



**UNIVERSIDADE ESTADUAL PAULISTA – UNESP
INSTITUTO DE QUÍMICA – ARARAQUARA
DEPARTAMENTO DE QUÍMICA ORGÂNICA**

ALAN CESAR PILON

***DESENVOLVIMENTO DE MÉTODOS ANALÍTICOS DE DESREPLICAÇÃO POR
RMN, EM E ANÁLISE MULTIVARIADA DO PERFIL METABOLÔMICO DE
ESPÉCIES DE SOLANACEAE COM POTENCIAL BIOLÓGICO.***

Arararaquara – São Paulo
2015



**UNIVERSIDADE ESTADUAL PAULISTA – UNESP
INSTITUTO DE QUÍMICA – ARARAQUARA
DEPARTAMENTO DE QUÍMICA ORGÂNICA**

ALAN CESAR PILON

***DESENVOLVIMENTO DE MÉTODOS ANALÍTICOS DE DESREPLICAÇÃO POR
RMN, EM E ANÁLISE MULTIVARIADA DO PERFIL METABOLÔMICO DE
ESPÉCIES DE SOLANACEAE COM POTENCIAL BIOLÓGICO.***

Tese apresentada ao Instituto de
Química, Universidade Estadual Paulista,
como parte dos requisitos para obtenção do
título de **Doutor em Química**

ORIENTADOR: Prof. Dr. Ian Castro-Gamboa

Araraquara – São Paulo
2015

ALAN CESAR PILON

Tese de Doutorado apresentada ao
Instituto de Química, Universidade
Estadual Paulista, como parte dos
requisitos para obtenção do título de
Doutor em Química.

Araraquara, 15 de maio de 2015.

BANCA EXAMINADORA



Prof. Dr. IAN CASTRO-GAMBOA (Orientador)
Instituto de Química / UNESP / Araraquara – SP



Dr. LUIZ ALBERTO COLNAGO
Embrapa Instrumentação Agropecuária / EMBRAPA / São Carlos – SP



Prof. Dr. RODRIGO RAMOS CATHARINO
Faculdade de Ciências Médicas / UNICAMP / Campinas – SP



Dr^a. PATRÍCIA VERARDI ABDELNUR
Embrapa Instrumentação Agropecuária / EMBRAPA / Brasília - DF



Prof^a. Dr^a. DENISE BRENTAN DA SILVA
Universidade Federal do Mato Grosso do Sul / UFMS / Campo Grande - MS

DADOS CURRICULARES

FORMAÇÃO ACADÊMICA

- 2011-atual** Doutorado em Química.
- Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, São Paulo, Brasil
- Título: Desenvolvimento de métodos analíticos de desreplicação por RMN, EM e análise multivariada do perfil metabolômico de espécies de Solanaceae com potencial biológico.
- Orientador: Prof. Dr. Ian Castro-Gamboa
- Bolsa: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)
-
- 2009 - 2011** Mestrado em Química.
- Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, São Paulo, Brasil
- Título: Estudo de *Jatropha gossypifolia* e *J. multifida* (Euphorbiaceae) aplicando métodos analíticos *in silico* e de desreplicação, visando a detecção e elucidação *in situ* dos constituintes micromoleculares com atividades acetilcolinesterásicas, cicatrizantes e antioxidantes.
- Orientador: Prof. Dr. Ian Castro-Gamboa
- Coorientadora: Prof^a. Dr.^a Marcia Nasser Lopes
- Bolsa: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)
-
- 2004 - 2009** Graduação em Licenciatura em Química.
- Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, São Paulo, Brasil

FORMAÇÃO ACADÊMICA COMPLEMENTAR

03/2013 -03/2014 Estágio de doutorado sanduiche.

Samuel Roberts Noble Foundation, Ardmore, Oklahoma, Estados Unidos

Título: Desenvolvimento de ferramentas de análise multivariada e bioinformática no estudo metabolômico de *Mendicago trucatula* e sub-espécies fazendo uso de CG e UHPLC-MS

Orientador: Prof. Dr. Lloyd W. Sumner

Bolsa BEPE: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

DISCIPLINAS CURSADAS DURANTE O DOUTORADO

1. RMN de produtos naturais (08 créditos), **conceito A**
2. Introdução a espectroscopia de RMN: processamento e análise de dados multivariáveis - UNESP (02 créditos), **conceito A**.
3. Tópicos em química analítica: análise multivariada de dados experimentais em química – UFSCAR (13 créditos), **conceito A**.
4. Fisiologia vegetal – Faculdade de Ciências Farmacêuticas – UNESP (06 créditos), **conceito B**.
5. Teoria e métodos de separação, isolamento e purificação de compostos orgânicos – UNESP (12 créditos), **conceito A**.
6. Quimiometria I - UFSCar, (04 créditos), **conceito B**
7. Quimiometria II - UFSCar, (04 créditos), **conceito A**
8. Quimiometria III - UFSCar, (04 créditos), **conceito A**
9. Quimiometria IV – UFSCAR, (04 créditos), **conceito A**

Total: 54 créditos

ARTIGOS COMPLETOS PUBLICADOS EM PERIÓDICOS

1. JING, J.; LEI, Z.; ZHANG, G.; PILON, A. C.; HUHMANN, D. V.; XIE, R.; XI, W.; ZHOU, Z.; SUMNER, L. W. Metabolite profiles of essential oils in citrus peels and their taxonomic implications. **Metabolomics**, 2015. doi:10.1007/s11306-014-0751-x, in press.

2. PILON, A. C., CARNEIRO, R. L., CARNEVALE NETO, F., BOLZANI, V. da S., CASTRO-GAMBOA, I. Interval Multivariate Curve Resolution in the Dereplication of HPLC-DAD Data from *Jatropha gossypifolia*. **Phytochemical Analysis**, v.24, p. 401 - 406, 2013.
3. CARNEVALE NETO, F., PILON, A. C., BOLZANI, V. S., CASTRO-GAMBOA, I. Chrysobalanaceae: secondary metabolites, ethnopharmacology and pharmacological potential. **Phytochemistry Reviews**, v.12, p. 121 - 146, 2013.
4. CARNEVALE NETO, F., SIQUITELLI, C. D., PILON, A. C., SILVA, D. H. S., BOLZANI, V. da S., CASTRO-GAMBOA, I. Dereplication of Phenolic Derivatives of *Qualea grandiflora* and *Qualea cordata* (Vochysiaceae) using Liquid Chromatography coupled with ESI-QToF-MS/MS. **Journal of the Brazilian Chemical Society**, v.24, p. 758 - 764, 2013.
5. CARNEVALE NETO, F., PILON, A. C., SILVA, D. H. S., BOLZANI, V. S., CASTRO-GAMBOA, I. Vochysiaceae: secondary metabolites, ethnopharmacology and pharmacological potential. **Phytochemistry Reviews**, v.10, p. 413 - 429, 2011.

PARTICIPAÇÃO EM EVENTOS CIENTÍFICOS:

1. FREIRE, R. T., PILON, A. C., CARNEVALE NETO, F., CASTRO-GAMBOA, I. Investigation of the NMR parameters analyses by a made-up MATLAB GUI (guide user interface) on an in-house database. International Conference on Omics Studies, 2013, Orlando.
2. CARNEVALE NETO, F., PILON, A. C., FREIRE, R. T., CARDOSO, P., GU, H., DJUKOVIC, D., RAFTERY, D., CASTRO-GAMBOA, I. Metabolomics as a tool for anticancer lead-finding from Natural Products. International Conference on Omics Studies, 2013, Orlando.
3. CARNEVALE NETO, F., PILON, A. C., FREIRE, R. T., CARDOSO, P., BOLZANI, V. S., CASTRO-GAMBOA, I. Dereplication of Brazilian Chrysobalanaceae plant species using Interval MCR-ALS based on recovery of pure compounds from HPLC-DAD-ESI-QToF-MS/MS metabolomics data. 8th Annual Meeting of the Metabolomics Society and the Plant Metabolomics Forum, 2012, Washington DC.
4. PILON, A. C., CARNEVALE NETO, F., CARNEIRO, R. L., FREIRE, R. T., BOLZANI, V. S., Castro-Gamboa, I. Chemometric-assisted tool on plant

metabolomics: Chemotypes distribution from *Solanum* species using PARAFAC-HCA method. 8th Annual Meeting of the Metabolomics Society and the Plant Metabolomics Forum, 2012, Washington DC.

5. CARNEVALE NETO, F., PILON, A. C., FREIRE, R. T., BOLZANI, V. S., CASTRO-GAMBOA, I. HPLC-DAD-ESI-QToF-MS/MS dereplication and chemotaxonomy approach of Chrysobalanaceae species. 3rd Brazilian Conference on Natural Products, 2011, Ouro Preto.
6. FREIRE, R. T., PILON, A. C., CARNEVALE NETO, F., CASTRO-GAMBOA, I. Construction of a versatile NMR in-house database: searching for adequate acquisition parameters, production of scripts and multivariate analysis. 3rd Brazilian Conference on Natural Products, 2011, Ouro Preto.
7. PILON, A. C., CARNEVALE NETO, F., FREIRE, R. T., CARNEIRO, R. L., CASTRO-GAMBOA, I. Development of a deconvolution method using Multivariate Curve Resolution (MCR) for dereplication studies on leaves extracts from *Jatropha multifida*. 3rd Brazilian Conference on Natural Products, 2011, Ouro Preto.
8. PILON, A. C., CARNEIRO, R. L., CARNEVALE NETO, F., CASTRO-GAMBOA, I. Interval multivariate curve resolution in the dereplication of HPLC-DAD data from *Jatropha gossypifolia* In: 23rd International Symposium on Pharmaceutical and Biomedical Analysis, 2011, João Pessoa.
9. CARNEVALE NETO, F., CASTRO-GAMBOA, I. Metabolomic analysis for the understanding of pharmacological response of bioactive complex matrices. São Paulo School of Advanced Science, 2011, Campinas.
10. CARNEVALE NETO, F., PILON, A. C., QUEIROZ, M. M. F., SILVA, D. H. S., BOLZANI, V. S., CASTRO-GAMBOA, I. Desreplificação de *Qualea grandiflora* e *Qualea cordata* (Vochysiaceae) através de HPLC-HRMS/MS e RMN. 34a Reunião Anual da Sociedade Brasileira de Química SBQ, 2011, Florianópolis.
11. PILON, A. C., CARNEVALE NETO, F., QUEIROZ, M. M. F., LOPES, M. N., SILVA, D. H. S., BOLZANI, V. S., CASTRO-GAMBOA, I. Planejamento e otimização de extração em espécies de *Jatropha* para estudos

metabolômicos. 34a Reunião Anual da Sociedade Brasileira de Química SBQ, 2011, Florianópolis.

1. Outros:

- Co-autor de painel premiado intitulado "Investigation of the NMR parameters analyses by a made-up MATLAB GUI (guide user interface) on an in-house database", International Conference on Omics Studies, 2013.
- Co-autor de painel premiado intitulado "Metabolomics as a tool for anticancer lead-finding from Natural Products", International Conference on Omics Studies, 2013.
- Trabalho premiado intitulado "Chemometric-assisted tool on plant metabolomics: Chemotypes distribution from Solanum species using PARAFAC-HCA method", The Noble Foundation Awards Best Young Investigator Oral Presentation in Plant Metabolomic, 2012.
- Co-autor de painel premiado intitulado "HPLC-DAD-ESI-QToF-MS/MS dereplication and chemotaxonomy approach of Chrysobalanaceae species". 3rd Brazilian Conference on Natural Products, 2011.
- Painel premiado intitulado "Development of a deconvolution method using Multivariate Curve Resolution (MCR) for dereplication studies on leaves extracts from *Jatropha multifida*". 3rd Brazilian Conference on Natural Products, 2011.
- Painel premiado intitulado "Interval multivariate curve resolution in the dereplication of HPLC-DAD data from *Jatropha gossypifolia*". 23rd International Symposium on Pharmaceutical and Biomedical Analysis, 2011.

Dedico esta tese a

Deus

Jesus

Thom Yorke

Brad Mehldau

Led Zeppelin

Fritjof Capra

Ana Paez-García

Fausto Carnevale Neto

Rafael Teixeira Freire

NuBBE

Ian Castro-Gamboa

Irmãos

Mãe e Pai

Por me ajudarem a encontrar o meu
caminho, e a Mônica,

que há muito tempo percorre comigo
nesta caminhada.

*“Só é lutador
quem sabe lutar consigo mesmo”*

Carlos Drummond de Andrade

AGRADECIMENTOS

Inicialmente agradeço à Deus por permitir essa caminhada.

Faço aqui um sincero agradecimento ao meu amigo-orientador Ian Castro-Gamboa, não só por ter aceitado orientar minha dissertação e em seguida minha tese, mas celebrar comigo nos momentos de vitória assim como demonstrar afeto, companheirismo, motivação e muitos outros adjetivos que serviram de inspiração nos momentos difíceis dessa trajetória.

Quero também deixar um agradecimento especial à todas as pessoas que tive a oportunidade de encontrar no NuBBE. Contatos diários com o pluralismo deste grupo enriqueceram não somente meu trabalho científico mas alimentaram a minha ânsia pelo aperfeiçoamento cultural, social e pessoal.

Um agradecimento em especial aos professores, Alberto Cavalheiro, Ângela Araújo, Maysa Furlan, Dulce Helena Siqueira Silva e Vanderlan da Silva Bolzani que de forma direta e indireta contribuíram para minha formação. Agradeço aos técnicos e funcionários pelos serviços e colaborações.

Faço um agradecimento profundo à minha família, em especial aos meus pais e irmãos que sempre me apoiaram ao longo da vida e sempre me proporcionaram incondicionalmente todas os requisitos necessários para chegar até aqui. Eles sempre foram e serão as minhas maiores inspirações do que é ser uma pessoa digna e guerreira.

Por fim, mas não menos importante, deixo um profundo e sincero agradecimento à minha namorada Mônica Pereira por todo o amor, carinho, amizade, dedicação, compreensão e apoio nos momentos mais difíceis, sem a qual atingir este objetivo não teria o mesmo sabor.

RESUMO DO PROJETO

O uso de extratos e tinturas de partir de fontes naturais para fins terapêuticos remonta o aparecimento da humanidade. A ação biológica proveniente destas misturas é justificado pela diversidade estrutural e química de seus compostos na qual, interação concomitante a diferentes alvos proteicos. No entanto, os últimos anos foram marcados pela substituição das metodologias clássicas de isolamento e teste biológicos por ferramentas de automação como *High Throughput Screening* (HTS) e o uso da química combinatória para a geração de substâncias. No entanto, os resultados até o momento foram insatisfatórios e a indústria tem retomado o uso dos produtos naturais. Para tanto, ela tem buscado por abordagens mais racionais de detecção e identificação estrutural de biomarcadores. Dentro dessa óptica, a inserção e o desenvolvimento de metodologias de desreplicação associadas a estudos metabolômicos sob a visão da biologia sistêmica torna-se crítica na busca por micromoléculas ativas. Recentemente, o uso de técnicas hifenadas como a cromatografia líquida de alta eficiência (CLAE), associadas a métodos espectroscópicos de RMN e a espectrometria de massas de alta resolução (EMAR), tem sido fundamentais na geração de métodos analíticos em estudos metabolômicos e de biologia sistêmica em plantas, fungos endofíticos, e outras organismo, acelerando a escolha de matrizes bioativas promissoras. Este projeto visa ao entendimento das relações moleculares presentes em matrizes dinâmicas gerando ferramentas analíticas para uma abordagem racional no estudo de produtos naturais bioativos. Para tanto, espécimes depositadas na nossa extratoteca, da família Solanaceae, pouco estudadas no que diz respeito a sua composição química serão o objeto de estudo deste projeto.

ABSTRACT

The use of drugs from natural sources is as old as humanity itself. However, due to the economic impact caused by globalization, many pharmaceutical industries searched for alternatives to speed-up assays aiming for simpler structural molecules and low-costs budgets. Therefore, the insertion and development of methods of dereplication associated to metabolomics studies through an approach in system biology is a key issue when searching for bioactive molecules. Recently, due to the advancement of technology, the use of hyphenated techniques such as high performance liquid chromatography (HPLC) coupled to spectroscopic methods like nuclear magnetic resonance (NMR) and high resolution mass spectrometry (HRMS) intensified the metabolomic and system biology research in plants and fungi thus, accelerating the selection of bioactive substances. This project aims to study the molecular dynamics in complex matrixes using multivariate analysis and in doing so generating analytical tools to apply as a rational approach towards the study of bioactive natural products and the understanding of their complexity. To accomplish this task, specimens deposited in our bank of extracts, from the Solanaceae family, little studied on their chemical composition but that showed interesting bioactivities from initial biological screenings, will be the object of study on this project.

LISTA DE FIGURAS

FIGURA 1. AGENTES TERAPÊUTICOS DESENVOLVIDOS ENTRE 1950-1970 A PARTIR DA QUÍMICA DE PRODUTOS NATURAIS - (1) STREPTOMICINA; (2) TETRACICLINA E (3) MEVINOLINA. O SORAFENIB (4) FOI DESENVOLVIDO DURANTE NO INÍCIO DA DÉCADA DE 2000 COM A INTRODUÇÃO DA QUÍMICA COMPUTACIONAL E COMBINATÓRIA.	31
FIGURA 2. ESQUEMA ILUSTRATIVO DO FENÓTIPO DE UM ORGANISMO A PARTIR DO PRODUTO DA CONEXÃO INTRA-ESPECÍFICA ENTRE SISTEMAS GENÔMICOS, TRANSCRIPTÔMICOS, PROTEÔMICOS E METABOLÔMICOS.	34
FIGURA 3. DISTRIBUIÇÃO GEOGRÁFICA DA FAMÍLIA SOLANACEAE (EM VERDE) NO MUNDO, ACESSADO DE HTTP://WWW.MOBOT.ORG/MOBOT/RESEARCH/AP/WEB/	37
FIGURA 4. ESTRUTURA BASE DAS CINCO PRINCIPAIS CLASSES DE GLICOALCALOIDES ENCONTRADOS EM ESPÉCIES DE <i>SOLANUM</i> . O PRIMEIRO ESQUELETO FOI INSERIDO A CRITÉRIO DE NOMENCLATURA E NUMERAÇÃO DOS ANÉIS DOS GLICOALCALOIDES ESTEROIDAI. 1 – SOLANIDANOS; 2 – ESPIROSOLANO; 3 – (22,26) EPIMINOCLOSTANO; 4 – 3-AMINOESPIROSTANO E 5 – α -EPIMINOCICLOEMICETAL.	38
FIGURA 5. PROPOSTA BIOSINTÉTICA DA FORMAÇÃO DE ALCALOIDES ESPIROSOLANOS A PARTIR DA ROTA DO COLESTEROL. ENTRÁIDO DE DEWICK, 2002, P.390.	38
FIGURA 6. PARÂMETROS UTILIZADOS PARA DESCREVER A PERFORMANCE DE ESPECTRÔMETROS DE MASSAS. O PODER RESOLUÇÃO DE MASSA (OU RESOLUÇÃO), $m/\Delta m$, PODE SER DESCRITO DUAS MANEIRAS: (1). PARA m SENDO A MÉDIA DA RAZÃO MASSA-CARGA ASSOCIADA COM DOIS SINAIS DE MASSAS ADJACENTES DE IGUAL TAMANHO E FORMA QUE SE SOBREPÕES POR 50% E Δm SENDO A DIFERENÇA EM ENTRE O SINAL DE MASSA-CARGA ENTRE OS DOIS SINAIS DE MASSAS ADJACENTES OU (2) m SENDO O SINAL DE MASSA NO ÁPICE E Δm SENDO A LARGURA A MEIA ALTURA (50%) DO SINAL DE MASSA, DESIGNADO POR FWHM (FULL WIDTH AT HALF HEIGHT OF MAXIMUM). A EXATIDÃO DE MASSA É DESCRITO PELA RAZÃO ENTRE O ERRO DE MASSA (DIFERENÇA ENTRE A MASSA REAL E A MEDIDA) E A MASSA TEÓRICA, FREQUENTEMENTE REPRESENTADO COMO PARTES POR MILHÃO (PPM). A SENSIBILIDADE É DESCRITA PELA RAZÃO ENTRE NÍVEL DO SINAL DE MASSA E O NÍVEL DO SINAL DO RUÍDO. A VARIAÇÃO DA DINÂMICA LINEAR É DESCRITO COMO O INTERVALO DE LINEARIDADE DO SINAL DO ÍON MEDIDO EM FUNÇÃO DA CONCENTRAÇÃO DO ANALITO.....	49
FIGURA 7. FOTOS ILUSTRATIVAS DAS ESPÉCIES ESTUDADAS: <i>S. PANICULATUM</i> ; <i>S. SWARTZIANUM</i> ; <i>S. AMERICANUM</i> E <i>S. EXCELSUM</i> . .	52
FIGURA 8. FLUXOGRAMA RELATIVO A PREPARAÇÃO DAS AMOSTRAS PARA OS EXPERIMENTOS EM CG-EM E CLAE-DAD-EM. OS METABÓLITOS DAS AMOSTRAS ANALISADAS POR CG FORAM EXTRAÍDOS COM AUXÍLIO DE UMA MISTURA DE MeOH:CHCl ₃ (1:1 v/v) E, UMA VEZ SECOS, OS EXTRATOS FORAM SUBMETIDOS A REAÇÕES DE METOXIMAÇÃO SEGUIDA DE SILILAÇÃO PARA REMOÇÃO DE AGENTES DE ADESÃO A COLUNA DE CG E MELHORAR A CAPACIDADE DE VOLATILIZAÇÃO A UMA MENOR TEMPERATURA. UMA MISTURA DE ÉSTERES METÍLICOS FOI ADICIONADO PARA AS SUBSEQUENTES ETAPAS DE ELUCIDAÇÃO ESTRUTURAL UTILIZANDO O SOFTWARE AMDIS E BASE DE DADOS. FORAM REMOVIDOS INTERFERENTES APOLARES DAS AMOSTRAS DE CLAE-DAD-EM ATRAVÉS DA ELUIÇÃO DAS AMOSTRAS SOB CARTUCHOS DE C-18 NA PROPORÇÃO DE MeOH:H ₂ O (8:2 v/v) E FILTRADOS EM MEMBRADA DE 0,45 MM.....	55
FIGURA 9. MECANISMO DE REAÇÃO DE METOXIMAS EM SISTEMAS CARBONÍLICOS.	56
FIGURA 10. MECANISMO DE REAÇÃO PARA SILILAÇÃO DE GRUPOS HIDROXIL. FUNÇÕES COMO TIÓIS E AMINAS TAMBÉM SE ENQUADRAM NESTE TIPO DE REAÇÃO.	56
FIGURA 11. CONSTRUÇÃO DE REDE MOLECULAR VIA COMPARAÇÃO ESPECTRAL, FIGURA ADAPTADA DO TRABALHO DE WATROUS J. ET AL., 2012. (A) REPRESENTAÇÃO ESQUEMÁTICA DE COMO O NETWORKING MOLECULAR É GERADO. OS VALORES REPRESENTAM A SIMILARIDADE DE COSSENOS QUE VARIA DE 0 A 1, ONDE 1 REPRESENTA ESPECTROS IDÊNTICOS. A ESPESSURA DA ARESTAS (LINHAS AZUIS) INDICAM O NÍVEL DE SIMILARIDADE. (B) VISUALIZAÇÃO PELO CYTOSCAPE DE COMO A REDE É FORMADA. (C) UM EXEMPLO DO CÁLCULO DA PONTUAÇÃO DE SIMILARIDADE USANDO QUATROS ESPECTROS DIFERENTES.	62
FIGURA 12. GRÁFICO DO CÁLCULO DA EQUAÇÃO DE REGRESSÃO (POLINÔMIO CÚBICO) PARA OS VALORES MÉDIOS DOS ÉSTERES METÍLICOS	65
FIGURA 13. GRÁFICOS DE RESPOSTA EM FUNÇÃO DOS VALORES DE FDC SOB MATCHING FATOR > 90 PARA AS AMOSTRAS F55, M951, M952, F052, R270 E R271. (A) JANELA DE COMPONENTES; (B) SUBTRAÇÃO DE PICOS ADJACENTES; (C) FATOR DE RESOLUÇÃO; (D) FATOR DE SENSIBILIDADE E (E) REQUERIMENTOS DE FORMA.	68
FIGURA 14. (A) E (C) SÃO PERFIS CROMATOGRÁFICO DO EXTRATO DE FOLHAS E RAMOS (RESPECTIVAMENTE) DA ESPÉCIE <i>S. PANICULATUM</i> (M985, M986) OBTIDOS A 254, 280, 330 E 360 NM ENQUANTO AS FIGURAS (B) E (D) REPRESENTAM OS CROMATOGRAMAS DE PICOS BASE (BPC). OS VALORES ENCONTRADOS EM (B) E (D) FAZEM REFERÊNCIA AOS COMPOSTOS ENCONTRADOS NA TABELA 7. A COLORAÇÃO DA NUMERAÇÃO REPRESENTA A CLASSE DE COMPOSTOS: VERDE (NUCLEOTÍDEOS),	

LARANJA (AMINOÁCIDOS E DERIVADOS), VERMELHO (POLIAMINAS), AZUL (FLAVONOIDES) E PRETO (GLICOALCALOIDES E SAPONINAS)	75
FIGURA 15. (A) PROPOSTA MECANÍSTICA PARA A FORMAÇÃO DO ÍON m/z 273 E 161 A PARTIR DO PERFIL DA FRAGMENTAÇÃO DA TOMATIDINA SEGUNDO CLAEYS <i>ET.AL.</i> , 1996; (B) MECANISMO DE FRAGMENTAÇÃO PARA PERDA DE ÁGUA; (C) MECANISMO DE FRAGMENTAÇÃO PARA CADEIA LATERAL (ANEL E E F) E PERDA NEUTRA DE 143 DA.	88
FIGURA 16. MECANISMOS DE FRAGMENTAÇÃO DIRIGIDO POR CARGA REMOTA (À DISTÂNCIA) PROPOSTO PARA A FORMAÇÃO DO ÍON 147 m/z	89
FIGURA 17. PERFIL METABOLÔMICO OBTIDO DO CROMATOGRAMA DE ÍONS TOTAIS (TIC) PARA AMOSTRAS DE <i>SOLANUM</i> : (A) FOLHAS <i>S. SWARTZIANUM</i> – R271; (B) RAMOS <i>S. SWARTZIANUM</i> – R272; (C) FOLHAS <i>S. SWARTZIANUM</i> – F052; (D) FOLHAS <i>S. AMERICANUM</i> – M951; (E) RAMOS <i>S. AMERICANUM</i> – M952; (F) FOLHAS <i>S. EXCELSUM</i> – F55. A NUMERAÇÃO REFERE-SE AOS COMPOSTOS PRESENTES NAS TABELAS 8, 9 E 10. A COLORAÇÃO DA NUMERAÇÃO REFERE-SE: PRETO (FLAVONOIDES <i>O</i> -GLICOSILADOS – TABELA 8); VERMELHO (FLAVONOIDES <i>C</i> -GLICOSILADOS – TABELA 9) E AZUL (GLICOALCALOIDES ESTEROIDAIIS – TABELA 10);	94
FIGURA 18. MAPA DA REDE MOLECULAR (<i>MOLECULAR NETWORKING</i>) OBTIDA PARA OS DADOS EM/EM DAS ESPÉCIES DE <i>SOLANUM</i> : 1- FLAVONÓIDES <i>O</i> -GLICOSILADOS (CINZA); 2 – GLICOALCALOIDES ESTEROIDAIIS – SAPONINAS (AZUL) E 3 – FLAVONOIDES <i>C</i> -GLICOSILADOS (VERMELHO).	95
FIGURA 19. MAPA DA REDE MOLECULAR ASSOCIADA AOS FLAVONOIDES <i>O</i> -GLICOSILADOS PRESENTES NAS FOLHAS E RAMOS DE <i>S. SWARTZIANUM</i> (R271, R272 E F052), <i>S. AMERICANUM</i> (M951 E M952) E <i>S. EXCELSUM</i> (F55). A FORÇA DE ARESTA REPRESENTA O GRAU DE SIMILARIDADE ENTRE DOIS NODOS. TONS AZUIS APRESENTAM O NÍVEL DE SIMILARIDADE ENTRE 95 – 99 %, VERMELHOS ENTRE 90-95% E PRETO ENTRE 70-90%. A COR DA LEGENDA (DENTRO DOS NODOS) REVELA ONDE FOI ENCONTRADO O ÍON (FOLHAS OU RAMOS). O FORMATO DO NODOS DESCREVE O TIPO DE AGLICONA (FLAVONA; FLAVONOL E SEUS DERIVADOS); A COR DO NODO DESCREVE A ESPÉCIE ONDE FOI ENCONTRADA O ÍON MOLECULAR. O TAMANHO DO NODO ESTÁ ASSOCIADO AO NÚMERO DE COMPOSTOS POR NODO E A COR DE BORDA É RELATIVO AO TIPO E NÚMERO DE GLICOSÍDEOS PRESENTES PARA UMA DADA AGLICONA.	100
FIGURA 20. MAPA DE REDE MOLECULAR ASSOCIADA AOS FLAVONOIDES <i>C</i> -GLICOSILADOS PRESENTES EM <i>S. AMERICANUM</i> (M951 E M952) A FORÇA DE ARESTA REPRESENTA O GRAU DE SIMILARIDADE ENTRE DOIS NODOS. O FORMATO DO NODOS DESCREVE O TIPO DE AGLICONA (APIGENINA – 271 m/z ; LUTEOLINA – 287 m/z E CRISOERIOL – 301 m/z). A COR DE BORDA DESCREVE A POSIÇÃO EM QUE O GLICOSÍDEO ESTÁ LIGADO A AGLICONA (8-C OU 6-C).	102
FIGURA 21. MAPA DA REDE MOLECULAR ASSOCIADA AOS FLAVONOIDES <i>C</i> -GLICOSILADOS PRESENTES NAS FOLHAS E RAMOS DE <i>S. AMERICANUM</i> (M951 E M952) E <i>S. SWARTZIANUM</i> (R271). A FORÇA DE ARESTA REPRESENTA O GRAU DE SIMILARIDADE ENTRE OS NODOS. O FORMATO DO NODOS DESCREVE O TIPO DE AGLICONA (APIGENINA – 271 m/z ; LUTEOLINA – 287 m/z E CRISOERIOL – 301 m/z). A COR DE BORDA DESCREVE O TIPO DE GLICOSÍDEO (PENTOSE OU HEXOSE).	102
FIGURA 22. COMPARAÇÃO ESPECTRAL ENTRE VALORES EXPERIMENTAIS E DADOS DA LITERATURA PARA ISOVITEXINA E VITEXINA. PERFIS DE FRAGMENTAÇÃO DOS ISÔMEROS m/z 433 A 20eV, VITEXINA (SUPERIOR ESQUERDO - VERMELHO) E ISOVITEXINA (SUPERIOR DIREITO - AZUL) OBTIDOS DO TRABALHO DE WARIDEL ET. AL. 2001. (A) PERFIL ESPECTRAL OBTIDO A 30 eV PARA AMOSTRA <i>S. AMERICANUM</i> (FOLHAS) RELATIVO A VITEXINA E (B) PERFIL ESPECTRAL DA ISOVITEXINA OBTIDO 30 eV PARA A MESMA AMOSTRA.	103
FIGURA 23. MAPA DA REDE MOLECULAR ASSOCIADA AOS GLICOALCALOIDES E SAPONINAS ESTEROIDAIIS PRESENTE NAS FOLHAS E RAMOS DE <i>S. SWARTZIANUM</i> (R271, R272 E F052). A FORÇA DE ARESTA REPRESENTA O GRAU DE SIMILARIDADE ENTRE OS NODOS. EM AZUL ESTÃO OS NODOS COM CORRELAÇÃO ACIMA DE 95% DE SIMILARIDADE, EM VERMELHO ENTRE 90-95% E EM PRETO ENTRE 70-90%. O FORMATO DO NODOS DESCREVE O TIPO DE AGLICONA. A COR DE BORDA REPRESENTA OS GLICOSÍDEOS E ÍONS FRAGMENTOS DA QUEBRA DO NÚCLEO TRITERPÊNICOS.	105
FIGURA 24. ILUSTRAÇÃO DO FENÔMENO DE RESSONÂNCIA MAGNÉTICA NUCLEAR PARA O NÚCLEO DE PRÓTIO (ADAPTADO DE LEVITT M. H. <i>SPIN DYNAMICS</i> , 2007). (1) O ÁTOMO DE HIDROGÊNIO APRESENTA DOIS TIPOS MOMENTO MAGNÉTICO: (i) REFERENTE A PROPRIEDADE DA MATÉRIA INTRÍNSECA AOS PRÓTONS E ELÉTRONS (SPIN NUCLEARES) E AO MOVIMENTO ORBITAL DOS ELÉTRONS EM TORNO DO NÚCLEO (GERAÇÃO DE CAMPO MAGNÉTICO A PARTIR DO MOVIMENTO DE CARGA). (2) RESULTANTE DA APLICAÇÃO DE UM CAMPO MAGNÉTICO EXTERNO (B) NO VETOR DE MOMENTO MAGNÉTICO NUCLEAR (PARALELA E ANTIPARALELA). DEVIDO A PEQUENAS FLUTUAÇÕES DE CAMPO CAUSADO PELA FORTE MOVIMENTAÇÃO ATÔMICA (EFEITO DE EQUILÍBRIO TÉRMICO) A RESULTANTE PARALELA POSSUI UM LIGEIRO EXCESSO POPULACIONAL RESULTANDO NUM VETOR MAGNÉTICO EM DIREÇÃO AO EIXO Z. (3) FENÔMENO DE RESSONÂNCIA ATRAVÉS DA APLICAÇÃO DE PULSO DE RADIOFREQUÊNCIA ESPECÍFICO PARA OS NÚCLEOS. ATRAVÉS DESSA EXCITAÇÃO É POSSÍVEL GIRAR A RESULTANTE MAGNÉTICA DO NÚCLEO (EIXO Z) EM 90° PERMITINDO SUA DETECÇÃO SEM A CONTRIBUIÇÃO DOS EFEITOS MAGNÉTICOS DOS ELÉTRONS. (4) AQUISIÇÃO DOS DADOS DE RESSONÂNCIA A PARTIR DE BOBINAS APONTADAS AO LONGO DO EIXO PERPENDICULAR A CAMPO EXTERNO (B). OS	

- DADOS SÃO ADQUIRIDOS NA FORMA DE ONDAS NA QUAL SÃO TRANSFORMADAS POR FUNÇÕES MATEMÁTICAS (TRANSFORMADA DE FOURRIER) EM GRÁFICOS EM FUNÇÃO DE FREQUÊNCIA. 110
- FIGURA 25. ILUSTRAÇÃO DO PROCESSO DE EXECUÇÃO DO RANSY. UMA VEZ SELECIONADO *DRIVING PEAK** (SINAL DE DESLOCAMENTO QUÍMICO ESPECÍFICO PARA O COMPOSTO A) OCORRE A DIVISÃO ENTRE VALORES DE INTENSIDADE DOS DEMAIS SINAIS DE RMN COM O VALOR DE INTENSIDADE DO *DRIVING PEAK*. ESSE PROCESSO (RANSY) SE REPETE PARA CADA ESPECTRO DE RESSONÂNCIA, A PARTIR DE AMOSTRAS SUBMETIDAS A DIFERENTES CONDIÇÕES. OS VALORES DE RAZÃO ENTRE OS SINAIS RELATIVOS AO COMPOSTO A PERMANECEM RELATIVAMENTE CONSTANTES (PERTENCEM AO MESMO SISTEMA DE SPIN) E CONSEQUENTEMENTE, O DESVIO PADRÃO DA MÉDIA DAS RAZÕES É RELATIVAMENTE BAIXO. EM CONTRAPARTIDA, PARA AS RAZÕES REFERENTES AO COMPOSTO B NÃO SÃO CONSTANTES E O DESVIO PADRÃO RESULTANTE DAS RAZÕES É ELEVADO. O ESPECTRO DE RANSY É O RESULTADO DA MÉDIA DAS RAZÕES (SINAL RMN DIVIDIDO PELO *DRIVING PEAK*) DIVIDIDO PELO SEU DESVIO PADRÃO. ASSIM, O ESPECTRO CALCULADO RANSY DESTACA OS SINAIS DO COMPOSTO A EM DETRIMENTO DO COMPOSTO B, PERMITINDO VISUALIZAR OS ÍONS CARACTERÍSTICAS APENAS PARA ESTE COMPOSTO. 115
- FIGURA 26. ILUSTRAÇÃO DE UM TÍPICO EXPERIMENTO DE DOSY. (1) NESTA CONDIÇÃO TEMOS A RESULTANTE DO VETOR MAGNÉTICO DECORRENTE DE PULSO R.F. A 90°_{xy} . TODOS OS SPINS SÓ POSSUEM A MESMA FREQUÊNCIA DE RESSONÂNCIA (INICIALMENTE APONTADA PARA O EIXO $-Y$) DEVIDO A HOMOGENEIDADE DO CAMPO MAGNÉTICO. (2) GRADIENTE DE CAMPO ADICIONAL (BOBINAS DA SONDA - BG) GERADO AO LONGO DO EIXO Z (TG – TEMPO DE APLICAÇÃO DO GRADIENTE DE CAMPO). EM (2) A FREQUÊNCIA DE RESSONÂNCIA DE CADA SPIN É DIFERENTE, O QUE RESULTA NUM SINAL BAIXO (FORA DE FASE). (3) GRADIENTE DE CAMPO INVERSO SOB O MESMO TEMPO (TG). SE NÃO HOUVESSE O FENÔMENO DE DIFUSÃO, TEORICAMENTE, O GRADIENTE INVERSO REFOCARIÁ OS SPINS DA ETAPA 2 RESULTANDO NO MESMO SINAL QUE NA ETAPA EM 1. NO ENTANTO, DEVIDO A DIFUSÃO MOLECULAR OS SPINS NUCLEARES NÃO “SENTEM” O MESMO GRADIENTE INVERSO, POIS SE LOCALIZAM EM POSIÇÕES DIFERENTES DA INICIAL, RESULTANDO NUM SINAL MENOR QUE A ETAPA 1..... 121
- FIGURA 27. PERFIL CROMATOGRÁFICO DO EXTRATO DE FOLHAS DE *S. PANICULATUM* (M985) OBTIDOS A 254 E 280 NM (VERMELHO E AZUL RESPECTIVAMENTE). AS LINHAS VERTICAIS (PONTILHADAS) REFEREM-SE AO INTERVALO DE COLETA DOS PICOS NO SISTEMA CLAE-SPE. AS SIGLAS ENTRE PARÊNTESES DESCREVEM EM QUAIS CARTUCHOS DE SPE FORAM COLETADOS OS COMPOSTOS. A CRITÉRIO DO TRABALHO, ESTAS SIGLAS FORAM UTILIZADAS PARA IDENTIFICAR OS COMPOSTOS. AS CAIXAS PROJETADAS A PARTIR DO ÁPICE DE CADA BANDA CROMATOGRÁFICA REPRESENTA ESPECTRO DE UV (220-400 NM) PARA O COMPOSTO COLETADO. 127
- FIGURA 28. ESPECTRO DE ^1H PARA O COMPOSTO PRESENTE NO CARTUCHO 1G3 - QUERCETINA-3-O-B-[GLC, (1-2) 2'''-O-B-GLC. AS AMPLIAÇÕES ENTRE 4,5 – 5,4 PPM REVELAM A PRESENÇA DE DOIS GLICOSÍDEOS – DUBLETOS (8Hz), ENQUANTO A AMPLIAÇÃO ENTRE 6,2 E 7,8 MOSTRAM O PADRÃO DE SUBSTITUIÇÃO TÍPICO DA QUERCETINA. 128
- FIGURA 29. COMPARAÇÃO ESPECTRAL ENTRE OS DADOS EXPERIMENTAIS DE ^1H (EM VERMELHO) E DADOS SIMULADOS (MESTRENOVA – EM AZUL) PARA O COMPOSTO PRESENTE NO CARTUCHO 1G3 - QUERCETINA-3-O-B-[GLC, (1-2) -O-B-GLC SOB O INTERVAL DE 2,5 A 8,5 PPM. 128
- FIGURA 30. ESPECTRO DE ^1H PARA O COMPOSTO PRESENTE NO CARTUCHO 1G5 - KAEMPFEROL-3-O-B-[GLC (1-2) 2'''-O-B-GLC]. AS AMPLIAÇÕES ENTRE 4,4 – 5,3 PPM REVELAM A PRESENÇA DE DOIS GLICOSÍDEOS – DUBLETOS (8Hz), ENQUANTO AMPLIAÇÃO ENTRE 6,2 E 7,8 MOSTRA O PADRÃO DE SUBSTITUIÇÃO TÍPICO DO KANFEROL. 130
- FIGURA 31. ESPECTRO DE $^1\text{H} - ^{13}\text{C}$ G-HSQC PARA O COMPOSTO PRESENTE NO CARTUCHO 1G5 - KAEMPFEROL-3-O-B-[GLC (1-2) 2'''-O-B-GLC]. A COLORAÇÃO VERMELHA DO NODOS REPRESENTA OS CARBONOS METÍNICOS (CH) ENQUANTO AS AZUIS ESTÃO ASSOCIADAS A CARBONOS METILÊNICOS (CH₂). A LEGENDA DOS NODOS FOI OBTIDA ATRAVÉS DA EXECUÇÃO DE *PEAK PICKING* NO SOFTWARE MESTRENOVA..... 130
- FIGURA 32. ESPECTRO DE ^1H PARA O COMPOSTO PRESENTE NO CARTUCHO 1G8 - NARINGENINA. AS AMPLIAÇÕES ENTRE 2,6 – 3,3 PPM REVELAM O PADRÃO DE MULTIPLICIDADE ENTRE H-2 E H-3 (ACOPLAMENTOS GEMINAIS E EQUATORIAL/AXIAL). A AMPLIAÇÃO ENTRE 6,2 E 7,8 MOSTRA O PADRÃO DE SUBSTITUIÇÃO TÍPICO DO NARINGENINA..... 131
- FIGURA 33. COMPARAÇÃO ESPECTRAL ENTRE OS DADOS EXPERIMENTAIS DE ^1H (EM VERMELHO) E DADOS SIMULADOS (MESTRENOVA – EM AZUL) PARA O COMPOSTO PRESENTE NO CARTUCHO 1G8 – NARINGENINA, SOB O INTERVAL DE 2,5 A 13 PPM. 132
- FIGURA 34. ESPECTRO DE ^1H PARA O COMPOSTO PRESENTE NO CARTUCHO 1G11 – DERIVADO METOXILADO DE NARINGENINA (7 OU 4') . AS AMPLIAÇÕES ENTRE 2,6 – 3,3 PPM REVELAM O PADRÃO DE MULTIPLICIDADE ENTRE H-2 E H-3 (ACOPLAMENTOS GEMINAIS E EQUATORIAL/AXIAL). A AMPLIAÇÃO ENTRE 6,2 E 7,8 MOSTRA O PADRÃO DE SUBSTITUIÇÃO TÍPICO DO NARINGENINA..... 132

FIGURA 35. COMPARAÇÃO ESPECTRAL ENTRE OS DADOS EXPERIMENTAIS DE ^1H (EM VERMELHO) E DADOS SIMULADOS (MESTRENOVA) DO DERIVADO 4'-O-METILNARINGENINA – VERDE E 7-O-METILNARINGENINA – AZUL) PARA O COMPOSTO PRESENTE NO CARTUCHO 1G8, SOB O INTERVAL DE 2,6 A 7,8 PPM.	133
FIGURA 36. PERFIL CROMATOGRÁFICO DO EXTRATO DE FOLHAS DE <i>S. PANICULATUM</i> (FCFAR – Ar) OBTIDO A 220 NM. AS LINHAS VERTICAIS (PONTILHADAS) REFEREM-SE AO INTERVALO DE COLETA (80 SEGUNDOS) DOS PICOS NO SISTEMA CLAE-SPE. AS SIGLAS ENTRE PARÊNTESES DESCREVEM EM QUAIS CARTUCHOS DE SPE FORAM COLETADOS OS COMPOSTOS. A CRITÉRIO DO TRABALHO, ESTAS SIGLAS FORAM UTILIZADAS PARA IDENTIFICAR OS COMPOSTOS.	135
FIGURA 37. ESPECTROS DE ^1H (0,5 A 4,7 PPM) PARA O EXTRATO METANÓLICO DE <i>S. PANICULATUM</i> (FCFAR – Ar) – (PRETO) E AS FRAÇÕES 1B7 (VERDE), 1B8 (VERMELHO) E 1B10 (AZUL). NOTA-SE QUE OS COMPOSTOS PRESENTES NAS FRAÇÕES SÃO ALTAMENTE CORRELACIONADOS E APRESENTAM-SE COMO PRODUTOS MAJORITÁRIOS DO EXTRATO BRUTO.	137
FIGURA 38. ESPECTROS DE ^1H GERADOS PELA FERRAMENTAS STOCSY E RANSY PARA AS AMOSTRA 1B7, 1B8 E 1B10. EM (A) O ESPECTRO DE ^1H STOCSY OBTIDO PARA JANELA ESPECTRAL DE 0-5 PPM. A CORRELAÇÃO DOS SINAIS FRENTE AO <i>DRIVING PEAK</i> OCORRE ATRAVÉS DA COLORORAÇÃO. OU SEJA, TONS DE MESMA COR REVELAM SINAIS DE UM MESMO SISTEMA DE SPIN. EM (B) É OBSERVADO O ESPECTRO ^1H GERADO PELO RANSY SOB INTERVALO DE 0-5PPM. A CORRELAÇÃO DOS SINAIS DE UM MESMO SISTEMA DE SPIN FRENTE AO DP SE DÁ PELA INTENSIDADE DOS SINAIS. EM (C) E (D) SÃO MOSTRADOS AS AMPLIAÇÕES DE 0 A 1,5 PPM PARA O STOCSY E RANSY, RESPECTIVAMENTE.....	138
FIGURA 39. . ESPECTROS DE ^1H GERADOS PELA FERRAMENTAS STOCSY E RANSY PELA COMPARAÇÃO DAS AMOSTRAS 1B7, 1B8 E 1B10. (A) E (B) REPRESENTAM AS AMPLIAÇÕES DE 1,5 A 2,5 PPM E, (C) E (D) PARA OS INTERVALOS DE 2,5 A 5 PPM PARA O STOCSY E RANSY, RESPECTIVAMENTE.....	139
FIGURA 40. ESPECTRO PSEUDOBIIDIMENSIONAL (^1H) DE DOSY PARA FRAÇÃO 1B8. NO EIXO HORIZONTAL SÃO MOSTRADOS OS VALORES DE DESLOCAMENTO QUÍMICO (PPM) ENQUANTO O EIXO VERTICAL DESCREVE O COEFICIENTE DE DIFUSÃO MOLECULAR. EM (A) TEMOS A VISÃO GERAL DO ESPECTRO DE DOSY. A LINHAS HORIZONTALMENTE TRACEJADAS REPRESENTAM OS COMPOSTO PRESENTE NA FRAÇÃO. EM (B) OBSERVA-SE APENAS AS SUBST. 1 E 2 E O CONTAMINANTE DO CARTUCHO.	141
FIGURA 41. ESPECTROS $^1\text{H} - ^1\text{H}$ TOCSY1D PARA A FRAÇÃO 1B8. EM (A) OBSERVA-SE A PRESENÇA DE 7 ESPECTROS SENDO: VERMELHO (INFERIOR – EXTRATO BRUTO DA FRAÇÃO 1B8) ; BEGE (SEGUNDO – IRRADIADO EM 4,39 PPM); VERDE CLARO (TERCEIRO – 4,23 PPM); VERDE ESCURO (QUARTO – 4,18PPM); AZUL (QUINTO – 3,81PPM); ROXO (SEXTO – 3,48 PPM) E ROSA (SÉTIMO – 2,27 PPM). EM (B) SÃO APRESENTADOS 5 ESPECTROS: VERMELHO (INFERIOR – EXTRATO BRUTO DE 1B8); BEGE (SEGUNDO – 1,88 PPM); VERDE (TERCEIRO – 1,19 PPM); AZUL (QUARTO – 1,12 PPM) E ROXO (QUINTO – 1,03 PPM).	142
FIGURA 42. ESPECTROS BIDIMENSIONAIS OBTIDOS PARA A FRAÇÃO 1B8. (A) $^1\text{H} - ^1\text{H}$ COSY; (B) $^1\text{H} - ^{13}\text{C}$ HSQC E (C) $^1\text{H} - ^{13}\text{C}$ HMBC.	143
FIGURA 43. ILUSTRAÇÃO DA ATRIBUIÇÃO DOS SINAIS DE ^1H E ^{13}C REFERENTE AO SISTEMA DE SPIN 4. A COLORAÇÃO DOS ÁTOMOS FAZ REFERÊNCIA AOS DADOS APRESENTADOS NA TABELA 11.	146
FIGURA 44. ILUSTRAÇÃO DA ATRIBUIÇÃO DOS SINAIS DE ^1H E ^{13}C REFERENTE AO CONJUNTO DE SISTEMAS DE SPIN 1, 4 E 6 A COLORAÇÃO DOS ÁTOMOS FAZ REFERÊNCIA AOS DADOS APRESENTADOS NA TABELA 12. A COLORAÇÃO AZUL REPRESENTA O ÁTOMO DE OXIGÊNIO.	148
FIGURA 45. ILUSTRA A ESTRUTURA ELUCIDADA A PARTIR DA ESTRATÉGIA QUE COMBINA STOCSY, RANSY DOSY E EXPERIMENTO MONO E BIDIMENSIONAIS UTILIZANDO OS NÚCLEOS DE ^1H E ^{13}C – PANICULONIN A.	148

LISTA DE TABELAS

TABELA 1. O PLANEJAMENTO DO PRESENTE ESTUDO.	43
TABELA 2. LOCAIS DE COLETA DAS ESPÉCIES DE <i>SOLANUM</i> . OS CÓDIGOS DAS AMOSTRAS FAZEM REFERÊNCIA AOS IDENTIFICADORES (ID) PRESENTES NO BANCO DE EXTRATOS DO NuBBE.	52
TABELA 3. MÉTODO CROMATOGRÁFICO APLICADO A ANÁLISE DE CLAE-DAD-EM DAS AMOSTRAS M985 E M986 (MODO GRADIENTE DE ELUIÇÃO).....	57
TABELA 4. MÉTODO CROMATOGRÁFICO APLICADO AS AMOSTRAS DE <i>SOLANUM</i> : M985, M986, R271, R272, F052, M951, M952 E F55.	58
TABELA 5. VALORES DO TEMPO DE RETENÇÃO PARA A MISTURA DE ÉSTERES METÍLICOS (FAME) PARA AS AMOSTRAS: R270, R271, F052, M951, M952, F55 E O CONTROLE (CTRL. - BRANCO CROMATOGRÁFICO). ADICIONALMENTE SÃO APRESENTADOS OS VALORES MÉDIOS DO TEMPO DE RETENÇÃO (AZUL) PARA CADA ÉSTER ASSIM COMO O DESVIO PADRÃO (DP - VERMELHO). O ÍNDICE DE RETENÇÃO LINEAR (IRL) DOS RESPECTIVOS ÉSTERES TAMBÉM É FORNECIDO.....	64
TABELA 6. PARÂMETROS DE OTIMIZAÇÃO DO PROCESSO DE DECONVULUÇÃO E COMPARAÇÃO ESPECTRAL NO SOFTWARE AMDIS.	66
TABELA 7. METABÓLITOS IDENTIFICADOS ATRAVÉS DA ANÁLISE DOS DADOS CG-EM NO SOFTWARE DE DECONVULUÇÃO AMDIS. A CRITÉRIO DE FORMAÇÃO, AS AMOSTRAS FORAM CODIFICADAS COMO SEGUE: 1 – RAMOS DE <i>S. SWARTZIANUM</i> (R272); 2 – FOLHAS DE <i>S. SWARTZIANUM</i> (F052); 3 – FOLHAS DE <i>S. SWARTZIANUM</i> (R271); 4 – FOLHAS DE <i>S. AMERICANUM</i> (M951) E 5 – RAMOS DE <i>S. AMERICANUM</i> (M952); 6 – FOLHAS DE <i>S. EXCELSUM</i> (F55). SÃO FORNECIDOS OS VALORES DO TEMPO DE RETENÇÃO ENCONTRADOS NO AMDIS (tR), ÍNDICE DE RETENÇÃO LINEAR TEÓRICO (IRLit), ÍNDICE DE RETENÇÃO LINEAR CALCULADO A PARTIR DA EQUAÇÃO DE REGRESSÃO FIGURA 12 (IRCalc), PORCENTAGEM DE ERRO DO IR (IRCalc – IRLit), VALORES MÉDIOS DOS MATCHING FACTOR DE CADA COMPOSTO (MF) E A MASSA MOLECULAR (SEM A PRESENÇA DOS SILANÓIS).	69
TABELA 8. METABÓLITOS IDENTIFICADOS ATRAVÉS DA ANÁLISE DOS DADOS CLAE-DAD-EM-EM/EM DO EXTRATO DE FOLHAS E RAMOS DE <i>S. PANICULATUM</i> (AMT - AMOSTRA 1 - M985 E AMOSTRA 2 - M986). SÃO FORNECIDOS O TEMPO DE RETENÇÃO (tR), A MASSA MONOISOTÓPICA PROTONADA(EM AZUL - [M+H] ⁺), ERRO EM PPM DO VALOR MEDIDO COMPARADO AO VALOR TEÓRICO DA FORMULA MOLECULAR DO COMPOSTO (EM VERMELHO) E OS ÍONS MAJORITÁRIOS DO ESPECTRO DE FRAGMENTAÇÃO EM ² (EM/EM).....	76
TABELA 9. METABÓLITOS IDENTIFICADOS ATRAVÉS DA ANÁLISE DOS DADOS CLAE-AUTOEM/EM DAS AMOSTRAS <i>S. SWARTZIANUM</i> (1 - R272, 2 – R271, E 3 - F052), <i>S. AMERICANUM</i> (4 - M951 E 5 - M952) E <i>S. EXCELSUM</i> (6- F55). O TEMPO DE RETENÇÃO (tR), A MASSA MONOISOTÓPICA PROTONADA(EM AZUL - [M+H] ⁺), ERRO EM PPM DO VALOR MEDIDO COMPARADO AO VALOR TEÓRICO DA FORMULA MOLECULAR DO COMPOSTO (EM VERMELHO) E OS ÍONS MAJORITÁRIOS DO ESPECTRO DE FRAGMENTAÇÃO EM ² (EM/EM) TAMBÉM SÃO MOSTRADOS NA TABELA.	96
TABELA 10. FLAVONOIDES C-GLICOSILADOS ENCONTRADOS NAS FOLHAS E RAMOS DE <i>S. AMERICANUM</i> (M951 E M952). A NUMERAÇÃO VERMELHA REFERE-SE AOS PICOS CROMATOGRÁFICOS ENCONTRADOS NA FIGURA 17. SÃO APRESENTADOS OS TEMPOS DE RETENÇÃO (tR), A MASSA MONOISOTÓPICA PROTONADA (EM AZUL - [M+H] ⁺) E ERRO EM PPM DO VALOR MEDIDO COMPARADO A FORMULA MOLECULAR DO COMPOSTO (EM VERMELHO). *(<i>S. SWARTZIANUM</i> - R271).	101
TABELA 11. GLICOALCALOIDES E SAPONINAS ESTEROIDAIAS ENCONTRADAS NAS FOLHAS E RAMOS DE <i>S. SWARTZIANUM</i> (R271, R272 E F052). A NUMERAÇÃO EM AZUL REFERE-SE AOS PICOS CROMATOGRÁFICOS ENCONTRADOS NA FIGURA 17. SÃO APRESENTADOS OS TEMPOS DE RETENÇÃO (tR), A MASSA MONOISOTÓPICA PROTONADA (EM AZUL - [M+H] ⁺) E ERRO EM PPM DO VALOR MEDIDO COMPARADO A FORMULA MOLECULAR DO COMPOSTO (EM VERMELHO).	104
TABELA 12. SISTEMAS DE SPIN OBTIDOS A PARTIR DA COMBINAÇÃO DE RESULTADOS DAS TÉCNICAS STOCSY, RANSY, DOSY E EXPERIMENTOS DE ¹ H - ¹ H COSY, ¹ H - ¹ H TOCSY1D, ¹ H - ¹³ C HSQC E ¹ H - ¹³ C HMBC. (FIGURAS 38 - 42). TODOS CONJUNTOS DE TABELA APRESENTAM VALORES DE DESLOCAMENTO QUÍMICO DE ¹ H COM A RESPECTIVO ACOPLAMENTO ¹ H (MULTIPLICIDADE), COSY E COM ¹³ C ATRAVÉS DE HSQC E HMBC. CADA COR REPRESENTA UM SISTEMA DE SPIN: PRETO - SISTEMA 1; AZUL - SISTEMA 2; ROSA - SISTEMA 3; VERDE - SISTEMA 4; LARANJA - SISTEMA 5 E ROXO - SISTEMA 6. A PRESENÇA DE DIFERENTES NOS COLUNAS DESTINADAS AO HMBC REFEREM-SE A CONEXÃO ENTRE DIFERENTES SISTEMAS DE SPIN PARA A SUBST.1. A COLORAÇÃO VERMELHA PODE TANTO REPRESENTAR CARBONOS QUATERNÁRIOS NÃO PRESENTES NOS SISTEMAS DE SPIN OU ESTAR RELACIONADOS A SISTEMAS DE SPIN NÃO OBSERVADOS NOS EXPERIMENTOS.	144

LISTA DE ESQUEMAS

ESQUEMA 1. PROPOSTA DE FRAGMENTAÇÃO DA ADENOSINA (58) EM MODO POSITIVO DE IONIZAÇÃO (ESI).	78
ESQUEMA 2. PROPOSTA DE FRAGMENTAÇÃO DA FRUTOSE-(ISO)-LEUCINA (57) E FRUTOSE-FENILALANINA (59) EM MODO POSITIVO DE IONIZAÇÃO (ESI).	78
ESQUEMA 3. PROPOSTA DE FRAGMENTAÇÃO PARA O CAFEOL E FERULOIL PUTRESCINA (60 E 61) EM ESI(+)-MS.	80
ESQUEMA 4. PROPOSTA DE FRAGMENTAÇÃO DOS DERIVADOS DE ESPERMINA EM MODO POSITIVO. (R=H ($\Delta^{16,17}$ OU $\Delta^{24,25}$ (63); R=H $\Delta^{16,17}$ E $\Delta^{24,25}$ (65); R=HEXOSE $\Delta^{16,17}$ OU $\Delta^{24,25}$ (66); R=HEXOSE $\Delta^{16,17}$ E $\Delta^{24,25}$ (64));	82
ESQUEMA 5. PROPOSTA DE FRAGMENTAÇÃO DE DERIVADOS DE CAFEOL ESPERMIDINA (67, 68 E 73) EM MODOS POSITIVO. VIA A E VIA B REPRESENTAM AS DUAS POSSÍVEIS ROTAS DE FRAGMENTAÇÃO PARA ESTES DERIVADOS DE ESPERMIDINA..	83
ESQUEMA 6. PROPOSTA DE FRAGMENTAÇÃO PARA OS DERIVADOS DE ALCALOIDES ESTEROIDAIIS DO TIPO ESPIROSSOLANO (74, 75, 77E 78) EM MODO POSITIVO DE IONIZAÇÃO (ESI).	90
ESQUEMA 7. PROPOSTA DE FRAGMENTAÇÃO PARA OS ALCALOIDES DO TIPO SOLANIDANO (79 – 81) EM MODO POSITIVO DE IONIZAÇÃO (ESI).	91
ESQUEMA 8. ILUSTRA A ESTRATÉGIA DE ELUCIDAÇÃO ESTRUTURAL A PARTIR DE COLETAS EM <i>TIME SLICE</i> . A PRIMEIRA ETAPA CONSISTE NO FRACIONAMENTO DO EXTRATO BRUTO ATRAVÉS DE COLETAS SUCESSIVAS A CADA 80 SEGUNDOS EXEMPLIFICADO EM 1B7; 1B8, 1B9 E 1B10. EM SEGUIDA, DUAS ABORDAGENS SÃO REALIZADAS: (i) USO DO CONJUNTO DE FRAÇÕES PARA O PROCESSO DE DECONVOLUÇÃO ESPECTRAL (1B7, 1B8 E 1B10) EM STOCYSY E RANSY E, (ii) USO SINGULAR DA FRAÇÃO B8 PARA EXPERIMENTOS DE DOSY (DECONVOLUÇÃO), ^1H , g-HSQC, g-HMBC, COSY E TOCSY1D (ELUCIDAÇÃO ESTRUTURAL). AO FINAL OS RESULTADOS SÃO COMBINADOS E É DETERMINADA UMA DAS ESTRUTURAS PRESENTE NA FRAÇÃO 1B8.	136

LISTA DE ABREVIATURAS

AMDIS	<i>Automated Mass Spectral Deconvolution and Identification System</i> (Sistema automatizado de deconvolução e identificação espectral)
BS	Biologia sistêmica (<i>systems biology</i>)
C18	Octadecilsilano
CAS-number	Número de registro único no banco de dados do <i>Chemical Abstracts Service</i> , (ACS)
CG	Cromatografia a gás
CID	<i>Collision Induced Dissociation</i> (Dissociação induzida por colisão)
CLAE	Cromatografia Líquida de Alta Eficiência
CLAE-DAD-EM/EM	Cromatografia líquida de alta eficiência acoplada a detector de arranjo de diodos e espectrômetro de massas sequencial
Da	unidade de medida em Daltons
Deoxi	Deoxihexose
DNA	Ácido desoxirribonucleico (DeoxyriboNucleic Acid)
DNP	Dicionário de Produtos Naturais (<i>Dictionary of Natural Products</i>)
DPM	Desvio padrão médio
DOSY	<i>Diffusion Ordered Spectroscopy</i> (Espectroscopia de difusão ordenada)
EFP	Extração via fluído pressurizado
EI	Ionização por elétrons
EM	Espectrometria de Massas
EM/EM	Espectrometria de Massas sequencial
ESI	Electrospray ionization (Ionização por <i>electrospray</i>)
eV	eletronvolt
FAME	Ésteres metílicos de ácidos graxos (fatty acid methyl esthers)

FDA	<i>Food and Drug Administration</i>
FDC	Fator de detecção de compostos
FM	Fórmula Molecular
GA	Glicoalcaloides
HOAc	ácido acético
Hex	hexose
HTS	<i>High-throughput screening</i> (triagem em alta escala)
IBG	Isolamento bio guiado
IE	Ionização por elétrons
InChI	Identificador químico internacional (IUPAC)
InChi Key	Identificador padronizado de compostos químicos (IUPAC)
IRcalc.	Índice de retenção linear calculado
IRlit.	Índice de retenção linear da literatura
IRL	Índice de retenção linear
<i>m/z</i>	Razão massa-carga
MF	Fator de similaridade (<i>matching factor</i>)
MSTFA	<i>N</i> -metil- <i>N</i> -[trimetilsilil]-trifluoroacetamida
<i>mRNA</i>	Ácido ribonucleico mensageiro (<i>Messenger ribonucleic acid</i>)
NaTFA	ácido trifluoroacetato de sódio
NuBBE	Núcleo de Bioensaios, Biossíntese e Ecofisiologia de Produtos Naturais
P&D	Pesquisa e desenvolvimento
Pen	pentose
PN	Produtos naturais
ppm	partes por milhão
Pubchem CID	Identificador de compostos da bases de dados Pubchem
QC	Química combinatória
QPN	Química de produtos naturais
QToF	Analísadores do tipo quadrupolo e de tempo de voo sequenciais

RANSY	Ratio Analysis of NMR Spectroscopy
RDA	Retro Diels-Alder
RMN	Ressonância Magnética Nuclear
SMILES	Sistema simplificado de linha entrada de dados moleculares lineares (<i>Simplified Molecular Input Line Entry System</i>)
SPE	Extração em fase sólida (<i>solid phase extraction</i>)
STOCSY	<i>Statistical Total Correlation Spectroscopy</i>
TCM	Medicina tradicional chinesa (<i>Traditional Chinese Medicine</i>)
TMCS	trimetilclorosilano
ToF	Analizador do tipo tempo de voo (<i>Time of Flight</i>)
tR	Tempo de Retenção
UV/Vis	Ultravioleta/Visível

SUMÁRIO

LISTA DE FIGURAS.....	XIV
LISTA DE TABELAS.....	XVIII
LISTA DE ESQUEMAS.....	XIX
LISTA DE ABREVIATURAS	XX
SUMÁRIO	23
PREFÁCIO	25
PREFACE	27
CAPÍTULO 1 - A IMPORTÂNCIA DOS PRODUTOS NATURAIS NO SÉCULO XXI E O ESTUDO DO TÁXON	
<i>SOLANUM</i>	29
A IMPORTÂNCIA DOS PRODUTOS NATURAIS NO SÉCULO XXI	30
TÁXON <i>SOLANUM</i>	36
DESENVOLVIMENTO DE HIPÓTESES	40
CAPÍTULO 2 – USO DA CROMATOGRAFIA ACOPLADA A ESPECTROMETRIA DE MASSAS NO ESTUDO	
METABOLÔMICO DE DE ESPÉCIES DE <i>SOLANUM</i>: MÉTODOS DE DECONVOLUÇÃO, DESREPLICAÇÃO	
E AGRUPAMENTO MOLECULAR PARA DETECÇÃO DE SUBSTÂNCIAS.....	44
INTRODUÇÃO	45
<i>Algumas considerações sobre coleta e preparo de amostras</i>	<i>46</i>
<i>Algumas considerações sobre análise em CG-EM e CLAE-EM.....</i>	<i>48</i>
<i>Algumas considerações sobre a identificação de compostos via espectrometria de massas</i>	<i>50</i>
MATERIAIS E MÉTODOS	51
<i>Coleta e classificação do material vegetal</i>	<i>51</i>
<i>Extração e secagem.....</i>	<i>52</i>
<i>Solventes e reagentes empregados.....</i>	<i>53</i>
<i>Preparo de amostras para as análises cromatográficas</i>	<i>53</i>
<i>Configuração Instrumental: CG-EM</i>	<i>54</i>
<i>Configuração Instrumental: CLAE-EM e CLAE autoEM/EM.....</i>	<i>56</i>
<i>Construção de uma base de dados de Solanum</i>	<i>59</i>
<i>Softwares</i>	<i>59</i>
<i>Desreplcação por CG-EM.....</i>	<i>60</i>
<i>Desreplcação por CLAE-EM</i>	<i>60</i>
<i>Desreplcação CLAE-autoEM/EM utilizando rede de interações moleculares (Molecular</i>	
<i>networking)</i>	<i>61</i>
RESULTADOS E DISCUSSÃO	63
<i>Perfil metabolômico e desreplcação de metabólitos primários através de CG-EM e AMDIS</i>	<i>63</i>
<i>Perfil metabolômico e desreplcação de metabólitos secundários através de CLAE-EM-EM/EM .</i>	<i>73</i>

<i>Nucleosídeos</i>	77
<i>Amino ácidos e Derivados</i>	78
<i>Flavonoide O-glicosilados</i>	79
<i>Poliaminas e Amidas</i>	79
<i>Alcaloides Esteroidais</i>	84
<i>Perfil metabolômico e desreplicação de metabólitos secundários através de CLAE-AutoEM/EM e redes de interações moleculares (molecular networking)</i>	92
<i>Flavonoides O-glicosilados</i>	95
<i>Flavonoides C-glicosilados</i>	101
<i>Glicoalcaloides esteroidais / Saponinas</i>	104
CONCLUSÃO	106
 CAPÍTULO 3 – USO DA CROMATOGRAFIA ACOPLADA A RESSONÂNCIA MAGNÉTICA NUCLEAR NO ESTUDO METABOLÔMICO EM PLANTAS: MÉTODOS DE DECONVOLUÇÃO, DESREPLICAÇÃO E ELUCIDAÇÃO ESTRUTURAL	
INTRODUÇÃO	109
<i>CLAE-SPE-RMN</i>	112
<i>Considerações sobre identificação de compostos via RMN em estudos metabolômicos</i>	113
<i>Ratio Analysis Nuclear Magnetic Resonance Spectroscopy (RANSY)</i>	114
<i>Statistical Total Correlation Spectroscopy (STOCSY)</i>	115
<i>Diffusion Ordered Spectroscopy (DOSY)</i>	116
MATERIAIS E MÉTODOS	117
<i>Coleta e classificação do material vegetal</i>	117
<i>Extração e secagem</i>	117
<i>Solventes e reagentes empregados</i>	117
<i>CLAE-SPE-RMN</i>	118
<i>Preparação de amostras em CLAE-SPE</i>	118
<i>Análises por CLAE-SPE</i>	118
<i>Análises por RMN</i>	119
<i>Experimentos DOSY – Princípio da técnica</i>	120
<i>Análise por RANSY</i>	122
<i>Análise por STOCSY</i>	123
<i>Softwares</i>	124
RESULTADOS E DISCUSSÃO	125
<i>CLAE-SPE-RMN em modo seleção</i>	125
<i>CLAE-SPE-RMN em modo time slice</i>	134
CONCLUSÕES	149
CONCLUSÕES GERAIS E PERSPECTIVAS	151
REFERÊNCIAS	154

PREFÁCIO

Uma nova era em bioquímica de plantas, ao nível de sistemas, vem sendo incorporada na qual fenômenos bióticos e abióticos são entendidos através da correlação multifatorial de dados moleculares. Este campo emergente é direcionado a caracterização de moléculas pequenas (metabólitos) que atuam como substratos, ligantes ou entidades de sinalização em células. No presente estudo, são estabelecidas metodologias e estratégias para triagem e identificação de metabólitos (desreplicação) com base em estudos metabolômicos. Foram analisados metabólitos das folhas e ramos de espécies *Solanum* (*S. paniculatum*, *S. americanum*, *S. swartzianum* e *S. excelsum*) fazendo uso de técnicas analíticas do estado da arte em espectrometria e espectroscopia. Essas plantas foram escolhidas devido aos atributos nutricionais e medicinais, além de serem geneticamente próximas ao tomate, o qual representa um modelo para a família Solanaceae.

Afim de ampliar toda a faixa de detecção e identificação de metabólitos, foram utilizadas tecnologias acopladas entre as cromatografias a gás e líquida com a espectrometria de massas (CG-EM e CLAE-EM) e a ressonância magnética nuclear (CLAE-SPE-RMN). Adicionalmente, foram construídas bases de dados de metabólitos a partir de dados moleculares da razão massa/carga e deslocamento químico, assim como ferramentas quimiométricas e computacionais foram utilizadas para o aperfeiçoamento das técnicas de desreplicação.

Uma introdução geral (capítulo 1) descreve a importância histórica dos produtos naturais e o seu atual contexto no mercado farmacêutico. Adicionalmente, informações biológicas, químicas e econômicas relativas ao gênero *Solanum* são discutidas em detalhes. Os capítulos 2 e 3 descrevem a evolução dos métodos de desreplicação e suas aplicações em metabolômica de plantas.

O segundo capítulo descreve o desenvolvimento e implementação de metodologias baseadas em espectrometria de massas, utilizando de CG-EM e CLAE-DAD-EM para a análise e identificação de metabólitos primários e secundários presentes nas espécies de *Solanum*. Adicionalmente são apresentados protocolos com informações sobre os reagentes utilizados, preparação de amostra, configuração instrumental, métodos de análise e estratégias de processamento de dados.

O terceiro capítulo traz informações quanto ao uso da técnica de ressonância magnética nuclear para a identificação de metabólitos secundários. Uma seleção de experimentos de RMN mono e bidimensionais (^1H , HSQC, HMBC e TOCSY1D) foram realizados para as frações obtidas no sistema CLAE-SPE-RMN. Adicionalmente a descrição de experimentos pseudo-bidimensionais como DOSY assim como a aplicação de ferramentas quimiométricas como RANSY e STOCYSY são apresentadas com o propósito de acelerar o processo de triagem e principalmente superar as dificuldades relativas a elucidação estrutural de compostos em misturas complexas.

PREFACE

A new era in biochemistry of plants at the level of systems has been incorporated in which biotic and abiotic phenomena are understood by multifactorial correlation of molecular data. This emerging field is oriented towards the characterization of small molecules (metabolites) that act as substrates, ligands or signaling entities in cells. In this study, methodologies and strategies are established for screening and identification of metabolites (dereplication) based on metabolomic studies. Metabolites were analyzed from branches and leaves of *Solanum* species (*S. paniculatum*, *S. americanum*, *S.* and *S. swartzianum excelsum*) using state of the art analytical techniques in spectrometry and spectroscopy. These plants were selected because of their nutritional and medicinal attributes,

In order to extend the full range of detection and identification of metabolites, coupling technologies were used between the gas and liquid chromatography with mass spectrometry (GC-MS and HPLC-MS) and nuclear magnetic resonance (HPLC-SPE-NMR). In addition, a metabolic database was built using molecular mass and chemical shift data. Chemometrics and computational tools were used for assist in all dereplication steps.

A general introduction (chapter 1) describes the historical importance of natural products and its current context into pharmaceutical market. Biological, chemical and economic information about *Solanum* genus is discussed as well. Chapters 2 and 3 describe the evolution of dereplication methods and their applications in metabolomics plants.

The second chapter describes the development and implementation of mass spectrometry-based methodologies (GC-MS and HPLC-DAD-MS) for the analysis and identification of primary and secondary metabolites present in the species of *Solanum*. Protocols with information about the reagents, sample preparation, instrumental configuration, analysis methods and data processing strategies are described.

The third chapter provides information on the use of nuclear magnetic resonance technique to identify the secondary metabolites. A selection of mono and two-dimensional NMR experiments (1H, HSQC, HMBC and TOCSY1D) were performed for fractions obtained from HPLC-SPE-NMR system. In addition the

description of pseudo-two-dimensional experiments DOSY as well as the application of chemometric tools as RANSY and STOCYSY are presented in order to speed up the screening process and mostly overcome the difficulties of the structural elucidation of compounds presents in complex mixtures.

CAPÍTULO 1 - A IMPORTÂNCIA DOS PRODUTOS NATURAIS NO SÉCULO XXI E O ESTUDO DO TÁXON *SOLANUM*



A IMPORTÂNCIA DOS PRODUTOS NATURAIS NO SÉCULO XXI

Os recursos naturais sempre foram imprescindíveis para o desenvolvimento e manutenção da qualidade de vida. A diversidade de produtos obtidos dessa fonte incluem principalmente os alimentos, óleos, carboidratos, proteínas, fibras e os biocombustíveis (BAKER et al., 2007; BARREIRO; BOLZANI, 2009; PINTO et al., 2002; LUBBE; VERPOORTE, 2011).

Apesar da grande importância dos metabólitos primários como matéria-prima para inúmeros processos, plantas, fungos, bactérias e animais são cultivados e/ou extraídos de seu habitat em busca de substâncias especiais, os chamados metabólitos especializados ou “secundários”. Essa química fina serve de base na indústria agroquímica para a produção de pesticidas, herbicidas, fabricação de produtos cosméticos, perfumes, corantes e principalmente no desenvolvimento de fármacos (BARREIRO; BOLZANI, 2009; LUBBE; VERPOORTE, 2011).

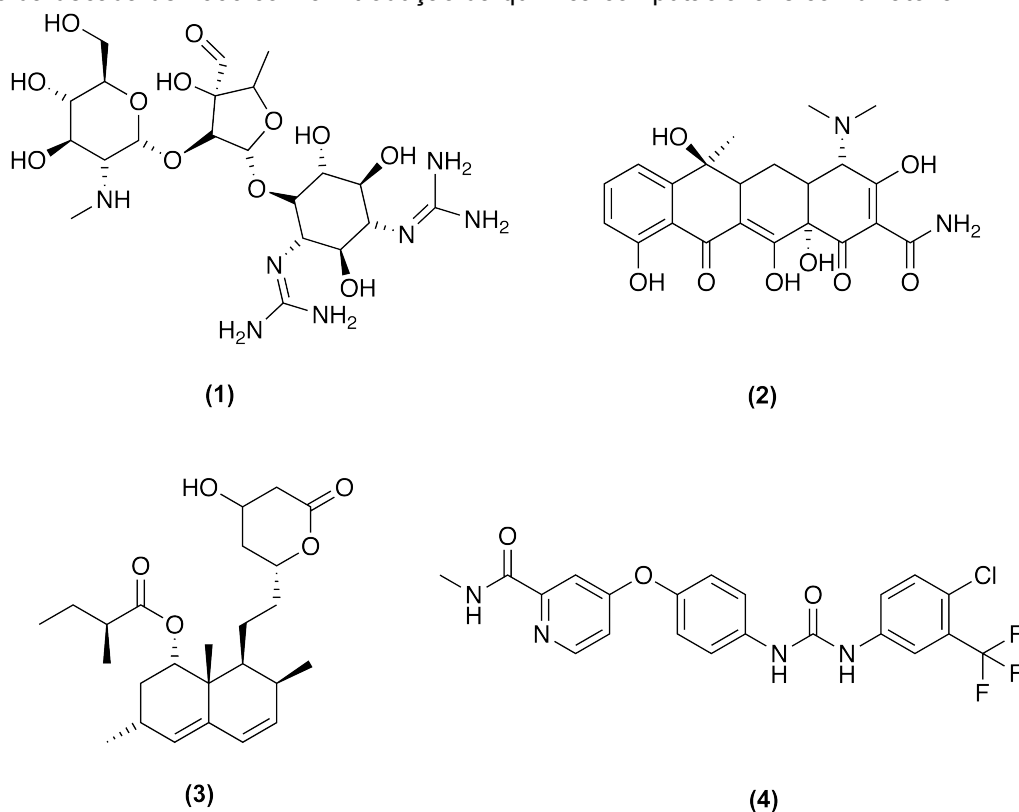
Dentro deste contexto, a química de produtos naturais (QPN) têm buscado e investigado por substâncias ativas em biotas terrestres e marinhos. Historicamente, a QPN possui um forte apelo terapêutico e não raro ela é a principal fonte de diversidade química para a descoberta de novos medicamentos (BARREIRO; BOLZANI, 2009; BUTLER, 2004; PUPO et al., 2007; NEWMAN; CRAGG, 2007, 2012, 2014).

De fato, o florescimento da QPN como ciência de fronteira quimio-farmacêutica ocorreu durante a segunda guerra mundial. A demanda por antibióticos impulsionou os programas de pesquisa dentro das companhias farmacêuticas por novas metodologias de triagem metabólica. A chamada *golden era* (1940-1980) ofereceu novas possibilidades no campo da medicina através da descoberta e produção de um arsenal terapêutico nunca visto antes. Para se ter uma ideia, a expectativa de vida era 44 anos no início do século XX e saltou para os 77 anos nos meados da década de 1990 (BAKER et al., 2007; LI; VEDERAS, 2009).

Descobertas como a estreptomicina (1), gentamicina, tetraciclina (2) e outros antibióticos não só auxiliaram no aumento da expectativa de vida mas serviram de estímulo para desenvolvimento de novos programas de pesquisa em torno dos produtos naturais. A compactina e mevinolina (3), inibidores da biossíntese do colesterol, são frutos da busca de protótipos nos programas P&D, figura 1. A partir

desses fármacos foram criadas as chamadas terapias de estatinas, na qual ainda hoje são um grande sucesso terapêutico e econômico (BAKER et al., 2007; BUTLER, 2004).

Figura 1. Agentes terapêuticos desenvolvidos entre 1950-1970 a partir da química de produtos naturais - (1) streptomicina; (2) tetraciclina e (3) mevinolina. O Sorafenib (4) foi desenvolvido durante no início da década de 2000 com a introdução da química computacional e combinatória.



Newman & Cragg lançaram três revisões (2000, 2007, 2012) da qual mostram que aproximadamente 50% dos fármacos de uso clínico presentes no mercado farmacêutico são origem natural. De modo geral, esses compostos são encontrados como purificados, derivados, análogos ou então obtidos através de processos de semi-síntese. Em último caso, os metabólitos encontrados em fontes naturais têm servido de inspiração para a síntese de novas entidades químicas (NEWMAN; CRAGG, 2007, 2012, 2014).

No caso dos medicamentos usados no tratamento do câncer ou doenças infecciosas esse percentual é ainda maior, cerca de dois terços desses agentes são de origem natural (NEWMAN; CRAGG 2007, 2012, 2014).

Apesar do sucesso dos PN, as décadas de 1990 e 2000 foram marcadas pela diminuição ou extinção de grande parte dos programas de P&D em PN. A introdução de ensaios em *Automated High Throughput Screening* (HTS), processo do qual um

grande número de amostras são testados frente a um alvo molecular, em grande parte contribuiu para desinteresse nos PN (LI; VEDERAS 2009; PINTO et al., 2002).

A rapidez dos ensaios e a baixa diversidade dos alvos proteicos exigia uma demanda por um *pool* de metabólitos puros na qual os PN não podiam atender. Além da limitação em número de compostos em comparação a bibliotecas de sintéticos, os *hits* de fontes naturais, em maioria, apresentavam estruturas mais complexas com numerosos substituintes com funções oxigenadas e uma abundância de centros estereogênicos. Isto dificultava o processo de identificação além de contribuir para problemas de escalonamento, suprimento e rendimento industrial (BUTLER, 2004; LI; VEDERAS 2009; NEWMAN; CRAGG, 2012).

Problemas relacionados a sazonalidade e as variações ambientais da composição dos organismos aumentaram ainda mais as barreiras para a pesquisa em PN (LI, VEDERAS 2009). Questões políticas sobre a propriedade intelectual, defesa da biodiversidade (patrimônio da fauna e flora) e os subsequentes *royalties* para as comunidades tradicionais nas quais os possíveis compostos bioativos seriam extraídos, finalmente determinou o abismo existente entre as grandes farmas e os PN (LI; VEDERAS, 2009).

A química combinatória (QC), em contrapartida, prometia preencher tal demanda criando uma imensidão de compostos para os testes em HTS sendo altamente reprodutíveis e de baixa complexidade. A QC seria totalmente transponível para escala industrial, em casos de aprovação pela FDA ou outros órgãos reguladores. Sob o ponto de vista político-econômico, a QC eliminava todas as questões diplomáticas e burocráticas envolvidas em qualquer tipo de extrativismo. Em resumo, o custo de investimento em descoberta e subsequente desenvolvimento operacional seria muito mais baixo e menos burocrático que o uso dos PN (BAKER et al. 2007; BUTLER 2004; LI; VEDERAS, 2009; NEWMAN; CRAGG 2012).

Notáveis exceções permaneceram com a abordagem baseada em PN como Novartis, Merck e Bayer enquanto GlaxoSmithKline, Sanofi-Aventis, Pfizer, Bristol-Myers Squibb, Eli Lilly, Hoffman-La Roche e Schering-Plough acataram a abordagem da QC e terminaram suas atividades em descobertas por meio de PN. (BAKER et al., 2007; BUTLER, 2004).

Contudo, os resultados da QC foram aquém das expectativas e, até meados da década de 1990 haviam dúvidas quanto a utilidade dessa metodologia na

descoberta de candidatos. Em uma entrevista publicada na revista Chemical e Engineering News, Lipinski chegou a afirmar:

- *The combinatorial libraries in the early years were so flawed that if you took the libraies across Pharma and stored them in dumpsters you would have improved productivity* (ROUCHI, 2003).

Até o momento, somente o fármaco Sorafenib (4), um inibidor de multiquinase indicado para câncer renal em estágio avançado foi obtido a partir derivado da química combinatória para uso clínico (NEWMAN; CRAGG, 2007, 2012).

O fato é que o uso da HTS combinado a QC aumentou a capacidade de triagem de compostos, mas o número de novas entidades químicas com valor medicinal caiu drasticamente. Em 2002, chegou-se ao nível mais baixo entre as entidades químicas em fase de testes pré-clínicos e clínicos desde o surgimento do HTS como metodologia de triagem (BUTLER, 2004 e 2005; LI; VEDERAS, 2009; ROUGUI, 2003).

A questão que permanece é: a queda é devido às limitações dos aspectos inovativos do processo de descoberta, por exemplo, geração de compostos utilizando um limitado espaço químico (ineficácia das biblioteca de compostos sintéticos e o baixo número de alvos proteicos) ou é devido à ausência da diversidade estrutural dos produtos naturais no processo de triagem?

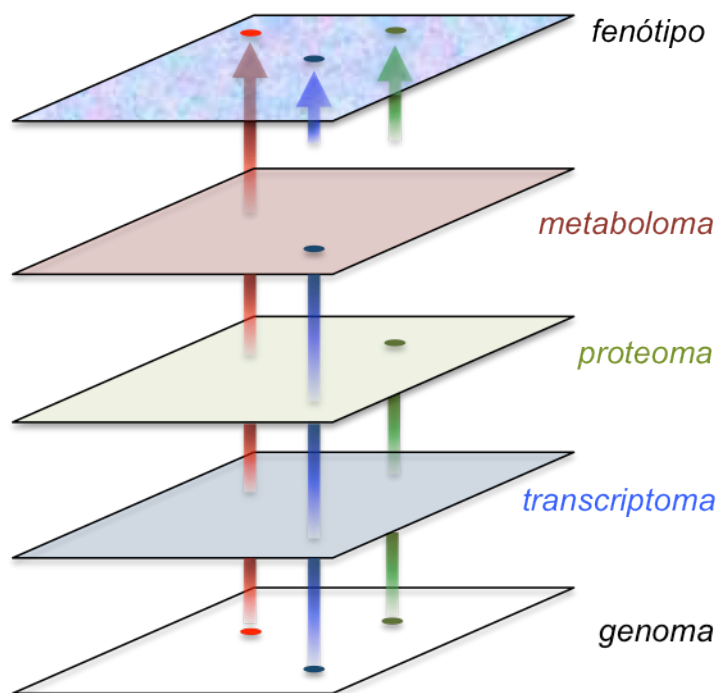
De todo modo, o que realmente pode-se afirmar é que metodologias baseadas em uma abordagem reducionista, na qual foca a pesquisa por mecanismos químicos através de um alvo terapêutico não deve ser a única opção. A combinação do conhecimento em nível molecular com concepções holísticas podem ser um atrativa alternativa para o uso dos PN. Um bom exemplo é o trabalho de Proudfoot (2002) no qual reportou que 8 das 29 fármacos lançados em 2000 foram derivadas de PN ou hormônios e, o HTS não apresentou impacto significativo para o processo de descoberta de tais estruturas (BUTLER, 2005; PROUDFOOT, 2002)

Portanto, a mudança do paradigma para uma visão mais holística, como são as filosofias da medicina tradicional chinesa (TCM – Traditional Chinese Medicine) e Ayurveda, podem servir de suporte para retomada dos PN.

Nesta concepção, um unguento (ou tecnicamente um extrato) na qual consiste de uma mistura de moléculas estruturalmente complexas é visto como uma entidade única com efeitos multifatoriais em um dado organismo. Sob uma perspectiva molecular, a descoberta de drogas envolveria questões integrativas aos

organismos, como as interações e interconversões moleculares entre os indivíduos e seus ambientes passando pelo DNA, mRNA, proteínas, metabólitos e, finalmente o fenótipo, figura 2 (BAKER et al., 2007; CAPRA, 2007; VERPOORTE et al., 2005).

Figura 2. Esquema ilustrativo do fenótipo de um organismo a partir do produto da conexão intra-específica entre sistemas genômicos, transcriptômicos, proteômicos e metabolômicos.



Dentro dessa perspectiva, a QPN tem buscado alternativas para o desenvolvimento de drogas potenciais e, fundamentalmente quatro novas abordagens veem sendo incorporadas: (i) a biologia de sistemas (BS), que considera as patologias e os mecanismos terapêuticos como processos multifatoriais; (ii) metabolômica, que busca, no perfil metabólico total (do doente e do agente terapêutico), informações e correlações relevantes ao desenvolvimento de novos agentes terapêuticos; (iii) quimiometria, que permite a análise dos complexos processos multifatoriais e sua interpretação, e; (iv) desreplicação, que simplifica o processo analítico através da identificação *in silico* de moléculas previamente conhecidas (HUR et al, 2012; LANG et al., 2008; LI; VEDERAS, 2009; NICHOLSON; LINDON, 2008; ULRICH-MERZENICH, et al. 2007).

Desta forma, a partir de estudos integrados dos componentes e comportamentos dinâmicos dos sistemas biológicos é possível determinar respostas multiparamétricas relativas a um dado fenômeno (LEVESQUE; BENFEY, 2004; FUNARI et al., 2013; NICHOLSON; LINDON, 2008;).

Utilizando de uma visão holística, sistemas celulares ou organismos íntegros (animais, órgãos, cultura de tecidos, etc.) são analisados através da integração entre ensaios clínicos e modelos experimentais (e.g. uso de animais), na qual observações clássicas de fisiologia (sangue, pressão ou alguma atividade biológica) são combinadas com ferramentas modernas de observação, como a expressão gênica e proteica.

A metabolômica por sua vez, tratou da análise de tais fenômenos como função de variação dinâmica do metaboloma (PATTI et al., 2012; NICHOLSON; LINDON, 2008; NIELSEN; JEWETT, 2007; VERPOORTE et al., 2010; YULIANA et al., 2011). Contudo, a exploração de moléculas biologicamente ativas em matrizes complexas apresenta um desafio às técnicas analíticas modernas devido a grande quantidade de dados e principalmente a complexidade intrínseca deles (WOLFENDER et al., 2009 e 2013). Assim, métodos quimiométricos são necessários para o planejamento experimental e, principalmente para tratamento de dados de forma inteligível.

No entanto, a tarefa mais complexa em qualquer estudo metabolômicos talvez esteja envolvida no processo de identificação estrutural. A complexidade e diversidade estrutural dos PN criam as principais dificuldades para a triagem de biomarcadores (WOLFENDER et al., 2009 e 2013).

Nesse contexto, a desreplicação vem sendo incorporada aos estudos metabolômicos como ferramenta de triagem molecular. Em geral, ela simplifica o processo analítico de identificação estrutural através do uso de ferramentas analíticas acopladas, como a HPLC-DAD-EM, juntamente a análise *in silico* de dados através da comparação em bases de dados. Neste processo, moléculas já reportadas são rapidamente identificadas resultando na priorização de moléculas bioativas inéditas relativas a um dado estímulo (LANG et al., 2008; LI; VEDERAS et al., 2009; NICHOLSON et al., 2008; ULRICH-MERZENICH et al., 2007).

Em suma, os avanços tecnológicos no último século permitiram não somente uma maior eficiência científica mas levantaram questionamentos sobre conceitos da visão mecanicistas do mundo atual. Vivemos num momento de transição conceitual em vários setores da sociedade e não diferente está para ciência. Abordagens sistêmicas e holísticas cada vez mais recorrentes em pesquisas de fronteira impulsionam o desenvolvimento de metodologias e ferramentas. Para o século XXI,

podemos esperar uma ciência sistêmica, integrativa, holística, colaborativa e principalmente multidisciplinar.

TÁXON *SOLANUM*

A família Solanaceae é constituída por 100 gêneros e mais de 2.500 espécies distribuídas nas regiões temperadas e tropicais, figura 3. As solanáceas representam uma das maiores culturas agrícolas do planeta, composta por batatas, tomates, pimentas e o tabaco. Sua importância econômica confere a ela estudos biológicos, químicos e tecnológicos. (OLMSTEAD et al., 2008).

O gênero *Solanum* destaca-se na família Solanaceae, composto por *c.a.* 1,400 espécies, o que representa cerca de 1% da flora de angiospermas. Sua alta taxa de especiação é rara nas angiospermas o que a torna alvo para estudos de adaptação ambiental e de processos evolutivos (POCZAI et al., 2008; WHALEN et al., 1979).

Inicialmente, a classificação taxonômica das espécies de *Solanum* foi baseada em características morfológicas como por exemplo a presença/ausência de espinhos ou características de anteras e cabelos (D'ARCY, 1979). No entanto, nenhum desses sistemas foi completamente satisfatório para compartimentar a diversidade morfológica do gênero. Somente em meados da década de 2000, o trabalho de Bohs et al. (2007) usando dados moleculares do gene de cloroplasto *ndhF* classificou os subgêneros e seções seguindo uma metodologia cladística. Atualmente, a filogenia do gênero é baseada na combinação de dados genômicos de *ndhF* de cloroplasto, *F-trnT* e *waxy* (BOHS et al., 2007).

Figura 3. Distribuição geográfica da família Solanaceae (em verde) no mundo, acessado de <http://www.mobot.org/mobot/Research/AP/web/>



A química do gênero *Solanum* é marcada pela presença de esteroides, saponinas, flavonoides, terpenos e glicoalcaloides (GA), sendo os dois últimos os quimiomarcadores do gênero (BREITHAUPT et al., 2002; MILNER et al., 2011; REN et al., 2009; SUN et al., 2006; SUTHAR et al., 2008; WEISSENBERG et al., 2001; ZHANG et al., 2010). De modo geral, as agliconas dos GA possuem um esqueleto colestano hidrofóbico de 27-carbonos com um nitrogênio incorporado ao anel F ou anel A na forma de anéis piperidínicos ou aminas primárias. As cinco principais classes de glicoalcaloides presentes no gênero são: solanidanos (1), nos quais apresentam anéis indolizidínicos fundidos, os espirosolanos (2), na qual apresentam um sistema oxa-azaespirodecano, (22,26)-epiminocolestanos (3), 3-aminoespirostanos (4) e R-epiminocicloemicetal (5), figura 4 (MILNER et al., 2011;).

Do ponto de vista biossintético estes compostos são análogos nitrogenados das sapogeninas e assim como elas, são biosintetizados a partir da rota do colesterol com subseqüentes modificações na cadeia lateral. Segundo Dewick 2002, o processo final de biossíntese dos espirosolanos ocorre através da aminação na posição C-26 do derivado 26-hidroxicolesterol fazendo uso da L-arginina como precursor. Após uma hidroxilação em C-22 ocorre a ciclização via uma substituição nucleofílica entre a amina primária e o carbono secundário em C-22 formando o anel piperidínico. Para a etapa de formação do núcleo esteroidal é proposto uma segunda hidroxilação em C-16 resultando numa adição a imina, figura 5.

Aparentemente todas as outras classes de glicoalcaloides presentes nas espécies de *Solanum* seguem o mesmo padrão da rota do colesterol e diferem uma das outras no processo de ramificações de suas cadeias laterais (DEWICK, 2002).

Figura 4. Estrutura base das cinco principais classes de glicoalcaloides encontrados em espécies de *Solanum*. O primeiro esqueleto foi inserido a critério de nomenclatura e numeração dos anéis dos glicoalcaloides esteroidais. 1 – solanidanos; 2 – espirosolano; 3 – (22,26) epiminocolostano; 4 – 3-aminoespirostando e 5 – α -epiminocicloemietal.

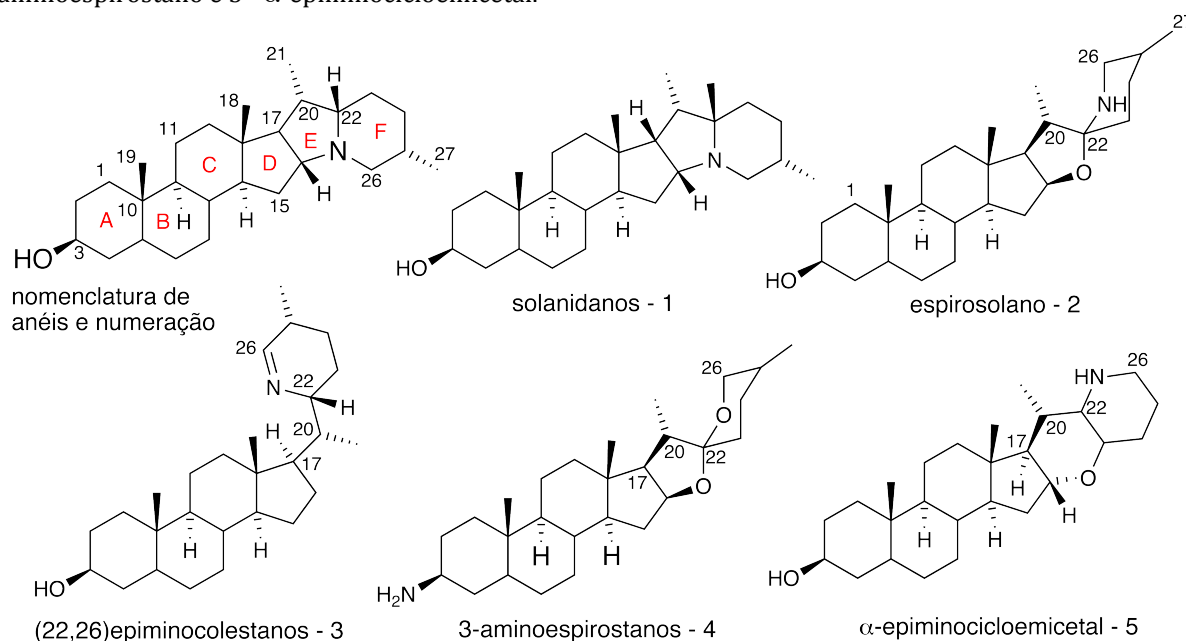
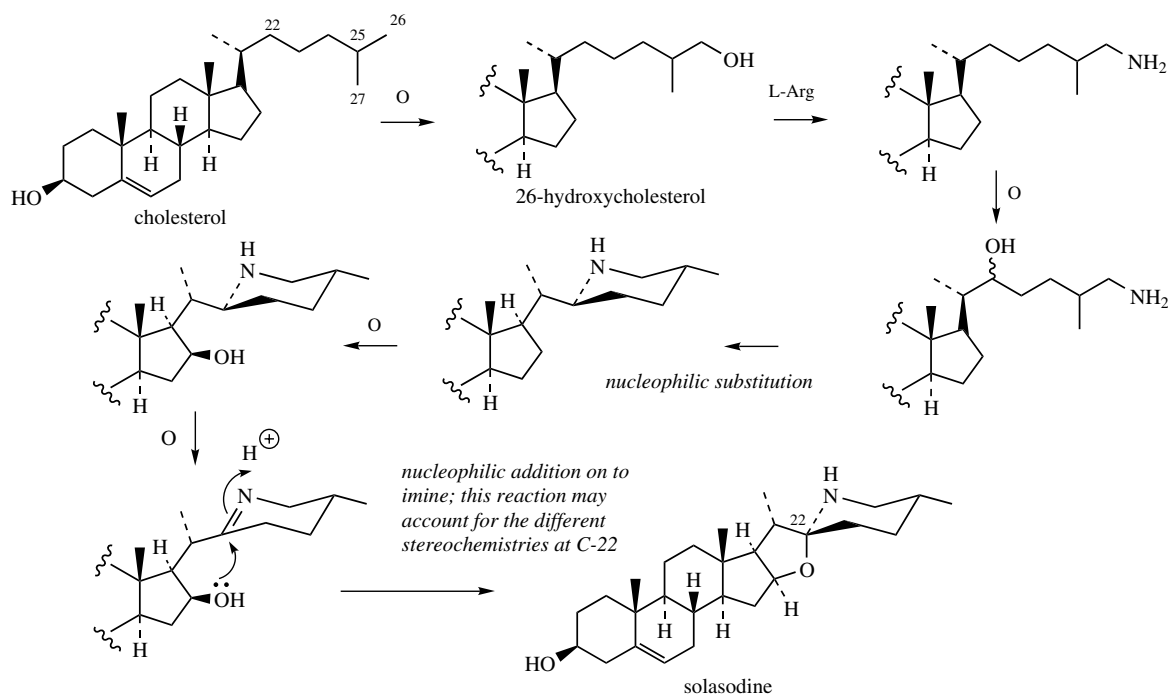


Figura 5. Proposta biossintética da formação de alcaloides espirosolanos a partir da rota do colesterol. extraído de Dewick, 2002, p.390.



Dentre todos os glicoalcaloides encontrados em *Solanum*, a solasodina é sem dúvida o composto mais recorrente. Na forma de glicosídeo, esse composto presente em folhas e frutos pode ser encontrado em até 2% em massa em determinadas espécies (*S. liciniatum* e *S. marginatum*). A alta concentração da solasodina e seus derivados tem despertado o interesse do mercado farmacêutico

por espécies de *Solanum*, uma vez que esses compostos podem ser convertidos a progesterona por meio de degradações de Marker (DEWICK, 2002).

Dentro da perspectiva terapêutica inúmeros exemplos mostram o potencial farmacológico das espécies de *Solanum*. Ikeda et al. (2003), avaliou atividade antitumoral de uma série de glicosídeos esteriodais encontrados em mais de 40 espécies de *Solanum*. O resultado mostrou que a desgalactotigonina, obtido de *Solanum lyratum* é o mais ativo sendo dez vezes mais potente que o padrão *cis*-diaminodicloroplatina (II) (CDDP) (IKEDA et al., 2003).

Mesia-Vela et al. (2002) utilizando de uma abordagem holística avaliou o extrato aquoso das raízes de *S. paniculatum* em ratos submetidos a ulceração via etanol. Os resultados mostraram uma atividade anti-úlceras via inibição de secreção de ácido gástrico (ED_{50} 448 mg.Kg⁻¹). A espécie *Solanum americanum* apresentou atividade anti-fúngica contra *Cryptococcus neoformans* (200 mg.mL⁻¹) e frente ao estágio epimastigota e tripomastigota de *Trypanosoma cruzi* (CÁCERES et al., 1998).

Inúmeros exemplos mostram que espécies de *Solanum* atuam como agentes terapêuticos para uma série de enfermidades, como câncer, doenças virais, disfunções hepatológicas e gástricas. Muitas delas ainda são ativas contra diversas espécies de bactérias, fungos e vírus (BARI, et al. 2010; NGUELEFACK et al., 2008; LI, et al., 2008; PEREZ, et al., 2006).

A potencial farmacológico de espécies de *Solanum* endêmicas do Brasil levou a inserção da espécie *Solanum paniculatum* (popularmente conhecida como jurubeba) na lista de 77 plantas medicinais lançadas em 2008 pela secretaria de Ciência, Tecnologia e Insumos Estratégicos do Brasil (RENISUS) (BRASIL, 2006).

De acordo com a portaria estabelecida, a fitoterapia aplicada no Sistema Único de Saúde (SUS) desta espécie (*Solanum paniculatum*) é relativa as atividades antiemética, anti-nauseante, anti-infecciosa, analgésica, anti-inflamatória e anti-pasmódica (BRASIL, 2006).

Portanto, a escolha do táxon Solanaceae teve como princípio explorar a biodiversidade ecológica e química de espécies brasileiras com potencial terapêutico e consequentemente, contribuir para o desenvolvimento da sociedade e a ciência brasileira.

DESENVOLVIMENTO DE HIPÓTESES

A diversidade estrutural dos produtos naturais tem constituído ao longo da história um dos pilares da descoberta e desenvolvimento de agentes terapêuticos. Ao mesmo tempo que a complexidade química e biológica dos PN permite uma maior afinidade aos alvos proteicos eles também criam barreiras e desafios no processo analítico de triagem molecular.

Tradicionalmente, o isolamento de compostos bioativos de fontes naturais tem seguido metodologias baseadas em isolamento bioguiado (IBG). Fazendo o uso de técnicas cromatográficas, compostos bioativos são repetidamente coletados e separados até a purificação. Em seguida, eles são analisados por técnicas espectrométricas e espectroscópicas e finalmente testados nos ensaios de HTS. Entretanto, esta abordagem clássica vem sendo questionada devido aos inúmeros desafios que a pesquisa moderna em PN tem encarado.

Em primeiro lugar, devemos nos perguntar se os resultados obtidos têm sido proporcionais aos tremendos esforços despendidos. O processo de purificação não somente requisita de uma intensiva mão-de-obra, mas também consome grandes quantidades de solventes e materiais caros e potencialmente impactantes para o meio ambiente. Em parte, estes problemas estão associados a complexidade química e a baixa concentração dos compostos bioativos nos PN. Afim de estabelecer uma metodologia mais eficiente e econômica para determinação de estruturas, **este estudo traz a hipótese que os processos convencionais de purificação podem ser acelerados e focados por métodos otimizados de preparação de amostra, separação e detecção de compostos (Hipótese A).**

Embora muitos compostos puros podem ser obtidos através IBG, os subsequentes resultados das análises espectroscópicas/espectrometria (por exemplo, UV-Vis, EM, RMN, etc.) podem trazer resultados decepcionantes como por exemplo, a elucidação de compostos que não apresentam interesse para as indústrias farmacêuticas ou re-isolamento de compostos previamente identificados.

A incerteza e o acaso de resultados tem encorajado vários grupos de pesquisa a adotar metodologias mais racionais na busca por moléculas de interesse. Entre elas destaca-se a desreplicação, na qual visa melhorar a eficiência do processo de isolamento e elucidação estrutural em PN. **Este estudo visa expandir**

a utilidade das ferramentas analíticas acopladas (CLAE, CG, EM e RMN) para a desreplicação de estruturas em PN. Adicionalmente, a combinação de EM e RMN com técnicas cromatográficas podem assistir em um planejamento mais eficiente de esquemas de fracionamento assim como gerar perfis metabolômicos das espécies em estudo. (Hipótese B).

O principal desafio nas análises de RMN de produtos naturais é o alto número de sobreposições de sinais de picos na qual muitas vezes interfere na interpretação dos dados de RMN. Em particular, a similaridade estrutural de congêneres de PN cria problemas no processo de desreplicação devido aos padrões de similaridade espectroscópica. Assim, torna-se difícil empregar metodologias convencionais, tal como a comparação de espectros obtidos em bases de dados com os resultados experimentais. **No entanto, este estudo traz a hipótese que o uso de ferramentas estatísticas de deconvolução permitirá a separação de sinais sobrepostos tornando mais fácil o processo de elucidação estrutural em RMN. (Hipótese C)**

Em estudos metabolômicos via CLAE-EM, a identificação molecular é baseada nos dados de massa de alta resolução e no perfil de fragmentação EM/EM. Assim, os resultados são comparados com bases de dados de massas como METLIN, MassBank, DNP, entre outras, levando a detecção de possíveis *Hit's* para um dado íon. Entretanto, devido a especificidade metabólica dos organismos muitas informações ainda não estão presentes nas bases de dados.

Deste modo, o que ocorre é uma interpretação manual dos dados de fragmentação, que pode levar entre minutos até várias horas dependendo da natureza e complexidade dos compostos. Do ponto de vista dinâmico (tempo) isto inviabiliza estudos metabolômicos, uma vez que centenas ou até milhares de compostos são detectados ao mesmo tempo. **Para este estudo, levanta-se a hipótese que o uso de ferramentas quimiométricas de correlação poderiam organizar de forma sistemática o conjunto de dados intra e inter-amostras. Para tanto, informações EM/EM deverão ser analisadas através de agrupamentos de similaridade. (Hipótese D)**

Nas próximas seções serão demonstrados uma variedade de ferramentas hoje disponíveis para melhorar e/ou acelerar a descoberta de agentes terapêuticos a partir da abordagem metabolômica. A conjunção entre os avanços analíticos e as ferramentas quimiométricas de deconvolução, agregação e comparação de dados

levarão ao desenvolvimento de estratégias rápidas e racionais para a elucidação estrutural em PN e assim podendo auxiliar na descoberta de novos agentes terapêuticos. A tabela 1 reporta um planejamento do estudo dividido em capítulos com os respectivos tópicos.

Tabela 1. O planejamento do presente estudo.

CAPÍTULO: TÓPICO	Aspecto da pesquisa em PNs	Aspecto da complexidade dos PNs	Hipóteses	Metodologia
CAPÍTULO 2 – USO DA CROMATOGRAFIA ACOPLADA A ESPECTROMETRIA DE MASSAS NO ESTUDO METABOLÔMICO DE PLANTAS: MÉTODOS DE DECONVOLUÇÃO, DESREPLICAÇÃO E AGRUPAMENTO MOLECULAR PARA DETECÇÃO DE SUBSTÂNCIAS	• Separação /Isolamento; • Perfil metabolômico • Desreplicação	• Diversidade estrutural & similaridade; • Variação do gradiente de concentração metabólica • Complexidade residual	Hipótese A Hipótese B Hipótese D	• Uso de softwares para adaptação e otimização de sistemas cromatográficos; • Deconvolução e desreplicação via CG-EM e CLAE-EM • Uso das mapas de rede molecular através de dados de massas sequencial para desreplicação
CAPÍTULO 3 – USO DA CROMATOGRAFIA ACOPLADA A RESSONÂNCIA MAGNÉTICA NUCLEAR NO ESTUDO METABOLÔMICO EM PLANTAS: MÉTODOS DE DECONVOLUÇÃO, DESREPLICAÇÃO E ELUCIDAÇÃO ESTRUTURAL	• Separação /Isolamento; • Desreplicação; • Elucidação;	• Diversidade estrutural & similaridade; • Variação do gradiente de concentração metabólica • Complexidade residual	Hipótese A Hipótese B Hipótese C	• Uso do acoplamento CLAE-SPE-TT-RMN para a desreplicação (modo seleção) • Estratégia de integração de ferramentas quimiométricas (RANSY e STOCYSY) com novas de sequências de pulso em RMN TCOSY1D e DOSY (modo <i>time slice</i>)

CAPÍTULO 2 – USO DA CROMATOGRAFIA ACOPLADA A ESPECTROMETRIA DE MASSAS NO ESTUDO METABOLÔMICO DE DE ESPÉCIES DE *SOLANUM*: MÉTODOS DE DECONVOLUÇÃO, DESREPLICAÇÃO E AGRUPAMENTO MOLECULAR PARA DETECÇÃO DE SUBSTÂNCIAS



INTRODUÇÃO

A metabolômica surgiu como tecnologia de análise e comparação da resposta metabólica multiparamétrica de seres vivos através de estímulos genéticos, ambientais ou patofisiológicos (NICHOLSON, 2008). Atualmente ela é usada numa infinidade de aplicações tanto em sistemas humanos, microbianos quanto em plantas (DIXON et al, 2006; HALL, 2006; NIELSEN; JEWETT, 2007).

Particularmente as plantas, como organismos sésseis, utilizam de sua maquinaria bioquímica para produção de uma panóplia de compostos estruturalmente complexos para proteção e adaptação ambiental (WENG, 2014). Essa multitude estruturalmente complexa (incluindo também a variação quantitativa) torna o processo analítico de extração e detecção um grande desafio para estudos metabolômicos (WOLFENDER et al., 2009; NICHOLSON, 2008).

Ao longo das últimas décadas, vários métodos voltados a metabolômica de plantas vem sendo desenvolvidos. A maioria deles fazendo uso de técnicas de acoplamento (separação e detecção), como por exemplo a cromatografia associada a espectrometria de massas e/ou ressonância magnética nuclear (CLAE-EM e CG-EM, EC-EM e CLAE-SPE-RMN) (BINGOL et al., 2015; DIXON et al., 2006; FILLET; FRÉDÉRICH, 2015; HALL, 2006; LISEC et al., 2006; PETERSON et al., 2014).

De fato, as plataformas baseadas em EM são atualmente as mais utilizadas em metabolômica de plantas devido principalmente a sensibilidade e as tecnologias de acoplamento, que em RMN ainda é um desafio (Hall, 2006). A CG-EM utilizando o analisador do tipo tempo de voo (time of flight – ToF) foi a primeira ferramenta usada no *screening* de metabólitos em larga escala (FIEHN et al., 2000). Um detalhado protocolo de CG-EM usado para extração, derivação e subsequente análise de dados foi descrito por LISEC et al. (2006).

A abordagem CG-EM utilizando de metodologias de derivação permite a análise de metabólitos não voláteis, principalmente aqueles envolvidos no metabolismo primário, como ácidos orgânicos, aminoácidos, carboidratos, álcoois, intermediários fosforilados (derivados da fração polar) além dos compostos lipofílicos tais como ácidos graxos e esteróis (fração apolar) (FIEHN et al., 2000; LISEC et al., 2006; PETERSON et al., 2014).

Por outro lado, a CLAE-EM agrega principalmente a bioquímica semi-polar, incluindo o grupo dos metabólitos especializados/secundários (GIKAA et al., 2014; GÓMEZ-ROMERO et al., 2010; HUR et al., 2013; JANDRIĆ et al., 2014). Dentro deste grupo de compostos podemos destacar os alcaloides, saponinas, ácidos fenólicos, fenilpropanoides, flavonoides, glicosinolatos, poliaminas e todos os seus derivados (GIKAA et al., 2014, BINGOL et al., 2015).

Algumas considerações sobre coleta e preparo de amostras

Embora não existam limitações quanto ao tipo de amostra, algumas recomendações devem ser consideradas para qualquer método de coleta e extração em metabolômica:

- **Padronização de coleta;**
- **Inibição enzimática;**
- **Homogeneidade amostral;**

A padronização de coleta remete a localidade e as condições de coleta. Assim, estudos metabolômicos devem considerar questões de sazonalidade, localidade e tempo. Qualquer amostragem deve conter no mínimo 6 indivíduos e ser realizada no mesmo período do ano, local/região e horário. Isso evita que os efeitos sazonais possam confundir ou interferir nos fenômenos de estudo (ÁLVAREZ-SÁNCHEZ et al., 2010; KIM; LISEC et al., 2006; VERPOORTE, 2010).

A inibição enzimática é outro aspecto fundamental para garantia da reprodutibilidade. No ato de coleta algumas enzimas desencadeiam diferentes reações que podem alterar o perfil metabólico. Deste modo, o uso de nitrogênio líquido ou metanol resfriado (-40°C) são necessários para a desnaturação de eventuais proteínas que podem levar a degradação química (ÁLVAREZ-SÁNCHEZ et al., 2010; KIM; VERPOORTE, 2010; MUSHTAQ et al., 2014).

A homogeneidade amostral tem por finalidade a padronização do material vegetal para a etapa de extração. Após submetido a imersão em nitrogênio líquido ou metanol (-40°C) o material vegetal é pulverizado com auxílio de um sistema de trituração, que pode ser desde gral e pistilo até sistemas de moinhos resfriados (KIM; VERPOORTE, 2010; ÁLVAREZ-SÁNCHEZ et al., 2010; MUSHTAQ et al., 2014). O importante nesses casos é que o material tenha uma uniformidade de

granulação na qual possibilite a entrada de solvente nas cavidades celulares resultando em maior eficiência e reprodutibilidade de extração (ÁLVAREZ-SÁNCHEZ et al., 2010).

Assim como no processo de amostragem, a etapa de extração é extremamente importante pois é ela que define qual tipo de metaboloma será observado. Em todos os casos, algumas recomendações devem ser tomadas durante a etapa de extração para estudos metabolômicos:

- **Rapidez**
- **Capacidade de Extração**
- **Reprodutibilidade**

Como em muitos estudos metabolômicos utilizam de grandes populações, métodos rápidos são necessários para que a execução do projeto seja viável em tempo quanto economicamente (MUSHTAQ et al., 2014).

A extensão talvez seja o aspecto mais importante de qualquer estudo metabolômico juntamente com a reprodutibilidade. A capacidade de extração de compostos estruturalmente diversos é determinante para a detecção da multitude de fenômenos desencadeados a partir de um dado estímulo. Assim, diversos métodos de extração tem sido avaliados como a maceração, ultrassonificação, microondas, extração em fase sólida (SPE) e fluído supercríticos. Contudo, a maceração continua sendo o método mais utilizado (ÁLVAREZ-SÁNCHEZ et al., 2010; MUSHTAQ et al., 2014; BECKONERT et al., 2007; HEYMAN et al., 2012).

Com o auxílio de soluções a base de álcool (metanol e/ou etanol) compostos semi-polares como ácidos fenólicos, flavonoides, alcaloides e esteróis glicosilados são eficientemente extraídos da matriz vegetal. Enquanto, carotenoides, ácidos graxos e terpenos são melhor extraídos em clorofórmio (MUSHTAQ et al., 2014; KIM; VERPOORTE, 2010). Contudo, o NuBBE tem optado pela substituição de solventes clorados por agentes menos tóxicos como acetato de etila (FUNARI et al., 2014).

A reprodutibilidade é determinante para o diagnóstico de alvos moleculares. De forma resumida, a metabolômica decorre através da comparação de dados sob diferentes condições. Assim, em sistemas não normalizados, existe a possibilidade de não observar qualquer efeito ou ainda pode levar a falsos-positivos, o que seria um desastre no desenvolvimento das justificativas subsequentes (ÁLVAREZ-SÁNCHEZ et al., 2010; MUSHTAQ et al., 2014).

Vale ressaltar que cada análise representa um simples momento do estado metabólico do organismo sem qualquer referência a variação biológica ou erros de medida. Para estimar tais variações, replicatas (no mínimo 6) biológicas e suficientes replicatas técnicas (no mínimo 3) de uma mesma espécie necessitam ser preparadas e analisadas.

Algumas considerações sobre análise em CG-EM e CLAE-EM

Além das tradicionais vantagens em eficiência de separação e reprodutibilidade (tempos de retenção) comparado a CLAE, a CG-EM possui outros aspectos positivos para o uso em estudos metabolômicos. A reprodutibilidade do perfil de fragmentação (EM/EM), decorrente da padronização do modo de ionização (ionização por elétrons) e da normalização das energias de ionização (70eV), possibilita a criação de bases de dados (a partir de padrões autênticos) para a comparação espectral intra e inter laboratoriais e, assim acelerando o processo desreplicação (FIEHN et al., 2000; LISEC et al., 2006; PETERSON et al., 2014).

Adicionalmente, reações de metoximação seguidas de sililação ampliaram a faixa de análise da CG-EM incluindo principalmente os metabólitos primários com funções hidroxil, aminas, amidas, fosfatos e tióis (LISEC et al., 2006). No entanto, alguns pontos negativos devem ser destacados no processo de derivação como: a formação de isômeros no caso de carboidratos e de degradação em certas classes de aminoácidos (FIEHN et al., 2000; LISEC et al., 2006).

No caso da CLAE-EM, a vantagem é que, em princípio, nenhuma etapa de derivação é necessária para a análise de compostos polares e semi-polares. Do ponto de vista instrumental, a CLAE-EM possui um enorme vantagem em relação ao número de analisadores disponíveis no mercado quando comparado a CG-EM.

De todo modo, a avaliação da performance de espectrômetros de massas é baseada nos seguintes critérios, (figura 6):

- **Poder de resolução (ou resolução);**
- **Exatidão de massa;**
- **Variação da dinâmica linear;**
- **Sensibilidade;**

Em geral, instrumentos que utilizam de analisadores do tipo quadrupolo (Q) possuem um poder de resolução de até 4 vezes menor que sistemas em tempo de voo (ToF) (BALOGH, 2004). Uma maior resolução em massa resulta numa maior distinção entre sinais de razão massa-carga relativamente próximos. O uso de soluções de *lock mass*, podem ajudar de forma complementar na correção de desvios no eixo de massa-carga (BALOGH, 2004; FIEHN, 2000).

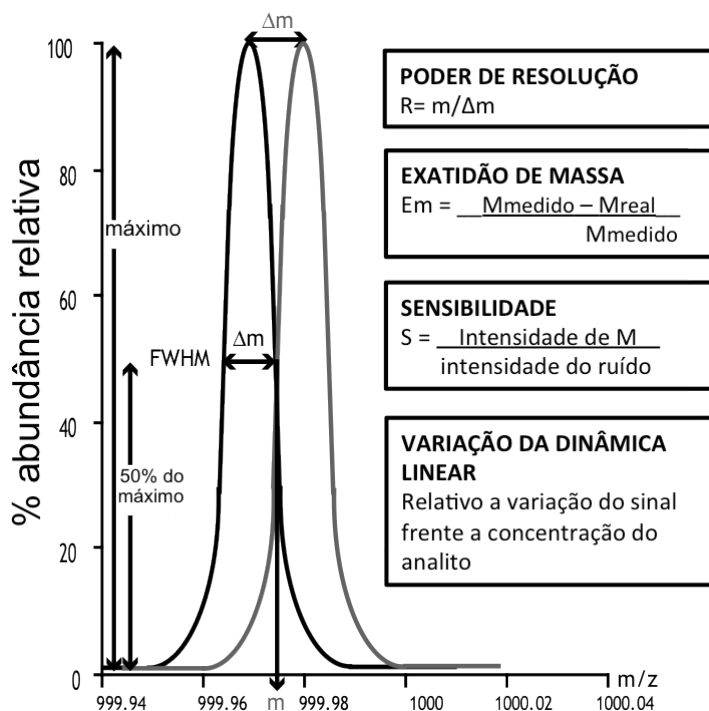
Instrumentos híbridos como QToF possuem alta sensibilidade e resolução, no entanto, em termos de variação da dinâmica linear eles são limitados por suas propriedades de tempo-digital que é somente capaz de gravar um íon por tempo morto. Assim, sinais intensos tornam-se saturados, mascarando sua real intensidade e levando a distorções na forma do pico e na exatidão dos valores de massa (VERHOEVEN et al., 2006).

Adicionalmente, aspectos relativos a composição da matriz como solventes e interferentes podem contribuir para o efeito de supressão de sinal ou efeito de matriz. Como resultado, metabólitos secundários minoritários são em grande parte não detectados. Esse fenômeno é frequentemente observado em sistemas que fazem uso de infusões diretas (PETERSON et al., 2014).

Já os instrumentos de FT-ICR-EM ou do tipo orbitrap permitem medidas em maior nível de exatidão e robustez na variação dinâmica (MAKAROV et al., 2006). Avanços na tecnologia de alguns instrumentos de QToF e sistemas em Orbitrap tem levado a maior velocidade de transmissão iônica permitindo a realização de experimentos do tipo AutoEM/EM (MAKAROV et al., 2006). Em modo automatizado o experimento AutoEM/EM realiza um experimento EM (EM^1) na qual seleciona os íons majoritários e, em seguida realiza experimentos de fragmentação do tipo EM^2 em tempo real. Isso permite que todos os picos da amostra sejam fragmentados numa única injeção. No caso dos sistemas Orbitraps, é possível analisar os íons em ambos os modos de ionização (positivo e negativo) num único experimento (PETERSON et al. 2014).

Figura 6. Parâmetros utilizados para descrever a performance de espectrômetros de massas. O poder resolução de massa (ou resolução), $m/\Delta m$, pode ser descrito duas maneiras: (1). para m sendo a média da razão massa-carga associada com dois sinais de massas adjacentes de igual tamanho e forma que se sobrepõem por 50% e Δm sendo a diferença em entre o sinal de massa-carga entre os dois sinais de massas adjacentes ou (2) m sendo o sinal de massa

no ápice e Δm sendo a largura a meia altura (50%) do sinal de massa, designado por FWHM (*full width at half height of maximum*). A exatidão de massa é descrito pela razão entre o erro de massa (diferença entre a massa real e a medida) e a massa teórica, frequentemente representado como partes por milhão (ppm). A sensibilidade é descrita pela razão entre nível do sinal de massa e o nível do sinal do ruído. A variação da dinâmica linear é descrito como o intervalo de linearidade do sinal do íon medido em função da concentração do analito.



Algumas considerações sobre a identificação de compostos via espectrometria de massas

Usualmente a identificação de metabólitos via espectrometria de massas é realizado segundo a combinação das seguintes informações:

- **Exatidão de massas;**
- **Distribuição isotópica;**
- **Padrões de Fragmentação;**

O cálculo de combinações químicas que ajustam a determinados valores da razão massa-carga é geralmente o primeiro passo para se obter um conjunto de alternativas que possam levar a identidade dos metabólitos detectados (KIND; FIEHN, 2010; NEUMANN; BÖCKER, 2010; WOLF et al., 2010).

Utilizando um equipamento com alta resolução, o número de possibilidades de formulas moleculares (FM) diminui drasticamente. De fato, o número de FM

aumenta em proporção ao valor de m/z e assim, uma pré-seleção de elementos químicos deve ser feita, evitando a geração de excessivas falsas alternativas (KIND; FIEHN, 2010).

Em geral, para a metabolômica de plantas, a maioria dos metais podem ser excluídos (exceto Na e K que são adutos comuns em EM), sendo que os elementos chave são o C, H, O e N e em segundo plano P e S (GIKKA et al., 2014; KIND; FIEHN, 2010)

Uma maneira eficiente de confirmação de fórmula molecular é fazer uso de padrões isotópicos a partir do íon molecular. Para todas as moléculas, a intensidade do segundo sinal revela relação da abundância isotópica da composição molecular do íon (KIND; FIEHN, 2010). Esta estratégia pode remover mais que 95% dos falsos positivos permitindo as subseqüentes etapas de elucidação com um número limitado de FM (KIND; FIEHN, 2007, 2010).

O padrão de fragmentação de um dado íon fornece informações estruturais daquele composto. O conhecimento das vias de fragmentação assim como o entendimento de quebras homolíticas ou heterolíticas tem permitido a elucidação de compostos via EM. Um exemplo típico são os flavonoides O-glicosilados. Ligações lábeis como por exemplo, entre a aglicona e o glicosídeo são facilmente rompidas através de perdas neutras. Assim, com baixas energia de colisão (como na fonte de ionização) é possível distinguir diferentes classes de açúcar como hexoses, deoxihexose, pentoses, etc.

Para maiores energias de colisão é notado quebras relativas ao esqueleto molecular como reações de retro-Diels-Alder (RDA), saídas de monóxido de carbono, quebras radicalares e perdas neutras de água. Com este arsenal de informações é possível retro-construir a molécula estudada permitindo, em alguns casos, determinar o padrão de substituição e a conectividade dos seus elementos (CLAEYS et al., 1996; CROTTI et al., 2006; CUYCKENS et al., 2000 e 2005; MARCH et al, 2006; PETERSON et al., 2014; WARIDEL et al., 2001).

MATERIAIS E MÉTODOS

Coleta e classificação do material vegetal

As amostras vegetais foram coletadas em diferentes localidades no estado de São Paulo, conforme tabela 2, e identificadas pela botânica Giselda Durigan. As exsiccatas foram depositadas no herbário do Instituto de Botânica da cidade de São Paulo. A figura 7 ilustra as espécies coletadas.

Tabela 2. Locais de coleta das espécies de *Solanum*. Os códigos das amostras fazem referência aos identificadores (ID) presentes no banco de extratos do NuBBE.

Código	Espécie	Parte	Local da coleta
M985	<i>S. paniculatum</i>	folhas	Estação Ecológica de Itirapina
M986	<i>S. paniculatum</i>	ramos	Estação Ecológica de Itirapina
R271	<i>S. swartzianum</i>	folhas	Estação Ecológica e Experimental de Assis
R272	<i>S. swartzianum</i>	ramos	Estação Ecológica e Experimental de Assis
F052	<i>S. swartzianum</i>	folhas	Parque Estadual da Serra do Mar / Núcleo Cunha
M951	<i>S. americanum</i>	folhas	Estação Ecológica de Itirapina
M952	<i>S. americanum</i>	ramos	Estação Ecológica de Itirapina
F55	<i>S. excelsum</i>	folhas	Parque Estadual da Serra do Mar / Núcleo Cunha

Figura 7. fotos ilustrativas das espécies estudadas: *S. paniculatum*; *S. swartzianum*; *S. americanum* e *S. excelsum*.



S. paniculatum



S. swartzianum



S. americanum



S. excelsum

Extração e secagem

O material vegetal obtido no horto da faculdade de ciências farmacêuticas de Araraquara – FCFAR foi previamente separado em ramos e folhas e submetido ao banho em N₂ líquido para inibição enzimática. Em seguida, o material foi seco e pulverizado em moinho de facas e extraído através de processo de fluido pressurizado - Dionex ASE 300 - Dionex Canada Ltd., Oakville, ON, CA. O protocolo de extração estabelecido entre Instituto de botânica (USP - São Paulo) e o NuBBE (Núcleo de Bioensaios, Biossíntese e Ecofisiologia de Produtos Naturais - UNESP)

utilizou de EtOH (92,8° INPM) a 60°C, 1500 psi sob a duração de 15 minutos. A eliminação dos solventes de extração foi realizada em evaporador rotativo Eppendorf Vacufuge (Eppendorf, Hauppauge, NY, EUA).

Solventes e reagentes empregados

Para a preparação das amostras e análise por cromatografia líquida de alta eficiência (CLAE) foram utilizados: ácido acético glacial (HOAc), metanol (MeOH), Clorofórmio (CHCl_3), piridina grau HPLC (Mallinckrodt, Paris, Kentucky, USA) e H_2O deionizada (18m Ω ; Milli-Q, Millipore). Os solventes foram degaseificados com auxílio de bomba de vácuo e banho de ultrassom. Para as reações de derivação, empregou-se cloridrato de metoxiamina, MSTFA (N-metil-N-[trimetilsilil]-trifluoracetamida) com 1% de TMCS (trimetilclorosilano) (Sigma Aldrich, St Louis, MO, USA).

Preparo de amostras para as análises cromatográficas

O preparo de amostras para CG-EM e CLAE-DAD-EM são ilustrados no fluxograma da figura 8. Para as análises de CG-EM, 4,0 mg de cada extrato foi dissolvido em MeOH/ CHCl_3 (0,4 mL - 1:1, v/v) e secos em Speedvac. Foram adicionados padrões internos de ácido mirístico- d^{27} (5 μL) e FAME (fatty acid methyl esters - 4 μL) (Fiehn GC/EM Metabolomics Standards Kit, KIND et al., 2009) com intuito de padronizar a pressão instrumental assim como estimar do índice de retenção linear.

A solução de FAME foi preparada através de uma mistura de homólogos de ésteres metílicos com diferentes tamanhos de cadeia lateral (C8, C9, C10, C12, C14, C16, C18, C20, C22, C24, C26, C28 e C30) dissolvidos em clorofórmio sob concentração de 0,8 mg.mL⁻¹ (C8-C16) e 0,4 mg.mL⁻¹ (C18-C30).

Para as reações de metoximação, 40 μL de cloridrato de metoxiamina (20 mg.mL⁻¹ dissolvido em piridina) foram adicionados ao material seco. Esta solução foi mantida sob 37°C durante 120 minutos para formação de oximas a partir de carbonilas cetônicas, figura 9. A etapa final consistiu na sililação de hidroxilas

(aminas, tóis e etc.) pela adição de 90 µL de MSTFA (N-metil-N-(trimetilsilil)-trifluoroacetamida) a 37°C durante 30 minutos, figura 10 (LISEC, et al., 2006).

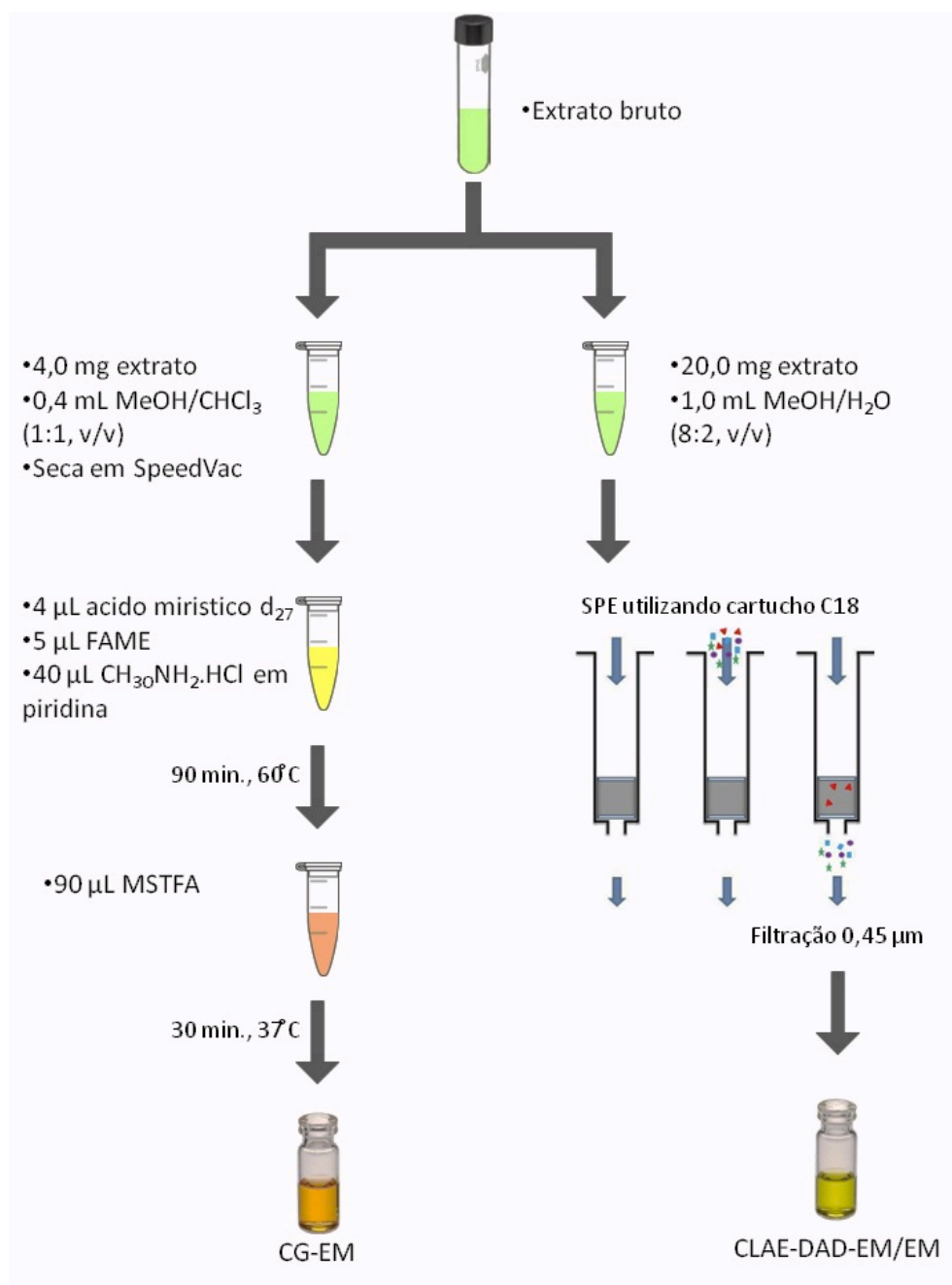
Para as análises por CLAE-DAD-EM, os extratos passaram por processo de remoção de compostos apolares através de extração em fase sólida (SPE) utilizando cartuchos de C18. Estes foram previamente ativados com MeOH (5,0 mL) e equilibrados em MeOH/H₂O (8:2, v/v). Cada amostra (20,0 mg) foi dissolvida em 1,0 mL da mistura MeOH/H₂O (8:2, v/v) e submetida a extração em fase sólida. Os compostos foram eluídos com a adição de 5,0 mL de MeOH/H₂O (8:2, v/v). As amostras foram secas em *speedvac* e, posteriormente redissolvidas em MeOH:H₂O (9:1 v/v) grau EM e filtradas em polipropileno de 0,45 µm de diâmetro de poro.

Configuração Instrumental: CG-EM

As amostras foram analisadas em cromatógrafo a gás Agilent 7890A GC-5975C acoplado a um espectrômetro de massas MSD com modo de ionização por elétrons (IE) e analisador do tipo quadrupolo (Q) (Agilent Technologies, Palo Alto, CA, EUA). Utilizou-se coluna capilar Agilent DB5ms (30 m x 0,25 mm x 0,25 mm + 10m pré-coluna Duraguard) (Agilent Technologies, Palo Alto, CA, EUA). Durante as análises usou-se hélio como gás de arraste. Os compostos foram detectados com o uso de *split* de saída do cromatógrafo na razão 1:10.

Foi injetado 1,0 µL das amostras derivadas no cromatógrafo a gás, utilizando método adaptado de Kind et al. (2009). As condições experimentais foram: modo de injeção *split* 1:10; temperatura no injetor 250°C; programa de aquecimento da coluna: 100°C isotérmico por 1 min, seguido por uma de temperatura 20°C.min⁻¹ até 200°C e então um acréscimo de 3°C.min⁻¹ até 325°C, totalizando um tempo de corrida de 47,7 min. A temperatura no detector foi mantida a 250°C utilizando hélio como gás de arraste a uma vazão de 0,7 mL.min⁻¹.

Figura 8. Fluxograma relativo a preparação das amostras para os experimentos em CG-EM e CLAE-DAD-EM. Os metabólitos das amostras analisadas por CG foram extraídos com auxílio de uma mistura de MeOH:CHCl₃ (1:1 v/v) e, uma vez secos, os extratos foram submetidos a reações de metoximação seguida de sililação para remoção de agentes de adesão a coluna de CG e melhorar a capacidade de volatilização a uma menor temperatura. Uma mistura de ésteres metílicos foi adicionado para as subseqüentes etapas de elucidação estrutural utilizando o software AMDIS e base de dados. Foram removidos interferentes apolares das amostras de CLAE-DAD-EM através da eluição das amostras sob cartuchos de C-18 na proporção de MeOH:H₂O (8:2 v/v) e filtrados em membrada de 0,45 µm.



Os parâmetros instrumentais utilizados no detector de EM foram: energia de ionização a 70 eV; detecção de 50-1000 m/z a 2 espectros.s⁻¹; tempo inicial de aquisição a 6,10 min (evitando a saturação do detector com subprodutos de derivação) e interrupção da detecção no intervalo 10,0-13,0 min (saturação causada

pela elevada concentração de monossacarídeos). Para a calibração e *tuning* utilizou-se de autotune FC43 (perfluorotributilamina), de acordo com as especificações do equipamento.

Figura 9. Mecanismo de reação de metoximas em sistemas carbonílicos.

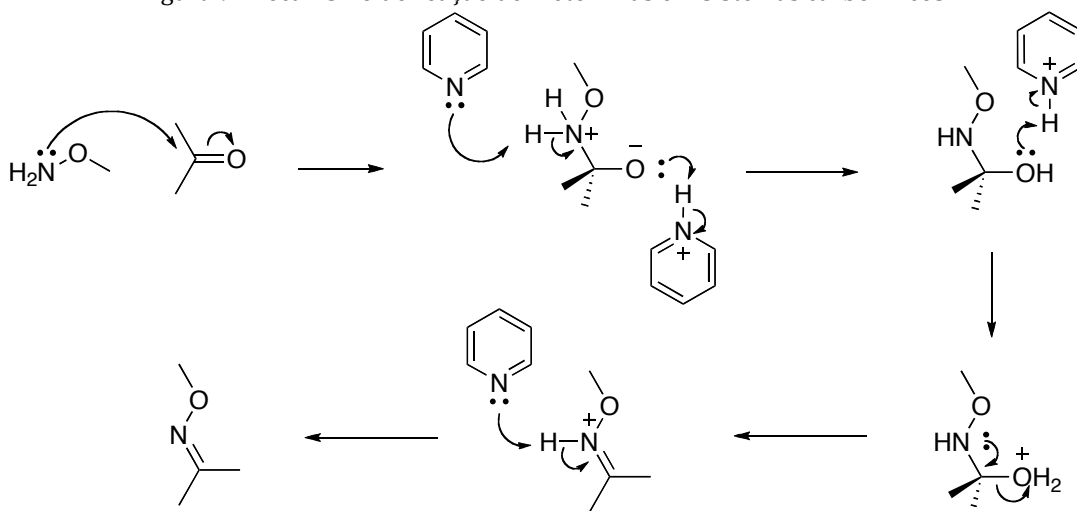
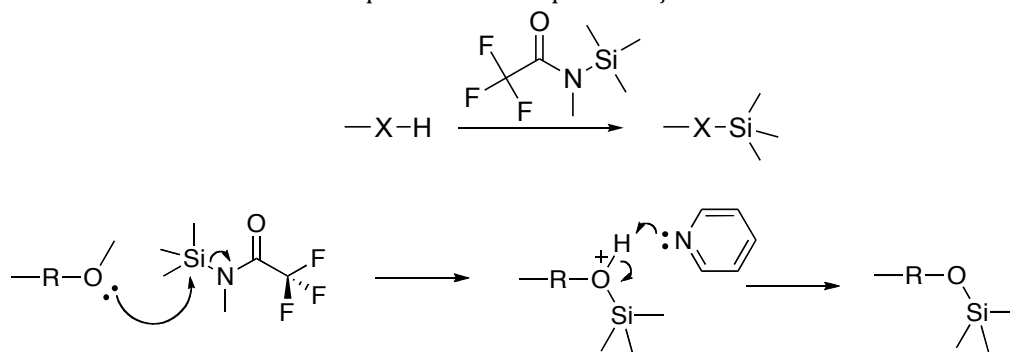


Figura 10. Mecanismo de reação para sililação de grupos hidroxil. Funções como tióis e aminas também se enquadram neste tipo de reação.



Configuração Instrumental: CLAE-EM e CLAE autoEM/EM

Dois equipamentos de CLAE-EM foram utilizados para a análise do perfil metabômico da espécie de *S. paniculatum* (M985 e M986). O primeiro consistiu em um cromatógrafo analítico ternário Shimadzu (Kyoto, Japan) modelo LC-20A acoplado a um detector UV-DAD Shimadzu CBM20A e um espectrômetro de massas UltrOTof (Bruker Daltonics, Billerica, MA, EUA) com modo de ionização por eletrospray (IES) e analisador do tipo QToF (analisadores do tipo quadrupolo e de tempo de voo, sequenciais). Para a condição cromatográfica foram utilizadas pré-coluna e coluna monolítica Onyx® C18, (5 X 4,6 mm e 100 X 4,6 mm, Phenomenex, Torrance, CA). As amostras analisadas passaram pelo processo de remoção de

interferentes apolares através de extração em fase sólida (SPE - *solid phase extraction*) utilizando cartuchos de C18 Strata de 55 μm , 5 g/20 mL, 70A (Phenomenex, Torrance, CA).

O método cromatográfico foi baseado no trabalho de Gómez-Romero, para estudo de desreplicação em *Solanum esculentum* (tomate) e ajustado segundo as configurações do instrumento e os seguintes parâmetros cromatográficos: fase estacionária, fase móvel, temperatura de coluna, vazão e volume de injeção (GÓMEZ-ROMERO et al., 2010). A conversão do método foi realizada segundo trabalho de GUILLARME et al. (2008) utilizando o software *HPLC Calculator v3.0*, disponível no site:

<http://www.unige.ch/sciences/pharm/fanal/lcap/index.htm>.

Alíquotas de 2,0 μL das amostras M985 e M986, sob concentração de 5,0 mg.mL^{-1} , foram injetadas no sistema cromatográfico utilizando como fase móvel ACN:HOAc (0,1%, v/v) e H₂O:HOAc (0,1%, v/v) na vazão de 0,300 mL.min^{-1} , em método de gradiente exploratório conforme descrito na tabela 3.

O eluente da coluna foi dividido fazendo uso de um *splitter* na proporção de 5:1, sendo o maior fluxo direcionado ao DAD e o restante ao espectrômetro de massas (vazão de 150 $\mu\text{L.min}^{-1}$). A detecção no UV-DAD foi realizada entre 200 a 800 nm. As condições instrumentais utilizadas no detector de EM foram: sonda capilar de 3kV e cone de 25V; gás N₂ para nebulização em fluxo de 345 L.h^{-1} e 27 L.h^{-1} , respectivamente, e temperatura de 180 °C. A cela de colisão foi preenchida com N₂ a pressão de 7 psi e a energia de colisão entre 15 a 40 eV para os experimentos de EM/EM.

Tabela 3. Método cromatográfico aplicado a análise de CLAE-DAD-EM das amostras M985 e M986 (modo gradiente de eluição).

Tempo (minutos)	% de B (ACN)
0,00	1
2,33	3
12,00	12
35,00	50
48,00	100
60,00	100
62,00	1
72,00	1

Para a calibração dos dados de massas de alta resolução, foi injetado uma solução de Na-TFA (10 mg.mL⁻¹) ao final de cada análise cromatográfica.

Já as amostras de *S. americanum*, *S. swartzianum*, *S. excelsum* e *S. paniculatum* foram analisadas num instrumento CLAE-EM/EM - Agilent 1200 SL acoplado com um espectrômetro de massas Agilent 6520 QToF (Agilent Technologies, Santa Clara, CA). As análises cromatográficas foram realizadas em coluna Zorbax Eclipse XDB C18 (100 X 3,0 mm, 1,8 µm) acoplada a coluna de guarda (4,6 × 5,0 mm) de material equivalente.

Foram injetados 10,0 µL no sistema CLAE-EM/EM segundo as conversões propostas pelo software *HPLC calculator v3.0* para o método desenvolvido de Gómez-Romero et al. (2010) (GUILLARME et al., 2008). A condição cromatográfica foi: fase móvel H₂O/HOAc (0,5%, v/v, solvente A) e ACN:HOAc (0,5%, v/v, solvente B) sob vazão de 0,7 mL.min⁻¹ e temperatura de coluna de 37 °C. O modo eluição é descrito na tabela 4.

Tabela 4. Método cromatográfico aplicado as amostras de *Solanum*: M985, M986, R271, R272, F052, M951, M952 e F55.

Tempo (minutos)	% de B (ACN)
0,00	3
1,00	3
14,00	9
32,00	35
39,00	100
45,00	100

Os espectros de autoEM/EM foram adquiridos no mesmo sistema *Agilent* em ESI (+)-MS e janela espectral de 100-1000 *m/z*. Os íons precursores foram selecionados a partir de analisador quadrupolo utilizando Δ 4,0 *m/z*. O sistema de fragmentação CID EM/EM utilizou o N₂ como gás de colisão e energia de colisão de 15, 30 e 45 eV. Os parâmetros do ESI-Q-ToF-MS foram: temperatura do gás 250 °C, gás de arraste 6.0 L.min⁻¹, pressão do nebulizador 35 psi, voltagem do capilar em 3500 V e taxa de aquisição de espectros 8 Hz. A calibração foi realizada com padrões internos purina 2 µM e HP-921.

O modo de aquisição autoEM/EM utilizou ciclos EM-EM/EM de 4,0s, com 2,0s para aquisição no modo EM e 1,0s para cada aquisição EM/EM para os 2 íons mais

intensos (acima de > 1000 *counts*), excluindo os íons interferentes determinados previamente no branco do método.

Construção de uma base de dados de *Solanum*

Os estudos de desreplicação se fundamentam na comparação de informações experimentais com bases de dados, incluindo informações espectroscópicas, espectrométricas, propriedades físico-químicas, desenho estrutural entre outras propriedades. Deste modo, foi necessário o desenvolvimento de um banco de dados de moléculas já reportadas para o gênero *Solanum*

Através de um levantamento bibliográfico nas principais fontes de busca, como *ScienceDirect*, *Web of Science*, *Scifinder* e o *Dictionary of Natural Products* (DNP), foram analisados 225 trabalhos científicos envolvendo artigos, teses e revisões em busca de substâncias presentes neste gênero. Ao todo 175 espécies de *Solanum* foram analisados resultando em 795 moléculas catalogadas. Dentre elas, 69 flavonoides, 205 alcaloides, 246 terpenoides e o restante relativo a metabólitos envolvidos no metabolismo primário ou em classes gerais de compostos como as poliaminas. A base de dados foi construída no formato *.xml* (excel) contendo informações de nomes IUPAC, nome trivial, CAS-Number, CID (Pubchem ID), SMILES - formato relacionado ao desenho estrutural, InChI e InChkey - formato relacionado ao desenho estrutural, massa monoisotópica, massa monoisotópica protonada e desprotonada e a fórmula molecular.

Softwares

Os dados foram analisados pelos softwares: *MassHunter Workstation software* (versão B.06.00), *Automated Mass Spectral Deconvolution and Identification System* - AMDIS, (versão 2.64, NIST, EUA) e *DataAnalysis* (versão 4.2, Bruker Daltonics, Billerica, CA) Os métodos quimiométricos foram desenvolvidos e aplicados no Software MATLAB ver.7.10 (R2010a) (*The MathWorks Inc.*, Natick, EUA).

Desreplicação por CG-EM

A identificação metabólica por CG-EM foi efetuada através da comparação ortogonal de valores de índice de retenção (IR) e perfil de fragmentação EM/EM.

Os índices de retenção foram calculados a partir de equação de regressão linear utilizando dos tempos de retenção de uma mistura de ésteres metílicos (solução FAME - padrão interno) com os respectivos valores tabelados de IR desses compostos (literatura).

A solução FAME é composta por ésteres metílicos de ácidos graxos (C8, C9, C10, C12, C14, C16, C18, C20, C22, C24, C26, C28 e C30), dissolvidos em clorofórmio na concentração de 0,8 mg.mL⁻¹ (C8-C16) e 0.4 mg.mL⁻¹ (C18-C30). (KIND et al., 2009).

Os perfis de fragmentação, obtidos após deconvolução cromatográfica realizado pelo software AMDIS, foram comparados com as bases de dados *Golm Metabolome Database* (GMD_20111121_VAR5_FAME, *Max Planck Institute, Potsdam–Golm*, Germany), *FiehnLib* (University of California, Davis, CA, EUA) e *Noble database* (Samuel Roberts Noble foundation, Ardmore, OK, EUA) afim de avaliar o nível de similaridade (*match factor* MF) e, conseqüentemente, apontar possíveis *Hits*. Somente os compostos com valores maiores que 80 no fator de similaridade e menores que 5% de erro no valor do índice de retenção foram considerados neste estudo.

Desreplicação por CLAE-EM

Os picos cromatográficos obtidos nas análises por CLAE-DAD-EM/EM foram identificados através de avaliação comparativa com as moléculas descritas na literatura para a família Solanaceae (Bases de dados online e “*home made*”) e por propostas mecanística de fragmentação. Para tanto, foram avaliados os espectros de UV, os valores de razão massa-carga de alta resolução (molécula protonada

[M+H]⁺) e espectros de íons produto gerados em experimentos EM/EM a partir de dissociação induzida por colisão. Considerou-se a detecção de substâncias cujo erro entre a massa observada e calculada (teórica) não ultrapassasse 5 ppm (FUNARI et al., 2012).

Desreplicação CLAE-autoEM/EM utilizando rede de interações moleculares (*Molecular networking*)

Baseado no trabalho de Watrous et al. (2012), dados obtidos CLAE-autoEM/EM foram convertidos a extensão *.mgf*, formato baseado em texto usado para representar dados de espectrometria de massas descrevendo o número de *scans*, íon precursor, *m/z* e intensidade de cada íon observado. Em seguida, os arquivos foram submetidos ao software *Spectral Networks* situado na plataforma online GNPS (WATROUS et al., 2012, YANG et al., 2013) - *Global Natural Products Molecular Networking* – homepage:

<https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp?redirect=auth>)).

Computacionalmente, os espectros foram convertidos em vetores de espaço *n*-dimensional e, em pares, estes foram comparados através de cálculo do produto vetorial (Dot Product). Por definição, o *Dot Product* permite calcular o nível de similaridade entre dois conjuntos de dados a partir do cosseno do ângulo formado entre estes vetores. A pontuação atribuída para cada par de vetores varia entre 0 a 1, onde 1 representa espectros idênticos - cosseno de 0° (WATROUS et al., 2012; YANG et al., 2013).

Neste experimento foram gravados espectros AutoEM/EM em ciclos de 4,0 s. De fato, grande parte dos espectros adquiridos fazem referencia a uma mesma molécula. Isto ocorre pois os picos cromatográficos levam em média 30 segundos para a eluição total.

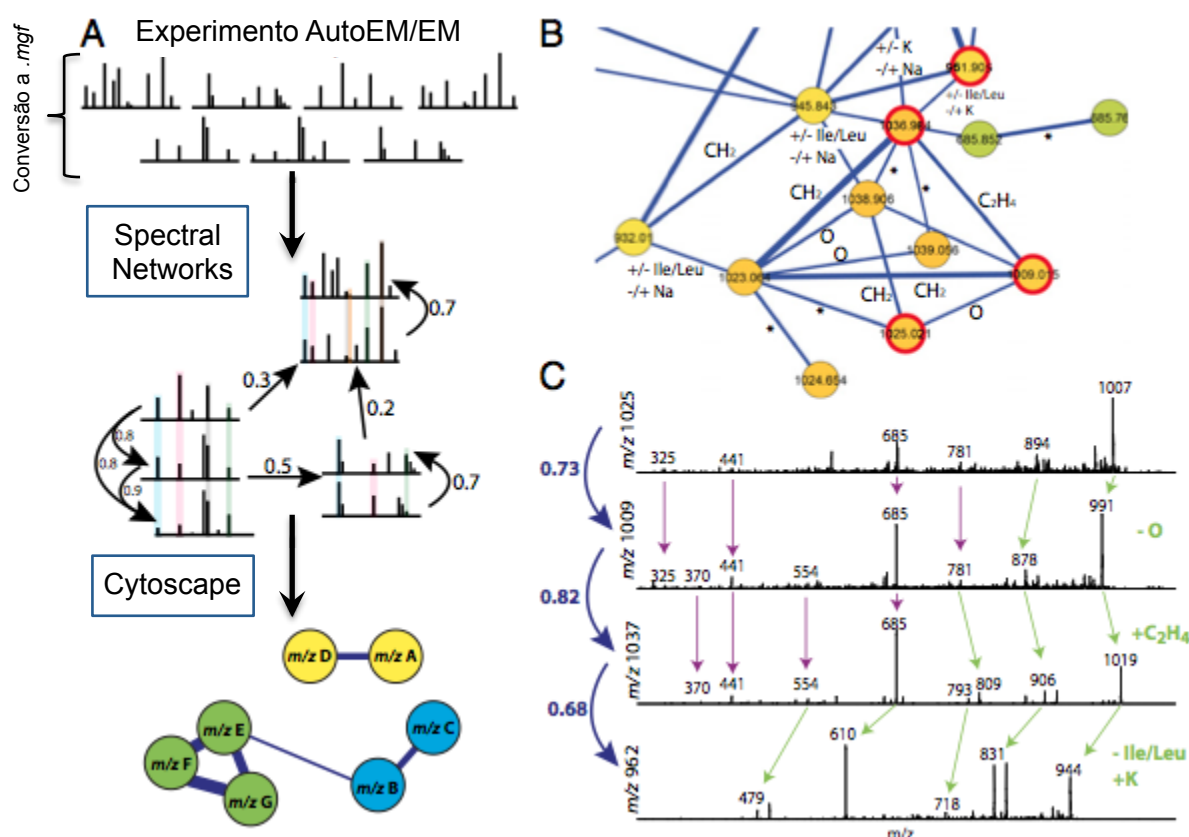
Assim, para reduzir a complexidade dos agrupamentos resultantes, espectros idênticos de EM/EM foram combinados através do algoritmo *MS-Clustering* levando em consideração os valores de massas dos íons precursores, dentro de um desvio de 0,5 Da. Em seguida, o *Spectral Networks* comparou todos os possíveis pares de vetores a partir dos espectros combinados. Os parâmetros do algoritmo incluem tolerância de massa para íons fragmentos de 0,3 Da, e desvio de 1,0 Da para íon

precursor. Para o cálculo de similaridade entre um par de vetores foram selecionados espectros com no mínimo 6 pares de íons (picos) em correspondência (WATROUS et al., 2012; YANG et al., 2013).

Em seguida, os resultados foram exportados em formato texto e importados dentro do software grátis Cytoscape (<http://www.cytoscape.org>) e transformados nos mapas de rede molecular (nodo e arestas - conexões), figura 11.

Os nodos foram organizados sob a forma de Force Direct Display. As cores dos nodos foram atribuídas a para a diferenciação das espécies; a largura das arestas assim como as cores das arestas (conexões) foram associadas aos valores de similaridade entre os nodos. A figura 11 ilustra o processo de construção dos mapas de redes metabólicas desde o processo de aquisição até a transformação no espaço do cytoscape.

Figura 11. Construção de rede molecular via comparação espectral, figura adaptada do trabalho de Watrous J. *et al.*, 2012. (A) representação esquemática de como o networking molecular é gerado. Os valores representam a similaridade de cossenos que varia de 0 a 1, onde 1 representa espectros idênticos. A espessura da arestas (linhas azuis) indicam o nível de similaridade. (B) Visualização pelo cytoscape de como a rede é formada. (C) Um exemplo do cálculo da pontuação de similaridade usando quatros espectros diferentes.



RESULTADOS E DISCUSSÃO

Três estratégias foram aplicadas para a geração de perfis metabolômicos e estudos de desreplicação dos metabólitos primários e secundários das espécies de *Solanum*: (1) CG-EM associado ao software de deconvolução e comparação espectral (AMDIS); (2) CLAE-DAD-EM-EM/EM em comparação com base de dados na busca por metabólitos secundários e (3) desreplicação via rede interação molecular fazendo uso de experimentos CLAE-AutoEM/EM.

Perfil metabolômico e desreplicação de metabólitos primários através de CG-EM e AMDIS

A estratégia de desreplicação por CG-EM e AMDIS combinou informações ortogonais da comparação entre perfis espectrais com o índice de retenção dos compostos.

Inicialmente o material vegetal seco e pulverizado passou por extração via fluido pressurizado (EFP) com auxílio de etanol. A alta pressão em conjunto com temperaturas elevadas favorece a uma maior solubilidade, razão de difusão e transferência de massa dos compostos. A nível molecular, a EFP favorece o rompimento de interações soluto-matriz que envolvem as ligações iônicas, ligações de hidrogênio, forças de Van der Waals e interações de dipolo (HASSAS-ROUDSARI, et al., 2009; JACQUES, 2005).

Em geral, os equipamentos de EFP são de fácil manuseio operacional e cumprem diversas premissas exigidas nos estudos metabolômicos. A EFP reduz o tempo médio gasto no processo de extração, sendo que o preparo de amostra é rápido e simples e a extensão de extração (diversidade metabólica) é alta. Utiliza-se pouca quantidade de solvente (eficiência) e todo o processo é automatizado resultando em alta reprodutibilidade (HASSAS-ROUDSARI et al., 2009; JACQUES, 2005).

Uma vez obtido os extratos, 4,0 mg desse material foi ressuspenso em 0,4 mL de uma solução MeOH:CHCl₃(1:1, v/v) visando a extração principalmente dos compostos de média polaridade. O solvente foi removido com auxílio de um

speedvac e iniciou-se o procedimento de metoximação e sililação dos grupamentos cetônicos e substituintes hidroxilados, aminas e tióis.

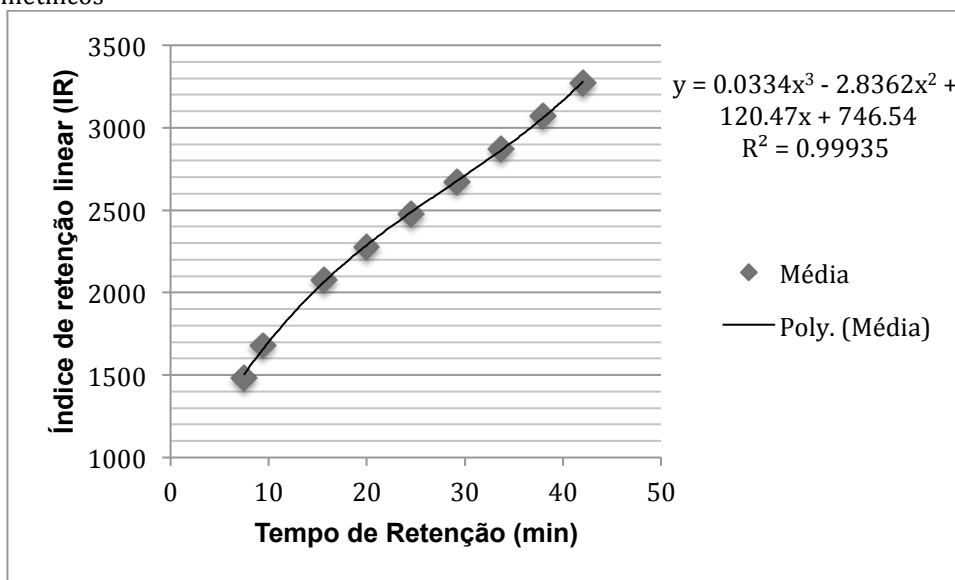
A etapa inicial consistiu na determinação do índice de retenção (IR) do perfil metabólico dos extratos observados no software AMDIS, tabela 5. O cálculo do IR foi realizado através da interpolação de dados (compostos) numa equação de regressão (figura 12). Esta equação foi gerada a partir do tempo de retenção (tR) dos padrões da solução FAME com os respectivos valores de IR catalogados na literatura, figura 12

Os padrões de ésteres C8, C9, C10 não foram detectados devido aos tRs serem inferiores ao início da aquisição de dados pelo EM (6,10 min). O padrão C16 também não está presente nas amostras devido a interrupção da detecção pelo EM no intervalo de 10,0-13,0 min devido a saturação do detector.

Tabela 5. Valores dos tempo de retenção para a mistura de ésteres metílicos (FAME) para as amostras: R270, R271, F052, M951, M952, F55 e o controle (Ctrl. - branco cromatográfico). Adicionalmente são apresentados os valores médios do tempo de retenção (azul) para cada éster assim como o desvio padrão (DP - vermelho). O índice de retenção linear (IRL) dos respectivos ésteres também é fornecido.

Éster	IRL	Ctrl.	R270	R271	F052	M951	M952	F55	Média	DP
C12	1481	7,49	7,49	7,49	7,49	7,50	7,49	7,49	7,49	0,01
C14	1680	9,37	9,37	9,37	9,37	9,37	9,38	9,37	9,37	0,00
C18	2077	15,65	15,65	15,65	15,65	15,66	15,66	15,65	15,65	0,01
C20	2276	19,94	19,94	19,94	19,94	19,93	19,93	19,94	19,94	0,01
C22	2475	24,53	24,53	24,54	24,54	24,54	24,54	24,54	24,54	0,00
C24	2674	29,18	29,18	29,17	29,17	29,16	29,18	29,16	29,17	0,01
C26	2872	33,67	33,67	33,66	33,66	33,65	33,66	33,65	33,66	0,01
C28	3071	37,95	37,95	37,94	37,94	37,95	37,94	37,92	37,94	0,01
C30	3270	42,01	42,01	42,02	42,02	41,99	42,00	41,99	42,00	0,01

Figura 12. Gráfico do cálculo da equação de regressão (polinômio cúbico) para os valores médios dos ésteres metílicos



A segunda etapa consistiu na avaliação dos dados espectrais gerados pelos experimentos CG-EM/EM. Devido a complexidade estrutural, os picos sobrepostos no sistema cromatográfico foram inicialmente deconvoluídos a partir do uso do software AMDIS e posteriormente, comparados com as bases de dados espectrais (GOLM Database; Fiehn Database e Noble Foundation Database).

Todo o processo computacional foi realizado no software AMDIS e baseou-se na otimização de cinco parâmetros: (1) janela de componentes, na qual refere-se ao número de *scans* utilizados para a formação de um pico cromatográfico, (2) subtração de pico adjacente, que relaciona o número de máximo de compostos em coeluição para cada janela espectral, (3) resolução, associada a detecção de sistemas em coeluição, (4) sensibilidade, valor mínimo de detecção (S/R) e (5) os requerimentos de forma, na qual correlaciona os *scans* com formato de um pico cromatográfico. Os níveis de cada fator mostrados acima são mostrados na tabela 6.

Supõe-se que condição ótima de configuração do sistema AMDIS levaria um maior número de propostas moleculares na comparação com as bases de dados (padrões autênticos). No entanto, uma avaliação puramente numérica dos resultados poderia levar a falsos-positivos. Isto porquê no ajuste dos parâmetros é possível diminuir o nível de sensibilidade ao nível do ruído, ou ainda superajustar os requerimentos de forma e resolução resultando na combinação de dados de dois ou mais picos na forma de um único sinal.

Tabela 6. Parâmetros de otimização do processo de deconvolução e comparação espectral no software AMDIS.

Variáveis AMDIS	NÍVEIS		
	Mínimo	Intermediário	Máximo
Janela de componentes	8	16	32
Subtração de pico adjacente	0	1	2
Resolução	<i>low</i>	<i>medium</i>	<i>high</i>
Sensibilidade	<i>very low/low</i>	<i>medium</i>	<i>high/very high</i>
Requerimentos de forma	<i>low</i>	<i>medium</i>	<i>high</i>

Deste modo foi desenvolvido, através de uma equação heurística, um critério de resposta que relaciona o número de compostos alvos (identificado nas bases) em função do número de compostos detectados (detectado no software AMDIS mas não presente nas bases de dados). Este critério, Fator de Detecção de Compostos (FDC) maximiza a identificação dos compostos nas base de dados com o mínimo de requerimentos. A equação 1 representa o que foi escrito acima,

$$FDC = A \times \left(\frac{A}{B}\right)^x \quad (1)$$

onde A é o número de compostos identificados e B é o número de picos detectados. Nessa equação, o fator determinante está associado ao termo A, enquanto a razão $(A/B)^x$ exprime um termo de restrição. O valor de x depende de quão importante é o fator de restrição para um dado estudo (tipicamente $x = 2$).

Uma vez que a limitante do processo de identificação são as bases de dados (fator relacionada ao termo A), a variação dos parâmetros instrumentais que levem somente o aumento do nível de detecção (relativo ao termo B) resulta num menor valor de FDC. Assim, a otimização foca no aumento concomitante dos termos A e B.

Os parâmetros de deconvolução e detecção foram otimizados utilizando o *matching fator* maior que 90 (*scores*) e analisados através do FDC. Os resultados do FDC entre os diferentes níveis dos fatores para cada amostra está representado na figura 13.

A configuração capaz de gerar o maior número de propostas utilizando o fator FDC rendeu os seguintes valores: janela de componentes = 8; subtração de pico

adjacente = 2; resolução = *high*; sensibilidade = *low*; e requerimentos de forma = *low*. A partir dos resultados obtidos da FDC é possível apontar que não houve qualquer tipo de *overfitting* do processo de deconvolução assim como de resolução. Na maioria, foram utilizados os valores mínimos ou intermediários dos parâmetros de deconvolução e identificação envolvidos no software AMDIS.

A identificação metabólica se deu através da análise de similaridade entre os espectros de massas experimentais e aqueles encontrados nas bases de dados através do software AMDIS. Considerou-se apenas os compostos que possuem erros menores que 4% para índice de retenção e similaridade igual ou acima de 80 unidades similaridade no AMDIS *matching factor* (FIEHN, 2000).

A Tabela 7 apresenta as propostas moleculares encontradas a partir do perfil metabolômico de espécies de *Solanum* analisadas via CG-EM. Nesta tabela são encontrados os índices de retenção calculados a partir da equação de regressão (IR_{cal}) e comparados com padrões autênticos (IR_{lit} – Golm database), erros na predição dos valores retenção (%), *matching fator scores* (MF), relativos ao nível de similaridade entre o perfil espectral do compostos encontrado nas bases e os valores de massa moleculares sem a presença dos silanaóis.

Figura 13. Gráficos de resposta em função dos valores de FDC sob matching fator > 90 para as amostras F55, M951, M952, F052, R270 e R271. (a) Janela de componentes; (b) subtração de picos adjacentes; (c) fator de resolução; (d) fator de sensibilidade e (e) requerimentos de forma.

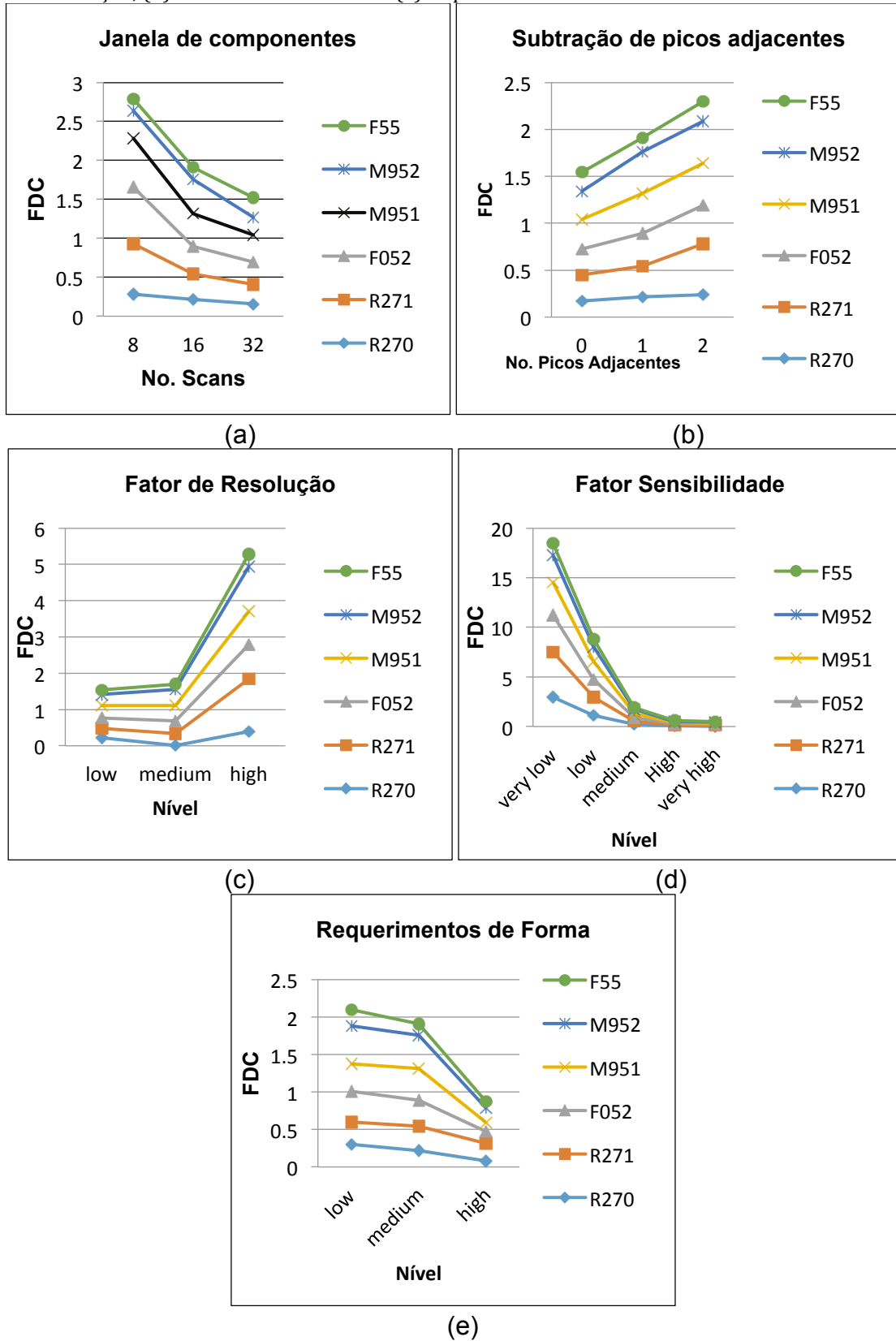


Tabela 7. Metabólitos identificados através da análise dos dados CG-EM no software de deconvolução AMDIS. A critério de formação, as amostras foram codificadas como segue: **1** – ramos de *S. swartzianum* (R272); **2** – folhas de *S. swartzianum* (F052); **3** – folhas de *S. swartzianum* (R271); **4** – folhas de *S. americanum* (M951) e **5** – ramos de *S. americanum* (M952); **6** – folhas de *S. excelsum* (F55). São fornecidos os valores do tempo de retenção encontrados no AMDIS (tR), Índice de retenção linear teórico (IR_{lit}), índice de retenção linear calculado a partir da equação de regressão figura 12 (IR_{calc}), porcentagem de erro do IR (IR_{calc} – IR_{lit}), valores médios dos matching factor de cada composto (MF) e a massa molecular (sem a presença dos silanóis).

Aminoácidos								
No.	Nome	Amostras	tR (min)	IR _{calc}	IR _{lit}	Erro (%)	MF (%)	Massa Molecular
1	serina (3 TMS)	2,3,4,5,6	6,22	1355	1354	0,1	93,6	105,0426
2	L-treonina (3 TMS)	3,4	6,39	1374	1377	-0,2	96,5	119,0582
3	beta-Alanina (3 TMS)	1,2,3, 4,5,6	6,74	1413	1424	-0,8	90,0	89,0477
4	N-metil trans-4-hidroxi-L-prolina (2S, 4R)-4- ácido hidroxi-1-metil pirrolidina-2-carboxílico (2 TMS)	1,2,3,4,5,6	6,78	1417	1435	-1,3	90,5	272,1372
5	Asparagina [-H ₂ O] (2 TMS)	3	7,31	1472	1499	-1,8	84,0	132,0535
6	L-ácido aspártico (3 TMS)	2	7,36	1478	1511	-2,2	96,0	133,0375
7	5-metil citosina (2 TMS)	6	7,57	1499	*	-	80,0	135,0354
8	L-ácido glutâmico 2 (2 TMS)	2,4,5	8,2	1560	1528	2,1	84,5	147,0532
9	L-phenilalanina (2 TMS)	1,4	8,38	1577	1629	-3,2	85,5	165,0790
10	L-asparagina (3 TMS)	4	8,7	1606	1666	-3,6	91,0	132,0535
11	ácido piroglutâmico (2 TMS)	1,2,3,4,5,6	7,47	1489	1522	-2,2	94,7	129,0426
Ácidos Graxos								
No.	Compostos	Amostras	tR (min)	IR _{calc}	IR _{lit}	Erro (%)	MF (%)	Massa
12	ácido hexadecanóico (1 TMS)	1,2,3,4,5,6	14,06	1994	2045	-2,5	99,3	256,2402
13	ácido linoleico (1 TMS)	2,3,4,5,6	17,35	2166	2219	-2,4	94,0	280,2402
14	ácido oleico (1 TMS)	1,2,3,5,6	17,72	2184	2225	-1,9	91,3	282,2559

15	ácido esteárico (1 TMS)	2,3,4,5,6	18,04	2198	2243	-2,0	99,0	284,2715
16	ácido n-eicosanóico (1 TMS)	2	22,46	2388	2453	-2,6	86,0	312,3028
17	ácido tricosanóico (1 TMS)	2	29,3	2677	2731	-2,0	93,0	354,3498
<i>Ácidos Orgânicos</i>								
No.	Compostos	Amostras	tR (min)	IR _{calc}	IR _{lit}	Erro (%)	MF (%)	Massa
18	ácido pipecólico (2 TMS)	1,2,3	6,36	1371	1365	0,4	91,3	129,0789
19	ácido málico (3 TMS)	1,2,3,4,5,6	7,1	1451	1479	-1,9	97,3	134,0215
20	ácido salicílico (1 TMS)	2	7,38	1480	*	-	84,0	180,0423
21	ácido 4-aminobutírico (3 TMS)	1,2,3,5,6	7,51	1493	1527	-2,2	85,8	103,0633
22	ácido 4-guanidinobutírico 2 (3 TMS)	1,2,4,5,6	7,52	1494	*	-	92,0	174,9147
23	ácido treônico (4TMS)	2,3,6	7,52	1494	1545	-3,3	88,3	136,0372
24	ácido trihidroxibutírico (4 TMS)	1,2,3,4,5,6	7,53	1495	*	-	91,0	136,0372
25	ácido 3-fenilático (3 TMS)	3	7,98	1539	1585	-2,9	92,0	166,0630
26	ácido p-hidroxibenzóico (2 TMS)	1,2,3,4,5,6	8,37	1576	1633	-3,5	95,0	138,0317
27	Ácido 4-hidroxifenilacético (2 TMS)	1,2	8,48	1586	1644	-3,5	86,7	152,0474
28	ácido 4-hidroxi-3-metoxibenzóico (2 TMS)	3,4	9,79	1700	*	-	95,7	168,0423
29	ácido vanilínico (2 TMS)	1,2,3,5,6	9,79	1700	1766	-3,7	95,3	168,0423
<i>Poliols</i>								
No.	Compostos	Amostras	tR (min)	IR _{calc}	IR _{lit}	Erro (%)	MF (%)	Massa

30	D-treitol (4 TMS)	1,2,3,4,5,6	7,16	1457	1485	-1,9	97,0	122,0591
31	eritritol (4 TMS)	1,2,3,4,5,6	7,23	1464	1493	-1,9	99,7	122,0579
32	ononitol (5 TMS)	4	13,17	1941	1945	-0,2	91,0	194,0790
33	mio-inositol (6 TMS)	1,2,3,4,5,6	14,83	2038	2080	-2,0	97,8	180,0634
<i>Açúcares</i>								
No.	Compostos	Amostras	tR (min)	IR _{calc}	IR _{lit}	Erro (%)	MF (%)	Massa
34	xilitol (5 TMS)	1,2,3,4,5,6	9,05	1637	1694	-3,4	89,4	152,0685
35	Arabitol (5 TMS)	1,2,3,4,5,6	9,21	1651	1707	-3,3	99,8	152,0685
36	ribitol (5 TMS)	1,2,3,4,5,6	9,25	1655	1712	-3,4	96,8	152,0685
37	D-(+)-manose (5 TMS)	4,5	15,22	2059	*	-	86,0	180,0634
38	D-alose 1(5 TMS)	1	15,23	2059	*	-	81,0	180,0634
39	Trehalose (6 TMS)	1,2,3,4,5,6	29,27	2676	2726	-1,8	98,2	342,1162
40	Rafinose (1 TMS)	5	42,91	3327	3350	-0,7	98,0	504,1690
41	ácido-ribônico gama-lactona 1 (4 TMS)	6	8,84	1619	1681	-3,7	89,0	148,1140
42	D-lixose 2 (6 TMS)	2,4,5	8,56	1593	1650	-3,4	90,7	150,0528
43	Ribose (4 TMS) MeOX	1,2	8,6	1597	1665	-4,0	93,5	150,0528
<i>Miscelânea</i>								
No.	Compostos	Amostras	tR (min)	IR _{calc}	IR _{lit}	Erro (%)	MF (%)	Massa
44	vanilina MeOX (1 TMS)	2,5	8,58	1595	*	-	85,0	152,0474
45	L-putrescina (4 TMS)	4,5	9,53	1678	1736	-3,3	91,0	88,1001
46	escopoletina	3,6	14,11	1997	2067	-3,4	93,0	192,0423
47	ácido Iminodiacético (3 TMS)	2	7,36	1478	1521	-2,9	93,0	133,0375
48	Fitol (1 TMS)	1,2,3	16,52	2126	2132	-0,3	95,5	295,3001
49	hidroquinona (3 TMS)	6	6,57	1394	1401	-0,5	88,0	110,0368
50	alfa-tocoferol (1 TMS)	1,2,6	37,84	3085	3161	-2,4	93,0	430,3811

51	glicerol 1-fosfato (1 TMS)	1,2,3,4,5,6	9,66	1689	*	-	93,2	172,0137
52	estigmasterol	2	40,59	3220	3319	-3,0	89,5	412,3705
53	beta-sitosterol (1 TMS)	2,6	41,81	3277	3385	-3,2	90,0	414,3862
54	ácido clorogênico	3	37,44	3065	3099	-1,1	96,0	354,0951
55	p-hidroxibenzaldeído MeOX (1 TMS)	1,2	7,29	1470	*	-	84,0	122,0368
56	ácido O-acetilsalicílico (1TMS)	2	7,37	1479	*	-	92,0	180,0423

* ausência de valores do índice de retenção na GOLM database

A utilização do pacote AMDIS junto ao cálculo de índice de retenção permitiu a determinação de 56 compostos, compreendidos em: aminoácidos, ácidos graxos, ácidos orgânicos, polióis, açúcares entre outros. Tal resultado só foi possível pela combinação de informações ortogonais entre a cromatografia e a espectrometria de massas, unindo as propriedades de separação, capacidade de pico e reprodutibilidade exclusivas a cromatografia a gás com a sensibilidade e caracterização estrutural obtida pelos padrões de fragmentação dos compostos.

Dois fatores relevantes incorporados nesta proposta: desenvolvimento de uma equação heurística na qual evitou o superajuste do processo de deconvolução e, identificação e a combinação de dados de diferentes bases de dados ampliando o número e a robustez da identificação para os compostos apresentados.

Perfil metabolômico e desreplicação de metabólitos secundários através de CLAE-EM-EM/EM

Nesta etapa foram realizados estudos de desreplicação a partir do perfil metabolômico em CLAE-EM-EM/EM visando a detecção dos compostos de média polaridade presentes na espécie *Solanum paniculatum* (M951 e M952), figura 14

O processo inicial de extração do material vegetal seguiu o esquema descrito na figura 8. A partir dos extratos brutos via EFP, foram removidos os interferentes para o sistema cromatográfico através eluição em cartuchos de fase sólida em C-18 na proporção 8:2 MeOH:H₂O. Esta proporção foi escolhida uma vez que a maior parte dos compostos de baixa polaridade estariam incluídos no sistema CG-EM. Desta forma, reduziu-se a complexidade química dos extratos visando exclusivamente aqueles compostos de média polaridade associados ao metabolismo secundário.

Os perfis metabólitos obtidos, figura 14 (a) e (c), destacam a eficiência na extração para compostos polifenólicos como observado nos altos valores de intensidade para os cromatogramas em 254, 280, 330 e 360 nm.

A determinação estrutural se deu através de metodologias clássicas de desreplicação. Inicialmente foram corrigidos os valores de massa de alta resolução através do calibrante Na.TFA. Em seguida, com auxílio do software *DataAnalysis* v. 4.0 (Bruker Daltonics) os espectros co-eluídos foram separados através da função *Dissect* (deconvolução), gerando uma lista de espectros “puros” relativo a cada composto detectado.

A partir destes dados foi avaliada as possíveis fórmulas moleculares através dos valores de massa de alta resolução utilizando o aplicativo *SmartFormula XR* (Bruker Daltonics). Neste processo, foram incluídas algumas recomendações desenvolvidas no trabalho por Kind et al. (2007) - as sete regras de ouro para identificação de compostos a partir de massas de alta exatidão.

Além das regras gerais da proporcionalidade entre o número de carbonos e hidrogênios citadas por Kind et al. (2007), foram incluídos filtros que determinam o número de nitrogênios, enxofre e metais (Na e K) no processo de triagem do software *SmartFormula XR*. A análise do erro em ppm entre os íons observados e as fórmulas moleculares assim como os valores de sigma, relativos ao padrão

isotópico, foram incluídos (SigmaFit – Bruker Daltonics). Uma vez reduzido o número de FM, estudos mecanísticos de fragmentação baseado em experimentos EM/EM foram propostos visando da estrutura.

Em paralelo, os dados de massas de alta resolução assim como os dados de EM/EM foram comparados às bases de dados Scifinder, Dicionário de Produtos Naturais, METLIN, METFRAG e a base construída para este trabalho (WOLF et al., 2010). Em geral, os dados de massas de alta resolução foram inseridos nas funções de busca retornando *hits* relacionados aos dados fornecidos. No caso do MetFrag além da massa do íon molecular foram adicionados os íons decorrentes da fragmentação.

Em seguida os *Hits* foram analisados a partir de filtros de origem biológica, partindo-se de espécies em direção a gênero e família (FUNARI et al., 2014; KIND et al., 2007). Deste modo foi possível direcionar a busca de estruturas baseado nos quimiotipos/quimiomarcadores em espécies relacionadas ao táxon *Solanum*.

A partir desta informação juntamente com as propostas mecanística de fragmentação que as substâncias foram determinadas. Ainda, o estudo do padrão de fragmentação facilitou a determinação de análogos estruturais de moléculas inicialmente detectadas. O resultado de todo processo de detecção com os respectivos tempos de retenção, dados de massas e os principais íons da fragmentação EM² são encontrados na tabela 8.

A coloração da legenda encontrada dentro das figuras 14 (b) e (d) refere-se a classe de cada substância: nucleotídeos – verde; aminoácido e derivados – laranja; poliaminas – vermelho; flavonoides – azul e, preto – glicoalcaloides e saponinas.

Ao todo foram detectados 26 substâncias (57-82) sendo em maioria derivados de polimaninas e flavonoides para as folhas enquanto nos ramos observa-se principalmente os glicoalcaloides.

Figura 14. (a) e (c) são perfis cromatográfico do extrato de folhas e ramos (respectivamente) da espécie *S. paniculatum* (M985, M986) obtidos a 254, 280, 330 e 360 nm enquanto as figuras (b) e (d) representam os cromatogramas de picos base (BPC). Os valores encontrados em (b) e (d) fazem referência aos compostos encontrados na tabela 7. A coloração da numeração representa a classe de compostos: verde (nucleotídeos), laranja (aminoácidos e derivados), vermelho (poliaminas), azul (flavonoides) e preto (glicoalcaloides e saponinas).

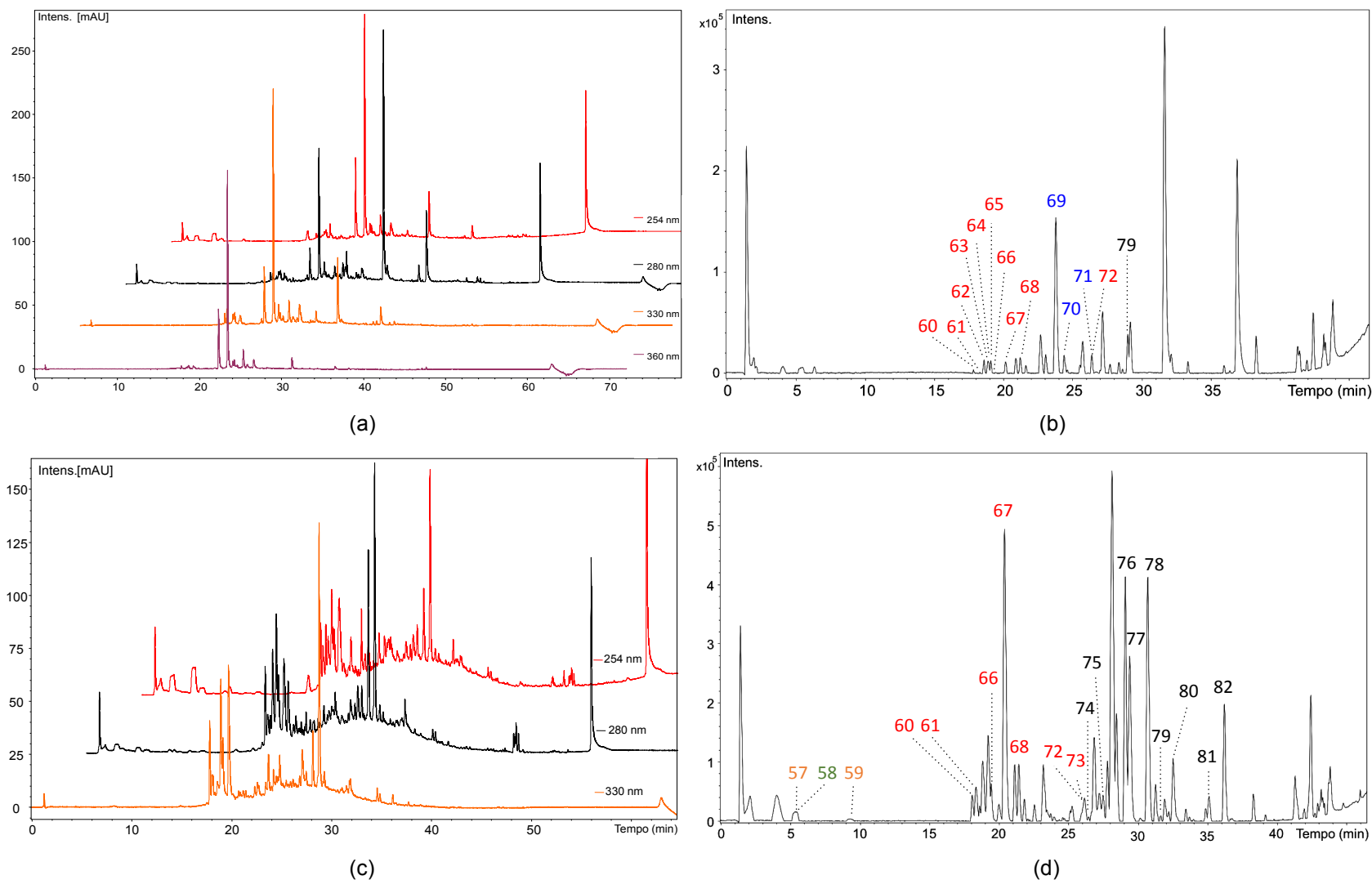


Tabela 8. Metabólitos identificados através da análise dos dados CLAE-DAD-EM-EM/EM do extrato de folhas e ramos de *S. paniculatum* (AMT - amostra 1 - M985 e amostra 2 - M986). São fornecidos o tempo de retenção (tR), a massa monoisotópica protonada(em azul - [M+H]⁺), erro em ppm do valor medido comparado ao valor teórico da formula molecular do composto (em vermelho) e os íons majoritários do espectro de fragmentação EM² (EM/EM).

No.	Nome	AMT	tR (min)	[M+H] ⁺ (Obs)	erro (ppm)	EM/EM (m/z)
57	frutose-L-(iso)leucina	2	5,1-5,2	294,1155	-1,5	276, 258 e 248
58	adenosina	2	5,6-5,8	268,1042	-0,5	224, 160 e 136
59	frutose-L-fenilalanina	2	9,5-9,7	328,1378	3,9	310, 292 e 282
60	cafeoil putrescina	1;2	17,8-18,0	251,1399	-3,7	234, 218, 163 e 135
61	feruloil putrescina	1;2	18,2-18,4	265,1527	7,3	248, 222, 177, 163 e 149
62	kukoamina A	1	18,4-18,6	531,3190	-2,5	367, 293, 261, 217, 169, 144 e 123,
63	N ¹ -cafeoil-N ⁴ - dihidrocafeoilespermina	1	18,8-19,0	529,3043	-4,3	367, 350, 293, 163, 149, 142 e 135
64	kukoamina B-hexose	1	19,1-19,3	693,3676	4,2	**
65	kukoamina B	1	19,1-19,4	527,2855	1,8	363, 346, 163
66	N ¹ -cafeoil-N ⁴ - dihidrocafeoilespermina- hexose	1;2	19,2-19,5	691,3566	-2,4	531, 529, 454, 367, 350, 293, 222 e 163
67	N ¹ ,N ³ - Bis(cafeoilespermidina)	1;2	20,5-20,6	474,2602	-0,7	457, 310, 293, 253, 236, 222, 165 e 123
68	N ¹ -cafeoil-N ³ - dihidrocafeoilespermidina	1;2	21,1-23,3	472,2422	4,4	457, 310, 293 236 e 222
69	QuercetinaO-di-hexose	1	22,5-22,8	627,1531	3,9	465, 303
70	Kaempferol-O-di-hexose	1	24,1-24,3	611,1597	1,5	449 e 287
71	Kaempferol-O-hexose+(160)	1	26,0-26,4	609,1400	1,6	449, 287
72	N-trans-cafeoil tiramina	2	26,1-26,3	300,1218	4,1	163
73	N ² -cafeoil-N ¹ ,N ³ Bis(cafeoilespermidina)	2	26,3-26,5	636,2929	-2,0	474, 472, 310, 235, 163
74	12-hidroxisolamargina (lobofrutosídeo)	2	26,4-26,7	884,5038	-4,1	738, 592, 430, 412, 394, 376, 287, 271, 269, 144, 126 e 114

75	12-hidroxi- γ_1 -solamarina	2	27,7-27,9	738,4423	0	592, 412, 394, 377, 376, 287, 269, 251, 144, 126 e 114,
76	<i>N</i> -trans-feruloiltiramina	2	29,0-29,1	314,1382	1,4	177, 149, 137, 121
77	δ -solanigrina ou leptinina I	2	29,7-30,0	868,5039	1,7	722, 414, 396, 379, 378, 271, 253, 144, 126 e 114,
78	β_2 -solamargina	2	30,0-30,1	722,4447	3,7	576, 414, 396, 379, 378, 271, 253, 126, 114,
79	Leptina I	2	31,5-31,7	910,5192	-3,7	764, 618, 522, 456, 271 e 253
80	23-OAc- β_2 -solamargina	2	32,3-32,4	764,4549	4,0	618 e 456
81	23-O-Ac-15-hidroxileptina I	2	35,0-35,2	926,5085	2,4	780, 634, 472, 454, e 436
82	23-O-Ac-15-hidroxi- β_2 -solamargina	2	36,0-36,1	780,4519	1,2	634, 472, 454 e 436

** não foi obtido o espectro de EM² para tal composto.

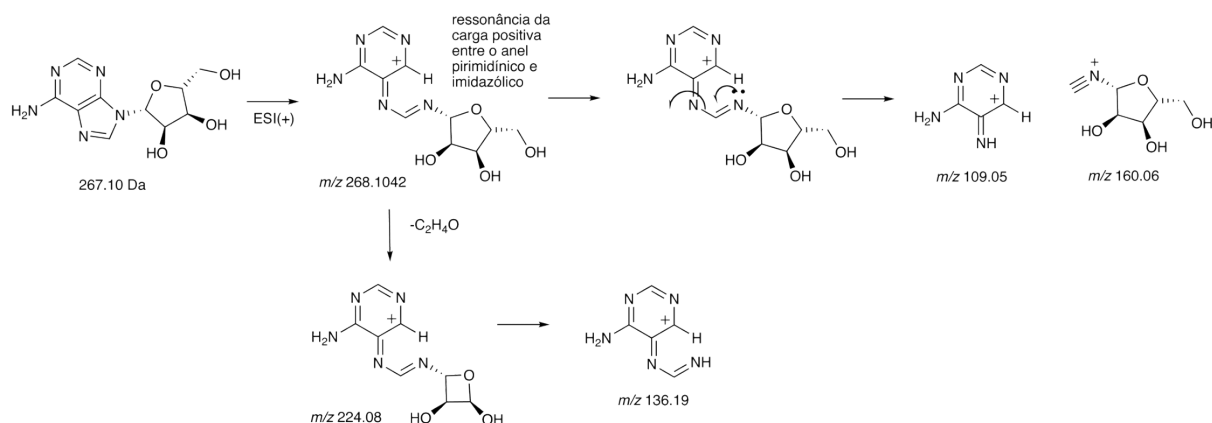
A seguir são mostrados os mecanismos de fragmentação que juntamente ao processo de desreplicação levou a determinação das estruturas encontradas na tabela 8.

Nucleosídeos

Os nucleosídeos são glicosilaminas formados por uma base nitrogenada e uma ribose. Estes compostos participam essencialmente de maior parte das rotas biossintéticas e apresentam, em sua maioria, características polares.

Adenosina (58) foi encontrada nos ramos de *S. paniculatum* (tR: 5,6 – 5,8 min) e determinada por duas rotas de fragmentação, esquema 1. A primeira via inicia pela quebra do anel imidazólico fundido ao anel pirimidínico gerando um íon contendo a ribose e um grupo ciano (*m/z* 160). A segunda rota procede pela saída de C₂H₄O (44 *u*) a partir do íon molecular, formando o íon *m/z* 224 seguido pela perda da ribose e formação do núcleo purínico (*m/z* 136) (GÓMEZ-ROMERO, et al., 2010; LIU, et al., 2008).

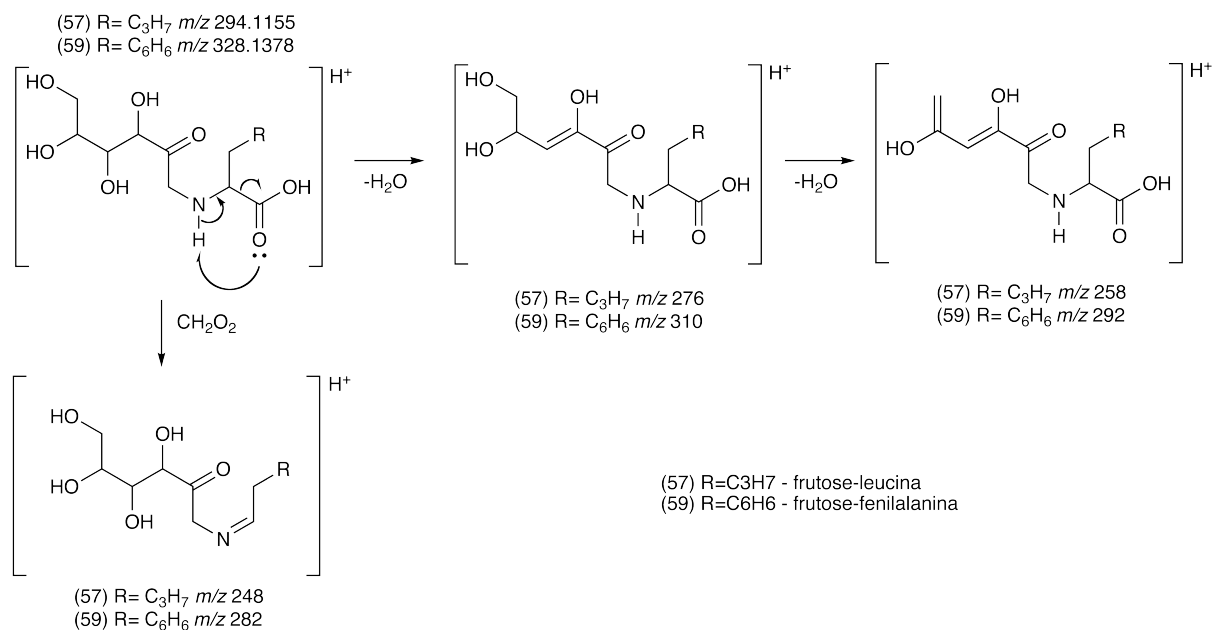
Esquema 1. Proposta de fragmentação da adenosina (58) em modo positivo de ionização (ESI).



Amino ácidos e Derivados

Dentre os principais aminoácidos presentes em plantas, os derivados da leucina (57) e fenilalanina (59) foram encontrados em maior concentração nos ramos de *S. paniculatum*. A determinação foi realizada com base em desidratações sucessivas e a perda do fragmento CH_2O_2 (46 u) esquema 2 (GÓMEZ-ROMERO et al. 2010).

Esquema 2. Proposta de fragmentação da frutose-(iso)-leucina (57) e frutose-fenilalanina (59) em modo positivo de ionização (ESI).



(57) R=C₃H₇ - frutose-leucina
(59) R=C₆H₆ - frutose-fenilalanina

Flavonoide O-glicosilados

Os compostos 69, 70 e 71 foram atribuídos aos derivados de kaempferol e quercetina O-glicosilados. O processo de desreplicação para estas estruturas foi baseado na comparação de dados de absorção de UV, massas de alta resolução e do perfil de fragmentação com dados da literatura (MARCH et al., 2006; VUKICS; GUTTMAN, 2010; WARIDEL et. al., 2001; YANG, W. et al., 2014).

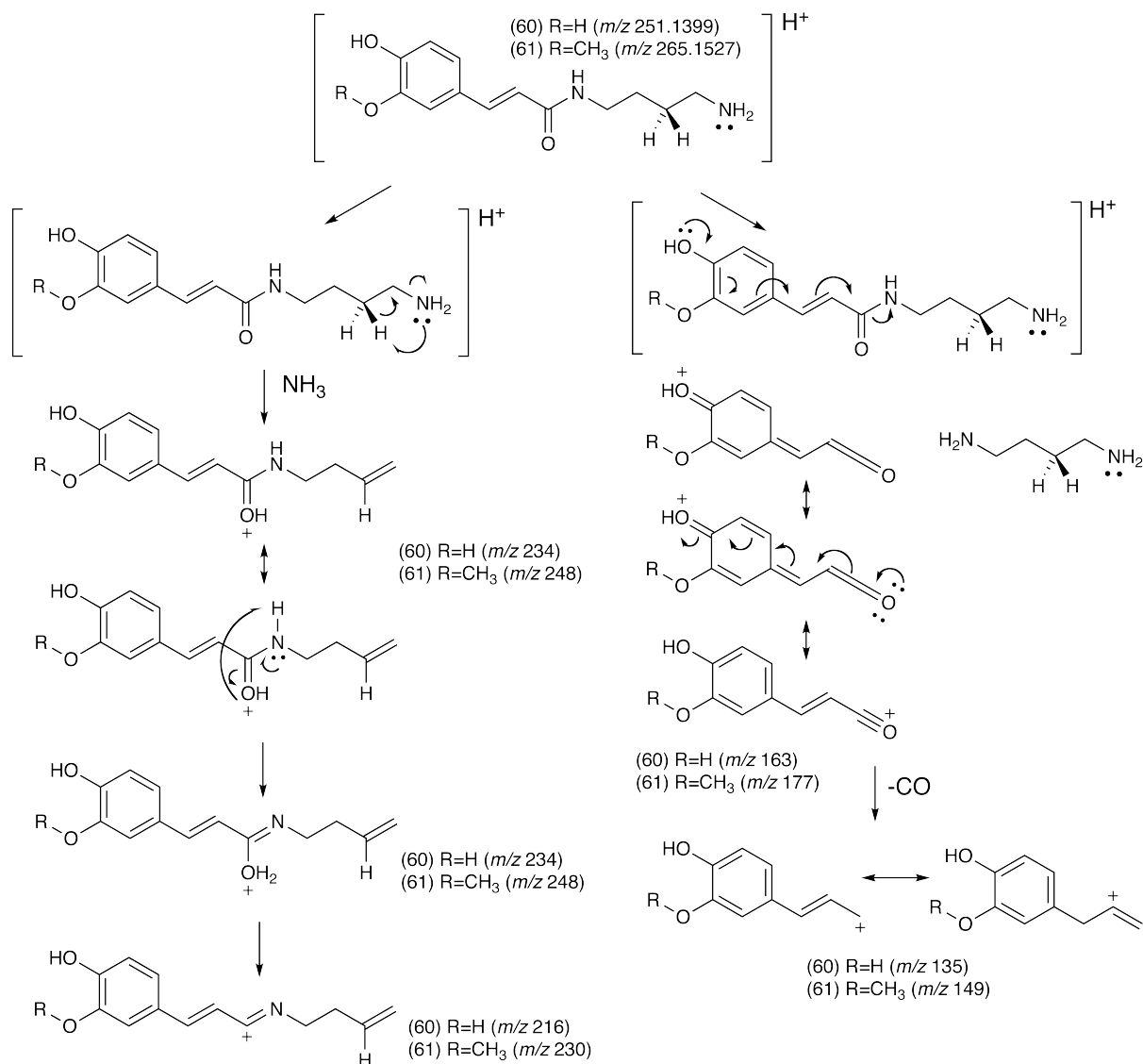
Poliaminas e Amidas

Em plantas, as poliaminas são comumente conjugadas com ácidos hidroxicinâmicos, sendo que a proporção entre as formas livre e conjugada varia muito entre as diferentes espécies de plantas (Silveira, 2008).

A conjugação se dá por meio da formação de uma ligação amídica, na qual ésteres de coenzima A (CoA) fornecem o grupo carboxil ativado. A putrescina e a espermidina são conjugadas por transferases, que diferem em sua especificidade para os derivados hidroxicinamóil-CoA. Assim, a reação entre cafeoil-CoA e a putrescina produz uma amida do ácido hidroxicinâmico denominada cafeoilputrescina (da Silveira, 2008). A putrescina pode formar monômeros ou dímeros de ácidos solúveis e insolúveis com os ácidos cumárico, caféico e ferúlico. As poliaminas também podem se ligar a componentes da parede celular, como a hemicelulose e a lignina (Silveira, 2008).

A partir da análise dos extratos de *S. paniculatum* foram encontrados derivados da putrescina, espermina e espermidina na forma de amidas como citado acima. A cafeiol putrescina (60) foi encontrada na folhas e ramos de *S. paniculatum* enquanto o feruloil putrescina (61) somente nas folhas. A caracterização molecular (composto 60 e 61) foi baseada pelo padrão de fragmentação. A clivagem da amina terminal gerou a saída de amônia (17 u) seguida por uma perda de água. Além disso, a quebra da ligação amídica, formação do íon acílio a partir do ácido (feruloil m/z 177 e cafeiol m/z 163) e perda de CO a partir do íon acílio (m/z 135 e 149) permitiram a confirmação dos derivados da putrescina assim como os derivados do ácido hidroxicinâmico, esquema 3.

Esquema 3. Proposta de fragmentação para o cafeoil e feruloil putrescina (60 e 61) em ESI(+)-MS.

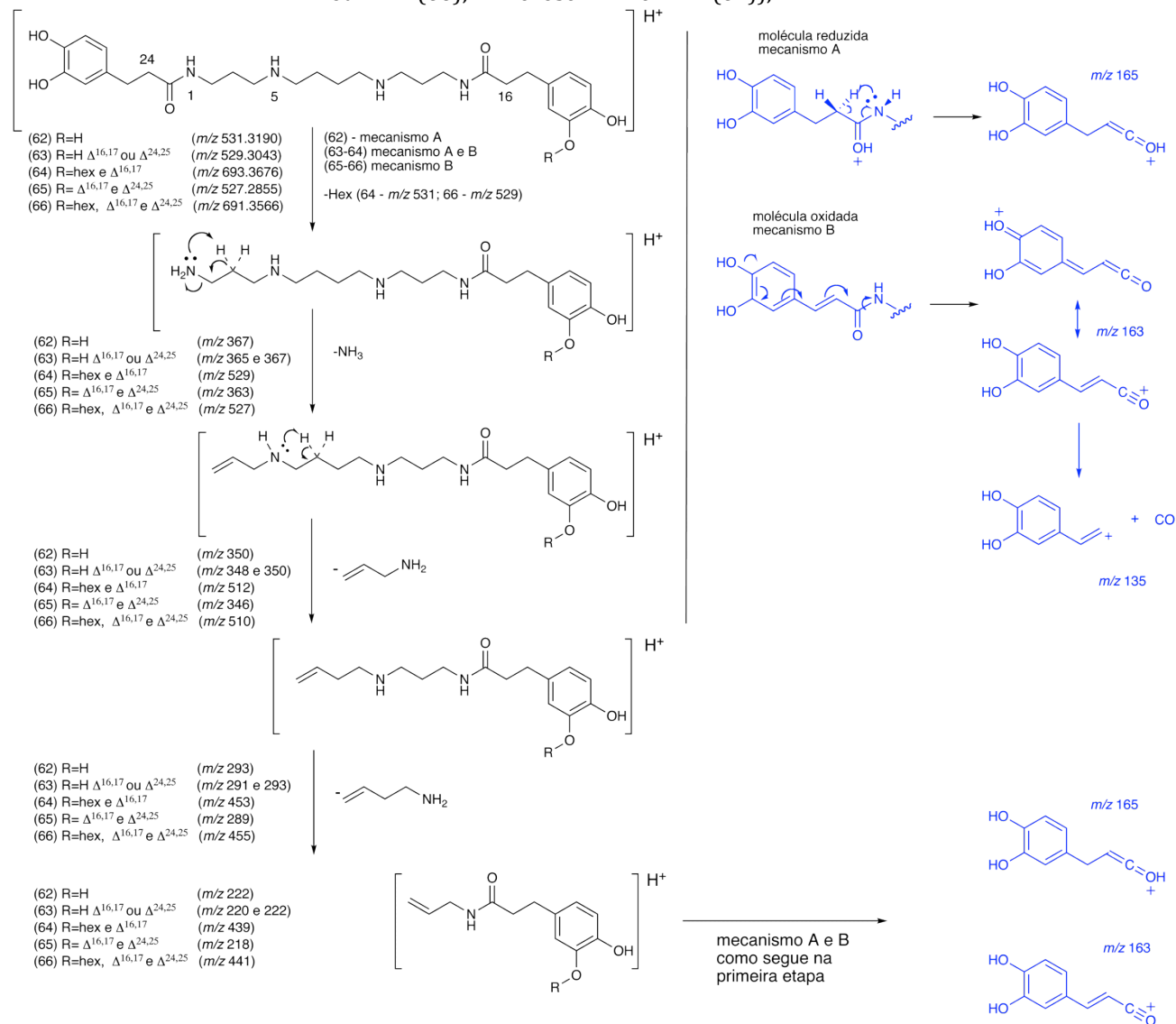


O composto 72 segue o mesmo padrão de fragmentação para os derivados de putrescina exceto pela perda neutra da amônia (PARK, et al., 2012). Para os derivados da espermina, substâncias 62 a 66, as rotas de fragmentação são as mesmas observadas para análogos da putrescina. De maneira geral observa-se a uma quebra heterolítica da ligação amídica (formação do derivado acílio do ácido hidroxicinâmico), perda de amônia (17 u) e fragmentos da cadeia amínica. A proposta de fragmentação dos adutos formados são mostrados no esquema 4 (PARR, et al., 2005).

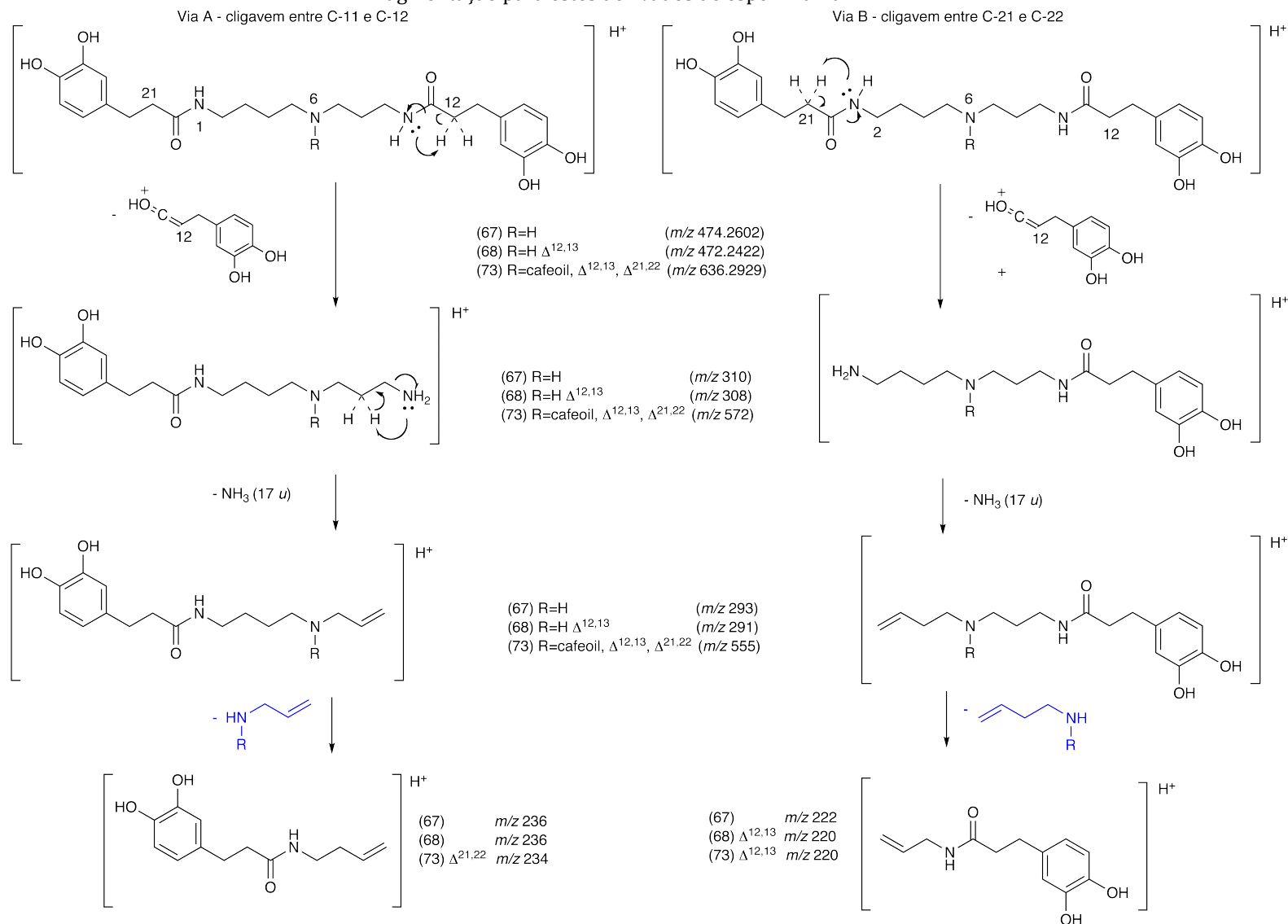
As substâncias 67, 68 e 73 (esquema 5) são derivados amídicos da espermidina ligados aos mesmo ácidos hidroxicinâmicos. A determinação estrutural foi proposta através mecanismos de fragmentação, semelhantes aos derivados da

espermina, esquema 5. Assim como outros derivados poliamínicos, observa-se a perda de amônia, alquilaminas primárias e fragmentos dos ácidos hidroxicinâmicos. Contudo, devido assimetria da disposição das aminas na espermidina a quebra da ligação amídica fornece fragmentos diferentes como mostrado no esquema 5 (Via A e Via B) (PARR, et al., 2005).

Esquema 4. Proposta de fragmentação dos derivados de espermina em modo positivo. (R=H (62)); R=H $\Delta^{16,17}$ ou $\Delta^{24,25}$ (63); R=H $\Delta^{16,17}$ e $\Delta^{24,25}$ (65); R=Hexose $\Delta^{16,17}$ ou $\Delta^{24,25}$ (66); R=Hexose $\Delta^{16,17}$ e $\Delta^{24,25}$ (64));



Esquema 5. Proposta de fragmentação de derivados de cafeoil espermidina (67, 68 e 73) em modos positivo. Via A e Via B representam as duas possíveis rotas de fragmentação para estes derivados de espermidina..



Alcaloides Esteroidais

A maioria das espécies de Solanaceae biossintetiza alcaloides esteroidais em C-27 baseados em esqueletos colestanos. Estes são essencialmente análogos nitrogenados de saponinas esteroidais e, na sua maioria, possuem a mesma estereoquímica em C-25 (metila equatorial). Contudo, existem isômeros em C-22 como a solasodina e a tomatidina. Geralmente, estes compostos estão presentes na forma de glicosilados e são tóxicos se ingeridos. Uma forma variante da cadeia lateral dos colestanos ocorre pela sua ciclização formando um sistema de anéis condensados com o nitrogênio na forma de ponte, os chamados solanidanos (DEWICK, 2002).

A estabilidade estrutural do esqueleto colestano dificulta a fragmentação destes glicoalcaloides. Assim, poucos íons são formados após o experimentos de MS/MS e muitos daqueles são inespecíficos para sua determinação estrutural. A literatura tem mostrado um baixo número de trabalhos e, ainda muitos deles incongruentes nas propostas de fragmentação e formação dos íons produto (CAHILL et al., 2010; CLAEYS et al., 1996).

Segundo Claeys et al. (1996) o processo de fragmentação (FAB-EM/EM de baixa energia) de glicoalcaloides esteroidais e algumas de suas agliconas e.g. tomatidina e solasodina (esqueleto do tipo espirossolano) e demissidina e solanidina (esqueleto do tipo solanidano) as dissociações envolvendo energias brandas resultam de dois mecanismos gerais: fragmentações dirigidas pela carga (*driven charge fragmentations*) ou por carga remota ou à distância (*remote charge fragmentations*).

As fragmentações dirigidas por carga parecem dominar o espectro de massas de compostos contendo anéis saturados. Essas reações resultam na perda de partes da cadeia lateral como espécies neutras e fornecem informações úteis sobre isômeros estruturais. Já as fragmentações por carga remota são mais intensas na presença de insaturações e também fornecem informações sobre as posições das mesmas ou de grupos hidroxil ou cetônicos nos anéis. Porém, durante as etapas de clivagem via carga remota os autores sugerem migrações e rearranjos de hidreto no anel C, o que é pouco provável para fragmentações do tipo CID, figura 15 A (via II) (CLAEYS et. al., 1996).

De acordo com Crotti e colaboradores (2006), a ionização para grande parte das substâncias orgânicas que não apresentam sistemas π conjugados, ocorre geralmente via reações ácido-base, como podemos sugerir para algumas classes de glicoalcaloides (CROTTI et al., 2006).

Especificamente, espirossolanos poderiam estabilizar um próton (H^+) através de uma quelação entre os heteroátomos espacialmente próximos, O e N nos anéis E e F respectivamente. Nesse caso, os íons produtos resultariam de duas vias fragmentação. Ora o íon H^+ fica retido no átomo de N, ora no átomo de O. A figura 15 (B) representa o processo de fragmentação via protonação do oxigênio. A assistência do par de elétrons não ligantes do nitrogênio (agente nucleófilo) em C-22 resulta na abertura do anel furânico (anel E) e deslocalização carga positiva entre o nitrogênio e o carbocátion em C-22 (estruturas canônicas de ressonância).

No caso onde o par de elétrons do oxigênio age como nucleófilo interno através de uma assistência anquimérica em C-22, ocorre a abertura do anel F seguida da perda cadeia lateral (anéis E e F) e na formação de um carbocátion alílico estável m/z 273, figura 15 (C). Adicionalmente, a presença de hidroxila em C-3 pode resultar na formação do íon m/z 255 através da perda de uma água, conforme representado na figura 15 (A).

Outros fragmentos característicos como m/z 161 e 147 foram determinados como resultado da perda de água e de múltiplas clivagens no anel C (mecanismos dirigidos por carga), resultando em íons alílicos, representados na figura 16. Outro íon característico dessa classe de glicoalcaloides é a perda de 143 u, gerada pela perda da cadeia lateral do anel F conforme mostrado na figura 15 (C).

Os metabólitos 74 a 78 foram atribuídos a glicoalcaloides do tipo espirosolanos em C-22. As proposta de fragmentação e determinação da posição dos substituintes está presente no esquema 6. Em todos os casos observa-se perdas sucessivas de hexose e pentoses (162 m/z e 146 m/z) até a formação da aglicona. A partir da massa de alta resolução foi possível atribuir as agliconas como os íons m/z 414 e 430.

Durante o processo de fragmentação dos glicoalcaloides são observados desidratações naturalmente associadas aos glicosídios, no entanto, uma vez formado a aglicona, este tipo de reação é notado somente na presença de grupos hidroxil no anel colestano. Os resultados da fragmentação da aglicona para os espirosolanos mostraram duas ou três desidratações dependendo dos substituintes na

molécula. Dessa forma, baseados em dados da literatura, traçamos uma proposta racional de fragmentação para os espirosolanos, partindo de uma clivagem no anel E gerando uma imina com uma carga positiva deslocizada entre no carbono terciário em C-22 e o nitrogênio, e formação de grupo hidroxil ligado ao anel D. A partir desses resultados foi possível determinar o número hidroxilas presentes em cada aglicona (Cahill, et al., 2010; Clayes, et al., 1995).

Os derivados de espirolanos apresentam o grupo pirimidínico ligado a um anel furânico através de um carbono quaternário, o qual é label para diferentes tipos clivagem, gerando assim diferentes vias de fragmentação para uma mesma aglicona. Dentro desta perspectiva foi observado uma perda neutra de 17 u, a qual foi atribuída a saída de amônia. Isto sugere uma via de fragmentação alternativa da mencionada acima, consistindo na abertura anel pirimidínico para a formação de uma amina primária terminal e uma dupla ligação no anel furânico. Assim, foi possível justificar o íon 379 m/z ou 395 m/z (íon hidroxilado) através saída da amônia e formação de uma ligação dupla terminal. Outros íons justificam tal proposta, como a desidratação a partir do fragmento 395 m/z gerando o 377 m/z , esquema 6 (Cahill, et al., 2010; Clayes, et al., 1995).

A estrutura condensada desses glicoalcaloides dificulta processo de fragmentação em modo suave de ionização como é o caso do eletrospray. No entanto, em experimentos EM/EM de alta voltagem (entre 50 e 60 eV) é possível observar algumas clivagem e assim determinar as posições de outros substituintes na molécula como os grupos hidroxil.

Assim, foram realizados experimentos *tandem* em 60 eV para os alcaloides esteroidais e, através da rota da imina, foi possível observar a clivagem do anel E gerando um íon alílico no anel D (m/z 271 e 287 para o composto hidroxilado) e formação de um derivado pirimidínico de m/z 126. Foi possível ainda determinar a posição da hidroxila por meio de uma desidratação a partir do íon alílico (de 287 para 269 m/z) juntamente com dados de biossíntese da solasodina, esquema 6 (DEWICK, 2002; CAHILL et al., 2010; CLAYES et al., 1995).

Nos experimentos EM/EM de alta enrgia ainda forneceram subsídios para uma segunda via de clivagem do núcleo terpênico. Foi possível observar através da quebra do anel pirimidínico, um intermediário carbonílico que, subsequentemente, levou a perda de um derivado amínico e formação do mesmo íon alílico (m/z 271).

No entanto, são gerados dois derivados amínicos, sendo o primeiro formado pela quebra da ligação do anel D e o carbono α carbonílico gerando um composto com uma amina terminal e uma carbonila protonada de razão massa-carga 144 Da. A segunda via se dá pela eliminação de um etano (C_2H_6) e formação de um derivado de íon acílio de razão massa-carga 114 m/z , esquema 6.

Os compostos 79 a 82 são glicoalcaloides do tipo solanidanos, o anel pirimidínico está condensando ao núcleo do triterpeno pelo nitrogênio. As propostas de fragmentação para estes metabólitos são mostrados no esquema 7. Estes compostos apresentam maior rigidez em sua estrutura condensada a fragmentação e consequentemente fornece um número menor de íons. Basicamente, foram gerados íons a partir da quebra de glicosídeos, desidratações, perda de ácido acético (60 u) e quebra do anel E, esquema 7 (CAHILL et al., 2010; CLAYES et al., 1995).

Figura 15. (A) Proposta mecanística para a formação do íon m/z 273 e 161 a partir do perfil da fragmentação da tomatidina segundo CLAEYS *et.al.*, 1996; (B) Mecanismo de fragmentação para perda de água; (C) mecanismo de fragmentação para cadeia lateral (anel E e F) e perda neutra de 143 Da.

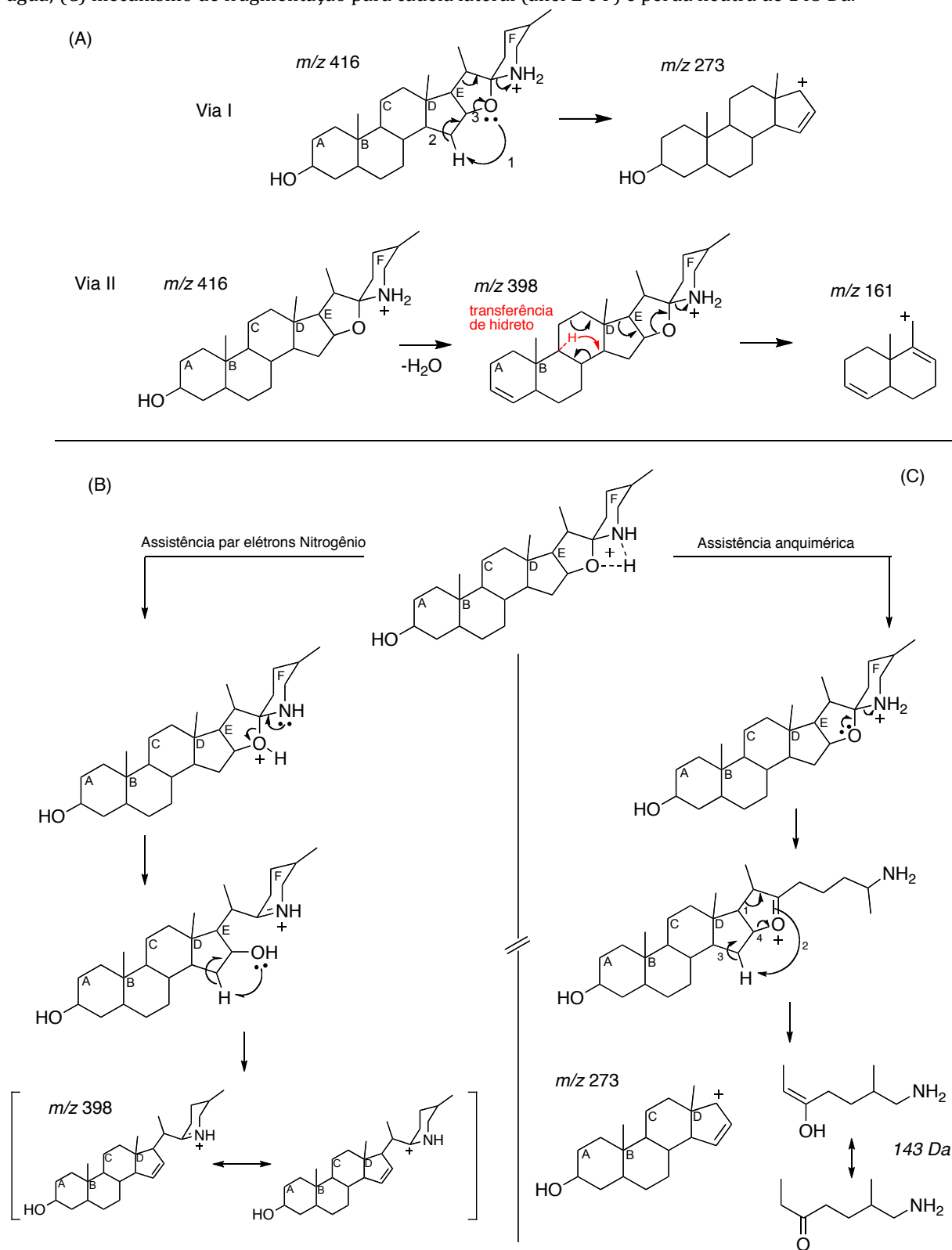
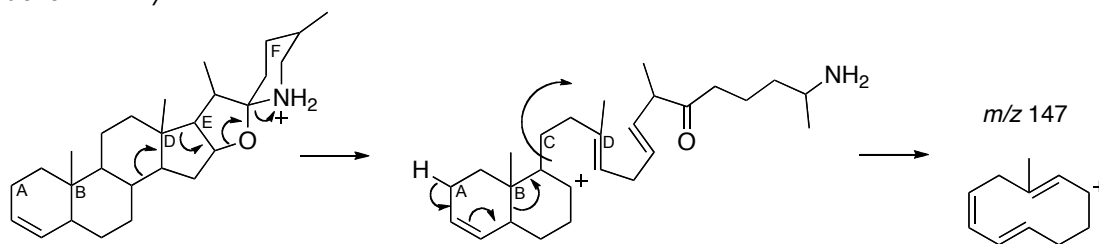
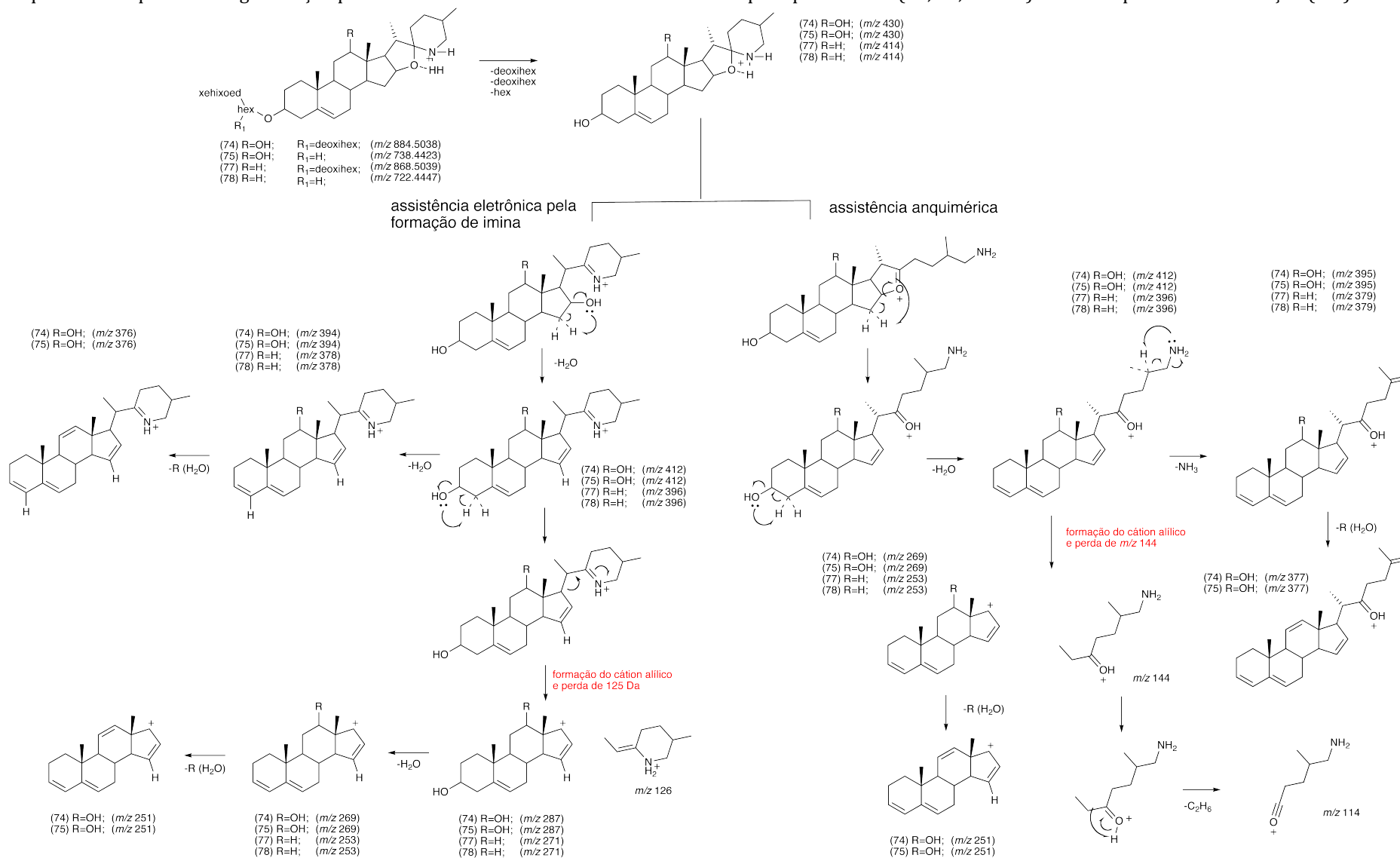


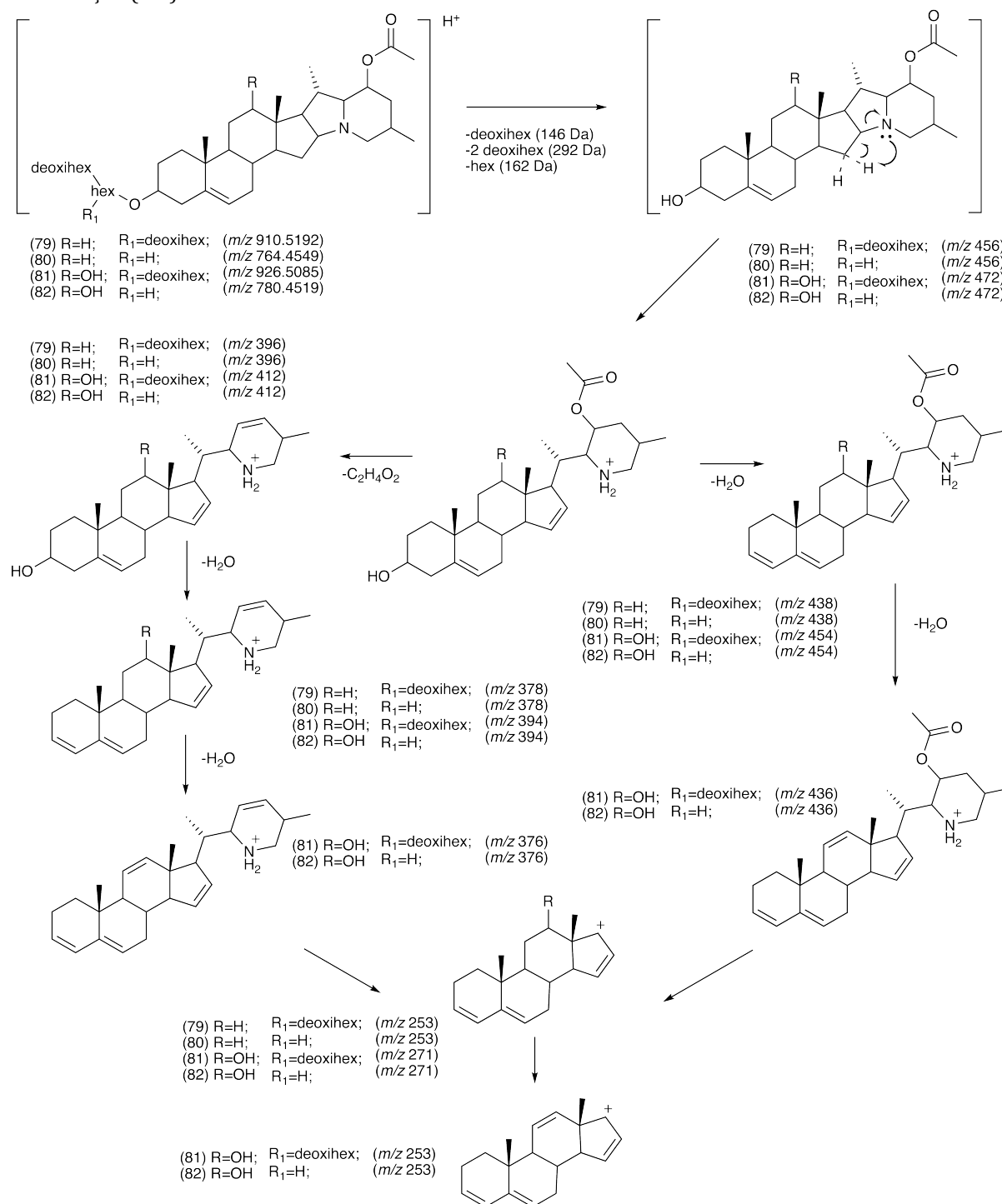
Figura 16. Mecanismos de fragmentação dirigido por carga remota (à distância) proposto para a formação do íon 147 m/z .



Esquema 6. Proposta de fragmentação para os derivados de alcaloides esteroidais do tipo espirossolano (74, 75, 77e 78) em modo positivo de ionização (ESI).



Esquema 7. Proposta de fragmentação para os alcaloides do tipo solanidano (79 – 81) em modo positivo de ionização (ESI).



Perfil metabolômico e desreplicação de metabólitos secundários através de CLAE-AutoEM/EM e redes de interações moleculares (*molecular networking*)

Em estudos metabolômicos o procedimento usual para a identificação de metabólitos secundários começa pela análise de perfil de fragmentação de massas *tandem* (CLAE-EM/EM) (WATROUS et al., 2012). Como observado nas seções anteriores, os resultados então são comparados com bases de dados (METLIN, MassBank, Dictionary of Natural Products (DNP), entre outras) e as possíveis estruturas são apontadas (METLIN: <https://metlin.scripps.edu/>; MassBank: [www.massbank.jp/?lang=en](http://www.massbank.jp/?lang=en;); DNP: dnp.chemnetbase.com/).

Contudo, o processo de desreplicação limita-se aos quimiotipos mais comuns e assim, impossibilitando a identificação de compostos inéditos e/ou incomuns. Em geral, a identificação destes compostos resulta da análise manual dos dados de fragmentação. Esse cenário é ainda pior quando se faz uso da abordagem metabolômica onde é necessário a análise e comparação de milhares de espectros (WATROUS et al., 2012, YANG et al., 2013).

Dessa forma é fundamental o desenvolvimento e aplicação de ferramentas que permitam organizar, navegar e analisar grandes conjuntos de dados. Dentro dessa perspectiva, a rede de interações moleculares (*Molecular networking*) surgiu como ferramenta para integração de elementos de similaridade em dados proteômicos (WATROUS et al., 2012). No entanto, essa ferramenta vem sendo aplicada a estudos metabolômicos por uso de dados de fragmentação EM/EM (realizados para cada molécula) congregando classes de metabólitos que partilham de um “mesmo” perfil espectral (WATROUS et al., 2012; YANG et al., 2013).

Para a aplicação das redes moleculares utilizando dados de massas são necessárias quatro etapas: (1) coleta dos espectros de massas EM/EM (configuração AutoEM/EM) e conversão em formato *.mgf* ou *.mzdata*; (2) agrupamento de espectros pertencentes a um mesmo pico e/ou mesmo composto (nível de similaridade >99%) (3) criação da rede molecular a partir dos valores de similaridade calculados a partir do ângulo de cosseno para um par de vetores (par de espectros); (4) análise da rede molecular (inclusão ou não de uma base de dados), figura 11.

A identificação de moléculas é mais rápida quando espectros EM/EM de padrões autênticos são incluídos embora ela não seja essencial. A biblioteca atua como uma “semente” espectral, isto é, serve como ponto de partida para determinação de compostos dentro de um agrupamento de dados.

A partir dos extratos brutos obtidos segundo esquema da figura 8 foram realizados experimentos em CLAE-AutoEM-EM/EM de seis amostras de *Solanum* (folhas e ramos de *S. americanum*, M951 e M952; folhas e ramos de *S. swartzianum*, R271, R272 e F052 e folhas de *S. excelsum*, F55), figura 17. Os mapas de rede de interação molecular foram gerados a partir da combinação de todos os dados de EM/EM para todas as espécies estudadas. A elucidação estrutural de cada nodo (composto) foi baseada na mesma metodologia descrita seção anterior utilizando: FM, perfil de fragmentação e dados quimiotaxômicos das espécies (bases de dados). Ao todo foram obtidos 19 agrupamentos (47 compostos), dos quais três foram identificados até o momento: (1) flavonoides O-glicosilados; (2) glicoalcaloides esteroidais/saponinas e (3) flavonoides C-glicosilados, figura 18.

Os agrupamentos foram organizados de acordo com 6 fatores: (1) força de aresta (largura), representa o nível de similaridade entre os perfis espectrais. Quanto mais grossa a largura de aresta maior a similaridade entre os nodos. Como forma de facilitar a visualização, as bordas foram coloridas de acordo critério de similaridade relativo a largura: azul>vermelho>preto; (2) A cor da legenda no interior de cada nodo refere-se a parte da planta onde o composto foi detectado, folhas e/ou ramos; (3) o formato de nodo foi associado a um valor massa que, tipicamente revelou íons relativos as agliconas; (4) cor do nodo, representa a(s) espécie(s) em que os compostos foram encontrados; (5) tamanho do nodo, demonstra a presença de 1 ou mais de isômeros para um dado m/z e (6) cor de borda foi associada ao número e ao tipo de substituinte ligado aos compostos, e.g. glicosídeos.

Figura 17. Perfil metabolômico obtido do cromatograma de íons totais (TIC) para amostras de *Solanum*: (A) folhas *S. swartzianum* – R271; (B) ramos *S. swartzianum* – R272; (C) folhas *S. swartzianum* – F052; (D) folhas *S. americanum* – M951; (E) ramos *S. americanum* – M952; (F) folhas *S. excelsum* – F55. A numeração refere-se aos compostos presentes nas tabelas 8, 9 e 10. A coloração da numeração refer-se: preto (flavonoides *O*-glicosilados – tabela 8); vermelho (flavonoides *C*-glicosilados – tabela 9) e azul (glicoalcaloides esteroidais – tabela 10);

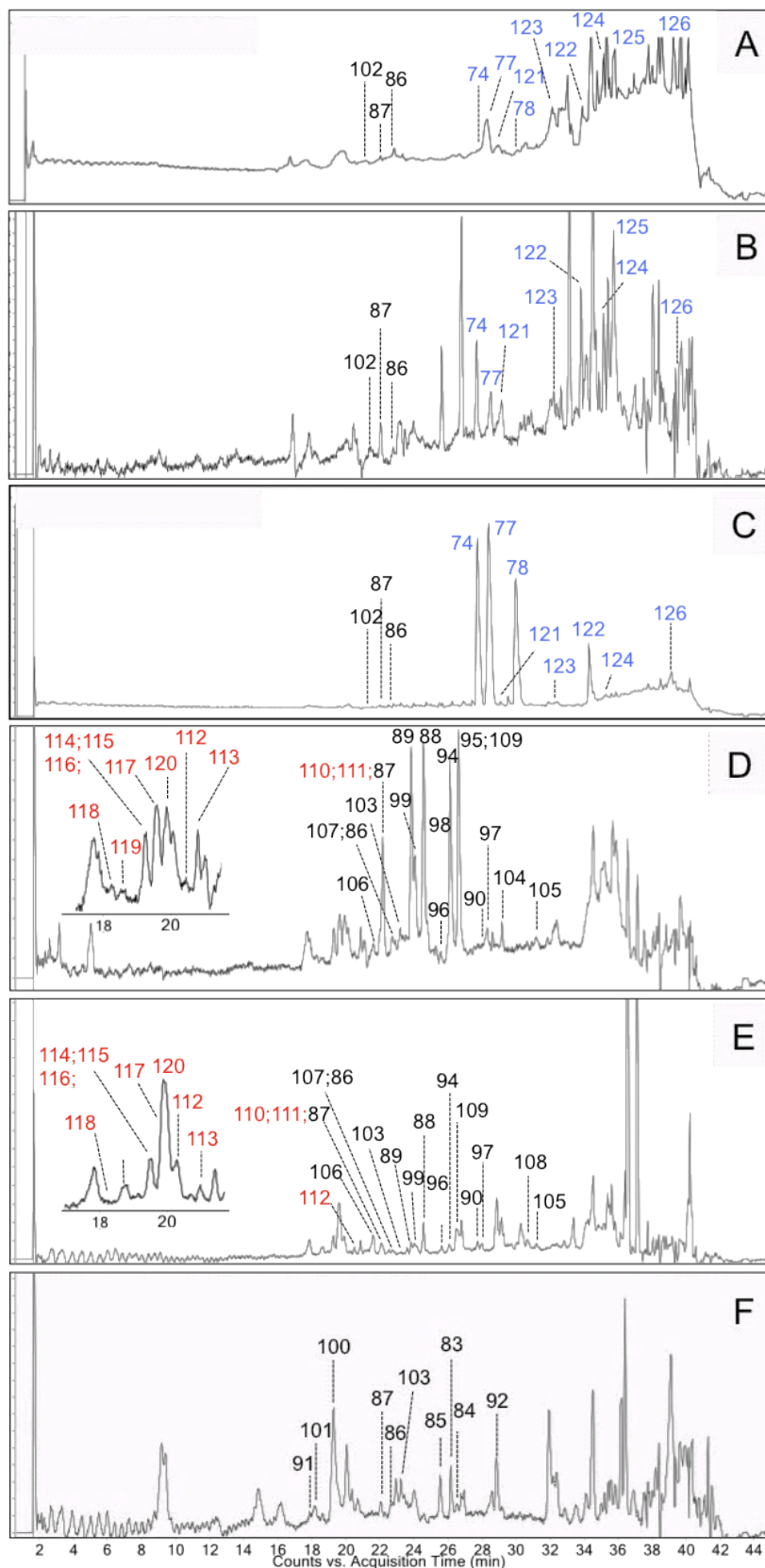
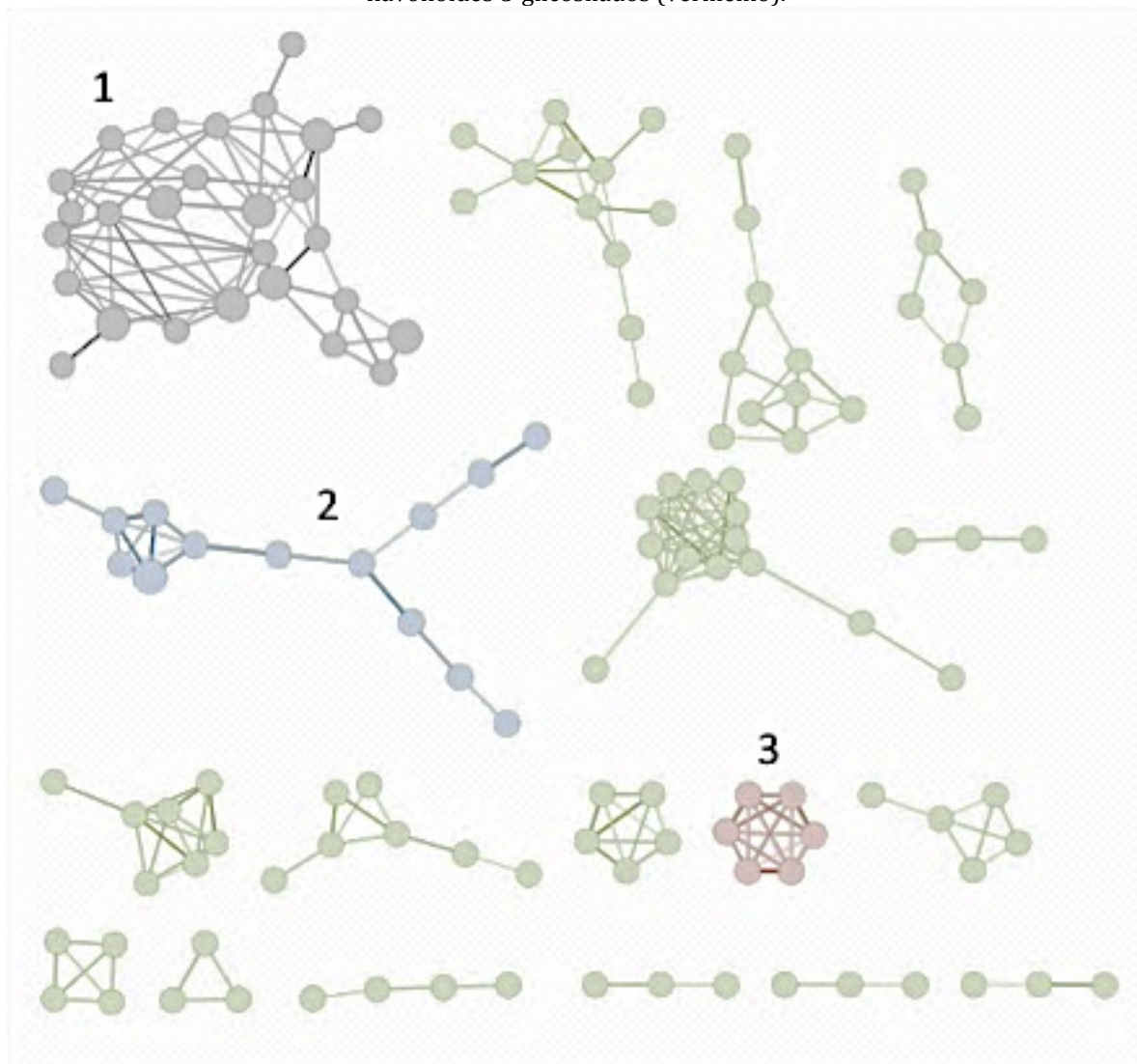


Figura 18. Mapa da rede molecular (*molecular networking*) obtida para os dados EM/EM das espécies de *Solanum*: 1- Flavonóides *O*-glicosilados (cinza); 2 – glicoalcaloides esteroidais – saponinas (azul) e 3 – flavonoides *C*-glicosilados (vermelho).



Flavonoides *O*-glicosilados

O gênero *Solanum* é caracterizado pela presença de flavonóis glicosilados, metoxilados e acilados usualmente de derivados do kaempferol e quercetina e raramente derivados de miricetina. A glicose e a ramnose são mais freqüentes enquanto a galactose, arabinose e a xilose são menos usuais. Adicionalmente, a glicosilação de seus compostos sucede principalmente em C-3 e, metilações e hidroxilações em C-3', C-4' e C-8 (SILVA et al., 2003).

A importância dessa classe é tal que Steinharter (1986) propôs uma classificação taxonômica para grupos infragenéricos de *Solanum* apoiado na

complexidade biossintética de flavonóis (EICH, 2008). Na proposta, a mono e diglicosilação em 3-O- e 3,7-O- foi usada para definir grupos basais e a metoxilação em 3, 4', 7 e/ou hidroxilação/metoxilação em C-8 para os subgêneros mais avançados na classificação filogenética.

A partir da análise do *molecular networking* e do espectro EM/EM, figura 18, foi detectado um grupo formado por 27 flavonoides O-glicosilados, figura 19.

Dentre as espécies estudadas, *S. americanum* apresentou 18 flavonoides derivados de flavonas (apigenina e luteolina) e somente um derivado metoxilado de flavonol *m/z* 317 (metilquercetina). Para espécie *S. excelsum* foram identificados 9 derivados metoxilados de flavonóis (kanferol e quercetina) e 4 derivados de kanferol para a espécie *S. swartzianum*, tabela 8.

Tabela 9. Metabólitos identificados através da análise dos dados CLAE-AutoEM/EM das amostras *S. swartzianum* (1 - R272, 2 - R271, e 3 - F052), *S. americanum* (4 - M951 e 5 - M952) e *S. excelsum* (6- F55). O tempo de retenção (tR), a massa monoisotópica protonada(em azul - [M+H]⁺), erro em ppm do valor medido comparado ao valor teórico da fórmula molecular do composto (em vermelho) e os íons majoritários do espectro de fragmentação EM² (EM/EM) também são mostrados na tabela.

Nº.	Composto	Amostra	tR (min.)	[M+H] ⁺ obs.	erro (ppm)	EM/EM (<i>m/z</i>)
83	Dimetilquercetina-O-hex	6	26,0-27,0	493,1349	-2,1	331 ^{b.p.} ; 316, 301, 273, 217, 177, 161, 153
84	Trimetilquercetina-O-hex	6	26,5-26,8	507,1506	-1,6	345 ^{b.p.} ; 330, 315; 301, 287, 259
85	Dimetilkanferol-O-hex	6	25,7-26,0	477,1401	-3,6	315 ^{b.p.} ; 300, 272, 257, 229 121
86	Dimetilquercetina-O-hex	1, 2, 3, 4, 5 e 6	22,7-23,7	479,1196	-4,1	317 ^{b.p.} ; 302, 288, 274, 257, 228, 137, 110
87	kanferol-O-hex	1, 2, 3, 4, 5 e 6	22,1-22,4	449,108	-4,2	287 ^{b.p.} ; 258, 241, 231, 165, 153, 147, 127, 121, 115
88	crisoeriol-O-hex	4 e 5	24,4-24,7	463,1229	-3,4	301 ^{b.p.} ; 286, 258, 229, 153
89	crisoeriol-O-hex-O-pent	4 e 5	24,1-24,5	595,1652	-6,7	301 ^{b.p.} ; 463, 286, 258, 229, 153, 133, 115

90	crisoeriol-O-hex-O-pent/O-hex-O-acetomalonil	4 e 5	27,7-28,2	721,1959	-8,0	301 ^{b.p.} ; 589, 506, 449, 439, 361, 325, 305, 286, 278, 271, 255, 253, 241, 207, 205, 199, 177, 145, 133, 127, 115
91	metilkanferol-O-hex-O-pent	6	17,6-18,0	625,1783	-7,1	301 ^{b.p.} ; 463, 286, 258, 247, 153, 145, 127, 109
92	metilkanferol-O-hex-O-ácido cafeico	6	29,2-29,4	625,1783	-6,0	301 ^{b.p.} ; 485, 387, 286, 229, 197, 185, 177, 163, 157, 145, 137
93	metilquercetina-O-hex-O-ácido cafeico	6	27,2-27,5	641,1513	-0,9	317 ^{b.p.} ; 538, 426, 359, 307, 302, 294, 286, 275, 255, 238, 234, 225, 218, 209, 207, 203, 177, 163
94	crisoeriol-O-hex-O-malonil	4	26,6-26,9	549,1242	-2,5	301 ^{b.p.} ; 286, 271, 264, 258, 231, 163, 155, 147
95	crisoeriol-O-hex-O-pent/O-hex-O-acil	4	26,9-27,1	637,1743	-9,0	301 ^{b.p.} ; 556, 473, 371, 286, 276, 258, 245, 227, 215, 187, 183, 177, 159, 149, 133, 128, 115
96	apigenina-O-hex-O-pent/O-hex-O-malonil	4 e 5	25,7-26,0	651,1547	-4,9	271 ^{b.p.} ; 519, 433, 313, 231, 213, 153, 145, 133, 127, 115
97	apigenina-O-hex-O-pent/O-hex-O-acil	4 e 5	28,2-28,5	607,167	3,0	271 ^{b.p.} ; 475, 205, 187, 153, 133, 127, 115
98	luteolina-O-hex-O-pent/O-hex-O-acil	4	26,4-26,7	623,1606	2,1	287 ^{b.p.} ; 542, 491, 329, 187, 153, 145, 133, 127, 115

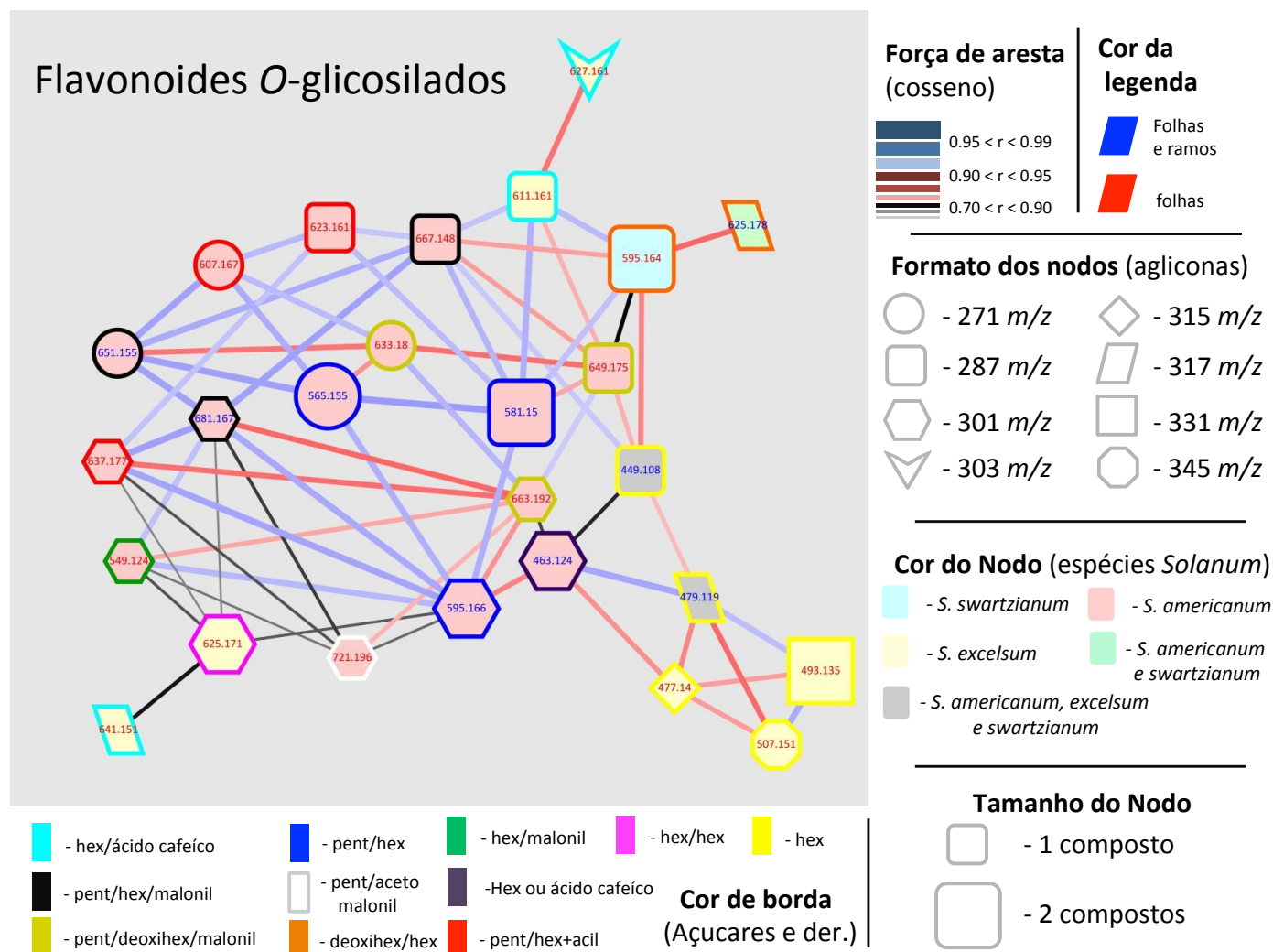
99	luteolina-O-hex-O-pent/O-hex-O-malonil	4 e 5	24,0-24,5	667,1482	4,3	287 ^{b.p.} ; 581, 535, 449, 257, 231, 159, 145, 133, 127, 115
100	kanferol-O-di-hex	6	19,2-20,2	611,1613	-3,0	287 ^{b.p.} ; 449, 431, 413, 353, 329, 213, 163, 145, 127, 115
101	quercetina-O-di-hex	6	18,0-18,4	627,1613	-2,4	303 ^{b.p.} ; 467, 289, 253, 235, 229, 170 145, 133;
102	kanferol-O-hex-O-deoxihex	1, 2 e 3	21,2-21,3	595,164	3,1	287 ^{b.p.} ; 433, 153, 145, 127, 115
103	metilquercetina-O-hex-O-deoxihex	1 e 4	23,2-23,4	625,1768	2,3	317 ^{b.p.} ; 595, 544, 502, 394, 356, 287, 241, 162, 153,145
104	luteolina-O-deoxihex-O-pent/O-deoxihex-O-malonil	4	29,1-29,4	649,1755	4,2	287 ^{b.p.} ; 231, 213, 187, 153, 133, 127, 115, 109
105	metilkanferol-O-deoxihex-O-pent/O-deoxihex-O-malonil	4 e 5	31,5-31,7	663,1917	5,3	301 ^{b.p.} ; 531, 409, 343, 286, 258, 231, 213, 145, 133, 127 115
106	luteolina-O-hex-O-pent	4 e 5	21,6-22,0	581,1502	-1,1	287 ^{b.p.} ; 449, 329, 153, 133 115
107	apigenina-O-hex-O-pent	4 e 5	23,4-23,7	565,155	-2,5	271 ^{b.p.} ; 433, 313, 153, 133, 115
108	apigenina-O-deoxihex-O-pent/O/deoxihex-O-malonil	4	31,0-31,2	633,1797	3,0	271 ^{b.p.} ; 501, 434, 231, 213, 145, 133, 127, 115
109	crisoeriol-O-hex-O-pent/O-hex-O-malonil	4 e 5	26,1-26,5	681,1665	4,9	301 ^{b.p.} ; 595, 549, 463, 343, 286, 258, 231, 213, 187, 159, 145, 133, 129, 127,

Foram encontrados mono e di-O-glicosídeos derivados de pentoses, deoxihexoses e hexoses assim como grupos acil, malonil, acetoacetil e acetomalonil ligados a estes açúcares. Dentre todas as espécies, *S. americanum* foi a única a apresentar pentoses e derivados acilados e malonados para seus metabólitos. A

espécie *S. excelsum* apresentou derivados mono e diglicosilados de hexose e um glicosídeo ligado a um ácido cafeíco. No caso da *S. swartzianum*, foram encontrados derivados de mono e diglicosilados de hexose e deoxihexose, tabela 9. Adicionalmente, a maioria dos derivados acil e malonil foram encontrados somente nas folhas de *S. americanum* enquanto o mono e diglicosídeos para os ramos.

Segundo a hipótese de Steinharter e colaboradores (1986), os resultados da figura 19 sugerem que as espécies *S. swartzianum* e *S. excelsum* possuem uma aproximação filogenética aos subgrupos basais como *Petota* e *Potate* enquanto a espécie *S. americanum* mostra-se ligada a subgrupos geneticamente evoluídos como *Leptostemonum* e *Solanum* (EICH, 2008).

Figura 19. Mapa da rede molecular associada aos flavonoides *O*-glicosilados presentes nas folhas e ramos de *S. swartzianum* (R271, R272 e F052), *S. americanum* (M951 e M952) e *S. excelsum* (F55). A força de aresta representa o grau de similaridade entre dois nodos. Tons azuis apresentam o nível de similaridade entre 95 – 99 %, vermelhos entre 90-95% e preto entre 70-90%. A cor da legenda (dentro dos nodos) revela onde foi encontrado o íon (folhas ou ramos). O formato do nodos descreve o tipo de aglicona (flavona; flavonol e seus derivados); A cor do nodo descreve a espécie onde foi encontrada o íon molecular. O tamanho do nodo está associado ao número de compostos por nodo e a cor de borda é relativo ao tipo e número de glicosídeos presentes para uma dada aglicona.



Flavonoides C-glicosilados

Além dos flavonols O-glicosilados foram relatados C-glicosídeos em *Solanum* derivados de apigenina e luteolina, como a vitexina e orientina (D'ARCY et al., 1979). Foram atribuídos a classe de flavonoides C-glicosilados dois grupos formados por 6 e 5 nodos, figuras 20 e 21, respectivamente e sumarizados na tabela 10. Os compostos derivados de apigenina (m/z 271), luteolina (m/z 281) e crisoeriol (m/z 301) foram encontrados nas espécies *S. americanum* e *S. swartzianum*.

Na figura 20 é possível observar a presença isotopômeros de m/z 433, 449 e 483. Devido as características do padrão de fragmentação em flavonoides C-glicosilados foi possível diferenciar as glicosilações em C-6 e C-8, figura 22 (isovitexina e vitexina). A figura 22 compara os valores experimentais de vitexina e isovitexina com dados obtidos dos trabalho de Waridel et al. (2001). A tabela 10 apresenta todos os flavonoides C-glicosilados detectados para os dois nodos

Tabela 10. Flavonoides C-glicosilados encontrados nas folhas e ramos de *S. americanum* (M951 e M952). A numeração vermelha refere-se aos picos cromatográficos encontrados na figura 17. São apresentados os tempos de retenção (tR), a massa monoisotópica protonada (em azul - $[M+H]^+$) e erro em ppm do valor medido comparado a formula molecular do composto (em vermelho). *(*S. swartzianum* - R271).

Nº.	Nome	tR (min)	$[M+H]^+$	erro (ppm)
110	scoparina	22,1-22,3	463,1223	-2,1
111	isoscoparina	22,3-22,5	463,1229	-1,6
112	vitexina	20,7-20,9	433,1136	-3,6
113	isovitexina	21,0-21,3	433,1133	-4,1
114	orientina	19,1-19,3	449,1080	-4,2
115	homorientina	19,2-19,4	449,1088	-3,4
116	schaftosideo	18,9-19,1	565,1550	-6,7
117	isoschaftosideo	19,5-19,6	565,1560	-3,7
118	lucenina	18,1-18,4	581,1502	-8,0
119	lucenina I	18,7-19,0	581,1510	-6,0
120*	Vicenina II	19,8-20,0	595,1652	-0,9

Figura 20. Mapa de rede molecular associada aos flavonoides C-glicosilados presentes em *S. americanum* (M951 e M952). A força de aresta representa o grau de similaridade entre dois nodos. O formato do nodos descreve o tipo de aglicona (apigenina – 271 *m/z*; luteolina – 287 *m/z* e crisoeriol – 301 *m/z*). A cor de borda descreve a posição em que o glicosídeo está ligado a aglicona (8-C ou 6-C).

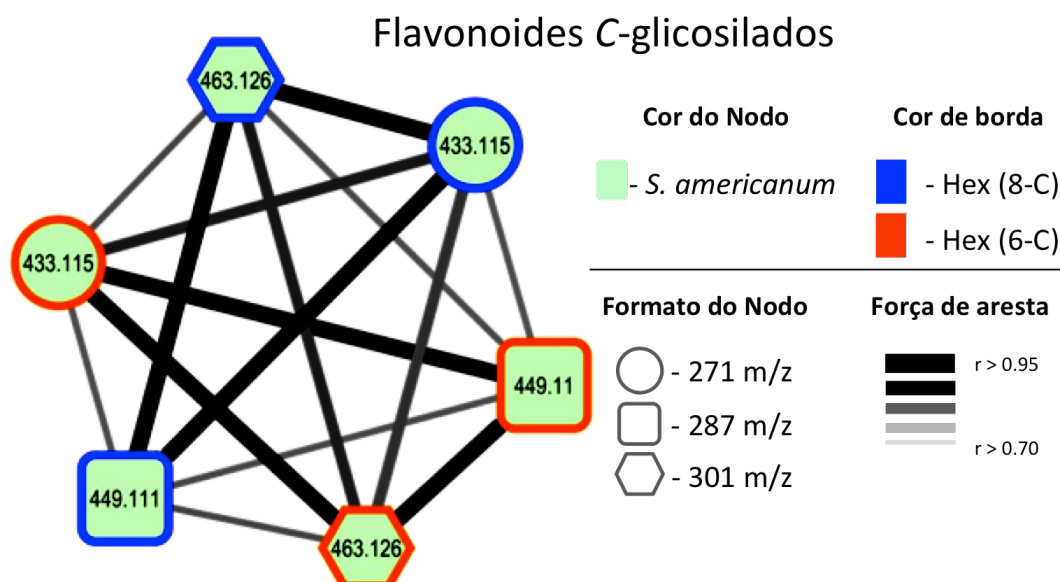


Figura 21. Mapa da rede molecular associada aos flavonoides C-glicosilados presentes nas folhas e ramos de *S. americanum* (M951 e M952) e *S. swartzianum* (R271). A força de aresta representa o grau de similaridade entre os nodos. O formato do nodos descreve o tipo de aglicona (apigenina – 271 *m/z*; luteolina – 287 *m/z* e crisoeriol – 301 *m/z*). A cor de borda descreve o tipo de glicosídeo (pentose ou hexose).

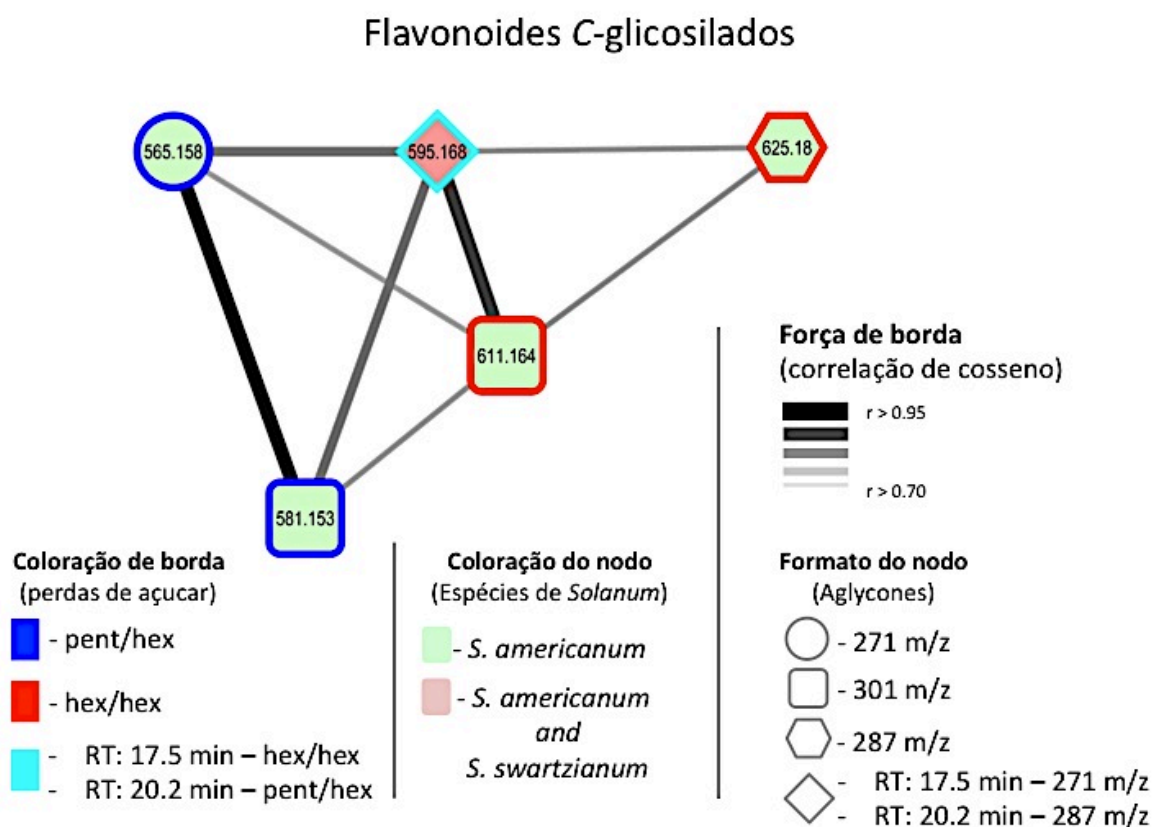
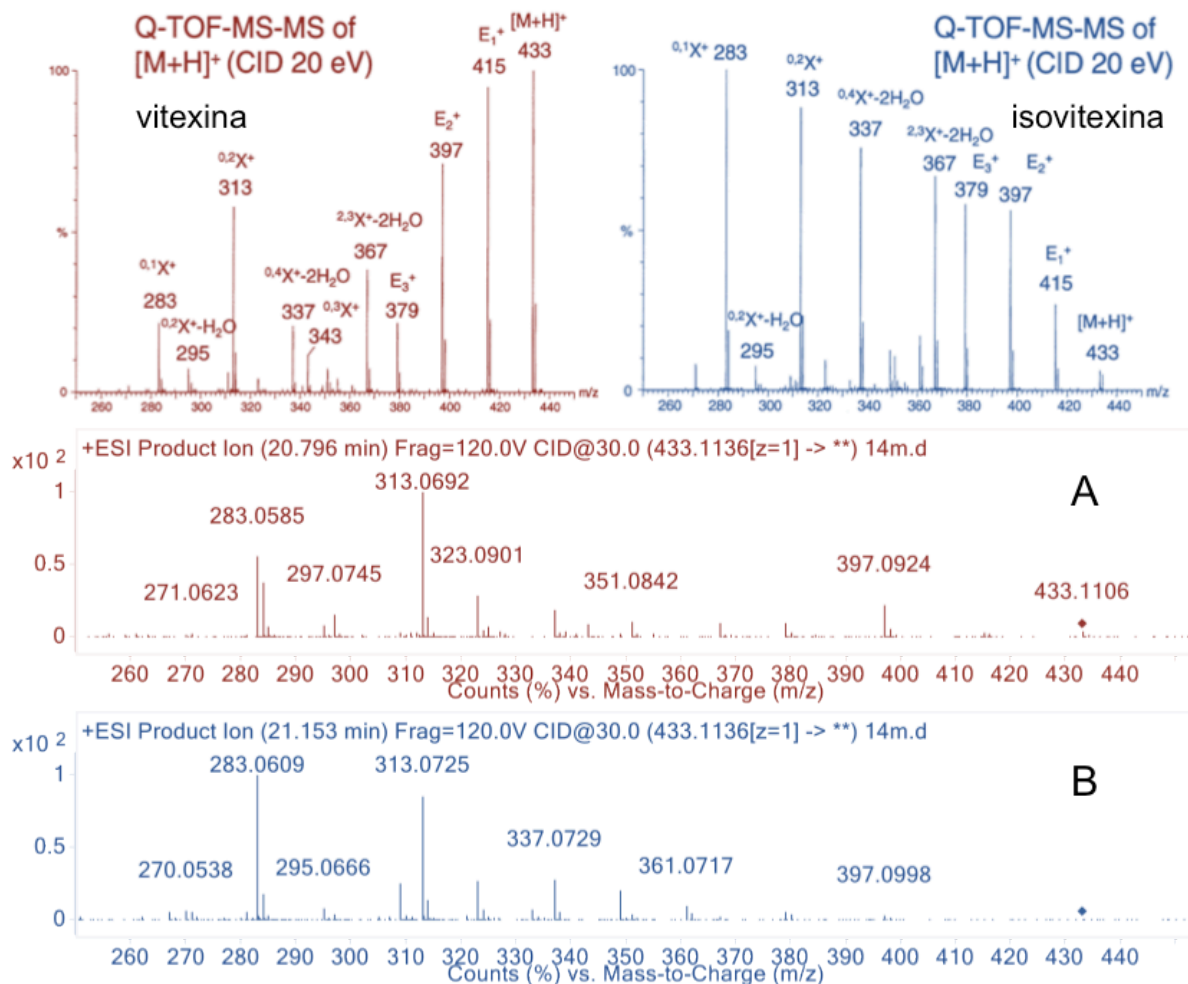


Figura 22. Comparação espectral entre valores experimentais e dados da literatura para isovitexina e vitexina. Perfis de fragmentação dos isômeros m/z 433 a 20eV, vitexina (superior esquerdo - vermelho) e isovitexina (superior direito - azul) obtidos do trabalho de Waridel et. al. 2001. (A) perfil espectral obtido a 30 eV para amostra *S. americanum* (folhas) relativo a vitexina e (B) perfil espectral da isovitexina obtido 30 eV para a mesma amostra.



Para a vitexina (figura 22 A), o íon molecular (m/z 433) e os sinais referentes a perdas sucessivas de água são mais pronunciados quando comparado a isovitexina (figura 22 B). Adicionalmente, o sinal m/z 313 (quebra $^{0,2}X^+$ do glicosídeo) é mais intenso que o íon m/z 283 (quebra $^{0,1}X^+$ da glicona) para vitexina, enquanto que na isovitexina observa-se o contrário (WARIDEL et al. 2001).

No caso da figura 21, o nodo referente ao valor m/z 595 representa duas substâncias, uma apigenina diglicosilada (duas hexoses) e luteolina diglicosilada (pentose-hexose). A determinação das substâncias foi baseada nos valores da aglicona assim como nos tempos de retenção dos compostos.

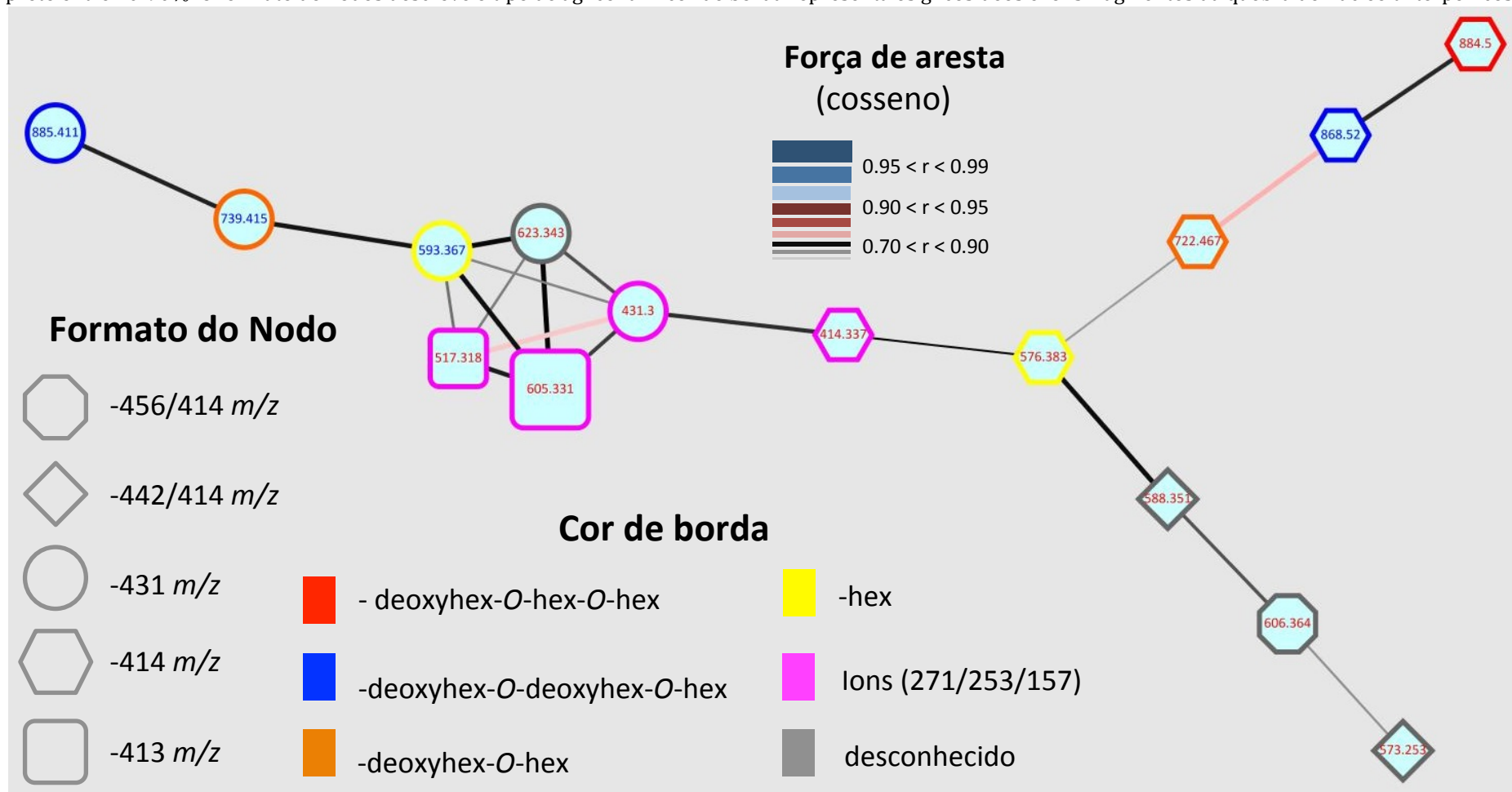
Glicoalcaloides esteroidais / Saponinas

Um grupo formado por 15 nodos foi atribuído aos glicoalcaloides esteroidais e derivados de saponinas, figura 23. Para metade dos compostos foi encontrado o íon m/z 414, associado ao esqueleto espirossolano, enquanto a outra metade foi caracterizada pelos íons m/z 413 e 431, associados a saponinas esteroidais. Adicionalmente, foram encontrados os íons m/z 273, 251 e 157, atribuídos a clivagem e perda da cadeia lateral dos anéis E e F, esquema 6 e 7. Os íons precursores, tempo de retenção e os íons referentes as agliconas estão representados na tabela 11.

Tabela 11. Glicoalcaloides e saponinas esteroidais encontradas nas folhas e ramos de *S. swartzianum* (R271, R272 e F052). A numeração em azul refere-se aos picos cromatográficos encontrados na figura 17. São apresentados os tempos de retenção (tR), a massa monoisotópica protonada (em azul - $[M+H]^+$) e erro em ppm do valor medido comparado a formula molecular do composto (em vermelho).

Nº.	Nome	tR (min)	$[M+H]^+$	erro (ppm)	Y ⁺
74	12-hidroxisolamargina	27,5-27,8	884,5045	-2,1	414
77	leptinina I	28,5-29,0	868,5025	-1,6	414
78	β-solamargina	29,6-30,0	722,4439	-3,6	414
121	tomatidenol-O-hex	29,2-29,8	576,3890	-4,1	414
122	tomatidenol	34,0-34,2	414,3373	-4,2	-
123	nuategenina-3-O-hex-O-hex-O-deoxi	32,6-32,9	885,4832	-3,4	431
124	nuategenina-3-O-hex-deoxi	34,9-35,3	739,4277	-6,7	431
125	nuategenina-3-O-hex	35,5-35,7	593,3680	-8,0	431
126	nuategenina	39,1-39,4	431,3147	-7,1	-

Figura 23. Mapa da rede molecular associada aos glicoalcaloides e saponinas esteroidais presente nas folhas e ramos de *S. swartzianum* (R271, R272 e F052). A força de aresta representa o grau de similaridade entre os nodos. Em azul estão os nodos com correlação acima de 95% de similaridade, em vermelho entre 90-95% e em preto entre 70-90%. O formato do nodos descreve o tipo de aglicona. A cor de borda representa os glicosídeos e íons fragmentos da quebra do núcleo triterpênicos.



CONCLUSÃO

Este capítulo introduziu conceitos e aplicações de ferramentas quimiométricas relativas a desreplicação de metabólitos em espécies de *Solanum* utilizando técnicas do estado da arte em espectrometria de massas.

O uso de algoritmos matemáticos para conversão de sistemas cromatográficos auxiliou na padronização das condições instrumentais (cromatografia líquida e a gás) da qual possibilitou a comparação dos resultados experimentais com trabalhos disponíveis na literatura. Isso não somente acelerou, mas trouxe uma maior robustez no desenvolvimento das propostas de desreplicação (Hipótese A).

Comparando os métodos de separação, a cromatografia a gás acoplada a espectrometria de massas limitou-se, em sua maioria, na desreplicação de compostos provenientes do metabolismo primário. Contudo, é importante salientar que esta restrição não está relacionada a ineficácia da CG-EM, mas ao pequeno número de compostos do metabolismo secundário presentes nas bases de dados. Assim, a inserção de mais compostos sililados relativos ao metabolismo secundário permitirá a desreplicação cada vez maior de metabólitos do metabolismo secundário via CG-EM (Hipótese A e B).

Adicionalmente foram gerados perfis metabolômicos de espécies de *Solanum* a partir da CLAE-EM/EM e CLAE-autoEM/EM. Os perfis de fragmentação via ESI geraram fragmentos característicos para diversas classes de substâncias como flavonoides O- e C-glicosilados, poliaminas, saponinas e glicoalcaloides. Outro aspecto destacado nesse capítulo é a falta de dados da literatura relativos ao perfil de fragmentação dos quimiomarcadores desse gênero, os glicoalcaloides esteroides. Assim, foram revisados e propostos novos mecanismos de fragmentação para a classe dos espirosolanos e solanidanos (Hipótese A e B).

No entanto, o elemento mais inovador presente no capítulo refere-se a aplicação dos mapas de rede de interação molecular via CLAE-AutoEM/EM. Os avanços das ferramentas de análise na espectrometria de massas tem permitido a realização de experimentos EM e EM/EM de modo automatizado, priorizando os íons majoritários. Esse aprimoramento juntamente com desenvolvimento de ferramentas quimiométricas tem permitido a análise dos dados de EM/EM através de

redes moleculares ou agrupamentos moleculares (Hipótese D).

Isso alterou os métodos convencionais de desreplicação baseados puramente na comparação de dados experimentais com bases de dados, por uma interface dinâmica de navegação e organização de dados através de mapas moleculares. Isto permitiu congrega análogos e classes semelhantes mesmo antes de qualquer etapa de elucidação estrutural. Assim, é possível tem uma ideia da relação entre sinais presentes (compostos) dentro de um perfil metabolômico (Hipótese D).

Adicionalmente, a elucidação de uma substância, dentro de uma rede, permite a descoberta parcial ou total das demais substância pertencentes ao grupo. Portanto, é possível não somente desreplicar mas elucidar substâncias inéditas a partir de substâncias previamente conhecidas (Hipótese D).

**CAPÍTULO 3 – USO DA CROMATOGRAFIA ACOPLADA A
RESSONÂNCIA MAGNÉTICA NUCLEAR NO ESTUDO
METABOLÔMICO EM PLANTAS: MÉTODOS DE DECONVOLUÇÃO,
DESREPLICAÇÃO E ELUCIDAÇÃO ESTRUTURAL**



INTRODUÇÃO

A ressonância magnética nuclear (RMN) é uma técnica espectroscópica da qual propriedades intrínsecas do núcleo atômico - momento magnético nuclear – são utilizadas para determinação de estruturas e/ou arranjos moleculares (LEVITT, 2007, SILVERSTEIN, 2005 ;NAKANISH, 1990).

Em geral, um experimento de ressonância magnética nuclear pode ser entendido através de 4 etapas, figura 24. A primeira etapa (1) consiste na escolha de isótopos sensíveis ao efeito de aplicação de um campo magnético (externo). Somente os núcleos que possuem spins nucleares diferentes de zero são sensíveis ao efeito de campo magnético. Dentre os principais núcleos utilizados no estudo de moléculas biorgânicas podemos destacar os núcleos ^1H , ^{13}C , ^{15}N , ^{17}O e ^{31}P (KIM; CHOI; VERPOORTE, 2010; LEVITT, 2007; VERPOORTE et al., 2010).

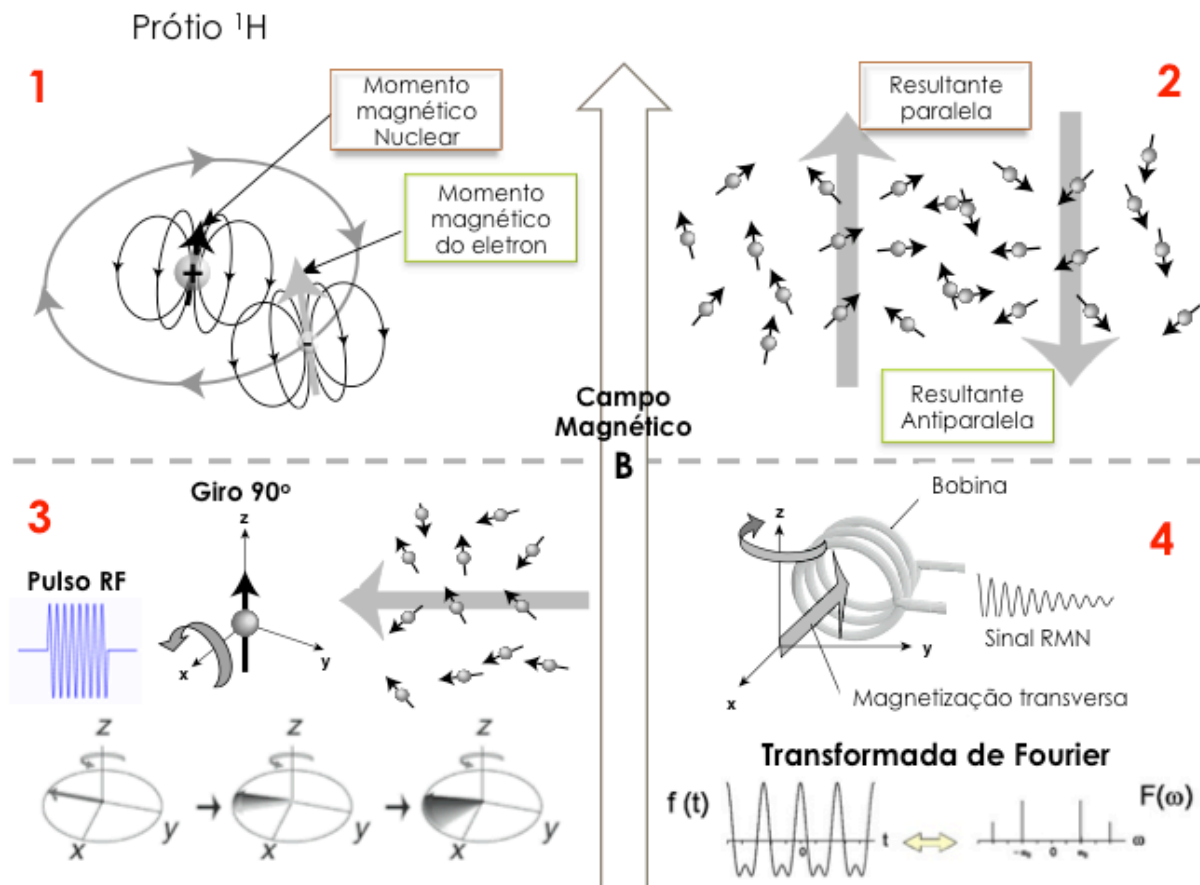
A segunda etapa (2) consiste na geração de uma resultante de campo magnético nuclear. Quando átomos são submetidos a aplicação de campo magnético externo, estes geram resultantes na direção paralela e antiparalela de forma isotrópica. No entanto, devido a diferença energética entre os dois níveis, ocorre ao longo do tempo a quebra equilíbrio magnético favorecendo a resultante paralela ao campo externo.

Uma vez alcançada tal resultante magnética segue a terceira etapa (3) na qual consiste na medição do sinal de origem magnética do núcleo. De fato, mesmo no exemplo mais simples como é o núcleo ^1H , os efeitos do campo magnético externo e momento magnético dos elétrons (rotacional e de spin) são maiores na ordem 10^4 em relação ao núcleo, impossibilitando a aquisição de informações no eixo longitudinal (campo externo)(LEVITT, 2007, SILVERSTEIN, 2005).

Deste modo, são irradiados os núcleos atômicos com pulsos específicos de radiofrequência (giro de 90° no eixo x) visando uma resultante magnética perpendicular ao campo externo (eixo Z - efeito de ressonar). Assim, elimina-se as contribuições magnéticas dos elétrons e do campo externo para obtenção de informações estruturais relacionada ao spin nuclear. A última etapa (4) consiste na conversão dos valores de onda (decaimento de indução livre – FID) para sinais discretos sob domínio de frequência (Hz). A operação matemática chamada transformada de Fourier consiste na multiplicação de FID por equações de onda

(cada equação é calculada em uma janela de frequências típicas do núcleo utilizado) e, em seguida as ondas resultantes são integradas gerando um sinal específico nos gráficos de deslocamento químico. Só a partir daí é que são realizados os estudos analíticos dos compostos, Figura 24, (LEVITT, 2007).

Figura 24. Ilustração do fenômeno de ressonância magnética nuclear para o núcleo de prótio (adaptado de Levitt M. H. **Spin dynamics**, 2007). (1) O átomo de hidrogênio apresenta dois tipos momento magnético: (i) referente a propriedade da matéria intrínseca aos prótons e elétrons (spin nucleares) e ao movimento orbital dos elétrons em torno do núcleo (geração de campo magnético a partir do movimento de carga). (2) Resultante da aplicação de um campo magnético externo (B) no vetor de momento magnético nuclear (paralela e antiparalela). Devido a pequenas flutuações de campo causado pela forte movimentação atômica (efeito de equilíbrio térmico) a resultante paralela possui um ligeiro excesso populacional resultando num vetor magnético em direção ao eixo z . (3) fenômeno de ressonância através da aplicação de pulso de radiofrequência específico para os núcleos. Através dessa excitação é possível girar a resultante magnética do núcleo (eixo z) em 90° permitindo sua detecção sem a contribuição dos efeitos magnéticos dos elétrons. (4) Aquisição dos dados de ressonância a partir de bobinas apontadas ao longo do eixo perpendicular a campo externo (B). Os dados são adquiridos na forma de ondas na qual são transformadas por funções matemáticas (transformada de Fourier) em gráficos em função de frequência.



Na área da química de PN, a RMN tem sido empregada na elucidação estrutural de compostos bioativos assim como em estudos de biossíntese. Atualmente a RMN é considerada como um dos pilares da tríade que envolve as técnicas analíticas usadas em metabolômica: cromatografia – EM – RMN (VERPOORTE et al., 2007, 2010). Ela não somente está envolvida na elucidação

mas é determinante para detecção de biomarcadores, variações de fluxo metabólico entre outras aplicações (RATCLIFFE; SHACHAR-HILL, 2005; KIM; CHOI; VERPOORTE, 2010; YANG, Z. et al., 2014; LARIVE et al., 2015; VERPOORTE et al., 2007 e 2010).

De modo geral, o núcleo de hidrogênio (^1H) é o mais utilizado em metabolômica devido a alta receptividade (99,9816-99.9974%) e a presença deste núcleo em quase todas as biomoléculas (LARIVE et al., 2015; SILVERSTEIN, 2005).

No entanto, as informações relativas aos espectros de ^1H muitas vezes são insuficientes para a determinação total de uma estrutura molecular e experimentos mais sofisticados como bidimensionais tornam-se necessários. ^1H - ^1H COSY (CORrelated SpectroscopY), ^1H - ^1H TOCSY (TOTAL Correlation SpectroscopY) e NOESY (Nuclear Overhauser Enhancement SpectroscopY) são experimentos homonucleares que revelam informações quanto a posição dos prótons numa molécula (LARIVE et al., 2015; KIM; CHOI; VERPOORTE, 2010; HEYMAN et al., 2012, VERPOORTE et al., 2010; MAHROUS; FARAG, 2015).

Já os experimentos ^1H - ^{13}C utilizando HSQC (Heteronuclear Single Quantum Coherence) e ^1H - ^{13}C HMBC (Heteronuclear Multiple-Bond Correlation spectroscopy) complementam as informações estruturais permitindo a observação direta entre a correlação entre hidrogênio e carbono (HSQC) e aquelas relativas a longa distância (HMBC - três ou mais ligações).

Uma importante aplicação da RMN em metabolômica é a possibilidade de análise em sistemas *in vivo*. Uma vez que a técnica não é invasiva ou destrutiva é possível monitorar, em séries de tempo, o comportamento de amostras frente a um determinado estímulo (BECKONERT et al., 2007; POWERS, 2009; DA SILVA et al., 2013). Neste caso, a análise de materiais sólidos ou semi-sólidos como tecidos ou linhagens celulares ocorre através do uso da técnica de HR-MAS (high resolution magic angle spinning). Esta técnica tem exercido um papel essencial na medicina como por exemplo, na análise de biópsias (LARIVE et al., 2015, BECKNONERT et al., 2007; TENG, 2013; VIZZOTO, 2004).

Abordagens baseadas em *fingerprints* tem sido outra opção em estudos metabolômicos via RMN. A aquisição de espectros de misturas complexas, como biofluidos ou extratos de plantas sob estudos metabolômicos (submetidos a estímulos ambientais ou induzidos) tem levado a caracterização de sinais

diagnósticos de doenças e a determinação de biomarcadores (RATCLIFFE AND SHACHAR-HILL, 2005; BECKONERT et al., 2007; HALABALAKI et al., 2014).

Ainda que a RMN seja menos sensível em termos de concentração em relação a espectrometria de massas, avanços no campo analítico, como os instrumentos de alto campo (até 1 GHz na frequência do ^1H) e as sonda criogênicas, tem aumentado acentuadamente a sensibilidade da técnica. Hoje é possível determinar estruturas na escala de miligramas ou mesmo de microgramas de metabólitos (KOVACS et al., 2005; LARIVE et al., 2015; SILVERSTEIN, 2005; LEVITT, 2007).

CLAE-SPE-RMN

Outra ferramenta que tem ajudado na elucidação de compostos minoritários é o acoplamento da CLAE com RMN. Em termos práticos, o acoplamento da CLAE com RMN ainda não ocorre de forma rotineira como em CLAE-EM mas vem se tornando uma ferramenta poderosa na identificação de metabólitos presentes em misturas complexas como ocorre nos PN. Existem diferentes configurações que podem ser usadas entre CLAE e RMN na qual incluem o acoplamento direto CLAE-RMN (sistema *on-flow* e *stop-flow*) e aqueles que incluem uma etapa de concentração CLAE-SPE-RMN (LARIVE et al., 2015; EXARCHOU et al., 2005; JOHANSEN et al., 2013; HALABALAKI et al., 2014; CASTRO et al., 2010; TANG et al., 2009).

Mais recentemente, o acoplamento *online* de CLAE-SPE-RMN tornou possível superar algumas barreiras analíticas de alguns modos prévios (*on-flow* e *stop-flow*). Nesta configuração, picos (bandas) cromatográficos são sequencialmente coletados em cartuchos de fase sólida por meio de múltiplas injeções utilizando solventes tradicionais grau CLAE. Logo que coletados, os cartuchos são submetidos a um fluxo de nitrogênio para remoção do solvente e, subsequentemente os compostos são eluídos com o auxílio de solventes deuterados (LARIVE et al., 2015; JOHANSEN et al. 2014; HALABALAKI et al. 2014; TANG et al., 2009).

Deste modo evita-se os problemas encontrados em sistemas *on-flow* como a baixa quantidade de massa que passa pela *probe* durante a aquisição e também economiza-se o uso de solventes deuterados. Além da questão de concentração, a

principal vantagem relativa ao sistema *stop-flow* é diminuição do fenômeno de fluxo encontrado no sistema CLAE-RMN. Sobretudo, a capacidade de uso de sondas criogênicas para o sistema CLAE-SPE-RMN eleva seu potencial da RMN para estudos metabolômicos (TANG et al., 2009).

Considerações sobre identificação de compostos via RMN em estudos metabolômicos.

A metabolômica como ferramenta de análise de “todos” os metabólitos de dado um organismo traz um verdadeiro desafio as pesquisas em RMN. Os obstáculos estão desde o processo de amostragem e extração até nas questões analíticas como detecção e análise (MUSHTAQ et al., 2014; NICHOLSON; LINDON, 2008).

Diferentemente dos experimentos realizados em EM, que na maioria, utilizam de plataformas hifenadas compostas por uma etapa de pré-purificação (cromatografia) seguida de aquisição de dados, os experimentos em RMN não apresentam uma etapa ortogonal de informação (exceto em CLAE-SPE-RMN), isto implica no estudo de moléculas em matrizes de alta complexidade (LARIVE et al, 2015; RATCLIFFE; SHACHAR-HILL, 2005).

Tipicamente um experimento ^1H para análise de um extrato bruto de uma tecido, planta ou microrganismo pode conter centenas de sinais que podem ou não estar correlacionados entre si, ou seja, não existe qualquer informação quanto a relação entre os sinais e conseqüentemente, ao número de compostos presentes. De modo geral, o que se observa é uma imensidão de sinais sobrepostos que pouco contribuem para elucidação estrutural (WOLFENDER, 2009; HALABALAKI et al., 2014).

Avanços no campo da extração de informação em sistemas complexos tem permitido superar tais barreiras através determinação sinais marcadores de classes ou entidades químicas específicas através de ferramentas quimiométricas como a PCA ou PLS. Para tanto, pesquisadores fazendo uso de padrões autênticos ou por meio de simulações determinam sinais “marcadores” em moléculas que

posteriormente servem de base na busca de metabólitos em matrizes reais (YANG et al., 2013; WOLFENDER et al., 2013; TENG, 2013).

No entanto, mesmo essa abordagem apresenta algumas falhas. Dependendo da estrutura os sinais gerados podem não ser determinantes para identificação da classe ou ainda dependendo da complexidade da matriz, tal sinal pode estar sobreposto por outros sinais, o que inviabiliza o uso dessa abordagem. Outro aspecto negativo é que ela faz uso de um único sinal como meio de identificação. Conhecendo a diversidade química de compostos orgânicos, a possibilidade de gerar falsos-positivos é muito grande por essa abordagem. Um típico exemplo são os sinais de hidrogênios metilênicos ou de glicosídeos os quais estão presentes numa infinidade de estruturas.

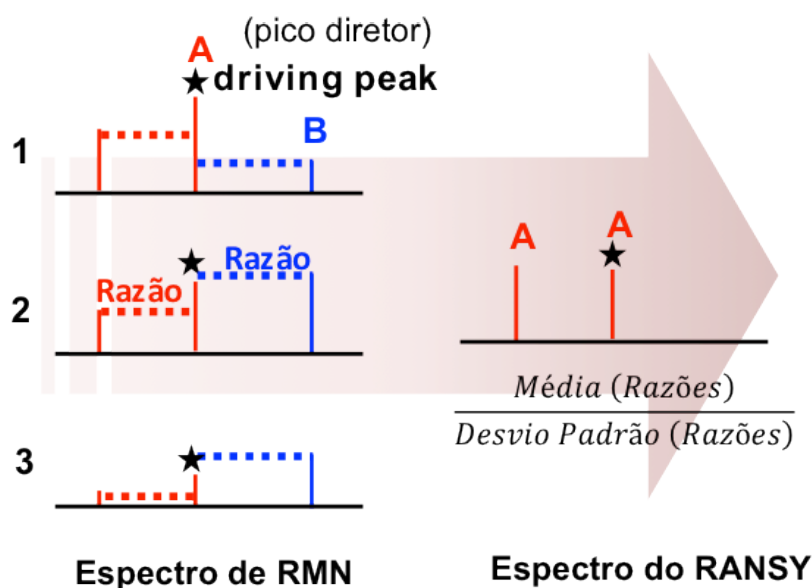
Como alternativa aos métodos univariados de identificação de sinais marcadores este trabalho implementou novos métodos de identificação estrutural baseados no uso ferramentas quimiométricas relativas a deconvolução de sinais em RMN. Dentre a ferramentas podemos destacar os algoritmos computacionais RANSY e STOCYSY e as novas de sequencia de pulsos em DOSY.

Ratio Analysis Nuclear Magnetic Resonance Spectroscopy (RANSY)

O RANSY foi desenvolvido por Gu et al. (2013) como ferramenta de auxílio na identificação de substâncias a partir de dados espectroscópicos ou espectrométricos. Em RMN, o princípio da ferramenta é baseado no fato que a razão entre dois sinais distintos (intensidades ou áreas) quando comparados em diferentes amostras devem se manter relativamente constantes se pertencerem a um mesmo composto (mesmo sistema de spin). Assim, o desvio padrão destas razões frente a diferentes amostras em teoria deveriam ser zero ou muito perto de zero caso fossem do mesmo sistema de spin. Por outro lado, se dois sinais (picos) são originados de metabólitos distintos, o correspondente desvio padrão irá ser grande, exceto em casos raros onde os metabólitos de interesse estão altamente correlacionados, figura 25.

Assim, é possível extrair informações relativas a estrutura de um único composto a partir da comparação espectral de amostras ligeiramente distintas.

Figura 25. ilustração do processo de execução do RANSY. Uma vez selecionado *driving peak** (sinal de deslocamento químico específico para o composto A) ocorre a divisão entre valores de intensidade dos demais sinais de RMN com o valor de intensidade do *driving peak*. Esse processo (RANSY) se repete para cada espectro de ressonância, a partir de amostras submetidas a diferentes condições. Os valores de razão entre os sinais relativos ao composto A permanecem relativamente constantes (pertencem ao mesmo sistema de spin) e consequentemente, o desvio padrão da média das razões é relativamente baixo. Em contrapartida, para as razões referentes ao composto B não são constantes e o desvio padrão resultante das razões é elevado. O espectro de RANSY é o resultado da média das razões (Sinal RMN dividido pelo *driving peak*) dividido pelo seu desvio padrão. Assim, o espectro calculado RANSY destaca os sinais do composto A em detrimento do composto B, permitindo visualizar os íons características apenas para este composto.



Statistical Total Correlation Spectroscopy (STOCSY)

Assim como RANSY, o STOCSY é um método de análise baseado na multicolinearidade de sinais para a identificação de potenciais biomarcadores em estudos metabolômicos baseados em dados de RMN (CLOAREC et al., 2005). No entanto, ao invés de um simples cálculo da nível de correlação entre o sinal escolhido e os demais sinais presentes no espectro, o STOCSY permite observar tendências em variações de concentrações entre diferentes substâncias (co-variância).

De fato os resultados gerados entre as duas ferramentas são muitos próximos, no entanto, o RANSY prioriza a determinação dos sistemas de spin referente a um sinal selecionado enquanto o STOCSY apresenta de forma mista tanto as relações entre os sinais como valores de co-variância. Do ponto de vista

prático, o STOCASY gera espectros mais complexos, no entanto, ele é mais robusto quanto à reprodutibilidade e a sobreposição de sinais. No caso do RANSY a sobreposição dos sinais possui uma maior influência nos resultados (GU et al., 2013; CLOAREC et al., 2005).

Diffusion Ordered Spectroscopy (DOSY)

O DOSY é uma técnica espectroscópica utilizada para medir o coeficiente de difusão de moléculas ou agregados em solução por espectroscopia de RMN (ANTALEK, 2002; LOENING et al., 2001; VIZZOTO, 2004). Esta técnica vem sendo utilizada principalmente: (i) na separação dos sinais de ressonância de diferentes componentes em uma mistura de acordo com seus coeficientes de difusão e (ii) na determinação do tamanho de agregados moleculares.

Desde o seu desenvolvimento sua principal aplicação ocorre na análise de misturas complexas, tais como em extratos de plantas, polímeros, combustíveis, alimentos entre outros (ANTALEK, 2002, VIZZOTO, 2004). Devido ao seu grande potencial em separação, a técnica DOSY representa uma ferramenta em potencial para estudos metabolômicos (VIZZOTO, 2004).

MATERIAIS E MÉTODOS

Coleta e classificação do material vegetal

Para os experimentos em RMN foram utilizadas apenas amostras de *S. paniculatum* coletadas na estação ecológica de Itirapina (amostra M985 vide tabela 1) e espécimes *S. paniculatum* encontrados no horto de plantas medicinais localizado na faculdade de ciências farmacêuticas da UNESP de Araraquara. As plantas foram identificadas pelo Prof. Luís Vitor Sacramento e as respectivas exsicatas foram depositadas no *Herbarium Rioclarense* (HRCB) do Instituto de Biociências da UNESP/ Campus Rio Claro.

Extração e secagem

A extração do material vegetal seco e pulverizado se deu por duas metodologias: a primeira segue as mesmas rotinas encontradas no capítulo I utilizando a amostra de *S. paniculatum* coletada na estação ecológica de Itirapina. O segundo método consistiu no uso de material fresco coletado no horto da faculdade de ciências farmacêuticas de Araraquara (FCFAR – Ar), previamente submetido a imersão em nitrogênio líquido e subsequentemente extraído através de maceração em MeOH e banho ultrassônico (Dionex ASE 300 (Dionex Canada Ltd., Oakville, ON, CA). A eliminação dos solventes foi realizada com ajuda de um speedvac (THERMO, NY, EUA).

Solventes e reagentes empregados

Para a preparação das amostras de RMN foram utilizados: acetonitrila deuterada (CD_3CN) e metanol deuterado (CD_3OD) grau 99,96% de pureza (Sigma Aldrich, St. Louis, MO, USA). Para o sistema CLAE-SPE-TT foram utilizados Metanol

(MeOH), Acetonitrila (CH_3CN) grau HPLC (Mallinckrodt, Paris, Kentucky, USA) e H_2O deionizada (18m Ω ; Milli-Q, Millipore).

CLAE-SPE-RMN

As coletas seguidas de análise por RMN foram obtidas através de um HPLC Agilent Technologies 1260 Infinity, equipado com uma bomba quaternária G1311B, amostrador automático e detector de arranjo de diodos G1329B. Um sistema extração em fase sólida Prospekt 2 (Spark, Emmen, Netherlands) equipado com módulo ACE (Automated Cartridge Exchange) para coleta e trapeamento dos analitos e duas bombas Smartline 100 (KNAUER). O sistema de transferência dos compostos trapeados no Prospekt para tubos de RMN (3mm) foi efetuado através de coletor Gilson WI 53562 (USA).

Preparação de amostras em CLAE-SPE

Para processo de preparação de amostra para análise por CLAE-SPE-TT, os extratos brutos de *S. paniculatum* foram dissolvidos em MeOH:H₂O (9:1) e submetidos ao processo de *clean-up* para remoção de interferentes apolares (cartuchos de C-18 seguidos de filtros de 0,45 μm). Uma vez secos os extratos foram redissolvidos sob concentração de 20,0 mg.mL⁻¹ e injetados no sistema de CLAE-SPE.

Análises por CLAE-SPE

Para a amostra M985 a eluição foi realizada empregando uma coluna Onyx® monolítica de 2-5 μm C-18, (100 X 4,6 mm, Phenomenex, Torrance, CA). A fase móvel consistiu de água deionizada (A) e acetonitrila (B) em sistema gradiente, nas seguintes condições cromatográficas: 5-15 % B (fase orgânica) em 25 minutos, seguido de 15-50% B em 25 minutos (total 50min) e de 50-100% em 2 minutos (total 52 min) e 10 minutos adicionais em decorrência de limpeza e estabilização de

coluna para condição inicial. Sob vazão de coluna $1,5 \text{ mL} \cdot \text{min}^{-1}$ foram realizados vinte injeções de $20 \text{ } \mu\text{L}$ em amostras de $20,0 \text{ mg} \cdot \text{mL}^{-1}$. Os analitos de interesse foram trapeados em cartuchos GP (general phase; polímeros esféricos de polidivilbenzeno de tamanho de particular entre $5 - 15 \text{ } \mu\text{m}$) e eluídos com CD_3CN para análise por RMN.

Para as análises de *S. paniculatum*, coletadas no horto da farmácia FCFAR - Ar, foi utilizada uma coluna Prontosil Eurobond C18 ($5,0 \text{ } \mu\text{m}$, $125 \times 4,0 \text{ mm d.i.}$, Bischoff Chromatography, Leonberg, Germany) sob modo gradiente de eluição. A fase inicial do sistema cromatográfico consistiu de 40% B (metanol) e 60% de A (água). Foi aplicado um gradiente linear de 40-100% de B durante em 45 min seguido de 10 min em 100% de B e recondicionamento a 40% de B durante 8 minutos. Os eluentes foram monitorados a 220, 254, 280 e 330 nm. Foram realizadas 11 injeções de $35 \text{ } \mu\text{L}$ sob concentração $20 \text{ mg} \cdot \text{mL}^{-1}$ e vazão $1,3 \text{ mL} \cdot \text{min}^{-1}$. Os metabólitos eluídos da coluna foram coletados em cartuchos do tipo GP.

Análises por RMN

Todos os experimentos de RMN foram realizados em um espectrômetro Bruker Avance III de 14,1 Tesla, (600,12 MHz para frequência do hidrogênio). As amostras foram solubilizadas em $160 \text{ } \mu\text{L}$ de solvente deuterado (CD_3OD ou CD_3CN). Os deslocamentos químicos foram referenciados pelos sinais do padrão tetrametilsilano (TMS). As medidas foram realizadas em criosonda tripla inversa ^1H $\{^{13}\text{C}$ e $^{15}\text{N}\}$ de 5 mm de diâmetro interno e com bobinas geradoras de gradiente de campo na coordenada z.

Os experimentos de ^1H foram baseados na sequência de pulso lc1pnf2 – LC1D12 usando supressão de sinal da água e do solvente (MeOH ou ACN). Foram utilizados os seguintes parâmetros nos experimentos: tempo de reciclagem 2,4 s; janela espectral de -5 a 15 ppm com 32k pontos e tempo de aquisição de 1,36 segundos. O processamento de todos os experimentos (exceto se mencionado) foi realizado através de multiplicação exponencial sob um fator de 0,3 Hz (Line Broadening - LB), correção manual de fase e de linha de base.

Os experimentos de ^1H – ^1H COSY (cosygpppqf) foram realizados utilizando uma janela espectral de 0 a 13ppm, tempo de aquisição de 0,127s, tempo de

reciclagem de 2,0s e 16 *scans* para 128 incrementos. Informações relativas ao acoplamento direto ^1H e ^{13}C foram obtidos por meio de experimentos ^1H - ^{13}C HSQC (hsqcedetgpsisp2.3), empregou janela espectral de -1 a 11ppm para a dimensão F_2 (^1H) e 0 a 165 ppm para a dimensão em F_1 (^{13}C). O tempo de aquisição foi de 0,14s, tempo de reciclagem (d1) de 2s e 16 *scans* para os experimentos com incremento de 256. Foi utilizado como acoplamento médio entre ^1H e ^{13}C o valor de 145Hz (Cnst2)

Mapas de correlação a longa distância de ^1H - ^{13}C HMBC foram obtidos sob janela espectral de 0 a 11,5 ppm na dimensão de F_2 (^1H) e de 0 a 210ppm para F_1 (^{13}C). Foram aplicados ciclos de 16 *scans* para cada incremento (256), utilizando o tempo de aquisição de 0,28s, tempo de reciclagem de 2s e constante de acoplamento a longa distância (Cnst13) de 8Hz.

Os experimentos de TOCSY1D utilizaram da sequência seldigp, com janela espectral de 20 ppm, tempo de relação de 1,0s e tempo de aquisição de 2,72s. Foram realizados 11 experimentos utilizando 128 *scans*, *mixing time* (d9) de 0,12s sob irradiação nos seguintes pontos (o1p): 4,39, 4,23 4,18, 3,81, 3,48, 2,27, 1,88, 1,19, 1,12 e 1,03 ppm.

Experimentos DOSY – Princípio da técnica

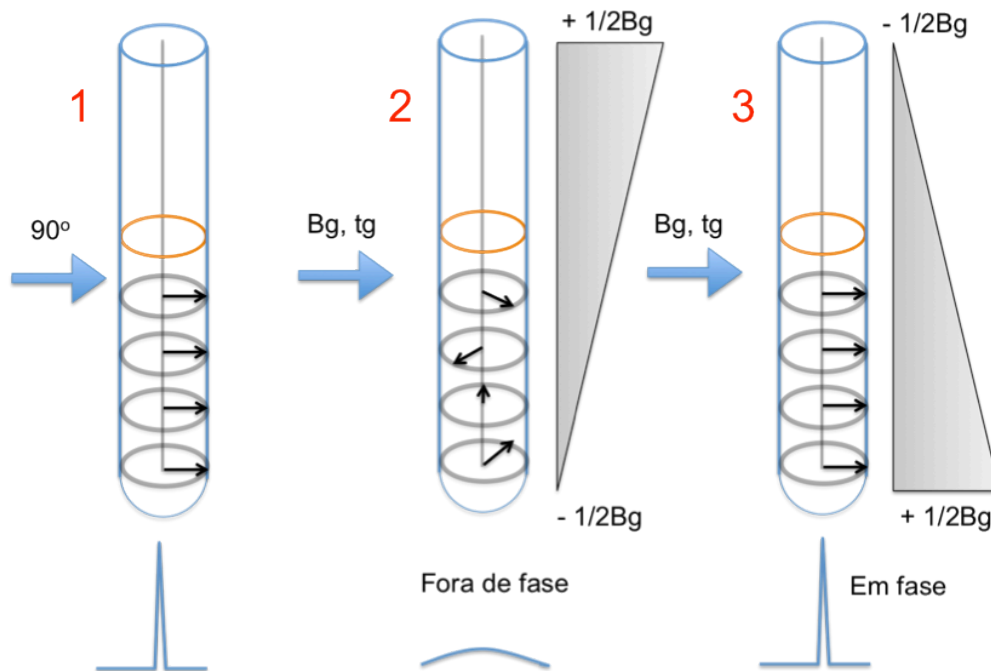
A auto-difusão é o processo de movimento translacional aleatório de matéria entre uma parte do sistema para outra como o resultado da energia térmica (movimento browniano). A medida de difusão é feita através cálculo do número de moléculas que atravessam uma seção transversal (área) por unidade de tempo, conhecida como auto-difusão, J ($\text{m}^2.\text{s}^{-1}$) (VIZZOTO, 2004). Para moléculas esse valor varia entre 10^{-10} a $10^{-10} \text{ m}^2.\text{s}^{-1}$.

Os experimentos de DOSY, usados na separação de sinais, foram possíveis através do desenvolvimento de sequências de pulso na qual aplicam gradiente de campo a partir das sondas no eixo longitudinal das amostras, os chamados pulsos de gradiente de campo (ANTALEK, 2002).

Em geral, os experimentos de DOSY são formados por uma série de experimentos que podem ser uni ou bidimensionais (^1H , ^{13}C , COSY, TOCSY entre outros) que diferem entre si pela amplitude (g) dos pulsos de gradiente aplicados,

que é aumentada gradativamente ao longo dos experimentos. O intensidade do sinal é inversamente proporcional a intensidade do campo aplicado. Assim, quando um conjunto de sinais é atenuado pelo aumento da intensidade do gradiente isso permite a identificação de quais sinais pertencem a uma dada molécula. Quanto maior a velocidade de decaimento dos sinais, maior é o coeficiente de difusão da molécula. A figura 26 ilustra o processo de DOSY. (ANTALEK, 2002; LOENING et al. 2001; VIZZOTO, 2004).

Figura 26. Ilustração de um típico experimento de DOSY. (1) nesta condição temos a resultante do vetor magnético decorrente de pulso r.f. a 90° . Todos os spins só possuem a mesma frequência de ressonância (inicialmente apontada para o eixo $-y$) devido a homogeneidade do campo magnético. (2) Gradiente de campo adicional (bobinas da sonda - B_g) gerado ao longo do eixo z (t_g – tempo de aplicação do gradiente de campo). Em (2) a frequência de ressonância de cada spin é diferente, o que resulta num sinal baixo (fora de fase). (3) Gradiente de campo inverso sob o mesmo tempo (t_g). Se não houvesse o fenômeno de difusão, teoricamente, o gradiente inverso refocaria os spins da etapa 2 resultando no mesmo sinal que na etapa em 1. No entanto, devido a difusão molecular os spins nucleares não “sentem” o mesmo gradiente inverso, pois se localizam em posições diferentes da inicial, resultando num sinal menor que a etapa 1.



A transformada inversa de Laplace deste conjunto de dados origina o espectro de DOSY; um espectro pseudo-bidimensional, onde os sinais de RMN 1D são apresentados em uma dimensão e a distribuição ordenada do coeficiente de difusão molecular (D) em outra.

Os experimentos de ^1H DOSY foram realizados a 22°C empregando-se uma criosonda de 5 mm de diâmetro interno, com detecção em modo inverso e equipadas com bobinas geradoras de gradientes de campo na coordenada z (campo

máximo de 53,5 Gauss.mm⁻¹). Os experimentos foram realizados sem auxílio de rotação. Para evitar problemas de convecção, as amostras foram preparadas em tubos de 3 mm. A sequência de pulso LEDBPGP2s foi aplicada, empregando-se uma largura espectral em ¹H de 9,9659 Hz, tempo de aquisição de 2,7s, tempo de reciclagem de 0s, duração dos pulsos de rf (90°C) de 6,72 µs e de gradiente de 600 ns. Foram acumulados 16 FIDs de ¹H, com o gradiente de campo aumentado linearmente de 2% a 95%. O processamento dos espectros de DOSY foram realizados no software TopSpin (Bruker Biospin). A escala do coeficiente de difusão está apresentada na forma logarítmica.

Análise por RANSY

A primeira etapa do cálculo consiste na seleção de pico diretor ou *driving peak* - DP (sinal da qual o usuário deseja monitorar). Em seguida, todos os outros pontos do espectro (intensidades) são divididos pela intensidade do pico selecionado *driving peak* (GU et al., 2013)

$$DP_{i,j} = \frac{X_{i,j}}{X_{i,k}} \quad (2)$$

onde i representa o i ésimo espectro de um conjunto de n espectros de RMN, j é o j ésimo ponto de um total de m pontos neste espectro. $X_{i,k}$ corresponde ao *driving peak* e DP representa a matriz $n \times m$.

Os valores de RANSY, denotado como um vetor R de m elementos, corresponde aos quocientes das médias e desvios-padrão das colunas da matriz DP. O desvio padrão é zero para o próprio *driving peak* e, portanto, seu valor RANSY é zero. Os outros valores de RANSY são calculados como elementos do vetor R como se segue:

$$R_j = \frac{\frac{1}{n} \sum_{i=1}^n DP_{i,j}}{\sqrt{\frac{1}{n} \sum_{i=0}^n (DP_{i,j} - \frac{1}{n} \sum_{i=0}^n DP_{i,j})^2}} \quad (4)$$

Uma vez que as razões dos desvios-padrão são utilizadas no denominador, um desvio-padrão pequeno irá produzir um grande valor de R , gerando um alto sinal que é relativo a presença no mesmo sistema de spin do *driving peak*, figura 25 (GU et al., 2013). Assim, para os sinais de RMN relativos ao mesmo composto, será gerado um sinal intenso de valor RANSY, enquanto sinais relacionados a outros compostos irão gerar grandes desvios-padrão, e consequentemente pequenos valores no RANSY.

Análise por STOCYSY

Os conceitos de correlação de sinais (colinearidade) do STOCYSY se assemelham as premissas encontradas no RANSY. No entanto, o STOCYSY não se restringe a simples razão de sinais multiplicadas pelo desvio padrão. No STOCYSY é realizado um cálculo de uma matriz de correlação C para um conjunto de n espectros contendo v variáveis espectrais (número de pontos) como segue a equação 5 (CLORAEC et al., 2005)

$$C = \left(\frac{1}{n} - 1\right) \left(\frac{1}{n} - 1\right) X_1^t X_2 \quad (5)$$

onde X_1^t and X_2 correspondem as matrizes transversa de $n \times v_1$ e $n \times v_2$ contendo dados de intensidade espectral. Se o conjunto de espectros são idênticos, uma matriz de autocorrelação é obtida, indicando a estrutura de correlação entre as variáveis. Quando as matrizes X_1 e X_2 descrevem diferentes espectros a matriz de correlação refletirá as correlações entre respostas (sinais) nos diferentes espectros. Assim, a matriz C descreve as correlações, intra e inter-metabólicas, em que uma dada resposta pode fornecer informações da estrutura molecular e conectividade assim como suas relações bioquímicas.

A força da correlação é dependente das propriedades inerentes de detecção da técnica. A colinearidade de sinais que surgem a partir da mesma molécula serão afetadas principalmente pela sobreposição de sinais, efeito de ruído e interferências (defeitos e artefatos) (CLORAEC et al. 2005).

Softwares

Os dados de CLAE-SPE-TT foram tratados por meio do software hystar 3.2 (Bruker, Fahrenheiststr, Bremen, Germany). Para o processamento fez-se o uso software Topspin (Bruker, Germany). Isto inclui o tratamento e aquisição de todos os dados de DOSY obtidos neste trabalho. Os algoritmos RANSY e STOCSY foram desenvolvidos na forma de script e função, respectivamente e, foram executados a partir da plataforma do software matemático MATLAB ver.7.10 (R2010a) (The MathWorks Inc., Natick, EUA). Foram gerados espectros simulados de possíveis estruturas utilizando o software NMR suíte MestreNova v. 8.1 (Mestrelab Research SL, Madri, Espanha).

RESULTADOS E DISCUSSÃO

Duas estratégias em RMN foram aplicadas para a geração de perfis metabólicos assim como na desreplicação de compostos presentes na espécie *S. paniculatum*: (i) CLAE-SPE-RMN em modo seleção de pico na qual cada pico foi coletado e identificado com base na desreplicação de metabólitos “puros” e (ii) CLAE-SPE-RMN em modo *time slice* onde frações são coletadas em intervalos iguais de tempo e analisadas através da combinação de ferramentas quimiométricas de deconvolução (RANSY e STOCYSY) e dados espectroscópicos de RMN (DOSY, TOCSY1D e aqueles regularmente utilizados em PN).

CLAE-SPE-RMN em modo seleção

A análise do metaboloma da espécie *S. paniculatum* coletado na estação ecológica de Itirapina (EEI) foi elaborado segundo protocolo desenvolvido por Gómez-Romero et al. (2010) e adaptado de acordo com as exigências instrumentais.

A amostra de *S. paniculatum* (M951) foi preparada de acordo com esquema apresentado na figura 8, (CLAE-EM). A escolha pelo modo seleção foi devida a alta capacidade de pico e resolução do sistema cromatográfico. Neste modo, substâncias puras ou parcialmente puras são repetidamente aprisionadas em cartuchos de SPE (GP) e posteriormente transferidos a tubos de 3 mm para subsequentes análises em RMN. A figura 27 representa o perfis cromatográficos obtido a 254 e 280 nm (vermelho e azul, respectivamente) da amostra de *S. paniculatum* (EEI).

Ao pico denominado 1G3 (figura 27) foram obtidos sinais de absorção máxima em UV de 260 e 352 nm correspondente à grupos fenólicos de alta conjugação, como as flavonas (conjugação entre os anéis C e B, banda I, 350nm e anel A, banda II, 255nm).

O espectro de RMN de ^1H para 1G3, figura 28, revelou a presença de um derivado de quercetina glicosilada (1G3 - Quercetina-3-O- β -diglicopiranosil).

O sinal em 12,6 ppm foi atribuído ao hidrogênio entre a hidroxila localizada na posição C-5 e a carbonila em C-4 (figura 28). O duplo dubleto a 7,47 ppm foi

atribuído ao hidrogênio H-6'. Neste caso é possível observar o acoplamento *orto* entre H-6' e H-5' e *meta* entre H-6' e H-2' (J - 8,3 e 2,6 Hz, respectivamente). O duplete em 6,93 ppm foi atribuído ao hidrogênio H-5' devido somente a um acoplamento *orto* de J - 8,3 Hz. Este padrão de substituição sugere a presença de sistema catecólico em 3' e 4' característico de derivados de quercetina. Dubletos em 6,44 e 6,23 ppm de acoplamento de J - 2,1Hz associados a correlação *meta* entre os hidrogênios H-6 e H-8 típico de flavonas diidroxiladas (C-5 e C-7) reforçam a hipótese da quercetina.

A ausência de um singleto entre 6 e 7 ppm indica uma substituição na posição C-3 da flavona. Segundo Markham e Gieger (1993), dubletos como em 5,22 e 4,66 ppm com constantes de acoplamento de aproximadamente J - 7,8 Hz indicam diglicosilações do tipo *O*- β na posição C-3 de flavonas (MARKHAM; GEIGER, 1993). Espectros simulados de ^1H (figura 29) foram gerados pela plataforma MestreNova afim de avaliar os possíveis sítios de glicosilação entre os açúcares. A comparação dos dados espectrais reais e simulados sugere uma glicosilação em C-2", como observado na estrutura da figura 28.

Para a substância 1G5 (figura 30) foram encontrados os mesmos os sinais em RMN e UV previamente descritos para a molécula de 1G3 (figuras 27 e 28). Além dos sistemas insaturados e aromáticos típicos das flavonas, nota-se a presença de diidroxilação no anel A (C-5, C-7) e as glicosilações na posição C-3 da flavona. A única diferença observada ocorre na porção do anel B, que ao invés de um sistema catecólico observa-se um anel *para* substituído. Esta diferença é caracterizada por dois dubletos 7,88 e 6,73 ppm (J - 8,8Hz) referentes ao acoplamento *orto* entre o H-2' e H-3' e H-6' e H-5' caracterizando o composto kaempferol-3-*O*- β -diglicopiranosil.

Figura 27. Perfil cromatográfico do extrato de folhas de *S. paniculatum* (M985) obtidos a 254 e 280 nm (vermelho e azul respectivamente). As linhas verticais (pontilhadas) referem-se ao intervalo de coleta dos picos no sistema CLAE-SPE. As siglas entre parênteses descrevem em quais cartuchos de SPE foram coletados os compostos. A critério do trabalho, estas siglas foram utilizadas para identificar os compostos. As caixas projetadas a partir do ápice de cada banda cromatográfica representa espectro de UV (220-400 nm) para o composto coletado.

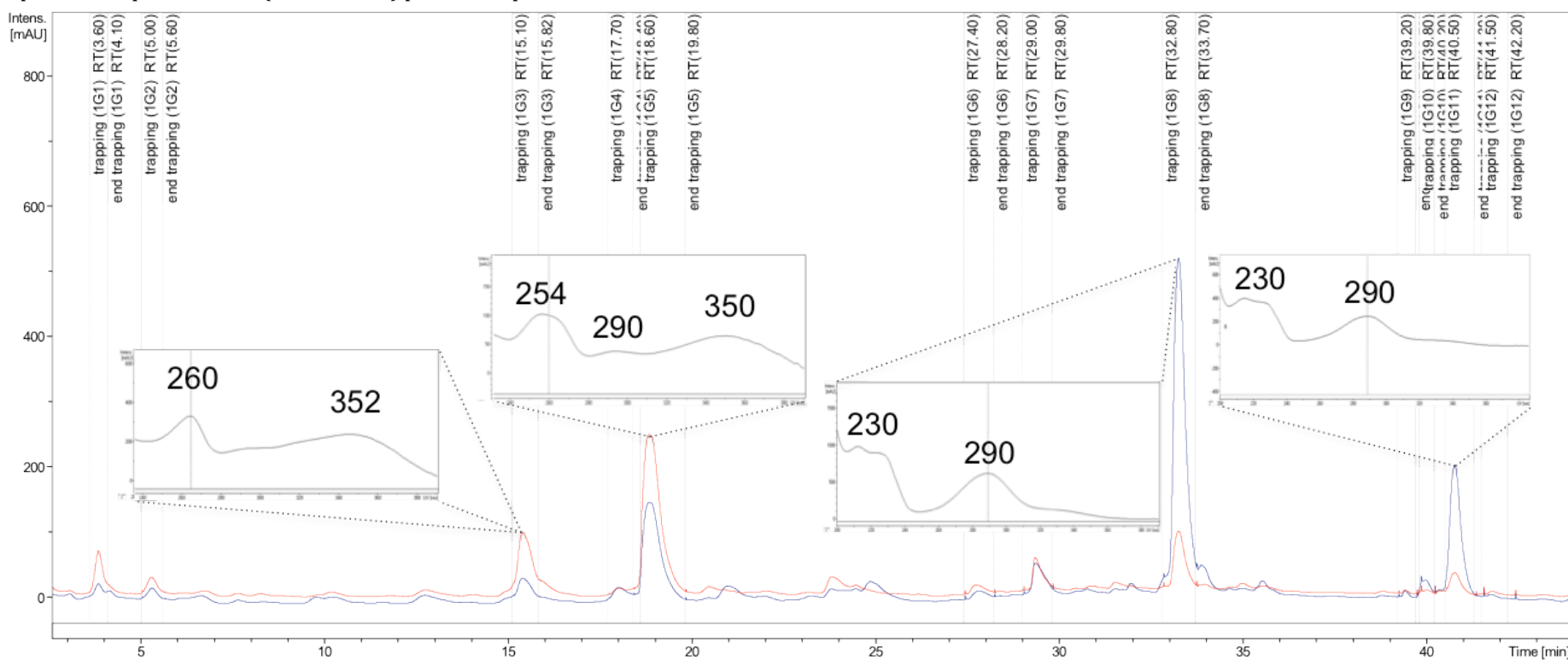


Figura 28. Espectro de ^1H para o composto presente no cartucho 1G3 - Quercetina-3- O - β -[glc, (1-2) 2''- O - β -glc. As ampliações entre 4,5 – 5,4 ppm revelam a presença de dois glicosídeos – dubletos (8Hz), enquanto a ampliação entre 6,2 e 7,8 mostram o padrão de substituição típico da quercetina.

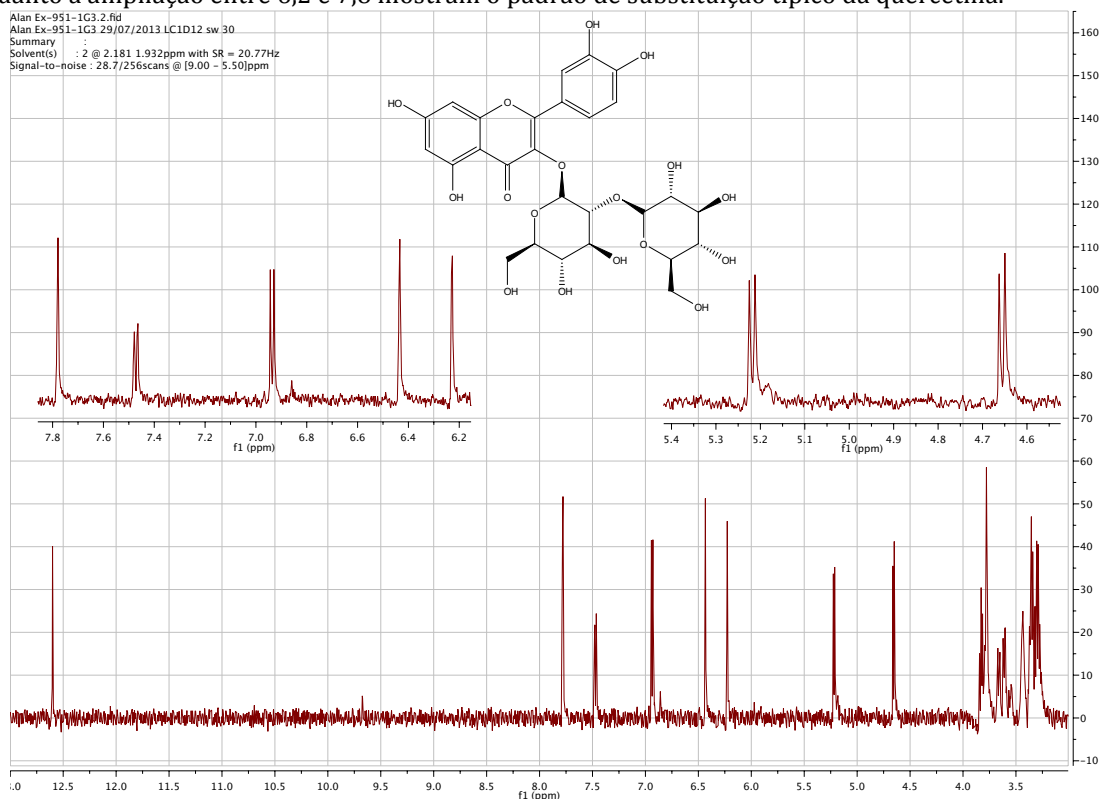
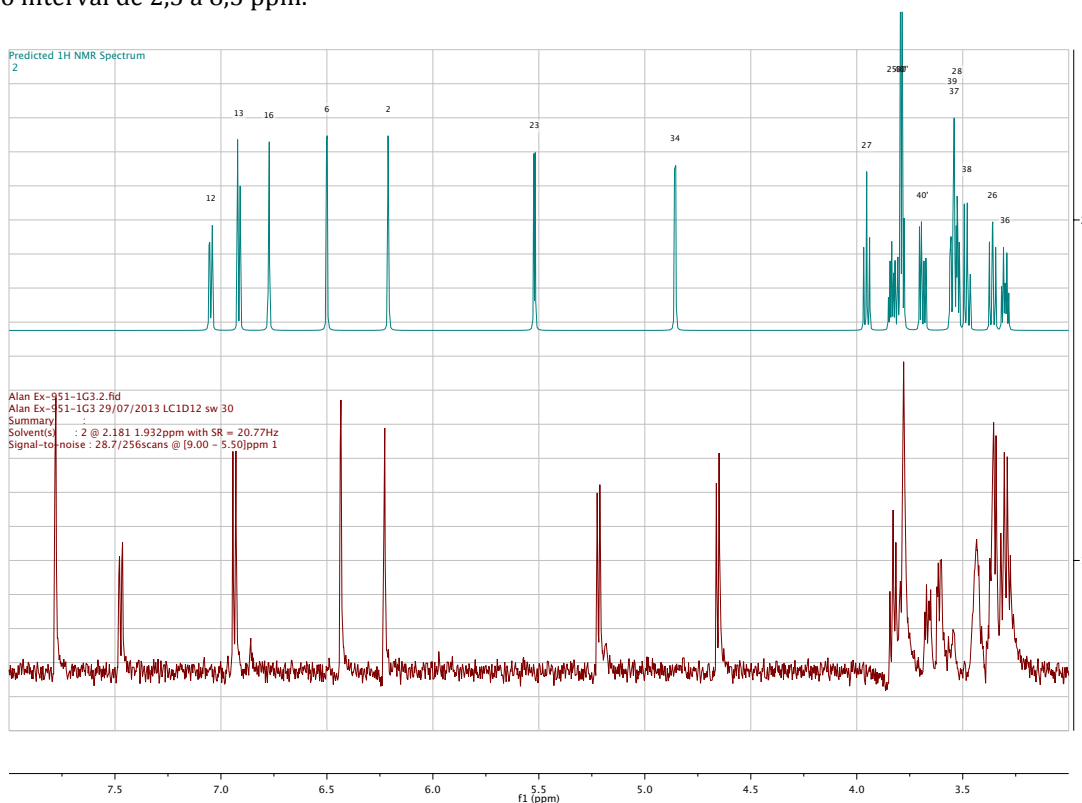


Figura 29. Comparação espectral entre os dados experimentais de ^1H (em vermelho) e dados simulados (MestreNova – em azul) para o composto presente no cartucho 1G3 - Quercetina-3- O - β -[glc, (1-2) - O - β -glc sob o intervalo de 2,5 a 8,5 ppm.



Experimentos em HSQC para a amostra 1G5 (figura 31) corrobora a elucidação estrutural do derivado de kaempferol. De acordo com Markham e Geiger (1993) o carbono anomérico (C-1") do kaempferol-3-O- β glc é encontrado em torno de 100-101 ppm, enquanto em sistemas do tipo 3-O- β -glc-[(1-2)-glc] são encontrados em torno de 103-104 ppm como observado no espectro da figura 31. Os sinais entre 79-82 ppm foram característicos das substituições em C-2", uma vez que os monoglicosídeos apresentam tal sinal entre 74-75 ppm. Os demais sinais estão de acordo com as propostas de elucidação dos derivados de kaempferol (MARKHAM; GEIGER, 1993).

No caso das substâncias 1G8 e 1G11, o espectro de UV revela bandas características de flavanonas (Banda II, 290nm). O experimento de ^1H para 1G8 (figura 32) indica uma flavanona sem a presença de glucosídeos. O deslocamento químico a 12,15 ppm refere-se ao hidrogênio quelado entre a hidroxila e carbonila em C-4 e C-5. Um sistema *para* substituído é observado a partir dos sinais 7,35 e 6,85 ppm (J - 8,7Hz) referentes ao acoplamento *orto* entre H-2' e H-3' e H-5' e H-6' (figura 32). Um duplete de J - 2,5Hz localizado em 5,96 ppm integrando para dois hidrogênios (2H) indica o mesmo deslocamento químico de ^1H para C-6 e C-8.

Contudo, são os sinais em 5,43, 3,18 e 2,75 ppm que marcam a presença de uma saturação entre C-2 e C-3 típico de dihidroflavonas. Os desdobramentos observados no duplo duplete de J - 17,2 e 3,0Hz para o sinal em 2,75 ppm estão associados ao acoplamento geminal entre os hidrogênios em C-3 e entre os hidrogênios em H-3 e H-2 do tipo pseudo equatorial-axial. Já o duplo duplete em 3,18 ppm (J - 17,2 e 13,0 Hz) refere-se ao mesmo acoplamento geminal em C-3 (axial-equatorial) e ao acoplamento pseudo axial-axial entre H-2 e H-3 (J - 13 Hz).

O sinal em 5,43 ppm (J - 13,0 e 3,0 Hz) está relacionado ao H-2 com outros dois hidrogênios em C-3, caracterizando a naringenina, figura 32. Como na elucidação da molécula em 1G3, foram realizados estudos de simulação *in silico* para a naringenina o qual revelou sinais muito próximos aos valores experimentais (figura 33).

Figura 30. Espectro de ^1H para o composto presente no cartucho 1G5 - kaempferol-3-*O*- β -[glc (1-2) 2''-*O*- β -glc]. As ampliações entre 4,4 – 5,3 ppm revelam a presença de dois glicosídeos – dubletos (8Hz), enquanto ampliação entre 6,2 e 7,8 mostra o padrão de substituição típico do kanferol.

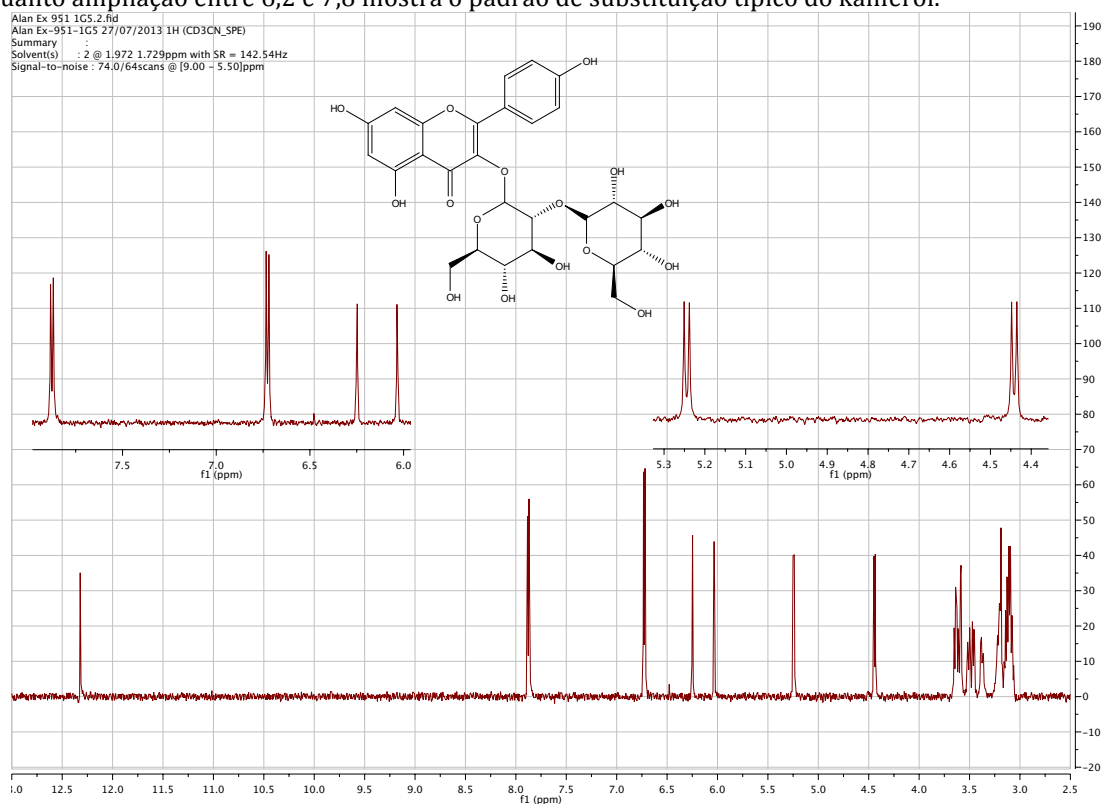
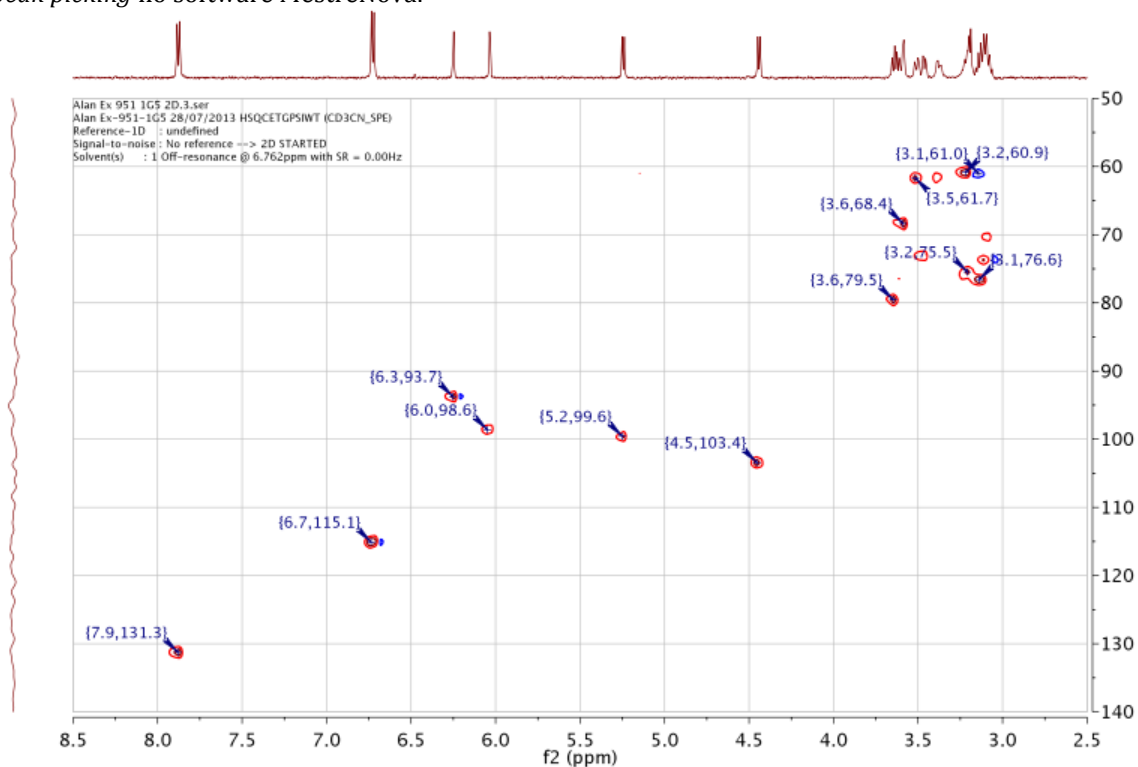


Figura 31. Espectro de ^1H – ^{13}C g-HSQC para o composto presente no cartucho 1G5 - kaempferol-3-*O*- β -[glc (1-2) 2''-*O*- β -glc]. A coloração vermelha dos nodos representa os carbonos metínicos (CH) enquanto as azuis estão associadas a carbonos metilênicos (CH_2). A legenda dos nodos foi obtida através da execução de *peak picking* no software MestreNova.



A partir do espectro de ^1H para a fração 1G11 nota-se a presença de uma dihidroflavona muito similar a naringenina, figura 34. Contudo, observa-se um sinal de deslocamento químico em 3,8 ppm (integrado para 3 hidrogênios – 3H) associado a uma metoxilação, que tanto pode ser em C-7 quanto em C-4'. Espectros simulados de ^1H (figura 35) demonstram um alto grau de similaridade entre C-7, C4' e o espectro real tornando difícil a sua determinação. Nesse caso, experimentos de HMBC e NOESY permitiriam através de correlações espaciais e de longa distância do tipo *J*-3 localizar a posição do grupo metoxi mas devido a baixa quantidade de massa não foi possível a realização.

Figura 32. Espectro de ^1H para o composto presente no cartucho 1G8 - naringenina. As ampliações entre 2,6 – 3,3 ppm revelam o padrão de multiplicidade entre H-2 e H-3 (acoplamentos geminais e equatorial/axial). A ampliação entre 6,2 e 7,8 mostra o padrão de substituição típico do naringenina.

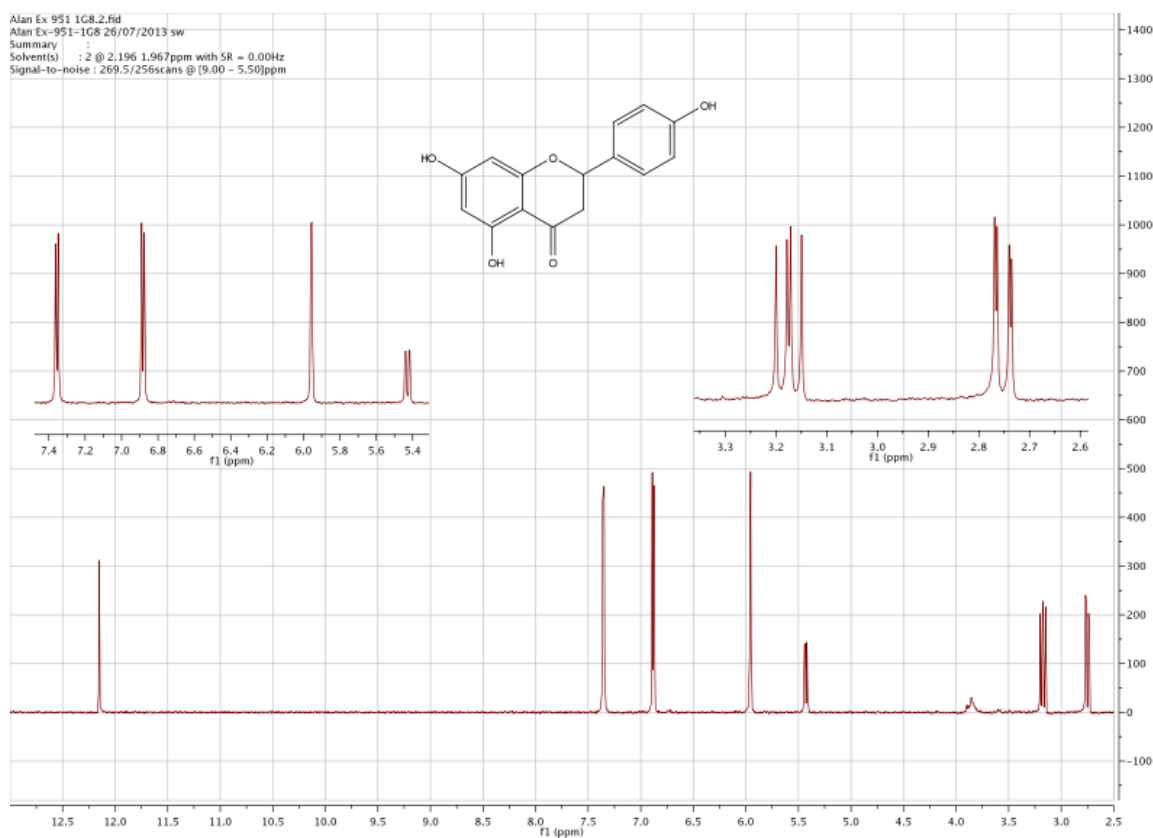


Figura 33. Comparação espectral entre os dados experimentais de ^1H (em vermelho) e dados simulados (MestreNova – em azul) para o composto presente no cartucho 1G8 – naringenina, sob o intervalo de 2,5 a 13 ppm.

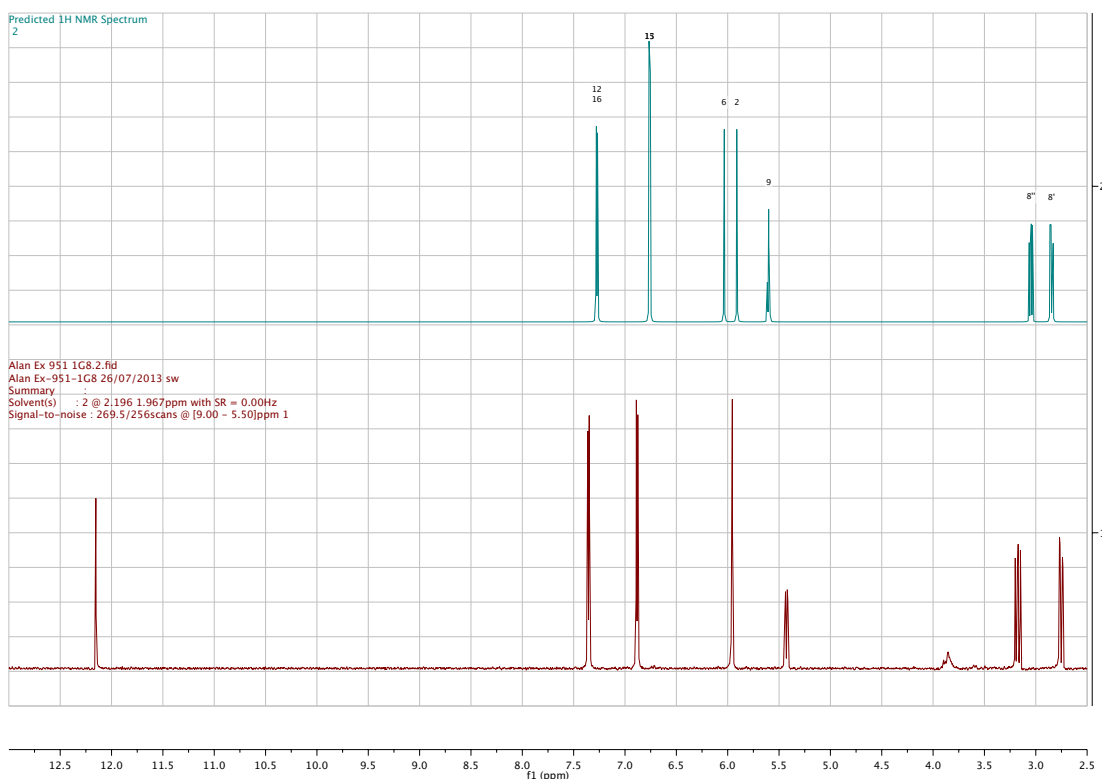


Figura 34. Espectro de ^1H para o composto presente no cartucho 1G11 – derivado metoxilado de naringenina (7 ou 4'). As ampliações entre 2,6 – 3,3 ppm revelam o padrão de multiplicidade entre H-2 e H-3 (acoplamentos geminais e equatorial/axial). A ampliação entre 6,2 e 7,8 mostra o padrão de substituição típico do naringenina.

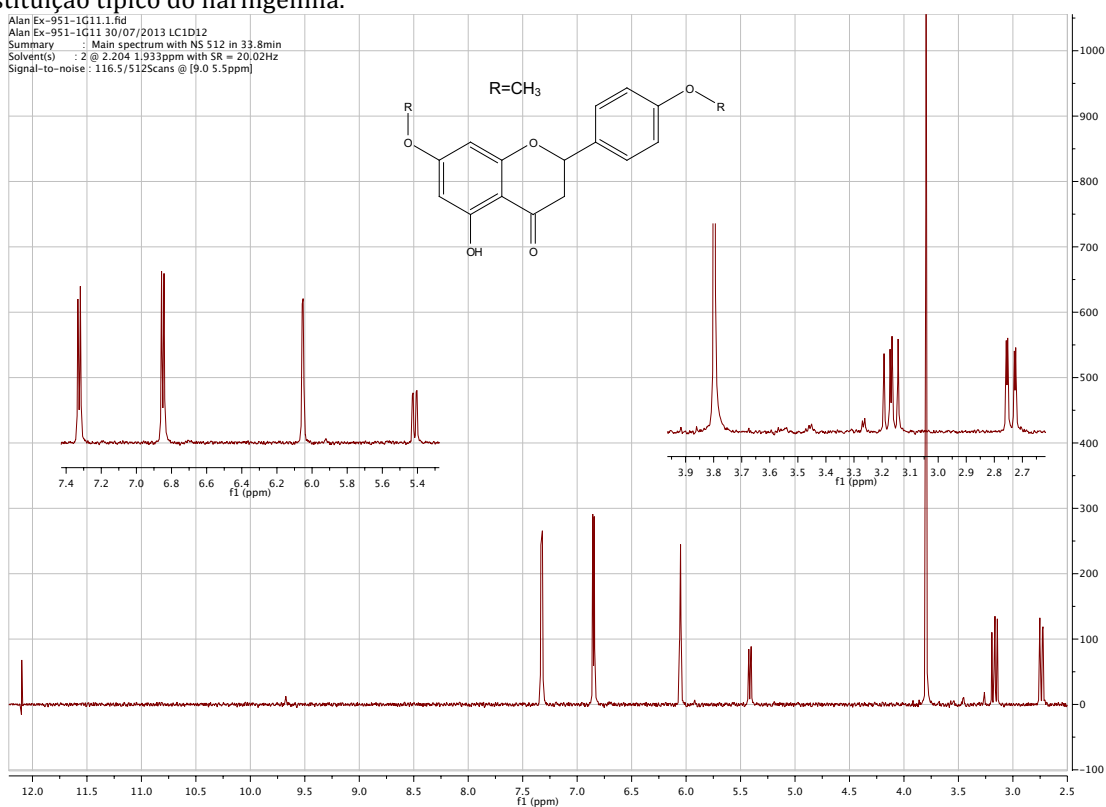
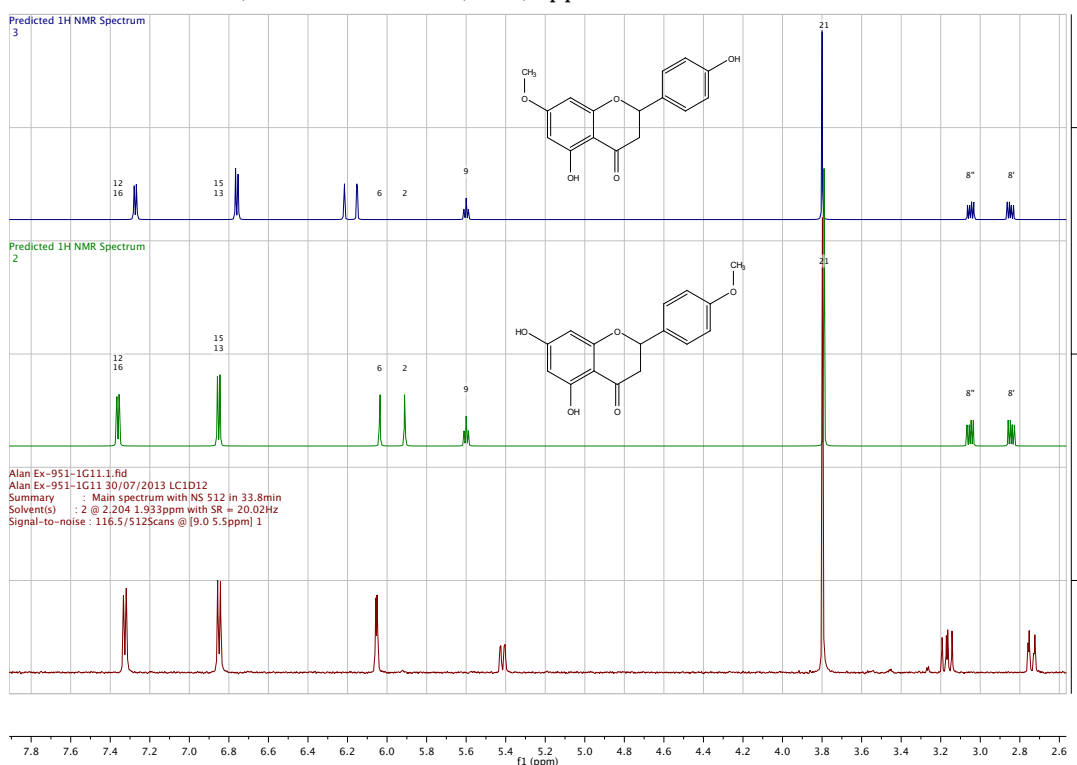


Figura 35. Comparação espectral entre os dados experimentais de ^1H (em vermelho) e dados simulados (MestreNova) do derivado 4'-O-metilnaringenina – verde e 7-O-metilnaringenina – azul) para o composto presente no cartucho 1G8, sob o intervalo de 2,6 a 7,8 ppm.



Com o auxílio do sistema CLAE-SPE-RMN foi possível coletar de forma automatizada os quatro picos majoritários do extrato bruto de folhas de *S. paniculatum* (M951). Utilizando de apenas de 8 mg de extrato bruto (20 injeções de 20 μL - 400 μL - sob uma solução de 20 $\text{mg} \cdot \text{mL}^{-1}$) foi possível elucidar 4 flavonóides (3 totais ou 1 parcial) derivados da quercetina e kaempferol glicosilados. Isso só foi possível graças as novas sequências de pulso para geração de experimentos de dupla radiação em RMN (supressão do sinal do solvente e da água).

Esta abordagem mostrou-se eficiente tanto na busca por métodos rápidos de separação e desreplicação de compostos quanto das premissas da química verde, como a miniaturização, baixo custo energético e principalmente a diminuição do uso de solventes tóxicos.

CLAE-SPE-RMN em modo *time slice*

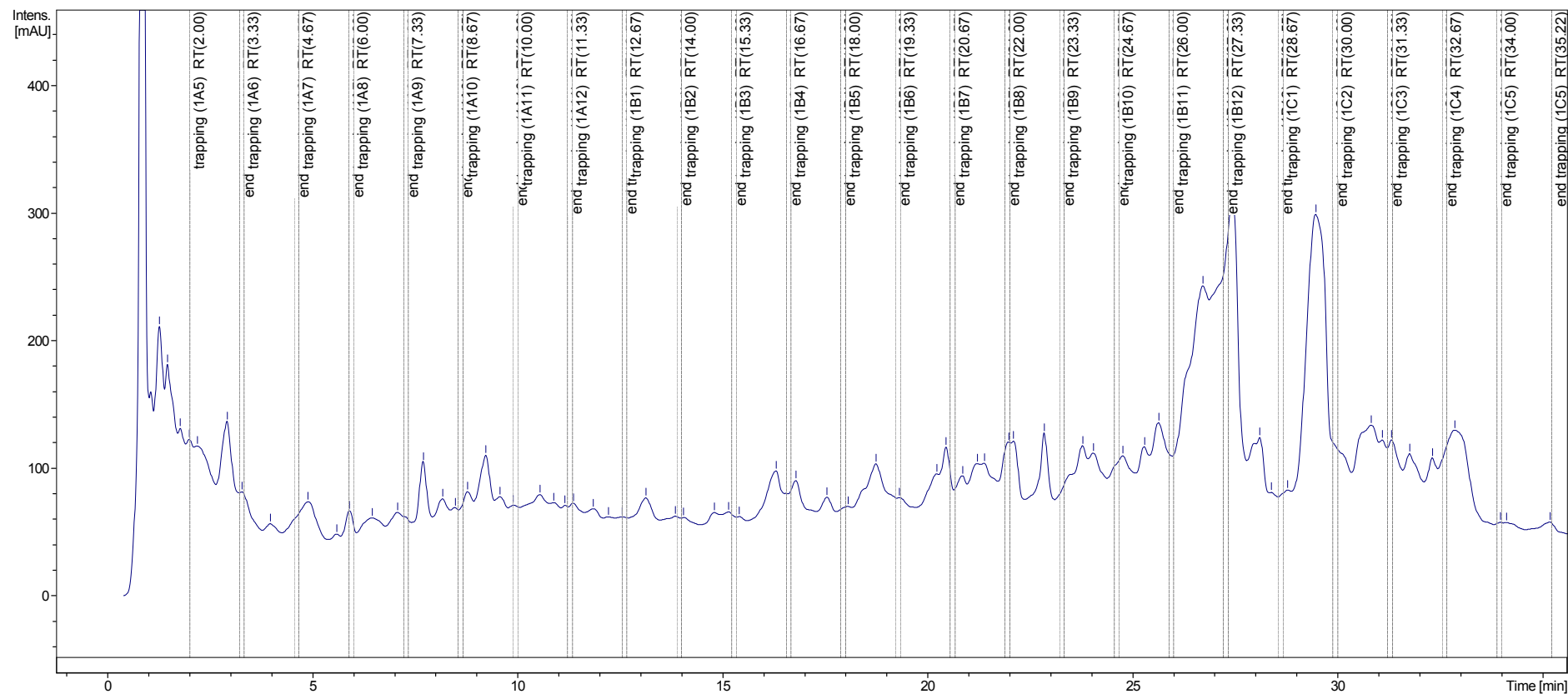
A análise do perfil metabolômico via CLAE-DAD do extrato metanólico das folhas de *S. paniculatum* (coletada no horto da faculdade de ciências farmacêuticas da UNESP de Araraquara) revelou uma baixa resolução cromatográfica impossibilitando o uso da abordagem em CLAE-SPE em modo seleção, figura 36. Nesse caso, frações semi-purificadas foram coletadas em intervalos iguais de tempo e, em seguida, submetidas a análise através de experimentos de ^1H com dupla radiação (LC1D12 - solvente e água). Devido a complexidade química, contaminantes do cartucho e baixa concentração dos compostos somente foi possível a detecção de sinais de RMN para as frações B7, B8 e B10, figura 37.

O principal desafio em abordagens de desreplicação em misturas via RMN, como ocorre, é a falta ou sobreposição de sinais marcadores de classes metabólicas. Certas classes de alcaloides e terpenos esteroidais pertencentes às espécies de *Solanum* se enquadram nesse caso. Correlações de segunda ordem, desdobramentos entre hidrogênios de carbonos quirais, correlações de sistemas cíclicos, efeitos anisotrópicos entre outros criam um ambiente altamente complexo e sobreposto numa estreita faixa de sinais (1 a 5 ppm). Nestes casos, fica impossibilitado a aplicação de ferramentas tradicionais de desreplicação em mistura.

Dentro deste cenário, este trabalho buscou alternativas para desreplicação de compostos de alta complexidade e, uma estratégia de abordagem vem sendo desenvolvida através da combinação entre experimentos homonucleares e heteronucleares: ^1H , COSY, DOSY, TOCSY1D, HSQC e HMBC com ferramentas de estatísticas de deconvolução em RMN como RANSY e STOCYSY.

O esquema 8 apresenta a estratégia de combinação entre os experimentos de deconvolução de sinais de RMN com aqueles relativos determinação estrutural de compostos. A fração 1B8 foi escolhida como modelo de aplicação devido a sua alta concentração e menor complexidade química. Para os experimentos de RANSY e STOCYSY foram utilizados as informações das frações 1B7, 1B8 e 1B10 visando a elucidação do composto majoritário em 1B8. Estas frações só foram selecionadas devido a alto grau de similaridade, ou seja, pertencem a mesma classe de compostos, figura 37. Os experimentos de ^1H , COSY, DOSY, TOCSY1D, HSQC e HMBC foram realizados apenas para a fração 1B8.

Figura 36. Perfil cromatográfico do extrato de folhas de *S. paniculatum* (FCFAR – Ar) obtido a 220 nm. As linhas verticais (pontilhadas) referem-se ao intervalo de coleta (80 segundos) dos picos no sistema CLAE-SPE. As siglas entre parênteses descrevem em quais cartuchos de SPE foram coletados os compostos. A critério do trabalho, estas siglas foram utilizadas para identificar os compostos.



Esquema 8. Ilustra a estratégia de elucidação estrutural a partir de coletas em *time slice*. A primeira etapa consiste no fracionamento do extrato bruto através de coletas sucessivas a cada 80 segundos exemplificado em 1B7; 1B8, 1B9 e 1B10. Em seguida, duas abordagens são realizadas: (i) uso do conjunto de frações para o processo de deconvolução espectral (1B7, 1B8 e 1B10) em STOCSY e RANSY e, (ii) uso singular da fração B8 para experimentos de DOSY (deconvolução), ^1H , g-HSQC, g-HMBC, COSY e TOCSY1D (elucidação estrutural). Ao final os resultados são combinados e é determinada uma das estruturas presente na fração 1B8.

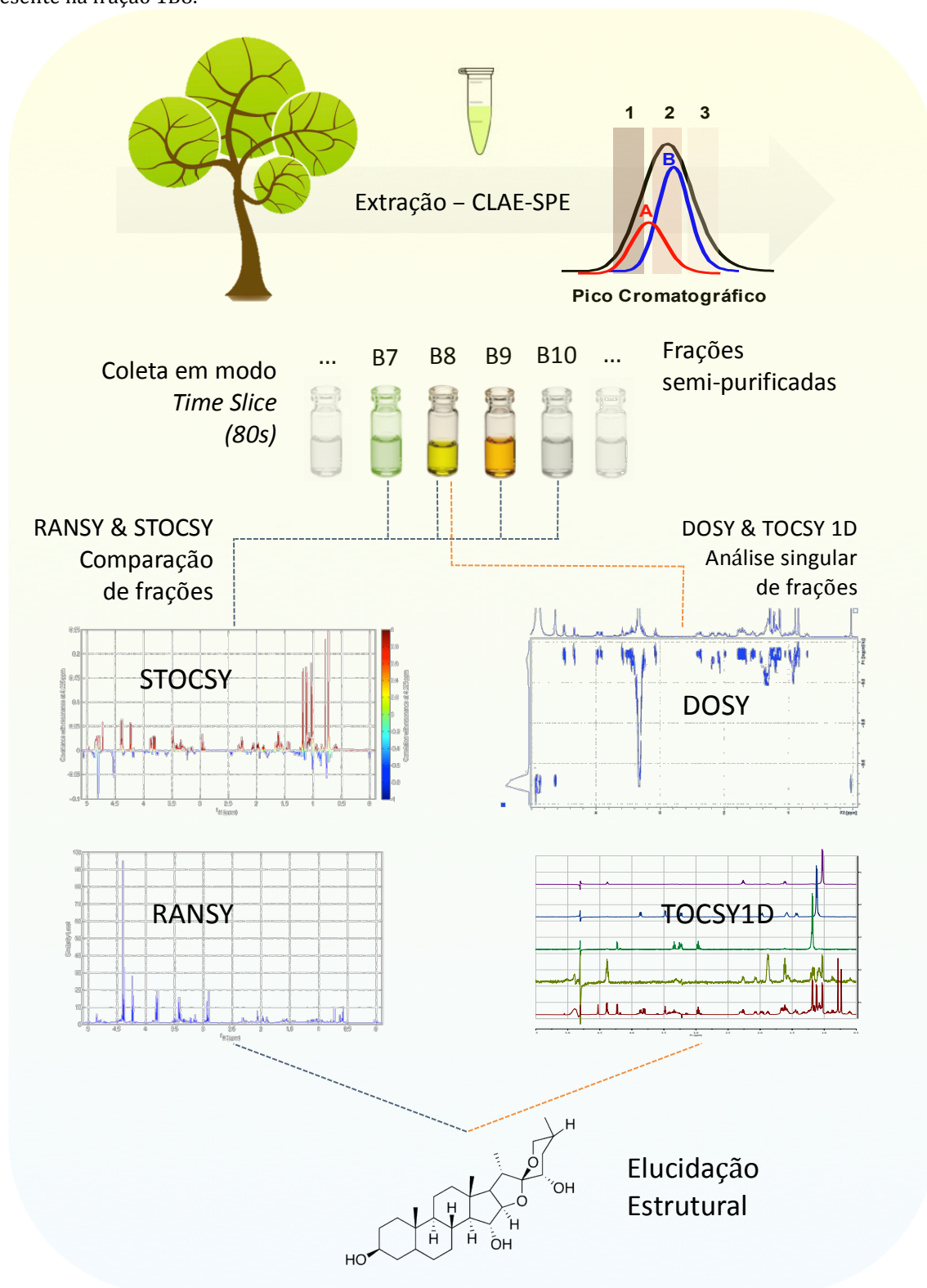
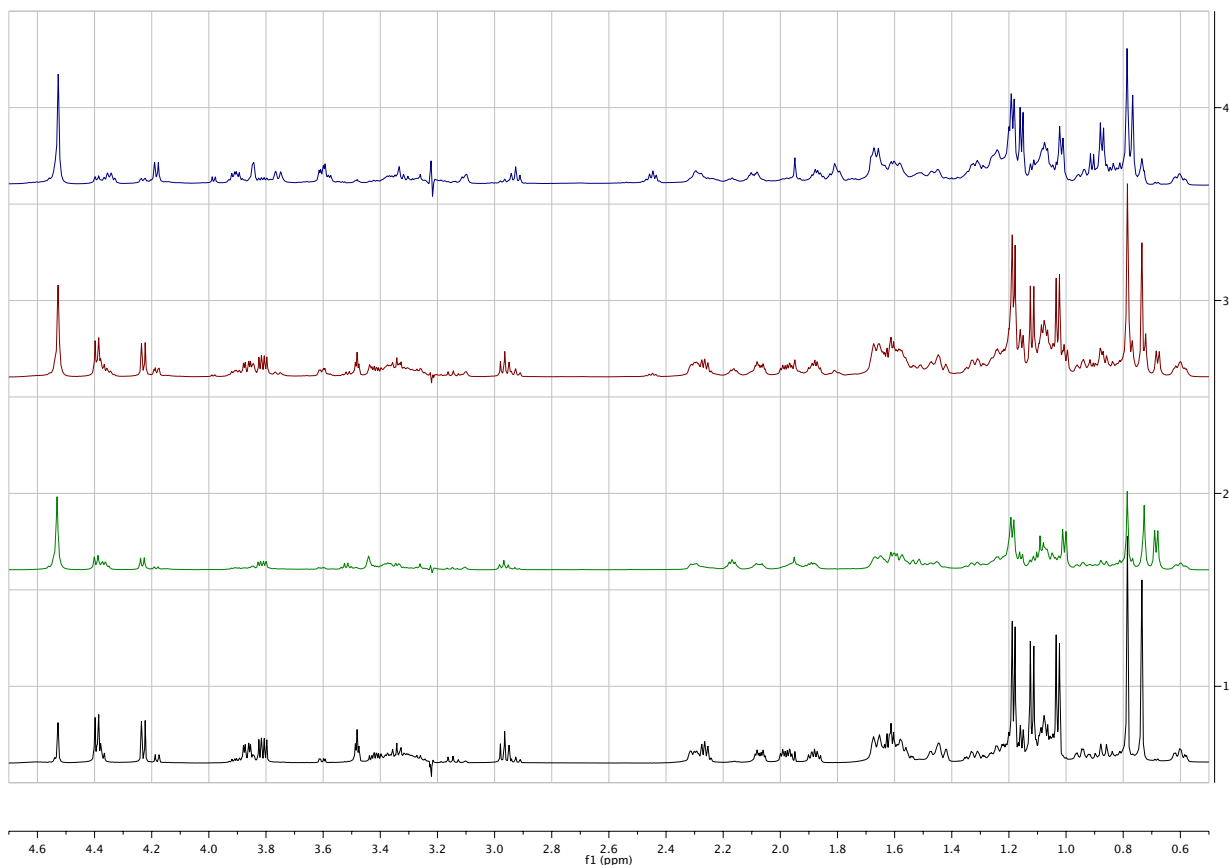


Figura 37. Espectros de ^1H (0,5 a 4,7 ppm) para o extrato metanólico de *S. paniculatum* (FCFAR – Ar) – (preto) e as frações 1B7 (verde), 1B8 (vermelho) e 1B10 (azul). Nota-se que os compostos presentes nas frações são altamente correlacionados e apresentam-se como produtos majoritários do extrato bruto.



Para o processo de deconvolução via RANSY e STOCOSY foram selecionados os sinais de *driving peak* (fração 1B8) a partir do sinal majoritário a cada intervalo de 0,5 ppm. Ao todo foram gerados 12 diferentes espectros para cada ferramenta de deconvolução. No entanto, devido a similaridade dos resultados somente são mostrados os espectros gerados pelo RANSY e STOCOSY para o *driving peak* a 4,4 ppm, figuras 38 e 39.

Os espectros gerados pelo STOCOSY, figuras. 38(A e C) e 39 (A e C) representam as correlação entre os sinais de RMN a partir de duas componentes: cor e intensidade. A cor representa o nível de correlação entre os sinais de RMN (sistemas de spin) frente ao *driving peak* selecionado. Quanto mais vermelho, maior é o grau de correlação entre o DP e os demais sinais, enquanto a coloração azul corresponde a premissa inversa. A intensidade revela quanto em “concentração” estes sinais estão relacionados. Assim é possível observar quem está relacionado ao DP e ainda quanto em concentração estes sinais estão associados.

Figura 38. Espectros de ^1H gerados pela ferramentas STOCSY e RANSY para as amostra 1B7, 1B8 e 1B10. Em (A) o espectro de ^1H STOCSY obtido para janela espectral de 0-5 ppm. A correlação dos sinais frente ao *Driving Peak* ocorre através da coloração, ou seja, tons de mesma cor revelam sinais de um mesmo sistema de spin. Em (B) é observado o espectro ^1H gerado pelo RANSY sob intervalo de 0-5ppm. A correlação dos sinais de um mesmo sistema de spin frente ao DP se dá pela intensidade dos sinais. Em (C) e (D) são mostrados as ampliações de 0 a 1,5 ppm para o STOCSY e RANSY, respectivamente.

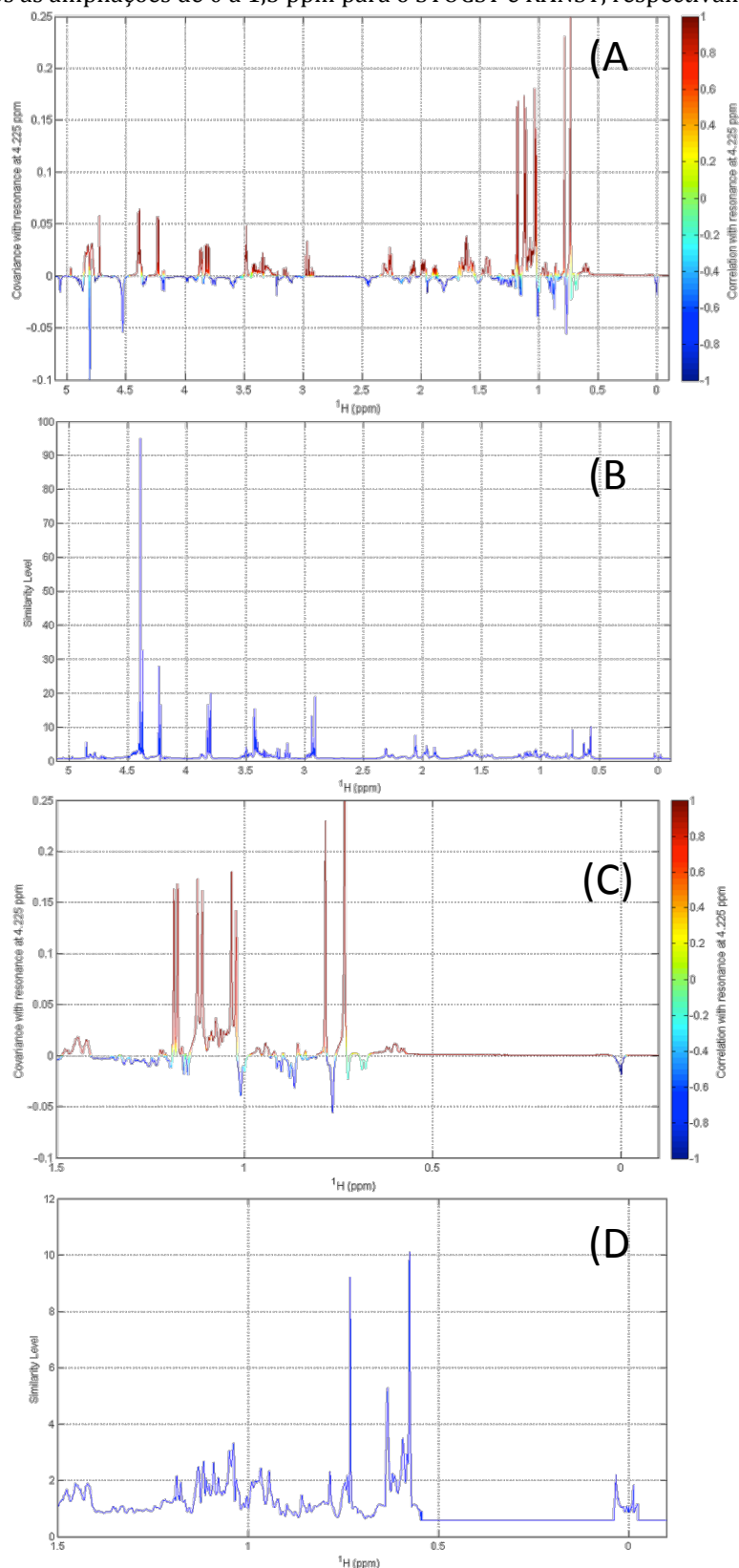
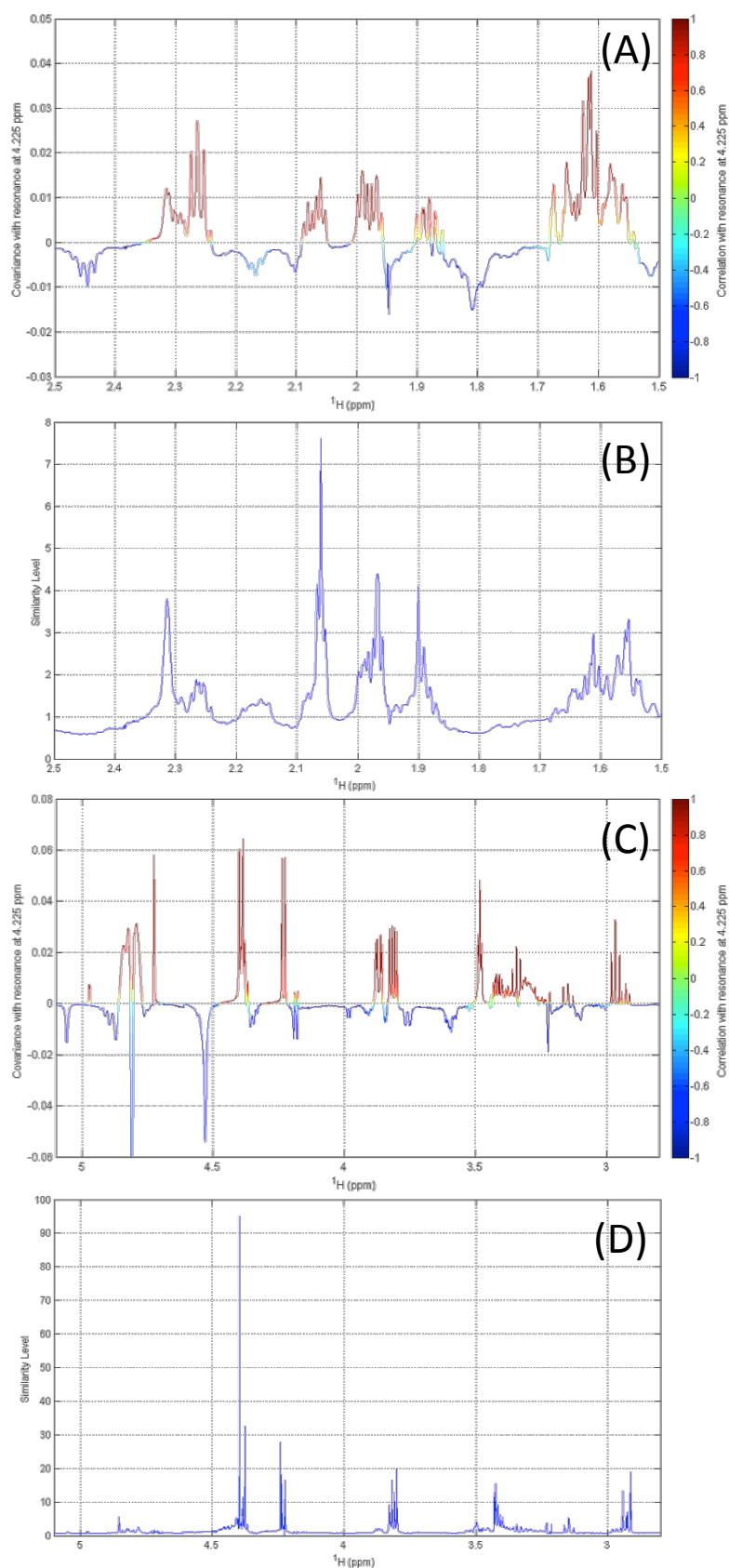


Figura 39. . Espectros de ^1H gerados pela ferramentas STOCSY e RANSY pela comparação das amostras 1B7, 1B8 e 1B10. (A) e (B) representam as ampliações de 1,5 a 2,5 ppm e, (C) e (D) para os intervalos de 2,5 a 5 ppm para o STOCSY e RANSY, respectivamente



As figuras 38 (B e D) e 39 (B e D) representam os espectros gerados a partir do RANSY. Ao analisar ambas metodologias de deconvolução é possível observar um padrão semelhante tanto para o STOCYSY quanto para o RANSY.

No entanto, o STOCYSY estabelece um critério de normalização das razões através da coloração, o RANSY por outro lado, divide as razões pelo desvio padrão entre os mesmos sinais nas diferentes frações gerando um espectro adimensional contendo os valores brutos. O RANSY não agrada visualmente mas apresenta uma relação fidedigna das relações entre os sinais de RMN e o DP.

De modo geral os espectros gerados pelo STOCYSY e RANSY revelaram a correlação de sistemas spin relativos a glicosídeos (região entre 3 e 3,5 ppm e hidrogênios anomérico entre 4 e 4,5 ppm), metilas ligadas a carbonos terciários ou quaternários (singletos e dubletos) em torno de 1,0 a 1,2 ppm e a um grande número de multipletos entre 1,3 a 3,0 ppm.

A segunda etapa da estratégia utilizou apenas a fração B8 para geração de espectros monodimensionais (^1H , TOCSY1D), bidimensionais (COSY, HSQC e HMBC) e pseudo-bidimensionais (DOSY).

O experimento de DOSY foi realizado afim de avaliar diferentes compostos presentes na fração B8 a partir do coeficiente de difusão molecular. A partir da figura 40, é possível identificar a presença de uma mistura de 7 compostos: TMS, MeOH, H_2O , HDO, um composto associado a resina do cartucho utilizado para o aprisionamento molecular e duas substâncias não identificadas (Subst. 1 e Subst.2).

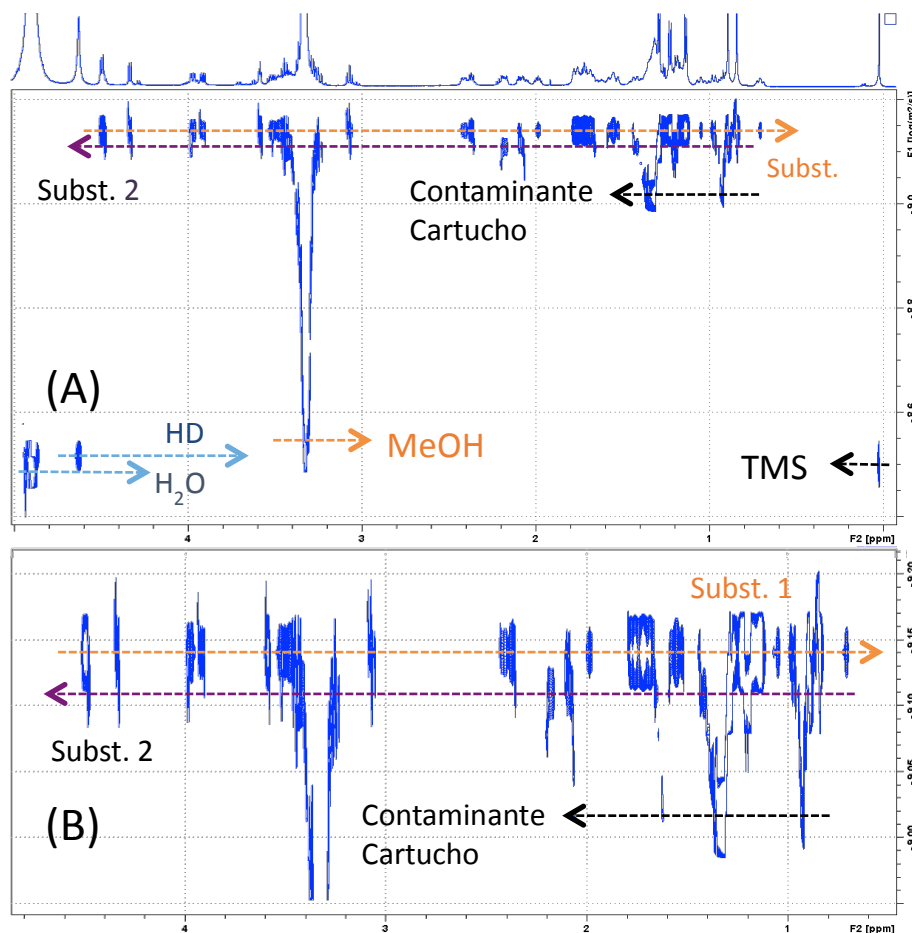
De fato, a substância 1 apresenta-se como composto majoritário na mistura. Combinando as informações obtidas pelos experimentos de RANSY, STOCYSY e DOSY com os dados provenientes da base de dados construída e dos principais quimiomarcadores da família Solanaceae podemos apontar que este composto pertence a um terpeno ou alcaloide esteroideal. Essa afirmação baseou-se na metilas quaternárias e terciárias e multipletos relativos a hidrogênios metilênicos, sem presença de insaturações nos anéis B e/ou C do triterpeno (ausência de dubletos na região entre 5 e 6 ppm) e sinais relativos a glicosídeos (dois dubletos associados a hidrogênios anoméricos).

Uma vez definido o conjunto de sinais pertencentes a Subst. 1 foram realizados experimentos de radiação específica (TOCSY1D) afim de separar aqueles sinais sobrepostos que não puderam ser definidos nas etapas anteriores.

Num típico experimento de TOCSY (2D) observa-se, através da transferência polarização entre os núcleos de ^1H , correlações de longa distância entre hidrogênios de um mesmo sistema de spin. Tais sistemas de spins são limitados pela presença de carbonos quaternários (insaturados) e/ou heteroátomos como o oxigênio e nitrogênio. Contudo, nos experimento de TOCSY(2D) todos os prótons são excitados ao mesmo tempo o que leva a um aumento do nível de sobreposição de sinais.

Entretanto, em experimentos do tipo TOCSY1D o usuário seleciona qual o sinal em Hz que será irradiado, delimitando a observação de apenas um sistema de spin. Assim, permite-se a separação de sinais em regiões sobrepostas que pertencem a diferentes sistemas de spin. Neste estudo foram realizados experimentos de TOCSY 1D a partir do conjunto de sinais pertencentes a Subst.1 obtidos através das técnicas de RANSY STOCYSY e DOSY, Figura 41.

Figura 40. Espectro pseudobidimensional (^1H) de DOSY para fração 1B8. No eixo horizontal são mostrados os valores de deslocamento químico (ppm) enquanto o eixo vertical descreve o coeficiente de difusão molecular. Em (A) temos a visão geral do espectro de DOSY. A linhas horizontalmente tracejadas representam os composto presente na fração. Em (B) observa-se apenas as Subst. 1 e 2 e o contaminante do cartucho.



A partir dos resultados foi possível determinar a presença de seis sistemas de spin, tabela 12. No entanto, alguns sistemas de spin não puderam ser determinados devido ao nível de sobreposição (desvio do centro do sinal) e pela complexidade espectral dos sinais (multipletos). Adicionalmente foram realizados experimentos de $^1\text{H} - ^{13}\text{C}$ HSQC e $^1\text{H} - ^{13}\text{C}$ HMBC para a fração B8, figura 42.

Figura 41. Espectros $^1\text{H} - ^1\text{H}$ TOCSY1D para a fração 1B8. Em (A) observa-se a presença de 7 espectros sendo: vermelho (inferior – extrato bruto da fração 1B8) ; bege (segundo – irradiado em 4,39 ppm); verde claro (terceiro – 4,23 ppm); verde escuro (quarto – 4,18ppm); azul (quinto – 3,81ppm); roxo (sexto – 3,48 ppm) e rosa (sétimo – 2,27 ppm). Em (B) são apresentados 5 espectros: vermelho (inferior – extrato bruto de 1B8); bege (segundo – 1,88 ppm); verde (terceiro – 1,19 ppm); azul (quarto – 1,12 ppm) e roxo (quinto – 1,03 ppm).

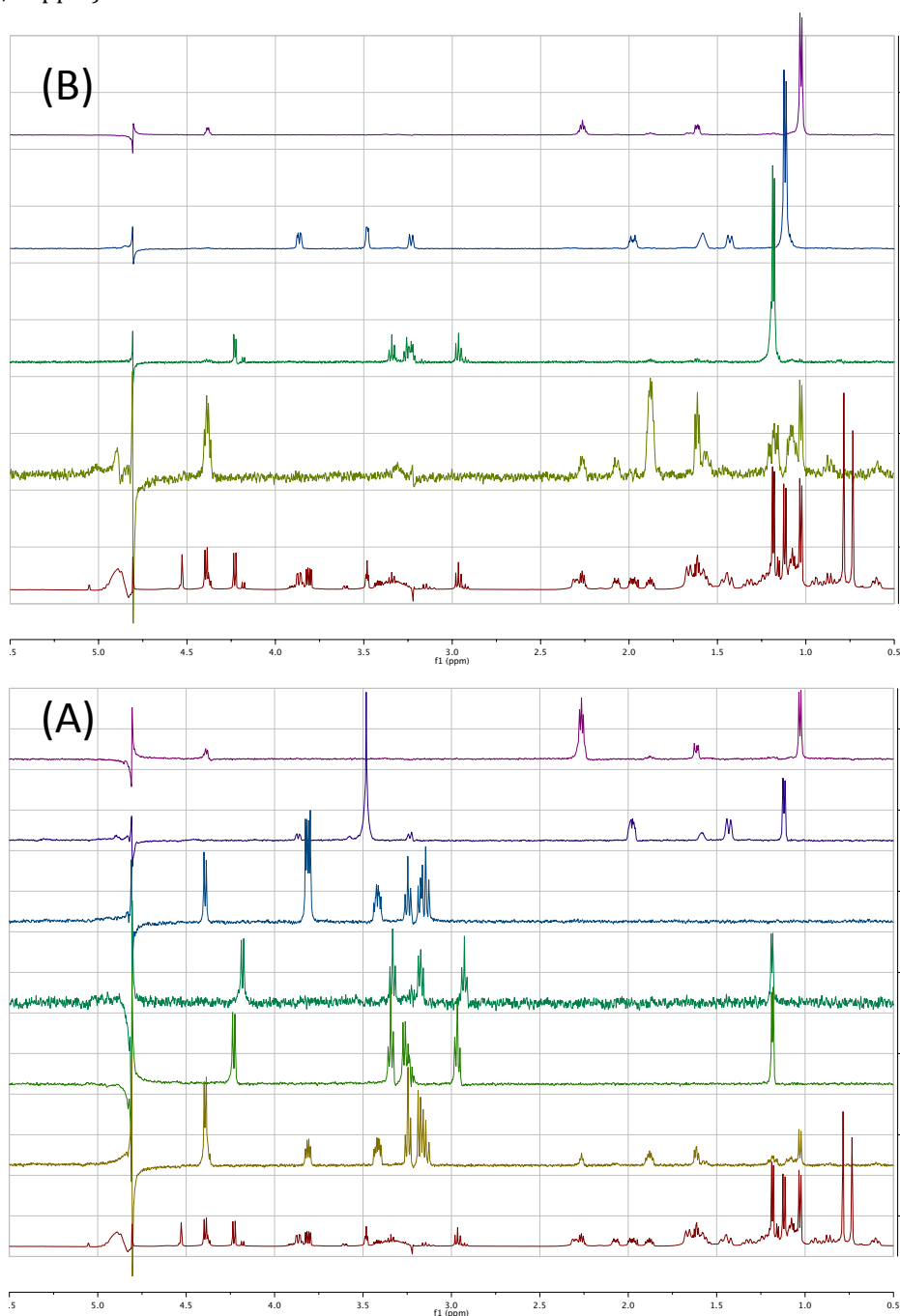


Figura 42. Espectros Bidimensionais obtidos para a fração 1B8. (A) $^1\text{H} - ^1\text{H}$ COSY; (B) $^1\text{H} - ^{13}\text{C}$ HSQC e (C) $^1\text{H} - ^{13}\text{C}$ HMBC.

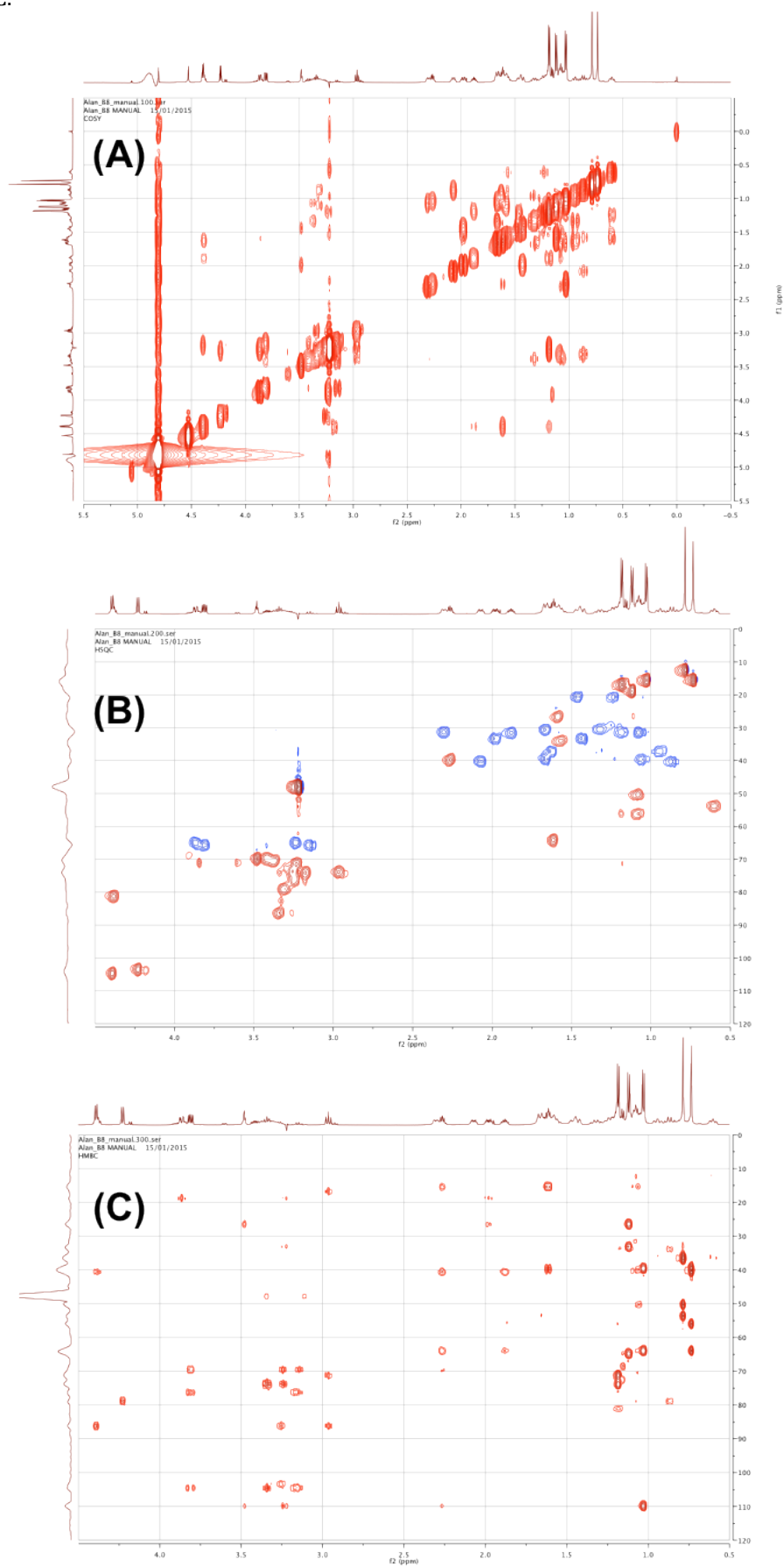


Tabela 12. Sistemas de spin obtidos a partir da combinação de resultados das técnicas STOCSY, RANSY, DOSY e experimentos de ^1H - ^1H COSY, ^1H - ^1H TOCSY1D, ^1H - ^{13}C HSQC e ^1H - ^{13}C HMBC. (figuras 38 - 42). Todos conjuntos de tabela apresentam valores de deslocamento químico de ^1H com a respectivo acoplamento ^1H (multiplicidade), COSY e com ^{13}C através de HSQC e HMBC. Cada cor representa um sistema de spin: preto - sistema 1; azul - sistema 2; rosa - sistema 3; verde - sistema 4; laranja - sistema 5 e roxo - sistema 6. A presença de diferentes nos colunas destinadas ao HMBC referem-se a conexão entre diferentes sistemas de spin para a subst.1. A coloração vermelha pode tanto representar carbonos quaternários não presentes nos sistemas de spin ou estar relacionados a sistemas de spin não observados nos experimentos.

SISTEMA DE SPIN 1 - Aglicona (2,27 ppm - espectro rosa: figura 41(A) e 1,03 ppm - espectro roxo: figura 41(B)).

#	^1H (J)	HSCQ ^{13}C	HMBC1	HMBC2	HMBC3	HMBC4	HMBC 5	COSY
1	4,38(dd)	81,5 (CH)		42,1				1,61
2	2,31(m)	31,3 (CH ₂)		40,8				1,06
3	2,26(qt)	40,8 (CH)	16,7	42,1	65,2	70,0	111,0	1,03
4	1,88(m)	31,3 (CH ₂)		40,8				4,38
5	1,61(dd)	65,2 (CH)	16,7	40,8		81,5		4,38
6	1,03(d)	16,7 (CH ₃)		40,8	65,2		111,0	2,26

SISTEMA DE SPIN 2 - O- β -Glicose [4,39 ppm - espectro bege: figura 41(A) e 3,81 ppm - espectro azul: figura 41(A)]

#	^1H	HSCQ ^{13}C	HMBC 1	HMBC 2	HMBC 3	HMBC 4	HMBC 5	COSY
1	4,39 (d)	106,0 (CH)				87,5		3,18
2	3,81 (dd)	66,8 (CH ₂)	70,8					3,42
3	3,42 (m)	70,8 (CH)			77,5			3,81
4	3,25 (t)	77,5 (CH)	70,8					
5	3,15 (t)	66,9 (CH ₂)	70,8					3,42
6	3,18 (t)	74,9 (CH)			77,5		106,0	4,40

SISTEMA DE SPIN 3 - O- β -deoxi-glicose [4,23 ppm - espectro verde claro: figura 41(A) e 1,19 ppm espectro verde : figura 41(B)]

#	^1H	HSCQ ^{13}C	HMBC 1	HMBC 2	HMBC 3	HMBC 4	HMBC 5	COSY
1	4,23 (d)	104,5 (CH)	79,9	72,6				3,25
2	3,34 (t)	87,5 (CH)			74,9		106,0	2,96
3	3,25 (m)	75,0 (CH)				87,5	104,5	1,18
4	2,96 (t)	75,0 (CH)	18,0	72,6		87,5		3,34
5	1,18 (d)	18,0 (CH ₃)		72,6	75,1			3,25*

SISTEMA DE SPIN 4 - Aglicona [3,81 ppm - espectro azul: figura 41(A) e 3,48 ppm espectro roxo: figura 41(A)]

#	^1H	HSCQ ^{13}C	HMBC 1	HMBC 2	HMBC 3	HMBC 4	HMBC 5	COSY
1	3,86 (dd)	66,1 (CH ₂)		34,4			111,1	3,23
2	3,48 (t)	70,0 (CH)	27,7	34,4			111,1	1,43
3	3,23 (d*)	66,1 (CH ₂)			75,2		111,1	3,86
								1,43
4	1,98 (dt)	34,4 (CH ₂)	27,7					3,48
5	1,58 (s*)	27,7 (CH)						
								1,98;
6	1,43 (d*)	34,4 (CH ₂)					111,1	3,48
7	1,12 (d)	19,1 (CH ₃)	27,7	34,4	66,1			1,58

* representa um sinal largo.

SISTEMA DE SPIN 5 - Aglicona [0,78 ppm]

#	¹ H	HSCQ ¹³ C	HMBC 1	HMBC 2	HMBC 3	HMBC 4	HMBC 5	COSY
1	0,78 (s)	12,6 (CH ₃)	37,7	51,3	54,8			

SISTEMA DE SPIN 6 - Aglicona [0,73 ppm]

#	¹ H	HSCQ ¹³ C	HMBC 1	HMBC 2	HMBC 3	HMBC 4	HMBC 5	COSY
1	0,73 (s)	15,5 (CH ₃)	40,8	42,1	56,2(CH)	65,2		

A elucidação da subst. 1 partiu da análise do sistema de spin 2, tabela 12. Neste sistema é possível observar a presença de um hidrogênio anomérico (dublete) a 4,39 ppm associado a uma xilose na forma beta (β -xilose) de anéis piranosídicos ($J - 7,7$ Hz). A presença de seis hidrogênios na região entre 3-5 ppm, sendo cinco metínicos (CH) e dois metilênicos (CH₂) colabora para a presença de tal pentose.

Os cinco sinais presentes no sistema de spin 3 sugere a presença de uma segunda xilose. Entretanto, o experimento de HSQC utilizando de filtros quânticos (Duplo filtro quântico - determina a presença de CH₃/CH ou CH₂ pelo uso de diferentes fases) determinou a presença de uma metila a 1,18 ppm, sugerindo uma β -D-deoxi-hexose. Os dados do hidrogênio anomérico (dublete) a 4,23 ppm com constante de acoplamento de ($J - 7,9$ Hz) reforçam a hipótese de um derivado de glicose. Em sistemas de anéis furanosídicos, os sinais de hidrogênio anomérico são deslocados para campo mais alto quando comparados aos sistemas piranosídicos (acima de 5 ppm)(SIVERSTEIN, 2005; MARKHAM; GEIGER, 1993).

Adicionalmente, sugerimos que os dois glicosídeos estejam ligados entre si através de um ponte glicosídica em 1 – 3 xilose – deoxi-glicose (1 –O – β xilose ligada a posição 3 – O – deoxi – β -glicose). Os sinais de HMBC revelam o acoplamento a longa distância entre o carbono anomérico da xilose (106,0 ppm) com o hidrogênio ligado 3,34 ppm (carbono 87,4 ppm – vide tabela de spins 2 e 3) associado a posição 3” da deoxi-glicose.

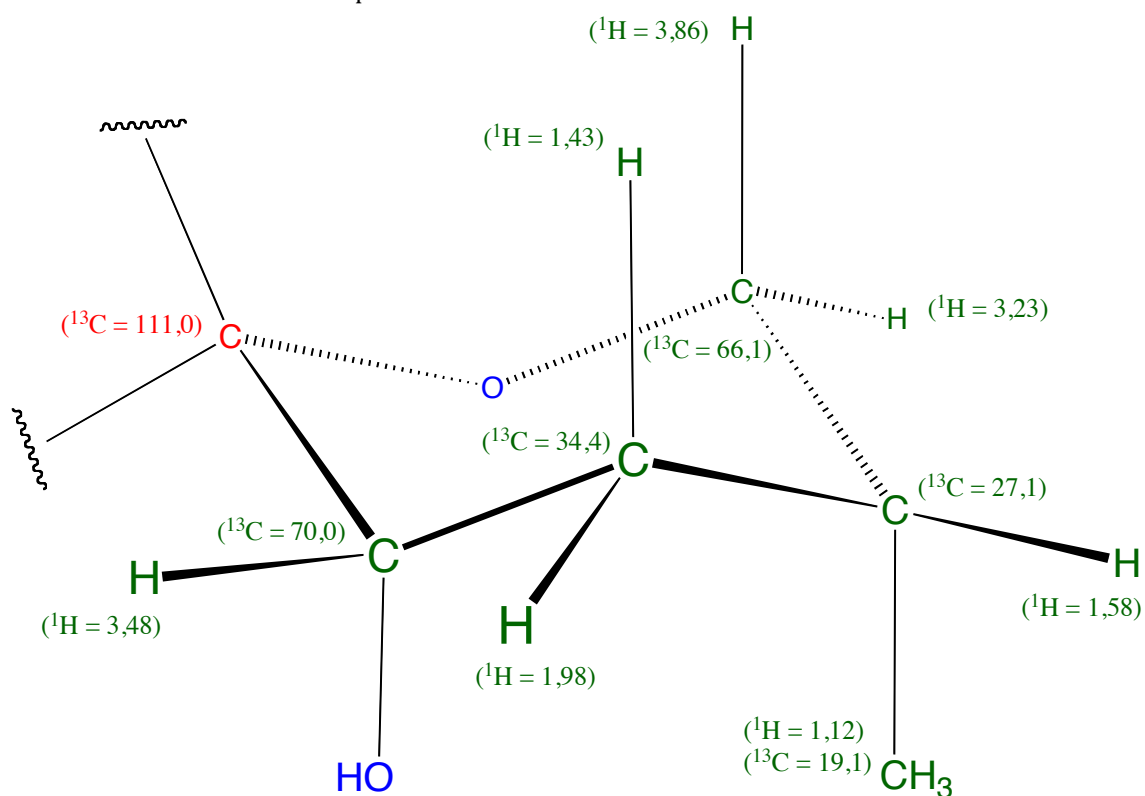
Nos demais sistemas de spin podemos observar as metilas associadas ao esqueleto triterpênico (aglicona). Especificamente observamos duas metilas quaternárias (sistemas 5 e 6) e duas metilas ligadas a carbonos metínicos (sistemas 1 e 4).

A elucidação estrutural da aglicona partiu do dublete relativo a metila do sistema de spin 4. O experimento HMBC revela um acoplamento desta metila a J2 com um carbono metínico a 27,2 ppm. A partir da análise dos experimentos de HSQC e TOCSY1D, este carbono (27,2 ppm) está ligado a multipletos largo a 1,58

ppm - ^1H (figura 41-A espectro roxo). Ainda o HMBC da metila (1,12 ppm) revela acoplamento a J_3 com dois outros carbonos metilênicos (34,4 e 66,1 ppm). Um dos hidrogênios ligado ao carbono em 66,1 ppm gera um duplo dubleto (3,4 e 11,7 Hz) a 3,86 ppm relativo ao acoplamento típico de anéis piranos esteroidais do tipo equatorial-axial e geminal, respectivamente.

O segundo hidrogênio a 3,23ppm gera um dubleto largo de $J = 11,3$ Hz o qual foi atribuído ao acoplamento do tipo geminal (11,3Hz) e equatorial-equatorial (≈ 1 Hz), Figura 43.

Figura 43. Ilustração da atribuição dos sinais de ^1H e ^{13}C referente ao sistema de spin 4. A coloração dos átomos faz referência aos dados apresentados na tabela 11.



Para os dois hidrogênio relativos ao δ de 34,4 ppm observa-se o mesmo tipo de acoplamento visto anteriormente para os hidrogênios ligados ao carbono 66,1ppm. No entanto, HMBC aponta que este carbono metilênico acopla com dois carbonos metínicos. Um deles já esperado pois está associado ao carbono ligado a metila previamente descrita (19,1 ppm) no entanto, para o outro carbono atribuímos a presença de uma hidroxila. O valor de deslocamento químico a 70,0 ppm juntamente com ausência de acoplamentos a longa distância com outros carbonos reforça a ideia de uma hidroxila figura 43.

Dois δ estão associados ao carbono 34,4 ppm, sendo 1,98 ppm (duplo tripleto de $J = 14\text{Hz}$) resultado de um equivalente acoplamento equatorial-equatorial entre seus vizinhos (3,48 e 1,58 ppm) e um acoplamento geminal de (14Hz). O segundo sinal a 1,43ppm apresenta um duplete largo (11,7 Hz) referente ao acoplamento axial-equatorial e um acoplamento geminal ($\approx 11\text{Hz}$) Figura 43.

A multiplicidade dos hidrogênios metilênicos não somente possibilitou a determinação da estrutura do anel pirano mas também forneceu subsídios para atribuição da estereoquímica do grupo metil (axial) e da hidroxila (axial) presentes no anel. Adicionalmente, o experimento de HMBC permitiu a observação do acoplamento a $J3$ dos hidrogênios metilênicos com o carbono quaternário em 111 ppm. Este carbono é responsável pela ligação espiro entre o anel pirano e o restante do triterpeno, tabela 12 e figura 43.

A sequência da elucidação estrutural partiu da busca do carbono quaternário a 111 ppm em outros sistemas de spin. De fato, o sistema de spin 1, apresenta entre os sinais de HMBC hidrogênio que acoplam para o carbono a 111 ppm.

Assim, a elucidação estrutural do sistema de spin 1 partiu do dubleto da metila ($^1\text{H} - 1,03\text{ ppm}$). Os sinais de HMBC mostram que esta metila está associada a dois carbonos metínicos $\delta\ ^{13}\text{C}$ 40,8 – ^1H 2,26 (quintopletto) e $\delta\ ^{13}\text{C}$ 65,2 – ^1H 1,61 duplo dubleto. Baseado nos triterpenos do gênero *Solanum* e na multiplicidade do sinal de ^1H , é possível afirmar que o carbono 40,8 está ligado diretamente a metila, tabela 12. Outro aspecto importante se dá pelo acoplamento a longa distância (HMBC) entre este sinal de hidrogênio (2,26 ppm) com os carbonos do sistema de spin 4 (70,0 e 111,0 ppm). Assim, o sistema de spin 1 está associado ao anel E do triterpenos, figura 44.

Assim como na etapa anterior, o HMBC permitiu a junção entre o sistema de spin 1 com o 5 através de uma correlação com o carbono quaternário em 42,1 ppm e com carbono metínico em 65,2 ppm. Adicionalmente o sistema de spin 6 fornece através do HSQC (CH) o deslocamento químico do carbono adjacente ao quaternário (^{13}C 56,2 e ^1H 1,08).

Através da histórico biossintético das saponinas esteroidais e das moléculas já relatadas para a espécie *S. paniculatum* juntamente com os dados de ^{13}C do sistema de spin 5 sugerimos que a ligação glicosídica ocorra na posição C-6 do

núcleo triterpênico resultando no composto – paniculonin A, Este composto já havia sido reportado no extrato de folhas de *S. paniculatum* (REF.), figura 45.

Figura 44. Ilustração da atribuição dos sinais de ^1H e ^{13}C referente ao conjunto de sistemas de spin 1, 4 e 6. A coloração dos átomos faz referência aos dados apresentados na tabela 12. A coloração azul representa o átomo de oxigênio.

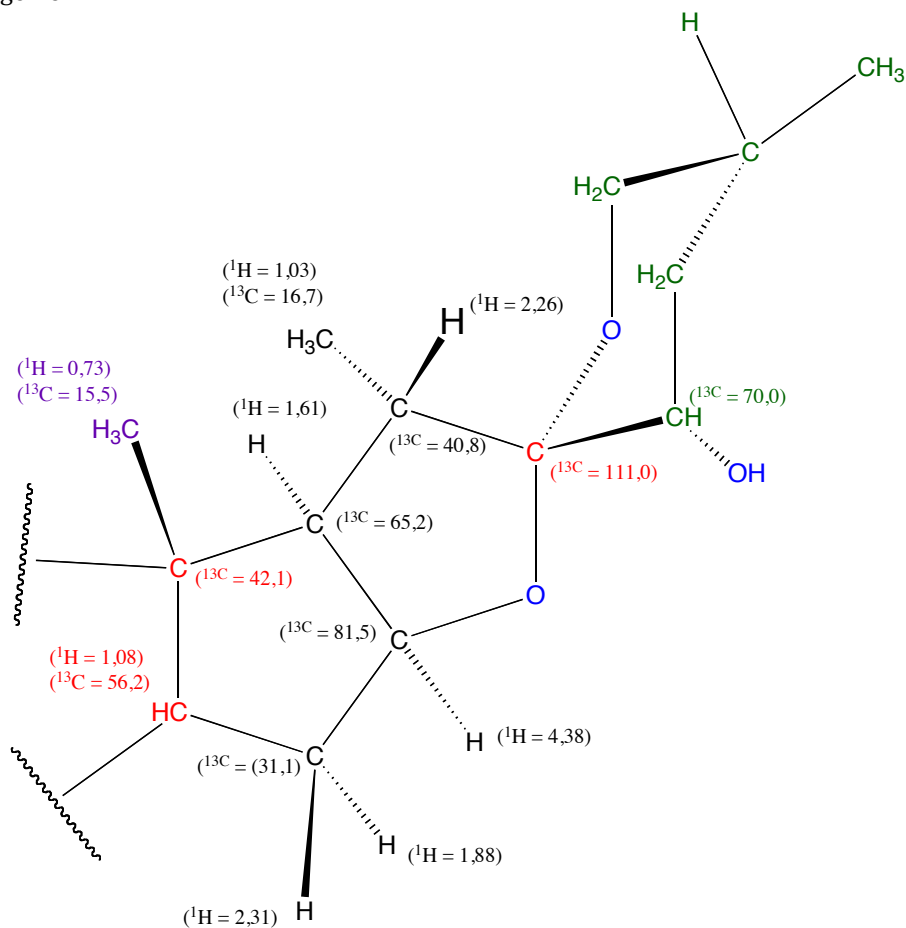
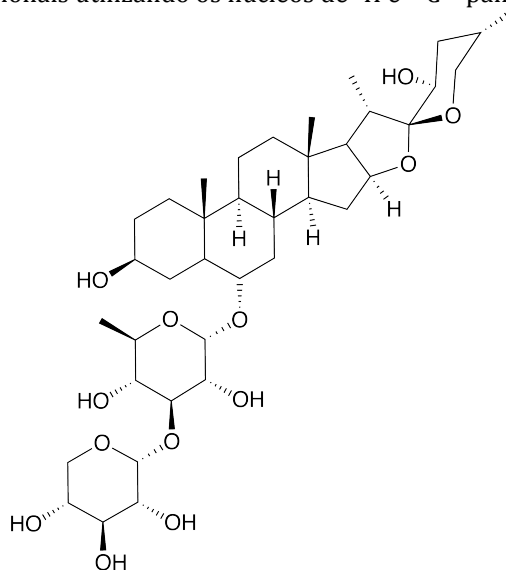


Figura 45. Ilustra a estrutura elucidada a partir da estratégia que combina STOCSY, RANSY DOSY e experimento mono e bidimensionais utilizando os núcleos de ^1H e ^{13}C – paniculonin A.



CONCLUSÕES

O capítulo 3 introduziu conceitos de como a RMN pode auxiliar no desenvolvimento de perfis metabolômicos assim como em estudos de desreplicação fazendo uso do acoplamento da CLAE-SPE-RMN.

Em geral, os processos da CLAE-SPE são de baixo custo em termos materiais de consumo visto que não utiliza solventes deuterados nas etapas de separação cromatográfica como ocorre nas configurações CLAE-RMN *on-flow* e *stop-flow*.

Adicionalmente, a CLAE-SPE-RMN em modo seleção, mostrou que a RMN assim como EM podem acelerar o processo de desreplicação de compostos previamente reportados na literatura. Contudo, a RMN fornece informações mais ricas para a identificação de compostos, como foi o caso dos derivados de flavonas e flavanonas. Foi possível determinar as posições das hidroxilas e glicosídeos dos flavonoides assim como a saturação entre C-2 e C-3 para as flavanonas. Contudo, vale ressaltar que este processo só foi possível graças a capacidade de pico do sistema cromatográfico, na qual separou efetivamente os compostos majoritários de seus vizinhos (Hipótese A e B).

Entretanto, o sistema CLAE-SPE-RMN possui muitas limitações. O uso de detector de arranjo de diodos prioriza uma faixa muito estreita de compostos. Adicionalmente, este tipo de configuração limita o uso modificadores nas fases móveis do sistema de CLAE-SPE. Isto ocorre devido ao número restrito de canais encontrados na probes na qual podem suprimir até dois sinais por experimento (tipicamente água e solvente – Fase móvel). Se caso fossem adicionados quaisquer modificadores, um sinal preponderante seria observado no espectro de RMN. Atualmente existem modificadores deuterados, no entanto, o uso destes solventes estaria em contra-senso em uma das premissa da desreplicação - redução de custos no processo de identificação.

A segunda estratégia vem de encontro com esta limitação. Em casos em que existe uma baixa resolução cromatográfica, métodos mais sofisticados são exigidos no processo de desreplicação ou identificação de compostos.

Assim, esta estratégia priorizou o uso de uma abordagem metabolômica não direcionada (non-target) visando uma rápida separação em pequenas frações que poderiam ser analisadas via experimentos de ^1H (Hipótese C).

A partir daí, fazendo uso de ferramentas quimiométricas de deconvolução como o RANSY e STOCSY foi possível extrair informações adicionais referente a um dado composto de interesse. Dentro de uma perspectiva elucidativa, os métodos tradicionais de determinação de compostos ocorrem pela análise exclusiva de um simples espectro. Nessa estratégia conseguimos extrair informações adicionais pela comparação de um conjunto de espectros referente a uma amostra.

Em paralelo, utilizamos o experimento de DOSY para fins de confirmação dos resultados do RANSY e STOCSY mas também para revelar sinais sobrepostos que tiveram problemas de deconvolução. Ao final, fizemos uso de experimentos de TCOSY1D, HSQC e HMBC para elucidação específica de cada parte da estrutura do composto paniculin A.

CONCLUSÕES GERAIS E PERSPECTIVAS

Ao longo da história, os PN tem provado ser uma fonte inesgotável de medicamentos, seja na forma de fitoterápicos seja na forma de compostos isolados. Durante décadas, enormes esforços foram dedicados ao desenvolvimento de métodos e instrumentos para a pesquisa em PN.

No entanto, os últimos anos foram marcados pela substituição dos programas de pesquisa e desenvolvimento em PN (farmacêuticas) por bibliotecas de compostos sintéticos em conjunto com a química combinatória. A justificativa era que a indústria farmacêutica havia se modernizado para sistemas automatizados e robotizados e, os métodos utilizados pela PN limitavam o processo de descoberta e consequentemente de produção.

Contudo, a promessa da QC por protótipos a fármacos foi abaixo das expectativas e muitos questionamentos foram realizados sobre a base da ciência de descoberta de fármacos. Neste intervalo, avanços nas tecnologias de análise assim como no campo da estatística permitiram a detecção de um grande número de metabólitos em matrizes complexas como plantas e microorganismos. De fato, é dentro deste cenário de transformação que a metabolômica surge no fim da década 1990.

Já sob uma abordagem de sistemas, a metabolômica superou muitas dificuldades práticas envolvidas na química de produtos naturais como a velocidade de triagem e complexidade estrutural dos compostos. De fato, a metabolômica tem sido a ponte entre propriedades intrínsecas dos compostos e a redescoberta de fármacos pelos PN.

Contudo, são os avanços da desreplicação que tem levado em grande parte o sucesso da abordagem metabolômica. Com o advento das bases de dados padronizadas, em conjunto com novas ferramentas quimiométricas e química computacional, a desreplicação saltou de uma condição manualmente comparativa, para análises automatizadas de alta performance visando não somente os compostos previamente relatados mas apontando novos compostos.

Dentro desta perspectiva o presente estudo desenvolveu novos conceitos, métodos e aplicações em desreplicação em PN utilizando as ferramentas analíticas

do estado da arte em cromatografia, espectrometria e espectroscopia (CG, CLAE, EM e RMN).

Métodos baseados no acoplamento CG-EM e CLAE-EM permitiram de forma rápida e sensível a desreplicação de 126 compostos distribuídos em metabólitos primários e secundários. Tal performance só foi possível graças a aplicação de ferramentas quimiométricas de deconvolução e comparação de dados (AMDIS, DataAnalysis e Spectral Networks). Em geral, estas ferramentas tem extraído e agrupado informações ortogonais de amostras levando a resultados mais robustos e confiáveis.

No caso da CLAE-SPE em associação com espectrômetros modernos de RMN (criosonda) foi possível acelerar o processo desreplicação com simples experimentos de ^1H como por exemplo os quatro flavonóides desreplicados. Adicionalmente, a estratégia para elucidação estrutural em meio a misturas de compostos tomou novamente as ferramentas quimiométricas como RANSY e STOCSY assim como novos experimentos em RMN como o DOSY e o TOCSY1D como aporte do processo desreplicação da saponina esteroideal, paniculin A.

De fato, o acoplamento da cromatografia às técnicas de detecção tem sido fundamental para o desenvolvimento de estudos que utilizam os PN. Este acoplamento tem substituído as metodologias clássicas de isolamento, por sistemas automatizados e principalmente reprodutíveis.

A CLAE-RMN, no entanto, é uma técnica recente quando comparada ao acoplamento com técnicas de espectrometria de massas. Assim, existe uma demanda por metodologias e principalmente instrumentos em RMN que sejam mais sensíveis a determinação estrutural de compostos presentes num dado perfil cromatográfico. As bases de dados de RMN é outro aspecto importante para desenvolvimento da área. Existe uma demanda por bancos de compostos que contenham uma extensa diversidade estrutural de forma padronizada na qual o usuário possa comparar seus experimentos com informações já relatadas na literatura.

Os PN sempre foram o centro das pesquisas para a descoberta de fármacos pois eles oferecem um arsenal de moléculas construído ao longo de milhões de anos de pressão evolutiva na qual os métodos atuais de síntese e química combinatória são incapazes de reproduzir.

Certa vez o Prof. Otto Gotlieb disse: *Cada planta possui centenas de substâncias e umas delas pode ser mais importante que uma galáxia.* Neste contexto, a nossa função é encontrar os atalhos para essas raridades de forma racional e principalmente consciente de que a natureza é finita.

REFERÊNCIAS

ÁLVAREZ-SÁNCHEZ, B.; PRIEGO-CAPOTE, F.; LUQUE DE CASTRO, M. D. Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation. **Trac, Trends Anal. Chem.**, v. 29, n. 2, p. 111-119, 2010.

ANTALEK, B. Using pulsed gradient spin echo NMR for chemical mixture analysis: how to obtain optimum results. **Concepts Magn. Reson.**, v. 14, n. 4, p. 225-258, 2002.

BAKER, D. D.; CHU, M.; OZA, U.; RAJGARHIA, V. The value of natural products to future pharmaceutical discovery. **Nat. Prod. Rep.**, v. 24, p. 1225-1244, 2007.

BALOGH, M. P. Debating resolution and mass accuracy. **LCGC North Am.**, v. 22, p. 118-125, 2004.

BARI, M. A.; ISLAM, W.; KHAN, A. R.; MANDAL, A. Antibacterial and antifungal activity of *Solanum torvum* (Solanaceae). **Int. J. Agric. Biol.**, v. 12, n. 3, p. 386-390, 2010.

BARREIRO, E. J.; BOLZANI, V. D. S. Biodiversidade: fonte potencial para a descoberta de fármacos. **Quim. Nova**, v. 32, n. 3, p. 679-688, 2009.

BECKONERT, O.; KEUN, H. C.; EBBELS, T. M. D.; BUNDY, J.; HOLMES, E.; LINDON, J. C.; NICHOLSON, J. K. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. **Nat. Protoc.**, v. 2, n. 11, p. 2692-2703, 2007.

BINGOL, K.; BRUSCHWEILER-LI, L.; YU, C.; SOMOGYI, A.; ZHANG, F.; BRÜSCHWEILER, R. Metabolomics beyond spectroscopic databases: a combined MS/NMR strategy for the rapid identification of new metabolites in complex mixtures. **Anal. Chem.**, v. 87, n. 7, p. 3864-3870, 2015.

BRASIL. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Assistência Farmacêutica. **A fitoterapia no SUS e o Programa de Pesquisa de Plantas Medicinais da Central de Medicamentos**. Brasília, DF, 2006. (Série B: textos básicos de saúde).

BREITHAUPT, D. E.; BAMEDI, A. J. Carotenoids and carotenoid esters in potatoes (*Solanum tuberosum* L.): new insights into an ancient vegetable. **Agric. Food Chem.**, v. 50, p. 7175-7181, 2002.

BUTLER, M. S. The role of natural product chemistry in drug discovery. **J. Nat. Prod.**, v. 67, p. 2141-2153, 2004.

BUTLER, M. S. Natural products to drugs: natural product derived compounds in clinical trials. **Nat. Prod. Rep.**, v. 22, p. 162-195, 2005.

CÁCERES, A.; LÓPEZ B.; GONZÁLEZ, S.; BERGER, I.; TADA, I.; MAKI, J. Plants used in Guatemala for the treatment of protozoal infections. I. Screening of activity to bacteria, fungi and american trypanosomes of 13 native plants. **J. Ethnopharmacol.**, v. 62, p. 195-202, 1998.

CAHILL, M. G.; CAPRIOLI, G.; VITTORI, S.; JAMES, K. J. Elucidation of the mass fragmentation pathways of potato glycoalkaloids and aglycons using Orbitrap mass spectrometry. **J. Mass Spectrom.**, v. 45, p. 1019-1025, 2010.

CAPRA, F. **O ponto de mutação: a ciência, a sociedade e a cultura emergente.** São Paulo: Cultrix, 2007.

CASTRO, A.; MOCO, S.; COLL, J.; VERVOORT, J. LC-MS-SPE-NMR for the isolation and characterization of neo-clerodane diterpenoids from *Teucrium luteum* subsp. *flavovirens*. **J. Nat. Prod.**, v. 73, n. 5, p. 962-965, 2010.

CLAEYS, M.; Van den HEUVEL, H.; CHEN, S.; DERRICK, P. J.; MELLON, F., A.; PRICE, K. R. Comparison of high- and low-energy collision-induced dissociation tandem mass spectrometry in the analysis of glycoalkaloids and their aglycons. **J. Am. Soc. Mass Spectrom.**, v. 7, p. 173-181, 1996.

CLOAREC, O.; DUMAS, M. E.; CRAIG, A.; BARTON, R. H.; TRYGG, J.; HUDSON, J.; BLANCHER, C.; GAUGUIER, D.; LINDON, J. C.; HOLMES, E.; NICHOLSON, J. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic ¹H NMR data sets. **Anal. Chem.**, v. 77, n. 5, p. 1282-1289, 2005.

CROTTI, A. E. M.; VESSECCHI, R.; LOPES, J. L. C.; LOPES, N. P. Espectrometria de massas com ionização por “electrospray”: processos químicos envolvidos na formação de íons de substâncias orgânicas de baixo peso molecular. **Quim. Nova**, v. 29, n. 2, p. 287-292, 2006.

CUYCKENS, F.; CLAEYS, M. Determination of the glycosylation site in flavonoid mono-O-glycosides by collision-induced dissociation of electrospray-generated deprotonated and sodiated molecules. **J. Mass Spectrom.**, v. 40, n. 3, p. 364-372, 2005.

CUYCKENS, F.; MA, Y. L.; POCSFALVI, G.; CLAEYS, M. Tandem mass spectral strategies for the structural characterization of flavonoid glycosides. **Analisis**, v. 28, n. 10, p. 888-895, 2000.

D'ARCY, W. G. The classification of the Solanaceae. In: HAWKES, J. G.; LESTER, R. N.; SKELDING, A. D. **The biology and taxonomy of the Solanaceae.** London: Academic Press, 1979. p. 3-48.

DEWICK, P. M. **Medicinal natural products: a biosynthetic approach.** 2nd ed. Chichester: John Wiley & Sons, 2002.

DIXON, R. A.; GANG, D. R.; CHARLTON, A. J.; FIEHN, O.; KUIPER, H. A.; REYNOLDS, T. L.; TJEERDEMA, R. S.; JEFFERY, E. H.; GERMAN, J. B.; RIDLEY, W. P.; SEIBER, J. N. Perspective - applications of metabolomics in agriculture. **J. Agric. Food Chem.**, v. 54, p. 8984-8994, 2006.

EICH, E. Phenylalanine-derived metabolites/ phenylpropanoids. In: _____. **Solanaceae and convolvulaceae**: secondary metabolites. Berlin: Springer-Verlag, 2008. Chap. 6, p. 304-308.

EXARCHOU, V.; KRUCKER, M.; van BEEK, T. A.; VERVOORT, J.; GEROTHANASSIS, I. P.; ALBERT, K. LC-NMR coupling technology: recent advancements and applications in natural products analysis. **Magn. Res. Chem.**, v. 43, p. 681-687, 2005.

FIEHN, O.; KOPKA, J.; DORMANN, P.; ALTMANN, T.; TRETHERWEY, R. N.; WILLMITZER, L. Metabolite profiling for plant functional genomics. **Nat. Biotechnol.**, v. 18 p. 1157-1161, 2000.

FILLET, M.; FRÉDÉRICH, M. The emergence of metabolomics as a key discipline in the drug discovery process. **Drug Discov. Today: Technol.**, 2015. doi:10.1016/j.ddtec.2015.01.006. In press.

FUNARI, C. S.; CASTRO-GAMBOA, I.; CAVALHEIRO, A. J.; BOLZANI, V. da S. Metabolomics, an optimized approach for the rational exploitation of Brazilian biodiversity: state of the art, new scenarios, and challenges. **Quim. Nova**, v. 36, n. 10, p. 1605-1609, 2013.

FUNARI, C. S.; CARNEIRO, R. L.; ANDRADE, A. M.; HILDER, E. F.; CAVALHEIRO, A. J. Green chromatographic fingerprinting: an environmentally friendly approach for the development of separation methods for fingerprinting complex matrices. **J. Sep. Sci.**, v. 37, n. 1/2, p. 37-44, 2014.

GIKAA, H. G.; THEODORIDIS, G. A.; PLUMBC, R. S.; WILSONC, I. D. Current practice of liquid chromatography–mass spectrometry in metabolomics and metabonomics. **J. Pharm. Biomed. Anal.**, v. 87, p. 12-25, 2014.

GÓMEZ-ROMERO, M.; SEGURA-CARRETERO, A.; FERNÁNDEZ-GUTIÉRREZ, A. Metabolite profiling and quantification of phenolic compounds in methanol extracts of tomato fruit. **Phytochemistry**, v. 71, n. 16, p. 1848-1864, 2010.

GU, H.; NAGANA GOWDA, G. A.; CARNEVALE NETO, F.; OPP, M. R.; RAFTERY, D. RAMSY: ratio analysis of mass spectrometry to improve compound identification. **Anal. Chem.**, v. 85, n. 22, p. 10771-10779, 2013.

GUILLARME, D.; NGUYEN, D. T.; RUDAZ, S.; VEUTHEY, J. L. Method transfer for fast liquid chromatography in pharmaceutical analysis: application to short columns packed with small particle. Part II: gradient experiments. **Eur. J. Pharm. Biopharm.**, v. 68, n. 2, p. 430-440, 2008.

HALABALAKI, M.; VOUGOGIANNPOULOU, K.; MIKROS, E.; SKALTSOUNIS, A. L. Recent advances and new strategies in the NMR-based identification of natural products. **Curr. Opin. Biotech.**, v. 25, p. 1-7, 2014.

HALL, R. D. Plant metabolomics: from holistic hope, to hype, to hot topic. **New Phytol.**, v. 169, p. 453-468, 2006.

HEYMAN, H. M.; MEYER, J. J. M. NMR-based metabolomics as a quality control tool for herbal products. **S. Afr. J. Bot.**, v. 82, p. 21-32, 2012.

HUR, M.; CAMPBELL, A. A.; ALMEIDA-de-MACEDO, M.; LI, L.; RANSOM, N.; JOSE, A.; CRISPIN, M.; NIKOLAU, B. J.; WURTELE, E. S. A global approach to analysis and interpretation of metabolic data for plant natural product discovery. **Nat. Prod. Rep.**, v. 30, n. 4, p. 565-583, 2013.

IKEDA, T.; TSUMAGARI, H.; HONBU, T.; NOHARA, T. Cytotoxic activity of steroidal glycosides from *solanum* plants. **Biol. Pharm. Bull.**, v. 26, n. 8, p. 1198-1201, 2003.

JACQUES, R. S. **Caracterização química da erva mate (*ilex paraguariensis*): aplicação de diferentes processos de extração e influência das condições de plantio sobre a composição química.** 2005. 139 f. Tese (Doutorado em Química) - Instituto de Química, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.

JANDRIĆ, Z.; ROBERTS, D.; RATHOR, M. N.; ABRAHIM, A.; ISLAM, M.; CANNAVAN, A. Assessment of fruit juice authenticity using UPLC-QToF MS: a metabolomics approach. **Food Chem.**, v. 148, p. 7-17, 2014.

JOHANSEN, K. T.; WUBSHET, S. G.; NYBERG, N. T. HPLC–NMR revisited: using time-slice high-performance liquid chromatography–solid-phase extraction–nuclear magnetic resonance with database-assisted dereplication. **Anal. Chem.**, v. 85, n. 6, p. 3183-3189, 2013.

KIM, H. K.; VERPOORTE, R. Sample preparation for plant metabolomics. **Phytochem. Anal.**, v. 21, p. 4-13, 2010.

KIM, H. K.; CHOI, Y. H.; VERPOORTE, R. NMR-based metabolomic analysis of plants. **Nat. Protoc.**, v. 5, n. 3, p. 536-549, 2010.

KIND, T.; FIEHN, O. Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. **BMC Bioinformatics**, v. 8, n. 105, 2007. doi:10.1186/1471-2105-8-105.

KIND, T.; FIEHN, O. Advances in structure elucidation of small molecules using mass spectrometry. **Bioanal. Rev.**, v. 2, p. 23-60, 2010.

KIND, T.; WOHLGEMUTH, G.; LEE, Y. D.; LU, Y.; PALAZOGLU, M.; SHAHBAZ, S.; FIEHN, O. FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. **Anal. Chem.**, v. 81, n. 24, p. 10038-10048, 2009.

KOVACS, H.; MOSKAU, D.; SPRAUL, M. Cryogenically cooled probes - a leap in NMR technology. **Prog. Nucl. Mag. Res. Sp.**, v. 46, p. 131-155, 2005.

LANG, G.; MAYHUDIN, N. A.; MITOVA, M. I.; SUN, L.; van der SAR, S.; BLUNT, J. W.; COLE, A. L.; ELLIS, G.; LAATSCH, H.; MUNRO M. H. Evolving trends in the dereplication of natural product extracts: new methodology for rapid, small-scale investigation of natural product extracts. **J. Nat. Prod.**, v. 71, n. 9, p. 1595-1599, 2008.

LARIVE, C. K.; BARDING, G. A. Jr.; DINGES, M. M. NMR spectroscopy for metabolomics and metabolic profiling. **Anal. Chem.**, v. 87, p. 133-146, 2015.

LEVESQUE, M. P.; BENFEY, P. N. Systems biology. **Curr. Biol.**, v. 14, n. 5, p. 179-180, 2004.

LEVITT, L. H. Nuclear magnetism. In: _____. **Spin dynamics**: basics of nuclear magnetic resonance. 2nd ed. Hoboken: John Wiley & Sons, 2008. Chap. 1, p. 3-59.

LI, J. W.-H.; VEDERAS, J. C. Drug discovery and natural products: end of an era or an endless frontier? **Science**, v. 325, n. 5937, p. 161-165, 2009.

LI, J.; LI, Q.; FENG, T.; LI, K. Aqueous extract of *Solanum nigrum* inhibit growth of cervical carcinoma (U14) via modulating immune response of tumor bearing mice and inducing apoptosis of tumor cells. **Fitoterapia**, v. 79, p. 548-556, 2008.

LISEC, J.; SCHAUER, N.; KOPKA, J.; WILLMITZER, L.; FERNIE, A. R. Gas chromatography mass spectrometry-based metabolite profiling in plants. **Nat. Protoc.**, v. 1, p. 387-396, 2006.

LOENING, N. M.; KEELER, J.; MORRIS, G. A. One-dimensional DOSY. **J. Magn. Res.**, v. 153, p. 103-112, 2001.

LUBBE, A.; VERPOORTE, R. Cultivation of medicinal and aromatic plants for specialty industrial materials. **Ind. Crop. Prod.**, v. 34, p. 785-801, 2011.

MAHROUS, E. A.; FARAG, M. A. Two dimensional NMR spectroscopic approaches for exploring plant metabolome: a review. **J. Advanc. Res.**, v. 6, p. 3-15, 2015.

MAKAROV, A.; DENISOV, E.; KHOLOMEEV, A.; BAISCHUN, W.; LANGE, O.; STRUPAT, K.; HORNING, S. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. **Anal. Chem.**, v. 78, p. 2113-2120, 2006.

MARCH, R. E.; LEWARS, E. G.; STADEY C. J., MIAO X.-S.; ZHAO, X.; METCALFE C. D. A comparison of flavonoid glycosides by electrospray tandem mass spectrometry International. **J. Mass Spec.**, v. 248, p. 61-85, 2006.

MARKHAM, K. R.; GEIGER, H. ^1H nuclear magnetic resonance spectroscopy of flavonoids and their glycosides in hexadeuterodimethylsulfoxide. In: HARBONE, J. B. **The flavonoids advances in research since 1986**. London: Chapman & Hall, 1993. Chap. 10, p. 443-510.

MESIA-VELA, S.; SANTOS, M. T.; SOUCCAR, C.; LIMA-LANDMAN, M. T. R.; LAPA, A. J. *Solanum paniculatum* L. (Jurubeba): potent inhibitor of gastric acid secretion in mice. **Phytomedicine**, v. 9, p. 508-514, 2002.

MILNER, S. E.; BRUNTON, N. P.; JONES, P. W.; O' BRIEN, N. M.; COLLINS, S. G.; MAGUIRE, A. R. Bioactivities of glycoalkaloids and their aglycones from *Solanum* species. **Agric. Food Chem.**, v. 59, p. 3454-3484, 2011.

MUSHTAQ, M. Y.; CHOI, Y. H.; VERPOORTE, R.; WILSON, E. G. Extraction for metabolomics: access to the metabolome. **Phytochem. Anal.**, v. 25, p. 291-306, 2014.

NAKANISH, K. **One-dimensional and two-dimensional NMR spectra by modern pulse techniques**. Tokyo: University Science Books, 1990.

NEUMANN, S.; BÖCKER, S. Computational mass spectrometry for metabolomics: identification of metabolites and small molecules. **Anal. Bioanal. Chem.**, v. 398, p. 2779-2788, 2010.

NEWMAN, D. J.; CRAGG, G. M. Natural products as sources of new drugs over the last 25 years. **J. Nat. Prod.**, v. 70, p. 461-477, 2007.

NEWMAN, D. J.; CRAGG, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. **J. Nat. Prod.**, v. 75, n. 3, p. 311-335, 2012.

NEWMAN, D. J.; GIDDINGS, L-A. Natural products as leads to antitumor drugs. **Phytochem. Rev.**, v. 13, p. 123-137, 2014.

NEWMAN, D. J.; CRAGG, G. M.; SNADER, K. M. The influence of natural products upon drug discovery. **Nat. Prod. Rep.**, v. 17, n. 3, p. 215-234, 2000.

NGUELEFACK, T. B.; FEUMEBO, C. B.; ATEUFACK, G.; WATCHO, P.; TATSIMO, S.; ATSAMO, A. D.; TANE, P.; KAMANYI, A. Anti-ulcerogenic properties of the aqueous and methanol extracts from the leaves of *Solanum torvum* Swartz (Solanaceae) in rats. **J. Ethnopharmacol.**, v. 119, p. 135-140, 2008.

NICHOLSON, J. K.; LINDON, J. C. Systems biology: metabonomics. **Nature**, v. 455, n. 23, p. 1054-1056, 2008.

NIELSEN, J. H.; JEWETT, M. C. (Ed.). **Metabolomics: a powerful tool in systems biology**. Berlin: Springer, 2007.

OLMSTEAD, R. G.; BOHS, L.; MIGID, H. A.; SANTIAGO-VALENTIN, E.; GARCIA, V. F.; COLLIER, S. M. A molecular phylogeny of the Solanaceae. **Taxon**, v. 57, n. 4, p. 1159-1181, 2008.

PARK, S.; JEONG, W. Y.; LEE, J. H.; KIM, Y. H.; JEONG, S. W.; KIM, G.-S.; BAE, D. W.; LIM, C.-S.; JIN, J. S.; LEE, S. J.; SHIN, S. C. Determination of polyphenol levels variation in *Capsicum annuum* L. cv. Chelsea (yellow bell pepper) infected by anthracnose (*Colletotrichum gloeosporioides*) using liquid chromatography–tandem mass spectrometry. **Food Chem.**, v. 30, p. 981-985, 2012.

PARR, A. J.; MELLON, F. A.; COLQUHOUN, I. J.; DAVIES, H. V. Dihydrocaffeoyl polyamines (kukoamine and allies) in potato (*Solanum tuberosum*) tubers detected during metabolite profiling. **Agric. Food Chem.**, v. 53, p. 5461-5466, 2005.

PATTI, G. J.; YANES, O.; SIUZDAK, G. Innovation: metabolomics: the apogee of the omics trilogy. **Nat. Rev. Mol. Cell.**, v. 13, n. 4, p. 263-269, 2012.

PEREZ, A. C.; FRANCA, V.; DALDEGAN, V. M. Jr.; DUARTE, I. D. G. Effect of *Solanum lycocarpum* St. Hill on various haematological parameters in diabetic rats. **J. Ethnopharmacol.**, v. 106, p. 442-444, 2006.

PETERSON, A. C.; BALLOON, A. J.; WESTPHALL, M. S.; COON, J. J. Development of a CG/quadrupole-orbitrap mass spectrometer, part II: new approaches for discovery metabolomics. **Anal. Chem.**, v. 86, p. 10044-10051, 2014.

PINTO, A. C.; SILVA, D. H. S.; BOLZANI, V. da S.; LOPES, N. P.; EPIFANIO, R. de A. Produtos naturais: atualidade, desafios e perspectivas. **Quim. Nova**, v. 25, n. 1, p. 45-61, 2002.

POCZAI, P.; TALLER, J.; SZABÓ, I. Analysis of phylogenetic relationships in the genus *Solanum* (Solanaceae) as revealed by RAPD markers. **Plant. Syst. Evol.**, v. 275, p. 59-67, 2008.

POWERS, R. NMR metabolomics and drug discovery. **Magn. Reson. Chem.**, v. 47, p. S2-S11, 2009.

PROUDFOOT, J. R. Drugs, leads, and drug-likeness: an analysis of some recently launched drugs. **Bioorg. Med. Chem. Lett.**, v. 12, n. 12, p. 1647-1650, 2002.

PUPO, M. T.; GALLO, M. B. C.; VIEIRA, P. C. Biologia química: uma estratégia moderna para a pesquisa em produtos naturais. **Quim. Nova**, v. 30, n. 6, p. 1446-1455, 2007.

RATCLIFFE, R. G.; SHACHAR-HILL, Y. Revealing metabolic phenotypes in plants: inputs from NMR analysis. **Biol. Rev.**, v. 80, p. 27-43, 2005.

REN, Y.; SHEN, L.; ZHANG, D. W.; DAI, S. J. Two new sesquiterpenoids from *Solanum lyratum* with cytotoxic activities. **Chem. Pharm. Bull.**, v. 57, n. 4, p. 408-410, 2009.

ROUCHI, A. M. Rediscovering natural products. **Chem. Eng. News**, v. 13, p. 77-91, 2003.

SILVA, T. M. S.; CARVALHO, M. G.; BRAZ FILHO, R.; AGRA, M. F. Ocorrência de flavonas, flavonóis e seus glicosídeos em espécies do gênero *Solanum* (solanaceae). **Quim. Nova**, v. 26, p. 517-522, 2003.

SILVA, L. da; GODEJOHANN, M.; MARTIN, F.-P. J.; COLLINO, S.; BÜRKLE, A.; MORENO-VILLANUEVA, M.; BERNHARDT, J.; TOUSSAINT, O.; GRUBECK-LOEBENSTEIN, B.; GONOS, E. S.; SIKORA, E.; GRUNE, T.; BREUSING, N.; FRANCESCHI, C.; HERVONEN, A.; SPRAUL, M.; MOCO, S. High-resolution quantitative metabolome analysis of urine by automated flow injection NMR. **Anal. Chem.**, v. 85, n. 12, p. 5801-5809, 2013.

SILVEIRA, T. M. L. da. **Aminas bioativas livres e conjugadas no café solúvel: metodologia de análise e influência do processamento**. 2008. 179 f. Tese (Doutorado em Química) - Faculdade de Farmácia, Universidade Federal de Minas Gerais, Belo Horizonte, 2008

SILVERSTEIN, R. M.; BASSLER, G. C.; MORRILL, T. C. **Spectroscopic identification of organic compounds**. 7th ed. New York: Wiley, 2005.

SUN, L. X.; FU, W. W.; REN, J.; XU, L.; BI, K. S.; WANG, M. W. Cytotoxic constituents from *Solanum lyratum*. **Arch. Pharm. Res.**, v. 29, n. 2, p. 135-139, 2006.

SUTHAR, A. C.; MULANI, R. M. A high performance thin layer chromatography method for quantitative estimation of Diosgenin in *Solanum Nigrum*. **Pharmacogn. Mag.**, v. 4, n. 14, p. 112-115, 2008.

TANG, H.; XIAO, C.; WANG, Y. Important roles of the hyphenated HPLC-DAD-MS-SPE-NMR technique in metabonomics. **Magn. Reson. Chem.**, v. 47, p. S157-S162, 2009.

TENG, G. **Structural biology: practical NMR applications**. 2nd ed. New York: Springer, 2013.

ULRICH-MEZENICH, G.; ZEITLER, H.; JOBST, D.; PANEK, D.; VETTER, H.; WAGNER, H. Application of the "-Omic-" technologies in phytomedicine. **Phytomedicine**, v. 14, p. 70-82, 2007.

VERHOEVEN H. A.; de VOS, R. C. H.; BINO R. J.; HALL R. D. Plant metabolomics strategies based upon quadrupole time of flight mass spectrometry (QToF-MS). In: SAITO, K.; DIXON, R. A.; WILLMITZER, L. (Ed.). **Plant metabolomics**. Berlin; New York: Springer, 2006. (Biotechnology in agriculture and forestry, 57).

VERPOORTE, R.; CHOI, Y.; KIM, H. K. Ethnopharmacology and systems biology: a perfect holistic match. **J. Ethnopharmacol.**, v. 100, n. 1, p. 53-56, 2005.

VERPOORTE, R.; CHOI, Y. H.; KIM, H. K. NMR-based metabolomics at work in phytochemistry. **Phytochem. Rev.**, v. 6, p. 3-14, 2007.

VERPOORTE, R.; CHOI, Y. H.; KIM, H. K. Metabolomics: will it stay? **Phytochem. Anal.**, v. 21, p. 2-3, 2010.

VIZZOTO, L. **Estudo e aplicações das técnicas de RMN HR-MAS e DOSY em plantas e extratos de plantas da ordem rutales**. 2004. 107 f. Tese (Doutorado em Ciências) - Centro de Ciências Exatas e de Tecnologia, Universidade Federal de São Carlos, São Carlos, 2004.

VUKICS, V.; GUTTMAN, A. Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. **Mass Spec. Rev.**, v. 29, n. 1, p. 1-16, 2010.

WARIDEL, P.; WOLFENDER, J.-L.; NDJOKO, K.; HOBBY, K. R.; MAJOR, H. J.; HOSTETTMANN, K. Evaluation of quadrupole time-of-flight tandem mass spectrometry and ion-trap multiple-stage mass spectrometry for the differentiation of C-glycosidic flavonoid isomers. **J. Chrom. A**, v. 926, p. 29-41, 2001.

WATROUS, J.; ROACH, P.; ALEXANDROV, T.; HEATH, B. S.; YANG, J.; KERSTEN, R. D.; van der VOORT, M.; POGLIANO, K.; GROSS, H.; RAAIJMAKERS, J.; MOORE, B. S.; LASKIN, J.; BANDEIRA, N.; DORRESTEIN, P. C. Mass spectral molecular networking of living microbial colonies. **Proc. Natl. Acad. Sci. U.S.A.**, v. 109, p. 1743-1752, 2012.

WEISSENBERG, M. Isolation of solasodine and other steroidal alkaloids and sapogenins by direct hydrolysis-extraction of *Solanum* plants or glycosides therefrom. **Phytochemistry**, v. 58, p. 501-508, 2001.

WENG, J.-K. The evolutionary paths towards complexity: a metabolic perspective. **New Phytol.**, v. 201, p. 1141-1149, 2014.

WHALEN, M. D.; MABRY, T. J. New 8-hydroxyflavonoids from *Solanum* section *Androceras*. **Phytochemistry**, v. 18, p. 263-265, 1979.

WOLF, S.; SCHMIDT, S.; MÜLLER-HANNEMANN, M.; NEUMANN, S. In silico fragmentation for computer assisted identification of metabolite mass spectra. **BMC Bioinformatics**, v. 11, 2010. doi:10.1186/1471-2105-11-148.

WOLFENDER, J.-L. HPLC in natural product analysis: the detection issue. **Planta Med.**, v. 75, n. 7, p. 719-734, 2009.

WOLFENDER, J.-L.; RUDAZ, S.; CHOI, Y. H.; KIM, H. K. Plant metabolomics: from holistic data to relevant biomarkers. **Curr. Med. Chem.**, v. 20, n. 8, p. 1056-1090, 2013.

YANG, J. Y.; SANCHEZ, L. M.; RATH, C. M.; LIU, X.; BOUDREAU, P. D.; BRUNS, N.; GLUKHOV, E.; WODTKE, A.; FELICIO, R. de; FENNER, A.; WONG, W. R.; LININGTON, R. G.; ZHANG, L.; DEBONSI, H. M.; GERWICK, W. H.; DORRESTEIN, P. C. Molecular networking as a dereplication strategy. **J. Nat. Prod.**, v. 76, p. 1686-1699, 2013.

YANG, W. Z.; QIAO, X.; BO, T.; WANG, Q.; GUO, D. A.; YE, M. Low energy induced homolytic fragmentation of flavonol 3-O-glycosides by negative electrospray ionization tandem mass spectrometry. **Rapid Commun. Mass Spectrom.**, v. 28, n. 4, p. 385-395, 2014.

YANG, Z.; NAKABAYASHI, R.; OKAZAKI, Y.; MORI, T.; TAKAMATSU, S.; KITANAKA, S.; KIKUCHI, J.; SAITO, K. Toward better annotation in plant metabolomics: isolation and structure elucidation of 36 specialized metabolites from *Oryza sativa* (rice) by using MS/MS and NMR analyses. **Metabolomics**, v. 10, p. 543-555, 2014.

YULIANA, N. D.; KHATIB, A.; CHOI, Y. H.; VERPOORTE, R. Metabolomics for bioactivity assessment of natural products. **Phytother. Res.**, v. 25, n. 2, p. 157-169, 2011.

ZHANG, D. W.; LI, G. H.; YU, Q. Y.; DAI, S. J. New anti-inflammatory 4-hydroxyisoflavans from *Solanum lyratum*. **Chem. Pharm. Bull.**, v. 58, n. 6, p. 840-842, 2010.