, termodinâmica) pode trazer um grande número de informações que auxiliariam na compreensão de inúmeros problemas biológicos. Um dos exemplos mais atraentes relacionados à interação entre proteínas pode ser visto na medicina, uma vez que o número de doenças humanas associadas a proteínas é superior ao número de genes no genoma, o que significa que muitas doenças podem estar relacionadas a ausência das condições necessárias para a formação de complexos protéicos; esse fato fez com que interações proteína-proteína tenham se tornado alvos de muitas

drogas nos últimos anos.<sup>13,15</sup> Além disso, a compreensão dessas interações pode elucidar algumas questões relacionadas aos projetos genoma, uma vez que o número de genes de um organismo tão complexo quanto o ser humano é próximo do número de genes de um dos eucariotos mais simples (*Caenorhabditis elegans*) e bem inferior a de alguns vegetais, como soja e arroz, o que indica que em parte a complexidade está relacionada ao modo como esses produtos gênicos interagem entre si para realizar suas funções biológicas.<sup>16</sup> Finalmente, outra grande importância relacionada ao conhecimento do modo de interação entre proteínas está na descoberta da função de outras proteínas. Uma vez que a grande maioria dos processos biológicos é mediada por interações desse tipo, se a função de pelo menos um dos parceiros de interação é conhecida o assinalamento da função da proteína de interesse pode ser facilitado.<sup>17,18,19</sup>

Em virtude dos fatos mencionados acima, os últimos anos presenciaram uma busca incessante pela rede total de interações moleculares em vários organismos, denominada interactoma.<sup>20</sup> Espécies como *S. cerevisiae*<sup>21</sup>, *D. melanogaster*<sup>22</sup>, *C. elegans*<sup>23</sup> e *H. sapiens*<sup>24</sup> tiveram suas redes de interação proteína-proteína (em alguns casos parcialmente) publicadas recentemente. A importância dessas redes é tão grande que estudos recentes demonstram que o tamanho do interactoma – e não o número de genes – está correlacionado com a complexidade dos organismos. Se por um lado, os seres humanos apresentam genoma somente 50% maior que o de *C. elegans*, seu interactoma é cerca de 300% maior (Figura 1).<sup>25</sup>



**Figura 1** - (A) Gráfico de correlação entre o número de proteínas e o tamanho estimado do interactoma (referência 25). (B), (C) e (D) Representação esquemática dos interactomas de *C. elegans*<sup>23</sup>, *D. melanogaster*<sup>22</sup> e *H. sapiens*<sup>24</sup> respectivamente.

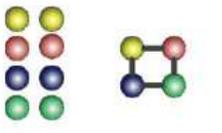
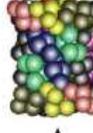
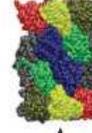
## I.2 MÉTODOS DE CARACTERIZAÇÃO ESTRUTURAL DE PROTEÍNAS E SUAS INTERAÇÕES

Interações proteína-proteína constituem uma das mais complexas, diversas e impactantes interações macromoleculares nas células. Em virtude da importância dessas interações, muito esforço tem sido dedicado ao desenvolvimento de drogas que atuem nesse sentido. Entretanto, são muitas as

dificuldades na determinação do modo de interação entre essas moléculas, dentre as quais está no grande número de possibilidades: um organismo simples como a levedura, capaz de expressar aproximadamente 6000 genes, pode apresentar potencialmente 18 milhões de interações binárias, sendo que somente uma pequena fração dessas ocorre nas células. Além disso, muitas dessas interações são transientes, instáveis ou reguladas por modificações pós-traducionais, o que dificulta sua detecção.<sup>26</sup> Em virtude desses fatores, estudos de interactoma realizados recentemente para a levedura levaram a resultados com muitas diferenças, principalmente no que diz respeito à composição desses complexos.

Do ponto de vista estrutural, a região de contato entre duas proteínas é cerca de 3 a 10 vezes maior que no caso de interações proteína-moléculas pequenas. Além disso, essa região de contato em geral é aproximadamente plana (sem canais ou bolsões) como no caso de interações com moléculas menores. Salvo algumas exceções, muitas das interações entre as proteínas não envolvem seqüências contínuas na cadeia primária, de modo que mesmo peptídeos presentes na superfície de interação não são capazes de romper essa interação.<sup>27</sup>

As técnicas de caracterização de complexos protéicos disponíveis atualmente podem ser divididas nos seguintes grupos<sup>28,29</sup> (Figura 2): técnicas de estrutura primária, alta resolução, média resolução, baixa resolução/sonda e computacionais. Técnicas de caracterização de estrutura primária não fornecem nenhum tipo de informação espacial acerca do complexo de interesse, limitando-se a definir somente quais são as proteínas constituintes. A MS é hoje uma das principais ferramentas para identificação de parceiros de interação, sendo os detalhes referentes à técnica discutidos na próxima seção (Seção I.3).

Composition Stoichiometry	Interaction types	Component interactions	Component configuration	Molecular architecture	Pseudo- atomic structure	Atomic structure	Dynamic processes
<b>Component types</b>		<b>Component instances</b>					
							
Composition Copy numbers	Interaction partners	Interaction instances	Component position	Component position and orientation	Residue positions	Atomic positions	Positions and time
<b>Non-spatial information</b>		<b>Spatial information</b>				<b>Spatial &amp; temporal information</b>	
<ul style="list-style-type: none"> <li>Overlay assays</li> <li>Affinity purification</li> <li>Yeast two-hybrid</li> <li>Genetic interactions</li> <li>Bioinformatics</li> <li>Quantitative immunoblotting</li> <li>PCA</li> <li>Surface plasmon resonance</li> <li>Calorimetry</li> <li>Mass spectrometry</li> </ul>		<ul style="list-style-type: none"> <li>Immuno-EM</li> <li>FRET spectroscopy</li> <li>Electron microscopy</li> <li>Symmetry information</li> <li>Mass spectrometry</li> </ul>	<ul style="list-style-type: none"> <li>Comparative modeling</li> <li>Ab initio prediction</li> <li>Comparative patch analysis</li> <li>Computational docking</li> <li>Electron spin resonance</li> <li>Bioinformatics</li> </ul>	<ul style="list-style-type: none"> <li>Cryo-EM</li> <li>Cryo-ET</li> <li>SAXS</li> <li>H/D exchange</li> <li>Footprinting</li> </ul>	<ul style="list-style-type: none"> <li>X-ray crystallography</li> <li>NMR spectroscopy</li> </ul>	<ul style="list-style-type: none"> <li>NMR spectroscopy</li> <li>FRET spectroscopy</li> <li>Cryo-ET</li> </ul>	

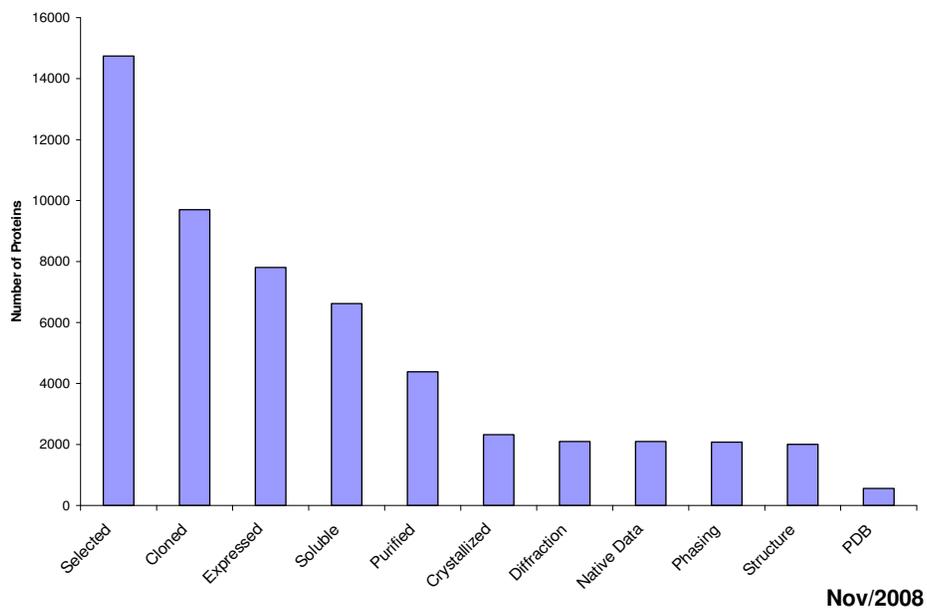
**Figura 2** – Classificação dos diversos métodos de caracterização de complexos protéicos (referência 28).

As técnicas de baixa resolução/sonda englobam as espectroscopias, onde a resposta é determinada pela presença de uma certa característica estrutural da proteína. Essa classe engloba as técnicas de Dicroísmo Circular (CD), Fluorescência no UV, Transferência de Energia Ressonante por Fluorescência (FRET) dentre outras. Embora essas técnicas sejam em geral sensíveis, rápidas, relativamente simples e de fácil acesso, a resposta obtida em geral é insuficiente por si só para caracterização mais profunda do sistema. Como técnicas de média resolução, aquelas que fornecem informações acerca da forma geral do complexo tais como posição e orientação dos componentes, destacam-se as microscopias e o Espalhamento de Raios-X a Baixos Ângulos (SAXS). Em geral tais técnicas são bem restritas a alguns usuários tanto pela necessidade de instrumentação muito específica e dispendiosa quanto pelo alto grau de dificuldade de interpretação de dados.

Atualmente existem duas técnicas estabelecidas para determinação estrutural de proteínas com alta resolução (resolução atômica): Ressonância Magnética Nuclear (RMN) e Cristalografia de Proteínas (CRP). Entretanto, ambas as técnicas apresentam um grande número de limitações que impedem seu uso irrestrito para qualquer tipo de proteína, dentre as quais se destacam a necessidade de uma grande quantidade de amostra (da ordem de vários miligramas) com alto grau de pureza. No caso de RMN, há ainda a necessidade da amostra ser estável em algum dos tampões que não interfiram na análise por um longo intervalo de tempo (dias a semanas) a temperatura ambiente. Além disso, as técnicas disponíveis atualmente restringem o tamanho máximo da amostra a aproximadamente 50 kDa, o que muitas vezes representa a massa de um único componente do complexo. A parte experimental é cara, devido à necessidade de meios de cultura enriquecidos com diferentes isótopos e tanto o experimento quanto a análise dos dados são processos longos e muito pouco automatizados. A CRP por sua vez requer que a amostra esteja na forma de monocristal, o que constitui a principal limitação da técnica. Mesmo que o cristal apresente boas características de difração, a obtenção da estrutura (que assim como no caso de RMN, é feita em grande parte de forma manual) não está garantida, uma vez que em vários casos não se faz possível a resolução do problema das fases. Finalmente, a estrutura obtida por CRP é uma estrutura estática e não necessariamente representativa do sistema em solução.<sup>30</sup>

Um grande exemplo das dificuldades dessas técnicas na obtenção de dados estruturais de proteínas e complexos protéicos pode ser encontrado nos dados do New York Structural Genomics Research Consortium - NYSGRC ([www.nysgrc.org](http://www.nysgrc.org)), um grupo de pesquisa constituído por cinco instituições diferentes, cujo objetivo geral consiste no estudo estrutural (por CRP) e

funcional de proteínas com alta capacidade de análise. A Figura 3 apresenta os resultados obtidos pelo NYSGRG desde a época de sua criação (Janeiro de 1999): é interessante notar que dos quase 15000 alvos selecionados para o estudo, somente cerca de 2000 apresentam estrutura resolvida em 9 anos, uma taxa de sucesso de aproximadamente 13%. Vale ressaltar ainda que o NYSGRG apresenta uma das maiores taxas de sucesso dentre todos os consórcios de genômica estrutural do mundo. Essa figura demonstra as limitações dos métodos atuais em biologia estrutural ao mesmo tempo em que suscita o desenvolvimento de novas metodologias para esse objetivo.



**Figura 3** – Estatísticas do NYSGRG: das aproximadamente 15000 proteínas selecionadas para o estudo, cerca de apenas 5% apresentam estruturas depositadas no PDB.

### I.3 MÉTODOS BASEADOS EM MS PARA CARACTERIZAÇÃO ESTRUTURAL DE PROTEÍNAS E SUAS INTERAÇÕES

Desde o início da década de 60, houve um interesse na utilização de MS para a análise de proteínas, principalmente devido às vantagens intrínsecas da técnica, tais como alta sensibilidade, rapidez de análise e especificidade. Embora tenha havido relativo sucesso, principalmente com a utilização de FAB (*Fast Atom Bombardment*) como fonte de íons a partir da década de 80, essas análises restringiam-se a pequenos peptídeos derivatizados e muitas vezes necessitavam de experimentos de Degradação de Edman para a confirmação dos resultados.<sup>31</sup>

No fim da década de 80, a análise de peptídeos e proteínas por MS deu um grande salto de qualidade, com o desenvolvimento das técnicas de ionização de ESI<sup>32</sup> e MALDI<sup>33</sup>, que permitiam a ionização e volatilização dessas macromoléculas com alta eficiência e sem a necessidade de derivatização.<sup>34</sup> Vale ressaltar nesse ponto que o desenvolvimento dessas técnicas também contribuiu significativamente para a melhoria dos analisadores de  $m/z$ , com o desenvolvimento na década de 90 de equipamentos comerciais dos tipos *ion trap*, Q-Tof e Tof-Tof.<sup>31</sup> Desse modo, esse período presenciou uma grande quantidade de artigos explorando os diferentes tipos de informação que poderiam ser obtidos pela análise de proteínas e peptídeos por MS.

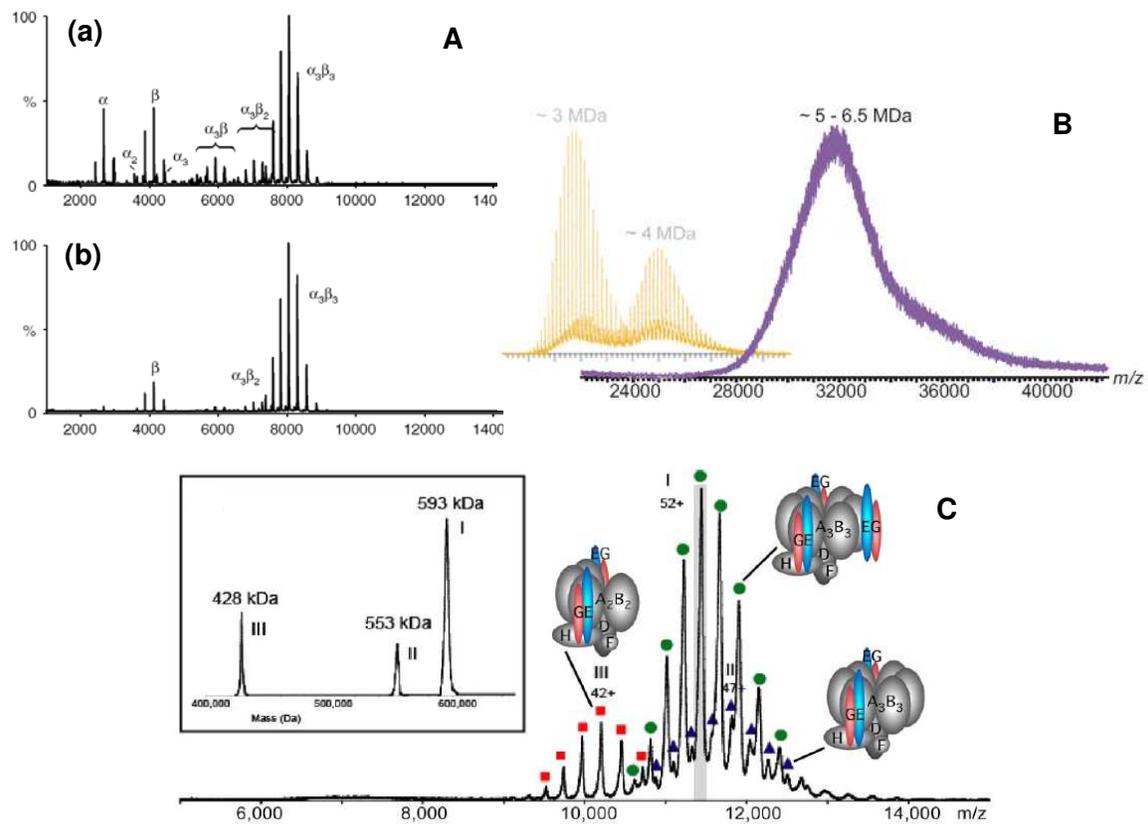
Em 1989, Hunt demonstrou a possibilidade de seqüenciamento de peptídeos gerados por ESI mediante Dissociação Induzida por Colisão (CID).<sup>35</sup> Esse trabalho foi seguido por outros demonstrando o seqüenciamento de peptídeos endógenos<sup>36</sup> e proteínas digeridas por tripsina e analisadas por

LC-MS/MS<sup>37</sup>. Simultaneamente, experimentos de seqüenciamento foram conduzidos em peptídeos gerados por MALDI/*Post-Source Decay* (PSD), onde inicialmente a dissociação dessas espécies não ocorria por meio de fragmentação, mas por decaimento de energia resultante do processo de ionização.<sup>38</sup> Esses achados levaram ao desenvolvimento de algoritmos de busca em banco de dados que permitiam a identificação de proteínas tanto em experimentos de MS<sup>39,40,41</sup> quanto MS/MS<sup>42,43</sup>. A partir desse momento, MS tornou-se a principal técnica de identificação de proteínas, em especial em experimentos de análise proteômica, onde se faz necessário a identificação de um grande número de proteínas em um pequeno intervalo de tempo. Com o desenvolvimento de métodos de dissociação de peptídeos, foi possível obter informações acerca da seqüência dos aminoácidos, sendo que os resultados permitiam a determinação de peptídeos modificados pós-traducionalmente (fosforilação e sulfonação, por exemplo) bem como a análise de espécies bloqueadas no N-terminal, o que era impossível até então por Degradação de Edman.<sup>44</sup> No fim da década de 90, surgiram métodos que permitiam a análise quantitativa diferencial de proteínas por MS.<sup>45,46</sup> Desse modo, as duas últimas décadas presenciaram uma enorme quantidade de trabalhos que elevaram a MS ao posto de principal técnica de análise de problemas relacionados à estrutura primária de proteínas: identificação de proteínas, determinação da massa dessas macromoléculas, seqüenciamento, identificação e localização de modificações pós-traducionais e quantificação.<sup>47,48,49,50,51,52,53</sup> Em virtude desse sucesso, seria muito interessante utilizar as vantagens intrínsecas da técnica de MS no estudo de interações proteína-proteína. Vários métodos foram desenvolvidos com o intuito de identificar parceiros de interação<sup>54,55,56,57</sup>, campo esse onde a MS rapidamente se destacou dentre as técnicas disponíveis, mas como esses métodos não fornecem informações

especiais acerca dos complexos protéicos, eles não serão discutidos nesse ponto.

Alguns anos após o desenvolvimento das técnicas de ionização de ESI e MALDI foi possível observar que essas técnicas eram suficientemente suaves para a manutenção, ao menos em parte, da estrutura de complexos protéicos em fase gasosa.<sup>58,59,60</sup> Inicialmente, os primeiros trabalhos mostravam que em muitos casos, interações proteína-ligante podiam ser mantidas em fase gasosa desde que o processo de ionização fosse realizado em condições controladas.<sup>61</sup> Anos mais tarde, foi possível observar que essas interações também poderiam ser mantidas quando as espécies fossem analisadas por MALDI.<sup>62</sup> Entretanto, embora a ionização por MALDI fosse um processo energeticamente suave, a preparação da amostra requer condições nas quais a grande maioria dos complexos protéicos tende a se dissociar. Apesar dos primeiros resultados animadores, esse método sofria de grandes limitações instrumentais, uma vez que complexos maiores tendem a se dissociar facilmente devido às grandes diferenças de pressão às quais são submetidos na análise por MS; além disso, detectores convencionais apresentavam limitações na faixa de  $m/z$  que impediam a detecção dessas espécies. O desenvolvimento de instrumentação dedicada a essa aplicação contribuiu significativamente para o estudo de complexos maiores e mais lábeis.<sup>63</sup> *Macromolecular MS*, denominação proposta para esses experimentos, tem sido atualmente utilizada para determinação da estequiometria de complexos<sup>64</sup>, determinação de força de ligação das espécies constituintes de complexos<sup>65,66</sup>, análise de mudanças conformacionais promovidas pela interação com ligantes<sup>67</sup>, cinética de enovelamento e desenovelamento<sup>68,69</sup> e determinação da topologia<sup>70,71</sup> (Figura 4). Embora experimentos de *Macromolecular MS* tenham obtido grande êxito em vários casos, algumas limitações impedem seu uso irrestrito. A principal

delas está na necessidade de instrumentos adaptados (não comerciais), cujas características dificultam sua utilização para outros fins. Além disso, mesmo de posse de instrumentação específica não são todos os complexos que suportam essas condições experimentais, podendo estes ser dissociados quando da formação dos íons. Finalmente, os resultados provenientes dessa técnica não apresentam informações acerca da superfície de interação entre os constituintes.



**Figura 4** – (A) Determinação do modo de enovelamento das unidades  $\alpha$  e  $\beta$  da urease de *H. pilory* em função do tempo (a – 15 minutos, b – 120 minutos) (referência 68); (B) Espectro de MS do capsídeo do vírus da hepatite B; os espectros em amarelo referem-se ao arranjo quaternário da proteína do capsídeo na forma truncada, demonstrando a presença de dois arranjos diferentes; o espectro em violeta apresenta o arranjo quaternário da proteína inteira, mostrando a formação de um arranjo espacial preferencial (referência 65); (C) Espectro de V<sub>1</sub>-ATPase de levedura, onde é possível observar a presença de outros dois subcomplexos devido à perda de espécies mais lábeis. (referência 70)

Um método de caracterização estrutural de proteínas que permite a determinação de superfícies de interação entre proteínas é a Troca Hidrogênio/Deutério (HDX).<sup>72,73,74</sup> Em 1991, foi proposta pela primeira vez a utilização de MS para o cálculo da cinética de troca entre os isótopos.<sup>75</sup> HDX baseia-se no fato de que quando proteínas são solubilizadas em excesso de solvente deuterado, os H da proteína trocam com o D do solvente. Os hidrogênios nos grupos funcionais dos aminoácidos trocam muito rapidamente e, portanto, não permitem determinação de sua cinética de troca por MS. Os H das amidas da cadeia principal da proteína, por sua vez, apresentam uma faixa de  $10^8$  ordens de grandeza de diferença de velocidade de troca, dependendo do perfil de ligações de hidrogênio e, principalmente, do grau de exposição desses átomos ao solvente.<sup>73</sup> Após o experimento de troca (realizado por diferentes intervalos de tempo para permitir o cálculo da cinética), o pH do meio reacional é mudado para 2-3 e a temperatura para 0° C de modo a evitar a troca reversa com H. Análises por MS podem ser realizadas tanto de modo global, afim de determinar o nível de incorporação de D, bem como na análise dos peptídeos, com o objetivo de determinar a cinética de troca em cada uma das regiões da proteína e, conseqüentemente, a superfície de interação proteína-ligante.

Assim como no caso anterior, HDX apresenta várias dificuldades técnicas que dificultam sua aplicação em muitos casos. Como descrito anteriormente, para diminuir a taxa de troca reversa, após a reação deve-se abaixar o pH e a temperatura. Essas condições são incompatíveis com a grande maioria das enzimas, sendo que atualmente a única utilizada em experimentos desse tipo é a pepsina. Essa enzima entretanto apresenta baixa especificidade e eficiência no pH utilizado, o que dificulta a identificação dos peptídeos obtidos por MS, além de diminuir a cobertura da proteína quando comparada a peptídeos trípticos. Isso implica em um limite máximo de

tamanho das amostras, mesmo porque boa parte da interpretação dos dados é feita de modo manual. Além disso, as condições cromatográficas devem ser adequadas de modo que a etapa de LC também seja realizada a baixas temperaturas, o que acarreta no aumento significativo da pressão do sistema. Finalmente, a interpretação dos espectros de MS/MS não permite a identificação inequívoca dos sítios de troca, devido ao efeito conhecido como embaralhamento da marcação isotópica (*Hydrogen Scrambling*).<sup>73</sup> Apesar das dificuldades apresentadas anteriormente, HDX tem sido aplicada com relativo sucesso para o estudo de enovelamento e estabilidade de proteínas<sup>76,77</sup>, mudanças conformacionais mediante interação com ligantes<sup>78,79</sup> e no estudo de interações proteína-proteína<sup>80</sup>.

A Espectrometria de Mobilidade Iônica (IMS) é mais uma técnica baseada em MS capaz de fornecer informações espaciais acerca de complexos protéicos. IMS foi desenvolvida na década de 60 por Cohen e Karasek, mas devido a limitações instrumentais ela só voltou a tona no fim da década de 90, sendo que o desenvolvimento de um instrumento comercial e o acoplamento com MS (IMMS) alavancou sua utilização na análise de macromoléculas biológicas. IMS se baseia na separação de íons através de um gás por meio de um campo elétrico, sendo essa separação dependente da massa, carga e seção de choque do analito. A separação por seção de choque é extremamente interessante do ponto de vista de MS, uma vez que permite a diferenciação de isóbaros pela conformação que estes adquirem em fase gasosa. Existem quatro tipos básicos de analisadores de mobilidade iônica, conhecidos como *Drift Time*, *Aspiration*, *Field-Asymmetric Waveform* (FAIMS) e *Travelling Wave* (TWIMS), sendo que a diferença entre esses se baseia no modo como esses instrumentos separam os analitos de interesse.<sup>81,82,83,84,85</sup> Embora o uso de IMMS seja relativamente recente (quando comparado a *Macromolecular MS* e

HDX), grande sucesso já foi alcançado, principalmente com equipamentos do tipo *home-made*, em estudos de estrutura/conformação de peptídeos e proteínas<sup>86,87,88,89</sup>, estudos de enovelamento e desenovelamento de proteínas<sup>87,90</sup>, interações proteína-proteína<sup>91,92,93</sup> e mudanças conformacionais<sup>94,95</sup>.

### I.3.1 *Cross-Linking*

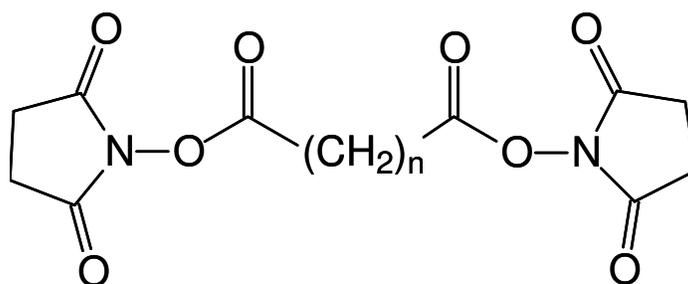
*Cross-linking* consiste na união de duas ou mais moléculas por meio de uma ligação covalente, mediada por um reagente denominado *cross-linker*. Tais componentes podem variar desde proteínas a peptídeos, fármacos e ácidos nucleicos até partículas sólidas. Os primeiros experimentos de *cross-linking* com proteínas datam da década de 20, onde a técnica era aplicada na determinação do sítio ativo de enzimas. Após a década de 50, houve um renascimento da técnica onde surgiram novas aplicações tais como aumento da estabilidade de proteínas, imobilização dessas macromoléculas em suportes sólidos, acoplamento de duas ou mais diferentes proteínas, identificação e caracterização da natureza e extensão de interações proteína-proteína.<sup>96,97,98</sup>

*Cross-linking* entre proteínas podem originar-se de várias formas distintas, sendo que na maioria dos casos estas ocorrem *in vivo*. Talvez o tipo de ligação cruzada mais comum seja a formação de pontes dissulfeto entre resíduos de cisteína. Entretanto, no que concerne a análise da interação entre proteínas, o mais interessante é a introdução de *cross-linking* em estudos *in vitro*, uma vez que escolhendo-se os reagentes adequados, pode-se obter as informações desejadas. Além disso, as interações de complexos não-

covalentes são em geral muito fracas e, desse modo, são facilmente dissociadas, o que torna sua análise muito difícil. Assim, do ponto de vista analítico, é muito interessante a estabilização desses complexos mediante a formação de ligações covalentes, uma vez que isso facilitaria a posterior purificação e caracterização do mesmo em condições mais adequadas, mas que ainda retém as informações espaciais desejadas. As informações obtidas podem então ser utilizadas como base para interrogação de um grande número de informações importantes desses complexos.<sup>99</sup>

*Cross-linkers* são em geral reagentes bifuncionais, ou seja, contém dois grupos reativos frente a cadeias laterais presentes na proteína separados por um conjunto de átomos conhecido como cadeia espaçadora. Eles podem ser classificados como homobifuncionais ou heterobifuncionais, dependendo do fato dos grupos reativos serem iguais ou diferentes. Há ainda uma terceira classe, conhecida como *zero-length*, reagentes que mediam a formação de ligações covalentes sem adição de átomos extrínsecos a proteínas (não contém a cadeia espaçadora). A escolha do reagente ideal é dependente do experimento e da informação desejada, dependendo de inúmeros fatores tais como (in)especificidade dos grupos reativos, solubilidade, tamanho e geometria dos reagentes e presença de grupos que auxiliem na detecção das moléculas ligadas.<sup>97,98</sup> Um dos itens mais importantes nessa escolha consiste na natureza dos grupos reativos, visto que são eles que definem a especificidade das ligações formadas. Dentre os *cross-linkers* disponíveis atualmente, os de maior destaque são aqueles reativos frente a aminas primárias. Isso se deve principalmente a alta ocorrência de grupos aminos (resíduos de lisina e o N-terminal da proteína). Os derivados da N-hidróxi succinimida (NHS) são os *cross-linkers* mais utilizados atualmente, devido aos rendimentos relativamente altos das reações e, principalmente, à facilidade

de sua obtenção, bastando para isso reagir um ácido carboxílico com o NHS.<sup>100</sup> Como são vários os diácidos disponíveis comercialmente, pode-se obter uma série de *cross-linkers* homobifuncionais homólogos; isso permite um estudo mais refinado da estrutura do complexo, uma vez que conhecendo-se o comprimento da cadeia espaçadora, pode-se inferir com mais precisão a real distância entre dois grupos reativos (Figura 5). Outro fator importante na escolha do *cross-linker* é a cadeia espaçadora, visto que além de determinar a distância entre os grupos unidos e, portanto, fornecer as informações estruturais desejadas, ela também afeta algumas das propriedades físico-químicas, como a solubilidade em água e o grau de penetração da molécula em diferentes ambientes.



n	Nome
3	Disuccinimidil glutarato
5	Disuccinimidil pimelato
6	Disuccinimidil suberato
8	Disuccinimidil sebacato

**Figura 5** – Série de *cross-linkers* homólogos derivados de NHS.

Em 2000, Young e colaboradores<sup>101</sup> propuseram pela primeira vez a utilização conjunta de *cross-linking* e MS para a obtenção de informações estruturais de proteínas. Tais informações seriam provenientes do tamanho da

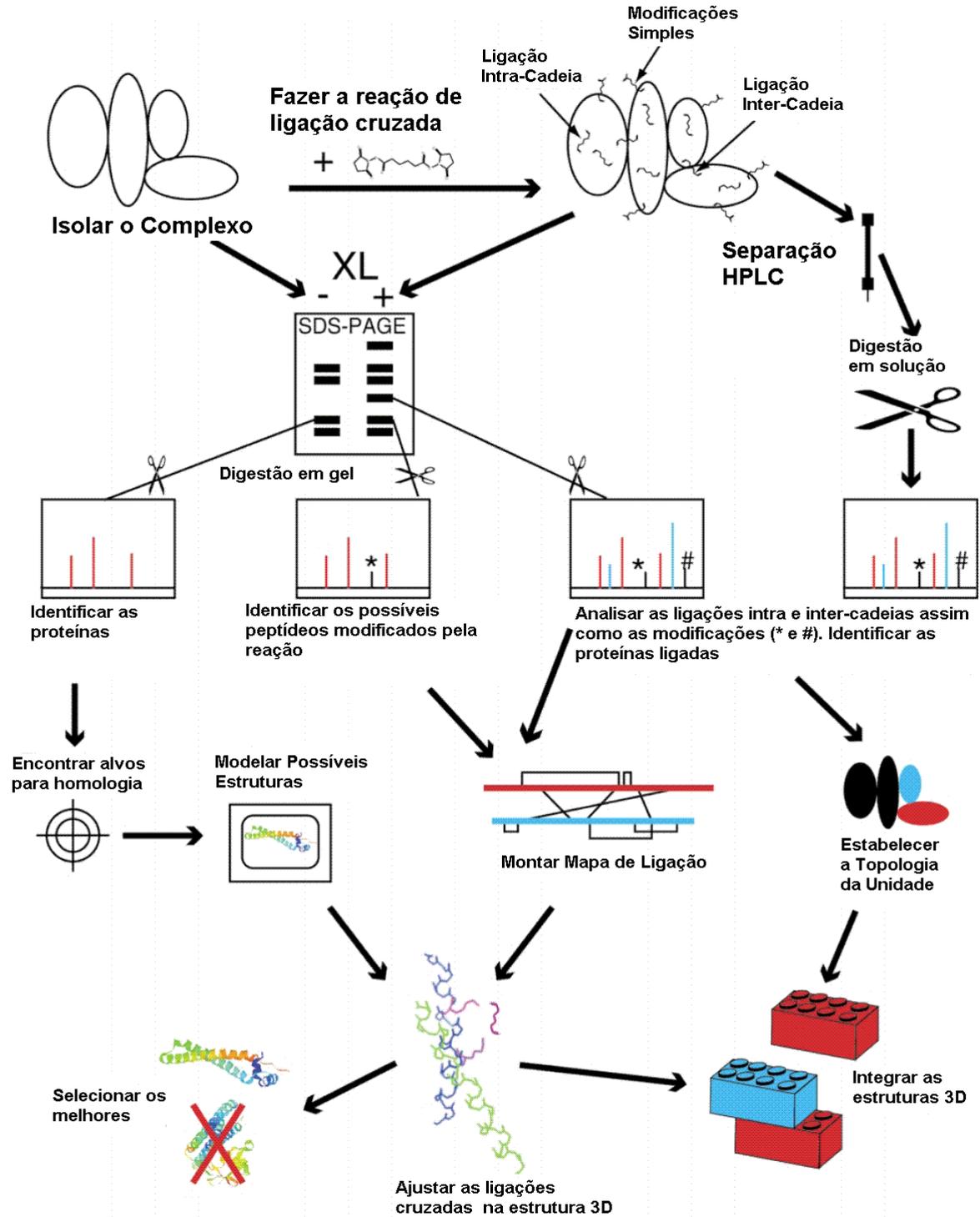
cadeia espaçadora do reagente utilizado, onde a distância máxima entre dois resíduos ligados pelo *cross-linker* corresponderia ao comprimento do reagente. Caso o número dessas restrições espaciais fosse igual a  $N/10$  ( $N$  igual ao número de resíduos da proteína), seria possível determinar o tipo de enovelamento por meio de métodos computacionais. Desse modo, foi proposta pela primeira vez a utilização de um método rápido e sensível baseado em *cross-linking* e MS para a obtenção de informações estruturais de proteínas.

O experimento típico de *cross-linking* e MS visando a obtenção de dados estruturais de proteínas e seus complexos pode ser visualizado na Figura 6.<sup>102</sup> Tipicamente, uma proteína ou complexo protéico isolado é submetido à reação de *cross-linking*, sendo o(s) reagente(s) escolhido(s) em função dos parâmetros mencionados anteriormente. Após a reação, vários diferentes tipos de produtos podem ser formados: i) *dead-end*, onde somente um dos grupos reativos reagiu com a proteína, enquanto o outro sofreu hidrólise; ii) intramolecular, onde os dois resíduos de aminoácidos ligados encontram-se na mesma cadeia polipeptídica; e iii) intermolecular, onde os resíduos ligados encontram-se em cadeias distintas. Os produtos de interesse são separados (eventualmente os peptídeos marcados podem ser enriquecidos) tanto por HPLC quanto por SDS-PAGE e digeridos com uma enzima.<sup>102,103</sup>

Os peptídeos resultantes são então analisados por MS, sendo que a partir dessa etapa as análises realizadas são dependentes do tipo de informação desejada: pode-se simplesmente identificar as proteínas contidas no complexo até a identificação exata de quais os peptídeos estão ligados entre si, o que permite definir a topologia do complexo bem como montar mapas de ligação entre as proteínas. Finalmente, no caso da obtenção de uma quantidade de dados compatível com o tamanho do complexo e com o auxílio de métodos computacionais, podem-se obter modelos tridimensionais das proteínas/

---

complexos de interesse.<sup>102,103</sup> Essa abordagem tem sido extensivamente utilizada nos últimos anos visando a obtenção de dados estruturais como determinação do tipo de enovelamento de proteínas<sup>104,105,106,107</sup>, identificação de parceiros de interação<sup>108,109,110</sup>, monitoramento de mudanças conformacionais mediante interação com ligante<sup>111,112,113</sup>, determinação de superfície de interações em complexos protéicos<sup>114,115,116,117,118,119,120</sup> e finalmente na determinação de acessibilidade ao solvente<sup>121,122,123,124</sup>.

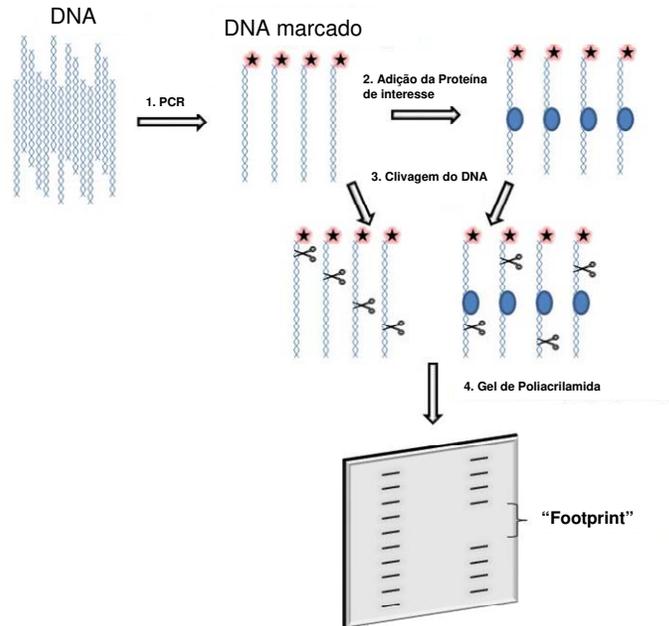


**Figura 6** – Esquema experimental da utilização de *cross-linking* associado a MS visando a obtenção de dados estruturais de proteínas. (Adaptado da referência 102)

### I.3.2 *Footprinting*

*Footprinting* é o termo que refere a um grupo de técnicas baseadas na determinação do grau de exposição ao solvente de partes de uma macromolécula por meio de reações de modificação química ou enzimática.<sup>125,126,127,128,129,130</sup> Modificações químicas são utilizadas para marcar as regiões da proteína expostas ao solvente e com isso determinar quais resíduos ou bases se encontram na superfície; reações enzimáticas por sua vez envolvem a clivagem da macromolécula e a posterior leitura dos resíduos clivados. Essas modificações, de natureza covalente, têm como objetivo avaliar mudanças conformacionais devido à interação da macromolécula de interesse com um determinado ligante.

A técnica de *footprinting* foi introduzida em 1972<sup>130</sup>, na análise dos sítios de interação de uma molécula de DNA com uma proteína. Nesse caso, o complexo proteína-DNA foi digerido com uma DNase e os produtos da digestão analisados por gel de agarose, juntamente com o DNA digerido na ausência da proteína. A raia do gel correspondente a amostra contendo DNA apresenta uma “pegada” (do termo *footprinting* em inglês), referente à porção do DNA que interage com a proteína e conseqüentemente não pode ser digerida pela enzima (Figura 7). *Footprinting* de DNA por meio de modificações químicas foi desenvolvido pouco tempo depois, onde a ligação da proteína com uma região do ácido nucléico impedia a marcação dessa região.



**Figura 7** – Experimento de *footprinting* de DNA visando a identificação das bases que interagem com a proteína.

A utilização de *footprinting* para mapear regiões da proteína que interagem com outras moléculas se deu a partir da década de 80, com a utilização de algumas proteases para digestão de porções expostas da proteína na presença do ligante.<sup>131</sup> Foi somente na década de 90<sup>132</sup> que a utilização de reagentes químicos específicos para alguns resíduos associado à degradação de Edman foi utilizado para determinação de resíduos expostos em proteínas, fato esse que alavancou a utilização de diversos reagentes para a avaliação de vários resíduos.

Embora as abordagens descritas acima tenham contribuído para a compreensão de alguns sistemas biológicos, estas sofriam de algumas limitações que dificultavam sua aplicação em muitos casos. Uma das principais dificuldades estava relacionada à especificidade tanto das proteases quanto dos reagentes químicos, que limitavam significativamente o número de

resíduos que poderiam ser sondados pela técnica. No caso das proteases, havia ainda um agravante relacionado ao seu tamanho, que muitas vezes acabava por superestimar a área de interação da proteína. Para solucionar esses problemas, Heyduk<sup>133</sup> propôs a utilização de radicais hidroxila como reagente de clivagem da cadeia principal, sendo seus principais atrativos seu pequeno tamanho, alta reatividade, inespecificidade e facilidade de geração em soluções aquosas.

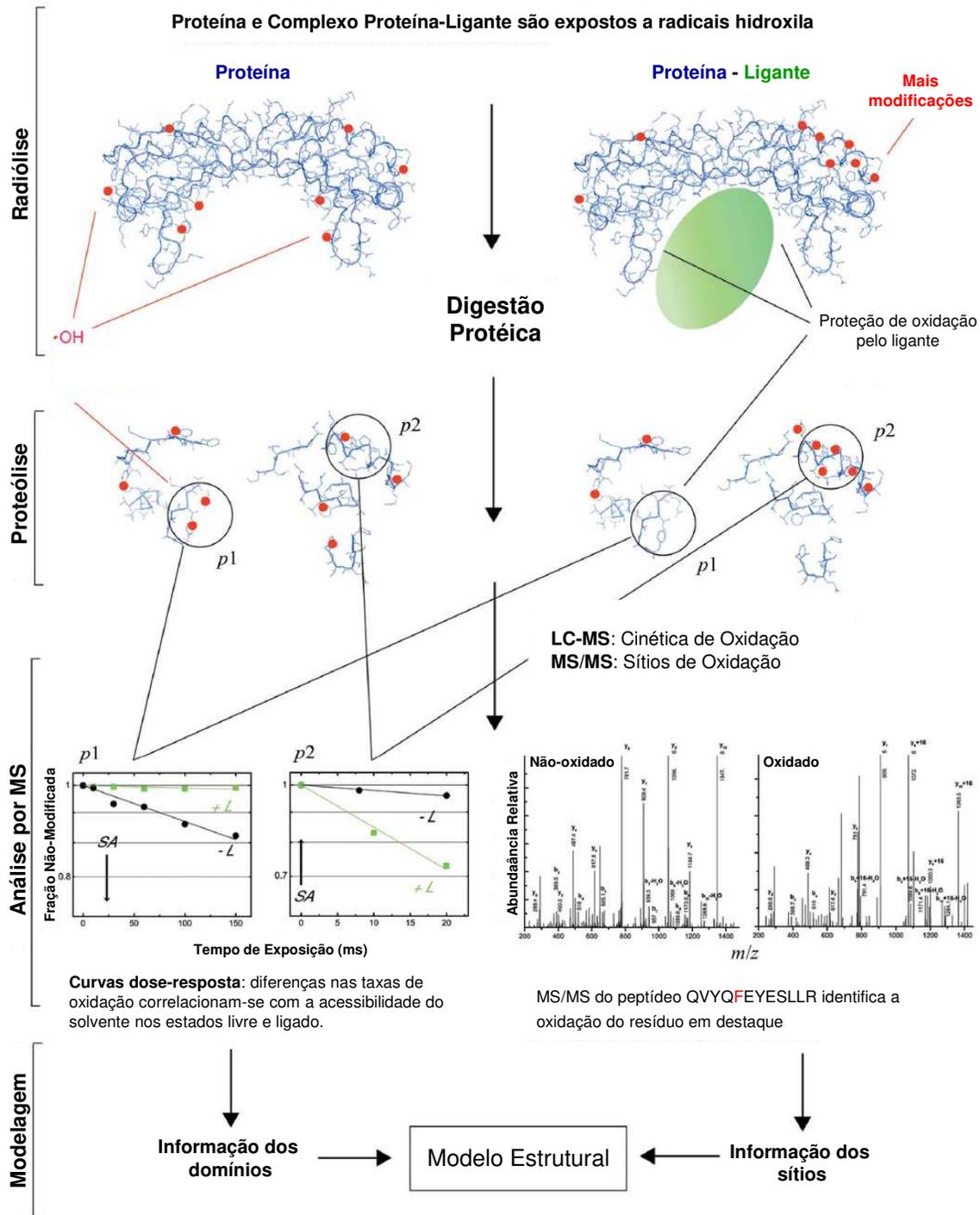
Além de ser reativo frente aos C $\alpha$  da cadeia principal, radicais hidroxila também apresentam reatividade frente à cadeia lateral dos aminoácidos, levando a formação de produtos de modificação covalente.<sup>134</sup> Desse modo, o grupo do pesquisador Mark Chance propôs a utilização dessas espécies como sonda para determinação de acessibilidade ao solvente dos diferentes resíduos de uma proteína.<sup>135</sup> Nesse estudo, foi possível determinar que controlando-se a dose de radicais  $\bullet\text{OH}$  às quais as proteínas são submetidas, é possível obter somente a formação de produtos de modificação (ausência de clivagem da cadeia polipeptídica). Em decorrência desse trabalho, novos estudos foram realizados com o objetivo de determinar quais seriam os produtos da reação de oxidação para os aminoácidos, sendo que diversos produtos de oxidação foram propostos para resíduos alifáticos/aromáticos<sup>135</sup>, básicos<sup>136</sup>, ácidos<sup>137</sup> e contendo enxofre<sup>138</sup>, bem como a ordem de reatividade intrínseca de cada aminoácido<sup>139</sup>.

A obtenção de radicais hidroxila em solução pode ser obtida de diversos modos. Os primeiros estudos estruturais envolvendo  $\bullet\text{OH}$  e proteínas envolveu a utilização da reação de Fenton, na qual a oxidação de íons  $\text{Fe}^{2+}$  por  $\text{H}_2\text{O}_2$  leva à formação dos radicais de interesse. Esse método, do fim do século XIX, foi aperfeiçoado em meados da década de 80, pela adição de EDTA e ascorbato ao meio reacional. Entretanto, a utilização de reagentes de Fenton

para geração de radicais sofre de algumas limitações comuns a outros métodos (fotólise de  $\text{H}_2\text{O}_2$  por laser<sup>140</sup> ou UV<sup>141</sup>, radiação  $\gamma$ <sup>142</sup>, métodos eletroquímicos<sup>143</sup> entre outros<sup>130</sup>): longos tempos de exposição do analito aos radicais (podendo chegar a algumas horas de exposição) e introdução de outros reagentes na amostra. Esses dois fatores podem levar a mudanças conformacionais na proteína ou no complexo antes ou durante o experimento, levando a resultados errôneos.

Em virtude desses fatos, Mark Chance propôs a utilização de um feixe de luz síncrotron de alto fluxo para a geração de radicais hidroxila em solução aquosa. O processo envolve a ejeção de elétrons da água devido à exposição a raios-X na faixa de alguns kV de energia; esses elétrons de alta energia ionizam outras moléculas de água, que por sua vez reagem com moléculas de água neutra levando a formação de radicais hidroxila. Em virtude desse “efeito cascata”, a concentração de radicais hidroxila mantém-se constante (aproximadamente 1  $\mu\text{M}$ ) ao longo do tempo de exposição.<sup>130,135</sup> Como o processo envolve somente a radiólise da água, não se faz necessária a adição de nenhum reagente além das proteínas de interesse. Além disso, como a concentração de radicais é muito alta e se mantém constante ao longo do tempo de exposição, as proteínas podem ser irradiada por intervalos de tempo muito curtos (da ordem de ms), tempo esse compatível com mudanças conformacionais de macromoléculas. Isso implica tanto na ausência de mudanças conformacionais significativas ao longo do tempo de exposição (revelando regiões da proteína que até então encontravam-se enterradas na estrutura da proteína) como na possibilidade da realização de experimentos resolvidos no tempo, onde pode-se determinar a dinâmica de interação proteína-ligante.

O esquema do experimento de *footprinting* acoplado a MS com o intuito de determinar a superfície de interação proteína ligante está mostrado na Figura 8. O experimento típico de *footprinting* consiste na exposição tanto da proteína isolada quanto do complexo proteína-ligante ao raio-X proveniente da fonte de luz síncrotron por diferentes intervalos de tempo. Após isso, as amostras são digeridas e analisadas por MS. Análises do tipo LC-MS/MS permitem identificar os sítios de oxidação, enquanto análises de LC/MS são utilizadas para quantificar a oxidação. O próximo passo consiste na determinação da cinética de oxidação dos peptídeos obtidos anteriormente e comparação dos dados: aqueles peptídeos que se encontram na superfície de interação proteína-ligante ficam protegidos da reação com  $\bullet\text{OH}$  e, portanto, tem sua cinética de oxidação diminuída em relação a amostra sem ligante. Métodos computacionais podem ser subseqüentemente utilizados para a construção de modelos tridimensionais do complexo.<sup>144</sup> Utilizando o método de *footprinting* com geração de radicais por meio de fontes de luz síncrotron, tem sido possível realizar estudos relacionados a estrutura de proteínas em solução<sup>145</sup> e o seu desenovelamento<sup>146,147</sup>, acompanhando a dinâmica desse processo em estudos resolvidos no tempo. Além disso, também foi possível determinar sítios de interação proteína-ligante, sendo o ligante íons metálicos<sup>148,149</sup>, peptídeos<sup>150</sup>, proteínas<sup>151,152,153</sup> e DNA/ RNA<sup>154,155</sup>.



**Figura 8** – Esquema experimental da técnica de *footprinting* acoplada a MS para determinação da superfície de interação proteína-ligante. (Adaptada da referência 156).

#### I.4 COMPLEXO TIF34/TIF35 DE *Saccharomyces cerevisiae*

O início do processo de síntese de proteínas consiste no recrutamento do complexo ribossomo-tRNA iniciador para o códon de iniciação do mRNA. Em eucariotos esse processo é muito complexo, uma vez que envolve várias interações proteína-RNA e proteína-proteína. Nesse caso, é necessária a participação coordenada de inúmeros fatores de iniciação de tradução (eIFs) para manter o correto funcionamento da maquinaria de síntese protéica. O fator de iniciação de tradução 3 (eIF3) foi primeiramente isolado e purificado como um complexo de alto peso molecular de reticulócitos de coelho. Desde sua caracterização inicial, sua estrutura e função são pouca compreendidas e mesmo a composição e a estequiometria das subunidades não foram determinadas rigorosamente. Estudos funcionais da eIF3 em mamíferos têm demonstrado que esse complexo tem papel fundamental na via de iniciação, organizando estruturalmente outros componentes traducionais na superfície da subunidade ribossomal 40S. O eIF3 liga-se diretamente à subunidade 40S do ribossomo na ausência de outros componentes e afeta a associação/dissociação dos ribossomos; além disso, ele também promove a ligação do tRNA e mRNA à subunidade 40S (Figura 9).<sup>157</sup>

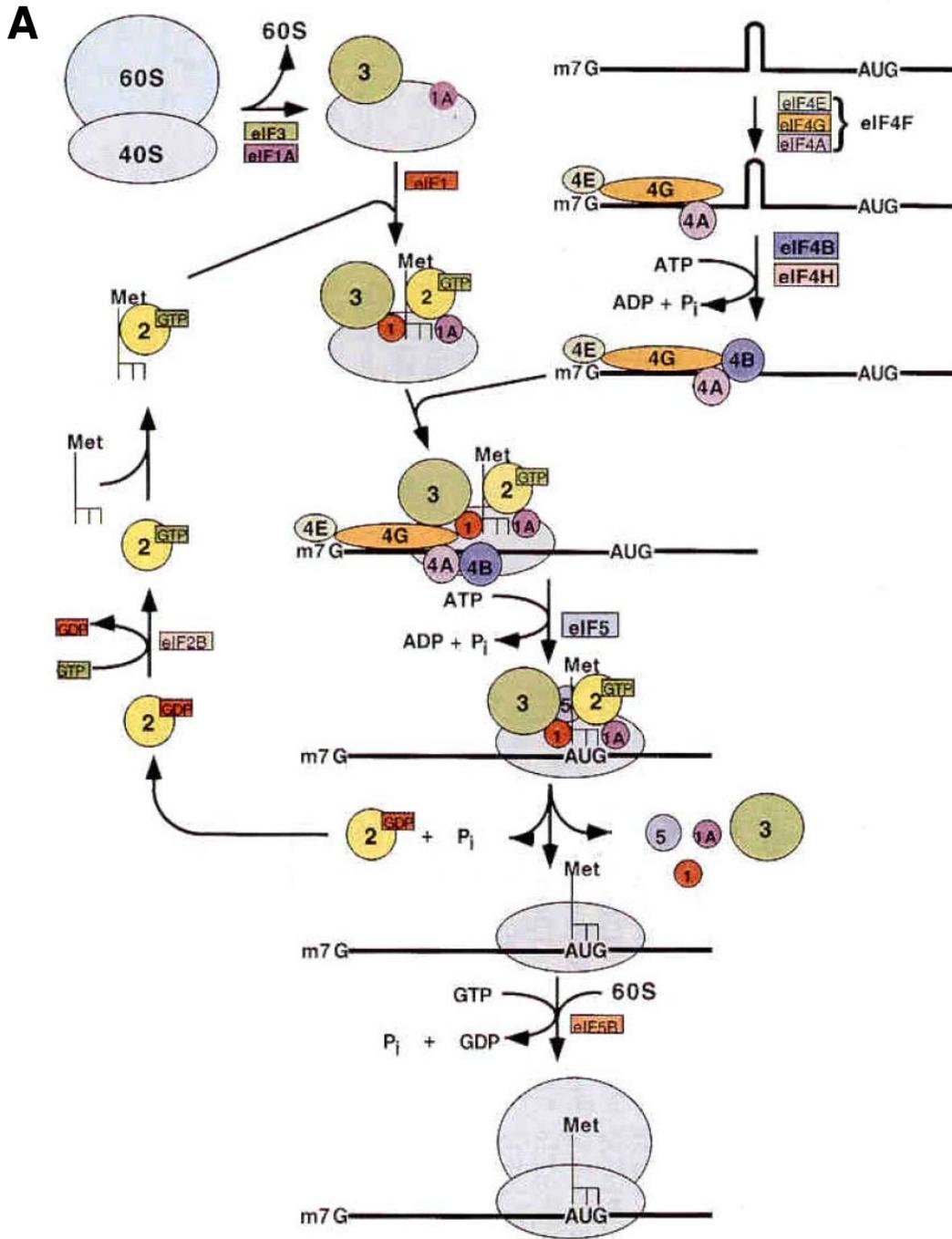
A levedura *Saccharomyces cerevisiae* tem sido muito utilizada no estudo da eIF3, sendo que análises de genética molecular têm contribuído para o entendimento da função desta neste organismo. Análises de interação entre as subunidades da eIF3 por métodos bioquímicos identificaram cinco unidades centrais componentes da eIF3 (Tif32p, Prt1p, Nip1p, Tif34p e Tif35p) além da subunidade Hcr1p, sendo todas essenciais para o crescimento celular. Estudos de interação entre as proteínas do complexo eIF3 indicaram que as três

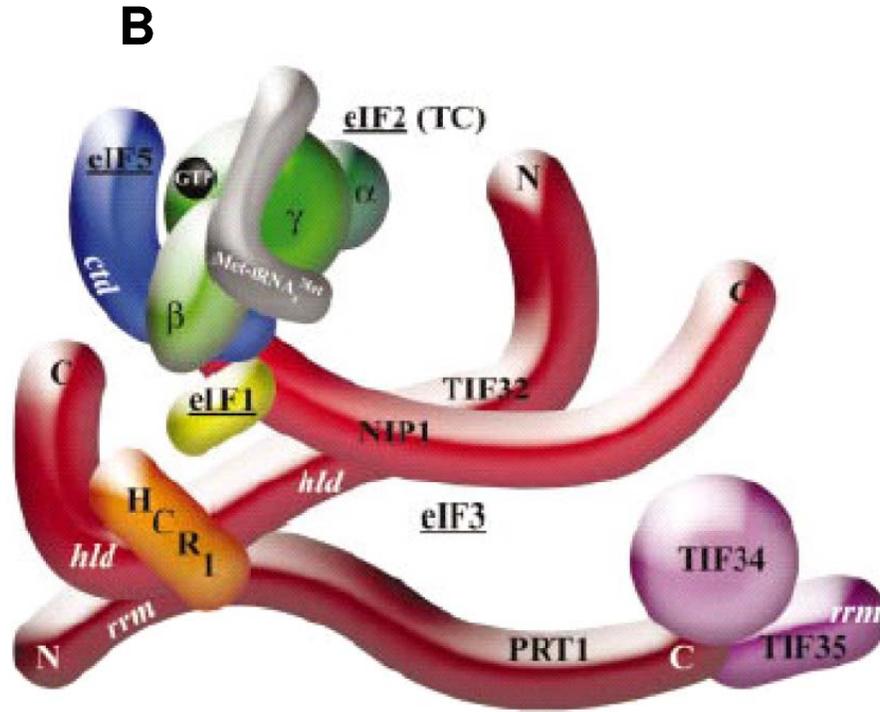
maiores subunidades (Tif32p, Prt1p, Nip1p) apresentam sítios de ligação para as outras duas e que a extremidade C-terminal da Prt1p adicionalmente interage com as subunidades Tif34p e Tif35p (Figura 9).<sup>158,159</sup>

A proteína Tif34p, de aproximadamente 39 kDa, é codificada pelo gene *TIF34*. Essa proteína é necessária para a viabilidade celular, além de ser importante na iniciação da tradução, possivelmente por ser necessária para a montagem e manutenção do complexo eIF3. Experimentos conduzidos com a deleção do gene *TIF34* resultaram na rápida degradação de todas as subunidades da eIF3, o que é condizente com o papel da Tif34p na integridade, estabilidade e manutenção do complexo eIF3. Estudos de interação (duplo-híbrido) indicaram que a proteína Tif34p interage com a proteína Tif35p. Essa proteína, de aproximadamente 30 kDa, é codificada pelo gene *TIF35* e é essencial na iniciação da síntese de proteínas. Além disso, ela também apresenta um domínio de ligação ao RNA na sua porção C-terminal. Aparentemente, a Tif35p tem um importante papel na organização estrutural e integridade funcional do complexo eIF3.<sup>160</sup>

Por análise de similaridade de seqüência primária, identificou-se a proteína Tif34p como pertencente à família de proteínas *WD-repeat*. Tal família é caracterizada pelo domínio WD, que compreende de 4 a 16 repetições conservadas de aproximadamente 40 aminoácidos, iniciadas com o dipeptídeo GH (11 a 24 resíduos do N-terminal) e apresentam o dipeptídeo WD na porção C-terminal. As proteínas *WD-repeat* são encontradas no citoplasma e no núcleo desempenhando várias funções como processamento e síntese de RNA e regulação do ciclo celular. A partir de estudos estruturais, revelou-se que os domínios WD formam folhas- $\beta$  antiparalelas conectadas por alças e compostas predominantemente por quatro fitas cada. Essas folhas- $\beta$

dobram-se em estruturas altamente ordenadas e simétricas, denominadas  *$\beta$ -propellers*. Apesar de serem estruturalmente importantes, até o momento sabe-se somente que as repetições WD são importantes para a interação entre proteínas, servindo como base para diversos processos vitais. Em vista dos dados funcionais sobre a Tif34p e outras proteínas da família WD, supõe-se que os domínios formadores da estrutura  *$\beta$ -propeller* estejam envolvidos na interação com outras subunidades da eIF3. Estes dados indicam que a Tif34p estaria relacionada à iniciação da tradução como uma proteína adaptadora, interagindo não somente com a Tif35p, mas com duas outras ao mesmo tempo, atuando como base para todo o complexo eIF3.<sup>161</sup>





**Figura 9** – (A) Modelo de iniciação de síntese de proteínas em eucariotos. Os números na Figura indicam os diferentes fatores de iniciação de tradução (eIF). M7G indica 7-metilguanossina. (referência 157) (B) Modelo de interação entre as unidades do eIF3 (referência 158). TC indica o complexo ternário eIF2/GTP/tRNA<sub>i</sub>, ctd indica o C-terminal, hld refere-se ao domínio tipo HCR1 e rrm significa sítio de reconhecimento de RNA.

## II OBJETIVOS

### II.1 OBJETIVO GERAL

O objetivo geral desta Tese foi o desenvolvimento de técnicas em Espectrometria de Massas visando a obtenção de informações de estruturas superiores de proteínas e complexos protéicos.

### II.2 OBJETIVOS ESPECÍFICOS

Nessa Tese foram utilizadas as técnicas de *cross-linking* e *footprinting* como métodos de caracterização de estrutura de proteínas e suas interações por Espectrometria de Massas. Os conhecimentos gerados foram posteriormente aplicados no estudo de interação de um complexo protéico. Com isso, os objetivos específicos dessa Tese foram:

- i) Estudo de fragmentação de peptídeos contendo *cross-linkers* intermoleculares;

- 
- ii) Utilização do modo de Varredura de Íons Precursores para identificação de peptídeos modificados por *cross-linkers*;
  - iii) Desenvolvimento de uma linha de luz no Laboratório Nacional de Luz Síncrotron para realização de experimentos de *footprinting*;
  - iv) Desenvolvimento de um método de quantificação baseado em LC-MS para cálculo de cinética de oxidação de peptídeos gerados nos experimentos de *footprinting*;
  - v) Utilização de *cross-linking* e *footprinting* para o estudo de interação entre as proteínas Tif34 e Tif35 do Fator de Iniciação de Tradução 3 de *Saccharomyces cerevisiae*.

## III PARTE EXPERIMENTAL

### III.1 *CROSS-LINKING*

#### III.1.1 Estudo de Fragmentação de Peptídeos Contendo DSS

##### III.1.1.1 Reação de Cross-Linking

Para o estudo de fragmentação de peptídeos contendo *cross-linking*, foram utilizadas as seguintes seqüências de peptídeos: AGAKGAERLVKAGVR (PX), Ac-ARKGCREVTKNDLR (P1), Ac-ARGKWPREVKIHR (P2) E Ac-YTKDLSQRAFKGMR (P3). Esses peptídeos (Proteimax, SP) foram utilizados conforme obtidos. O reagente de *cross-linking* utilizado foi o diéster de N-hidroxisuccinimida do ácido subérico (DSS, Pierce). Os peptídeos foram dissolvidos em tampão fosfato de sódio 50 mM pH 7.0 a uma concentração final de 40 µM; DSS por sua vez foi dissolvido em DMF em concentração final de 1 mg/mL e adicionado em 50 vezes de excesso molar à solução anterior. A reação foi mantida a temperatura ambiente por duas horas, e após isso adicionou-se tampão Tris•HCl (1 M, pH 7.6) a concentração final de 50 mM. Após 15 minutos, adicionou-se tripsina

(razão tripsina/peptídeo 1:50 em massa) e a digestão enzimática foi conduzida por 3 horas a 37° C. Após a digestão os produtos da reação foram purificados usando o cartucho Oasis HLB (Waters Co.) e o solvente foi evaporado no SpeedVac.

### III.1.1.2 Análise dos Produtos de *Cross-Linking* por MS

Para as análises por MALDI-MS e MS/MS, as amostras foram resuspensas em 1 mL de solução aquosa contendo 50% ACN e 0.1% TFA. A amostra foi preparada pelo método *dried droplet*, utilizando-se como matriz o CHCA (10 mg/mL, na mesma solução do peptídeo). Os espectros foram adquiridos no equipamento Q-Tof Premier (Waters Co.), que opera com um laser de estado sólido de 200 Hz. Os parâmetros de análise foram 250 a.u. (energia do laser), 10 V (sample plate) e 2200 V (MCP). O gás de colisão utilizado foi o Argônio e os espectros de fragmentação foram adquiridos manualmente, aumentando-se a energia de colisão até a completa dissociação do precursor.

As análises de ESI-MS e MS/MS foram realizadas após resuspender as amostras em 1 mL de solução aquosa contendo 50% ACN e 0.1% ácido fórmico, utilizando-se para isso o equipamento Q-Tof Ultima (Waters Co.). Os experimentos foram realizados por infusão direta da amostra com fluxo de 2  $\mu$ L/min, com voltagens do cone e capilar de 100 V e 3.5 kV, respectivamente. Como no caso anterior, os espectros de fragmentação foram adquiridos aumentando-se a energia de colisão até completa dissociação do precursor. Os espectros adquiridos foram deconvoluídos utilizando-se o

algoritmo MaxEnt3 (Waters Co.). Em todos os casos, os espectros foram calibrados internamente utilizando-se como padrão a massa do precursor.

### III.1.2 Varredura de Íons Precursores (PIS)

#### III.1.2.1 Preparação das Amostras

A amostra denominada PX+Hemoglobina foi preparada misturando-se quantidades equimolares da amostra PX (cuja preparação foi descrita no item anterior) com um padrão de digestão trípico de Hemoglobina bovina (Waters Co.).

O experimento de cross-linking de Lisozima (Sigma) com DSS foi realizado adicionando-se o *cross-linker* dissolvido em DMF (1 mg/mL) à solução de 10 µM de proteína em tampão fosfato de sódio 50 mM pH 7.5, na proporção molar de 1:50 (proteína/DSS). Após 30 minutos de reação, adicionou-se Tris•HCl 1M pH 7.5 de modo que sua concentração final fosse 50 mM. Após 15 minutos, adicionou-se DTT (razão tripsina/proteína 1:50 em massa) e a reação foi conduzida por 1 hora a 60° C. Subseqüentemente, adicionou-se IAA (1:10 m/m), sendo que a reação ocorreu por 30 minutos no escuro; após esse período, tripsina foi adicionada (1:50 m/m) e a digestão enzimática foi conduzida a 60° C por 1 hora.

### III.1.2.2 Análise por LC-MS/MS

A análise de PIS foi realizada no equipamento Q-Tof Ultima (Waters Co.) acoplado *online* com o sistema cromatográfico UPLC nanoAcquity (Waters Co.). Em geral, 1  $\mu\text{L}$  de amostra foi injetada e os peptídeos foram dessalinizados na coluna Symmetry C18 (Waters Co.) por 5 minutos. Após esse período, os peptídeos foram direcionados para a coluna BEH130 C18 (Waters Co.), onde foi feito um gradiente (45 minutos) de água/0.1% ácido fórmico (Sol. A) e ACN/0.1% ácido fórmico (Sol.B), fluxo 1.1  $\mu\text{L}/\text{min}$ , na coluna com o seguinte perfil: 100%  $\rightarrow$  70% A de 0 a 30 minutos; 70%  $\rightarrow$  50% de 30 a 32 minutos; 50% até 38 minutos; 50%  $\rightarrow$  100% de 38 a 40 minutos. Os parâmetros utilizados no equipamento foram os seguintes: 2 s (tempo de aquisição do espectro), 3.5 kV (voltagem do capilar), 100 V (voltagem do cone), 10 eV (energia de colisão no modo de baixa energia), 70-85 eV (energia de colisão no modo de alta energia) e 2200 V (MCP). Os espectros de fragmentação foram calibrados utilizando-se a massa do trímero de ácido fosfórico injetado ao longo da corrida ( $m/z$  294.9385).

## III.2 FOOTPRINTING

### III.2.1 Desenvolvimento da Linha D06B

Os primeiros experimentos de *footprinting* foram realizados na linha de luz DFX do Laboratório Nacional de Luz Síncrotron (LNLS). Essa linha contém um *shutter* (dispositivo que controla o tempo de exposição da amostra) com resolução de aproximadamente 1 s. Após os primeiros resultados positivos, os experimentos passaram a ser realizados na linha D06B do LNLS, onde foi desenvolvido um novo *shutter* controlado por computador, com resolução temporal de 100 ms. Também foi desenvolvido no LNLS um porta amostra controlado remotamente com capacidade para 6 amostras, sendo que a distância entre a saída do feixe de luz ( $10^{-9}$  bar) e a amostra (pressão ambiente) é de aproximadamente 2 cm. Para alinhar o feixe de luz com o porta-amostra, colocou-se uma câmera CCD dentro da estação experimental e adaptou-se um pedaço de cerâmica (material que fluoresce quando irradiado) à esse dispositivo; o *shutter* foi então aberto e a posição ajustada de modo a maximizar a área de incidência de radiação na amostra.

Os experimentos para otimização das condições experimentais (tempo de exposição e posição do porta-amostra) foram realizados com uma solução aquosa saturada de diclorofluoresceína (DCF). Após a exposição, as amostras foram analisados no fluorímetro K2 (ISS Co.), sendo irradiadas em 512 nm e sua fluorescência medida em 530 nm.

### III.2.2 Quantificação *Label-Free* da Cinética de Oxidação

Visando quantificar as taxas de oxidação pelo método *label-free*, 10  $\mu\text{L}$  de Mioglobina de coração de cavalo (Sigma) em tampão fosfato 50 mM pH 7.5 foram irradiados pelos seguintes intervalos de tempo (em quintuplicata): 400, 500, 600, 700 e 800 ms. Após isso, adicionou-se metionina em concentração final de 10 mM e tripsina (1:50 m/m) e a digestão enzimática foi conduzida a 37° C *overnight*. Os peptídeos resultantes foram injetados no equipamento Q-Tof Ultima (Waters Co.) acoplado *online* ao cromatógrafo nanoAcquity UPLC (Waters Co.). 1  $\mu\text{L}$  de amostra foi injetada e os peptídeos foram dessalinizados na coluna Symmetry C18 (Waters Co.) por 5 minutos. Após esse período, os peptídeos foram direcionados para a coluna BEH130 C18 (Waters Co.), onde foi feito um gradiente (60 minutos) de água/0.1% ácido fórmico (Sol. A) e ACN/0.1% ácido fórmico (Sol.B), fluxo de 1.1  $\mu\text{L}/\text{min}$ , com o seguinte perfil: 90%  $\rightarrow$  70% A de 0 a 35 minutos; 70%  $\rightarrow$  50% de 35 a 45 minutos; 50% até 50 minutos; 50%  $\rightarrow$  90% de 50 a 55 minutos. Primeiramente realizou-se uma análise de LC-MS/MS, com o objetivo de identificar os peptídeos gerados pela digestão da proteína. Para isso, foi realizada uma análise do tipo DDA (*Data Dependent Analysis*), onde a cada segundo o equipamento adquire um espectro de MS. No caso da presença de espécies multicarregadas, as 3 mais intensas são fragmentadas na câmara de colisão (energia de colisão definida pela  $m/z$  e carga do precursor). Os dados foram posteriormente processados no *software* ProteinLynx (Waters Co.) e a busca no banco de dados foi realizada no programa Mascot (Matrix Science), com os seguintes parâmetros: MSDB (banco de dados), 0.1 Da (erro nos modos MS e MS/MS), tripsina (protease), permitindo peptídeos com 1

sítio de clivagem não-digerido e oxidação de metionina (modificação variável).

Uma vez que os peptídeos gerados pela digestão foram identificados, as quintuplicatas das amostras com diferentes tempos de irradiação foram submetidas a análises de LC-MS para quantificação, utilizando-se o mesmo gradiente descrito anteriormente. Após as corridas, extraiu-se a corrente iônica para cada um dos peptídeos identificados anteriormente e a área do pico cromatográfico correspondente foi integrada no programa MassLynx (Waters Co.). Os valores das áreas dos picos cromatográficos foram utilizados para contruir uma curva de intensidade do peptídeo em função do tempo, da qual foi extraída a constante de velocidade de oxidação para cada um dos peptídeos.

### **III.3 ESTUDO DE INTERAÇÃO DO COMPLEXO TIF34/TIF35**

#### **III.3.1 Expressão e Purificação de HisTif34p, HisTif35N e Tif34p/HisTif35N<sup>160</sup>**

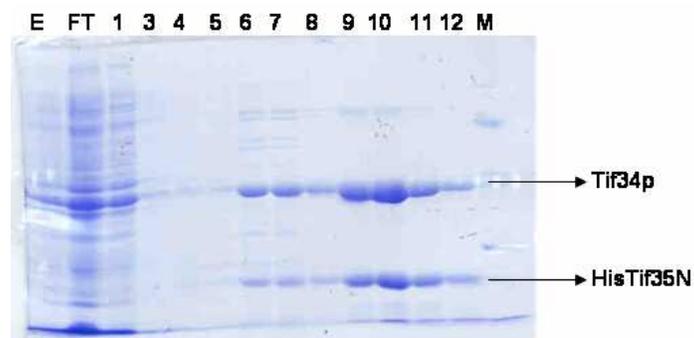
##### **III.3.1.1 Expressão das Proteínas em *Escherichia coli***

Para a expressão dos sistemas de interesse em *E. coli*, utilizou-se a cepa BL21(DE3) $\Delta$ slyD<sup>-</sup> (gene *slyD* deletado) e os plasmídeos pET28a/HisTif34p (proteína HisTif34p), pET28a/HisTif35N (proteína HisTif35N) e pET23a/TIF34 + pET28a/HisTIF35N (complexo Tif34p/HisTif35N). A indução do complexo foi realizada em meio de cultura LB (1% de peptona, 0,5% de extrato de levedura e 1% de NaCl) contendo os antibióticos

adequados (ampicilina para pET23a e canamicina para pET28a). As células foram incubadas nesse meio a 30° C sob agitação de 200 rpm até atingir OD<sub>600</sub> entre 0,8 e 1,0. Após isso, adicionou-se solução de IPTG na concentração final de 0,5 mM. Essa mistura foi mantida no shaker sob agitação de 200 rpm a 30° C por mais 4 horas e posteriormente centrifugada a 6000 rpm e 4° C por 10 minutos. As células foram ressuspensas em uma solução contendo tampão fosfato 50 mM pH 7,2, 100 mM de NaCl e 5% de glicerol, e a seguir adicionou-se lisozima e a mistura foi mantida em um banho de gelo por 30 minutos. O material foi então sonicado (10 vezes de 20 segundos) e centrifugado por 30 minutos a 4° C e 15000 rpm para isolar o sobrenadante.

### **III.3.1.2 Cromatografia de Afinidade**

As proteínas solúveis contendo cauda de histidina foram isoladas do extrato obtido anteriormente por cromatografia de afinidade com níquel imobilizado, utilizando-se para isso a resina Ni-NTA Superflow (Qiagen). Os extratos foram incubados com a resina por 1 hora a 4° C e posteriormente eluídos com tampão da amostra contendo concentrações crescentes de imidazol (5 a 200 mM). As frações coletadas foram então submetidas a análise por SDS-PAGE, sendo que as frações nas quais foram detectadas as proteínas de interesse foram combinadas (Figura 10).

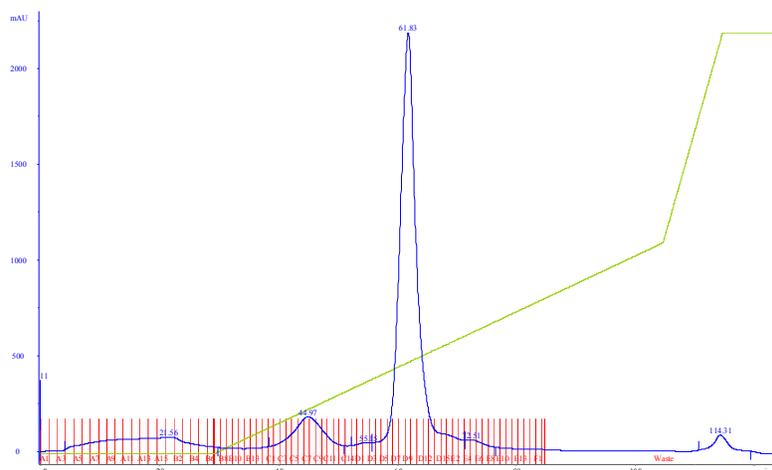


**Figura 10** - SDS-PAGE da purificação por afinidade do complexo Tif34p/HisTif35N. E: extrato total; FT: fração não retida na resina; 1-3: tampão sem imidazol; 4-5: 5mM imidazol; 6-8: 20 mM imidazol; 9-11: 50 mM imidazol; 12: 100 mM imidazol; M: Marcador de peso molecular. As frações 6-12 foram combinadas para o próximo passo da purificação.

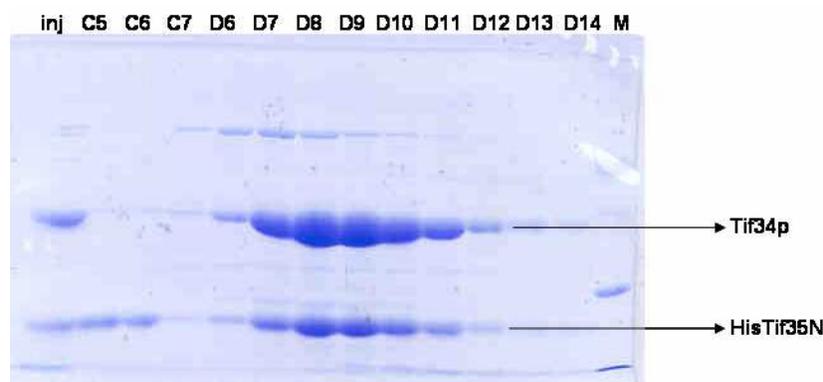
### III.3.1.3 Cromatografia de Troca Iônica

As frações coletadas na etapa anterior foram dialisadas a 4° C contra uma solução contendo 20 mM de tampão fosfato de sódio pH 7,2, 50 mM de NaCl, 5% de glicerol, 0,5 mM de DTT e 0,5 mM de PMSF. A cromatografia foi realizada no sistema AKTA-FPLC (Amersham-Biosciences), utilizando-se a coluna HiTrap Q Sepharose 5 mL (Amersham-Biosciences) com gradiente de 120 minutos (fluxo de 1 mL/min, frações de 1 mL), utilizando como tampão A a mesma solução da diálise e tampão B sendo constituído de tampão A com NaCl 1M (Figura 4). Antes da corrida a amostra foi centrifugada a 14000 rpm e 4° C por 10 minutos. Ao fim da cromatografia, as frações foram analisadas por SDS-PAGE (Figura 5) e as mais concentradas foram combinadas e concentradas para um volume de 300 µL utilizando-se a

membrana Amicon Ultra-PL5 5 kDa (Millipore), centrifugando-se a 4000 rpm e 4° C no concentrador UltraFree4 (Millipore).



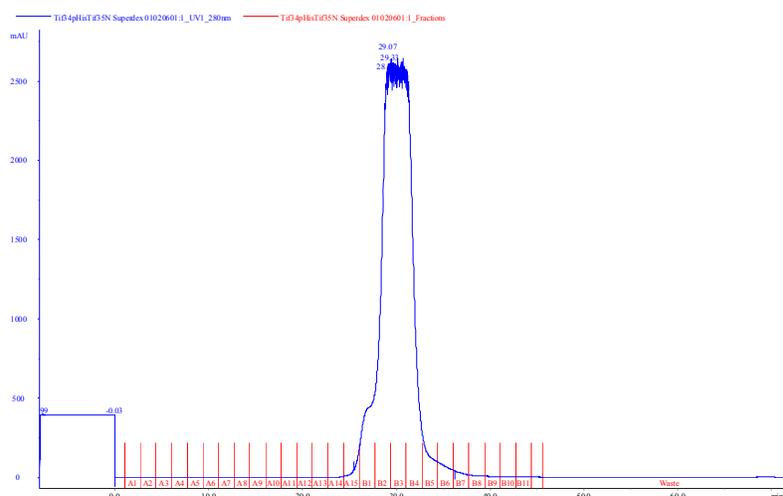
**Figura 11** - Cromatograma do complexo fracionado utilizando-se a coluna HiTrap Q Sepharose 5 mL (troca iônica).



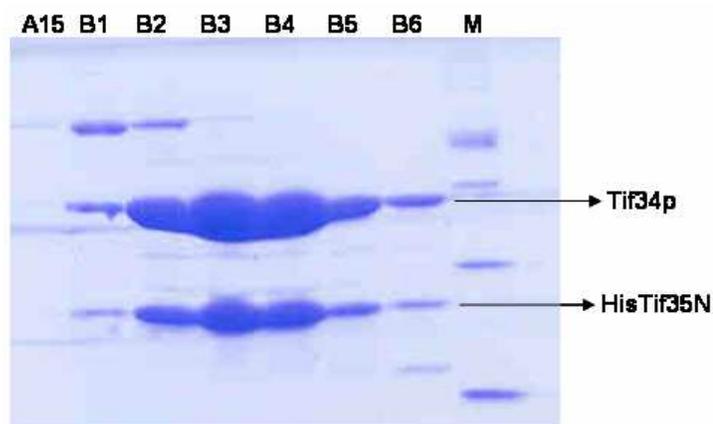
**Figura 12** - Exemplo de gel de SDS-PAGE da purificação por troca iônica do complexo Tif34p/HisTif35N. inj: amostra injetada; C5-C7 e D6-D14: frações coletadas, de acordo com o cromatograma da figura anterior; M: Marcador de peso molecular. As frações D6-D13 foram combinadas para o próximo passo da purificação.

### III.3.1.4 Cromatografia de Gel Filtração

Antes da cromatografia, a amostra foi centrifugada a 14000 rpm e 4° C por 10 minutos. O eluente utilizado foi uma solução contendo tampão fosfato de sódio 20 mM pH 7,2, 100 mM de NaCl e 0,5 mM de DTT e a coluna usada foi a Superdex 75 HR 10/30 - Amersham-Biosciences - (fluxo de 0,3 mL/min, frações de 0,5 mL).



**Figura 13** - Cromatograma do complexo fracionado utilizando-se a coluna Superdex 75 HR 10/30 (gel filtração).



**Figura 14** – Exemplo de gel de SDS-PAGE da purificação por gel filtração do complexo Tif34p/HisTif35N. A15 e B1-B6: frações coletadas, de acordo com o cromatograma da figura anterior; M: Marcador de peso molecular. As frações B1-B6 foram combinadas.

### III.3.2 *Cross-Linking*

O complexo Tif34p/HisTif35N purificado foi submetido à reação de *cross-linking* com o reagente DSS, adicionado na proporção molar de 1:50, por 30 minutos. Após esse período adicionou-se tampão Tris•HCl 1M pH 7.6 à concentração final de 50 mM. A mistura reacional foi separada por SDS-PAGE (12.5% de poliacrilamida) e a banda correspondente aos produtos de *cross-linking* foi recortada do gel, reduzida, alquilada e digerida com tripsina.<sup>162</sup> As amostras foram analisadas por LC-MS/MS, do mesmo modo que a mioglobina (descrito na seção III.2.2).

### ***III.3.3 Footprinting***

Os experimentos de *footprinting* com a proteína HisTif34p e o complexo Tif34p/HisTif35N foram realizados na linha de luz XRL (seção III.2.1) do LNLS. As amostras foram irradiadas por intervalos de tempo entre 400-800 ms, posteriormente digeridas e analisadas por LC-MS, conforme descrito anteriormente (Seção III.2.2).

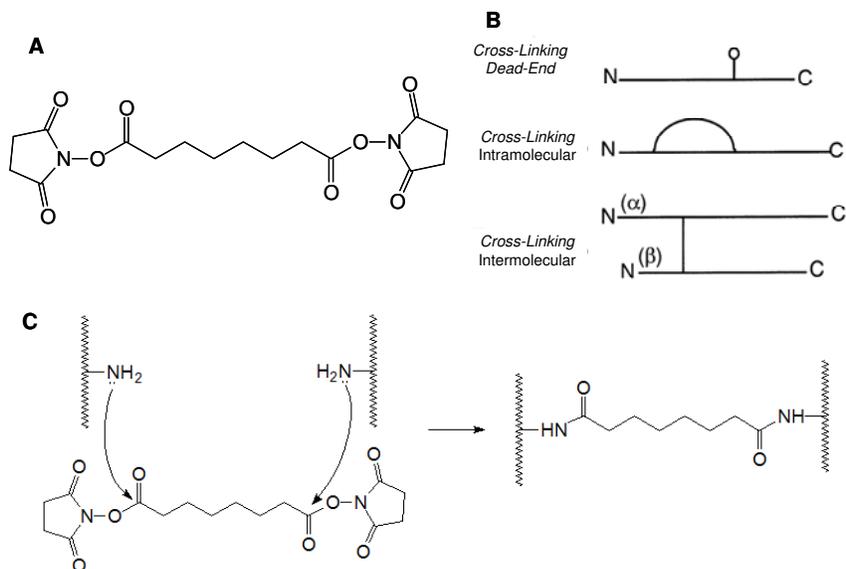
## IV RESULTADOS E DISCUSSÃO

### IV.1 *CROSS-LINKING*

#### IV.1.1 Introdução

Conforme descrito anteriormente, o acoplamento das técnicas de MS e *cross-linking*, também conhecido como MS3D (*Three-Dimensional Protein Structure Using Mass Spectrometry*), apresenta grande potencial na área de Genômica Estrutural. As técnicas de caracterização estrutural de proteínas de alta resolução disponíveis atualmente (RMN e CRP) necessitam de condições de amostra incompatíveis com um grande número de proteínas: alto grau de pureza, grandes quantidades de amostra e alta estabilidade, além das necessidades específicas de cada técnica. MS3D por sua vez é uma técnica extremamente versátil, uma vez que a amostra não apresenta limitações de tamanho e tampão (desde que esse seja compatível com o *cross-linker*); além disso, os experimentos são rápidos (podendo ser aplicados a espécies instáveis, proteínas que degradam rapidamente ou a complexos lábeis) e sensíveis, necessitando pequenas quantidades de proteína.

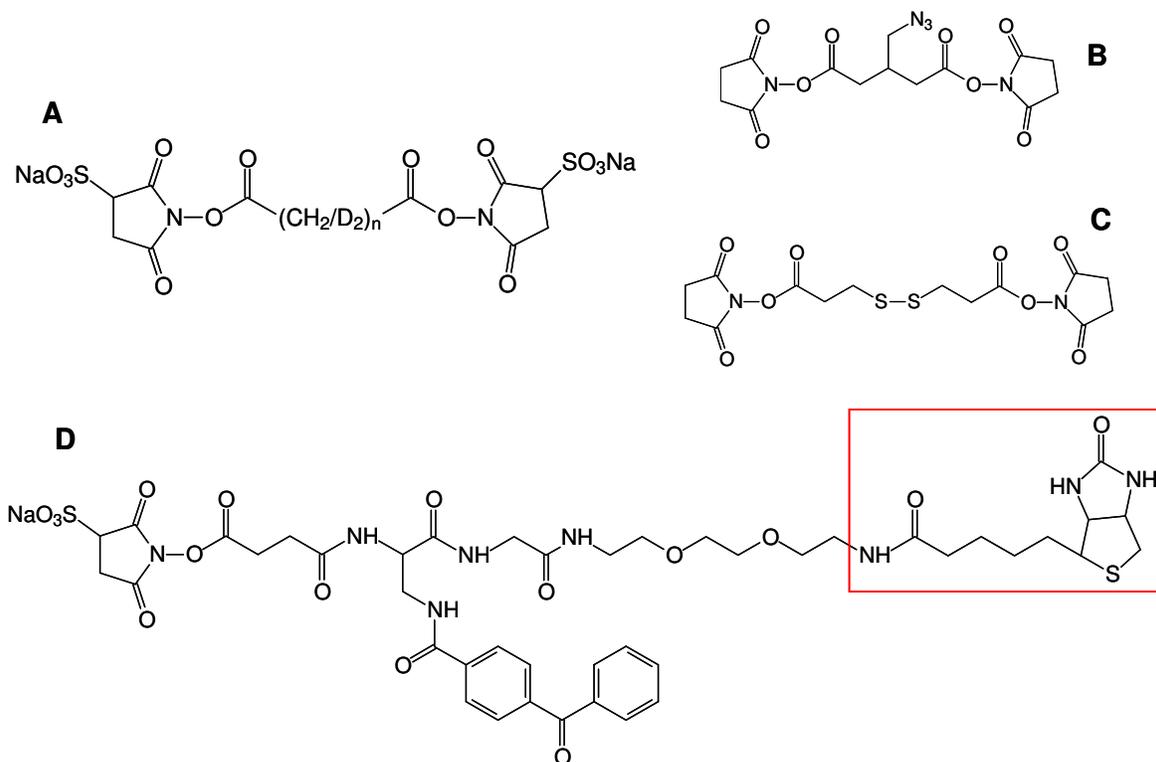
Como mostrado na Figura 6, o fluxograma experimental de um experimento de MS3D consiste na reação da proteína com o *cross-linker* de interesse. Dentre os inúmeros reagentes disponíveis, os *cross-linkers* mais utilizados são os derivados da N-hidroxisuccinimida (NHS), em especial o DSS (DiSuccinimidil Suberato – Figura 15A), devido principalmente à sua disponibilidade comercial e ao tamanho e flexibilidade de sua cadeia espaçadora. Ésteres de NHS são suscetíveis a ataque nucleofílico, levando a perda do anel succinimida e formação da ligação entre a cadeia espaçadora e o resíduo de aminoácido (Figura 15C). Como o DSS é um reagente homobifuncional, os produtos da reação de *cross-linking* podem ser do tipo intramolecular (resíduos conectados estão na mesma cadeia), intermolecular (resíduos conectados encontram-se em cadeias distintas) ou ainda *dead-end*, onde somente um dos grupos reativos se ligou a cadeia peptídica (Figura 15B).



**Figura 15** – (A) Estrutura do DSS. (B) Possíveis produtos de reação entre um peptídeo/proteína e um *cross-linker* bifuncional. (C) Exemplo de reação do DSS com a cadeia lateral de duas lisinas: para que esse produto seja formado, a distância entre esses resíduos deve ser compatível com o tamanho da cadeia espaçadora do DSS.

O próximo passo para a obtenção de informação estrutural no experimento de MS3D consiste na análise por MS dos produtos obtidos, com o intuito de detectar esses peptídeos modificados e anotar sua estrutura. Essa etapa constitui o maior gargalo atual na aplicação de MS3D para o estudo de estrutura de proteínas e suas interações. A detectabilidade desses peptídeos é muito difícil, uma vez que experimentalmente evita-se a formação de mais de um *cross-linking* por molécula de proteína (o que poderia, em princípio, distorcer a estrutura original, gerando restrições espaciais irreais). Isso faz com que após a digestão enzimática do complexo de interesse, a quantidade de peptídeos não-modificados seja muito maior do que a de peptídeos ligados entre si. Conseqüentemente, o sinal dos outros peptídeos acaba por suprimir a ionização das espécies de interesse. Várias abordagens têm sido descritas na literatura com o objetivo de contornar essa dificuldade, sendo que uma das

mais utilizadas consiste na marcação isotópica desses peptídeos tanto pela utilização de *cross-linkers* marcados<sup>163,164,165</sup> (Figura 16A) ou pela incorporação de átomos de  $^{18}\text{O}$  quando da digestão trípica da amostra em  $\text{H}_2^{18}\text{O}$ <sup>166,167</sup>. Outra abordagem consiste na utilização de *cross-linkers* trifuncionais, onde a funcionalidade extra é um grupo (em geral, a biotina) que pode ser utilizado para enriquecimento da amostra após a reação.<sup>168,169</sup> Embora o método pareça promissor, pouco sucesso foi obtido até o momento, devido em grande parte ao tamanho desses reagentes que dificulta a acessibilidade do *cross-linker* aos resíduos reativos (Figura 16D). Finalmente, pode-se destacar a utilização de reagentes cliváveis, onde após a reação de *cross-linking*, a cadeia espaçadora pode ser clivada por uma reação química que levará a formação das cadeias peptídicas separadas (Figura 16B e C).<sup>170,171</sup> Todos esses métodos baseiam-se no desenvolvimento de reagentes químicos com diferentes funcionalidades. Entretanto, nem sempre esses reagentes são de fácil síntese e posterior aplicação em soluções contendo proteínas, além de que em muitos casos o grande tamanho desses reagentes dificulta a acessibilidade desses aos resíduos.



**Figura 16** – Exemplos de *cross-linkers* descritos na literatura especificamente desenhados para auxiliar na detecção de peptídeos modificados: reagentes marcados isotopicamente (A) (referência 163), cliváveis com TCEP (B) (referência 171) e DTT (C) (referência 170) e com biotina como grupo de afinidade (D) (a biotina está destacada com o quadrado vermelho) (referência 169).

Com relação à identificação desses peptídeos modificados, tem sido proposto na literatura que a utilização de equipamentos do tipo ICR é de extrema importância, devido às altas resolução e exatidão de massas desse tipo de equipamento.<sup>105,106,119,172,173,174</sup> Isso se deve ao fato de que após a reação de *cross-linking*, um grande número de espécies podem ser geradas, acarretando em uma explosão combinatória onde praticamente todos os valores de  $m/z$  nominal poderiam corresponder a uma espécie presente na amostra. Desse modo, a utilização de equipamentos de alta exatidão aumentaria significativamente a confiabilidade dos dados gerados, diminuindo o teor de

falsos positivos que poderiam levar a uma estrutura que não condiz com a realidade.

Essa abordagem no entanto não é válida para casos onde dentro de um peptídeo exista mais de um possível sítio de ligação ao *cross-linker*. No caso de ésteres de NHS, quando as lisinas são os resíduos modificados, a tripsina perde a capacidade de digestão nesse sítio. Desse modo, caso não exista nenhum sítio de clivagem não digerido, a única lisina no meio da cadeia peptídica seria o resíduo ligado. Ocorre que mesmo no caso de digestão de proteínas isoladas em solução, é razoavelmente comum observar eventos de digestão incompleta. Em geral, isso ocorre pelo fato de que em solução não são todos os sítios de clivagem que estão expostos ao solvente e, conseqüentemente, à protease. No caso de proteínas ou complexos protéicos com *cross-linking*, esse caso é ainda mais crítico, uma vez que a reação de ligação cruzada torna a estrutura mais compacta, o que em alguns casos pode dificultar a digestão enzimática em algumas regiões da proteína. Em virtude desse fato, é comum observar em experimentos de *cross-linking* peptídeos grandes, em que dois ou mais sítios de clivagem não foram digeridos pela tripsina.<sup>173</sup> Com isso, no caso de um peptídeo hipotético modificado por um éster de NHS de seqüência XXXKXXXKXXKXXK, pode-se afirmar que com certeza o cross-linker não se encontra na última lisina, visto que ela foi digerida com tripsina; entretanto, não se pode dizer em qual das outras lisinas ele se encontra mesmo utilizando um equipamento com altíssima exatidão de massa (caso do FT-ICR), uma vez que os peptídeos modificados em cada uma das lisinas seriam isóbaros.

Outro caso onde a utilização de instrumentos de alta exatidão não seria suficiente para definir exatamente o sítio de ligação seria no caso de o *cross-linker* se ligar a um aminoácido diferente de lisina. Os principais softwares de

interpretação de dados gerados a partir de experimentos de MS desconsideram muitos dos possíveis sítios de ligação da proteína ao *cross-linker*.<sup>175,176</sup> Nos últimos anos, têm sido descrito na literatura que, embora ésteres de NHS reajam preferencialmente nas lisinas e N-terminal, vários outros resíduos também podem reagir, tais como serina, tirosina, treonina e histidina.<sup>177,178,179</sup> Embora esse grande número de possibilidades possa aumentar a dificuldade na interpretação dos dados, utilizando-se a versatilidade desses *cross-linkers* pode-se aumentar significativamente o número de restrições espaciais obtidas em um único experimento. Mais uma vez, a determinação inequívoca do sítio de modificação no peptídeo é dependente da análise por MS/MS.

Ocorre que não existem na literatura trabalhos sistemáticos investigando quais são os mecanismos de fragmentação e estruturas dos fragmentos de peptídeos conectados por *cross-linkers*. Muitos trabalhos reportam a interpretação de espectros de MS/MS de peptídeos ligados, utilizando para isso vários tipos de equipamentos e várias formas de fragmentação, tais como CID, IRMPD e ECD.<sup>165,168,180,181,182</sup> Entretanto, em todos esses casos a interpretação do espectro de MS/MS restringe-se a observação de íons comumente encontrados quando da fragmentação de peptídeos lineares (íons *y* e *b*) e longes do sítio de ligação, sendo que em geral boa parte do espectro não está interpretada, inclusive com íons intensos não anotados.

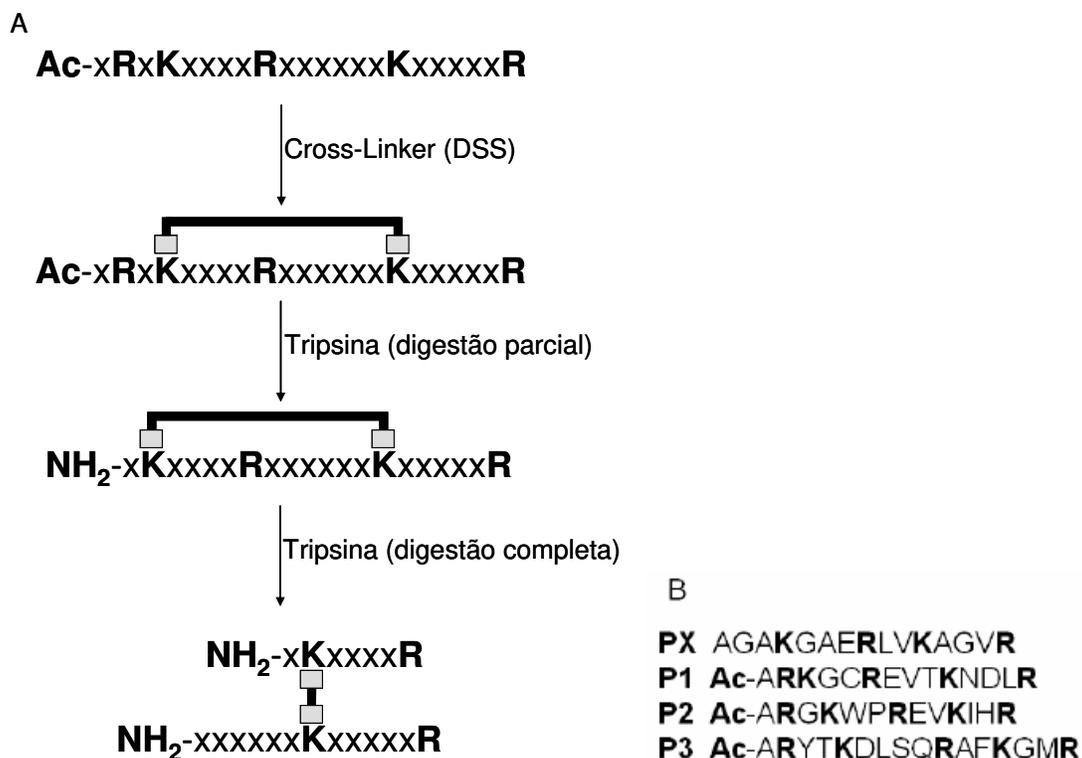
Em 2003, foi proposta pela primeira vez a possibilidade de implementação de uma ferramenta de bioinformática capaz de interpretar os dados de experimentos de MS3D automaticamente.<sup>175</sup> O desenvolvimento dessas ferramentas seria extremamente interessante, uma vez que aumentaria consideravelmente a velocidade de análise de dados e permitiria desse modo a elucidação de inúmeras estruturas em curto intervalo de tempo. Essa ferramenta foi dividida em duas partes, ASAP e MS2Assign. Enquanto o

ASAP seria responsável pela atribuição dos *cross-linkings* pelos dados de MS, o MS2Assign seria utilizado para validar essas identificações pelos dados de MS/MS. Desse modo, vários mecanismos de fragmentação foram propostos, incluindo a formação de vários íons cuja denominação foi proposta. Entretanto, nenhum espectro de fragmentação apresentado continha as espécies propostas, sendo estes dominados pela presença de íons do tipo *y* e *b*.

Posteriormente, o mesmo grupo realizou um estudo sistemático de fragmentação de peptídeos contendo diferentes *cross-linkers* derivados de ésteres de NHS.<sup>183</sup> Para realizar esses estudos, foram utilizados peptídeos contendo múltiplas repetições de um aminoácido (A ou G), contendo uma ou duas K (para o estudo de espécies inter ou intramoleculares, respectivamente) e o N-terminal acetilado. A acetilação do N-terminal foi realizada para bloquear esse sítio de reação com o *cross-linker*, dirigindo a reação somente à lisina. Isso evitaria a formação de espécies isóbaras que impediriam o estudo de fragmentação dos peptídeos de interesse. Entretanto, a acetilação do N-terminal de peptídeos altera significativamente o perfil de fragmentação dessas espécies, visto que essa modificação exclui um sítio de protonação móvel da molécula.<sup>184</sup> Além disso, a utilização de seqüências baseadas na repetição de aminoácidos em muitos casos levou a formação de fragmentos isóbaros, de modo que suas estruturas não puderam ser atribuídas. Em vista disso, os mecanismos de fragmentação desses peptídeos não poderiam ser aplicados em experimentos de MS3D, onde proteínas com seqüências aleatórias são modificadas e posteriormente digeridas levando a formação de peptídeos trípticos. Desta forma, um dos objetivos desta Tese consistiu no estudo do padrão de fragmentação de peptídeos trípticos contendo *cross-linkers*.

### IV.1.2 Desenho dos Peptídeos Modelos para Estudos de Fragmentação

Com o intuito de obter peptídeos que mimetizassem as espécies encontradas em experimentos de MS3D, foi proposta a utilização de peptídeos modelos que após a reação com DSS e digestão com tripsina gerassem peptídeos modificados realmente trípticos.<sup>185</sup> Esses peptídeos apresentam a seqüência comum **Ac-XXRXXXKXXXRXXXKXXXR**, onde X representa qualquer aminoácido que não lisina e arginina (Figura 17). A única exceção é o peptídeo PX, que não apresenta o N-terminal acetilado; o objetivo nesse caso foi avaliar a fragmentação de peptídeos intramoleculares com o N-terminal.

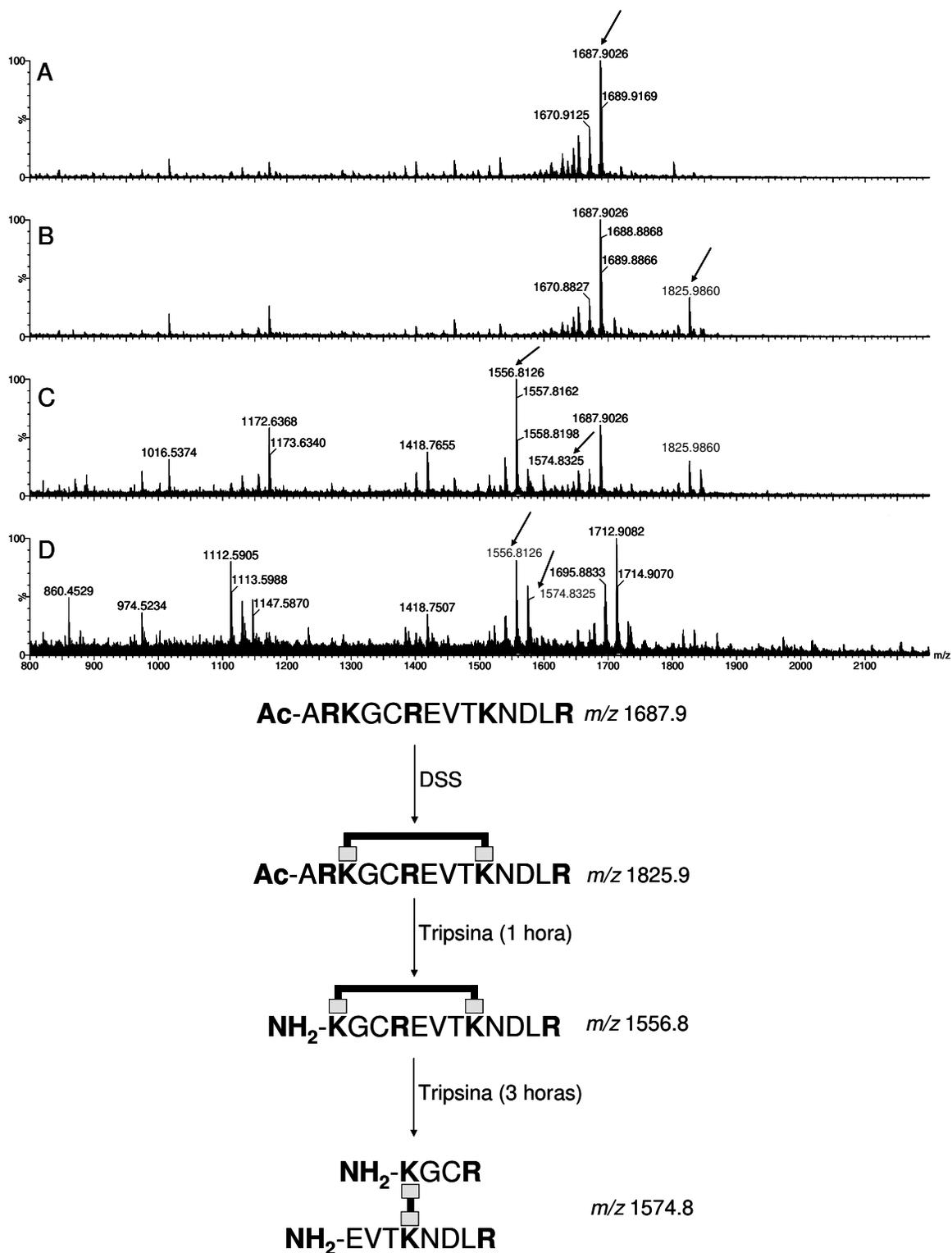


**Figura 17** - (A) Esquema geral da reação dos peptídeos com *cross-linking* (DSS nesse caso) para a formação de espécies intra e intermoleculares. (B) Sequências dos peptídeos utilizadas nesse estudo.

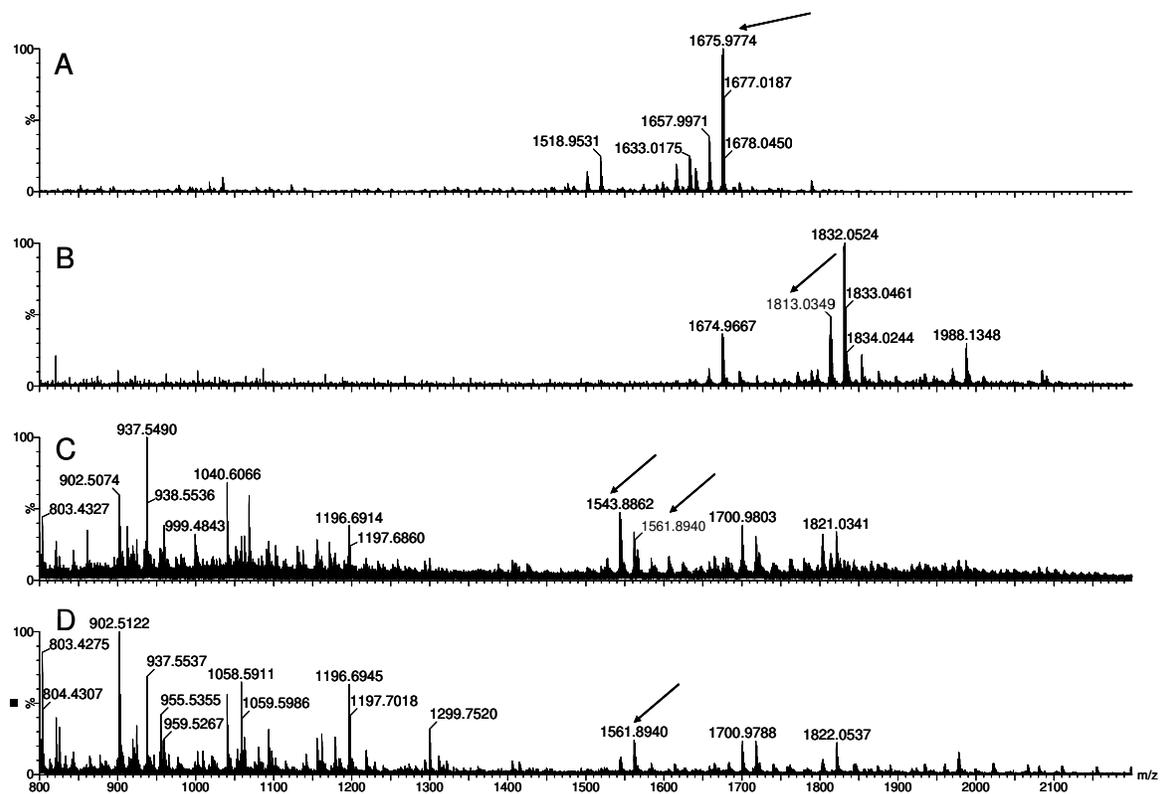
A utilização de diferentes peptídeos com seqüências aleatórias permitiu a atribuição inequívoca de vários íons fragmentos que seriam isóbaros no caso da utilização de seqüências com repetições de um mesmo aminoácido. Esses peptídeos também permitiram o estudo de fragmentações dependente da seqüência, ao mesmo tempo em que as posições fixas dos resíduos de lisina e arginina possibilitaram a formação tanto de espécies intra e intermoleculares em um único experimento e a partir de um único peptídeo precursor. Além disso, como as espécies intermoleculares são geradas a partir dos peptídeos intramoleculares, não é necessária a utilização de grandes concentrações de peptídeos nos experimentos, o que reduz significativamente os custos; de fato, a concentração de peptídeo utilizada nesse trabalho é cerca de 200 vezes

menor que o de experimentos descritos na literatura.<sup>183</sup> A estratégia experimental para isso é extremamente simples (Figura 17): os peptídeos reagem com DSS por 2 horas e após esse período é realizada a digestão enzimática com tripsina, onde o tempo da proteólise favorece a formação das espécies intra ou intermoleculares.

As Figuras Figura 18, Figura 19, Figura 20 e Figura 21 apresentam respectivamente os espectros de MALDI-MS dos peptídeos P1, P2, P3 e PX antes e depois da reação com o DSS, digestão com tripsina por 1 hora (P1, P2 e P3) e 3 horas. Pelos espectros é possível notar que o rendimento da reação de *cross-linking* é dependente da seqüência dos peptídeos, sendo que enquanto para P1 o rendimento foi de aproximadamente 30%, a reação com PX converteu praticamente todo o material de partida no produto desejado. A formação de outras espécies, como peptídeos tipo *dead-end* também é afetada pela seqüência do peptídeo. Além disso, é possível observar que com diferentes tempos de digestão pode-se formar preferencialmente as espécies intra ou intermolecular, conforme esperado. Nota-se ainda que o controle do tempo de digestão com tripsina determina a presença majoritária de espécies intra ou intermoleculares, sendo que também esses rendimentos se mostraram dependentes da seqüência dos peptídeos. Esse comportamento é esperado visto que a cinética de digestão enzimática é mais rápida no resíduo de arginina exposto, gerando preferencialmente a espécie intramolecular não-bloqueada em menores intervalos de digestão.



**Figura 18 e Esquema 1** - Espectros de MALDI-MS do peptídeo P1 intacto (A), após a reação com DSS (B) e digestão com tripsina por 1 hora (C) e 3 horas (D). Os íons assinalados com uma seta correspondem às espécies de interesse.



**Ac-ARGKWPREVKIHR**  $m/z$  1674.9

DSS

**Ac-ARGKWPREVKIHR**  $m/z$  1813.0

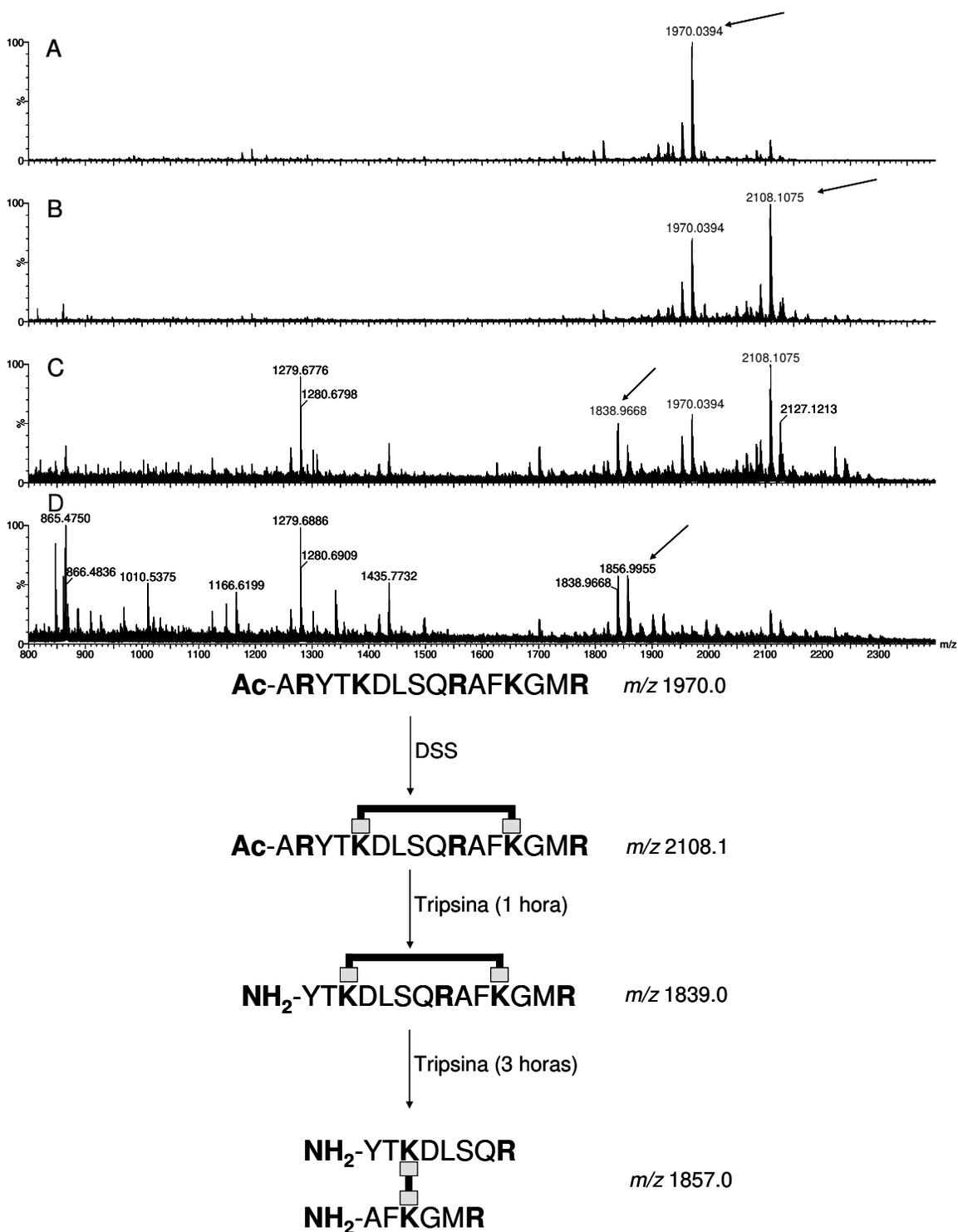
Tripsina (1 hora)

**NH<sub>2</sub>-GKWPREVKIHR**  $m/z$  1543.9

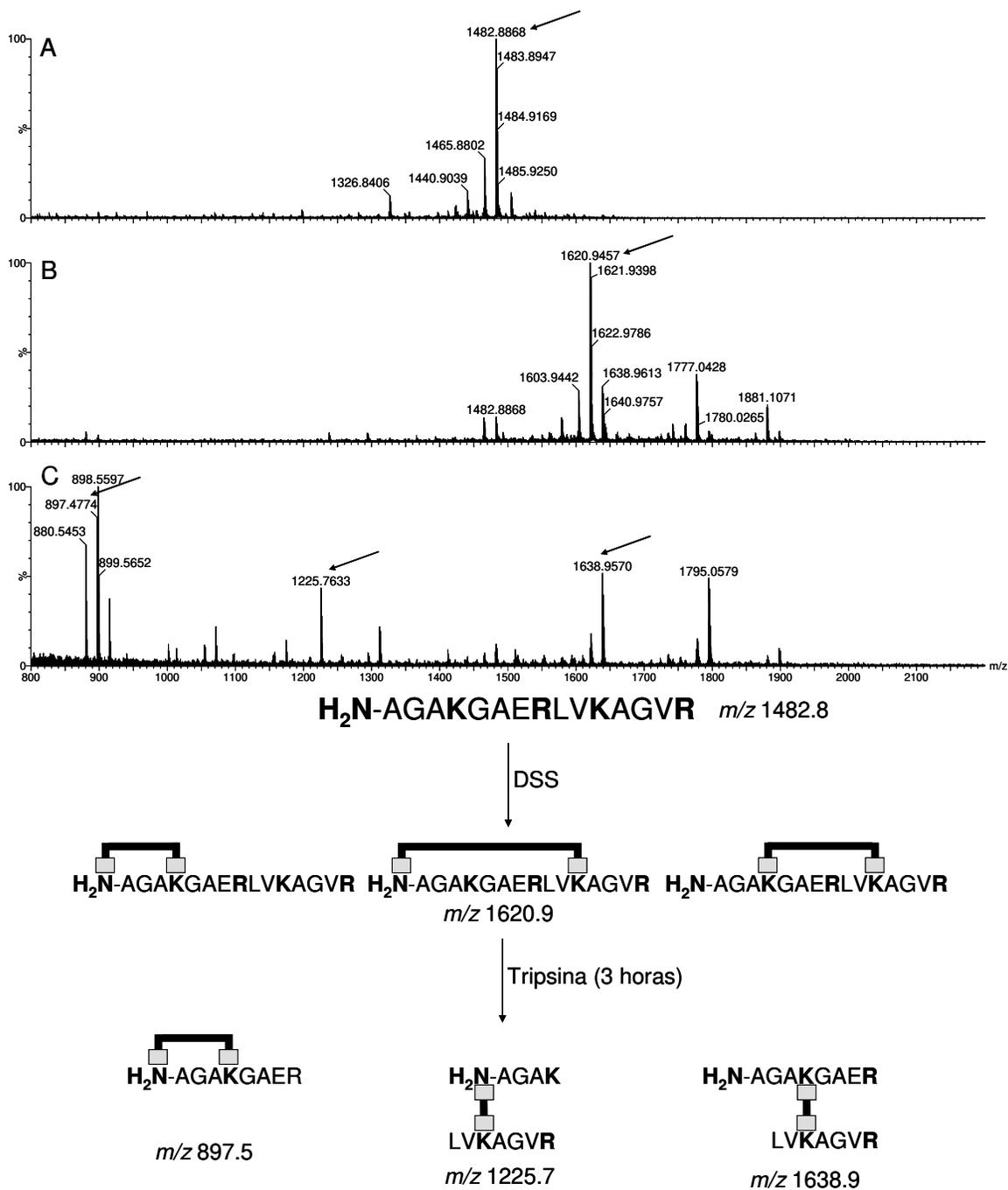
Tripsina (3 horas)

**NH<sub>2</sub>-GKWPR**  $m/z$  1561.9  
**NH<sub>2</sub>-EVKIHR**

**Figura 19 e Esquema 2** - Espectros de MALDI-MS do peptídeo P2 intacto (A), após a reação com DSS (B) e digestão com tripsina por 1 hora (C) e 3 horas (D). Os íons assinalados com uma seta correspondem às espécies de interesse.

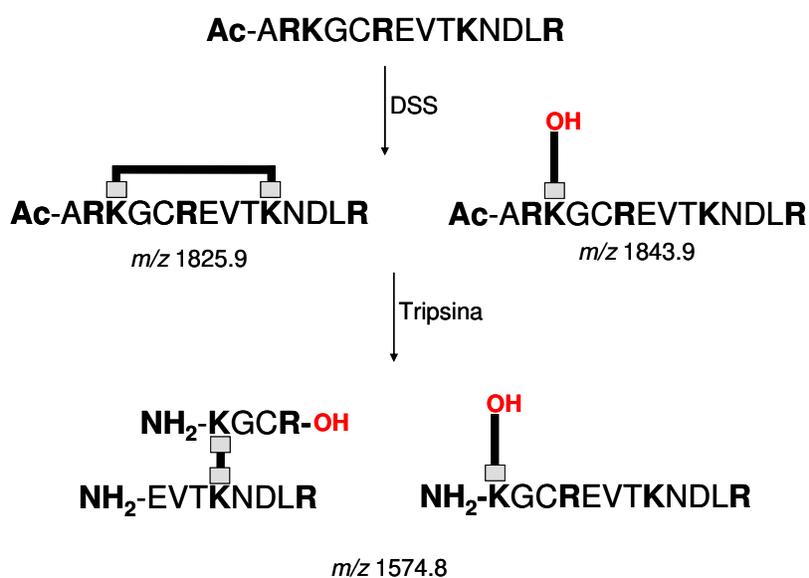


**Figura 20 e Esquema 3** - Espectros de MALDI-MS do peptídeo P3 intacto (A), após a reação com DSS (B) e digestão com tripsina por 1 hora (C) e 3 horas (D). Os íons assinalados com uma seta correspondem às espécies de interesse.



**Figura 21 e Esquema 4** - Espectros de MALDI-MS do peptídeo PX intacto (A), após a reação com DSS (B) e digestão com tripsina por 1 hora (C) e 3 horas (D). Os íons assinalados com uma seta correspondem às espécies de interesse.

Nesse ponto, é importante notar que existe a possibilidade da formação de espécies isóbaras que poderiam interferir no estudo de fragmentação das moléculas de interesse. No caso da formação das espécies intermoleculares, pode haver coincidência de massas para todos os peptídeos entre as seguintes espécies: i) peptídeos intermoleculares, cuja digestão ocorreu em todos os sítios trípticos não modificados; e ii) peptídeo *dead-end*, onde a modificação se encontra em uma das lisinas e na qual a digestão ocorreu somente no primeiro resíduo tríptico (R em todos os casos). Essa coincidência está exemplificada na Figura 22 para o P1.



**Figura 22** – Exemplo de possível coincidência de massas para o peptídeo P1 entre as espécies intermolecular e *dead-end*.

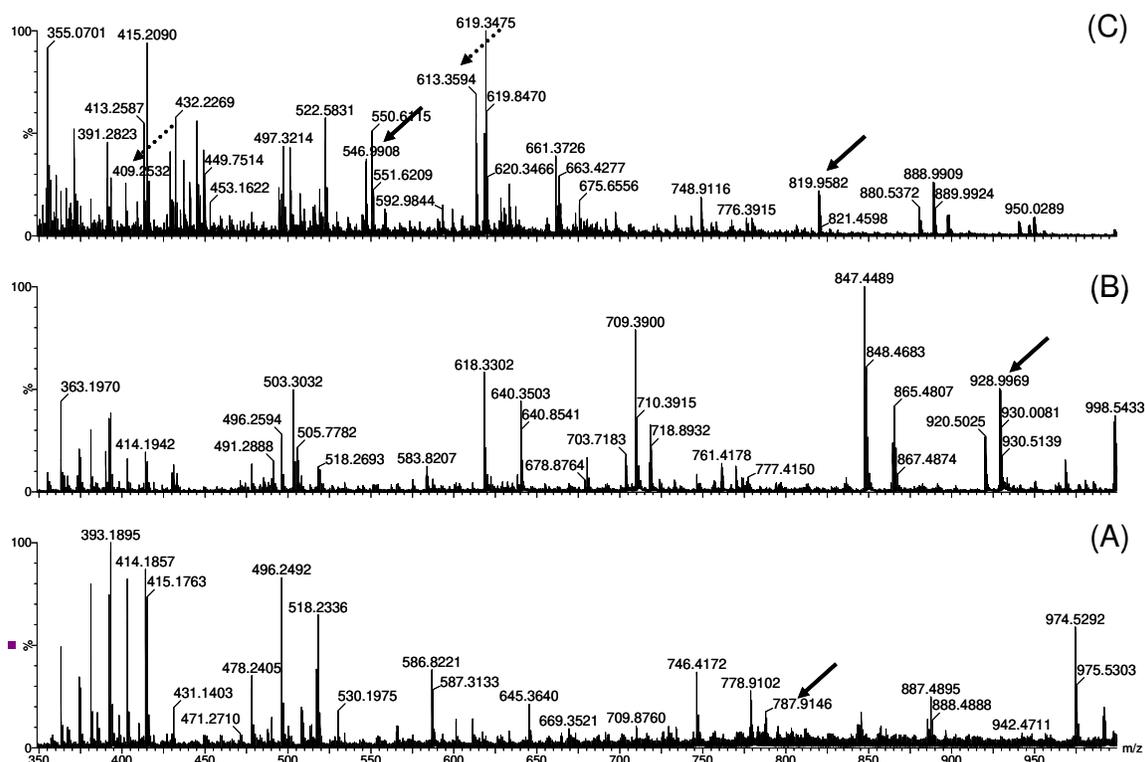
A princípio essa possibilidade não pode ser descartada, uma vez que em todos os espectros após a reação com DSS (Figura 18B a Figura 21B) é possível observar a formação de peptídeos do tipo *dead-end*, com intensidades relativas que variam de 10% (P1) a 100% (P2). Entretanto, existem vários

motivos para que se possa afirmar que essas coincidências não acontecem utilizando o procedimento experimental descrito anteriormente. No caso de P1 por exemplo, de modo a obter a espécie *dead-end* isóbara ao peptídeo intermolecular de interesse ( $m/z$  1574.8) seria necessário que o peptídeo *dead-end* inicial ( $m/z$  1843.9) mantivesse intacto à digestão com tripsina por 3 horas tendo como único sítio de clivagem digerido a primeira arginina (Figura 22). Qualquer outro produto de digestão em um sítio tríptico levaria a formação de uma espécie *dead-end* com diferente  $m/z$ . Além disso, o espectro de dissociação do peptídeo intermolecular após 16 horas de digestão é idêntico ao espectro de digestão após 3 horas, o que indica a ausência de fragmentos de outras espécies que não a de interesse.

#### **IV.1.3 Determinação dos Mecanismos de Fragmentação de Peptídeos Intermoleculares Contendo DSS**

Peptídeos contendo *cross-linking* intermoleculares foram reportados na literatura como os de mais difícil interpretação, visto que os fragmentos podem ser oriundos de uma ou ambas as cadeias, além da possibilidade da ruptura das ligações do próprio *cross-linker*.<sup>183</sup> Esses peptídeos são extremamente interessantes do ponto de vista estrutural, uma vez que conectam resíduos de lisinas distantes na estrutura primária da proteína ou de diferentes proteínas. Desse modo, essas informações são extremamente úteis tanto na determinação do tipo de enovelamento de uma proteína<sup>101</sup> bem como na elucidação da topologia de complexos protéicos<sup>102</sup>. Nesse caso, foram utilizados os peptídeos PX, P1 e P3, já que P2 apresenta, após a digestão, um

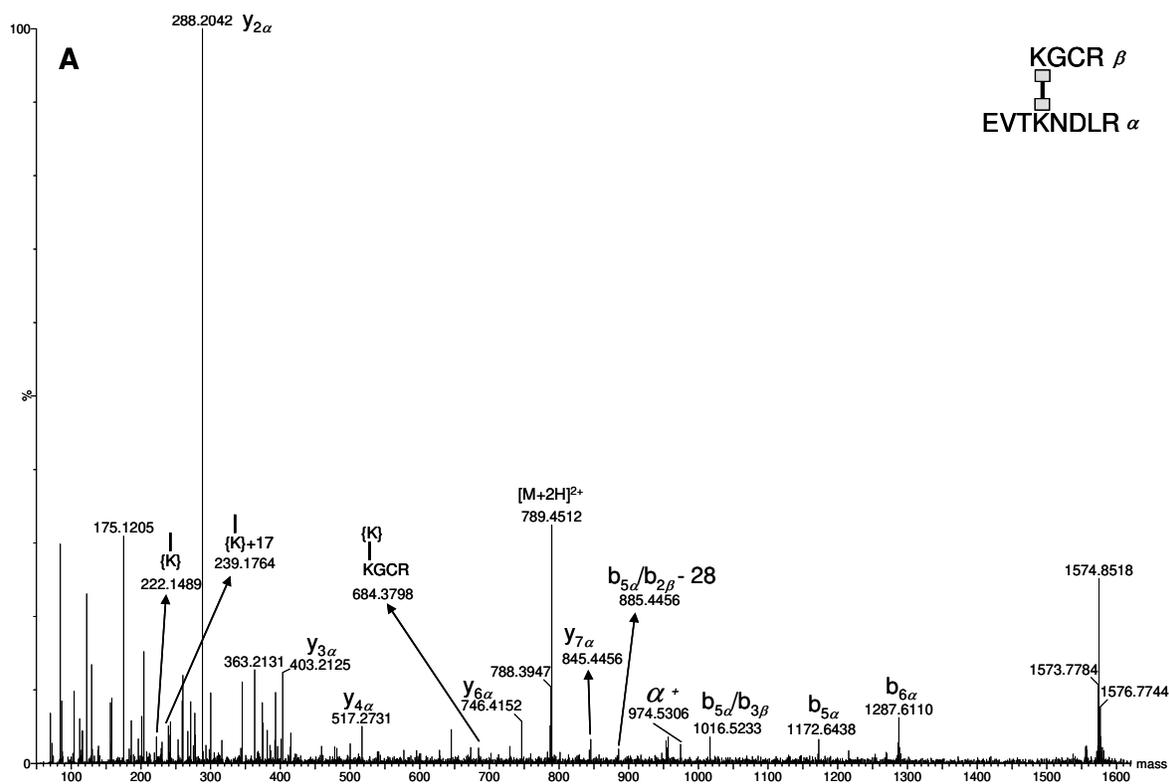
ion de mesma  $m/z$  que a espécie contendo *cross-linking*. P1 e P3 apresentam o N-terminal acetilado e conseqüentemente os sítios contendo DSS no peptídeo são as lisinas. No caso do PX, o N-terminal livre permite a formação de duas diferentes espécies intermoleculares após a digestão, que serão denominadas PX1 (peptídeo intermolecular conectado por duas K com  $m/z$  1638.9) e PX2 (peptídeo intermolecular conectado via K e N-terminal,  $m/z$  1225.7). PX2 representa um tipo de *cross-linking* entre o N-terminal da proteína e um resíduo de lisina; embora a fragmentação desse tipo tenha sido negligenciada na literatura, sua ocorrência em experimentos de MS3D deve ser muito comum, visto que em geral o N-terminal das proteínas é muito flexível. Os espectros de ESI-MS desses peptídeos estão mostrados na Figura 23 (Os espectros de MALDI-MS foram apresentados nas Figuras Figura 18 a Figura 21).

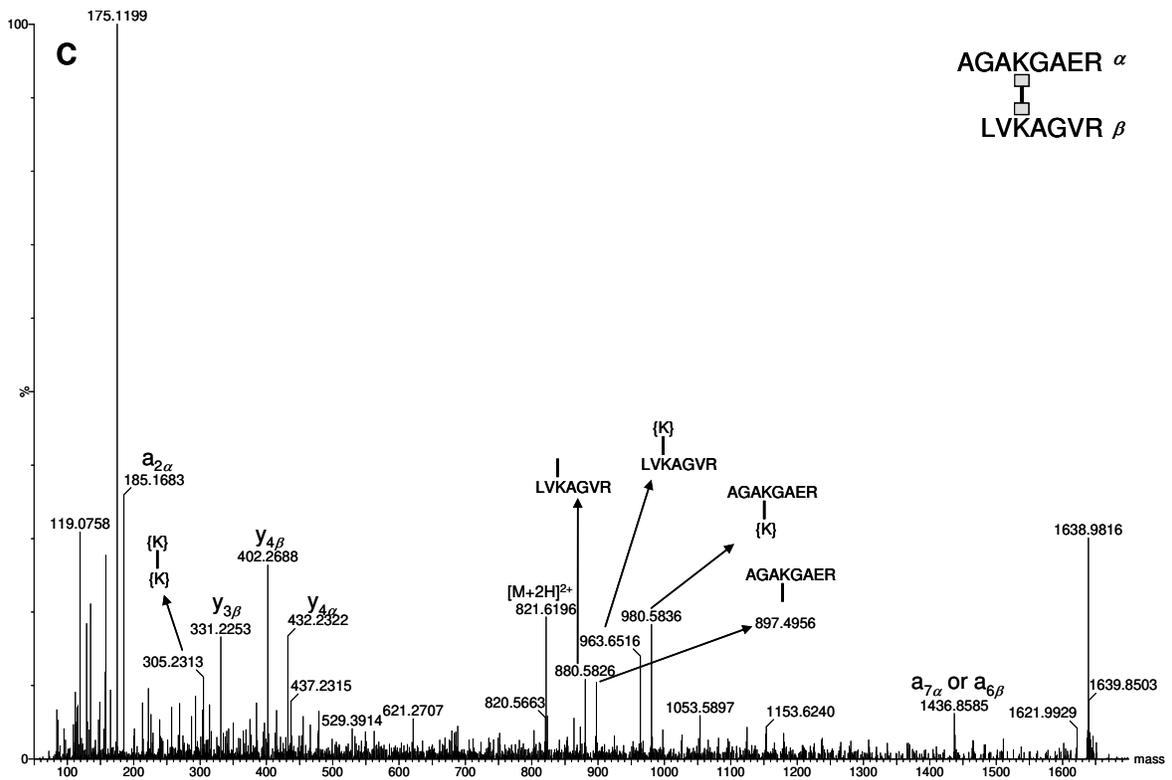
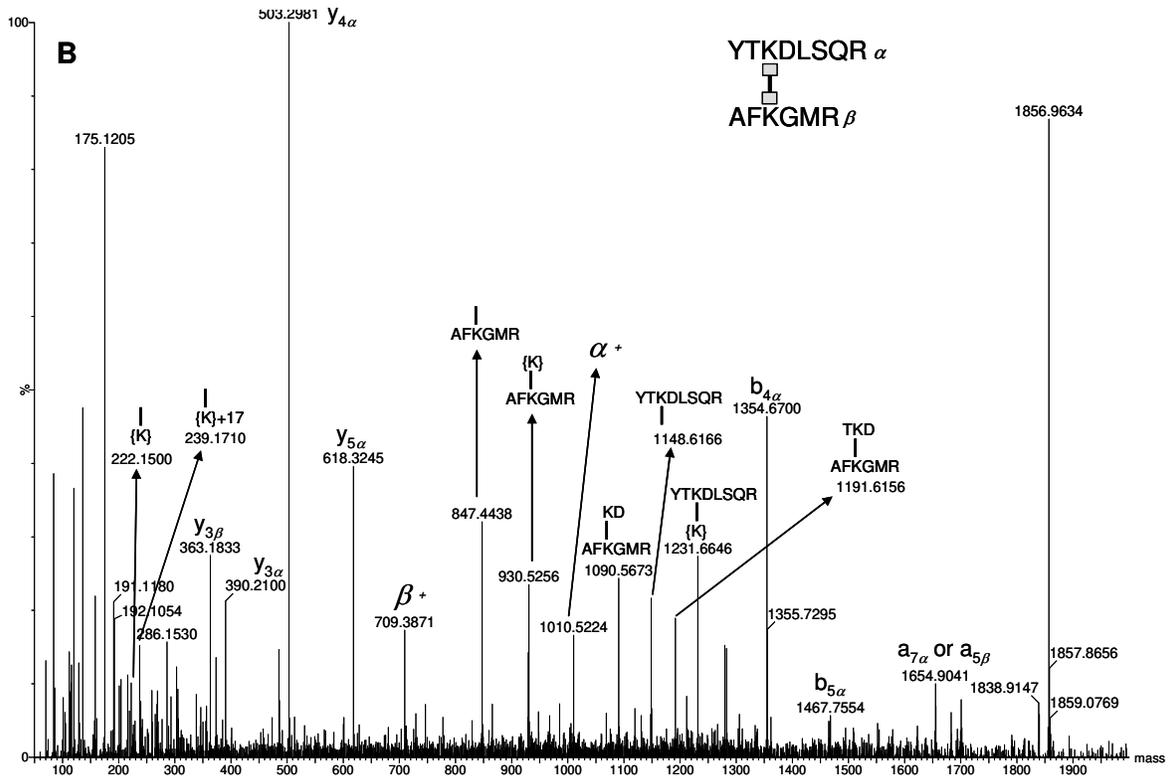


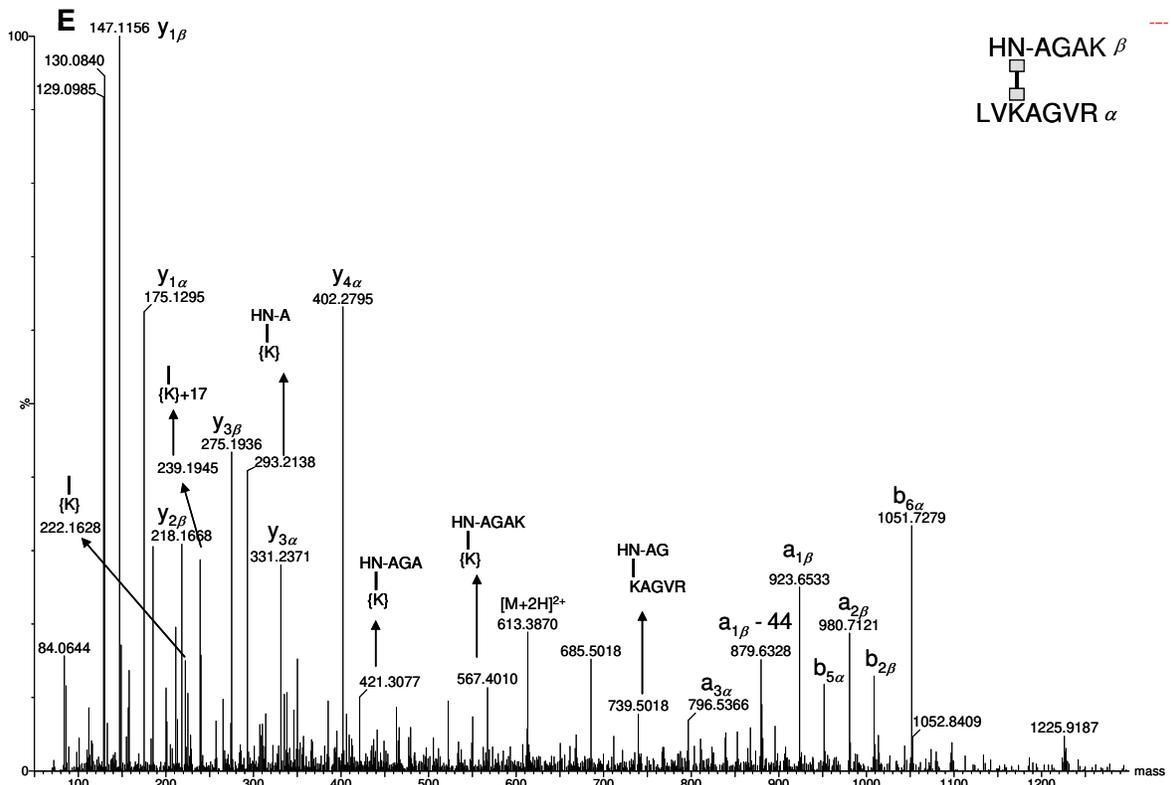
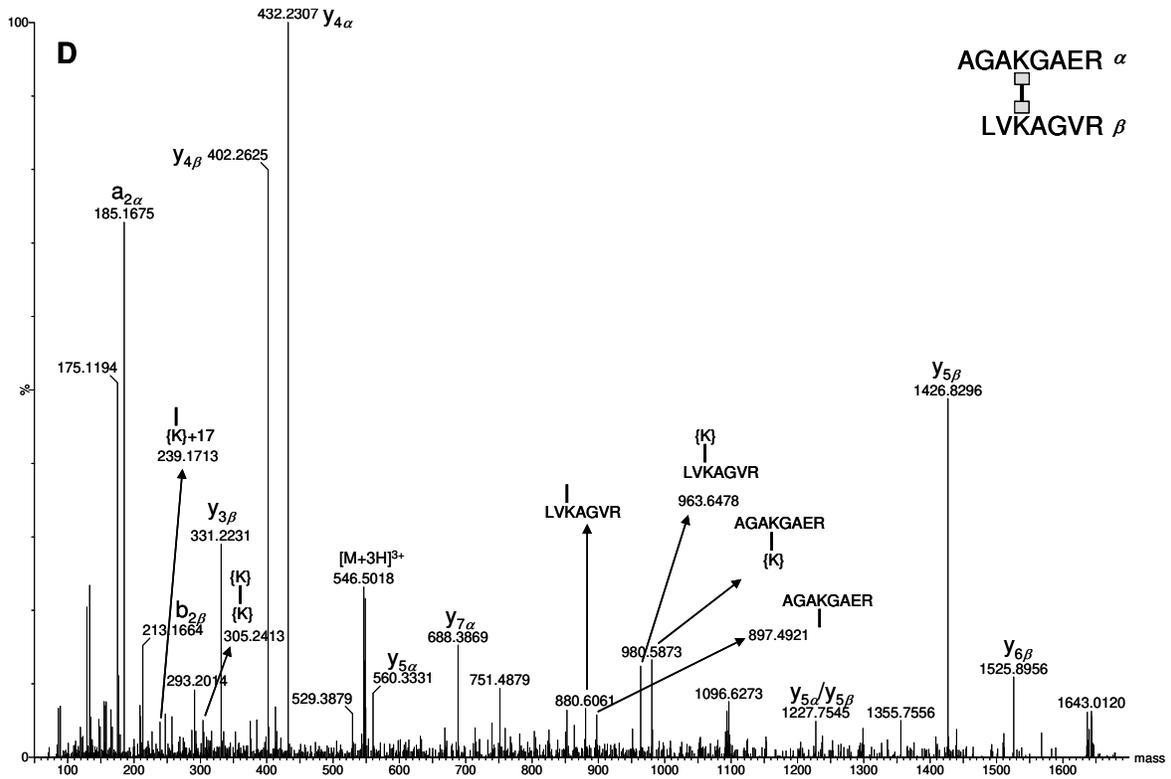
**Figura 23** – Espectros de ESI-MS dos produtos de reação dos peptídeos P1 (A), P3 (B) e PX (C) com DSS seguido da digestão com tripsina. Os íons indicados com as setas correspondem às espécies intermoleculares. No caso de PX, as setas contínuas indicam os íons referentes à PX1 e as setas pontilhadas apontam aos íons referentes à PX2.

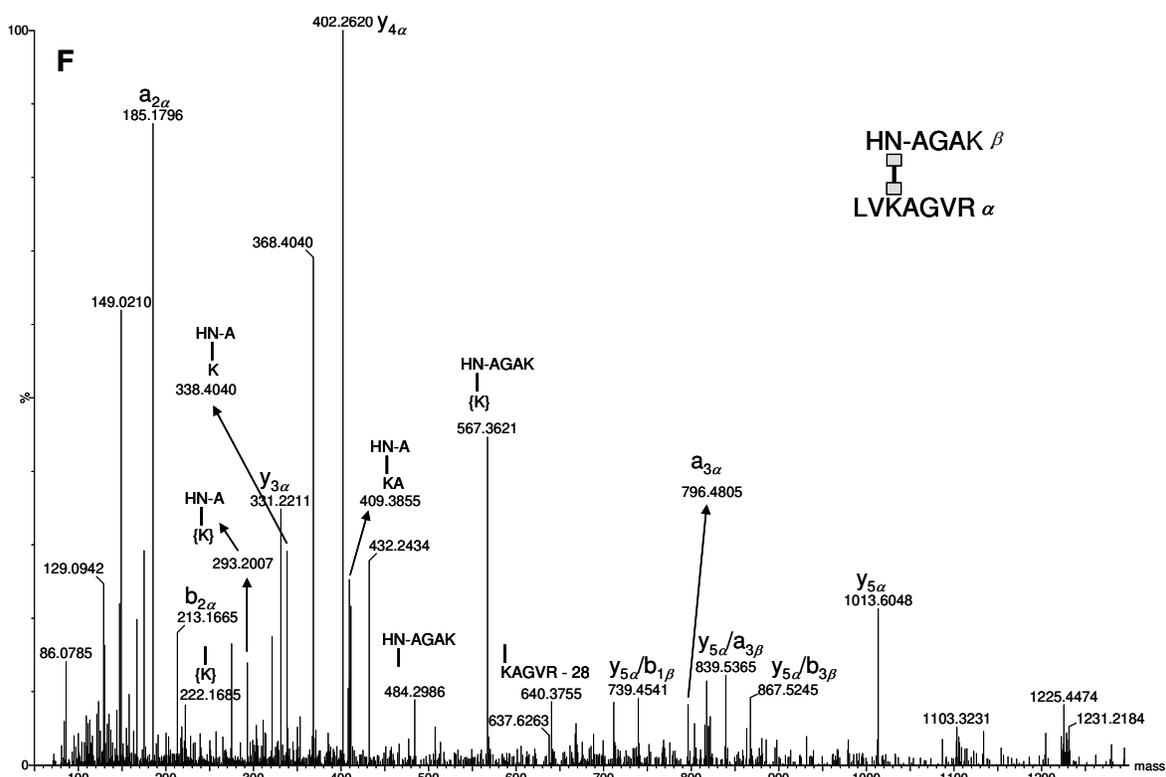
Os espectros de ESI-MS de P1 e P3 são caracterizados pela presença somente dos peptídeos no estado de carga 2+ ( $m/z$  787.9 para P1 e 928.9 para P3), enquanto que tanto PX1 quanto PX2 apresentam as espécies 2+ ( $m/z$  819.9 e 546.9) e 3+ ( $m/z$  613.4 e 409.2). A princípio, seria esperada a presença de espécies com maior número de cargas, em virtude de que cada peptídeo apresenta ao menos 4 sítios básicos e conseqüentemente passíveis de protonação (2 grupos N-terminal e 2 argininas no C-terminal). A presença de íons com menor número de cargas que o esperado pode ser resultado tanto da seqüência desses íons quanto da conformação desses em fase gasosa, impedindo maior número de prótons devido à repulsão eletrostática. De modo

a compreender a fundo os mecanismos de fragmentação de peptídeos intermoleculares, foram realizados experimentos de dissociação dessas espécies utilizando tanto MALDI quanto ESI como fonte de íons. Os espectros anotados de ESI-MS/MS dos peptídeos P1, P3, PX1 e PX2 estão mostrados na Figura 24.





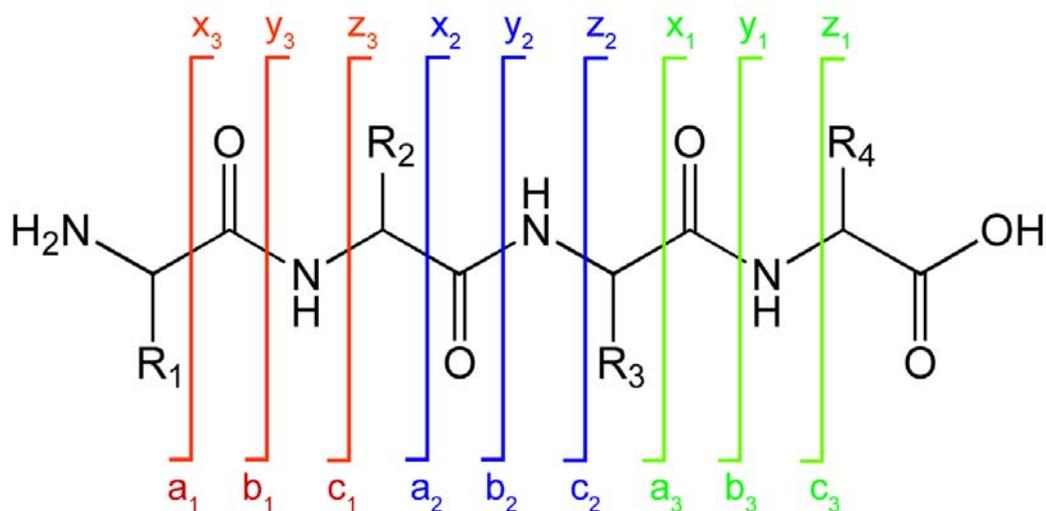




**Figura 24** - Espectros de ESI-MS/MS dos peptídeos intermoleculares de P1 (A), P3 (B), PX1 2+ (C), PX1 3+ (D), PX2 2+ (E) e PX2 3+ (F).

As anotações nos espectros de fragmentos comumente observados na dissociação de peptídeos lineares estão de acordo com a nomenclatura proposta por Roepstorff<sup>186</sup> e modificada por Biemann<sup>187</sup> (Figura 25). A grande maioria dos íons gerados é proveniente de clivagens da cadeia principal do peptídeo (clivagens na cadeia lateral são mais observadas em regimes energia de colisão muito alta). Nessa nomenclatura, íons *a*, *b* e *c* são aqueles que retêm a carga na porção N-terminal do peptídeo original, enquanto as espécies *x*, *y* e *z* apresentam a carga localizada no C-terminal. A diferenciação entre esses íons é dada pela ligação que se rompeu na dissociação: ligação C $\alpha$ -C (*a* e *x*), C-N (*b* e *y*) e N-C $\alpha$  (*c* e *z*). Íons não previstos na nomenclatura anterior (formados em virtude da reação de *cross-linking*) estão de acordo com o

proposto na referência 175, onde os índices  $\alpha$  e  $\beta$  indicam as cadeias de maior e menor massa respectivamente.



**Figura 25** – Nomenclatura utilizada para a estrutura de fragmentos observados na dissociação de peptídeos em fase gasosa.

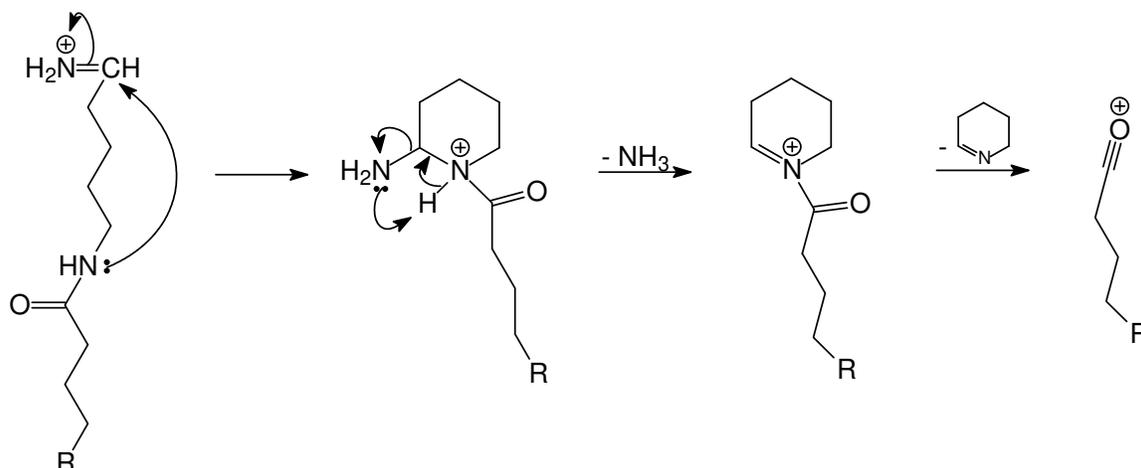
A análise dos espectros de ESI-MS/MS de P1, P3, PX1 e PX2 obtidos indicam a presença de fragmentos do tipo *b* e *y*, comumente encontrados na dissociação de peptídeos lineares. Em todos os casos é possível observar a formação de fragmentos nos quais uma das cadeias tem uma ligação peptídica clivada enquanto a outra cadeia se encontra intacta (por exemplo, no espectro de P1 o íon  $b_{6\alpha}$  indica a clivagem da ligação D-L, com a manutenção do *cross-linking* K-K, estando a cadeia  $\beta$  intacta); isso significa que uma das cadeias pode funcionar como se fosse uma modificação no aminoácido da outra cadeia. No entanto, é interessante notar que para os peptídeos conectados via lisinas (P1, P3 e PX1) a origem desses fragmentos (cadeia  $\alpha$  ou  $\beta$ ) depende de cada uma dessas espécies. No caso de P1, é possível notar somente a formação de fragmentos da cadeia  $\alpha$ , sendo possível observar os íons  $y_{7\alpha}$ ,  $y_{6\alpha}$ ,  $y_{4\alpha}$ ,  $y_{3\alpha}$ ,

$y_{2\alpha}$ ,  $b_{6\alpha}$  e  $b_{5\alpha}$ . Do mesmo modo, no caso de P3 é possível observar a formação preferencial de íons das séries  $y$  e  $b$  para a cadeia  $\alpha$  ( $y_{5\alpha}$ ,  $y_{4\alpha}$ ,  $y_{3\alpha}$ ,  $b_{5\alpha}$  e  $b_{4\alpha}$ ) com relação à cadeia  $\beta$  (somente  $y_{3\beta}$  é observado). Por outro lado, os espectros de dissociação das espécies  $2+$  e  $3+$  de PX1 apresentam uma distribuição uniforme de íons provenientes de ambas as cadeias: para a espécie  $2+$  é possível observar 2 íons provenientes de cada uma das cadeias ( $a_{2\alpha}$ ,  $y_{4\alpha}$ ,  $y_{3\beta}$  e  $y_{4\beta}$ ), enquanto que para o precursor com carga  $3+$  nota-se a formação de 4 fragmentos da cadeia  $\alpha$  e 5 da cadeia  $\beta$  ( $a_{2\alpha}$ ,  $y_{4\alpha}$ ,  $y_{5\alpha}$ ,  $y_{7\alpha}$ ,  $b_{2\beta}$ ,  $y_{3\beta}$ ,  $y_{4\beta}$ ,  $y_{5\beta}$  e  $y_{6\beta}$ ). Essa distribuição, também observada na fragmentação de espécies geradas por MALDI, está provavelmente relacionada à diferença de tamanho das cadeias conectadas: nos casos de P1 e P3, onde as cadeias  $\alpha$  (973.5 e 1009.5 Da respectivamente) apresentam massa muito maiores que as das cadeias  $\beta$  (462.2 e 708.3 Da respectivamente), a fragmentação é oriunda quase que exclusivamente da cadeia mais pesada. No caso de PX1 por outro lado, onde as cadeias apresentam massas muito semelhantes (758.4 e 741.4 Da), os fragmentos originam-se de ambas as cadeias. Uma possível explicação para esses fatos pode estar na localização do próton quando da dissociação desses peptídeos: como a fragmentação dessas espécies é dirigida pela carga, a maior cadeia tende a se fragmentar com maior frequência, visto que é nesta que o próton se localizará a maior parte do tempo. Isso explica o porquê da distribuição homogênea dos fragmentos quando as cadeias apresentam massas muito próximas (caso de PX1).

Outra característica observada nesses espectros que ainda não havia sido descrita na literatura está relacionada à formação do íon correspondente a uma das duas cadeias intactas (ruptura da ligação amida entre uma cadeia peptídica e o *cross-linker*). Como exemplo, pode-se observar a formação dos íons de  $m/z$  974.5 referente ao íon  $\alpha^+$  de P1, e 1010.5 e 709.3 referentes às espécies  $\alpha^+$

e  $\beta^+$  de P2 respectivamente. Esse tipo de fragmentação era esperado, visto que a ligação entre o *cross-linker* DSS e o N da lisina é do tipo amida, ligação essa mais comumente rompida em espectros de CID de peptídeos (levando à formação dos íons *b* e *y*).

Uma outra característica interessante nesses espectros é a formação preferencial de íons derivados do imônio da lisina modificada com DSS. Imônios são fragmentos gerados pela dissociação de duas ligações: ligação amida entre o N-terminal do resíduo e o C-terminal do resíduo anterior e a ligação entre o carbono  $\alpha$  e o carbono da carbonila dentro do mesmo resíduo. Embora a formação preferencial dessas espécies já houvesse sido descrita na literatura, suas estruturas e mecanismos de fragmentação permaneciam desconhecidos (esses íons foram denominados de  $K^{L\alpha/\beta} - 17$ , porém nenhuma estrutura foi associada a essa espécie).<sup>183</sup> É interessante notar que sempre que houve a formação do imônio da lisina contendo DSS, a outra cadeia conectada ao *cross-linker* permaneceu intacta. Essas espécies podem ser observadas nos espectros de P1 ( $m/z$  684.3), P3 ( $m/z$  930.5 e 1148.6), PX1 2+ e 3+ ( $m/z$  897.4 e 963.6), PX2 2+ e 3+ ( $m/z$  567.4). Além disso, uma característica interessante associada à formação dessas espécies está no fato de que quase sempre esses íons estão acompanhados de uma perda neutra de 83 Da, como no caso de P3 ( $m/z$  847.4 e 1148.6), PX1 2+ e 3+ ( $m/z$  880.5 e 897.4) e PX2 3+ ( $m/z$  484.2). A proposta do mecanismo de fragmentação para formação das espécies mencionadas está apresentada no Esquema 5.

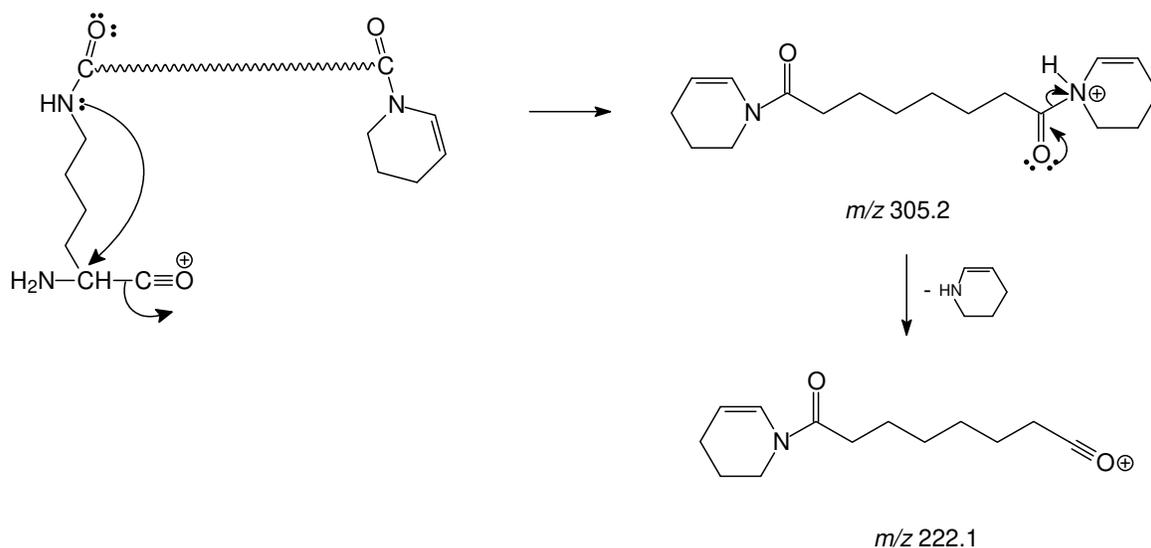


Esquema 5

Nessa proposta, ocorre um ataque do nitrogênio da lisina conectada ao DSS ao carbono  $\alpha$  do mesmo resíduo. Esse ataque é favorecido tanto pela eletrofilicidade desse carbono, uma vez que este está ligado a um átomo de nitrogênio positivo por uma ligação dupla, quanto pela formação de um anel estável de 6 membros. Esse íon pode facilmente perder um grupo amônia, levando a formação de um anel tetrahidropiridina (THP, indicado por {K} nos espectros de fragmentação); desse modo, essa é a estrutura proposta para o íon previamente relatado como  $K^{L\alpha/\beta} - 17$ . A perda neutra de 83 Da por sua vez é decorrente da clivagem da ligação C-N, com a saída do anel THP, levando a formação do íon acílio correspondente. É possível notar que a estrutura do anel THP é composta somente por átomos da lisina; desse modo, essa proposta de fragmentação é suportada pelo fato de que em espectros de fragmentação de peptídeos contendo esse resíduo, sempre que há a formação do imônio da lisina, ocorre também a formação de outro íon de  $m/z$  84, referente ao anel THP protonado. Como no caso do peptídeo modificado o nitrogênio do anel, após a perda do grupo amônio, não se encontra ligado a nenhum outro átomo de hidrogênio, a THP sai como uma espécie neutra,

deixando a carga no oxigênio. A perda neutra de THP no entanto é extremamente interessante, uma vez que ela é específica para a presença de DSS; desse modo, experimentos de Varredura de Perda Neutra podem ser utilizados para auxiliar na detecção de peptídeos modificados por esse *cross-linker*.

Ainda com relação a formação do anel THP a partir do rearranjo da cadeia lateral da lisina, é possível notar que caso a outra extremidade do *cross-linker* se ligue a outra lisina (e não ao N-terminal), também é possível que esta se rearranje do mesmo modo. Com isso, haveria a formação de um íon cuja estrutura seria composta por dois anéis THPs separados pela cadeia espaçadora do *cross-linker* (Esquema 6).



Esquema 6

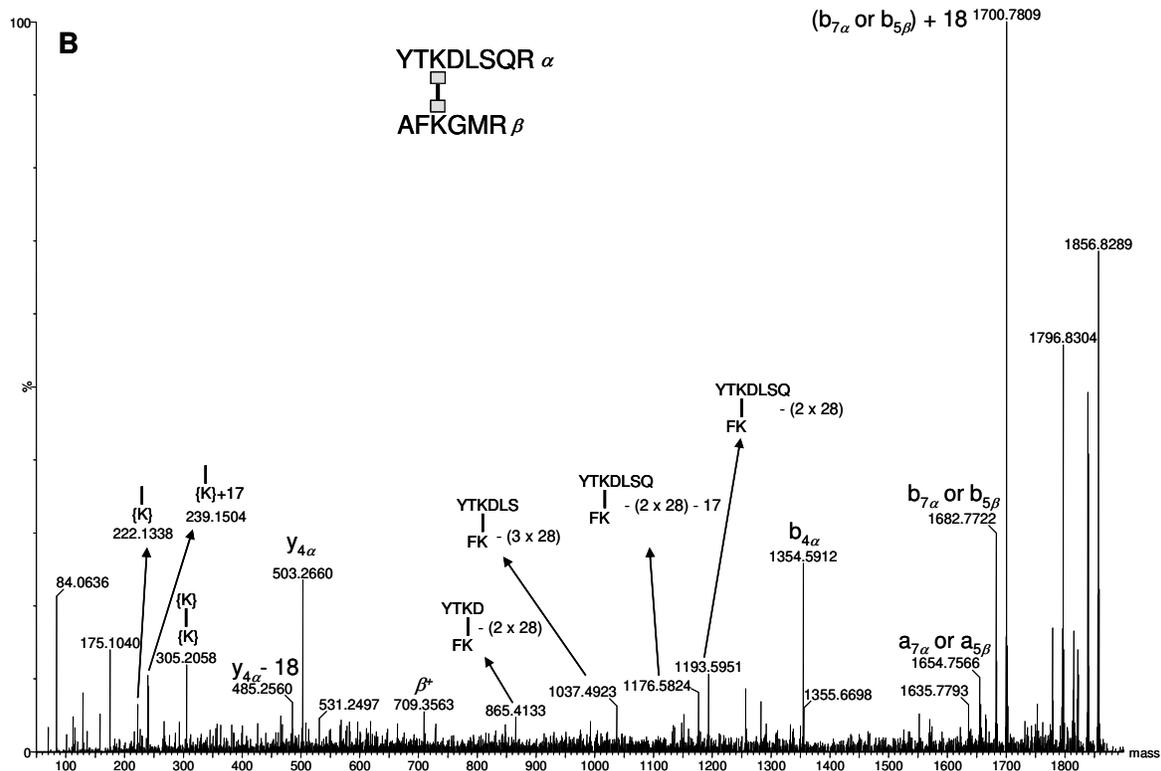
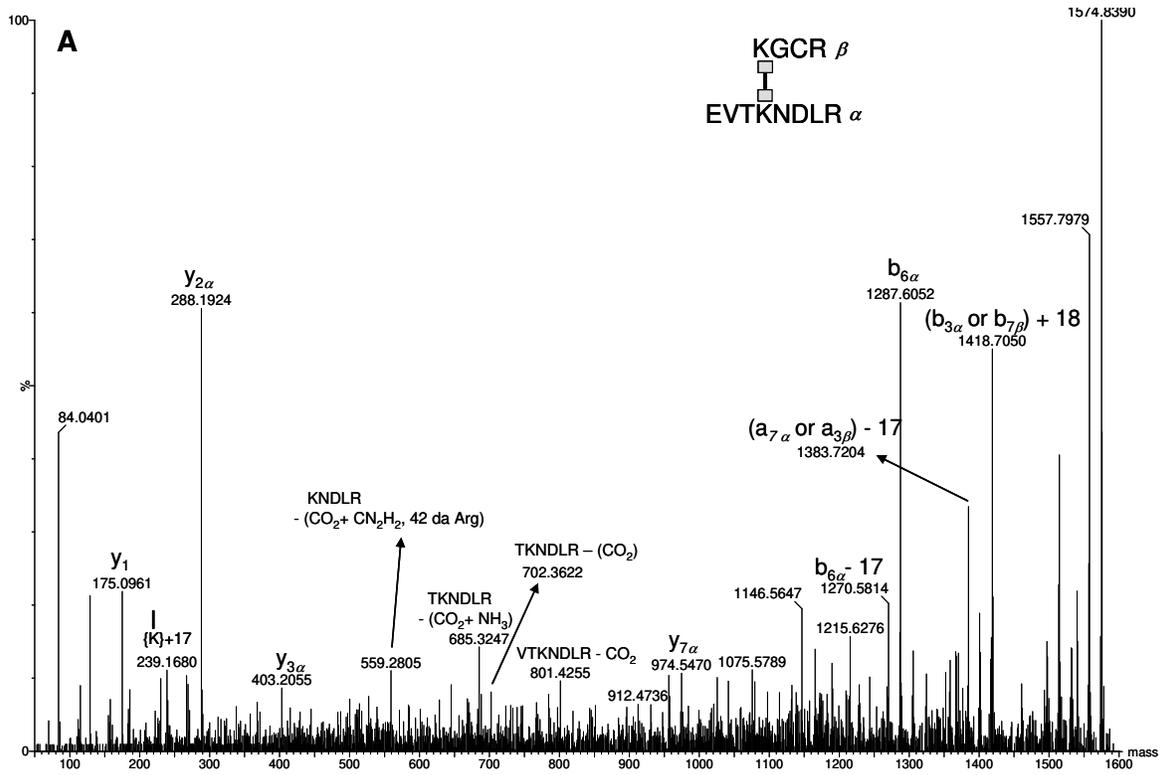
No caso do DSS, cuja cadeia espaçadora é composta por 6 grupos metílenos, a *m/z* desse íon seria 305.2. De fato, esse íon está presente em todos os espectros de fragmentação de peptídeos contendo DSS como *cross-linker* (indicado como {K}-{K} nos espectros), sendo sua intensidade dependente da

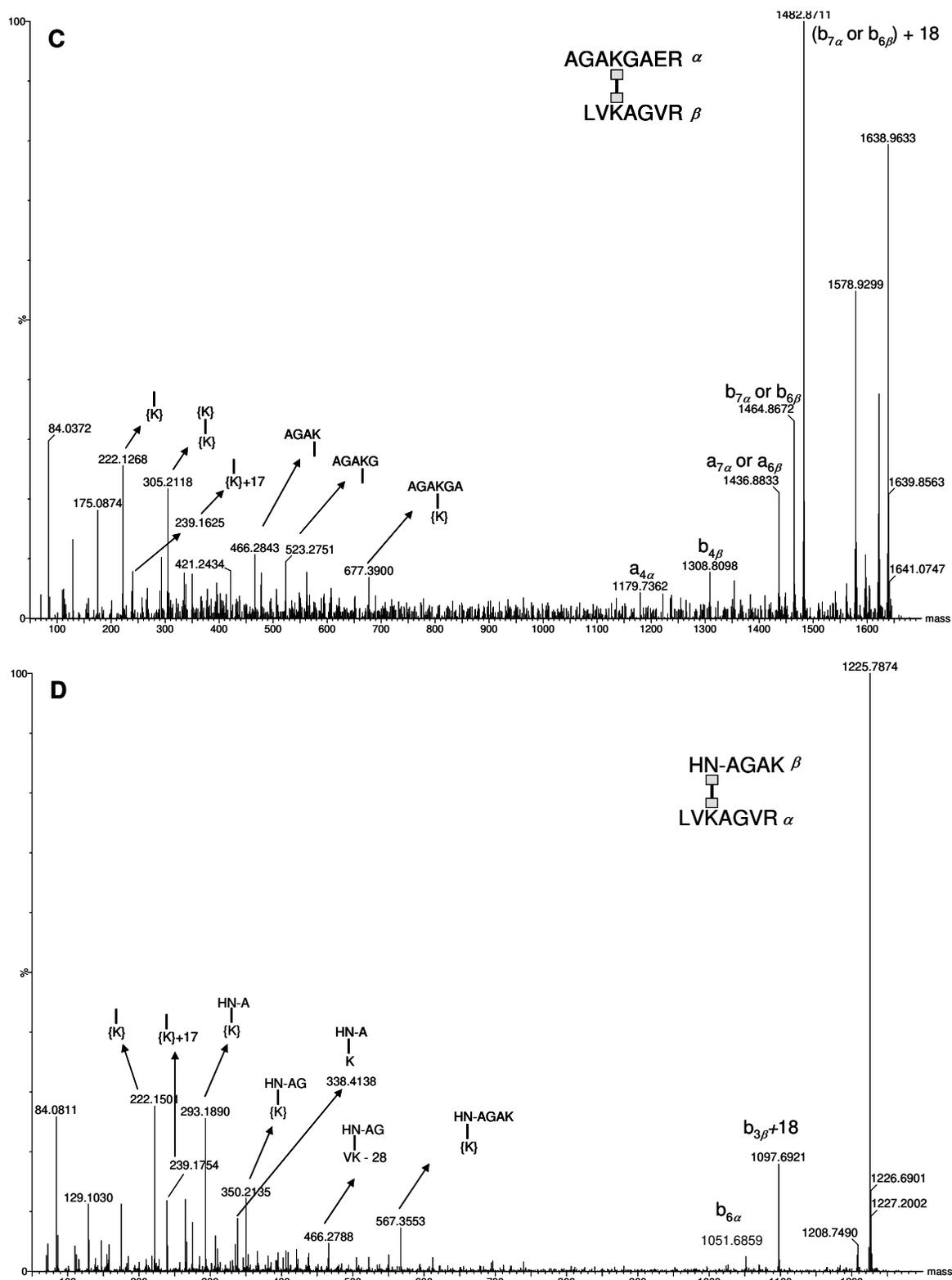
seqüência das cadeias peptídicas conectadas. Além disso, a formação dessa espécie independe da fonte de ionização utilizada para gerar o precursor (ESI ou MALDI, como mostrado posteriormente) e do tipo de *cross-linking* formado, intra ou intermolecular. Desse modo, esse íon pode ser utilizado como íon diagnóstico da presença de DSS, sendo este ainda mais interessante uma vez que sua massa nominal não coincide com nenhum fragmento tipo a, b ou y dos 20 aminoácidos naturais. Finalmente, como a estrutura desse íon é composta pelas cadeias laterais das duas lisinas mais a cadeia espaçadora do reagente utilizado, íons característicos para diferentes *cross-linkers* derivados de NHS (DSG, DST, DTSSP) podem ser obtidos pelo mesmo mecanismo. Acompanhando a presença do íon de  $m/z$  305.2, é possível observar nos espectros de fragmentação de todos os peptídeos contendo DSS os íons de  $m/z$  222.1 e 239.1. A formação do íon de  $m/z$  222.1 ocorre a partir da perda de um dos anéis THP do íon anterior, levando a formação de um íon acílio, conforme demonstrado no Esquema 6. O íon de  $m/z$  239.1 apresenta mesmo mecanismo de formação, porém sem a perda de amônia quando do rearranjo da cadeia lateral da lisina para formação do anel THP. Assim como no caso anterior, ambos os íons não apresentam coincidência de massas com nenhum fragmento  $a_2$ ,  $b_2$  ou  $y_2$  dos 20 aminoácidos naturais e conseqüentemente podem também vir a ser usados como íons marcadores para a presença de DSS.

Com relação à fragmentação das espécies intermoleculares conectadas via N-terminal (PX2 2+ e 3+), pode-se observar que, em geral, essas espécies apresentam características de dissociação semelhantes às das espécies conectadas por lisinas: formação extensa de íons do tipo  $y$  e  $b$  e formação dos íons marcadores da presença de DSS ( $m/z$  222 e 239). Uma das principais diferenças entre os espectros das espécies 2+ e 3+ consiste na maior energia de fragmentação requerida para completa dissociação do íon dicarregado

(45 eV x 25 eV). No caso da fragmentação da espécie 3+, foi possível observar o mesmo efeito da diferença de massa entre as cadeias  $\alpha$  e  $\beta$  descrito anteriormente: foram observados 6 íons provenientes da fragmentação da cadeia  $\alpha$  ( $y_{5\alpha}$ ,  $y_{4\alpha}$ ,  $y_{3\alpha}$ ,  $a_{3\alpha}$ ,  $a_{2\alpha}$  e  $b_{2\alpha}$ ) e nenhum da cadeia  $\beta$ . Entretanto, na fragmentação do mesmo peptídeo contendo 2 cargas, foi observado o efeito contrário: 4 íons provenientes da cadeia  $\alpha$  ( $a_{3\alpha}$ ,  $b_{5\alpha}$ ,  $b_{6\alpha}$  e  $y_{1\alpha}$ ) e 8 da cadeia  $\beta$  ( $b_{2\beta}$ ,  $a_{2\beta}$ ,  $a_{1\beta}$ ,  $a_{1\beta} - 44$ ,  $y_{1\beta}$ ,  $y_{2\beta}$  e  $y_{3\beta}$ ). Outra diferença observada nos espectros de fragmentação de PX2, independente da carga, foi a formação abundante de espécies resultantes de clivagens nas duas cadeias. Nesse caso, pôde-se observar espécies que continham ou não a estrutura do DSS ( $m/z$  739.5 na Figura 24E e  $m/z$  867.5 na Figura 24F, respectivamente). Como esses estudos foram realizados com somente um peptídeo desse tipo, não é possível determinar se esse é um efeito dependente da seqüência, carga ou tamanho dessa cadeia ou se esse é realmente um efeito comum de peptídeos conectados via N-terminal.

Os espectros anotados de MALDI-MS/MS dos peptídeos P1, P3, PX1 e PX2 estão mostrados na Figura 26. Assim como no caso anterior, as anotações nos espectros estão de acordo com o proposto na referência 175.





**Figura 26** - Espectros de MALDI-MS/MS dos peptídeos intermoleculares de P1 (A), P3 (B), PX1 (C) e PX2 (D).

Comparando-se com os espectros de ESI-MS/MS, os espectros de fragmentação de MALDI não são tão informativos quanto à seqüência de aminoácidos, visto que esses contêm um menor número de íons tipo  $y$  e  $b$ . Entretanto, nos três casos onde o *cross-linking* ocorre entre lisinas (P1, P3 e PX1), pode-se observar o mesmo efeito da diferença de massa entre as duas cadeias na extensão da fragmentação proveniente de  $\alpha$  e  $\beta$ . P1 ( $b_{6\alpha}$ ,  $b_{6\alpha} - 17$ ,  $y_{7\alpha}$ ,  $y_{3\alpha}$  e  $y_{2\alpha}$ ) e P3 ( $b_{4\alpha}$ ,  $y_{4\alpha}$  e  $y_{4\alpha} - 18$ ) apresentam fragmentos provenientes somente da cadeia  $\alpha$ , enquanto PX1 apresenta um íon oriundo de cada cadeia ( $a_{4\alpha}$  e  $b_{4\beta}$ ). Somente no caso de P3 foi possível observar a ruptura da ligação DSS-peptídeo, levando a formação da cadeia intacta  $\beta^+$ .

Diferentemente dos espectros de ESI, somente em um caso foi possível observar a formação do imônio da lisina modificada, com a formação e posterior perda do anel THP (PX2). Provavelmente essas espécies são formadas mas rapidamente dissociadas dando origem a outros fragmentos, em virtude do regime mais energético de fragmentação de íons monocarregados. Essa natureza mais energética pode ser evidenciada pela formação de um grande número de fragmentos internos, onde um número maior de ligações necessita ser rompido.

Finalmente, vale ressaltar que do mesmo modo que nos espectros de ESI, em todos os espectros de MALDI foi possível observar a formação dos íons marcadores da presença de DSS, de  $m/z$  222.1, 239.1 e 305.2. A presença dessas espécies nos espectros de fragmentação utilizando ambas as técnicas de ionização aumenta ainda mais o potencial da utilização destes como diagnósticos para a presença de DSS.

Em resumo, pode-se concluir que os espectros de fragmentação de MALDI e ESI mostraram-se complementares na completa caracterização do mecanismo de fragmentação de peptídeos contendo *cross-linking*

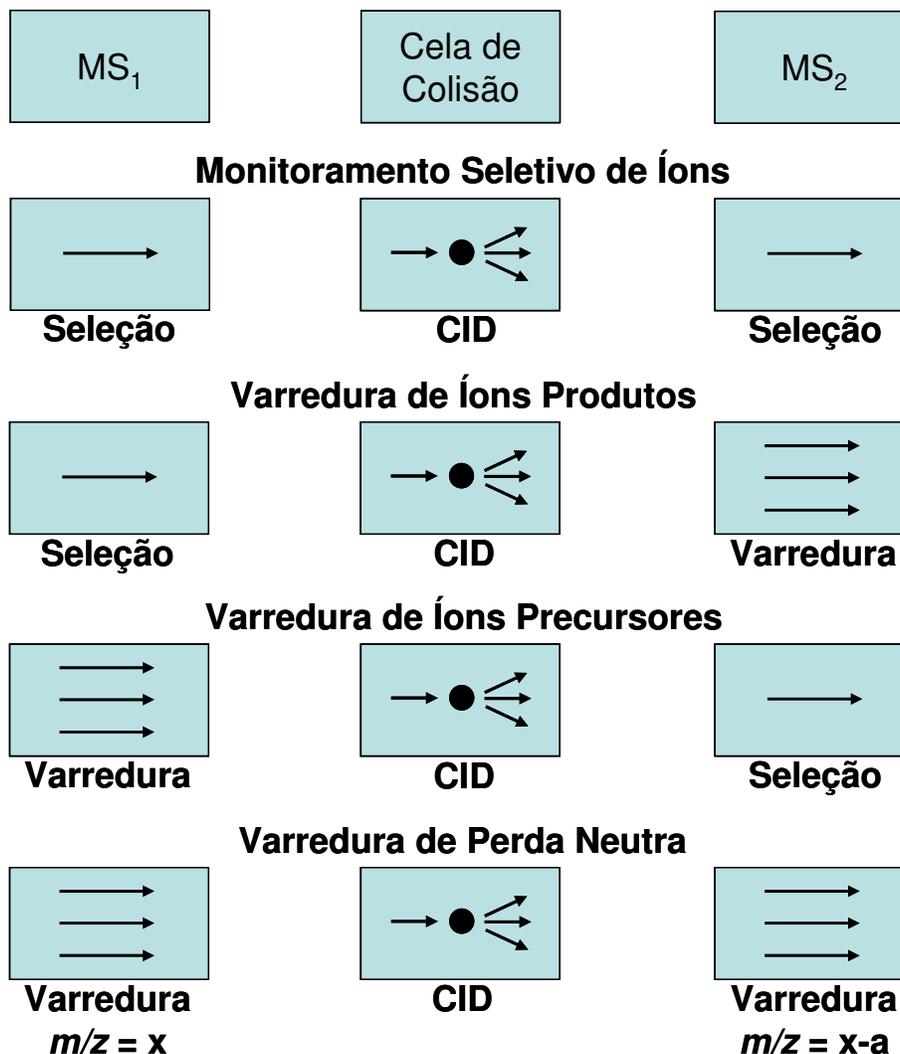
intermolecular. Apesar de algumas diferenças, como por exemplo a presença massiva de fragmentos internos em MALDI, os espectros apresentam várias características em comum, como por exemplo a presença majoritária de fragmentos tipo *b*. Além disso, diferentemente do caso de peptídeos intramoleculares,<sup>185</sup> os espectros de fragmentação de peptídeos intermoleculares são muito mais dependentes das seqüências, no que diz respeito à quantidade de íons *y* e *b*. Por outro lado, uma característica marcante em comum com todas as fontes de ionização e estados de carga é a quantidade maior de fragmentação da cadeia  $\alpha$  (maior cadeia) mantendo a cadeia  $\beta$  intacta. Além disso, quanto maior a diferença de massa entre as duas cadeias, mais pronunciado é esse efeito: no caso de P1, onde a cadeia  $\alpha$  é quase o dobro da  $\beta$  o espectro é dominado por fragmentos da primeira cadeia, enquanto no caso de PX1, onde as cadeias apresentam massas muito próximas, o número de fragmentos provenientes de ambas as cadeias é muito próximo. Isso é interessante uma vez que pode-se considerar a menor cadeia como uma modificação da outra, o que pode facilitar na futura implementação desses resultados em *softwares* que automatizem a identificação dessas espécies.

#### **IV.1.4 Identificação de Espécies Contendo *Cross-Linkers* por Varredura de Íon Precursor**

Dentro da metodologia de *cross-linkers* aliada a MS para o estudo de estrutura e interação de proteínas, uma das maiores dificuldades reside na detecção e identificação de peptídeos contendo cross-linking. Isso ocorre

devido ao fato de que peptídeos contendo *cross-linkers* estão em quantidades subestequiométricas na amostra quando comparados aos peptídeos não modificados. Como mostrado nas seções anteriores, os peptídeos modificados por esses reagentes podem comportar-se de modo semelhante a modificações pós-traducionais (PTM) no que diz respeito à fragmentação em fase gasosa. Desse modo, seria interessante a utilização de metodologias baseadas em MS tradicionalmente utilizadas para PTM de modo a auxiliar na detecção desses peptídeos. Isso tornaria os experimentos de *cross-linking* mais baratos, rápidos, simples e versáteis e com menor manipulação de amostras. Um dos métodos mais tradicionais de MS para detectar peptídeos modificados consiste na Varredura de Íons Precusores (PIS) em equipamentos tipo Triplo Quadrupolo. PIS é um experimento de MS/MS onde o primeiro analisador de  $m/z$  realiza a varredura de todos os íons gerados na fonte dentro de uma determinada faixa, sendo esses fragmentados simultaneamente na câmara de colisão. O segundo analisador por sua vez encontra-se fixo em uma ou mais  $m/z$ , permitindo somente a passagem de fragmentos de determinada  $m/z$  (Figura 27). Desse modo, esse tipo de análise é extremamente vantajoso quando se deseja buscar uma determinada classe de compostos em uma matriz complexa, desde que esses gerem um fragmento em comum. Vale ressaltar nesse ponto que PIS é o experimento oposto ao típico Varredura de Íons Produtos, onde o primeiro analisador seleciona um íon de determinada  $m/z$ , sendo somente esse fragmentado e todos os seus fragmentos medidos no segundo analisador (Figura 27). PIS em equipamentos tipo Triplo Quadrupolo (e mais recentemente em equipamentos do tipo *ion trap* linear) tem sido utilizado para determinação de sítios de fosforilação<sup>188,189,190</sup>, modificação por aldeídos<sup>191</sup>, identificação e diferenciação de oligossacarídeos ligados por N ou

O em glicoproteínas<sup>192</sup>, bem como na detecção de glicopeptídeos<sup>193</sup> e glicoproteínas<sup>194</sup>.



**Figura 27** – Representação esquemática dos diferentes tipos possíveis de experimentos de MS/MS.

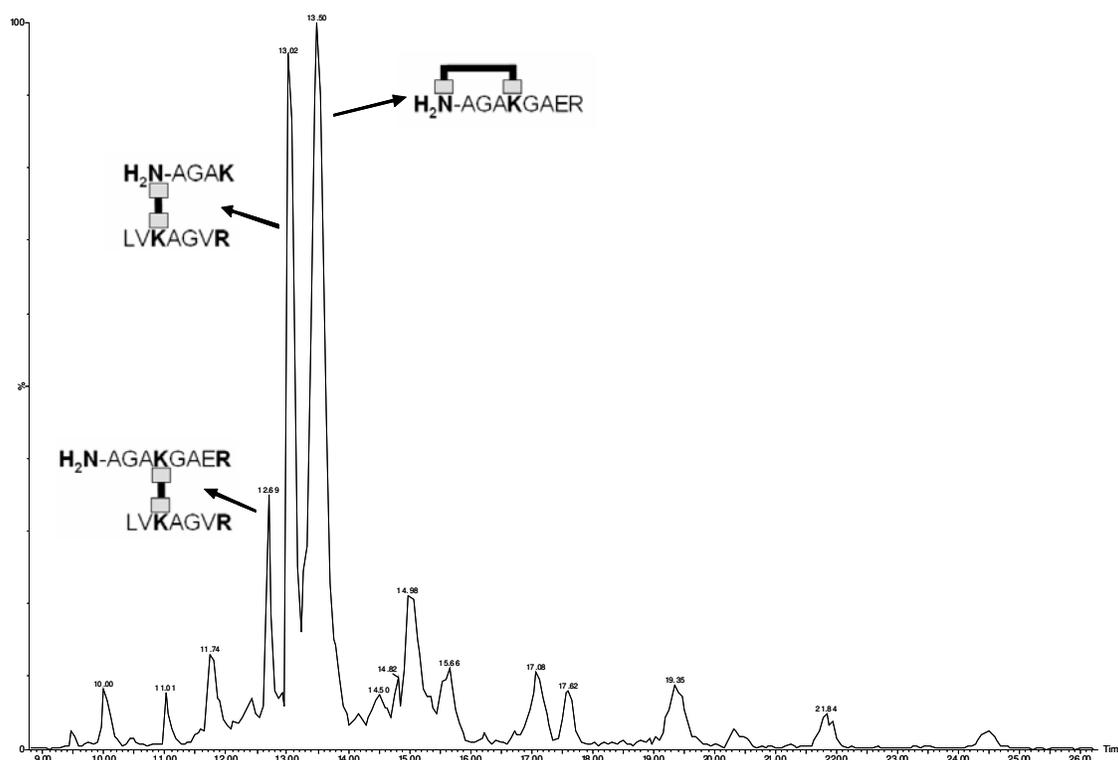
Em 2002, Bateman propôs a utilização do equipamento Q-Tof em um experimento similar ao PIS.<sup>195</sup> A princípio, o experimento de PIS não poderia ser realizado em equipamentos tipo Q-Tof, uma vez que analisadores Tof não tem capacidade de filtrar íons de determinada  $m/z$ . O método proposto

consistia em um experimento no qual o quadrupolo transmitia íons continuamente, enquanto o Tof adquiria espectros alternados com e sem fragmentação (isso é feito alterando-se a energia de colisão com o gás). Caso o fragmento de interesse fosse gerado na colisão, um espectro de varredura de íons produtos era adquirido para identificar o precursor que gerou o fragmento. Esse tipo de experimento traria todas as vantagens de equipamentos tipo Q-Tof quando comparados a triplo quadrupolos, como por exemplo a alta velocidade de aquisição de espectros aliada a alta sensibilidade e exatidão de massas. Recentemente, Niggeweg contribuiu para a melhoria do método tornando-o compatível com separação cromatográfica.<sup>196</sup> Desde então, PIS em equipamentos tipo Q-Tof tem sido utilizado para identificação e quantificação de fosfopeptídeos<sup>197,198</sup>, detecção e diferenciação de dimetilação em arginina<sup>199</sup> e na detecção de peptídeos glicosilados<sup>200</sup>.

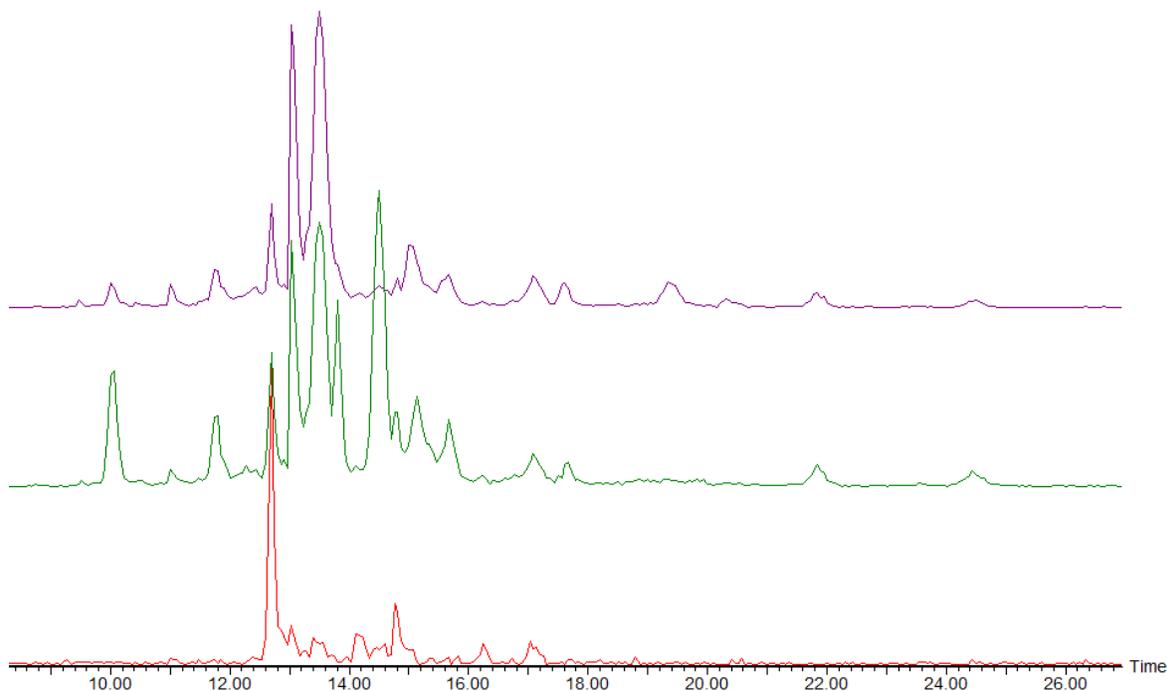
Como mencionado acima, para realização de experimentos de PIS, é necessário a presença de um íon marcador no espectro de Varredura de Íons Produtos, ou seja, um fragmento característico de uma determinada classe de compostos. De acordo com os estudos de fragmentação de peptídeos contendo DSS como *cross-linking* descritos anteriormente, foi possível identificar 3 íons característicos: 222.1494, 239.1759 (peptídeos contendo DSS) e 305.2229 (peptídeos contendo DSS conectados por duas lisinas).<sup>185</sup> Esses íons apresentam grande potencial para sua utilização em experimentos de PIS, uma vez que suas *m/z* não coincidem com nenhuma combinação de íons tipo a, b ou y (íons mais comuns na fragmentação de peptídeos por colisão) dos 20 aminoácidos naturais. Esse dado, aliado à alta exatidão de massas do Tof, pode conferir alta especificidade para esse tipo de análise.

O primeiro experimento realizado consistiu na análise da amostra PX. A Figura 28 apresenta o cromatograma da amostra PX com baixa energia de

colisão (sem fragmentação), sendo as espécies mais intensas (indicadas nos espectros) os peptídeos utilizados nos estudos de fragmentação. A Figura 29 apresenta os cromatogramas de alta energia de colisão de íons selecionados para os íons marcadores com alta exatidão ( $m/z$  222.1494, 239.1759 e  $305.2229 \pm 0.02$  Da).



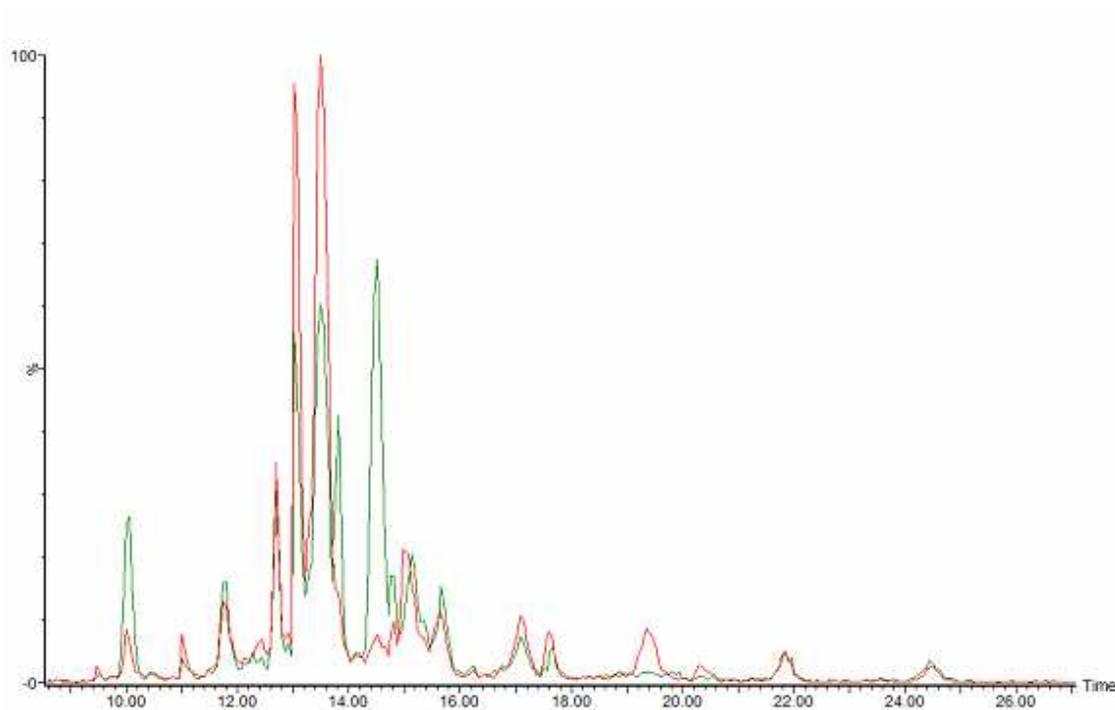
**Figura 28** - Cromatograma da amostra PX com baixa energia de colisão, indicando a presença dos íons de interesse.



**Figura 29** - Cromatograma de íons selecionados da amostra PX com alta energia de colisão, indicando a presença dos íons de  $m/z$  222.1494 (linha roxa), 239.1759 (linha verde) e 305.2229 (linha vermelha) (Erro de 0.02 Da).

É possível observar no cromatograma de baixa energia de colisão a presença de várias outras espécies além das indicadas na figura. Além disso, quando fragmentadas essas espécies dão origem aos fragmentos dos íons marcadores, o que indica que todos esses sinais são referentes a espécies contendo DSS em sua estrutura. Isso já era esperado, uma vez que após a reação com o *cross-linker* e digestão com tripsina, várias espécies foram formadas (como pode-se ver no espectro de MALDI-MS do produto da reação, Figura 21) além das mencionadas anteriormente: espécies contendo somente *dead-end* ou *dead-end* mais ligações intra ou intermolecular, produtos de digestão incompleta, dentre outros). Desse modo, é interessante notar que a potencialidade da técnica para detecção de peptídeos contendo *cross-linking*.

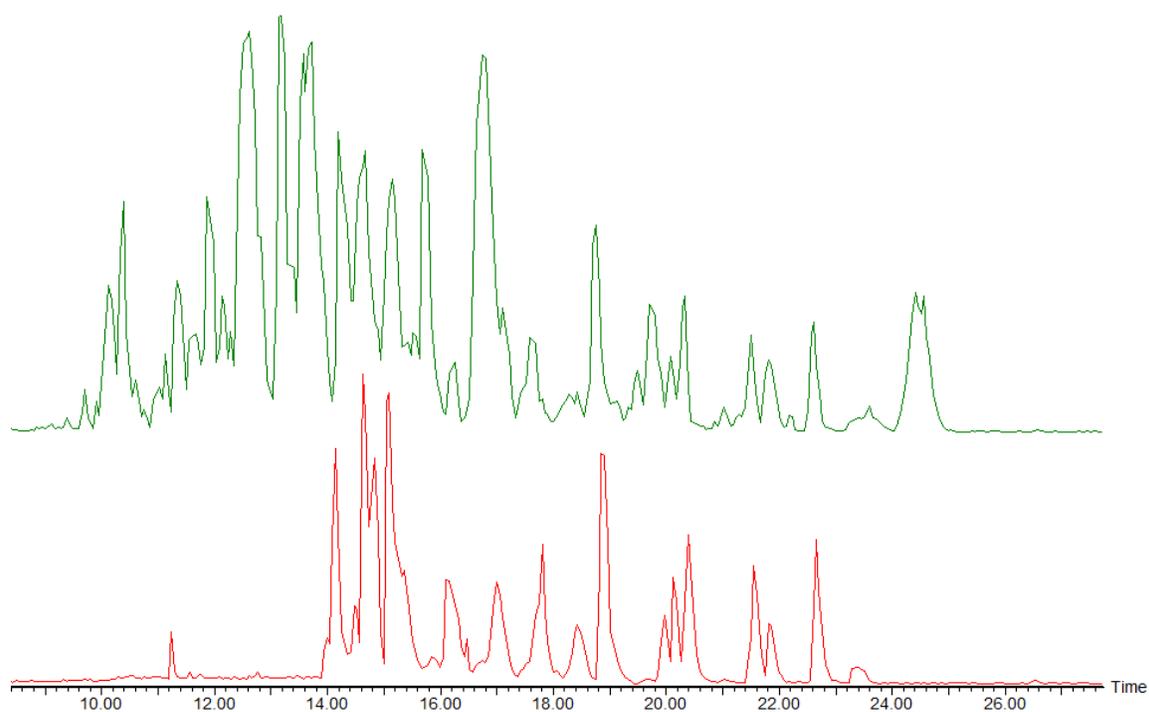
De acordo com os estudos de fragmentação apresentados anteriormente, ambos os íons de  $m/z$  222.1 e 239.2 seriam candidatos para utilização como marcadores, uma vez que ambos indicam a presença de peptídeos ligados por DSS (o íon 305.2 é mais específico, indicando peptídeos contendo DSS ligados por duas lisinas). Desse modo, seria interessante avaliar qual desses seria de melhor aplicabilidade para a técnica. A Figura 30 apresenta em detalhes o cromatograma de íons selecionados para os dois íons. Por essa figura, é possível observar que para as espécies mais intensas o íon 222.1 é um pouco mais intenso que o íon 239.2; entretanto, essa tendência é invertida para os sinais menos intensos. Isso indica que o íon 239.2 tende a ser mais sensível que o 222.1 com relação a presença de peptídeos contendo DSS. Entretanto, o único caso em que não há sobreposição corresponde ao íon com tempo de retenção de 19.35, sendo que esse deu origem somente ao marcador 222.1. Desse modo, a utilização de ambos os marcadores em conjunto é a melhor opção para detectar esses peptídeos modificados; a detecção de ambos os íons no mesmo experimento é possível em equipamentos Q-Tof sem perda de sensibilidade (diferentemente de equipamentos triplo quadrupolos, onde cada canal adicionado diminui a sensibilidade).



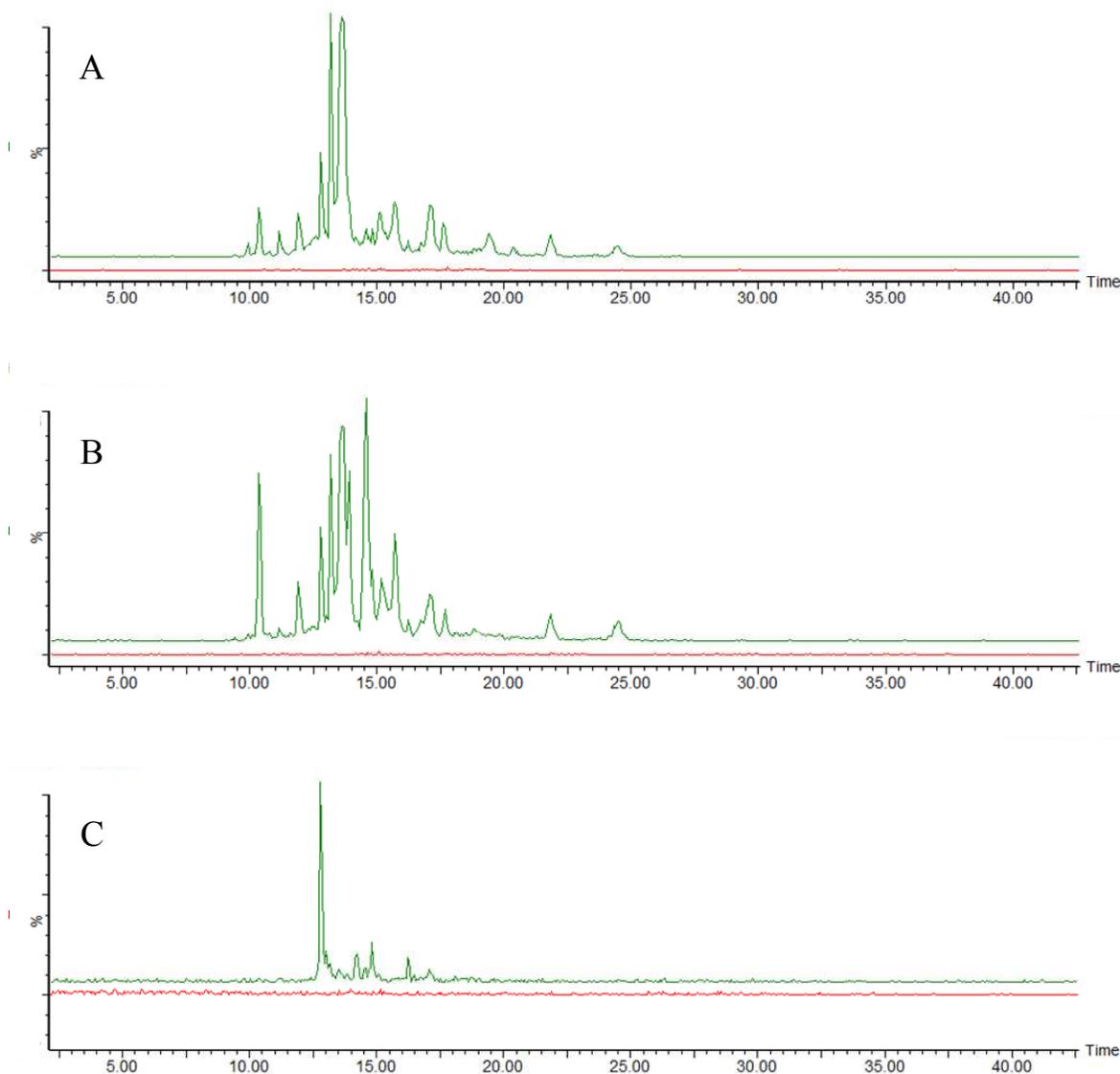
**Figura 30** - Cromatograma de íons selecionados com alta energia de colisão (linhas em vermelho e verde correspondem aos íons de  $m/z$  222.1 e 239.1).

Com o intuito de simular uma amostra que se assemelhasse a experimentos com proteínas, foi realizado outro experimento onde uma alíquota de PX foi adicionada a um digesto da proteína hemoglobina em quantidades estequiométricas. A Figura 31 apresenta o cromatograma de baixa energia de colisão das amostras controle (digesto de hemoglobina) e experimento (digesto de hemoglobina + PX). É clara a presença de um maior número de picos cromatográficos devido a amostra PX. A Figura 32 apresenta os cromatogramas de íons selecionados (íons marcadores de presença de DSS) de alta energia de colisão. Uma das principais características desses cromatogramas refere-se à especificidade dos íons marcadores, ou seja, com a alta exatidão de massas de um Tof, nenhum peptídeo da hemoglobina quando fragmentado gera um dos íons marcadores. Isso pode ser observado também

pela comparação dos cromatogramas de íons selecionados desse experimento quando comparado com o experimento realizado somente com o PX.



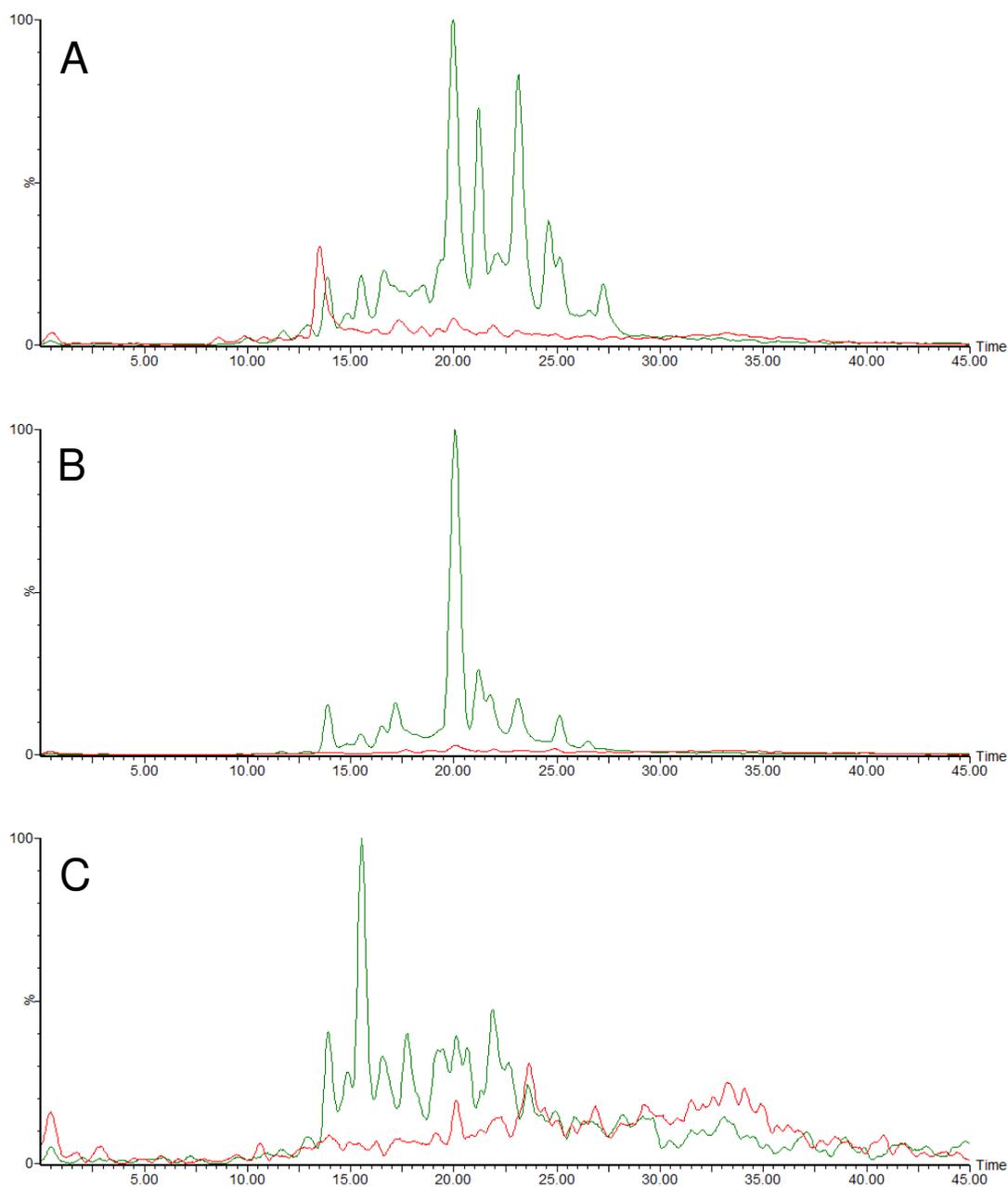
**Figura 31** - Cromatogramas com baixa energia de colisão para a hemoglobina digerida com (linha verde) e sem PX (linha vermelha).



**Figura 32** - Cromatogramas de íons selecionados com alta energia de colisão para as amostras contendo hemoglobina + PX (linha verde) e somente hemoglobina (linha vermelha), indicando a presença dos íons marcadores de  $m/z$  222.1 (A), 239.1 (B) e 305.2 (C).

O próximo passo na utilização de PIS como método de detecção de peptídeos contendo DSS foi realizar um experimento de *cross-linking* com proteína. Com isso, o alvo escolhido foi a proteína lisozima, uma proteína modelo pequena e com estrutura resolvida tanto por Cristalografia de Proteínas como por Ressonância Magnética Nuclear. Além disso, essa

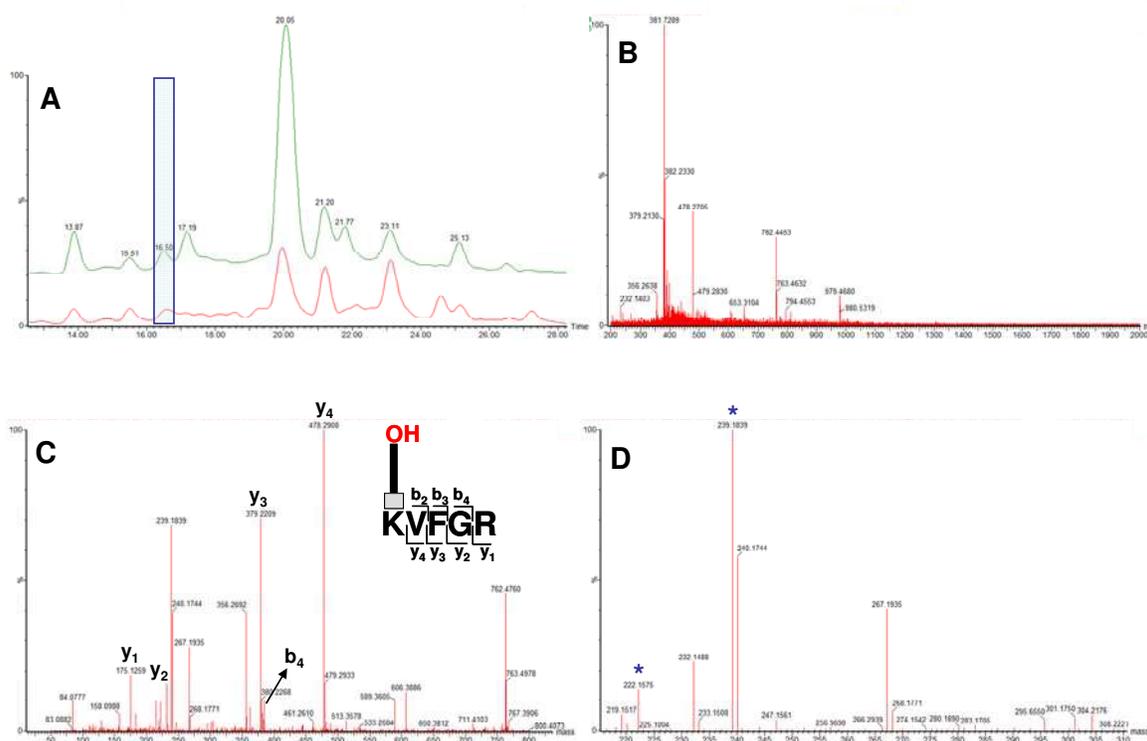
proteína tem um grande número de resíduos de lisina, o que a torna um bom alvo para a validação do método apresentado. A Figura 33 indica os cromatogramas referentes aos íons marcadores tanto do controle (lisozima digerida) quanto do experimento (lisozima + DSS digerida). Mais uma vez é notável a excelente especificidade dos íons marcadores, visto que são muito poucos os íons no controle que geram os marcadores. Essa figura indica também que a utilização dos 3 íons simultaneamente é extremamente interessante, uma vez que aumenta a especificidade da técnica sem a perda de sensibilidade (em equipamentos tipo Q-Tof).



**Figura 33** - Cromatograma com alta energia de colisão para o controle (vermelho) e experimento (verde), indicando a presença dos íons marcadores de  $m/z$  222.1 (A), 239.1 (B) e 305.2 (C).

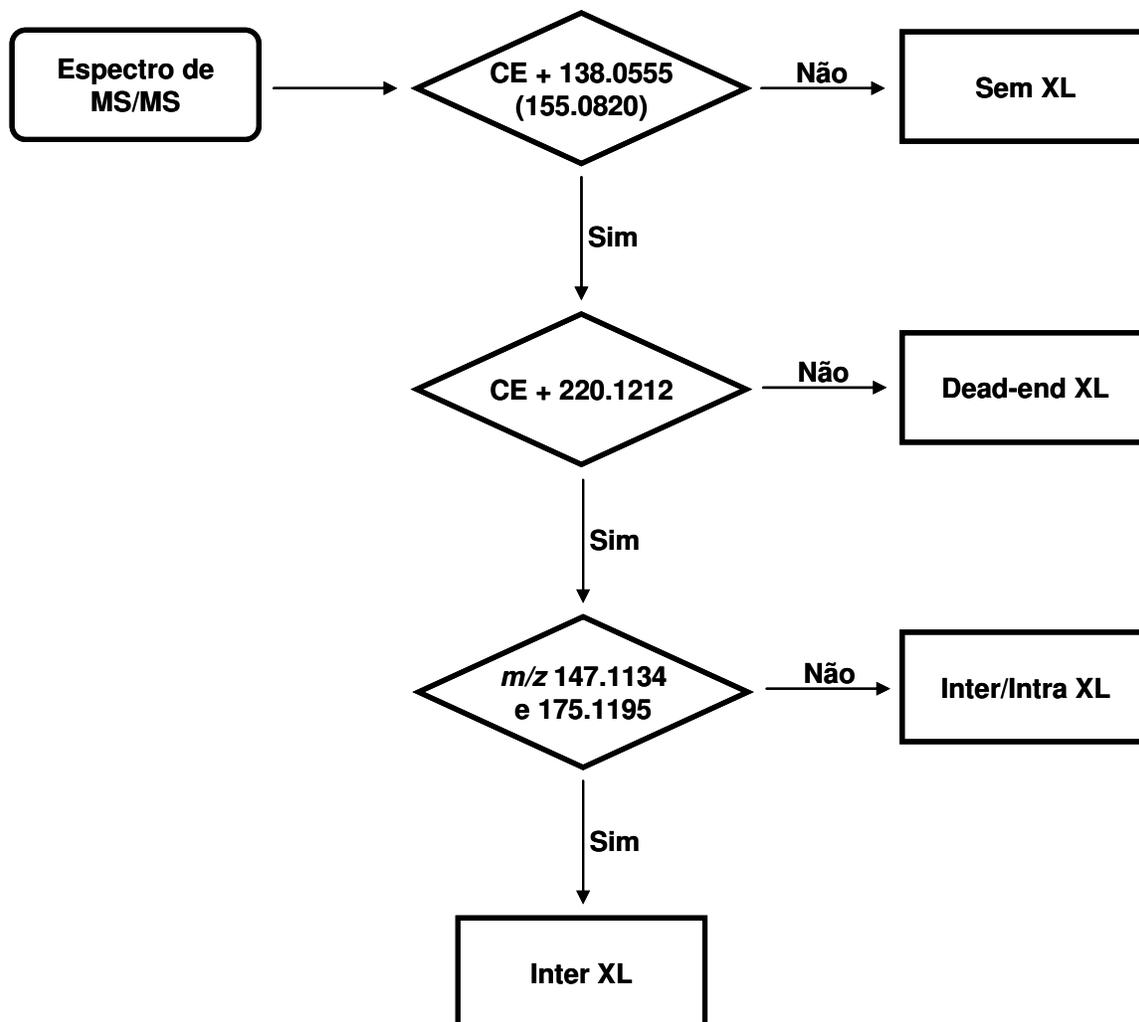
Para cada um dos picos cromatográficos acima nos quais havia a presença dos íons marcadores 222.1 e 239.1, os respectivos espectros de MS de baixa energia de colisão foram processados para identificação de peptídeos

modificados com DSS. Esses espectros foram deconvoluídos e as  $m/z$  dos possíveis precursores monocarregados foram carregados no *software* xBobcat<sup>176</sup>, que cria uma lista teórica de possíveis peptídeos modificados baseados somente em dados de MS. Os peptídeos atribuídos tiveram sua identificação manualmente confirmada por espectros de MS/MS. A Figura 34 apresenta um exemplo de identificação de um peptídeo do tipo *dead-end*; nesse caso, é interessante notar a presença dos íons 222.1 e 239.1 e a ausência do íon 305.2, fato esse esperado para peptídeos desse tipo.



**Figura 34** – (A) Cromatogramas de alta energia de colisão de íons selecionados para as  $m/z$  222.1494 e 239.1759 ( $\pm 0.02$  Da), em vermelho e verde respectivamente. O retângulo destaca o pico cromatográfico referente ao espectro de MS (baixa energia de colisão) apresentado em (B). (C) Espectro anotado de MS/MS do íon de  $m/z$  381.7, peptídeo N-terminal da lisozima modificado com DSS (*dead-end*). Em (D), ampliação do espectro (C), destacando a presença dos íons marcadores.

De modo a tornar a detecção e identificação de *cross-linking* um processo automatizado e, conseqüentemente, com velocidade compatível com as necessidades da área de Genômica Estrutural, seria interessante o desenvolvimento de ferramentas de bioinformática capazes de auxiliar esses processos. Embora existam alguns *softwares* disponíveis<sup>175,176</sup>, esses são extremamente dependentes da análise manual dos dados, além de considerarem como únicos critérios de seleção a *m/z* e seqüências de fragmentos independentes da presença do *cross-linker*. Desse modo, foi proposta uma rotina para implementação de um *software* que levasse em consideração os achados descritos até o momento acerca da fragmentação de espécies contendo *cross-linkers*. Primeiramente, foi proposta a utilização dos íons marcadores para localizar no cromatograma os tempos de retenção dos possíveis peptídeos contendo DSS. Conforme descrito na seção anterior, a formação desses íons diagnósticos é dependente somente da cadeia espaçadora do *cross-linker* e da cadeia lateral da lisina. Desse modo, qualquer *cross-linker* derivado de diésteres de NHS poderia formar essas espécies. Com isso, foi desenvolvido um fluxograma a ser implementado em um *software*, conforme apresentado na Figura 35.



**Figura 35** – Fluxograma de interpretação de dados obtidos de um experimento de PIS para detecção de peptídeos contendo *cross-linkers* derivados de diésteres de NHS (CE indica a massa da cadeia espaçadora do *cross-linker* utilizado no experimento).

A primeira etapa do fluxograma consiste na avaliação da presença das espécies análogas aos íons de  $m/z$  222.1 e 239.1. Conforme descrito anteriormente, esses íons são gerados quando da fragmentação de peptídeos contendo DSS, independente do tipo de *cross-linking* formado: intra ou intermolecular e *dead-end*. Caso nenhuma dessas espécies esteja presente no espectro de fragmentação, a presença do *cross-linker* pode ser descartada. Para calcular a massa do análogo do 222.1 para outro *cross-linker*, deve-se somar a

massa do anel THP com dois grupos carbonilas (que formam a ligação amida com o N da lisina) e da cadeia espaçadora (CE). O íon análogo ao 239.1 é calculado do mesmo modo, apenas adicionando-se a massa do grupo amônia. Para diferenciar as espécies conectadas via duas lisinas (intra ou intermolecular) das espécies *dead-end*, deve-se avaliar a presença do análogo ao 305.2: esse íon só é formado quando ocorre o rearranjo de duas cadeias laterais de lisinas, ou seja, somente quando duas lisinas presentes estão conectadas na estrutura do peptídeo. A massa desse íon pode ser calculada pela soma da massa de dois anéis THP, dois grupos carbonila e a CE do *cross-linker*. Finalmente, pode-se tentar diferenciar as espécies intra e intermoleculares pela presença de íons  $y_1$ . Esses íons são resultantes da fragmentação do resíduo C-terminal do peptídeo e comumente observados quando da fragmentação dessas moléculas por CID. Caso seja possível observar no espectro de dissociação as  $m/z$  147.1 e 175.1, referentes aos  $y_1$  de lisina e arginina respectivamente, isso indicaria a presença de dois resíduos diferentes na porção C-terminal. Esse fato só seria possível se dois peptídeos diferentes estivessem conectados, ou seja, se essa fosse uma espécie intermolecular. Por outro lado, a presença de somente um desses íons seria inconclusiva, uma vez que poderia indicar tanto a presença de somente um C-terminal (e conseqüentemente uma espécie intramolecular) quanto dois C-terminais idênticos (peptídeo intermolecular). Após essa análise preliminar, somente os espectro de dissociação dos íons de interesse seriam analisados para verificar outras características que permitissem a identificação inequívoca dos sítios de ligação ao *cross-linker* utilizado.

### IV.1.5 Conclusões

A elucidação dos mecanismos de fragmentação de peptídeos contendo *cross-linkers* é um passo fundamental para a utilização desses reagentes como “régua molecular” na determinação de estruturas superiores de proteínas. Não só esse conhecimento permite a identificação inequívoca dos resíduos que estão espacialmente próximos, mas também abre caminho para o desenvolvimento de ferramentas de bioinformática que permitam automatizar esse processo. Os resultados apresentados mostram que a dissociação de peptídeos contendo *cross-linking* intermoleculares gera íons tipo *b* e *y*, sendo que ocorre formação preferencial de íons provenientes da cadeia  $\alpha$ .

Essas espécies apresentam ainda algumas outras clivagens específicas, como por exemplo formação preferencial do imônio da lisina modificada e clivagem da ligação peptídeo-reagente. A partir da formação desse imônio, pode ocorrer o rearranjo da cadeia lateral dessa lisina, levando a formação de um anel tetrahidropiridina. Caso a outra extremidade do reagente esteja conectada a uma lisina (e não ao N-terminal), esse rearranjo também pode ocorrer na outra extremidade, levando à formação de um íon de  $m/z$  305.2. Esse íon foi observado em todos os espectros de MS/MS de peptídeos contendo DSS conectados via lisinas e conseqüentemente pode ser utilizado como íon diagnóstico. Além disso, fragmentação subsequente deste pode levar a formação dos íons de  $m/z$  222.1 e 239.1, que também foram observados em todos os espectros de fragmentação de espécies contendo DSS. A formação destes íons característicos permitiu a realização de experimentos de PIS para detecção de peptídeos contendo DSS. Esses experimentos demonstraram grande seletividade desses íons, principalmente quando esses foram utilizados

conjuntamente. Isso se deve ao fato de que nenhuma dessas  $m/z$  coincidam, nominalmente, com a  $m/z$  de nenhuma combinação de íons a, b ou y dos 20 aminoácidos. Esses achados permitiram criar uma rotina a ser implementada em um *software* capaz de automatizar os processos de detecção e identificação de peptídeos contendo *cross-linkers*, o que futuramente contribuirá significativamente na diminuição do tempo necessário para a interpretação dos dados provenientes de experimentos de MS3D.

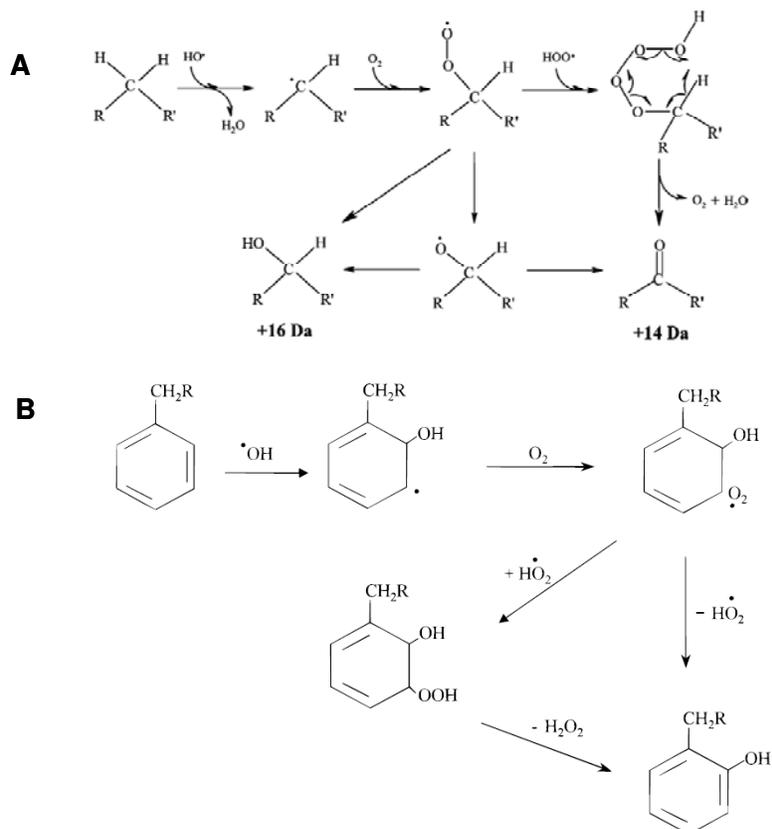
## IV.2 FOOTRPINTING

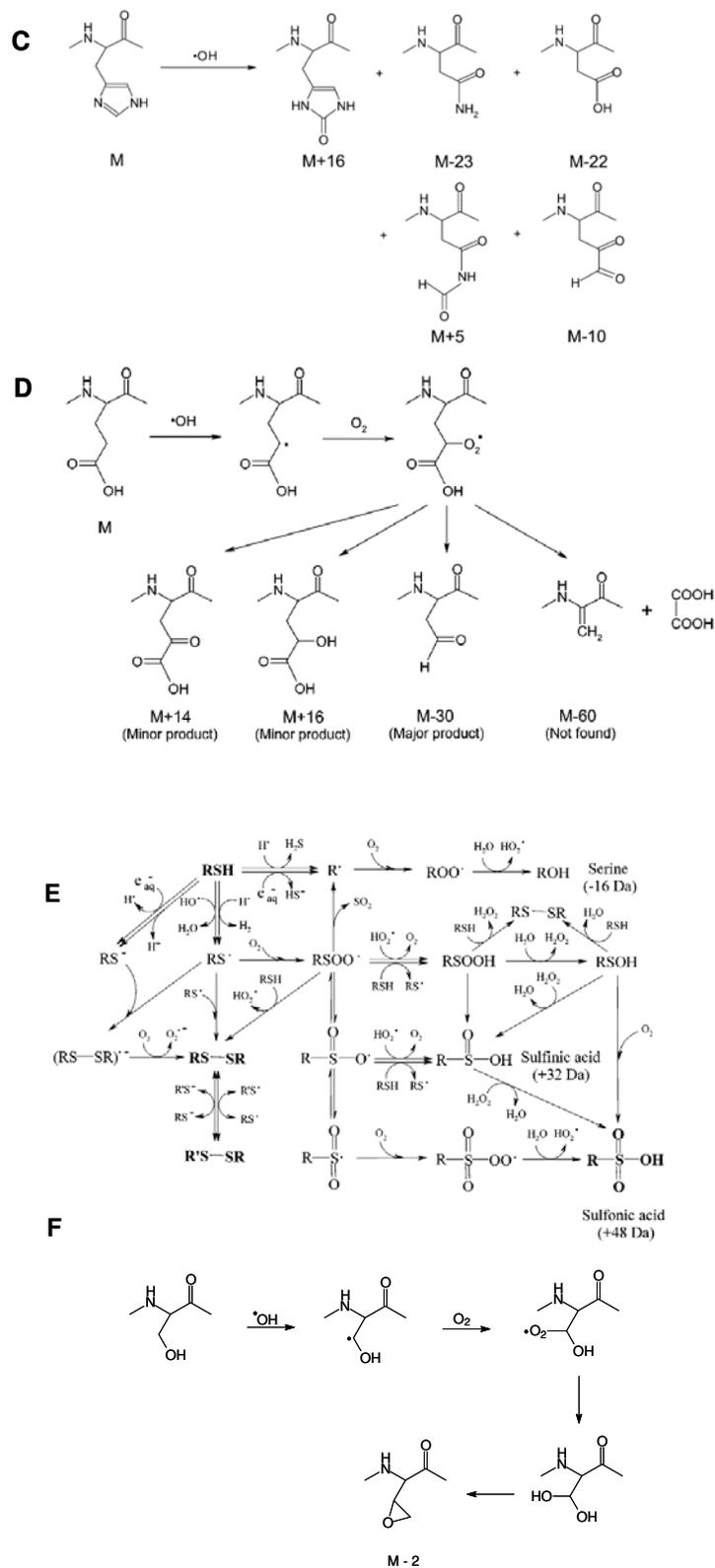
### IV.2.1 Introdução

Em 1999, o grupo de Mark Chance propôs a utilização de radicais hidroxilas provenientes da radiólise da água para o estudo de estrutura de proteínas e suas interações em solução.<sup>135</sup> A idéia por trás desse experimento estava no sucesso obtido no estudo de estrutura de ácidos nucleicos em solução utilizando o mesmo aparato experimental.<sup>201</sup> Embora a utilização de  $\bullet\text{OH}$  aplicada a estudos de estrutura de proteínas não fosse uma inovação, a utilização de fontes de luz síncrotron trouxe como grande vantagem para esses experimentos o fato das modificações ocorrerem na escala de ms, compatível com o tempo de grandes mudanças conformacionais nas proteínas. Entretanto, existia uma diferença fundamental nos experimentos com relação ao tipo de modificação causado pelos radicais: enquanto no caso de RNA e DNA ocorria a clivagem da cadeia principal (ligação fosfodiéster), no caso de proteínas as reações ocorrem majoritariamente na cadeia principal. Isso se deve ao fato de que a cinética da reação nesses sítios é cerca de 5 a 100 vezes maior que no carbono  $\alpha$ , dependendo da reatividade da cadeia lateral.<sup>135</sup>

A partir desse trabalho, vários outros foram realizados com o intuito de quantificar a cinética de oxidação dos diferentes aminoácidos e caracterizar seus produtos de oxidação. Primeiramente, determinou-se os produtos da reação de  $\bullet\text{OH}$  com aminoácidos alifáticos e aromáticos.<sup>135</sup> Resíduos alifáticos reagem mediante abstração de próton pelo radical, levando a formação de uma espécie radicalar na qual o elétron desemparelhado localiza-se no átomo de carbono. Após essa etapa, ocorre a reação com o  $\text{O}_2$  dissolvido na água

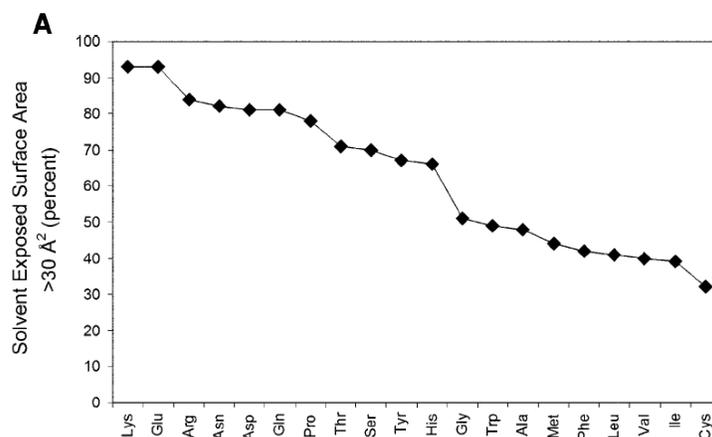
levando a formação de um radical peroxil, que pode levar a formação dos produtos de adição de hidroxila (mais comum) ou a formação de uma carbonila (Figura 36).<sup>130</sup> Resíduos aromáticos por sua vez reagem por meio de abstração de um hidrogênio do anel, criando um sítio radicalar no carbono do anel; esse sítio sofre adição de O<sub>2</sub>, reagindo posteriormente como espécies alifáticas (Figura 36).<sup>130</sup> Em ambos os casos é possível observar a formação de produtos de múltipla oxidação, dependendo dos parâmetros experimentais (em especial tempo de exposição), sendo esses muito comumente observados em resíduos aromáticos. Posteriormente, foram descritos os mecanismos propostos para a oxidação de resíduos básicos<sup>136</sup>, ácidos<sup>137</sup> e contendo enxofre<sup>138</sup>.





**Figura 36** – Mecanismo e principais produtos de oxidação de resíduos alifáticos<sup>130</sup> (A), aromáticos<sup>135</sup> (B), básicos<sup>136</sup> (C), ácidos<sup>137</sup> (D), contendo enxofre<sup>138</sup> (E) e álcoois<sup>130</sup> (F) em experimentos de *footprinting*.

Após a determinação das possíveis espécies obtidas quando da oxidação de proteínas por radicais hidroxila, o grupo de Chance avaliou as taxas de oxidação relativa, juntamente com a facilidade de detecção dessas espécies por MS, para determinar quais resíduos seriam sondas úteis nesses experimentos.<sup>139</sup> Os dados obtidos utilizando soluções diluídas de aminoácidos ( $\mu\text{M}$ ) e tempos de exposição da ordem de ms estavam de acordo com o estabelecido anteriormente, onde a ordem de reatividade dos aminoácidos varia da seguinte forma: contendo S > aromáticos > básicos > alifáticos ~ ácidos. Em virtude desse fato, Gly, Asn, Asp e Ala foram descartados como sondas devido à sua baixa reatividade (Figura 37). Por outro lado, Ser e Thr também foram descartadas como sondas nesses experimentos, em virtude da dificuldade de detecção dos produtos de oxidação por MS. Isso ocorre porque após a adição de  $\text{O}_2$  ao carbono próximo a hidroxila da cadeia lateral desses resíduos, pode haver a liberação do radical  $\bullet\text{OOH}$ , gerando um produto cuja diferença de massa com o precursor é de somente 2 Da.<sup>139</sup>

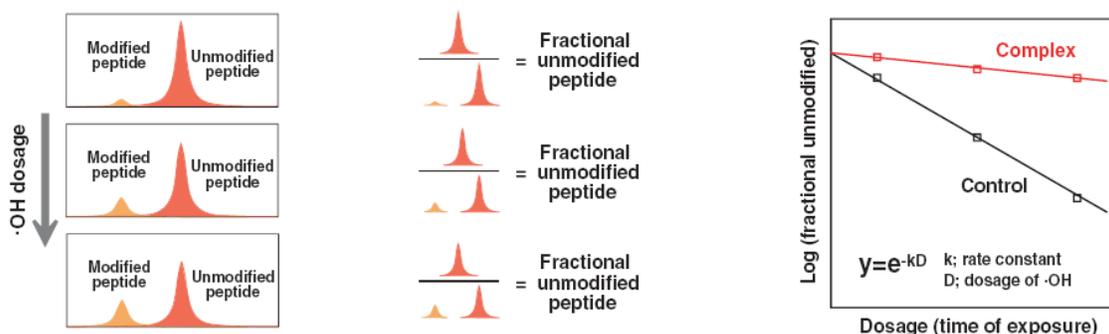


<b>B</b> amino acid	rate of reaction with $\cdot\text{OH}$ ( $\text{M}^{-1} \text{s}^{-1}$ ) <sup>a</sup>	frequency in proteins (%) <sup>b</sup>	fraction observed with SEA $>30 \text{ \AA}^2$ <sup>c</sup>	SEA $\times$ frequency <sup>d</sup>
Cys	$3.4 \times 10^{10}$	2.8	0.32	0.9
Tyr	$1.3 \times 10^{10}$	3.5	0.67	2.3
His	$1.3 \times 10^{10}$	2.1	0.66	1.4
Met	$8.3 \times 10^9$	1.7	0.44	0.7
Phe	$6.5 \times 10^9$	3.5	0.42	1.5
Arg	$3.5 \times 10^9$	4.7	0.84	3.9
Leu	$1.7 \times 10^9$	7.5	0.41	3.1
Ile	na	4.6	0.39	1.8
Trp	$1.3 \times 10^9$	1.1	0.49	0.5
Pro	na	4.6	0.78	3.6
Val	$7.6 \times 10^8$	6.9	0.4	2.8
Thr	$5.1 \times 10^8$	6.0	0.71	4.3
Ser	$3.2 \times 10^8$	7.1	0.70	5.0
Glu	$2.3 \times 10^8$	6.2	0.93	5.8
Gln	na	3.9	0.81	3.2
Ala	$7.7 \times 10^7$	9.0	0.48	4.3
Asp	$7.5 \times 10^7$	5.5	0.81	4.5
Asn	$4.9 \times 10^7$	4.4	0.82	3.6
Lys	$3.5 \times 10^7$	7.0	0.93	6.5
Gly	$1.7 \times 10^7$	7.5	0.51	3.8

**Figura 37** – (A) Área superficial exposta ao solvente para os 20 aminoácidos em proteínas depositadas no PDB.<sup>98</sup> (B) Cinética de reação dos aminoácidos frente ao radical hidroxila, frequência de cada aminoácido em proteínas, fração na qual esses aminoácidos apresentam área superficial exposta (SEA) maior que  $30 \text{ \AA}^2$  e o produto entre a frequência dos aminoácidos em proteínas e a SEA.

Baseados nos dados obtidos anteriormente, foi posteriormente proposto o diagrama apresentado na Figura 8 para a obtenção de dados estruturais de complexos protéicos por experimentos de *footprinting*. Após a irradiação do controle e experimento por diferentes intervalos de tempo, as amostras são digeridas e analisadas por LC-MS/MS. O método utilizado para quantificar a taxa de oxidação de peptídeos consiste na construção de uma curva tipo dose-resposta. Essa curva é construída sendo que para cada intervalo de tempo de exposição à radiação, a fração de peptídeo não-modificado é calculada dividindo-se a área do pico cromatográfico do peptídeo não-modificado pela soma das áreas do peptídeo modificado e não modificado. Esses valores são ajustados para uma reta, considerando cinética de pseudo-primeira ordem de oxidação (Figura 38).<sup>129</sup>

### Calculation of rate constants from dose response curve



**Figura 38** – Cálculo da cinética de oxidação de peptídeos em experimentos de *footprinting*: para cada intervalo de tempo de exposição, a fração de peptídeo não modificado é calculada pela divisão entre peptídeo não modificado e peptídeo total (modificado + não modificado).<sup>129</sup>

A grande limitação dessa metodologia consiste na determinação da quantidade de peptídeo modificado. Em todos os trabalhos descritos, considera-se como produto de oxidação peptídeos que apresentam diferença de massa +14 ou +16 Da em relação ao precursor, visto que essa é razoavelmente comum para aminoácidos alifáticos. Ocorre que, como descrito na Introdução, a reatividade dos resíduos em experimentos de *footprinting* é dependente de dois parâmetros: reatividade intrínseca e grau de exposição ao solvente. Desse modo, considerando uma contribuição igual para esses dois parâmetros, os resíduos mais suscetíveis a oxidação seriam na ordem: lisina, ácido glutâmico, serina, ácido aspártico e treonina. É interessante notar que para esses aminoácidos os produtos de oxidação +14 e +16 Da são minoritários ou até mesmo inexistentes. Além disso, em muitos casos pode ocorrer de em um determinado peptídeo haver mais de um aminoácido exposto/protegido, e portanto passível de oxidação. Isso deve ser extremamente comum uma vez que a interação entre proteínas não ocorre via

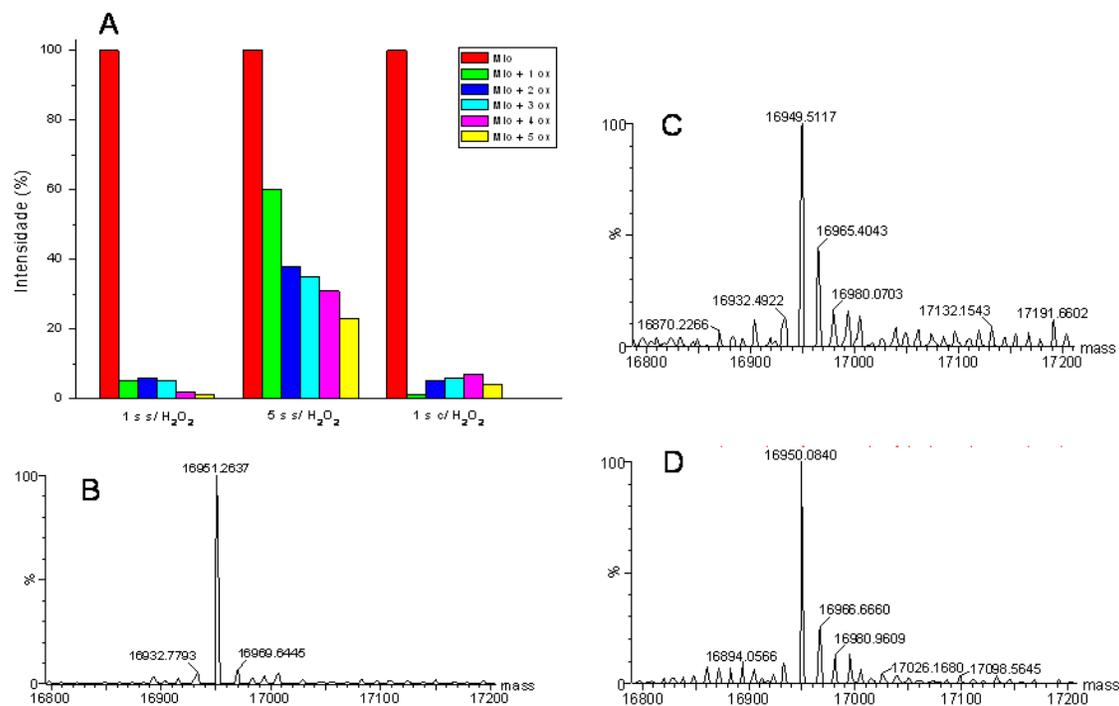
contato entre alguns poucos aminoácidos, mas sim entre superfícies, englobando vários resíduos contidos em vários peptídeos.<sup>27</sup> Finalmente, essa abordagem não considera a possibilidade de formação de outros produtos de oxidação ainda não descritos na literatura. Desse modo, considerar somente essas modificações no cálculo da cinética de oxidação em muitos casos acarretaria em subestimar tais taxas.

#### IV.2.2 Experimentos de *Footprinting* na Linha de Luz DFX do LNLS

Os primeiros experimentos de *footprinting* foram realizados na linha de luz DFX do Laboratório Nacional de Luz Síncrotron (LNLS). A DFX é uma linha de luz utilizada para realizar testes diagnósticos do feixe de luz (intensidade, estabilidade), que não apresenta nenhum elemento óptico que cortasse a região espectral de interesse ( $0.9 \text{ \AA} < \lambda < 4.0 \text{ \AA}$ ). Vale ressaltar nesse ponto que, diferentemente da maioria dos outros experimentos utilizando luz síncrotron, nesse caso não é necessário a utilização de nenhum elemento para seleção de comprimento de onda. Na realidade, a ausência desses elementos constitui a situação ideal, uma vez que a função da radiação é gerar radicais hidroxila, sendo que para isso é desejável a maior densidade de fótons possível (com isso, diminui-se o tempo de exposição necessário para causar as modificações desejadas, evitando-se assim outras reações, como clivagem da cadeia polipeptídica ou modificações após mudanças conformacionais induzidas pela oxidação). Isso faz com que, dentre as linhas de luz de um anel síncrotron, a instrumentação necessária para a realização desse tipo de experimento seja uma das mais simples, sendo necessário somente um *shutter*

(mecanismo que controla o tempo de exposição da amostra a radiação) com a resolução temporal necessária.

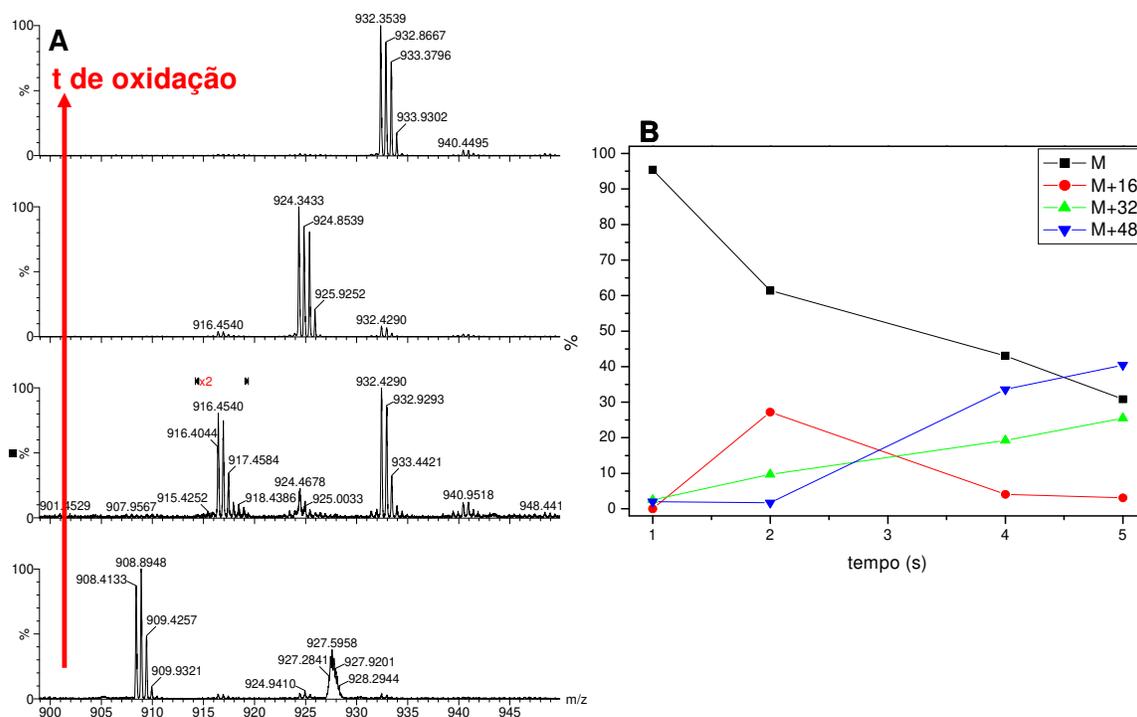
Os primeiros experimentos de *footprinting* foram realizados com o objetivo de determinar se a radiação gerada no LNLS seria suficiente para causar oxidações na proteína suficientemente detectáveis por MS e em um curto intervalo de tempo. A linha DFX dispõe de um *shutter* com resolução temporal de 1 s e a amostra fica localizada a cerca de 50 cm da saída do feixe da região de alto vácuo. Além disso, também foi avaliado se a adição de peróxido de hidrogênio interferiria nessa escala de tempo, ou seja, se devido à facilidade dessa espécie em formar radicais hidroxila, seria possível utilizar menores intervalos de exposição. A Figura 39 apresenta um gráfico da intensidade dos produtos de oxidação formados por diferentes intervalos de tempo e com a presença ou ausência de H<sub>2</sub>O<sub>2</sub> da proteína mioglobina. Por esse gráfico é possível observar que a presença de H<sub>2</sub>O<sub>2</sub> altera o perfil de oxidação, tanto quali como quantitativamente, ou seja, a adição de peróxido induz tanto a formação de maior quantidade de proteína oxidada como maior número de produtos de oxidação distintos. A exposição da amostra contendo peróxido por intervalos de tempo maiores que 1 s levaram a formação de produtos de degradação, decorrentes da clivagem da cadeia principal da proteína.



**Figura 39** – (A) Gráfico da formação de produtos de oxidação da mioglobina em função do tempo de exposição e da presença de peróxido. (B), (C) e (D) apresentam os espectros de MS deconvoluídos para a mioglobina para 1 s sem peróxido, 5 s sem peróxido e 1 s com peróxido, respectivamente.

Após isso, foram realizados experimentos em que o tempo de exposição das proteínas variou de 1 a 5 segundos (em intervalos de 1 segundo), para determinar a taxa e os sítios de oxidação dos peptídeos. Um dos peptídeos identificados pela análise de LC-MS/MS foi o GLSDGEWQQVLNVWGK, que se encontra no N-terminal da mioglobina. Esse peptídeo aparece no espectro como um íon duplamente carregado ( $m/z$  908.4); além desse, foi possível também identificar os íons de  $m/z$  916.4, 932.4 e 940.4, que correspondem respectivamente a uma, duas e três adições de hidroxila a esse peptídeo. Como era esperado, os peptídeos modificados apresentam tempo de retenção menores na cromatografia de fase reversa, uma vez que a oxidação os torna mais polares. Além disso, outro fato interessante refere-se à abundância dos diferentes estados de oxidação: o peptídeo monooxidado tem abundância

mais alta com menores tempos de exposição, mas a partir de 4 segundos essa tendência se inverte e os peptídeo di e trioxidados passam a ser os mais abundantes, confirmando a sua transformação em outras espécies com maior número de oxidações.



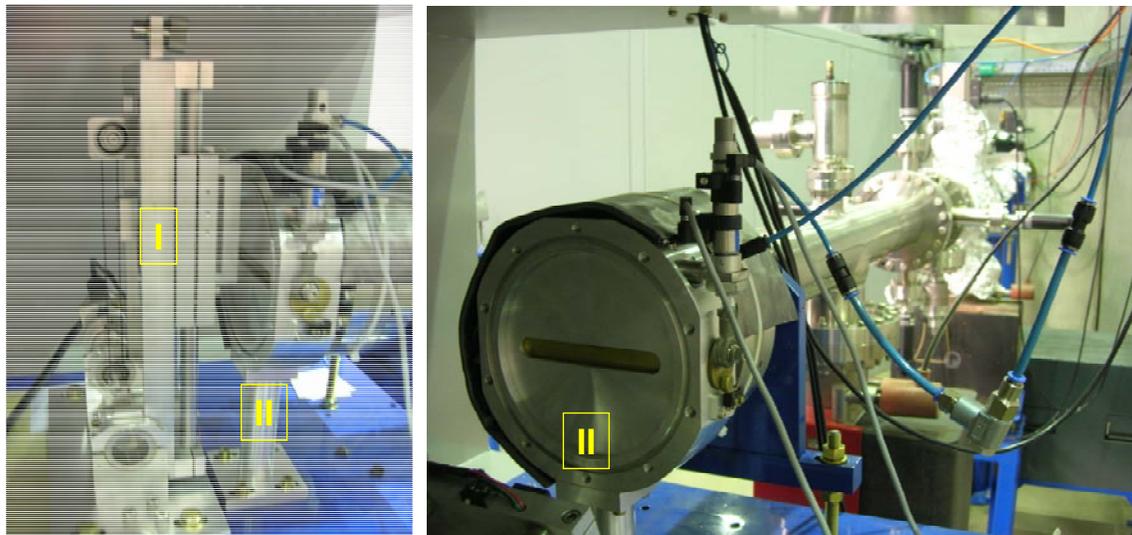
**Figura 40 -** (A) Espectros de MS da amostra com diferentes intervalos de tempo de exposição. Os íons de  $m/z$  908.2, 916.2, 924.2 e 932.2 indicam respectivamente o precursor e as espécies mono, di e trihidroxilada. (B) Cinética de formação dos produtos apresentados no espectro anterior.

Embora os resultados obtidos até então estivessem de acordo com os dados da literatura<sup>146,147</sup>, em especial com relação ao grau de cobertura dos resíduos modificados, a linha DFX apresentava limitações que dificultavam sua aplicação para a identificação de sítios de interação entre proteínas. Essa linha não apresentava um suporte fixo para as amostras, o que dificultava o mesmo posicionamento para todas as irradiações. Além disso, esse suporte

localizava-se a uma grande distância da saída do feixe (aproximadamente 50 cm), sendo que a essa distância grande parte da radiação sofria espalhamento e desse modo não alcançava a amostra. Finalmente, o *shutter* disponível para o controle do tempo de exposição era lento e acionado manualmente, o que aumentava o erro no tempo de exposição a tal modo que impedia o cálculo da cinética de oxidação desses peptídeos.

### IV.2.3 Experimentos de *Footprinting* na Linha de Luz D06B do LNLS

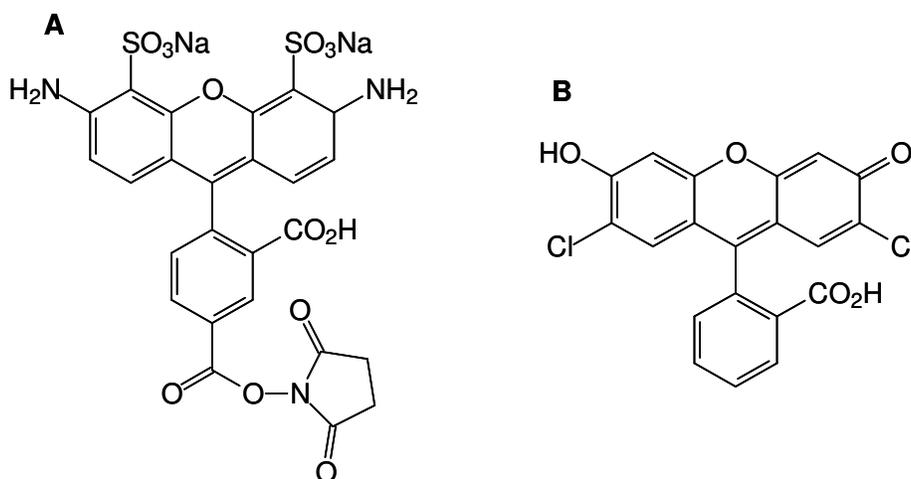
Para solucionar essas limitações, novos experimentos de *footprinting* foram realizados na linha de luz D06B do LNLS. A linha D06B era utilizada para experimentos de litografia de raios-X e, portanto, não contava com nenhum dispositivo óptico que cortasse fótons na região espectral de interesse. Desse modo, três mudanças foram implementadas nessa linha para torná-la ainda mais adequada para a realização dos experimentos de interesse: i) substituição do porta-amostras anterior por um sistema controlado remotamente por computador, que permite o posicionamento de até 6 amostras sem a necessidade de reabrir a estação experimental; ii) instalação de um *shutter* rápido, com resolução da ordem de dezenas de ms; iii) adequação do *software* de controle dos dispositivos da linha. A Figura 41 apresenta uma foto da estação experimental da linha XRL; é interessante notar que nessa nova linha de luz a amostra fica localizada a poucos centímetros da saída do feixe, o que maximiza a eficiência do processo.



**Figura 41** – Estação experimental da linha de luz XRL, onde foram realizados os experimentos de *footprinting*. Os índices I e II indicam respectivamente o porta-amostras (capacidade para 6 amostras) e o *shutter* rápido.

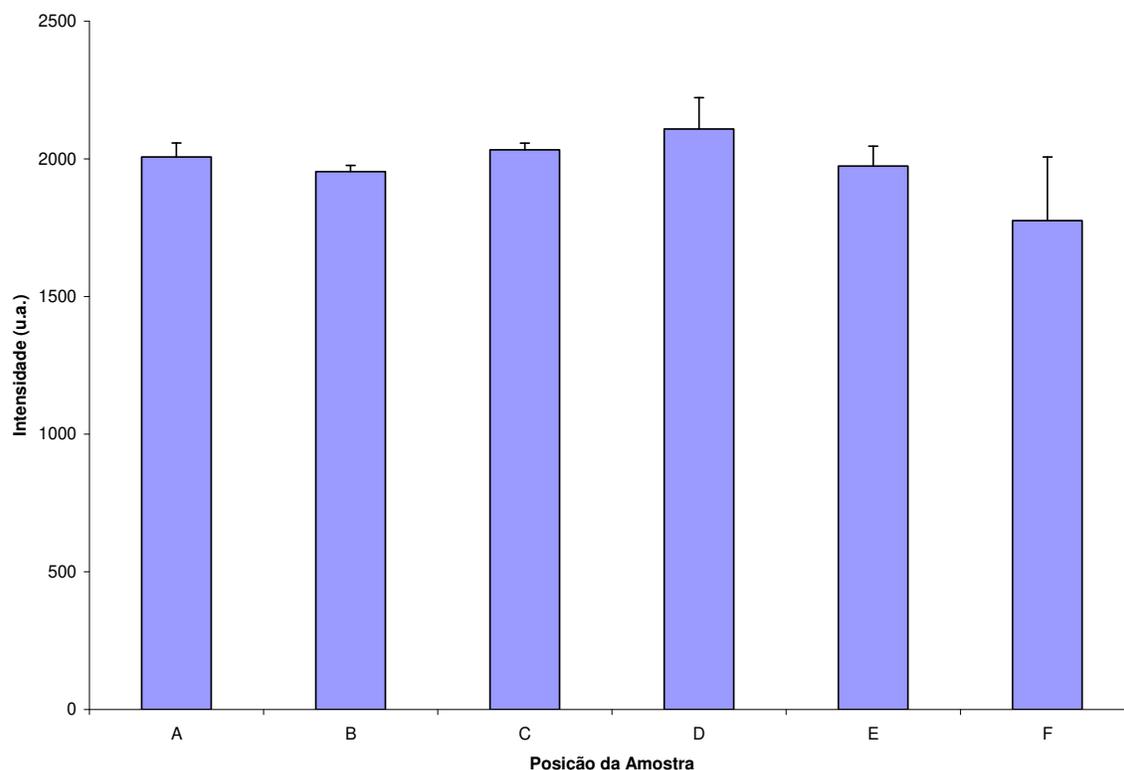
Antes da realização de experimentos com proteínas, a linha XRL passou por dois testes de caracterização: posicionamento do porta-amostras e determinação do intervalo de tempo no qual a dosagem de radicais hidroxila variava de acordo com o tempo de exposição. Foi descrito na literatura a utilização de moléculas “dosímetras”, utilizadas para quantificar a dose de radicais às quais a amostra era submetida.<sup>130</sup> Soluções contendo esses fluoróforos eram irradiadas por diferentes intervalos de tempo e após isso, sua fluorescência era medida. A reação dos radicais com essas moléculas levam ao *quenching* de sua fluorescência, de modo que esse tipo de experimento permite quantificar o teor de modificação e, conseqüentemente, de exposição da amostra a radiação. Em geral, a molécula utilizada para esse fim era a Alexa Fluor 488, um derivado da fluoresceína altamente solúvel em água; os nossos experimentos foram realizados com a diclorofluoresceína (DCF), que

além de ser mais barato apresentou as mesmas propriedades da molécula anterior (Figura 42).



**Figura 42** – Estruturas do Alexa Flúor 488 (A) e Diclorofluoresceína (B).

O primeiro experimento na linha consistiu na caracterização do porta-amostras. Essa peça apresenta 6 posições possíveis para o posicionamento de tubos tipo Eppendorf 200  $\mu$ L. Foram realizados experimentos para verificar se a mesma amostra posicionada nas diferentes posições e irradiadas pelo mesmo intervalo de tempo, recebiam a mesma dosagem de radiação. Esses experimentos foram realizados alinhando-se o porta-amostras na primeira posição (A) e irradiando em todas as posições de modo automático (controlado pelo computador). Os resultados estão apresentados na Figura 43. É possível notar pelo gráfico que uma das posições (F) aparentemente recebe uma quantidade de radiação diferente das outras posições e, portanto, essa posição não foi utilizada para os outros experimentos.



**Figura 43** – Intensidade da fluorescência da amostra irradiada nas diferentes posições do porta-amostras (n = 3).

A próxima etapa consistiu na caracterização dos intervalos de tempo de exposição da amostra. Os objetivos desse experimento foram determinar se dose de exposição é proporcional ao tempo de exposição (ou seja, se não ocorre saturação do meio a partir de um intervalo de tempo, onde não haveria mais oxidação) e se essa relação mantinha sua tendência ao longo do experimento. Soluções aquosas de qualquer analito (tanto a DCF quanto peptídeos e proteínas) em baixas concentrações ( $\mu\text{M}$ ) podem ter sua cinética de oxidação aproximada a de pseudo-primeira ordem. Isso resulta do fato da água estar em concentrações muito superiores ao analito, de modo que esse é o fator limitante da reação. Desse modo, pode-se considerar a reação de radiólise como:

$$M \rightarrow P_1 + P_2 + P_3 + \dots = \sum_{i=1}^n P_i$$

onde M e P são respectivamente o analito e o produto. Por definição, a concentração inicial de peptídeo é 1, M é a fração do analito não-modificado e  $P_i$  é a fração do  $i$ -ésimo produto de oxidação. Ao longo do experimento, considerando reações somente de modificação (ausência de clivagem), a concentração total do analito permanece inalterada, ou seja:

$$M + \sum_{i=1}^n P_i = 1$$

Como a velocidade da reação é limitada pela presença de analito (a concentração de radical permanece inalterada), a taxa de decaimento de M é dada pela equação:

$$dM/dt = -kM$$

ou seja, ocorre um decaimento exponencial de M em função do tempo (k é a constante de velocidade da reação). A concentração de M pode ser obtida em qualquer momento do experimento pela integração da equação anterior:

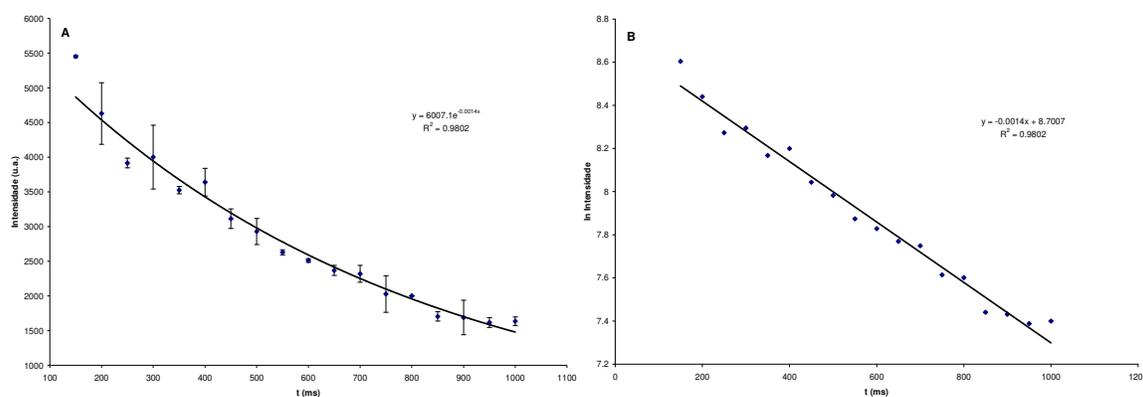
$$M = e^{-kt}$$

A constante de velocidade da oxidação pode ser obtida pelo ajuste não-linear da equação anterior:

$$\ln M = -kt$$

Isso significa que traçando-se uma curva de M em função do tempo pode-se obter o valor de k, que será subsequentemente utilizado na comparação de peptídeos em duas condições distintas: uma diminuição de k de um peptídeo após a interação da proteína com seu ligante significa que sua cinética de oxidação foi diminuída em função da proteção a oxidação.

A curva dose-resposta para a solução de DCF irradiada por diferentes intervalos de tempo está mostrada na Figura 44A. A fluorescência de DCF diminui exponencialmente com o aumento do tempo de exposição, fato esse compatível com uma cinética de pseudo-primeira ordem. A Figura 44B apresenta o ajuste linear para a curva anterior; o fato de a linearidade ser mantida no intervalo de 100-1000 ms indica que esse intervalo de tempo pode ser utilizado em experimentos de *footprinting* com proteínas.



**Figura 44** – (A) Intensidade da fluorescência de DCF em função do tempo de exposição. (B) Linearização da curva (A) de acordo com as equações anteriores.

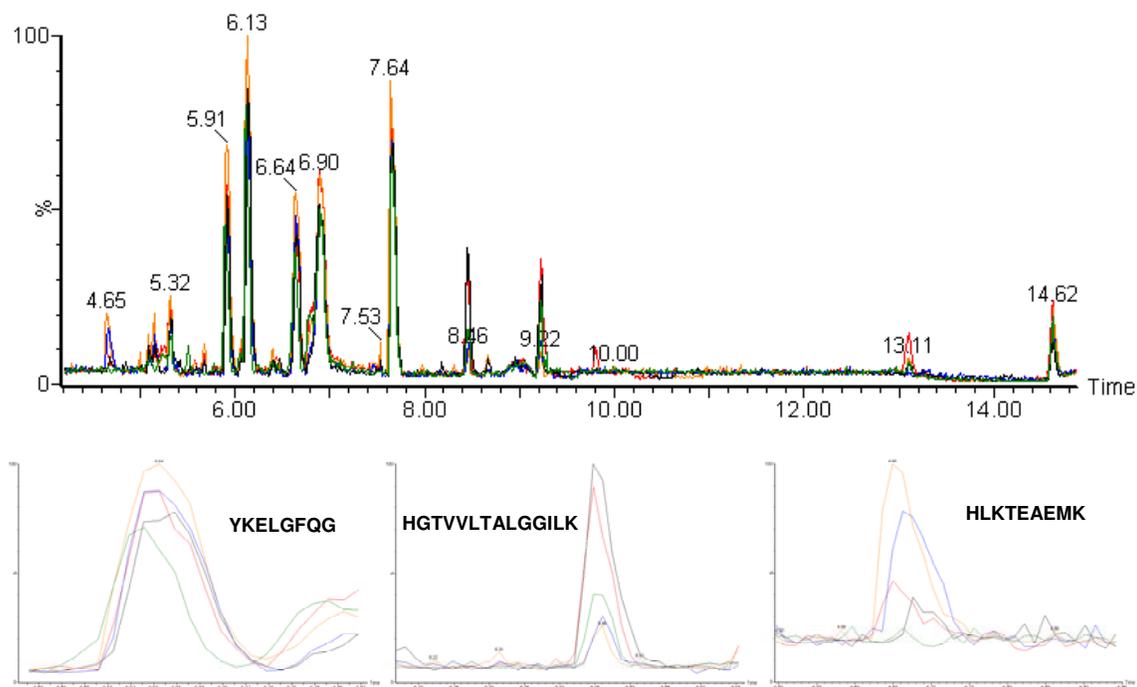
#### IV.2.4 Quantificação de Cinética de Oxidação pelo Método *Label-Free*

Conforme descrito anteriormente, o método utilizado para o cálculo de cinética de oxidação de peptídeos em experimentos de *footprinting* pode subestimar tais taxas. Isso ocorre em virtude do fato de serem considerados como produtos de oxidação somente as espécies +14 e +16. Com o objetivo de aumentar a confiabilidade nas taxas de oxidação obtidas nesse experimento, foi proposto um novo método de quantificação, conhecido como *label-free*

(sem adição de marcação isotópica).<sup>202,203,204</sup> Embora esse método tenha sido aplicado há muitos anos para moléculas pequenas, os primeiros exemplos com peptídeos datam do início dessa década, em virtude de dificuldades experimentais principalmente relacionadas com a etapa cromatográfica. Nesse caso, ao invés de considerar somente algumas modificações específicas para realizar a quantificação, todas as modificações são consideradas, inclusive oxidações múltiplas do peptídeo e até mesmo eventuais modificações ainda não descritas na literatura. Experimentalmente, a quantificação é realizada calculando-se somente a área do pico cromatográfico referente ao peptídeo não-modificado. Com isso, considera-se que a diminuição de intensidade do pico cromatográfico corresponde a alguma modificação. Do ponto de vista experimental, além desse método considerar todas as possíveis modificações em um determinado peptídeo, ele é ainda mais simples de ser implementado, visto que para cada íon é necessária calcular a área de somente um pico cromatográfico e não mais três (peptídeo não-modificado e os produtos +14 e +16). Além disso, existem alguns softwares com a capacidade de automatizar todo esse processo, o que aumentaria significativamente a velocidade de análise.

A implementação da rotina do método de *label-free* passa pela necessidade de um sistema cromatográfico extremamente reprodutível, tanto do ponto de vista quali como quantitativo. Em geral, a análise quantitativa é realizada em experimentos de LC-MS e não LC-MS/MS; isso é feito para obtenção de um maior número de pontos no pico cromatográfico, o que aumenta a exatidão dessa análise. Desse modo, a identidade dos peptídeos não é mais associada ao seu espectro de dissociação. Os únicos parâmetros que permitem identificar uma espécie nessas análises são o tempo de retenção cromatográfico, a  $m/z$  e o perfil do padrão isotópico. Além disso, do ponto de

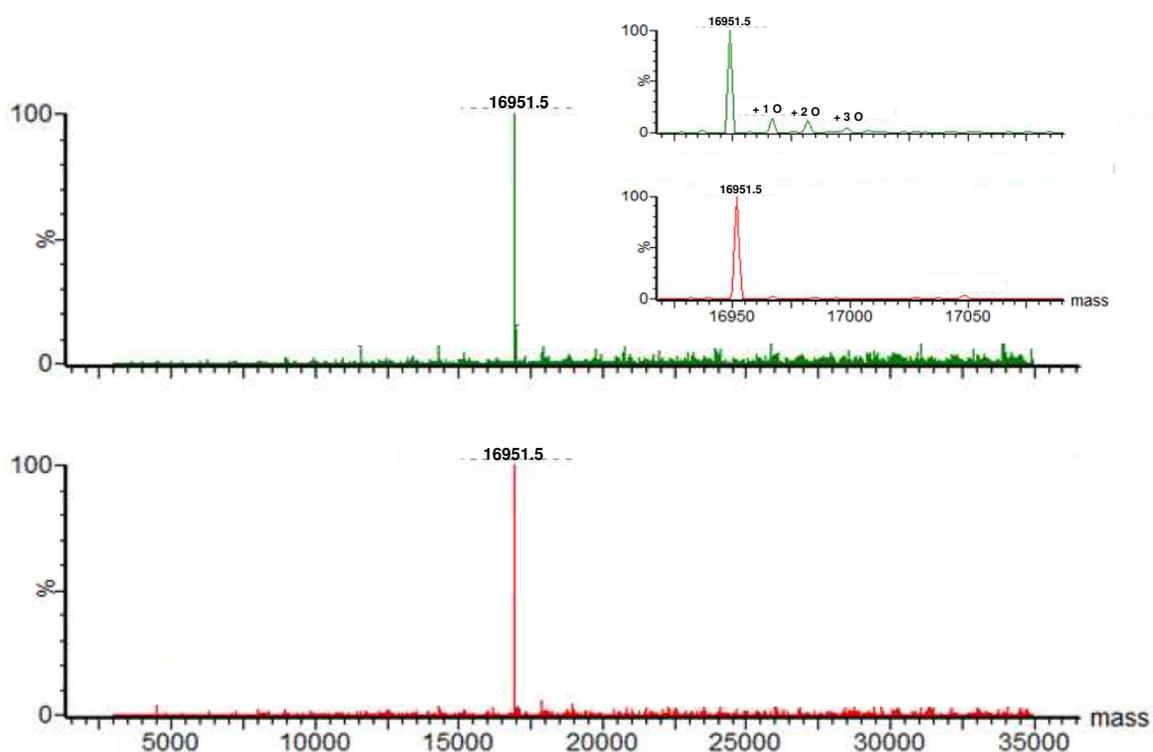
vista quantitativo, a análise realizada considerando-se os produtos +14 e +16 eram internamente normalizadas, ou seja, caso houvesse algum erro na injeção ou na própria cromatografia, idealmente todas as espécies seriam afetadas proporcionalmente, o que permitia a normalização e conseqüentemente a eliminação desse erro; por outro lado, o método proposto não permite essa normalização, visto que somente uma espécie é quantificada. A Figura 45 apresenta 5 cromatogramas de um digesto de mioglobina exposta por diferentes intervalos de tempo. É importante notar nessa figura a altíssima reprodutibilidade de tempo de retenção, aliada a uma grande resolução cromatográfica que garante um mínimo de co-eluição de peptídeos; isso é extremamente desejável quando se realiza uma análise quantitativa para evitar a supressão iônica. Além disso, é possível observar pelas correntes iônicas extraídas para 3 diferentes peptídeos as diferentes cinéticas de oxidação: o peptídeo YKELGFQG apresenta suas cadeias laterais mais suscetíveis a oxidação protegidas na estrutura da mioglobina (corrente iônica permanece praticamente constante com o aumento do tempo de exposição), enquanto o peptídeo HLKTEAEMK apresenta grande diminuição da intensidade com o aumento do tempo, o que indica que esse peptídeo encontra-se exposto na estrutura da mioglobina.



**Figura 45** – Cromatogramas correspondentes a digestos de mioglobina após 5 intervalos de tempo de exposição diferentes e corrente iônica extraída para 3 peptídeos diferentes da mioglobina.

Outro fato que merece destaque na análise pelo método *label-free* consiste na necessidade de se garantir que a diminuição do pico cromatográfico referente ao peptídeo não-modificado se deve exclusivamente a fenômenos oxidativos. Foi descrito na literatura que dependendo das condições experimentais (em especial com extensa dosagem de radicais hidroxila), pode ocorrer a formação de *cross-linking* entre cadeias polipeptídicas e clivagem nas ligações da cadeia carbônica principal da proteína (mediadas pela formação de um sítio radicalar nos carbonos da cadeia principal).<sup>135</sup> Para comprovar que as condições experimentais utilizadas não favoreceram esse tipo de modificação, foram adquiridos espectros de MS da proteína intacta, antes e após o maior tempo de exposição utilizado, 800 ms (Figura 46). É possível observar nessa figura que os espectros de MS das duas

espécies são praticamente idênticos, não sendo observados na amostra irradiada produtos de massa menor (degradação) ou muito maior (*cross-linking*) que da mioglobina.



**Figura 46** – Espectro de MS deconvolvido para as amostras controle (linha vermelha) e irradiada por 800 ms (linha verde). A ampliação mostra a formação dos produtos de oxidação com o aumento do tempo de exposição.

## IV.2.5 Conclusões

Dentre os diversos modos de geração de radicais hidroxila em solução, a utilização de luz síncrotron é uma das mais interessantes, visto que devido à sua alta densidade de fótons é possível utilizar tempos de irradiação da ordem de ms. Esses curtos tempos de exposição garantem que não ocorram mudanças

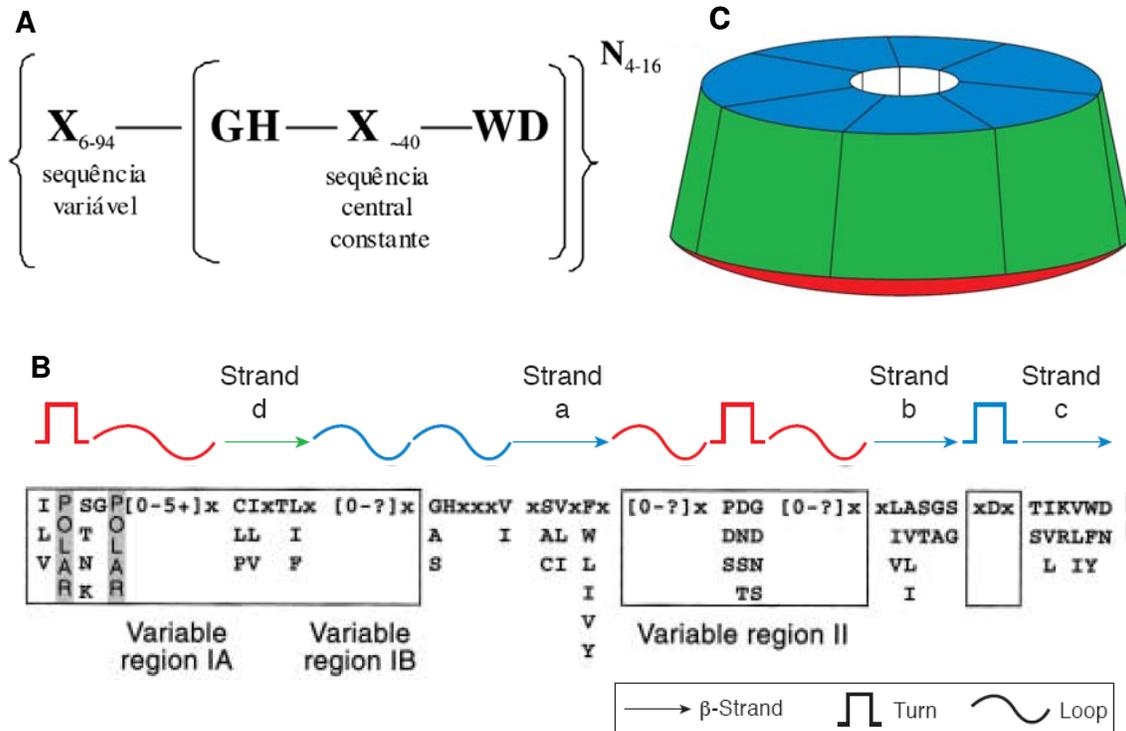
conformacionais enquanto os radicais são gerados, o que acarretaria na oxidação de sítios anteriormente protegidos na estrutura da proteína ou na superfície de interação em complexos protéicos. Após os primeiros experimentos em uma linha de diagnósticos de feixe no LNLS, foi desenvolvida a instrumentação necessária para a realização desses experimentos com resolução temporal. Isso incluiu o desenvolvimento de um *shutter* com capacidade de permitir a passagem da radiação por intervalos de tempo da ordem de ms, um porta-amostras localizada logo na saída do feixe, capaz de suportar até 5 amostras simultaneamente e um sistema controlado por computador capaz de controlar esses dispositivos, bem como alinhar o feixe de luz com a amostra, garantindo maior aproveitamento dos fótons. Para determinação de quais tempos de exposição poderiam ser utilizados para a amostra, utilizou-se o reagente diclorofluoresceína como “dosímetro molecular”, o que permitiu concluir que nos experimentos conduzidos no intervalo de 100-1000 ms obtém-se quantidade proporcional de radicais hidroxila com o tempo de exposição.

Para determinação da cinética de oxidação dos peptídeos em experimentos de *footprinting*, foi proposta uma nova metodologia de quantificação conhecida como *label-free*. Nesse caso, a quantidade de peptídeo não-modificado foi calculada somente de acordo com a diminuição da intensidade do íon correspondente no espectro de MS. Essa abordagem, além de mais simples do ponto de vista experimental, considera todos os possíveis produtos de oxidação gerados, inclusive produtos de oxidação múltiplos e ainda não descritos na literatura. Essa abordagem é diferente no sentido de não considerar somente os produtos de oxidação mais comuns para aminoácidos alifáticos (+14 e +16), o que pode subestimar significativamente a cinética de oxidação dos peptídeos.

## IV.3 ESTUDO DE INTERAÇÃO ENTRE AS PROTEÍNAS TIF34 E TIF35 DE *SACCHAROMYCES CEREVISIAE*

### IV.3.1 Introdução

Conforme descrito na Introdução, análises de similaridade de seqüência apontaram a proteína Tif34 como pertencente à família de proteínas conhecidas como *WD-repeat*. Essa classe de proteínas é conhecida por apresentar domínios conservados de aproximadamente 40 a 60 aminoácidos, começando com o dipeptídeo GH, 11-24 resíduos a partir do N-terminal e terminando com o dipeptídeo WD. Entre esses dipeptídeos encontra-se a região central conservada contendo aproximadamente 40 aminoácidos. As proteínas dessa família apresentam de 4 a 16 repetições dos domínios WD (Figura 47A).<sup>205</sup> A primeira estrutura de uma proteína dessa família, subunidade  $\beta$  da proteína G, foi resolvida em 1995.<sup>206</sup> A estrutura terciária dessa proteína consiste em uma  $\beta$ -propeller (hélice), estrutura essa altamente simétrica constituída de vários *WD-repeats*. Cada uma dessas repetições adota um arranjo de 4 folhas  $\beta$  anti-paralelas, de modo que cada seqüência corresponde a uma repetição estrutural dessas folhas  $\beta$  (Figura 47B). Entretanto, cada repetição não corresponde a uma “lâmina” da hélice: cada lâmina tem 4 folhas  $\beta$ , sendo que na realidade a primeira delas faz parte do último trecho da repetição WD (Figura 47C). Essa divisão de folhas entre as lâminas apresenta grande potencial de estabilização para essa estrutura.

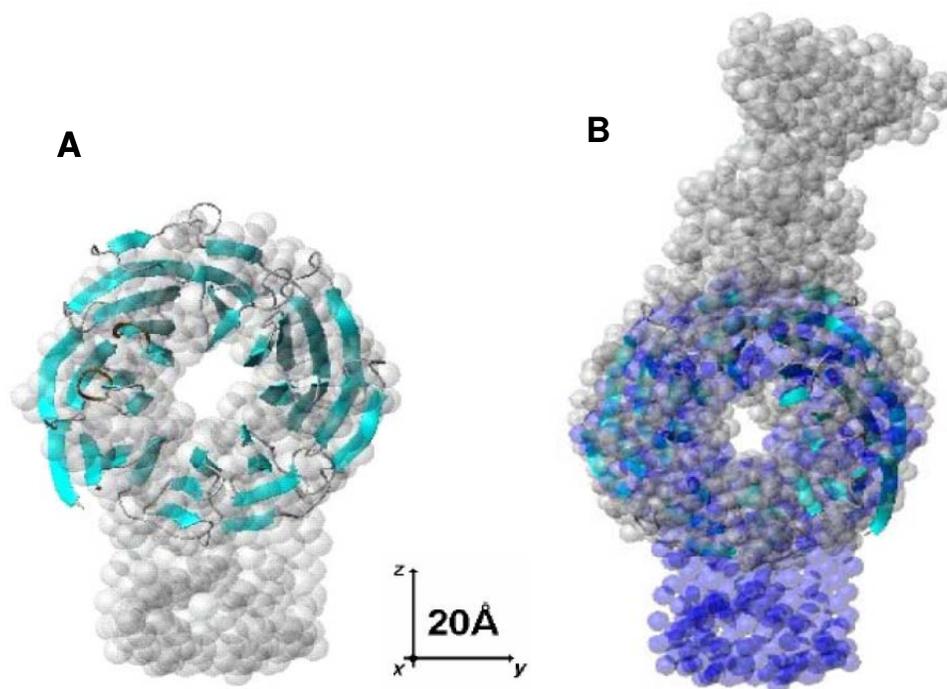


**Figura 47** – (A) Sequência característica de proteínas da família *WD-repeat*. (B) Regiões da sequência que definem as características estruturais das proteínas dessa família; os resíduos em destaque referem-se às regiões conservadas nessa estrutura. (C) Representação esquemática da estrutura tipo  $\beta$ -propeller, sendo as diferentes cores as regiões por onde as proteínas dessa família podem interagir.

Embora as repetições WD não apresentem nenhuma função enzimática, foram atribuídas inúmeras funções às proteínas dessa família, tais como síntese e processamento de RNA, controle de ciclo celular, apoptose, dentre outras. O fato aceito na literatura hoje é que os processos evolutivos atuaram no sentido da conservação da estrutura do tipo  $\beta$ -propeller, uma vez que essa estrutura apresenta um alto grau de exposição da sua superfície (regiões superior, inferior e lateral, além do canal interno). Esse fato indica que aparentemente essa estrutura serviria como uma plataforma eficaz para um grande número de interações proteína-proteína, sendo possivelmente muitas delas simultaneamente.<sup>205</sup>

Com relação à proteína Tif34 e seu complexo com a Tif35, estruturas de alta resolução não foram obtidas para esses sistemas até o momento. Tif34 é uma proteína razoavelmente grande para os padrões atuais de RMN (39 kDa); o complexo com Tif35 apresenta massa aproximada de 59 kDa. Além disso, a estrutura de Tif35 é altamente desordenada, o que dificulta significativamente análise por RMN. Um outro agravante para a utilização dessa técnica consiste na baixa estabilidade do sistema, mesmo em condições de alta força iônica (o complexo precipita quando armazenado por mais de alguns dias a 4° C em presença de tampão fosfato com 150 mM de sal). Esse complexo também foi submetido a milhares de tentativas de cristalização ao longo dos últimos anos, sendo que em nenhum caso foi possível obter cristais de proteína.<sup>160</sup>

Os dados de estrutura desses sistemas de mais alta resolução disponíveis até o momento são provenientes da técnica de Espalhamento de Raios-X a Baixos Ângulos (SAXS, Figura 48).<sup>160</sup> Os modelos obtidos por homologia permitem concluir que o envelope da proteína Tif34 é coerente com a estrutura tipo *β-propeller*, onde há uma estrutura circular com uma cavidade central. Além disso, é possível observar no modelo do complexo que a interação ocorre pela porção lateral do anel (e não pelas partes inferior e superior) e que não há grandes mudanças conformacionais em Tif34 induzidas pelo acoplamento do ligante. Entretanto, devido à alta simetria da estrutura dessa proteína, não é possível determinar pelo modelo obtido por SAXS qual o sítio de interação entre essas proteínas.

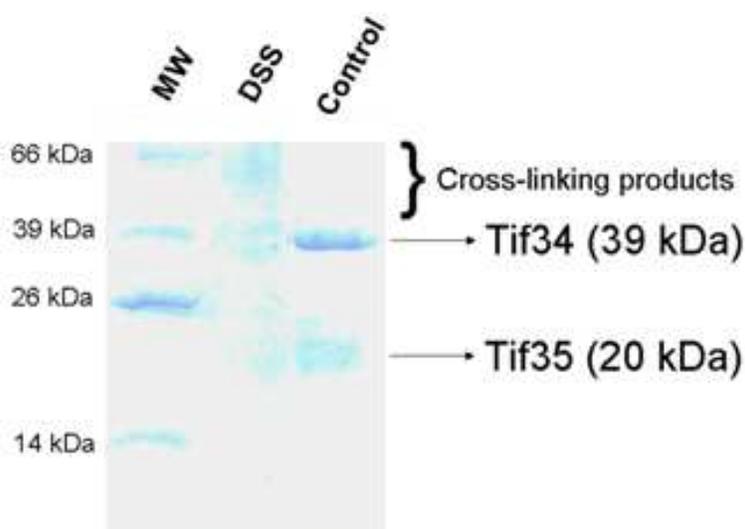


**Figura 48** – Sobreposição dos modelos obtidos por SAXS e homologia para a proteína Tif34 (A) e o complexo Tif34/Tif35 (B).<sup>160</sup>

Em virtude das características do complexo Tif34/Tif35, seria extremamente interessante a utilização de métodos de caracterização estrutural de proteínas que permitissem a identificação do sítio de interação entre essas espécies utilizando para isso pequenas quantidades de amostras e técnicas experimentais rápidas e que envolvessem a menor manipulação possível da amostra. Com isso, esse complexo foi analisado utilizando-se as técnicas descritas nas seções anteriores, *cross-linking* e *footprinting* acoplados a Espectrometria de Massas.

### IV.3.2 *Cross-Linking* do Complexo Tif34/Tif35

Após a última etapa de purificação (cromatografia de gel filtração), o complexo Tif34/Tif35 foi submetido à reação de *cross-linking* com o reagente DSS. A Figura 49 apresenta o gel de SDS-PAGE do material de partida e dos produtos de reação.



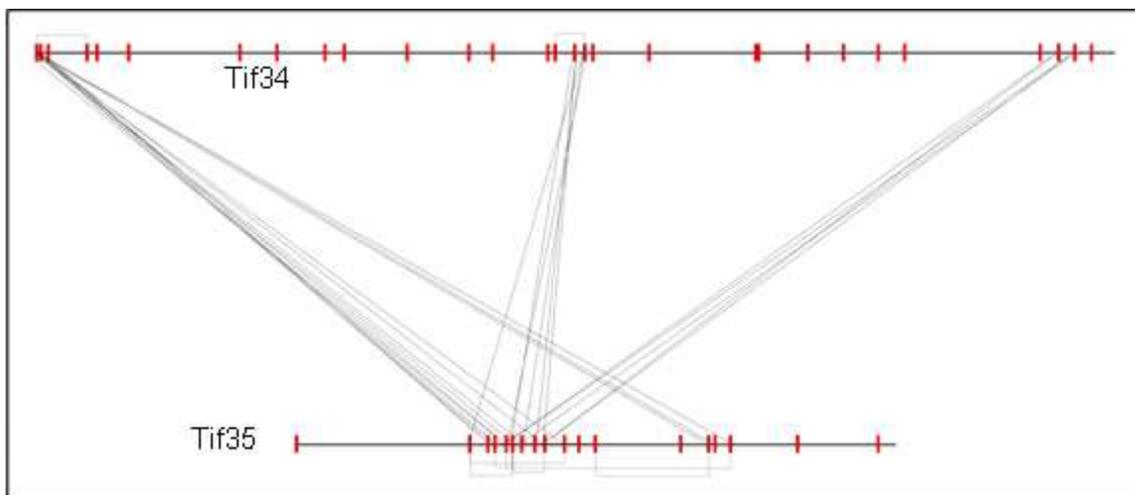
**Figura 49** – SDS-PAGE da reação de *cross-linking* entre Tif34, Tif35 e DSS.

É possível notar que o complexo nativo se dissocia nas condições em que o gel é realizado, sendo possível observar as bandas isoladas das proteínas Tif34 (39 kDa) e Tif35 (20 kDa). Por outro lado, é nota-se que a adição de DSS resulta na diminuição da intensidade das bandas referentes às proteínas isoladas bem como no aparecimento de uma banda difusa na região de 60 kDa, correspondente ao complexo covalente formado mediante a adição do *cross-linker*. A presença de uma mancha na região correspondente aos produtos de reação (e não uma banda bem definida) é esperada, em virtude da

possibilidade da formação de vários produtos da reação, tais como *dead-end* e produtos de múltiplas modificações. A região do gel referente aos produtos de interesse foi recortada, digerida com tripsina e analisada por LC-MS/MS. Embora a digestão de proteínas em solução resulte em geral em maior grau de cobertura e maior eficiência da proteólise, essa abordagem foi preterida em relação a anterior com o objetivo de enriquecer os peptídeos modificados. Como descrito nas seções anteriores, uma das grandes dificuldades da identificação dessas espécies está no fato de que os peptídeos de interesse encontram-se em quantidades subestequiométricas na amostra; com isso, mesmo com a análise de MS acoplada a LC a supressão de ionização dessas espécies é comumente observada, principalmente para complexos de aproximadamente 60 kDa. Como controle, as bandas referentes às proteínas Tif34 e Tif35 também foram recortadas, digeridas e analisadas via LC-MS/MS.

A análise dos dados foi realizada utilizando-se o *software* ASAP (<http://rosweel.ca.sandia.gov/~mmyoung>).<sup>175</sup> O ASAP cria uma lista teórica de possíveis *m/z* baseado na seqüência das proteínas utilizadas dentro de um determinado erro. Além disso, o programa permite também estipular a enzima utilizada para a digestão, bem como a presença de modificações fixas ou variáveis. Para reduzir consideravelmente a taxa de falso positivo, um íon só foi considerado como contendo *cross-linker* se apresentasse as seguintes características: não fosse encontrado no experimento controle (Tif34p/HisTif35N digeridas em solução sem o *cross-linker*) e se o erro calculado para a espécie modificada estivesse dentro da faixa de erro dos peptídeos não ligados, valores esses obtidos pelo programa MASCOT (tipicamente da ordem de 10 ppm). Os dados obtidos permitiram a construção

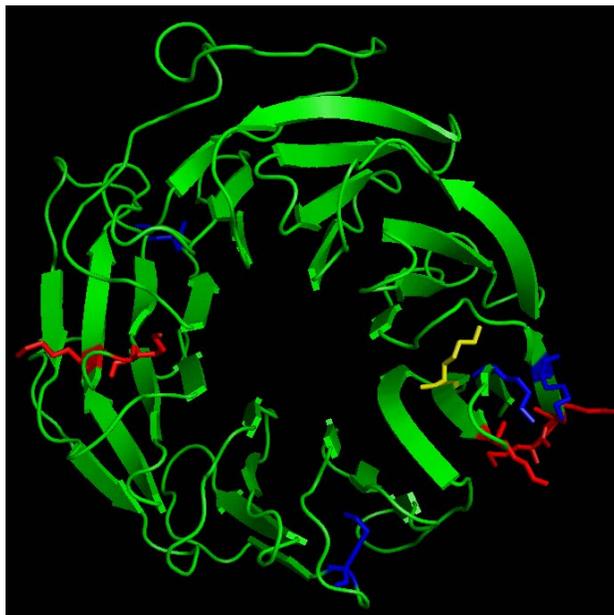
de um mapa de ligações cruzadas entre essas proteínas, apresentado na Figura 50.



**Figura 50** – Representação esquemática das ligações obtidas pela análise de MS do complexo Tif34/Tif35 ligado por DSS. Os retângulos vermelhos representam as posições dos resíduos de lisinas ao longo da seqüência das proteínas.

Nota-se pela figura acima uma grande especificidade frente algumas lisinas na formação de *cross-linking*. No caso da Tif34, há três regiões apenas onde *cross-linking* são observados: no N-terminal (resíduo N-terminal e primeira lisina da cadeia), na região central (duas lisinas vizinhas na estrutura de Tif34) e no C-terminal (duas das três últimas lisinas da cadeia). As outras lisinas, apesar de numerosas, não apresentam ligações cruzadas, mostrando a especificidade do método em relação à localização das lisinas na estrutura da proteína. Além de peptídeos com *cross-linking* intra e intercadeia, o ASAP também gerou uma lista contendo alguns peptídeos *dead-end*. No caso de modificações *dead-end*, a extensão da modificação é bem maior, uma vez que o *cross-linker* necessita conectar-se somente a um resíduo. Mais uma vez é importante ressaltar a importância da identificação desses íons, uma vez que

os *dead-end* indicam quais os resíduos encontram-se na superfície exposta da proteína. A Figura 51 apresenta o modelo de homologia da proteína Tif34p, onde foram destacadas as lisinas mais reativas com o DSS.



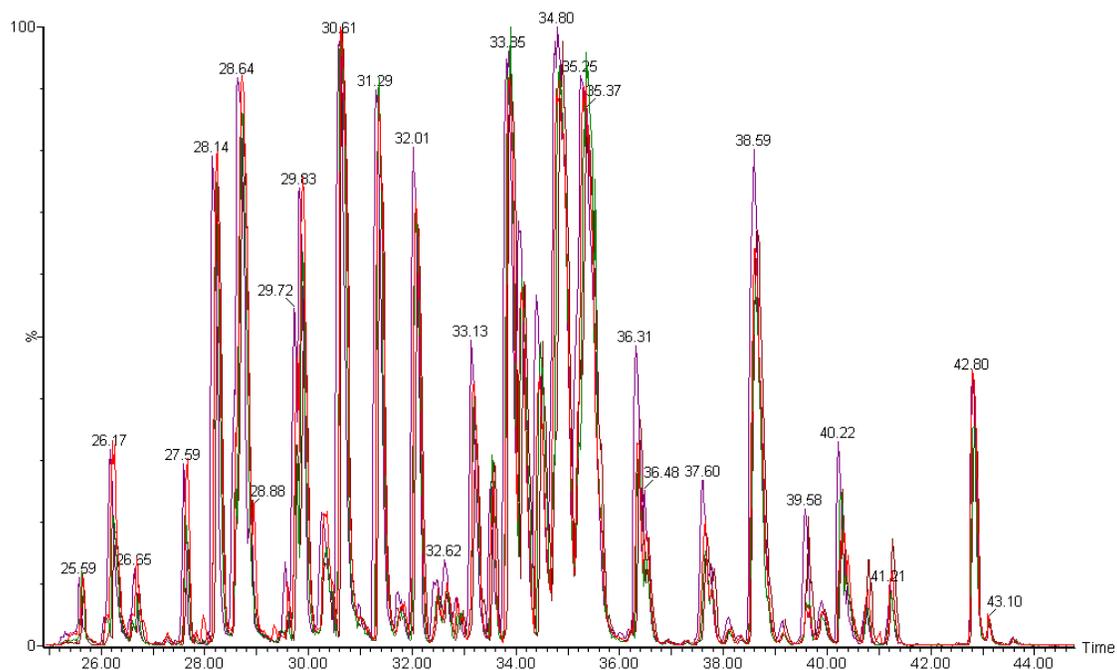
**Figura 51** – Modelo de homologia da proteína Tif34p, onde as lisinas marcadas em vermelho, azul e amarelo representam respectivamente resíduos com maior tendência a formação de peptídeos com *cross-linking* intra ou intermolecular; resíduos onde ocorre formação preferencial de *dead-end*; e um resíduo que formou tanto *cross-linking* intermolecular quanto *dead-end*.

A presença de somente três regiões contendo DSS muito bem definidas em Tif34p indica que essas são as superfícies de interação dessa proteína com Tif35. Isso é muito significativo no caso da Tif34p, uma proteína razoavelmente grande (cerca de 39 kDa) e que apresenta um grande número de lisinas. Caso os *cross-linking* observados fossem falso-positivos, seria de se esperar uma distribuição aleatória de resíduos conectados e não uma especificidade na seqüência tão grande quanto esta. Além disso, a observação do modelo de homologia dessa proteína indica que o N- e o C-terminal estão

espacialmente bem próximos, enquanto a outra região central encontra-se simetricamente oposta a essa face. Finalmente, a maioria dos resíduos modificados com *dead-end* encontra-se localizada em regiões adjacentes àquelas que apresentam *cross-linking*.

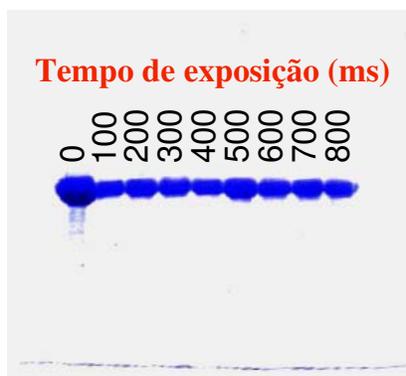
### IV.3.3 *Footprinting* do Complexo Tif34/Tif35

Experimentos de *footprinting* foram realizados para a proteína Tif34 e o seu complexo com Tif35, com o intuito de desvendar o sítio de interação entre essas proteínas. Esses experimentos foram realizados na linha de luz D06B do LNLS, cuja instrumentação foi descrita na seção IV.2.3. A quantificação das taxas de oxidação dos peptídeos gerados pela digestão trípica de Tif34 foi realizada pelo método *label-free* descrito anteriormente. Como mencionado anteriormente, a técnica de *label-free* requer uma alta reprodutibilidade cromatográfica, tanto do ponto de vista quali como quantitativo. A Figura 52 apresenta a sobreposição de cromatogramas de 5 tempos de exposição diferentes para a proteína Tif34. Como no caso da mioglobina, pode-se notar a excelente reprodutibilidade cromatográfica, sendo os tempos de retenção do mesmo peptídeo nas diferentes corridas praticamente idênticos. Além disso, é interessante notar que alguns peptídeos (tempo de retenção de 28.14 e 28.64 min) apresentam intensidades praticamente constantes em função do tempo de exposição, enquanto outras espécies (tempo de retenção de 36.31 e 38.59 min) apresentam grande decréscimo na intensidade com o aumento do intervalo de exposição a radicais hidroxila. Esses fatos demonstram claramente a alta sensibilidade da técnica para peptídeos com diferentes cinéticas de oxidação.



**Figura 52** – Sobreposição de cromatogramas da proteína Tif34 exposta à radiação síncrotron por diferentes intervalos de tempo (400-800 ms).

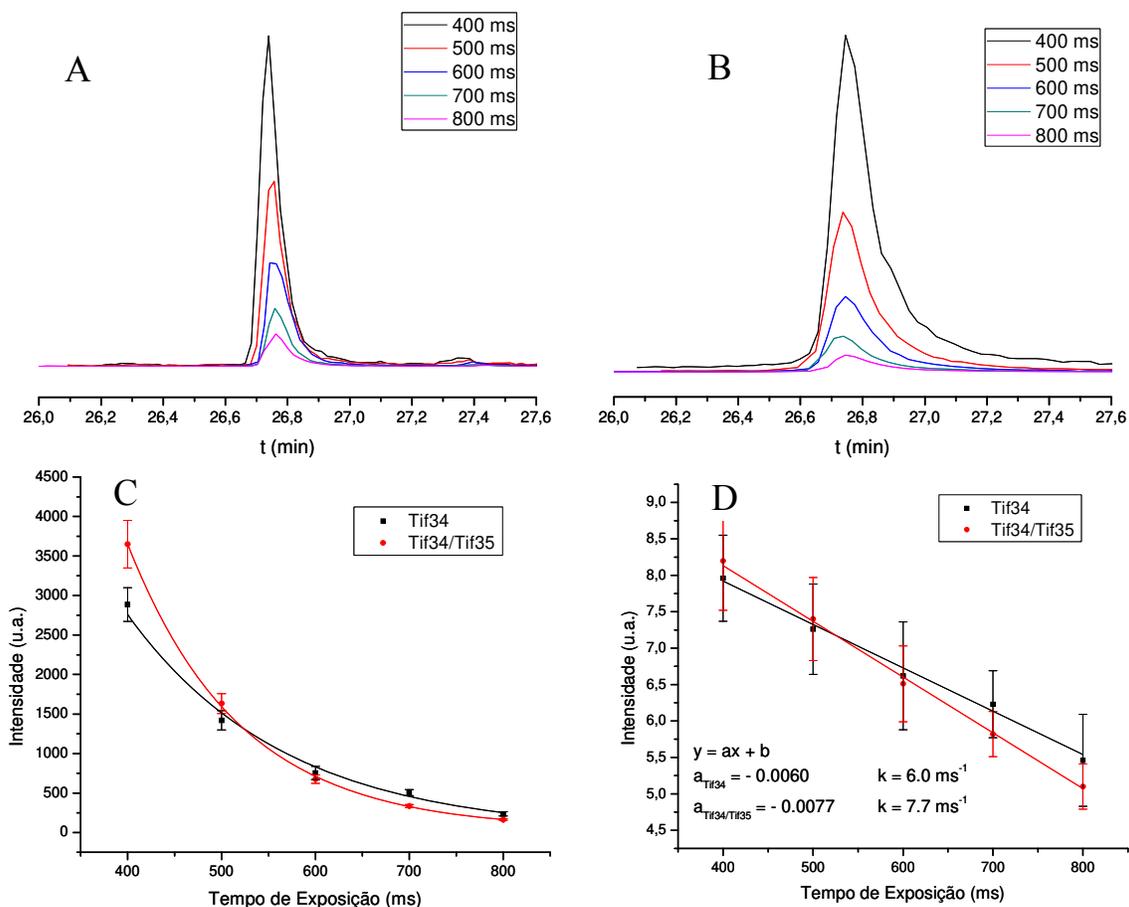
Outro pré-requisito para a utilização de *label-free* na quantificação dessas espécies está no fato de que a diminuição da intensidade dos peptídeos deve ser relacionada somente a processos oxidativos, ou seja, não deve estar relacionada a degradação da proteína. Com o intuito de provar que a estrutura de Tif34 mantinha-se intacta após os diferentes intervalos de tempo de exposição, as amostras foram analisadas por SDS-PAGE (Figura 53). É possível notar que mesmo após o maior tempo de irradiação da amostra, bandas referentes a degradação ou produtos de *cross-linking* das proteínas estão ausentes no gel. Com isso, demonstrou-se que *label-free* poderia ser utilizado na quantificação da cinética de oxidação dos peptídeos de interesse.



**Figura 53** – SDS-PAGE da proteína Tif34 irradiada por diferentes intervalos de tempo.

Como esse sistema apresenta as características necessárias, os peptídeos gerados da digestão triptica e posteriormente identificados por análise de LC-MS/MS foram quantificados pelo método *label-free*. Desse modo, a cinética de oxidação foi aproximada para pseudo-primeira ordem, onde a concentração de peptídeo era o fator limitante da reação. A intensidade do sinal referente a cada peptídeo foi utilizada para o cálculo dessa constante, sendo que uma curva do tipo dose-resposta foi obtida pela integração do sinal de cada espécie em função do tempo. O coeficiente linear dessa reta resultou nos valores que foram posteriormente comparados para a proteína isolada e proteína na presença do ligante. As Figura 54A e B apresentam os cromatogramas alinhados para o peptídeo que engloba os aminoácidos 18-32 (GSHMASMTGGQQMGR) da seqüência da HisTif34 ( $m/z$  512.5, 3+). É possível observar que em ambos os casos a intensidade desses íons decresce em função do aumento de exposição ao raio-X, fato esse esperado em virtude da formação de diferentes produtos de oxidação. A Figura 54C apresenta o gráfico dessas intensidades em função do tempo, sendo que a linearização dessa curva (Figura 54D) permite obter a cinética de oxidação para esse peptídeo nas duas condições distintas. É interessante notar que esse peptídeo

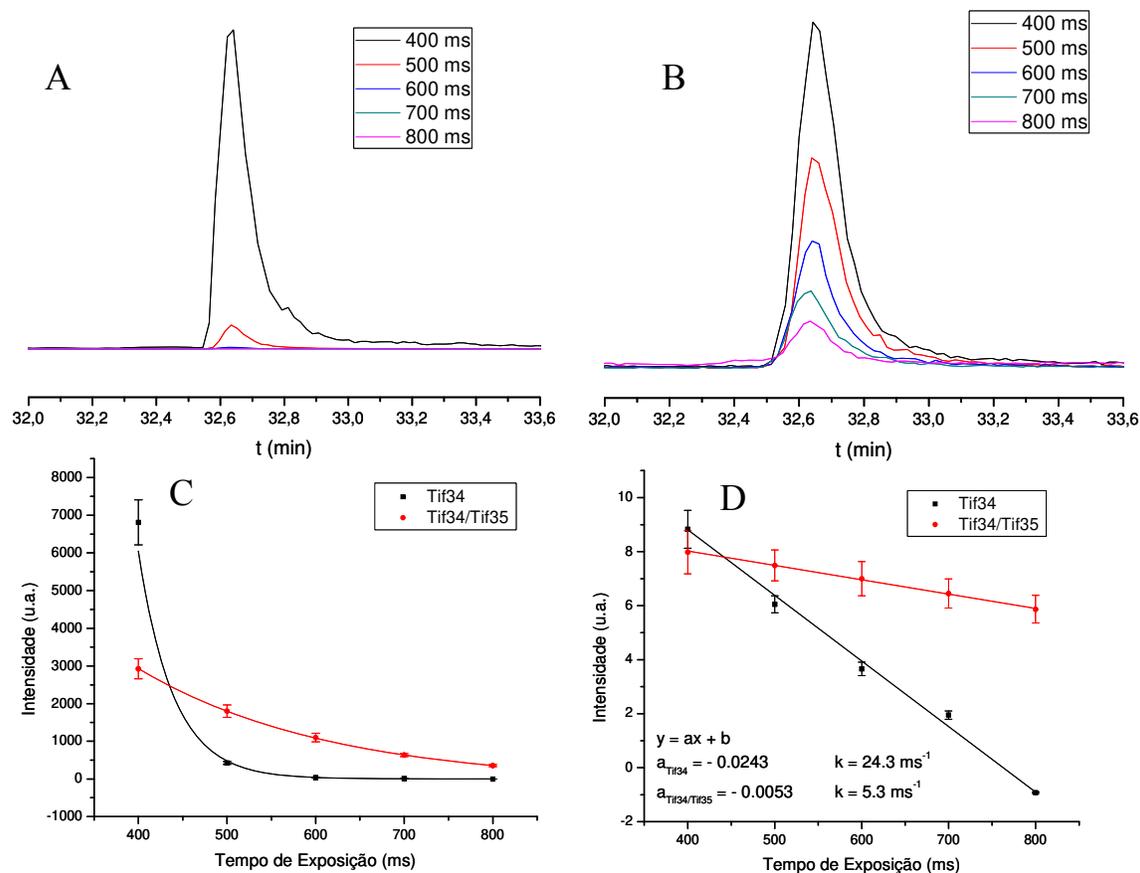
apresenta maior cinética de oxidação quando a proteína se encontra na sua forma complexada. Esse fato seria possível caso a Tif34 viesse a sofrer alguma mudança conformacional induzida por Tif35 que expusesse mais esse peptídeo ao solvente. Entretanto, conforme descrito na literatura, proteínas com estrutura tipo  *$\beta$ -propeller* não apresentam grandes mudanças conformacionais induzidas por ligantes devido à rigidez de sua estrutura.<sup>206</sup> A explicação proposta para esse fato é que no caso da proteína isolada, ela é expressa com a cauda de afinidade (*HisTag*); por outro lado, no complexo esse grupo encontra-se na proteína Tif35, deixando o N-terminal de Tif34 livre e mais suscetível à oxidação. Isso mais uma vez demonstra a sensibilidade da técnica, uma vez que pequenas alterações no ambiente do peptídeo podem ser detectadas pela alteração do valor da taxa de oxidação.



**Figura 54** – Perfil cromatográfico do íon de  $m/z$  512.5 (3+) em função do tempo de exposição para a proteína Tif34 isolada (A) e complexada com Tif35 (B). (C) Gráfico da intensidade do íon em função do tempo de exposição. (D) Linearização do gráfico anterior, de onde é possível obter as diferentes taxas de oxidação.

As Figura 55A e B apresentam os cromatogramas alinhados para o peptídeo 179-196 (IITHEGLDAATVAGWSTK) de HisTif34 ( $m/z$  624.0, 3+). Diferentemente do caso anterior, esse peptídeo apresenta perfis cromatográficos bem diferentes para a proteína isolada e complexada: no caso da proteína livre em solução, esse peptídeo é rapidamente oxidado, de modo que após 500 ms praticamente não é possível observá-lo mais no cromatograma. Isso indica que os sítios de oxidação dessa espécie estão muito expostos ao solvente. Por outro lado, quando na presença de Tif35 sua cinética

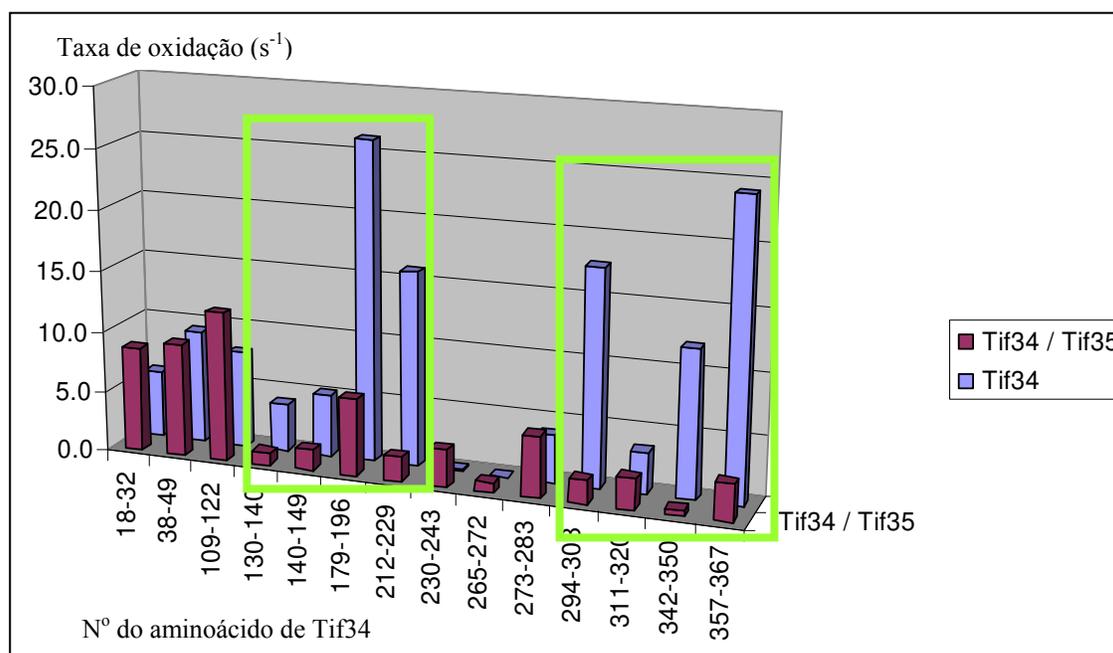
de oxidação é drasticamente reduzida (Figura 55C e D), indicando alto grau de proteção ao solvente induzida pela complexação com o ligante. Desse modo, pode-se concluir por esses dados que esse peptídeo se encontra na superfície de interação entre essas duas proteínas.



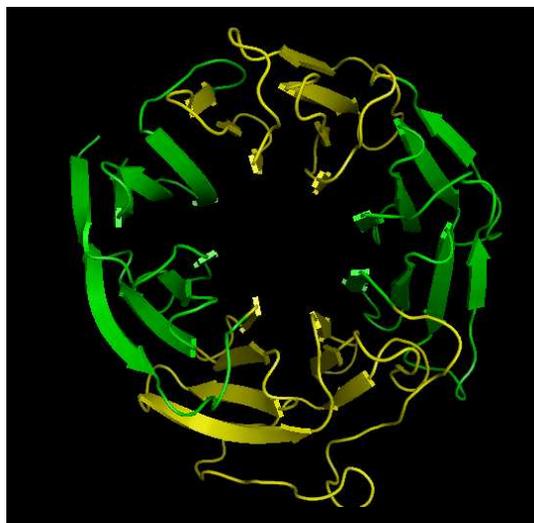
**Figura 55** - Perfil cromatográfico do íon de  $m/z$  624.0 (3+) em função do tempo de exposição para a proteína Tif34 isolada (A) e complexada com Tif35 (B). (C) Gráfico da intensidade do íon em função do tempo de exposição. (D) Linearização do gráfico anterior, de onde é possível obter as diferentes taxas de oxidação.

Os mesmos cálculos apresentados acima foram realizados para os outros peptídeos identificados pela análise de Tif34. A Figura 56 apresenta as taxas de oxidação desses peptídeos tanto na proteína isolada quanto na presença de Tif35, onde é possível notar a presença de duas regiões em Tif34 que

apresentam cinéticas de oxidação consideravelmente diminuídas mediante interação com Tif35: aminoácidos 130-196 e 294-367. Conforme descrito anteriormente, a diminuição dessa taxa está associada a proteção de oxidação pela interação com o ligante, o que significa que Tif35 interage com Tif34 nessas duas regiões. A Figura 57 apresenta o modelo de homologia de Tif34, estando as regiões protegidas destacadas na estrutura. É interessante notar que essas regiões encontram-se espacialmente opostas uma a outra na estrutura simétrica de Tif34, além de serem constituídas do mesmo *motif* estrutural: três folhas  $\beta$  seguidas de quatro folhas  $\beta$ .



**Figura 56** – Comparação da cinética de oxidação dos peptídeos de Tif34 na presença e ausência de Tif35. As regiões com maior diferença de taxas de oxidação na presença e ausência do ligante estão marcadas com o quadrado verde.



**Figura 57** – Modelo de homologia de Tif34, onde os sítios protegidos de oxidação pela ligação com Tif35 estão mostrados em amarelo.

É interessante notar a complementaridade das técnicas: os peptídeos de Tif34 que foram encontrados ligados a Tif35 pela técnica de *cross-linking* foram aqueles que apresentam alto grau de exposição ao solvente, determinado pela técnica de *footprinting*. Esse fato era esperado visto que, do mesmo modo que radicais hidroxila necessitam de acessibilidade para atacar a cadeia lateral dos aminoácidos, *cross-linkers* também necessitam que a cadeia lateral da lisina esteja exposta para que a reação possa ocorrer (Figura 58).

*Footprinting*

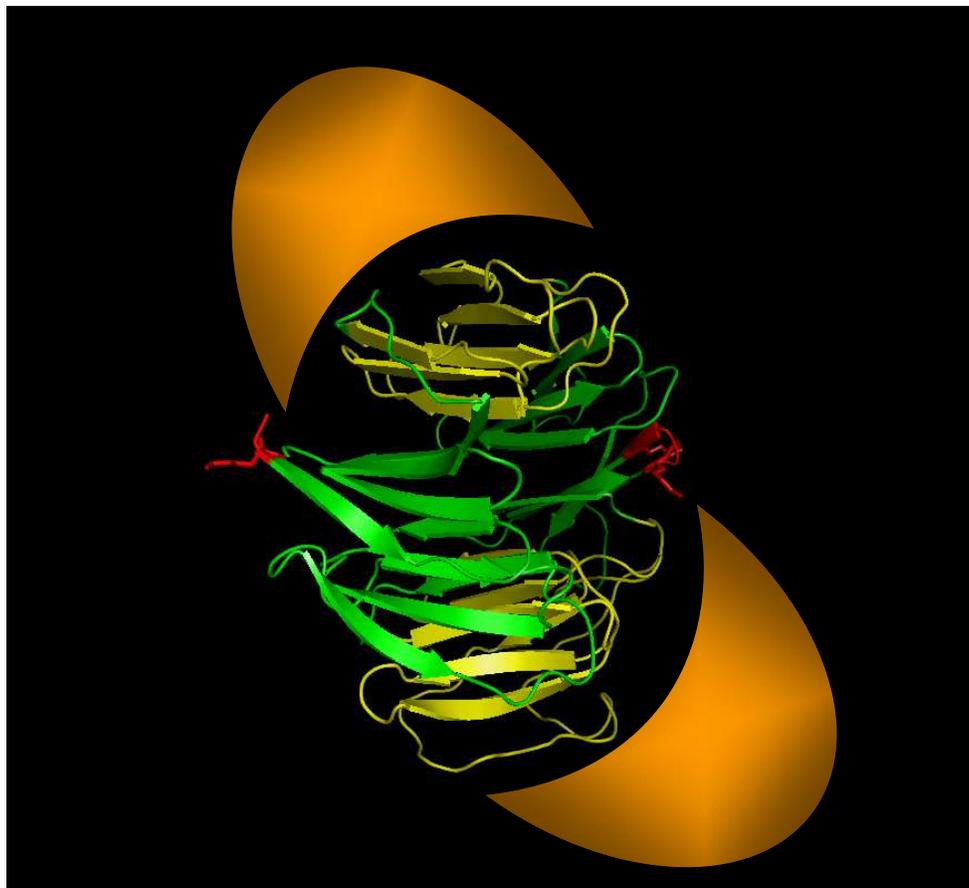
**MKAIKLTGHERPLTQVKYNKEGDLFSCSKDSSASVWYS**  
 LNERLGTLDGHTGTIWSIDVDCFTKYCVTGSADYSIKLW  
 DVSNQGCVATW**KSPVPVKRVEFSPCGNYFLAILDNVMKN**  
 PGSINIYEIERDSATHELTKVSEEPIHKITHEGLDAATVAG  
 WSTKGYIIAGHKD**GKISKYDVSNNYEYVDSIDLHEKSISD**  
 MQFSPDLTYFITSSRDTNSFLVDVSTLQVLKKYETDCPLNT  
 AVITPLKEFIILGGGQ**EAKDVTTSANEGKFEARFYHKIFEE**  
**EIGRVQGHFGPLNTVAISPQGTSYASGGEDGFIRLHHFEK**  
**SYDFKDYDVEKAAEAKEHMQEAN**

*Cross-Linking*

**MKAIKLTGHERPLTQVKYNKEGDLFSCSKDSSASVWYS**  
 LNERLGTLDGHTGTIWSIDVDCFTKYCVTGSADYSIKLW  
 DVSNQGCVATW**KSPVPVKRVEFSPCGNYFLAILDNVMKN**  
 PGSINIYEIERDSATHELTKVSEEPIHKITHEGLDAATVAG  
 WST**KGKYIIAGHKD**GKISKYDVSNNYEYVDSIDLHEKSISD  
 MQFSPDLTYFITSSRDTNSFLVDVSTLQVLKKYETDCPLNT  
 AVITPLKEFIILGGGQ**EAKDVTTSANEGKFEARFYHKIFEE**  
**EIGRVQGHFGPLNTVAISPQGTSYASGGEDGFIRLHHFEK**  
**SYDFKDYDVEKAAEAKEHMQEAN**

**Figura 58** - Regiões de Tif34 oxidadas nos experimentos de *footprinting* (em vermelho) e modificadas com DSS nos experimentos de *cross-linking* (em vermelho).

Além disso, em ambos os casos é possível notar que os dados de interação demonstraram a presença de duas regiões distintas em Tif34 responsáveis pela ligação entre as proteínas. Essas duas regiões encontram-se opostas no modelo de homologia de Tif34, sendo constituídas pelo mesmo *motif* estrutural: quatro folhas  $\beta$  seguidas de três folhas  $\beta$ . Desse modo, foi proposta a existência de dois sítios de interação entre essas proteínas, fato esse que explica os dados complementares obtidos pelas duas técnicas utilizadas (Figura 59). Esses dados reforçam a idéia de que estruturas do tipo  *$\beta$ -propeller* são excelentes suportes para múltiplas interações tipo proteína-proteína, principalmente pela fato de apresentar uma grande área superficial que favorece contatos múltiplos e simultâneos.<sup>205</sup>



**Figura 59** - Modelo de interação entre as proteínas Tif34 (modelo de homologia) e Tif35 (em laranja), baseado nos dados provenientes de experimentos de *cross-linking* e *footprinting*.

#### IV.3.4 Conclusões

O complexo Tif34/Tif35 apresenta características que dificultam consideravelmente seu estudo por técnicas de alta resolução: baixa estabilidade, tamanho relativamente grande, dificuldades para obtenção em larga escala, agregação quando muito concentrado, dentre outros. Em virtude desses fatos, CRP e RMN não foram capazes de fazer a caracterização estrutural desse composto. Com isso, as técnicas de *cross-linking* e

*footprinting* descritas nas seções anteriores mostraram-se extremamente adaptadas ao complexo: são técnicas que necessitam pequenas quantidades de amostra (de fato, em uma indução em 2 L de meio de cultura de bactéria cerca de 5% do complexo obtido após a purificação era utilizado nos experimentos), com manipulação mínima e rápida.

Não só as técnicas mostraram-se capazes de lidar com essas dificuldades, mas os resultados obtidos nos experimentos de *footprinting* e *cross-linking* mostraram-se extremamente complementares. Em ambos os casos, os dados obtidos indicaram a presença de dois sítios em Tif34 de possível interação com Tif35. Esses sítios encontram-se em regiões opostas na estrutura simétrica de Tif34, sendo compostos pelo mesmo *motif* estrutural: quatro folhas  $\beta$  seguidas de três folhas  $\beta$ . A possibilidade de dois sítios de interação em Tif34 está de acordo com sua estrutura tipo  *$\beta$ -propeller*, reconhecidamente importante para interações proteína-proteína devido à sua alta área superficial. Vale destacar ainda que esses dados estão de acordo com os previamente obtidos pela técnica de SAXS, onde foi possível notar que a interação entre essas proteínas se dá pela face lateral de Tif34 (e não pela parte superior ou inferior). Pelos dados de SAXS entretanto não foi possível definir conclusivamente a real localização dos sítios de interação bem como a presença de mais de um sítio, em virtude da natureza simétrica da estrutura de Tif34.

## V CONCLUSÕES GERAIS

Um dos principais desafios para a compreensão mais detalhada dos fenômenos biológicos está relacionado ao modo como as proteínas interagem. Inúmeros projetos de interactoma estão sendo conduzidos atualmente com esse objetivo, sendo que um grande número de interações proteína-proteína tem sido relatado na literatura. Embora o conhecimento da identidade das proteínas seja fundamental, a seqüência primária não permite a elucidação dos mecanismos químicos pelos quais ocorrem esses processos. Desse modo, a Genômica Estrutural visa a elucidação em larga escala da estrutura espacial de proteínas e suas interações. Entretanto, os métodos de caracterização de estrutura de proteínas de alta resolução não são compatíveis com as necessidades atuais, em virtude do grande número de alvos de interesse.

Em vista desse fato, essa Tese objetivou o desenvolvimento de técnicas em Espectrometria de Massas aplicadas a obtenção de dados estruturais de proteínas e suas interações. MS constitui a principal técnica de caracterização de estrutura primária de proteínas devido às suas vantagens, tais como alta sensibilidade aliada a velocidade de análise e relativa facilidade. Essas vantagens seriam indiscutivelmente atraentes quando aplicadas também a estruturas superiores de proteínas.

A primeira técnica estudada foi a de *cross-linking*, onde ligações covalentes são formadas entre a cadeia lateral de determinados aminoácidos com reagentes conhecidos como *cross-linkers*. O conhecimento do comprimento da cadeia desse reagente permite sua utilização como régua molecular que fornece restrições espaciais para a obtenção de modelos da estrutura da proteína. Dentre os vários reagentes disponíveis, o diéster de N-hidroxisuccinimida do ácido subérico (DSS) é o mais utilizado, devido à sua disponibilidade comercial e características da cadeia espaçadora. Com isso, foram desenhados peptídeos modelos que após reação com DSS e digestão com tripsina gerassem espécies que mimetizassem experimentos com proteínas. Esses peptídeos foram utilizados para a elucidação dos mecanismos de fragmentação de peptídeos contendo *cross-linkers*, uma vez que esse é o único modo possível para a atribuição inequívoca dos sítios ligados. Esses mecanismos foram propostos utilizando os dados provenientes de experimentos utilizando tanto MALDI quanto ESI. Dentre as várias propostas de fragmentação, uma se mostrou extremamente interessante, uma vez que apontou a formação de íons característicos para a presença do *cross-linker*. Esses íons característicos foram posteriormente utilizados em experimentos do tipo Varredura de Íons Precusores. A presença de três íons marcadores não só aumentou a seletividade do método, como também permitiu a elaboração de uma rotina a ser implementada em um *software* que vise a identificação automatizada de espécies ligadas.

Outro método envolvendo MS utilizado foi o *footprinting*, no qual são utilizados sondas para avaliar o grau de exposição de resíduos ao solvente. Com esse fim, foi desenvolvida no Laboratório Nacional de Luz Síncrotron uma fonte de luz que permitia a obtenção de radicais hidroxila na solução de proteína. A taxa de modificação de determinado peptídeo em dois ambientes

distintos pode ser correlacionada com alguma mudança conformacional da proteína. Para o cálculo dessa cinética de oxidação, foi proposto um método alternativo ao descrito na literatura, conhecido como *label-free*. Diferente do proposto em alguns artigos, esse método avalia todos os produtos de oxidação possíveis, como produtos de múltiplas oxidações e até mesmo produtos ainda não descritos na literatura. Desse modo, esse método permite a quantificação mais exata das taxas de oxidação dos peptídeos de interesse.

As técnicas de *cross-linking* e *footprinting* foram posteriormente aplicadas ao complexo Tif34/Tif35 de *Saccharomyces cerevisiae*, proteínas constituintes do fator de iniciação de tradução 3 (eIF3). Esse complexo, de aproximadamente 60 kDa, apresenta características que impossibilitaram a utilização de técnicas de alta resolução para a sua caracterização, tais como instabilidade e impossibilidade da obtenção de monocristais. Desse modo, *cross-linking* e *footprinting* mostraram-se extremamente adequadas para a obtenção de dados estruturais desse complexo. Além disso, os dados obtidos pelas duas técnicas mostraram-se totalmente complementares, tendo sido possível a determinação de dois sítios de interação entre essas proteínas. Esses sítios encontram-se em regiões opostas na estrutura simétrica de Tif34, além de apresentarem grande similaridade em suas características estruturais. A possibilidade de mais de um sítio de interação em Tif34 é resultante das características estruturais de proteínas com enovelamento do tipo  *$\beta$ -propeller*, o qual possui grande área superficial que permite múltiplas interações, inclusive simultaneamente.

Finalmente, a técnica de Espectrometria de Massas pôde ser aplicada nesses estudos com todas suas vantagens intrínsecas, de modo que o estudo de estrutura de proteínas e suas interações podem se beneficiar ainda mais desta técnica.

## VI BIBLIOGRAFIA

---

- <sup>1</sup> Nelson D. L.; Cox M. M. *Lehninger Principles of Biochemistry*, Fourth Edition. Cambridge, W. H. Freeman, 2004.
- <sup>2</sup> Montelione G. T.; Anderson S. *Nature Struct. Biol.*, 6, 11-12, 1999.
- <sup>3</sup> Adams M. A.; Suits M. D. L.; Zheng J.; Jia Z. *Proteomics*, 7, 2920-2932, 2007.
- <sup>4</sup> Lesk A. M. *Introduction to Protein Architecture*. New York, Oxford University Press, 2001.
- <sup>5</sup> Fersht A. *Structure and Mechanism in Protein Science*. Cambridge, W. H. Freeman, 1999.
- <sup>6</sup> Jones S.; Thornton J. M. *Proc. Natl. Acad. Sci. USA*, 93, 13-20, 1996.
- <sup>7</sup> Alberts B. *Cell* 92, 291-294, 1998.
- <sup>8</sup> Krause F. *Electrophoresis*, 27, 2759-2781, 2006.
- <sup>9</sup> Robinson C. V.; Sali A.; Baumeister W. *Nature*, 450, 973-982, 2007.
- <sup>10</sup> Kuriyan J.; Eisenberg D. *Nature*, 450, 983-990, 2007.
- <sup>11</sup> Bai H.; Ma W.; Liu S.; Lai L. *Proteins*, 70, 1323-1331, 2008.
- <sup>12</sup> Pang C. N. I.; Krycer J. R.; Lek A.; Wilkins M. R. *Proteomics*, 8, 425-434, 2008.
- <sup>13</sup> Keskin O.; Gursoy A.; Ma B.; Nussinov R. *Chem. Rev.*, 108, 1225-1244, 2008.
- <sup>14</sup> Boehr D. D.; Wright P. E. *Science*, 320, 1429-1430, 2008.
- <sup>15</sup> Archakov A. I.; Govorun V. M.; Dubanov A. V.; Ivanov Y. D.; Veselovsky A. V.; Lewi P.; Janssen P. *Proteomics*, 3, 380-391, 2003.

---

<sup>16</sup> Gavin A. C.; Bösche M.; Krause R.; Grandi P.; Marzioch M.; Bauer A.; Schultz J.; Rick J. M.; Michon A. M.; Cruciat C. M.; Remor M.; Höfert C.; Schelder M.; Brajenovic M.; Ruffner H.; Merino A.; Klein K.; Hudak M.; Dickson D.; Rudi T.; Gnau V.; Bauch A.; Bastuck S.; Huhse B.; Leutwein C.; Heurtier M. A.; Copley R. R.; Edelman A.; Querfurth E.; Rybin V.; Drewes G.; Raida M.; Bouwmeester T.; Bork P.; Seraphin B.; Kuster B.; Neubauer G.; Furga G. *Nature*, 415, 141-147, 2002.

<sup>17</sup> Vazquez A.; Flammini A.; Maritan A.; Vespignani A. *Nature Biotech.*, 21, 697-700, 2003.

<sup>18</sup> Zhang S.; Jin G.; Zhang X. S.; Chen L. *Proteomics*, 7, 2856-2869, 2007.

<sup>19</sup> Droit A.; Poirier G. G.; Hunter J. M. *J. Mol. Endocrinol.*, 34, 263-280, 2005.

<sup>20</sup> Sanchez C.; Lachaize C.; Janody F.; Bellon B.; Röder L.; Euzenat J.; Rechenmann F.; Jacq B. *Nucleic Acid Res.*, 27, 89-94, 1999.

<sup>21</sup> Gavin A. C.; Aloy P.; Grandi P.; Krause R.; Boesche M.; Marzioch M.; Rau C.; Jensen L. J.; Bastuck S.; Dumpelfeld B.; Edelman A.; Heurtier M. A.; Hoffman V.; Hoefert C.; Klein K.; Hudak M.; Michon A. M.; Schelder M.; Schirle M.; Remor M.; Rudi T.; Hooper S.; Bauer A.; Bouwmeester T.; Casari G.; Drewes G.; Neubauer G.; Rick J. M.; Kuster B.; Bork P.; Russell R. B.; Superti-Furga G. *Nature*, 440, 631-636, 2006.

<sup>22</sup> Giot L.; Bader J. S.; Brouwer C.; Chaudhuri A.; Kuang B.; Li Y.; Hao Y. L.; Ooi C. E.; Godwin B.; Vitols E.; Vijayadamodar G.; Pochart P.; Machineni H.; Welsh M.; Kong Y.; Zerhusen B.; Malcolm R.; Varrone Z.; Collis A.; Minto M.; Burgess S.; McDaniel L.; Stimpson E.; Spriggs F.; Williams J.; Neurath K.; Ioime N.; Agee M.; Voss E.; Furtak K.; Renzulli R.; Aanensen N.; Carrolla S.; Bickelhaupt E.; Lazovatsky Y.; DaSilva A.; Zhong J.; Stanyon C. A.; Finley Jr. R. L.; White K. P.; Braverman M.; Jarvie T.; Gold S.; Leach M.; Knight J.; Shimkets R. A.; McKenna M. P.; Chant J.; Rothberg J. M. *Science*, 302, 1727-1736, 2003.

<sup>23</sup> Li S.; Armstrong C. M.; Bertin N.; Ge H.; Milstein S.; Boxem M.; Vidalain P. O.; Han J. D. J.; Chesneau A.; Hao T.; Goldberg D. S.; Li N.; Martinez M.; Rual J. F.; Lamesch P.; Xu L.; Tewari M.; Wong S. L.; Zhang L. V.; Berriz G. F.; Jacotot L.; Vaglio P.; Reboul J.; Kishikawa T. H.; Li Q.; Gabel H. W.; Elewa A.; Baumgartner B.; Rose D. J.; Yu H.; Bosak S.; Sequerra R.; Fraser A.; Mango S. E.; Saxton W. M.; Strome S.; Heuvel S.; Piano F.;

- Vandenhoute J.; Sardet C.; Gerstein M.; Stamm L. D.; Gunsalus K. C.; Harper J. W.; Cusick M. E.; Roth F. P.; Hill D. E.; Vidal M. *Science*, 303, 540-543, 2004.
- <sup>24</sup> Stelzl U.; Worm U.; Lalowski M.; Haenig C.; Brembeck F. H.; Goehler H.; Stroedicke M.; Zenkner M.; Schoenherr A.; Koeppen S.; Timm J.; Mintzlaff S.; Abraham C.; Bock N.; Kietzmann S.; Goedde A.; Toksöz E.; Droege A.; Krobitsch S.; Korn B.; Birchmeier W.; Lehrach H.; Wanker E. E. *Cell*, 122, 957-968, 2005.
- <sup>25</sup> Stumpf M. P. H.; Thorne T.; Silva E.; Stewart R.; An H. J.; Lappe M.; Wiuf C. *Proc. Natl. Acad. Sci. USA*, 105, 6959-6964, 2008.
- <sup>26</sup> Sprinzak E.; Sattath S.; Margalit H. *J. Mol. Biol.*, 327, 919-923, 2003.
- <sup>27</sup> Wells J. A.; McClendon C. L. *Nature*, 450, 1001-1009, 2007.
- <sup>28</sup> Alber F.; Förster F.; Korkein D.; Topf M.; Sali A. *Annu. Rev. Biochem.*, 77, 443-477, 2008.
- <sup>29</sup> Sali A.; Glaeser R.; Earnest T.; Baumeister W. *Nature*, 422, 216-225, 2003.
- <sup>30</sup> Puglise J. D. *Nature*, 450, 1171-1172, 2007.
- <sup>31</sup> Kinter M.; Sherman N. E. *Protein Sequencing and Identification Using Tandem Mass Spectrometry*, New York, John Wiley & Sons, 2000.
- <sup>32</sup> Yamashita M.; Fenn J. B. *J. Phys. Chem.*, 84, 4451-4459, 1984.
- <sup>33</sup> Karas M.; Bachmann D.; Bahr U.; Hillenkamp F. *Int. J. Mass Spectrom. Ion Proc.*, 78, 53-68, 1987.
- <sup>34</sup> Fenn J. B.; Mann M.; Meng C. K.; Wong S. F.; Whitehouse C. F. *Science*, 246, 64-71, 1989.
- <sup>35</sup> Hunt D. F.; Zhu N. Z.; Shabanowitz J. *Rapid Commun. Mass Spectrom.*, 3, 122-124, 1989.
- <sup>36</sup> Hunt D. F.; Henderson R. A.; Shabanowitz J.; Sakaguchi K.; Michel H.; Sevilir N.; Cox A. L.; Appella E.; Engelhard V. H. *Science*, 255, 1261-1263, 1992.
- <sup>37</sup> Covey T. R.; Huang E. C.; Henion J. D. *Anal. Chem.*, 63, 1193-1200, 1991.
- <sup>38</sup> Patterson D. H.; Tarr G. E.; Regnier F. E.; Martin S. A. *Anal. Chem.*, 67, 3971-3978, 1995.
- <sup>39</sup> James P.; Quadroni M.; Carafoli E.; Gonnet G. *Proc. Natl. Acad. Sci. USA*, 195, 58-64, 1993.

- 
- <sup>40</sup> Yates III J. R.; Speicher S.; Griffin P. R.; Hunkapiller T. *Anal. Biochemistry*, 214, 397-408, 1993.
- <sup>41</sup> Henzel W. J.; Billeci T. M.; Stults J. T.; Wong S. C.; Grimley C.; Watanabe C. *Proc. Natl. Acad. Sci. USA*, 90, 5011-5015, 1993.
- <sup>42</sup> Eng J. K.; McCormack A. L.; Yates III J. R. *J. Am. Soc. Mass Spectrom.*, 5, 976-989, 1994.
- <sup>43</sup> Pappin D. J. C.; Hojruo P.; Bleasby A. J. *Curr. Biol.*, 3, 327-332, 1993.
- <sup>44</sup> Bieman K.; Scoble H. A. *Science*, 237, 992-998, 1987.
- <sup>45</sup> Oda Y.; Huang K.; Cross F. R.; Cowburn D.; Chait B. T. *Proc. Natl. Acad. Sci. USA*, 96, 6591-6596, 1999.
- <sup>46</sup> Gygi S. P.; Rist B.; Gerber S. A.; Turecek F.; Gelb M. H.; Aebersold R. *Nature Biotechnol.*, 17, 994-999, 1999.
- <sup>47</sup> Aebersold R.; Goodlett D. R. *Chem. Rev.*, 101, 269-295, 2001.
- <sup>48</sup> Aebersold R.; Mann M. *Nature*, 422, 198-207, 2003.
- <sup>49</sup> Patterson S. D.; Aebersold R. *Nature Genet.*, 33, 311-323, 2003.
- <sup>50</sup> Steen H.; Mann M. *Nature Rev. Mol. Cell Biol.*, 5, 699-711, 2004.
- <sup>51</sup> Yates III J. R. *Annu. Rev. Biophys. Biomol. Struct.*, 33, 297-316, 2004.
- <sup>52</sup> Domon B.; Aebersold R. *Science*, 312, 212-218, 2006.
- <sup>53</sup> Cravatt B. F.; Simon G. M.; Yates III J. R.; *Nature*, 450, 991-1000, 2007.
- <sup>54</sup> Gingras A. C.; Gstaiger M.; Raught B.; Aebersold R. *Nature Rev. Mol. Cell Biol.*, 8, 645-654, 2007.
- <sup>55</sup> Berggard T.; Linse S.; James P. *Proteomics*, 7, 2833-2842, 2007.
- <sup>56</sup> Zhou M.; Veenstra T. D. *Proteomics*, 7, 2688-2697, 2007.
- <sup>57</sup> Wodak S. J.; Pu S.; Vlasblom J.; Séraphin B. *Mol. Cell. Proteomics*, 8, 3-18, 2009.
- <sup>58</sup> Benesch J. L. P.; Ruotolo B. T.; Simmons D. A.; Robinson C. V. *Chem. Rev.*, 107, 3544-3567, 2007.
- <sup>59</sup> Loo J. A. *Mass Spectrom. Rev.*, 16, 1-23, 1997.
- <sup>60</sup> Heck A. J. R.; van den Heuvel R. H. H. *Mass. Spectrom. Rev.*, 23, 368-389, 2004.
- <sup>61</sup> Ganem B.; Li Y. T.; Henion J. D. *J. Am. Chem. Soc.*, 113, 7818-7819, 1991.
- <sup>62</sup> Juhasz P.; Biemann K. *Proc. Natl. Acad. Sci. USA*, 91, 4333-4337, 270.

- 
- <sup>63</sup> Heuvel R. H. H.; Duijn E.; Mazon H.; Synowsky S. A.; Lorenzen K.; Versluis C.; Brouns S. J. J.; Langridge D.; Oost J.; Hoyes J.; Heck A. J. R. *Anal. Chem.*, 78, 7473-7483, 2006.
- <sup>64</sup> Lorenzen K.; Olia A. S.; Uetrecht C.; Cingolani G.; Heck A. J. R. *J. Mol. Biol.*, 379, 385-396, 2008.
- <sup>65</sup> Uetrecht C.; Versluis C.; Watts N. R.; Roos W. H.; Wuite G. J. L.; Wingfield P. T.; Stevens A. C.; Heck A. J. R. *Proc. Natl. Acad. Sci. USA*, 105, 9216-9220, 2008.
- <sup>66</sup> Rose R. J.; Verger D.; Daviter T.; Remaut H.; Paci E.; Waksman G.; Ashcroft A. E.; Radford S. E. *Proc. Natl. Acad. Sci. USA*, 105, 12873-12878, 2008.
- <sup>67</sup> Mazon H.; Gábor K.; Leys D.; Heck A. J. R.; Oost J.; Heuvel R. H. H. *J. Biol. Chem.*, 282, 11281-11290, 2007.
- <sup>68</sup> Pinske M. W. H.; Maier C. S.; Kim J. I.; Oh B. H.; Heck A. J. R. *J. Mass Spectrom.*, 38, 315-320, 2003.
- <sup>69</sup> Duijn E.; Bakkes P. J.; Heeren R. M. A.; Heuvel R. H. H.; Heerikhuizen H.; Vies S. M.; Heck A. J. R. *Nature Meth.*, 2, 371-376, 2005.
- <sup>70</sup> Zhang Z.; Zheng Y.; Mazon H.; Milgron E.; Kitagawa N.; Trier E. K.; Heck A. J. R.; Kane P. M.; Wilkens S. *J. Biol. Chem.*, 283, 35983-35995, 2008.
- <sup>71</sup> Synowsky S. A.; Heuvel R. H. H.; Mohammed S.; Pijnappel W. W. M. P.; Heck A. J. R. *Mol. Cell. Proteomics*, 5, 1581-1592, 2006.
- <sup>72</sup> Hoofnagle A. N.; Resing K. A.; Ahn N. G. *Annu. Rev. Biophys. Biomol. Struct.*, 32, 1-25, 2003.
- <sup>73</sup> Yan X.; Watson J.; Ho P. S.; Deinzer M. L. *Mol. Cell. Proteomics*, 3, 10-23, 2004.
- <sup>74</sup> Busenlehner L. S.; Armstrong R. N. *Arch. Biochemistry Biophys.*, 433, 34-46, 2005.
- <sup>75</sup> Katta V.; Chait B. T. *Rapid Commun. Mass Spectrom.*, 5, 214-217, 1991.
- <sup>76</sup> McLafferty F. W.; Guan Z.; Haupts U.; Wood T. D.; Kelleher N. L. *J. Am. Chem. Soc.*, 120, 4732-4740, 1998.
- <sup>77</sup> Carulla N.; Caddy G. L.; Hall D. R.; Zurdo J.; Gairí M.; Feliz M.; Giralt E.; Robinson C. V.; Dobson C. M. *Nature*, 436, 554-558, 2005.
- <sup>78</sup> Wang F.; Li W.; Emmett M. R.; Hendrickson C. L.; Marshal A. G.; Zhang Y. L.; Wu L.; Zhang Z. Y. *Biochemistry*, 37, 15289-15299, 1998.

- 
- <sup>79</sup> Eyles S. J.; Speir J. P.; Kruppa G. H.; Gierasch L. M.; Kaltashov I. A. *J. Am. Chem. Soc.*, 122, 495-500, 2000.
- <sup>80</sup> Mandell J. G.; Falick A. M.; Komives E. A. *Proc. Natl. Acad. Sci. USA*, 95, 14705-14710, 1998.
- <sup>81</sup> Hyzer C. S. H.; Counterman A. E.; Clemmer D. E. *Chem. Rev.*, 99, 3037-3080, 1999.
- <sup>82</sup> Jarrold M. F. *Annu. Rev. Phys. Chem.*, 51, 179-207, 2000.
- <sup>83</sup> McLean J. A.; Ruotolo B. T.; Gillig K. J.; Russell D. H. *Int. J. Mass Spectrom.*, 240, 301-315, 2005.
- <sup>84</sup> Ruotolo B. T.; Robinson C. V. *Curr. Opin. Chem. Biol.*, 10, 402-408, 2006.
- <sup>85</sup> Kanu A. B.; Dwivedi P.; Tam M.; Matz L.; Hill Jr. H. H. *J. Mass Spectrom.*, 43, 1-22, 2008.
- <sup>86</sup> Shelimov K. B.; Clemmer D. E.; Hudgins R. R.; Jarrold M. F. *J. Am. Chem. Soc.*, 119, 2240-2248, 1997.
- <sup>87</sup> Shelimov K. B.; Jarrold M. F. *J. Am. Chem. Soc.*, 119, 2987-2994, 1997.
- <sup>88</sup> Koeniger S. L.; Merenbloom S. I.; Clemmer D. E. *J. Phys. Chem. B*, 110, 7017-7021, 2006.
- <sup>89</sup> Shvartsburg A. A.; Li F.; Tang K.; Smith R. D. *Anal. Chem.*, 78, 3304-3315, 2006.
- <sup>90</sup> Jarrold M. F. *Acc. Chem. Res.*, 32, 360-367, 1999.
- <sup>91</sup> Ruotolo B. T.; Giles K.; Campuzano I.; Sandercock A. M.; Bateman R. H.; Robinson C. V. *Science*, 310, 1658-1661, 2005.
- <sup>92</sup> Bernstein S. L.; Wytttenbach T.; Baumketner A.; Shea J. E.; Bitan G.; Teplow D. B.; Mowers M. T. *J. Am. Chem. Soc.*, 127, 2075-2084, 2005.
- <sup>93</sup> Loo J. L.; Berhane B.; Kaddis C. S.; Wooding K. M.; Xie Y.; Kaufman S. L.; Chernushevich I. V. *J. Am. Soc. Mass Spectrom.*, 16, 998-1008, 2005.
- <sup>94</sup> Wytttenbach T.; Helden G.; Bowers M. T. *J. Am. Chem. Soc.*, 118, 8355-8364, 1996.
- <sup>95</sup> Bernstein S. L.; Liu D.; Wytttenbach T.; Bowers M. T.; Lee J. C.; Gray H. B.; Winkler J. R. *J. Am. Soc. Mass Spectrom.*, 15, 1435-1443, 2004.
- <sup>96</sup> Means G. E.; Feeney R. E. *Bioconjugate Chem.*, 1, 2-12, 1990.
- <sup>97</sup> Wong S. S. *Chemistry of Protein Conjugation and Cross-Linking*. CRC Press, 1991.
- <sup>98</sup> Hermanson G. T. *Bioconjugate Techniques*. 2<sup>nd</sup> Edition, Elsevier, 2008.

- 
- <sup>99</sup> Steen H.; Jensen O. N. *Mass Spectrom. Rev.*, 21, 163-182, 2002.
- <sup>100</sup> Staros J. V. *Biochemistry*, 21, 3950-3955, 1982.
- <sup>101</sup> Young M. M.; Tang N.; Hempel J. C.; Oshiro C. M.; Taylor E. W.; Kuntz I. D.; Gibson B. W.; Dollinger G. *Proc. Natl. Acad. Sci. USA*, 97, 5802-5806, 2000.
- <sup>102</sup> Back J. W.; Jong L.; Muijsers A. O.; Koster C. G. *J. Mol. Biol.*, 331, 303-313, 2003.
- <sup>103</sup> Sinz A. *Mass Spectrom. Rev.* 25, 663-682, 2006.
- <sup>104</sup> Hung B. X.; Kim H. Y.; Dass C. *J. Am. Soc. Mass Spectrom.*, 15, 1237-1247, 2004.
- <sup>105</sup> Dihazi G. H.; Sinz A. *Rapid Commun. Mass Spectrom.* 17, 2005-2014, 2003.
- <sup>106</sup> Lee Y. J.; Lackner L. L.; Nunnari J. M.; Phinney B. S. *J. Proteome Res.* 6, 3908-3917, 2007.
- <sup>107</sup> Nielsen T.; Thaysen-Andersen M.; Larsen N.; Jorgensen F. S.; Houen G.; Hojrup P. *Int. J. Mass Spectrom.* 268, 217-226, 2007.
- <sup>108</sup> Rappsilber J.; Siniosoglou S.; Hurt E. C.; Mann M. *Anal. Chem.* 72, 267-275, 2000.
- <sup>109</sup> El-Shafey A.; Tolic N.; Young M. M.; Sale K.; Smith R. D.; Kery V. *Protein Sci.* 15, 429-440, 2006.
- <sup>110</sup> Kitatsuji C.; Kuroguchi M.; Nishimura S. I.; Ishimori K.; Wakasugi K. *J. Mol. Biol.*, 368, 150-160, 2007.
- <sup>111</sup> Carven G. J.; Stern L. J. *Biochemistry* 44, 13625-13637, 2005.
- <sup>112</sup> Bhat S.; Sorci-Thomas M. G.; Alexander E. T.; Samuel M. P.; Thomas M. J. *J. Biol. Chem.* 280, 33015-33025, 2005.
- <sup>113</sup> Huang B. X.; Kim H. Y. *Mol. Cell. Proteomics* 5, 1045-1053, 2006.
- <sup>114</sup> Gertz M.; Seelert H.; Dencher N. A.; Poetsch A. *Biochim. Biophys. Acta* 1774, 566-574, 2007.
- <sup>115</sup> Suchanek M.; Radzikowska A.; Thiele C. *Nat. Methods* 2, 261-268, 2005.
- <sup>116</sup> Taverner T.; Hall N. E.; O'Hair R. A. J.; Simpson R. J. *J. Biol. Chem.* 277, 46487-46492, 2002.
- <sup>117</sup> Chang Z.; Kuchar J.; Hausinger R. P. *J. Biol. Chem.* 279, 15305-15313, 2004.
- <sup>118</sup> Back J. W.; Sanz M. A.; Jong L.; Koning L. J.; Nijtmans L. G. J.; Koster C. G.; Grivell L. A.; Spek H.; Muijsers A. O. *Protein Sci.* 11, 2471-2478, 2002.

- 
- <sup>119</sup> Maiolica A.; Cittaro D.; Borsotti D.; Sennels L.; Ciferri C.; Tarricone C.; Musacchio A.; Rappsilber J. *Mol. Cell. Proteomics* 6, 2200-2211, 2007.
- <sup>120</sup> Ahrman E.; Lambert W.; Aquilina J. A.; Robinson C. V.; Emanuelsson C. S. *Protein Sci.* 16, 1464-1478, 2007.
- <sup>121</sup> Azim-Zadeh O.; Hillebrecht A.; Linne U.; Marahiel M. A.; Klebe G.; Lingelbach K.; Nyalwidhe J. *J. Biol. Chem.* 282, 21609-21617, 2007.
- <sup>122</sup> Pimenova T.; Nazabal A.; Roschitzki B.; Seebacher J.; Rinner O.; Zenobi R. *J. Mass Spectrom.* 43, 185-195, 2008.
- <sup>123</sup> Gabant G.; Augier J.; Armengaud J. *J. Mass Spectrom.* 43, 360-370, 2008.
- <sup>124</sup> Baker D. L.; Seyfried N. T.; Li H.; Orlando R.; Terns R. M.; Terns M. P. *Biochemistry* 47, 1500–1510, 2008.
- <sup>125</sup> Maleknia S. D.; Ralston C. Y.; Brenowitz M. D.; Downard K. M.; Chance M. R. *Anal. Biochemistry*, 289, 103-115, 2001.
- <sup>126</sup> Maleknia S. D.; Downard K. *Mass Spectrom. Rev.*, 20, 388-401, 2001.
- <sup>127</sup> Guan J. Q.; Almo S. C.; Chance M. R. *Acc. Chem. Res.*, 37, 221-229, 2004.
- <sup>128</sup> Guan J. Q.; Chance M. R. *Trends Biochem. Sci.*, 30, 583-592, 2005.
- <sup>129</sup> Takamoto K.; Chance M. R. *Annu. Rev. Biophys. Biomol. Struct.*, 35, 251-275, 2006.
- <sup>130</sup> Xu G.; Chance M. R. *Chem. Rev.*, 107, 3514-3543, 2007.
- <sup>131</sup> Jay D. G. *J. Biol. Chem.*, 259, 15572-15578, 1984.
- <sup>132</sup> Hanai R.; Wang J. C. *Proc. Natl. Acad. Sci. USA*, 91, 11904-11908, 1994.
- <sup>133</sup> Heyduk E.; Heyduk T. *Biochemistry*, 33, 9643-9650, 1994.
- <sup>134</sup> Garrison W. M. *Chem. Rev.*, 87, 381-398, 1987.
- <sup>135</sup> Maleknia S. D.; Brenowitz M.; Chance M. R. *Anal. Chem.*, 71, 3965-3973, 1999.
- <sup>136</sup> Xu G.; Takamoto K.; Chance M. R. *Anal. Chem.*, 75, 6995-7007, 2003.
- <sup>137</sup> Xu G.; Chance M. R. *Anal. Chem.*, 76, 1213-1221, 2004.
- <sup>138</sup> Xu G.; Chance M. R. *Anal. Chem.*, 77, 2437-2449, 2005.
- <sup>139</sup> Xu G.; Chance M. R. *Anal. Chem.*, 77, 4549-4555, 2005.
- <sup>140</sup> Aye T. T.; Low T. Y.; Sze S. K. *Anal. Chem.*, 77, 5814-5822, 2005.
- <sup>141</sup> Sharp J. S.; Becker J. M.; Hettich R. L. *Anal. Chem.*, 76, 672-683, 2004.
- <sup>142</sup> Tong X.; Wren J. C.; Konermann L. *Anal. Chem.*, 80, 2222-2231, 2008.

- 
- <sup>143</sup> McClintock C.; Kertesz V.; Hettich R. L. *Anal. Chem.*, 80, 3304-3317, 2008.
- <sup>144</sup> Kamal J. K. A.; Chance M. R. *Protein Sci.*, 17, 79-94, 2008.
- <sup>145</sup> Zheng X.; Wintrode P. L.; Chance M. R. *Structure*, 16, 38-51, 2008.
- <sup>146</sup> Maleknia S. D.; Downard K. M. *Eur. J. Biochem.*, 268, 5578-5588, 2001.
- <sup>147</sup> Chance M. R. *Biochem. Biophys. Res. Commun.*, 287, 614-621, 2001.
- <sup>148</sup> Guan J. Q.; Almo S. C.; Reisler E.; Chance M. R. *Biochemistry*, 42, 11992-12000, 2003.
- <sup>149</sup> Kiselar J. G.; Janmey P. A.; Almo S. C.; Chance M. R. *Proc. Natl. Acad. Sci. USA*, 100, 3942-3947, 2003.
- <sup>150</sup> Gupta S.; Mangels W. F.; McGrath W. J.; Perek J. L.; Lee D. W.; Takamoto K.; Chance M. R. *Mol. Cell. Proteomics*, 3, 950-959, 2004.
- <sup>151</sup> Xu G.; Liu R.; Zak O.; Aisen P.; Chance M. R. *Mol. Cell. Proteomics*, 4, 1959-1967, 2005.
- <sup>152</sup> Liu R.; Guan J. Q.; Zak O.; Aisen P.; Chance M. R. *Biochemistry*, 42, 12447-12454, 2003.
- <sup>153</sup> Guan J. Q.; Vorobiev S.; Almo S. C.; Chance M. R. *Biochemistry*, 41, 5765-5775, 2002.
- <sup>154</sup> Gupta S.; Cheng H.; Mollah A. K. M. M.; Jamison E.; Morris S.; Chance M. R.; Khrapunov S.; Brenowitz M. *Biochemistry*, 46, 9886-9898, 2007.
- <sup>155</sup> Rashidzadeh H.; Khrapunov S.; Chance M. R.; Brenowitz M. *Biochemistry*, 42, 3655-3665, 2003.
- <sup>156</sup> Gupta S.; Sullivan M.; Toomey J.; Kiselar J.; Chance M. R. *J. Synchrotron Rad.*, 14, 233-243, 2007.
- <sup>157</sup> Sonenberg N.; Hershey J. W. B.; Mathews M. B. *Translational Control of Gene Expression*, New York, Cold Spring Harbor Laboratory Press, 2000.
- <sup>158</sup> Valásek L.; Phan L.; Schoenfeld L. W.; Valaskova V.; Hinnebusch A. G. *EMBO J.*, 20, 891-904, 2001.
- <sup>159</sup> Valásek L.; Nielsen K. H.; Hinnebusch A. G. *EMBO J.*, 21, 5886-5898, 2002.
- <sup>160</sup> Costa C. C. Caracterização da proteína Tif34p e do sub-complexo Tif34p/Tif35p do fator de tradução eIF3 de *Saccharomyces cerevisiae*. Tese de Mestrado, Instituto de Biologia – UNICAMP, 2004.

- 
- <sup>161</sup> Naranda T.; Kainuma M.; MacMillan S.; Hershey J. W. B. *Mol. Cell Biol.*, 17, 145-153, 1997.
- <sup>162</sup> Wilm M.; Shevchenko A.; Houthaeve T.; Breit S.; Schweigerer L.; Fotsis T.; Mann M. *Nature*, 379, 466-469, 1996.
- <sup>163</sup> Muller D. R.; Schindler P.; Towbin H.; Wirth U.; Voshol H.; Hoving S.; Steinmertz M. *O. Anal. Chem.*, 73, 1927-1934, 2001.
- <sup>164</sup> Maiolica A.; Cittaro D.; Borsotti D.; Sennels L.; Ciferri C.; Tarricone C.; Musacchio A.; Rappsilber J. *Mol. Cell. Proteomics*, 6, 2200-2211, 2007.
- <sup>165</sup> Petrotchenko E. V.; Olkhovik V. K.; Borchers C. H. *Mol. Cell. Proteomics*, 4, 1167-1179, 2005.
- <sup>166</sup> El-Shafey A.; Tolic N.; Young M. M.; Sale K.; Smith R. D.; Kery V. *Protein Sci.*, 15, 429-440, 2006.
- <sup>167</sup> Gao Q.; Doneanu C. E.; Shaffer S. A.; Adman E. T.; Goodlett D. R.; Nelson S. D. *J. Biol. Chem.*, 281, 20404-20417, 2006.
- <sup>168</sup> Chu F.; Mahrus S.; Craik C. S.; Burlingame A. L. *J. Am. Chem. Soc.*, 128, 10362-10363, 2006.
- <sup>169</sup> Sinz A. *Angew. Chem. Int. Ed.*, 46, 660-662, 2007.
- <sup>170</sup> Davidson W. S.; Hilliard G. M. *J. Biol. Chem.*, 278, 27199-27207, 2003.
- <sup>171</sup> Kasper P. T.; Back J. W.; Vitale M.; Hartog A. F.; Roseboom W.; Koning L. J.; Maarseveen J. H.; Muijsers A. O.; Koster C. G.; Jong L. *ChemBioChem*, 8, 1281-1292, 2007.
- <sup>172</sup> Kruppa G. H.; Schoeniger J.; Young M. M. *Rapid Commun. Mass Spectrom.*, 17, 155-162, 2003.
- <sup>173</sup> Jacobsen R. B.; Sale K. L.; Ayson M. J.; Novak P.; Hong J.; Lane P.; Wood N. L.; Kruppa G. H.; Young M. M.; Schoeniger J. S. *Protein Sci.*, 15, 1303-1317, 2006.
- <sup>174</sup> Fabris D.; Chaudhari P.; Hagan N.; Turner K. *Eur. J. Mass Spectrom.*, 13, 29-33, 2007.
- <sup>175</sup> Schilling B.; Row R. H.; Gibson B. W.; Guo X.; Young M. M. *J. Am. Soc. Mass Spectrom.*, 14, 834-850, 2003.
- <sup>176</sup> Rinner O.; Seebacher J.; Walzthoeni T.; Mueller L.; Beck M.; Schmidt A.; Mueller M.; Aebersold R. *Nature Meth.*, 5, 315-318, 2008.

- 
- <sup>177</sup> Leavell M. D.; Novak P.; Behrens C. R.; Schoeniger J. S.; Kruppa G. H. *J. Am. Soc. Mass Spectrom.*, 15, 1604-1611, 2004.
- <sup>178</sup> Kalkhof S.; Sinz A. *Anal. Bioanal. Chem.*, 392, 305-312, 2008.
- <sup>179</sup> Mädler S.; Bich C.; Touboul D.; Zenobi R. *J. Mass Spectrom.*, *In Press* (DOI: 10.1002/jms.1544).
- <sup>180</sup> Chu F.; Shan S.; Moustakas D. T.; Alber F.; Egea P. F.; Stroud R. M.; Walter P.; Burlingame A. L. *Proc. Natl. Acad. Sci. USA*, 101, 16454-16459, 2004.
- <sup>181</sup> Novak P.; Haskins W. E.; Ayson M. J.; Jacobsen R. B.; Schoeniger J. S.; Leavell M. D.; Young M. M.; Kruppa G. H. *Anal. Chem.*, 77, 5101-5106, 2005.
- <sup>182</sup> Silva R. A. G.; Hilliard G. M.; Li L.; Segrest J. P.; Davidson W. S. *Biochemistry*, 44, 8600-8607, 2005.
- <sup>183</sup> Gaucher S. P.; Hadi M. Z.; Young M. M. *J. Am. Soc. Mass Spectrom.*, 17, 395-405, 2006.
- <sup>184</sup> Barlow C. K.; O'Hair R. A. J. *J. Mass Spectrom.*, 43, 1301-1319, 2008.
- <sup>185</sup> Iglesias A. H.; Santos L. F. A.; Gozzo F. C. *J. Am. Soc. Mass Spectrom.*, *In Press*, (doi:10.1016/j.jasms.2008.11.009).
- <sup>186</sup> Roepstorff P.; Fohlman J. *Biomed. Mass Spectrom.*, 11, 601, 1984.
- <sup>187</sup> Biemann K. *Methods Enzymol.*, 193, 886-887, 1990.
- <sup>188</sup> Huddleston M. J.; Annan R. S.; Bean M. F.; Carr S. A. *J. Am. Soc. Mass Spectrom.*, 4, 710-717, 1993.
- <sup>189</sup> Neubauer G.; Mann M. *Anal. Chem.*, 71, 235-242, 1999.
- <sup>190</sup> Williamson B. L.; Marchese J.; Morrice N. A. *Mol. Cell. Proteomics*, 5, 337-346, 2006.
- <sup>191</sup> Bolgar M. S.; Gaskell S. J. *Anal. Chem.*, 68, 2325-2330, 1996.
- <sup>192</sup> Carr S. A.; Huddleston M. J.; Bean M. F. *Protein Sci.*, 2, 183-196, 1993.
- <sup>193</sup> Huddleston M. J.; Bean M. F.; Carr S. A. *Anal. Chem.*, 65, 877-884, 1993.
- <sup>194</sup> Haynes P. A.; Aebersold R. *Anal. Chem.*, 72, 5402-5410, 2000.
- <sup>195</sup> Bateman R. H.; Carruthers R.; Hoyes J. B.; Jones C.; Langridge J. I.; Millar A.; Vissers J. P. C. *J. Am. Soc. Mass Spectrom.*, 13, 792-803, 2002.
- <sup>196</sup> Niggeweg R.; Köcher T.; Gentzel M.; Buscaino A.; Taipale M.; Akhtar A.; Wilm M. *Proteomics*, 6, 41-53, 2006.

- 
- <sup>197</sup> Steen H.; Kuster B.; Fernandez M.; Pandey A.; Mann M. *Anal. Chem.*, 73, 1440-1448, 2001.
- <sup>198</sup> Paradela A.; Albar J. P. *J. Proteome Res.*, 7, 1809-1818, 2008.
- <sup>199</sup> Rappsilber J.; Friesen W. J.; Paushkin S.; Dreyfuss G.; Mann M. *Anal. Chem.*, 75, 3107-3114, 2003.
- <sup>200</sup> Frolov A.; Hoffmann P.; Hoffmann R. *J. Mass Spectrom.*, 41, 1459-1469, 2006.
- <sup>201</sup> Sclavi B.; Woodson S A.; Sullivan M.; Chance M. R.; Brenowitz M. *J. Mol. Biol.*, 266, 144-159, 1997.
- <sup>202</sup> Bondarenko P. V.; Chelius D.; Shaler T. A. *Anal. Chem.*, 74, 4741-4749, 2002.
- <sup>203</sup> Chelius D.; Bondarenko P. V. *J. Proteome Res.*, 1, 317-323, 2002.
- <sup>204</sup> Wang W.; Zhou H.; Lin H.; Roy S.; Shaler T. A.; Hill L. R.; Norton S.; Kumar P.; Anderle M.; Becker C. H. *Anal. Chem.*, 75, 4818-4826, 2003.
- <sup>205</sup> Li D.; Roberts R. *Cell. Mol. Life Sci.*, 58, 2085-2097, 2001.
- <sup>206</sup> Wall M. A.; Coleman D. E.; Lee E.; Iniguez-Lluhi J A.; Posner B. A.; Gilman A. G.; Sprang S. R. *Cell*, 83, 1047-1058, 1995.