UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

TESE DE DOUTORADO

Uma nova abordagem para resolver deficiência de posto, desalinhamento de pico e interferências na análise de aminas biogênicas em peixes usando HPLC-DAD

José Licarion Pinto Segundo Neto

João Pessoa – PB - Brasil
Fevereiro/2018

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

TESE DE DOUTORADO

Uma nova abordagem para resolver deficiência de posto, desalinhamento de pico e interferências na análise de aminas biogênicas em peixes usando HPLC-DAD

José Licarion Pinto Segundo Neto*

Tese apresentada ao Programa de Pós-Graduação em Química da Universidade Federal da Paraíba como parte dos requisitos para obtenção do título de Doutor em Química, área de concentração Química Analítica.

Orientador: **Prof. Dr. Mário César Ugulino de Araújo**

* Bolsista

Conselho Nacional de Desenvolvimento Científico e Tecnológico

João Pessoa – PB – Brasil

Fevereiro/2018

Catalogação na publicação
Seção de Catalogação e Classificação

S456n Segundo Neto, José Licarion Pinto.
Uma nova abordagem para resolver deficiência de posto, desalinhamento de pico e interferências na análise de aminas biogênicas em peixes usando HPLC-DAD / José Licarion Pinto Segundo Neto. - João Pessoa, 2018.
141 f. : il.

Orientação: Mário César Ugulino de Araújo.
Tese (Doutorado) - UFPB/CCEN.

1. Química analítica. 2. Aminas biogênicas. 3. Deficiência de posto. 4. HPLC-DAD. 5. MCR-ALS. 6. Coshift. I. Araújo, Mário César Ugulino de. II. Título.

UFPB/BC

Uma nova abordagem para resolver deficiência de posto, desalinhamento de pico e interferências na análise de aminas biogênicas em peixes usando HPLC-DAD.

Tese de Doutorado apresentada pelo aluno José Licarion Pinto Segundo Neto e aprovada pela banca examinadora em 05 de fevereiro de 2018.

Prof. Dr. Mário César Ugulino de Araújo
Orientador/Presidente

Profa. Dra. Patrícia Kaori Soares
Examinadora

Profa. Dra. Maria Fernanda Pimentel Avelar
Examinadora

Profa. Dra. Kátia Messias Bichinho
Examinadora

Prof. Dr. Márcio José Coelho de Pontes
Examinador

DEDICO A TODOS OS QUE ME AJUDARAM NO DECORRER DESTE TRABALHO.

"*Estou convencido das minhas próprias limitações - e esta convicção é minha força*"

Mahatma Gandhi

AGRADECIMENTOS

A Deus pela saúde, disposição, sabedoria concedida e pelo seu infinito amor.

A minha família, em especial meus pais José Licarion Pinto Neto e Maria Betânia Lima Diniz Pinto por todo amor e carinho dedicado, por ter acreditado e me incentivado a sempre prosseguir e nunca desistir. Aos meus irmãos Ricardo e Juliana, pelo apoio. A minha tia Roseane pelo apoio, amizade, carinho, incentivo e ajuda, em especial durante a graduação. Aos meus tios (a), primos (a) e avô (a) pelo carinho e pelos ótimos momentos juntos.

Ao meu orientador de mestrado e doutorado, Prof. Dr. Mário Cesar Ugulino de Araújo, por todo ensinamento, dedicação, ajuda, pelas oportunidades de crescimento profissional e pelo trabalho árduo dedicado a melhorar e facilitar o aprendizado no LAQA ao longo desses anos.

Ao meu orientador de iniciação científica, Prof. Dr. Sherlan Guimarães Lemos, por todos os ensinamentos e por proporcionar um crescimento pessoal e profissional durante os meus primeiros passos na vida acadêmica.

Ao Prof. Dr. Hector Cassimiro Goicoechea e aos amigos argentinos do LADAQ, pelos ensinamentos e acolhimento ao longo de minha estadia na cidade de Santa Fé – Argentina.

A Adriano e a Emanuella, por toda troca de conhecimento, e ajuda no desenvolvimento deste trabalho. A Cesar Diaz Nieto, pela parceria e colaboração neste trabalho.

A Adriano, Emanuella, Welma, Amanda, Anabel, David, Valber, Mayara, Edilene, Paulo, Wellington, Daniel, Mayara, Karla. Enfim, a todos os integrantes do LAQA que tornaram o aprendizado e a convivência no laboratório mais tranquila, divertida e prazerosa.

Aos amigos da graduação Mayara, Helenise, Marcel, Larissa, Kaline, Rafael, Zé Luiz, Raquel Cardoso, Vivian, Luana, Fernaninha, Anallyne, Raque Gomes, Katharinne, Sanierlly, Jean, Lucas, Ciro e Anderson.

Aos amigos feitos ao longo desses anos e que tornam os dias do doutorado ainda melhor, Jussieu Júnior, João Batista Júnior, Alberto, Felipe, Caio, Josemar, Edjan, Tarcizo, Danilo, Marcio, Poliana, Julio, Lucas, Anja, Jéssica Medeiros e Alvaro.

A Profª. Drª. Marcia Cristina Breitkreitz pelas oportunidades e recepção na UNICAMP.

Aos Professores do PPGQ-UFPB pelos ensinamentos nas disciplinas cursadas. Em especial as Profªs. Drª. Elizete, Teresa e Liliana pelo carinho e ensinamentos no âmbito acadêmico, pessoal e da docência.

A banca de qualificação pelas contribuições significativas.

Ao PPGQ-UFPB e ao CNPq pela bolsa concedida.

RESUMO

As aminas biogênicas (ABs) são usadas como indicadores de degradação em alimentos e em elevadas concentrações podem causar danos ao organismo. As ABs são produzidas pela ação de microrganismos através da descarboxilação enzimática de aminoácidos naturais. Existem diversas ABs, no entanto, apenas a histamina é controlada por órgãos regulamentadores como ESFA (*European Food Safety Authority*) e FAO/WHO (*Food and Agriculture Organization of the United Nations/World Health Organization*). Alguns inconvenientes estão presentes nas análises de ABs. A elevada energia de transição eletrônica requer uma etapa de derivatização previamente à análise usando HPLC-DAD (*high-performance liquid chromatography – diode array detector*). A derivatização mais comum é a que utiliza o cloreto de dansil. No entanto, esta é uma reação não seletiva que ocorre com qualquer amina primária ou secundária o que leva a interferentes coeluidos com perfis muito correlacionados, resultando em deficiência de posto. Além da deficiência de posto, os dados apresentam quebra da trilinearidade devido ao deslocamento da retenção frequentemente presente em dados de cromatografia. A forma mais comum de resolver deficiência de posto é o uso de um detector mais seletivo, no entanto, essa abordagem eleva o custo da análise e a complexidade no tratamento de dados, além de nem sempre esse detector mais seletivo existir para a análise em questão. A metodologia proposta neste trabalho faz uso de uma eluição isocrática de apenas quatro minutos, o que eleva a frequência analítica e está de acordo com os princípios da química analítica verde por reduzir o volume de solvente gasto. Com esta abordagem desenvolvida foi possível obter limites de quantificação entre 0,08 e 0,15 mg kg^{-1} em amostras sintéticas e entre 0,04 e 0,17 mg kg^{-1} em amostras de peixe, que estão bem abaixo do limite máximo de 50 mg kg^{-1} estabelecido pela FAO/WHO e pela ESFA para histamina. Foram alcançados resultados exatos na predição das dezesseis amostras sintéticas com REP (*root error of prediction*) variando entre 4,0 e 8,0 %. Foi possível identificar as ABs nas amostras de peixe recém-pescado em baixas concentrações, no entanto o aumento desta concentração pode ser observado ao realizar um estudo cinético de degradação das quatro amostras acondicionadas a 4 ℃ por treze dias. Neste trabalho, mostrou-se a eficiência do alinhamento de pico com o *coshift* em restaurar a trilinearidade dos dados HPLC-DAD. Desse modo, foi possível alcançar a vantagem de segunda ordem usando o MCR-ALS (*Multivariate curve resolution - alternating least squares*) aumentado no modo espectral. Até o momento não havia proposta para alcançar a vantagem de segunda ordem em dados com essas características sem recorrer à separação dos constituintes que causam deficiência de posto ou ao uso de outro detector sincronizado que seja mais seletivo.

Palavras-chaves: aminas biogênicas, deficiência de posto, HPLC-DAD, MCR-ALS, coshift.

ABSTRACT

Biogenic amines (BAs) are used as food spoilage marker and can cause damage to the body at high concentrations levels. BAs are produced by the action of microorganisms through the enzymatic decarboxylation of natural amino acids. There are several BAs, however only histamine is monitored by a monitoring agency as ESFA (*European Food Safety Authority*) and FAO/WHO (*Food and Agriculture Organization of the United Nations/World Health Organization*). Some drawbacks are present in the BAs analyzes. The high electron transition energy requires a derivatization step prior to the HPLC-DAD (*high-performance liquid chromatography – diode array detector*) analysis. The most common derivation is the one using dansyl chloride. However, this is a non-selective reaction that occurs with any primary or secondary amine, leading to interferers with highly correlated profiles, and resulting in rank deficiency. In addition to rank deficiency, the data present trilinearity break due to time misalignment frequently present in chromatographic data. The most common way of solving rank deficiency is by using a more selective detector, however, this approach increases the cost of analysis and the complexity of data processing, besides not always this selective detector exists. The proposed methodology uses of a four minutes' isocratic elution, which raises the analytical frequency and is in agreement with the principles of the green analytical chemistry by reducing the solvent consumption. Quantification limits were between 0.08 and 0.15 mg kg^{-1} for synthetic samples and between 0.04 and 0.17 mg kg^{-1} for fish samples were obtained with the developed approach, which are way below the maximum limit of 50 mg kg^{-1} established by FAO/WHO and by the EFSA for histamine in fish. Accuracy results was obtained in the prediction of the sixteen synthetic samples with REP (*root error of prediction*) varying between 4.0 and 8.0%. BAs were initially identified is fresh fish samples at low concentration levels, however the increase of this concentration can be observed when a kinetic study of degradation of the four samples conditioned at 4 ° C for thirteen days was performed. In this work, the efficiency of the peak alignment with the coshift on restoring the trilinearity of the HPLC-DAD data was shown. So, it was then possible to achieve the second-order advantage using the MCR-ALS (*Multivariate curve resolution - alternating least squares*) augmented in the spectral mode. So far there was no proposal to achieve the second-order advantage in data with these characteristics without the previously separation of the constituents that cause rank deficiency or the use of another more selective detector.

Keywords: biogenic amines, rank deficiency, HPLC-DAD, MCR-ALS, coshift.

LISTA DE FIGURAS

Figura 1 − Estrutura molecular da TRI (**a**), FEN (**b**), PUT (**c**), CAD (**d**), HIS (**e**) e dos reagentes derivatizantes DNS (**f**) e OPA (**g**). Esferas brancas, cinzas, azuis, vermelhas, verde e amarela, indicam hidrogênios, carbonos, nitrogênios, oxigênios, cloro e enxofre, respectivamente. Fonte: autor. ---------- 26

Figura 2 − Reação de derivatização da PUT com o Dansil. Esferas brancas, cinzas, azuis, vermelhas, verde e amarela indicam hidrogênios, carbonos, nitrogênios, oxigênios, cloro e enxofre, respectivamente. Fonte: autor. ---------- 27

Figura 3 − Reação de derivatização da PUT com o OPA e o 2-mercaptoetanol. Esferas brancas, cinzas, azuis, vermelhas, verde e amarela indicam hidrogênios, carbonos, nitrogênios, oxigênios, cloro e enxofre, respectivamente. Fonte: autor. ---------- 27

Figura 4 − Organização dos tipos de dados e exemplos para calibração univariada e multivariada de ordem 1 e 2. Adaptado de (Olivieri, 2014). ---------- 29

Figura 5 − Decomposição trilinear de um tensor de 3 vias *X* em 3 matrizes de pesos A, B e C adicionado de um tensor de resíduos *E*. Adaptado de (Sena, Trevisan, Poppi, 2005). ------- 34

Figura 6 − Esquema para desdobrar uma amostra em um vetor (**a**) e um conjunto de amostra em vetores (b) para construção de um modelo UPLS. Adaptado de (de la Peña, Goicoechea, Escandar, Olivieri, 2015). ---------- 36

Figura 7 − Fluxograma do RBL adaptado de (de la Peña, Goicoechea, Escandar, Olivieri, 2015; Pinto, 2014). ---------- 37

Figura 8 − Ilustração de como funciona o RBL para a amostra de teste contendo o analito e um interferente. Fonte: autor. ---------- 38

Figura 9 − Decomposição MCR-ALS de dados LC-DAD aumentado no modo de eluição para 1 amostra contendo 1 analito (**a**) e para um conjunto de amostras contendo 4 analitos (**b**). Adaptado de (de la Peña, Goicoechea, Escandar, Olivieri, 2015; Pinto, 2014). ---------- 40

Figura 10 − Decomposição MCR-ALS de dados LC-DAD aumentado no modo espectral para 1 amostra contendo 1 analito e para um conjunto de amostras contendo 4 analitos. Adaptado de (Pinto, 2014). ---------- 44

Figura 11 − Procedimento para derivatização das aminas biogênicas nos padrões e amostras. Fonte: autor. ---------- 51

Figura 12 − Peixes analisados (**a**) e região de amostragem na posta do peixe (**b**). Fonte: autor. 53

Figura 13 − Cromatogramas das médias de cada nível de concentração do conjunto de calibração, (**a**) registrado em 215 nm, (**b**) entre as linhas vermelhas está a região selecionada para a modelagem, onde aparece as cinco aminas biogênicas derivatizadas analisadas, TRI (▬),

FEN (—), PUT (—), CAD (—) e HIS (—), (**c**) espectros dos analitos com o comprimento de onda de 215 nm destacado com uma linha preta e (**d**) os mesmos espectros normalizados. 59

Figura 14 − Cromatogramas das amostras de teste (—) alinhados por intervalos (**a-d**) e por deslocamento completo (**e**) sobreposto a amostra de calibração de 20 mg L^{-1} (área colorida representando cada analito TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—)). Cromatograma em 3D da amostra de teste 2 alinhado por intervalos (**a1-d1**) e por deslocamento completo (**e1**). Os intervalos testados foram 20 (**a**), 40 (**b**), 60 (**c**) e 120 (**d**).62

Figura 15 − Cromatogramas das amostras de testes (—) alinhados por intervalos (**a-d**) e por deslocamento completo (**e**) sobreposto à amostra de calibração de 20 mg L^{-1} (área colorida representando cada analito). Cromatograma em 3D da amostra de teste 2 alinhado por intervalos (**a1-d1**) e por deslocamento completo (**e1**). Os intervalos testados foram 5 (**a**), 10 (**b**), 15 (**c**) e 30 (**d**). ---------- 63

Figura 16 − Cromatogramas das amostras do conjunto de calibração sem alinhar (**a**) e alinhando (**c**). Cromatograma da amostra de calibração de 20 mg L^{-1} (colorida de acordo com o analito) sobreposta aos das amostras do conjunto de teste (—) com o conjunto de calibração não alinhado (**b**) e alinhado (**d**), as 4 regiões modeladas são mostradas entre linhas coloridas (Tabela 1S). ---------- 64

Figura 17 − Perfis recuperados para o PARAFAC com 1 (**a**), 2 (**b**) e 3 (**c**) componentes para os dados não alinhados. O perfil com 1 componente é representado por uma linha preta cheia (—), com 2 por uma linha tracejada cinza (---) e com 3 por uma linha cheia cinza (—). O sinal normalizado da amostra de calibração de 20 mg L^{-1} em cinza é sobreposto ao sinal normalizado recuperado pelo PARAFAC apenas para o sinal da TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—) (d).---------- 67

Figura 18 − Cromatograma em 3D da amostra de teste 5 registrado na faixa dos dois analitos juntos (**a**), e separadamente para a CAD (**b**) e para a HIS (**c**).---------- 68

Figura 19 − Perfis recuperados para o PARAFAC com 1 (**a**), 2 (**b**) e 3 (**c**) componentes para os dados alinhados. Os perfis recuperados com 1, 2 e 3 componentes são respectivamente representados por uma linha preta cheia (—), uma linha tracejada cinza (---) e uma linha cheia cinza (—). O sinal normalizado da amostra de calibração de 20 mg L^{-1} em cinza é sobreposto ao sinal normalizado recuperado pelo PARAFAC apenas para o sinal da TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—) (**d**).---------- 70

Figura 20 − Perfil recuperado pelo RBL no modo de eluição e comprimento de onda para a FEN (**a e d**) e PUT (**b e e**), para as amostras de teste 8 (—) e 9 (—). Em (**c**) é mostrado o sinal normalizado em 215 nm para as amostras de teste 8 (—) e 9 (—) e em cinza o perfil da amostra de calibração não alinhada e normalizada em 215nm da FEN e PUT no modo de

eluição. Em vermelho e verde são mostrados o sinal no modo espectral da PUT e FEN (**f**), respectivamente. ---------- 73

Figura 21 − Perfil recuperado pelo RBL no modo de eluição e comprimento de onda da FEN (**a e d**) para as amostras de teste 8 (—) e 9 (—) e da PUT (**b e e**) para as amostras de teste 7 (—), 8 (—) e 9 (—). Em (**c**) é mostrado o sinal normalizado em 215 nm para as amostras de teste 7 (—), 8 (—) e 9 (—) e em cinza o perfil da amostra de calibração alinhada e normalizada em 215 nm da FEN e PUT no modo de eluição. Em (**f**) são mostrados os espectros da PUT e FEN, respectivamente, em vermelho e verde.---------- 75

Figura 22 − Perfil dos pesos recuperado pelo RBL no modo de eluição e comprimento de onda para a CAD (**a e c**), para a amostra de teste 6 (—). (**b**) Sinal normalizado da CAD em 215 nm para a amostra de teste 6 (—) e em cinza o perfil da amostra de calibração de 20 mg L^{-1} alinhada e normalizada. (**d**) Espectros normalizado da CAD registrado no tempo de retenção para a amostra de calibração de 20 mg L^{-1} (—).---------- 77

Figura 23 − Perfis recuperados no modo de eluição para as amostras de teste 7 (**a, a1**), 8 (**b, b1**) e 9 (**c, c1**) com um (**a-c**) e dois (**a1-c1**) perfis recuperados para os dados não alinhados. Apenas a TRI foi recuperado com um só perfil em ambos os casos. TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—). ---------- 82

Figura 24 − Autovalor versus número de componentes para os dados aumentados no modo de eluição (**a**), zoom do gráfico a (**a1**), modo espectral (**b**) e zoom gráfico b (**b1**). TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—).---------- 85

Figura 25 − Perfil de eluição recuperado para os dados alinhados e aumentados no modo espectral, com 1 componente para TRI, CAD e HIS e 2 para FEN e PUT (**a**) e com 2 componentes para CAD e HIS (**b**), mantendo iguais o dos outros analitos. Em cinza o perfil normalizado da amostra de calibração mais concentrada. TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—).---------- 86

Figura 26 − Perfil aumentado no modo espectral para o modelo de TRI (**a**), FEN (**b**), PUT (**c**), CAD (**d**) e HIS (**e**) com (linha preta) e sem (linha colorida) restrição de trilinearidade. Em (**f**) são mostrados os perfis normalizados do modo de eluição para os modelos com (linha preta) e sem (linha colorida) restrição de trilinearidade. TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—).---------- 88

Figura 27 − Cromatogramas registrados em 215 nm para as amostras de *Salminus maxillosus* (**a e a1**), *Scomber scombrus* (**b e b1**), *Seriola dumerili* (**c e c1**) e *Lutjanus analis* (**d e d1**), sem fortificação (—) e fortificadas no nível 1 (—), 2 (—), 3 (—) e 4 (—), segundo o procedimento descrito na seção 3.6. Em cinza é mostrado a amostra de calibração de 20 mg L^{-1} alinhada (**a-d**) e sem alinhamento (**a1-d1**). ---------- 93

Figura 28 − Cromatograma em 2 D (**a**) e em superfície de contorno da amostra de *Seriola dumerili* sem fortificação (—) (**b**) e fortificadas no nível 1 (—) (**c**), 2 (—) (**d**), 3 (—) (**e**) e 4 (—) (**f**), segundo o procedimento descrito na seção 3.6. Os analitos e interferentes (**INT**) são apontados com setas. Em (**a**) o pico cinza representa o sinal do analito na concentração de 20mg L^{-1}. ---------- 94

Figura 29 − Perfil cromatográfico recuperado pelo MCR-ALS para as amostras *Salminus maxillosus* (**a**), *Scomber scombrus* (**b**), *Seriola dumerili* (**c**) e *Lutjanus analis* (**d**). Os analitos são representados como linhas coloridas, TRI (—), FEN (—), PUT (—), CAD (—), HIS (—) e os interferentes como linhas cinzas (—). ---------- 95

Figura 30 − Perfil espectral aumentado recuperado pelo MCR-ALS para a TRI (**a1-a4**) e para o interferente (**b1-b4**) nas fortificações das amostras de *Salminus maxillosus* (**1**), *Scomber scombrus* (**2**), *Seriola dumerili* (**3**) e *Lutjanus analis* (**4**) realizadas no primeiro dia de análise. O espectro da TRI e do interferente em uma concentração foi normalizado para comprimento 1 e sobreposto para comparação (**c1-c4**). ---------- 96

Figura 31 − Perfil espectral recuperado da TRI (**a**), FEN (**b**), PUT (**c**), CAD (**d**) e HIS (**e**) para o conjunto de calibração (—) e para as amostras de *Salminus maxillosus* (—), *Scomber scombrus* (—), *Seriola dumerili*(—) e *Lutjanus analis*(—). ---------- 98

Figura 32 − Valor predito versus o nominal fortificado de TRI (**a**), FEN (**b**), PUT (**c**), CAD (**d**) e HIS (**e**) para o conjunto de calibração (—) e para as amostras de *Salminus maxillosus* (—), *Scomber scombrus* (—), *Seriola dumerili* (—) e *Lutjanus analis*(—).---------- 100

Figura 33 − Evolução da concentração da TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—) ao longo de 13 dias de análise das amostras de *Salminus maxillosus* (**a**), *Scomber scombrus* (**b**), *Seriola dumerili* (**c**) e *Lutjanus analis* (**d**). A barra de erro mostrada em cada concentração predita foi calculada de acorda na interface MVC2 (Olivieri, Wu, Yu, 2009; Olivieri, 2014). A Linha roxa (—) em 50 mg kg^{-1} representa o limite máximo estabelecido para o peixe estar seguro para o consumo.---------- 103

Figura 34 − Cromatograma registrado em 215 nm para a amostra de *Scomber scombrus* registrado nos dias 1(—), 3(—), 5(—), 7(—), 9(—), 11(—), 13(—) e os cromatogramas preenchidos foram registrados para TRI (Azul), FEN (Verde), PUT (Vermelho), CAD (Ciano) e HIS (Margenta) na amostra de calibração de10 mg L^{-1}. ---------- 105

LISTA DE TABELAS

Tabela 1 − Conjunto de teste das aminas biogênicas em quarto níveis de concentração (mg L^{-1}). ---------- 52

Tabela 2 − Modelos construídos com o conjunto de calibração não alinhados, para o cromatograma dividido em regiões das misturas das ABs presentes no conjunto teste. ---------- 66

Tabela 3 − Modelos construídos com o conjunto de calibração alinhados para o cromatograma dividido em regiões das misturas das ABs presentes no conjunto teste. ---------- 69

Tabela 4 − Métricas de desempenho analítico calculadas para os cinco modelos PARAFAC aplicadas ao conjunto de teste com o de calibração alinhado. As concentrações dos conjuntos de teste e de calibração variaram respectivamente nas faixas de 0,4 a 16 mg L^{-1} e 0,05 a 20 mg L^{-1} para todos os 5 analitos. ---------- 71

Tabela 5 − Variância explicada e ajuste residual para os modelos construídos para cada analito com uma VL com base nos dados não alinhados. ---------- 72

Tabela 6 − Variância explicada e ajuste residual para os modelos construídos para cada analito com uma VL com os dados alinhados. ---------- 74

Tabela 7 − Métricas de desempenho analítico calculadas para os cinco modelos UPLS-RBL construído com o conjunto de calibração alinhado. As concentrações dos conjuntos de teste e de calibração variaram respectivamente nas faixas de 0,4 a 16 mg L^{-1} e 0,05 a 20 mg L^{-1} para todos os 5 analitos. ---------- 76

Tabela 8 − Autovalores para a matriz de dados de teste e de calibração, não alinhados, aumentado no modo de eluição nas respectivas faixas selecionadas para cada AB. ---------- 79

Tabela 9 − Variância explicada e ajuste residual dos dados não alinhados aumentados no modo de eluição para a decomposição MCR-ALS dos cinco analitos nas respectivas faixas selecionadas. ---------- 80

Tabela 10 − Variância explicada e ajuste residual do resultado otimizado pelo MCR-ALS com os dados não alinhados. ---------- 82

Tabela 11 − Variância explicada e falta de ajuste dos modelos bilineares e trilineares construídos para as ABs com os dados alinhados. ---------- 84

Tabela 12 − Autovalores para os dados alinhados das ABs aumentados no modo de eluição e dos espectros de 1 a 5 componentes. ---------- 85

Tabela 13 − Variância explicada e ajuste residual dos dados alinhados aumentados no modo espectral para a decomposição MCR-ALS dos cinco analitos nas respectivas faixas selecionadas. ---------- 86

Tabela 14 − Métricas de desempenho analítico calculadas para os cinco modelos MCR-ALS com o conjunto de calibração alinhado, a restrição de trilinearidade não foi aplicada. ---------- 87

Tabela 15 − Efeito da restrição de trilinearidade no modelo das amostras de teste para os dados alinhados.---------- 89

Tabela 16 − Ambiguidade rotacional do MCR-ALS no modelo da amostra de Seriola dumerili. ---------- 99

Tabela 17 − Teste de recuperação das aminas biogênicas nas amostras de Seriola dumerili usando o MCR-ALS. ---------- 102

Tabela 18 − Metodologias cromatográficas reportadas na literatura para quantificação de aminas biogênicas em peixes.---------- 107

LISTA DE EQUAÇÕES

(1) $\mathbf{x_{IxJxK}} = \sum_1^{\mathbf{I}} \prod_1^{\mathbf{3}} \mathbf{a} + \mathbf{e_{IxJxK}} = \sum_1^{\mathbf{I}} \mathbf{a_{IxD}}\, \mathbf{b_{JxE}} \mathbf{c_{KxF}} + \mathbf{e_{IxJxK}}$ ---------- 33

(2) $\boldsymbol{X} = \mathbf{A}(\mathbf{B}| \otimes |\mathbf{C})^{\boldsymbol{t}} + \boldsymbol{E}$ ---------- 33

(3) $\mathbf{X_{Cal}} = \mathbf{P}\mathbf{T}^{\mathbf{T}} + \mathbf{E_{Cal}}$ ---------- 36

(4) $\mathbf{y_{Cal}} = \mathbf{T}\mathbf{v} + \mathbf{e_y}$ ---------- 36

(5) $\mathbf{x_u} = (\mathbf{P}\, \mathbf{t_u^T}) + \mathbf{B}\mathbf{G}\mathbf{C}^{\mathbf{T}} + \mathbf{e_u}$ ---------- 37

(6) $\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathbf{T}} + \mathbf{E}$ ---------- 38

(7) $\mathbf{D} = \mathbf{C}\mathbf{T}\mathbf{T}^{-\mathbf{1}}\mathbf{S}^{\mathbf{T}} + \mathbf{E}$ ---------- 42

(8) $\boldsymbol{f_n} = \frac{\|\mathbf{C_n}\mathbf{S_n^T}\|}{\|\mathbf{C}\mathbf{S}^{\mathbf{T}}\|}$ ---------- 42

(9) $\mathbf{y} = \mathbf{a}\mathbf{x} + \mathbf{b}$ ---------- 46

(10) $\mathbf{S_{12}} = \frac{\|\mathbf{s_1^T}\mathbf{s_2}\|}{\|\mathbf{s_1}\|\|\mathbf{s_2}\|}$ ---------- 97

LISTA DE ACRÔNIMOS, SIGLAS E SÍMBOLOS

γ– Sensibilidade analítica;

AB - Amina biogênica;

ALS - Mínimos quadrados alternados, do inglês *Alternanting Least Squares*;

C18– Octadecil;

CAD– Cadaverina;

CORCONDIA - Diagnóstico de consistência de core, do inglês *CORe CONsistency DIAgnostic*;

Coshift - Deslocamento correlacionado, do inglês *COrrelation SHIFTing*;

COW - Alinhamento por Correlação Otimizada, do inglês *Correlation Optimized Warping*;

DAD - Detector de Arranjo de Diodos, do inglês *Diode Array Detector*;

DNS– Cloreto de dansil;

ESFA - *European Food Safety Authority*;

FAO/WHO - *Food and Agriculture Organization of the United Nations and World Health Organization*;

FEN– Feniletilamina;

FLU - Fluorescência molecular;

HIS– Histamina;

HPLC - Cromatografia líquida de alta performance, do inglês *High Performance Liquid Chromatography*;

i*Coshift* - Deslocamento correlacionado por intervalos, do inglês *Interval COrrelation SHIFTing*;

LC-DAD - Cromatografia Líquida com Detector de Arranjo de Diodos, do inglês *Liquid Chromatography Diode Array Detector*;

LOD - Limite de detecção, do inglês *Limit of Detection*;

LOQ - Limite de quantificação, do inglês *Limit of Quantification*;

MCR-ALS - Resolução de curvas multivariadas com mínimos quadrados alternados, do inglês *Multivariate Curve Resolution Alternating Least Square*;

MS - Espectrometria de massas, do inglês *Mass Spectrometry*;

NAS - Sinal Analítico Líquido, do inglês *Net Analyte Signal*;

OPA - o-ftaldialdeido;

PARAFAC - Análise de fatores paralelos, do inglês *PARAllel FACtor analysis*;

PARAFAC2 - Análise de fatores paralelos 2, do inglês *PARAllel FACtor analysis 2*;

PARALIND - Perfis Paralelos com Dependência Linear, do inglês *PARAllel profiles with LINear Dependencies*;

PCA - Análise de Componentes Principais, do inglês, *Principal Component Analysis*;

PLS – Mínimos quadrados parciais, do inglês *Partial Least Squares*;

PUT – Putrescina;

REP – Erro Relativo de Predição do inglês, *Relative Error of Prediction*;

SEN – Sensibilidade;

SIMPLISMA – Análise de mistura auto modelável iterativa simples de usar, do inglês *Simple to use interactive self modeling mixture analysis*;

SVD - Decomposição de Valores Singulares, do inglês *Single Value Decomposition*;

TRI – Triptamina;

UPLS-RBL -Mínimos quadrados parciais desdobrados com bilinearização residual, do inglês *Unfolded Partial Least-Squares with Residual Bilinearization*;

UV– Ultravioleta.

SUMÁRIO

RESUMO vi

ABSTRACT vii

LISTA DE FIGURAS viii

LISTA DE TABELAS xii

LISTA DE EQUAÇÕES xiv

LISTA DE ACRÔNIMOS, SIGLAS E SÍMBOLOS xv

1 INTRODUÇÃO 20

1.1 Caracterização da problemática 20

1.2 Objetivo 23

1.2.1. Objetivo geral 23

1.2.2. Objetivos específicos 23

2 FUNDAMENTAÇÃO TEÓRICA 25

2.1 Aminas biogênicas e sua quantificação por cromatografia 25

2.2 Derivatização 27

2.3 Calibração univariada e multivariada: vantagens e limitações 28

2.4 Planejamento Taguchi 32

2.5 PARAFAC 33

2.6 UPLS-RBL 35

2.7 MCR-ALS 38

2.8 Deficiência de posto 43

2.9 Algoritmos de alinhamento de pico 45

2.10 Métricas de desempenho analítico 46

3 METODOLOGIA 50

3.1 Reagentes 50

3.2 Eluição e parâmetros cromatográficos 50

3.3 Derivatização 50
3.4 Conjunto de calibração e de teste 51
3.5 Amostras, amostragem e preparo da amostra 52
3.6 Teste de recuperação e estudo de degradação 53
3.7 Procedimento quimiométrico e software 54
3.7.1. PARAFAC 54
3.7.2. MCR-ALS 55
3.7.3. UPLS-RBL 55
3.7.4. Software 56
4 RESULTADOS E DISCUSSÃO 58
4.1 Conjunto de calibração 58
4.2 Avaliação do algoritmo de alinhamento nos conjuntos de calibração e teste 60
4.3 Tratamento dos dados com o PARAFAC 65
4.4 Tratamento dos dados com o UPLS-RBL 71
4.5 Tratamento dos dados com o MCR-ALS 78
4.6 Considerações parciais 90
4.7 Quantificação das aminas biogênicas em amostras de peixe 92
4.8 Estudo cinético de degradação 103
4.9 Comparação da metodologia proposta com outras publicadas na literatura 105
5 CONCLUSÕES 109
5.1 Propostas futuras 109
REFERÊNCIAS 111
APÊNDICE 116
a) Tratamento dos resíduos gerados 116
b) Figuras 117
c) Tabelas 138
d) Artigos 140

Capítulo 1

INTRODUÇÃO

1 INTRODUÇÃO

1.1 Caracterização da problemática

As aminas biogênicas (ABs) são naturalmente produzidas pela ação de enzimas na descarboxilação de aminoácidos naturais. A formação de ABs é usada como marcador de degradação sendo a histamina (HIS), putrescina (PUT), cadaverina (CAD), triptamina (TRI) e a feniletilamina (FEN), as mais frequentemente identificadas em alimentos. Assim sua determinação é importante no controle de qualidade desses produtos (Erim, 2013; EFSA, 2011; FAO/WHO, 2013). Apesar de fundamentais no desempenho de funções biológicas no organismo humano, em especial no sistema nervoso central, em elevadas concentrações podem causar diversos problemas no organismo como: reações alérgicas que levam à dificuldade para respirar, erupções cutâneas, vômito, febre e hipertensão que podem levar até à hemorragia cerebral (EFSA, 2011; FAO/WHO, 2013). Entre as aminas biogênicas presentes em alimentos, apenas a HIS é controlada por órgão regulamentadores como a ESFA (*European Food Safety Authority*) e a FAO/WHO (*Food and Agriculture Organization of the United Nations/World Health Organization*), sendo sua concentração máxima aceitável de 50 mg kg^{-1}. A quantificação de outras ABs em alimentos é importante, pois sua presença pode potencializar a toxicidade da histamina (Erim, 2013; EFSA, 2011; FAO/WHO, 2013). Assim, é evidente a importância da quantificação das ABs, para o controle de qualidade de alimentos e a consequente manutenção da saúde humana.

As ABs podem ser encontradas em bebidas fermentadas como cerveja e vinho, em produtos lácteos como leite, iogurte e queijos e em produtos cárneos como linguiça, carnes e peixes. Todas essas amostras são complexas, pois contém outros constituintes em concentração maiores do que as ABs, requerendo manipulação da amostra como extração, limpeza e/ou derivatização. A análise das amostras de carne e peixe, além dessa complexidade, é ainda mais desafiadora devido à heterogeneidade da distribuição das ABs na amostra. Vale ressaltar a importância da quantificação das ABs em peixes por estes apresentarem grande quantidade de histidina livre nos músculos e por conterem microrganismos que liberam enzimas que aceleram a descarboxilação, o que facilita a conversão desta amina em histamina (EFSA, 2011; FAO/WHO, 2013; Ordóñez, Troncoso, García-Parrilla, Callejón, 2016).

Diversas metodologias são utilizadas para a quantificação de histamina em peixes, destacando-se as semi quantitativas como ensaios imunológicos do tipo ELISA e colorimétricos que são feitos em menos de uma hora por amostra; e as quantitativas como os métodos fluorimétricos, em que a detecção é rápida, mas requer uma etapa de derivatização seletiva da amostra. Entretanto, devido à quantidade de aminas e outros constituintes presentes na amostra,

não é possível a análise de outras ABs por estas metodologias, sendo necessário o uso de técnicas cromatográficas para a detecção de múltiplas aminas (EFSA, 2011; FAO/WHO, 2013).

A quantificação simultânea das ABs pode ser feita tanto por cromatografia gasosa quanto por cromatografia líquida, sendo a mais comum por cromatografia líquida acoplada à espectrometria no ultravioleta (UV), de fluorescência molecular (FLU) e de massas (MS, do inglês *Mass spectrometry*). Na análise por cromatografia líquida, após a extração das ABs presentes nas amostras é necessário realizar uma derivatização pré ou pós-coluna, pois muitas aminas biogênicas não possuem cromóforos para serem identificadas por detectores ópticos como UV ou FLU (Mohammed, Bashammakh, Alsibaai, Alwael, El-Shahawi, 2016). A reação de derivatização pós coluna, em geral, é feita com o reagente OPA (o-ftaldialdeido) e só altera a estrutura do analito após a separação na coluna. Por apresentarem estruturas pequenas e possuírem baixa quantidade de carbonos, as ABs apresentam baixa retenção na coluna em fase reversa de octadecil (C18), o que pode levar à coeluição com outros constituintes presentes no extrato.

A derivatização pré-coluna geralmente é feita com cloreto de dansil (DNS), este tem a vantagem de aumentar a estrutura das ABs antes da separação cromatográfica, aumentando a retenção dos analitos, o que minimiza a possibilidade de coeluição com outros constituintes (Erim, 2013; Mohammed, Bashammakh, Alsibaai, Alwael, El-Shahawi, 2016; Ordóñez, Troncoso, García-Parrilla, Callejón, 2016). Mesmo aumentando a retenção das ABs na coluna é possível aparecer interferência na análise devido à sua coeluição com outras aminas, uma vez que a reação de derivatização com o DNS é seletiva a aminas primária e secundária, e na amostra pode haver outras não contempladas no estudo (Erim, 2013; Ordóñez, Troncoso, García-Parrilla, Callejón, 2016; Mohammed, Bashammakh, Alsibaai, Alwael, El-Shahawi, 2016). Nesses casos em que são identificadas interferências na análise, abordagens quimiométricas podem ser usadas permitindo a quantificação exata, mesmo na presença de interferentes, usando um algoritmo apropriado na modelagem do conjunto de dados (Arancibia, Damiani, Escandar, Ibañez, Olivieri, 2012; Escandar, Goicoechea, Peña, Olivieri, 2014).

O detector mais utilizado para a modelagem com algoritmo de segunda ordem é o DAD, devido à elevada frequência de aquisição de espectros para cada tempo de retenção, quando comparado a detectores de FLU e ao baixo custo de manutenção e análise frente ao MS. No entanto, como a maior parte da estrutura molecular dos analitos e interferentes derivatizados é referente ao DNS, os espectros registrados no UV são muito similares e em alguns casos idênticos, causando deficiência de posto ao se tentar resolver o problema de interferência com calibração de segunda ordem (PINTO et al, 2016).

A deficiência de posto refere-se a uma quebra da trilinearidade do conjunto de dados no modo espectral devido à presença de perfis espectrais muito correlacionados e pode ser resolvida

usando os algoritmos PARALIND (do inglês, *PARAllel profiles with LINear Dependencies*), UPLS-RBL (do inglês, *Unfolded Partial Least-Squares with Residual Bilinearization*) ou o MCR-ALS (do inglês, *Multivariate Curve Resolution and Alternating Least Square*) aumentado no modo espectral ou com restrição de trilinearidade (Culzoni, et al., 2011; Culzoni, et al., 2008; Parastar, 2015; Jaumot, de Juan, Tauler, 2015; Amrhein, Srinivasan, Bonvin, Schumacher, 1996; Bro, Harshman, Sidiropoulos, Lundy, 2009; Borraccetti, Damiani, Olivieri, 2009). No entanto, para dados de cromatografia líquida há também a quebra da trilinearidade no modo da eluição devido aos deslocamentos do tempo de retenção dos analitos. Dessa forma, nenhum desses algoritmos, por si só, apresentariam resultados satisfatórios para dados de cromatografia líquida com deficiência de posto.

Para resolver esse problema a literatura reporta que é necessário adicionar à modelagem dos dados uma informação proveniente de um detector mais seletivo (Peré-Trepat Tauler, 2006; Parastar, 2015). Esta estratégia nem sempre é a mais efetiva, uma vez que irá elevar o custo da análise, a complexidade da modelagem e por não haver um detector mais seletivo que o usado.

Sendo assim, neste trabalho é proposto resolver a quebra da trilinearidade ao alinhar os picos com o algoritmo Coshift (do inglês, *COrrelation SHIFTing*) e para contornar o problema de interferência foram testados os algoritmos mais usados no tratamento de dados de segunda ordem, PARAFAC (do inglês, *PARAllel FACtor analysis*), UPLS-RBL (do inglês, *Unfolded Partial Least-Squares with Residual Bilinearization*) e o MCR-ALS aumentado no modo espectral e com restrição de trilinearidade. Para confirmar a trilinearidade dos dados após o alinhamento, foram avaliadas a variância explicada e a falta de ajuste dos modelos bilineares e trilineares na modelagem usando o MCR-ALS. A eficiência de cada um desses algoritmos foi avaliada na modelagem dos dados com deslocamento de pico e deficiência de posto em amostras sintéticas, e o que apresentou o melhor desempenho foi usado na modelagem das amostras de peixe.

1.2 Objetivo

1.2.1. Objetivo geral

Alcançar a vantagem de segunda ordem em dados com deficiência de posto e desalinhamento na retenção dos analitos, na quantificação de aminas biogênicas em peixes, usando cromatografia líquida de alta eficiência com detecção por arranjo de diodos, alinhamento com o *coshift* e calibração com o algoritmo MCR-ALS.

1.2.2. Objetivos específicos

1 Identificar nas amostras de teste e de peixe os analitos cuja retenção está desalinhada em relação ao conjunto de calibração;

2 Identificar os constituintes em cada região modelada que causam a deficiência de posto nas amostras sintéticas e de peixe;

3 Avaliar o desempenho do alinhamento de picos usando *coshift* e o i*coshift* para as amostras sintéticas e do *coshift* para as amostras de peixe;

4 Confirmar a trilinearidade dos dados após esse alinhamento;

5 Avaliar e comparar o desempenho dos algoritmos de calibração de segunda ordem quanto as figuras de mérito: PARAFAC, MCR-ALS e UPLS-RBL na quantificação de cinco aminas biogênicas, triptamina, 2-fenileltilamina, putrescina, cadaverina e histamina, em amostras sintéticas antes e após o alinhamento de picos usando o algoritmo *coshift* e i*coshift*;

6 Quantificar as aminas biogênicas triptamina, 2-fenileltilamina, putrescina, cadaverina e histamina em amostras de peixes com o alinhamento de picos usando o algoritmo *coshift* e o MCR-ALS aumentado no modo espectral;

7 Realizar um estudo de degradação dos peixes acondicionados a 4 ℃, durante 13 dias, com análise da amostra de peixe a cada dois dias.

Capítulo 2

FUNDAMENTAÇÃO TEÓRICA

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Aminas biogênicas e sua quantificação por cromatografia

As aminas biogênicas (ABs) são compostos alcalinos de estrutura molecular pequena, formadas, principalmente, pela descarboxilação de aminoácidos livres, realizada por microrganismos exógenos, resultante das condições de baixa higienização (Erim., 2013; Pinto, Díaz Nieto, Zón, Fernández, Araújo, 2016; Mohammed, Bashammakh, Alsibaai, Alwael, El-Shahawi, 2016). A quantidade e o tipo de ABs depende fortemente do tipo do alimento e do microrganismo, sendo encontrados em peixes, carnes, produtos fermentados e seus derivados (Mohammed, Bashammakh, Alsibaai, Alwael, El-Shahawi, 2016). A elevada quantidade de histidina livre em peixes faz com que o monitoramento das ABs nesses alimentos seja de grande importância, uma vez que a sua descarboxilação gera histamina (HIS) que é a única AB controlada por órgãos regulamentadores como a FAO/WHO e a ESFA (Erim, 2013).

Para a proteção da saúde pública, a FAO/WHO e a ESFA estabelecem 50 mg L^{-1} como sendo o limite máximo de histamina para a maioria dos alimentos (EFSA, 2011; FAO/WHO, 2013). Apesar da HIS ser a única AB cujo controle é regulamentado, outras como triptamina (TRI), putrescina (PUT), 2-feniletilamina (FEN) e cadaverina (CAD) também são destacadas como sendo de importante controle. Essas ABs elevam o potencial toxicológico da HIS, fazendo com que quantidades menores sejam nocivas à saúde humana (Erim, 2013; Mohammed, Bashammakh, Alsibaai, Alwael, El-Shahawi, 2016). Pelo exposto, é de fundamental importância para a saúde da população o monitoramento dessas ABs. As técnicas mais utilizadas para o monitoramento de ABs em amostras de peixe são as cromatográficas com detector de arranjo de diodos, fluorescência e massas (Erim, 2013; Mohammed, Bashammakh, Alsibaai, Alwael, El-Shahawi, 2016; Herrero, Sanllorente, Reguera, Ortiz, Sarabia).

A cromatografia líquida é uma técnica de separação físico-química, na qual ocorre a separação dos constituintes presentes na amostra através do seu equilíbrio com duas fases imiscíveis, uma estacionária (coluna) e uma móvel (solvente) (Collins , 2006; Lanças, 2009). A cromatografia líquida em fase reversa com uma coluna C18 é amplamente utilizada em diversas aplicações e análises, sendo apropriado para separar constituintes com caráter mais apolar e com carbonos em sua estrutura molecular.

Quanto mais carbono tiver na estrutura molecular do constituinte a ser analisado, maior será a retenção do analito nesta coluna. Uma vez que a estrutura molecular das ABs, mostrada na **Figura 01 a-e**, não possui uma grande quantidade carbonos, a retenção na coluna é baixa, sendo necessário uma coluna grande para distanciar seu sinal do tempo morto da coluna e de outros

constituintes, para não haver coeluição. Ainda assim é elevada a probabilidade de coeluições com interferentes.

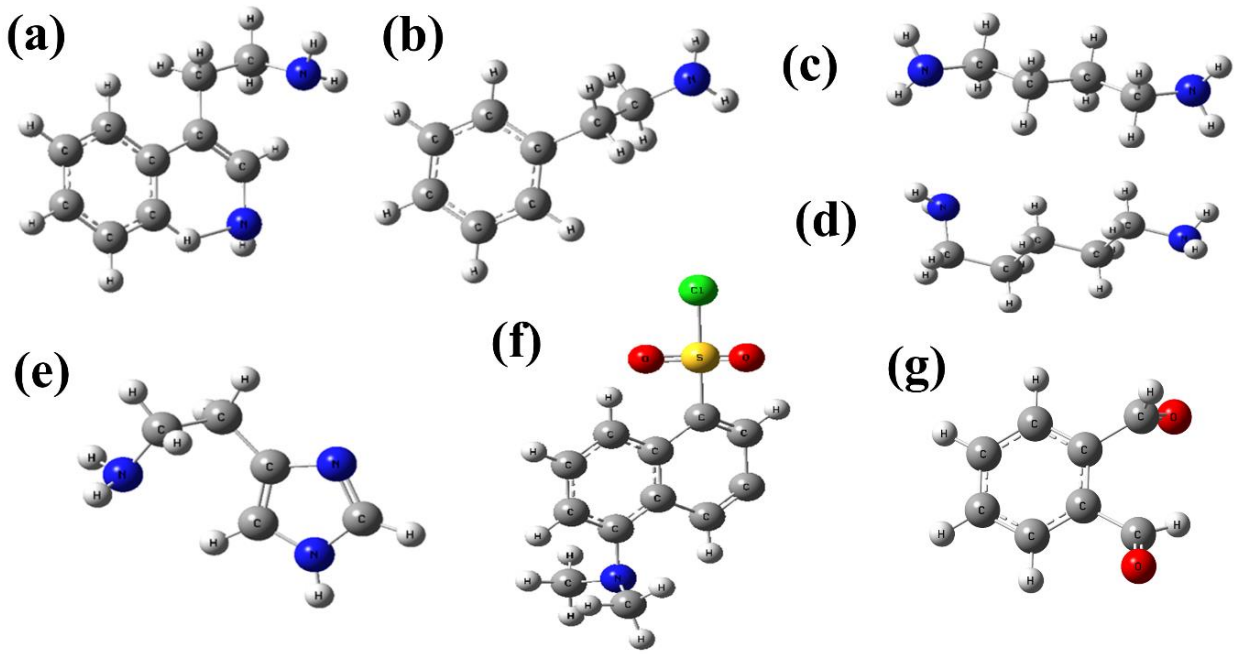

Figura 1 − Estrutura molecular da TRI (**a**), FEN (**b**), PUT (**c**), CAD (**d**), HIS (**e**) e dos reagentes derivatizantes DNS (**f**) e OPA (**g**). Esferas brancas, cinzas, azuis, vermelhas, verde e amarela, indicam hidrogênios, carbonos, nitrogênios, oxigênios, cloro e enxofre, respectivamente. Fonte: autor.

Outra característica da estrutura das ABs é a elevada energia de transição eletrônica proveniente dos cromóforos existentes, em especial a PUT e a CAD, em que a principal transição eletrônica é a da ligação amina (-NH_2), o que dificulta sua quantificação utilizando detectores convencionais de ultravioleta (UV). Para a análise por detector de fluorescência molecular (FLU) é necessário registrar a emissão do fóton quando a molécula excitada voltar ao seu estado fundamental. Isso ocorre em moléculas com estruturas rígidas, ou seja, com menor liberdade vibracional e ricas em ligações π. Nem a PUT nem a CAD possuem ligações π, e todas as ABs possuem liberdade vibracional, o que reduz o seu sinal fluorescente (Erim, 2013).

No caso da espectrometria de massas é possível a análise de qualquer substância ionizável, uma vez que o sinal registrado é a razão massa carga, assim é possível a análise com elevada sensibilidade dessas ABs (Mohammed, Bashammakh, Alsibaai, Alwael, El-Shahawi, 2016; Erim, 2013). Este detector não é tão popular devido ao seu elevado custo de aquisição e, principalmente, de manutenção. Diante disso, outras alternativas, usando detectores de UV e FLU, vem sendo utilizadas na análise das ABs. A derivatização pré-coluna usando o dansil (estrutura molecular na **Figura 1 f**) é a estratégia mais utilizada para a análise com ambos os detectores óticos.

2.2 Derivatização

Há duas formas de derivatizar as ABs, uma usando o reagente cloreto de dansil (DNS) **(Figura 1f)** e outra usando o-ftaldialdeido (OPA) **(Figura 1g)**, ambas são eficientes em aumentar a sensibilidade das ABs com detecção por UV e FLU. A derivatização pode ser feita antes da eluição (pré-coluna) com ambos os reagentes e após a eluição (pós-coluna) apenas com o OPA, pois a reação com o DNS é mais lenta, podendo ser necessário 1 hora para se completar, o que torna inviável para acoplar em uma análise em fluxo.

Apesar de ser mais lenta, a derivatização pré-coluna com o DNS é a mais utilizada devido a sua facilidade de implementação, uso de reagentes menos tóxicos, possibilitando uma maior retenção do analito derivatizado na coluna e por resultar em uma molécula com mais cromóforos e insaturações, como pode ser visto nas **Figuras 02 e 03**. Esta última característica é bastante desejável uma vez que a sensibilidade do detector para a molécula derivatizada com o DNS será maior do que com o OPA (Chen, et al., 2010; Latorre-Moratalla, et al., 2009; Mey, et al., 2012; Zhai, et al., 2012; Hu, Huang, Li, Yang, 2012; Simat Dalgaard, 2011).

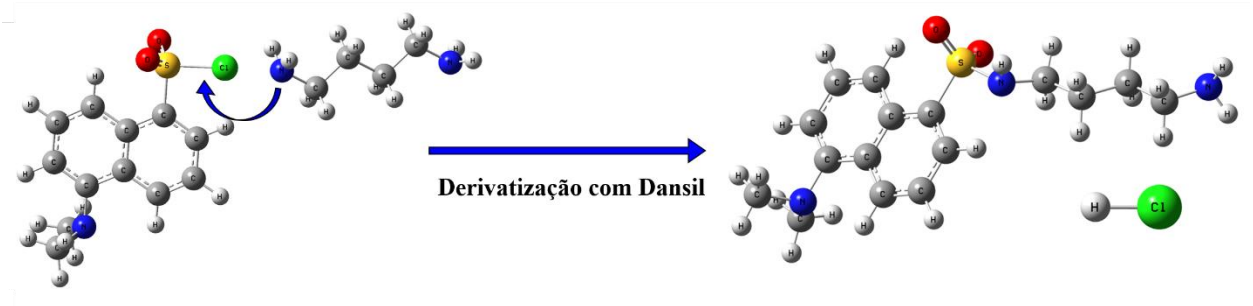

Figura 2 – Reação de derivatização da PUT com o Dansil. Esferas brancas, cinzas, azuis, vermelhas, verde e amarela indicam hidrogênios, carbonos, nitrogênios, oxigênios, cloro e enxofre, respectivamente. Fonte: autor.

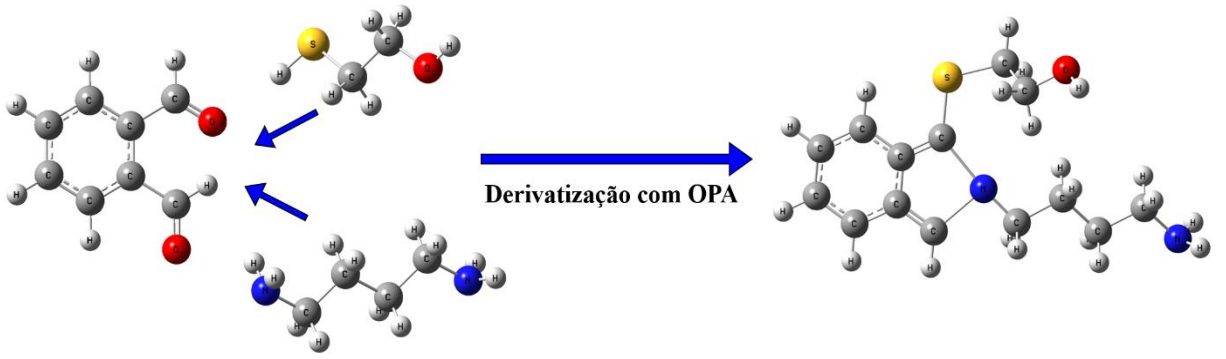

Figura 3 – Reação de derivatização da PUT com o OPA e o 2-mercaptoetanol. Esferas brancas, cinzas, azuis, vermelhas, verde e amarela indicam hidrogênios, carbonos, nitrogênios, oxigênios, cloro e enxofre, respectivamente. Fonte: autor.

Nas **Figuras 02 e 03** são mostradas as reações de derivatização da PUT com o DNS e com o OPA, respectivamente. A derivatização com o OPA é rápida podendo feita com sistema em fluxo

fechado e acoplado após o módulo da coluna. Uma vez que esta reação necessita do uso de reagentes tóxicos como 2-mercaptoetanol, deve-se realizá-la em sistema fechado ou em uma capela, o que torna a automação desta reação desejável por reduzir o contato deste reagente com o analista (Latorre-Moratalla, et al., 2009).

Algumas implicações decorrem da derivatização usada. A corrida é reduzida ao se optar pelo uso do OPA no pós-coluna, pois as ABs possuem uma retenção bem menor antes da derivatização. Isso fará com que seja gasto menos solvente e a frequência analítica aumente. No entanto, a probabilidade de interferências é elevada, uma vez que as ABs serão eluídas mais próximas do volume morto da coluna, o que comprometerá a sua quantificação (Chen, et al., 2010; Latorre-Moratalla, et al., 2009).

A estrutura molecular do analito que será eluído é bem maior quando a derivatização ocorre pré-coluna, isso resultará em corridas mais longas e maior consumo de solventes, uma vez que a força da fase móvel deve ser maior para eluí-lo da coluna. Se por um lado o aumento de retenção acarreta no elevado consumo de solventes, por outro lado o analito se distancia do tempo morto ao aumentar sua retenção na coluna reduzindo a probabilidade de coeluição com interferentes. Na derivatização com o DNS a estrutura molecular da AB derivatizada será maior e terá mais cromóforos do que quando o OPA é usado, o que pode aumentar a sensibilidade analítica na detecção por espectrometria molecular na região do ultravioleta (Skoog, Crouch, Holler, 2009).

Ao derivatizar com o OPA é necessário sincronizar esta reação com a injeção da amostra e com o sistema de detecção. Essa não é uma tarefa simples, pois os softwares cromatográficos são fechados, não sendo disponibilizados códigos de interação com outras interfaces ou outros softwares. Essa falta de sincronização pode dificultar atingir a vantagem de segunda ordem no tratamento de dados, devido à quebra da trilinearidade no modo de eluição.

2.3 Calibração univariada e multivariada: vantagens e limitações

Calibração é uma operação que relaciona uma propriedade que se deseja conhecer com uma propriedade medida, geralmente, em um instrumento. Na química analítica o instrumento é calibrado para se obter a concentração de um constituinte a partir de uma propriedade medida no detector utilizado. O pHmetro é o equipamento mais usado para esse fim e a calibração é feita com soluções padrões cujo valor de pH é correlacionado de forma linear com o potencial medido (Olivieri, 2008; 2014).

Há diversos equipamentos que podem ser usados para correlacionar a concentração de um analito com uma propriedade instrumental medida. A escolha do detector é crucial para uma quantificação exata e depende, principalmente, do analito, da amostra e de como e qual

concentração deste analito se encontra na amostra. Neste trabalho, foi usada a espectrometria molecular na região do ultravioleta por arranjo de diodos com prévia separação em um sistema de cromatografia em fase líquida (LC-DAD, do inglês *Liquid Chromatography Diode Array Detector*). Esta detecção foi escolhida por permitir a quantificação do analito com exatidão em concentrações na ordem de µg L^{-1}, uma vez que os constituintes da mistura são separados em uma coluna cromatográfica previamente à detecção (Olivieri, 2008).

Este tipo de detecção gera uma quantidade significativa de informação, que pode acarretar em vantagens para o processo de calibração (Olivieri, 2008; 2014). Na **Figura 4** é mostrada a forma de organização para diversos tipos de dados, dependendo do tipo ou ordem da calibração. Na cromatografia, a área do pico de um padrão analítico eluído pode ser correlacionada com a sua concentração e assim construir um modelo de calibração univariada. Para quantificar a concentração do analito em uma amostra desconhecida, é necessário apenas interpolar a área do pico do constituinte de interesse no modelo previamente ajustado, com padrões, em geral usando o método dos mínimos quadrados.

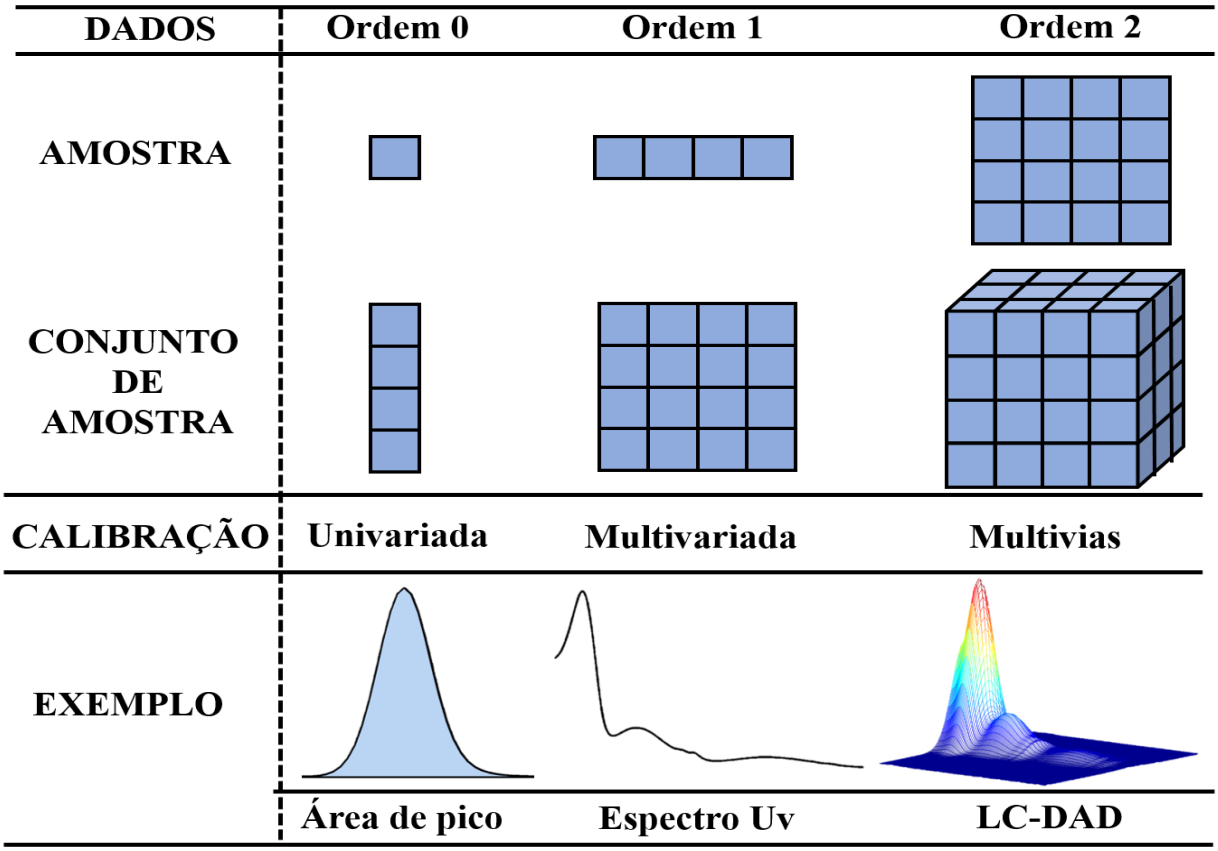

Figura 4 − Organização dos tipos de dados e exemplos para calibração univariada e multivariada de ordem 1 e 2. Adaptado de (Olivieri, 2014).

A calibração univariada é simples, demanda um esforço matemático pequeno e a predição da concentração numa amostra pode ser feita sem auxílio de um computador. No entanto, é

necessário ter 100% de seletividade no detector, ou seja, o sinal do analito medido pelo detector tem que estar puro. Caso haja algum constituinte não esperado se sobrepondo ao sinal do analito, não é possível nem identificar a presença deste interferente nem realizar quantificações exatas na calibração univariada. Na prática é difícil garantir, mesmo para cromatografia líquida, que o sinal do analito será sempre puro em uma amostra, especialmente quando se usa detectores não seletivos como o DAD. Para resolver esses problemas de interferências podem-se usar abordagens multivariadas (Barros Neto, Pimentel, Araújo, 2002; Olivieri, 2014).

A calibração multivariada é caracterizada pelo uso de mais de uma informação (variável) para relacionar com a concentração de um analito, podendo ser de ordem 1 ou superior. Dados de ordem 1 portam duas vantagens: (**1**) possibilitar a identificação do interferente; (**2**) a de quantificação exata do analito, caso se inclua o sinal do constituinte não esperado sobreposto ao conjunto de calibração. Mas esse tipo de abordagem possui alguns inconvenientes: (**1**) a necessidade de grande quantidade de amostras de modo a incluir todos os possíveis constituintes no conjunto de calibração; (**2**) caso apareça um interferente (constituinte não calibrado), o modelo é invalidado requerendo uma atualização. Para a modelagem desse tipo de dados é necessário um esforço matemático maior que para dados de ordem 0, ou seja, a matemática para esse cálculo é mais complexa e requer o uso de computador para os cálculos (Olivieri, 2014; 2015).

Para resolver futuras interferências sem recorrer à análise de um grande número de amostras ou à inclusão de mais amostras no modelo de calibração, pode-se recorrer a abordagens multivias, em que informações em mais de um modo de aquisição de dados podem ser usadas para modelar um analito. É possível distinguir analito de interferente através de diferenças no sinal registrado em pelo menos um dos modos de detecção (Olivieri, 2014).

Instrumentação do tipo LC-DAD gera dados multivariados de ordem 2, que portam algumas vantagens, como possibilitar a identificação de constituintes não esperados coeluindo com o de interesse, além de realizar quantificações exatas do analito na presença dessas interferências, característica conhecida como vantagem de segunda ordem. Essa propriedade leva a diversos benefícios: (**1**) a possibilidade de realizar a calibração com padrões dos analitos de forma similar à univariada. Isso implica que os mesmos dados usados para calibração multivariada de ordem 2 podem ser usados para a calibração univariada, o que não ocorre na calibração de ordem 1 a qual requer um grande número de amostras de modo que os interferentes sejam inclusos no conjunto de calibração. Como os interferentes são removidos matematicamente da amostra antes de realizar a quantificação, não é necessário incluí-los na calibração. Por isso não é necessário a análise de dezenas ou centenas de amostras para calibrar, como ocorre na calibração com dados de ordem 1; (**2**) não importa qual o interferente que esteja presente na amostra, sua contribuição no sinal é matematicamente removida, desde que haja diferença entre o seu sinal e o do analito em ao menos

um dos modos de aquisição, seja de eluição ou espectrais. Essa é uma condição muito mais fácil de ser atingida do que a separação total de analitos e interferentes em todas as análises com um mesmo programa de eluição (de la Peña et al., 2015).

Para a modelagem de dados de ordem 2, o esforço computacional é bem maior do que o necessário para a calibração de ordem 1, uma vez que a quantidade de dados a serem tratados é bem maior. Dependendo do tamanho dos dados, é necessário um computador com maior capacidade de processamento e memória RAM. Percebe-se que o esforço químico, seja ele de preparo da amostra ou maior complexidade instrumental, é reduzido ao passo que o esforço matemático aumenta. Assim, as vantagens alcançadas com a calibração multivias são interessantes mesmo considerando-se todo o esforço matemático necessário (de la Peña et al., 2015).

Apesar de inúmeras vantagens, os dados de segunda ordem devem ser calibrados com um algoritmo apropriado, uma vez que cada tipo de dado possui uma característica particular. Outras dificuldades na implementação de metodologias de ordem superior estão relacionadas com a necessidade de um maior tempo de cálculo, mais experiência por parte do analista, e a quantidade reduzida de instrumentação disponível para geração desses dados quando comparado com dados de ordem 1 (Escandar et al., 2014; de la Peña, Goicoechea, Escandar, Olivieri, 2015; Escandar, Goicoechea, Peña, Olivieri, 2014).

Os principais instrumentos usados para gerar dados de ordem superior são a fluorescência molecular de excitação e emissão e cromatografia acoplada a um detector de aquisição rápida como o DAD, espectrometria de massas (MS, do inglês *mass espectrometer*) e de fluorescência de emissão com o modo de excitação fixo (de la Peña, Goicoechea, Escandar, Olivieri, 2015; Escandar, Goicoechea, Peña, Olivieri, 2014; Arancibia, Damiani, Escandar, Ibañez, Olivieri, 2012).

Um espectrofluorímetro gera dados que portam a característica da trilinearidade, em que é possível remontar uma matriz de dados como sendo o produto entre três vetores linearmente independentes, cada um referente a um dos modos, adicionado de uma matriz de resíduo. Quimicamente, essa propriedade diz que o formato ou perfil do sinal em um modo independe do outro, a única dependência é com relação à intensidade do sinal. À semelhança da fluorescência molecular, independentemente do comprimento de onda de excitação, o analito emitirá o sinal da mesma forma em cada comprimento de onda de emissão registrado, alterando apenas a intensidade da fluorescência dependendo do seu perfil de excitação. A intensidade de fluorescência emitida varia de forma linear com a variação da concentração do analito (Escandar, Goicoechea, Peña, Olivieri, 2014; Olivieri A. C., 2014; Parastar, 2015). Essa propriedade garante à modelagem a característica de unicidade, ou seja, só haverá um único vetor referente a cada modo de aquisição que consiga remontar um tensor de dados com um baixo resíduo.

Os dados de cromatografia líquida não portam essa característica de trilinearidade, uma vez que a separação depende de caminhos preferenciais que o analito percorrerá na coluna e de mecanismo de partição que pode variar por diferentes fatores, sendo o principal dele a força iônica da amostra (Collins , 2006; Lanças, 2009; Fanali, Haddad, Poole, Schenmakers, Lloyd, 2013). Essa variação no tempo de retenção é uma das formas que leva a quebra de trilinearidade. Esses dados bilineares quando modelados por PARAFAC conduz a resultados errôneos de perfis recuperados, no entanto, apresentam resultados satisfatórios quando modelados com MCR-ALS. Duas variações do PARAFAC foram implementadas para resolver quebras de trilinearidade. O PARAFAC2 foi desenvolvido, especificamente, para dados de cromatografia que possuem deslocamento de tempo de retenção. Outra variação do PARAFAC é o PARALIND em que a quebra de trilinearidade ocorre devido à dependência linear em um dos modos de aquisição. A dependência linear é causada por sinais bastante sobrepostos, isso faz com que algoritmos de recuperação de perfis como PARAFAC e MCR-ALS apresentem problemas de deficiência de posto, ou seja, o algoritmo recupera um único perfil para dois ou mais analitos com espectros idênticos. O UPLS-RBL também é reportado como um algoritmo capaz de resolver bem quando há forte sobreposição espectral, embora não tenha sido testado quando a seletividade no modo espectral é zero no caso de espectros exatamente iguais (Arancibia, Damiani, Escandar, Ibañez, Olivieri, 2012; Borraccetti, Damiani, Olivieri, 2009; Bro, Harshman, Sidiropoulos, Lundy, 2009; Escandar, Goicoechea, Peña, Olivieri, 2014; Parastar, 2015); Pinto, Díaz Nieto, Zón, Fernández, Araújo, 2016).

2.4 Planejamento Taguchi

O planejamento Taguchi é usado para preparar misturas em que os níveis e os fatores são representados, respectivamente, pela concentração e analitos. Neste planejamento, as concentrações preparadas para um analito são ortogonais aos outros. É importante usar um planejamento de mistura ortogonal para não modelar nenhuma influência referente às concentrações do planejamento e, assim, afetar a qualidade da modelagem pelos algoritmos utilizados (Gaitonde et al., 2006; Maghsoodloo et al., 2004).

O planejamento Taguchi possui apenas a restrição de ortogonalidade, não sendo tão adequado para ser usado na busca de um valor ótimo em um espaço estudado como o planejamento composto central, por não ser rotacionável e não permitir o estudo de interações quadráticas entre os fatores (Barros Neto, Benício; Scarminio, Ieda Spacino; Bruns, Roy Edward, 2010; Hibbert, 2012). O planejamento Taguchi foi escolhido para gerar as misturas de teste por resultar em um número bem menor de experimentos e devido ao intuito do seu uso, neste trabalho, não ser a otimização e, sim, garantir que as concentrações dos analitos não sejam dependentes.

O planejamento Taguchi é definido de acordo com o número de níveis, fatores e experimentos. Para reduzir ao máximo o número de experimentos mantendo a ortogonalidade, apenas, alguns arranjos predefinidos são possíveis. O utilizado nesta tese foi o de quatro níveis, cinco fatores resultando em dezesseis experimentos (Gaitonde et al., 2006; Maghsoodloo et al., 2004).

2.5 PARAFAC

A análise de fatores paralelos – PARAFAC é um algoritmo de decomposição trilinear, em que três matrizes de pesos são estimadas e ajustadas de modo a minimizar a soma dos quadrados dos resíduos. Uma breve descrição do PARAFAC é feita nesta tese. A equação do PARAFAC é mostrada na **Equação 1**.

$$\boldsymbol{x_{\mathrm{ixjxk}}} = \sum_{\mathbf{1}}^{\mathbf{I}} \prod_{\mathbf{1}}^{\mathbf{3}} \boldsymbol{a} + \boldsymbol{e_{\mathrm{IxJxK}}} = \sum_{\mathbf{1}}^{\mathbf{I}} \boldsymbol{a_{\mathrm{IxD}}}\, \boldsymbol{b_{\mathrm{JxE}}} \boldsymbol{c_{\mathrm{KxF}}} + \boldsymbol{e_{\mathrm{IxJxK}}} \quad \textbf{(1)}$$

Onde x é um escalar que representa um valor pontual no tensor de dados $\boldsymbol{X}$, adquiridos em um sistema LC-DAD; *a*, *b* e *c* são escalares que representam, respectivamente, o escore de PARAFAC, um valor pontual do sinal normalizado do tempo de retenção e da absorbância; E *e* é um escalar que representa o resíduo.

Outra forma de representar o PARAFAC é usando o produto de Khatri-Rao (Sena, Trevisan, Poppi, 2005; de la Peña, Goicoechea, Escandar, Olivieri, 2015), como mostrado na **Equação 2**.

$$\boldsymbol{X} = \mathbf{A}(\mathbf{B}| \otimes |\mathbf{C})^{\boldsymbol{t}} + \boldsymbol{E} \quad \textbf{(2)}$$

Onde $\boldsymbol{X}$ é o tensor de dados adquiridos em um sistema LC-DAD; **A** é a matriz de escores para i amostras e d analitos, **B** é a matriz que contém os sinais de eluição normalizados para j tempos de retenção e e analitos; **C** é a matriz que contém os sinais de absorbância para k comprimentos de onda e f analitos. $\boldsymbol{E}$ é o tensor de resíduos.

O PARAFAC se assemelha a um modelo Tucker 3, no entanto os índices da matriz de core d, e e f são iguais, enquanto que no modelo Tucker 3 podem ser diferentes (Sena, Trevisan, Poppi, 2005; Escandar, Goicoechea, Peña, Olivieri, 2014; Bro, 1998). Essa restrição faz com que cada constituinte apresente apenas um valor de escore, um perfil de eluição e espectral, numa decomposição de um tensor de três vias gerado pelo registro de um conjunto de amostra em um sistema LC-DAD. Analisando as **Equações 1 e 2** e a **Figura 5**, pode-se dizer também que o PARAFAC é uma extensão da análise por componentes principais - PCA (do inglês, *Principal Component Analysis*) para dados de ordem superior ou uma restrição de um modelo Tucker 3, em

que os índices da matriz de core são iguais (Olivieri, Escandar, 2014; de la Peña, Goicoechea, Escandar, Olivieri, 2015).

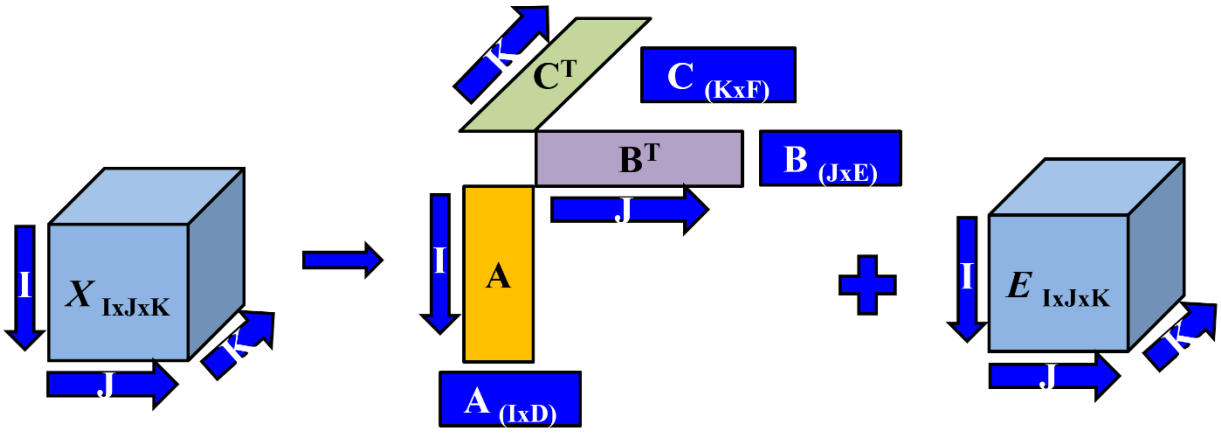

Figura 5 – Decomposição trilinear de um tensor de 3 vias ***X*** em 3 matrizes de pesos **A**, **B** e **C** adicionado de um tensor de resíduos ***E***. Adaptado de (Sena, Trevisan, Poppi, 2005).

O PARAFAC, por ser um algoritmo de decomposição trilinear, apresenta solução única. Para esta solução ser a correta e quimicamente interpretável, algumas restrições podem ser aplicadas nos escores e nos pesos dos dois modos de aquisição de dados na etapa de otimização com mínimos quadrados alternados – ALS (do inglês *Alternanting Least Squares*). As restrições mais comuns são: inicialização com os espectros puros de todos os constituintes, não negatividade, unimodalidade, normalização e ortogonalidade. Antes de otimizar os perfis com o ALS, deve-se incialmente estimar os perfis em cada um dos modos. As formas mais comuns de inicialização desse algoritmo são por decomposição trilinear direta, com valores randômicos ou com o melhor ajuste de cinco inicializações randômicas em conjunto com a decomposição trilinear direta. Os passos do ALS para o PARAFAC são (Bro, 1998; Sena, Trevisan, Poppi, 2005; de la Peña, Goicoechea, Escandar, Olivieri, 2015):

1- Definir o número de fatores;

2- Inicializar **B** e **C** (pesos);

3- Estimar **A** por um ajuste dos mínimos quadrados, usando ***X***, **B** e **C**;

4- Estimar **B**;

5- Estimar **C**;

6- Repetir as etapas de 3 a 5 até a convergência (soma dos quadrados dos resíduos $\leq 10^{-6}$ ou um número predeterminado de iterações).

Para confirmar que o algoritmo convergiu para o resultado correto, avalia-se a falta de ajuste, a variância explicada, os perfis otimizados em ambos os modos e o valor de consistência de core – CORCONDIA (do inglês, *CORe CONsistency DIAgnostic*). Esses diagnósticos serão mais bem explicados nos resultados. O CORCONDIA consiste em avaliar a matriz de core de modo a inferir se há deficiência ou inconsistência de trilinearidade.

O PARAFAC2 é uma modificação do PARAFAC que possibilita a quebra da trilinearidade em um modo. Isso ocorre por que é permitido para os pesos em um dos modos mudar de uma amostra para outra. Já no PARALIND (Perfis paralelos com dependência linear, do inglês, *PARAllel profiles with LINear Dependencies*), outra variação do PARAFAC, a matriz dos escores é multiplicada por uma matriz de dependência **H**, em que os últimos perfis estimados são forçados a serem idênticos no caso de deficiência de posto. Nesse caso, será calculada mais de uma concentração na matriz de escores para o mesmo perfil na matriz de pesos (Bro, 1998; Bahram, Bro, 2007; Bordagaray, Amigo, 2015). Nesta tese, os cálculos não foram feitos para essas variações do PARAFAC, uma vez que as métricas de desempenho analítico não estão implementadas devido à dificuldade em calcular a sensibilidade (Olivieri A. C., 2014).

2.6 UPLS-RBL

A calibração usando UPLS-RBL consiste em desdobrar a matriz de dados de cada amostra em um vetor, similar ao modelo Tucker 1, e construir um modelo PLS clássico (do inglês, *Partial Least Squares*) (de la Peña, Goicoechea, Escandar, Olivieri, 2015; Borraccetti, Damiani, Olivieri, 2009; Olivieri, Escandar, 2014). Uma breve descrição do UPLS-RBL é feita nesta tese.

Na **Figura 6** é mostrada a forma de desdobrar uma matriz de dados do tipo HPLC-DAD em um vetor para cada amostra. No esquema do *unfold* mostrado para uma amostra, a dimensão J corresponde aos tempos de eluição e a dimensão K aos comprimentos de onda registrados em cada tempo de eluição. Dessa forma, o vetor desdobrado para uma amostra consiste em vários cromatogramas concatenados lado a lado, sendo cada cromatograma referente a um comprimento de onda registrado no detector DAD. Situação similar ocorre ao desdobrar-se um conjunto de amostra como mostrado no esquema da **Figura 6**.

Após a etapa de desdobramento dos dados, é construído um modelo PLS de forma similar ao feito tradicionalmente. Inicialmente, o conjunto de dados de calibração registrado por um instrumento ($\mathbf{X}_{Cal}$) é desdobrado e constrói-se um modelo PLS da forma como tradicionalmente é feita usando validação cruzada removendo uma amostra por vez (do inglês, *leave-one-out cross validation*). Decompõe-se essa matriz desdobrada em termos de pesos (**P**) e escores (**t**), onde **P** e

t são calculados de forma iterativa objetivando maximizar a variância de $\mathbf{X_{Cal}}$ e sua covariância com o vetor de resposta **ycal**. A **Equação 3** representa um modelo PLS.

$$\mathbf{X_{Cal}} = \mathbf{P}t + \mathbf{E_{Cal}} \quad \textbf{(3)}$$

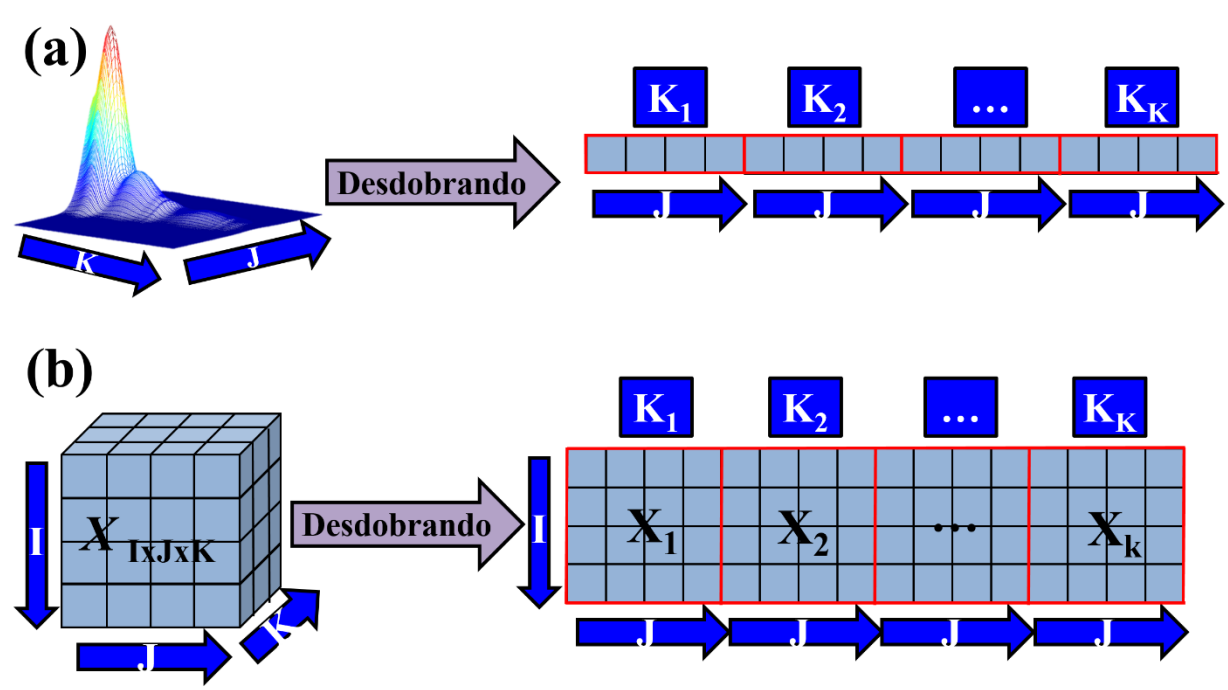

Figura 6 – Esquema para desdobrar uma amostra em um vetor (**a**) e um conjunto de amostra em vetores (**b**) para construção de um modelo UPLS. Adaptado de (de la Peña, Goicoechea, Escandar, Olivieri, 2015).

Encontrados os valores de escores **t** ao construir o modelo PLS com o número de variáveis latentes corretos, é feita uma regressão com o vetor de concentração nominal **ycal** para encontrar os vetores de regressão **v** que minimizem o erro $\mathbf{e_y}$, como mostrado na **Equação 4**. A construção do modelo é encontrar esses vetores de regressão e o valor de escore da amostra de teste é usado para prever a concentração do analito na mesma equação.

$$\mathbf{y_{Cal}} = \mathbf{tv} + \mathbf{e_y} \quad \textbf{(4)}$$

É necessário testar os escores da amostra de teste antes de prever a concentração do analito, a fim de identificar a presença de interferências. O resíduo de UPLS na amostra de teste sendo maior que o de calibração indica a presença de algum constituinte não calibrado no resíduo, e isto ocorre devido à presença de interferentes na amostra. Nesse caso, é necessária uma etapa pós-calibração de RBL para remover a contribuição dessa interferência do valor de escore da amostra de teste. Na **Figura 7** é mostrado um fluxograma de como essa etapa de RBL funciona.

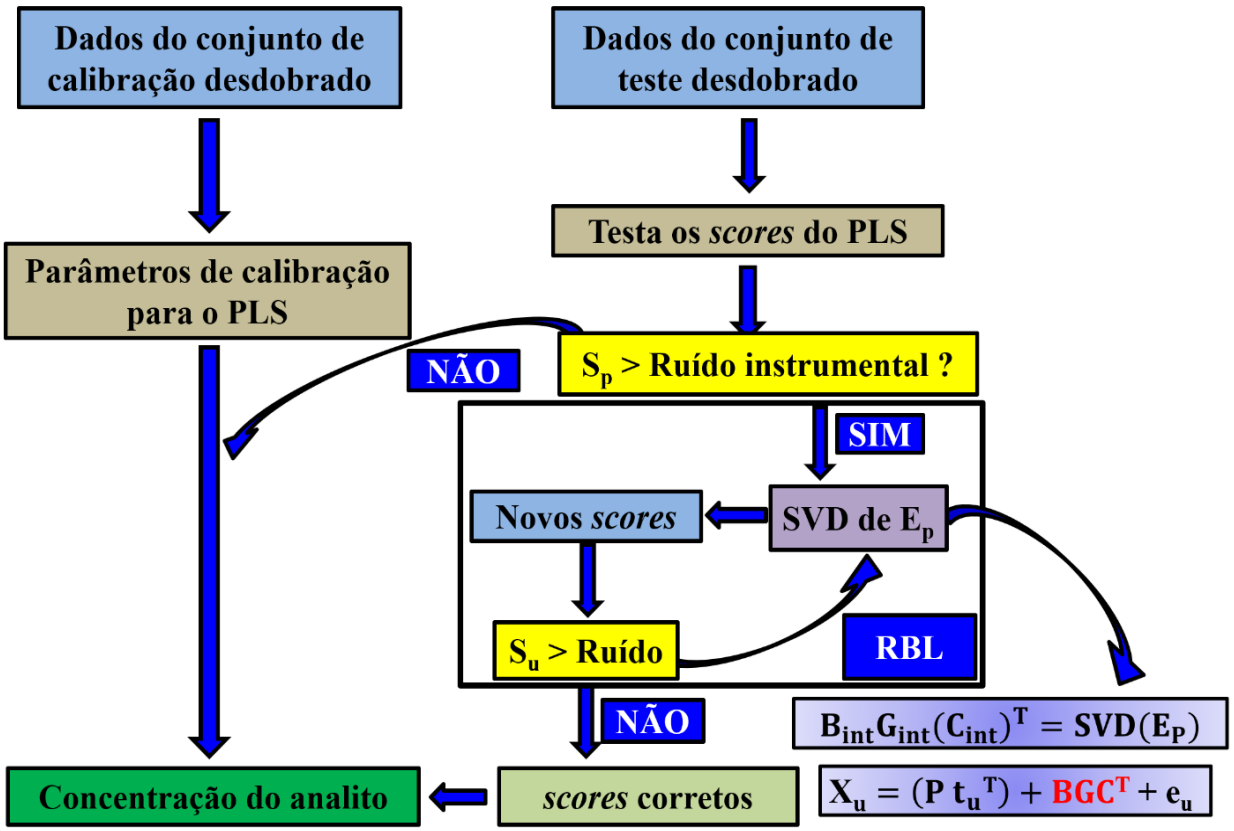

Figura 7 − Fluxograma do RBL adaptado de (de la Peña, Goicoechea, Escandar, Olivieri, 2015; Pinto, 2014). S_p é o ruído da amostra de teste, E_p é o resíduo do modelo PLS, S_u é o ruído após aplicar o RBL, P são os *loadings* de PLS e t_u os escores, e_u é o erro do RBL e os vetores B, G e C são as matrizes de escores, conectora e de *loadings* obtidos pelo SVD da matriz do resíduo de PLS.

Na etapa de RBL o vetor de resíduo de UPLS é remontado em uma matriz com as dimensões originais e então é realizado um SVD (do inglês *Single Value Decomposition*) na matriz de resíduo da amostra de teste remontada ($\mathbf{E_p}$). A quantidade de RBL necessária é avaliada com base no resíduo após o SVD, o valor correto é quando esse ruído da amostra de teste se aproxima do valor do de calibração. A equação usada na amostra de teste é a **Equação 3** de um PLS clássico acrescida da contribuição do interferente, encontrada na etapa de RBL adicionado do resíduo de RBL desdobrado (e_u). Essa contribuição do interferente é usada para recalcular os escores da amostra de teste que são, então, usados para prever a concentração do analito. Essa etapa é feita para cada amostra de forma individual, uma vez que as amostras podem apresentar interferências distintas. Na **Equação 5** está representado o modelo UPLS-RBL de uma amostra de teste.

$$\mathbf{x_u} = (\mathbf{P}\,\mathbf{t_u^T}) + \boldsymbol{BGC^T} + \mathbf{e_u} \qquad \textbf{(5)}$$

Um modelo de calibração com um analito é usado para prever a concentração de uma amostra de teste contendo dois constituintes, um analito e um interferente. Para uma correta quantificação, a equação da amostra de teste $\mathbf{x_u}$ é representada por 3 frações: (**1**) referente ao sinal do analito ($\mathbf{P}tu^T$); (**2**) ao interferente (BGC^T), esses dois sem ruído uma vez que foram modelados;

e (**3**) a contribuição final referente ao ruído aleatório presente na amostra de teste que deve ser de uma magnitude próxima ao das amostras de calibração. A **Figura 8** ilustra o modelo da **Equação 5**, onde a amostra de teste é $\mathbf{x_u}$, o modelo de calibração explica o analito em $\mathbf{P}t_u^{\mathrm{T}}$, os interferentes são recuperados em BGC^{T} como mostrado na **Figura 7** e $\mathbf{e_u}$ é o erro aleatório após a etapa de RBL. Assim, a etapa RBL consiste em calcular e remover a contribuição do interferente, representada como BGC^{T}, do ruído após a modelagem UPLS. Ao separar a contribuição do interferente do ruído os escores usados para estimar a concentração do analito são corrigidos e assim podem ser usados para realizar quantificações exatas.

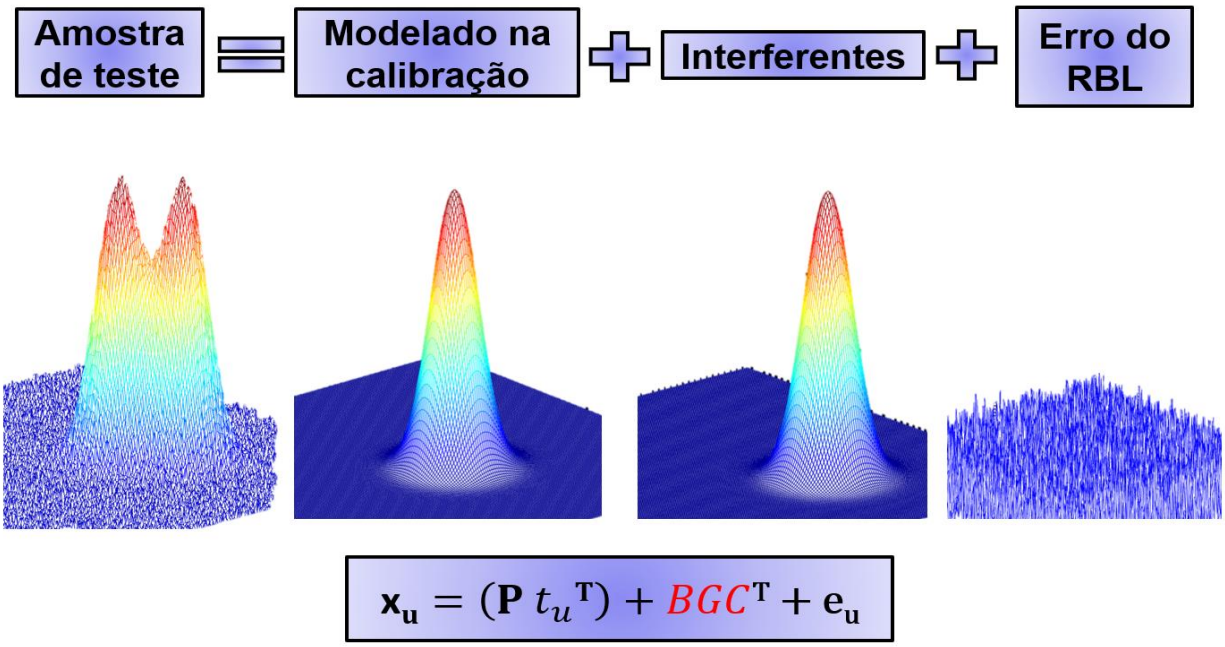

Figura 8 – Ilustração de como funciona o RBL para a amostra de teste contendo o analito e um interferente. Fonte: autor.

2.7 MCR-ALS

A resolução de curvas multivariadas - MCR-ALS (do inglês *Multivariate Curve Resolution*) é um algoritmo bilinear de deconvolução de curvas. Uma breve descrição do MCR-ALS é feita nesta tese.

Tradicionalmente, este algoritmo foi desenvolvido para explicar sistemas bilineares a exemplo da lei de Beer-Lambert. É possível ver a semelhança com a lei de Beer na representação matemática do MCR-ALS mostrado na **Equação 6**.

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathbf{T}} + \mathbf{E} \qquad \mathbf{(6)}$$

Onde **D** é a matriz de uma amostra, **C** e **S** são matrizes contendo os perfis individuais de cada constituinte presente na amostra para o modo de eluição e dos espectros, respectivamente; e **E** é a matriz de resíduo da amostra.

Desde 1993, o MCR-ALS tem ganhado destaque na resolução de curvas em dados de ordem superior a fim de resolver interferências em amostras complexas, em especial, em dados de imagens hiperespectrais, fluorescência molecular, cinética de uma reação química e, principalmente, cromatografia (Tauler, Barceló, 1993; de la Peña, Goicoechea, Escandar, Olivieri, 2015).

O MCR-ALS é, especialmente, útil para tratar dados de ordem superior advindos de separações cromatográficas, uma vez que os dados de HPLC-DAD são bilineares para uma amostra, mas não são trilineares para um conjunto de amostra (Ahmadvand, Parastar, Sereshti, Olivieri, Tauler, 2016). Isso ocorre devido ao tempo de retenção do analito poder mudar entre eluições por causa de caminhos preferenciais distintos que ocorrem dentro da coluna, durante uma separação cromatográfica. Essas mudanças no tempo de retenção do analito entre injeções, causam a quebra da trilinearidade, pois o perfil de eluição muda entre amostras, inviabilizando o uso de algoritmos de deconvolução trilinear a exemplo do PARAFAC. Com o uso de algoritmos baseados no PLS é possível atingir a vantagem de segunda ordem para pequenas variações no modo de eluição. No entanto, para maiores variações os perfis de pesos serão mais distintos levando ao analito ser recuperado como interferente na etapa de RBL, como será mostrado nos resultados desta tese.

O MCR-ALS tem sido bastante usado para dados de cromatografia líquida pela capacidade de resolver a quebra de trilinearidade no modo de eluição ao deconvoluir a matriz aumentada neste modo, uma vez que não são esperadas mudanças no formato do espectro de um analito entre injeções. Embora a maioria dos dados de cromatografia líquida sejam tratados por MCR-ALS com matriz aumentada no modo de eluição, alguns trabalhos aumentam no modo espectral para resolver quebra de trilinearidade devido à elevada sobreposição espectral entre analitos e interferentes (Culzoni, et al., 2008; Escandar, Goicoechea, Peña, Olivieri, 2014; Pinto, Díaz Nieto, Zón, Fernández, Araújo, 2016).

Na **Figura 9** é mostrada a forma como é aumentada a matriz para uma amostra com um analito registrada em um sistema LC-DAD, onde a dimensão **j** é referente ao modo de eluição, a **k** ao dos espectros e **i** às várias amostras. O tensor de respostas ***D*** é aumentado nas amostras no modo que há a quebra da trilinearidade, geralmente o de eluição para dados LC-DAD. Na **Figura 9** é observado que cada analito possui um perfil espectral distinto e que não varia entre injeções, isso torna a matriz bilinear, permitindo ao MCR-ALS recuperar corretamente o perfil de cada constituinte, mesmo com deslocamento de tempo de retenção e mudança no perfil cromatográfico.

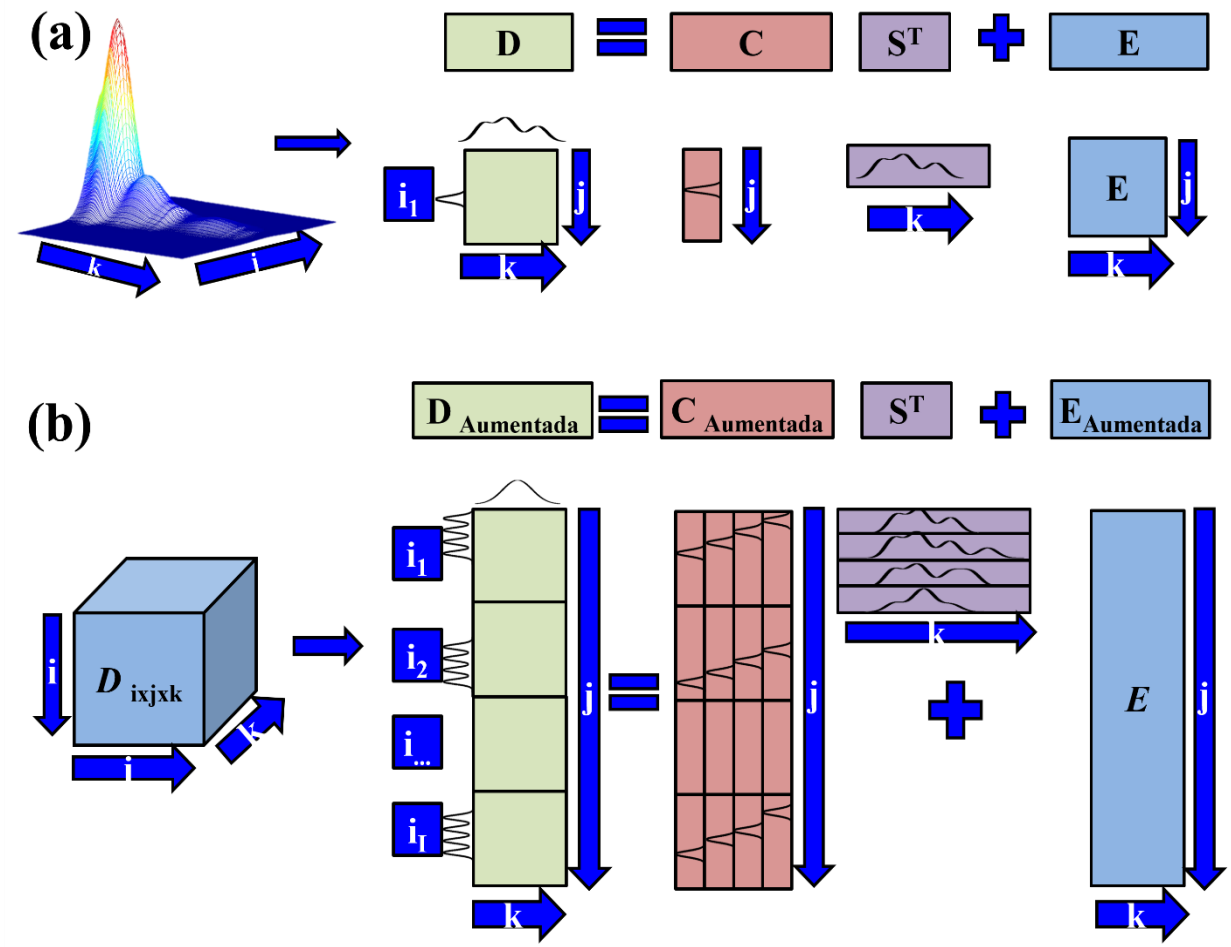

Figura 9 – Deconvolução MCR-ALS de dados LC-DAD aumentado no modo de eluição para 1 amostra contendo 1 analito (**a**) e para um conjunto de amostras contendo 4 analitos (**b**). Adaptado de (de la Peña, Goicoechea, Escandar, Olivieri, 2015; Pinto, 2014).

Similar ao PARAFAC, o MCR é um algoritmo de resolução de curvas que requer uma otimização dos perfis inicialmente estimados, o mesmo algoritmo ALS é usado para otimização de ambos os algoritmos, no entanto o ajuste pelos mínimos quadrados é feito de forma distinta, uma vez que o MCR-ALS é um algoritmo bilinear. As etapas para a deconvolução usando o MCR-ALS são:

1- Decidir o número de componentes com base numa PCA feita nos dados aumentados;

2- Inicializar **C** ou **S**;

3- Estimar **C** ou **S**, com base na inicialização do item 2, ao reconstruir a matriz aumentada **D** com um ajuste por mínimos quadrados;

4- Recalcular uma nova matriz para **C** e **S** de modo a minimizar o resíduo;

5- Repetir a etapa 4 até a convergência (variação entre interações $\leq 0{,}1\%$ ou até atingir o critério de iterações máxima).

Para a convergência a perfis corretos e quimicamente explicáveis, algumas restrições são impostas durante a otimização do ALS. As quantidades de restrições implementadas para o MCR-ALS são inúmeras, diferentemente do PARAFAC que possui menos restrições. As restrições mais comuns são: (**1**) inicialização com os espectros puros de todos os constituintes; (**2**) não-negatividade; (**3**) unimodalidade; (**4**) organização de perfis por ordem de variável pura; (**5**) normalização; (**6**) correspondência entre as espécies; (**7**) multilinearidade; (**8**) clausura; (**9**) igualdade; (**10**) cinética (*hard modeling*); e (**11**) correlação (de la Peña, Goicoechea, Escandar, Olivieri, 2015; Jaumot, de Juan, Tauler, 2015).

Devido ao MCR-ALS ser um algoritmo de deconvolução de dados bilineares seu uso pode levar a problemas de ambiguidade durante a modelagem. Isso ocorre, pois, matematicamente, podem existir várias matrizes **C** que ao multiplicar por uma matriz $\mathbf{S}^T$ resulte na matriz aumentada **D** com um resíduo **E** baixo e homocedástico, diferentemente do PARAFAC, que possui solução única devido à deconvolução ser feita em uma multiplicação de três matrizes.

São três as ambiguidades que o MCR-ALS pode apresentar: a ambiguidade de posição em que o perfil de concentração de um analito é relacionado com o perfil espectral de outro constituinte. Isso é facilmente resolvido ao organizar os perfis por ordem de variável pura. Outra restrição é a de intensidade, onde um perfil de um analito pode ser muito mais intenso que o de outro constituinte, tanto no modo de concentração quanto no espectral. Ao normalizar um dos modos isso é resolvido, em geral, é normalizado o modo dos espectros. A terceira ambiguidade é a ambiguidade rotacional, esta é a mais complicada de ser resolvida, é caracterizada por recuperar um perfil que não corresponde ao do analito, seja no modo de eluição ou no modo dos espectros. Para resolver esse tipo de ambiguidade deve-se aplicar restrições. As mais efetivas para dados LC-DAD seriam a inicialização com os espectros puros de todos os constituintes, não negatividade, normalização, organização de perfis por ordem de variável pura, unimodalidade e correspondência entre as espécies. Essas restrições podem ser facilmente aplicadas nas interfaces gráficas disponíveis na literatura (Jaumot, de Juan, Tauler, 2015; Olivieri, Wu, Yu, 2009).

Apesar de ser mais comum usar algoritmos de seleção de variáveis puras (de la Peña, Goicoechea, Escandar, Olivieri, 2015) para inicializar o ALS com esses perfis puros dos constituintes presentes na amostra analisada, é possível inicializar o ALS com o perfil dos analitos e interferentes, que além de auxiliar na inicialização acelerando a convergência, é importante para comparar com o resultado final otimizado. O grau de semelhança pode ser calculado pelo cosseno entre as normas dos vetores correspondentes aos perfis recuperados. Nos resultados, mais detalhes serão fornecidos sobre esse cálculo (Culzoni, et al., 2008). Uma vez que os perfis do analito, tanto de concentração quanto espectral, são semelhantes aos otimizados não há ambiguidade rotacional.

Caso não se tenha os perfis dos analitos a avaliação da ambiguidade rotacional pode ser feita com os cálculos de bandas de confiança. Esse cálculo consiste em multiplicar as matrizes **C** e $\mathbf{S}^T$ otimizadas por uma matriz inversível (**T**), geralmente uma matriz identidade, como mostrado na **Equação 7**.

$$\mathbf{D} = \mathbf{C}\mathbf{T}\mathbf{T}^{-1}\mathbf{S}^T + \mathbf{E} \quad (7)$$

Multiplicar **C** por **T** e $\mathbf{S}^T$ por $\mathbf{T}^{-1}$, resulta em novas matrizes $\mathbf{C}_{new}$ e $\mathbf{S}_{new}^T$ com perfis distintos, no entanto, para um mesmo valor de **T**, a relação entre a norma do sinal do analito normalizado sobre o sinal de todos os componentes do sistema (norma de Frobenious, mostrado na **Equação 8**) não deve diferir (Jaumot & Tauler, 2010).

$$f_n = \frac{\|\mathbf{C}_n\mathbf{S}_n^T\|}{\|\mathbf{C}\mathbf{S}^T\|} \quad (8)$$

onde f_n é um escalar que representa a norma de Frobenious (do inglês, *Frobenious norm*), $\mathbf{C}_n$ e $\mathbf{S}_n^T$ são os perfis normalizados de um analito, **C** e $\mathbf{S}^T$ são referentes ao sinal completo de todos os constituintes presentes na mistura.

Os cálculos para f_n são feitos para novos $\mathbf{C}_{new}$ e $\mathbf{S}_{new}^T$ao multiplicar os antigos por distintos **T** e $\mathbf{T}^{-1}$, isso é feito para vários valores de **T** até encontrar um valor de f_{max} e f_{min}. A diferença entre os valores de f_{max} e f_{min} deve ser bem próxima de zero para todos os analitos de modo a não haver ambiguidade de rotação (Jaumot, Tauler, 2010; Março, Valderrama, Alexandrino, Poppi, Tauler, 2014).

Para usar o MCR-ALS a matriz deve ser bilinear ao aumentar no modo que cause a quebra da trilinearidade. Uma vez que a maioria dos sistemas químicos são bilineares, o MCR-ALS é aplicável para tratamento de dados proveniente de diversos instrumentos (de la Peña, Goicoechea, Escandar, Olivieri, 2015).

Os conjuntos de dados LC-DAD são bilineares, mas pode haver situações em que além da quebra da trilinearidade no modo da eluição, os espectros registrados sejam muito correlacionados, gerando deficiência de posto. Um caso que isso pode ocorrer é separação de constituintes quirais, em que há variações no tempo de retenção e os espectros são idênticos (Grisales, Arancibia, Castells, Olivieri, 2012). Neste trabalho, apesar dos analitos não serem quirais, os espectros são extremamente sobrepostos, e em alguns casos idênticos, devido à derivatização das aminas que podem ser analitos ou interferentes levando à deficiência de posto (Pinto, Díaz Nieto, Zón, Fernández, Araújo, 2016). Em alguns casos é possível separar os analitos em faixas para modelar cada constituinte idêntico individualmente (Pérez Escandar, 2014). No estudo de caso desta tese, isto é impossível de ser feito, uma vez que na amostra podem existir distintas aminas que ao serem

derivatizadas resultem em constituintes cujos espectros são iguais, e que podem coeluir no programa de eluição usado.

Nesse caso, ao aumentar a matriz no modo de eluição o MCR-ALS acarretará em deficiência de posto: caso a matriz seja aumentada no modo espectral para resolver a elevada correlação espectral, as diferenças no tempo de eluição serão entendidas como mais de um analito similar ao que ocorre com o PARAFAC. Dessa forma, a matriz não será bilinear nem se for aumentada no modo espectral nem no modo de eluição. Então, não é possível atingir a vantagem de segunda ordem apenas com a modelagem com o MCR-ALS seja a matriz aumentada no modo de eluição ou no modo espectral devido ao problema de desalinhamento de pico e de deficiência de posto.

2.8 Deficiência de posto

Deficiência de posto ou dependência linear são denominações usadas para se referir quando um ou mais modos de aquisição instrumental possuem dois ou mais constituintes com sinais muito correlacionados (Amrhein, Srinivasan, Bonvin, Schumacher, 1996; Bordagaray, Amigo, 2015). A consequência da deficiência de posto é usar menos componentes do que o necessário para explicar os constituintes do sistema, ou seja, o posto matemático difere do químico.

Em um dado do tipo LC-DAD com dois constituintes com espectros idênticos, será necessário apenas um único componente para recuperar o sinal no modo espectral, enquanto no modo cromatográfico haverá dois sinais referentes a cada constituinte. Quando esses dados são tratados com o MCR-ALS aumentado no modo tempo sem nenhuma restrição, o sinal recuperado no modo de eluição será a soma dos dois picos cromatográficos. Caso se use a restrição de unimodalidade, será recuperado o perfil referente a um dos dois constituintes, e o outro se somará ao resíduo. Assim, a variância explicada será muito baixa e a falta de ajuste alta, além de a sensibilidade analítica reduzir bastante devido ao sinal não aleatório na matriz de resíduos. Há também uma possibilidade de ora recuperar o perfil referente ao primeiro pico, ora ao segundo, o que levará a erros na quantificação.

PARAFAC, PARALIND, MCR-ALS e UPLS podem resolver esses problemas, embora cada um possua limitações. O PARAFAC recupera bem o perfil, caso os picos do analito estejam alinhados. No entanto, a seletividade é prejudicada devido à elevada correlação entre os espectros. Em casos mais extremos nos quais em um modo há dois perfis idênticos e no outro modo o sinal referente a esses dois constituintes esteja muito sobreposto, apenas o PARALIND recuperaria corretamente o sinal. Contudo, devido à forma em que foi implementado, não há na literatura forma de calcular a sensibilidade para um modelo construído com o PARALIND, o que dificulta a sua comparação com outros modelos (Olivieri, 2014; Olivieri, Escandar, 2014; de la Peña,

Goicoechea, Escandar, Olivieri, 2015; Bordagaray, Amigo, 2015; Bro, Harshman, Sidiropoulos, Lundy, 2009).

O UPLS-RBL pode resolver deficiência de posto originada pela sobreposição espectral, no entanto, para dados de HPLC-DAD com deslocamento de tempo de retenção mais pronunciado, esse algoritmo falha em recuperar o perfil do interferente corretamente e, às vezes, confunde o sinal do analito com o do interferente, como será mostrado em detalhes nos resultados desta tese (Borraccetti, Damiani, Olivieri, 2009; de la Peña, Goicoechea, Escandar, Olivieri, 2015; Escandar, Goicoechea, Peña, Olivieri, 2014).

O MCR-ALS resolve a deficiência de posto ao aumentar o tensor LC-DAD para várias amostras no modo espectral, como mostrado na **Figura 10**. Ao aumentar os dados no modo espectral, deve-se tomar cuidado para que os picos referentes aos analitos estejam bem alinhados, ou seja, não pode haver deslocamentos no tempo de retenção, uma vez que este modo não é mais o modo aumentado. Caso isso não ocorra, será preciso recuperar mais componentes do que o necessário para explicar os dados, como ocorre no PARAFAC (de la Peña, Goicoechea, Escandar, Olivieri, 2015).

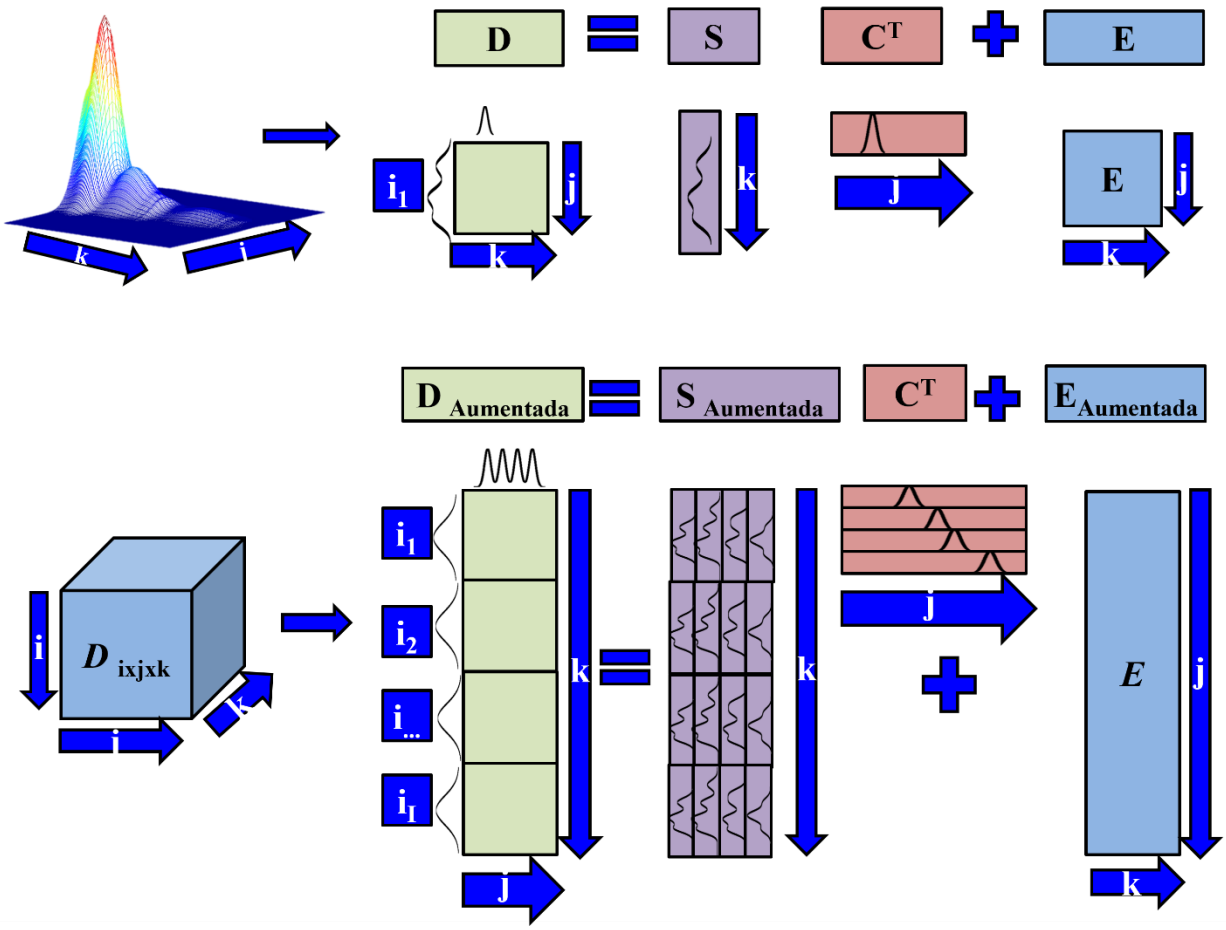

Figura 10 − Decomposição MCR-ALS de dados LC-DAD aumentado no modo espectral para 1 amostra contendo 1 analito e para um conjunto de amostras contendo 4 analitos. Adaptado de (Pinto, 2014).

2.9 Algoritmos de alinhamento de pico

Sabe-se que o deslocamento de tempo de retenção é intrínseco às medidas cromatográficas. Assim, diversos algoritmos foram desenvolvidos para alinhamento de pico em cromatografia líquida, sendo estes, em quase sua totalidade, voltados para dados de ordem 1, ou seja, para um cromatograma registrado em um único canal do detector. Os algoritmos mais usados para dados de primeira ordem são o COW (do inglês, *Correlation Optimized Warping*), e o *coshift* (do inglês, *COrrelation SHIFTing*) (Tomasi, Savorani, Engelsen, 2011; van den Berg, Tomasi, Viereck, 2005; Vest Nielsen, Carstensen, Smedsgaard, 1998; Tomasi, van den Berg, Andersson, 2004).

O COW baseia-se na função de compressão/expansão que assume que o deslocamento tem correlação com o alargamento do pico. Isso era muito comum ocorrer na cromatografia antigamente, uma vez que o deslocamento pode acontecer por mudanças na adsorção na coluna devido à falta de reprodutibilidade no gradiente ou na mistura de solventes. Como o controle do gradiente e da mistura dos solventes são mais refinados nos equipamentos mais modernos e as partículas quimicamente ligadas na coluna são mais reprodutíveis e sem vazios em seu interior, esta situação é mais difícil de ocorrer. Para um pico gaussiano, é mais comum ocorrer deslocamento sem necessariamente ter um alargamento do pico para eluições nas mesmas condições, uma vez que esses deslocamentos podem estar bem correlacionados com mudança da força iônica do meio no qual o analito se encontra (Tomasi, Savorani, Engelsen, 2011; Fanali, Haddad, Poole, Schenmakers, Lloyd, 2013; Collins , 2006).

Quando não houver alargamento nos picos, podem-se usar algoritmos de alinhamento baseados na inserção/supressão, em que se assume que a largura de pico se mantém invariável com o deslocamento de tempo de retenção. Tanto o *coshift* (do inglês *COrrelation SHIFTing*) quanto o i*coshift* (do inglês *Interval COrrelation SHIFTing*) usam essa função e fazem a correção do deslocamento usando a transformada de Fourier acelerada. O i*coshift* maximiza a correlação cruzada com o segmento do cromatograma alvo e o *coshift*, do cromatograma completo (Tomasi, Savorani, Engelsen, 2011).

É importante ressaltar que o i*coshift* e o *coshift* para alinhar os dados LC-DAD de segunda ordem, resultam em menor esforço computacional que o COW e não requerem modificações no algoritmo para alinhar dados de ordem superior, como será mostrado melhor nos resultados desta tese. Uma vez que tanto o i*coshift* quanto o *coshift* foram desenvolvidos para alinhar dados de uma dimensão, alguns cuidados devem considerados, como será mostrado neste trabalho, para o seu uso ao alinhar dados de segunda ordem.

Algumas estratégias para alinhar dados de ordem superior foram propostas, no entanto, nenhuma é eficiente para dados de cromatografia líquida com espectros muito sobrepostos

(Bortolato, Arancibia, Escandar, Olivieri, 2010; Parastar, Akvan, 2014). Ambas se baseiam em alinhar os perfis recuperados, uma usando o PARAFAC e outra, o MCR-ALS, com o COW implementado para dados de ordem 1. Apesar de terem se mostrado eficientes em estudo, essas abordagens falham quando se deparam com constituintes cujos espectros são muito correlacionados, ou seja, quando os dados levam à deficiência de posto ao serem modelados.

Uma vez resolvido esse deslocamento de tempo de retenção, a estrutura trilinear dos dados é restaurada e mais algoritmos podem ser usados para tratar esses dados, inclusive alcançando a propriedade de unicidade. Além disso, será mostrado no presente trabalho que alinhar os cromatogramas é essencial para alcançar a vantagem de segunda ordem para onde o segmento cromatográfico modelado possua mais de um constituinte com sinal muito correlacionado.

2.10 Métricas de desempenho analítico

As métricas de desempenho analítico, conhecidas como figuras de mérito (do inglês *figure of merit*), são valores numéricos que expressam a qualidade de um modelo construído. Essas métricas ajudam a comparar e a caracterizar dois instrumentos ou dois modelos construídos para os mesmos dados com algoritmos diferentes (Olivieri A. C., 2014). A maioria dessas métricas pode ser calculada em função da sensibilidade e do ruído instrumental, a dificuldade é encontrar uma forma de determinar isso corretamente. As principais métricas analíticas são: (1) sensibilidade; (2) sensibilidade analítica; (3) limite de detecção; (4) limite de quantificação; (5) erro relativo de predição. Essas métricas de desempenho são calculadas de forma similar levando em consideração se há interferência no sinal do analito e o tipo de dado e algoritmo usado para a modelagem. Como dito na **seção 2.3**, a aquisição dos dados deve ser feita de acordo com a complexidade da amostra e da instrumentação disponível para análise.

A forma mais simples e mais bem estabelecida de calcular essas métricas é a calibração univariada, na qual o modelo construído obedece a **Equação 9**. Para usar esse tipo de modelagem, o sistema de detecção deve ser seletivo o suficiente para garantir que o sinal medido seja obtido com 100% de seletividade. Na prática, ao analisar uma amostra complexa, isso só ocorre quando há uma separação física (analise cromatográfica) ou química (complexação com um reagente mascarante) e o sistema de detecção seja seletivo o suficiente para detectar apenas o analito (Olivieri, 2015).

$$\boldsymbol{y} = \boldsymbol{a}\boldsymbol{x} + \boldsymbol{b} + e \qquad \textbf{(9)}$$

Onde $\boldsymbol{y}$ é o vetor que contem os sinais medido em um instrumento, $\boldsymbol{x}$ é o vetor contendo o parâmetro calibrado (concentração do analito), a e b são os coeficientes angular e linear obtidos

pelo ajuste dos dados experimentais a um modelo linear por mínimos quadrados e *e* é o valor do resíduo.

Em um modelo univariado a sensibilidade é o coeficiente angular (a) obtido na **Equação 9**, e o ruído instrumental pode ser medido como sendo o desvio padrão entre um determinado número de repetições de uma amostra com menor concentração do analito, ou o desvio padrão do sinal, em que o analito seria registrado com o máximo de intensidade, em uma amostra branco (Eurachem, 2014). A sensibilidade analítica é a relação entre a sensibilidade e o ruído instrumental.

Os limites de detecção (LOD, do inglês *limit of detection*) e quantificação (LOQ, do inglês *limit of quantification*) são calculados respectivamente como sendo 3,3 e 10 vezes o inverso da sensibilidade analítica, onde esses valores estão relacionados aos erros do tipo I e II (Bauza, Ibañez, Tauler, Olivieri, 2012; Ortiz, Sarabia, Sánchez, 2010). Caso leve-se em consideração a incerteza no preparo dos padrões e soluções, o valor de LOD e LOQ deve ser calculado levando em consideração o valor de influência (do inglês *leverage*) na amostra branco ou de menor concentração (Olivieri, 2014).

A forma com que a sensibilidade e o ruído são calculados depende se a calibração é univariada, multivariada ou multivias. A univariada já foi mostrada anteriormente e já é bem estabelecida. Já a abordagem usada para quantificar a sensibilidade pelas metodologias multivariadas e multivias ainda estão em desenvolvimento e depende do algoritmo utilizado. E no caso das metodologias multivias, além do algoritmo utilizado, também depende da amostra, em que podem surgir constituintes em distintos perfis e quantidades (Olivieri, 2014; Allegrini Olivieri, 2014; Fragoso, Allegrini, Olivieri, 2016).

Inicialmente o conceito de sinal analítico líquido (NAS, do inglês *Net Analyte Signal*) foi usado para a calibração de primeira ordem usando PLS. Entretanto, recentemente Olivieri propôs uma forma de calcular a sensibilidade baseada na propagação da incerteza na calibração e desenvolveu uma equação geral usada para calcular a sensibilidade para calibração de ordem zero até a multivias (Olivieri, 2014).

Por último, há a métrica chamada seletividade, para a calibração univariada ela deve ser igual a 1, seletividade total. Para calibração de ordem 1 não há como calcular, uma vez que depende da intensidade do sinal do constituinte com sinal sobreposto ao analito. Na calibração multivias pode ser calculada apenas para os algoritmos que recuperam perfis, como o PARAFAC e o MCR-ALS.

De forma geral, a seletividade é a razão entre a sensibilidade do analito na mistura pela sensibilidade do analito puro (Olivieri, 2014). Nos algoritmos de ordem superior, a sensibilidade da mistura é calculada para as amostras e a do analito puro no conjunto de calibração. Como a

sensibilidade do analito na mistura é menor que ele puro, quanto maior a seletividade, maior será a sensibilidade, resultando em uma melhor detectabilidade do analito. Assim, no caso dos algoritmos que recuperam perfis, o ideal é recuperar a menor quantidade de perfis de modo a se obter uma maior seletividade, ou seja, as métricas de desempenho serão melhores se a modelagem de cada analito for feita individualmente e os interferentes forem recuperados com o menor número de perfis. Neste trabalho, os dados foram modelados por região e os interferentes recuperados como sendo um único perfil resultante da soma de todos. De acordo com o exposto, espera-se que, com isso, os parâmetros de desempenho analítico sejam mais condizentes com os dados experimentais tanto nas amostras de teste quanto nas de peixe.

Capítulo 3

METODOLOGIA

3 METODOLOGIA

3.1 Reagentes

Neste trabalho, foram utilizados padrões hidroclorados de triptamina (TRI), putrescina (PUT), 2-feniletilamina (FEN), cadaverina (CAD) e histamina (HIS), todos de grau cromatográfico adquiridos da Sigma Aldrich. Como reagente derivatizante, foi usado o cloreto de dansil (DNS) grau cromatográfico adquirido da Sigma Aldrich. Acetonitrila de grau cromatográfico foi adquirido da J. T. Baker. Hidróxido de amônia, bicarbonato de sódio e ácido clorídrico de grau analítico foram adquiridos da Vetec. Ácido perclórico, acetona e hidróxido de sódio de grau analítico foram adquiridos da Dinâmica. Em todos os experimentos foram utilizados água destilada e recém-purificada em um deionizador milli-Q da Millipore com filtro de 0,22 µm.

3.2 Eluição e parâmetros cromatográficos

As eluições cromatográficas foram realizadas utilizando um cromatógrafo a líquido de alta eficiência modelo Ultimate 3000 da Dionex, consistindo de uma bomba quaternária, um injetor manual com um loop fixo de 20 µL, um forno de coluna e um detector Ultravioleta-visível de arranjo de diodos. Foi utilizada uma coluna C18 de 150 mm x 4,6 mm, com partículas de sílica de 5 µm e tamanho de poro de 120 Å, modelo Acclaim^{TM} 120 da Dionex.

A eluição ocorreu em modo isocrático com uma mistura binária de 73% de acetonitrila (solvente B) e 27% de água (solvente A) numa vazão de 2 mL min^{-1} e com temperatura controlada em 30℃. A eluição total dos constituintes ocorreu em quatro minutos. Os dados foram coletados usando a versão 6.80 do *software* Chromeleon com uma frequência de aquisição dos espectros na região do UV de 2 Hz.

3.3 Derivatização

As derivatizações das aminas com o cloreto de dansil foram realizadas de acordo com o método desenvolvido por Chen e colaboradores (Chen, et al., 2010). Os passos desse procedimento são mostrados na **Figura 11**.

A 1 mL de uma solução contendo as aminas, foram adicionados 0,2 mL de hidróxido de sódio 2 mol L^{-1}, 0,3 mL de uma solução saturada de hidrogenocarbonato de sódio, e 2 mL de uma solução de 1% (m v^{-1}) de cloreto de dansil dissolvido em acetona. Esta solução foi

homogeneizada em vortex e deixada reagir a 40℃ por 45 minutos em capela de exaustão de gases. Após o tempo da reação, com a solução ainda quente, adicionou-se 0,1mL de hidróxido de amônia concentrado e deixou-se em repouso por trinta minutos. Ao final da reação aferiu-se com acetonitrila até um volume final de 5 mL em balão volumétrico. Esta solução foi centrifugada a 4.000g por dez minutos e o sobrenadante filtrado em um filtro seringa de nylon de 4mm com poro de 0,22 µm e analisado no sistema cromatográfico.

1 - 1 mL da solução contendo as aminas biogênicas
2 - 0,2 mL de NaOH 2 mol L^{-1}
3 - 0,3 mL de $NaHCO_3$ saturado
4 - 2 mL da solução de Dansil 1%
5 - Reage a 40℃ por 45 minutos
6 - 0,1 mL de NH_4OH e aguarda 30 minutos em repouso
7 - Completar para 5 mL com acetonitrila
8 - Centrifugar a 4.000g por 10 minutos
9 - Filtrar o sobrenadante com filtro seringa de nylon de 0,22 µm

Figura 11 – Procedimento para derivatização das aminas biogênicas nos padrões e amostras. Fonte: autor.

3.4 Conjunto de calibração e de teste

Soluções estoques dos padrões individuais de 1.000 mg L^{-1} foram preparados ao medir 100 mg do padrão e solubilizados em 100 mL de uma solução de ácido clorídrico 0,1 mol L^{-1}. Os estoques dos padrões foram armazenados a 4℃ até o uso.

Os conjuntos de calibração foram preparados individualmente para cada amina em triplicata autêntica e em 12 níveis de concentração, (1) 0,05; (2) 0,10; (3) 0,30; (4) 0,50; (5) 0,75; (6) 1,00; (7) 3,00; (8) 5,00; (9) 7,50; (10) 10,00; (11) 15,00; (12) 20,00 mg L^{-1}, por diluições a partir da solução estoque. Essa faixa de trabalho foi previamente definida com base em um estudo de linearidade realizado para cada AB. As injeções no HPLC-DAD foram feitas em três dias, sendo cada dia para uma das triplicatas autênticas preparadas para cada analito derivatizado. Assim incluíram-se no erro puro o erro de injeção, de preparo das amostras e as variações do dia de análise. Os conjuntos de calibração foram preparados individualmente devido à forma em que foi realizado o alinhamento nesta tese.

O conjunto de teste foi preparado seguindo o planejamento Taguchi mostrado na **Tabela 1**. Este planejamento consiste de dezesseis amostras contendo uma mistura das cinco aminas em quatro níveis de concentração, (1) 0,4; (2) 4,0; (3) 8,0 e (4) 16,0 mg L^{-1} (Yadav, Dasgupta, Singh, Varshney, Singh, 2015). Após preparar as soluções dos padrões, tanto para o conjunto de calibração quanto para o de teste, ambas foram submetidas à derivatização de acordo com o procedimento detalhado na **seção 3.3**. As amostras de teste foram submetidas, também, à extração da forma como foi feita para as amostras de peixe, como descrito na **seção 3.5**.

Tabela 1 − Conjunto de teste das aminas biogênicas em quarto níveis de concentração (mg L^{-1}).

Exp.	TRI	FEN	PUT	CAD	HIS
1	0,40	0,40	0,40	0,40	0,40
2	4,00	4,00	4,00	4,00	0,40
3	8,00	8,00	8,00	8,00	0,40
4	16,00	16,00	16,00	16,00	0,40
5	16,00	8,00	0,40	4,00	4,00
6	8,00	16,00	4,00	0,40	4,00
7	4,00	0,40	8,00	16,00	4,00
8	0,40	4,00	16,00	8,00	4,00
9	4,00	16,00	0,40	8,00	8,00
10	0,40	8,00	4,00	16,00	8,00
11	16,00	4,00	8,00	0,40	8,00
12	8,00	0,40	16,00	4,00	8,00
13	8,00	4,00	0,40	16,00	16,00
14	16,00	0,40	4,00	8,00	16,00
15	0,40	16,00	8,00	4,00	16,00
16	4,00	8,00	16,00	0,40	16,00

3.5 Amostras, amostragem e preparo da amostra

Quatro amostras de peixe fresco, recém-pescado, com aproximadamente 300g cada, foram adquiridas no mercado local de João Pessoa, Paraíba, Brasil. As amostras correspondem às espécies de peixes de água salgada típica da região; (1) Dourado (*Salminus maxillosus*), (2) Cavala (*Scomber scombrus*), (3) Arabaiana ou Olho-de-boi (*Seriola dumerili*) e (4) Cioba (*Lutjanus analis*), mostrados na **Figura 12 (a)**.

Cada uma das quatro amostras de peixes foi armazenada em saco do tipo zip, a 4℃ até a análise. Inicialmente, coletaram-se porções de cinco partes espaçadas da posta do peixe, conforme o ilustrado na **Figura 12 (b)**. As porções coletadas do mesmo peixe, ainda geladas, foram trituradas em um processador de alimentos até a completa homogeneização, com pulsações periódicas de modo a não aquecer a amostra por atrito.

Mediu-se uma porção de 5 gramas do peixe triturado em um béquer de 50 mL. A extração ocorreu com 30 mL de ácido perclórico 0,6 mol L^{-1} por quinze minutos com o auxílio de um agitador magnético. O extrato foi filtrado com um papel de filtro Nº. 2 (Whatman, Maidstone, England). Em seguida, o filtrado foi levado a um balão volumétrico e aferido a 50 mL com o ácido perclórico 0,6 mol L^{-1}, conforme descrito por Chen e colaboradores (Chen, et al., 2010). Uma alíquota de 1 mL desse extrato foi derivatizada, de acordo com o método descrito na **seção 3.3** (**Figura 11**) e analisada no sistema cromatográfico.

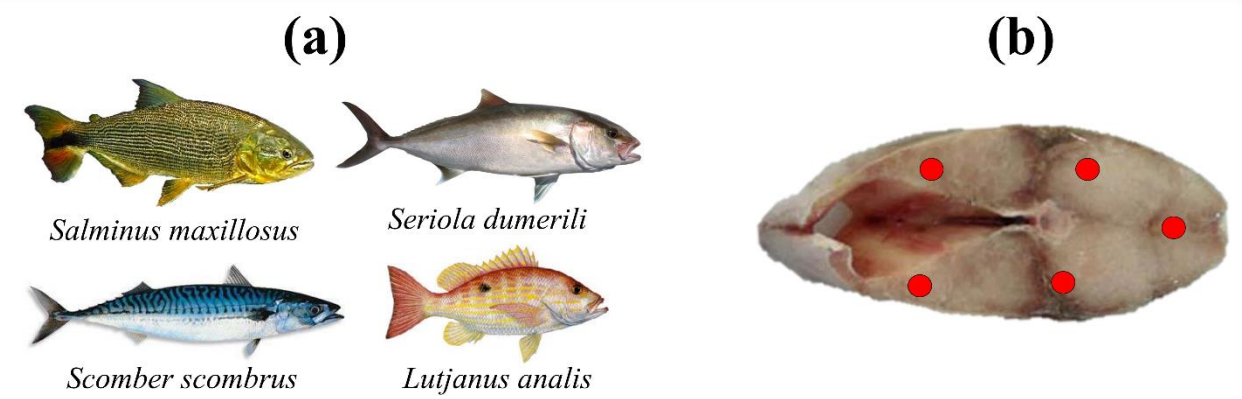

Figura 12 − Peixes analisados (**a**) e região de amostragem na posta do peixe (**b**). Fonte: autor.

3.6 Teste de recuperação e estudo de degradação

No primeiro dia de análise, as amostras de peixe foram fortificadas para avaliar a recuperação da metodologia proposta e a presença de efeito de matriz. Foram feitas cinco análises da mesma amostra no primeiro dia, (1) uma amostra sem fortificação e em quatro níveis de fortificação com (2) 4,00, (3) 40,00, (4) 80,00 e (5) 160,00 mg kg^{-1} do padrão diretamente na amostra triturada, antes de realizar a extração. Essa concentração fortificada corresponde a 10 vezes os níveis de concentração das misturas de teste. Assim, ao final da extração a concentração da alíquota da amostra fortificada foi a mesma que a do conjunto de teste mais a concentração inicial da amostra. Dessa forma, o efeito de matriz pode ser calculado comparando o coeficiente

angular da curva linear ajustada aos dados de calibração, teste e das amostras de peixe (Moreno-González, García-Campaña, 2017).

As ABs são indicadores de degradação de alimentos. Então, para avaliar a cinética de degradação, armazenaram-se as quatro amostras de peixe a 4 ℃ durante treze dias. Nesse período, a cada dois dias realizaram-se as medidas da concentração das ABs.

3.7 Procedimento quimiométrico e software

3.7.1. PARAFAC

O PARAFAC é um algoritmo de resolução de curvas. Dessa forma, o primeiro passo é definir o número de componentes que serão recuperados pelo algoritmo. Esse é um fator crucial, tendo em vista que informar o algoritmo que há uma quantidade de analitos inferior ou superior ao que está na matriz de dados irá resultar em perfis recuperados que não correspondem ao seu sinal individual. O número de componentes foi definido com base na variância explicada, o ajuste do modelo e o valor de CORCONDIA. Quanto mais próximo de 100% melhor é a consistência do modelo PARAFAC e valores abaixo de 50% indicam inconsistência com a trilinearidade dos dados (Bro, 1998; Sena, Trevisan, Poppi, 2005; de la Peña, Goicoechea, Escandar, Olivieri, 2015). A interpretação dos perfis recuperados, também, é de grande ajuda nessa escolha.

Os perfis são estimados usando mínimos quadrados alternados. Para guiar o algoritmo na convergência a perfis quimicamente explicáveis ao longo das iterações foram aplicadas as restrições (**1**) não negatividade no modo dos espectros, de eluição e escore; (**2**) unimodalidade no modo de eluição; (**3**) correspondência entre as espécies. Como inicialização, foi testado o melhor ajuste inicial baseado em decomposição trilinear direta, SVD e três tentativas com valores randômicos. A concentração da amostra é calculada ao estimar os perfis das amostras de teste juntamente com as de calibração e, em seguida, interpolar os escores das amostras de teste na curva construída com os escores das amostras de calibração, numa etapa conhecida como calibração pseudo univariada (de la Peña, Goicoechea, Escandar, Olivieri, 2015).

Nesta tese, serão mostrados os resultados de todos os algoritmos apenas para a modelagem de cada analito dividido em faixas. Essa abordagem foi usada, pois ao se recuperar menos componentes na mistura a sensibilidade analítica aumenta, como discutido na **seção 2.9**. Para todas as modelagens foram usadas a seguinte faixa: entre 1,8 e 2,2 minutos para TRI, entre 2,4 e 2,84 minutos para FEN e PUT, entre 2,84 e 3,2 para CAD e entre 3,1 e 3,6 para HIS, como mostrado na **Tabela 1S**.

3.7.2. MCR-ALS

O MCR-ALS é um algoritmo de deconvolução bilinear de curvas e como tal, inicialmente, requer que seja definido o número de componentes presentes nos dados aumentados. Pode-se estimar esse valor pelo conhecimento químico do sistema ou pela inspeção do autovalor de um modelo de componentes principais. Apesar de ser um indicativo adequado, é necessário ter conhecimento do sistema de modo a ser possível definir se os perfis recuperados condizem com o esperado. Dessa forma, neste trabalho, são avaliados os autovalores e os perfis recuperados para estimar corretamente o número de componentes necessários para construir um modelo MCR-ALS.

Os perfis são otimizados usando mínimos quadrados alternados. Assim, as matrizes de concentração (C) e espectrais (S) são estimadas alternadamente de modo a minimizar o resíduo calculado pela diferença da matriz aumentada original. Para reduzir a ambiguidade rotacional e acelerar a convergência do algoritmo a perfis quimicamente condizentes com o dos dados originais, a inicialização foi feita usando o algoritmo PURE no modo não aumentado e foram utilizadas, durante a iteração do ALS, as restrições: (**1**) organização de perfis por ordem de variável pura; (**2**) não negatividade nos modos de eluição e espectros; (**3**) unimodalidade no modo de eluição; (**4**) correspondência entre as espécies; (**5**) normalização do modo não aumentado; (**6**) Trilinearidade. Modelos usando restrição de trilinearidade foram comparados a resultados dos que não utilizaram essa restrição. A concentração da amostra é calculada ao estimar os perfis das amostras de teste juntamente com as de calibração e, em seguida, interpolar os escores das amostras de teste na curva construída com os escores das amostras de calibração, numa etapa conhecida como calibração pseudo univariada (de la Peña, Goicoechea, Escandar, Olivieri, 2015).

3.7.3. UPLS-RBL

O modelo UPLS-RBL é construído de forma similar ao modelo PLS. Inicialmente, desdobra o tensor de dados de três vias em uma matriz bidimensional e constrói-se um modelo PLS apenas com as amostras de calibração, diferentemente dos modelos PARAFAC e MCR-ALS nos quais as amostras de teste e de calibração são modeladas juntas. No caso desta tese, os modelos foram construídos com uma variável latente cada, uma vez que se sabe que há apenas uma fonte de variação nesses dados, o analito preparado individualmente.

Antes de realizar as predições da concentração das amostras, testa-se o resíduo da amostra de teste e aplica-se a etapa de RBL apenas nas amostras que apresentarem resíduos maiores que

os das amostras de calibração. Nessas amostras há interferência, sendo necessário remover a contribuição do interferente para corrigir os escores e realizar predições confiáveis. Dessa forma, é avaliado o resíduo de cada amostra individualmente e o número de RBLs usados é escolhido com base no resíduo da amostra analisada. Ao usar o número de RBLs apropriado, o valor de escore de PLS é adequado para realizar o procedimento de quantificação das amostras de teste similar como é feito no PLS de ordem 1 (de la Peña, Goicoechea, Escandar, Olivieri, 2015)

3.7.4. Software

Os dados foram tratados em ambiente MATLAB R2010a. Os cálculos envolvendo PARAFAC, MCL-ALS, UPLS-RBL, bem como os cálculos dos parâmetros de desempenho analítico, foram realizados usando a interface gráfica mvc2 desenvolvida por Olivieri e colaboradores (Olivieri, Wu, Yu, 2009), atualizada em 2016 e disponível em www.iquir-conicet.gov.ar/descargas/mvc2.rar.

Os cálculos envolvendo o MCR-ALS com restrição de trilinearidade e os cálculos de ambiguidade rotacional foram realizados usando a interface gráfica mcr_main desenvolvida por Tauler e colaboradores (Tauler, Barceló, 1993; Jaumot, de Juan, Tauler, 2015), atualizada em 2016 e disponível em www.mcrals.info/.

Os cálculos envolvendo o alinhamento de pico com o i*coshift* versão 1.1.1 foram realizados com a rotina implementada por Tomasi e colaboradores (Tomasi, Savorani, Engelsen, 2011), atualizada em 2010 e disponível em http://www.models.life.ku.dk/i*coshift*.

Capítulo 4

RESULTADOS E DISCUSSÃO

4 RESULTADOS E DISCUSSÃO

4.1 Conjunto de calibração

Na **Figura 13 (a)** são apresentados os cromatogramas médios para cada analito em cada um dos 12 níveis do conjunto de calibração, registrados em 215 nm. A eluição dos analitos ocorreu segundo o procedimento descrito na **seção 3.2**, que foi otimizado levando em conta o melhor compromisso entre os parâmetros cromatográficos como assimetria, resolução, largura de pico e o tempo total de análise. A vazão, apesar de alta, resultou em uma pressão inferior a 100 bar devido a quantidade de acetonitrila na fase móvel. As condições de temperatura e volume de injeção foram definidas visando a melhor interação entre o analito, a fase móvel e a fase estacionária de modo que não haja distorção do pico.

Os conjuntos de calibração foram preparados com cada analito individualmente, pois o alinhamento da mistura das ABs não foi eficiente, então mesmo com um maior esforço de análise a melhor estratégia foi alinhar o conjunto de calibração o que requereu que este fosse preparado individualmente, como mostrado na **seção 4.2**. Assim, inicialmente, foram preparadas misturas das ABs para o conjunto de calibração. No entanto, devido ao alinhamento não satisfatório dos cromatogramas com o icoshift, como mostrado na **seção 4.2**, e ao deslocamento da retenção de cada analito ser diferente entre si. Assim, nesta tese a abordagem utilizada para atingir um alinhamento eficiente foi preparar um conjunto de calibração com cada analito individualmente e alinhar o conjunto de calibração em detrimento do das amostras que se deseja predizer a concentração.

No tempo morto da coluna, em aproximadamente 0,6 minutos, são eluídas uma grande quantidade de constituintes. Esses constituintes possuem baixa retenção, um caráter mais polar, e são referentes às espécies químicas que possuem cromóforos provenientes da etapa de derivatização (**seção 3.3**). Em aproximadamente 1,2 minutos é eluído o reagente derivatizante, cloreto de dansil (DNS), que foi adicionado em excesso.

Na **Figura 13 (b)** as linhas vermelhas demarcam a região entre 1,8 e 3,6 minutos onde são eluídas as aminas biogênicas derivatizadas analisadas: triptamina em azul, 2-feniletilamina em verde, putrescina em vermelho, cadaverina em ciano e histamina em magenta.

Na **Figura 13 (c)** estão mostrados os espectros referentes a cada um dos analitos na concentração de 20 mg L^{-1}, identificados nas suas respectivas cores. A linha vertical preta em 215 nm marca o comprimento de onda que representa a maior relação sinal ruído para todos os analitos. Ainda é possível visualizar nesta mesma figura que o perfil dos espectros é semelhante.

Isso se deve à similaridade entre as estruturas dos analitos derivatizados, uma vez que a maior parte dos cromóforos com transições menos energéticas refere-se ao reagente derivatizante DNS. Essa similaridade dificulta a identificação apenas pelo comprimento de onda e pode vir a confundir algoritmos quimiométricos que façam uso da informação contida em todo o espectro.

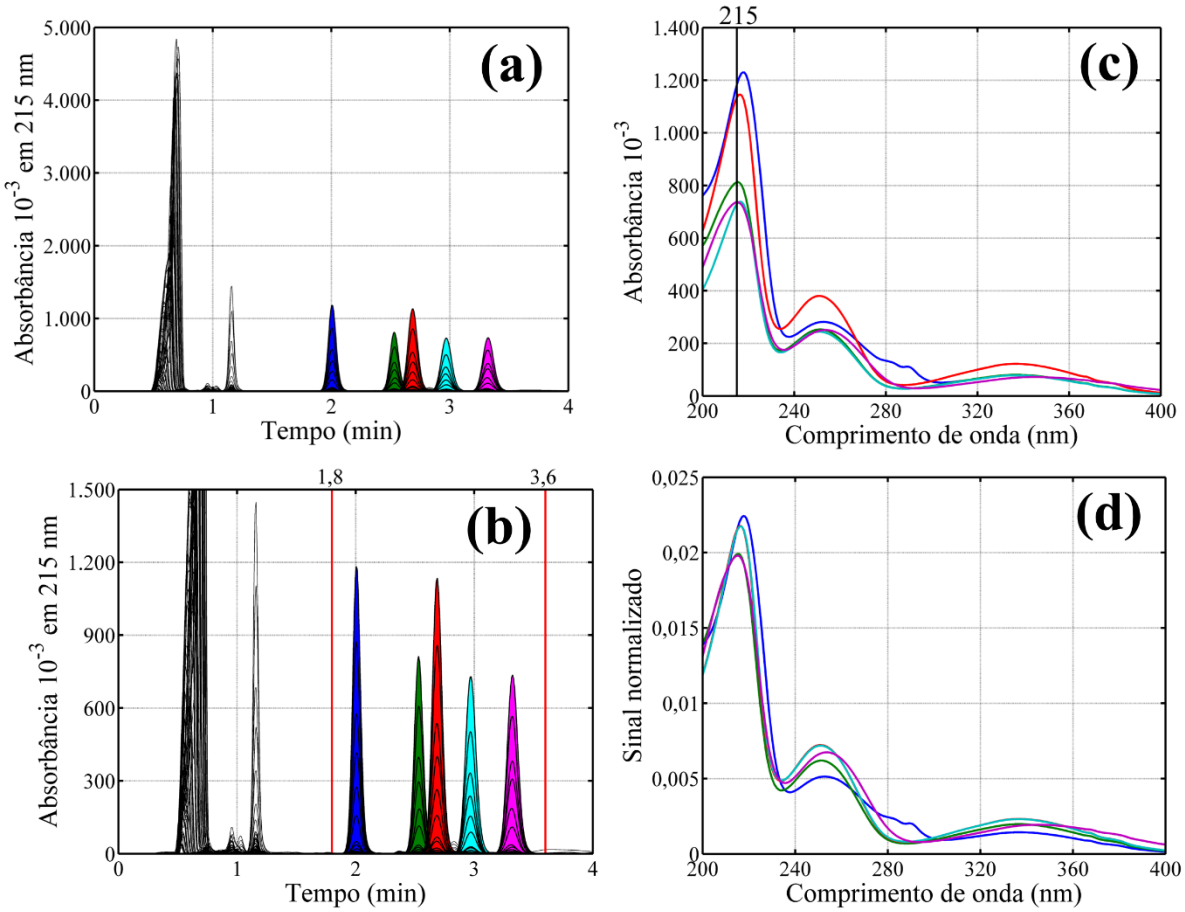

Figura 13 – Cromatogramas das médias de cada nível de concentração do conjunto de calibração, (**a**) registrado em 215 nm, (**b**) entre as linhas vermelhas está a região selecionada para a modelagem, onde aparece as cinco aminas biogênicas derivatizadas analisadas, TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—),(**c**) espectros dos analitos com o comprimento de onda de 215 nm destacado com uma linha vertical preta e (**d**) os mesmos espectros normalizados.

Na **Figura 13 d** está mostrado o sinal normalizado das ABs, a similaridade entre os espectros de PUT e CAD são mais acentuadas, sobrepondo-se em toda região. Percebe-se, também, que a HIS e a FEN são quase idênticas e apenas a TRI difere-se das outras ABs em comprimentos de onda próximos a 280 nm.

As estruturas das ABs podem ser visualizadas na **Figura 1**. A comparação das estruturas químicas dessas ABs justifica a similaridade dos espectros de absorção molecular mostradas na **Figura 13 d**. A diferença entre a PUT e a CAD é de apenas uma ligação CH_2 a mais na estrutura da CAD. Ambos os espectros serão idênticos, uma vez que esta ligação não apresenta cromóforos

e, assim, não apresentando sinal no UV. No entanto, após a derivatização com o DNS, a diferença desta ligação CH_2 permite separar esses analitos na coluna cromatográfica. Essa pequena diferença na retenção (**Figura 13 a e 13 b**) possibilita alcançar uma quantificação exata com os modelos de segunda ordem usados.

4.2 Avaliação do algoritmo de alinhamento nos conjuntos de calibração e teste

Ao eluir o conjunto de teste, percebe-se um pequeno deslocamento do tempo de retenção para os analitos, em especial para a PUT e para a HIS, quando comparado com o conjunto de calibração. Como esperado, algoritmos trilineares como o PARAFAC não apresentam resultados satisfatórios frente à quebra da trilinearidade no modo de eluição, característico dos métodos cromatográficos de análise.

A fim de superar esta característica do método de análise, deve-se manter esse modo flexível a pequenas variações ao modelar com algoritmos apropriados como o PARAFAC2, UPLS-RBL, ou o MCR-ALS. No entanto, neste trabalho foi identificada outra característica que dificulta a modelagem: os espectros dos analitos são muito correlacionados sendo os da PUT e da CAD idênticos como ilustrados na **Figura 13 (d)**.

Mesmo algoritmos que permitem a quebra da trilinearidade no modo de eluição, levarão à deficiência de posto quando espectros idênticos são modelados, ou seja, esses analitos serão recuperados como sendo um único perfil no modo de eluição. Por consequência, o posto químico irá diferir do matemático, levando à deficiência de posto, também conhecida como dependência linear ou quebra da trilinearidade no modo dos espectros (Parastar, 2015; Amrhein, Srinivasan, Bonvin, Schumacher, 1996; Bro, Harshman, Sidiropoulos, Lundy, 2009).

Quebra de trilinearidade no modo de eluição e deficiência de posto devido à baixa seletividade do detector é uma situação que nenhum algoritmo de calibração de ordem superior implementado, até hoje, consegue resolver completamente. Assim, é possível seguir dois caminhos para garantir uma quantificação exata desses analitos: (**1**) usar um detector mais seletivo ou (**2**) resolver uma dessas quebras de trilinearidade. Para o caso de diferenças mínimas como a de uma única ligação CH_2 é difícil encontrar um detector que possa ser sincronizado no HPLC e que gere sinais seletivos para ambas as moléculas derivatizadas de PUT e CAD. A alternativa encontrada para resolver esse problema e alcançar a vantagem de segunda ordem foi resolver a quebra da trilinearidade no modo de eluição ao realizar um alinhamento prévio dos cromatogramas (Pinto, Díaz Nieto, Zón, Fernández, Araújo, 2016).

Como mostrado anteriormente na descrição da teoria do algoritmo de alinhamento, é possível ver na literatura diversas aplicações do COW e do *coshift* e suas variações em dados de ordem 1. Uma vez que nos dados desta tese não é percebida nenhuma deformação do pico como alargamento ou mudança da assimetria (**Figura 1S**), basta resolver o deslocamento na retenção dos analitos. Neste trabalho, foi testado o *coshift* e o i*coshift*, que apesar de ser proposto para dados de ordem 1, pode alinhar dados de segunda ordem da forma que está implementado, sendo necessário se ater a alguns detalhes (Pinto, Díaz Nieto, Zón, Fernández, Araújo, 2016; Tomasi, Savorani, Engelsen, 2011).

No algoritmo i*coshift*, o alinhamento por seção do cromatograma ou do cromatograma completo (*coshift*) pode ser feito, assim, o mesmo algoritmo foi usado em ambos os casos. No i*coshift* o analista pode entrar com a matriz do tipo **m**x**n**, onde **m** é a dimensão das amostras e **n** as variáveis dos dados de ordem 1. Dessa forma, o algoritmo alinha a dimensão **n** para as várias amostras de **m**. É possível, ainda, alinhar um tensor do tipo **m**x**n**x**o**, onde **m** é a dimensão das amostras, **n** as variáveis de um modo de aquisição, e **o** as do outro modo, no caso de dados de ordem 2. Entrando com os dados em forma de um tensor de três vias do tipo **m**x**n**x**o**, o algoritmo irá alinhar a dimensão **n** registrada na primeira variável de **o** para as amostras de **m**. Na prática, isso significa alinhar cada cromatograma registrado em cada comprimento de onda individualmente para as várias amostras.

Na **Figura 14 a-e** são mostradas, respectivamente, as amostras de teste alinhadas por deslocamento completo (**e**) e com 20 (**a**), 40 (**b**), 60 (**c**) e 120 (**d**) intervalos, com o alinhamento alvo na amostra de calibração de 20 mg L^{-1}, destacado com área colorida. Para melhor visualização de cada uma desses cromatogramas ver **Figura 2S a-e** e **2S a1-e1** no **Apêndice**.

O total de variáveis nesta região é de 240. Apesar do alinhamento melhorar a partir de 60 intervalos nos cromatogramas, quando comparado com o alinhamento completo o resultado ainda não é satisfatório para restaurar a trilinearidade dos dados. Observando os cromatogramas em 3D na **Figura 14 a1-d1** é possível notar que ocorrem distorções quando o alinhamento é feito por intervalos. O alinhamento por deslocamento total (modo *whole*) representa o algoritmo *coshift* e resulta em um perfil tridimensional da matriz LC-DAD sem distorção, como pode ser visto na **Figura 14 e** e **e1**.

Apesar de sutil, há uma melhora nos cromatogramas registrados em 215 nm quando se alinha por intervalos (**Figura 14 a-d**), em especial para os analitos FEN (verde) e CAD (ciano). No entanto, quando comparados aos dados originais os resultados são piores, pois os cromatogramas para o mesmo conjunto de amostras, calibração ou teste, perdem o alinhamento original (**Figura 1S**).

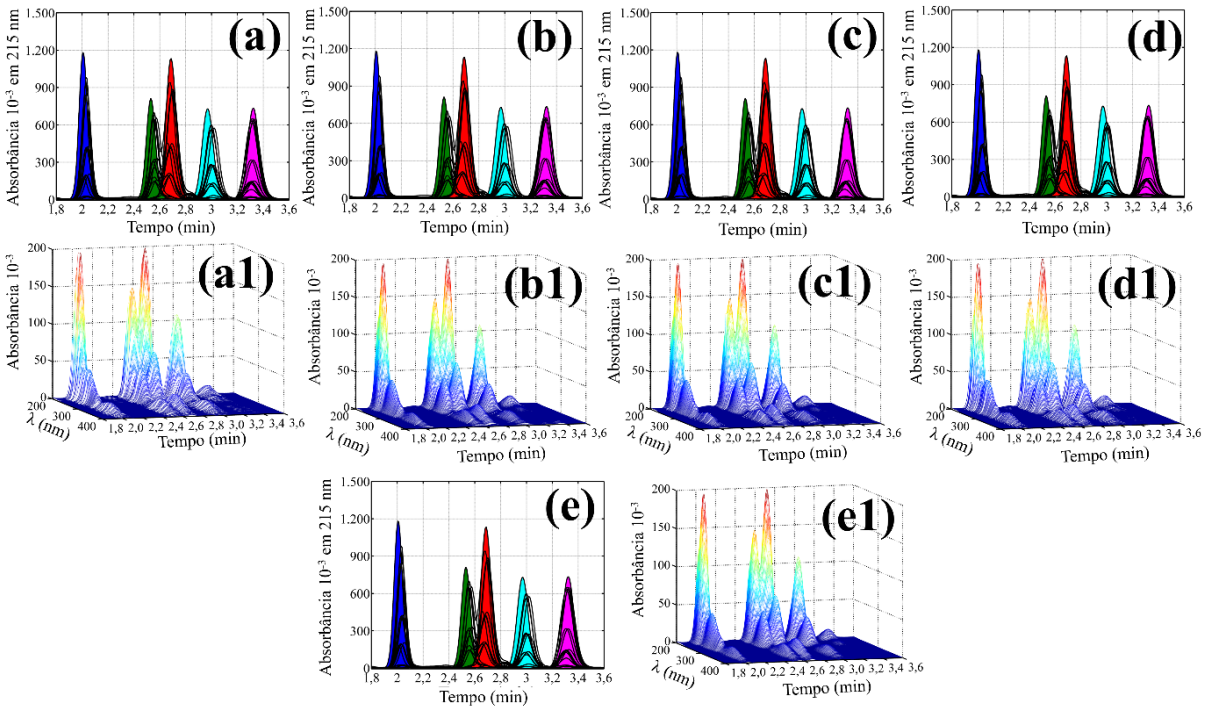

Figura 14 – Cromatogramas das amostras de teste (—) alinhados por intervalos (**a-d**) e por deslocamento completo (**e**) sobreposto a amostra de calibração de 20 mg L^{-1} (área colorida representando cada analito TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—)). Cromatograma em 3D da amostra de teste 2 alinhado por intervalos (**a1-d1**) e por deslocamento completo (**e1**). Os intervalos testados foram 20 (**a**), 40 (**b**), 60 (**c**) e 120 (**d**). Todas essas figuras são mostradas individualmente e ampliada na **Figura 2S** no apêndice.

Mesmo se os cromatogramas forem divididos em regiões e alinhados, os resultados não são satisfatórios. Na **Figura 15**, pode ser visto o deslocamento do tempo de retenção para a PUT, enquanto a FEN apresenta-se alinhada. Foi realizado o alinhamento com o *coshift* e com o i*coshift* divididos em 5, 10, 15 e 30 intervalos. Da mesma forma que para o cromatograma completo, o alinhamento não foi eficiente em nenhum dos casos e, nos dados tridimensionais, **Figura 15 a1-d1**, o cromatograma foi distorcido ao alinhar por intervalos, enquanto que por deslocamento completo não, como pode ser visto na **Figura 15**. As figuras que compõem a **Figura 15** podem ser vistas ampliadas nas **Figuras 3S a-e** e **a1-e1** no **Apêndice**.

Os cromatogramas registrados em 215 nm para os dados de teste originais são mostrados na **Figura 16** e **1S**. Na **Figura 16 a**, são mostrados os cromatogramas originais referentes ao conjunto de calibração e na **Figura 16 b** os cromatogramas originais referentes ao conjunto de calibração sobreposto ao conjunto de teste sem alinhamento, onde se pode perceber bem o deslocamento do tempo de retenção. Nota-se que todos os analitos possuem um deslocamento do tempo de retenção quando comparados ao conjunto de calibração e teste, e que para alguns analitos (PUT e HIS) o deslocamento é mais acentuado. Por causa dessas diferenças no deslocamento que varia de analito para analito, torna-se mais difícil a tarefa de alinhamento dos

dados com mais de um pico na mesma região usando o deslocamento total, ainda mais os analitos que estão parcialmente coeluídos. Para manter a estrutura dos dados tridimensionais sem distorção, optou-se por usar o *coshift* em detrimento do i*coshift*, uma vez que foram observadas distorções nos dados tridimensionais com o alinhamento por intervalos. Isso pode ser agravado quando há coeluições, o que é esperado para a amostra de peixe, impossibilitando alinhar os picos e alcançar a trilinearidade para os dados.

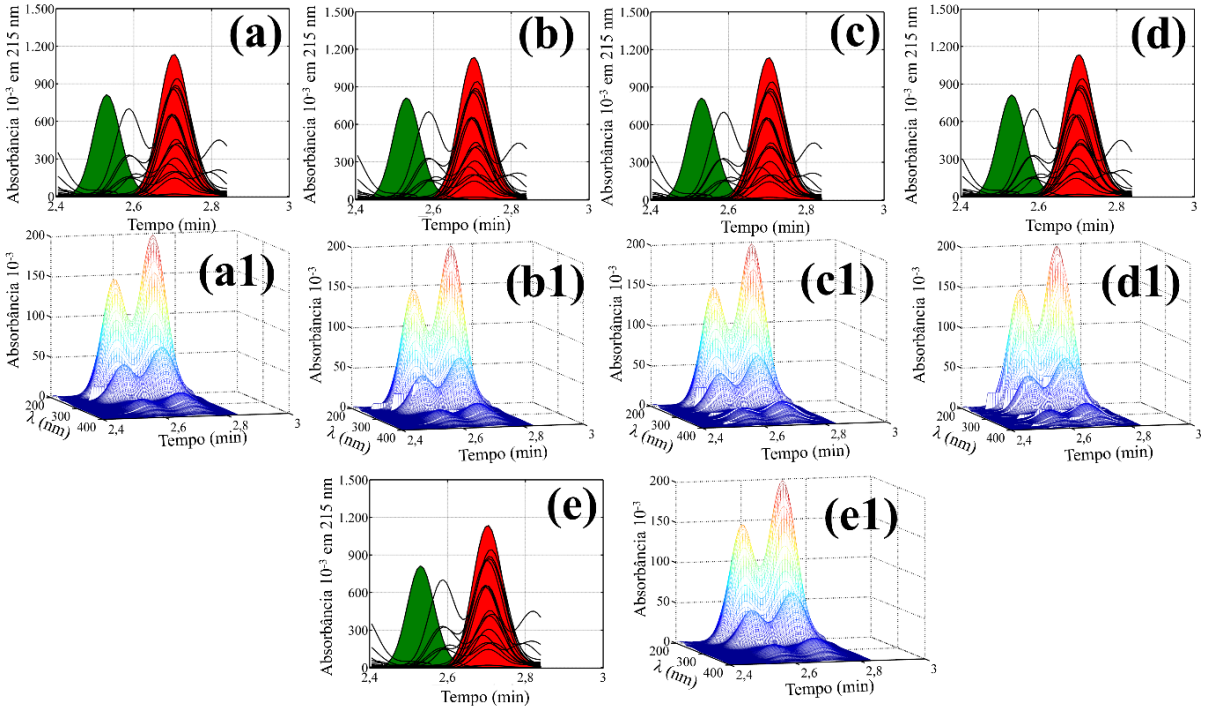

Figura 15 – Cromatogramas das amostras de testes (**—**) alinhados por intervalos (**a-d**) e por deslocamento completo (**e**) sobreposto à amostra de calibração de 20 mg L^{-1} (área colorida representando cada analito). Cromatograma em 3D da amostra de teste 2 alinhado por intervalos (**a1-d1**) e por deslocamento completo (**e1**). Os intervalos testados foram 5 (**a**), 10 (**b**), 15 (**c**) e 30 (**d**). Todas essas figuras são mostradas individualmente e ampliada na **Figura 3S** no apêndice.

Para que o alinhamento fosse eficiente, neste trabalho, optou-se por fazê-lo no conjunto de calibração e não no de teste. Nessa abordagem, há como alinhar de forma satisfatória o conjunto de calibração, uma vez que não haverá coeluições, o que torna este conjunto mais fácil de ser alinhado do que o que contém amostras com constituintes coeluídos. Como a trilinearidade no modo de eluição somente é perdida entre os conjuntos de calibração e teste, a estrutura trilinear é alcançada nos dados ao alinhar os sinais dos analitos, não requerendo que os sinais dos interferentes sejam alinhados, o que torna viável esse tratamento apenas no conjunto de calibração.

Na **Figura 16 c** e **d** são mostrados os cromatogramas de calibração alinhados e os mesmos sobrepostos ao conjunto de teste. Apesar de sutil percebe-se um deslocamento para a esquerda do cromatograma de calibração da PUT e da HIS quando comparado às **Figuras 16a** e **c** e **Figura**

16b e **d**. Assim, foi possível resolver a quebra de trilinearidade dos dados que será comprovada matematicamente ao se comparar os modelos bilineares e trilineares construídos no tratamento com o MCR-ALS. O cromatograma de calibração em 3D alinhado não foi distorcido, conforme pode ser observado na **Figura 4S a** e em suas respectivas faixas modeladas na **Figura 4S b e c** no **Apêndice**.

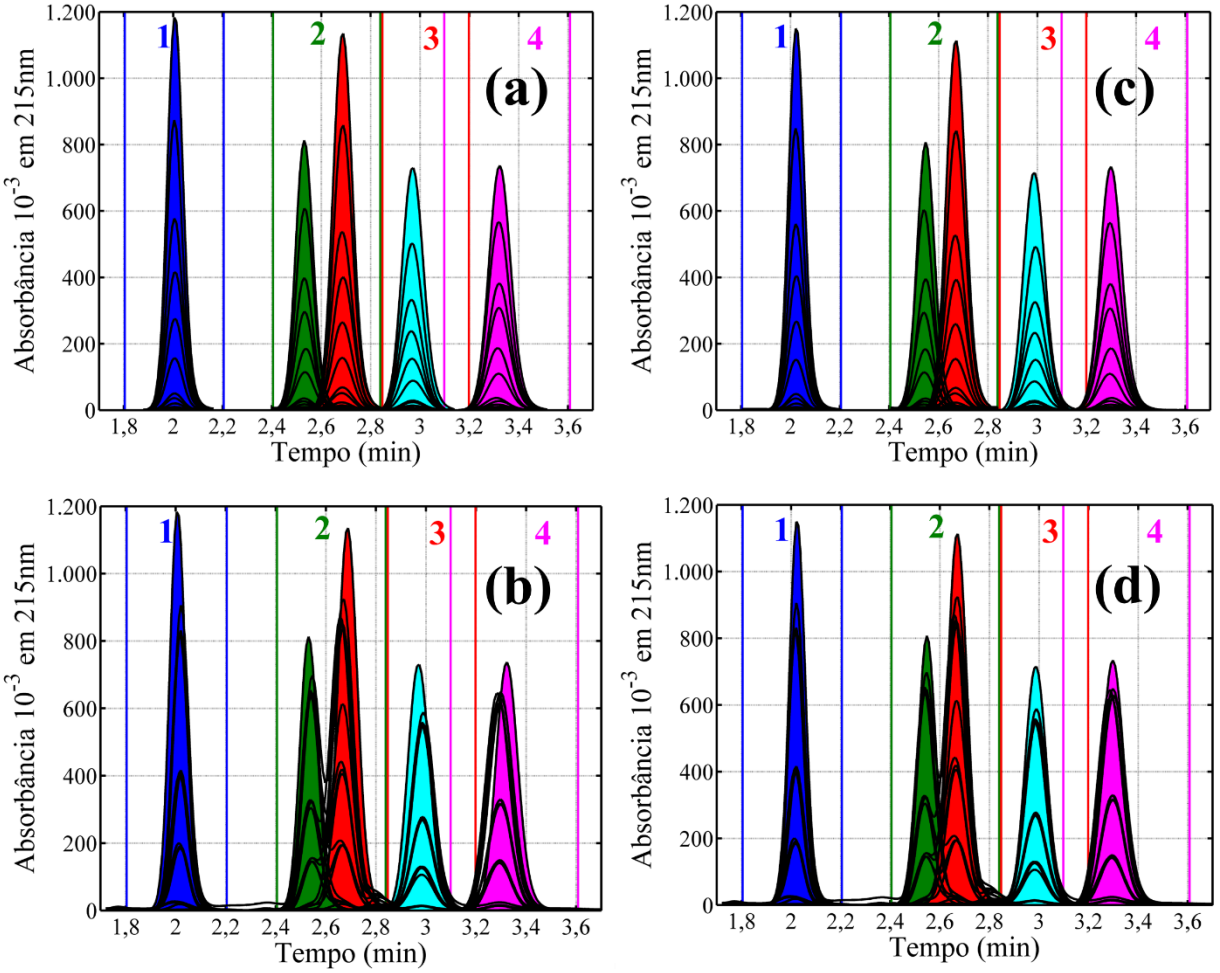

Figura 16 – Cromatogramas das amostras do conjunto de calibração sem alinhar (**a**) e alinhando (**c**). Cromatograma da amostra de calibração de 20 mg L^{-1} (colorida de acordo com o analito) sobreposta aos das amostras do conjunto de teste (**—**) com o conjunto de calibração não alinhado (**b**) e alinhado (**d**), as 4 regiões modeladas são mostradas entre linhas coloridas (**Tabela 1S**).

É importante notar que o uso do *coshift*, nesse caso, somente é possível por quatro motivos: (**1**) o formato do espectro não se altera, apenas a sua intensidade cresce até atingir o máximo no tempo de retenção do analito e decresce até o analito ser completamente eluído, formando assim o pico cromatográfico; (**2**) as amostras alinhadas serão as do conjunto de calibração onde estará presente apenas um analito por conjunto, ou seja, não há interferente. Caso se tente alinhar um cromatograma com constituintes diferentes em um determinado comprimento de onda, a intensidade do pico leva o algoritmo a confundir os sinais entre constituintes distintos, resultando

em um alinhamento com um espectro distorcido; (**3**) não há distorção no pico como alargamento ou mudança na sua assimetria, isso é importante, pois, também, levaria à quebra de trilinearidade no modo de eluição; (**4**) a composição do solvente em que a amostra de teste e de peixe estão preparadas é a mesma, uma vez que ambos são submetidos ao mesmo procedimento de extração e derivatização, a força iônica será a constante. Como esse deslocamento é proveniente de alterações na força iônica e não devido à distorção no pico, o tempo de retenção não varia quando um mesmo conjunto de amostra é analisado, e, sim, entre os conjuntos. Ou seja, entre as amostras de calibração não há deslocamento de pico, no entanto como as amostras de teste e de peixe foram previamente extraídas com ácido perclórico sua força iônica se difere da do conjunto de calibração e assim identificou-se um deslocamento de pico. Dessa forma, foi possível alinhar o sinal do conjunto de calibração com a amostra com a maior concentração do analito como amostra alvo do alinhamento, uma vez que não foi observado deslocamento de pico no mesmo conjunto.

4.3 Tratamento dos dados com o PARAFAC

Na **Tabela 2** estão mostrados os parâmetros usados para estimar o número de componentes para o PARAFAC. Uma vez que o conjunto de teste consiste em uma mistura dos padrões das cinco ABs, preparadas de acordo com o planejamento exposto na **Tabela 1**, é esperado que o posto químico seja igual ao número de analitos por região modelada. No entanto, percebe-se que a variância explicada aumenta e o ajuste residual reduz, de forma significativa para PUT e HIS quando se modela com um componente a mais, 3 e 2 respectivamente, indicando, desse modo, uma inconsistência dos cálculos com os dados experimentais mostrados na **Figura 16**. Levando em consideração esses dois parâmetros, observa-se que o posto químico difere do posto matemático, indicando que nos dados há uma quebra da trilinearidade.

Ao analisar o valor de consistência de core, é visto que há trilinearidade para a TRI e HIS mesmo recuperando um componente a mais devido à diferença no tempo de retenção, que pode ser visto na **Figura 17**, em que são mostrados os perfis dos analitos recuperados com 1, 2 e 3 componentes. Para a PUT, o valor de consistência de core indica corretamente a quebra da trilinearidade dos dados. No entanto, o valor de ajuste residual a partir de onde não há mais redução significativa, é alcançado com três componentes. Isso ocorre devido à quebra da trilinearidade proveniente do deslocamento na retenção mostrado na **Figura 16**. Apesar de o CORCONDIA ser usado para escolher o número de componentes, os parâmetros mais utilizados para essa decisão são a variância explicada do modelo e os perfis recuperados após a modelagem, em especial, para esse caso onde se sabe que há a quebra da trilinearidade.

Tabela 2 − Modelos construídos com o conjunto de calibração não alinhados, para o cromatograma dividido em regiões das misturas das ABs presentes no conjunto teste.

Comp*	Parâmetros	TRI	FEN	PUT	CAD	HIS
1	Consistência de Core	100,00			100,00	100,00
	Ajuste residual	3,77			3,41	14,50
	Variância explicada	99,62			99,46	90,65
	Iterações	2			2	2
2	Consistência de Core	81,38	98,86	98,25	99,51	98,67
	Ajuste residual	1,35	5,11	18,05	2,31	1,86
	Variância explicada	99,95	99,45	94,14	99,51	99,85
	Iterações	14	2	2	2	2
3	Consistência de Core	11,66	34,35	31,23	6,29	29,22
	Ajuste residual	1,04	3,70	3,19	1,17	1,03
	Variância explicada	99,97	99,71	99,82	99,94	99,95
	Iterações	316	25	35	95	264

*Número de componentes recuperados pelo PARAFAC

O modelo de segunda ordem construído com o PARAFAC requer que os dados sejam trilineares para a modelagem. No caso em estudo, há a quebra da trilinearidade no modo de eluição, o que inviabiliza o uso do PARAFAC nos dados que não foram alinhados. Ao recuperar o perfil referente a cada analito com o PARAFAC, em dados com quebra da trilinearidade, é necessário um componente a mais para explicar bem o modelo. Isso ocorre porque o algoritmo entende que há dois constituintes, e não um cujo tempo de eluição deslocou, como mostrado na **Tabela 2** e na **Figura 17**.

Os algoritmos de resolução de curvas permitem a visualização dos perfis recuperados que serão utilizados para a modelagem, isso é muito útil na inspeção da qualidade da modelagem dos dados. A inspeção desses perfis recuperados pelo algoritmo é de grande ajuda para se saber se a modelagem é feita com o sinal correto do analito. Na **Figura 17** é mostrado o perfil cromatográfico recuperado pelo PARAFAC quando decomposto com 1 (**a**), 2 (**b**) e 3 (**c**) componentes nas regiões definidas na **Tabela 1S**. O resultado do analito recuperado pelo PARAFAC é comparado com o sinal normalizado do padrão de calibração na **Figura 17 d**. A deconvolução para FEN e PUT foi feita a partir de 2 componentes, uma vez que já se sabe que essa região deve conter, ao menos, o sinal referente a esses dois analitos.

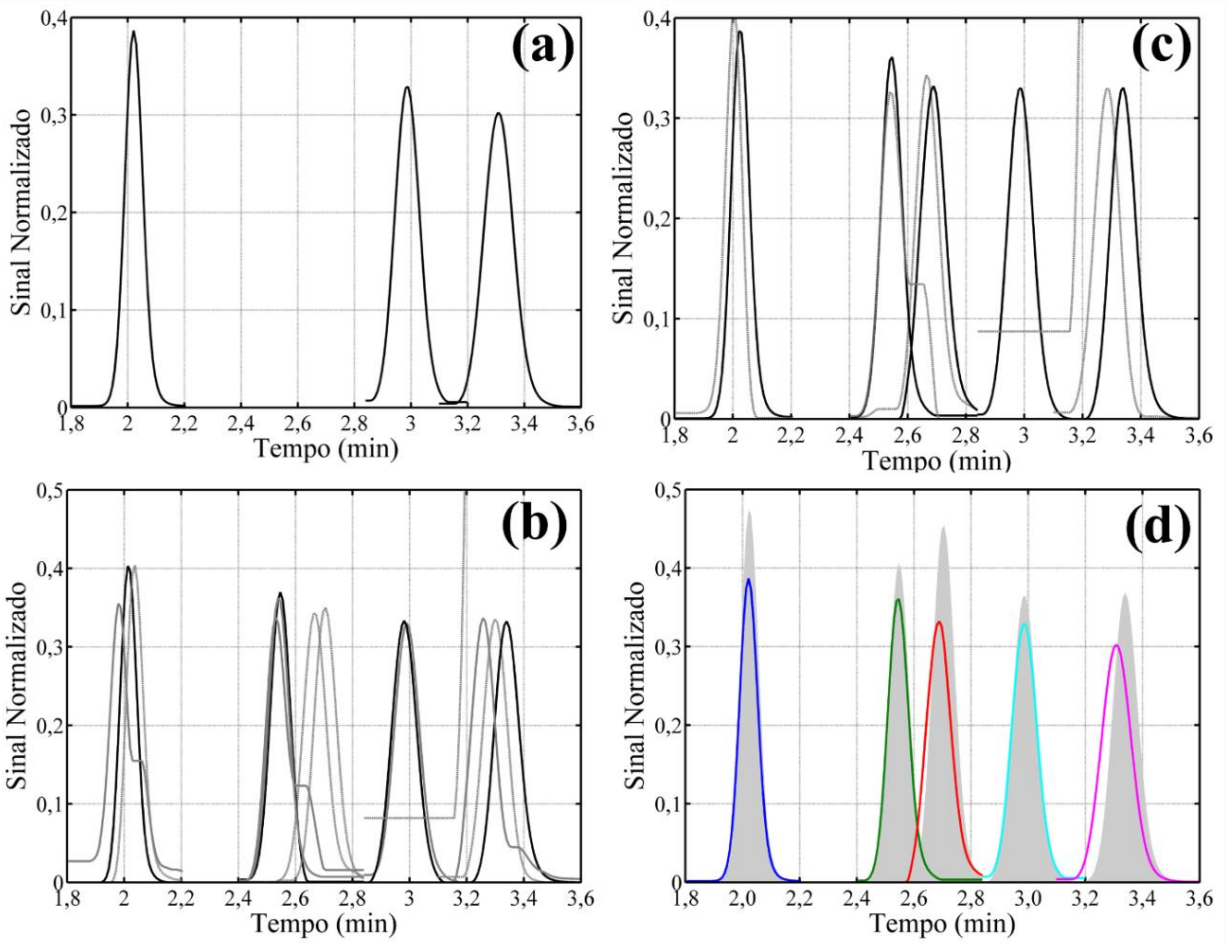

Figura 17 – Perfis recuperados para o PARAFAC com 1 (**a**), 2 (**b**) e 3 (**c**) componentes para os dados não alinhados. O perfil com 1 componente é representado por uma linha preta cheia (—), com 2 por uma linha tracejada cinza (···) e com 3 por uma linha cheia cinza (—). O sinal normalizado da amostra de calibração de 20 mg L^{-1} em cinza é sobreposto ao sinal normalizado recuperado pelo PARAFAC apenas para o sinal da TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—) (**d**).

Ao modelar a TRI com um perfil, obtém-se resultados satisfatórios, com REP de 8,0% ressaltando a exatidão do modelo. Quando se realiza os cálculos recuperando-se dois perfis, percebe-se que o PARAFAC é forçado a recuperar outro sinal que não condiz com os dados experimentais. Situação similar é vista quando se recupera um componente a mais que o presente na região dos outros analitos.

Apesar da CAD não apresentar um deslocamento tão acentuado quanto a PUT e a HIS, foi necessário recuperar um componente a mais referente ao sinal da HIS que se deslocou para a região de modelagem da CAD no conjunto de teste. Na **Figura 18**, é mostrado o cromatograma para a faixa da CAD e HIS (**a**) e individualmente em suas respectivas regiões modeladas (**b** e **c**). É possível visualizar que ainda há sinal da HIS na região selecionada para a CAD, mas o contrário ocorre em menor intensidade. Isso explica o porquê do PARAFAC requerer dois componentes

para recuperar os perfis de forma satisfatória nos dados da CAD, mas só um para a região da HIS.

Uma vez que os dados foram bem alinhados para a CAD, o valor de REP obtido para o modelo foi de 3,8% quando duas componentes foram recuperadas. A FEN apresentou um valor de REP de 3,1% quando os dados foram decompostos em dois perfis. No caso da HIS e da PUT que tiveram deslocamento, os modelos apresentaram baixa exatidão com valores de REP de 18,6 e 20,5 %, respectivamente.

O sinal de eluição normalizado da amostra de calibração de 20 mg L^{-1} foi comparado com os perfis recuperados pelo PARAFAC com 1 componente para TRI e HIS e 2 componentes para FEN, PUT e CAD. Na **Figura 17 d**, é possível observar o deslocamento do cromatograma em cores e sombreado para a PUT e a CAD, o que indica a quebra da trilinearidade e justifica o resultado da baixa exatidão. Assim, como esperado, o PARAFAC sofre com a quebra da trilinearidade e a única forma de aplicá-lo a dados de cromatografia é realizando um alinhamento. Os dados de calibração foram então alinhados com o *coshift* como mostrado na **seção 4.2**, e, dessa, forma passou a haver uma consistência do posto químico com o matemático.

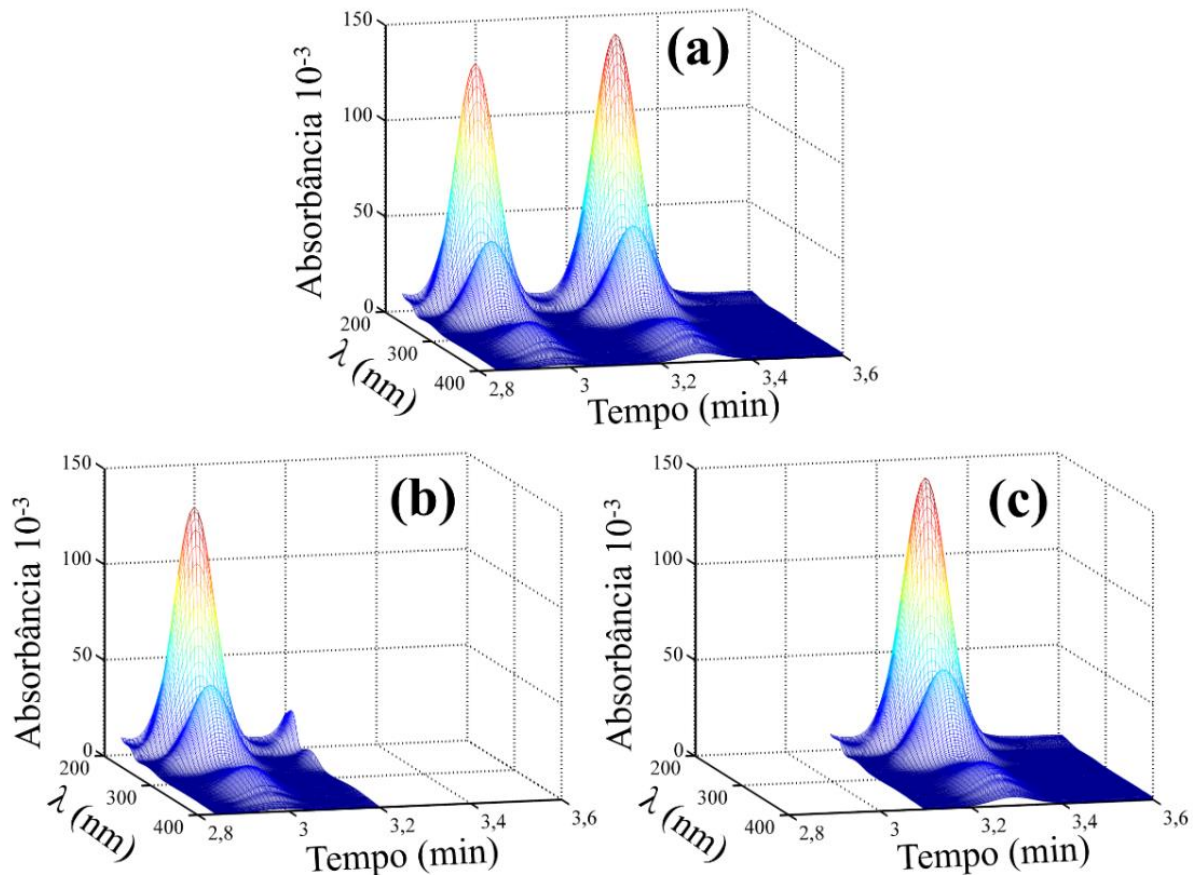

Figura 18 – Cromatograma em 3D da amostra de teste 5 registrado na faixa dos dois analitos juntos (**a**), e separadamente para a CAD (**b**) e para a HIS (**c**).

Na **Tabela 3** são apresentados os parâmetros usados na escolha do número de componentes necessários para a construção do modelo PARAFAC. Diferentemente de quando os dados não estavam alinhados, a variância explicada não aumenta significativamente a partir do número de componentes igual ao de constituintes presentes na região estudada, sendo 1 para TRI e HIS e 2 para FEN, PUT e CAD.

Mesmo realizando o alinhamento do conjunto de calibração, o sinal da HIS continua aparecendo na região da CAD, como mostrado na **Figura 16 d**. No entanto, da mesma forma para os dados não alinhados, a contribuição desse sinal não resulta em uma mudança significativa na variância explicada, sendo necessária a análise dos perfis recuperados para identificar corretamente o número de constituintes presentes na região modelada. Outro fator interessante é a necessidade de mais iterações para a convergência do algoritmo quando se testa recuperar um perfil a mais que o presente na região modelada (**Tabela 3**). Além dos perfis recuperados (**Figura 19**) não serem condizentes com o registrado pelo HPLC-DAD. Nos perfis recuperados na **Figura 19 d**, é perceptível a eficiência do alinhamento ao comparar com os perfis não alinhados na **Figura 17d**.

Tabela 3 – Modelos construídos com o conjunto de calibração alinhados para o cromatograma dividido em regiões das misturas das ABs presentes no conjunto teste.

Comp*	Parâmetros	TRI	FEN	PUT	CAD	HIS
1	Consistência de core	100			100	100
	Ajuste residual	3,77			3,64	3,7
	Variância explicada	99,62			99,39	99,39
	Iterações	2			2	2
2	Consistência de core	82,6	99,14	98,71	99,52	34,98
	Ajuste residual	1,35	4,77	6,55	2,65	1,65
	Variância explicada	99,95	99,52	99,22	99,67	99,88
	Iterações	15	2	2	2	25
3	Consistência de core	12,34	19,60	63,78	39,93	23,65
	Ajuste residual	1,04	3,46	4,24	1,2	0,95
	Variância explicada	99,97	99,75	99,67	99,93	99,96
	Iterações	324	21,00	63,00	38	83

*Número de componentes recuperados pelo PARAFAC para o cromatograma completo dividido em faixas

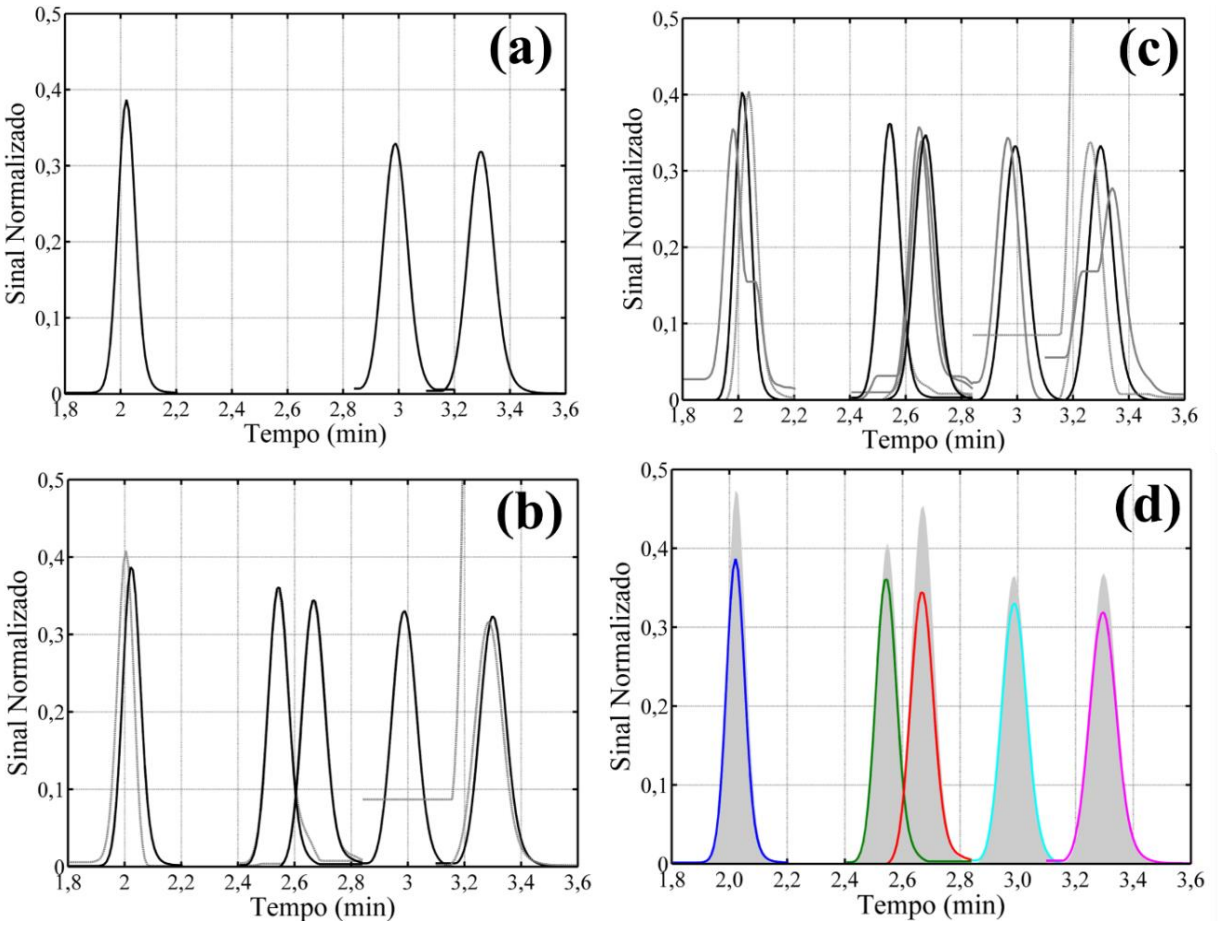

Figura 19 − Perfis recuperados para o PARAFAC com 1 (**a**), 2 (**b**) e 3 (**c**) componentes para os dados alinhados. Os perfis recuperados com 1, 2 e 3 componentes são respectivamente representados por uma linha preta cheia (—), uma linha tracejada cinza (- - -) e uma linha cheia cinza (—). O sinal normalizado da amostra de calibração de 20 mg L^{-1} em cinza é sobreposto ao sinal normalizado recuperado pelo PARAFAC apenas para o sinal da TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—) (**d**).

Com o modelo alinhado e bem ajustado, foram calculadas as métricas de desempenho analítico para os cinco modelos construídos para cada analito. Como pode ser visto na **Tabela 4**, os valores de REP foram todos inferiores a 10%, o que indica modelos exatos. No entanto, a seletividade para a FEN, PUT e CAD, onde foram recuperados dois perfis, foi muito baixa. Isso ocorre devido à elevada correlação entre os perfis espectrais da FEN com a PUT e da CAD com a HIS. Como a sensibilidade analítica depende da seletividade do modelo, os resultados de LOD e LOQ foram substancialmente afetados. Para os modelos da TRI e HIS, onde só um componente foi recuperado, a seletividade é igual a 1. Uma vez que não há interferências, os valores de LOQ foram condizentes com a curva univariada construída e o sinal experimental observado no conjunto de calibração.

Apesar dos dados estarem bem alinhados resultando em modelos exatos para o PARAFAC, as métricas de calibração como sensibilidade analítica e os limites de detecção e quantificação não são condizentes com o observado nos dados experimentais. Isso ocorre devido à deficiência

de posto causada pela similaridade entre os perfis espectrais. Tal dependência linear pode ser contornada usando algoritmos apropriados como o PARALIND. No entanto, modelos com o PARALIND não foram construídos neste trabalho devido às métricas de desempenho ainda não estarem implementadas, limitando-se a analisar apenas o REP para comparação entre os modelos construídos.

Tabela 4 − Métricas de desempenho analítico calculadas para os cinco modelos PARAFAC aplicadas ao conjunto de teste com o de calibração alinhado. As concentrações dos conjuntos de teste e de calibração variaram respectivamente nas faixas de 0,4 a 16 mg L^{-1} e 0,05 a 20 mg L^{-1} para todos os 5 analitos.

Métricas de desempenho PARAFAC	TRI (1)	FEN (2)	PUT (2)	CAD (2)	HIS (1)
RMSEP (mg L^{-1})	0,80	0,36	0,61	0,39	0,98
REP (%)	8,0	3,6	6,1	3,9	9,9
Sensibilidade (L mg^{-1})	723,17	46,73	74,38	53,92	563,62
Seletividade	1,00	0,09	0,10	0,11	1,00
Sensibilidade analítica (L mg^{-1})	191,75	9,80	11,37	20,38	152,29
LOD (mg L^{-1})	0,02	0,34	0,29	0,16	0,02
LOQ (mg L^{-1})	0,05	1,02	0,88	0,49	0,07

Entre parênteses é mostrado o número de constituintes estimados. **LOD**: limite de detecção; **LOQ**: limite de quantificação; **RMSEP**: Raiz quadrada do erro médio de predição (do inglês *Root mean square error of prediction*); **REP**: Erro relativo de predição (do inglês *Relative error of prediction*). Todas as métricas de desempenho foram calculadas de acordo com Olivieri e coautores (Olivieri, 2014, Bauza, Ibañez, Tauler, Olivieri, 2012).

4.4 Tratamento dos dados com o UPLS-RBL

O procedimento para construir o modelo UPLS-RBL se inicia ao escolher o número de variáveis latentes (VL). Uma vez que foram usados padrões puros derivatizados na calibração, o número de VL corresponde ao de constituintes presentes na calibração. Como o conjunto de calibração foi construído para cada analito individualmente, com base no conhecimento químico do sistema, é lógico inferir que é necessária apenas uma VL para explicar bem cada modelo.

Assim, modelos de calibração linear foram construídos para cada analito que apresentaram variância explicada acima de 99%, tanto para a matriz de dados experimentais, quanto para o vetor de concentrações. Além disso, foi obtido um valor de ajuste residual baixo e próximo entre os modelos construídos para os cinco analitos, como mostrado na **Tabela 5**, indicando a necessidade de apenas uma VL para explicar o conjunto de calibração.

Ao analisar os cromatogramas divididos por regiões, é necessário aplicar RBL apenas para a região da FEN e PUT que estão parcialmente coeluidos. Isso foi necessário, pois o conjunto de

calibração possui apenas um analito. O valor de REP obtido foi de 8,6%, 26,6%, 197,6%, 4,5% e 12,3% para TRI, FEN, PUT, CAD e HIS, respectivamente. Resultados não exatos foram obtidos para a HIS e a PUT, devido a esses analitos não estarem alinhados. No entanto, esperava-se que os resultados na predição da concentração da FEN fossem melhores uma vez que perfis correlacionados não deveriam causar deficiência de posto em dados desdobrados (Borraccetti, Damiani, Olivieri, 2009). Essa não exatidão na predição da concentração da FEN pode ser explicada ao avaliar o perfil recuperado pelo RBL, onde foi identificado que o algoritmo estava modelando o perfil do próprio analito como sendo interferente na amostra de teste, como pode ser visto na **Figura 20**.

Tabela 5 − Variância explicada e ajuste residual para os modelos construídos para cada analito com uma VL com base nos dados não alinhados.

Analito	Variância (x)[a]	Variância (y)[b]	Resíduo (x)[c]	Resíduo (y)[d]
Triptamina	99,92	99,95	1,72	0,18
Fenilamina	99,93	99,97	1,11	0,13
Putrescina	99,93	99,96	1,54	0,16
Cadaverina	99,88	99,70	1,48	0,46
Histamina	99,91	99,90	1,28	0,26

a: variância explicada da matriz de dados; **b**: variância explicada do vetor de resposta; **c**: ajuste residual da matriz de dados; **d**: ajuste residual do vetor de resposta.

Na **Figura 20 a** e **d** são mostrados, respectivamente, o perfil do interferente recuperado pelo RBL no modo de eluição e espectral, na amostra de teste 8 (—) e 9 (—) para a quantificação da FEN. Percebe-se que o perfil recuperado na amostra 8 para a quantificação da FEN está relacionado ao sinal da PUT indicando que o interferente foi recuperado corretamente no resíduo. No entanto, para a amostra de teste 9 foi recuperado um perfil correspondente ao sinal da FEN indicando que o sinal do analito estava no resíduo o que prejudicou a quantificação do mesmo, situação similar ocorreu quando a PUT foi quantificada. Isso ocorre devido ao deslocamento na retenção, correlação entre os perfis espectrais e ao fato de nessas amostras a diferença na concentração entre as duas ABs preparada para a mistura de teste ser mais pronunciada, como pode ser visto na **Tabela 1**.

Essa confusão pode ser mais bem identificada quando comparado o perfil original (**Figura 20 c**) com o recuperado na modelagem da FEN (**Figura 20a**) e da PUT (**Figura 20b**). Na amostra de teste 8 para a modelagem da FEN percebe-se, pelo tempo de retenção, no perfil da PUT recuperada como interferente, que o sinal é referente ao do conjunto de teste. No entanto, o sinal recuperado para a mesma amostra na modelagem da PUT é referente ao do conjunto de

calibração, indicando que, além da similaridade entre os espectros, variações no tempo de retenção também prejudicaram a modelagem.

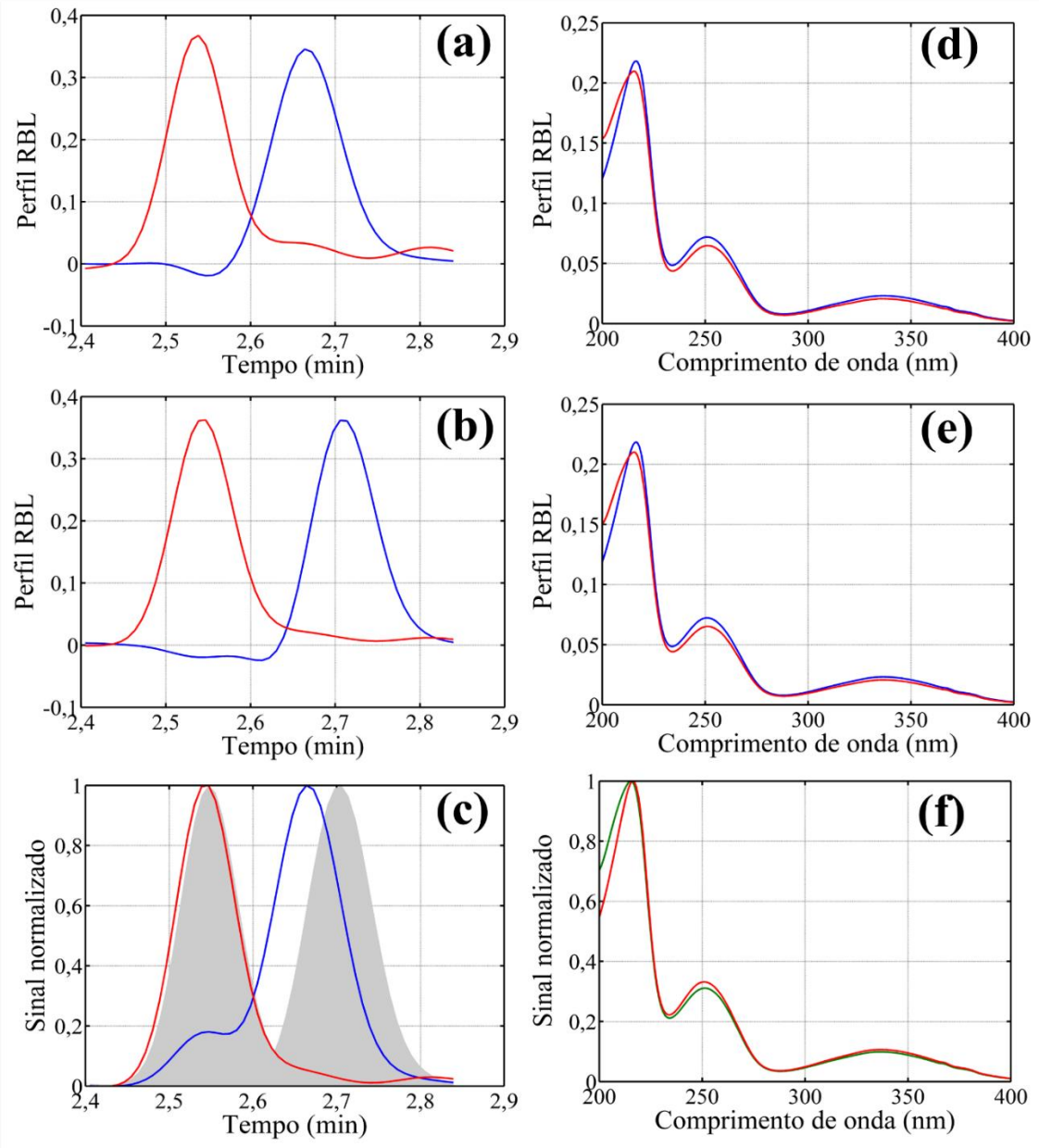

Figura 20 − Perfil recuperado pelo RBL no modo de eluição e comprimento de onda para a FEN (**a e d**) e PUT (**b e e**), para as amostras de teste 8 (—) e 9 (—). Em (**c**) é mostrado o sinal normalizado em 215 nm para as amostras de teste 8 (—) e 9 (—) e em cinza o perfil da amostra de calibração não alinhada e normalizada em 215nm da FEN e PUT no modo de eluição. Em vermelho e verde são mostrados o sinal no modo espectral da PUT e FEN (**f**), respectivamente.

Observando o sinal da amostra na **Figura 20 c**, percebe-se que a contribuição da PUT no sinal da amostra de teste 8 é maior do que a FEN, e na amostra 9 o inverso ocorre. Devido à similaridade dos espectros o algoritmo não modelou corretamente o analito. Isso pode ser observado pelo sinal representado por um pico sobreposto no cromatograma em azul entre 2,5 e

2,6 minutos na **Figura 20 c**, e em vermelho na mesma figura um sinal menor entre 2,6 e 2,7 minutos. Na amostra real, isso seria uma situação que comprometeria a quantificação exata, uma vez que é esperado um número maior de interferentes.

A literatura reporta que ao usar a estrutura de variáveis latentes nos dados desdobrados, é possível a quantificação dos analitos usando dados de cromatografia, desde que as variações na retenção não sejam tão pronunciadas (Borraccetti, Damiani, Olivieri, 2009, Montemurro, et al., 2016). Outros trabalhos relatam quantificações exatas na presença da deficiência de posto (Borraccetti, Damiani, Olivieri, 2009). No entanto, nenhum trabalho foi estudado com o UPLS-RBL quando ambos ocorrem. Assim, os dados de calibração foram, então, alinhados como mostrado na **seção 4.2**, alinhando as amostras de calibração com o *coshift*. Dessa forma, um desses problemas seria resolvido, permanecendo apenas a deficiência de posto.

Apenas uma VL é necessária para construir o modelo de calibração. Os valores de variância explicada foram todos superiores a 99% tanto para os cromatogramas registrados, quanto para o vetor de concentrações. Os valores de ajuste residual foram baixos e próximos aos resultados dos modelos desalinhados (**Tabela 5**), como pode ser visto na **Tabela 6**, indicando que o alinhamento não comprometeu a variância explicada nem o ajuste do modelo.

Tabela 6 − Variância explicada e ajuste residual para os modelos construídos para cada analito com uma VL com os dados alinhados.

Analito	Variância (x)[a]	Variância (y)[b]	Resíduo (x)[c]	Resíduo (y)[d]
Triptamina	99,92	99,95	1,72	0,18
Fenilamina	99,60	99,98	2,64	0,12
Putrescina	99,93	99,96	1,53	0,16
Cadaverina	99,78	99,70	1,98	0,45
Histamina	99,94	99,90	1,05	0,26

a: variância explicada da matriz de dados; **b**: variância explicada do vetor de resposta; **c**: ajuste residual da matriz de dados; **d**: ajuste residual do vetor de resposta.

Analisaram-se, então, os perfis recuperados pelo RBL para a quantificação da FEN e PUT novamente. As amostras de teste 8 e 9 apresentaram resultados satisfatórios para a PUT, uma vez que o perfil recuperado pelo RBL foi referente a FEN, como pode ser visto na **Figura 21 b** e **e**. Porém, outras amostras apresentaram um resultado não exato como foi o caso da 7 em que mostra que a PUT está no resíduo do próprio modelo novamente. Percebe-se que o alinhamento permitiu a correta distinção entre os sinais da FEN e da PUT para a amostra 8, no entanto, quando a diferença de concentração é maior, como nas amostras 7 e 9, o algoritmo ainda falha em

distinguir os sinais referentes a cada AB, provavelmente, devido à elevada correlação entre os perfis espectrais mostrados nas **Figuras 21 d-f**.

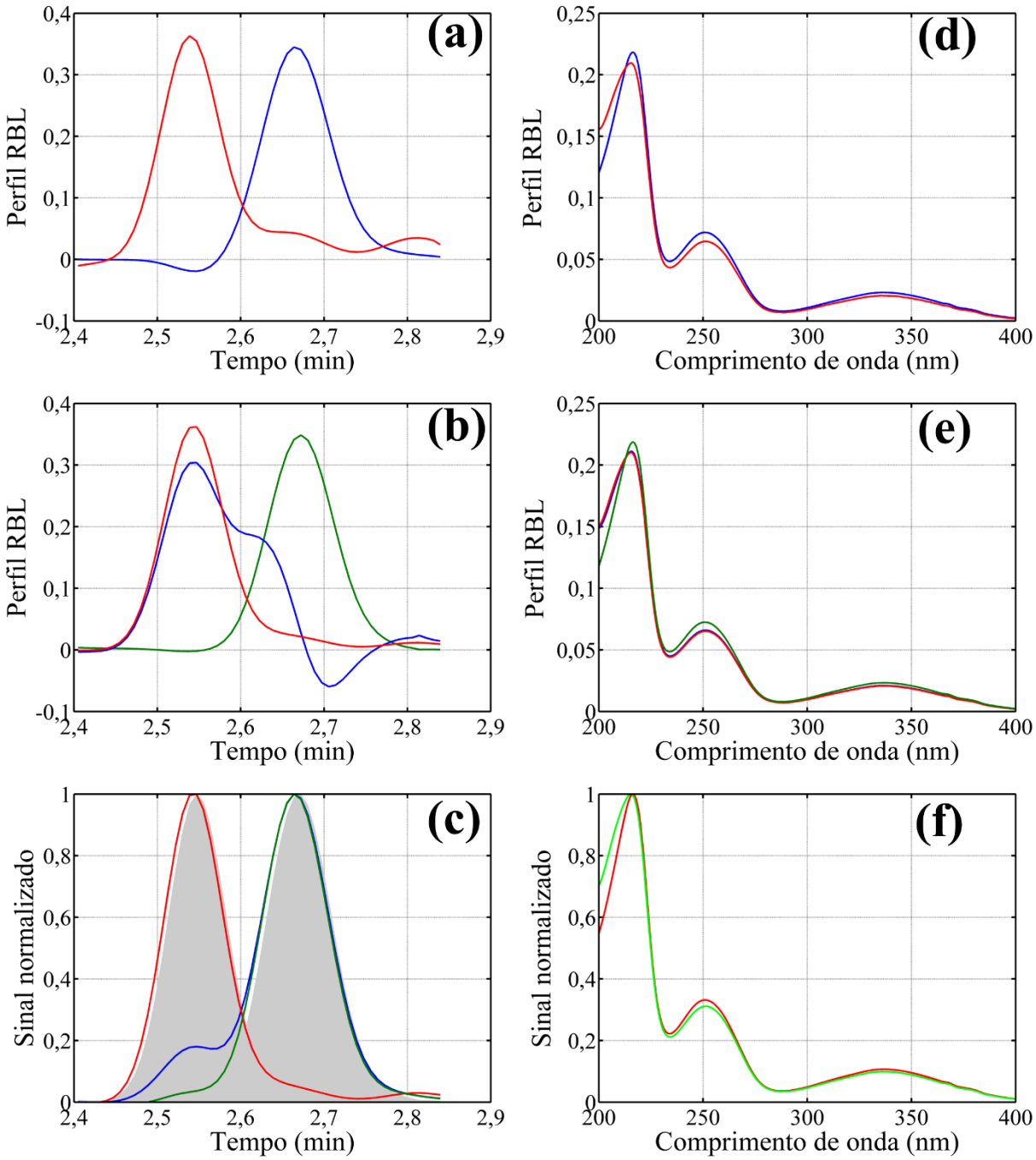

Figura 21 – Perfil recuperado pelo RBL no modo de eluição e comprimento de onda da FEN (**a e d**) para as amostras de teste 8 (—) e 9 (—) e da PUT (**b e e**) para as amostras de teste 7 (—), 8 (—) e 9 (—). Em (**c**) é mostrado o sinal normalizado em 215 nm para as amostras de teste 7 (—), 8 (—) e 9 (—) e em cinza o perfil da amostra de calibração alinhado e normalizado em 215 nm da FEN e PUT no modo de eluição. Em (**f**) são mostrados os espectros da PUT e FEN, respectivamente, em vermelho e verde.

No caso da FEN, os sinais dos analitos continuaram a serem confundidos da mesma forma que para os dados não alinhados. Ao analisar os sinais normalizados de calibração e teste na

Figura 21 c e **f**, percebe-se que os dados estão bem alinhados, mas, apesar disso, a modelagem com esse algoritmo não apresenta resultados satisfatórios quando mais de um constituinte com perfil espectral correlacionado é modelado na mesma faixa.

Apesar de alinhados, os dados apresentam elevada correlação entre os espectros, o que afetou os resultados do cálculo de sensibilidade analítica e, consequentemente, os valores de LOD e LOQ, como pode ser visto na **Tabela 7**. Percebe-se que os resultados só foram exatos para a modelagem em que há apenas um analito por região, onde não foi necessário aplicar a etapa de RBL. Na presença de outro constituinte com sinal muito correlacionado, o algoritmo apresenta resultados não exatos na quantificação, apesar de serem melhores do que com a modelagem dos dados não alinhados.

Tabela 7 − Métricas de desempenho analítico calculadas para os cinco modelos UPLS-RBL construído com o conjunto de calibração alinhado. As concentrações dos conjuntos de teste e de calibração variaram respectivamente nas faixas de 0,4 a 16 mg L^{-1} e 0,05 a 20 mg L^{-1} para todos os 5 analitos.

Métricas de desempenho UPLS-RBL	TRI	FEN	PUT	CAD	HIS
RMSEP (mg L^{-1})	0,86	2,34	6,26	0,46	0,94
REP (%)	8,6	23,5	62,8	4,6	9,4
Sensibilidade	724,37	34,22	71,88	485,62	564,76
Sensibilidade Analítica (L mg^{-1})	277,94	88,14	133,01	185,89	371,51
LOD min. (mg L^{-1})	0,02	0,13	0,10	0,03	0,02
LOD max. (mg L^{-1})	0,08	0,14	0,12	0,34	0,11
LOQ min. (mg L^{-1})	0,07	0,40	0,31	0,10	0,07
LOQ max. (mg L^{-1})	0,24	0,42	0,36	1,03	0,32

LOD min e LOD max: limite de detecção mínimo e máximo; **LOQ min e LOQ max**: limite de quantificação mínimo e máximo; esses limites foram calculados com base em erros do tipo 1 e 2 de acordo com (Ortiz, Sarabia, Sánchez, 2010; Allegrini, Olivieri, 2014); **RMSEP**: raiz quadrada do erro médio de predição (do inglês, *Root mean square error of prediction*); **REP**: Erro relativo de predição (do inglês, *Relative error of prediction*). Todas as métricas de desempenho foram calculadas de acordo com Olivieri e coautores (Olivieri, 2014; Bauza, Ibañez, Tauler, Olivieri, 2012; Allegrini, Olivieri, 2014).

Com base no conhecimento químico do sistema e do elevado resíduo da amostra de teste, sabe-se que na modelagem da FEN e da PUT é necessário 1 RBL. No entanto, ao se aplicar 1 RBL, os resultados não foram satisfatórios (**Tabela 7**) devido ao perfil RBL recuperado não condizer com o interferente. Os resultados de REP para a CAD e a HIS foram adequados, como mostrado na **Tabela 7**. No entanto, sabe-se que há uma contribuição da HIS na região modelada da CAD e vice-versa, como mostrado na **Figura 16**. Uma vez que a contribuição do interferente não é tão pronunciada em todas as amostras, o REP não foi tão afetado, no entanto, percebe-se

que ainda há sinal no resíduo que não foi incluído no modelo da CAD, fato este ressaltado pelo elevado valor de LOQ max. A presença de constituintes não modelados no resíduo altera o valor de escore, o que resulta em quantificações inexatas (**seção 2.5**). Na **Tabela 7** são mostrados os resultados do modelo UPLS aplicando a etapa RBL apenas a modelagem da FEN e da PUT.

Devido à presença do interferente na região modelada da CAD foi necessário aplicar uma etapa de RBL. No entanto, a exatidão foi comprometida resultando em um valor de REP de 159,4%. Isso ocorreu pelo mesmo motivo que para a PUT, a elevada correlação entre os espectros, ou seja, o perfil de pesos recuperado para o interferente contém o sinal da CAD, como pode ser visto na **Figura 22**.

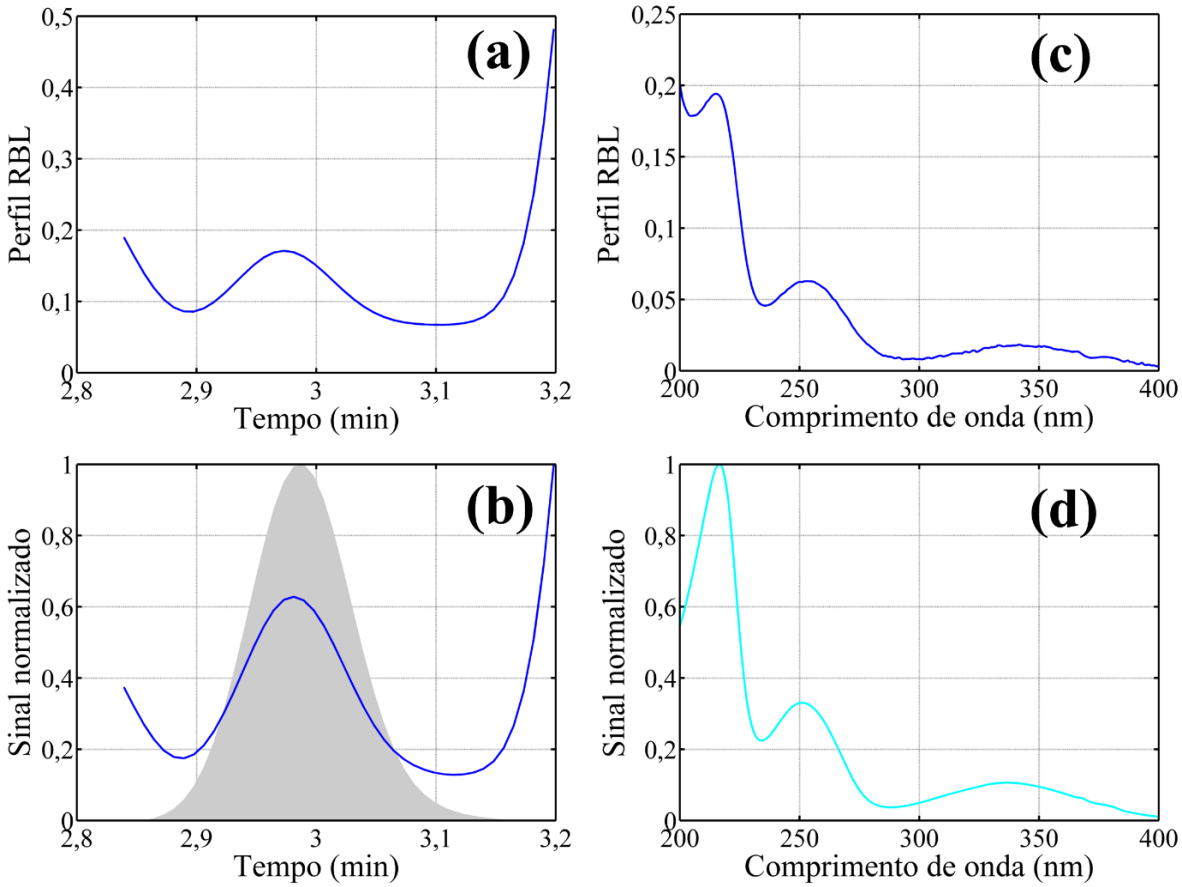

Figura 22 – Perfil dos pesos recuperado pelo RBL no modo de eluição e comprimento de onda para a CAD (**a** e **c**), para a amostra de teste 6 (—). (**b**) Sinal normalizado da CAD em 215 nm para a amostra de teste 6 (—) e em cinza o perfil da amostra de calibração de 20 mg L^{-1} alinhada e normalizada. (**d**) Espectros normalizado da CAD registrado no tempo de retenção para a amostra de calibração de 20 mg L^{-1} (—).

Na **Figura 22** é possível notar que o sinal recuperado do resíduo RBL em um dos modos é referente ao somatório dos cromatogramas da CAD e da HIS (**Figura 22 a**) e no outro modo é recuperado um peso (**Figura 22 c**) que contém informação referente ao espectro da CAD (**Figura 22 d**). Na **Figura 22 b** percebe-se pelo peso, indicado pelo perfil azul, que não só a contribuição

da HIS está presente no resíduo modelado pelo RBL, mas também o sinal da CAD. O sinal da CAD é eluído em 3 minutos e foi identificado ao sobrepor com o seu sinal registrado na amostra de calibração de 20 mg L^{-1} normalizada.

Devido ao sinal do analito ser identificado como interferente quando a etapa RBL é aplicada na região onde há constituintes com espectros correlacionados, o UPLS-RBL não se mostra eficaz para a análise desses dados, uma vez que é esperado um número maior de interferentes nas amostras de peixe.

4.5 Tratamento dos dados com o MCR-ALS

O MCR-ALS é um algoritmo de resolução de curvas como o PARAFAC. Apesar de ambos usarem o ALS na etapa de otimização dos perfis inicialmente estimados, são algoritmos bem distintos, sendo a principal dessas diferenças que o MCR-ALS faz deconvolução de uma matriz bilinear. Uma das consequências é a não unicidade do resultado final, pois o MCR-ALS é afetado por ambiguidade rotacional (**seção 2.6**). Por outro lado, os dados não precisam ser trilineares, ou seja, diferenças no tempo de retenção não afetam o MCR-ALS de forma significativa como no caso do PARAFAC e do UPLS-RBL (Jaumot, de Juan, Tauler, 2015; Escandar, Goicoechea, Peña, Olivieri, 2014; de la Peña, Goicoechea, Escandar, Olivieri, 2015). No caso desse trabalho, há a quebra da trilinearidade em dois modos: (**1**) eluição, característico dos dados de cromatografia; e (**2**) nos espectros advindos da derivatização das ABs.

Os dados gerados são de três vias e como o MCR-ALS faz deconvolução de uma matriz de dados e não de um tensor, como o PARAFAC, é necessário que os dados sejam aumentados. A forma de aumentar os dados depende da quebra da trilinearidade. Caso se aumente os dados no modo de eluição, o algoritmo entende que o perfil cromatográfico pode variar de amostra para amostra, enquanto o perfil dos espectros deve ser imutável e seletivo. Por esse motivo, o MCR-ALS é muito usado em dados de cromatografia em que a quebra de trilinearidade na eluição é frequente. Caso se aumente os dados no modo dos espectros, o algoritmo entende que os espectros podem mudar ou que não seja necessariamente seletivo, já o modo de eluição deve ser imutável e seletivo para distinguir os constituintes da amostra. Em ambos os casos, os dados são aumentados no modo coluna, uma vez que aumentar no modo linha indica a inclusão da informação referente a outro detector (de la Peña, Goicoechea, Escandar, Olivieri, 2015; Escandar, Goicoechea, Peña, Olivieri, 2014; Jaumot, Tauler, 2010; Jaumot, de Juan, Tauler, 2015).

Os dados deste estudo possuem particularidades que comprometem a modelagem com esse algoritmo em ambos os modos de aquisição: cromatográfico e espectral. Inicialmente, serão apresentados os resultados para a matriz aumentada no modo de eluição e espectral para os dados não alinhados e, em seguida, para os dados alinhados. Nos dados não alinhados é esperado que ao aumentar a matriz no modo de eluição seja resolvida a quebra de trilinearidade neste modo, no entanto, haverá deficiência de posto. Aumentando no modo espectral, mais de um pico deve ser recuperado no mesmo perfil.

Inicialmente, o MCR-ALS requer que seja definido o número de componentes distintos presentes na matriz de dados. Essa é uma etapa crucial da mesma forma que é para o PARAFAC, uma vez que resultados insatisfatórios podem ser obtidos caso se tente recuperar uma quantidade de componentes diferente do que realmente está presente na matriz de dados modelada. A forma mais comum é por meio de uma decomposição de valores singulares (SVD, do inglês *Singular Value Decomposition*).

Na **Tabela 8** são mostrados os autovalores para os dados não alinhados aumentados no modo de eluição. Define-se o número de componentes como sendo aquele que resulta em um autovalor elevado, onde em seguida esse valor reduz e se mantém constante ou com uma variação baixa. De acordo com os valores apresentados na **Tabela 8**, a escolha de um componente para explicar os casos parece ser adequada. Isso está de acordo com a informação química da região da TRI e da HIS, no entanto, sabe-se que para a região da FEN, PUT e CAD são necessários recuperar sinais referentes a dois constituintes.

Tabela 8 − Autovalores para a matriz de dados de teste e de calibração, não alinhados, aumentado no modo de eluição nas respectivas faixas selecionadas para cada AB.

AUTOVALOR					
Componentes	**TRI**	**FEN**	**PUT**	**CAD**	**HIS**
1	32.393,2	37.286,8	40.297,4	23.054,3	27.178,4
2	1.690,6	2.215,4	2.292,8	1.042,5	894,9
3	458,0	541,7	504,4	180,5	212,4
4	265,9	193,3	191,8	90,1	70,1
5	94,5	105,9	107,9	42,6	57,6

Outra forma de definir esse valor em um conjunto de dados é avaliando a variância explicada e o ajuste residual dos dados após a inclusão de um componente no modelo. Na **Tabela**

9 é mostrada a variância explicada para cada analito nas respectivas faixas selecionadas para os dados não alinhados aumentados no modo de eluição.

Tabela 9 − Variância explicada e ajuste residual dos dados não alinhados aumentados no modo de eluição para a decomposição MCR-ALS dos cinco analitos nas respectivas faixas selecionadas.

	TRI		FEN		PUT		CAD		HIS	
COMP*	VAR[a]	RES[b]	VAR[a]	RES[b]	VAR[a]	RES[b]	VAR[a]	RES[b]	VAR[a]	RES[b]
1	99,984	0,785	99,808	3,038	99,840	2,995	99,992	0,421	99,997	0,252
2	0,009	0,530	0,190	0,348	0,159	0,340	0,007	0,167	0,002	0,148
3	0,005	0,279	0,001	0,233	0,001	0,221	0,001	0,100	0,001	0,097
4	0,001	0,190	0,001	0,151	0,000	0,144	0,000	0,066	0,000	0,061
5	0,001	0,126	0,000	0,097	0,000	0,078	0,000	0,051	0,000	0,048

a: variância explicada da matriz de dados; b: ajuste residual da matriz de dados. * número de componentes.

Para a TRI, CAD e HIS a variância explicada é próxima de 100% e o ajuste residual não reduz de forma significativa indicando que o modelo se ajusta bem com 1 componente, nesse caso o posto químico é igual ao matemático. Já na modelagem da FEN e da PUT, apesar da variância explicada ser próxima de 100%, o ajuste residual é elevado e isso indica que o modelo não se ajusta bem com apenas 1 componente. Nesse caso, sabe-se que o posto químico é dois, pois há o sinal da FEN e da PUT no conjunto de teste. No entanto, a variância explicada foi próxima a 100% para a FEN e para a PUT devido à elevada correlação entre os espectros. É interessante analisar os autovalores em conjunto com a variância explicada e o ajuste residual, por ter mais informação de como será o ajuste dos dados após a modelagem. Essas ferramentas, em geral, funcionam bem e indicam o número correto de componentes presentes em amostras, em especial as sintéticas. Embora a variância explicada e os autovalores indiquem 1 componente nos dados, sabe-se que o sistema químico apresenta dois componentes na região da FEN e PUT e da CAD com parte da HIS. Isso indica problema de deficiência de posto nos dados aumentados no modo cromatográfico. No caso da CAD, a diferença no ajuste residual não foi muito expressiva, uma vez que o sinal da HIS na região da CAD é bem menor que o da PUT na região da FEN e vice-versa. A fim de visualizar o efeito da deficiência de posto no tratamento dos dados serão recuperados 1 e 2 componentes para PUT e FEN. Uma vez que a curva de calibração pseudo-univariada é construída apenas com o sinal recuperado do analito, o perfil que estiver no resíduo não compromete a exatidão quando o analito é recuperado corretamente, entretanto, as métricas de desempenho analítico serão comprometidas. Assim, para a CAD, foi recuperado

apenas um perfil, pois a elevada correlação com os espectros da HIS prejudica a exatidão da modelagem.

Após definir o número de componentes deve-se fazer uma estimativa inicial do perfil dos analitos. Uma forma é usando um algoritmo de seleção de variáveis puras, baseado no SIMPLISMA e chamado de PURE, outra é por análise de fatores evolutivos, e a terceira forma é usar o espectro registrado experimentalmente com um padrão puro (Jaumot, de Juan, Tauler, 2015).

Na maioria dos trabalhos com cromatografia, é usado o PURE por resultar em melhores estimativas iniciais dos perfis químicos de cada componente presente na amostra. Sendo assim, em todos os casos, seja aumentado no modo de eluição ou dos espectros, será feita a estimativa inicial usando o PURE no modo em que se espera maior seletividade, ou seja, no modo dos espectros quando a matriz for aumentada no modo de eluição e no modo de eluição quando for aumentada no modo dos espectros.

Para os perfis convergirem para o sinal do analito reduzindo a ambiguidade de rotação, as restrições aplicadas na otimização ALS foram: (**1**) organização de perfis por ordem de variável pura; (**2**) não negatividade nos modos de eluição e espectros; (**3**) unimodalidade no modo de eluição; (**4**) correspondência entre as espécies; e (**5**) normalização do modo não aumentado.

O resultado otimizado do perfil no modo de eluição com 1 e 2 componentes é mostrado na **Figura 23** e na **Tabela 10**. Com apenas um componente a TRI, CAD e a HIS recuperam corretamente os seus perfis. É interessante ressaltar aqui que apesar da HIS apresentar quebra de trilinearidade no modo de eluição a modelagem foi satisfatória, pois o MCR-ALS é um algoritmo bilinear e ao aumentar no modo de eluição esse problema foi contornado. Destaca-se, ainda, a contribuição da HIS na faixa modelada da CAD e vice-versa. O ajuste residual foi reduzido de forma significativa ao passar para a modelagem com dois componentes, apesar de a variância explicada ser maior que 99%, indicando que serão necessários dois componentes para explicar os dados modelados nessa região.

Na **Tabela 10** é mostrada a variância explicada e o ajuste residual dos dados otimizados pelo MCR-ALS. Como foi discutido, para o tratamento de dados do PARAFAC ainda há sinal da HIS na faixa modelada para CAD, como o sinal só é perceptível nas amostras onde o interferente está em maior concentração, o ajuste residual não foi tão elevado quanto para FEN e PUT.

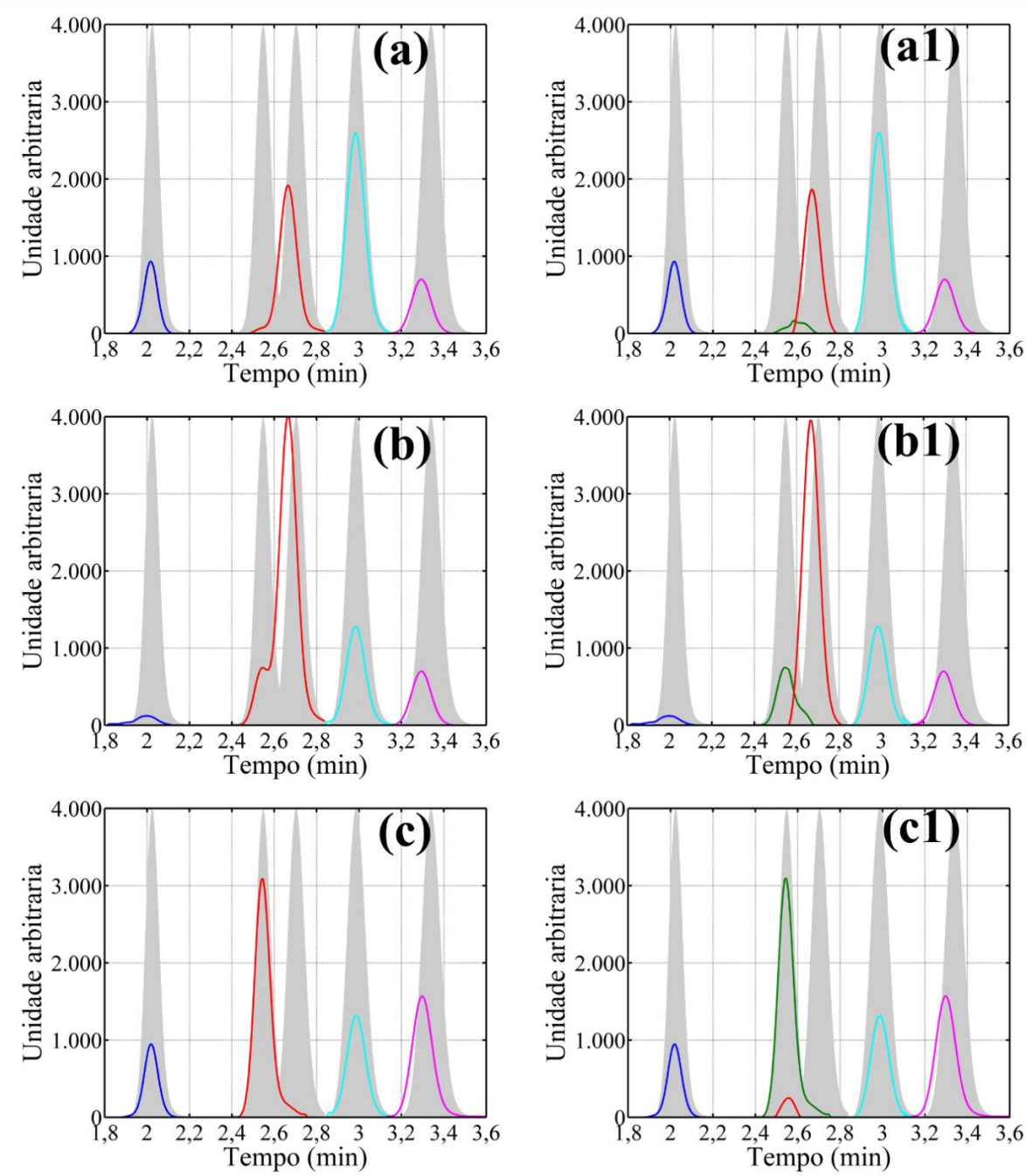

Figura 23 − Perfis recuperados no modo de eluição para as amostras de teste 7 (**a, a1**), 8 (**b, b1**) e 9 (**c, c1**) com um (**a-c**) e dois (**a1-c1**) perfis recuperados para os dados não alinhados. Apenas a TRI foi recuperado com um só perfil em ambos os casos. TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—).

Tabela 10 − Variância explicada e ajuste residual do resultado otimizado pelo MCR-ALS com os dados não alinhados.

	TRI		FEN		PUT		CAD		HIS	
COMP	**VAR[a]**	**RES[b]**	**VAR[a]**	**RES[b]**	**VAR[a]**	**RES[b]**	**VAR[a]**	**RES[b]**	**VAR[a]**	**RES[b]**
1	99,98	0,90	88,22	23,70	89,91	23,68	99,75	2,32	99,94	1,13
2			99,82	2,93	99,86	2,81	99,99	0,50	99,99	0,52

a: variância explicada da matriz de dados; b: ajuste residual da matriz de dados. * Número de componentes.

Para a FEN e CAD, sabe-se (**Tabela 9**) que serão necessários dois componentes para modelar corretamente os dados. Assim, quando modelado em ambos os casos com um único componente, o mesmo perfil é recuperado referente à soma dos sinais dos dois analitos. Ao se otimizar com dois componentes, são recuperados os perfis de cada analito separadamente. No entanto, quando um analito está em concentração baixa, como é o caso da PUT na amostra 9 (**Figura 23 c1**), o algoritmo erra devido à baixa seletividade no modo dos espectros causada pela elevada correlação. A deficiência de posto ocorre no MCR-ALS quando os espectros são idênticos, como é o caso da PUT e da CAD, ou em casos de correlação, como citado anteriormente para a FEN e PUT (Pinto, Díaz Nieto, Zón, Fernández, Araújo, 2016; Parastar, 2015). Além disso, mesmo quando se recupera o sinal dos analitos separadamente, o perfil não é gaussiano como esperado de um pico cromatográfico. A exemplo da amostra 8 (**Figura 23 b1**), na qual o perfil da FEN está claramente distorcido. Mesmo com perfis distintos, a similaridade dos espectros prejudica a correta recuperação dos perfis de eluição. Isso é agravado quando o número de componentes recuperados é maior, como no caso da amostra de peixe.

Pode-se inferir que, devido aos sinais instrumentais da FEN e PUT serem muito correlacionados eles são recuperados incorretamente dependendo do nível de concentração (**Figura 23 c1**), o que faz com que parte do sinal do analito não seja modelado. Mesmo com esses problemas, a exatidão da modelagem para o MCR-ALS foi melhor do que para os outros algoritmos, resultando em valores de REP de 7,9%, 12,7%, 13,15%, 4,62% e 16,11 % para TRI, FEN, PUT, CAD e HIS, respectivamente.

Valores insatisfatórios foram obtidos para a FEN, PUT e HIS, uma vez que o perfil recuperado não corresponde ao analito quando em baixas concentrações na presença do interferente, com concentrações altas devido à elevada correlação entre os espectros. Assim, pode-se concluir que, quando a concentração do analito está baixa e a do interferente, intensa não é necessário que os perfis espectrais sejam idênticos para que o MCR-ALS sofra com a deficiência de posto. Isso pode prejudicar a quantificação nas amostras de peixe que pode conter outros constituintes não esperados em concentrações mais elevadas.

Para resolver esse problema, as amostras de calibração foram alinhadas com o *coshift* e modeladas com o MCR-ALS aumentando-se no modo espectral. Também foi testado usar a restrição de trilinearidade, uma vez que após o alinhamento os dados serão trilineares, o que resultará em perfis sem ambiguidade de rotação. Inicialmente, será mostrado o MCR-ALS com os dados alinhados e com restrição de trilinearidade e, em seguida, aumentado no modo espectral. No caso em que a restrição de trilinearidade não é aplicada, as bandas de confiança serão calculadas e os perfis recuperados comparados para averiguar se há ambiguidade rotacional.

Antes de incluir na modelagem que os dados são trilineares, deve-se verificar a eficiência do alinhamento. No PARAFAC, isso foi feito avaliando-se a consistência do core, a variância explicada e a falta de ajuste do modelo, uma vez que o modelo é essencialmente trilinear. Isso não foi testado no UPLS-RBL. Já no MCR-ALS, isso pode ser testado de duas formas. A primeira é (**1**) construindo um modelo MCR-ALS com a restrição de trilinearidade, e avaliando a variância explicada e a falta de ajuste do modelo construído com os dados na forma bilinear e trilinear. Na **Tabela 11** são mostrados os valores de variância explicada e falta de ajuste para os modelos bilineares e trilineares para os mesmos dados. Observa-se que os valores são idênticos para ambos os modelos. Assim, pode-se afirmar que os dados cumprem com a restrição de trilinearidade. A segunda forma é (**2**) comparando os autovalores dos dados aumentados em ambos os modos, cromatográfico e espectral. Na **Tabela 12** são mostrados os autovalores para os dados aumentados no modo espectral e de eluição. O número de componentes necessários para a modelagem está destacado em vermelho.

Tabela 11 – Variância explicada e falta de ajuste dos modelos bilineares e trilineares construídos para as ABs com os dados alinhados.

	TRI		FEN		PUT		CAD		HIS	
Modelo	**VAR[a]**	**LOF[b]**	**VAR[a]**	**LOF[b]**	**VAR[a]**	**LOF[b]**	**VAR[a]**	**LOF[b]**	**VAR[a]**	**LOF[b]**
Bilinear	99,62	6,19	99,37	7,97	99,50	7,08	99,51	7,0161	99,39	7,81
Trilinear	99,62	6,19	99,37	7,97	99,50	7,08	99,51	7,0161	99,39	7,81

a: Variância explicada do modelo; b: LOF: Falta de ajuste do modelo, do inglês *lack of fit.*

Na **Tabela 12** é perceptível que o número de componentes necessários para explicar os dados é igual ao número de analitos presentes neles, ou seja, o posto químico é igual ao matemático tanto aumentando no modo de eluição quanto no modo espectral. Pode-se chegar à mesma conclusão do número de componentes inspecionando a **Figura 24**, que representa graficamente o resultado da **Tabela 12**. A necessidade de dois componentes para explicar os dados da FEN e PUT é observada, em especial, se aumentar os dados no modo espectral.

Ao aumentar no modo dos espectros, o número de componentes presentes na mistura é mais perceptível, pois a seletividade estará nas diferenças na retenção observada no cromatograma, em que os sinais dos analitos são menos correlacionados do que no modo dos espectros. Espera-se resolver o problema de deficiência de posto apenas ao tratar os dados aumentados no modo dos espectros e com restrição de trilinearidade, por isso será mostrado o tratamento dos dados apenas para esses dois casos.

Tabela 12 – Autovalores para os dados alinhados das ABs aumentados no modo de eluição e dos espectros de 1 a 5 componentes.

AUTOVALOR						
Modo aumentado	**Componentes**	**TRI**	**FEN**	**PUT**	**CAD**	**HIS**
Eluição	1	**32.005,0**	37.687,2	40.672,8	**23.087,2**	**27.996,4**
	2	304,5	**1.643,5**	**1.620,5**	191,8	120,3
	3	235,2	141,0	141,0	66,1	65,5
	4	107,0	95,7	90,7	36,9	44,3
	5	73,6	61,6	65,4	21,1	21,8
Espectro	1	**31.946,9**	32.997,1	37.159,8	**23.018,5**	**27.911,6**
	2	1.851,1	**18.135,3**	**16.423,1**	1.387,7	1.958,8
	3	510,1	1.903,4	2.125,9	1065,0	824,6
	4	397,4	1.450,9	955,8	323,4	410,6
	5	156,4	914,7	875,8	182,2	253,3

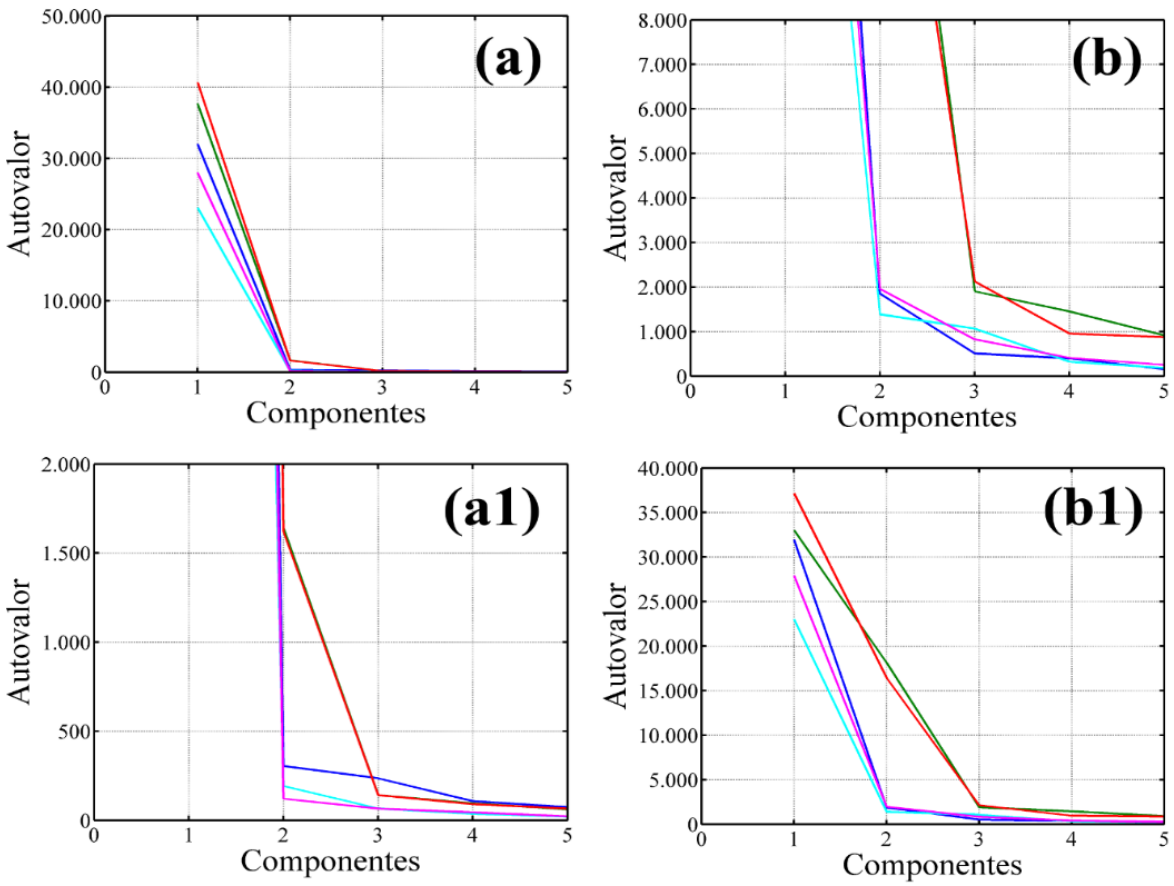

Figura 24 – Autovalor versus número de componentes para os dados aumentados no modo de eluição (**a**), zoom do gráfico a (**a1**), modo espectral (**b**) e zoom gráfico b (**b1**). TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—).

Como mostrado na **Figura 16**, espera-se que sejam necessários 2 componentes para a região da CAD e da HIS. No entanto, ao avaliar os autovalores na **Tabela 12**, e a variância explicada e o ajuste residual na **Tabela 13**, não é possível inferir que serão necessários dois componentes. Isso ocorre porque o sinal da CAD é muito mais intenso do que o da HIS nessa região. Sendo assim, os cálculos foram realizados com 1 e 2 componentes para a região da CAD e de acordo com o perfil recuperado foi escolhido o número de constituintes adequados.

Tabela 13 − Variância explicada e ajuste residual dos dados alinhados aumentados no modo espectral para a decomposição MCR-ALS dos cinco analitos nas respectivas faixas selecionadas.

	TRI		FEN		PUT		CAD		HIS	
COMP	VAR	RES	VAR	RES	VAR	RES	VAR	RES	VAR	RES
1	99,62	3,79	76,42	33,86	92,97	14,44	99,40	3,64	99,39	3,72
2	0,34	1,32	23,11	4,83	6,92	1,85	0,36	2,33	0,49	1,65
3	0,02	0,88	0,26	3,30	0,07	1,11	0,21	0,18	0,09	0,86
4	0,02	0,400	0,15	1,86	0,02	0,70	0,02	0,43	0,02	0,48
5	0,00	0,25	0,06	0,66	0,01	0,36	0,01	0,24	0,01	0,19

Na **Figura 25** é possível verificar que ao recuperar com um componente, o perfil da CAD apresenta um pouco do sinal da PUT e parte do sinal da HIS. Isso não foi observado ao tratar os dados com o PARAFAC, mas explica o sinal recuperado no peso do RBL no modo de eluição apresentado na **Figura 24 a** e **b**.

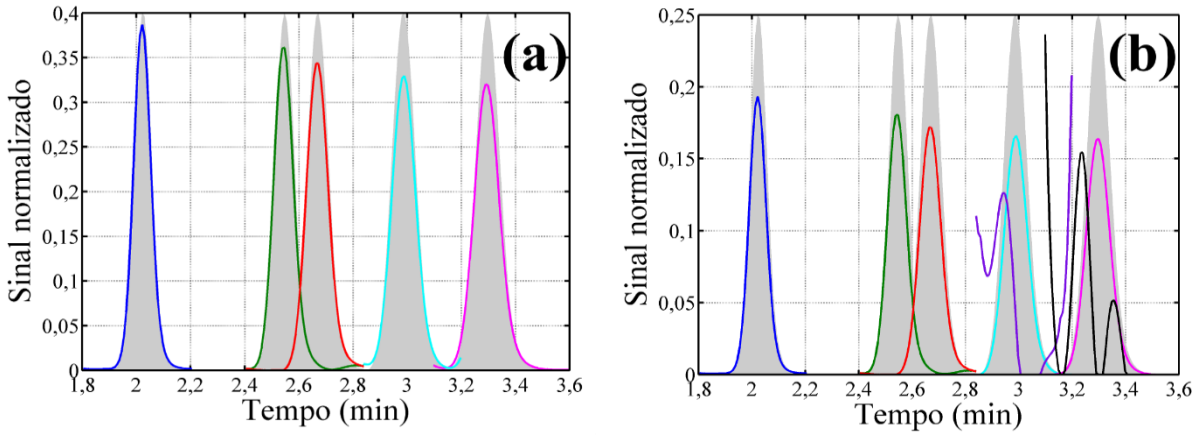

Figura 25 − Perfil de eluição recuperado para os dados alinhados e aumentados no modo espectral, com 1 componente para TRI, CAD e HIS e 2 para FEN e PUT (**a**) e com 2 componentes para CAD e HIS (**b**), mantendo iguais o dos outros analitos. Em cinza o perfil normalizado da amostra de calibração mais concentrada. TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—).

Entretanto, ao recuperar dois componentes na região de modelagem da CAD, o sinal do analito se encontra em um perfil e no outro, o da PUT e da HIS somados, restando apenas o ruído instrumental no resíduo. Foi possível identificar a presença do sinal da PUT e HIS apenas quando se modelou com o MCR-ALS e não, previamente, com os autovalores, isso ocorreu devido à contribuição desses analitos na faixa modelada ser muito menor que o sinal da CAD. Situação similar pode ser observada para a HIS na **Figura 25 b**, espera-se recuperar dois perfis nessa região. Como pode ser visto na **Figura 16 d**, é interessante notar que apenas o MCR-ALS distinguiu esse sinal baixo na CAD na região da HIS.

As métricas de desempenho para os modelos da TRI recuperado com um perfil e da FEN, PUT, CAD e HIS com dois perfis são mostradas na **Tabela 14**, na qual pode ser visto que os resultados são exatos para a modelagem de todos os analitos, indicado pelos valores de REP inferiores a 10%. O valor de sensibilidade analítica está alto o suficiente para manter o LOQ abaixo de 0,4 mg L^{-1} que é a menor concentração do conjunto de teste.

Tabela 14 − Métricas de desempenho analítico calculadas para os cinco modelos MCR-ALS com o conjunto de calibração alinhado, a restrição de trilinearidade não foi aplicada.

Métricas de desempenho MCR-ALS	**TRI (1)**	**FEN (2)**	**PUT (2)**	**CAD (2)**	**HIS (2)**
RMSEP (mg L^{-1})	0,79	0,44	0,49	0,40	0,54
REP (%)	8,0	4,4	4,9	4,0	5,4
Sensibilidade (L mg^{-1})	436,30	322,61	473,11	297,13	258,88
Seletividade	1,00	0,99	0,99	0,95	0,72
Sensibilidade Analítica (L mg^{-1})	115,93	67,02	76,45	121,41	125,12
LOD (mg L^{-1})	0,03	0,05	0,04	0,03	0,04
LOQ (mg L^{-1})	0,09	0,15	0,13	0,08	0,12

Entre parênteses é mostrado o número de constituintes estimados. **LOD**: limite de detecção; **LOQ**: limite de quantificação; **RMSEP**: Raiz quadrada do erro médio de predição (do inglês, *Root mean square error of prediction*); **REP**: Erro relativo de predição (do inglês, *Relative error of prediction*). Todas as métricas de desempenho foram calculadas de acordo com Olivieri e coautores (Olivieri, 2014; Bauza, Ibañez, Tauler, Olivieri, 2012).

Os valores de REP para modelos com restrição de trilinearidade foram de 8,0%, 4,3%, 5,2%, 4,8% e 5,5% para TRI, FEN, PUT, CAD e HIS, respectivamente. Essa similaridade na exatidão já é esperada, uma vez que os dados alinhados são trilineares, como mostrado anteriormente. Os resultados foram muito parecidos, como pode ser visto também na **Figura 26**, onde são mostrados os perfis recuperados pelo MCR-ALS com e sem restrição de trilinearidade para os modos de eluição e dos espectros. Os sinais são sobrepostos e não é possível distinguir o

perfil recuperado com a restrição de trilinearidade, resultando em uma ambiguidade rotacional quase nula para os dados aumentados no modo espectral.

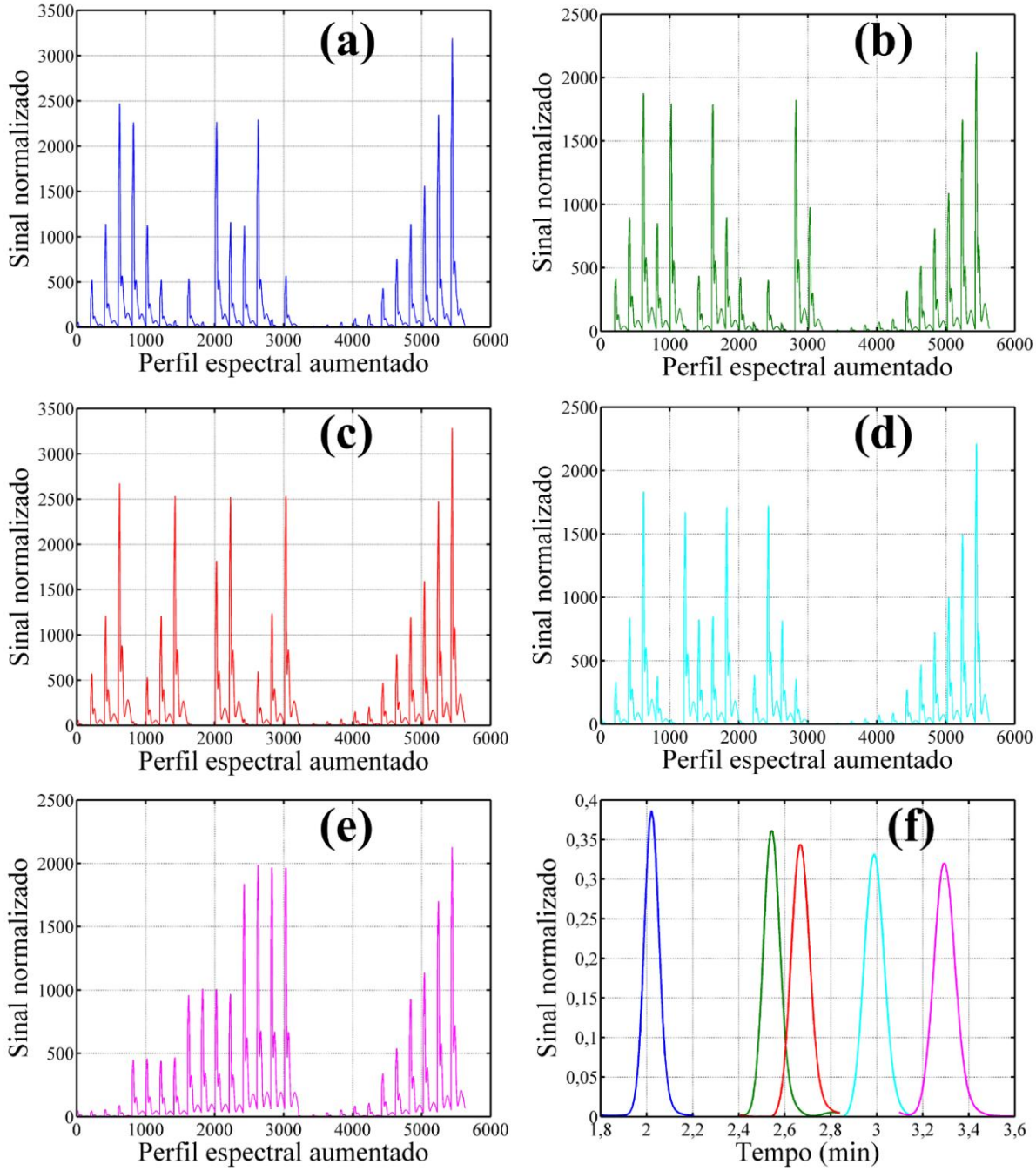

Figura 26 – Perfil aumentado no modo espectral para o modelo de TRI (**a**), FEN (**b**), PUT (**c**), CAD (**d**) e HIS (**e**) com (linha preta) e sem (linha colorida) restrição de trilinearidade. Em (**f**) são mostrados os perfis normalizados do modo de eluição para os modelos com (linha preta) e sem (linha colorida) restrição de trilinearidade. TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—).

Além da similaridade dos perfis recuperados em ambos os modos, a variância explicada e a falta de ajuste foram muito similares entre os modelos com restrição de trilinearidade e sem. Ambas as abordagens irão apresentar os mesmos resultados, uma vez que ao alinhar os picos

cromatográficos e tratar os dados aumentados no modo dos espectros com o MCR-ALS, a quebra de trilinearidade em ambos os modos é resolvida, resultando em modelos similares aplicando ou não a restrição de trilinearidade.

Na **Tabela 15** é mostrada a falta de ajuste e a variância explicada para o modelo MCR-ALS com e sem a restrição de trilinearidade, em que se percebe que ambos os ajustes são muito parecidos. Apesar do uso de restrições prejudicar a falta de ajuste e a variância explicada do modelo por reduzir o grau de liberdade, a variação, neste caso, foi mínima uma vez que os mesmos resultados de perfil recuperado em ambos os modos foram obtidos.

Tabela 15 − Efeito da restrição de trilinearidade no modelo das amostras de teste para os dados alinhados.

Analito*	Iterações [a]	Iterações [b]	LOF (%)[a]	LOF (%)[b]	VAR (%)[a]	VAR (%)[b]
TRI (1)	3	3	6,19	6,17	99,62	99,62
FEN (2)	6	7	6,89	6,86	99,52	99,53
PUT (2)	9	10	6,26	6,24	99,61	99,61
CAD (2)	3	3	5,53	5,52	99,69	99,70
HIS (1)	3	3	7,80	7,79	99,39	99,39

* Entre parênteses é mostrado o número de componentes do modelo.
VAR: Variância explicada pelo modelo. LOF: Falta de ajuste do modelo, do inglês *lack of fit.*
[a] Restrições: normalização; não-negatividade; unimodalidade; correspondência entre as espécies; trilinearidade.
[b] Restrições: normalização; não-negatividade; unimodalidade; correspondência entre as espécies.

A calibração usando o MCR-ALS foi feita usando a área sobre o perfil espectral de cada analito, uma vez que o sinal normalizado foi o do modo de eluição. Na **Figura 5S** são mostrados os perfis recuperados pelo MCR-ALS para os modos de eluição normalizado e espectral aumentado e sobrepostos. Apesar de não ser algo comum para dados cromatográficos, matematicamente, a área sob o perfil espectral recuperado pelo MCR-ALS é proporcional à concentração do analito e pode ser usada para predizer a sua concentração (Culzoni, et al., 2008; Pinto, Díaz Nieto, Zón, Fernández, Araújo, 2016). A ambiguidade rotacional foi calculada para os perfis recuperados do modelo com e sem restrição de trilinearidade, os valores da norma de Frobenious são mostrados na **Tabela 2S**.

Apesar de essa restrição garantir a propriedade de unicidade, como no PARAFAC, a comparação com os outros algoritmos será feita com o MCR-ALS nos dados aumentados no modo espectral. Optou-se por isso devido às métricas de desempenho analítico para calibração de segunda ordem estarem implementadas apenas na interface MVC2 (Olivieri, Wu, Yu, 2009), e a restrição de trilinearidade do MCR-ALS não está implementada nesta interface.

4.6 Considerações parciais

A necessidade de derivatização gera espectros muito parecidos, prejudicando a seletividade na modelagem com algoritmos que não sejam desenvolvidos para tratar dados com dependência linear. Inicialmente, decidiu-se modelar os dados divididos em faixas mostradas na **Tabela 1S**, cada uma contendo a menor quantidade de analitos possível, por não comprometer a seletividade e, consequentemente, a sensibilidade analítica.

Além do problema de dependência linear, em uma eluição cromatográfica podem ocorrer deslocamentos de tempo de retenção, que levam à quebra da trilinearidade do conjunto de dados quando se decompõe um tensor de três vias de um conjunto de amostras. Se modelar dados com essas duas características já é uma tarefa complexa, atingir a vantagem de segunda ordem com interferentes cujo perfil espectral apresentam uma elevada sobreposição com o do analito, é impossível com as abordagens existentes na literatura.

Poucos são os algoritmos capazes de tratar dados de cromatografia líquida cujo deslocamento do tempo de retenção leva à quebra de trilinearidade. O PARAFAC 2 é uma modificação do PARAFAC desenvolvido para dados de cromatografia líquida que permite a quebra de trilinearidade. No entanto, não há parâmetros de desempenho analítico desenvolvido para este algoritmo dificultando a sua avaliação, além disso, mesmo permitindo a quebra de trilinearidade, ele não resolve a dependência linear. O PARALIND é outra modificação do PARAFAC que foi criado para dados com problemas de dependência linear, no entanto, ele não resolve a quebra de trilinearidade no modo de eluição e, também, não possui parâmetros de desempenho implementados dificultando a avaliação da qualidade dos modelos gerados.

O UPLS-RBL foi testado e apresentou resultados satisfatórios em dados cuja sobreposição espectral é acentuada (Borraccetti, Damiani, Olivieri, 2009). No entanto, não resulta em uma modelagem satisfatória dos dados com dependência linear causado por espectros idênticos presente nos dados desta tese, e ainda não resolve a quebra de trilinearidade no modo de eluição, como mostrado para os resultados desta tese.

O MCR-ALS pode resolver tanto a quebra da trilinearidade no modo de eluição quanto a dependência linear. Porém, é necessário aumentar a matriz no modo que apresenta esse problema (Culzoni, et al., 2008; Arancibia, Damiani, Escandar, Ibañez, Olivieri, 2012; Parastar, 2015; Pinto, Díaz Nieto, Zón, Fernández, Araújo, 2016). Nos dados desse trabalho há artefatos, nos dois modos, que impossibilitam o MCR-ALS atingir a vantagem de segunda ordem e recuperar corretamente os perfis dos analitos.

Previamente a análise com MCR-ALS é necessário resolver um desses problemas antes da modelagem e, em seguida, usar um algoritmo que seja capaz de resolver o outro problema. Uma alternativa presente na literatura é o uso de algoritmos de alinhamento dos picos cromatográficos (Tomasi, Savorani, Engelsen, 2011). Ao alinhar os dados resolve-se a quebra de trilinearidade no modo de eluição. Já a quebra no modo espectral pode ser resolvida com o MCR-ALS e com o PARAFAC, no entanto, a seletividade do PARAFAC é comprometida, como mostrado nos resultados desta tese, e assim tornando o valor de LOD e LOQ não condizente com o valor de concentração previsto pelo algoritmo.

Para esses dados, a única forma encontrada para uma modelagem satisfatória das amostras de teste e de peixe foi alinhar o conjunto de calibração com o *coshift* e modelar com o MCR-ALS aumentado no modo espectral. Apesar de ser uma forma pouco usual para o MCR-ALS, há artigos na literatura que fazem uso dessa abordagem (Culzoni, et al., 2008), embora, neste caso, não apresentasse quebra de trilinearidade no outro modo.

Com os dados bem alinhados pode-se aplicar a restrição de trilinearidade no MCR-ALS. No entanto, nas análises das amostras de peixes não será abordado o uso dessa restrição devido ao fato de ela reduzir a variância explicada e aumentar a falta de ajuste, como visto na **Tabela 15**, embora apresente perfis recuperados praticamente idênticos mostrados na **Figura 26**. Além disso, sem a restrição de trilinearidade as concentrações preditas foram praticamente idênticas, e, uma vez que se sabe o espectro e o tempo de retenção do analito, é fácil identificar a ambiguidade rotacional mesmo sem realizar o teste de f_{min} e $f_{max,}$ o que faz com que a unicidade não seja um fator limitante. Outra razão foi devido aos parâmetros de desempenho não estarem implementados na interface MCR de Tauler e colaboradores (Jaumot, de Juan, Tauler, 2015), e a ausência da restrição de trilinearidade na interface MVC2 de Olivieri e colaboradores (Olivieri, Wu, Yu, 2009).

Uma vez que para esses dados apenas o MCR-ALS nos dados alinhados e aumentado no modo espectral apresenta resultados satisfatórios nas amostras de teste, optou-se por analisar as amostras de peixe apenas com esta abordagem e aplicando na otimização da modelagem MCR-ALS as seguintes restrições: (1) não negatividade no modo espectral e de eluição; (2) unimodalidade no modo de eluição; (3) correspondência entre as espécies; e (4) normalização do perfil no modo de eluição. Para identificar se há ambiguidade rotacional, os perfis otimizados serão avaliados usando a interface MCR-*bands* (Jaumot, Tauler, 2010) e onde será comparado o perfil recuperado em ambos os modos com o do padrão puro.

4.7 Quantificação das aminas biogênicas em amostras de peixe

A quantificação das ABs nas amostras de peixe seguiu o procedimento descrito na **seção 3.5**. No procedimento utilizado para a análise das ABs duas diluições são feitas: (1) ao final da extração as aminas biogênicas contidas nas amostras foram diluídas dez vezes e (2) a diluição devido a derivatização. Esta segunda diluição pode ser desconsiderada nos cálculos, pois ela é realizada no conjunto de calibração, ou seja, ela já está sendo considerada ao construir o modelo com o conjunto de calibração.

Na **Figura 27** é mostrado um cromatograma das quatro amostras de peixe com as quatro fortificações realizadas segundo o procedimento descrito na **seção 3.6**. A amostra de 20 mg L^{-1} do conjunto de calibração antes e após o alinhamento foi sobreposta a fim de comparação. Pela análise da **Figura 27** e da **Figura 16**, percebe-se que o deslocamento não foi igual ao do conjunto de teste, sendo necessário alinhar, novamente, o conjunto de calibração para cada amostra.

Foi possível alinhar os dados com facilidade, pois o cromatograma possui poucos constituintes não esperados na região de eluição dos analitos. Isso já era esperado, principalmente, por dois motivos, um é que a força da fase móvel é bem elevada, sendo ela constituída por 73% de acetonitrila, então muitos constituintes que não serão analisados serão eluídos no tempo morto da coluna. Outro motivo, é que os analitos foram extraídos e derivatizados, então é esperado que na amostra os únicos constituintes que apareçam depois do DNS (ver **Figura 13 a e b**) sejam as aminas derivatizadas e outros constituintes com grande quantidade de carbono em sua composição. Essas aminas possuem elevado teor de carbono, e consequentemente tem menor afinidade com solventes polares como ácidos inorgânicos, a exemplo do ácido perclórico usado na extração das ABs abordadas nesta tese. Assim, os únicos possíveis interferentes na análise serão as aminas capazes de derivatizar com o DNS e que não foram contempladas no conjunto de calibração, não sendo necessariamente biogênicas, porém geram espectros similares, ou até idênticos ao dos analitos contemplados neste estudo.

Interferentes foram identificados em todas as amostras, alguns com intensidades maiores que outros. Na **Figura 28** são indicados os interferentes presentes na amostra de *Seriola dumerili*. Os cromatogramas exibidos foram registrados em 215nm com a amostra de 20 mg L^{-1} em cinza, e na superfície de contorno com fortificações em 4 níveis de concentração. Os mesmos cromatogramas das amostras de *Salminus maxillosus*, *Scomber scombrus*, e *Lutjanus analis* são mostrados respectivamente nas **Figuras 6S, 7S** e **8S** no apêndice.

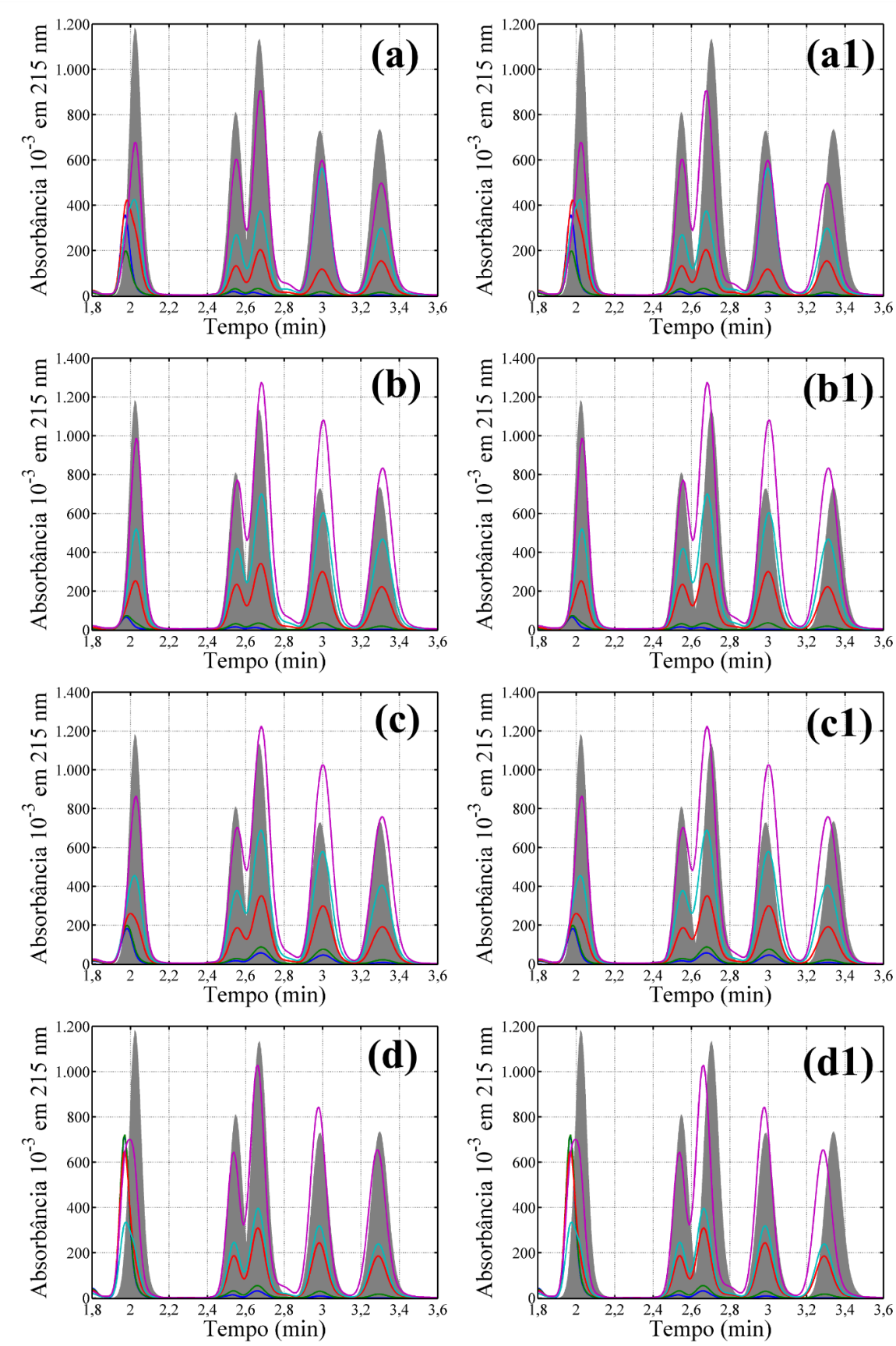

Figura 27 − Cromatogramas registrados em 215 nm para as amostras de *Salminus maxillosus* (**a e a1**), *Scomber scombrus* (**b e b1**), *Seriola dumerili* (**c e c1**) e *Lutjanus analis* (**d e d1**), sem fortificação (—) e fortificadas no nível 1 (—), 2 (—), 3 (—) e 4 (—), segundo o procedimento descrito na **seção 3.6**. Em cinza é mostrado a amostra de calibração de 20 mg L^{-1} alinhada (**a-d**) e sem alinhamento (**a1-d1**).

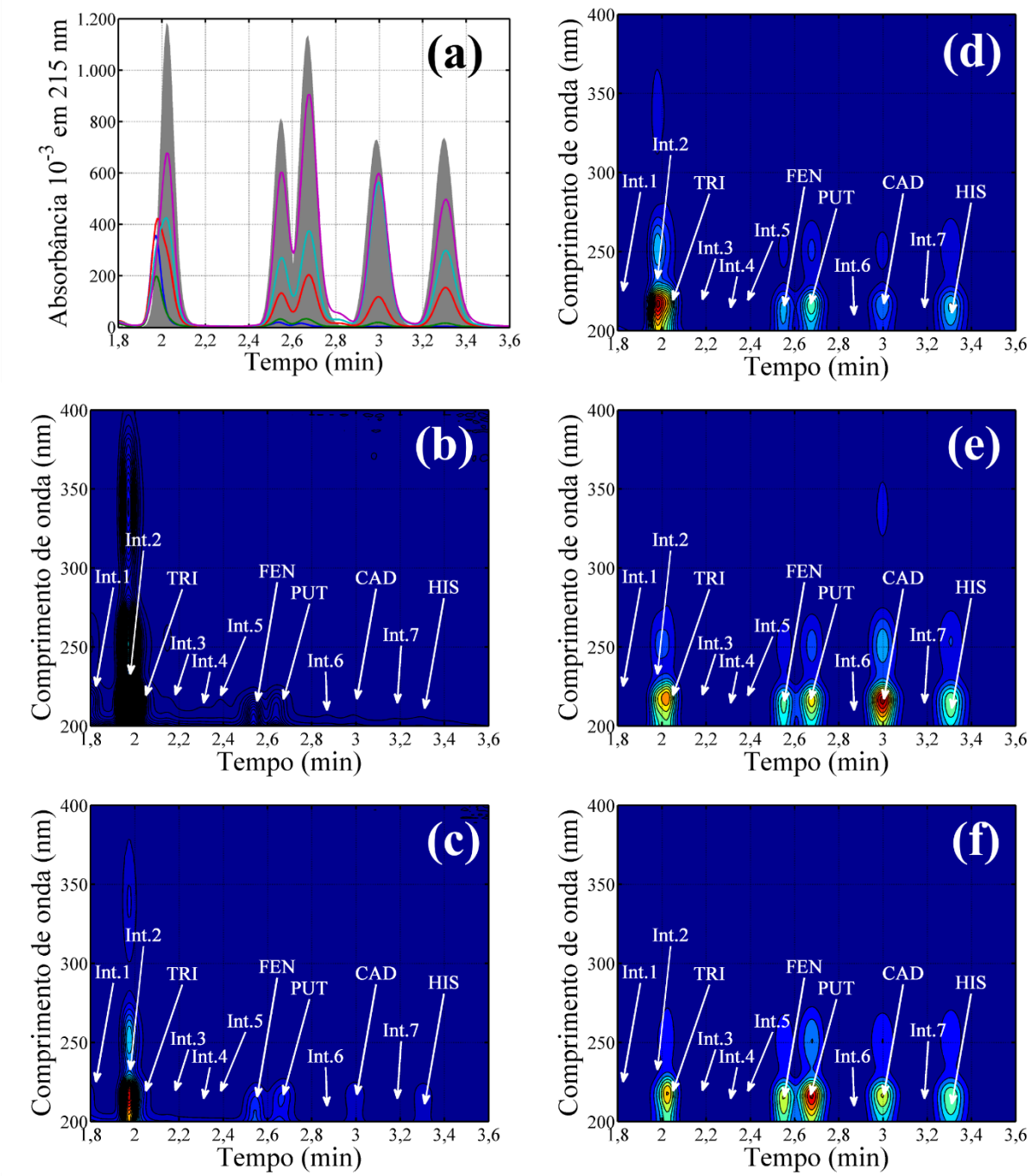

Figura 28 – Cromatograma em 2 D (**a**) e em superfície de contorno da amostra de *Seriola dumerili* sem fortificação (—) (**b**) e fortificadas no nível 1 (—) (**c**), 2 (—) (**d**), 3 (—) (**e**) e 4 (—) (**f**), segundo o procedimento descrito na **seção 3.6**. Os analitos e interferentes (**INT**) são apontados com setas. Em (**a**) o pico cinza representa o sinal do analito na concentração de 20mg L^{-1}.

O sinal do interferente da TRI é bem intenso e facilmente notado nos cromatogramas apresentados, na **Figura 28**, percebe-se que o espectro desse constituinte é parecido com o do analito e a coeluição é quase total, o que dificulta atingir a vantagem de segunda ordem com outras abordagens. Essa interferência é percebida devido a um alargamento do pico da TRI, a

quebra da assimetria do pico visto nos níveis de fortificação 2 e 3 (**Figura 28 c** e **d**) e pelo teste de pureza de pico. As setas nas **Figuras 28**, **6S**, **7S** e **8S** foram mantidas no mesmo local para melhor comparação com as outras fortificações mostradas nas superfícies de contorno de cada amostra e, assim, ser possível identificar melhor analitos e interferentes.

Nas **Figuras 28 a-d**, **30**, **6S**, **7S** e **8S**, percebe-se que em todas as quatro amostras há um interferente no sinal da TRI, no entanto, não é fácil perceber as interferências dos outros analitos. As pequenas diferenças no tempo de retenção que podem ser vistas tanto na **Figura 29 a-d** quanto na **Figura 28** apresentam seletividade o suficiente para o MCR-ALS alcançar a vantagem de segunda ordem almejada neste trabalho, em que os dados foram alinhados e aumentados no modo dos espectros.

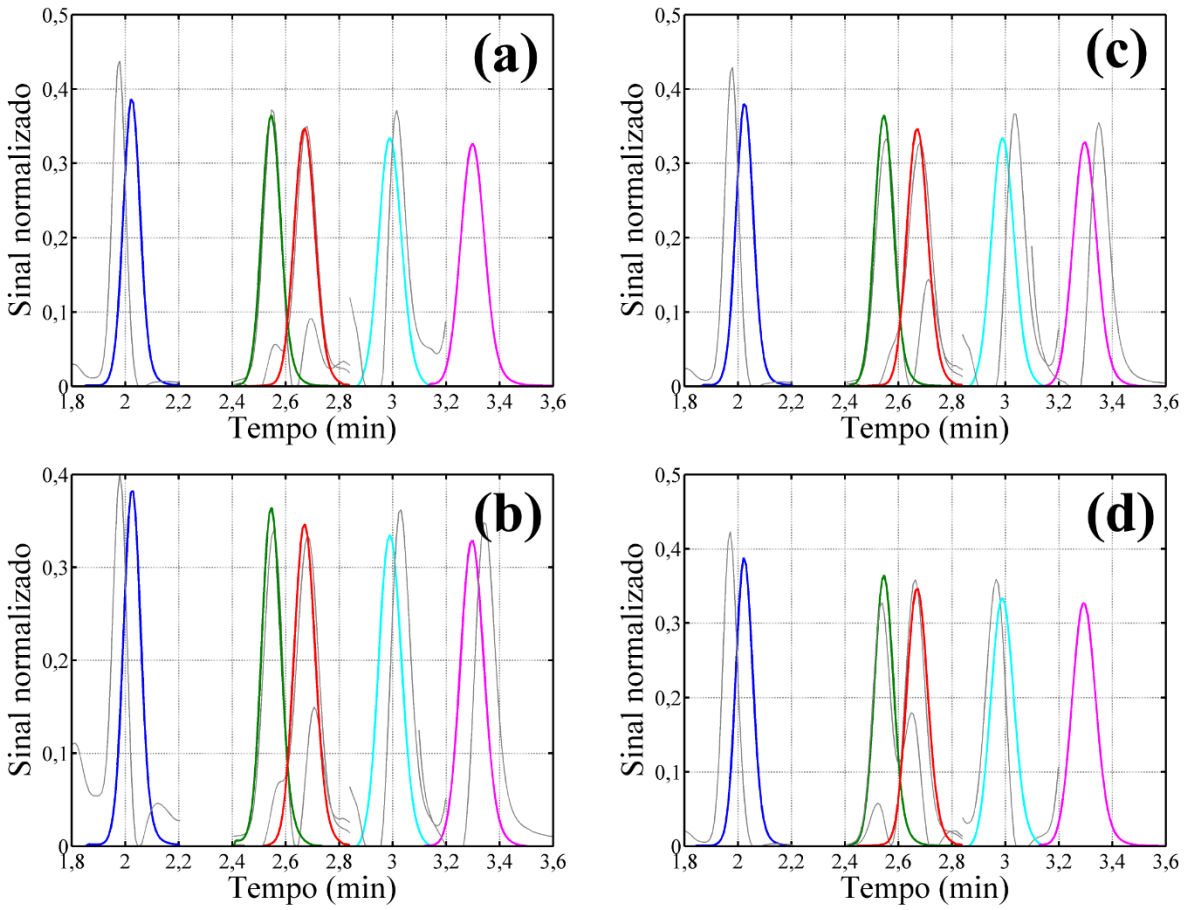

Figura 29 – Perfil cromatográfico recuperado pelo MCR-ALS para as amostras *Salminus maxillosus* (**a**), *Scomber scombrus* (**b**), *Seriola dumerili* (**c**) e *Lutjanus analis* (**d**). Os analitos são representados como linhas coloridas, TRI (—), FEN (—), PUT (—), CAD (—), HIS (—) e os interferentes como linhas cinzas (—).

Na **Figura 30** são mostrados os sinais normalizados dos analitos e dos interferentes recuperados pelo MCR-ALS no modo cromatográfico para as quatro amostras de peixe. Ao

comparar os perfis cromatográficos recuperados com o MCR-ALS na **Figura 30** com os cromatogramas mostrados nas **Figuras 29 a-d**, **30**, **6S**, **7S** e **8S**, é possível perceber que os alargamentos e desalinhamentos observados para as amostras de peixe e suas fortificações ocorrem devido à interferências.

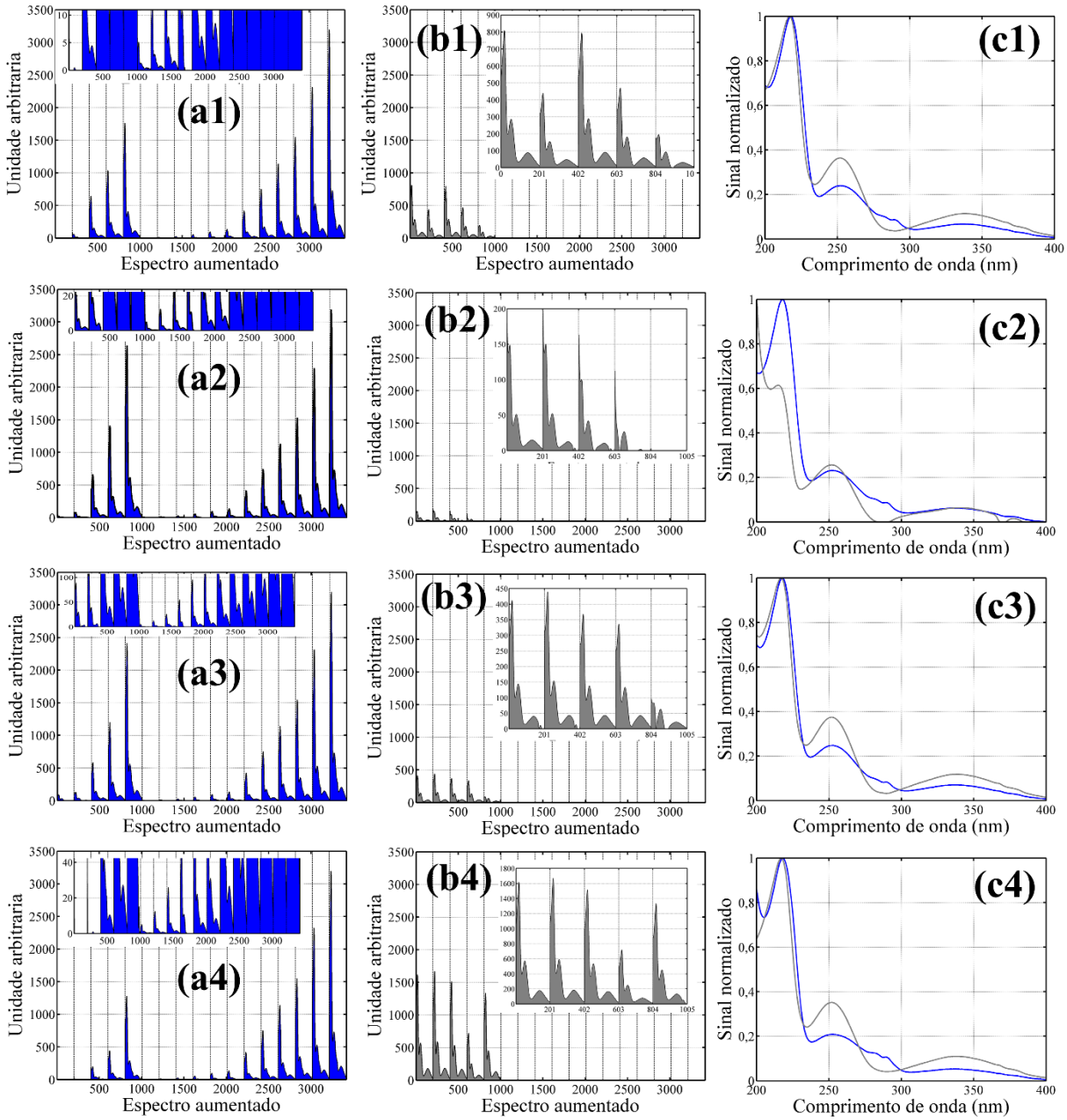

Figura 30 − Perfil espectral aumentado recuperado pelo MCR-ALS para a TRI (**a1-a4**) e para o interferente **(b1-b4**) nas fortificações das amostras de *Salminus maxillosus* (**1**), *Scomber scombrus* (**2**), *Seriola dumerili* (**3**) e *Lutjanus analis* (**4**) realizadas no primeiro dia de análise. O espectro da TRI e do interferente em uma concentração foi normalizado para comprimento 1 e sobreposto para comparação (**c1-c4**).

As interferências percebidas como um deslocamento ou um alargamento do sinal do analito nas amostras fortificadas na **Figura 29 a-d** foram recuperadas como interferentes na

Figura 31. Ao analisar apenas essas duas figuras pode-se inferir que esse problema ocorra devido ao alinhamento, e que o MCR-ALS está recuperando o sinal do mesmo analito deslocado. Entretanto, ao analisar as superfícies de contorno nas **Figuras 30**, **6S**, **7S** e **8S** pode-se identificar que os interferentes recuperados pelo MCR-ALS estão presentes nos cromatogramas para as amostras de peixe.

Nas amostras 2 (*Scomber scombrus*) e 3 (*Seriola dumerili*), é percebida uma interferência no sinal da HIS como um deslocamento para a direita nas amostras fortificadas. Essa interferência pode ser visualizada nas **Figuras 30 b** como a seta mostrando o interferente 6 e na **Figura 6S b** na forma de um sinal semelhante depois da HIS. O alongamento visto nessas figuras para o sinal da HIS também foi recuperado pelo MCR-ALS.

O perfil recuperado pelo MCR-ALS no modo espectral aumentado é mostrado nas **Figuras 30**, **9S**, **10S**, **11S** e **12S** respectivamente para os analitos TRI, FEN, PUT, CAD e HIS e os interferentes. As linhas verticais pontilhadas separam cada amostra, onde as cinco primeiras são da amostra sem fortificação e fortificando em quatro níveis, informadas na **seção 3.6**, respectivamente. Na sequência, são mostrados os perfis recuperados para as amostras de calibração. Tanto para o perfil dos analitos quanto para o dos interferentes, é apresentado um zoom, dentro da própria figura, para melhor visualização das amostras com concentrações mais baixas.

Antes de realizar a quantificação dos analitos é necessário saber se há ambiguidade rotacional nos perfis recuperados pelo MCR-ALS. Há duas formas de identificar ambiguidade rotacional, a primeira é comparando o perfil do analito registrado no conjunto de calibração e recuperado pelo MCR-ALS. Na **Figura 31** é mostrado o espectro normalizado para comprimento 1 dos analitos recuperados pelo MCR-ALS para cada amostra e registrado na amostra de calibração.

Matematicamente essa similaridade pode ser vista quando se compara os cossenos dos vetores normalizados de acordo com a **Equação 10** (Culzoni, et al., 2008).

$$\mathbf{S_{12}} = \frac{\|\mathbf{s_1^T s_2}\|}{\|\mathbf{s_1}\|\|\mathbf{s_2}\|} \qquad \textbf{(10)}$$

Onde, S_{12} é a similaridade entre os vetores S_1 e S_2; S_1 é o sinal normalizado do analito recuperado pelo algoritmo e S_2, o registrado pelo equipamento, também, normalizado.

Essa similaridade foi calculada tanto para o perfil no modo de eluição quanto no modo dos espectros, resultando em um valor de S_{12} superior a 0,99 para todos os analitos em ambos os modos. S_{12} igual a 1 significa perfis idênticos, e igual a 0 sinais ortogonais.

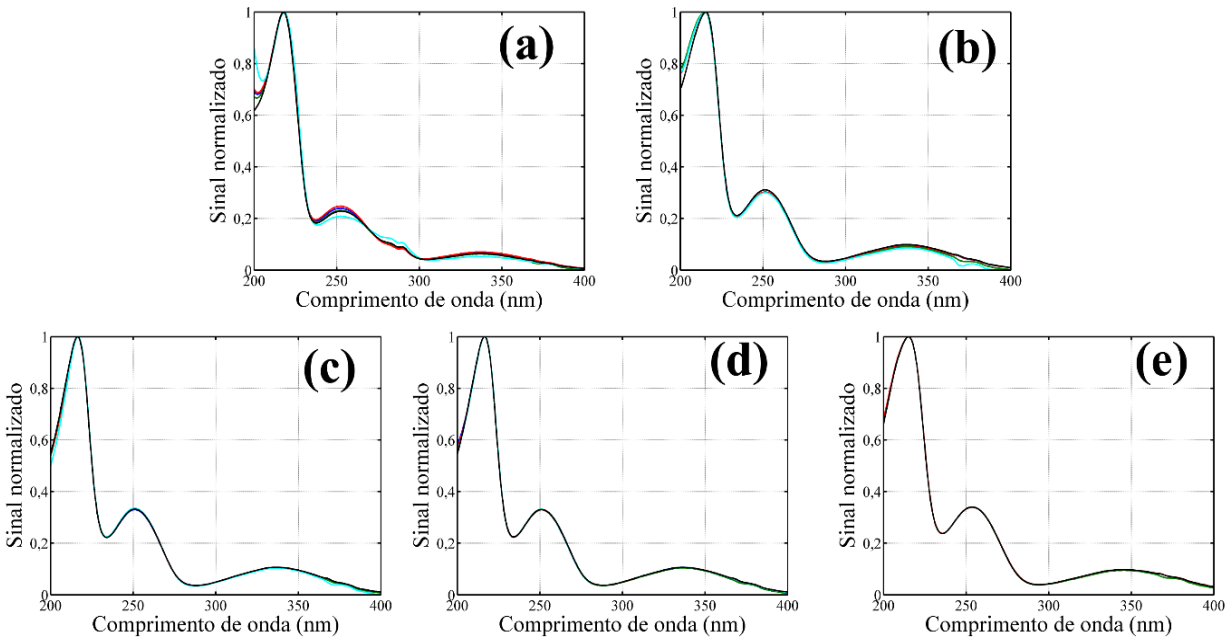

Figura 31 − Perfil espectral recuperado da TRI (**a**), FEN (**b**), PUT (**c**), CAD (**d**) e HIS (**e**) para o conjunto de calibração (**—**) e para as amostras de *Salminus maxillosus* (—), *Scomber scombrus* (—), *Seriola dumerili*(—) e *Lutjanus analis*(—).

A outra forma é calculando o valor da diferença entre a norma de Frobenious máxima e mínima, mostrado na **Equação 8**, essa diferença deve ser próxima de 0 para que não haja ambiguidade de rotação. Apesar das duas abordagens serem mostradas no conjunto de teste e para as amostras de peixe, apenas umas dessas comparações é suficiente para confirmar que o perfil recuperado não possui ambiguidade rotacional.

Na **Tabela 16** são mostrados os valores da norma de Frobenious inicial, máxima, mínima e a diferença entre a máxima e a mínima para a amostra de *Seriola dumerili*. O resultado para as outras amostras é mostrado na **Tabela 3S**. A diferença entre as normas foram 0 para TRI, FEN, PUT e CAD e próxima de 0 para HIS. Na **Figura 31** pode-se ver que os espectros recuperados de todas as amostras são bem parecidos com os registrados para o conjunto de calibração. Assim, pode-se dizer que não há ambiguidade rotacional para nenhum dos analitos.

Pode-se visualizar na **Figura 30** que a concentração da TRI nas amostras 2 (*Scomber scombrus*) e 3 (*Seriola dumerili*), foi maior que a menor concentração do conjunto de calibração tornando assim a sua quantificação possível. Percebe-se que, apesar das fortificações serem sempre as mesmas, há uma diferença na sensibilidade do sinal da TRI nas quatro amostras de

peixe. Situação semelhante ocorreu para os outros analitos, por isso foi feito o teste de efeito de matriz para todos os analitos e amostras.

Tabela 16 − Ambiguidade rotacional do MCR-ALS no modelo da amostra de *Seriola dumerili*.

Amostra [a]	f	TRI (2)*	FEN (2) *	PUT (2) *	CAD (2) *	HIS (2) *
Seriola dumerili	f_{ini}	0,802	0,632	0,869	0,903	0,934
	f_{max}	0,802	0,632	0,869	0,903	0,934
	f_{min}	0,802	0,632	0,869	0,903	0,862
	$\mathbf{f_{max-min}}$	0	0	0	0	0,072

[a] Restrição: 1 normalização; 2 não-negatividade; 3 unimodalidade; 4 correspondência entre as espécies;
$f_{max} - f_{min}$ é a diferença entre a norma de Frobenious máxima e mínima.
* Entre parênteses é mostrado o número de constituintes recuperados no modelo.

Existem dois tipos de efeito de matriz, um fixo referente a um constituinte interferente ou linha de base, e um proporcional que altera a sensibilidade do analito (Eurachem, 2014). O efeito de matriz fixo, no caso deste trabalho, foi removido pela calibração de ordem superior. No entanto, o proporcional só pode ser resolvido pelo método de adição de padrão, sendo assim é necessário para algumas amostras realizar este procedimento a fim de realizar a quantificação exata do analito.

Na **Figura 32** é mostrado o ajuste linear entre o valor predito e a concentração do analito fortificada, o valor apresentado expresso em mg kg^{-1} foi ajustado pelo fator de diluição explicado na **seção 3.6** e pela massa da amostra pesada antes da extração e derivatização.

Para identificar o efeito de matriz para cada analito foi comparado o coeficiente angular do conjunto de calibração e da amostra, onde a diferença entre os dois não deve ser maior que 20% (Eurachem, 2014; Moreno-González García-Campaña, 2017; Martins, et al., 2016). Tendo confirmado que o perfil recuperado não possui ambiguidade, foi possível alcançar a vantagem de segunda ordem e identificar as amostras que contém efeito de matriz, e, assim, a concentração inicial foi predita com os modelos gerados.

Na **Tabela 17** é mostrado o valor predito pelo modelo para cada analito na amostra de *Seriola dumerili*, o resultado para as demais amostras pode ser visto na **Tabela 4S** no Apêndice. No intuito de verificar a presença do efeito de matriz proporcional foi realizada a predição na amostra sem e com fortificação nos níveis mostrados na **seção 3.6**.

Uma vez que os interferentes para esses analitos não foram intensos como o interferente da TRI, já era possível perceber o efeito de matriz ao analisar os dados, antes de serem tratados

com o MCR-ALS, nas **Figuras 27 a-d**, **28**, **6S**, **7S** e **8S** e Apêndice. Assim, na amostra de *Seriola dumerili* foi identificado efeito de matriz para a PUT e a CAD. O único analito que não apresentou efeito de matriz em nenhuma amostra foi a HIS.

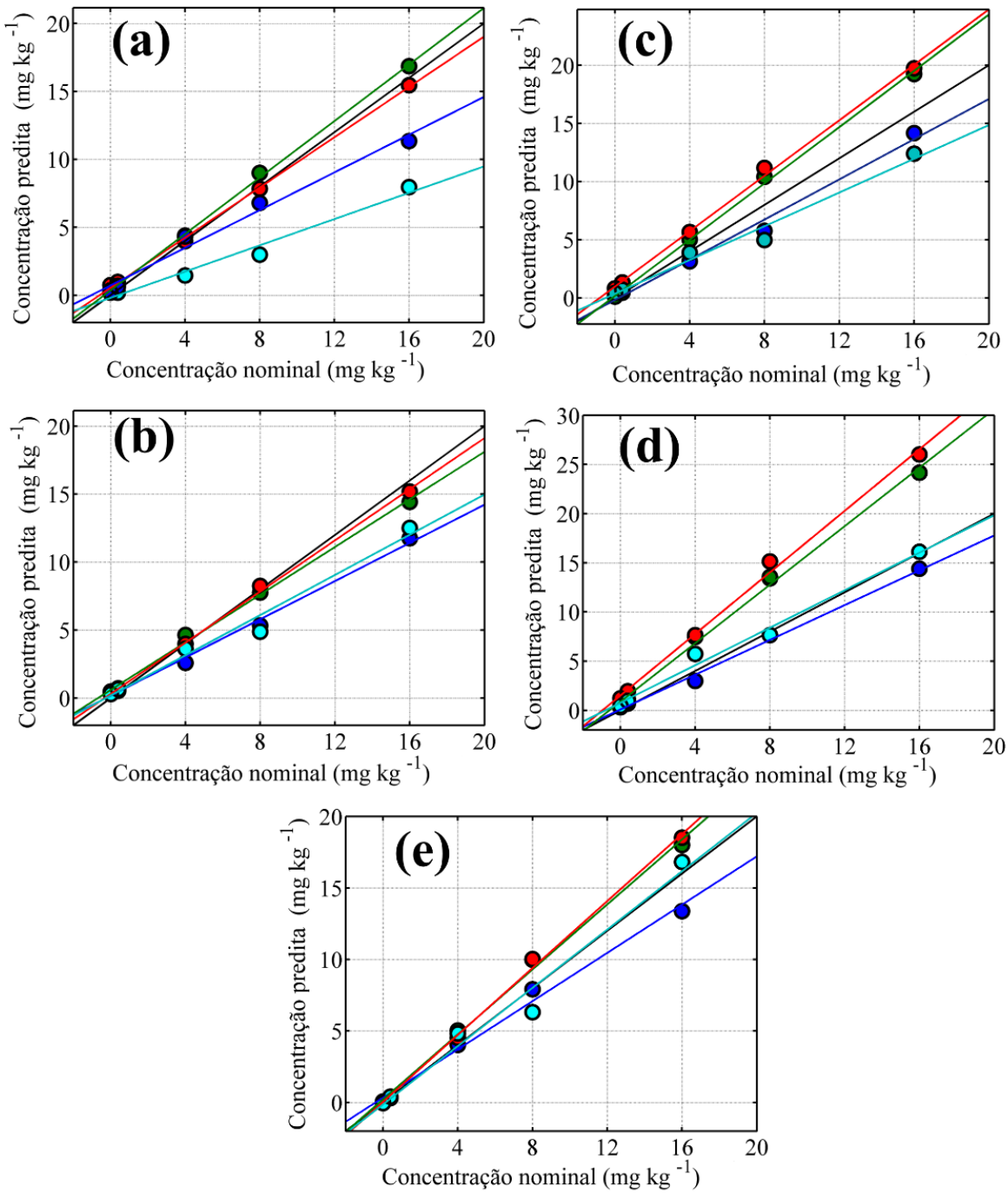

Figura 32 − Valor predito versus o nominal fortificado de TRI (**a**), FEN (**b**), PUT (**c**), CAD (**d**) e HIS (**e**) para o conjunto de calibração (—) e para as amostras de *Salminus maxillosus* (—), *Scomber scombrus* (—), *Seriola dumerili* (—) e *Lutjanus analis*(—).

Quando o efeito de matriz proporcional estava presente, não foi mostrado o valor de recuperação, uma vez que essa alteração na sensibilidade do analito torna errôneo o uso dessa métrica. Entre as amostras que não apresentaram esse efeito de matriz, a recuperação média variou entre 76,3% e 111,26%. Levando em consideração a complexidade da matriz e da determinação que envolve extração e derivatização, esta metodologia apresentou resultados com exatidão adequada. Vale ressaltar que foi possível identificar, praticamente, todas as aminas biogênicas nas amostras de peixe, e felizmente em todos os casos a concentração foi inferior à máxima estabelecida pelos órgãos regulamentadores como sendo seguro para o consumo (EFSA, 2011; FAO/WHO, 2013).

Já era esperado que os níveis de concentração das ABs não fossem elevados, uma vez que as amostras de peixe adquiridas no mercado de peixe local, eram frescas e recém-pescadas. As amostras foram acondicionadas a 4 ºC por duas semanas e analisadas a cada dois dias, nesse tempo acompanhou-se a evolução no crescimento da concentração das ABs indicando a deterioração do peixe.

Tabela 17 – Teste de recuperação das aminas biogênicas nas amostras de *Seriola dumerili* usando o MCR-ALS.

Peixe	TRI[a]			FEN[a]			PUT[a]			CAD[a]			HIS[a]		
	N[b]	P[c]	R(%)	N[b]	P[c]	R(%)	N[b]	P[c]	R(%)	N[b]	P[c]	R(%)	N[b]	P[c]	R(%)
Seriola dumerili	0	0,75(0,06)		0	0,35(0,04)		0	0,84(0,06)		0	1,23(0,18)		0	0,08(0,11)	
	0,4	0,98(0,06)	59,65	0,4	0,54(0,04)	47,15	0,4	1,34(0,06)		0,4	1,94(0,17)		0,4	0,32(0,11)	60,10
	4	3,99(0,05)	81,07	4	3,99(0,03)	91,04	4	5,66(0,05)		4	7,66(0,16)		4	4,53(0,09)	111,37
	8	7,84(0,06)	88,68	8	8,26(0,03)	98,81	8	11,18(0,07)		8	15,16(0,28)		8	10,04(0,11)	124,51
	16	15,46(0,10)	91,99	16	15,20(0,06)	92,78	16	19,74(0,13)		16	26,04(0,52)		16	18,52(0,20)	115,26
Inicial		**7,46**	**80,35**		**3,52**	**82,44**		**8,28**			**9,66**			**0,78**	**102,81**

[a] Concentração corrigida levando em consideração o procedimento de extração e a massa da amostra em mg kg^{-1}. Recuperação (R) é expressa em porcentagem. A recuperação não foi mostrada no caso onde o efeito de matriz proporcional estava presente. Em vermelho é representado a média das recuperações. [b] Concentração fortificada na amostra. [c] Concentração predita para as amostras de peixe com os modelos MCR-ALS. Desvio padrão é mostrado entre parênteses para cada valor predito. [c] A concentração na amostra sem fortificação é expressa em mg kg^{-1} e destacada em azul. O Método de adição de padrão foi usado para calcular a concentração na amostra quando o efeito de matriz proporcional estava presente.

4.8 Estudo cinético de degradação

As ABs são indicadores de degradação e são naturalmente formadas por bactérias em amostras de peixe. Assim, foi realizado um estudo cinético ao longo de treze dias para as quatro amostras de peixe, sem fortificação, em que foi avaliado o crescimento na concentração das cinco ABs. A quantificação foi realizada de acordo com o método descrito na **Metodologia**, e foi acompanhado o crescimento na concentração das cinco ABs ao longo dos dias de análise, como mostrado na **Figura 33**.

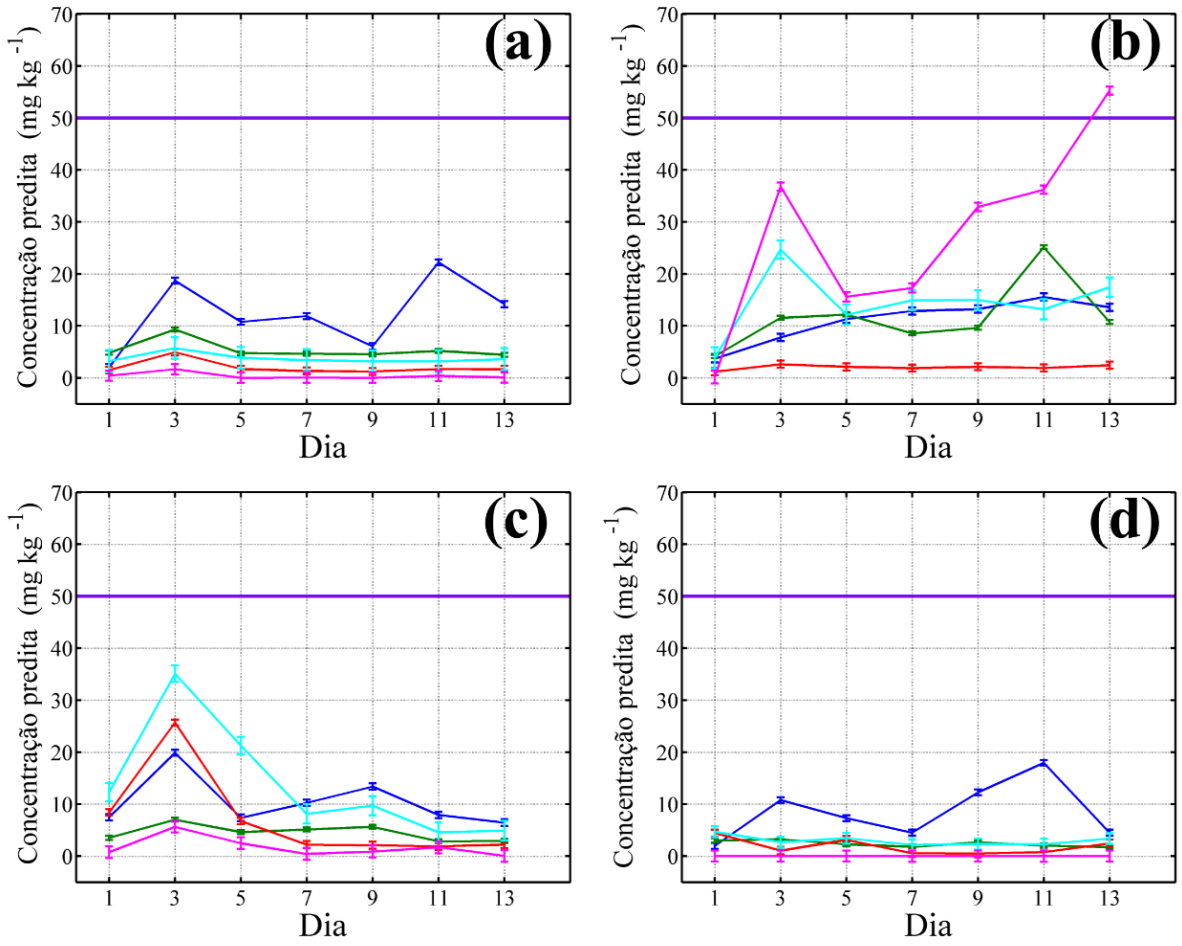

Figura 33 − Evolução da concentração da TRI (—), FEN (—), PUT (—), CAD (—) e HIS (—) ao longo de 13 dias de análise das amostras de *Salminus maxillosus* (**a**), *Scomber scombrus* (**b**), *Seriola dumerili* (**c**) e *Lutjanus analis* (**d**). A barra de erro mostrada em cada concentração predita foi calculada de acorda na interface MVC2 (Olivieri, Wu, Yu, 2009; Olivieri, 2014). A linha roxa (—) em 50 mg kg^{-1} representa o limite máximo estabelecido para o peixe estar seguro para o consumo.

Na **Figura 33** é mostrado um gráfico da concentração de cada ABs nos treze dias de análise para as quatro amostras de peixe acondicionadas a 4 ℃. Percebe-se que o crescimento não é linear, essa variação ocorre devido à cinética de formação das ABs depender da distribuição dos microrganismos e das proteínas percussoras no peixe. Apesar de a amostragem ter sido feita a

fim de melhorar a representatividade da amostra, há partes do peixe nas quais o crescimento das ABs é maior devido à presença de microrganismos e substrato em concentrações variadas, resultando em um crescimento não linear da curva de degradação. Essa é uma característica de amostras cujos analitos são formados através de processos biológicos.

Apenas a amostra de *Scomber scombrus* atingiu a concentração superior a 50 mg kg^{-1} no último dia de análise, ultrapassando o valor estabelecido como seguro para o consumo (EFSA, 2011; FAO/WHO, 2013). Esse é um resultado que pode alarmar a população, no entanto, vale ressaltar que não se pode generalizar esse resultado para todas as amostras de cavala, uma vez que foi feita a análise de, apenas, uma amostra deste peixe. Além disso, a produção das ABs aumenta com a presença de microrganismos que realizam o processo de descarboxilação da amina percussora, o que aumenta a variabilidade desse resultado. Outro fator a se considerar é que essas amostras foram acondicionadas na temperatura em que os peixes devem ser mantidos nos mercados. Geralmente, a população armazena esse produto a temperaturas inferiores, o que eleva o tempo necessário para degradá-las. Por outro lado, a cinética de formação das ABs aumenta a temperatura maiores que 4℃, o que reforça o cuidado que se deve ter ao armazenar esse produto nos mercados. Ressalta-se, também, a necessidade de analises sensíveis, seletivas e exatas para a quantificação das ABs, visando o controle fitossanitário de peixes. Dados da literatura ressaltam que a 25℃ o peixe torna-se impróprio para o consumo em até 6 horas (FAO/WHO, 2013).

Na **Figura 34** são mostrados os cromatogramas para a amostra de *Scomber scombrus* registrados em 215 nm, a variação na concentração é perceptível, em especial para a HIS. Para os outros analitos essa variação não é tão nítida devido à presença de interferências. A variação das concentrações mostrada na **Figura 34** pode ser mais bem visualizada na **Figura 13S**, na qual são apresentados os espectros aumentados recuperados pelo MCR-ALS para cada analito, em que a área do espectro recuperado é proporcional à concentração. Na **Figura 13S** é possível visualizar, facilmente, o crescimento na concentração de histamina, em especial, nas três últimas análises, a concentração permaneceu praticamente constante para as outras aminas biogênicas. É de se esperar que a cinética de formação da histamina seja acelerada nesta amostra, uma vez que ela possui um elevado teor de histidina, amina precursora da histamina.

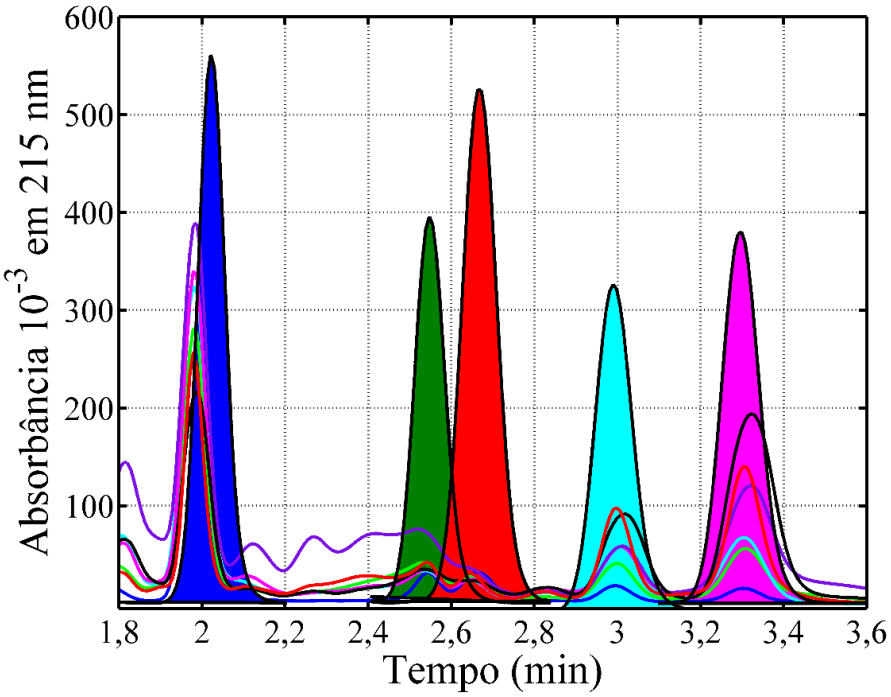

Figura 34 − Cromatograma registrado em 215 nm para a amostra de *Scomber scombrus* registrado nos dias 1(—), 3(—), 5(—), 7(—), 9(—), 11(—), 13(—) e os cromatogramas preenchidos foram registrados para TRI (**Azul**), FEN (**Verde**), PUT (**Vermelho**), CAD (**Ciano**) e HIS (**Margenta**) na amostra de calibração de10 mg L^{-1}.

4.9 Comparação da metodologia proposta com outras publicadas na literatura

A grande vantagem deste trabalho, em comparação a outros de quantificação de ABs em peixe, é o uso da calibração de ordem superior. Com a vantagem de segunda ordem, o tempo de eluição pode ser reduzido sem comprometer a exatidão da análise, pois os interferentes são resolvidos com a modelagem dos dados aumentados no modo espectral com o MCR-ALS realizado nos dados alinhados.

Na **Tabela 18** é feita uma comparação entre este e mais oito trabalhos publicados para a análise de peixes e seus derivados usando cromatografia líquida. Como pode ser verificado, a etapa de extração é realizada com ácido perclórico ($HClO_4$), ou com ácido tricloroacético (TCA). No entanto, levando em consideração o valor de ambos os ácidos e a necessidade do uso de uma quantidade mais elevada de TCA nos trabalhos publicados, optou-se por usar o $HClO_4$, devido ao seu uso ser em concentrações baixas. Neste trabalho foram testados ambas as extrações que apresentaram resultados semelhantes. Vale destacar a eficiência frente à simplicidade da extração, que requer, apenas, uma agitação da amostra triturada. O uso desse solvente com caráter mais polar é eficiente para extrair as ABs, mantendo a carga orgânica apolar, cuja retenção na coluna é maior, na amostra. Essa forma de extração reduz drasticamente a quantidade de interferências nas análises, além de ser eficiente para a extração das ABs. Outro fator importante

é que o uso de $HCLO_4$ em baixa concentração não promove a digestão da amostra à temperatura ambiente, isso é muito importante tendo em vista a preservação da integridade da amostra.

A maioria dos trabalhos faz uso do DNS como reagente derivatizante, pois em geral ele é usado no pré-coluna e não demanda uma instrumentação tão elaborada (Latorre-Moratalla et al., 2009). Em todos os trabalhos a HIS foi analisada, pois apenas esta possui um limite máximo estabelecido por órgãos regulamentadores. A eluição gradiente foi a mais usada, e apenas neste trabalho foi possível quantificar cinco ABs com uma eluição isocrática menor que quatro minutos, o que resultou em uma frequência analítica duas vezes maior até que metodologias que fazem uso de UPLC-MS/MS (Romero-González, Alarcón-Flores, Vidal, Frenich, 2012). A faixa de calibração, que variou de 0,05 a 20 mg L^{-1}, foi ampla e permitiu a análise das ABs na faixa necessária para identificar a degradação, 50 mg L^{-1} que gerou um sinal de 5 mg L^{-1} na curva de calibração devido ao fator de diluição de 10 vezes. O LOQ foi baixo e compatível com o sinal observado experimentalmente, isso é importante de ser estudado e avaliado uma vez que as métricas de desempenho analítico para calibração de ordem superior ainda estão em desenvolvimento e sendo aprimoradas (Olivieri A. C., 2014).

Com a metodologia proposta foi possível a quantificação de cinco ABs em um tempo inferior a outras metodologias, em especial, as que usam eluição isocrática que requereram catorze minutos para a quantificação de uma única amina em atum (Tahmouzi, S., Khaksar, R. Ghasemlou, M., 2011). A faixa de recuperação e o LOQ variaram entre 76,3 e 111,3 %, e 0,0 e 0,1, respectivamente para a análise das amostras de peixe. Resultado muito promissor, em especial, quando comparado a outras abordagens que fazem uso de eluições em gradiente e detecção mais complexa e de elevado custo de manutenção, como a espectrometria de massas. Além disso, vale destacar que tal resultado foi obtido mesmo com a complexidade da amostra e da análise, fazendo uso de uma etapa de derivatização e contendo outros constituintes majoritários em sua composição, como: proteínas, lipídeos, carboidratos, aminas, entre outros compostos que podem interferir na análise e podem estar em concentrações mais elevadas que o analito.

Apesar das inúmeras vantagens, vale ressaltar a necessidade de um analista experiente na área de calibração de ordem superior. Também é necessário ter a disponibilidade de um equipamento que gere esse tipo de dados. No entanto, quando se compara com as metodologias propostas na literatura referenciadas na **Tabela 18**, em todas elas usam um sistema LC que na maioria dos casos dispõe como detector básico o DAD que foi usado neste trabalho, o qual gera dados de ordem superior de forma rápida e simplificada.

Tabela 18 − Metodologias cromatográficas reportadas na literatura para quantificação de aminas biogênicas em peixes.

Amostra	Extração[a]	R. Der.[b]	ABs[c]	Método[d]	Calibração (mg l^{-1})[e]	Tempo de Eluição e modo[f]	LOQ (mg l^{-1})[g]	RECc (%)[h]	Referência[i]
Atum	5% TCA	OPA	5	HPLC-FLD	5 - 100	14 min (Isocrático)	4,50	95,73 - 104,78	(Tahmouzi, 2011)
Peixe e derivados	5% TCA	DNS	1-6, 12-13	HPLC-FLD	1 - 40	26 min (Gradiente)	0,06 - 0,72	80,90 - 97,40	(Zhai, 2012)
Derivados de peixe	0,4 mol L^{-1} $HClO_4$	DNS	1-6, 12-13	HPLC-DAD	2,5 - 400	20 min (Gradiente)	1,98–5,46	68,90 - 108,70	(Köse, 2012)
Fruto do mar	0,4 mol L^{-1} $HClO_4$	DNS	1-7, 12-13	HPLC-DAD	0,5 - 100	12 min (Gradiente)	0,40 - 2,40	87,00 - 121,00	(Simat, 2011)
Peixe	5% TCA	DNS	1, 3-6, 12-13	HPLC-DAD	0,5 - 100	20 min (Gradiente)	0,06 - 0,78	73,20 - 107,00	(Hu, 2012)
Peixe	0,6 mol L^{-1} $HClO_4$	OPA	1-7, 9, 12-15	UHPLC-FLD	0,1 - 50	7 min (Gradiente)	0,15 - 0,9	83,68 - 98,81	(Latorre-Moratalla, 2009)
Anchova	0,6 mol L^{-1} $HClO_4$	-	3 - 6	UPLC-MS/MS	0,01 – 0,75	8,5 min (Gradiente)	0,03 - 0,06	71,40 - 108,40	(Romero-González, 2012)
Cubos de peixe	6% TCA	DNS	1-7, 11-12	HPLC-DAD	0 - 20	30 min (Gradiente)	0,15 – 1,5	55,6- 113,7	(Chen, 2010)
Peixe	0,6 mol L^{-1} $HClO_4$	DNS	1 - 5	HPLC-DAD-MCR-ALS	0,05 - 20	4 min (Isocrático)	0,04 - 0,11	76,3-111,26	[Esta Tese]

[a] TCA –ácido tricloroacético,$HClO_4$–ácido perclorico. [b] Reagente derivatizante: OPA - o-ftalaldeido, DNS–cloreto de dansil. [c] aminas biogênicas (ABs) triptamina (1), 2-feniletilamina (2), putrescina (3), cadaverina (4), histamina (5), tiramina (6), agmatina (7), metilamina (8), dopamina (9), etilamina (10), isopentilamina (11), spermidina (12), spermina (13), serotonina (14), octopamina (15).[d] Método de análise. [e] Faixa de calibração das ABs. [f] Tempo gasto na eluição e modo de eluição. [g] Limite de quantificação na amostra de peixe. [h] Faixa de recuperação. [i] Referências.

Capítulo 5

CONCLUSÕES

5 CONCLUSÕES

Neste trabalho foi possível alcançar a vantagem de segunda ordem em um sistema complexo, tanto do ponto de vista químico quanto do quimiométrico.

A parte química se torna mais complexa devido à detecção requerer extração e derivatização dos padrões e amostras e a quantificação ser feita em amostra de carne de peixe, que além de complexa e possuir constituintes que não são de interesse em maior quantidade, é uma amostra em que o analito está distribuído de forma heterogênea.

Do ponto de vista quimiométrico, a complexidade de alcançar a vantagem de ordem superior ocorre devido à falta de algoritmos específicos para tratar dados com essa problemática, requerendo associações de estratégias para alcançar quantificações exatas. Foi necessário construir um conjunto de calibração para cada analito e alinhá-lo com base no sinal da amostra analisada, além de modelar os dados com o MCR-ALS aumentado no modo espectral.

Com a abordagem proposta, foi possível realizar a modelagem de dados contendo quebra da trilinearidade devido ao deslocamento do tempo de retenção e deficiência de posto pela similaridade dos espectros dos analitos. Mostrou-se que o alinhamento com o algoritmo *coshift* foi eficiente e não distorceu a matriz dos dados. No entanto, é necessário realizar o alinhamento no conjunto de calibração contendo apenas um analito. Foi possível atingir resultados satisfatórios com REP inferior a 8% para todas as ABs analisadas e valores de LOQ bem abaixo do limite máximo estabelecido para esses analitos em amostras de carne de peixe. Foi possível detectar as ABs nas amostras sem fortificação e acompanhou-se a cinética de formação delas, o que é interessante devido a estes analitos serem marcadores de degradação da amostra analisada.

O estudo cinético permitiu a análise das ABs e avaliar o crescimento da concentração nas amostras analisadas. É importante notar que a temperatura de armazenamento das amostras é a regulamentada para o armazenamento do peixe em mercados (4 ℃), e apenas após 13 dias uma única amostra atingiu o valor de 50 mg Kg^{-1} que indica que a amostra não está adequada para o consumo. No entanto é importante ressaltar que o consumidor não deve se chocar com esse curto prazo de conservação da amostra, uma vez que o armazenamento pelo consumidor ocorre em congeladores ou freezers que atingem temperaturas mais baixas que 4 ℃ o que aumenta o tempo em que a amostra apresenta esse teor de degradação, pois reduz a atividade bacteriana.

5.1 Propostas futuras

Apesar do *coshift* ser um algoritmo idealizado para dados de primeira ordem, foi possível utilizá-lo para alinhar os picos cromatográficos. No entanto, é importante o desenvolvimento de

algoritmos de alinhamento específicos para dados de ordem superior, uma vez que nem sempre será viável alinhar dados que contêm um único sinal, como foi o caso deste trabalho.

Alguns alinhamentos foram propostos para dados de ordem superior (Bortolato, Arancibia, Escandar, Olivieri, 2010; Parastar, Akvan, 2014), no entanto, nenhum deles permite alinhar corretamente o analito quando há interferentes com espectros muito correlacionados coeluidos, ou na mesma região. Apesar de não resolver o problema de deslocamento do tempo de retenção quando há dependência linear nos dados, a estratégia de alinhar os perfis recuperados pelo MCR-ALS é interessante (Parastar, Akvan, 2014), e também se mostrou possível alinhar usando perfis recuperados pelo PARAFAC (Bortolato, Arancibia, Escandar, Olivieri, 2010).

Propõe-se alinhar os perfis recuperados usando o PARALIND com o COW e o *coshift*, e comparar os resultados da mesma forma que nesta tese. Os resultados do alinhamento serão comparados com os desta tese para a modelagem das aminas biogênicas. Esta abordagem será testada também na quantificação de quatro hormônios femininos com perfis idênticos coeluidos. Em ambos os casos, os espectros são correlacionados apresentando deficiência de posto e interferências, além do desalinhamento de pico.

REFERÊNCIAS

Ahmadvand, M., Parastar, H., Sereshti, H., Olivieri, A., Tauler, R. (2016). A systematic study on the effect of noise and shift on multivariate figures of merit of second-order calibration algorithms. *Analytica Chimica Acta, pp. 18-31*.

Allegrini, F., Olivieri, A. C. (2014). IUPAC-Consistent Approach to the Limit of Detection in Partial Least-Squares Calibration. *Analytical Chemistry*, pp. 7858-7866.

Amrhein, M., Srinivasan, B., Bonvin, D., Schumacher, M. M. (1996). On the rank deficiency and rank augmentation of the spectral measurement matrix. *Chemometrics and Intelligent Laboratory Systems*, pp. 17-33.

Arancibia, J. A., Damiani, P. C., Escandar, G. M., Ibañez, G. A., Olivieri, A. C. (2012). A review on second- and third-order multivariate calibration applied to chromatographic data. *Journal of Chromatography B*, pp. 22-30.

Bahram, M., Bro, R. (2007). A novel strategy for solving matrix effect in three-way data using parallel profiles with linear dependencies. *Analytica Chimica Acta*, pp. 397–402.

Barros Neto, B., Pimentel, M. F., Araújo, M. C. (2002). Recomendações para calibração em química analítica - parte i. Fundamentos e calibração com um componente (calibração univariada). *Química Nova*, pp. 856-865.

Barros Neto, B.; Scarminio, I. S.; Bruns, R. E. *Como Fazer Experimentos: Aplicações na Ciência e na Indústria*. p. 414, 2010.

Bauza, M. C., Ibañez, G. A., Tauler, R., Olivieri, A. C. (2012). Sensitivity Equation for Quantitative Analysis with Multivariate Curve Resolution-Alternating Least-Squares: Theoretical and Experimental Approach. *Analitycal Chemistry*, pp. 8697-8706.

Bordagaray, A., Amigo, J. M. (2015). Modelling highly co-eluted peaks of analytes with high spectral similarity. *Trends in Analytical Chemistry*, pp. 107-118.

Borraccetti, M. D., Damiani, P. C., Olivieri, A. C. (2009). When unfolding is better: unique success of unfolded partial least-squares regression with residual bilinearization for the processing of spectral–pH data with strong spectral overlapping. Analysis of fluoroquinolones in human urine based on flow-injectio. *Analyst*, pp. 1682–1691.

Bortolato, S. A., Arancibia, J. A., Escandar, G. M., Olivieri, A. C. (2010). Time-alignment of bidimensional chromatograms in the presence of uncalibrated interferences using parallel factor analysis Application to multi-component determinations using liquid-chromatography with spectrofluorimetric detection. *Chemometrics and Intelligent Laboratory Systems*, pp. 30–37.

Bro, R. (1998). Multi-way Analysis in the Food Industry: Models, Algorithms, and Applications. Denmark.

Bro, R., Harshman, R. A., Sidiropoulos, N. D., Lundy, M. E. (2009). Modeling multi-way data with linearly dependent pesos. *Journal of Chemometrics*, pp. 324-340.

Chen, H.-C., Huang, Y.-R., Hsu, H.-H., Lin, C.-S., Chen, W.-C., Lin, C.-M., Tsai, Y.-H. (2010). Determination of histamine and biogenic amines in fish cubes (Tetrapturus angustirostris) implicated in a food-borne poisoning. *Food Control*, pp. 13-18.

Collins , C. H. (2006). *Fundamentos de Cromatografia.* Campinas: Unicamp.

Culzoni, M. J., Goicoechea, H. C., Ibañez, G. A., Lozano, V. A., Marsili, N. R., Olivieri, A. C., Pagani, A. P. (2008). Second-order advantage from kinetic-spectroscopic data matrices in

the presence of extreme spectral overlapping A multivariate curve resolution-Alternating least-squares approach. *Analytica chimica ACTA*, pp. 46-57.

Culzoni, M. J., Llanos, A. M., De Zan, M. M., Espinosa-Mansilla, A., Cañada-Cañada, F., Peña, A. M., Goicoechea, H. C. (2011). Enhanced MCR-ALS modeling of HPLC with fast scan fluorimetric detection second-order data for quantitation of metabolic disorder marker pteridines in urine. *Talanta*, pp. 2368– 2374.

de la Peña, A. M., Goicoechea, H. C., Escandar, G. M., Olivieri, A. C. (2015). *Fundamentals and Analytical Applications of Multiway Calibration.* Amsterdam: Elsevier.

EFSA. (2011). Scientific Opinion on risk based control of biogenic amine formation in fermented foods. *European Food Safety Authority*, 2393-2486.

Erim, F. B. (2013). Recent analytical approaches to the analysis of biogenic amines in food samples. *Trends in Analytical Chemistry*, pp. 239–247.

Escandar, G. M., Goicoechea, H. C., Peña, A. M., Olivieri, A. C. (2014). Second- and higher-order data generation and calibration: A tutorial. *Analytica Chimica Acta*, pp. 8-26.

Eurachem. (2014). *The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics.* Fonte: Eurachem: www.eurachem.org

Fanali, S., Haddad, P. R., Poole, C. F., Schenmakers, P., Lloyd, D. (2013). *LIQUID CHROMATOGRAPHY:FUNDAMENTALS AND INSTRUMENTATION.* Amsterdam: Elsevier.

FAO/WHO. (2013). *Public Health Risks of Histamine and other Biogenic Amines from Fish and Fishery Products. Meeting report.* Rome Italy: FAO headquarters.

Fragoso, W., Allegrini, F., Olivieri, A. C. (2016). A new and consistent parameter for measuring the quality of multivariate analytical methods: Generalized analytical sensitivity. *Analytica Chimica Acta*, pp. 43-49.

Gaitonde, V. N., Karnik, S R Achyutha, B. T., Siddeswarappa, B (2006) Multi-response optimization in drilling using Taguchi's quality loss function. *Indian Journal of Engineering and Materials Sciences*, pp. 484–488.

Grisales, J. O., Arancibia, J. A., Castells, C. B., Olivieri, A. C. (2012). Determination of enantiomeric composition of ibuprofen in pharmaceutical formulations by partial least-squares regression of strongly overlapped chromatographic profiles. *Journal of Chromatography B*, pp. 78-83.

Herrero, A., Sanllorente, S., Reguera, C., Ortiz, M. C., Sarabia, L. A. (s.d.). A new multiresponse optimization approach in combination with a D-Optimal experimental design for the determination of biogenic amines in fish by HPLC-FLD. *Analytica Chimica Acta*, pp. 31e38.

Hibbert, D. B. Experimental design in chromatography: A tutorial review. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, pp. 2–13, 2012.

Hu, Y., Huang, Z., Li, J., Yang, H. (2012). Concentrations of biogenic amines in fish, squid and octopus and their changes during storage. *Food Chemistry*, pp. 2604–2611.

Jaumot, J., Tauler, R. (2010). MCR-BANDS: A user friendly MATLAB program for the evaluation of rotation ambiguities in Multivariate Curve Resolution. *Chemometrics and Intelligent Laboratory Systems*, pp. 96–107.

Jaumot, J., de Juan, A., Tauler, R. (2015). MCR-ALS GUI 2.0: New features and applications. *Chemometrics and Intelligent Laboratory Systems*, pp. 1-12.

Köse, S., Koral, S., Tufan, B., Pompe, M., Scavnic¸ar, A., Koc¸ar, D. (2012). Biogenic amine contents of commercially processed traditional fish products originating from European countries and Turkey. *European Food Research and Technology*, pp. 669–683.

Lanças, F. M. (2009). *CROMATOGRAFIA LIQUIDA MODERNA - HPLC/CLAE*. Atomo.

Latorre-Moratalla, M. L., Bosch-Fusté, J., Lavizzari, T., Bover-Cid, S., Veciana-Nogués, M. T., Vidal-Carou, M. C. (2009). Validation of an ultra high pressure liquid chromatographic method for the determination of biologically active amines in food. *Journal of Chromatography A*, pp. 7715–7720.

Maghsoodloo, S., Ozdemir, G., Jordan, V., Huang, C. H.. (2004) Strengths and Limitations of Taguchi's Contributions to Quality, Manufacturing, and Process Engineering. *Journal of Manufacturing Systems*, pp. 73–126.

Março, P. H., Valderrama, P., Alexandrino, G. L., Poppi, R. J., Tauler, R. (2014). Resolução multivariada de curvas com mínimos quadrados alternantes: descrição, funcionamento e aplicações. *Química nova*, pp. 1525-1532.

Martins, M. L., Rizzetti, T. M., Kemmerich, M., Saibt, N., Prestes, O. D., Adaime, M. B., Zanella, R. (2016). Dilution standard addition calibration: A practical calibration strategy for multiresidue organic compounds determination. *Journal of Chromatography A*, pp. 84-91.

Mey, E. d., Drabik-Markiewicz, G., Maere, H. d., Peeters, M. C., Derdelinckx, G., Paelinck, H., Kowalska, T. (2012). Dabsyl derivatisation as an alternative for dansylation in the detection of biogenic amines in fermented meat products by reversed phase high performance liquid chromatography. *Food Chemistry*, pp. 1017–1023.

Mohammed, G. I., Bashammakh, A. S., Alsibaai, A. A., Alwael, H., El-Shahawi, M. S. (2016). A critical overview on the chemistry, clean-up and recent advances in analysis of biogenic amines in foodstuffs. *Trends in Analytical Chemistry*, pp. 84-94.

Montemurro, M., Pinto, L., Véras, G., Gomes, A. d., Culzoni, M. J., Araújo, M. C., Goicoechea, H. C. (2016). Highly sensitive quantitation of pesticides in fruit juice samples by modeling four-way data gathered with high-performance liquid chromatography with fluorescence excitation-emission detection. *Talanta*, pp. 208-218.

Moreno-González, D., García-Campaña, A. M. (2017). Salting-out assisted liquid–liquid extraction coupled to ultra-high performance liquid chromatography–tandem mass spectrometry for the determination of tetracycline residues in infant foods. *Food Chemistry* , pp. 1763–1769.

Olivieri, A. C. (2008). Analytical Advantages of Multivariate Data Processing. One, Two, Three, Infinity? *Analytical Chemistry*, pp. 5713–5720.

Olivieri, A. C. (2014). Analytical Figures of Merit: From Univariate to Multiway Calibration. *Chemical Reviews*, pp. 5358-5378.

Olivieri, A. C. (2015). Practical guidelines for reporting results in single- and multi-component analytical calibration: A tutorial. *Analytica Chimica Acta*, pp. 10–22.

Olivieri, A. C., Escandar, G. M. (2014). *Practical Three-Way Calibration.* Amsterdam: Elsevier.

Olivieri, A. C., Wu, H.-L., Yu, R.-Q. (2009). MVC2: A MATLAB graphical interface toolbox for second-order multivariate calibration. *Chemometrics and Intelligent Laboratory Systems*, pp. 246-251.

Ordóñez, J. L., Troncoso, A. M., García-Parrilla, M. D., Callejón, R. M. (2016). Recent trends in the determination of biogenic amines in fermented beverages - A review. *Analytica Chimica Acta*, pp. 10-25.

Ortiz, M. C., Sarabia, L. A., Sánchez, M. S. (2010). Tutorial on evaluation of type I and type II errors in chemical analyses: From the analytical detection to authentication of products and process control. *Analytica Chimica Acta*, pp. 123-142.

Parastar, H. (2015). Multivariate Curve Resolution Methods for Qualitative and Quantitative Analysis in Analytical Chemistry. Em A. M. Peña, H. C. Goicoechea, G. M. Escandar, A. C. Olivieri, *Data Handling in Science and Technology: Fundamentals and Analytical Applications of Multiway Calibration* (p. 591). Amsterdam,: Elsevier.

Parastar, H., Akvan, N. (2014). Multivariate curve resolution based chromatographic peak alignment combined with parallel factor analysis to exploit second-order advantage in complex chromatographic measurements. *Analytica Chimica Acta*, pp. 18–27.

Peré-Trepat, E., Tauler, R. (2006). Analysis of environmental samples by application of multivariate curve resolution on fused high-performance liquid chromatography–diode array detection mass spectrometry data. *Journal of Chromatography A*, pp. 85-96.

Pérez, R. L., Escandar, G. M. (2014). Liquid chromatography with diode array detection and multivariate curve resolution for the selective and sensitive quantification of estrogens in natural waters. *Analytica Chimica Acta*, pp. 19-28.

Pinto, L. (2014). Uma metodologia analítica rápida para quantificação simultânea de estrógenos em águas usando HPLC-DAD e calibração de segunda ordem. João Pessoa, Paraíba, Brasil.

Pinto, L., Díaz Nieto, C. H., Zón, M. A., Fernández, H., Araújo, M. C. (2016). Handling time misalignment and rank deficiency in liquid chromatography by multivariate curve resolution: Quantitation of five biogenic amines in fish. *Analytica Chimica Acta*, pp. 59-69.

Romero-González, R., Alarcón-Flores, M. I., Vidal, J. L., Frenich, A. G. (2012). Simultaneous Determination of Four Biogenic and Three Volatile Amines in Anchovy by Ultra-High-Performance Liquid Chromatography Coupled to Tandem Mass Spectrometry. *Journal of Agricultural and Food Chemistry*, pp. 5324-5329.

Sena, M. M., Trevisan, M. G., Poppi, R. J. (2005). Parafac: uma ferramenta quimiométrica para tratamento de dados multidimensionais. Aplicações na determinação direta de fármacos em plasma humano por espectrofluorimetria. *Química Nova*, pp. 910-920.

Simat, V., Dalgaard, P. (2011). Use of small diameter column particles to enhance HPLC determination of histamine and other biogenic amines in seafood. *LWT - Food Science and Technology*, pp. 399-406.

Skoog, D. A., Crouch, S. R., Holler, F. J. (2009). *Principios De Analise Instrumental.* Bookman.

Tahmouzi, S., Khaksar, R., Ghasemlou, M. (2011). Development and validation of an HPLC-FLD method for rapid determination of histamine in skipjack tuna fish (Katsuwonus pelamis). *Food Chemistry*, pp. 756–761.

Tan, G. H., Chai, M.-K. (2011). Sample Preparation in the Analysis of Pesticides Residue in Food by Chromatographic Techniques. Em M. Stoytcheva, *Pesticides - Strategies for Pesticides Analysis* (pp. 27-58). InTech.

Tauler, R., Barceló, D. (1993). Multivariate curve resolution applied to liquid chromatography-diode array detection. *Trends in Analytical Chemistry*, pp. 319-327.

Tomasi, G., Savorani, F., Engelsen, S. B. (2011). i*coshift*: An effective tool for the alignment of chromatographic data. *Journal of Chromatography A*, pp. 7832– 7840.

Tomasi, G., van den Berg, F., Andersson, C. (2004). Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *Journal of chemometrics*, pp. 231–241.

van den Berg, F., Tomasi, G., Viereck, N. (2005). Warping: investigation of NMR pre-processing and correction. Em S. B. Engelsen, P. S. Belton, H. J. Jakobsen , *The Multivariate Challenge* (pp. 131–138). Cambridge: Royal Society of Chemistry.

Vest Nielsen, N. P., Carstensen, J. M., Smedsgaard, J. (1998). Aligning of Single and Multiple Wavelength Chromatographic Profiles for Chemometric Data Analysis Using Correlation Optimised Warping. *Journal of Chromatography A*, pp. 17-35 .

Yadav, K. K., Dasgupta, K., Singh, D. K., Varshney, L., Singh, H. (2015). Dysprosium sorption by polymeric composite bead: Robustparametric optimization using Taguchi method. *Journal of Chromatography A*, pp. 37-43.

Zhai, H., Yang, X., Li, L., Xia, G., Cen, J., Huang, H., Hao, S. (2012). Biogenic amines in commercial fish and fish products sold in southern China. *Food Control*, pp. 303-308.

APÊNDICE

a) Tratamento dos resíduos gerados

Os resíduos gerados na análise podem ser divididos em três soluções com misturas de constituintes distintos: (**1**) o extrato ácido da amostra que contém as ABs; (**2**) solução da reação de derivatização; e (**3**) descarte da análise pelo HPLC. A solução ácida foi neutralizada com hidróxido de sódio e descartada na pia sob fluxo de água, uma vez que em sua composição só há o peixe e ácido perclórico em baixa concentração. Na etapa de derivatização, a solução derivatizada e aferida com acetonitrila foi armazenada em galões cedidos pela UFPB para posterior descarte (incineração). Neste mesmo galão foi realizado o descarte do eluente do HPLC.

O volume de descarte para incineração foi de 5 mL do extrato derivatizado, e 8 mL da eluição, totalizando 13 mL por amostra. Um volume inferior quando comparado a outras metodologias que não fazem uso da calibração de ordem superior, em que a eluição é realizada em corridas longas, como mostrado na **Tabela 2** da **seção 4.9**. A exemplo da metodologia de (Chen, et al., 2010), cuja derivatização foi seguida, o gasto de solvente da eluição cromatográfica foi de 30 mL, enquanto o da metodologia proposta nesta tese foi de 8 mL. Além disso, o tempo de eluição é 7,5 vezes menor, uma redução de trinta para quatro minutos, o que eleva bastante a frequência analítica e reduz o gasto com solventes e incineração.

Outros materiais como papel de filtro, seringa e filtro seringa, uma vez que só contém material sólido proveniente da carne do peixe da etapa de extração e pequena quantidade de sais não tóxicos como $NaHCO_3$ provenientes da etapa de derivatização, foram descartados em embalagens descarpack de papelão, de forma similar a material médico hospitalar.

b) Figuras

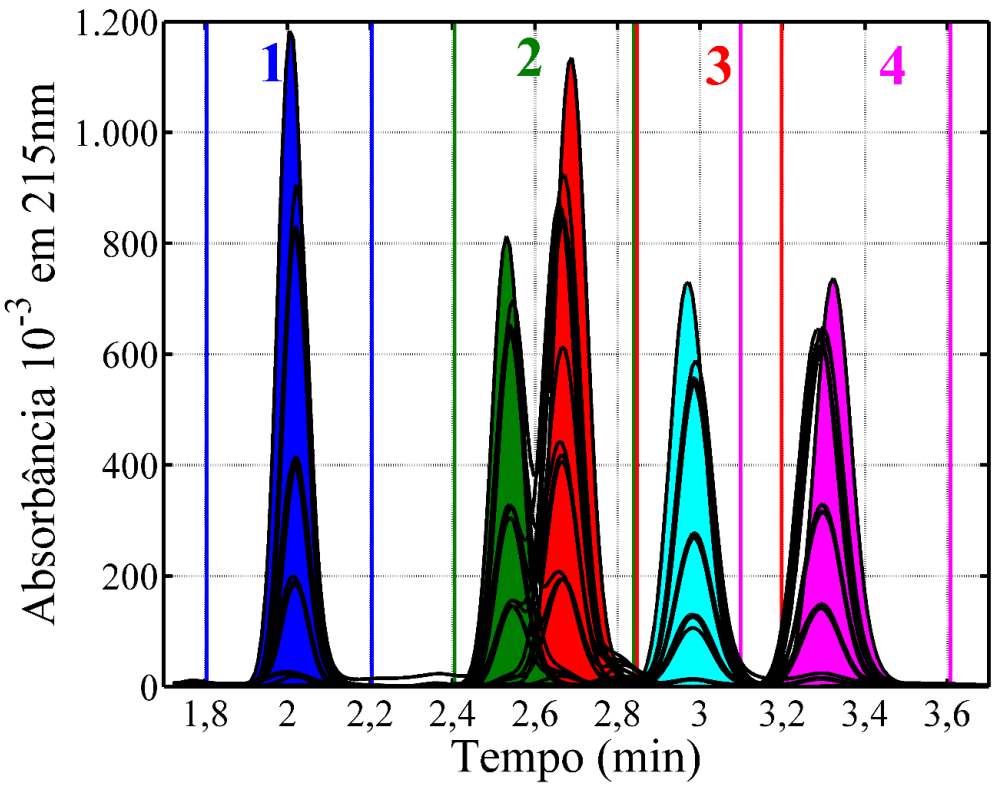

Figura 1S − Cromatograma em registrado em 215 nm para as amostras de teste (—) a amostra de calibração de 20 mg L-1 mostrado como um pico com área colorida. As 4 respectivas faixas de modelagem são definidas como a região entre linhas coloridas para **TRI (—), FEN** e **PUT (—), CAD (—)** e **HIS (—)**.

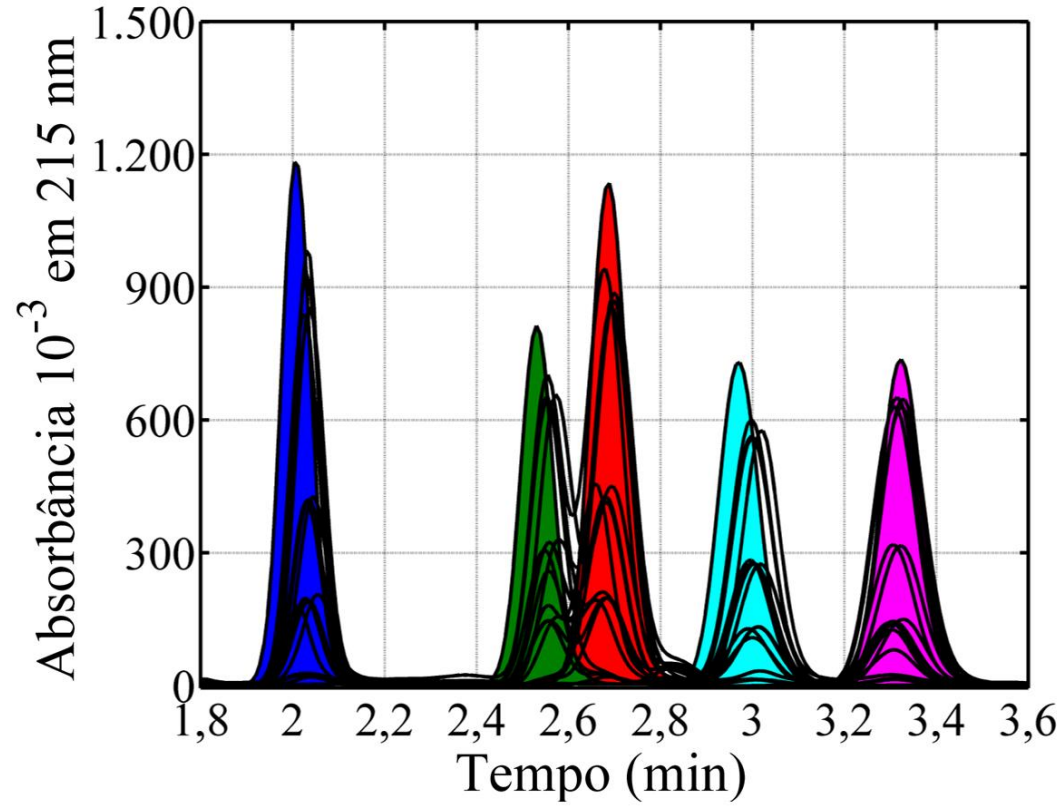

Figura 2S a − Cromatograma em 2D alinhado usando o i*coshift* com 20 intervalos.

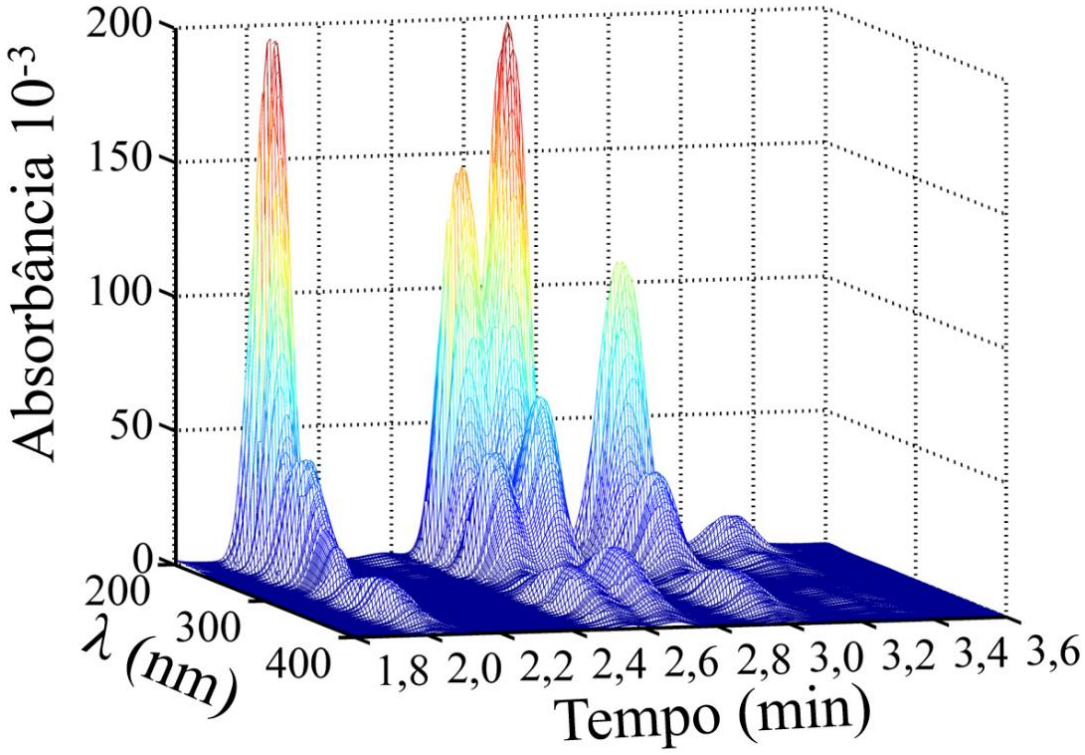

Figura 2S a1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o i*coshift* com 20 intervalos,

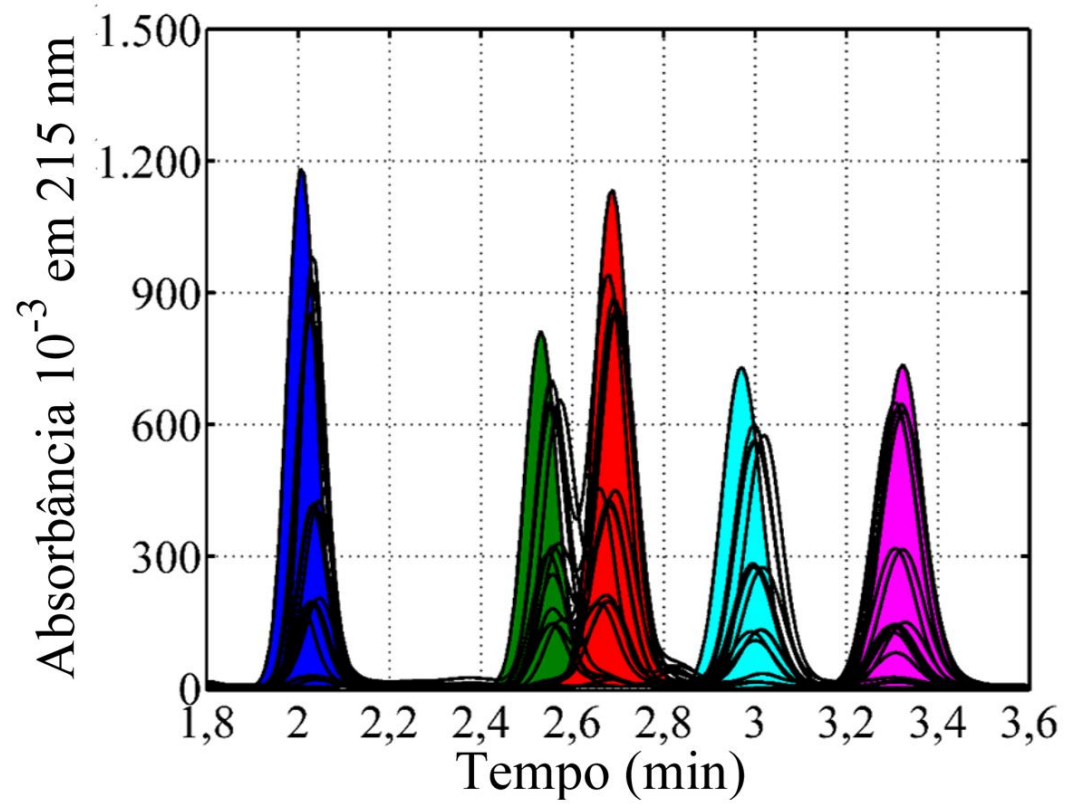

Figura 2S b − Cromatograma em 2D alinhado usando o i*coshift* com 40 intervalos.

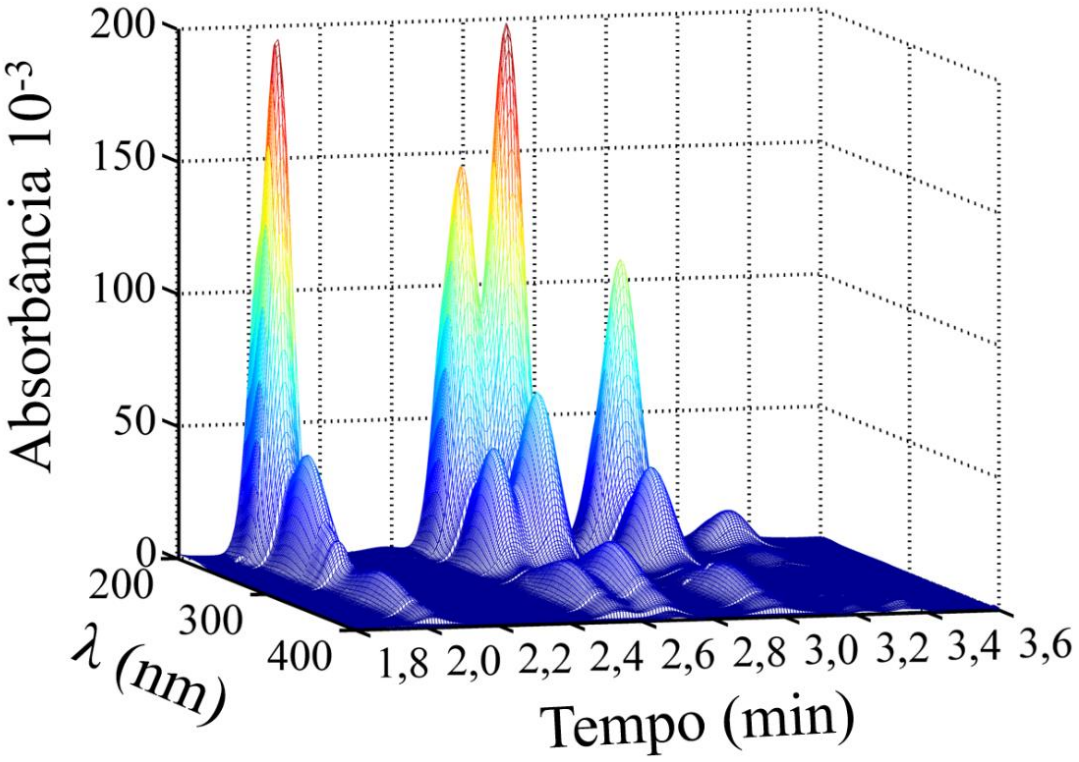

Figura 2S b1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o i*coshift* com 40 intervalos.

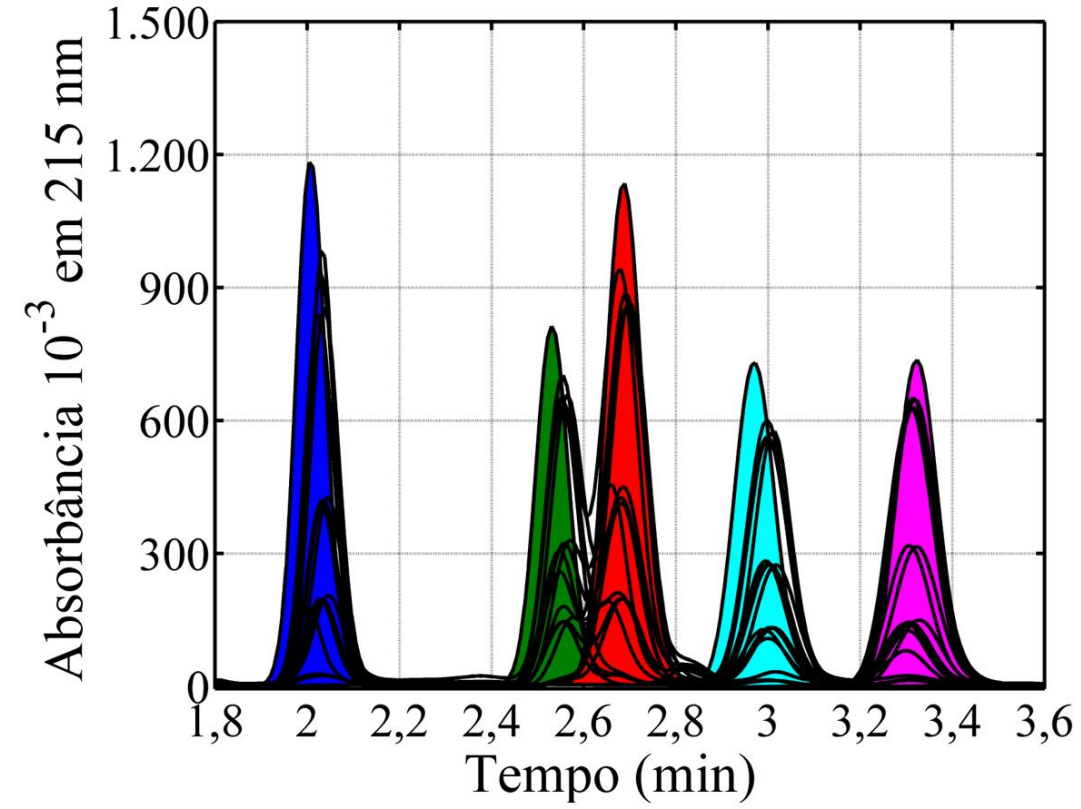

Figura 2S c − Cromatograma em 2D alinhado usando o i*coshift* com 60 intervalos.

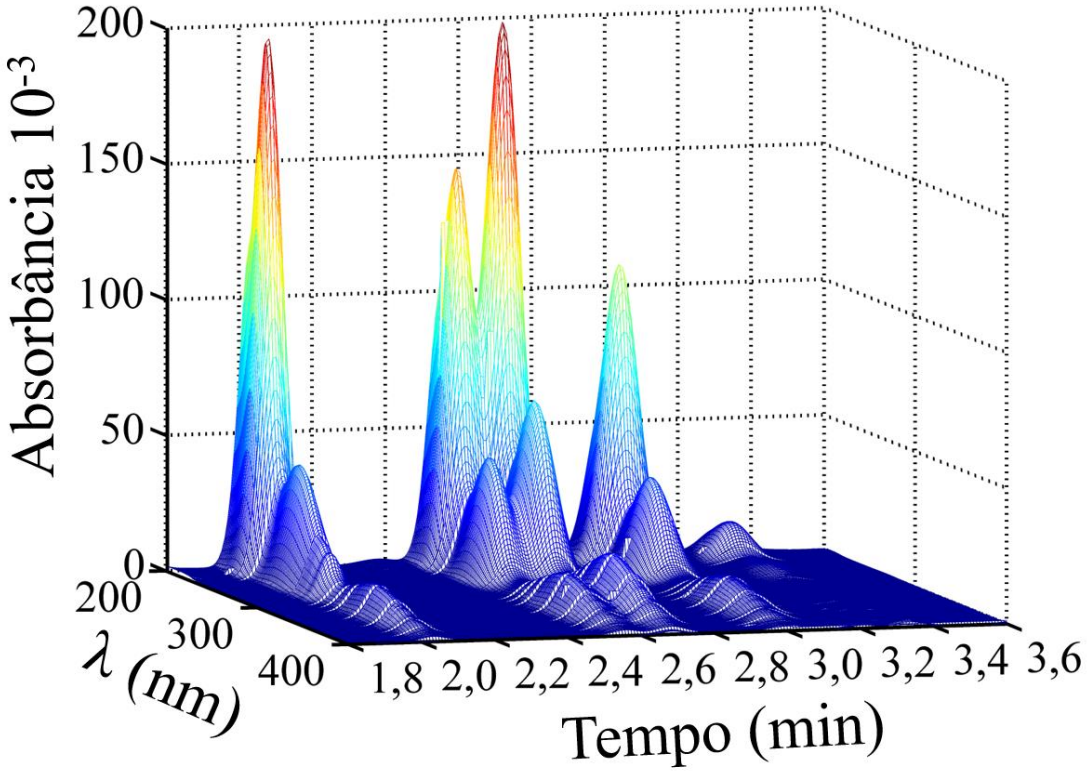

Figura 2S c1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o i*coshift* com 60 intervalos.

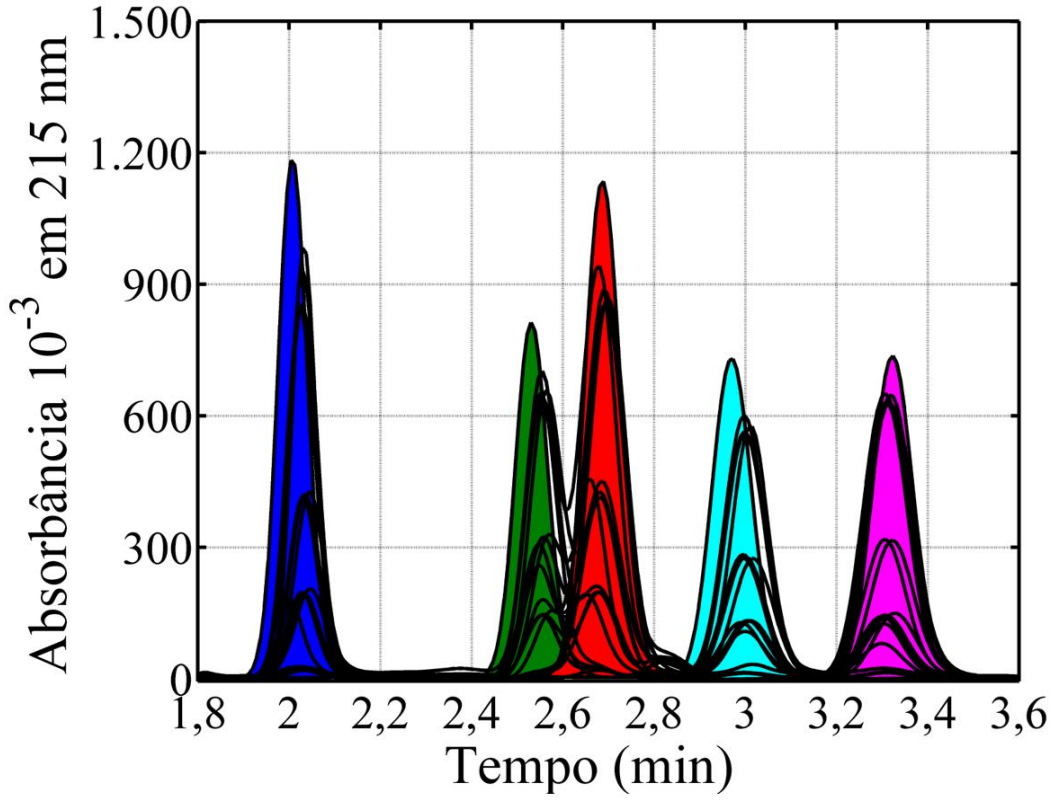

Figura 2S d − Cromatograma em 2D alinhado usando o i*coshift* com 120 intervalos.

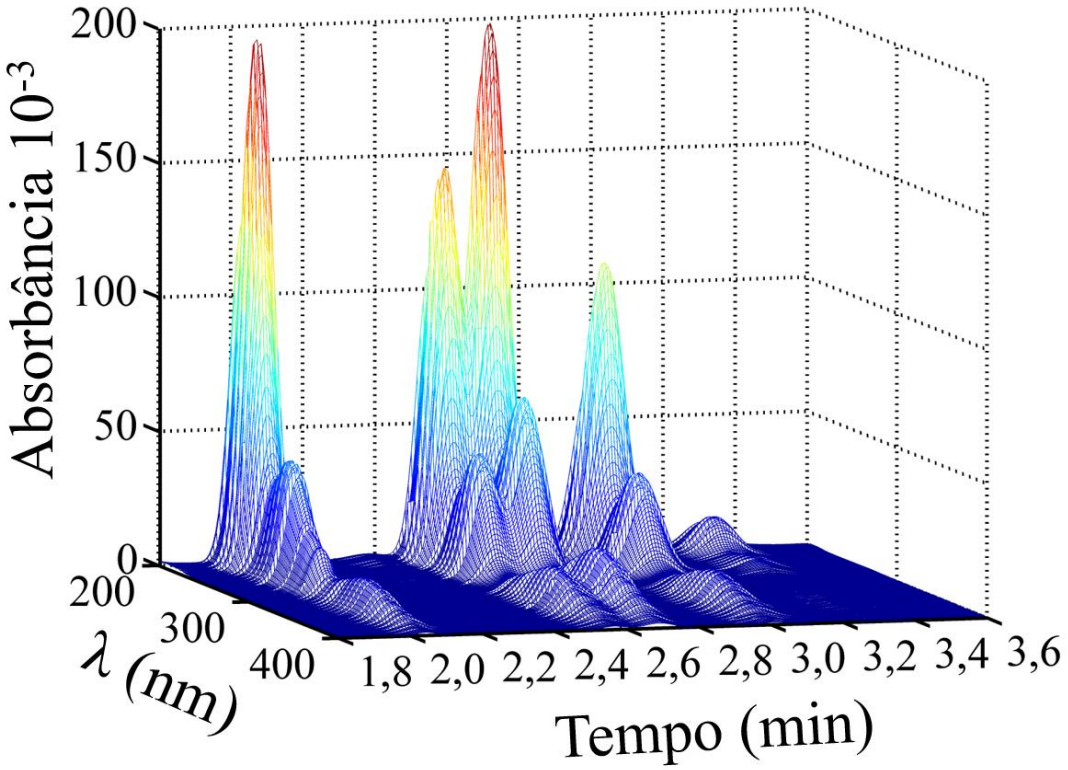

Figura 2S d1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o i*coshift* com 120 intervalos.

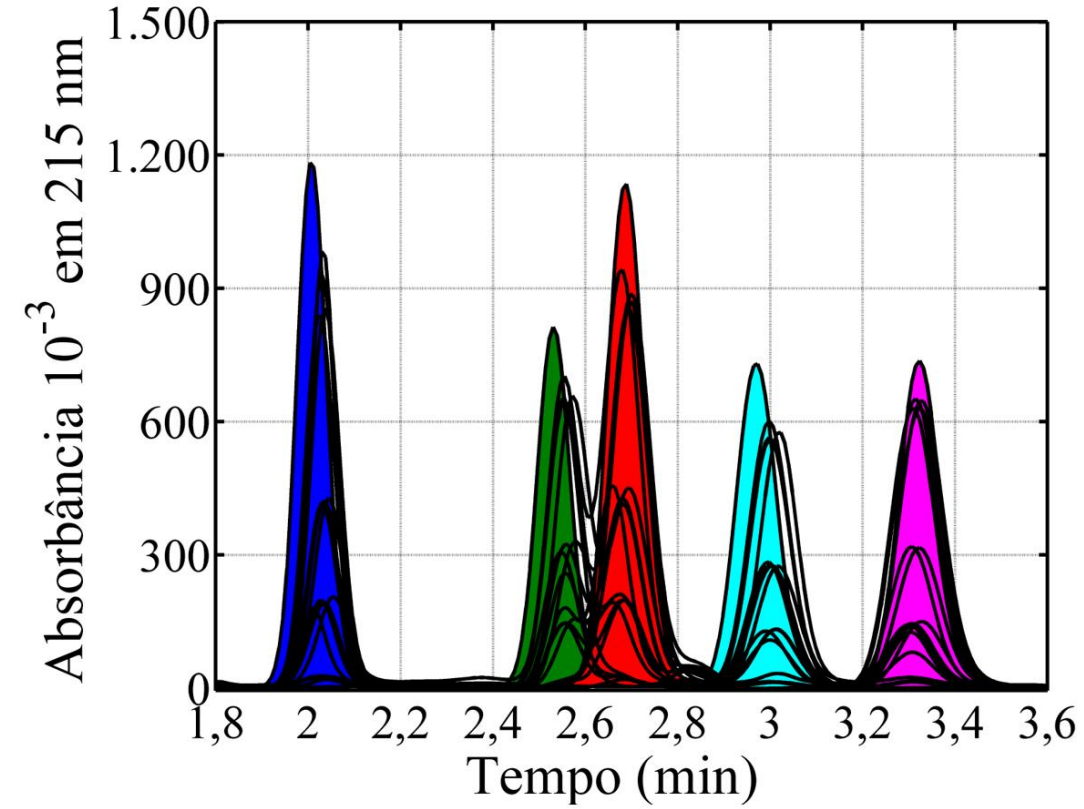

Figura 2S e − Cromatograma em 2D alinhado usando o *coshift* com deslocamento completo do cromatograma.

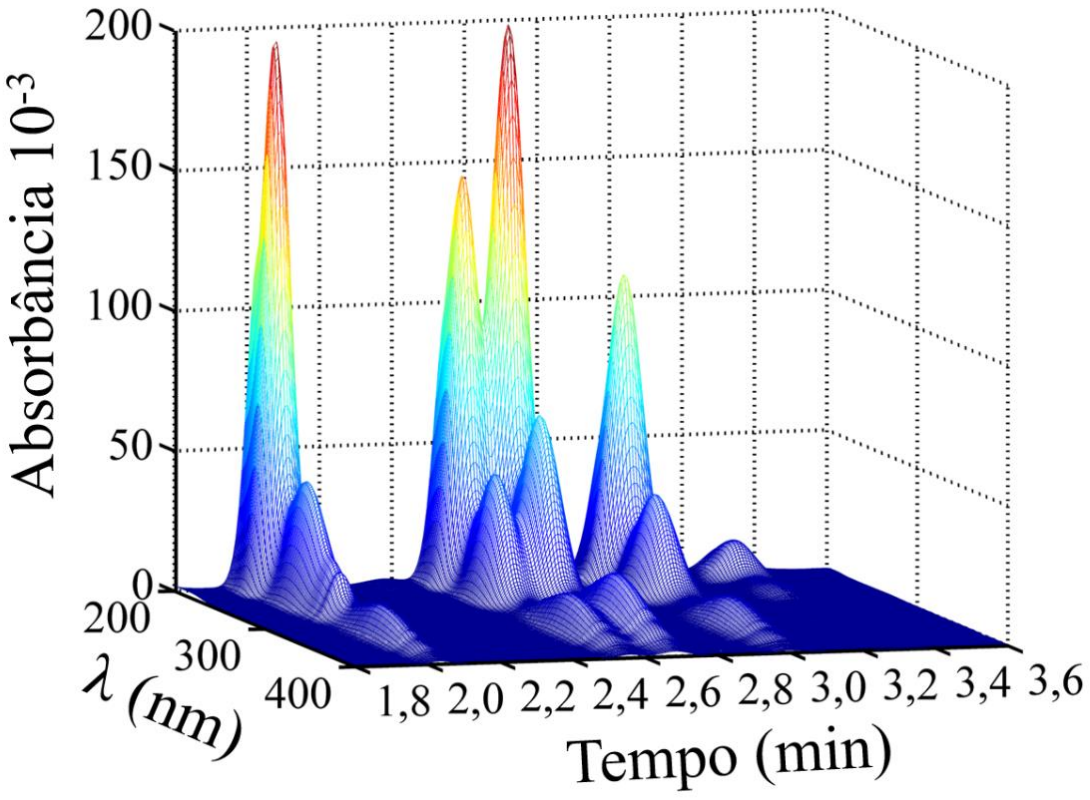

Figura 2S e1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o *coshift* com deslocamento completo do cromatograma,

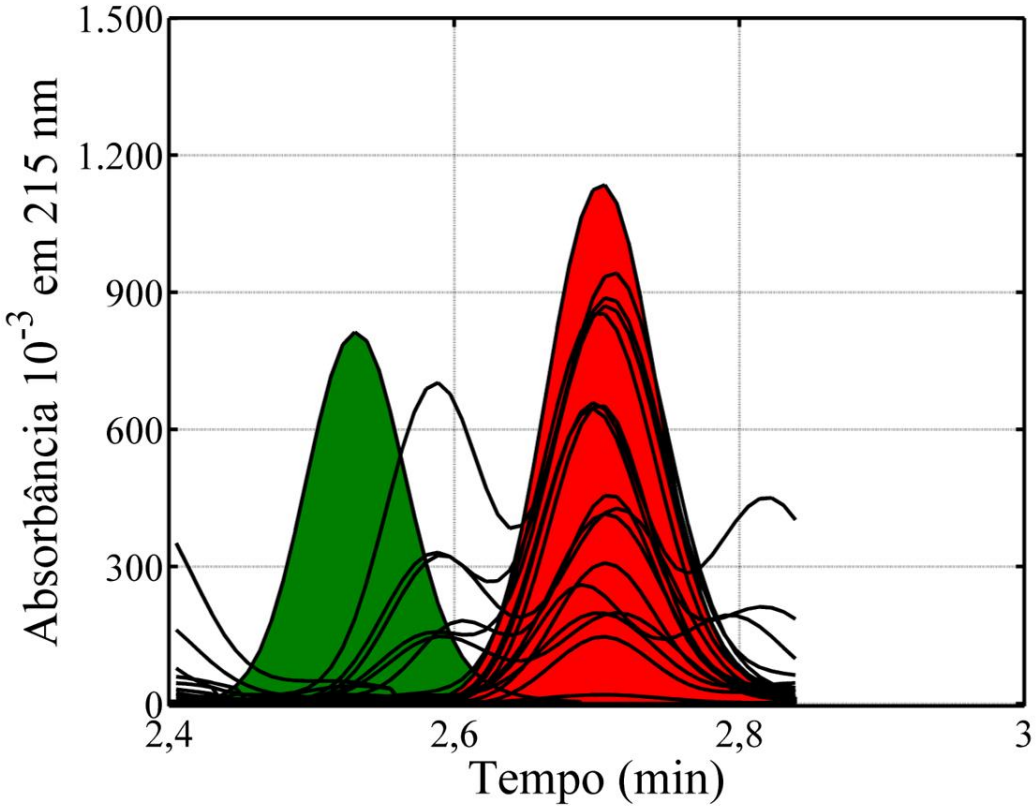

Figura 3S a − Cromatograma em 2D alinhado usando o i*coshift* com 5 intervalos.

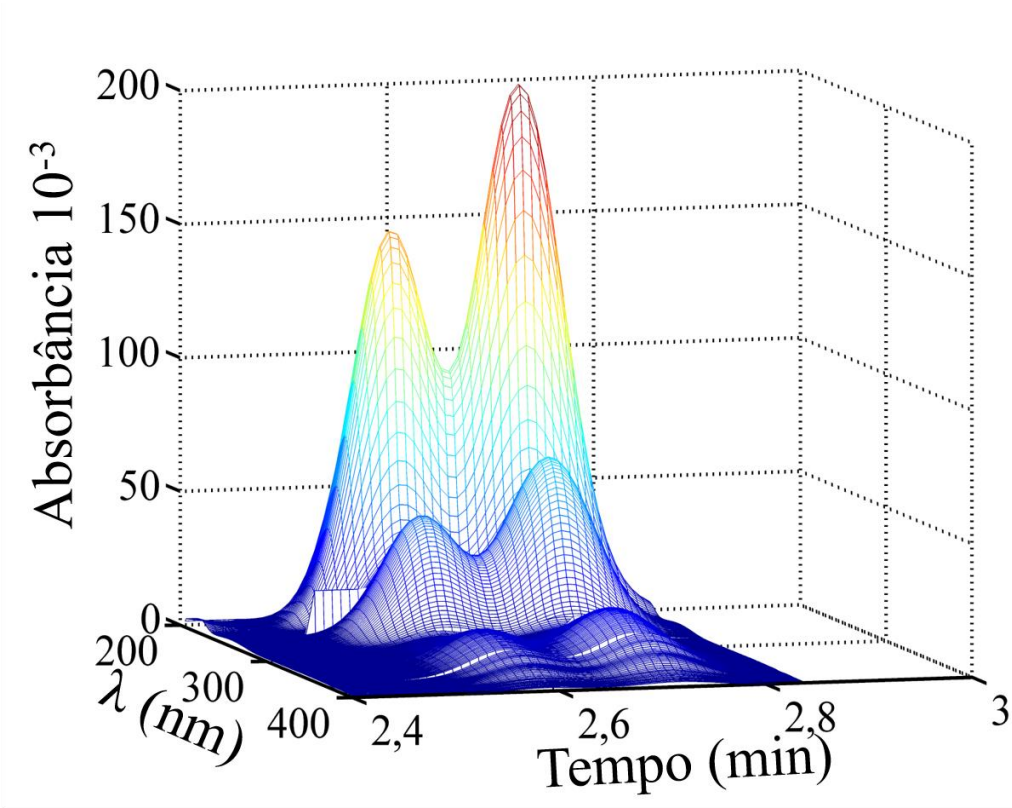

Figura 3S a1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o i*coshift* com 5 intervalos.

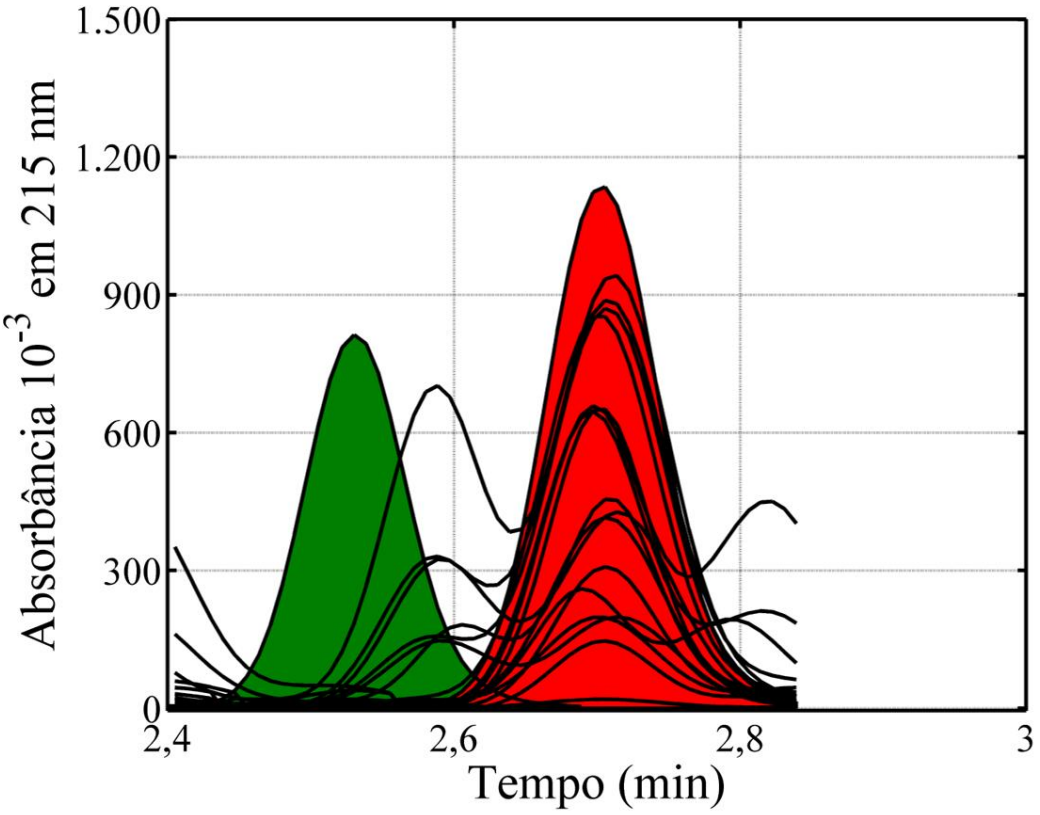

Figura 3S b − Cromatograma em 2D alinhado usando o i*coshift* com 10 intervalos.

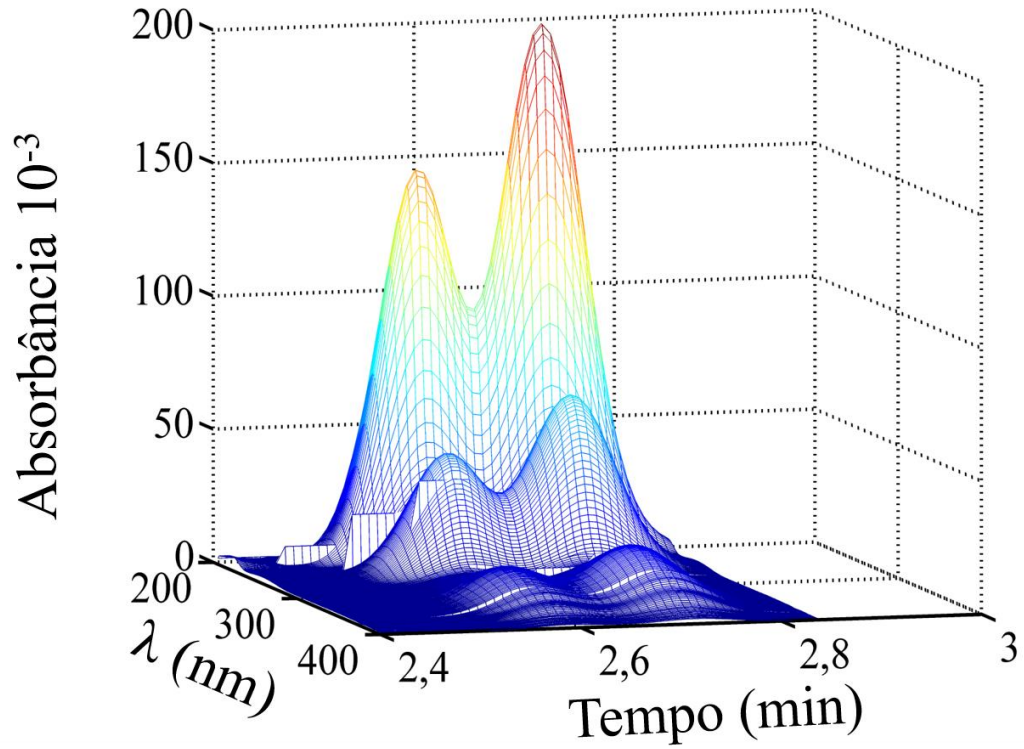

Figura 3S b1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o i*coshift* com 10 intervalos.

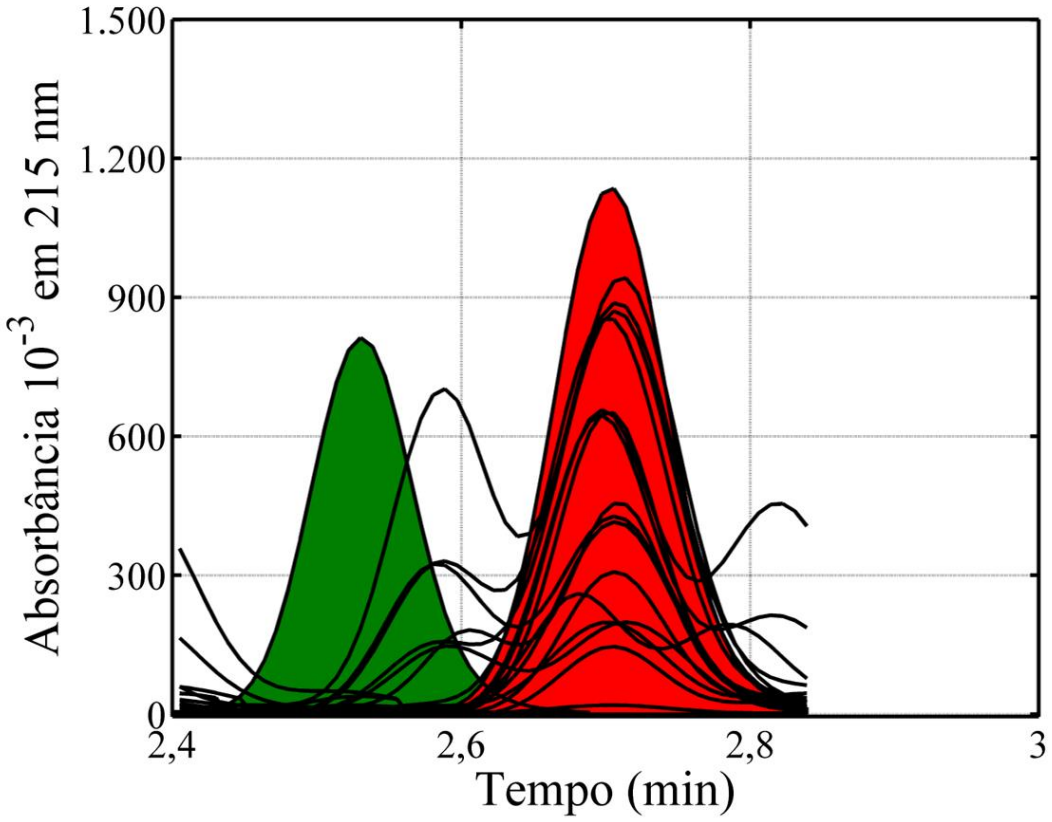

Figura 3S c − Cromatograma em 2D alinhado usando o i*coshift* com 15 intervalos.

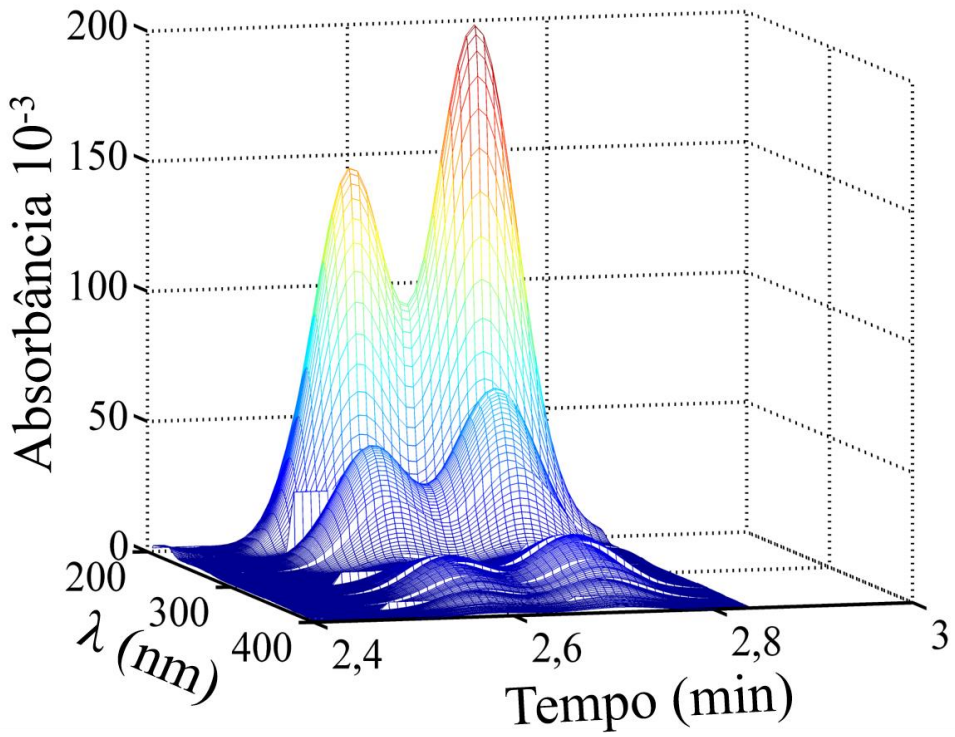

Figura 3S c1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o i*coshift* com 15 intervalos.

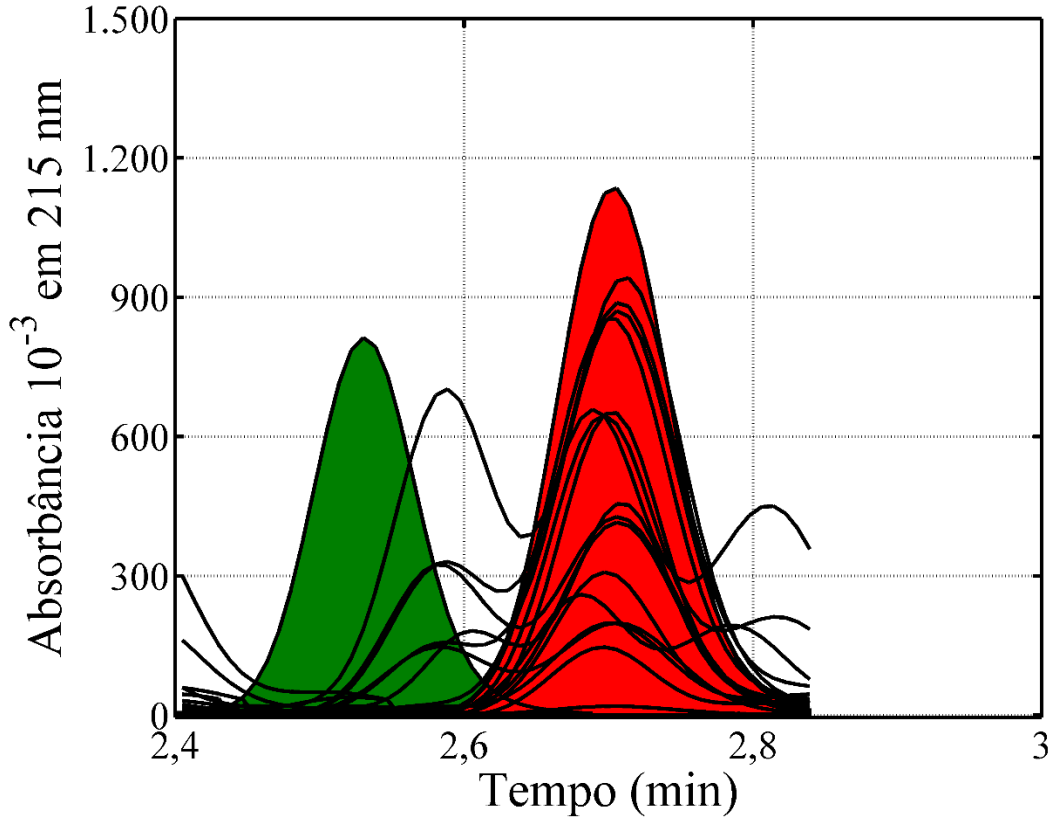

Figura 3S d − Cromatograma em 2D alinhado usando o i*coshift* com 30 intervalos.

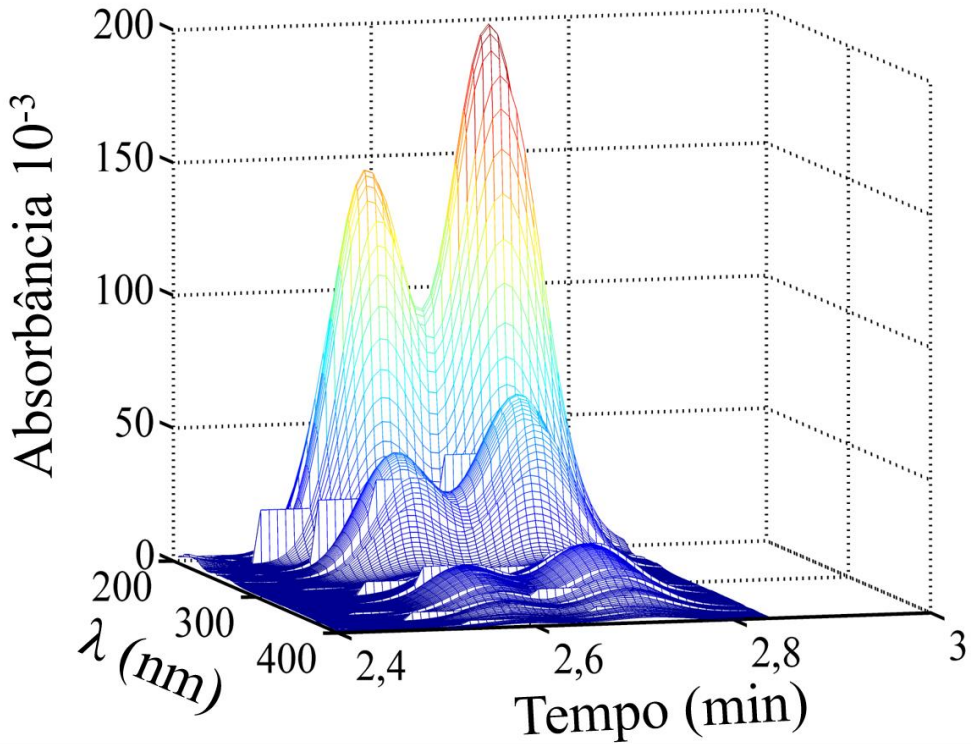

Figura 3S d1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o i*coshift* com 30 intervalos.

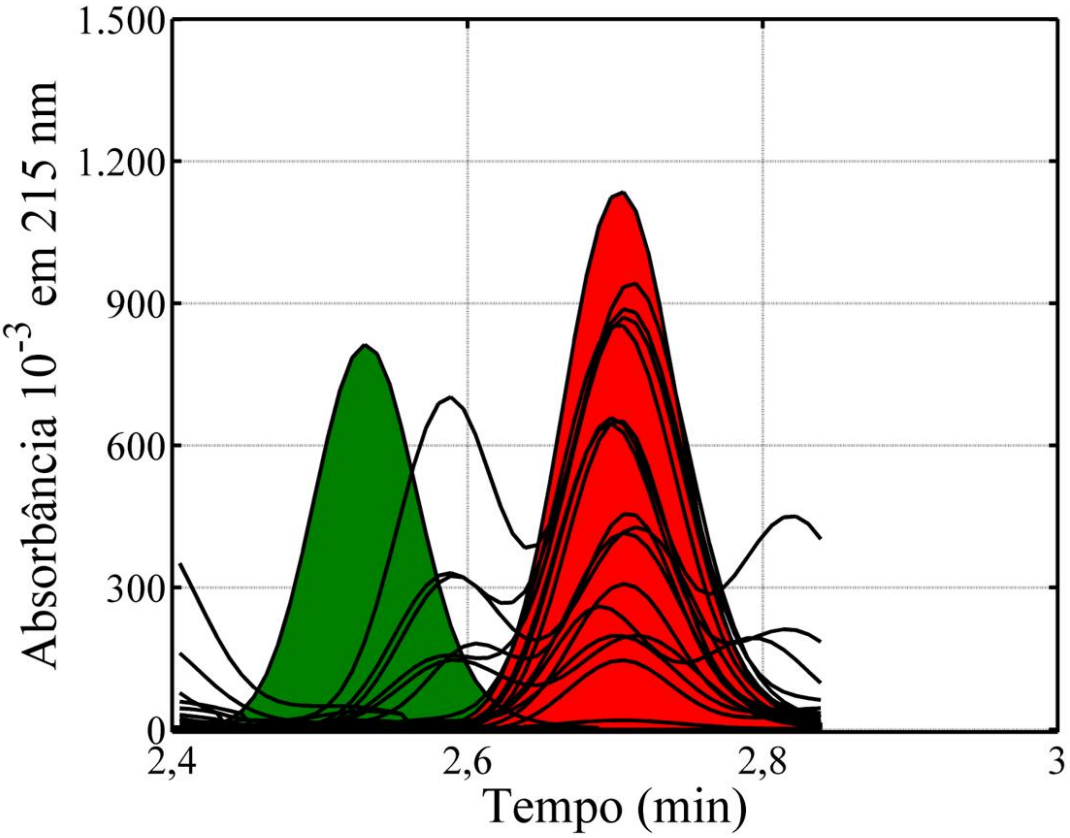

Figura 3S e − Cromatograma em 2D alinhado usando o *coshift* com deslocamento completo do cromatograma.

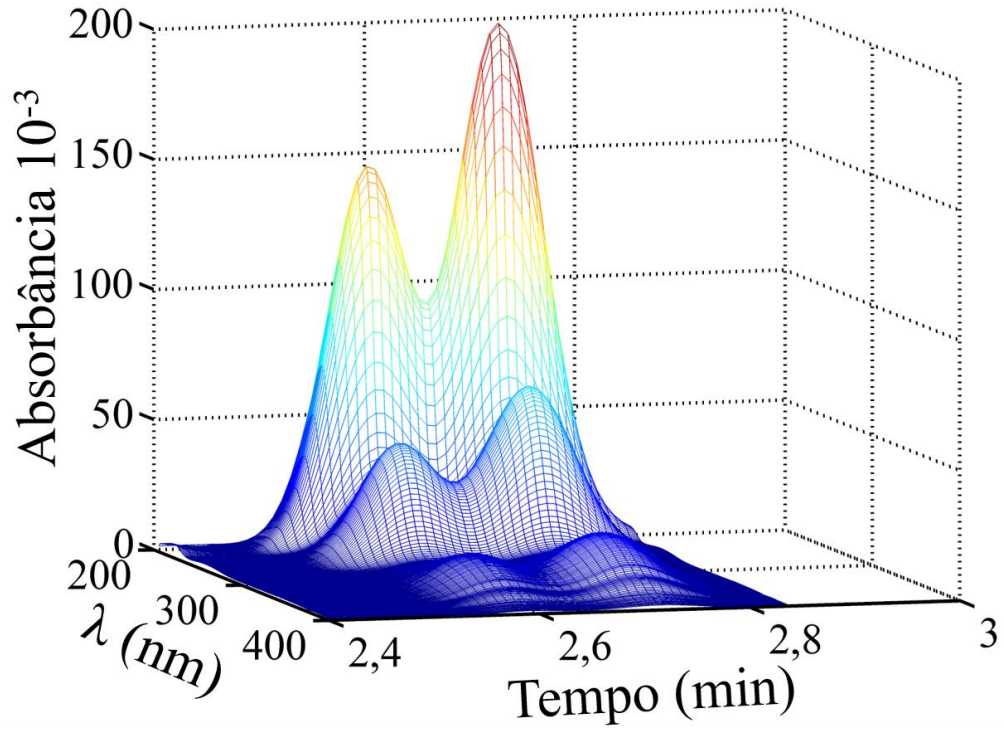

Figura 3S e1 − Cromatograma LC-DAD da amostra de teste 2, alinhado usando o *coshift* com deslocamento completo do cromatograma,

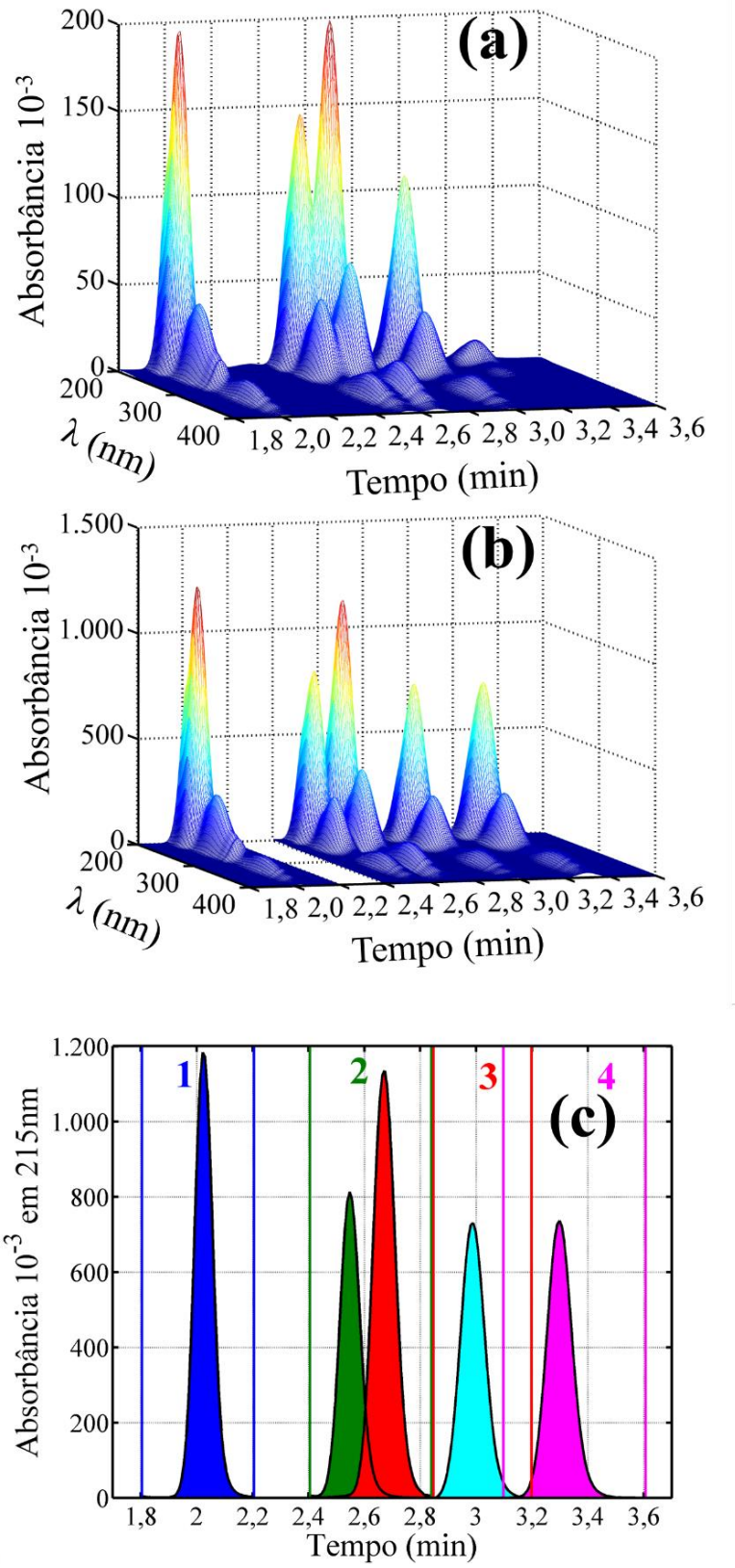

Figura 4S − (**a**) Cromatograma em 3D da amostra de teste 2 bruta registrado entre 1,8 e 3,6 minutos e na faixa espectral de 200 a 400 nm.(**b**)Cromatograma em 3D da amostra de calibração de 20 mg L^{-1} alinhados de acordo com cada faixa modelada das aminas biogênicas, (**c**) Cromatograma da amostra de calibração de 20mg L^{-1} alinhada registrada em 215nm,e as 4 respectivas faixas de modelagem é definida como a região entre linhas coloridas para **TRI (—), FEN** e **PUT (—), CAD (—)** e **HIS(—)**.

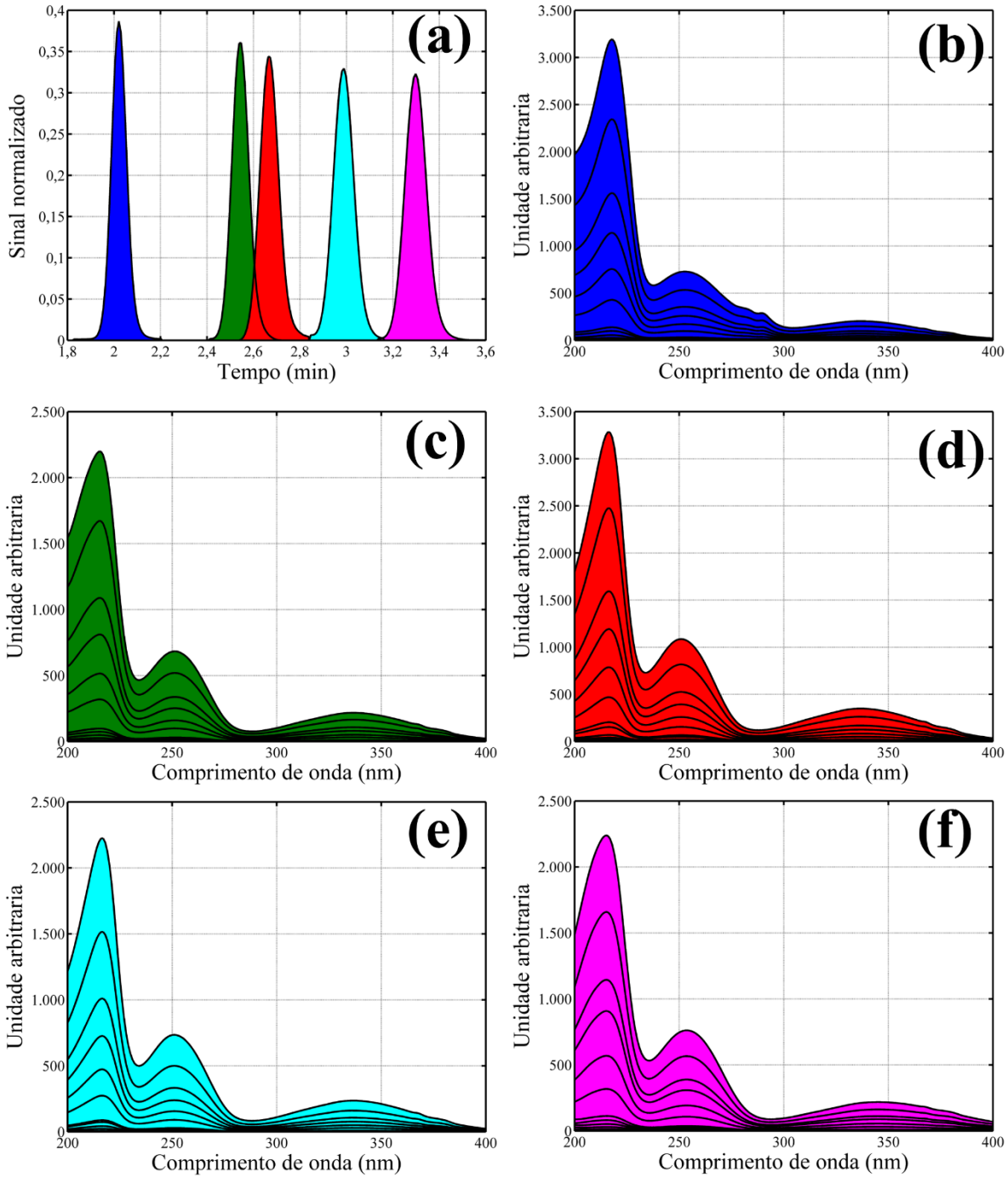

Figura 5S − Perfis otimizados pelo MCR-ALS para o conjunto de calibração com os dados aumentados no modo dos espectros. Em (**a**) é mostrado os cromatogramas normalizados, e os espectros sobrepostos para TRI, FEN, PUT, CAD e HIS são mostrados respectivamente em (**b**), (**c**), (**d**), (**e**) e (**f**).

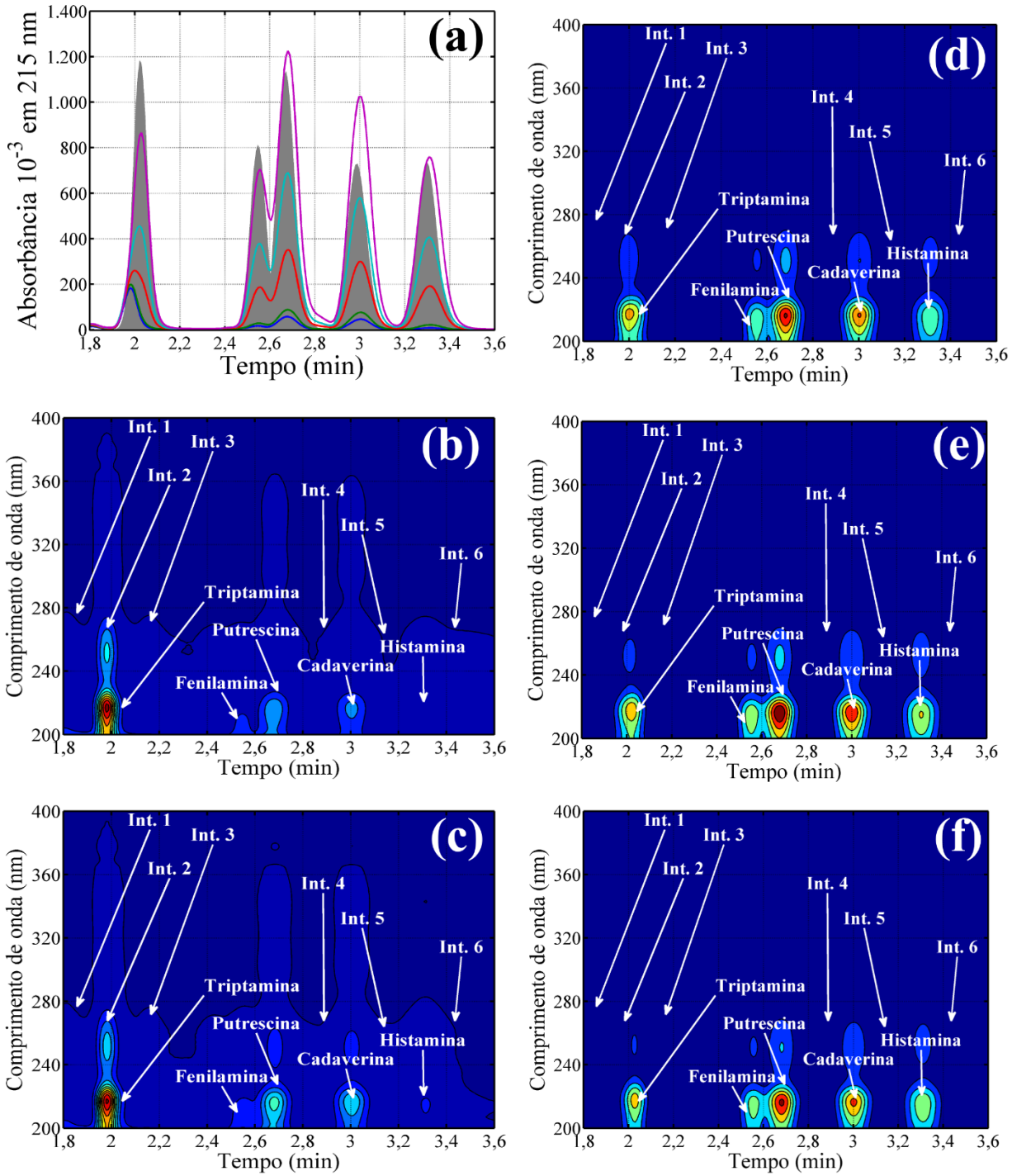

Figura 6S – Cromatograma em 2 D (**a**) e em superfície de contorno da amostra de *Salminus maxillosus* sem fortificação (—) (**b**) e fortificadas no nível 1 (—) (**c**), 2 (—) (**d**), 3 (—) (**e**) e 4 (—) (**f**), segundo o procedimento descrito na **seção 3.6**. Os analitos e interferentes (INT) são apontados com setas. Em (**a**) o pico cinza representa o sinal do analito na concentração de 20mg L^{-1}.

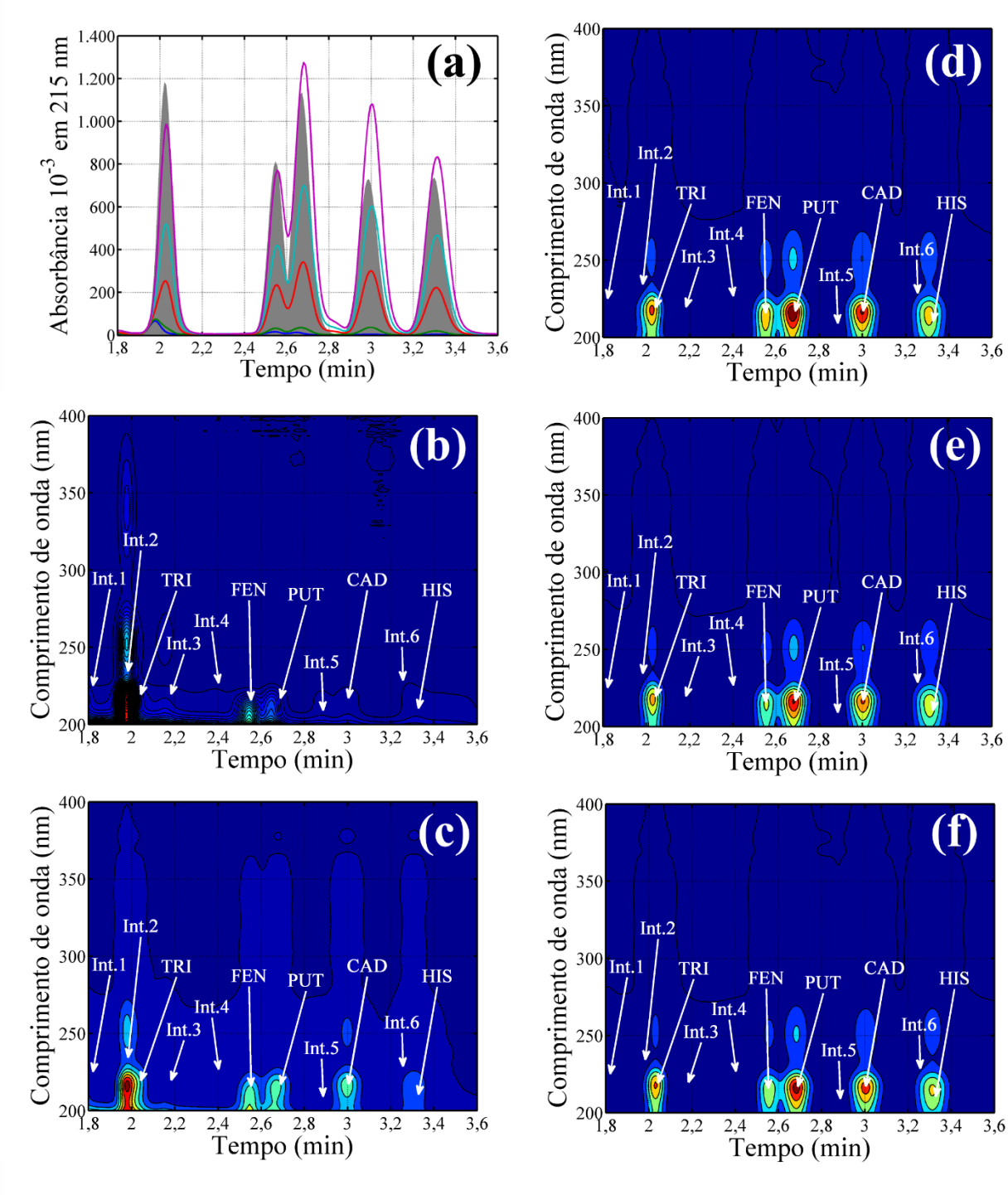

Figura 7S – Cromatograma em 2 D (**a**) e em superfície de contorno da amostra de *Scomber scombrus* sem fortificação (—) (**b**) e fortificadas no nível 1 (—) (**c**), 2 (—) (**d**), 3 (—) (**e**) e 4 (—) (**f**), segundo o procedimento descrito na **seção 3.6**. Os analitos e interferentes (INT) são apontados com setas. Em (**a**) o pico cinza representa o sinal do analito na concentração de 20mg L^{-1}.

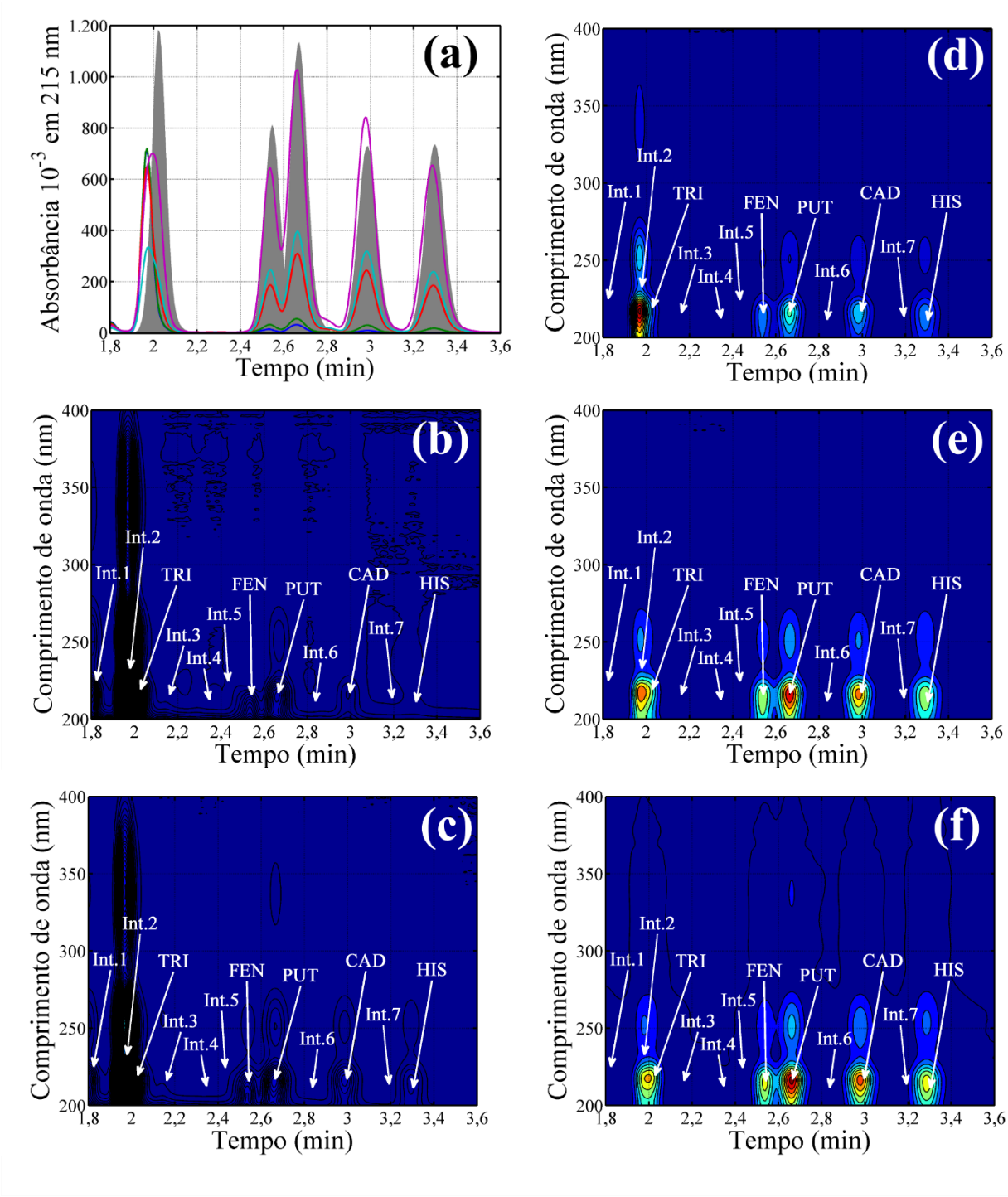

Figura 8S – Cromatograma em 2 D (**a**) e em superfície de contorno da amostra de *Lutjanus analis* sem fortificação (—) (**b**) e fortificadas no nível 1 (—) (**c**), 2 (—) (**d**), 3 (—) (**e**) e 4 (—) (**f**), segundo o procedimento descrito na **seção 3.6**. Os analitos e interferentes (INT) são apontados com setas. Em (**a**) o pico cinza representa o sinal do analito na concentração de 20mg L^{-1}.

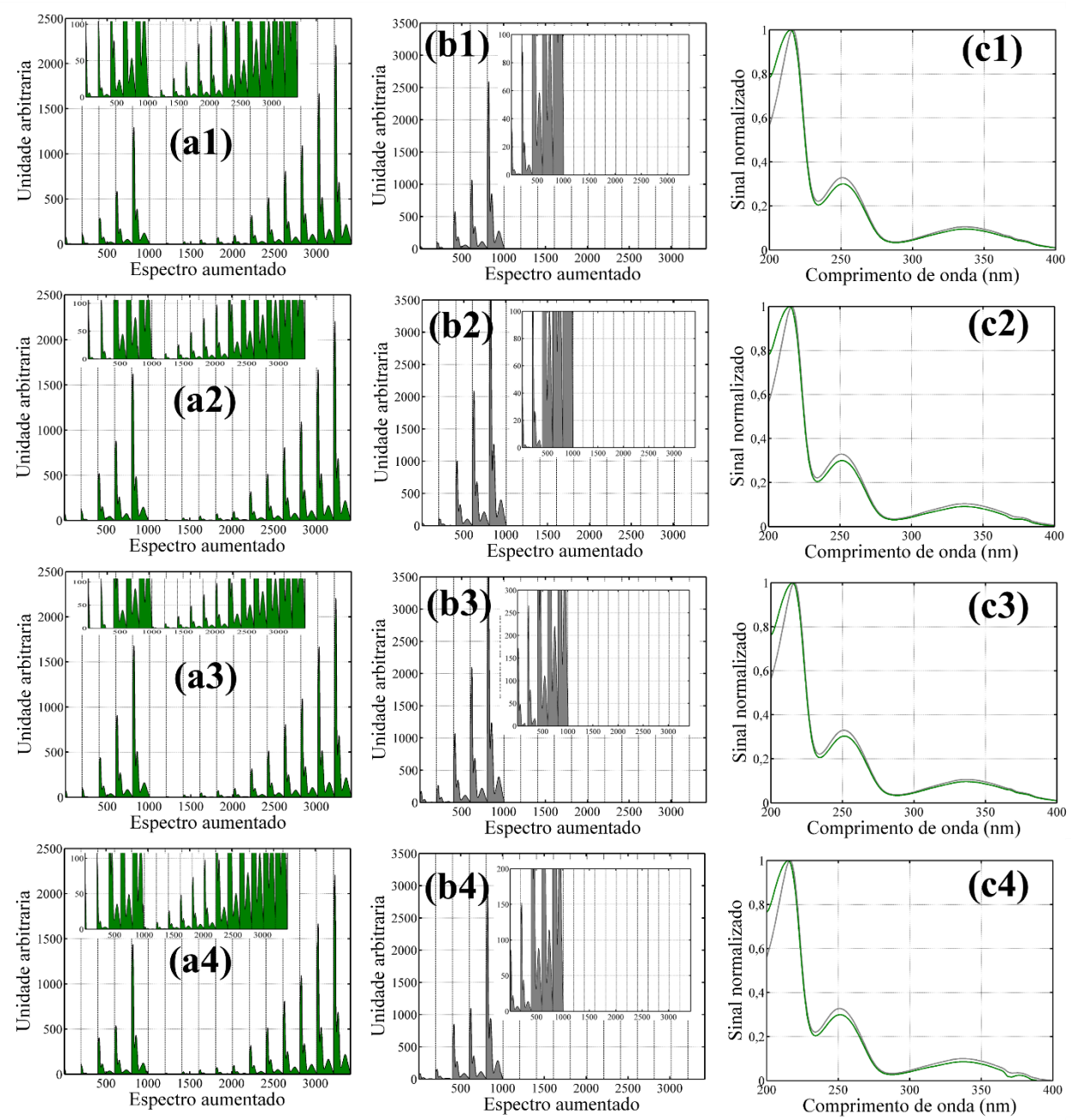

Figura 9S − Perfil espectral aumentado recuperado pelo MCR-ALS para a FEN (**a1-a4**) e para o interferente (**b1-b4**) nas fortificações das amostras de *Salminus maxillosus* (1), *Scomber scombrus* (2), *Seriola dumerili* (3) e *Lutjanus analis* (4) realizadas no primeiro dia de análise. O espectro da FEN e do interferente em uma concentração foi normalizado para comprimento 1 e sobreposto para comparação (**c1-c4**).

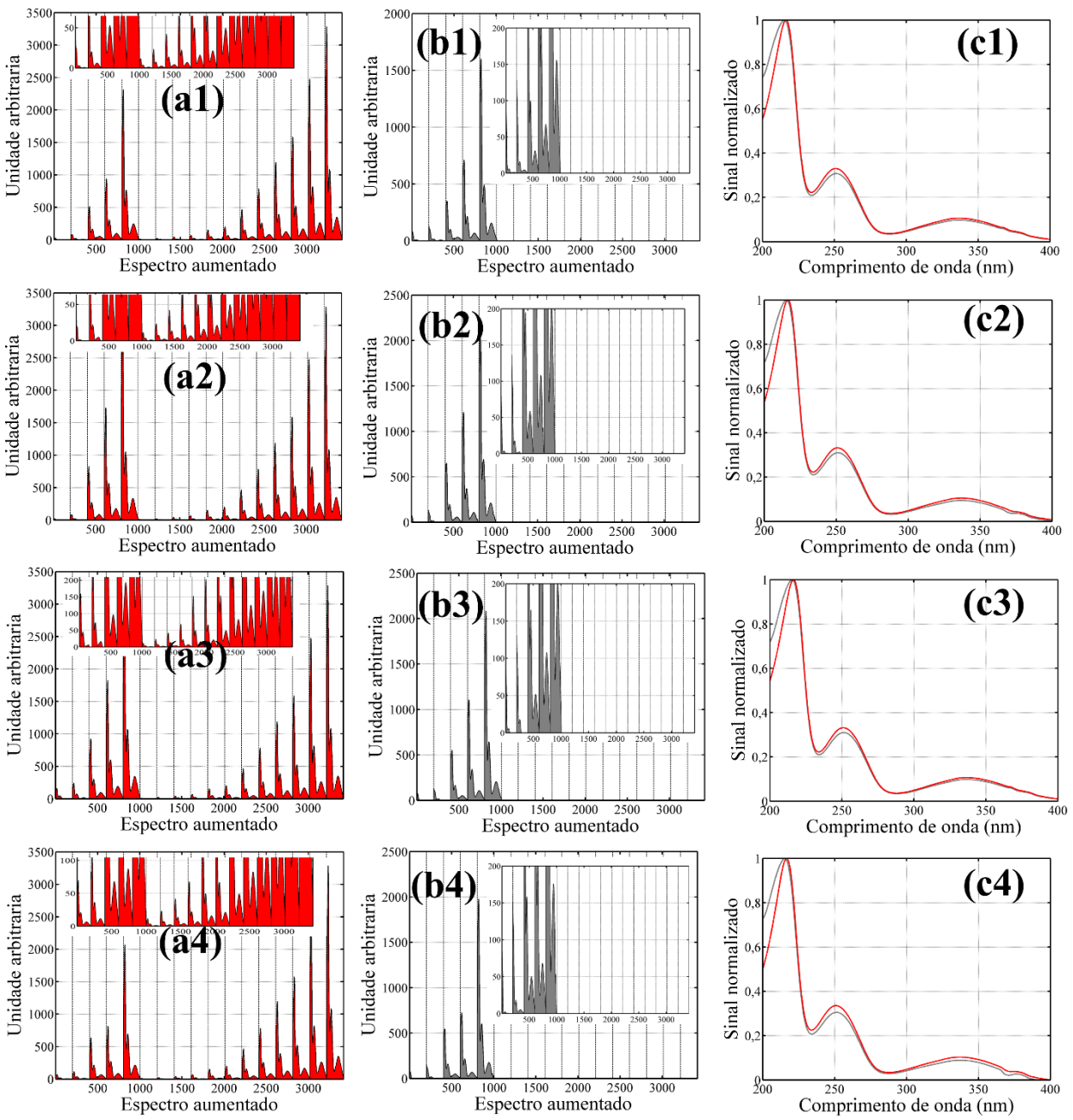

Figura 10S − Perfil espectral aumentado recuperado pelo MCR-ALS para a PUT (**a1-a4**) e para o interferente (**b1-b4**) nas fortificações das amostras de *Salminus maxillosus* (1), *Scomber scombrus* (2), *Seriola dumerili* (3) e *Lutjanus analis* (4) realizadas no primeiro dia de análise. O espectro da PUT e do interferente em uma concentração foi normalizado para comprimento 1 e sobreposto para comparação (**c1-c4**).

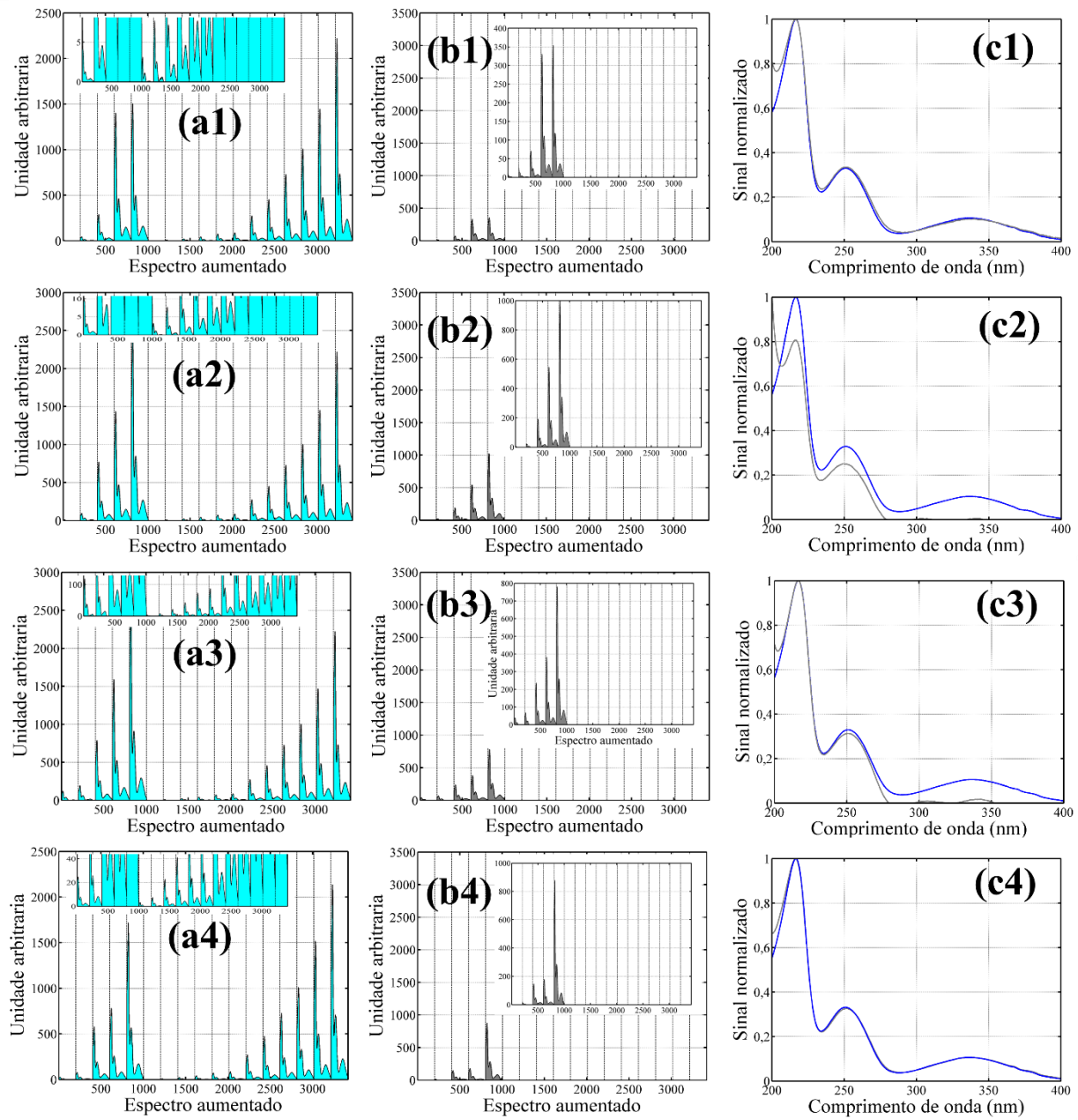

Figura 11S − Perfil espectral aumentado recuperado pelo MCR-ALS para a CAD (**a1-a4**) e para o interferente (**b1-b4**) nas fortificações das amostras de *Salminus maxillosus* (1), *Scomber scombrus* (2), *Seriola dumerili* (3) e *Lutjanus analis* (4) realizadas no primeiro dia de análise. O espectro da CAD e do interferente em uma concentração foi normalizado para comprimento 1 e sobreposto para comparação (**c1-c4**).

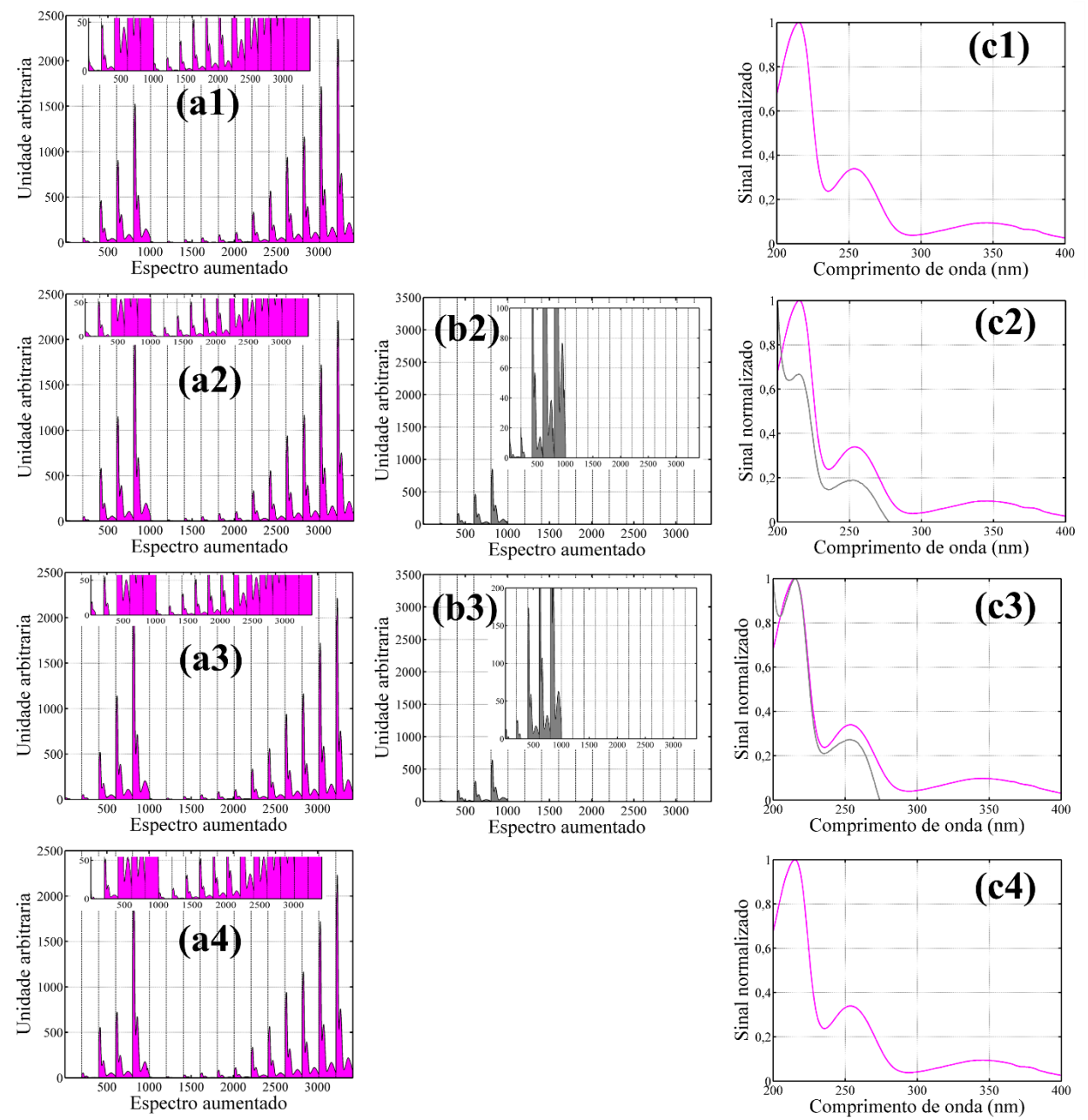

Figura 12S − Perfil espectral aumentado recuperado pelo MCR-ALS para a HIS (**a1-a4**) e para o interferente (**b2-b3**) nas fortificações das amostras de *Salminus maxillosus* (1), *Scomber scombrus* (2), *Seriola dumerili* (3) e *Lutjanus analis* (4) realizadas no primeiro dia de análise. O espectro da HIS e do interferente em uma concentração foi normalizado para comprimento 1 e sobreposto para comparação (**c1-c4**).

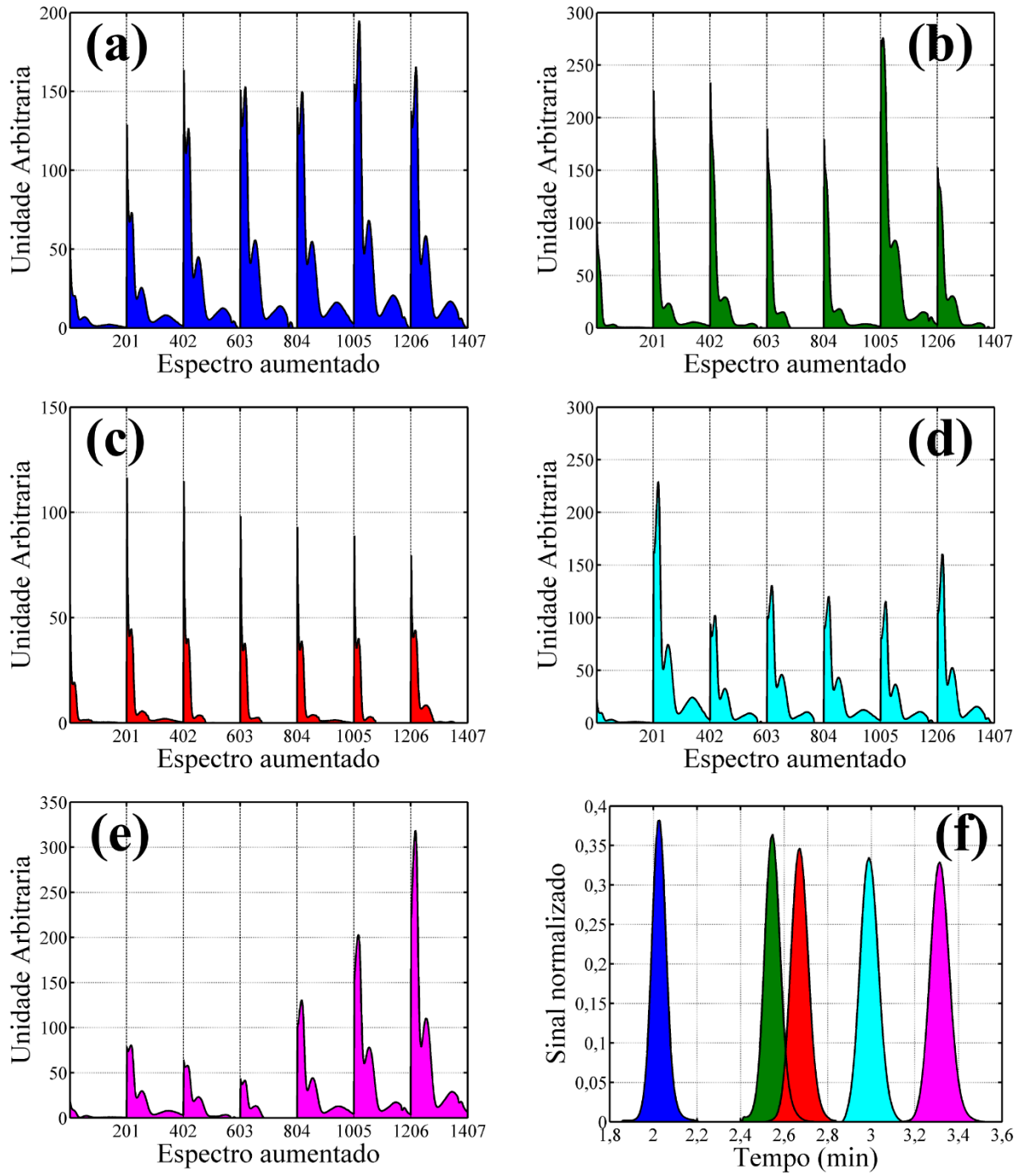

Figura 13S − Espectros aumentados da TRI (**a**), FEN (**b**), PUT (**c**), CAD (**d**) e HIS (**e**), e cromatogramas normalizado (**f**), todos recuperados pelo MCR-ALS para a amostra de *Scomber Scombus* nos 13 dias de estudo cinético mostrado nos cromatogramas da **Figura 36**.

c) TABELAS

Tabela 1S − Região modelada para todas as amostras e todos os algoritmos.

Analitos	TRI	FEN	PUT	CAD	HIS
Região modelada (min)	1,8-2,2	2,4-2,84	2,4-2,84	2,84-3,2	3,1-3,6

Tabela 2S: Efeito da restrição de trilinearidade na ambiguidade rotacional do MCR-ALS no modelo das amostras de teste.

Restrição [a]	f	TRI (1)*	FEN (2) *	PUT (2) *	CAD (2) *	HIS (1) *
1,2,3,4	f_{max}	1	0,6580	0,8504	0,9825	1
	f_{ini}	1	0,6580	0,8504	0,9825	1
	f_{min}	1	0,6577	0,8504	0,9825	1
	$\mathbf{f_{max-min}}$	0	0,0003	0,0000	0,0000	0
1,2,3,4,5	f_{max}	1	0,6583	0,8503	0,9825	1
	f_{ini}	1	0,6583	0,8503	0,9825	1
	f_{min}	1	0,6583	0,8503	0,9825	1
	$\mathbf{f_{max-min}}$	0	0,0000	0,0000	0,0000	0

[a] Restrição: 1 normalização; 2 não-negatividade; 3 unimodalidade; 4 correspondência entre as espécies; 5 trilinearidade. f_{max}– f_{min} corresponde à diferença entre a norma de Frobenious máxima e mínima.
* Entre parênteses é mostrado o número de constituintes recuperados no modelo.

Tabela 3S − Ambiguidade rotacional do MCR-ALS no modelo da amostra de peixe.

Amostra [a]	f	TRI (2)*	FEN (2) *	PUT (2) *	CAD (2) *	HIS (2) *
Salminus maxillosus	f_{ini}	0,7719	0,7563	0,9138	0,9365	1
	f_{max}	0,7719	0,7563	0,9138	0,9365	1
	f_{min}	0,7719	0,6895	0,9138	0,9365	1
	$\mathbf{f_{max-min}}$	0	0,0668	0	0	0
Scomber scombrus	f_{ini}	0,9326	0,6243	0,8472	0,8566	0,8891
	f_{max}	0,9326	0,6243	0,8472	0,8566	0,8891
	f_{min}	0,9326	0,6243	0,8472	0,7257	0,8424
	$\mathbf{f_{max-min}}$	0	0	0	0,1309	0,0467
Lutjanus analis	f_{ini}	0,6853	0,7261	0,8592	0,877	1
	f_{max}	0,6853	0,7268	0,8592	0,877	1
	f_{min}	0,6853	0,7261	0,8516	0,877	1
	$\mathbf{f_{max-min}}$	0	0,0007	0,0076	0	0

[a] Restrição: 1 normalização; 2 não-negatividade; 3 unimodalidade; 4 correspondência entre as espécies; f_{max}– f_{min} é a diferença entre a norma de Frobenious máxima e mínima.
* Entre parênteses é mostrado o número de constituintes recuperados no modelo. Para a amostra de *Salminus maxillosus e Lutjanus analis* foi necessário apenas 1 componente na modelagem da HIS.

Tabela 4S − Teste de recuperação das aminas biogênicas nas amostras de peixe usando o MCR-ALS.

Peixe	TRI[a]			FEN[a]			PUT[a]			CAD[a]			HIS[a]		
	N[b]	P[c]	R(%)	N[b]	P[c]	R(%)	N[b]	P[c]	R(%)	N[b]	P[c]	R(%)	N[b]	P[c]	R(%)
Salminus maxillosus	0	0,20(0,06)		0	0,48(0,04)		0	0,15(0,07)		0	0,32(0,21)		0	0,04(0,10)	
	0,4	0,61(0,06)		0,4	0,71(0,04)		0,4	0,47(0,06)	81,03	0,4	0,69(0,20)	91,65	0,4	0,37(0,10)	81,08
	4	4,33(0,05)		4	2,60(0,03)		4	3,15(0,05)	74,98	4	3,00(0,18)	66,93	4	4,02(0,08)	99,48
	8	6,81(0,05)		8	5,33(0,03)		8	5,76(0,05)	70,19	8	13,44(0,27)	163,91	8	7,92(0,08)	98,52
	16	11,35(0,07)		16	11,75(0,04)		16	14,17(0,09)	87,67	16	14,40(0,29)	88,00	16	13,38(0,13)	83,35
Inicial		**2,02**			**4,80**			**1,48**	**78,47**		**3,23**	**82,19**		**0,42**	**90,60**
Scomber scombrus	0	0,37(0,07)		0	0,42(0,04)		0	0,12(0,07)		0	0,38(0,21)		0	0,01(0,12)	
	0,4	0,70(0,07)	82,00	0,4	0,69(0,04)	69,58	0,4	0,46(0,07)		0,4	1,13(0,20)		0,4	0,33(0,12)	79,83
	4	4,40(0,06)	100,83	4	4,64(0,03)	105,66	4	5,01(0,05)		4	7,47(0,17)		4	5,04(0,09)	125,73
	8	9,00(0,07)	107,91	8	7,76(0,04)	91,80	8	10,42(0,07)		8	13,57(0,27)		8	9,99(0,12)	124,70
	16	16,86(0,12)	103,07	16	14,43(0,06)	87,59	16	19,26(0,13)		16	24,16(0,51)		16	18,01(0,21)	112,46
Inicial		**3,68**	**98,45**		**4,16**	**88,66**		**1,55**			**5,99**			**0,13**	**110,68**
Lutjanus analis	0	0,20(0,06)		0	0,30(0,04)		0	0,44(0,07)		0	0,47(0,10)		0	0(0,11)	
	0,4	0,20(0,06)		0,4	0,65(0,04)	88,43	0,4	0,72(0,07)		0,4	0,97(0,10)	125,68	0,4	0,40(0,10)	100,75
	4	1,45(0,06)		4	3,63(0,03)	83,20	4	3,90(0,06)		4	5,73(0,08)	131,66	4	4,83(0,08)	120,66
	8	2,99(0,05)		8	4,88(0,03)	57,27	8	4,99(0,06)		8	7,65(0,09)	89,77	8	6,32(0,08)	78,95
	16	7,97(0,05)		16	12,51(0,05)	76,31	16	12,41(0,08)		16	16,13(0,16)	97,92	16	16,82(0,17)	105,14
Inicial		**1,95**			**3,01**	**76,30**		**5,14**			**4,67**	**111,26**		**0**	**101,37**

[a] Concentração corrigida levando em consideração o procedimento de extração e a massa da amostra em mg kg^{-1}. Recuperação (R) é expressa em porcentagem. A recuperação não foi mostrada no caso onde o efeito de matriz proporcional estava presente. Em vermelho é representado a média das recuperações. [b] Concentração fortificada na amostra. [c] Concentração predita para as amostras de peixe com os modelos MCR-ALS. Desvio padrão é mostrado entre parênteses para cada valor predito. [c] A concentração na amostra sem fortificação é expressa em mg kg^{-1} e destacada em azul. O Método de adição de padrão foi usado para calcular a concentração na amostra quando o efeito de matriz proporcional estava presente.

d) ARTIGOS

Artigo da Tese

Analytica Chimica Acta 902 (2016) 59–69

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Handling time misalignment and rank deficiency in liquid chromatography by multivariate curve resolution: Quantitation of five biogenic amines in fish

Licarion Pinto[a], César Horacio Díaz Nieto[b], María Alicia Zón[b], Héctor Fernández[b], Mario Cesar Ugulino de Araujo[a,*]

[a] Laboratório de Automação e Instrumentação em Química Analítica e Quimiometria (LAQA), Universidade Federal da Paraíba, CCEN, Departamento de Química, Caixa Postal 5093, CEP 58051-970, João Pessoa, PB, Brazil
[b] Departamento de Química, Facultad de Ciencias Exactas, Físico-Químicas y Naturales, Universidad Nacional de Río Cuarto, 5800, Río Cuarto, Argentina

Outros artigos publicados durante o doutorado relacionados a calibração de ordem superior

Calibração de segunda ordem usando HPLC-DAD

http://dx.doi.org/10.5935/0103-5053.20150125

J. Braz. Chem. Soc., Vol. 26, No. 8, 1573-1582, 2015.
Printed in Brazil - ©2015 Sociedade Brasileira de Química
0103 - 5053 $6.00+0.00

Article

A Fast Chromatographic Method for Determination of Daidzein and Genistein in Spiked Water River Samples Using Multivariate Curve Resolution

Edilene D. T. Moreira,[a] Licarion Pinto,[a] Adriano A. Gomes,[a] Héctor C. Goicoechea[,b] and Mário C. U. Araújo*[,a]*

[a]*Laboratório de Automação e Instrumentação em Química Analítica e Quimiometria (LAQA), Departamento de Química, Universidade Federal da Paraíba, CCEN, CP 5093, 58051-970 João Pessoa-PB, Brazil*

[b]*Laboratorio de Desarrollo Analítico y Quimiometría (LADAQ), Cátedra de Química Analítica I, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral-CONICET, Ciudad Universitaria, 3000 Santa Fe, Argentina*

Microchemical Journal 134 (2017) 131–139

Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

A chemometric cleanup using multivariate curve resolution in liquid chromatography: Quantification of pesticide residues in vegetables

Emanuella Santos Sousa, Licarion Pinto, Mario Cesar Ugulino de Araujo *

Laboratório de Automação e Instrumentação em Química Analítica e Quimiometria (LAQA) Universidade Federal da Paraíba, CCEN, Departamento de Química, Caixa Postal 5093, CEP 58051-970 João Pessoa, PB, Brazil

Calibração de terceira ordem usando HPLC-FLU

Talanta 154 (2016) 208–218

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Highly sensitive quantitation of pesticides in fruit juice samples by modeling four-way data gathered with high-performance liquid chromatography with fluorescence excitation-emission detection

Milagros Montemurro [a], Licarion Pinto [b], Germano Véras [c], Adriano de Araújo Gomes [b], María J. Culzoni [a], Mário C. Ugulino de Araújo [b,*], Héctor C. Goicoechea [a,*]

[a] Laboratorio de Desarrollo Analítico y Quimiometría (LADAQ), Cátedra de Química Analítica I, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Ciudad Universitaria, 3000 Santa Fe, Argentina

[b] Laboratório de Automação e Instrumentação em Química Analítica e Quimiometria (LAQA) Universidade Federal da Paraíba, CCEN, Departamento de Química, Caixa Postal 5093, CEP 58051-900 João Pessoa, PB, Brazil

[c] Departamento de Química, Centro de Ciências e Tecnologia, Universidade Estadual da Paraíba, 58.429-500 Campina Grande, PB, Brazil