Universidade Estadual de Campinas
Instituto de Química
Departamento de Química Orgânica

Tese de Doutorado

Petroleômica por espectrometria de massas de altíssima resolução e exatidão: informação química precisa e seletiva através de novas estratégias de processamento e visualização de dados

Yuri Eberlim de Corilo.

Orientador: Prof. Dr. Marcos Nogueira Eberlin.

Campinas
17 de Fevereiro de 2011

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO INSTITUTO DE QUÍMICA DA UNICAMP

C813p

Corilo, Yuri Eberlim de.
Petroleômica por espectrometria de massas de altíssima resolução e exatidão: informação química precisa e seletiva através de novas estratégias de processamento e visualização de dados / Yuri Eberlim de Corilo. -- Campinas, SP: [s.n], 2011.

Orientador: Marcos Nogueira Eberlin.

Doutorado - Universidade Estadual de Campinas, Instituto de Química.

1. Petroleômica. 2. Espectrometria de massas. 3. Ressonância ciclotrônica de íons. I. Eberlin, Marcos Nogueira. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

Título em inglês: Petroleomics by ultra-high resolving power and accuracy mass spectrometry: selective and accurate chemical information by new procedures of data processing

Palavras-chaves em inglês: Petroleomics, Mass spectrometry, Ion cyclotron resonance

Área de concentração: Química Orgânica

Titulação: Doutor em Ciências

Banca examinadora: Prof. Dr. Marcos Nogueira Eberlin (orientador), Prof. Dr. Luiz Alberto Beraldo Moraes (DQ-USP-Ribeirão Preto), Dr. Eugenio Vaz dos Santos Neto (CENPES-Geoquímica-Petrobras), Profa. Dra. Carol Hollingworth Collins (IQ-UNICAMP), Prof. Dr. Sebastião Ferreira Fonseca (IQ-UNICAMP)

Data de defesa: 17/02/2011

Agradecimentos

A todos aqueles que acreditavam na realização final deste trabalho e que me deram forças para continuar frente a tantas dificuldades e obstáculos. A fidelidade e cumplicidade da minha esposa que me apoiou e me fortaleceu durante todo o período de desenvolvimento, pois sem ela nada disso teria sido alcançado. A todos do grupo de pesquisa do Laboratório ThoMSon de Espectrometria de Massas, que direta ou indiretamente colaboraram com este trabalho.

Agradeço ao Professor Marcos N. Eberlin pela oportunidade de fazer parte de seu grupo de pesquisa durante todos esses anos e pela confiança depositada em minha pessoa para a realização deste e de outros trabalhos. E finalmente agradeço a ANP pelo apoio financeiro concedido através da empresa Petrobrás e da rede Geoquímica.

Curriculum Vitae

Formação Acadêmica (Graduação e Pós-Graduação)

Bacharelado em Química pela Universidade Estadual de Maringá com início em Agosto de 2002 e término em Fevereiro de 2006.

Mestrado em Química Orgânica pela Universidade Estadual de Campinas com início em Agosto de 2006 e término em Março de 2009

Doutorado em Química pela Universidade Estadual de Campinas com início em Março de 2009 e término em Fevereiro de 2011.

Produção científica

Trabalho em Congressos:

1 - Corilo, Yuri E.; Catharino, Rodrigo R.; Abdelnur, Patricia V.; Eberlin, Marcos N. **Reactive Fused Droplet Electrospray Ionization Mass Spectrometry.** Oral talk, 55th Conference on Mass Spectrometry & Allied Topics, Indianapolis, *IN,* June 3-7 (2007).

2 - Corilo, Yuri E.; Eberlin, Marcos N. **Elucidação Estrutural de Ésteres Borônicos através de reações Íon/Molécula seletivas**. Poster, 30th RASBQ, Águas de Lindóia SP, Brazil, June 31-03 (2007).

3 **-** Corilo, Yuri E.; Eberlin, Marcos N. **Distinction of posicional isomers by Ion/Molecule reaction.** 2nd Brazilian Society for Mass Spectrometry (BrMASS) Conference, Poster, Campinas, SP, Brazil, December 09-11 (2007).

4 - Vaz, Boniek G.; Corilo, Yuri E.; Eberlin, Marcos N. **The Fingerprint Reaction: Rapid Derivatization of beer carbonyl compounds and detection by ESI-MS.** 2nd Brazilian Society for Mass Spectrometry (BrMASS) Conference, Poster, Campinas, SP, Brazil, December 09-11 (2007).

5 - Campo, F. R.; Machado, M. W.; Neto, C.S.; Salgado, J.; Zaffari, G.; Biavatti, M. W.; Barison, A.; Corilo, Y. E.; Eberlin, M. N. **Alkaloids investigation on callus of Passiflora alata by HPLC and ESI(+)-MS/MS.** 2nd Brazilian Society for Mass Spectrometry (BrMASS) Conference, Poster, Campinas, SP, Brazil, December 09-11 (2007).

6 - Corilo, Yuri E.; Eberlin, Marcos N. **Recognizing α, β or δ-Substitution in Pyridines by Mass Spectrometry**. Poster , 56th ASMS Conf. on Mass Spectrometry & Allied Topics, Denver, *CO,* June 1-5 (2008).

7 - Romão, Wanderson; Franco, Marcos F.; Corilo, Yuri E.; Eberlin, Marcos. N.; De Paoli,Marco A. **Poly (ethylene terephthalate) thermo-mechanical and thermo-oxidative degradation mechanisms**. Poster, The Polymer Processing Society 24th Annual Meeting, Salerno, SA, Italy, June 15-19 (2008).

8 - Nachtigall, Fabiane M.; Corilo, Yuri E.; Abdelnur, Patricia V.; Eberlin, Marcos N.; Dupont, Jairton. **Charged Carbenes**. Poster, 57th ASMS Conference on Mass Spectrometry & Allied Topics, Philadelphia, PA, May 31 - June 4 (2009).

9 - Corilo, Yuri E.; Vaz, Boniek G.; Eberlin, Marcos N. **Petroleomics by EASI-MS**. Poster, 7th North American *FT MS Conference*. Key West, FL, April 19-23 (2009).

10 - Corilo, Yuri E.; Eberlin, Marcos N**. Charged (di)-radicals and Charged (di)-arynes: Formation and characterization of unprecedented species in the gas phase.** Poster, 18th International Mass Spectrometry Conference. Bremen, Germany, Aug 30 – Sep 4 (2009).

11 - Corilo, Yuri E.; Vaz, Boniek G.; Simas, Rosineide C.; Lopes, Heliara D.; Klitzke, Clécio F.; Pereira, Rosana C. L.; Bastos, Wagner L.; Eberlin, Marcos N. **PetroMassSpec: Simplifying Petroleomic Data Analysis via Software Processing and Graphics.** Invited Talk, 3rd Brazilian Society for Mass Spectrometry (BrMASS) Conference, Campinas, SP - Brazil. December 12-15 (2009).

12 - Vaz, Boniek G.; Corilo, Yuri E. ; Simas, Rosineide C.; Lopes, Heliara D.; Klitzke, Clécio F.; Pereira, Rosana C. L.; Bastos, Wagner L.; Eberlin, Marcos N. **Petroleomic by Easy Ambient Sonic Spray ionization Mass Spectrometry (EASI-MS).** Invited Talk, 3rd Brazilian Society for Mass Spectrometry (BrMASS) Conference, Campinas, SP-Brazil. December 12-15 (2009).

13 - Klitzke, Clécio F.; Corilo, Yuri E.; Lopes, Heliara D.; Pereira Vaz, Boniek G.; Simas, Rosineide C.; Pereira, Rosana C. L.; Bastos, Wagner L.; Eberlin, Marcos N. **Petroleomic via FAB-MS versus ESI FT-ICR MS: Characterization of Naphthenic Acids from Brazilian Crude Oil Samples;** Poster, 58th ASMS Conference on Mass Spectrometry & Allied Topics, Salt Lake City, UT, May 23 - 27 (2010).

14 – Lopes, Heliara D.; Vaz, Boniek G.; Corilo, Yuri E. ; Moraes, Reinaldo J. R.; Sant'Ana, Antônio E. G.; Reboucas, Lucia M.C.; Eberlin Marcos N.; **Naphthenic Acids Profiles via ESI-FT ICR MS as Markers of Crude Oil Sources**; Poster, 58th ASMS Conference on Mass Spectrometry & Allied Topics, Salt Lake City, UT, May 23 - 27 (2010).

15 - Vaz, Boniek G.; Corilo, Yuri E. ; Simas, Rosineide C.; Lopes, Heliara D.; Klitzke, Clécio F.; Pereira, Rosana C. L.; Bastos, Wagner L.; Eberlin, Marcos N. **Petroleomics by EASI(±) FT-ICR MS**; Poster, 58th ASMS Conference on Mass Spectrometry & Allied Topics, Salt Lake City, UT, May 23 - 27 (2010).

16 - Simas, Rosineide C.; Lopes, Heliara D.; Vaz, Boniek G.; Klitzke, Clécio F.; Corilo, Yuri E.; Pereira, Rosana C. L.; Bastos, Wagner L.; Eberlin, Marcos N. **Fatorial Design and Optimization of the Petroleomic MS Analysis: Accessing Data for Precision Evaluation**. Poster, 58th ASMS Conference on Mass Spectrometry & Allied Topics, Salt Lake City, UT, May 23 - 27 (2010).

17 – Lalli, Priscila M.; Corilo, Yuri E.; Sa, Gilberto; Daroda, Romeu J.; Souza, Vanderlea; Ebeling, Gunter; Dupont, Jairton; Eberlin, Marcos N. **Intrinsic Mobility of Gaseous Cationic and Anionic Aggregates of Ionic Liquids by Ion Mobility Mass**

Spectrometry; Poster, 58th ASMS Conference on Mass Spectrometry & Allied Topics, Salt Lake City, UT, May 23 - 27 (2010).

18 - Corilo, Yuri E.; Vaz, Boniek G.; Simas, Rosineide C.; Lopes, Heliara D.; Klitzke, Clécio F.; Pereira, Rosana C. L.; Bastos, Wagner L.; Eberlin, Marcos N. **PetroMassSpec: Simplifying Petroleomic Data Analysis via Software Processing and Graphics**. Poster, 58th ASMS Conference on Mass Spectrometry & Allied Topics, Salt Lake City, UT, May 23 - 27 (2010).

19 – Lemões, J. S.; Monteiro, A. P. L.; Farias, S.P.; Viêgas, C.V.; D´Oca, M. G. M.; Klitzke, Clécio F.; Corilo, Yuri .E. **Determinação do perfil graxo de Chlorella pyrenoidosa por GC-FID e ICR-MS.** Pôster, 33th RASBQ, Águas de Lindóia SP, Brazil, May 28-31 (2010).

20 - Rebouças Lucia MC, Rita CC Silva Reinaldo JR Moraes, Agciel B Alves, Boniek G Vaz, Yuri E Corilo, Heliara D L Nascimento, Clécio F Klitzke, Marcos N Eberlin, Antonio E G. **FT-ICR MS STUDY OF NAPHTHENIC ACIDS IN CRUDE OILS FROM SUB-BASIN ALAGOAS**. Poster, Rio Oil & Gas Expo and Conference , Rio de Janeiro, RJ Brazil, Sep 13- 16(2010).

21 - Rebouças, Lucia M.C.; Silva, Rita C.C.; Moraes, Reinaldo J.R.; Alvez, Agciel B.; Vaz, Boniek G.; Corilo, Yuri E.; Nascimento, Heliara D. L.; Klitzke, Clécio F.; Simas, Rosineide C.; Eberlin, Marcos N.; Sant´Ana, Antonio E.G. **Asphaltene present in crude oils of Alagoas sub-basin studied by ESI FT-ICR MS.** Poster, Petrophase XI, Jersey City, NJ, EUA, June 13-17 (2010).

22 - Eberlin, Marcos N.; Corilo, Yuri E.; Klitzke, Clecio F.; Lalli, Priscila M. **Petroleomics via Traveling wave ion mobility mass spectrometry using polarizable drift gases**. PacifiChem 2010, Honolulu, Hawai, EUA December 15-20 (2010)

Publicações

Artigos publicados

1 - Corilo, Yuri E.; Vaz, Boniek G.; Simas, Rosineide C.; Lopes Nascimento, Heliara D.; Klitzke, Clecio F.; Pereira, Rosana C. L.; Bastos, Wagner L.; Santos Neto, Eugenio V.; Rodgers, Ryan P.; Eberlin, Marcos N. **Petroleomics by EASI FT-ICR MS.** Analytical Chemistry (2010), 82(10), 3990-3996 – CAPA.

2 - Vijayakrishnan, B.; Issaree A.; Corilo Y. E.; Ferreira C R.; Eberlin, M. N. Eberlin, Peter, M. G. **MS^n of the six isomers of (GlcN)2(GlcNAc)2 aminoglucan tetrasaccharides (diacetylchitotetraoses): Rules of fragmentation for the sodiated molecules and application to sequence analysis of hetero-chitooligosaccharides.** *Carbohydrate Polymers* (2011), *84*, 713-726

3 - Lalli, Priscila M.; Corilo, Yuri E.; Abdelnur, Patricia V.; Eberlin, Marcos N.; Laali, Kenneth K. **Intrinsic acidity and electrophilicity of gaseous propargyl/allenyl carbocations.** Organic & Biomolecular Chemistry (2010), 8(11), 2580-2585

4 - Sanvido, Gustavo B.; Garcia, Jerusa S.; Corilo, Yuri E.; **Fast Screening and Secure Confirmation of Milk Powder Adulteration with Maltodextrin via Electrospray Ionization Mass Spectrometry [ESI(+)-MS] and Selective Enzymatic Hydrolysis.** Journal of Agricultural and food chemistry (2010), 58(17), 9407-9412

5 - Albores, Pablo; Slep, Leonardo D.; Eberlin, Livia S.; Corilo, Yuri E.; Eberlin, Marcos N.; Benitez, Guillermo; Vela, Maria E.; Salvarezza, Roberto C.; Baraldo, Luis M. **From Monomers to Geometry-Constrained Molecules: One Step Further Toward Cyanide Bridged Wires.** Inorganic Chemistry (2009), 48(23), 11226-11235

6 - Romao, Wanderson; Franco, Marcos F.; Corilo, Yuri E.; Eberlin, Marcos N.; Spinace, Marcia A. S.; De Paoli, Marco-A. **Poly (ethylene terephthalate) thermo-mechanical and thermo-oxidative degradation mechanisms.** Polymer Degradation and Stability (2009), 94(10), 1849-1859

7 - Wu, Lianming; Eberlin, Marcos N.; Corilo, Yuri E.; Liu, David Q.; Yin, Hao. **Dimerization of ionized 4-(methyl mercapto)-phenol during ESI, APCI and APPI mass spectrometry.** Journal of Mass Spectrometry (2009), 44(9), 1389-1394

8 - Buskuhl, Humberto; Freitas, Rilton A.; Delle Monache, Franco; Barison, Andersson; Campos, Francinete R.; Corilo, Yuri E.; Eberlin, Marcos N.; Biavatti, Maique W. **A new polyacetylene from Vernonia scorpioides (Lam.) Pers. (Asteraceae) and its in vitro antitumoral activity.** Journal of the Brazilian Chemical Society (2009), 20(7), 1327-1333

9 - Benassi, Mario; Corilo, Yuri E.; Uria, Diana; Augusti, Rodinei; Eberlin, Marcos N. **Recognition and Resolution of Isomeric Alkyl Anilines by Mass Spectrometry.** Journal of the American Society for Mass Spectrometry (2009), 20(2), 269-277

10 - Corilo, Yuri E.; Eberlin, Marcos N. **Recognizing -α, -β or -δ substitution in pyridines by mass spectrometry.** Journal of Mass Spectrometry (2008), 43(12), 1636-1640

11 - Buskuhl, Humberto; Freitas, Rilton A.; Delle Monache, Franco; Barison, Andersson; Campos, Francinete R.; Corilo, Yuri E.; Eberlin, Marcos N.; Biavatti, Maique W. **Sesquiterpene lactones from Vernonia scorpioides and their in vitro cytotoxicity.** Phytochemistry (2010), *71*, 1539-1544

12 - Nachtigall, Fabiane M.; Corilo, Yuri E.; Cassol, Claudia C.; Ebeling, Guenter; Morgon, Nelson H.; Dupont, Jairton; Eberlin, Marcos N. **Multiply charged (di-)radicals.** Angewandte Chemie, International Edition (2008), 47(1), 151-154

Cursos Realizados:

LTQ FT Ultra Operator Training – Carga horária de 24 horas - Bremen – Alemanha – de 12 à 14 de novembro de 2007 .

"Petroleomics by Mass Spectrometry: MS Returns to Its Roots" – 3° Congresso Brasileiro de Espectrometria de Massas – Carga horária de 8 horas. Campinas – SP.

Cursos Proferidos:

"Petroleômica por Espectrometria de Massas" – Universidade Federal do Alagoas – UFAL, Instituto de Química e Biotecnologia – IQB, 14 de junho de 2010 – Carga horária de 8 horas.

Resumo

Amostras de petróleos bruto e derivados podem ser caracterizadas com altíssima rapidez e eficiência através da técnica denominada petroleômica por espectrometria de massas por ressonância ciclotrônica de íons e transformada de Fourier (FT-ICR MS). Milhares de marcadores polares são resolvidos e identificados em questão de 6 ~ 7 minutos com muita agilidade e confiabilidade. A técnica produz, porém, uma enormidade de dados que precisam ser tratados com eficiência, através de um *software* que combine estratégias eficientes de programação e inteligência química no processamento e visualização. Neste trabalho, toda uma estratégia já aprovada em testes preliminares foi desenvolvida com esta finalidade. Esta estratégia levou a atribuição da composição elementar inequívoca para todos os sinais adquiridos. As etapas envolveram: a atribuição da *m/z* para cada sinal do espectro, a atribuição do limiar ótimo de intensidade para o ruído de cada espectro individualmente, a construção de um banco de dados teórico de correlação massa e fórmula molecular para os possíveis constituintes polares do petróleo, a atribuição da composição elementar para cada sinal através da comparação com este banco teórico, a eliminação de duplicidade de atribuições através da confirmação do padrão isotópico e comparação com séries homólogas, a organização dos dados por classes (heteroátomos) e tipo (*Double bond equivalent - DBE*) e a construção de bibliotecas gráficas para a visualização e interpretação das atribuições (Diagramas de Kendrick e van Krevelen). A combinação destes algoritmos e a execução automática deles formam o *software* de altíssima eficiência e alta tecnologia para tratamento e visualização de dados de petroleômica por FT- ICR MS. Este *software* foi e está sendo utilizado pelo grupo de pesquisa como ferramenta para diferentes linhas de pesquisa em petroleômica por FT-ICR MS aplicada a petróleo e seus derivados, entre elas: a) acompanhamento do processo de eliminação de compostos polares de diesel através de hidrotratamento e oxitratamento, b) atribuição e detalhamento de ácidos naftênicos presentes no petróleo, a sua correlação com índice de corrosão associado, c) o acompanhamento rápido para processos visando sua remoção e d) análises de frações de asfaltenos e a identificação de contaminantes.

Abstract

Crude oil samples and their derivatives can be characterized with ultra high speed and efficiency through the technique called petroleomics by Fourier transform ion cyclotron resonance mass spectrometry (FT ICR-MS). Thousands of polar markers are resolved and identified in ~ 6-7 minutes with great reliability and agility. This technique provides a large amount of data that requires efficient procedures for data processing. This may be achieved with software that combines efficient programming strategies and chemical intelligence in processing and graphical displays. In this work, all strategies already approved by preliminary data were applied with this goal. Through these strategies all unequivocal elementary compositions were achieved for all resolved signals in the spectra. The steps performed were: attribution of m/z value for each signal; attribution of the best intensity threshold for noise for each individual spectrum; theoretical data base elaboration with mass and molecular formula for all possible polar compounds present in petroleum; elemental composition assignments for each signal through data base comparison; duplicated assignment elimination by isotopic pattern confirmation and homologous series comparison; data organization by heteroatom classes and type (DBE); and graphical library elaboration needed to assist assignment interpretation (Kendrick and van Krevelen Diagrams). The combination of all these algorithms with automatic execution establishes highly efficient software and highly technological petroleomics data processing by FT- ICR MS. This software is being used by our research group as a study tool in different research areas and for different goals, specifically in Petroleomics by FT-ICR MS applied to crude oil and its derivatives. Some of them are (a) monitoring polar compound removal from diesel by hydro and oxy treatments, (b) detailed naphthenic acid assignment, and monitoring of their correlation with and (c) their removal processes and (d) identification of asphaltene fraction and identification of contaminants.

Índice

LISTA DE ABREVIATURAS XIX

LISTA DE TABELAS XX

LISTA DE FIGURAS XXI

1. INTRODUÇÃO 1

1.1 PETRÓLEO 3
1.2 ESPECTROMETRIA DE MASSAS 5
1.3 IONIZAÇÃO A PRESSÃO ATMOSFÉRICA POR *ELECTROSPRAY* (ESI). 6
1.4 O ANALISADOR DE MASSAS POR RESSONÂNCIA CICLOTRÔNICA DE ÍONS 9
1.5 HISTÓRICO 9
1.6 CARACTERÍSTICAS DOS ANALISADORES FT-ICR 10
1.7 PRINCÍPIOS BÁSICOS 12
1.8 EXCITAÇÃO E DETECÇÃO 13
1.9 ANÁLISE DO PETRÓLEO POR INSERÇÃO DIRETA UTILIZANDO ESI(±)-FT-ICR MS 15
1.10 MASSA NOMINAL 18
1.11 MASSA EXATA 18
1.12 DEFEITO DE MASSA 19
1.13 EXATIDÃO DE MASSAS 19

2. OBJETIVOS 25

3. PARTE EXPERIMENTAL 26

3.1 ANÁLISE DAS AMOSTRAS POR FT ICR MS 26
3.2 LTQ FT ULTRA MASS SPECTROMETER 28
3.2.1 CARACTERÍSTICAS DO FTQ FT MS ULTRA 29
3.3 DESENVOLVIMENTO DO *SOFTWARE* 32

4. RESULTADOS E DISCUSSÃO 33

4.1 ENTRADA DE DADOS E TRANSFORMAÇÃO PARA MODO CENTRÓIDE 34
4.2 FILTRAGEM DE SINAIS PELA RELAÇÃO SINAL/RUÍDO 40
4.3 DEFINIÇÃO DE FAIXA DE ERRO EXPERIMENTAL 45
4.4 RECALIBRAÇÃO DE MASSAS 47
4.5 ATRIBUIÇÃO DE FÓRMULA MOLECULAR 54
4.5.1 CLASSIFICAÇÃO DE ESPÉCIES POLARES NO PETRÓLEO 55
4.5.2 BANCO DE DADOS TEÓRICOS 57

4.6 CONFIRMAÇÃO POR PADRÃO ISOTÓPICO 63
4.7 CONFIRMAÇÃO POR SERIALIZAÇÃO 68
4.8 ORGANIZAÇÃO E VISUALIZAÇÃO DOS DADOS 71
4.8.1 VISUALIZAÇÃO ATRAVÉS DE DIAGRAMAS CLÁSSICOS. 79
4.8.1.1 Diagrama de Kendrick 81
4.8.1.2 Diagrama de número de carbonos vs. DBE 84
4.8.1.3 Diagrama de van Krevelen 86
4.8.1.4 Diagrama de van Krevelen modificado 88

5. CONCLUSÕES 90

6. PERSPECTIVAS FUTURAS 92

Lista de Abreviaturas

ICR – Ressonância ciclotrônica de íons do inglês *Ion Cyclotron Resonance*.

FT-ICR – Ressonância ciclotrônica de íons com transformada de Fourier do inglês *Fourier Transform Ion Cyclotron Resonance.*

ESI – Ionização por *Electrospray do* inglês *ElectroSpray Ionization.*

MS – Espectrometria de Massas do inglês *Mass Spectrometry*

GC-MS – Cromatografia a gás acoplada à espectrometria de massas.

LC-MS – Cromatografia líquida acoplada a espectrometria de massas.

TOF – Tempo de Vôo do inglês *Time of Flight*

FT – Transformada de Fourier do inglês *Fourier Transform*

LTQ – *Linear Trap Quadrupole*

LIT – *Linear Ion Trap*

AGC – Controle automático de íons do inglês *automatic gain control*

DBE – *double bond equivalent*

KNM – Massa nominal de Kendrick do inglês Kendrick *Nominal Mass*

KMD – Defeito de massa de Kendrick do inglês *Kendrick defect mass*

Lista de Tabelas

Tabela 1. *m/z* teórica calculada a partir de combinações dos átomos de carbono, hidrogênio, nitrogênio, oxigênio e enxofre com suas respectivas fórmulas moleculares e os erros correspondente à *m/z* experimental de 417.13352..... 2

Tabela 2. . *m/z* teórica calculada a partir de combinações dos átomos de carbono, hidrogênio, nitrogênio, oxigênio e enxofre com suas respectivas fórmulas moleculares e os erros correspondente à *m/z* experimental de 417.13352..... 3

Tabela 3. Valores de *m/z* e intensidade para o sinal adquirido mostrado na **Figura 14**. Valores em vermelho utilizados para o calculo do desvio padrão (σ). Valor de *m/z* em verde correspondente ao máximo do sinal..... 8

Lista de Figuras

Figura 1. Ilustração da análise direta de petróleo por FT-ICR MS utilizando a técnica de ionização à pressão atmosférica por *electrospray* (ESI): espectros adquiridos nos modos positivo e negativo de íons com 30 mil espécies detectadas no total. 2

Figura 2. Atribuição das classes de compostos polares encontrados no petróleo através do uso da técnica de ionização direta ESI após tratamento por *software* específico de estudos em petroleômica por FT-ICR MS. 2

Figura 3. Esquema mostrando o processo de ionização da técnica de ionização à pressão atmosférica por *electrospray* (ESI)..... 8

Figura 4. Trajetória circular de um cátion na presença de um campo magnético (B) constante. 12

Figura 5. Esquema de uma cela de ICR mostrando os eletrodos de excitação e confinamento/detecção e suas disposições eletrônicas. 14

Figura 6. Da esquerda para a direita – Imagem da corrente induzida no domínio de tempo – Espectro de freqüência após FT – Espectro de massas depois da transformada de freqüência para *m/z*. 14

Figura 7. (a) Espectro de ESI(+)-ICR FT MS para amostra de petróleo bruto com mais de 25 mil sinais resolvidos (b) expansão para o intervalo de 0,5 *m/z* mostrando 7 sinais resolvidos e caracterizados a nível molecular. 16

Figura 8. Procedimento experimental mostrando as etapas necessárias para a obtenção do espectro de ESI-FT-ICR MS no modo positivo de íons, de uma amostra de petróleo bruto contendo 16 mil sinais resolvidos. 27

Figura 9. Esquema do espectrômetro de massas LTQ FT MS Ultra mostrando os dois analisadores de massas utilizados, a disposição da fonte de ionização e os quadrupolos e octapolos de transferência de íons. 29

Figura 10. Cela de ICR do LTQ FT MS Ultra mostrando as grades utilizadas como eletrodos de excitação (esquerda) e a maior homogeneidade do campo elétrico criado mostrado pelas linhas de campo isopotencial (direita - acima). 30

Figura 11. Esquema geral de processamento de dados para análise de dados de Petroleômica por FT-ICR MS. 33

Figura 12. Pagina principal do *software* mostrando o espectro de ESI FT-ICR MS adquirido (acima) e a sua expansão para um intervalo de 0,022 *m/z* mostrando a forma do sinal adquirido. 34

Figura 13. Dois sinais adquiridos por ESI FT-ICR MS mostrando a) um sinal com a forma gaussiana perfeita b) sinal assimétrico devido a baixa amplitude da intensidade (pouco pontos).....35

Figura 14. Sinal simétrico (azul), distância a meia altura (vermelho) e o ponto máximo calculado pela mediana da distância da meia altura (triangulo verde). 38

Figura 15. Sinal assimétrico (azul), distância a meia altura (vermelho) e o ponto máximo calculado pela mediana da distância da meia altura (ponto verde) 39

Figura 16. Espectro com alta relação sinal/ruído, mostrando o limiar estabelecido utilizando intensidade fixa de 5 % (linha vermelha – acima) e o limiar estabelecido utilizando a distribuição normal (histogramas – abaixo) e 3 vezes o desvio padrão (linha verde nos espectros e linha pontilhada vermelha nos histogramas). 42

Figura 17. Espectro com média relação sinal/ruído, mostrando o limiar estabelecido utilizando intensidade fixa de 5 % (linha vermelha – acima) e o limiar estabelecido utilizando a distribuição normal (histogramas – abaixo) e 3 vezes o desvio padrão(linha verde nos espectros e linha pontilhada vermelha nos histogramas). 43

Figura 18. Espectro com baixa relação sinal/ruído, mostrando o limiar estabelecido utilizando intensidade fixa de 5 % (linha vermelha – acima) e o limiar estabelecido utilizando a distribuição normal (histogramas – abaixo) e 3 vezes o desvio padrão (linha verde nos espectros e linha pontilhada vermelha nos histogramas). 44

Figura 19. Expansão para os espectro de ESI(+)-FT ICR-MS mostrando (acima) unidades de repetição CH_2 e (abaixo) unidades de repetição de 2H.....49

Figura 20. Dispersão de erro de massas para espectro de ESI(+)- FT ICR-MS para o modo positivo de íons antes e depois da recalibração. 52

Figura 21. Espectro de ESI(+)-FT-ICR MS mostrando os sinais nos quais houve a atribuição de sua composição elementar (linhas azuis), seus isotopólogos de ^{13}C e ^{34}S (linhas verdes) e os não identificados (linhas vermelhas) antes e após o processo de recalibração. 53

Figura 22. Principais classe de compostos polares presentes no petróleo detectados por ESI(±)-FT-ICR MS 54

Figura 23. Fluxograma mostrando o início do processo de construção do banco de dado teórico para a classe N (fórmulas moleculares com DBE de 0 até 10). 59

Figura 24. Interface gráfica do módulo de criação dos bancos de dados teóricos mostrando a fórmula molecular inicial, o respectivo valor de DBE calculado dinamicamente e as opções de máximo de DBE e o modo de íons a ser utilizado. 60

Figura 25. Fluxograma mostrando a segunda parte do processo de construção do banco de dado teórico para a classe N (fórmulas com DBE de 11 até 40). 61

Figura 26. Expansão de um espectro de ESI(+)-FT ICR-MS mostrando as atribuições das fórmulas moleculares para as *m/z* 334,25293 e *m/z* 336,26858 como $C_{24}H_{31}N$ e $C_{24}H_{33}N$ respectivamente. Os isotopólogos de ^{13}C [M + 1] ($C_{23}H_{31}N^{13}C$ e $C_{23}H_{33}N^{13}C$) correspondentes são então caracterizados com as intensidades relativas 24% para ambos mostrando os sinais com as respectivas fórmulas moleculares mono isotópicas. 65

Figura 27. Espectro de ESI(+)-FT ICR-MS mostrando os sinais caracterizados (linhas azuis) e o sinais identificados para os respectivos isotopólogos (linhas verdes).....67

Figura 28. Expansão para os espectro de ESI(+)-FT ICR-MS mostrando unidades de repetição -CH_2 (acima) e unidades de repetição de 2 H (abaixo). 69

Figura 29. Fórmulas estruturais sugeridas para íons mostrados na Figura 28, exemplificando a serialização por insaturações e por alquilações. 70

Figura 30. Espectro de ESI(+)-FT-ICR-MS obtido para a amostra de petróleo 1 mostrando mais de 17000 sinais resolvidos. 71

Figura 31. Interface gráfica do *software* desenvolvido mostrando a aba de visualização das atribuições realizadas separadas por classe de heteroátomos. 72

Figura 32. Sinais demarcados no espectro mostrando a distribuição de cada classe de heteroátomos (N, NS, NO, OS, N2) identificadas para o modo negativo de íons e os sinais caracterizados, não caracterizados, os isotopólogos e os sinais atribuídos como parte do ruído 73

Figura 33. Gráfico de barras mostrando a proporção relativa para cada classe de heteroátomo caracterizada para o espectro de ESI(+)-FT-ICR MS mostrado na **Figura 32**. 75

Figura 34. Espectro de ESI(-)-FT-ICR-MS obtido para a amostra de petróleo 1 mostrando mais de 27000 sinais resolvidos. 76

Figura 35. Sinais demarcados no espectro mostrando a distribuição de cada classe de heteroátomos (N, NO, NO2, NO3, O, O2, O3S) identificada para o modo negativo de íons e os sinais caracterizados, não caracterizados, os isotopólogos e os sinais atribuídos como parte do ruído. 77

Figura 36. Gráfico de barras mostrando a proporção relativa para cada classe de heteroátomo caracterizada para o espectro de ESI(-)-FT-ICR MS mostrado na Figura 35. 78

Figura 37. Interface para seleção de diagramas (Kendrick, van Krevelen, nº carbonos vs. DBE e gráfico de barras) e criação dos gráficos para classes selecionadas pelo usuário. 80

Figura 38. Diagramas de Kendrick do espectro mostrado na **Figura 32** para a classe N (homólogos de piridinas – esquerdos) e para as classes N e NO (homólogos de piridinas(N) e amidas/quinolinas (NO) - direita). 82

Figura 39. Diagrama de Kendrick em três dimensões do espectro mostrado na **Figura 38** para a classe N (homólogos de piridínicos) e suas respectivas fórmulas estruturais propostas para a série homóloga de maior e menor KMD. 83

Figura 40. Diagrama de número de carbonos versus DBE para a classe N detectada no modo positivo de íons (acima), e para a classe O2 no modo negativo de íons e suas

respectivas fórmulas estruturais sugeridas para a série homóloga de maior e menor valor de DBE. 84

Figura 41. Diagrama de van Krevelen para o acompanhamento de hidrotratamento em diesel para as classes NO, NO2 e NO3 no modo negativo de íons. 87

Figura 42. Diagrama de van Krevelen 3D modificado para o eixo da abscissa (número de carbono) utilizando a razão H/C para a ordenada, e intensidade como mapa de cor, realizado no modo de imagem pelo *software* desenvolvido para a fração de maltenos analisada por FT ICR-MS no modo positivo de íons. 89

1. Introdução

A investigação de novos indicadores moleculares de petróleo e seus derivados e o estudo de suas diversas aplicações em Geologia, Química e Geoquímica do Petróleo, bem como em diferentes áreas da sua indústria através de técnicas rápidas, modernas e abrangentes através da Petroleômica por transformada de Fourier e ressonância ciclotrônica de íons (FT-ICR MS) de altíssima resolução e exatidão. Essa tecnologia se coloca hoje como uma estratégia na vanguarda no conhecimento para análise de petróleo, sendo objeto de estudos e conhecimento de alguns grupos pioneiros de pesquisa ao redor do mundo, na academia e nas principais empresas de petróleo e energia no mundo.[1] Através da técnica de FT-ICR MS, é possível determinar as fórmulas moleculares de milhares de componentes polares do petróleo e de seus derivados (**Figura 1**), e assim ordená-los e classificá-los conforme suas principais classes e seus principais heteroátomos (como as classes N, NO, NS, O2 e OS) (**Figura 2**), e mais detalhadamente de acordo com seus graus de insaturação/ciclização (DBE) e seus números de carbonos. [2]

[1] Marshall, A.G.; Hendrickson, C. L.; Jackson, G. S. **Fourier transform ion cyclotron resonance mass spectrometry: a primer.** *Mass Spectrometry Reviews* (1998), 17, 1-35.

[2] Marshall, A. G.; Rodgers, R. P. **Petroleomics: The Next Grand Challenge for Chemical Analysis.** *Accounts of Chemical Research.* (2004), 37, 53-59.

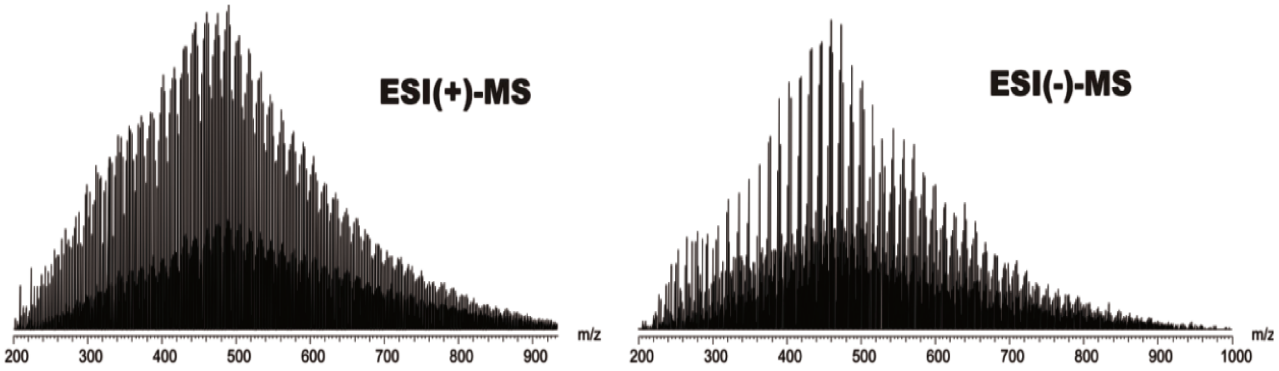


Figura 1. Ilustração da análise direta de petróleo por FT-ICR MS utilizando a técnica de ionização à pressão atmosférica por *electrospray* (ESI): espectros adquiridos nos modos positivo e negativo de íons com 30 mil espécies detectadas no total.

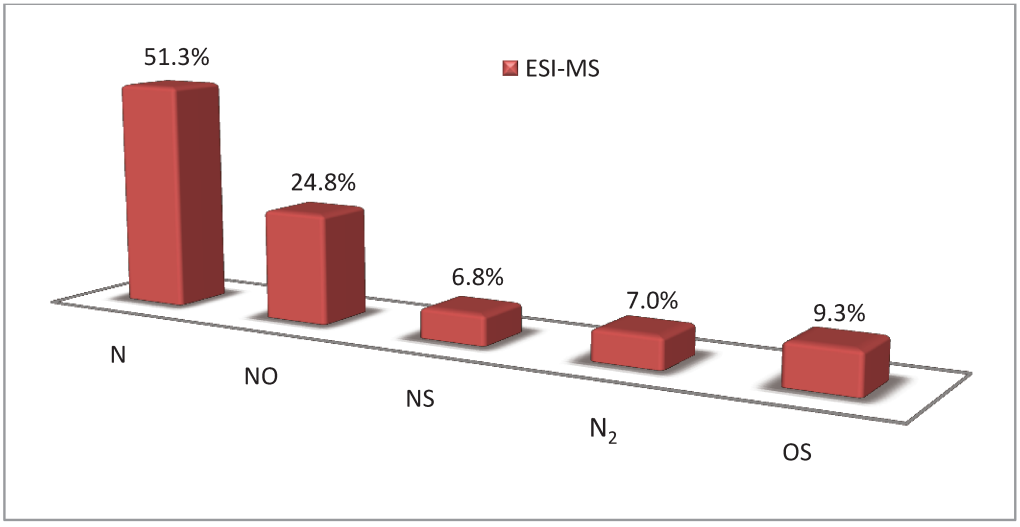


Figura 2. Atribuição das classes de compostos polares encontrados no petróleo através do uso da técnica de ionização direta ESI após tratamento por *software* específico de estudos em petroleômica por FT-ICR MS.

1.1 Petróleo

O petróleo bruto é sem dúvida a mistura mais complexa em termos de composição em misturas orgânicas e em termos de número de constituintes quimicamente distintos. É composto principalmente por hidrocarbonetos insaturados e saturados, além de moléculas hétero atômicas (N, S e O), com uma pequena porcentagem de metais, em particular o vanádio, níquel, ferro e cobre (<100 ppb). A fração de hidrocarbonetos pode ser tão alta quanto 90% em massa nos óleos leves, e de cerca de 70% no petróleo pesado. As espécies hétero atômicas polares estão em torno de 5%, podendo chegar a 15% nos óleos pesados, e os compostos aromáticos em torno de 30%[3].

Compostos hetero-atômicos representam uma parcela relativamente pequena do petróleo, porém a determinação dos compostos contendo os heteroátomos de nitrogênio, enxofre e oxigênio é de extrema importância, já que estas espécies contribuem para, entre outros, a deposição de sólidos, floculação e desativação de catalisadores, refletindo em problemas de armazenamento e corrosão na exploração e nas refinarias[4].

[3] Speight, J. G. **Handbook of Petroleum Analysis**. John Wiley and Sons, Inc.: New York, 2001.

[4] Altgelt, K. H. B. **Composition and Analysis of Heavy Petroleum Fractions.** Marcel Dekker: New York, 1994.

As espécies contendo enxofre são encontradas tanto em compostos não-polares (tiofenos) como nos polares (sulfóxidos). O percentual total de enxofre no petróleo varia de 0,04% até cerca de 5% para óleos pesados. O número de átomos de enxofre por molécula aumenta acentuadamente em frações de maior ponto de ebulição. Os compostos de enxofre são considerados, entre os compostos hetero-atômicos, os mais importantes por causa dos efeitos prejudiciais, tais como corrosão e formação de depósitos e emissão de poluentes após a combustão de seus derivados[4].

As espécies contendo oxigênio são encontradas na forma de fenóis e ácidos carboxílicos. O conteúdo total de oxigênio é normalmente inferior a 2% em massa, porém, o teor de oxigênio pode aumentar drasticamente com o ponto de ebulição, resultando em resíduo não-volátil com teor de até 8%. Estes compostos oxigenados são conhecidos por serem os principais contribuintes para a corrosão dos equipamentos durante o processo de extração e refinamento de produtos petrolíferos [4].

Os compostos de nitrogênio básicos são principalmente piridinas e ocorrem durante quase todos os intervalos de ponto de ebulição, com tendência a existirem em maior abundância nas frações de alto ponto de ebulição e nos resíduos. Os compostos de nitrogênio não-básicos, que são geralmente pirróis, indóis e carbazóis, estão presentes em faixas com maior ponto de ebulição e nas frações dos resíduos. Os compostos nitrogenados são geralmente encontrados em pequenas concentrações, de no máximo 0,9% em massa do petróleo bruto [4].

1.2 Espectrometria de Massas

A espectrometria de massas (MS) através de inserção direta[5] ou em combinação com diferentes técnicas de separação (ou seja, GC-MS [6] ou LC-MS[7]) tem sido utilizada na caracterização abrangente de destilados de petróleo. Porém a complexidade da composição do petróleo bruto e seus derivados pesados têm requerido muito mais da química moderna, uma vez que a resolução e exatidão das medidas nos atuais analisadores de massas dos espectrômetros de massas fornecem informações limitadas da composição em análises direta de petróleo. Entre os desenvolvimentos recentes em espectrometria de massas destaca-se a evolução das técnicas de ionização preconizadas pelo surgimento do *electrospray* (ESI) capaz de ejetar para fase gasosa misturas de uma forma branda, e o desenvolvimento de analisadores de massas com altíssima resolução como o ICR. A combinação destas técnicas (ESI e ICR) permitiu o desenvolvimento de análises de misturas complexas, como o petróleo, sem a necessidade de utilizar métodos prévios de separação[8].

---

[5]Zhan, D.; Fenn, J. B. **Electrospray mass spectrometry of fossil fuels.** *International Journal of Mass Spectrometry* (2000), 194, 197-208.

[6] Killops, S. D.; Readman, J. W. **Hplc Fractionation and GC-MS Determination of Aromatic-Hydrocarbons from Oils and Sediments.** *Organic Geochemistry* (1985), 8, (4), 247-257.

[7] Felix, G.; Bertrand, C.; Vangastel, F. **A New Caffeine Bonded Phase for Separation of Polyaromatic Hydrocarbons and Petroleum Asphaltenes by High-Performance Liquid-Chromatography**. *Chromatographia* (1985) 20, (3), 155-160.

[8] Rodgers, R.P.; Schaub, T. M.; Marshall A. G. **Petroleomics: MS returns to its roots.** *Analytical Chemistry* (2005), 77, 20A.

1.3 Ionização a pressão atmosférica por *electrospray* (ESI).

Sendo uma das técnicas de ionização em maior expansão, a ionização por ESI passou por duas fases distintas de investigação e desenvolvimento. A primeira decorreu antes de 1970 e centrou-se mais nos aspectos fundamentais do processo de produção de carga assim como no modo experimental de concretizá-lo, podendo salientar o trabalho realizado por Dole *et al*.[9] A segunda fase deu-se a partir de 1970 com destaque para o trabalho desenvolvido em 1984 por Yamashita e Fenn,[10] considerado pioneiro na utilização da ionização por *ESI* na espectrometria de massas. A partir deste trabalho a técnica sofreu um incremento notório com o desenvolvimento e construção de fontes de ionização comerciais baseadas no princípio de carregar gotas eletricamente.

Há essencialmente três características que fazem com que ESI destaque-se em relação a outras técnicas de ionização. A primeira destas características é a capacidade de produzir íons multiplamente carregados, com número de carga elevado, reduzindo assim a razão *m/z*, de tal modo que é possível analisar compostos de elevada massa molar até centenas de kDa. Uma segunda característica é que as amostras a serem analisadas devem ser introduzidas em solução, o que torna possível o acoplamento com muitas técnicas de separação. Por último e não menos importante, o fato de ser uma técnica de ionização branda permite a análise de mistura complexas

[9] Dole, M.; Mack, L.L.; Hines, R.L.; Mobley, R.C.; Ferguson L.D.; Alice, M.B. **Molecular beams of macroions.** Journal of Chemical Physics (1968), 49(5) , 2240-9.

[10] Yamashita, M.; Fenn, J.B. **Electrospray ion source. Another variation on the free-jet theme.** Journal of Chemical Physics (1984)**,** 88(20), 4451-9.

em uma única etapa, uma vez que a fragmentação das espécies transferidas para a fase gasosa é minimizada.

A produção de íons por *electrospray* requer essencialmente duas etapas: dispersão das gotas altamente carregadas à pressão atmosférica seguida de condições que permitam a evaporação da gota.

As soluções são primeiramente aspergidas eletrostaticamente com formação de gotas pequenas e altamente carregadas eletricamente. A nebulização da solução é em alguns casos facilitada pela ajuda de um gás auxiliar (N_2). Posteriormente as moléculas do analito devem de alguma forma ser separadas do solvente e ejetadas para a fase gasosa. Este passo de formação de íons, como em muitas das técnicas de ionização consideradas suaves, é provavelmente o menos compreendido no processo global do ESI. Alguns mecanismos têm sido propostos para a dessorção dos íons a partir de gotas carregadas sendo que o modelo de resíduo de carga de Dole,[9] aplicado a macromoléculas, foi talvez o primeiro a servir de base para a atual técnica de ESI. Neste modelo é considerado que à medida que o solvente se evapora a densidade de carga na superfície aumentará até que as forças repulsivas de Coulomb entre as cargas superficiais excederão a tensão superficial levando à divisão da gota inicial. Se este processo de divisão continuar e se a solução original for suficientemente diluída será alcançado um estado no qual cada gota conterá uma única molécula que reterá parte da carga inicial, ou seja, formará macro íons.

Outro mecanismo para a geração de pequenos íons, o da evaporação iônica, foi proposto por Iribarne e Thomson,[11] onde sugerem que a evaporação do solvente conduz a uma instabilidade das gotas com razões elevadas de densidade de carga superficial/ raio da gota. A energia eletrostática associada com a gota carregada torna-se então suficientemente grande para ejetar os íons do analito para a fase gasosa.

Este mecanismo foi aplicado a macromoléculas por Fenn e colaboradores,[12] o qual propôs que uma parte da molécula carregada teria a capacidade de penetrar a superfície da gota devido a movimento Browniano. A existência de repulsão coulombiana entre esta parte da molécula e a superfície da gota ejetará a molécula carregada para fora da gota (**Figura 3)**.

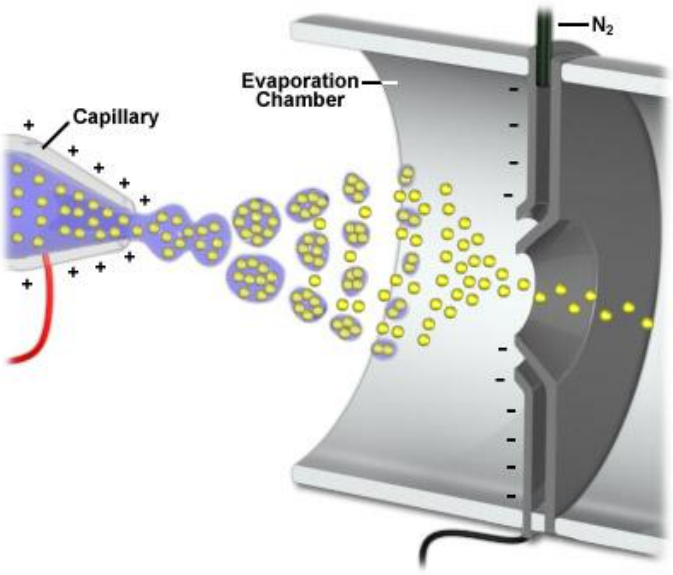


Figura 3. Esquema mostrando o processo de ionização da técnica de ionização à pressão atmosférica por *electrospray* (ESI).

[11] Iribarne, J.V.; Thomson, B.A. **On the evaporation of small ions from charged droplets.** *Journal of Chemical Physics* (1976), 64(6), 2287-94.

[12] Wong, S.F.; Meng, C.K.; Fenn, J.B. **Multiple charging in electrospray ionization of polyethylene glycols.** *Journal of Chemical Physics* (1988)**,** 92-2, 546-50.

1.4 O analisador de massas por ressonância ciclotrônica de íons

A análise de matrizes complexas como o petróleo tornou-se possível com o desenvolvimento dos analisadores de massas por Ressonância Ciclotrônica de Íons (ICR, do inglês *ion cyclotron ressonance*)[1], capaz de produzir espectros de massas com altíssimo poder de resolução (> 750.000 $\frac{m}{\Delta m}$) e com altíssima exatidão de massas (menor que 1.0 ppm).

1.5 Histórico

ICR-MS foi desenvolvido por Alan G. Marshall e Melvin B. Comisarow e sua primeira publicação foi realizada em 1974[13]. Atualmente é o analisador de massas capaz de atingir os níveis mais altos em termos de poder de resolução ($m/z/\Delta m/z_{50\%}$, onde $\Delta m/z_{50\%}$ é a diferença em *m/z* a meia altura do sinal) tendo como recorde de 3.000.000 e de altíssima exatidão com erros médios de 200 ppb[14].

Analisadores de massas como o ICR são baseados em medidas de frequência dos íons em campo magnético constante, em contraste com outros de deflação (setores magnéticos/elétricos), estabilidade de íons (quadrupolos e *ions traps*) ou tempo de vôo (TOF). Essas freqüências são correlacionadas diretamente com a *m/z*

[13] Comisarow, M.B; Marshall A.G. *Chemical Physics Letters.* (1974), *25*, 282.

[14] Schaub, T. M.; Hendrickson, C.L.; Horning, S.; Marshall, A.G. **High-Performance Mass Spectrometry: Fourier Transform Ion Cyclotron Resonance at 14,5 Tesla**. *Analytical Chemistry.* (2008), 11, 3985-3990.

sendo tal relação possível devido às freqüências dos íons serem independentes de suas energias e velocidade.

1.6 Características dos analisadores FT-ICR

Existem algumas características peculiares do analisador ICR devido à redução de dados por FT [15]. Entre as mais importantes estão:

1) Como a corrente induzida pelos pacotes de íons de uma larga faixa de *m/z* pode ser detectada simultaneamente, analisadores do tipo FT-ICR MS oferecem várias vantagens ao obter o espectro de massas de uma só vez ao invés de analisar cada *m/z* separadamente.

2) Como a magnitude de resposta é linearmente proporcional à magnitude de excitação (quanto mais perto do detector e mais coerente forem os pacotes de íons melhor é a indução), quanto mais seletiva e em fase for à excitação, ou o pré empacotamento, melhor a resposta.

3) O poder de resolução varia diretamente com o tempo de aquisição T [1]. Entretanto T é limitado geralmente por dois motivos: a) A introdução de amostra por métodos cromatográficos de separação as quais requerem aquisições rápidas. b) O sinal de freqüência em domínio de tempo decai exponencialmente com o tempo devido

[15] Marshall A.G.; Verdun, F.R. **Fourier Transforms in NMR, Optical, and Mass Spectrometry**. A User's Handbook, Elsevier, New York (1990), p. 460.

à perda de energia por radiação e por colisão. Logo a relação sinal/ruído diminui com o tempo e a aquisição de dados não deve ser prolongada [16].

Outra característica dos analisadores ICR é que estes requerem aproximadamente 1s para a aquisição de dados. Entretanto, como as técnicas de ionização produzem íons constantemente, é vantajoso que os íons sejam acumulados externamente ao analisador e ejetados assim que o ciclo de aquisição de dados esteja completo. Para o ICR os íons são comumente acumulados em um *ion trap* linear [17].

[16] – Marshall, A.G.; Comisarow, M.B.; Parisod, G. **Relaxation and spectral line shape in Fourier transform ion cyclotron resonance spectroscopy**. *Journal of Chemical Physics*. (1979), 71, 4434–44.

[17] Belov, M.E.; Nikolaev, E.N.; Anderson, G.A.; Udseth, H.R.; Conrads, T.P.; Veenstra, T.D.; Masselon, C.D.; Gorshkov, M.V.; Smith, R.D. **Design and Performance of an ESI Interface for Selective External Ion Accumulation Coupled to a Fourier Transform Ion Cyclotron Mass Spectrometer**. *Analytical Chemistry* (2001), *73*, 253-261.

1.7 Princípios Básicos

Como dito anteriormente o analisador de massas por ressonância ciclotrônica de íons (ICR-MS) é utilizado para a determinação da *m/z* baseado na freqüência ciclotrônica de íons em um campo magnético fixo[1] (**Figura 4**).

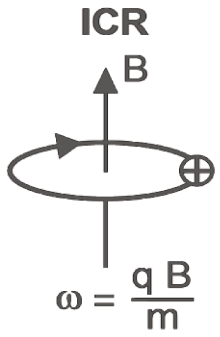


Figura 4. Trajetória circular de um cátion na presença de um campo magnético (B) constante.

O principio básico de operação é fundamentado na relação entre a frequência ciclotrônica e a *m/z* de um dado íon em um campo magnético uniforme, que se apresenta na forma ideal como[1]:

$$f = \frac{qB}{2\pi m} \qquad \textbf{(Equação 1)}$$

onde q é a carga do íons (q = z.e) , B é o campo magnético, m é a massa do íon e f é a freqüência ciclotrônica. Como a freqüência angular é definida como ω = 2πf, temos que a freqüência angular ciclotrônica é dada por:

$$\omega = \frac{qB}{m} \qquad \textbf{(Equação 2)},$$

que é a equação da freqüência ciclotrônica não perturbada (ideal). Logo, todos os íons de uma mesma m/q terão a mesma freqüência ciclotrônica em um campo magnético constante independente de sua velocidade[1].

1.8 Excitação e detecção

Em temperatura ambiente os raios de orbita ciclotrônica dos íons encontram-se em comprimentos na ordem de sub milímetros e longe das placas de detecção das celas que variam de 1 a 2 cm. Além disso, íons de mesma *m/z* oscilam em fases distintas. Logo, para que o sinal gerado atinja uma amplitude que seja detectável, é necessário excitar os íons de forma ressonante através de um campo elétrico oscilante ou rotatório em fase com o campo magnético, a fim de aumentar o raio da orbita de oscilação e formar pacotes de íons de mesma *m/z* que sejam os mais compactos e coerentes espacialmente possíveis[1].

Após a excitação, o movimento desses pacotes de íons gera um sinal no domínio de tempo o qual consiste na diferença na corrente induzida pelos pacotes nos dois eletrodos (**Figura 5)**.

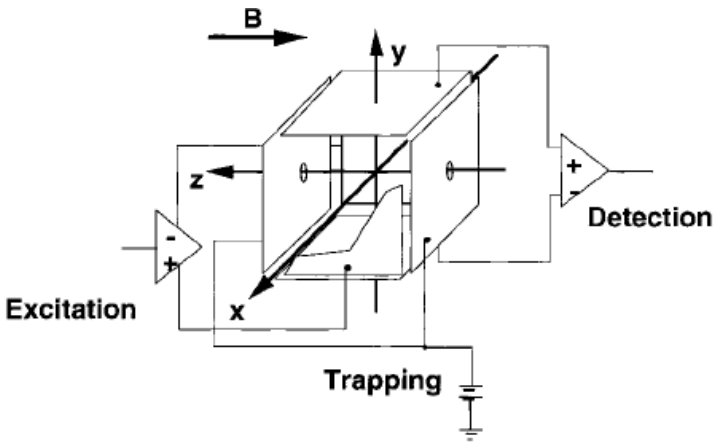


Figura 5. Esquema de uma cela de ICR mostrando os eletrodos de excitação e confinamento/detecção e suas disposições eletrônicas.

Esse sinal é registrado, digitalizado e submetido à transformada de Fourier discreta gerando um espectro das freqüências ciclotrônicas dos íons. A frequência é então convertida para a unidade de razão *m/z* através da equação ciclotrônica (equação 2 - **Figura 6**) resultando no espectro de massas.

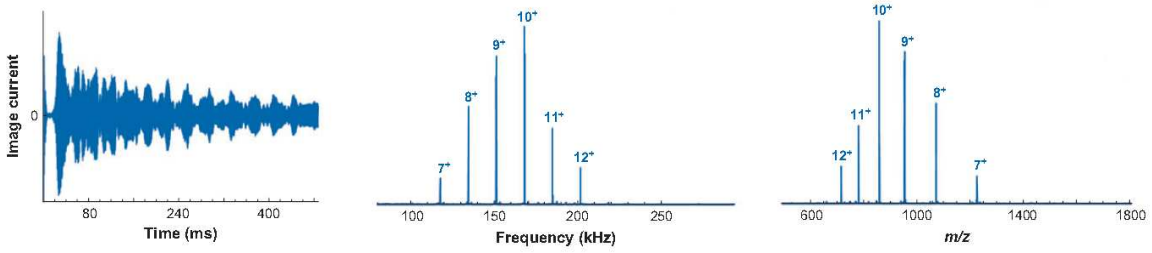


Figura 6. Da esquerda para a direita – Imagem da corrente induzida no domínio de tempo – Espectro de freqüência após FT – Espectro de massas depois da transformada de freqüência para *m/z*.

1.9 Análise do petróleo por inserção direta utilizando ESI(±)-FT-ICR MS

A **Figura 7** mostra o espectro no modo positivo de íons de ESI(+)-FT ICR-MS obtido para uma amostra de petróleo brasileiro com mais de 15 mil sinais resolvidos. A expansão para o intervalo de 0,5 *m/z* (**Figura 2. Atribuição** das classes de compostos polares encontrados no petróleo através do uso da técnica de ionização direta ESI após tratamento por software *específico* de estudos em petroleômica por FT-ICR MS. (**Figura 7**–espectro inferior) mostra a resolução e identificação de 7 sinais e sua correspondente fórmula molecular.

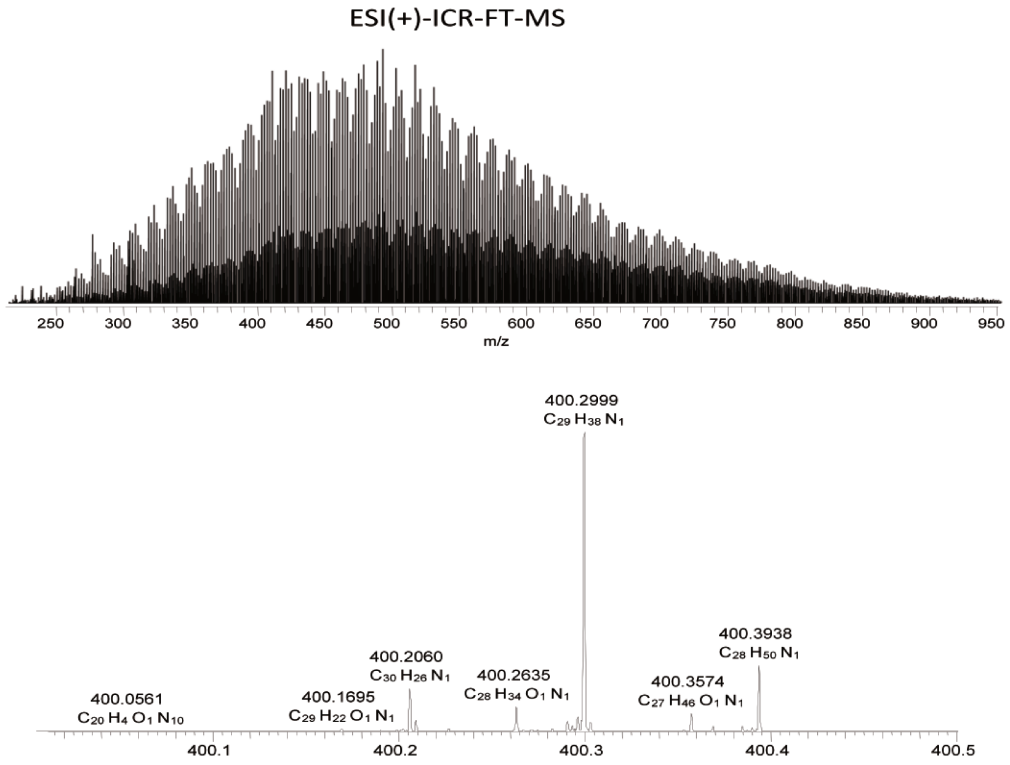


Figura 7. (a) Espectro de ESI(+)-ICR FT MS para amostra de petróleo bruto com mais de 25 mil sinais resolvidos (b) expansão para o intervalo de 0,5 *m/z* mostrando 7 sinais resolvidos e caracterizados a nível molecular.

Esta nova tecnologia tem sido utilizada com muito sucesso na caracterização total do petróleo bruto e suas várias frações incluindo os asfaltenos, provendo assim um método inovador, preciso e de elevado nível de detalhamento molecular, e com extrema rapidez (< 6 minutos)[8,18,19]. Séries de compostos, com ênfase nos polares, mas também

[18] Klein, G.C.; Ansgtrom, A.; Rodgers, R.P.; Marshall, A.G. **Use of Saturates/Aromatics/Resins/Asphaltenes (SARA) Fractionation To Determine Matrix Effects in Crude Oil Analysis by Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.** Energy & Fuels, (2006), 20(2), 668-72.

[19] Klein, G. C.; Kim, S.; Rodgers, R. P.; Marshall, A. G.; Yen, A. **Mass Spectral Analysis of Asphaltenes. II. Detailed Compositional Comparison of Asphaltenes Deposit to Its Crude Oil Counterpart for Two Geographically Different Crude Oils by ESI FT-ICR MS**. Energy & Fuels (2006), 20(5), 1973-1979.

os semi-polares e até apolares, podem ser analisadas em poucos minutos diretamente de uma amostra bruta determinando-se a presença e quantidade relativa de classes, suas séries homólogas, número de carbonos e insaturações/anéis. O método dispensa assim, a utilização de uma bateria de procedimentos bem mais dispendiosos de preparação de amostras e outras técnicas diversas que podem totalizar horas de análise [20,21]. As áreas da indústria petrolífera em que esta tecnologia se aplica incluem a exploração, o refino, o controle ambiental e a eficiência na inovação tecnológica de métodos analíticos com a implantação e desenvolvimento de procedimentos na vanguarda do conhecimento científico em petróleo e seus derivados, com o conseqüente aumento de produtividade, rapidez e qualidade dos serviços prestados.

A atribuição inequívoca da composição elementar pode ser realizada para moléculas com baixa massa molar (< 400 Da) utilizando o princípio de massa exata através de seus defeitos de massas e atribuições baseadas em massa de Kendrick vem sido utilizada[22]. Para melhor entendimento alguns termos devem ser definidos, entre eles estão: massa nominal, massa exata, defeito de massa e exatidão de massas.

---

[20] Speight, J. G.; Pancirov, R. J. **Structural types in petroleum asphaltenes as deduced from pyrolysis/gas chromatography/mass spectrometry**. Liquid Fuels Technology (1984), 2(3), 287-305.

[21] Killops, S. D.; Readman, J. W. **HPLC fractionation and GC-MS determination of aromatic hydrocarbons from oils and sediments.** Organic Geochemistry (1985), 8(4), 247-57.

[22] Kendrick, E. **Mass scale based on CH2 = 14,0000 for high-resolution mass spectrometry of organic compounds.** *Analytical Chemistry* (1963), *35*, 2146–2154.

1.10 Massa Nominal

A massa nominal de uma molécula é definida como a soma das massas unitárias dos isótopos mais abundantes em uma molécula. Dois íons apresentando a mesma *m/z* unitária (mesma massa nominal) são então denominados isóbaros. Por exemplo, a molécula N_2 contém massa nominal de 28 Da, da mesma forma que a molécula C_2H_4. Analisadores de massas com insuficiente poder de resolução, como o quadrupolo (poder de resolução unitária), não são capazes de distinguir estas duas moléculas através apenas da aferição de *m/z*.

1.11 Massa Exata

A massa mono isotópica de uma molécula, também conhecida como massa exata, difere da massa nominal, pois ao invés da massa unitária é utilizada a soma das massas exatas dos isótopos naturalmente mais abundantes. A massa exata de um átomo é o resultado da soma da massa do núcleo (levando em consideração a energia de ligação nuclear) e dos seus elétrons.

Utilizando o exemplo anterior, a massa exata para N_2 equivale a 28.00614 Da e para C_2H_4 equivale a 28.03130 Da. Fica claro que estas moléculas diferenciam-se através da massa exata. Logo, fórmulas moleculares empíricas diferentes com mesma massa nominal terão necessariamente massas exatas distintas e únicas.

1.12 Defeito de Massa

O defeito de massa é simplesmente a diferença entre a massa exata e a massa nominal. Esse defeito é característico para cada átomo em particular. Por exemplo, o átomo ^{16}O é composto de oito prótons, oito nêutrons e oito elétrons, sendo as massas individuais destes 1,007276470, 1,008664904, e 0,000548579903 Da respectivamente.

Somando-se estes valores e multiplicando pela suas quantidades, obtém-se o valor de 16,131919633 Da, que é diferente da massa exata experimental para este átomo (15,994915 Da). Essa diferença é explicada através da energia de ligação nuclear. A energia de ligação é necessária para manter prótons e nêutrons no núcleo atômico, e o defeito de massa observado surge a partir da perda relativística de massa ($E = m\ c^2$) quando esta energia de ligação é liberada durante a formação do núcleo atômico. Como esta energia liberada é diferente para cada tipo de átomo, o defeito de massa é específico para cada átomo, assim cada composição elementar terá um defeito de massa único.

1.13 Exatidão de massas

O termo exatidão de massas caracteriza quão perto é a medida de massas experimental de sua massa teórica. Idealmente, se a medida de massa pode ser realizada com altíssima exatidão, através da sua massa exata, proveniente dos defeitos de massas específicos dos átomos, chega-se a uma fórmula molecular empírica única. Entretanto, a identificação inequívoca da fórmula molecular de espectros adquiridos de

amostras complexas, como o petróleo bruto, é uma tarefa complexa e trabalhosa, uma vez que existe mais de vinte mil composições químicas elementares distintas possíveis[2] para os compostos polares de petróleo.

A caracterização molecular realizada, baseada apenas na massa exata e exatidão de massas, pode levar a resultados errôneos apresentando combinações de átomos quimicamente improváveis ou impossíveis. Por exemplo, um sinal de *m/z* 417,13352 aferida em modo positivo de íons sendo caracterizado como íon na forma protonada $[M + H]^+$ apresenta, dentro da faixa de erro de 1,0 ppm, 13 combinações de átomos possíveis, sendo 9 delas quimicamente impossíveis ou improváveis (asterisco Tabela 1).

Tabela 1. *m/z* teórica calculada a partir de combinações dos átomos de carbono, hidrogênio, nitrogênio, oxigênio e enxofre com suas respectivas fórmulas moleculares e os erros correspondente à *m/z* experimental de 417,13352.

Fórmula Molecular	*m/z* Teórica	Erro de massas
$C_{25}H_{21}O_6$	417,133265	0,6
$C_{24}H_{15}N_7O$	417,13326	0,6
$C_{18}H_{29}N_2O_3S_3$	417,133481	0,1
$^*C_{12}H_{148}N_2O_6$	417,133192	0,8
$^*C_{11}H_{37}N_4S_6$	417,133698	-0,4
$^*C_{11}H_{142}N_9O$	417,133187	0,8
$C_{10}H_{25}N_8O_6S_2$	417,133298	0,5
$^*C_6H_{154}NO_9S$	417,133883	-0,9
$^*C_5H_{156}N_4O_3S_3$	417,133408	0,3
$^*C_5H_{148}N_8O_4S$	417,133877	-0,9
$^*C_4H_{39}N_3O_8S_5$	417,13352	0,0
$^*C_3H_{33}N_{10}O_3S_5$	417,133515	0,0
$^*H_{287}S_4$	417,133518	0,0

Mesmo com uma população reduzida de quantidade de átomos e retirando as fórmulas moleculares improváveis ou impossíveis quimicamente, ainda restam quatro combinações de átomos que podem ser consideradas como possíveis fórmulas moleculares.

Algumas restrições adicionais podem ajudar na atribuição como os valores de DBE, de DBE/C e com limites mais estreitos na quantidade de átomos de oxigênio (0 - 4), nitrogênio (0 – 3), enxofre (0 – 4) (Tabela 2).

Tabela 2 – *m/z* teórica calculada a partir de combinações dos átomos de carbono, hidrogênio, nitrogênio, oxigênio e enxofre com suas respectivas fórmulas moleculares, DBE e os erros correspondente à *m/z* experimental de 417,13352.

Fórmula Molecular	m/z Teórica	Erro de massas	DBE	DBE/C
*$C_{25}H_{21}O_6$	417,133265	0,6	16.0	0,6
*$C_{24}H_{15}N_7O$	417,133260	0,6	21,5	0,9
$C_{18}H_{29}N_2O_3S_3$	417,133481	0,1	6,0	0,3
*$C_{12}H_{148}N_2O_6$	417,133192	0,8	-59,5	-5,0
*$C_{11}H_{37}N_4S_6$	417,133698	-0,4	-4,0	-0,4
*$C_{11}H_{142}N_9O$	417,133187	0,8	-54,0	-4,9
*$C_{10}H_{25}N_8O_6S_2$	417,133298	0,5	3,0	0,3
*$C_6H_{154}NO_9S$	417,133883	-0,9	-69,0	-11,5
*$C_5H_{156}N_4O_3S_3$	417,133408	0,3	-69,5	-13,9
*$C_5H_{148}N_8O_4S$	417,133877	-0,9	-63,5	-12,7
*$C_4H_{39}N_3O_8S_5$	417,133520	0,0	-12,5	-3,1
*$C_3H_{33}N_{10}O_3S_5$	417,133515	0,0	-7,0	-2,3
*$H_{287}S_4$	417,133518	0,0	-142,0	-

Com os limites estabelecidos excluem-se todas as fórmulas com valores negativos de DBE e com valores fora da faixa de 0 a 1,0 da razão DBE/C. Entre as fórmulas restantes apenas uma encaixa-se nos limites de átomos estabelecidos. Logo resta apenas uma fórmula molecular possível para a *m/z* experimental de 417,13352 que é $C_{18}H_{29}N_2O_3S_3$.

Entretanto, quanto maior a *m/z*, a quantidade de fórmulas moleculares possíveis, mesmo com as restrições realizadas, cresce acentuadamente e assim existe

a possibilidade de duplicidade nas atribuições. Por exemplo, a *m/z* experimental de 698.65950 pode ser atribuída como $C_{50}H_{48}N$ de *m/z* 698,659828, com erro de -0,5 ppm ou também como $C_{42}H_{88}N_3O_2S$ de *m/z* 698,659176 com erro de 0,5 ppm. O DBE correspondente às formas neutras são 10 e 1, e a razão DBE/C 0,2 e 0,02, respectivamente. Assim, fica evidente a necessidade da utilização de outras estratégias para a obtenção da fórmula molecular inequívoca. Entre elas estão: definição de faixa de erro experimental, recalibração de massas, confirmação de padrão isotópico e correlação de séries homólogas.

Assim, a rapidez de aquisição dos dados por FT ICR-MS se perde por completo devido ao tempo demasiadamente prolongado do processamento e interpretação dos dados em uma análise manual. Isso se deve ao grande volume de espécies de compostos hetero-atômicos presentes no petróleo e nos seus derivados (até ~ 16000 compostos, considerando apenas as moléculas com seus isótopos mais abundantes, e até ~ 50000 considerando os seus isotopólogos de ^{13}C ^{34}S). Além disso, muitas vezes somos confrontados com espectros com características não ideais (baixa exatidão de massa (entre 1-2 ppm) e/ou baixa relação sinal-ruído) que tornam o processamento dos dados manual uma tarefa ainda mais complexa e portanto inviável.

Dessa forma, é evidente a indispensabilidade da utilização de um *software* que seja capaz de automaticamente diferenciar os sinais dos espectros adquiridos, fornecer de maneira confiável as atribuições das composições moleculares elementares correspondentes utilizando estratégias inteligentes de processamento.

2. Objetivos

Realizar a análise de dados oriundos da caracterização de amostras complexas, como o petróleo e seus derivados, por ESI(±)-FT-ICR MS, através da automatização via algoritmos capazes de filtrar, classificar e atribuir as composições elementares inequívocas dos dados, e organizá-los de forma quimicamente lógica. Apresentar ferramentas gráficas utilizadas para facilitar a interpretação e visualização dos dados, que permita ao final a elaboração de um *software* eficaz que permita o acesso a todos estes procedimentos e ferramentas gráficas de forma amigável, maximizando assim o tratamento quimicamente eficiente de dados através da automatização.

3. Parte Experimental

3.1 Análise das amostras por FT ICR MS

As amostras de petróleo bruto e diesel utilizadas foram pesadas (10 mg) e dissolvidas em 10 mL de tolueno e então a solução foi diluída em 10 mL de metanol. Para análises no modo positivo de íons, um total de 2 µL de uma solução aquosa 1% de ácido fórmico foi adicionada aos 20 mL da solução final a fim de servir como agente provedor de prótons para os compostos nitrogenados básicos resultando em moléculas protonadas $[M + H]^+$. Para o modo negativo de íons um total de 2 µL de uma solução aquosa de hidróxido de amônio 1% foi adicionado para o mesmo volume de amostra preparado no modo positivo. As amostras foram inseridas no espectrômetro de massas por infusão direta, ionizadas por ESI e analisadas utilizando o espectrômetro de massas em um magneto de *7,2T, LTQ FT Ultra mass spectrometer* (ThermoScientific, Bremen, Germany). Para a inserção da amostra no modo positivo de íons foi usado o sistema automatizado "chip-based nano-ESI-MS" através do equipamento Triversa NanoMate 100 system (Advion BioSciences, Ithaca, NY , USA) sendo as amostras dispostas em pratos de 96 posições (volume total de 100 µL em cada posição). Já no modo negativo de íons foi utilizada a fonte comercial convencional de ESI. As condições gerais para ESI foram: pressão do gás de 0,3 psi e voltagem capilar de 1,55 kV. Os espectros de massas foram obtidos pela soma de 100 transientes totalizando 6 minutos de aquisição com o poder de resolução de 400.000 $\frac{m}{\Delta m}$ (**Figura 8)**

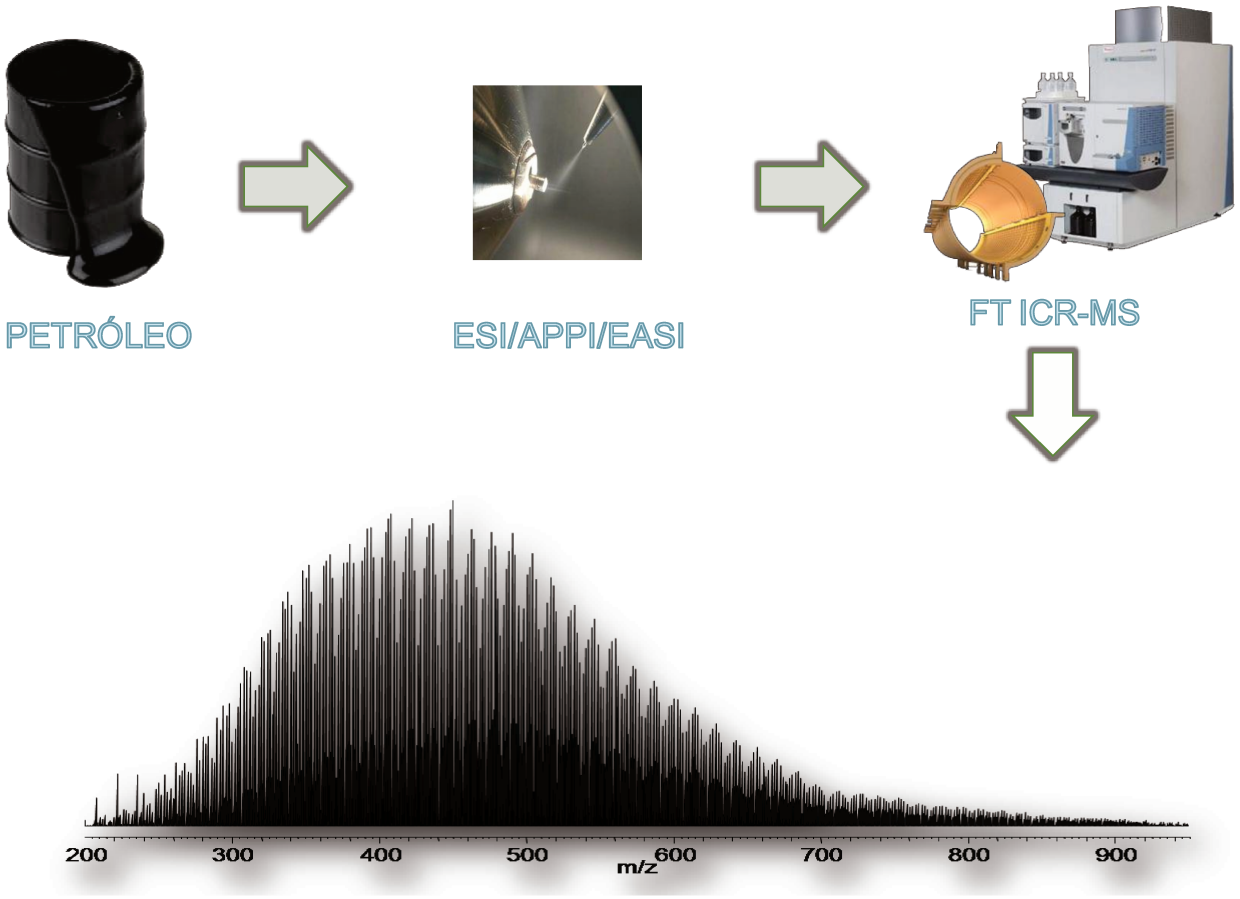


Figura 8. Procedimento experimental mostrando as etapas necessárias para a obtenção do espectro de ESI-FT-ICR MS no modo positivo de íons, de uma amostra de petróleo bruto contendo 16 mil sinais resolvidos.

3.2 LTQ FT Ultra mass spectrometer

Diferentes tipos de analisadores podem ser acoplados para formar novos instrumentos com o intuito de melhorar o poder de resolução e exatidão. É neste contexto, que o LTQ FT Ultra é inserido, uma combinação de dois analisadores tipo *ion trap* diferentes: um do tipo linear (LIT) e um ICR [23]. O objetivo desta combinação é obter maiores coeficientes de resolução e exatidão do que os fornecidos por instrumentos que empregam um TOF como analisador no segundo estágio, ou até mesmo um Orbitrap.

O primeiro espectrômetro de massas comercial empregando este arranjo foi lançado pela empresa Thermo Scientific, o qual fornece altíssima resolução e exatidão. O atributo de empregar um analisador *ion trap linear* (LIT) (**Figura 9**) como analisador de massas de primeiro estágio é que este possui a finalidade de acumular íons no seu interior, manipulando corrente alternada e de radiofrequência simultaneamente. A captura permite a liberação controlada de íons, um requisito para análise em ICR, visto que há um limite na quantidade de íons que poderá ser transferidos para o referido analisador. Além disso, é útil para experimentos múlti-acoplados (MS^n), que permitem estudar mecanismos de fragmentação de moléculas relativamente complexas, além de fornecerem espectros com alta sensibilidade. A possibilidade de capturar íons por períodos que variam de milissegundos a dias permite a análise de reações de fragmentação extremamente incomuns, dando origem a espectros de massas que não podem ser obtidos através de nenhuma outra técnica.

[23] Guan, S.; Marshall, A.; International Journal of Mass Spectrometry and Ion Processes (1995) 146/147, 261.

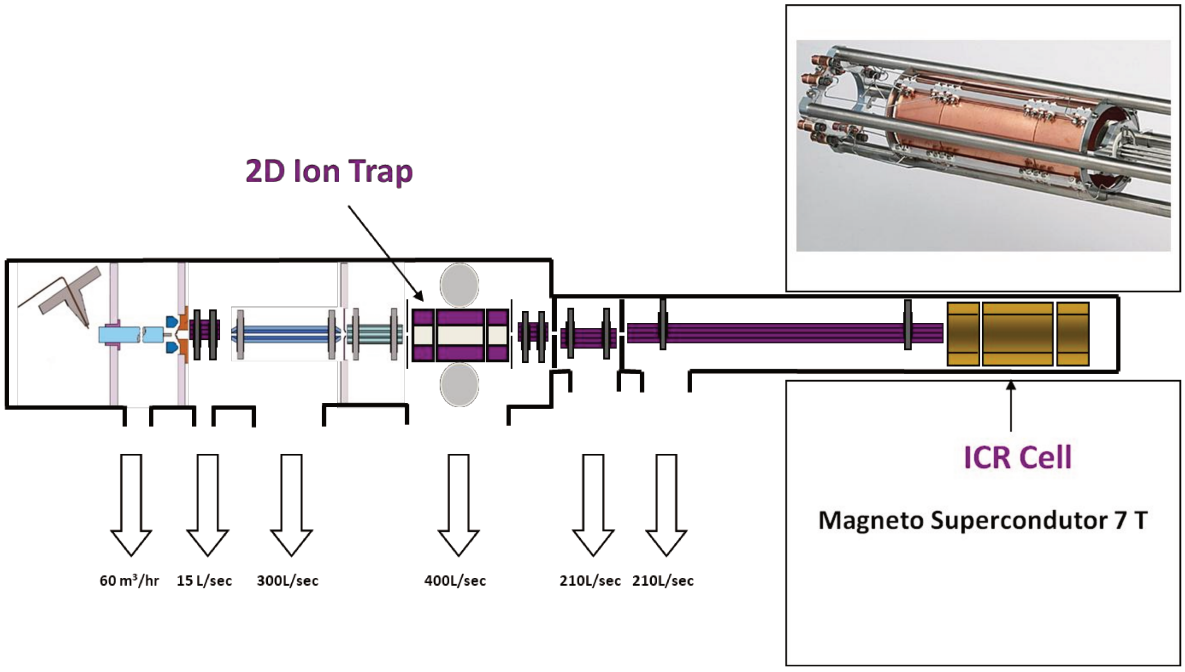


Figura 9. Esquema do espectrômetro de massas LTQ FT MS Ultra mostrando os dois analisadores de massas utilizados, a disposição da fonte de ionização e os quadrupolos e octapolos de transferência de íons.

3.2.1 Características do FTQ FT MS Ultra

O LTQ FT MS Ultra fornece uma exatidão de massa menor do que 1 ppm e a calibração externa é estável por vários dias. Assim, a calibração interna não é necessária. Na espectrometria de massas FT ICR, a exatidão de massa é dependente do número total de íons capturados na céla de ICR. Um elevado número de íons no interior da cela produz o efeito "space charge", sendo a razão de deslocamentos de massas com magnitudes em algumas centenas de ppm. No LTQ FT Ultra há uma

ferramenta, desenvolvida pela Thermo Scientific, para solucionar tal problema, que é o Controle Automático de Ganho (AGC). A função desta ferramenta é regular o número de íons de um experimento no *ion trap* e na cela de ICR. Com uma pré-varredura o LIT determina o total de íons que estão entrando no espectrômetro de massas e com este total de íons é estabelecida a quantidade de íons que deve ser transmitida para a cela de ICR. Além deste controle da população de íons a cela do LTQ FT Ultra [24] (**Figura 10)** utiliza grades como eletrodo de excitação reduzindo os componentes axiais devido a um campo elétrico mais uniforme e homogêneo, permitindo que maiores raios de excitação sejam atingidos, resultando em uma amplitude maior de detecção, sensibilidade e reduzindo os efeitos indesejáveis de "*space charge*".

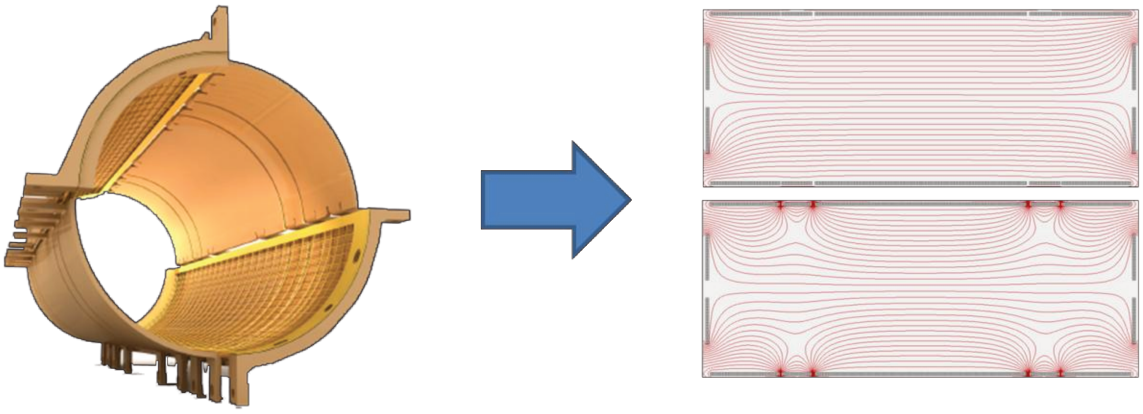


Figura 10. Cela de ICR do LTQ FT MS Ultra mostrando as grades utilizadas como eletrodos de excitação (esquerda) e a maior homogeneidade do campo elétrico criado mostrado pelas linhas de campo isopotencial (direita - acima).

[24] Wieghaus, A., Fröhlich, U., Malek , R., and Horning, S. **The Grid Cell: a New Cell Design for Reduced Z-Axis Ejection in Fourier Transform Ion Cyclotron Mass Spectrometry.** *In Proceedings of the 54th ASMS Conference for Mass Spectrometry and Allied Topics: Seattle, WA, USA,* 2006

A **Figura 10** mostra os contornos de isopotencial da Ultra Cell e da cela padrão de ICR. Essas melhorias adicionais são significativas para a aquisição de transientes em períodos de tempos longos em espectros de ultra-alto poder de resolução (1.000,00 para *m/z* 400).

3.3 Desenvolvimento do *software*

Os algoritmos desenvolvidos foram escritos utilizando a linguagem de programação Python 2.7[25] para os protótipos iniciais e a linguagem Java SE 6[26] para o produto final. Todos os parâmetros peculiares do processamento serão ajustados para as condições necessárias tendo como base amostras reais. O arquivo de entrada utilizado será a lista de correlação *m/z* e intensidade exportada do *software* de processamento de dados Xcalibur 2.2 SP2 (ThermoScientific, Bremen, Germany).

[25] http://docs.python.org/ - © Copyright 1990-2011, Python Software Foundation.

[26] http://download.oracle.com/javase/6/docs/ - Copyright© 1995, 2010, Oracle and/or its affiliates.

4. Resultados e Discussão

A **Figura 11** mostra as etapas propostas de processamento de dados desde a entrada até sua classificação e organização.

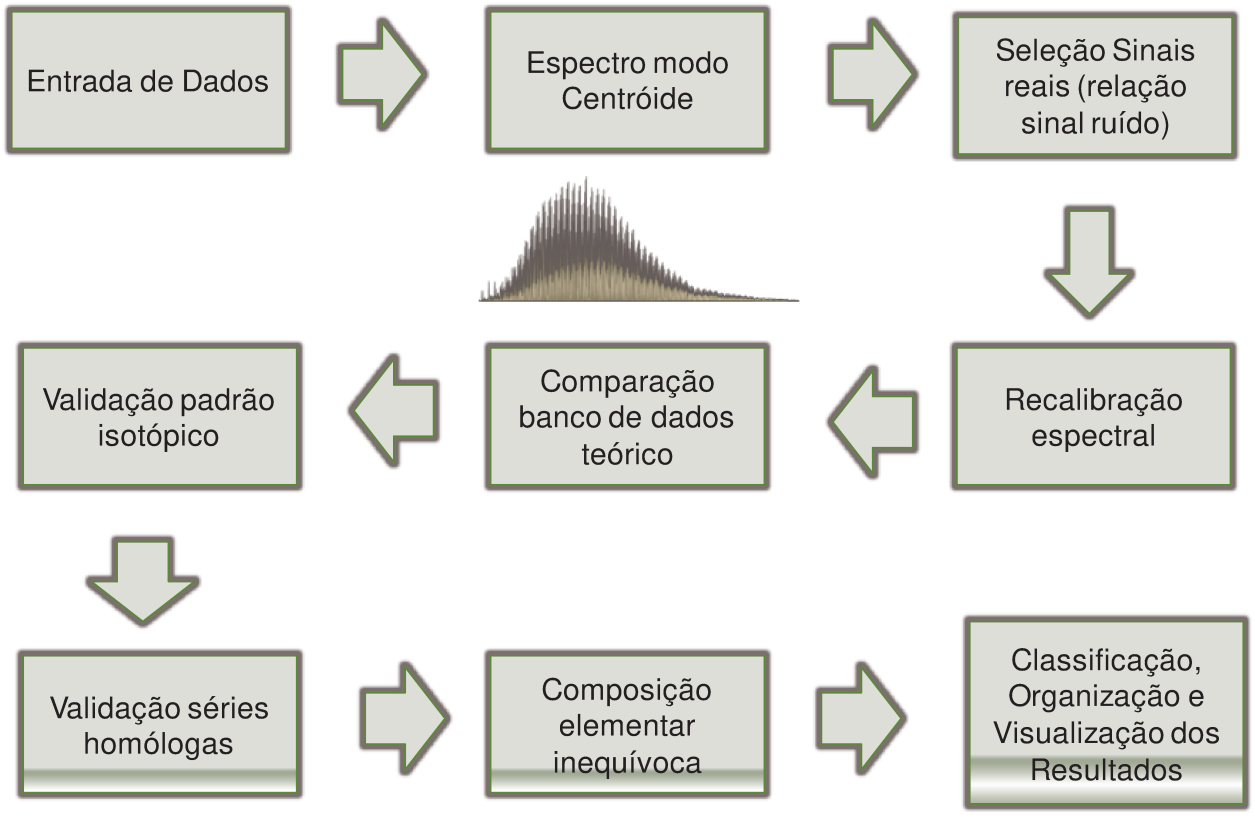


Figura 11. Esquema geral de processamento de dados para análise de dados de Petroleômica por FT-ICR MS.

4.1 Entrada de Dados e Transformação para Modo Centróide

Inicialmente os dados obtidos pelo equipamento foram exportados em modo de texto no modo contínuo onde duas colunas representam a *m/z* e as intensidades correspondentes detectadas pelo analisador ICR. Admite-se que os sinais detectados apresentam-se em forma gaussiana, onde na abscissa encontram-se valores de *m/z* e nas ordenadas os valores de intensidade referentes a populações iônicas correspondentes (**Figura 12**).

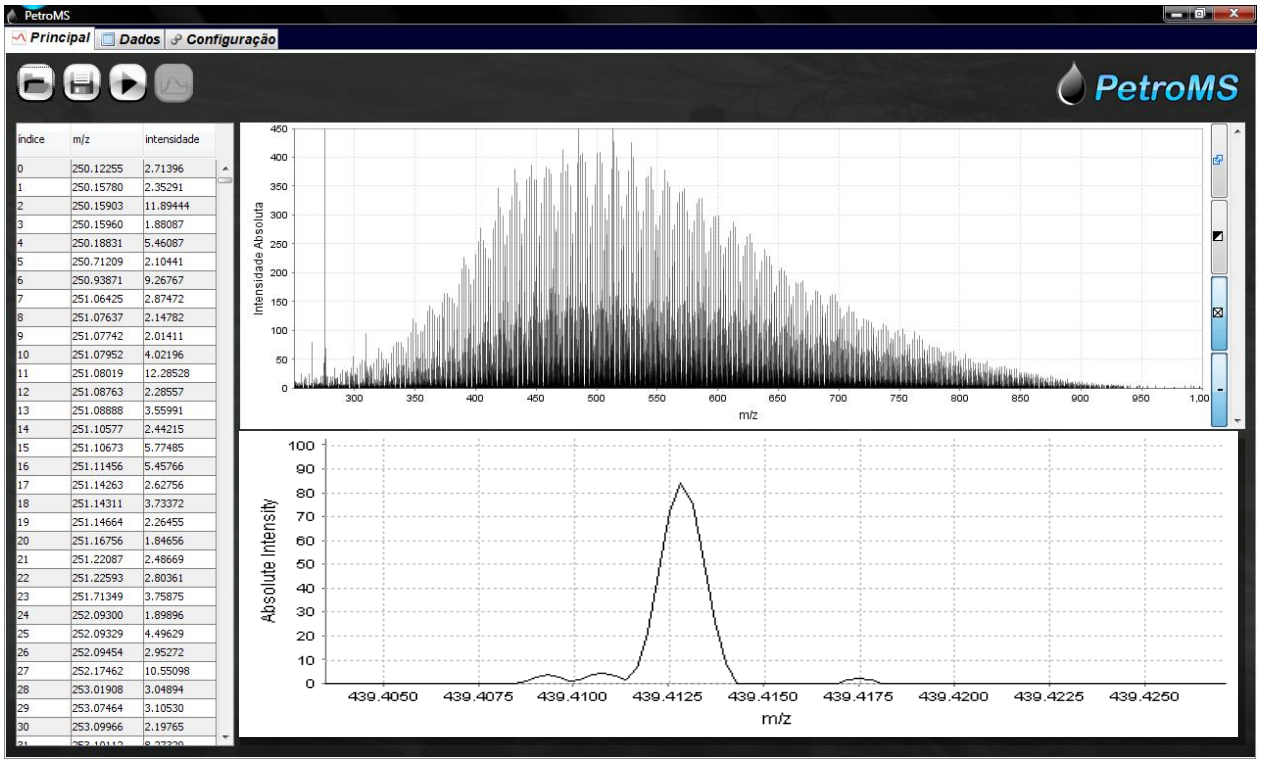


Figura 12. Pagina principal do *software* mostrando o espectro de ESI FT-ICR MS adquirido (acima) e a sua expansão para um intervalo de 0,022 *m/z* mostrando a forma do sinal adquirido.

Devido a possíveis assimetrias nessa forma gaussiana (**Figura 13**) do sinal detectado, utilizaremos a mediana da distância da abscissa (*m/z*) na meia altura do valor de intensidade (ordenada) do sinal detectado para a obtenção da *m/z*. Com isso, obtemos valores mais precisos para cada sinal detectado pelo equipamento.

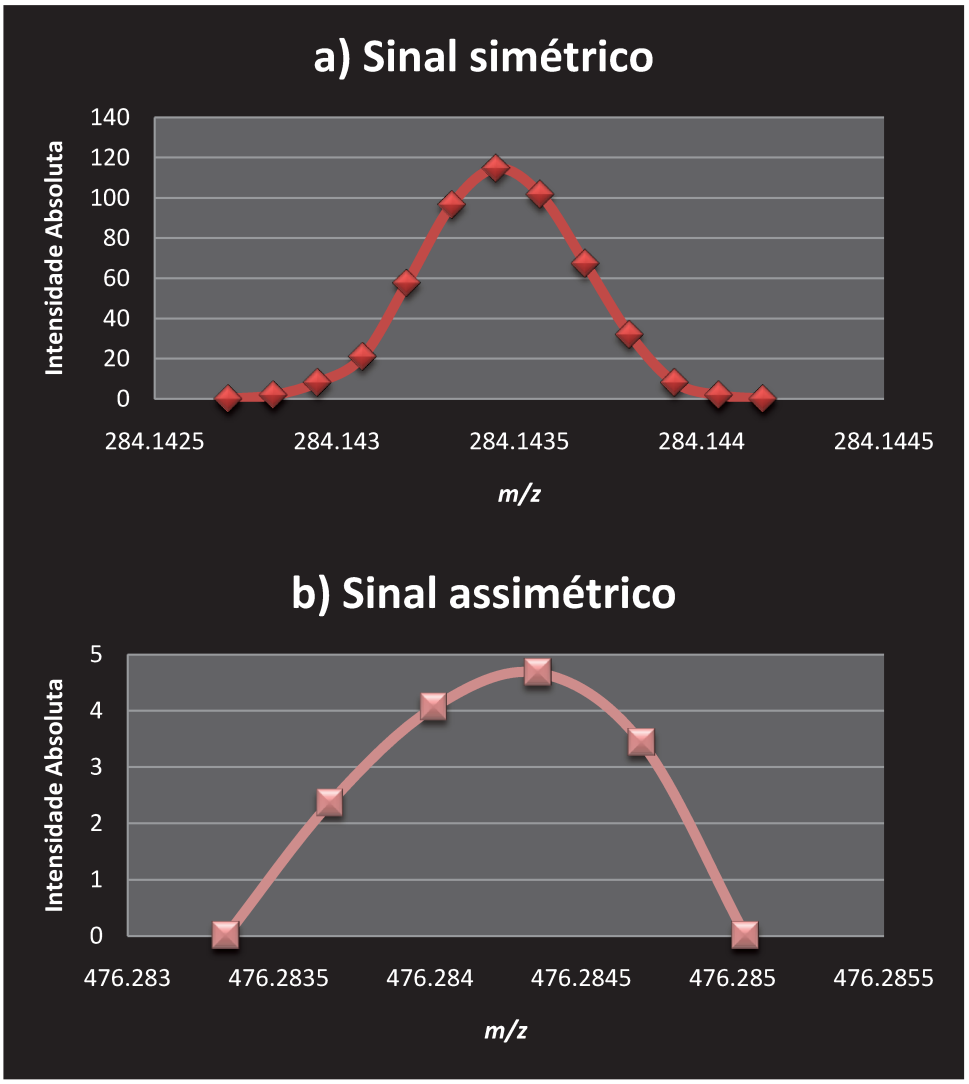


Figura 13. Dois sinais adquiridos por ESI FT-ICR MS mostrando a) um sinal com a forma gaussiana perfeita b) sinal assimétrico devido a baixa amplitude da intensidade (pouco pontos).

Para isso calculamos a distância na meia altura (Γ) do sinal partindo da função de Gauss:

$$e^{\frac{-(x_0-\mu)^2}{2\sigma^2}} = \frac{1}{2} f(x_{max}) \qquad \textbf{(Equação 3)}$$

onde $f(x_{max})$ ocorre quando $(x_{max}) = \mu$, logo:

$$e^{\frac{-(x_0-\mu)^2}{2\sigma^2}} = \frac{1}{2} f(\mu) = \frac{1}{2} \qquad \textbf{(Equação 4)}$$

resolvendo,

$$e^{\frac{-(x_0-\mu)^2}{2\sigma^2}} = 2^{-1} \qquad \textbf{(Equação 5)}$$

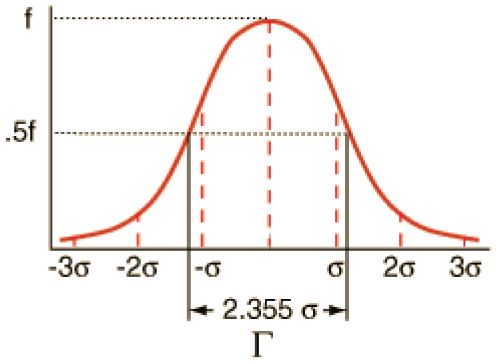
multiplicando pelo logaritmo natural (ln):

$$\frac{-(x_0-\mu)^2}{2\sigma^2} = -\ln 2 \qquad \textbf{(Equação 6)}$$

$$(x_0-\mu)^2 = 2\sigma^2 \ln 2 \qquad \textbf{(Equação 7)}$$

para média zero e como a distancia a meia altura (Γ) é igual a $2x_0$ temos:

$$\Gamma = 2\sqrt{2\ln 2}\,\sigma \approx 2{,}3548\,\sigma \qquad \textbf{(Equação 8)}$$



Utilizando o sinal simétrico apresentado na Figura 13 como exemplo calculou-se o desvio padrão a partir de um sexto da diferença do ultimo (x_2) e o primeiro ponto (x_1) do sinal (valores em vermelho na **Tabela 3**).

$$\sigma = (x_2 - x_1)\frac{1}{6} = (\mathbf{284{,}143437 - 284{,}142703})\frac{1}{6} = 0{,}0002445 \ \textit{m/z}.$$

Tabela 3 – Valores de *m/z* e intensidade para o sinal adquirido mostrado na **Figura 14**. Valores em vermelho utilizados para o calculo do desvio padrão (σ). Valor de *m/z* em verde correspondente ao máximo do sinal.

m/z	Intensidade
284,142703	0,00000
284,142826	1,77984
284,142948	8,133862
284,143070	20,86226
284,143192	57,58983
284,143315	96,61409
284,143437	114,73705
284,143559	101,45697
284,143681	67,37982
284,143803	32,05024
284,143926	8,13386
284,144048	1,77984
284,144170	0,00000
284,143437	0,00000

Com os valores do desvio padrão calcula-se os dois pontos da distancia h_1 e h_2 da meia altura como $h_1 = x_1 + \sigma$ e $h_2 = x_2 - \sigma$.

$$h_1 = (x_1 + 0{,}0002445) = \mathbf{284{,}142703} + 0{,}0002445 = 284{,}143126$$

$$h_2 = (x_2 - 0{,}0002445) = \mathbf{284{,}143437} - 0{,}0002445 = 284{,}143749$$

Assim a mediana da distância é igual a **284,143437** (valor em verde **Tabela 3**) que é exatamente, como esperado, o valor obtido experimentalmente para o máximo do sinal **(Figura 14)**. O mesmo procedimento foi realizado para todos os sinais adquiridos no espectro de massas. A **Figura 15** mostra um exemplo de correção do valor de *m/z* para o sinal assimétrico apresentado na **Figura 13.**

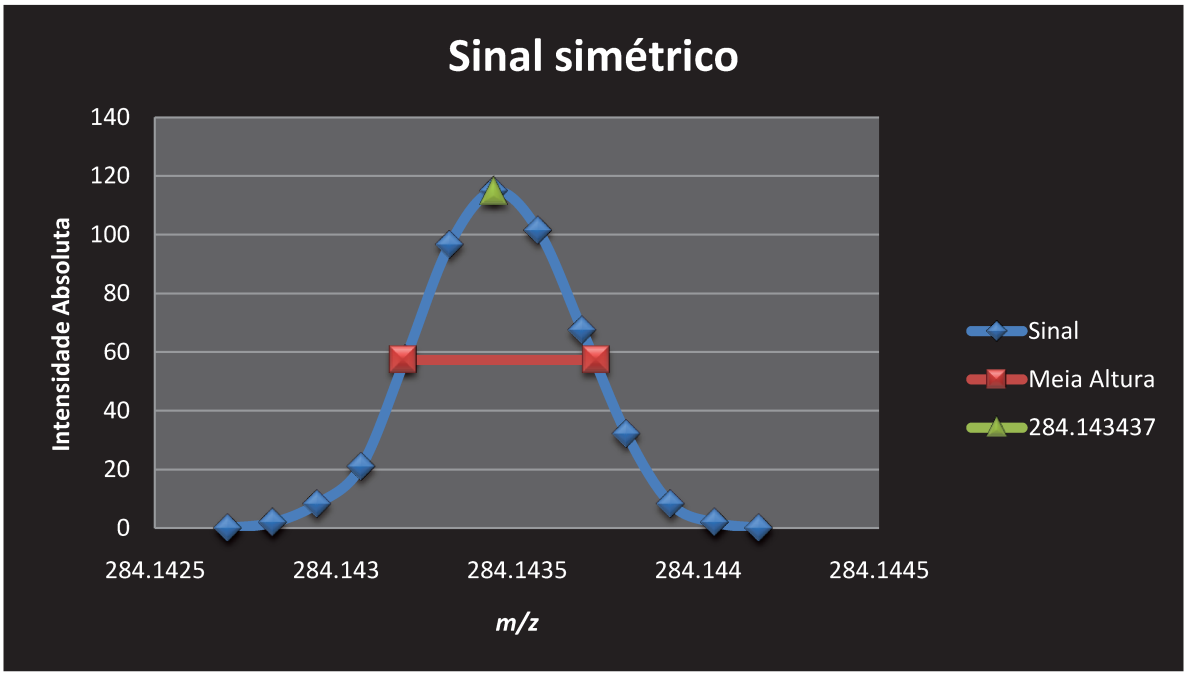


Figura 14. Sinal simétrico (azul), distância a meia altura (vermelho) e o ponto máximo calculado pela mediana da distância da meia altura (triangulo verde).

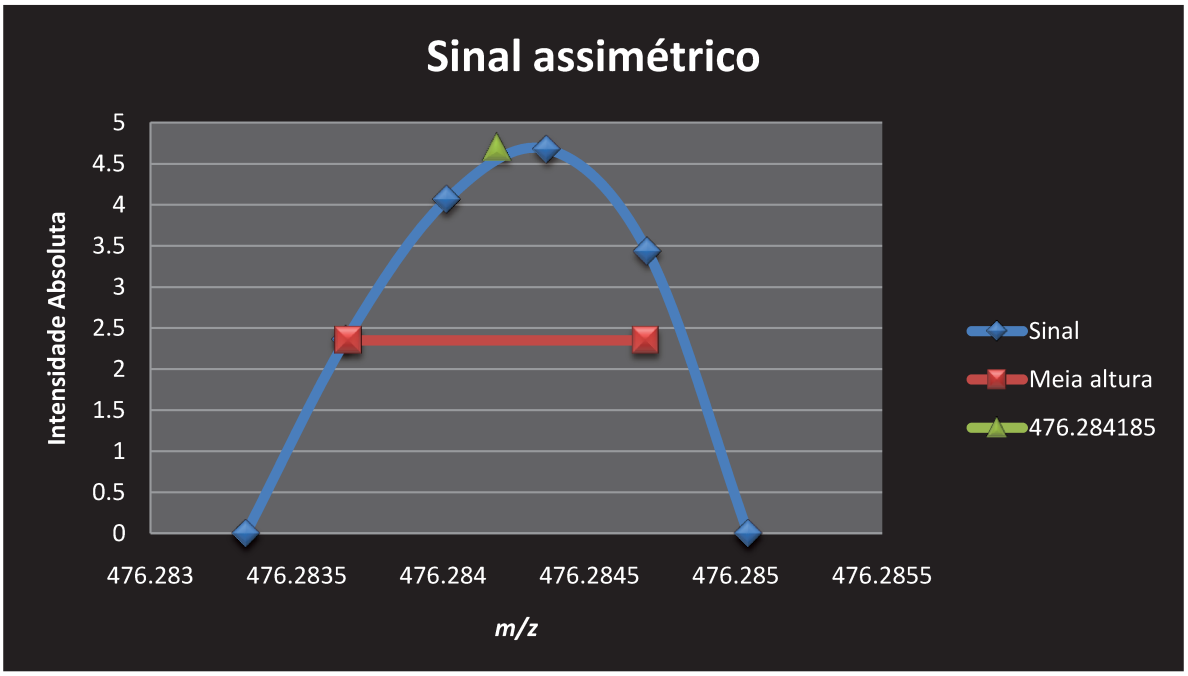


Figura 15. Sinal assimétrico (azul), distância a meia altura (vermelho) e o ponto máximo calculado pela mediana da distância da meia altura (ponto verde)

4.2 Filtragem de sinais pela relação sinal/ruído

Com a diminuição da amplitude da intensidade do sinal detectado pelo analisador ICR, é maior a possibilidade de distorção do sinal e deste sinal não corresponder a um íon real e sim a uma interferência eletrônica nos detectores do analisador (ruído). Essa interferência é de difícil análise manual, e quando os sinais estão em intensidade muito baixa fica impossível a distinção entre sinais reais e as prováveis interferências. Corriqueiramente se desprezam sinais com intensidade relativa abaixo, de aproximadamente, 5% dependendo do equipamento e das condições utilizadas.

Entretanto esta estratégia pode ser arriscada, pois a relação sinal ruído varia entre amostras, conseqüentemente as informações importantes sobre a amostra em questão podem ser perdidas. Sabendo-se disso, foi desenvolvido o procedimento de estabelecer o limiar de intensidade de sinais, que é capaz de atribuir para cada espectro individualmente qual é a intensidade relativa mínima aceitável onde se pode obter a confiabilidade de atribuição de *m/z*.

Primeiramente, admite-se que o comportamento das intensidades dos sinais detectados referentes às interferências siga uma distribuição normal com média igual a zero em intensidade. Toma-se então o menor sinal a cada três sinais em ordem crescente de *m/z* por todo o espectro. Com estes valores calcula-se o desvio padrão populacional a partir da fórmula:

$$\boldsymbol{\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}} \qquad \textbf{(Equação 9)}$$

mas como $\boldsymbol{\mu = 0}$ e a população está reduzida pela metade, temos:

$$\boldsymbol{\sigma = \sqrt{\frac{1}{2N}\sum_{i=1}^{N}(x_i)^2}} \qquad \textbf{(Equação 10)}$$

Como o valor de $\boldsymbol{\sigma}$ calculado estabelece-se o limiar da intensidade fazendo três vezes o valor de $\boldsymbol{\sigma}$**.**

As **Figuras 16, 17** e **18** mostram respectivamente três exemplos de boa, média e péssima relação sinal/ruído de diferentes espectros de petróleo adquirido através de ESI(+)-FT ICR-MS. Os histogramas mostram a porção lateral direita da distribuição normal dos ruídos, uma vez que o sinal foi pré-processado pelo *software* XCalibur onde a média dos menores sinais foi estabelecida como zero e os sinais abaixo desta media (negativos) foram previamente retirados. As linhas vermelhas (limiar fixo) e as linhas verdes (limiar normalizado) mostram claramente que a utilização do limiar fixo causaria no caso da **Figura 16** uma perda de informação por considerar sinais que provavelmente são íons e não interferências, como ruído. No caso da **Figura 18** levaria a utilização de prováveis ruídos como sinais de íons reais.

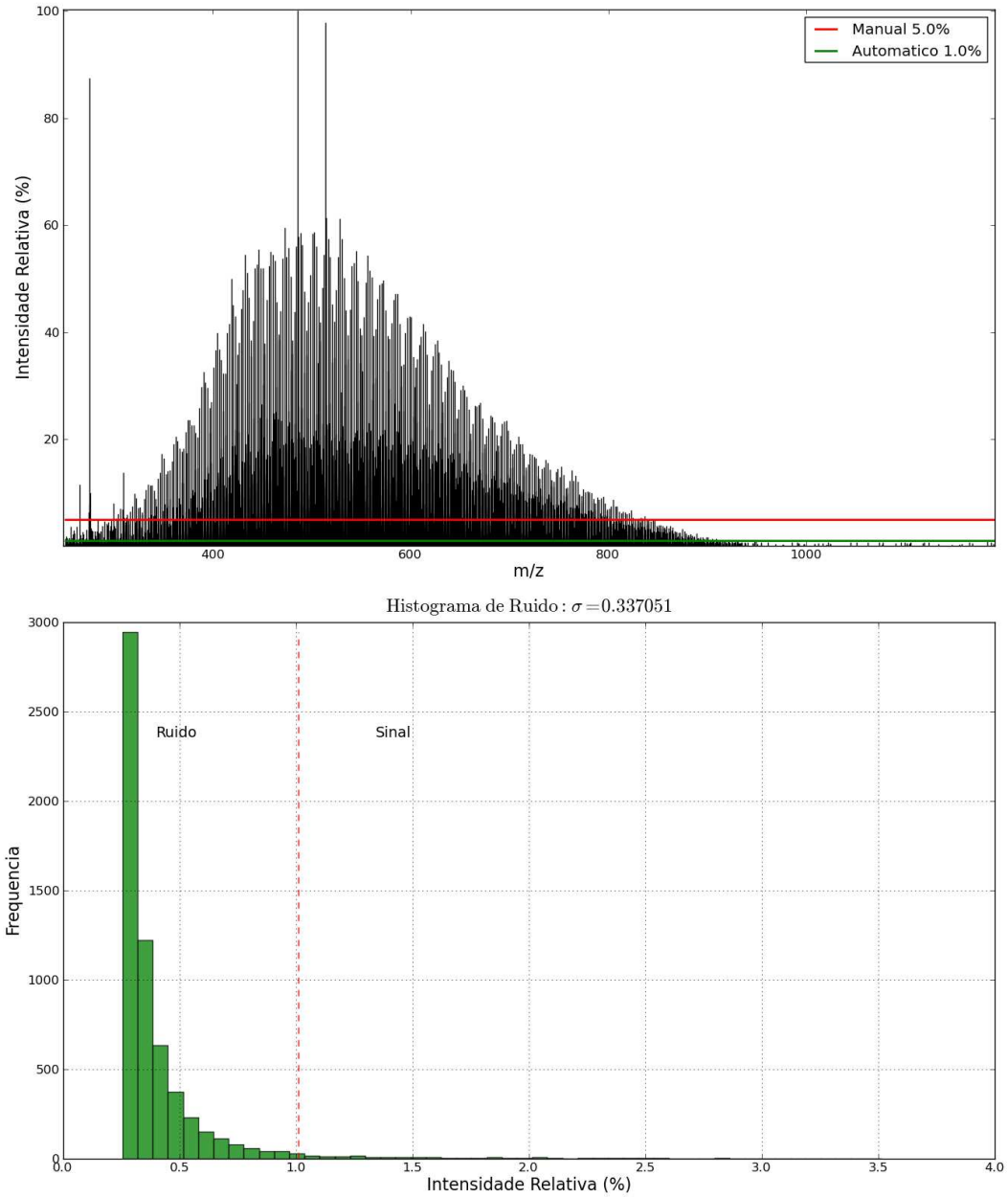


Figura 16. Espectro com alta relação sinal/ruído, mostrando o limiar estabelecido utilizando intensidade fixa de 5 % (linha vermelha – acima) e o limiar estabelecido utilizando a distribuição normal (histogramas – abaixo) e 3 vezes o desvio padrão (linha verde nos espectros e linha pontilhada vermelha nos histogramas).

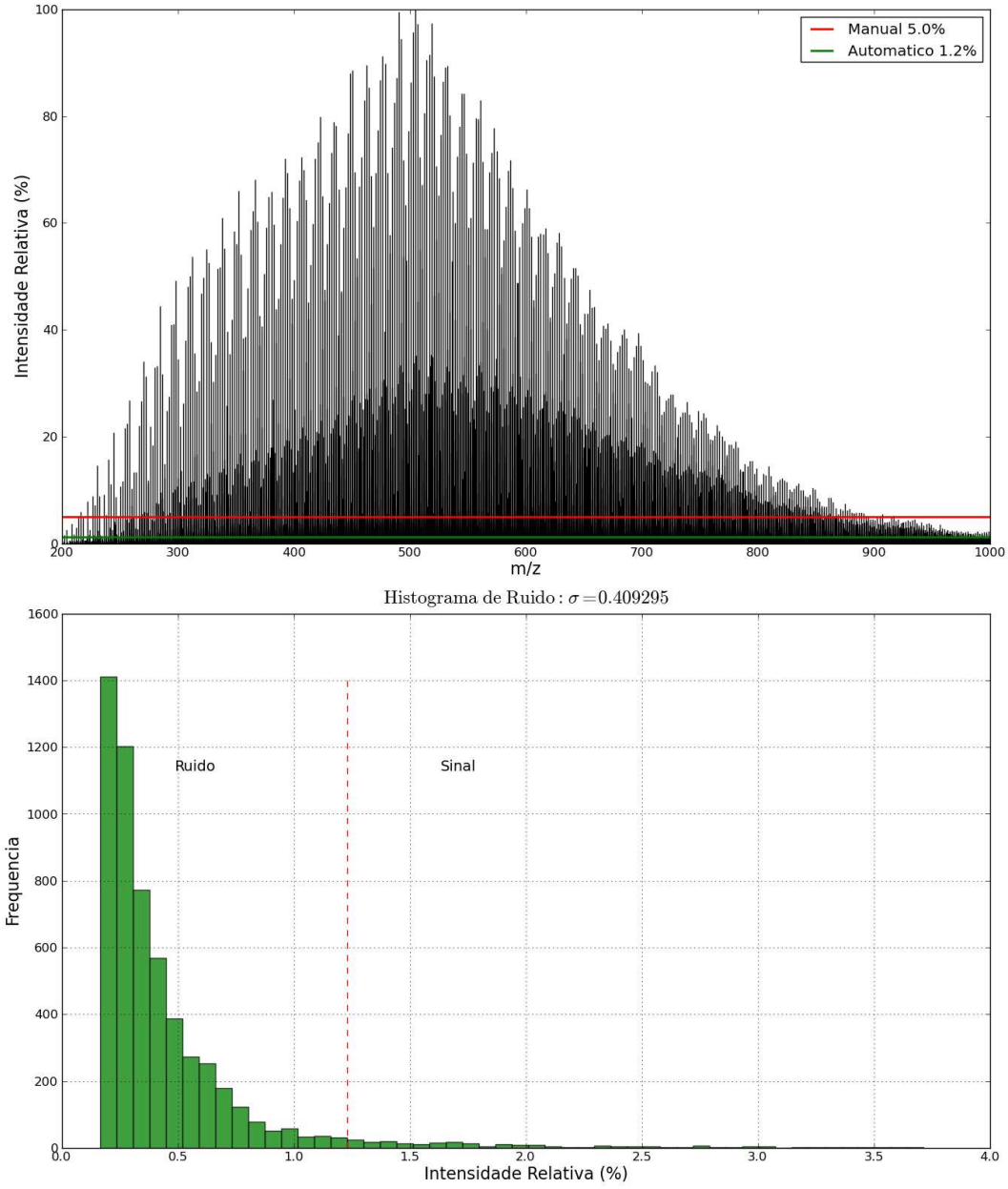


Figura 17. Espectro com média relação sinal/ruído, mostrando o limiar estabelecido utilizando intensidade fixa de 5 % (linha vermelha – acima) e o limiar estabelecido utilizando a distribuição normal (histogramas – abaixo) e 3 vezes o desvio padrão(linha verde nos espectros e linha pontilhada vermelha nos histogramas).

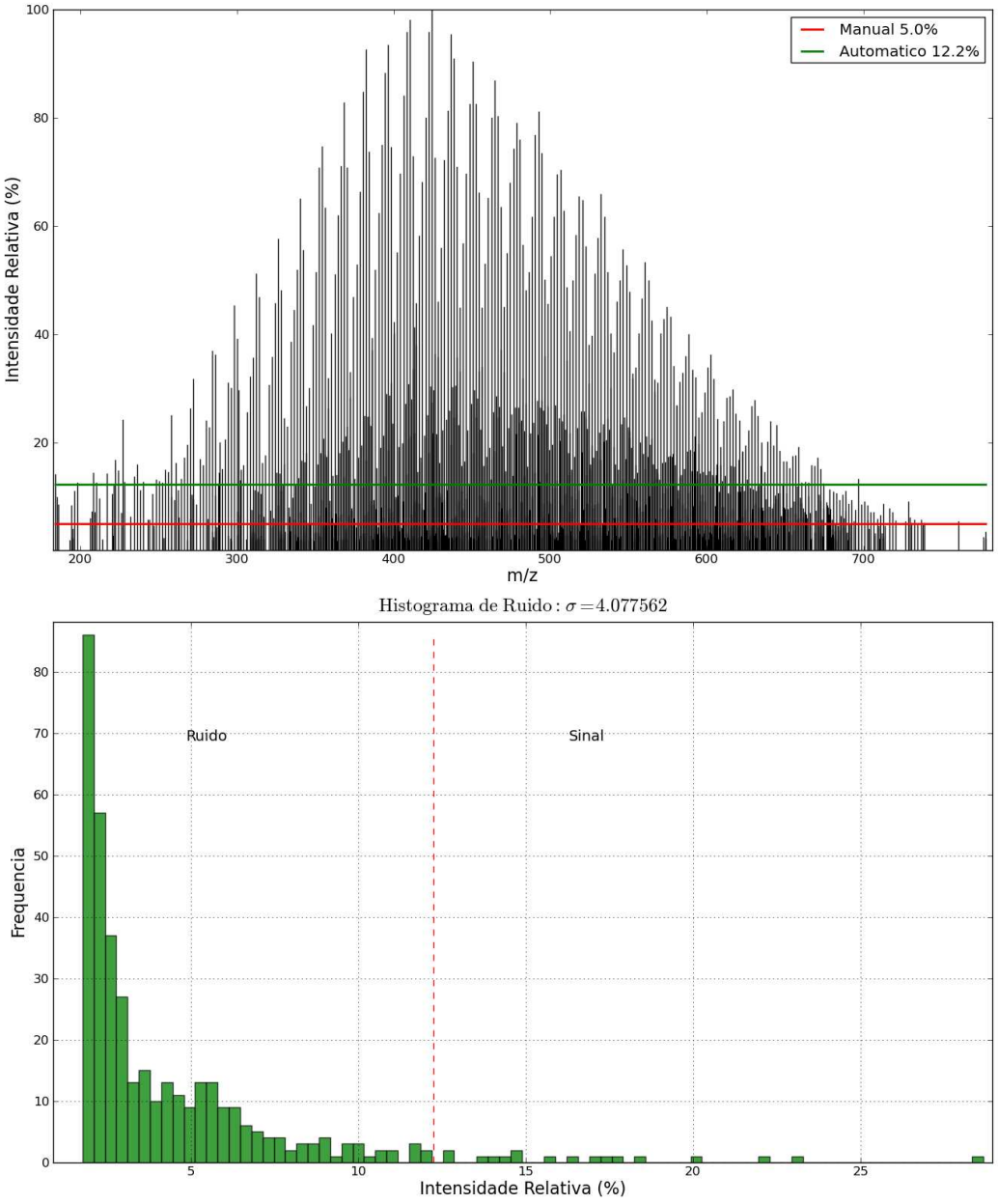


Figura 18. Espectro com baixa relação sinal/ruído, mostrando o limiar estabelecido utilizando intensidade fixa de 5 % (linha vermelha – acima) e o limiar estabelecido utilizando a distribuição normal (histogramas – abaixo) e 3 vezes o desvio padrão (linha verde nos espectros e linha pontilhada vermelha nos histogramas).

4.3 Definição de faixa de erro experimental

Como mostrado na Tabela 2, a definição da faixa do erro experimental na aferição da *m/z* é fundamental para a atribuição inequívoca da fórmula molecular correspondente. Quanto menor essa faixa de erro menor a quantidade de fórmulas moleculares cabíveis. Porém esta faixa de erro é afetada devido a vários fatores operacionais, entre eles destacam-se: o excesso de íons dentro da cela de ICR, calibração de massas mal sucedida e/ou antiga e baixa população de íons gerada através da ionização. A atribuição da faixa de erro aceitável é normalmente fixada de acordo com as especificações providas pelo fabricante do equipamento (< 1ppm para o caso do FT LQT MS Ultra). Entretanto esta faixa de erro varia a cada experimento e geralmente encontra-se em limites mais estreitos do que 1,0 ppm.

Para uma atribuição mais precisa desta faixa de erro, executa-se o procedimento em etapas descrito a seguir:

1) Verifica-se qual é a classe majoritária esperada para a amostra (geralmente N (piridínicos) para o modo positivo e N (carbazóis) ou O_2 (ácidos carboxílicos) para o modo negativo).

2) Atribui-se o valor de 1,0 ppm para a faixa máxima de erro de massas e elimina-se sinais com intensidade menores que 5% ou com a intensidade calculada a partir da relação sinal/ruído.

3) Os valores experimentais restantes são comparados com os valores teóricos do banco de dados para as séries homólogas da classe hetero-atômica especificada.

4) A partir destas atribuições se obtém o erro experimental em ppm para todas as observações e a faixa de erro para atribuição das classes restantes é atribuída pelo menor e maior valor encontrado.

4.4 Recalibração de Massas

Visando corrigir fatores que influenciam a exatidão das medidas de *m/z* pelo analisador ICR, várias equações foram propostas para permitir a correção destes valores [27]. Em 2010 Marshall e colaboradores[28] propuseram a extensão destas equações incluindo o termo ($\boldsymbol{I}$) relativo a população iônica:

$$\boldsymbol{m/z = A + B/f + C/f^2 + D\ I\ /f^2} \qquad \textbf{(Equação 11)}$$

onde ($\boldsymbol{A}$) equivale ao termo de correção relativo a não linearidade do campo magnético ao longo da cela de ICR, (**B)** equivale ao termo de correção relativo ao campo elétrico (oscilação em relação ao eixo Z), (**C**) equivalente a população iônica , ($\boldsymbol{f}$) é a freqüência ciclotrônica dos íons dentro da cela de ICR e ($\boldsymbol{I}$) corresponde a intensidade relativa desta frequência observada.

Como a frequência ciclotrônica é inversamente proporcional a *m/z* obtemos a equação

$$\boldsymbol{m/z = A + B(m/z) + C(m/z)^2 + D\ I\ (m/z)^2} \qquad \textbf{(Equação 12)}$$

[27]Ledford Jr, E. B.; Rempel, Don L.; Gross, M. L. **Space charge effects in Fourier transform mass spectrometry. Mass calibration**. *Analytical Chemistry (*1984)**,** 56 (14), 2744–2748

[28] Savory, J.J.; Kaiser, N.K.; McKenna, A.M.; Blakney, G.T.; Hendrickson, C.L.; Rodgers, R.P.; Xian, F. and Marshall, A.G. ***A Walking Mass Calibration Equation for Complex Mixture Analysis by FT-ICR Mass Spectrometry.*** 58th Annual Conference on Mass Spectrometry & Allied Topics, Salt Lake City, UT, May 23-27 (2010)

A partir desta equação foi construída uma curva de calibração entre os valores de *m/z* experimental e os valores de *m/z* teóricos, sendo os valores teóricos estipulados como a série homóloga mais abundante do espectro (geralmente N para modo positivo e O2 ou N para o modo negativo de íons - **Figura 19**). A solução da equação foi realizada utilizando o método de regressão por mínimos quadrados, através da equação de matricial [29]:

$$\mathbf{y = X\ b + e} \qquad \textbf{(Equação 13)}$$

onde a solução para os coeficientes utilizando o método dos mínimos quadrados:

$$\sum \mathbf{e^2 = 0} \qquad \textbf{(Equação 14)}$$

equivale a:

$$\boldsymbol{b} = (\mathbf{X^t\ X})^{-1}\ \mathbf{X^t\ y} \qquad \textbf{(Equação 15)}$$

onde ($\boldsymbol{b}$) é a matriz resultante 1x3 dos coeficientes A, B, C e D;

$$\boldsymbol{b} = \begin{vmatrix} \boldsymbol{A} \\ \boldsymbol{B} \\ \boldsymbol{C} \\ \boldsymbol{D} \end{vmatrix} \qquad \textbf{(Equação 16)}$$

[29] Neto, B. B; Scarminio, I. S.; Bruns, R. E. **Como fazer experimentos**. Editora Unicamp, Campinas – SP, 2008.

(X) é a matriz contendo os valores das ***(n)*** variáveis (***m/z* e intensidade (I)**)

$$X = \begin{vmatrix} 1 & m/z & (m/z)^2 & (m/z)^2 * I \\ \vdots & \vdots & \vdots & \vdots \\ 1_n & m/z_n & (m/z)^2{}_n & (m/z)^2{}_n * I_n \end{vmatrix} \qquad \textbf{(Equação 17)}$$

e (y) é a matriz dos valores de **m/z** teóricas [11].

$$y = \begin{vmatrix} m/z \\ \vdots \\ m/z_n \end{vmatrix} \qquad \textbf{(Equação 18)}$$

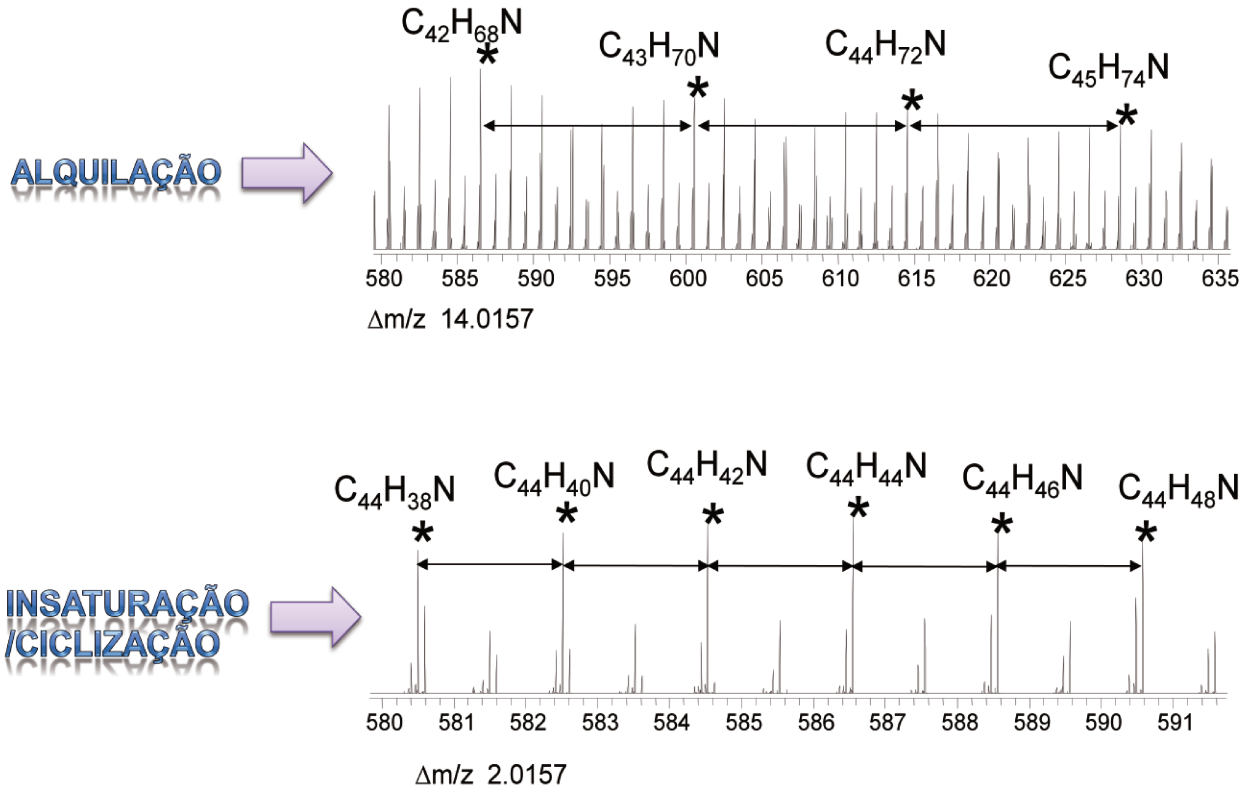
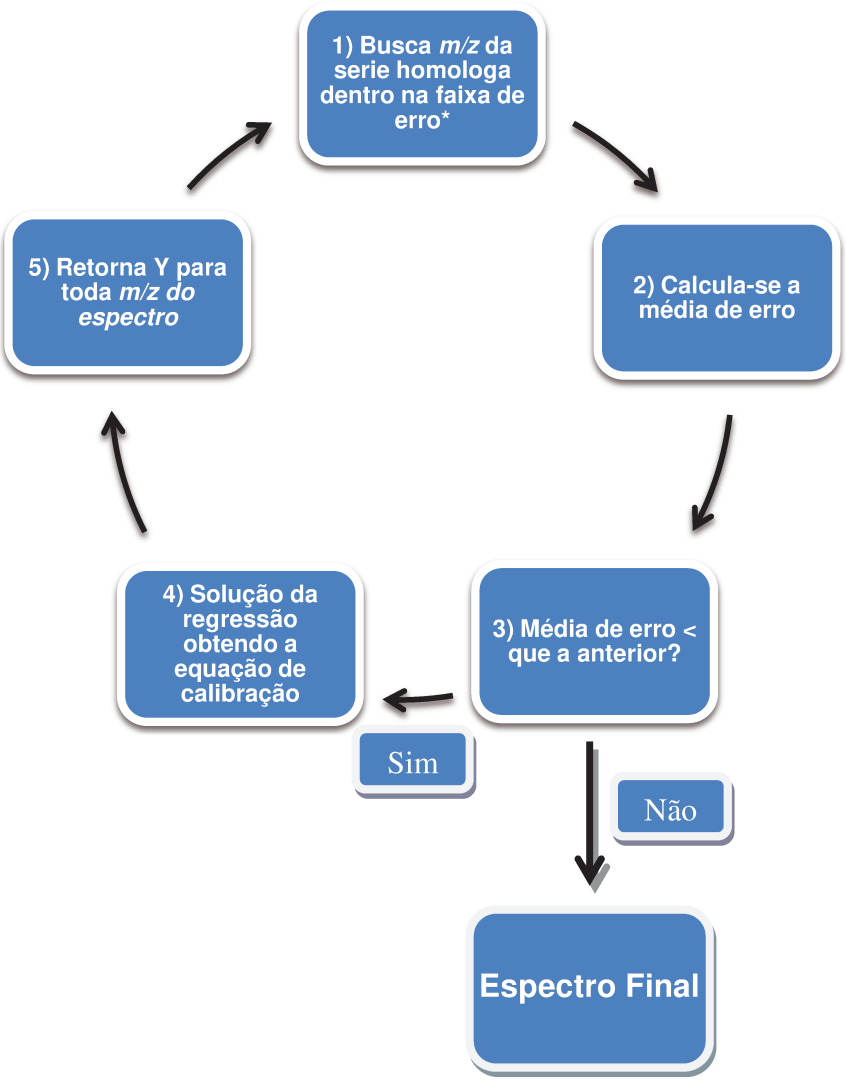


Figura 19. Expansão para os espectro de ESI(+)-FT ICR-MS mostrando (acima) unidades de repetição CH_2 e (abaixo) unidades de repetição de 2H.

Com os valores da equação da curva de calibração, obtêm-se os valores de *m/z* recalibrados para todo o espectro. O **Fluxograma 1** mostra as etapas envolvidas no processo de recalibração onde:



Fluxograma 1 – Etapas do processo de recalibração para todo o espectro de massas.

1) Procura-se por valores de *m/z* experimentais dentro de um banco teórico com valores de *m/z* correspondentes a componentes das séries homólogas mais abundantes (Figura 7) que se encontrem dentro de uma faixa de erro (ppm):

$$\text{erro} = \left(\frac{\text{m/z}_{\text{teórico}} - \text{m/z}_{\text{experimental}}}{\text{m/z}_{\text{teórico}}}\right) 10^6 \qquad \textbf{(Equação 19)}$$

sendo esta estipulada pelo usuário ou utiliza-se a faixa de erro encontrada no processo de definição automática .

2) Com os valores de erro que se enquadrarem nesta faixa calcula-se a média destes.

3) Se a média de erro calculada for menor que a média de erro do processo anterior segue o fluxo de processamento para obtenção da equação de regressão. Caso contrário as *m/z* atuais são consideradas as *m/z* finais da recalibração. Para o primeiro ciclo estabelece-se um valor alto de erro (50 ppm) como artifício para prosseguir o fluxo de processamento.

4) Com os valores encontrados de *m/z* resolve-se a equação de regressão para a obtenção dos valores dos coeficientes A, B, C e D e por final obtém-se a equação de regressão:

$$\boldsymbol{m/z} = \boldsymbol{A} + \boldsymbol{B}(\boldsymbol{m/z}) + \boldsymbol{C}(\boldsymbol{m/z})^2 + \boldsymbol{D\ I}\ (\boldsymbol{m/z})^2 \qquad \textbf{(Equação 12)}$$

3) Calcula-se o novo valor de *m/z* resolvendo a equação para todos os valores de *m/z* experimentais. Estes valores são novamente inseridos no processo até que a média de erro dos valores de *m/z* chegue a um mínimo ou se iguale a zero ou o valor do desvio padrão aumente em relação a medida anterior.

A **Figura 20** (esquerda) mostra a dispersão do erro em relação a *m/z* para um espectro de ESI(+)-MS onde observa-se uma baixa exatidão de massas, sendo a média de erro de 0,88 ppm e o desvio padrão de 0,41 ppm. Após a realização da recalibração foi obtido uma média de erro de 0,0003 ppm ou 0,3 ppb e um desvio padrão de 0,35 ppm (Figura 20 – direita).

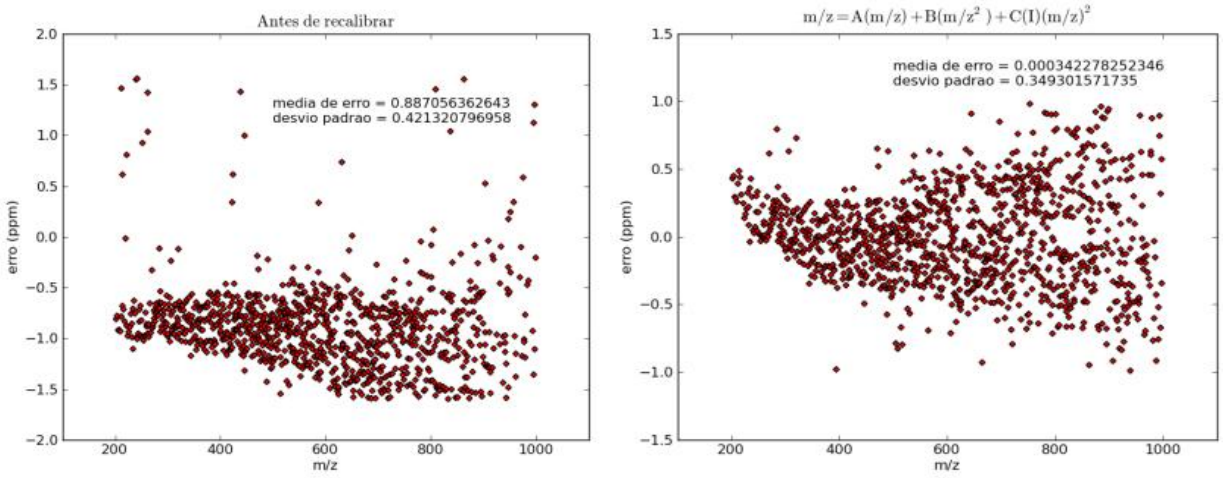


Figura 20. Dispersão de erro de massas para espectro de ESI(+)- FT ICR-MS para o modo positivo de íons antes e depois da recalibração.

Com estes novos valores minimiza-se a possível duplicidade de atribuição da composição elementar para cada sinal registrado no espectro de massas e aumenta a quantidade de espécies com suas respectivas composições elementares estabelecidas após procura destas no banco de dados teóricos (**Figura 21**).

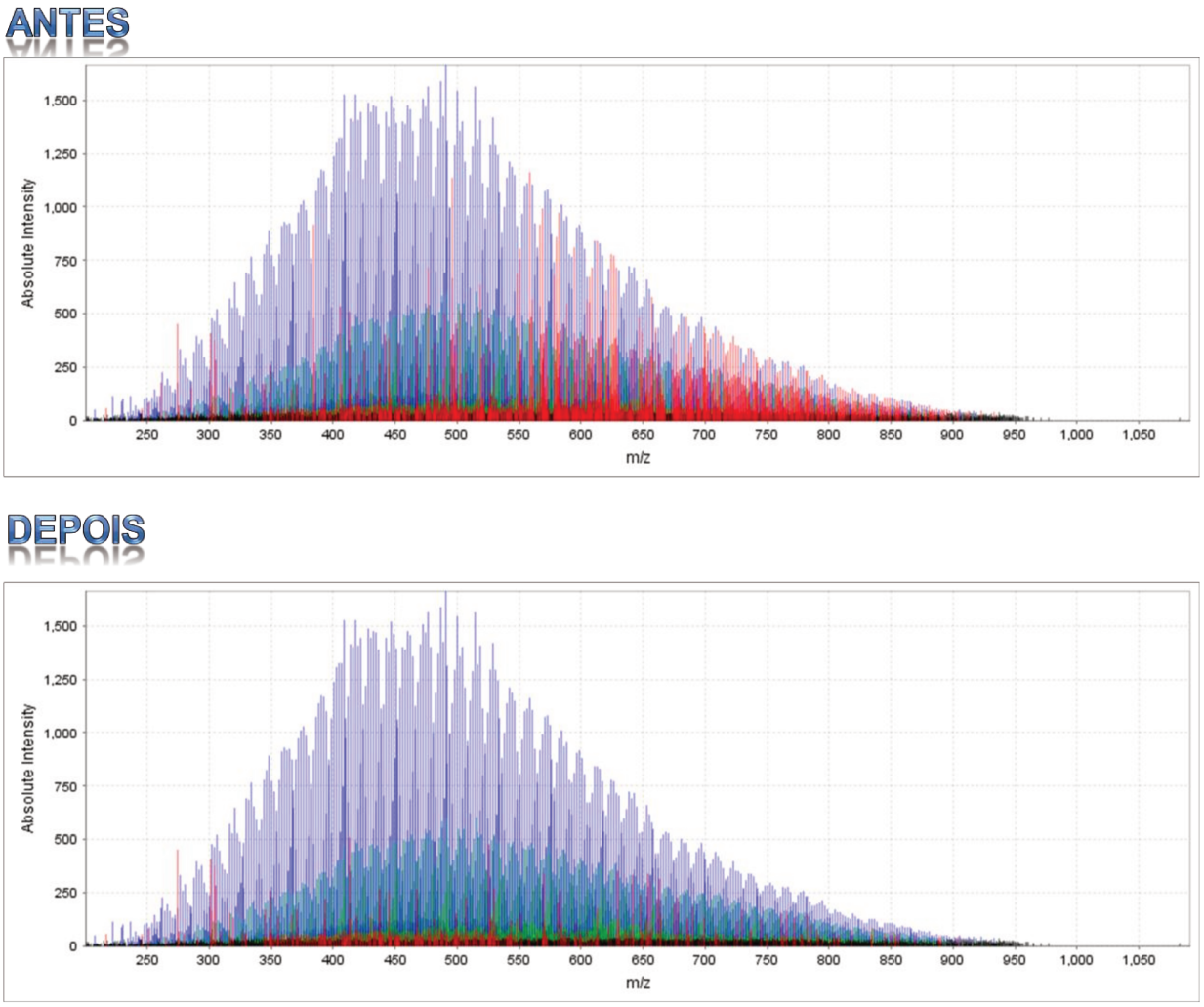


Figura 21. Espectro de ESI(+)-FT-ICR MS mostrando os sinais nos quais houve a atribuição de sua composição elementar (linhas azuis), seus isotopólogos de ^{13}C e ^{34}S (linhas verdes) e os não identificados (linhas vermelhas) antes e após o processo de recalibração.

4.5 Atribuição de Fórmula Molecular

A etapa crucial para a identificação a nível molecular dos mais de vinte mil sinais resolvidos pelo espectrômetro de massas é a comparação dos valores de *m/z* experimentais com valores de *m/z* teóricos relacionados com a respectiva fórmula molecular através da comparação de erro de massas a partir da **(Equação 19)**. Para isso foi construído um banco de dados contendo todas as possíveis composições elementares utilizando como base os átomos de carbono, hidrogênio, nitrogênio, oxigênio e enxofre, sendo estes os principais átomos componentes das espécies presentes na parte polar do petróleo e seus derivados (**Figura 22**)[3].

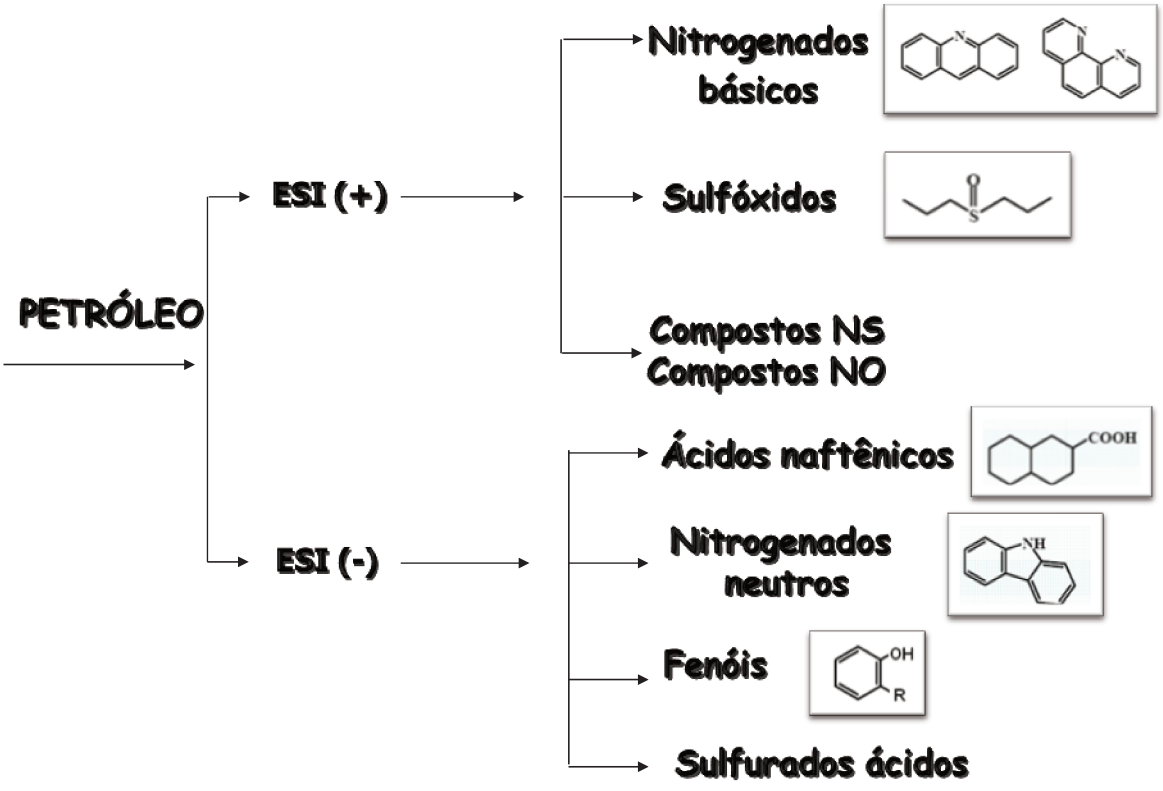


Figura 22. Principais classe de compostos polares presentes no petróleo detectados por ESI(±)-FT-ICR MS

4.5.1 Classificação de espécies polares no petróleo

A classificação adotada refere-se à classificação por classes e tipo. Estas classes são denominadas em relação ao heteroátomo X presente na espécie partindo da fórmula empírica $C_cH_hX_x$, onde a quantidade X de heteroátomos na molécula determina a sua classe. Por exemplo, a molécula com fórmula molecular $C_{10}H_{17}N$ pertencerá à classe N, da mesma forma que a molécula com fórmula molecular $C_{10}H_{21}N$. A molécula $C_{10}H_{18}O_2S$ será definida como pertencente à classe O_2S da mesma forma que a molécula de fórmula molecular $C_{10}H_{22}O_2S$ e assim por diante.

O tipo é definido a partir do valor da DBE, expressão que vem do inglês *double bond equivalent* (DBE), sendo este definido a partir da fórmula empírica $C_cH_hN_nO_oS_s$ como:

$$\boldsymbol{DBE = c - \frac{h}{2} + \frac{n}{2} + 1} \qquad \textbf{(Equação 20)}$$

Este valor é diretamente correlacionado com a quantidade de anéis e insaturações da molécula em questão. Por exemplo: a molécula de fórmula molecular $C_{10}H_{17}N$ será classificada como classe N e de tipo 3 e a molécula de fórmula molecular $C_{10}H_{21}N$ como classe N e tipo 1. O mesmo acontece para moléculas pertencentes à classe O_2S as quais são representadas pelas fórmulas moleculares $C_{10}H_{18}O_2S$ e $C_{10}H_{22}O_2S$ sendo estas pertencentes aos tipos 2 e 0 respectivamente. Para facilitar a visualização uma representação compacta é adotada como tipo mais classe. Logo a molécula com formula molecular $C_{10}H_{22}O_2S$ será representada como 0-O_2S assim como a molécula $C_{10}H_{17}N$ será representada como 3-N.

Esta classificação é útil para a visualização de dados entre diferentes tipos, de amostras de petróleo mostrando a variação da série homóloga alquílica, comum neste produto.

O banco foi dividido em classes formando sub-bancos com todas as fórmulas moleculares possíveis e suas respectivas massas exatas teóricas. Com os dados do espectro de massas importado, a correlação massa/fórmula molecular é realizada por comparação de *m/z* experimentais com as *m/z* teóricas em todos os sub-bancos, um por vez. Ao final, caso a *m/z* experimental obtiver apenas uma atribuição possível dentro de uma faixa de erro aceitável (três vezes o desvio padrão de erro calculado na recalibração ou estipulado pelo usuário), essa *m/z* é correlacionada com a respectiva fórmula molecular e o resultado é salvo em vetores para futura manipulação:

$$\boldsymbol{Vetor = [m/z_{exp}\{\,[classe,\ m/z_{teo}\ ,\ fórmula]^{n}\ \}\ ]}$$

sendo **(n)** equivalente a quantidade de possíveis atribuições correlacionadas com esta *m/z* experimental.

4.5.2 Banco de Dados Teóricos

Diferentemente da atribuição de fórmulas moleculares a partir de um universo ilimitado de combinações de átomos e das possíveis combinações entre eles, a estratégia utilizada para a atribuição parte da comparação com um banco de dados teórico, no qual restrições em relação à massa mínima e máxima, número mínimo e máximo do valor de DBE, número mínimo e máximo de átomos de carbono, hidrogênio, nitrogênio e enxofre foram realizados. Para isso algumas condições devem ser cumpridas. Entre elas estão:

1) A faixa de *m/z* deve ser abrangente a ponto de ser capaz de analisar moléculas de massa 150 até 1400 Da (Observação experimental)

2) O banco de dados deve abranger todas as classes de heteroátomos (NOS) e todas as suas combinações possíveis (ex: N, N2, O, O2, O3, OS, O2S e assim por diante).

3) A restrição do valor mínimo de DBE parte do princípio de que para moléculas neutras o número máximo de hidrogênio por número de carbonos é delimitado, a partir da fórmula empírica de hidrocarbonetos saturados $C_aH_{2a+2,}$ por:

$$\mathbf{DBE = a - \frac{2a+2}{2} + 1 = 0} \qquad \textbf{(Equação 21)}$$

Assim a quantidade máxima de hidrogênio, em moléculas totalmente saturadas, correlacionada ao número de carbonos leva a um DBE mínimo de valor igual

a zero. Tal resultado já era esperado uma vez que o valor de DBE reflete diretamente o número de anéis e insaturações de uma molécula, sendo que esta não poderia conter valores negativos.

4) Moléculas maiores podem potencialmente conter mais anéis e insaturações que moléculas pequenas, o valor máximo de DBE cresce com o valor de átomos de carbono. Além disso, como a existência de ligações com hibridização ***sp*** é rara em compostos presentes no petróleo (assim como também o é para outras substâncias naturais) a razão DBE/C máxima estabelecida foi equivalente a 1,0 (compostos poli-aromáticos) e o valor máximo de DBE estabelecido para 40 unidades.

5) Todas as espécies transferidas para a fase gasosa através da técnica de *electrospray* encontram-se protonadas $[M + H]^+$ para o modo positivo de íons e desprotonadas $[M – H]^-$ para o modo negativo de íons. Nos dois modos as espécies estão presentes em suas formas mono carregadas.

6) As fórmulas moleculares das espécies são sempre feitas em relação à molécula neutra, e a sua massa representada pelo valor das espécies protonadas/desprotonadas. Ou seja, a soma do valor individual de cada átomo constituinte da fórmula molecular mais a soma de um elétron para o modo negativo de aquisição e a subtração do valor de um elétron para o modo positivo.

Essas condições foram elaboradas a partir da observação experimental de espectros adquiridos de 10 amostras de petróleo. Estas amostras foram escolhidas com a finalidade de representar um universo diversificado. Este universo é abrangente para

variáveis como grau de maturação, biodegradabilidade, diferentes poços e diferentes profundidades.

As **Figuras 23** e **25** mostram os fluxogramas do processo de criação do banco de dados teórico sendo este dividido em duas etapas (Parte I e II) para a classe N, ou seja, espécies contendo apenas um átomo de nitrogênio.

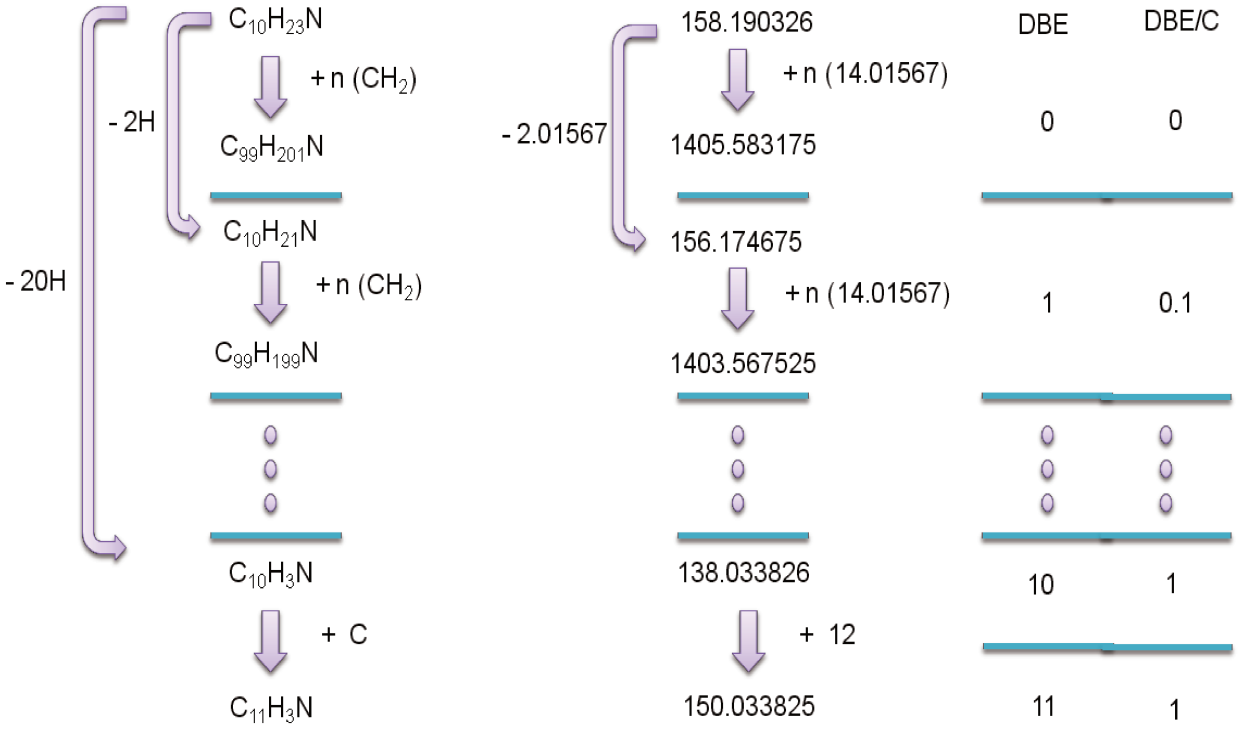


Figura 23. Fluxograma mostrando o início do processo de construção do banco de dado teórico para a classe N (fórmulas moleculares com DBE de 0 até 10).

Para a inicialização do algoritmo deve-se primeiro introduzir a fórmula base da classe hetero-atômica a qual se pretende criar ($C_{10}H_{23}N$ – **Figura 24**) obedecendo duas condições:

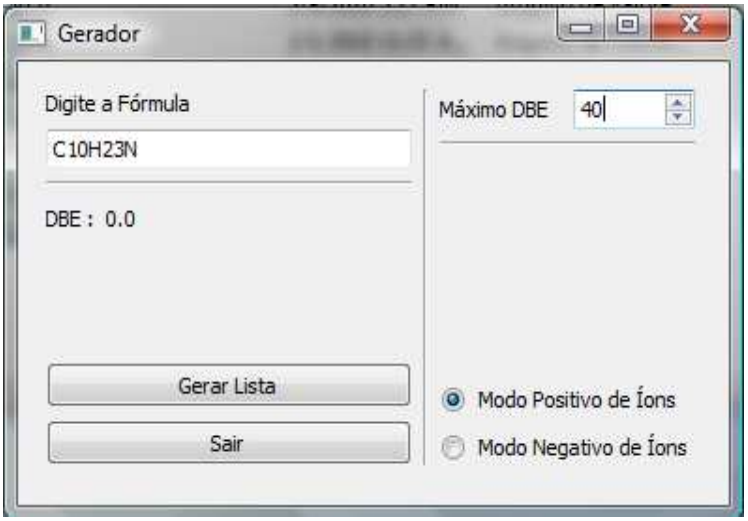


Figura 24. Interface gráfica do módulo de criação dos bancos de dados teóricos mostrando a fórmula molecular inicial, o respectivo valor de DBE calculado dinamicamente e as opções de máximo de DBE e o modo de íons a ser utilizado.

1) Inicializar com fórmula molecular com o valor de DBE igual a 0

2) Iniciar com fórmula molecular observando a regra do nitrogênio onde moléculas com numero impar de nitrogênio tem numero impar de hidrogênios.

A partir desta fórmula molecular inicial somam-se consecutivos valores de 14,01567, correspondentes à adição de uma unidade metilênica ($–CH_2$), até que o valor de *m/z* correspondente ultrapasse *m/z* 1400 ($C_{101}H_{183}N$ – *m/z* 1411.442325). Quando este primeiro limite é alcançado temos todas as possíveis fórmulas moleculares para espécies protonadas com valores de DBE correspondentes a zero, ou seja, todas as espécies totalmente saturadas.

Para obter as fórmulas moleculares com DBE igual a um, ou seja, espécies apresentando uma hibridização sp^2 entre os átomos de carbono e hidrogênio, ou

formados por um anel, retiram-se dois átomos de hidrogênio da fórmula molecular inicial. No caso da **Figura 23** temos a fórmula molecular $C_{10}H_{21}N$. A partir desta repete-se a etapa de retirada de unidades metilênicas até atingir o limiar estabelecido em relação à *m/z*.

Este procedimento se repete até que a razão DBE/C seja igual a 1 e então inicia-se a segunda parte deste algoritmo que é exemplificado na **Figura 25**.

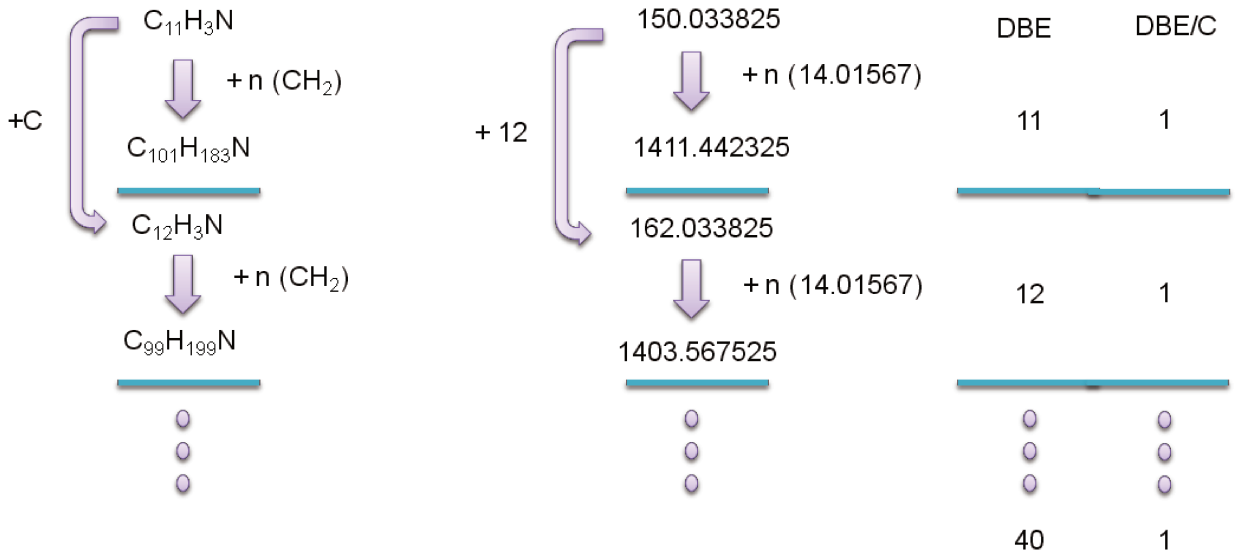


Figura 25. Fluxograma mostrando a segunda parte do processo de construção do banco de dado teórico para a classe N (fórmulas com DBE de 11 até 40).

Depois de alcançado o valor de contorno, no caso da razão DBE/C igual a 1, ao invés de retirar duas unidades de hidrogênio da fórmula inicial, acrescenta-se uma unidade de carbono, obtendo um ganho de uma unidade de DBE. Assim como na parte I do algoritmo acrescentam-se então unidades metilênicas a fim de criar todas as fórmulas possíveis com este DBE, e novamente segue-se até que a fórmula molecular

alcance um limite de *m/z* 1400. Quando este limite é alcançado adiciona-se novamente uma unidade de C (12.00000 Da) a fórmula inicial anterior e repete-se o acréscimo de unidade metilênicas. Este procedimento é repetido até atingir o limite máximo estabelecido para valores de DBE (40 unidades).

Ao final tem-se um arquivo de todas as fórmulas moleculares possíveis com suas massas correspondentes calculadas para a classe hétero atômica em questão. Este procedimento foi realizado para todas as classes de heteroátomos possíveis levando em consideração os átomos de N, O e S onde os números máximos estabelecidos foram 4, 10 e 4 respectivamente.

Entretanto, se mesmo com a faixa de erro limitada e a recalibração realizada, houver mais de uma possibilidade de fórmula molecular compatível com um valor de *m/z* aferido, são então realizados procedimentos para tentar eliminar esta duplicidade. São eles em ordem do fluxo de processamento:

1) procura de padrão isotópico para ^{13}C e ^{34}S. Caso os sinais dos isotopólogos estejam abaixo da intensidade mínima aceitável (ruído) realiza-se:
2) uma busca por melhor enquadramento deste sinal dentro de séries homólogas já definidas.

4.6 Confirmação por padrão isotópico

Francis William Aston foi um físico inglês que, entre outras realizações cientificas, descobriu uma grande quantidade de isótopos. Por esta descoberta ganhou o prêmio Nobel de química do ano de 1922. Desde então varias pesquisas demonstraram que os isótopos ocorriam com abundância relativa específica para cada elemento e consequentemente moléculas são encontradas na natureza com uma quantidade específica dos isótopos dos seus elementos constituintes. A quantidade relativa de moléculas com diferente quantidade de isótopos pode ser calculada utilizando a expansão soma:

$$(\boldsymbol{a}+\boldsymbol{b})^{\boldsymbol{n}} = \mathbf{1} + \frac{\boldsymbol{n}(\boldsymbol{a}^{\boldsymbol{n}-\mathbf{1}})\boldsymbol{b}}{\mathbf{1}!} + \frac{\boldsymbol{n}(\boldsymbol{n}-\mathbf{1})(\boldsymbol{a}^{\boldsymbol{n}-\mathbf{2}})\boldsymbol{b}^{\mathbf{2}}}{\mathbf{2}!} + \frac{\boldsymbol{n}(\boldsymbol{n}-\mathbf{1})(\boldsymbol{n}-\mathbf{2})(\boldsymbol{a}^{\boldsymbol{n}-\mathbf{3}})\boldsymbol{b}^{\mathbf{3}}}{\mathbf{3}!} \ldots \quad \textbf{(Equação 21)}$$

onde ($\boldsymbol{a}$) equivale a intensidade do sinal, sendo que esta pode variar entre 0-100% (como artifício de cálculo utiliza-se o valor de 1). ($\boldsymbol{b}$) equivale à porcentagem relativa da abundância natural do isótopo em questão e ($\boldsymbol{n}$) a quantidade de átomos. Resolvendo o primeiro termo temos a intensidade relativa à (a) do isotopólogo M +1, o segundo termo temos a intensidade relativa do isotopólogo M + 2 e assim por diante.

Esta correlação vem sendo utilizada desde então na espectrometria de massas como uma ferramenta útil para atribuição da composição elementar dos sinais dos espectros adquiridos pelos analisadores de massas.

Nos espectros de petróleo as abundâncias dos isotopólogos são influenciadas principalmente pelos números de ^{13}C e também pelo ^{34}S, quando

presente, os quais ocorrem com abundâncias naturais de 1,112% e 4,21%, respectivamente.

As definições de composições elementares realizadas apenas com base no defeito e exatidão de massas podem não ser inequívocas. Logo, realiza-se o cálculo de intensidade relativa e de massa para os isotopólogos M + 1, M + 2 (utilizando-se as equações **22 e 23**) e procura-se por esses valores nos espectros de massas. Então, as fórmulas moleculares em que não forem encontrados esses valores correspondentes ao padrão isotópico são excluídas.

$$\text{Intensidade (M + 1)} = n\ b \qquad \textbf{(Equação 22)}$$

$$\text{Intensidade (M + 2)} = \frac{\left(n(n-1)b^2\right)}{2} \qquad \textbf{(Equação 23)}$$

onde, (n) equivale ao número de elementos e (b) à abundância relativa natural do elemento.

Com estes resultados pode-se eliminar uma atribuição equivocada da formula molecular de um valor de *m/z.* Por exemplo:

Um sinal adquirido pelo analisador ICR (primeiro sinal no espectro mostrado na **Figura 26**) de *m/z* 334,252600 pode ser proveniente de duas espécies de fórmula moleculares distintas (apresentadas nas formas protonadas- A e B) considerando uma margem de erro de -1,0 até 1,0 *ppm*:

A) $C_{24}H_{32}N$ ➔ *m/z* = 334,252926 - 1,0 ppm DBE = 10,0

B) $C_{16}H_{36}N_3O_2S$ ➔ *m/z* = 334,252275 +1,0 ppm DBE = 1,0

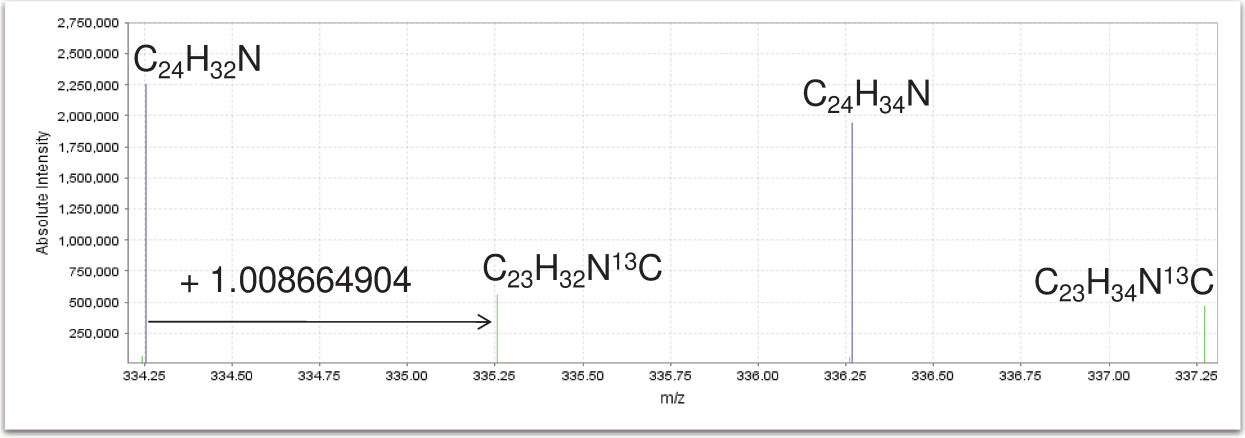


Figura 26. Expansão de um espectro de ESI(+)-FT ICR-MS mostrando as atribuições das fórmulas moleculares para as *m/z* 334,25293 e *m/z* 336,26858 como $C_{24}H_{31}N$ e $C_{24}H_{33}N$ respectivamente. Os isotopólogos de ^{13}C [M + 1] ($C_{23}H_{31}N^{13}C$ e $C_{23}H_{33}N^{13}C$) correspondentes são então caracterizados com as intensidades relativas 24% para ambos mostrando os sinais com as respectivas fórmulas moleculares mono isotópicas.

Calculando as intensidades relativas esperadas para o primeiro isotopólogo (^{13}C) em relação aos átomos de carbonos destas espécies [M + 1] através da **Equação 22**, obtemos:

Intensidade M + 1 teórico (A) = (24) (0,011) = 0,264

Intensidade M + 1 teórico (B) = (16) (0,011) = 0,176

Com as intensidades experimentais dos dois primeiros sinais mostrados na **Figura 26** calculamos razão entre eles:

Intensidade M + 1 experimental= 540/2.250 = 0,24

sendo o limite aceitável para o desvio deste valor de 3 pontos percentuais para mais ou para menos, temos os dois intervalos aceitáveis para os valores de intensidade relativas como:

A) 0,256 >= Intensidade experimental <= 0,272

B) 0,171 >= Intensidade experimental <= 0,181

Assim conclui-se que o sinal de *m/z* 334,252600 é proveniente da espécie com fórmula molecular $C_{24}H_{32}N$.

Se a intensidade calculada dos isotopólogos estiver abaixo do limiar de intensidade estabelecido para o espectro, a duplicidade é eliminada procurando-se por séries homólogas destas espécies.

Além de eliminar a duplicidade, uma vez que o banco de dados teórico foi construído utilizando apenas as espécies mono-isotópicas (espécies com os isótopos com maior abundancia natural), é realizada uma busca por todos os isotopólogos de ^{13}C e ^{34}S em todos os sinais caracterizados no espectro. Todos os termos da **Equação 21** são calculados até que próximo termo apresente um valor menor que o valor estabelecido para o limiar de intensidade. A **Figura 27** mostra o espectro de massas para uma amostra de petróleo leve obtida em 6 minutos de aquisição, depois do processamento de atribuições. Os sinais em azul correspondem aos sinais com as fórmulas moleculares definidas (mono isotópicas) e os sinais verdes correspondem aos isotopólogos de ^{13}C, totalizando 3 mil espécies caracterizadas. Todas as atribuições neste espectro correspondem às moléculas pertencentes à classe N.

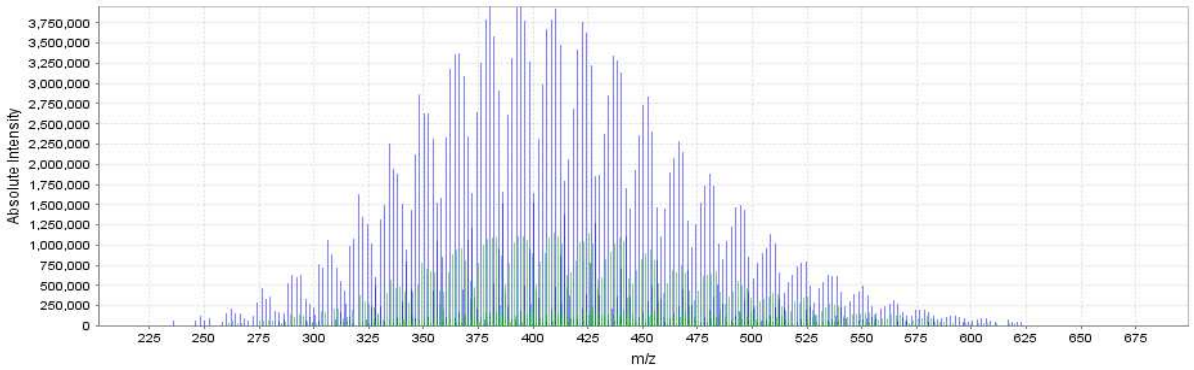


Figura 27. Espectro de ESI(+)-FT ICR-MS mostrando os sinais caracterizados (linhas azuis) e o sinais identificados para os respectivos isotopólogos (linhas verdes).

4.7 Confirmação por serialização

Como mostrado anteriormente é observada uma serialização ao longo de todo espectro de massas relativas às repetições de unidades alquílicas e os diferentes níveis de insaturação/ciclização. Detalhes de como estes compostos estão distribuídos podem ser observados em intervalos de *m/z* escolhidos convenientemente (**Figura 28**). Por exemplo, os sinais identificados na expansão no intervalo de *m/z* 580−640 (**Figura 28**, superior) representam membros de uma mesma série homóloga onde diferença entre eles é de *m/z* 14,0156, correspondente a repetições de unidades metilênicas (**Figura 28,** superior). Os sinais marcados na expansão para o intervalo de *m/z* 580-591 têm uma diferença de *m/z* 2,0157 (**Figura 28**, inferior) representando a diferença na quantidade de insaturações ou de anéis para uma mesma classe, sendo esta diferença correspondente a massa de dois átomos de hidrogênio.

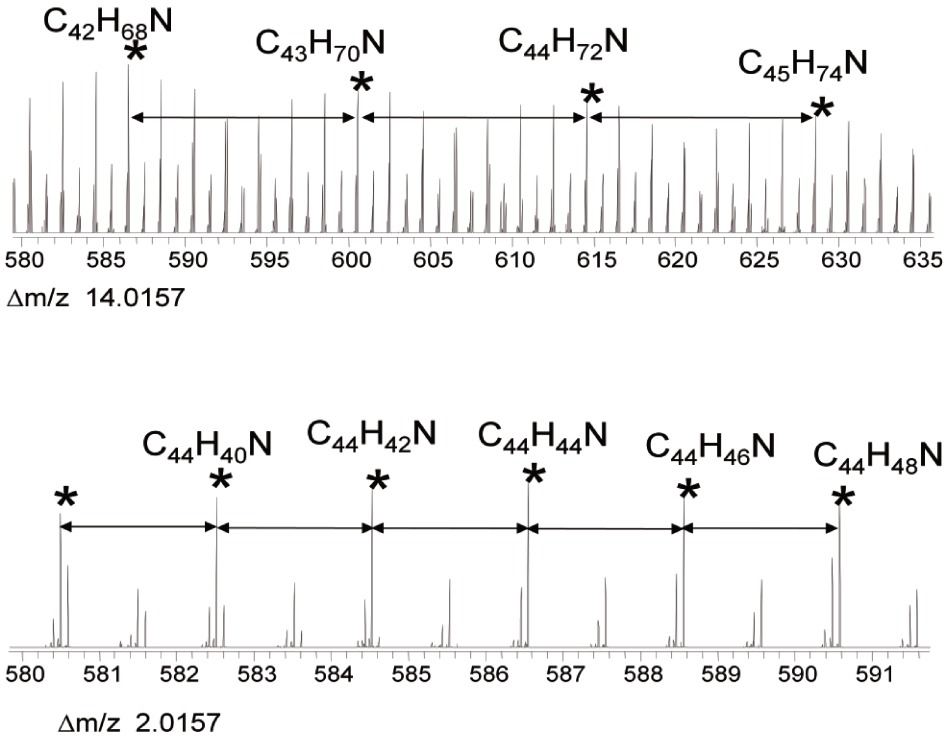


Figura 28. Expansão para os espectro de ESI(+)-FT ICR-MS mostrando unidades de repetição -CH_2 (acima) e unidades de repetição de 2 H (abaixo).

A **Figura 29** exemplifica a serialização discutida acima, mostrando algumas fórmulas estruturais sugeridas para a primeira e ultima *m/z* da parte inferior e superior mostradas na **Figura 28**.

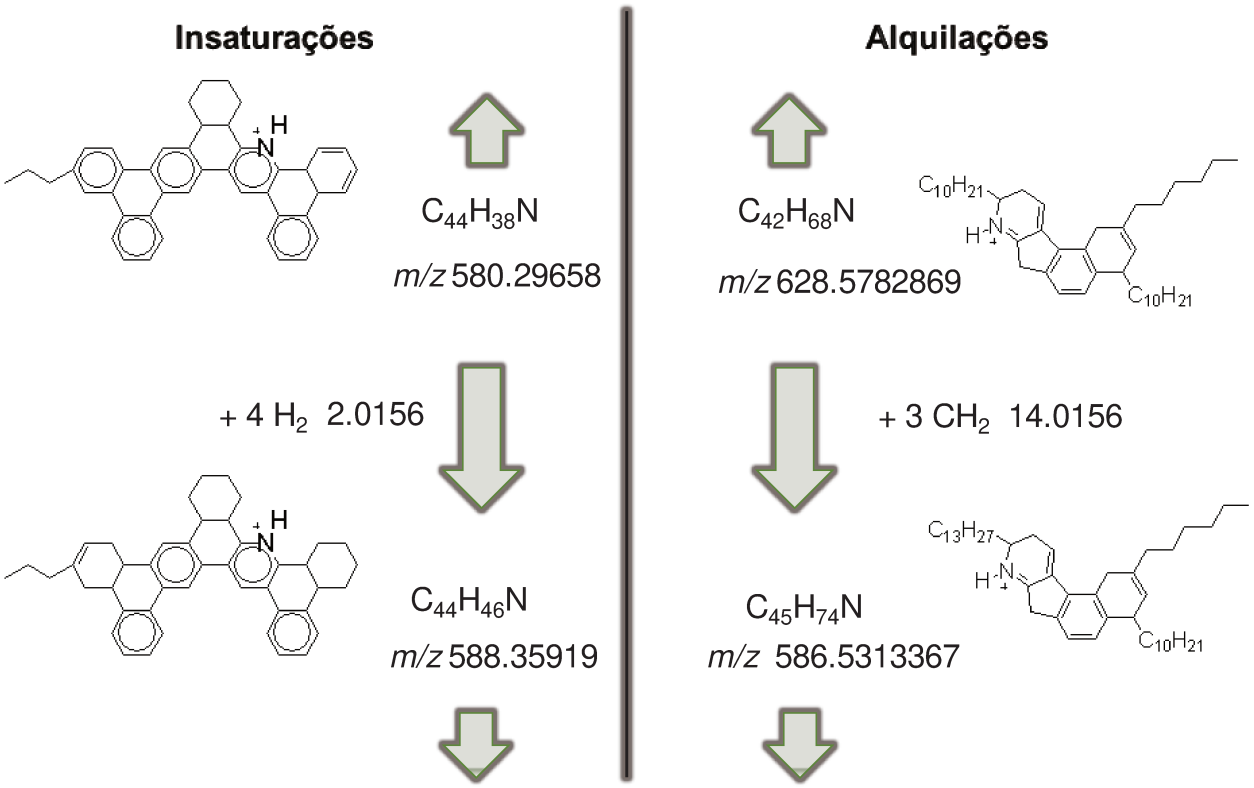


Figura 29. Fórmulas estruturais sugeridas para íons mostrados na Figura 28, exemplificando a serialização por insaturações e por alquilações.

Esta propriedade é utilizada para eliminação de possíveis duplicidades de atribuição em uma *m/z* específica. Para isso as atribuições duplicadas são comparadas com as outras fórmulas já identificadas através da procura no banco de dados teóricos e com a confirmação por padrão isotópico. Confirma-se que a fórmula molecular correta é aquela que obtiver uma maior quantidade de compostos homólogos (repetições de unidade metilênicas) e correlatos (compostos com mesma quantidade de carbono, porém com quantidades de insaturações distintas) no espectro.

4.8 Organização e visualização dos dados

A **Figura 30** mostra a aba principal do programa desenvolvido mostrando a representação gráfica e também os valores de *m/z* e intensidade absoluta em uma tabela (lado esquerdo) de um espectro de massas adquirido no modo positivo de íons utilizando a técnica ESI(+)-FT-ICR. Observa-se um total de 17.000 sinais resolvidos provenientes de íons detectados em sua forma protonada $[M + H]^+$ entre uma faixa de *m/z* 200-950 com um poder de resolução de 400000 na *m/z* 400.

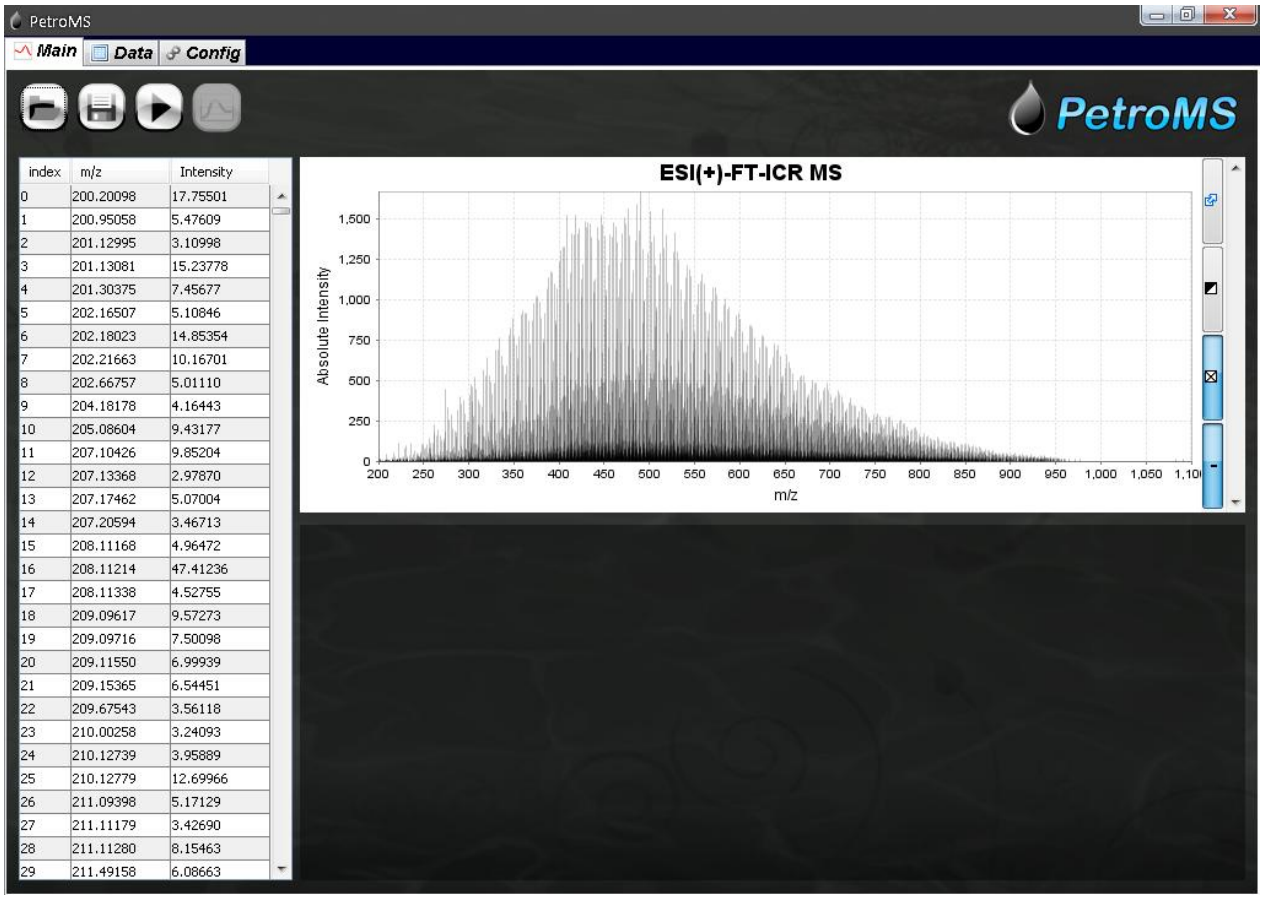


Figura 30. Espectro de ESI(+)-FT-ICR-MS obtido para a amostra de petróleo 1 mostrando mais de 17000 sinais resolvidos.

Após os procedimentos realizados em 3 segundos de atribuição de faixa de erro experimental, busca por composição elementar para todos os sinais detectados e eliminação de duplicidade de atribuições, os dados são organizados em tabelas representativas, separadas por classes e mostrando todas as correlações de *m/z* experimental, *m/z* teórica, erro de massas, DBE, intensidades e sua fórmula molecular possibilitando a ordenação crescente e decrescente por cada uma destas variáveis, além da identificação visual pontual no espectro de massas (**Figura 31**).

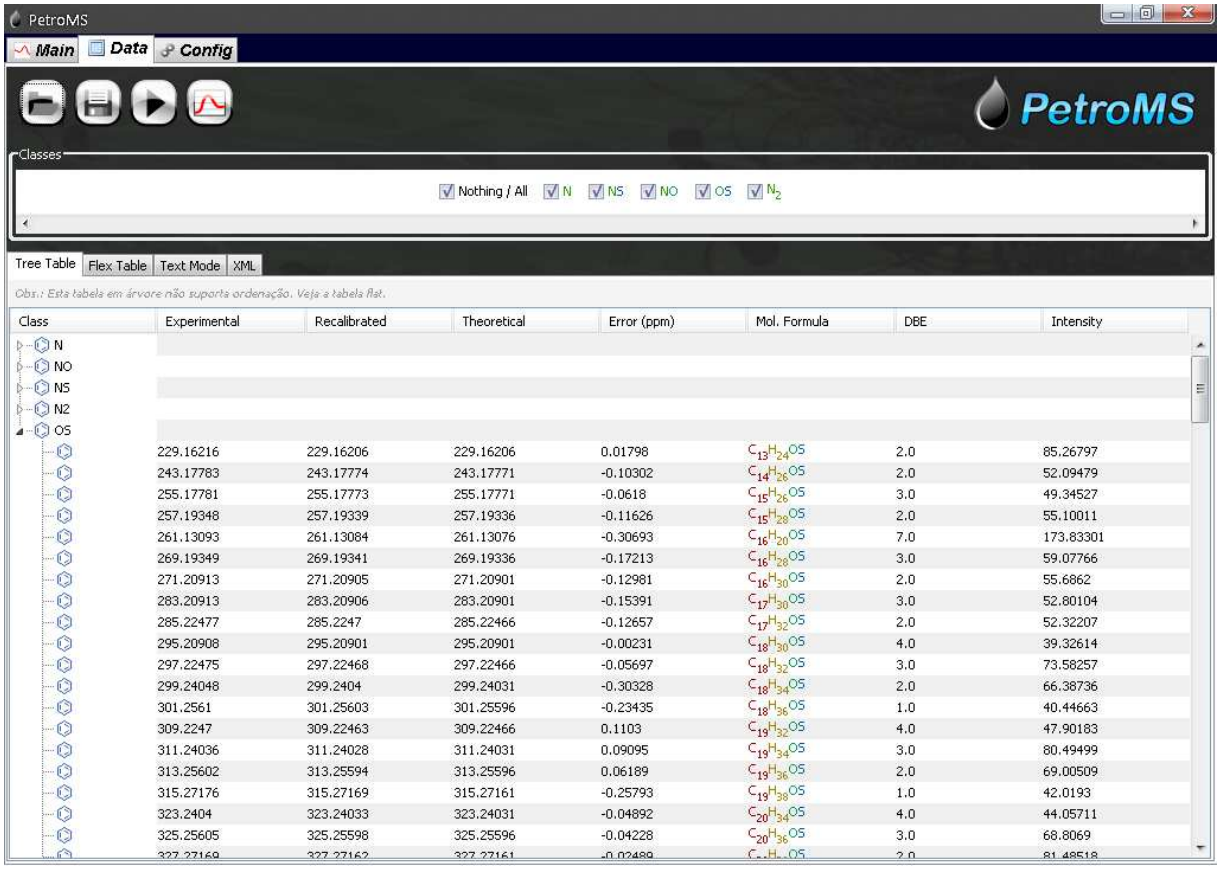


Figura 31. Interface gráfica do *software* desenvolvido mostrando a aba de visualização das atribuições realizadas separadas por classe de heteroátomos.

Na aba principal do *software* é possível visualizar, de forma dinâmica através de botões localizados à direita, os espectros correspondentes dos sinais caracterizados (sinais em azul), não caracterizados (sinais em vermelho), isotopólogos (sinais em verde) e os sinais eliminados por estarem abaixo do limiar de intensidade calculado (**Figura 32 – parte inferior)**

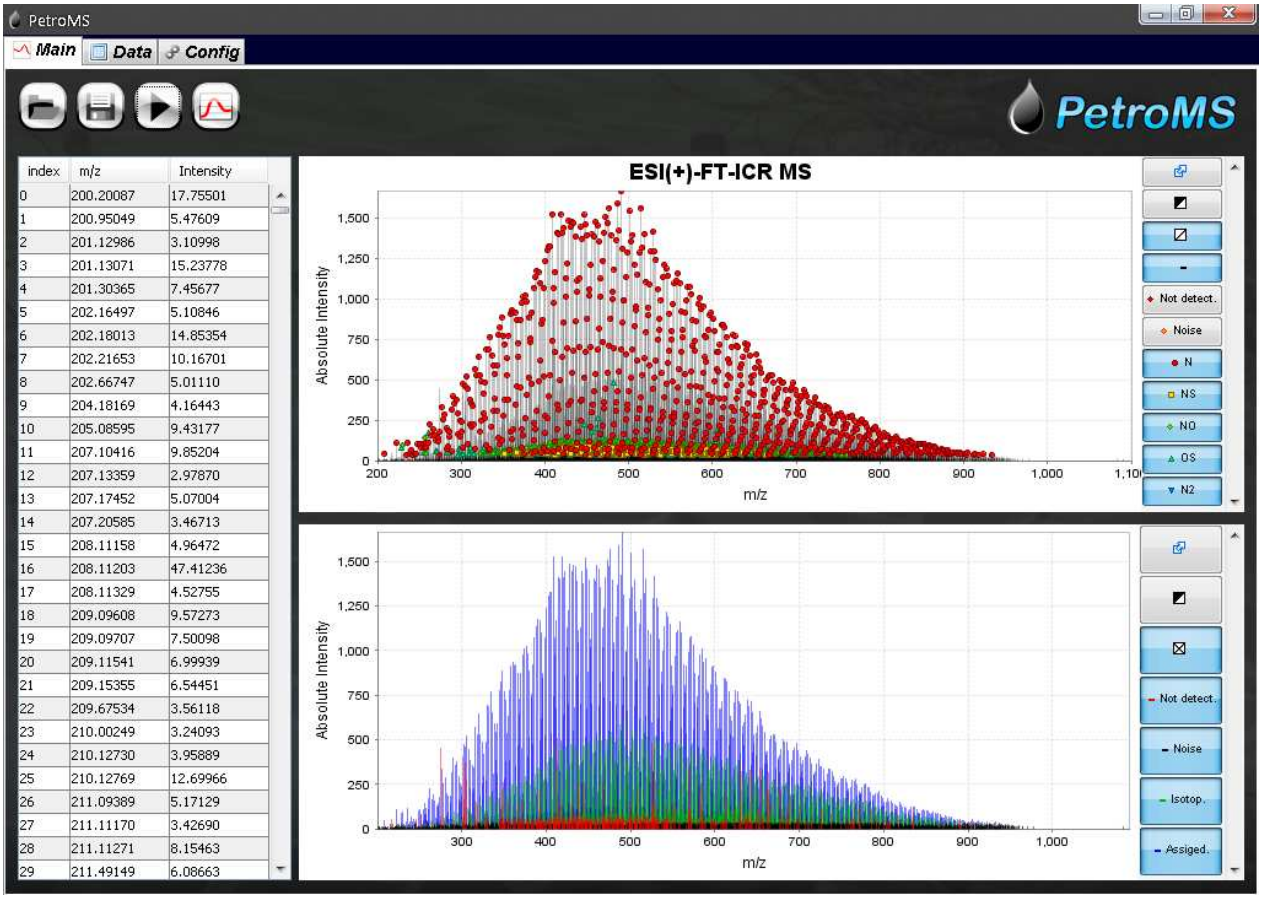


Figura 32. Sinais demarcados no espectro mostrando a distribuição de cada classe de heteroátomos (N, NS, NO, OS, N2) identificadas para o modo negativo de íons e os sinais caracterizados, não caracterizados, os isotopólogos e os sinais atribuídos como parte do ruído

Além desta visualização gráfica geral, todos os sinais do espectro original (**Figura 32** acima) são marcados com suas respectivas classes de heteroátomos demarcados por pontos coloridos fornecendo uma visualização da massa média amostral, e sua distribuição total para as espécies caracterizadas e também para cada classe individualmente sendo este também um painel dinâmico onde o usuário pode escolher qual classe de heteroátomos pretende visualizar (botões à direita do espectro).

Para uma melhor correlação entre as classes detectadas, é conveniente a análise da proporção relativa entre elas. Esta visualização pode ser realizada de forma rápida utilizando a ferramenta de criação de gráficos que, entre outros, provê um gráfico de barras para tal comparação com a possibilidade de escolha das classes a serem utilizadas. As classes encontradas para o espectro mostrado na **Figura 32** no modo positivo de íons são N (85%), NO (7,5%), N2 (2,7% íons), NS(1,6%) e OS (3,1%) (**Figura 33**).

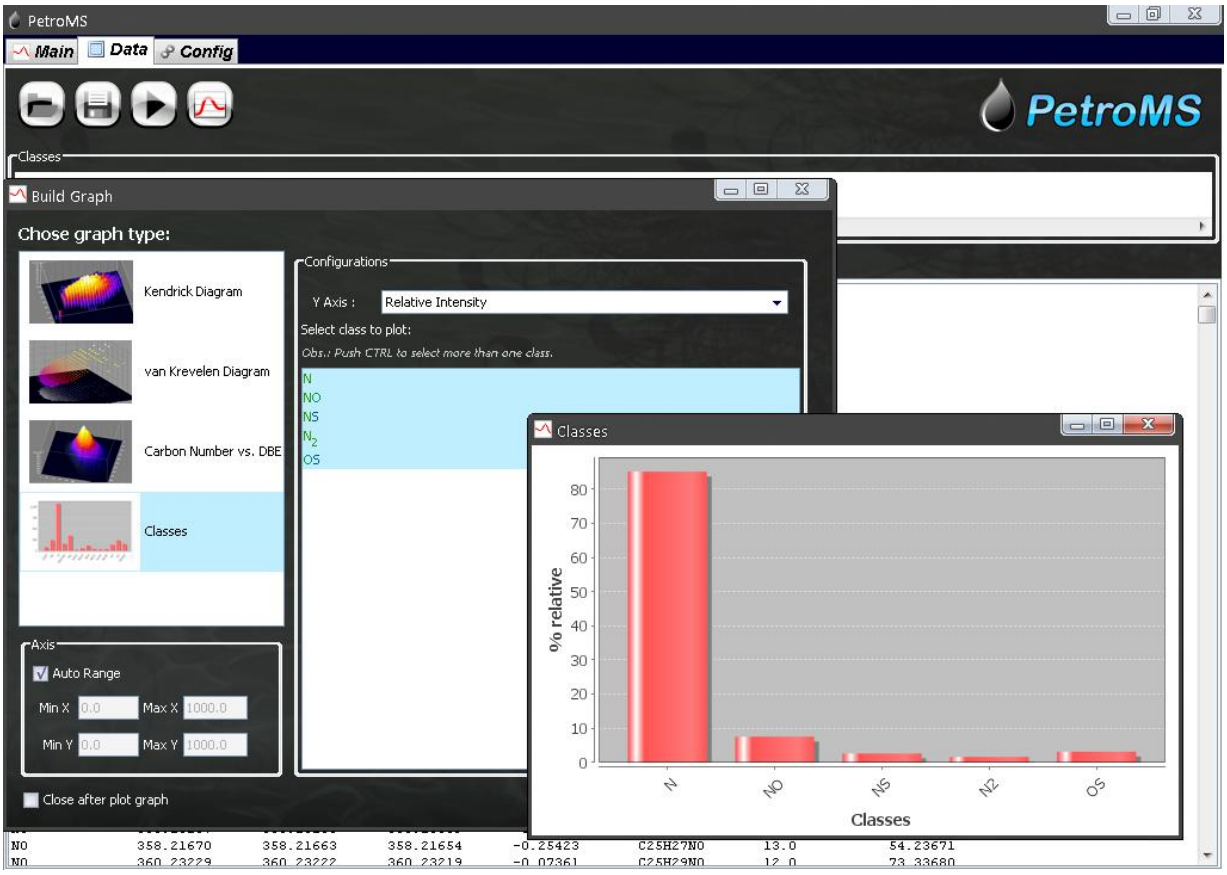


Figura 33. Gráfico de barras mostrando a proporção relativa para cada classe de heteroátomo caracterizada para o espectro de ESI(+)-FT-ICR MS mostrado na **Figura 32**.

A amostra também foi adquirida no modo negativo de íons e submetida ao processamento. A **Figura 34** mostra o espectro ESI(-)-FT-ICR MS obtido após a importação pelo *software*. Observa-se um total de mais de 20000 sinais resolvidos entre uma faixa de 200-1000 *m/z* e um poder de resolução > 400.000 na *m/z* 400. Os íons são detectados na forma desprotonada $[M - H]^-$.

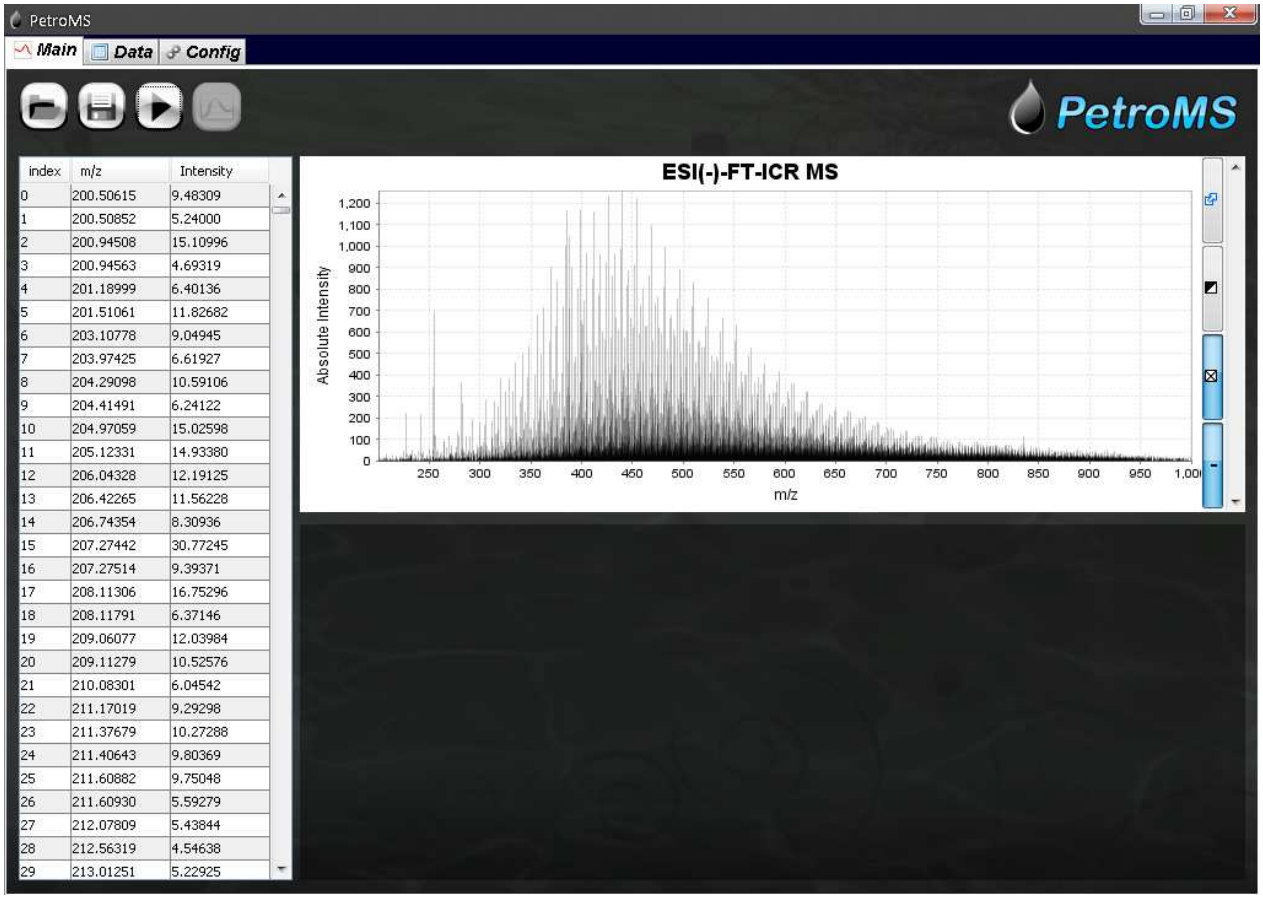


Figura 34. Espectro de ESI(-)-FT-ICR-MS obtido para a amostra de petróleo 1 mostrando mais de 27000 sinais resolvidos.

A **Figura 35** mostra os sinais caracterizados, não caracterizados, isotopólogos e atribuídos como ruído (espectro parte inferior) e as classes atribuídas ao longo do espectro (espectro parte superior) após o processamento dos dados e a **Figura 36** assim como o gráfico de barras mostrando a proporção relativa das classes.

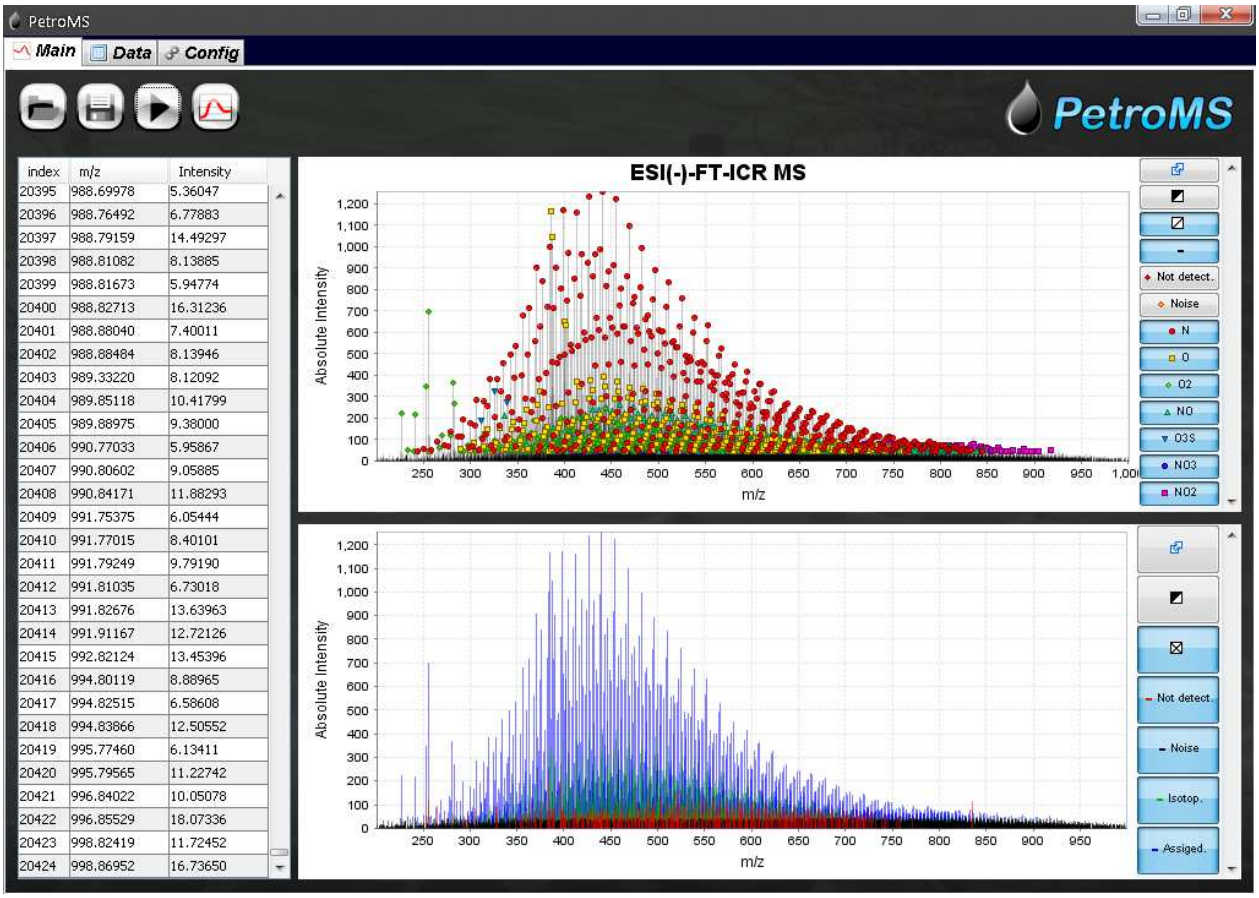


Figura 35. Sinais demarcados no espectro mostrando a distribuição de cada classe de heteroátomos (N, NO, NO2, NO3, O, O2, O3S) identificada para o modo negativo de íons e os sinais caracterizados, não caracterizados, os isotopólogos e os sinais atribuídos como parte do ruído.

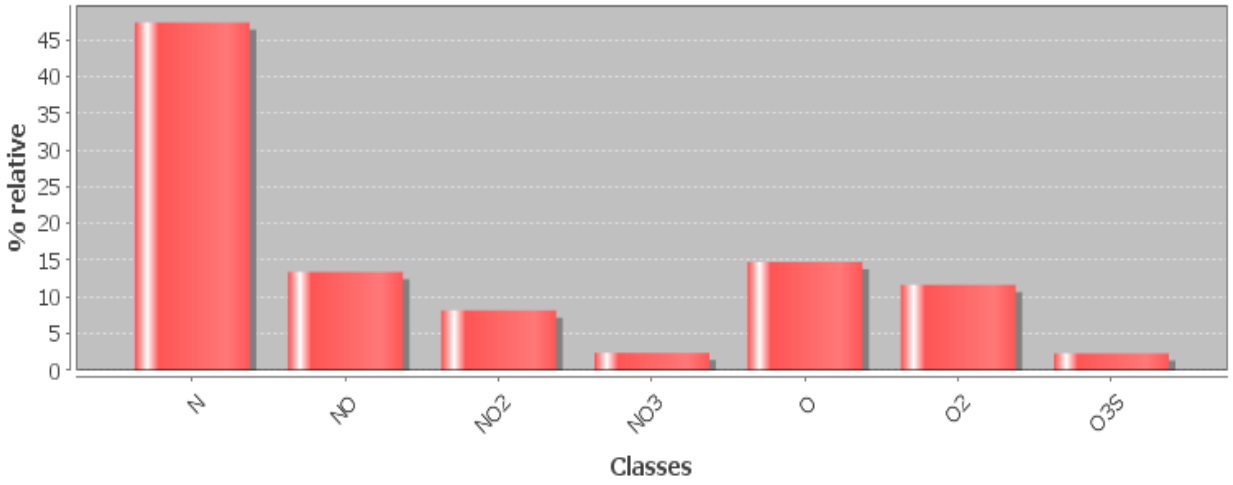


Figura 36. Gráfico de barras mostrando a proporção relativa para cada classe de heteroátomo caracterizada para o espectro de ESI(-)-FT-ICR MS mostrado na Figura 35.

4.8.1 Visualização através de diagramas clássicos.

Para auxiliar a visualização e interpretação dos dados adquiridos, devido a sua grande complexidade, alguns diagramas clássicos podem ser construídos, tais como os diagramas de Kendrick [22] e de van Krevelen[30], número de carbonos vs. DBE[8] e gráficos de barras para comparação de intensidade relativa. Assim, elaboramos uma plataforma de criação destes gráficos de forma automática, sendo possível selecionar o tipo de gráfico, as dimensões, os eixos e as classes de heteroátomos que se pretende visualizar e analisar, como mostra a **Figura 37**.

[30] Krevelen, D. W. van. **Hydrogen distribution in coal**. Fuel (1963), 42(5), 427-30.

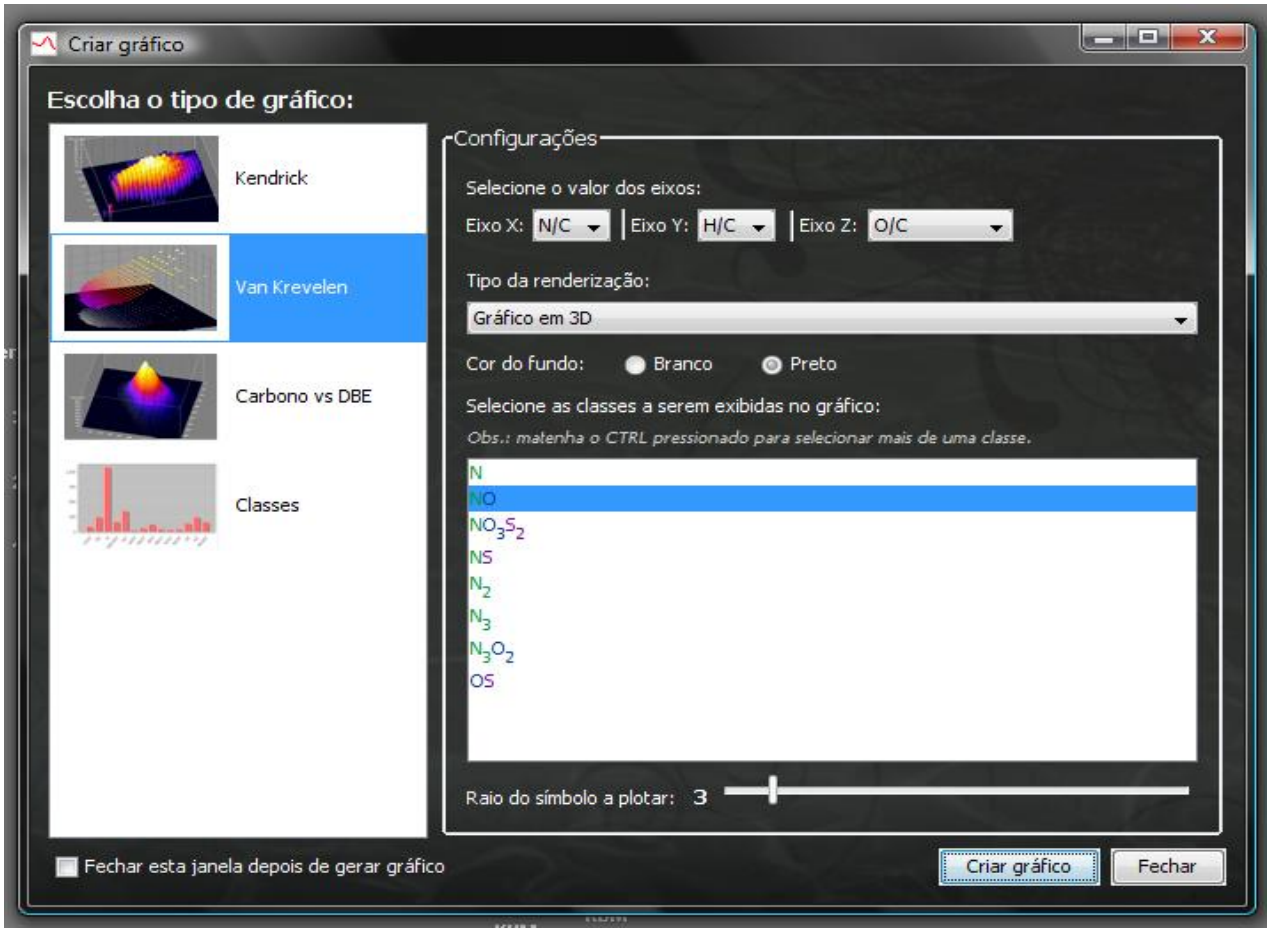


Figura 37. Interface para seleção de diagramas (Kendrick, van Krevelen, nº carbonos vs. DBE e gráfico de barras) e criação dos gráficos para classes selecionadas pelo usuário.

4.8.1.1 Diagrama de Kendrick

Para os diagramas de Kendrick, uma massa nominal de Kendrick (KNM, do inglês *Kendrick nominal mass*) é calculada convertendo as massas de todas as unidades de CH_2 de 14,015 65(IUPAC) para 14,0000. Dessa maneira, o defeito de massas das unidades de repetição CH_2 é cancelado, ou seja, séries homólogas de compostos com a mesma contribuição de heteroátomos e mesmo número de ligações insaturações/anéis (DBE), mas cadeias alquílicas de tamanhos distintos terão o mesmo defeito de Kendrick (KMD do inglês *Kendrick mass defect*) e estarão dispostos ao longo de uma mesma linha. Uma insaturação (duplas ligações ou anéis, valor de DBE) aumentará o KMD em 0,01340 Da (ou seja, o valor de KMD para dois átomos de H). Assim, compostos de uma mesma classe que diferem pelo tipo (valor de DBE) estarão dispostos em linhas paralelas separadas por 0,01340 KMD e quanto maior o valor de DBE mais alto o valor de KMD (**Figura 38**).

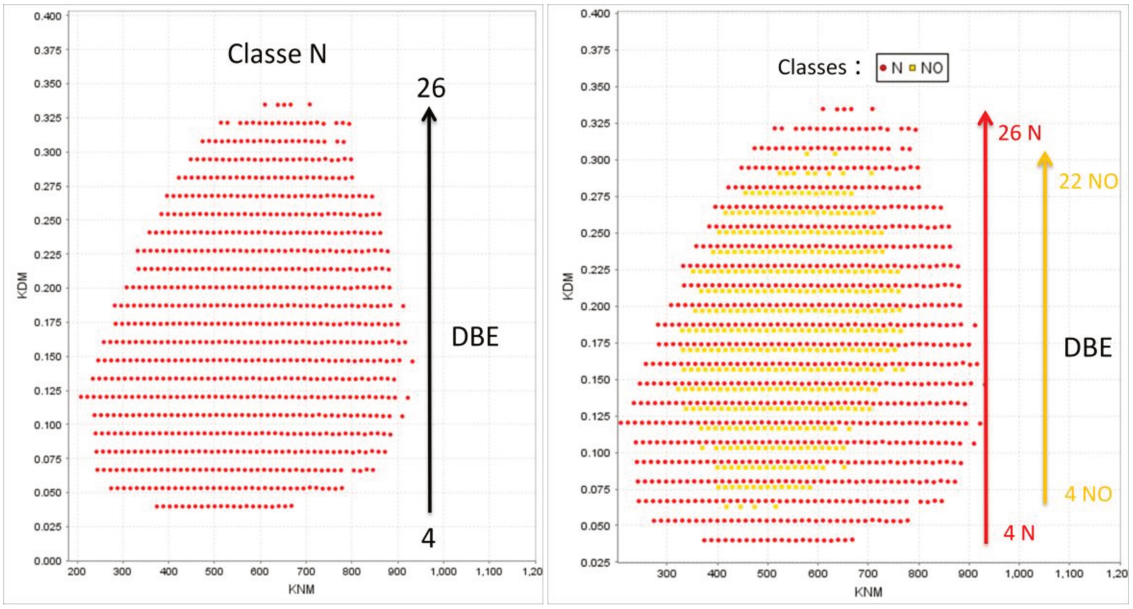


Figura 38. Diagramas de Kendrick do espectro mostrado na **Figura 32** para a classe N (homólogos de piridinas – esquerdos) e para as classes N e NO (homólogos de piridinas(N) e amidas/quinolinas (NO) - direita).

Quando um terceiro eixo (z intensidade) é adicionado ao diagrama de Kendrick, o gráfico provê a distribuição de massas de uma classe específica em função do grau de insaturação. O ponto mais alto da superfície (área vermelha da **Figura 39**) mostra extensão da insaturação e em qual região de massas está concentrada. Essa informação será útil para determinação de grau de maturação e de biodegradabilidade [8].

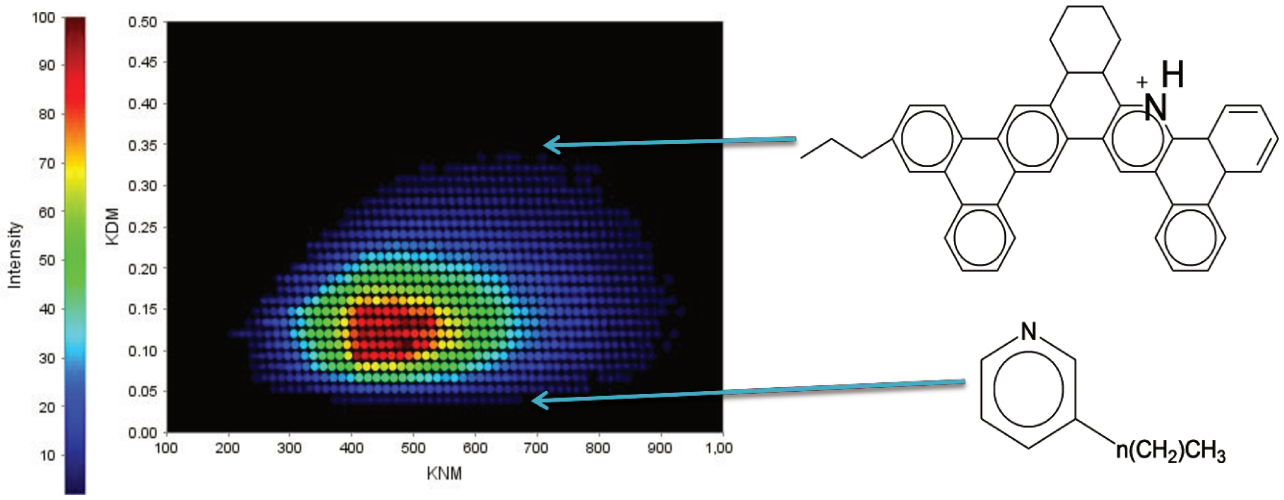


Figura 39. Diagrama de Kendrick em três dimensões do espectro mostrado na **Figura 38** para a classe N (homólogos de piridínicos) e suas respectivas fórmulas estruturais propostas para a série homóloga de maior e menor KMD.

4.8.1.2 Diagrama de número de carbonos vs. DBE

Quando não existe a necessidade da visualização de duas ou mais classes de heteroátomos em um mesmo gráfico a elaboração de diagramas de carbono por DBE provê uma relação mais direta para correlacionar número de carbonos (massa) e grau de insaturação/ciclização **(Figura 40)**.

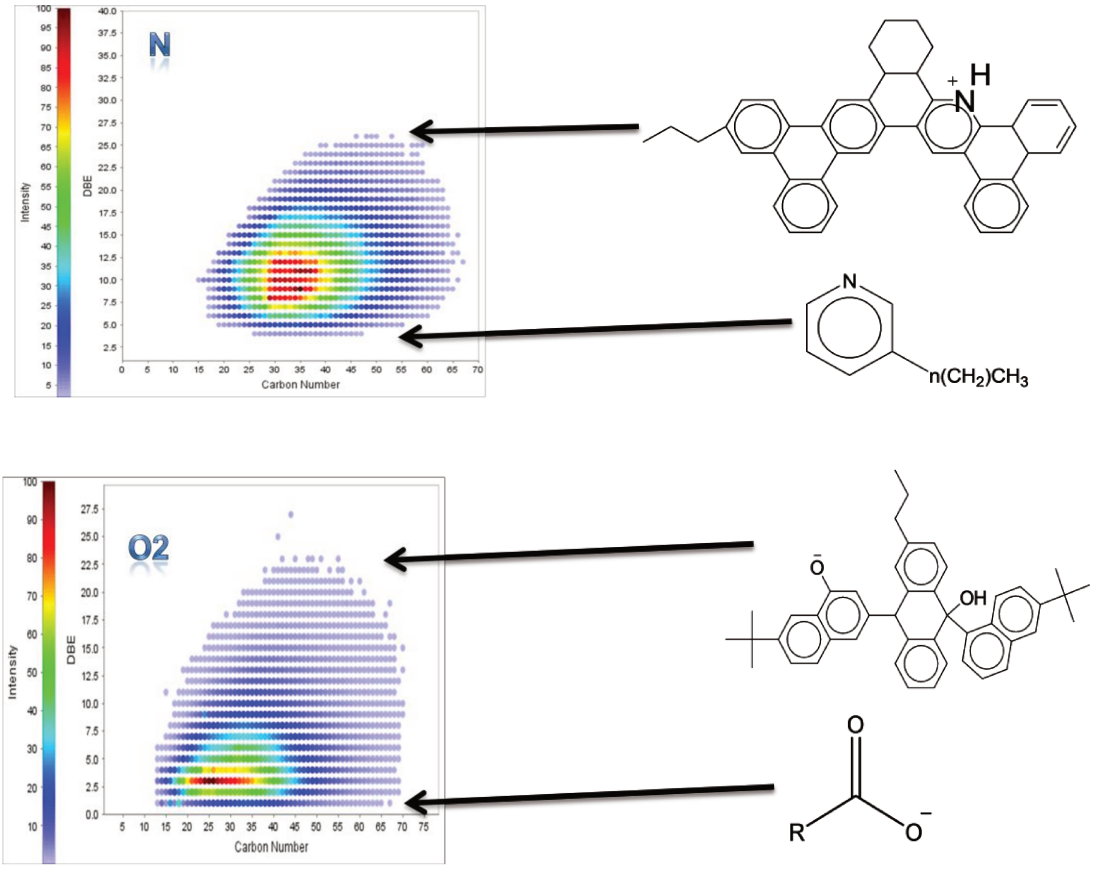


Figura 40. Diagrama de número de carbonos versus DBE para a classe N detectada no modo positivo de íons (acima), e para a classe O2 no modo negativo de íons e suas respectivas fórmulas estruturais sugeridas para a série homóloga de maior e menor valor de DBE.

Esses diagramas se mostram bastante úteis e serão utilizados para análise detalhada (nível molecular) de ácidos naftênicos (DBE > 2) e também para o acompanhamento de processos de remoção dos mesmos mostrando possíveis seletividades em relação à quantidade de carbonos e numero de anéis, auxiliando em pesquisas sobre corrosão, uma vez que estas espécies estão diretamente correlacionadas aos índices de corrosão do petróleo. Pretende-se também utilizá-lo para o acompanhamento da remoção das distintas classes de heteroátomos e a possível seletividade a uma ou algumas delas nos processos de hidro e oxitratamento. Assim, essa ferramenta desenvolvida é capaz de detalhar quimicamente as diferentes fases e condições destes processos em apenas alguns minutos de aquisição e segundos de processamento, permitindo a visualização dos dados de forma fácil e clara, tornando o estudo e desenvolvimento destes processos mais simples e ágeis.

4.8.1.3 Diagrama de van Krevelen

Nos diagramas de van Krevelen[30] a composição elementar dos componentes do petróleo é visualizada projetando as razões atômicas, como H/C, O/C, N/C, S/C em dois ou três eixos. A razão H/C separa os compostos de acordo com seu grau de insaturação, enquanto que as razões O/C, N/C e S/C os separam de acordo com a quantidade de O N ou S. Isso permite uma comparação visual de classes como função do teor de heteroátomos e do teor de insaturação. Esses gráficos são úteis para acompanhamento de processos de tratamento de petróleo e derivados. Por exemplo, o processo de hidrotratamento de diesel foi acompanhado utilizando o diagrama de van Krevelen para as classes NO, NO2, e NO3 utilizando as razões O/C e H/C no modo negativo de íons (**Figura 41**).

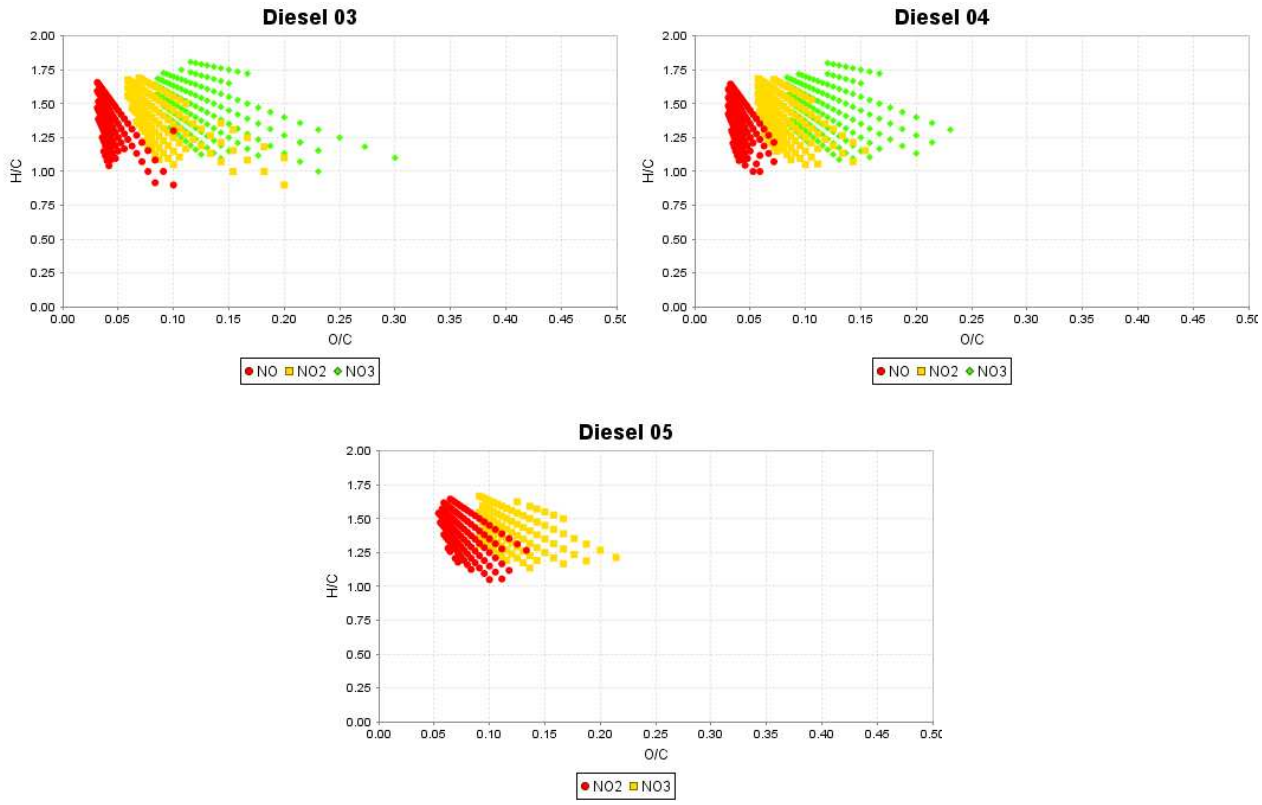


Figura 41. Diagrama de van Krevelen para o acompanhamento de hidrotratamento em diesel para as classes NO, NO2 e NO3 no modo negativo de íons.

Observa-se a eliminação gradual das três classes em razões distintas, sendo eliminados seletivamente, como esperado, os compostos mais insaturados (menor H/C) e pertencentes à classe NO (menor quantidade de heteroátomos). Com esta metodologia pretende-se estender esta ferramenta para análise de diferentes cortes, sendo possível a comparação direta entre as diferentes classes de heteroátomos para diferentes pontos de ebulição.

4.8.1.4 Diagrama de van Krevelen modificado

Além dos gráficos já discutidos, diagramas de van Krevelen modificados, que utilizam o número de carbono e a razão H/C, tem sido demonstrado de grande utilidade na análise de asfaltenos[31] e maltenos. A razão H/C para os asfaltenos está abaixo de 0,8 (apesar de haver a discussão sobre a existência de uma fase de transição entre 0,8 e 1,2) enquanto que a razão H/C para os maltenos está acima de 0,8 até 2.0 (totalmente saturado). Assim o gráfico de C vs H/C apresenta-se como ferramenta ideal para diferenciar a fração de maltenos das frações de asfaltenos, podendo ser utilizado para o acompanhamento de otimizações e na remoção mais efetivas em processos de separação (**Figura 42**).

[31] J. M.; Merdrignac, I.; Rodgers, R. P.; Marshall, A. G.; Gauthier T.; Guibard, I. **Stepwise Structural Characterization of Asphaltenes during Deep Hydroconversion Processes Determined by Atmospheric Pressure Photoionization (APPI) Fourier Transform Ion Cyclotron Resonance (FT-ICR) Mass Spectrometry.** Energy & Fuels (2010), 24(4), 2257-2265.

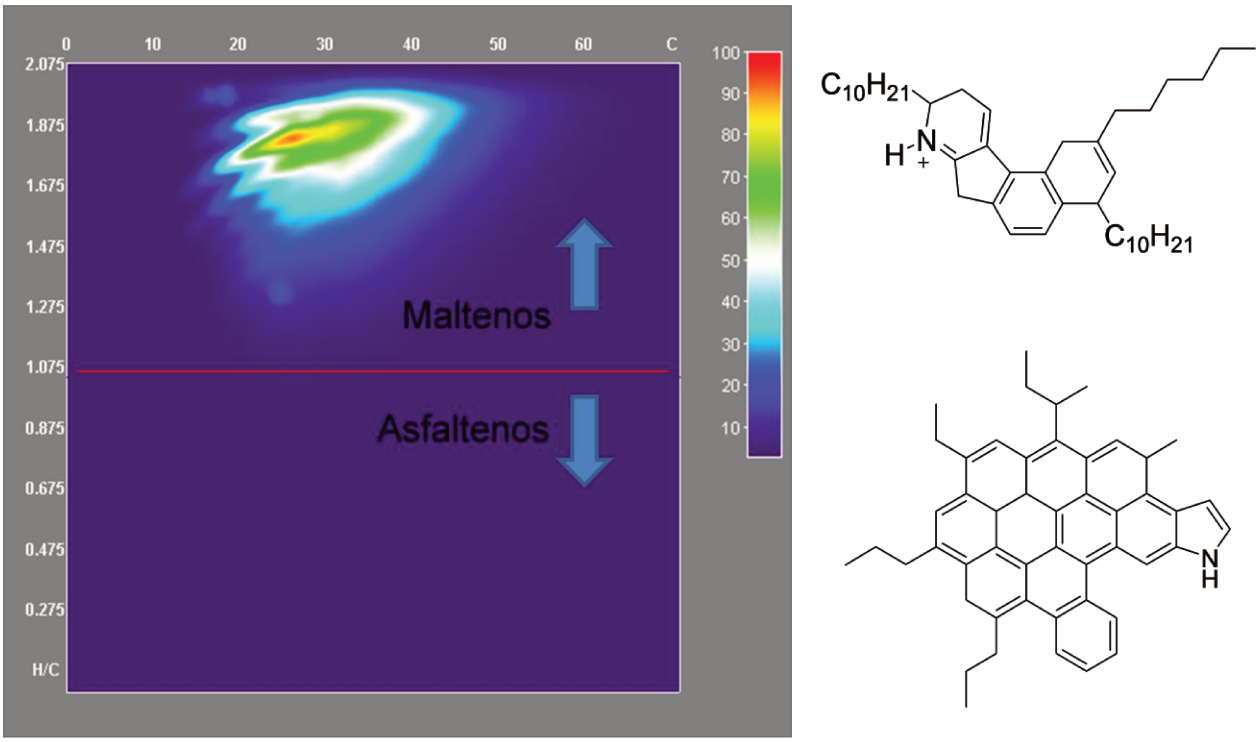


Figura 42. Diagrama de van Krevelen 3D modificado para o eixo da abscissa (número de carbono) utilizando a razão H/C para a ordenada, e intensidade como mapa de cor, realizado no modo de imagem pelo *software* desenvolvido para a fração de maltenos analisada por FT ICR-MS no modo positivo de íons.

5. Conclusões

Amostras de petróleo bruto e seus derivados podem ser caracterizadas com altíssima rapidez e eficiência através da técnica denominada de petroleômica por FT ICR-MS. Milhares de marcadores polares são resolvidos e identificados em questão de ~ 6-7 minutos com muita agilidade e elevada confiabilidade. A técnica produz, porém, uma enormidade de dados que precisam ser tratados com eficiência através de *software* que combine estratégias eficientes de programação e inteligência química no processamento e visualização. Neste trabalho, toda uma estratégia já aprovada em testes preliminares foi desenvolvida com esta finalidade. Esta estratégia levou a atribuição da composição elementar inequívoca para todos os sinais adquiridos. As etapas envolveram: a atribuição da *m/z* para cada sinal do espectro otimizado, a atribuição do limiar ótimo de intensidade para o ruído de cada espectro individualmente, a construção de um banco de dados teórico de correlação massa e fórmula molecular para os possíveis constituintes polares do petróleo, a atribuição da composição elementar para cada sinal através da comparação com este banco teórico, a eliminação de duplicidade de atribuições através da confirmação do padrão isotópico e comparação com séries homólogas, a organização dos dados por classes (heteroátomos) e tipo (DBE) e a construção de bibliotecas gráficas para a visualização e interpretação das atribuições (Diagramas de Kendrick, van Krevelen e numero de carbonos por DBE). O sequenciamento lógico destes procedimentos constituiu a construção de algoritmos baseados nestes procedimentos. A combinação destes algoritmos e a execução automática deles formam um *software* de altíssima eficiência e

alta tecnologia essencial para o tratamento e visualização de dados de petroleômica por FT- ICR MS.

6. Perspectivas futuras

O *software* desenvolvido nesta pesquisa está e continuará sendo utilizado como ferramenta para diferentes linhas de pesquisa em petroleômica por FT-ICR MS aplicada a petróleo e seus derivados, entre elas: a) acompanhamento do processo de eliminação de compostos polares de diesel através de hidrotratamento e oxitratamento, b) atribuição e detalhamento de ácidos naftênicos presentes no petróleo, a sua correlação com índice de corrosão associado e c) o acompanhamento rápido para processos visando sua remoção, e d) análises de frações de asfaltenos e a identificação de contaminantes. Em suma, um *software* que agrupa estratégias de alta eficiência, inteligência química em processamento de dados foi desenvolvido para a análise de dados de petroleômica por FT-ICR MS, uma técnica na fronteira dos conhecimentos em caracterização de petróleo e seus derivados. Este *software* agilizará estudos em petroleômica aplicados a várias etapas da exploração, produção, refino.