Weliton Pedro Batiston

Explorando os fundamentos da cromatografia líquida e espectrometria de massas para integração entre proteômica, microfluídica e quimiometria

Tese apresentada ao Instituto de Química de São Carlos da Universidade de São Paulo como parte dos requisitos para a obtenção do título de doutor em ciência

Área de concentração: Química Analítica e Inorgânica

Orientador: Prof. Dr. Emanuel Carrilho

Orientador no Exterior: Prof. Dr. John R. Yates III

Exemplar revisado

O exemplar original encontra-se em
acervo reservado na Biblioteca do IQSC-USP

São Carlos

2020

DEDICATÓRIA

Primeiramente a Deus,

Dedico este trabalho à toda minha família, em especial à minha mãe Cecília, ao meu pai Pedrão e meu irmão Denis.

À tia “Madre” (Irmã Maria do Carmo), à minha madrinha Aparecida, ao meu padrinho Antônio (in memoriam) e aos meus tios Belmiro e Luzia.

AGRADECIMENTOS

Agradeço ao Professor Dr. Emanuel Carrilho pela oportunidade de usufruir de um laboratório com tantos conhecimentos diversificados em química analítica e áreas correlatas, pela troca de experiência, liberdade, amadurecimento, visão, paciência e pensamento científico. O meu muito obrigado!

Ao Professor Dr. John R. Yates III pela disponibilidade, aceitação do projeto BEPE (Bolsa de Estágio de Pesquisa no Exterior) e oportunidade na realização do meu projeto, assim como a receptividade e preocupação por toda a minha estadia nos Estados Unidos. Thank you very much!

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, (processo 140928/2015-4) pelo aporte financeiro no primeiro semestre deste doutorado.

À Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP, (processos 2015/16025-8 e 2017/17368-1), sou imensamente grato por todas as oportunidades concedidas, a qual espero ter retribuído para o crescimento científico e perpétuo do nosso querido país.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes, pelo aporte financeiro na formação docente (código de financiamento 001).

À Camila Fontes Neves da Silva, pela companhia e paciência nesta minha trajetória.

A todos os integrantes de ambos os grupos de pesquisa (Brasil e Estados Unidos).

Ao Ferminio Cesar Polachini da empresa Microtube, Araraquara – SP e Professor Dr. Sydney Jose Lima Ribeiro da Universidade Estadual de São Paulo – Unesp – SP, pela disponibilidade e doação dos capilares de sílica fundida com revestimento de acrilato.

Aos pesquisadores do Laboratório de Microfabricação do LNNano (Laboratório Nacional de Nanotecnologia Brasileiro) do CNPEM (Centro Nacional de Pesquisa em Energias e Materiais), em Campinas – SP, Dr. Renato Sousa Lima, Maria Helena Piazzetta, Angelo Gobbi e Rui Murer. Pela atenção e discussões sobre o desenvolvimento do microchip multidimensional.

Á Dra. Alessandra Maffei Monteiro pela disponibilidade e gentileza de conceder o protocolo de selagem térmica vidro/vidro para a confecção dos microchips.

A todos os Professores que me acompanharam na minha trajetória acadêmica.

A todos os meus amigos e colegas de trabalho.

A todos os funcionários do Instituto de Química de São Carlos por toda disposição e auxílio nos momentos que precisei.

RESUMO

Esta tese preocupa-se em proporcionar aos leitores tendências científicas decorrentes da química analítica aplicada à solução e desenvolvimento de instrumentação analítica para análises proteômicas. Assim, o **Capítulo 1** descreve uma nova estratégia quimiométrica para otimização de técnicas analíticas hifenizadas, a qual permite maior identificação de espécies amostrais (> 80%). Como consequência deste estudo, no **Capítulo 2** foi proposto três modelos com informações tridimensionais validadas: *i*) químicas, *ii*) físicas e *iii*) instrumentais; as quais permitem melhoras nas figuras analíticas de mérito e identificação de isoformas para análises proteômicas *shotgun* e *bottom-up*. Descobriu-se que dentre todas as propriedades físico-químicas dos peptídeos, a hidrofobicidade da biomolécula em conjunto com parâmetros físicos instrumentais nos processos de partição ^{LC} e dessolvatação ^{MS} permitiram a identificação recorde de peptídeos (> 65k). No **Capítulo 3** foi abordado o desenvolvimento de colunas cromatográficas capilares monolíticas híbridas de nanofluxo, para separações de biomoléculas (aminoácidos, peptídeos e proteínas intactas). Demonstrou-se a multifuncionalidade deste produto, tais como: aplicações para coluna de armadilha (*trap*), em elevada percentagem de solvente orgânico (95%) ou utilização em sistemas cromatográficos multidimensionais, evidenciando análises rápidas por meio da separação de 7 peptídeos em menos de 10 minutos. No **Capítulo 4** foi experimentado e discutido estratégias para o desenvolvimento de *μchip* para separações multidimensionais, acoplado ao espectrômetro de massas. Por fim, no **Capítulo 5** elaborou-se uma equação que prevê por meio de princípios termodinâmicos, o alargamento de bandas de biomoléculas em separações provenientes de colunas de troca iônica. Esta abordagem matemática acrescenta conhecimento à ciência de separações e justifica as análises proteômicas eficientes de sistemas cromatográficos multidimensionais, com parada de fluxo na primeira dimensão. Esta tese tem caráter multidisciplinar com agregação de conhecimentos nas áreas de química, matemática, engenharia e biologia. Em especial nos **Capítulos 3 e 4** propôs-se um viés de mercado, por meio do desenvolvimento de produtos de inovação tecnológica.

ABSTRACT

This thesis proposes to the readers scientific trends in analytical chemistry applied to the solution and instrumental development to proteomics analysis. Thus, in **Chapter 1** we described a new chemometrics strategy to the optimization of hyphenated analytical techniques achieving greater identification of sample compounds (> 80%). Consequently, in **Chapter 2**, we proposed three models with validated tridimensional information: *i*) chemistry; *ii*) physics, and *iii*) instrumental; which one of them improved some analytical figures of merit and isoforms identification applied to shotgun and bottom-up proteomics analysis. Our findings indicate that among all physical-chemistry properties from peptides, the hydrophobicity of the biomolecules linked to specific instrumental parameters in the step of partition ^{LC} and desolvation ^{MS} achieved record identification of peptides (> 65k). **Chapter 3** approaches the development of hybrid monolithic capillary columns for nanoflow in liquid chromatography, applied to the separation of biomolecules (amino acids, peptides, and intact proteins). We tested the multifunctionality of this product, such as a trap application using a high percentage of organic solvent (95%) or using it as a multidimensional chromatography system, showing the fast separation of seven peptides in less than 10 minutes. In **Chapter 4** we discussed strategies for the development of μ chip for multidimensional separation coupled to a mass spectrometer. At last, **Chapter 5**, we elaborated on a thermodynamic equation to predict peak broadening of biomolecules during analytical separation by ion exchange mode. This mathematical approach adds knowledge in separation science and justifies the efficient proteomics analysis from multidimensional chromatography systems with stop-flow in the first dimension. The purpose of this thesis is to add a multidisciplinary character in the areas of chemistry, mathematics, engineering, and biology. In special, **Chapters 3** and **4**, we propose a market bias in terms of the development of technology innovation products.

LISTA DE FIGURAS

Figura 1 - Interdisciplinaridade da quimiometria	25
Figura 2 – Uma analogia para um típico sistema de aplicação do planejamento experimental. ...	26
Figura 3 - Alguns exemplos da seleção dos intervalos dos fatores significativos na modelagem da superfície de resposta, no <i>DoE</i> com dois fatores significativos..	28
Figura 4 - Ilustração da estratégia quimiométrica de multi-planejamentos experimentais com multi-respostas (m-DoE) para a otimização do método LC×LC-MS/MS.....	35
Figura 5 - Análise multidimensional do proteoma em células e tecidos.	45
Figura 6 – Diferentes estratégias de cromatografia líquida em relação à complexidade de uma amostra.	49
Figura 7 - Representação de uma separação proveniente de duas colunas cromatográfica (bidimensional ou multidimensional).....	50
Figura 8 - Representação teórica e prática da ortogonalidade em cromatografia líquida multidimensional..	52
Figura 9 - Resumo do funcionamento do analisador de massas tipo quadrupolo de armadilha linear..	57
Figura 10 - Resumo do funcionamento do analisador de massas tipo Orbitrap.....	58
Figura 11 - Métodos de fragmentações com maior aplicação para peptídeos e proteínas, em espectrometria de massas sequencial.....	61
Figura 12 - Configuração e equipamentos do sistema LC-MS/MS para avaliação dos parâmetros de eletronebulização..	70
Figura 13 - Planejamento experimental composto central, para proteínas digeridas	75
Figura 14 - Gráfico 2D com tempo de retenção normalizado do método SCX×RP-Q-ToF-MS/MS.	77
Figura 15 - Cromatograma de íon total da mistura de proteínas digeridas da melhor condição instrumental da Figura 13 e proveniente do resultado da Tabela 5.....	77
Figura 16 - Modelo químico-matemático BSA, aplicado ao método SCX×RP-MS/MS.....	84
Figura 17 - Comandos do MATLAB para a geração do modelo químico-matemático	85
Figura 18 - Câmara da eletronebulização e ionização do Q-ToF.....	86

Figura 19 - Modelo químico-matemático para a mistura proteolítica de sete proteínas para o método SCX×RP-MS/MS..	90
Figura 20 - Comandos do MATLAB para a geração do modelo químico-matemático.	91
Figura 21 - Distribuição gaussiana dos peptídeos identificados na região hidrofílica do método SCX×RP-MS/MS.	92
Figura 22 - Avaliação da contribuição dos efeitos do <i>DoE</i> .	96
Figura 23 - Cromatogramas de íons provenientes do método MudPIT-LTQ-Orbitrap-MS/MS.	98
Figura 24 - Planejamento composto central para o método MudPIT-LTQ-Orbitrap-MS/MS....	100
Figura 25 - Modelo químico-matemático para maximizar a cobertura do proteoma de <i>S. cerevisiae</i> .	101
Figura 26 - Linguagem do MATLAB para plotar o modelo químico-matemático.	102
Figura 27 - Microscopia de varredura eletrônica evidenciando as diferenças da morfologia de monolitos poliméricos tipos orgânico e inorgânico.	113
Figura 28 - Imagens de MEV dos monolitos em capilares e μ chip.	122
Figura 29 - Imagens de microscopia ótica para colunas capilares monolíticas híbridas organo-sílica em diferentes concentrações dos reagentes de partida da reação de polimerização..	125
Figura 30 - Gráfico normal dos contrastes e tabela da percentagem de contribuição dos contrastes na resposta ótima do monolito fotopolimerizado no capilar da <i>Polymicro</i> .	126
Figura 31 - Gráfico normal dos contrastes e tabela da percentagem de contribuição dos contrastes na resposta ótima do monolito fotopolimerizado no capilar da <i>Microtube</i> .	126
Figura 32 - Fotografia do sistema cromatográfico nano-LC com coluna monolítica acoplada ao detector UV-Vis capilar.	127
Figura 33 - Curva analítica para quantificação de histidina utilizando-se detector capilar UV-Vis (Lab Alliance).	128
Figura 34 - Curva de <i>van Deemter</i> , cujo o analito foi o aminoácido triptofano.	130
Figura 35 - Cromatogramas obtidos para eluição de triptofano com vazão de 250 nL min ⁻¹ .	130
Figura 36 - Curva de <i>van Deemter</i> simétricas cujo analito foi a mioglobina.	131
Figura 37 - Cromatograma da mioglobina referente a coluna monolítica DoE 02.	132
Figura 38 - Superfície de resposta para a otimização das colunas monolíticas referentes as medidas para o triptofano.	135

Figura 39 - Superfície de resposta para a otimização das colunas monolíticas referentes as medidas para a mioglobina.	135
Figura 40 - Curva de <i>van Deemter</i> para as melhores coluna monolíticas obtidas.....	136
Figura 41 - Cromatogramas das coluna capilares monolíticas DoE 16.....	137
Figura 42 - Gráficos para a obtenção do tempo morto (t_m) do sistema nano-cromatográfico com vazão de 250 nL min^{-1} ou velocidade linear de $0,24 \text{ mm s}^{-1}$	138
Figura 43 - Cromatogramas indicando o tempo de retenção distinto das moléculas na coluna monolítica híbrida organo-sílica DoE 14..	139
Figura 44 - Cromatogramas de íons para mistura de 14 padrões de peptídeos obtidos na coluna monolítica híbrida organo-sílica utilizando um nLC-LTQ-Orbitrap XL..	140
Figura 45 - Cromatogramas de íons para 14 peptídeos obtidos na coluna C_{18} utilizando nLC-LTQ-Orbitrap XL com intervalo de massas observado de 400 a 1000 com $z = +2$	141
Figura 46 - Cromatograma de íons e espectro de massas para uma mistura de 3 padrões de biomoléculas para duas condições cromatográficas diferentes.	142
Figura 47 - Eluição direta de 14 peptídeos em nLC-MS, para quatro réplicas de colunas/traps cromatográficas monolíticas, denominadas como A, B, C e D.	143
Figura 48 -Esquema do método cromatográfico multidimensional para a separação rápida de sete peptídeos.....	144
Figura 49 - Cromatogramas de íons referente a separação de sete peptídeos com m/z ($z = +2$)..	144
Figura 50 - Ideia do <i>μchip</i> com separação cromatográfica multidimensional.	153
Figura 51 - Processo de selagem térmica vidro/vidro para confecção dos <i>μchips</i> multidimensionais.....	160
Figura 52 - Projeto do software Autocad para a confecção da máscara de fotolito para a fotogração dos canais microfluídicos em plataformas de vidro, referente a microfabricação do <i>μchip</i> -LC×LC.	161
Figura 53 - Imagens dos canais (azul) obtidas do perfilômetro.	164
Figura 54 - Diferentes imagens de microscopia óptica (4× de magnitude) ao longo dos microcanais microfabricados e selados..	164
Figura 55 - Visualização em diferentes ângulos do <i>μchip</i> multidimensional com suporte de aço-inóx.	165

LISTA DE TABELAS

Tabela 1 - Exemplo de uma matriz ortogonal do Planejamento Experimental <i>Plackett-Burman Design</i> para 12-experimentos com 11-fatores (A - L).....	30
Tabela 2 - Fatores codificados para os primeiros experimentos de um planejamento <i>Plackett-Burman Design</i> , segundo a regra $4n-1$ (número de fatores) para $4n$ (número de experimentos). 31	
Tabela 3 - Experimentos para análise das diferentes dimensões do proteoma.....	46
Tabela 4 - Programa para o corte dos capilares de sílica fundida acima de 100 μm diâmetro interno (d.i.) $\times$ 365 μm de diâmetro externo (d.e.).....	66
Tabela 5 - Dados da melhor condição experimental da Figura 12B para uma mistura proteolítica de 13 espécies digeridas (entre proteínas, fosfoproteínas e isoformas) submetidas ao método SCX $\times$ RP-Q-ToF-MS/MS.....	76
Tabela 6 - Dados da segunda melhor condição experimental da Figura 12B para uma mistura proteolítica de 13 espécies digeridas (entre proteínas, fosfoproteínas e isoformas) submetidas ao método SCX $\times$ RP-Q-ToF-MS/MS	76
Tabela 7 - Valores do modelo químico matemático.....	83
Tabela 8 - Fatores e intervalos independentes para otimização do método nanoLC(C_{18} column)MS/MS	93
Tabela 9 - Níveis de fatores codificados e respostas de saída da matriz, para encontrar os fatores significativos na etapa de ionização do espectrômetro de massas	94
Tabela 10 - Fatores avaliados no planejamento composto central e valores de resposta da matriz	97
Tabela 11 - Ensaios adicionais para avaliação da região máxima de detecção dos peptídeos no planejamento composto central	100
Tabela 12 - Fatores e intervalos para a otimização quimiométrica do monolito em capilares de sílica fundida da <i>Polymicro</i> e <i>Microtube</i>	122
Tabela 13 - Dados dos experimentos referentes ao monolito híbrido organo-sílica em capilar de sílica fundida com revestimento de TSU (<i>Polymicro</i>)	123
Tabela 14 - Dados dos experimentos referentes ao monolito híbrido organo-sílica em capilar de sílica fundida com revestimento de acrilato (<i>Microtube</i>).....	124
Tabela 15 - Valores obtidos referente a aplicação do composto central na otimização das colunas monolíticas híbridas organo-sílica para análise cromatográfica e física do triptofano	133

Tabela 16 - Valores obtidos referente a aplicação do composto central na otimização das colunas monolíticas híbridas organo-sílica para análise cromatográfica e física da mioglobina	133
Tabela 17 - Medidas e dimensões do canal microfluídico multidimensional, obtido no perfilômetro	163
Tabela 18 - Estimativas dos possíveis valores da variação da energia livre de Gibbs de retenção (ΔG_{ret})	176

LISTA DE ABREVIATURAS E SIGLAS

γ	Valor questionado da função desejabilidade da análise quimiométrica do planejamento experimental com múltiplas respostas
<i>μchip-CE</i>	Microchip com separação por eletroforese capilar
<i>μchip-LC</i>	Microchip com separação por cromatografia líquida
<i>ID</i>	Primeira dimensão cromatográfica (primeira coluna)
<i>2D</i>	Segunda dimensão cromatográfica (segunda coluna)
<i>A</i>	Valor alvo da função desejabilidade
<i>AC</i>	<i>Alternating current</i> (Corrente alternada)
<i>ACN</i>	Acetonitrila
<i>ANOVA</i>	Análise da variância
<i>autoMS/MS</i>	Espectrometria de massas sequencial automatizada (nomenclatura da <i>Bruker</i>)
<i>A_v</i>	Termo dos múltiplos caminhos (equação de <i>van Deemter</i>)
<i>B</i>	Termo da difusão molecular longitudinal (equação de <i>van Deemter</i>)
<i>Bottom-up</i>	Análise proteômica de peptídeos, provenientes da digestão proteolítica de mistura menos complexa
<i>BSA</i>	<i>Bovine serum albumin</i> (proteína de albumina de soro bovino)
<i>C</i>	Termo da transferência de massa do analito (equação de <i>van Deemter</i>)
<i>C_I</i>	Contrastes de primeira ordem referente ao planejamento experimental
<i>C₁₈</i>	Grupo hidrofóbico octadecil
<i>CCD</i>	<i>Central Composite Design</i> (Planejamento Composto Central)

<i>CE</i>	<i>Capillary electrophoresis</i> (Eletroforese capilar)
<i>C_m</i>	Termo da resistência de massa do analito com a fase estacionária (equação de <i>van Deemter</i>)
<i>C_s</i>	Termo da transferência de massa da fase móvel (equação de <i>van Deemter</i>)
<i>d</i>	Função de desejabilidade descrita por <i>Derringer e Suich</i>
<i>D</i>	Média geométrica das desejabilidades
<i>DC</i>	<i>Direct current</i> (Corrente contínua)
<i>DESI</i>	<i>Desorption electrospray ionization</i> (Ionização por dessorção e eletronebulização)
<i>DoE</i>	<i>Design of Experiments</i> (planejamento experimental)
<i>ESI</i>	<i>Electrospray</i> (eletronebulização)
<i>f</i>	Número de fatores
<i>FA</i>	<i>Formic acid</i> (ácido fórmico)
<i>FDR</i>	<i>False Discovery Rate</i> (taxa de descoberta falsa) métrica para avaliação da confiança global de um conjunto de dados proteômicos
<i>F_m</i>	Vazão/fluxo da fase móvel
<i>FTICR</i>	<i>Fourier-transform ion cyclotron resonance</i> (Ressonância com ciclotrônica de íon com transformada de Fourier)
<i>GC</i>	<i>Gas chromatography</i> (Cromatografia em fase gasosa ou cromatografia gasosa)
<i>Gravy</i>	<i>Grand average of hydropathicity</i> (Medida de hidrofobicidade para peptídeos e proteínas)
<i>H</i>	Altura equivalente de um prato teórico

<i>Ho</i>	Homogeneidade do monolito no capilar
<i>HPLC</i>	<i>High performance liquid chromatography</i> (cromatografia líquida de alta eficiência)
<i>ids</i>	Identificações
<i>IRMPD</i>	<i>Infrared multiphoton dissociation</i> (dissociação de multifóton no infravermelho)
<i>k</i>	Fator de retenção
<i>L</i>	Comprimento da coluna
<i>Label-free</i>	Uma das modalidades de quantificação de proteínas em proteômica <i>shotgun</i>
<i>LC</i>	<i>Liquid Chromatography</i> (cromatografia líquida)
<i>LC×LC</i>	Cromatografia líquida multidimensional abrangente
<i>LI</i>	Limite inferior do cálculo da desejabilidade
<i>LIT</i>	<i>Linear ion trap</i> (Analisador de massas tipo armadilha de ion linear)
<i>LS</i>	Limite superior do cálculo da desejabilidade
<i>LTQ</i>	<i>Linear trap quadrupole</i> (Quadrupolo de armadilha linear)
<i>m</i>	Desejabilidades individuais
<i>MALDI</i>	<i>Matrix assisted laser desorption ionization</i> (Dessorção-ionização a laser assistida por matriz)
<i>MEV</i>	Microscopia eletrônica de varredura
<i>MPTs</i>	Modificações pós-traducionais
<i>MQM</i>	Modelo Químico-Matemático
<i>MS</i>	<i>Mass spectrometry</i> (espectrometria de massas)

<i>MS/MS</i>	<i>Tandem mass spectrometry</i> (espectrometria de massas sequencial)
<i>MudPIT</i>	<i>Multidimensional protein identification technology</i> (tecnologia multidimensional de identificação de proteínas)
<i>N</i>	Número de pratos (eficiência cromatográfica)
<i>n</i>	Capacidade de pico
<i>n₁</i>	Capacidade de pico da primeira coluna cromatográfica
<i>n₂</i>	Capacidade de pico da segunda coluna cromatográfica
<i>Native</i>	Análise de proteínas no estado nativo
<i>n_c</i>	Capacidade de pico de uma modalidade cromatográfica multidimensional
<i>P</i>	Permeabilidade do monolito
<i>PBD</i>	<i>Plackett-Burman Design</i>
<i>PFC</i>	Planejamento Fatorial Completo
<i>PFF</i>	Planejamento Fatorial Fracionário
<i>Q-ToF</i>	<i>Quadrupole time-of-flight mass spectrometry</i> (Espectrômetro de massas tipo híbrido com analisadores quadrupolo acoplado ao analisador por tempo de voo)
<i>r</i>	Raio
<i>RF</i>	<i>Radiofrequency</i> (Ondas eletromagnéticas de rádio frequência)
<i>RP</i>	<i>Reversed-phase</i> (fase reversa)
<i>R_s</i>	Resolução
<i>RT_(máx)</i>	Tempo de retenção normalizado do último pico
<i>RT_i</i>	Tempo de retenção de peptídeo (analito) questionado

$RT_{i(norm)}$	Tempo de retenção normalizado
RT_{min}	Tempo de retenção em minutos
s	Magnitude para uma resposta específica na equação de desejabilidade
SCX	<i>Strong cation exchange</i> (trocadora catiônica forte)
$SCX \times RP$	Cromatografia líquida multidimensional abrangente cuja 1D é SCX e 2D trata-se de RP
SEC	<i>Size-exclusion chromatography</i> (cromatografia de exclusão por tamanho)
<i>shotgun</i>	Análise de proteínas digeridas em peptídeos provenientes de mistura complexa (proteomas)
T	Temperatura
t	Magnitude para uma resposta específica na equação de desejabilidade
t_1	Tempo de retenção do primeiro pico
TFA	Trifluoroacético
t_g	Tempo do gradiente
ToF	<i>Time-of-flight</i> (tempo de voo)
<i>Top-Down</i>	Análise de proteínas intactas provenientes de proteomas
t_R	Tempo de retenção
t_{R1}	Tempo de retenção do primeiro pico do cromatograma
<i>Trap</i>	Coluna de aprisionamento (armadilha) sem separação cromatográfica
t_{Rn}	Tempo de retenção do último pico do cromatograma
TSU	Revestimento polimérico à base de fluor
$UHPLC$	<i>Ultra High-pressure liquid chromatography</i> (cromatografia líquida de ultra

pressão)

<i>UVPD</i>	<i>Ultraviolet photodissociation</i> (fotodissociação por ultra violeta)
<i>W_h</i>	Largura da meia altura do pico cromatográfico
<i>ΔG_{ret}</i>	Variação da energia livre de Gibbs de retenção
<i>Δp</i>	Variação de pressão
<i>η</i>	Viscosidade
<i>σ²</i>	Variância de banda em unidades de tempo ou desvio padrão
<i>σ²_{ifc}</i>	Total de variância do desvio padrão (pico) do método com fluxo contínuo
<i>σ²_{tpf}</i>	Total de variância do desvio padrão (pico) do método com parada de fluxo
<i>φ</i>	Relação de fases da coluna (móvel e estacionária)

SUMÁRIO DOS CAPÍTULOS

Capítulo 1: Quimiometria

Múltiplos planejamentos experimentais com múltiplas respostas para a otimização de 2D LC-MS/MS em análises proteômicas

Capítulo 2: Proteômica

Modelos químico-matemáticos para aumentar a cobertura da sequência peptídica em proteômica *shotgun*

Capítulo 3: Microfluídica

Desenvolvimento de colunas monolíticas cromatográficas de nanofluxo para separação de biomoléculas

Capítulo 4: Lab-on-a-chip

Desenvolvimento de μ chip para separações cromatográficas multidimensionais

Capítulo 5: Equação para Cromatografia Líquida

Elaboração de uma equação termodinâmica para previsão do alargamento de banda

SUMÁRIO

CAPÍTULO 1	23
1.1. INTRODUÇÃO	24
1.1.1. As contribuições científicas da quimiometria para a química analítica moderna	24
1.1.2. Resumo da fundamentação teórica do planejamento experimental (<i>DoE</i>)	25
1.1.3. Possibilidades de planejamentos experimentais para otimização de processos químicos	28
1.1.3.1. Planejamento fatorial	28
1.1.3.2. Planejamento fatorial fracionário saturado	29
1.2. OBJETIVO	33
1.2.1. Objetivos específicos	33
1.3. PARTE EXPERIMENTAL	34
1.4. RESULTADOS	35
1.4.1. <i>m-DoE</i> - Multi-planejamentos experimentais com multi-respostas	35
1.5. CONCLUSÕES	38
1.6. REFERÊNCIAS	39
CAPÍTULO 2	41
2.1. INTRODUÇÃO	42
2.1.1. A importância dos estudos sobre proteínas: implicações na medicina à economia	42
2.1.2. O estado da arte: proteômica multidimensional para biologia celular	43
2.1.3. Técnicas aplicadas à proteômica	44
2.1.4. Os desafios analíticos da proteômica	47
2.1.5. O estado da arte da instrumentação analítica para proteômica	48
2.1.6. Cromatografia líquida multidimensional ou bidimensional (LC-LC e LC×LC)	48
2.1.7. Modos de configuração em cromatografia líquida multidimensional	53
2.1.8. Espectrometria de massas sequencial	54

2.2.	OBJETIVO.....	62
2.2.1.	Objetivos específicos	62
2.3.	PARTE EXPERIMENTAL.....	63
2.3.1.	Padrões e reagentes.....	63
2.3.2.	Equipamentos e acessórios	63
2.3.3.	Procedimento experimental	65
2.4.	RESULTADOS.....	74
2.4.1.	Superfície de resposta para uma mistura proteolítica de sete biomoléculas.....	74
2.4.2.	Avaliação da otimização quimiométrica do método LC×LC-MS/MS	75
2.4.3.	Estudo do sequenciamento dos peptídeos em proteínas	79
2.4.4.	Desenvolvimento do modelo-químico matemático	81
2.4.5.	Interpretação do modelo químico-matemático para o aumento da cobertura de proteínas em proteômica <i>shotgun</i> e <i>bottom-up</i>	84
2.4.6.	Comprovação da hidrofobicidade dos peptídeos nos modelos químico-matemáticos	88
2.4.7.	Robustez sobre a hidrofobicidade dos peptídeos nos modelos químico-matemáticos	91
2.4.8.	Superfície de resposta para um proteoma de levedura (<i>Saccharomyces cerevisiae</i>)..	92
2.5.	CONCLUSÕES.....	102
2.6.	REFERÊNCIAS	104
CAPÍTULO 3		110
3.1.	INTRODUÇÃO	111
3.1.1.	Uma breve história da evolução da cromatografia líquida	111
3.1.2.	Entendendo a química dos materiais porosos de fases monolíticas	112
3.1.3.	Monolitos poliméricos de sílica ou inorgânico.....	113
3.1.4.	Monolitos poliméricos orgânicos	114
3.1.5.	Monolitos poliméricos híbridos organo-sílica	116
3.1.6.	Entendendo os padrões morfológicos de monolitos híbridos para as separações cromatográficas.....	117
3.2.	OBJETIVOS.....	120

3.2.1.	Objetivos específicos	120
3.3.	RESULTADOS E DISCUSSÃO	121
3.3.1.	Testes iniciais das fases monolíticas em capilares e $\mu chip$	121
3.3.2.	Otimização da síntese dos monolitos	121
3.3.3.	Validação da instrumentação para avaliação da eficiência das colunas capilares monolíticas	126
3.3.4.	Obtenção da curva de <i>van Deemter</i>	128
3.3.5.	Avaliação da polimerização do material monolítico nas colunas otimizadas por meio da superfície de resposta	131
3.3.6.	Refinamento da otimização das colunas capilares monolíticas para nanofluxo	134
3.4.	Avaliação das colunas capilares monolíticas para proteômica <i>bottom-up</i> ou <i>shotgun</i>	139
3.5.	Avaliação de colunas capilares monolíticas híbridas organo-sílica para proteômica <i>top-down</i>	140
3.6.	Alternativas para as colunas monolíticas híbridas organo-sílica desenvolvidas	142
3.6.1.	CONCLUSÕES	145
3.4.	REFERÊNCIAS	147
CAPÍTULO 4		150
4.1.	INTRODUÇÃO	151
4.2.	OBJETIVOS	156
4.2.1.	Objetivos específicos	156
4.3.	PARTE EXPERIMENTAL	157
4.3.1.	Substratos e reagentes	157
4.3.2.	Equipamentos	157
4.3.3.	Fabricação dos microcranis em vidro	157
4.3.4.	Selagem térmica vidro/vidro do $\mu chip$	159
4.3.4.3.	Programação de temperatura	160
4.4.	Resultados e Discussões	161
4.4.1.	Microfabricação do $\mu chip$ multidimensional	161

4.5.	CONCLUSÕES.....	165
4.6.	REFERÊNCIAS	167
CAPÍTULO 5		169
5.1.	INTRODUÇÃO	170
5.1.1.	A cromatografia líquida multidimensional com parada de fluxo na 1D ocasiona alargamentos de bandas?	170
5.2.	OBJETIVO.....	173
5.3.	RESULTADOS	174
5.3.1.	Elaboração de uma equação termodinâmica para a previsão do alargamento de banda em separações cromatográficas provenientes de colunas de troca catiônica.....	174
5.4.	CONCLUSÕES.....	178
5.5.	REFERÊNCIAS	179
ANEXOS		180
m-DoE – Multi-Design of Experiments with Multi-Responses for Optimization of on-line 2D LC-MS/MS in Proteomic Analysis		192

Capítulo 1

Quimiometria

Ciência a qual utiliza-se de ferramentas matemáticas e estatísticas para a solução de dados complexos, provenientes de experimentos laboratoriais. Este capítulo é dedicado ao físico e matemático que fundamentou grande parte da física clássica:

“Science progresses through observation.”

Isaac Newtown

1.1. INTRODUÇÃO

1.1.1. As contribuições científicas da quimiometria para a química analítica moderna

Se algum dia me efetuarem uma proposta para resumir a QUÍMICA ANALÍTICA, eu direi que se trata de “*Um Sinal Interpretável*” pois, em suma é isso que ela representa. Ao colocar uma amostra em um equipamento, e considerando que um dos analitos emita um sinal reprodutível, estável e linear, proporcional ao aumento da concentração, provavelmente teremos a presença essencial da química analítica. Portanto, por trás dos valores qualitativos e quantitativos presente ao nosso redor existem os fundamentos desta ciência.

A química analítica emergiu para um novo patamar de processos modernos de análise instrumental, com o advento dos avanços tecnológicos da informática. Como resultado, as análises de rotina nos laboratórios tornaram-se mais simples, automatizadas, rápidas, eficientes e, principalmente informativas. A química instrumental moderna possibilitou a aquisição de dados analíticos com nível extremo de aprofundamento e riqueza de informação química. Um exemplo é a possibilidade de analisar amostras altamente complexas, como é o caso das ciências ômicas. Em proteômica surge a possibilidade de identificação de 20k de proteínas (analitos) os quais são as evidências para a compreensão de questões biológicas relevantes, tais como o tratamento de inúmeras doenças. [1,2]

Se por um lado, entender todas as espécies químicas da matéria torna-se essencial, por outro lado, a capacidade humana de interpretação abrangente do sistema de estudo é sem dúvida limitada. É neste cenário que surge a quimiometria, esta utiliza-se de ferramentas matemáticas e estatísticas capazes de converter os conjuntos de dados abrangentes na análise dos experimentos químicos, para o máximo de informação útil interpretável. Em suma, o consenso científico apresenta a quimiometria como a intersecção de três grandes áreas, conforme representado pela **Figura 1**.

O ramo da quimiometria com maior aplicação é denominado planejamento experimental, cuja nomenclatura em inglês é *Design of Experiments*, (DoE). O DoE faz uso de matrizes matemáticas cuja as colunas são os fatores químicos avaliados, como por exemplo: pH, pressão, vazão, temperatura, *etc*, e as linhas da matriz são os diferentes experimentos supervisionados. A

última coluna da matriz possui caráter especial, pois trata-se da resposta do sistema avaliado. O tratamento estatístico dos dados permite a visualização da melhor resposta e do(s) fator(es) responsáveis para tal desempenho. Diferente aos métodos de otimização univariada, o *DoE* permite economia de custos e tempo, pois necessita de menos experimentos. [3-6]

Figura 1 - Interdisciplinaridade da quimiometria

Fonte: Adaptação de Batiston (2019) a partir de FERREIRA, M.M.C., Quimiometria: Conceitos, métodos e aplicações. Unicamp, Brasil, 2015, p. 25.

O meu interesse a respeito da quimiometria se iniciou na minha graduação. Assim, no mestrado tive a oportunidade de aplicar o *DoE* para o desenvolvimento do método analítico; cromatografia líquida multidimensional capilar acoplada à espectrometria de massas sequencial para proteômica *shotgun*. Desta forma, no primeiro ano do doutoramento foi desenvolvido uma nova estratégia quimiométrica, a qual é baseada nas limitações dos planejamentos saturados, tais como *Plackett-Burman Design*. [7] Para esta nova estratégia, nós denominamos como *m-DoE* – *Multi Design of Experiments with Multi Responses* (**Anexo**). Para entender a proposta *m-DoE* é apresentado abaixo os fundamentos necessários do *DoE*.

1.1.2. Resumo da fundamentação teórica do planejamento experimental (*DoE*)

A primeira etapa para aplicação do *DoE* é a escolha dos fatores, *i.e.*, os parâmetros instrumentais que podem interferir significativamente nos efeitos da resposta do sistema analítico. Neste momento, um bom conhecimento dos fatores sobre o comportamento químico do

sistema é essencial, pois pretende-se selecionar somente os fatores mais significativos para a otimização da resposta do sistema investigado. Geralmente, esta tarefa é realizada por meio da literatura em conjunto com a experiência do analista, cujo o objetivo é conhecer o sistema avaliado. A **Figura 2** ilustra um típico *DoE*, onde poucos fatores são responsáveis por grande parte da otimização, conforme o *Princípio de Pareto*. [3,4]

Figura 2 - Uma analogia para um típico sistema aplicado ao *DoE*. A ilustração menciona a seleção dos fatores significativos e seu peso. Seguindo o *Princípio de Pareto*, poucos fatores são significativos na resposta e, portanto, possuem maior peso na resposta. Por outro lado, a maioria dos fatores possuem menor peso na resposta, os quais não apresentam significado na otimização do sistema.

Fonte: Autoria própria

O número de experimentos aumenta de forma exponencial de acordo com o número de fatores (f), devido a relação 2^f . Assim, a melhor estratégia seria selecionar poucos f e os mais significativos na otimização. No entanto, às vezes isso não é possível. Por exemplo, no campo da indústria e engenharia é comum haver muitas variáveis para otimização de condições totalmente desconhecidas, como no desenvolvimento de um produto inovador. Além disso, o analista pode não ter experiência em um dado assunto, de modo que a situação necessita de uma solução rápida. Para tal, a leitura exhaustiva sobre o tema específico será limitada. Em ambas as situações, encontrar os fatores significativos pode ser um verdadeiro desafio. [3,4,8]

A segunda etapa para otimização eficiente do *DoE*, é a escolha do intervalo adequado de cada fator, o qual é o ponto crítico do *DoE* e o bom discernimento das condições químicas é essencial para o seu sucesso na otimização. [8,9] A **Figura 3** ilustra alguns desafios para o

analista. Com o intuito de simplificar a lógica da referida ilustração, esta mostra superfícies de respostas com dois fatores plotados. No entanto, nesta etapa, o *DoE* geralmente possui mais fatores (3-5), aspecto que aumenta o desafio na seleção adequada dos intervalos dos fatores. Assim, a **Figura 3A** mostra um cientista que tem conhecimento dos intervalos dos fatores (pontos α e β), uma vez que ele selecionou tais valores para a matriz. No entanto, o cientista não possui conhecimento da superfície de resposta e, provavelmente, analisará os dados em outro momento. As tabelas sobre as figuras mostram as coordenadas x,y de um modelo que está vinculado aos fatores 1 e 2, respectivamente. O eixo z significa a resposta em função do fator 1 (eixo x) e fator 2 (eixo y). Portanto, os valores destas tabelas são os experimentos selecionados (x,y) e as respostas (z) obtidas pelo cientista, ou seja, as tabelas significam a conclusão do cientista. Desta forma, a resposta para o ponto α é z_α e a resposta para o ponto β é z_β . Em **A**, o pesquisador selecionou dois pontos da superfície com intervalos inadequados. Devido a esta decisão, ele não conseguirá identificar o aumento da resposta na otimização, pois os experimentos selecionados não indicam a direção máxima da resposta do modelo, porque $z_\alpha = z_\beta$. Em **B**, o pesquisador selecionou um intervalo adequado, para obter as informações do aumento de resposta na otimização, pois um dos experimentos terá uma melhor resposta. Portanto, é possível prever a direção de uma melhor resposta na superfície porque $z_\alpha > z_\beta$. Em **C**, o pesquisador selecionou níveis de intervalo inadequados, pois a melhor resposta não será identificada, uma vez que os experimentos selecionados indicam uma resposta máxima e existem duas respostas máximas. Em **D**, o pesquisador obteve uma boa escolha dos valores dos intervalos, pois pode-se obter uma conclusão melhor para a seleção das duas respostas máximas na otimização.

A terceira etapa é a escolha do melhor planejamento de modo a modelar os experimentos para obtenção da(s) resposta(s), onde é possível evidenciar os fatores significativos e não significativos, assim como as interações. Aspecto que atribui simplificação na otimização permitindo a modelagem matemática com poucos fatores, o qual geralmente é explicado numa superfície de resposta. Existem dois conceitos estatísticos sobre o *DoE*, os quais fortalecem tais pressupostos: *i)* geralmente, mas nem sempre, os principais fatores são mais significativos do que as interações de ordem superior, esta é uma tendência que ocorre naturalmente na matemática, denominada de esparsidade; [10,11] *ii)* a escolha de um planejamento adequado depende do número de fatores e o tempo necessário para realizar os diferentes experimentos [3,4].

Figura 3 - Alguns exemplos da seleção dos intervalos dos fatores significativos na modelagem da superfície de resposta, no *DoE* com dois fatores significativos. Em A e B, a superfície de resposta possui somente um máximo para otimização. C e D mostram outro exemplo de superfície de resposta, em que existem duas regiões de otimização máxima. As tabelas sobre as figuras indicam a conclusão do cientista, uma vez que o mesmo não possui conhecimento da superfície de resposta. Uma explicação detalhada do racional está descrita no texto.

Fonte: Autoria própria

1.1.3. Possibilidades de planejamentos experimentais para otimização de processos químicos

1.1.3.1. Planejamento fatorial

Existem duas possibilidades nos planejamentos fatoriais: *i)* quando o *DoE* possui de 2 a 4 fatores, o planejamento fatorial completo (PFC) torna-se o mais adequado, pois é possível obter todas as combinações possíveis dos fatores nos níveis designados. Portanto, um *DoE* com 2 níveis e 4 fatores terá 16 experimentos ($2^4=16$); *ii)* quando o *DoE* possui de 5 a 7 fatores, o planejamento fatorial fracionário (PFF) é o mais apropriado, pois o número de experimentos aumenta exponencialmente conforme o número de fatores selecionados. Desta forma, uma consideração matemática é feita, onde somente uma fração de todas as combinações dos níveis dos fatores é avaliada, e parte da matriz de um projeto completo são selecionados para reduzir o número de experimentos. [3,4]

Por exemplo, para um planejamento fatorial completo com 5 fatores e 2 níveis $2^5 = 32$ experimentos serão necessários. Por outro lado, considerando o mesmo número de fatores e níveis em um planejamento fatorial fracionário, com uma resolução de 5, $2^{5-1} = 16$ experimentos serão necessários. Neste último caso, os contrastes estão confundidos com as interações de 4 fatores, de modo que o quinto fator é calculado pela relação múltipla de 5 = 1234 ou geratriz (função geradora) $I = 12345$. Isso significa que os efeitos da matriz (resposta de saída) de mesmo valor foram negligenciados, exceto pelo sinal. [3] A redução do número de experimentos no PFF implica diminuir o conteúdo informativo para resposta mais precisa da otimização do sistema investigado. No entanto, para uma triagem inicial do sistema, o PFF pode ser uma boa estratégia para o sucesso da otimização.

A explicação da questão de confusão entre os fatores de um PFF merece destaque. O PFF com resolução de 5 (2^{5-1}) significa que os efeitos principais são confundidos com as interações de quarta ordem e as interações de segunda ordem são confundidas com as interações de terceira ordem (o termo confundido, significa uma relação matemática de multiplicidade matricial dentre os fatores, para a sua diminuição, ou seja, no caso mencionado, os 5 fatores se transformam em 4 fatores). No PFF com resolução de 4 (2^{4-1}), significa que os principais efeitos dos contrastes são confundidos com as interações de terceira ordem, tendendo assim ser menos significativo para o planejamento. Além disso, os principais efeitos dos contrastes não são confundidos com os contrastes das interações de segunda ordem. [3]

1.1.3.2.Planejamento fatorial fracionário saturado

Fragmentar ou confundir as interações do PFC em PFF, para obter menos experimentos com resultados consistentes têm um limite em função do aumento do número de fatores. Assim, o planejamento experimental atinge a condição saturada. [3,10,11] Para planejamentos experimentais ≥ 7 fatores, a estratégia mais utilizada para planejamentos que visam menos experimentos é o planejamento saturado *Plackett-Burman Design* (PBD). O PBD utiliza um padrão matemático de uma matriz ortogonal (**Tabela 1**), onde somente os fatores principais são avaliados, em outras palavras, não é possível avaliar as interações dos fatores. [7] Esta é uma consideração válida devido ao efeito de esparsidade. [10,11] Desta forma, as considerações matemáticas da matriz ortogonal do PBD exige um número de experimentos e fatores pré-estabelecidos conforme a **Equação 1**, onde n é um número inteiro qualquer.

$$4n - 1 = \text{número de fatores} \quad \text{Equação 1}$$

Assim, podemos supor que quando a matriz ortogonal do PBD exigir um número de $4n$ experimentos, ela poderá investigar no máximo $4n-1$ fatores em dois níveis. Portanto, para um número de 11 fatores (**Tabela 1**), o número de experimentos será 12, pois $n = 3$. [7]

Tabela 1 - Exemplo de uma matriz ortogonal do planejamento experimental *Plackett-Burman Design* para 12-experimentos com 11-fatores (A - L).

Fatores	A	B	C	D	E	F	G	H	I	J	L
Experimentos											
1	+1	+1	-1	+1	+1	+1	-1	-1	-1	+1	-1
2	+1	-1	+1	+1	+1	-1	-1	-1	+1	-1	+1
3	-1	+1	+1	+1	-1	-1	-1	+1	-1	+1	+1
4	+1	+1	+1	-1	-1	-1	+1	-1	+1	+1	-1
5	+1	+1	-1	-1	-1	+1	-1	+1	+1	-1	+1
6	+1	-1	-1	-1	+1	-1	+1	+1	-1	+1	+1
7	-1	-1	-1	+1	-1	+1	+1	-1	+1	+1	+1
8	-1	-1	+1	-1	+1	+1	-1	+1	+1	+1	-1
9	-1	+1	-1	+1	+1	-1	+1	+1	+1	-1	-1
10	+1	-1	+1	+1	-1	+1	+1	+1	-1	-1	-1
11	-1	+1	+1	-1	+1	+1	+1	-1	-1	-1	+1
12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

Se a primeira linha dos níveis da matriz de PBD é conhecida, as linhas restantes podem ser geradas, deslocando-se ciclicamente de forma sequencial a partir da primeira linha. [7] Portanto, o primeiro fator codificado da linha, será o último fator codificado para a linha subsequente; este arranjo segue até a penúltima linha da matriz, de forma que na última linha é adicionado somente fatores codificados de sinal negativo. Este arranjo confere uma matriz simétrica para o PBD, onde o número de sinais em cada coluna é o mesmo. Por exemplo, na **Tabela 1**, em todas as colunas há 6 sinais +1 e 6 sinais -1. Em relação as linhas da matrix todas exceto o experimento/linha 12 possuem 6 sinais +1 e 5 sinais -1, logo os experimentos/linhas de 1 até 11 possuem $+1 \times 11 = +11$ sinais positivos a mais que os sinais negativos. Por esta razão adiciona-se na última linha da matriz da **Tabela 1** somente sinais negativos, ou seja -11 sinais negativos, de modo a balancear os +11 sinais positivos a mais das linhas superiores à última linha. Em outras palavras, a matrix de PBD possui a mesma quantidade de cada código (+1 e -1). Portanto o arranjo é perfeitamente simétrico. A **Tabela 2**, indica as possibilidades de primeira linha para matrizes ortogonais, nos diferentes valores de n . Portanto, para tais possibilidades, se a linha mudar ciclicamente na matriz de PBD, esta possuirá a propriedade de uma matriz ortogonal.

Tabela 2 - Fatores codificados para os primeiros experimentos de um planejamento *Plackett–Burman Design*, segundo a regra de $4n-1$ (número de fatores) para $4n$ (número de experimentos).

nº Inteiro e positivo	Número de experimentos	Número de fatores	Primeira linha da matriz PBD
n	$4n$	$4n-1$	Fatores codificados
2	8	7	+ + + - - + -
3	12	11	+ + - + + + - - - + -
4	16	15	+ + + + - - - - - + + + - +
5	20	19	+ - + + - - - - + - + - + + + - - +
6	24	23	+ + + + + - + - + + - - + + - - + - - - -

Podemos mencionar que o PBD segue um padrão de condições matemáticas bem definidas. [7] No entanto, o discernimento químico deve ser adequado para o sucesso desta abordagem. Pois o PBD possui duas limitações: *i*) preocupações com os padrões de confusão ou a resposta das interações dos fatores, uma vez que não são investigadas e *ii*) as suposições necessárias para evitar a escolha de um intervalo inadequado entre os níveis dos fatores

selecionados (raciocínio da **Figura 3**). [12,13,14] Portanto, se as interações dos fatores possuírem efeitos significativos na otimização ou a seleção dos intervalos dos níveis dos fatores forem inadequados, o PBD terá resultados seriamente tendenciosos. Além disso, o PBD é limitado a um número específico de fatores, conforme indicado nas **Tabela 1** e **Tabela 2**, onde o valor de n controla o número de fatores devido a regra $4n-1$. Esse comportamento intrínseco é muito complexo; pois a matriz de PBD se torna dependente de um número exato de parâmetros (fatores). Certamente, em alguns casos será necessário adicionar ou remover alguns dos fatores da matriz de PBD, de modo a satisfazer a condição de $4n-1$.

As limitações de PBD tem recebido destaque nas áreas afins do conhecimento. Em revistas especializadas em estatística, a atenção tem sido relevante para os esforços de novas estratégias para melhorar a abordagem matemática [11-15]. No entanto, a comunidade química não tem investigado as limitações da abordagem de PBD. Estratégias, para melhorar o discernimento da informação química relevante do sistema investigado, poderiam ser um caminho. Com essas premissas, propõe-se aqui a alternativa *m-DoE*, a qual permite aumentar as informações do sistema, por meio da escolha de um intervalo adequado de cada fator selecionado, de modo a selecionar os fatores verdadeiramente significativos de um sistema com ≥ 7 fatores.

Nesta tese, discutiu-se o racional dos experimentos, para condução da estratégia *m-DoE*. De modo que mais detalhes analíticos podem ser obtidos no artigo (**Anexo**).

1.2. OBJETIVO

Propor uma alternativa quimiométrica aos planejamentos experimentais saturados.

1.2.1. Objetivos específicos

1. Evidenciar as limitações dos planejamentos experimentais saturados, tais como o *Plackett-Burman Design*.
2. Desenvolver um novo raciocínio de otimização quimiométrica.
3. Propor os conceitos teóricos da alternativa multi-planejamentos experimentais com multi-respostas (*m-DoE*).

1.3. PARTE EXPERIMENTAL

O *m-DoE*, reportado no artigo em questão, consiste em três planejamentos fatoriais fracionários PFF (*DoE* 1, 3 e 5), dois PFC (*DoE* 2 e 4) e a superfície de resposta pelo planejamento composto central. Os experimentos, bem como suas respostas, foram planejados para cada dimensão avaliada com as seguintes premissas: se possível experimentos rápidos e respostas avaliadas em multiplicidade de mesma magnitude. Em todos os experimentos, a seleção dos fatores significativos foram avaliados com 95% do intervalo de confiança e analisados. A seleção ocorreu de maneira randomizada. Os programas Excel, Statistica [16] e MATLAB [17] foram utilizados para a elaboração de tabelas, superfícies de resposta, análise de variância e equações.

1.4. RESULTADOS

1.4.1. *m-DoE* - Multi-planejamentos experimentais com multi-respostas

O *m-DoE* consiste em avaliar o sistema por módulos individuais. Portanto, o *m-DoE* é recomendado para técnicas analíticas hifenizadas, onde o método pode ser compartimentalizado. A otimização por meio de *m-DoE* pode ser obtida pelo PFC ou PFF para cada módulo de estudo. Desta maneira, tanto os efeitos principais quanto as interações cruzadas da matriz são avaliadas em cada módulo. A **Figura 4** ilustra o esquema da divisão para a otimização do método da cromatografia líquida multidimensional acoplado a espectrometria de massas sequencial para proteômica *shotgun*.

Figura 4 - Ilustração da estratégia quimiométrica de multi-planejamentos experimentais com multi-respostas (*m-DoE*) para a otimização do método LC×LC-MS/MS. **A)** Cada cubo representa uma técnica do método analítico. Assim, o cubo azul representa a coluna trocadora catiônica forte, o cubo verde representa a coluna de fase reversa e o cubo roxo representa a plataforma espectrometria de massas sequencial. Verifica-se, nas faces de cada cubo, o número das diferentes respostas avaliadas para cada dimensão. **B)** Estratégia de *m-DoE* para calcular as múltiplas respostas do método analítico. As dimensões significam diferentes técnica utilizada. Assim, a otimização da primeira dimensão, cujo o objetivo é aumentar a cobertura de amostra, significa otimizar indiretamente a segunda dimensão, onde o objetivo é aumentar o número de picos (peptídeos) e, conseqüentemente, otimizar a terceira dimensão, onde o objetivo é aumentar a cobertura das proteínas. A otimização de cada módulo aumenta o conhecimento sobre a informação química do sistema avaliado, portanto a estratégia é mais precisa para a seleção exata dos níveis dos intervalos em relação ao PBD.

Fonte: Autoria própria

Verifica-se que diferentes respostas podem ser utilizadas em cada módulo por meio das funções de desejabilidade de *Derringer* (**Equação 2**). [18] Esta é aplicável para quando se deseja obter valores de máxima resposta. Caso se deseja obter valores de mínima resposta aplica-se a

Equação 3. Em seguida, calcula-se a média geométrica dos valores (**Equação 4**). Onde A é o valor alvo, L é o menor valor experimental aceitável dentre os experimentos realizados, U é o maior valor experimental aceitável dentre os experimentos realizados, y é o valor questionado (obtido na resposta) e t é a magnitude para uma resposta específica, caso se deseja aumentar o peso de uma resposta em relação a outra. [3,4,18]

$$d = \left(\frac{\gamma - L}{A - L} \right)^t \quad \text{Para } L \leq \gamma \leq A; \quad \text{Equação 2}$$

$$d = \left(\frac{U - \gamma}{U - A} \right)^t \quad \text{Para } A \leq \gamma \leq U; \quad \text{Equação 3}$$

$$D = \sqrt[m]{d_1 d_2 \dots d_m} \quad \text{Equação 4}$$

$$d = 0 \quad \text{Para } \gamma \text{ fora do intervalo } L \quad \text{Equação 5}$$

Como o *m-DoE* utiliza várias respostas no planejamento, a resposta de cada módulo pode ser diferente, mas elas devem ter uma correlação analítica para melhorar a resposta dos outros módulos. Por exemplo, a resposta do módulo 1 (caixa azul da **Figura 4A**), possui um sentido analítico para melhorar a resposta no módulo 2 (caixa verde da **Figura 4A**). Este conceito está resumido na **Figura 4B**.

Podemos inferir que a estratégia *m-DoE* possui mais informações sobre o sistema avaliado. Consequentemente, a estratégia é mais precisa na seleção dos fatores e níveis dos intervalos em relação ao PBD. No entanto, no *m-DoE* é necessário aumentar o número de experimentos, uma vez que as interações dos fatores são consideradas. Por outro lado, o *m-DoE* se torna vantajoso em relação à abordagem univariada, em termos da necessidade de menor número de experimentos. Além disso, quando o PBD é inadequado, devido a falta de familiaridade dos fatores e seus intervalos, o *m-DoE* supera o PBD. [9]

Por meio da alternativa *m-DoE*, a otimização do método cromatografia líquida multidimensional acoplada a espectrometria de massas sequencial para proteômica *shotgun*, permitiu a cobertura inédita de BSA e mistura de 7 proteínas, cujo os valores foram: 93% e 83%, respectivamente. Adicionalmente, a estratégia *m-DoE* foi aplicada em instrumentos análogos,

estes métodos multidimensionais possuem facilidade para a perda de amostragem no processo de separação e detecção e *m-DoE* permitiu a identificação recorde de números de peptídeos proveniente de proteoma. Detalhes que serão discutidos no próximo capítulo desta tese.

1.5. CONCLUSÕES

Neste tópico, demonstrou-se o raciocínio para aplicação de uma nova alternativa quimiométrica. Assim, o *m-DoE* proporciona alternativa às limitações dos planejamentos saturados (PBD). Portanto, esta proposta ataca em dois princípios nos quais o PBD é limitado: *i*) detecção eficaz dos efeitos principais, assim como, as interações significativas dos fatores utilizados no modelo e *ii*) melhor seleção dos fatores e seus intervalos. Assim, o *m-DoE* pode conter um número maior de experimentos em relação ao PBD. No entanto, as informações químicas no sistema são aumentadas, permitindo maiores chances de sucesso para a condução de uma otimização verdadeira. Podemos inferir que o PBD é a melhor estratégia quando os intervalos dos fatores são bem conhecidos, ou seja, as condições químicas são claras. Quando este fato é limitado, recomenda-se o uso de *m-DoE*. Em suma, este estudo é útil para otimizações ≥ 7 fatores, provenientes de um sistema com pouca informação química, em relação os fatores e níveis dos intervalos. É importante mencionar que no *m-DoE* torna-se necessário que as respostas dos módulos tenham uma estratégia de conexão para melhorar a resposta dos outros módulos. Assim, quando um módulo está sendo otimizado, este deve ter uma otimização indireta para melhorar os módulos subsequentes. Verificou-se que o *m-DoE* permitiu a identificação recorde de números de peptídeos em métodos cromatográficos multidimensionais acoplados à espectrometria de massas sequencial de alta resolução para a proteômica *shotgun*. Dados que serão fundamentados e demonstrados experimentalmente, com maior abrangência no próximo capítulo desta tese.

1.6. REFERÊNCIAS

- 1 FERREIRA, M. M. C. **Quimiometria: conceitos, métodos e aplicações**. Campinas: Unicamp, 2015. p. 25.
- 2 ZHANG, Y.; FONSLow, B.R.; SHAN, B.; BAEK, M.C.; YATES, J.R. Protein analysis by shotgun/bottom-up proteomics. **Chemical Reviews**, Washington, v. 113, n. 4, p. 2343-2394, 2013.
- 3 BARROS NETO, B; SCARMINIO, I. S.; BRUNS, R. E. **Como fazer experimentos: pesquisa e desenvolvimento na ciência e na indústria**. Porto Alegre: Bookman, 2010. 401 p.
- 4 PEREIRA FILHO, E. R. **Planejamento fatorial em química: maximizando a obtenção de resultados**. São Carlos: EdUFSCar, 2015. 88. p.
- 5 HIBBERT, D. B. Experimental design in chromatography: a tutorial review. **Journal of Chromatography B**, Amsterdam, v. 910, p. 2-13, 2012.
- 6 FERREIRA, S. L. C.; BRUNS, R. E.; FERREIRA, H. S.; MATOS, G. D.; DAVID, J. M.; BRANDÃO, G. C.; SILVA, E. G. P.; PORTUGAL, L. A.; REIS, P. S.; SOUZA, A. S.; SANTOS, W. N. L. Box-Behnken design: an alternative for the optimization of analytical methods. **Analytica Chimica Acta**, Amsterdam, v. 597, p. 179-186, 2007.
- 7 PLACKETT, R. L.; BURMAN, J. P. The design of optimum multifactorial experiments. **Biometrika**, Oxford, v. 33, n. 4, p. 305-325, 1946.
- 8 LEARDI, R. Experimental design in chemistry: a tutorial. **Analytica Chimica Acta**, Amsterdam, v. 652, p. 161-172, 2009.
- 9 RITER, L. S; VITEK, O; GOODING, K. M.; HODGE, B. D.; JULIAN, R. K. Statistical design of experiments as a tool in mass spectrometry. **Journal of Mass Spectrometry**, Oxford, v. 40, p. 565-579, 2005.
- 10 BOX, G. E. P.; MEYER, R. D. Dispersion effects from fractional designs. **Technometrics**, Milwaukee, v. 28, n. 1, p. 19-27, 1986.
- 11 QUINLAN, K. R.; LIN, D. K. J. Run order considerations for Plackett and Burman designs. **Journal of Statistical Planning and Inference**, Amsterdam, v. 165, p. 56-62, 2015.
- 12 AHMAD, T.; AKHTAR, M.; GILMOUR, S. G. Multilevel augmented pairs second-order response surface designs and their robustness to missing data. **Communications in Statistics - Theory and Methods**, New York, v. 41, n. 3, p. 437-452, 2012.
- 13 BARRENTINE, L. B. Illustration of confounding in plackett-burman plackett-burman

designs. **Quality Engineering**, New York, v. 9, n. 1, p. 11-20, 1996.

14 MULHOLLAND, M.; WATERHOUSE, J. Investigation of the limitations of saturated fractional factorial experimental designs, with confounding effects for an HPLC ruggedness test. **Chromatographia**, Heidelberg, v. 25, n. 9, p. 2-7, 1988.

15 CHEN, J.; LIU, M-Q; FANG, K-T; ZHANG, D. A cyclic construction of saturated and supersaturated designs. **Journal of Statistical Planning and Inference**, Amsterdam, v. 143, n. 12, p. 2121-2127, 2013.

16 STATSOFT. **Statistica**: (data analysis software system), Version 13.5. Tulsa: Statsoft Power Solutions, Inc., 2013. 1v.

17 MATHWORKS. **Matlab**, Version R2015a. Natick: MathWorks, 2015. 1v.

18 DERRINGER, G.; SUICH, R. Simultaneous optimization of several response variables. **Journal of Quality Technology**, Milwaukee, v. 12, p. 214-219, 1980.

Capítulo 2

Proteômica

Ramo da biologia que estuda todas as possibilidades de interpretações das proteínas nos sistemas biológicos. Dedico este capítulo ao descobridor dos benefícios da vitamina C, desvendador da anemia falciforme e, principal cientista sobre a teoria das ligações químicas:

*“Facts are the air of the scientists.
Without them you can never fly.”*

Linus Pauling

2.1. INTRODUÇÃO

2.1.1. A importância dos estudos sobre proteínas: implicações na medicina à economia

Se algum dia me efetuem a escolher uma molécula para representar a CIÊNCIA, eu provavelmente escolherei uma proteína. Esta macromolécula tem instigado a curiosidade por décadas em nós cientistas. Primeiramente por sua complexidade estrutural, de centenas a milhares de resíduos de aminoácidos, proporcionando assim longas cadeias poliméricas, aspecto que exige supremo conhecimento químico para efetiva compreensão. Tal estrutura permite flexibilidade molecular fantástica, porém desafiadora. Uma vez que as diversas ligações químicas tipo *sigma* permitem livre rotação. Em princípio, tal característica possibilita numerosas conformações estruturais. Entretanto, em meados de 1920, baseado na estereoquímica permitida e consequente estabilidade da macromolécula, o consenso científico sugeriu que cada proteína tinha certa especificidade química ou função estrutural, sugerindo única estrutura tridimensional. Foi a era das elucidações cristalográficas de macromoléculas, fato que revolucionou os achados sobre as proteínas e suas funções. Porém, com o avanço na área, os estudos mostraram que as informações obtidas eram incompletas para responder o dinamismo das questões biológicas. Portanto, a estrutura das proteínas são sempre maleáveis e na maioria das vezes apresentam questões inesperadas. [1,2]

Um segundo argumento, sobre a importância das proteínas é a sua versatilidade nos sistemas vivos, proporcionando funções essenciais em praticamente todos os processos biológicos. Sendo elas agentes catalisadores de reações bioquímicas, transportadoras e armazenadoras de outras moléculas, transmissoras de sinais, constituintes essenciais na estrutura de vários órgãos biológicos, *etc.* Assim, desvendar de maneira confiável os processos biológicos das proteínas na célula é o caminho para a solução e entendimentos de várias doenças, sugerindo avanços na química, farmácia, biologia, bioquímica e medicina. [1,2]

Na prática, cientistas em medicina molecular comparam amostras sadias com as amostras que contém a doença questionada, por meio de técnicas analíticas e bioinformática, eles monitoram o surgimento de proteínas análogas na célula e/ou a variação de sua quantidade ou ainda modificações biomoleculares tridimensionais, que podem representar o mecanismo da

origem da doença. A relação do surgimento da doença com a variação das proteínas permite a criação de medicamentos que inibem macromoléculas alvos, proporcionando a remediação da doença em questão. [3]

Em contrapartida, algumas proteínas possuem a função de controlar o estado normal das células, assim sua caracterização e o entendimento da quantidade adequada permite a criação de suplementos/medicamentos para as necessidades básicas do indivíduo, permitindo assim a funcionalidade normal ou impedindo a proliferação de doenças. Estas são basicamente o funcionamento dos medicamentos biológicos e biossimilares, os quais em sua maioria são constituídos de anticorpo monoclonais. Tais medicamentos revolucionaram o tratamento de uma série de doenças, previamente intratáveis. Tornando-a um dos principais produtos da indústria farmacêutica moderna.

Na década de 90 a *BASF Pharma* em parceria com a *Cambridge Antibody Technology* desenvolveram o medicamento constituído da proteína Adalimumabe cujo nome comercial é Humira®. Este é destinado para o tratamento anti-inflamatório de diversas doenças crônicas, tais como: *i)* artrite; *ii)* espondilite anquilosante; *iii)* doença de Crohn; *iv)* colite e *v)* psoríase. O Instituto de Pesquisa Econômica Aplicada (Ipea) menciona que este produto gera US\$ 10 bilhões por ano para os cofres da empresa química alemã. [4] Por meio de uma simples pesquisa no google, no Brasil atualmente, uma embalagem do medicamento custa em torno de R\$ 10.000,00 reais. Várias farmacêuticas brasileiras têm iniciado a produção de biossimilares e biológicos na última década. No Brasil, apesar da crise este setor da economia representou crescimento anual de 10%. O recente aumento da faixa etária de idade em países emergentes, necessidades de constante inovação e recente expiração de patentes para medicamento biossimilares, permitindo assim a cópia legal, são as apostas das grandes farmacêuticas para expansão do mercado. [4]

Portanto, verifica-se que os estudos sobre proteínas implicam na solução dos desafios da medicina moderna, desenvolvimento de novos fármacos e, consequente atribuições econômicas para ganhos de capital.

2.1.2. O estado da arte: proteômica multidimensional para biologia celular

O proteoma representa o conjunto de todas as proteínas expressas em uma célula, tecido ou organismo, resultantes da expressão funcional de um determinado genoma. Assim, o termo proteômica engloba as estratégias para identificação e quantificação de todas as proteínas de um dado proteoma. [5]

O proteoma é um sistema dinâmico onde cada proteína possui propriedades interconectadas na célula, ou seja, dimensões, as quais quando analisadas em conjunto descrevem o fenótipo da célula. Apesar da possibilidade de obter informações relevantes, devido a diversidade e natureza dinâmica do proteoma, medir essas propriedades se mostra desafiador. Assim, a bioanalítica e bioinformática buscam estratégias para compreensão sistêmica do proteoma, a qual inclui diferentes experimentos para elucidar as diferentes propriedades das proteínas, tais como: *i)* abundância, *ii)* expressão das isoformas, *iii)* renovação (*turnover*), *iv)* localização subcelular, *v)* modificações pós-traducionais e *vi)* interações moleculares (**Figura 5**). [5]

Desta forma, a conexão dos resultados sobre as diferentes propriedades das proteínas e adequada interpretação bioquímica é atualmente denominada de proteômica multidimensional. Apesar dos desafios inerentes, ela permite uma visão global do proteoma, sugerindo assim diversos avanços científicos, principalmente quando se trata no desenvolvimento de medicamentos ou tratamento/compreensão de doenças atualmente incuráveis. [5]

Na **Tabela 3** os Professores Mark Larance e Angus I. Lamond da Universidade de Dundee, da Grã-Bretanha descrevem os principais experimentos utilizados para as medidas das diferentes propriedades (dimensões) do proteoma. Em destaque, verifica-se que a alta cobertura das proteínas permite maior identificação de isoformas específicas.

2.1.3. Técnicas aplicadas à proteômica

Os experimentos de proteômica são classificados de duas maneiras: *i)* análise de peptídeos, os quais são produtos da digestão proteolítica provenientes de algumas enzimas específicas, classificadas entre *bottom-up*, *shotgun* e *middle-down* ou *ii)* análise de proteínas intactas, classificadas entre *top-down* e *native*. A análise de peptídeos permite maior identificação de espécies, sendo atualmente a técnica mais adequada para caracterização global do proteoma e

comparação entre tratamentos amostrais. As análises de proteínas intactas permitem melhor caracterização estrutural das macromoléculas e, geralmente, provêm de misturas menos complexas. [6-11]

Figura 5 - Análise multidimensional do proteoma em células e tecidos. A) As proteínas podem ter diferentes propriedades (dimensões), as quais podem ser fisicamente (área sombreada em verde), quimicamente (área sombreada em azul) ou biologicamente (área sombreada em laranja) relevantes. Nesta figura é mostrada algumas das propriedades que os pesquisadores do artigo em questão consideram como as mais importantes para a pesquisa em biologia celular e aquelas que precisam ser levadas em consideração no desenvolvimento de novos métodos de separação e caracterização em análise multidimensional. B) É mostrado uma série de cubos empilhados, cada um contendo um conjunto discreto de dados proteômicos, os quais correspondem ao valor de cada dimensão (localização, fase do ciclo celular e taxa de renovação - *turnover*). Para cada cubo (veja cubos expandidos), podem-se analisar outras dimensões, como atividade proteica, abundância total das proteínas e estequiometria da fosforilação. Juntos, elas representam visualmente uma abordagem para a análise multidimensional dos dados de proteínas. As esferas no interior dos cubos expandidos A e B representam uma proteína específica de interesse, a qual na fase G1 do ciclo celular pode existir no citosol ou no núcleo, com diferentes taxas de renovação - *turnover*. Esses diferentes conjuntos (cubos) da mesma proteína possuem propriedades distintas, incluindo aumento da abundância das proteínas, fosforilação e atividade no conjunto nuclear (cubo B) em comparação com o conjunto no citosol (cubo A).

Fonte: LARANCE, M. e LAMOND, A.I., Multidimensional proteomics for cell biology. *Nature Reviews – Molecular Cell Biology*, Reino Unido, v. 16, 2015, p. 270.

A proteômica *bottom-up* trata-se da digestão proteolítica feita sob diferentes frações do proteoma ou proteína(s) de interesse. A proteômica *shotgun* trata-se da digestão proteolítica feita sob todo o proteoma (sem prévio fracionamento da amostra). Tanto a *bottom-up* quanto a *shotgun* buscam completa clivagem das ligações peptídicas, portanto, produtos de baixa massa molecular.

Esta característica facilita o manuseio da amostra e processo para a separação e detecção nos instrumentos analíticos. Contudo, após a análise dos peptídeos torna-se necessário a utilização da bioinformática. Ferramenta esta que fornece o sequenciamento dos peptídeos e propõem a(s) suposta(s) proteína(s). Este método é designado como a cobertura da proteína e a condição ideal seria sequenciar todos os peptídeos das proteínas, ou seja, 100% da cobertura. Contudo, perante as limitações da técnica, este fato tem se mostrado muito desafiador. [12-14]

Tabela 3 - Experimentos para análise das diferentes dimensões do proteoma.

Dimensão	Exemplos de técnicas utilizadas
Abundância (quantificação absoluta e relativa)	Quantificação <i>Label-free</i> <i>SILAC</i> Marcação com ^{15}N <i>NeuCode SILAC</i> Marcação com Dimetil <i>TMT</i> <i>iTRAQ</i>
Regulação do ciclo celular	Elutriação centrífuga Inibidores químicos da regulação do ciclo celular <i>FACS</i> (para conteúdo de DNA ou marcadores de fase-específica)
Distribuição no tecido	Dissecação <i>FACS</i> (para marcadores específicos do tipo celular)
Interações	Enriquecimento por afinidade (ensaio de imunoprecipitação endógenas ou ensaios <i>pull-down</i> marcado com proteína de fusão) Perfil de correlação de proteínas Marcação por proximidade
Modificações pós-traducionais	Enriquecimento por afinidade: TiO_2 Enriquecimento por afinidade: <i>IMAC</i> Antibióticos por modificações específicas Cromatografia: <i>IEX</i> Cromatografia: <i>HILIC</i> Cromatografia: <i>ERLIC</i>
Localização	Centrifugação Perfil de correlação de proteínas Marcação por proximidade Solubilidade em detergente
Renovação (<i>turnover</i>)	Marcação metabólica de pulso Tratamento cicloheximida
Expressão da isoforma	<ul style="list-style-type: none"> <u>Alta sequência da cobertura das proteínas para identificação de isoformas específicas dos peptídeos</u> Análise por espectrometria de massas direcionada, pode ser utilizada para detectar isoformas específicas de peptídeos
Solubilidade	Desnaturação térmica seguida de centrifugação diferencial
Atividade	Sensitividade análoga de quinases Domínios de ligação, dependentes de atividade
Estrutura terciária	Sensitividade de proteases Crosslinking

A proteômica *top-down* é o estudo das proteoformas, portanto análises de proteínas intactas, geralmente com massas moleculares de 30 kDa. Esta tolera mudanças estrutural proveniente das condições do método ou preparo de amostra. Em *native-protein mass spectrometry* as interações não covalentes entre as proteínas são preservadas. Para tal, a técnica utiliza condições do método semelhantes ao meio biológico de origem. Assim, se concentram na análise do dinamismo estrutural e funcional das interações que ocorrem no complexo proteico, ou seja, a estrutura quaternária. Análises de proteínas intactas norteiam maiores desafios nos processos de amostragem, separação e detecção. [9,10]

A proteômica *middle-down* trata-se da análise de peptídeos de massa molecular maior em relação à *shotgun* e *bottom-up* (> 5 kDa) ou clivagem peptídica limitada. Aspecto que permite estudos tanto para identificação de espécies biomoleculares, quanto relativa caracterização de proteoformas. Pode-se exemplificar que esta técnica é o meio termo entre *bottom-up/shotgun* e *top-down/native*. [6]

2.1.4. Os desafios analíticos da proteômica

Apesar dos avanços consideráveis relacionados às técnicas de separação e identificação de biomoléculas, a análise de proteínas intactas provenientes de mistura complexa, perdura como um desafio à química analítica. Isso se deve a extrema complexidade das proteínas, em termos de tamanho, modificações e propriedades físico-químicas. Podemos mencionar que as proteínas possuem todos os tipos de interações químicas possíveis, sejam intramolecular e intermolecular. Aspecto propenso para análises com baixa reprodutibilidade, tanto em sistemas cromatográficos quanto aos espectrométricos de análise. Além disso, as proteínas possuem elevada faixa dinâmica de expressão proteica, acredita-se numa ordem de grandeza superior a 10^{10} . Adicionalmente, as proteínas possuem heterogeneidade conformacional decorrentes das modificações pós-traducionais (MPTs). E, ainda, extrema diversidade de espécies em termos qualitativos e quantitativos. Portanto, a nossa capacidade de entender um proteoma por meio da análise de proteínas intactas, atualmente configura-se como limitada. [15]

É evidente que a química analítica tem avançado significativamente para a análise de moléculas pequenas. Em tese, estas possuem complexidade inferior, assim seu comportamento

químico em sistemas analíticos de separação e detecção são mais conclusivos para o sucesso de análise. Certamente, a proteômica *shotgun* e *bottom-up*, têm-se apoderado do conhecimento das técnicas analíticas para a análise de moléculas pequenas. Consequentemente, a *shotgun* tem permitido maior caracterização do proteoma, com números expressivos na identificação de proteínas, antes jamais reportado. Entretanto, é factível que a análise de um proteoma proveniente de uma mistura complexa de diversas proteínas digeridas em peptídeos faz com que aumente a dependência na exatidão da bioinformática, para que o sequenciamento de peptídeos em proteínas não seja tendencioso. Assim, a proteômica *shotgun* caracteriza-se como uma técnica relativa para a identificação de proteínas, assim como para a quantificação. Além disso, espera-se que a cobertura das proteínas atinja 100%, aspecto desafiador para a cromatografia líquida e espectrometria de massas sequencial de alta resolução devido principalmente a perda de peptídeos no processo de análise. [16,17]

2.1.5. O estado da arte da instrumentação analítica para proteômica

Como verificamos, as análises proteômicas são complexas, demoradas e dispendiosas. De modo a resolver os constituintes biomoleculares do proteoma, antes da detecção, torna-se indispensável o uso de técnicas com elevado poder de separação, preferencialmente automatizada e reprodutível. Atualmente, a cromatografia líquida multidimensional (LC×LC) possui os atributos mais adequados. Em contrapartida, uma técnica de detecção molecular permite caracterização precisa. Para tal, a espectrometria de massas é a ferramenta chave por sua versatilidade. Esta, permite tanto a detecção em MS¹ quanto caracterização após fragmentação no MS² de maneira rápida e, o mais importante, a nível do íon molecular. Além disso, esta técnica proporciona a separação de íons por diferença de m/z . Tal versatilidade da técnica acrescenta o poder de separação da LC×LC. Portanto, o estado da arte em estudos proteômicos é a cromatografia líquida multidimensional acoplada a espectrometria de massas sequencial de alta resolução. [18-22]

2.1.6. Cromatografia líquida multidimensional ou bidimensional (LC-LC e LC×LC)

A cromatografia multidimensional é designada para as técnicas que combinam duas ou mais colunas cromatográficas, assim para cada coluna faz-se a analogia de uma dimensão de separação. Sua aplicabilidade torna-se necessária para amostras com elevada complexidade, geralmente > 400 compostos. A **Figura 6** resume quando a cromatografia bidimensional (2D-LC), assim também denominada, torna-se necessária ao método de separação. Verifica-se historicamente que o modo 1D isocrático permite separação de até 10 compostos, numa média de 10 minutos de análise. Em contrapartida, para amostras com aproximadamente 100 compostos, torna-se imprescindível a utilização de gradiente, conforme indicado pelas setas A e B da **Figura 6**. Para amostras acima de 1000 compostos torna-se necessário a utilização da 2D LC, seta C. A grande vantagem da 2D-LC é a elevada capacidade de pico, conforme representado pela **Figura 7**, onde dois picos na 1D são separados em 4 picos na 2D. Este fenômeno ocorre devido às diferenças de seletividades no processo de partição dos solutos entre as fases estacionárias das dimensões. [22-24]

Figura 6 – Representação da evolução das diferentes estratégias de cromatografia líquida em relação à complexidade de uma amostra. As elipses indicam o domínio histórico dos diferentes modos de separação; as setas indicam mudanças nos limites observados dos domínios como resultado das pesquisas fundamentais da área.

Fonte: Adaptação de Batiston (2019) a partir de STOLL, D.R.; LI, X.; WANG, X.; CARR, P.W.; PORTER, S.E.G.; RUTAN, S.C. Fast, comprehensive two-dimensional liquid chromatography. **Journal of Chromatography A**, Holanda, v. 1168, 2007, p. 8.

Figura 7 - Representação de uma separação proveniente de duas colunas cromatográfica (bidimensional ou multidimensional).

Fonte: Adaptação de Batiston (2019) a partir de STOLL, D.R.; LI, X.; WANG, X.; CARR, P.W.; PORTER, S.E.G.; RUTAN, S.C. Fast, comprehensive two-dimensional liquid chromatography. **Journal of Chromatography A**, Holanda, v. 1168, 2007, p. 5,

J.W. Dolan *et al.* [25] discutem as diferentes maneiras de mensurar a capacidade de pico (n), cuja finalidade é propor de maneira precisa ao cálculo todos os fatores inerentes aos desvios das análises cromatográficas. Todavia, as **Equação 6** e **Equação 7** são as mais utilizadas devido as suas simplicidades, onde, t_R e t_1 são o tempo de retenção do último e primeiro pico cromatográfico, respectivamente, W é a média da largura da base de todos os picos e t_g é o tempo do gradiente. [24,25]

$$n = \frac{(t_R - t_1)}{W_b} \quad \text{Equação 6}$$

$$n = 1 + \frac{t_g}{W_b} \quad \text{Equação 7}$$

Portanto, n representa o número máximo de componentes que pode ser teoricamente separada numa dada coluna cromatográfica em determinado intervalo de tempo do gradiente. Com este pressuposto, o esperado é obter elevados valores de n , assim as estratégias em LC para tal são: *i)* diminuir a taxa do gradiente de fase orgânica; *ii)* utilizar colunas de comprimento maior, com o aumento proporcional no tempo de gradiente e consequente aumento da pressão no sistema; *iii)* empregar colunas empacotadas com partículas de sorvente menores. Para melhor desempenho da cromatografia multidimensional é necessário que as separações nas diferentes colunas apresentem ortogonalidade, ou seja, espera-se que os mecanismos físico-químicos no processo de partição dos analitos de cada dimensão cromatográfica sejam independentes, em outras palavras, forneçam diferentes seletividades. [24-26]

A **Figura 8** define o conceito prático e significado teórico de ortogonalidade. Podemos facilmente notar a origem da palavra ortogonalidade, a qual é proveniente dos conceitos trigonométricos, ou seja, cada cromatograma é projetado em um ângulo de 90°. Assim, quando os analitos estão distribuídos de forma aleatória na área cromatográfica, estes possuem pouca correlação e elevada ortogonalidade, portanto, ótima capacidade de pico. De forma análoga, quando os analitos não possuem aleatoriedade e elevada correlação, há baixa ortogonalidade e, portanto, ineficiência em capacidade de pico. [24]

Em 1995, J.C. Giddings formulou que a capacidade de pico teórica da cromatografia multidimensional é o produto múltiplo das capacidades de pico de cada dimensão de separação, ou seja, ao combinar os modos LC ortogonais da capacidade de pico de $n_1 = 60$ e $n_2 = 100$ o resultado é uma n_c de 6000, conforme a **Equação 8**. [27]

$$n_c = n_1 \times n_2 \times n_3 \times \dots n_n \quad \text{Equação 8}$$

De modo que n_c é a capacidade de pico total e n_1 , n_2 e n_n são as capacidades de pico de cada dimensão cromatográfica. Portanto, a capacidade de pico, depende da ortogonalidade das dimensões de separação. Caso a seletividade dos modos de separação não sejam totalmente ortogonais (diferentes) a capacidade de pico máxima atingível é menor do que o esperado.

Figura 8 - Representação teórica e prática da ortogonalidade em cromatografia líquida multidimensional. A) baixa correlação e alta ortogonalidade. B) alta correlação e baixa ortogonalidade. C) Separação proveniente de óleo de limão. D) Separação de esteroides.

Fonte: Adaptação de Batiston (2019) a partir de FRANÇOIS, I.; SANDRA, K.; SANDRA, P. Comprehensive liquid chromatography: fundamental aspects and practical considerations – A review. *Analytica Chimica Acta*, Holanda, p.17, 2009, v. 641.

2.1.7. Modos de configuração em cromatografia líquida multidimensional

A técnica em questão geralmente permite duas modalidades de configurações: *off-line* ou *on-line*. A configuração *off-line* é baseada na coleta de frações da 1D e injeção das frações na 2D. O modo *on-line* realiza a transferência direta das frações do eluente da primeira dimensão para a segunda dimensão, sem interrupção na vazão da fase móvel, geralmente por meio de válvulas de comutação. [23,28]

A estratégia *off-line* oferece maior simplicidade instrumental e flexibilidade quando as fases móveis entre as dimensões são diferentes, pois as frações podem ser: diluídas, concentradas ou dissolvidas em solventes apropriados. As desvantagens estão relacionadas com maior risco de perda e contaminação da amostra, devido à manipulação e dificuldade de automatização do sistema, o que geralmente resulta em maiores tempos de análise. [19,28]

No modo *on-line*, a manipulação da amostra é reduzida, o que contribui para evitar o volume morto, perda da amostra e redução no tempo de análise. No entanto, essa técnica apresenta requisito rigoroso na escolha dos solventes da fase móvel, uma vez que devem ser compatíveis em ambas as dimensões; caso contrário, surge a necessidade de uma coluna de aprisionamento (*trap*) para limpeza da amostra e troca do eluente, antes da transferência dos analitos para a segunda coluna. [23,28]

A abreviatura LC-LC significa que somente partes relevantes e devidamente escolhidas da análise feita na 1D são injetadas na 2D e o restante do eluato segue para o descarte. Esta fração está relacionada com interferentes ou compostos que contribuem com o efeito de matriz. Deste modo, a LC-LC é frequentemente utilizada quando se necessita de uma maior resolução para examinar um pequeno segmento de um determinado pico, em uma amostra complexa, neste caso a nomenclatura da técnica é *heart-cut*, ou seja, somente parte(s) de interesse(s) da separação na 1D é direcionada para a 2D. [23,28]

Outro segmento da técnica é a cromatografia líquida multidimensional abrangente (do inglês *comprehensive*), na qual todas as frações separadas na 1D são submetidas para a 2D. Desta forma, tem-se a análise de todos os compostos químicos da amostra e conseqüentemente todos são de interesse do analista. Geralmente, esta modalidade é utilizada para misturas com mais de

mil compostos onde o objetivo final é caracterizar a amostra por completo. Normalmente, ela é utilizada em estudos sobre as ciências ômicas, tais como: *i)* proteômica; *ii)* metabolômica; *iii)* lipidômica; *etc.*

As dificuldades no desenvolvimento desta técnica estão relacionadas com *(i)* a compatibilidade das fases móveis da primeira dimensão com a segunda dimensão, *(ii)* a resolução obtida na 1D não deve ser perdida no processo de eluição para a 2D e *(iii)* o tamanho das conexões e colunas devem ser planejados, principalmente na modalidade *on-line*, devido a quantidade de fração eluída da 1D para a 2D. [23,24,26]

Em termos de nomenclatura/abreviação quando se trata da modalidade *heart-cut* a abreviatura é LC-LC, em contrapartida quando se trata da modalidade abrangente a abreviatura é LC×LC. O símbolo de multiplicação significa o produto múltiplo da capacidade de pico entre a 1D e 2D, a qual representa a área de separação. [23,26,28]

2.1.8. Espectrometria de massas sequencial

Conforme as teorias do eletromagnetismo, partículas carregadas (íons) em movimento (aceleradas) geram campo eletromagnético, a ação de um campo externo, uso da radiofrequência (RF) e corrente DC ou AC sob os íons, permite o seu controle. Consequentemente, partículas submetidas a aceleração e deflexão (força) provenientes do campo eletromagnético externo, possibilitam o fenômeno físico para mensurar a m/z dos íons. Isso significa que; o tipo de força aplicada e adequada equação permite obter uma solução em m/z .

A proteômica tem impulsionado o avanço da espectrometria de massas. Embora a precisão de massa aliada a resolução sejam as figuras de mérito mais importantes, devido a discriminação entre massas (m) com valores < 1 ppm ($10^{-6} m/\Delta m$), verifica-se também avanços consideráveis em: sensibilidade, velocidade de aquisição e qualidade da fragmentação das biomoléculas. Para atingir estes atributos, as pesquisas têm focado no desenvolvimento de: *i)* fontes de ionização em micro e nano-fluxo; *ii)* inovações nos formatos dos analisadores; *iii)* subsequentes configurações híbridas; *iv)* dispositivos óticos para controlar o foco ou deflexão dos feixes de íons no vácuo, em função do campo eletromagnético e *v)* novas técnicas de fragmentação de peptídeos e proteínas. [10,19]

É interessante mencionar que as propriedades de separação, cujos íons são governados por um campo eletromagnético e, diferenciado por sua razão massa carga (m/z), torna a espectrometria de massas uma técnica extremamente rápida, quando comparada com a cromatografia, cuja separação é governada por propriedades físico-químicas de separação. Contudo, esta última perdura como a mais eficiente técnica em termos da ortogonalidade em separação para solução de amostras complexas. [18,26]

Sabe-se que proteínas e peptídeos em geral são polares, não voláteis e termicamente instáveis, diante destas propriedades, um dos maiores avanços científicos na área foi à introdução de métodos brandos de ionização, sem a degradação aleatória e abrupta da biomolécula, tais como a ionização por eletronebulização (ESI), [29] a ionização de dessorção a laser assistida por matriz (MALDI) [30] e, recentemente, a ionização por dessorção e eletronebulização (DESI). [31] Ao contrário da MALDI e DESI, a fonte de ionização ESI produz íons em solução, aspecto que torna compatível com a cromatografia líquida. Um ganho substancial, em termos de sensibilidade para a espectrometria de massas, foi a criação da nano eletronebulização (nano-ESI). [32] A nano-ESI não necessita de gases (N_2 ou He) para a dessolvatação da amostra, uma vez que ela usufrui de vazões em torno de 20 nL min^{-1} , o qual permite a formação de microgotículas com área superficial maior e raio da gota menor, o qual contribui com a explosão Coulômbica e consequente ganho de sensibilidade, devido a menor diluição dos analitos na fase móvel.

Contudo, o nanofluxo usual do acoplamento da LC com nano-ESI é de 250 nL min^{-1} . Isto se deve a dois fatores que essencialmente necessitam trabalhar em cooperação: *i*) eficiência cromatográfica conforme os parâmetros da equação de *van Deemter* na coluna cromatográfica e *ii*) eficiente formação do aerossol das biomoléculas no estado ionizado (ionização), para atingirem de maneira satisfatória a câmara de vácuo do espectrômetro de massas.

Em especial, comprovou-se que no processo ESI as ligações não-covalentes ou interações moleculares entre biomoléculas são preservadas. Esta versatilidade adicional da técnica permite a ionização e infusão de complexos proteicos no espectrômetro de massas. Aspecto que tornou possível a análise da estrutura quaternária das proteínas, emergindo as colaborações sobre os estudos das conformações de proteínas em suas formas nativas. [9,10]

É no analisador de massas que o fenômeno físico para medir a m/z dos íons acontece. Atualmente, os analisadores utilizados em proteômica ou outras demandas bioanalíticas complexas, se concentram nos sistemas híbridos entre: *i*) quadrupolo (Q); *ii*) tempo de voo (TOF); *iii*) armadilha de íon linear (LIT); *iv*) Orbitrap e *v*) Ressonância com Ciclotrônica de íon com Transformada de Fourier (FTICR). Sistemas híbridos em espectrometria de massas permitem melhores figuras analíticas de méritos para a detecção de íons, provenientes de mistura complexa.

Entre os analisadores de massas, o FTICR oferece maior resolução. Entretanto, o analisador Orbitrap, além da considerável resolução, oferece melhores resultados em termos de espectros de massas sequencial (MS/MS), com menor custo de manutenção do equipamento. Ambos os instrumentos, Orbitrap e FTICR, fazem uso da transformada de Fourier para converter o sinal de domínio de tempo da frequência da corrente iônica em espectros m/z . [6,15,19]

A configuração híbrida do íon trap com Orbitrap tem sido a ferramenta de escolha para os estudos biomoleculares provenientes de amostras com elevada complexidade, a qual inclui principalmente estudos dirigidos pela proteômica *shotgun*. O íon trap é considerado um analisador de massas de alto rendimento em proteômica. Este possui significativa versatilidade pois apresenta: *i*) alta capacidade de estocar íons; *ii*) aspecto que permite o enriquecimento da sensibilidade; *iii*) ótima faixa de seleção de massas; *iv*) rapidez de escaneamento; *v*) eficiente ciclo de serviço (*high-duty cycle*); *vi*) moderada resolução de massas (largura total à meia altura – FWHH – de 2000); *vii*) precisão de massa em torno de 100 ppm ($1 \times 10^{-4} m/\Delta m$); *viii*) robustez e *ix*) simplicidade na produção e miniaturização. [7,33]

O quadrupolo de armadilha linear (LTQ) é composto de 4 hastes hiperbólicas sobrepostas simetricamente formando um raio de 4 mm entre elas. Cada haste é dividida em três seções axiais de 12, 37 e 12 mm de comprimento (**Figura 9A**). Cada seção possui nível de DC discreta. Esta geometria foi primeiramente simulada pela técnica Simion, predizendo a possibilidade de confinamento dos íons ao longo do eixo da seção central do dispositivo, evitando possíveis distorções do campo eletromagnético. A habilidade desta configuração de evitar distorção do campo é mostrada na simulação da **Figura 9B**. [33]

Figura 9 - Resumo do funcionamento do analisador de massas tipo quadrupolo de armadilha linear. A) Característica básica do analisador; B) Simulação Simion da ressonância de excitação do campo eletromagnético sem distorção; C) A radiofrequência (RF) aplicada varia em função do tempo e potencial. De forma simétrica e sincronizada, o potencial positivo em dado momento é convertido em potencial negativo em um segundo momento. A visualização final assemelha-se a uma rotação circular de sinal/potencial; D) Momento instantâneo do campo gerado pelo quadrupolo ou superfície do potencial, em função de uma RF aplicada. O círculo aberto representa um íon na superfície do potencial.

Fonte: Adaptação de Batiston (2019) a partir de SCHWARTZ, J.C.; SENKO, M.W.; SYKA, J.E.P. A Two-Dimensional Quadrupole Ion Trap Mass Spectrometer. **J Am Soc Mass Spectrom**, Estados Unidos. 2002, v. 13, p. 660 e 661 (A, B e C). MARCH, R.E. Quadrupole Ion Trap Mass Spectrometer. **Encyclopedia of Analytical Chemistry**, Estados Unidos. 2000, p. 4 (D).

Embora a sistemática do funcionamento do LTQ requer maiores detalhes na explicação, basicamente o fenômeno físico de confinamento dos íons acontece pela rotação de potencial aplicado na seção central, em conjunto com aplicação de DC discreta em cada segmento do analisador, conforme descrito pela **Figura 9C**. Assim, quando os íons positivos são atraídos pelos polos negativos, em um curto intervalo de tempo, os polos negativos são convertidos em polos

positivos ocasionando a repulsão dos íons positivos. A mudança temporal e espacial do potencial positivo para negativo gera uma superfície de potencial em forma de sela (**Figura 9D**) oscilante, reprodutível e simétrica, ocasionando a captura ou estocagem de íons no vácuo.

A comprovação de que um eletrodo central em forma de fuso poderia gerar um potencial puramente harmônico na direção do eixo z foi primeiramente idealizada por Kingdon em 1923, comprovada por Knight em 1981 pela simulação Simion, conforme representado **Figura 10A** e refinada por Makarov em 2000, como o *trap* ideal de Kingdon, o que seria posteriormente chamado de Orbitrap. [19,34]

Figura 10 - Resumo do funcionamento do analisador de massas tipo Orbitrap. A) Gráficos Simion das linhas equipotenciais para parâmetros ideais da armadilha de Kingdon. Os números referem-se aos potenciais das referidas curvas de contornos. Verifique que os contornos centrais se assemelham a trajetória dos íons da Figura B. B) Vista em corte do analisador Orbitrap. Os íons são injetados no Orbitrap em um ponto onde $z = 0$ e perpendicular ao eixo z , onde iniciam as oscilações axiais de radiofrequência dos íons, sem a necessidade de excitação adicional. O cilindro em vermelho trata-se da área de trajetória dos íons.

Fonte: PERRY, R. H.; COOKS, R.G.; NOLL, R.J. Orbitrap mass spectrometry: Instrumentation, ion motion and applications. **Mass Spectrometry Reviews**, Estados Unidos, 2008, p. 664. (A). SCIGELOVA, M.; HORNSHAW, M.; GIANNAKOPOULOS, A.; MAKAROV, A. Fourier transform mass spectrometry. **Molecular & Cellular Proteomics**, Estados Unidos, 2011, p. 11.

O analisador Orbitrap, também considerado como armadilha de íons (*trap*), assim como o FTICR e LIT, é composto por eletrodo anódico em forma coaxial (externo) o qual envolve o eletrodo catódico em forma de fuso (central), esta geometria confere um potencial harmônico, estável e puramente governado por forças eletrostáticas (**Figura 10B**). A tensão elétrica é aplicada de forma axial entre o eletrodo central, proporcionando assim aos íons estocados uma

trajetória orbital estável e oscilação na direção axial entre os eletrodos. Assim, o Orbitrap utiliza a frequência de oscilação dos íons confinados ao longo do eletrodo central, para determinar sua razão m/z . Estas propriedades do Orbitrap conferem à elevada resolução do equipamento, em contrapartida a velocidade de escaneamento das m/z é considerada como moderada. Como via de regra, o uso da alta resolução prejudica a velocidade de escaneamento e vice-versa. [19,34]

Em termos de aprimoramento dos analisadores para a proteômica *top-down*, cientistas têm investido no desenvolvimento de novas células de harmonização do FTICR, geometria do Orbitrap para campo alto e avanços no processamento do sinal. Como resultado, a resolução em unidade de base foi recentemente estendida para proteínas com massas de 150 kDa, utilizando os analisadores FTICR ou Orbitrap. [15] Devido a maior resolução do FTICR, este tem sido o analisador de escolha para *top-down*, entretanto, o sistema híbrido LTQ-Orbitrap mostrou-se como uma estratégia de sucesso.

Geralmente a qualidade da fragmentação dos íons determina, em termos gerais, o sucesso da identificação dos peptídeos e proteínas. A fragmentação em espectrometria de massas trata-se de um processo de ativação da molécula, cuja a finalidade é gerar de maneira reprodutível “impressões digitais” biomoleculares de uma matriz diversificada de moléculas. O objetivo de todos os métodos de ativação é essencialmente o mesmo: depositar energia nos íons para causar clivagens nas ligações químicas, de maneira reprodutível e que produzam fragmentos de íons interpretáveis, ou seja, que revelem informações estruturais ou a sequência da molécula de interesse. [35]

As técnicas de fragmentação utilizadas em proteômica baseiam-se em: *i*) transferência de energia por colisão (dissociação induzida por colisão – CID, dissociação colisional de alta energia – HCD, dissociação por captura de elétron – ECD e dissociação por transferência de elétron – ETD; dissociação induzida na superfície – SID; reações íon-íon) ou *ii*) baseada em fótons (dissociação de multifóton no infravermelho – IRMPD; e fotodissociação por Ultravioleta – UVPD). [6,35]

Os métodos de fragmentação por transferência de energia por colisão são mais efetivos para peptídeos com massa molecular menor e menos eficientes na fragmentação para peptídeos com massa molecular elevada, assim como para proteínas intactas ou espécies contendo MPTs lábeis, tais como fosfopeptídeos.

Os métodos de fragmentação baseados em elétrons geram padrões de fragmentação de moléculas maiores, aspecto que favorece a interpretação de uma espécie com MPTs lábeis, como por exemplo, fosforilação e glicosilação. Assim, as fragmentações tipo ECD e ETD possuem resultados eficientes na caracterização de proteínas intactas. [10,15]

SID é a alternativa que provém a mais elevada energia aos métodos de colisão em fase gasosa, nos quais os íons são ativados e fragmentados após a colisão com uma superfície (o qual funciona com um grande alvo). Além de sua capacidade de gerar ricos padrões de fragmentações, para muitas classes de íons, devido a sua maior energia de deposição. A SID também tem sido utilizada na caracterização da estrutura de complexos proteicos. Nas técnicas de fotodissociação, os íons acumulam energia por meio da absorção de um ou mais fótons, consequentemente se obtém a fragmentação da espécie iônica. [34]

Em especial, dois métodos de fragmentação têm contribuído consideravelmente com os avanços da proteômica *top-down*, ETD e ultravioleta de fotodissociação UVPD. Atualmente são os métodos mais eficientes de fragmentações para proteínas intactas, especialmente quando utilizadas em combinação. [10,15]

Os métodos baseados por colisão CID/CAD são as técnicas de fragmentação mais comum em proteômica. Estas geralmente fornecem a fragmentação na ligação amida do peptídeo/proteína. Como resultado íons tipo *b*- e *y*- são formados. O método HCD fornece fragmentações que resultam formações de íons tipo *a*- e *x*-. Os métodos baseados por elétron, tais como ECD e ETD, resultam geralmente na clivagem da ligação N-C_∞ e produzem fragmentos de íons tipo *c*- e *z*-, **Figura 11**. [6,10,15,35]

Figura 11 - Métodos de fragmentações com maior aplicação para peptídeos e proteínas, em espectrometria de massas sequencial.

Deteção por captura de elétrons (ECD)

Dissociação por transferência de elétrons (ETD)

Dissociação induzida por colisão (CID) e dissociação por colisão de alta energia (HCD)

Fonte: Adaptação de Batiston (2019) a partir de LC/GC's CHROMacademy. Disponível em: <https://www.chromacademy.com/lms/sco1264/07-MS-MS-at-the-Protein-Level.html?fChannel=27>. Acesso em: 02 dez. 2019.

2.2. OBJETIVO

Propor modelos com informações tridimensionais (químicas, físicas e instrumentais) para o aumento da cobertura da sequência peptídica em proteômica *shotgun* e *bottom-up*. Consequentemente, proporcionar maior identificação de proteínas, assim como melhor diferenciação das isoformas.

2.2.1. Objetivos específicos

1. Plotar superfícies de respostas provenientes da otimização do método analítico e encontrar significado químico para as regiões de máxima resposta, as quais correspondem com maior número de peptídeos identificados.
2. Aplicar digerido proteico de maior complexidade no método, de modo a validar o modelo.
3. Aplicar um proteoma real no modelo utilizando cromatografia líquida multidimensional de nanofluxo, acoplado a espectrometria de massas sequencial de alta resolução (nLC×nLC-MS/MS).
4. Realizar interpretações das propriedades físico-químicas dos peptídeos por meio dos modelos, de modo a encontrar um significado simplificado na otimização e máxima cobertura das proteínas.
5. Propor para a comunidade científica a teoria do “modelo químico-matemático para aumentar a cobertura em proteômica *shotgun*”.

2.3. PARTE EXPERIMENTAL

2.3.1. Padrões e reagentes

Os padrões para a mistura de proteínas e fosfoproteínas, albumina de soro bovino (BSA) - A7906, mioglobina - M0630, lisozima - L6876, caseína e suas isoformas (α S1, α S2, β , κ) - C3400, hemoglobina humana e suas isoformas (HBB, HbD, HBG2, HbA) - H7379, lactalbumina - L7252, foram adquiridas da Sigma-Aldrich (São Paulo, SP, Brasil). O padrão de extrato das proteínas de levedura (células de *Saccharomyces cerevisiae*) – V734A foi adquirido da Promega (Madison, WI, EUA).

Os capilares de sílica fundida com revestimento externo de poliimida foram adquiridos da Polymicro Technologies (Phoenix, AZ, EUA). Os capilares de sílica fundida desativado e revestimento externo de poliimida foram adquiridos da Agilent Technologies (Santa Clara, CA, EUA). As partículas de fase estacionária C₁₈ foram adquiridas da Phenomenex. As partículas de fase estacionária da trocadora catiônica forte (SCX) foram adquiridas da Whatman (Clifton, NJ, EUA). Os reagentes para confecção dos frits nas colunas foram: Kasil® 1624 da The PQ Corporation (Valley Forge, PA, EUA) e formamida da Sigma-Aldrich (St. Luis, Missouri, EUA).

Os reagentes utilizados para o preparo de amostra foram: ureia (Sigma-Aldrich), bicarbonato de amônio (Sigma-Aldrich), ditiotreitól (DTT) (Fluka Analytical), cloreto de cálcio (Mallinckrodt GenAR), iodoacetamida (GE Healthcare), Tris – (hidroximetil)aminometano (Sigma-Aldrich), TCEP – (tris(2-carboxietil)fosfina) (Sigma-Aldrich), e tripsina (Sigma-Aldrich e Promega). Os solventes, sais e ácidos utilizados na composição das fases móveis foram: acetonitrila grau HPLC (Tedia), ácido fórmico 98% de pureza (Fluka Analytical), ácido trifluoroacético 99% de pureza (Sigma-Aldrich), cloreto de sódio (Mallinckrodt) e acetato de amônio (Sigma-Aldrich). A água utilizada para preparar todas as amostras e soluções foi purificada com sistema Milli-Q da Milipore.

2.3.2. Equipamentos e acessórios

- Balança analítica Adventurer™ Ohaus®

- pHmetro de bancada Tecnal Tec-3MP
- Laser puller Model P-2000 da Sutter Instrument CO.
- Eppendorf Centrifuge 5415R
- Thermomixer da eppendorf
- Agitador de tubos tipo vortex da Scientific Industries
- Coluna cromatográfica empacotada com partículas hidrofóbica de octadecil da *Thermo Scientific PepMap®* (75 μm d.i. $\times$ 15 cm, C₁₈, 3 μm , 100 Å)
- Partículas hidrofóbicas de octadecil da *Phenomenex* resina Aqua (C₁₈, 5 μm de diâmetro e 125 Å de tamanho de poro)
- Partículas trocadoras catiônicas com grupo de ácido sulfônico benzeno *Partisphere Whatman*, 5 μm de diâmetro e 120 Å de tamanho de poro.
- Cromatógrafo líquido de nanofluxo Easy-nLC II da *Thermo Scientific*
- Espectrômetro de massas sequencial híbrido quadrupolo de armadilha linear com Orbitrap da *Thermo Scientific - LTQ Orbitrap Velos Pro*

2.3.3. Procedimento experimental

2.3.3.1. Modelos: BSA digerido e Mix de proteínas digeridas

Estes modelos foram conduzidos nos equipamentos: *i)* cromatógrafo líquido micro/capilar *Ultra high pressure liquid chromatography* (UHPLC capilar), série UltiMate 3000 da Dionex nanoRSLC acoplado ao *ii)* espectrômetro de massas sequencial micrOTOF-QII Bruker. Todo o procedimento experimental é detalhado em Batiston (2015), disponível na Biblioteca Digital da USP: <<https://teses.usp.br/teses/disponiveis/75/75135/tde-07052015-110233/pt-br.php>> [36] Assim, nesta tese será norteada toda a interpretação química destes modelos.

2.3.3.2. Modelo: proteoma digerido de leveduras

Este modelo foi conduzido nos equipamentos: *i)* cromatógrafo líquido de nanofluxo *Easy-nLC II* acoplado ao *ii)* espectrômetro de massas sequencial *LTQ Orbitrap Velos Pro*. A identificação e sequenciamento dos peptídeos foram conduzidas pelo algoritmo *ProLuCID* qual é o principal componente do software *Integrated Proteomics Pipeline* (IP2.) Portanto, o procedimento experimental será retratado nesta tese. [14]

2.3.3.3. Preparação das colunas capilares analíticas de fase reversa

2.3.3.3.1. Confecção da ponta para ionização por eletronebulização

1. Por meio de uma lâmina de cerâmica, cortou-se aproximadamente 40 cm de capilar de sílica fundida, contendo revestimento externo de poliimida com 100 µm de diâmetro interno (d.i.).
2. Com o uso de um castiçal de lamparina, queimou-se uma área de aproximadamente 2 cm do segmento central do capilar, assim retirou-se o revestimento de poliimida do capilar e limpou-se a região com etanol ou metanol.

3. Para efetivo corte, o capilar foi alinhado de maneira retilínea no *laser puller*, proporcionando assim, focalização do feixe de laser na parte sem o revestimento de poliimida do capilar.
4. Selecionou-se o programa do equipamento conforme os valores da **Tabela 4** abaixo:

Tabela 4 - Programa para o corte dos capilares de sílica fundida acima de 100 µm de diâmetro interno (d.i.) × 365 µm de diâmetro externo (d.e.). Por conveniência as siglas da programação estão na língua inglesa.

Cycle1:	Heat=280,	FIL=0, VEL=40,	DEL=200,	PUL=0
Cycle2:	Heat=270,	FIL=0, VEL=30,	DEL=200,	PUL=0
Cycle3:	Heat=260,	FIL=0, VEL=25,	DEL=200,	PUL=0
Cycle4:	Heat=250,	FIL=0, VEL=20,	DEL=200,	PUL=0
Cycle: ciclo; Heat: aquecimento; FIL: arquivo; VEL: velocidade; DEL: delta				

5. Como resultado final, para cada capilar de 40 cm, foram confeccionadas duas colunas capilares com pontas de aproximadamente 5 µm d.i., para a eletronebulização no espectrômetro de massas.

2.3.3.3.2. Empacotamento das colunas analíticas de fase reversa

1. Em um tubo de 1,5 mL contendo 100% de metanol (MeOH), suspendeu-se uma pequena quantidade de partículas com grupo C₁₈ na superfície e agitou-se em *vortex* até completa homogeneização.
2. Colocou-se o tubo contendo MeOH e fase estacionária, no interior de uma bomba de empacotamento de colunas *homemade*, a qual permitiu pressão de arraste de He ou N₂ a 800 psi. A bomba de fluxo foi hermeticamente selada e o capilar com ponta para ionização foi alocado no orifício central da tampa da bomba, de modo que a base do capilar atingisse a solução MeOH/fase estacionária e a ponta para eletronebulização funcionasse como saída do solvente MeOH.
3. Abriu-se a válvula para liberar o gás de arraste e os capilares foram preenchidos com a fase estacionária. Realizou-se o monitoramento do empacotamento a olho nu e com

auxílio de régua. Após 12 cm de comprimento no empacotamento, a pressão da bomba foi cessada.

4. Em seguida, equilibrou-se a coluna cromatográfica analítica com a fase móvel B (80% de ACN, 20% água deionizada e 0,1% FA).
5. Posteriormente, equilibrou-se a mesma com fase móvel A (5% de ACN, 95% água deionizada e 0,1% FA).
6. Um procedimento semelhante deste protocolo é descrita por Richards, A. L. *et al.* [37]

2.3.3.4. Preparação das colunas cromatográficas estilo MudPIT

2.3.3.4.1. Polimerização para criação de frits nos capilares

1. Cortou-se por meio de cerâmica vários capilares de sílica fundida desativado (grupo silanol desprotonado) com revestimento de poliimida e 250 µm d.i. O comprimento final dos capilares foi de 20 cm.
2. Em um vial de vidro de 1,5 mL misturou-se 300 µL de silicato de potássio, Kasil® 1624 e 100 µL de formamida e agitou-se até completa homogeneização da solução.
3. Emergiu-se os capilares de sílica fundida desativado (grupo silanol desprotonado) no tubo contendo a mistura e, imediatamente, foram retirados (questão de segundos) e colocados em béquer de 50 mL.
4. Levou-se o béquer contendo os capilares para polimerização na estufa à 100 °C por 4 horas, como resultado obteve-se os frits (polímero poroso) na ponta dos capilares.
5. Cortou-se a região polimerizada excedente dos capilares e obteve-se *frit* com comprimento máximo de 2 cm.

2.3.3.4.2. Empacotamento das colunas MudPIT

1. Selecionou-se os capilares que possuem o *frit*.
2. O processo é semelhante ao empacotamento para as colunas analíticas de fase reversa.

3. Em um tubo de 1,5 mL contendo 100% de metanol, suspendeu uma pequena quantidade de resina de trocadora catiônica forte (partículas da fase estacionária SCX) e agitou-se até completa solubilização.
4. Por meio da bomba *homemade* para gás de arraste de He ou N₂, preencheu-se a fase estacionária nos capilares até 2,5 cm de comprimento.
5. Desligou-se a pressão do gás, para troca de solução contendo a fase estacionária.
6. Em um tubo de 1,5 mL contendo MeOH, suspendeu-se uma pequena quantidade de resina de fase estacionária hidrofóbica (partículas de C₁₈) e agitou-se até completa solubilização.
7. O procedimento 4 foi repetido, contudo o empacotamento foi realizado com partículas de C₁₈, até atingir 5 cm de comprimento no capilar. Portanto, no mesmo capilar empacotou-se 2,5 cm para partículas de SCX e 2,5 cm para partículas de C₁₈, totalizando 5 cm de fases estacionárias.
8. Em seguida, equilibrou-se o MudPIT com a fase móvel B (80% de ACN, 20% água deionizada e 0,1% FA).
9. Posteriormente, equilibrou-se com a fase móvel A (5% de ACN, 95% água deionizada e 0,1% FA).

2.3.3.5. Preparo das amostras

1. 100 µg do extrato de proteínas intactas de levedura (células de *Saccharomyces cerevisiae*) foi dissolvido em solução de 8 mol L⁻¹ de ureia gerando aproximadamente 60 µL, 100 mmol L⁻¹ de tampão Tris pH 8,5 - (hidroximetil)aminometano - o ajuste do pH foi realizado por meio da adição de HCl diluído.
2. Em seguida, realizou-se a redução das proteínas, desta forma, foi adicionado uma solução de TCEP de 100 mmol L⁻¹ para uma concentração final de 5 mmol L⁻¹. A solução foi agitada e centrifugada cuidadosamente (de modo a prevenir que não houvessem gotas da solução na parede do tubo na parte superior). Posteriormente, incubou-se em temperatura ambiente por 20 minutos.

3. A alquilação foi realizada por meio da adição da solução de ClAA 500 mmol L⁻¹, para concentração final de 50 mmol L⁻¹. A solução foi submetida a agitação, centrifugação cuidadosa e incubada a temperatura ambiente por 15 minutos sob proteção da luz.
4. As amostras foram submetidas a diluição, por um fator de 4× com 100 mmol L⁻¹ de Tris pH 8,5. Assim, obteve-se uma concentração final de ureia de no máximo 2 mol L⁻¹.
5. Adicionou-se solução de CaCl₂ 100 mmol L⁻¹ para obter uma concentração final de 1 mmol L⁻¹.
6. A digestão das proteínas foi realizada por meio da solução de tripsina 0,5 µg µL⁻¹ na proporção de 1:50 enzima/substrato, seguido por incubação durante 16 horas a 37 °C em *Thermomixer* sob agitação de 300 rpm.
7. A reação enzimática foi inativada pela adição de ácido fórmico para uma concentração final de 5%.
8. Centrifugou-se as amostras em alta velocidade, 2500 rcf (força centrífuga relativa) ou força g (gravitacional), por 15 minutos em temperatura ambiente e o sobrenadante foi transferido para um novo tubo com descarte do *pallet*.
9. As amostras foram analisadas no método analítico ou estocadas em freezer -80 °C, para posteriores análises.

2.3.3.5.1. Configurações adotadas no cromatógrafo *Easy-nLC 1000* e espectrômetro de massas *LTQ-Orbitrap*

Estes experimentos foram realizados no período do estágio no exterior por meio da Bolsa Estágio de Pesquisa no Exterior (BEPE) concedida pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). O laboratório anfitrião foi do Professor Ph.D John R. Yates III do The Scripps Research Institute, CA, Estados Unidos.

2.3.3.5.2. Método cromatográfico unidimensional acoplado à espectrometria de massas sequencial de alta resolução

A instrumentação do método consiste de coluna analítica da *Thermo Scientific PepMap®* (75 μm d.i. $\times$ 15 cm, C₁₈, 3 μm , 125 Å), cromatógrafo líquido de nanofluxo *EASY n-LC 1000* da *Thermo Scientific*, fonte de eletronebulização *EASY-Spray Thermo Scientific* e espectrômetro de massas sequencial tipo híbrido LTQ-Orbitrap *Thermo Scientific*, conforme ilustrado pela **Figura 12**. As fases móveis utilizadas como gradiente na coluna analítica foram: A) 5% ACN, 95% H₂O e 0,1% FA e B) 80% ACN, 20% H₂O e 0,1% FA.

Figura 12 - Configuração e equipamentos do sistema LC-MS/MS para avaliação dos parâmetros de eletronebulização. A) Fonte de eletronebulização *EASY-Spray Thermo Scientific*. B) Cromatógrafo líquido de nanofluxo *EASY n-LC 1000*, com coluna cromatográfica e fonte de eletronebulização acoplado ao espectrômetro de massas sequencial *LTQ-Orbitrap Velos Pro*.

Fonte: Adaptação de Batiston (2019) a partir de Thermo Fisher SCIENTIFIC. Estados Unidos. Disponível em: <<https://www.thermofisher.com/order/catalog/product/ES081>>. Acesso em: 02 dez. 2019.

2.3.3.5.3. Método cromatográfico multidimensional (MudPIT) acoplado à espectrometria de massas sequencial de alta resolução

A instrumentação do método consiste de coluna analítica *homemade* (100 μm d.i. $\times$ 15 cm, C₁₈, 5 μm , 100 Å), coluna MudPIT (250 μm d.i. $\times$ 2,5 cm, SCX, 5 μm , 120 Å com 2,5 cm, C₁₈, 5 μm , 125 Å), cromatógrafo líquido de nanofluxo *EASY n-LC 1000* da *Thermo Scientific*, aplicação de potencial para eletronebulização na base inferior da coluna e espectrômetro de massas sequencial tipo híbrido *LTQ-Orbitrap Thermo Scientific*. As fases móveis utilizadas como gradiente na coluna analítica foram: A) 5% ACN, 95% H₂O e 0,1% FA e B) 80% ACN, 20% H₂O e 0,1% FA. Preparou-se como solução estoque, o tampão acetato de amônio em pH 4,75. Os plugues de solução salina volátil foram preparados a partir desta solução estoque, por meio da

diluição com água deionizada em 7 concentrações distintas (5%, 15%, 30%, 45%, 60%, 75%, 90% e 100% de acetato de amônio) em frasco de polipropileno de 200 µL. Paralelamente e dependendo do experimento, cada método de plugue de solução salina foi ajustado para diferentes concentrações de ACN no tampão (5%, 9%, 13%, 17%, 21% de ACN), conforme o estudo quimiométrico.

2.3.3.5.4. Configuração do método espectrométrico

O espectrômetro de massas sequencial tipo híbrido *LTQ-Orbitrap Velos* foi configurado com resolução de escaneamento de 60.000 no analisador Orbitrap, no modo de escaneamento total (*full scan*) e polaridade positiva, seguido pela seleção dos três picos mais intensos para fragmentação. O número de *microscan* tanto para MS¹ (Orbitrap) quanto para MS² (íon trap) foi de 1. O tempo de injeção máxima dos íons no MS¹ foi de 10 ms, enquanto que para o MS² foi de 100 ms. O tipo de ativação para fragmentação das biomoléculas foi o CID, com valor mínimo de intensidade tolerável de 1.000, largura do pico mínimo tolerável de 2, energia de colisão normalizada de 35(%), estado de carga +2 a +4 e intervalo de escaneamento de massas de 300 a 2000 *m/z*. As configurações de exclusão dinâmica foram: contagem de repetição 1; duração da repetição 30 s; tamanho da lista de exclusão 500; duração da exclusão 1 min. A exclusão de massas por meio da largura de massa inferior e superior foram de 0,51 e 1,5, respectivamente. A temperatura do capilar e tensão elétrica para eletronebulização para a formação do spray foram avaliadas conforme o estudo quimiométrico no intervalo de 225 a 375 °C e 2,0 a 2,5 kV, respectivamente.

2.3.3.5.4. Sequenciamento e identificação dos peptídeos em proteínas e parâmetros do Software IP2 – Integrated Proteomics Pipeline Ver. 6.5.

A interpretação dos espectros adquiridos do proteoma de levedura foi realizada pelo algoritmo computacional *ProLuCID* o qual é o principal componente do Software *Integrated Proteomics Pipeline* (IP2, Integrated Proteomics Applications, Inc. San Diego, CA), versão

6.5. IP2 é uma plataforma operacional em desenvolvimento no *Yates Lab* o qual permite estudo abrangente do proteoma. [14]

O banco de dados de proteínas de leveduras (*Saccharomyces cerevisiae*) foi adquirido na plataforma *online* Uniprot. No software IP2 selecionou-se as condições instrumentais idênticas ao método do espectrômetro de massas de alta resolução (*LTQ Orbitrap Velos Pro*), assim como no preparo da amostra. Desta forma, na aba “*parâmetros básicos*” selecionou-se as seguintes opções: tipo de fragmentação ou método de ativação CID/HCD; massa monoisotópica; seleção de massa do fragmento no modo monoisotópica; tolerância de massa do peptídeo precursor de 50 ppm ($10^{-6} \text{ m}/\Delta m$). O intervalo de massa do precursor foi de 600 a 6.000, o intervalo de carga do precursor foi de 0 a 1.000 a tolerância de massa do fragmento foi de 600 ppm ($10^{-6} \text{ m}/\Delta m$). Na aba “*informação da enzima*” a opção especificidade foi de um fim (*one end*), número máximo de clivagem perdida de 3; a protease selecionada foi a tripsina, resíduos KR (lisina e arginina), posição de clivagem em C-terminal (carboxi-terminal). Na aba “*modificações estáticas/fixas*”, selecionou somente a modificações estática específica do resíduo de aminoácido como 57,02146 para uma modificação carboximetil. Na aba “*parâmetros avançados*” o número máximo de varreduras em arquivos espectrais divididos foi de 4.000. Na aba “*modificações diferenciais*” selecionou-se como 2 para número máximo de modificações diferentes internas, para modificação diferencial selecionou-se 15,994915 M a qual trata-se da oxidação no resíduo metionina. Na aba de pesquisa denominada “*rotulagem metabólica*” não se utilizou nenhum parâmetro. Na aba “*parâmetros avançados de pesquisa ProLuCID*”, o tipo de score principal foi o XCorr, tipo de score secundário foi Zscore, selecionou-se *não* para utilização de íons de baixa massa de fragmento, selecionou-se *zero* para modo de ativação de vários estágios, o comprimento mínimo de peptídeos no sequenciamento de uma proteína foi de 6, o limite (*threshold*) para o candidato a peptídeos foi de 300 ppm ($10^{-6} \text{ m}/\Delta m$) e as diferenças terminais peptídicas ou modificações variáveis foram zero tanto para peptídeo N-terminal, quanto para C-terminal. Na aba “*análise das localizações PTM*”, não se utilizou o cálculo de pontuação de localização das modificações pós-traducionais. Na aba “*parâmetros básicos do DTASelect 2.0*” em número mínimo de peptídeos por proteína digitou-se 2, para número mínimo de extremidade trípica por peptídeo digitou-se 1, taxa de falsos positivos por proteína foi 0.01, corte delta em massa do precursor foi 10, deslocamento em massa delta do precursor foi “*não utilizado*”, para o pré-filtro

de ruído foi “não utilizado”. Na aba “*parâmetros avançados do DTASelect 2.0*” dentre as possibilidades de parâmetro, selecionou-se apenas o modo de exibição dos espectros como 1 e selecionou *sim* para as seguintes opções: somente peptídeos único, estatísticas com modificações, estatística com status trípico e incluir subconjunto de proteínas, sendo as outras opções selecionadas como não.

2.4. RESULTADOS

2.4.1. Superfície de resposta para uma mistura proteolítica de sete biomoléculas

Os resultados obtidos neste tópico fazem referência ao método desenvolvido no período do mestrado, contudo validado e analisado no primeiro ano do doutoramento, uma vez que demandaram de maiores experimentações e interpretações para comprovação e entendimento dos achados científicos.

As aplicações deste tópico são provenientes do método LC×LC-MS/MS, denominada pela literatura como abrangente *on-line* com parada de fluxo na 1D. A primeira dimensão cromatográfica é uma trocadora catiônica forte (SCX) seguida por um *trap* de dessalinização e limpeza das frações da amostra, coluna capilar hidrofóbica com fase estacionária de partículas de C₁₈ (fase reversa - RP), separação por *m/z* e detecção por espectrometria de massas tipo híbrido, cujo analisadores foram quadrupolo e tempo de voo (Q-ToF) - SCX×RP-Q-ToF-MS/MS.

Em trabalho anterior, o método SCX×RP-Q-ToF-MS/MS foi quimiometricamente otimizado utilizando como padrão o digerido trípico de BSA. [36] Após a aplicação da alternativa quimiométrica *m-DoE* (**Capítulo 1**), obteve-se superfície de resposta por meio de dois parâmetros instrumentais, onde as respostas avaliadas e otimizadas foram *score* (3816,29), número de peptídeos (73) e cobertura da proteína (92,92%, **Figura 13A**). [36]

De modo a validar o modelo, aplicou-se amostra de maior complexidade sob as mesmas condições experimentais. Desta forma, validou-se o método por meio de 13 espécies digeridas entre proteínas, fosfoproteínas e isoformas. Assim, verificou-se que ambas as superfícies de respostas possuem similaridades (**Figura 13A e Figura 13B**). Esta observação permite afirmar com maior respaldo que a percentagem de acetonitrila no tampão fosfato, fase móvel da primeira dimensão cromatográfica, assim como a pressão do gás nebulizador, referente a etapa de eletronebulização, são de fato os fatores significativos de todo o método SCX×RP-Q-ToF-MS/MS.

As **Tabela 5** e **Tabela 6** indicam os valores provenientes da mistura proteolítica das sete biomoléculas analisadas, referente aos dois melhores resultados experimentais da superfície de resposta. Verificou-se cobertura média de 83,07% para as espécies identificadas (entre proteínas,

fosfoproteínas e isoformas), atingindo até 100% de cobertura para 45,45% das espécies analisadas. Adicionalmente, notou-se que quanto maior a cobertura das proteínas, melhor é a diferenciação das isoformas. Estes achados podem contribuir para a proteômica multidimensional de biologia celular, em termo de uma melhor caracterização das expressões de isoformas em proteômica *shotgun* e *bottom-up*, conforme relatadas na **Tabela 3**. [5]

Figura 13 - Planejamento experimental composto central, para proteínas digeridas A) Superfície de resposta referente a BSA digerida. [36] B) Superfície de resposta referente a mistura de 13 espécies entre proteínas, fosfoproteínas e isoformas. Ambas aplicadas no método SCX×RP-Q-ToF-MS/MS.

Fonte: Adaptação de Batiston (2019) a partir de BATISTON, W.P. Quimiometria aplicada à cromatografia líquida multidimensional capilar hifenizada a espectrometria de massas sequencial para proteômica *shotgun*. Dissertação de mestrado, Brasil, 2015, p. 92 (A).

2.4.2. Avaliação da otimização quimiométrica do método LC×LC-MS/MS

Podemos afirmar que a elevada cobertura das proteínas e número de peptídeos identificados (**Tabela 5** e **Tabela 6**), se devem a dois pontos essencialmente analíticos: *i*) a elevada ortogonalidade da plataforma SCX×RP (**Figura 14**) e *ii*) o excelente perfil, assim como intensidade do sinal analítico do método MS/MS (**Figura 15**).

Tabela 5 - Dados da melhor condição experimental da **Figura 13B** para uma mistura proteolítica de 13 espécies digeridas (entre proteínas, fosfoproteínas e isoformas) submetidas ao método SCX×RP-Q-ToF-MS/MS.

#	Acesso Uniprot	Descrição	Score	Cobertura	Peptídeos*	Peptídeos únicos
1	P02769	Albumina de Soro Bovino (<i>Bos taurus</i>)	2143,02	100,00	153	61
2	P68871	β-Hemoglobina (<i>Homo sapiens</i>)	13,08	100,00	19	4
2	P02042	δ-Hemoglobina (<i>Homo sapiens</i>)	10,27	100,00	22	4
2	P69892	γ-Hemoglobina gamma-2 (<i>Homo sapiens</i>)	8,30	100,00	19	1
2	P69905	α-Hemoglobina (<i>Homo sapiens</i>)	1,61	59,86	17	3
3	P68082	Mioglobina (<i>Equus caballus</i>)	4,70	88,96	36	3
4	P02668	Kappa-caseína (<i>Bos taurus</i>)	2,76	37,89	13	3
4	P02663	Alfa-S2-caseína (<i>Bos taurus</i>)	1,93	79,28	47	2
4	P02662	Alfa-S1-caseína (<i>Bos taurus</i>)	1,75	79,44	25	1
4	P02666	Beta-caseína (<i>Bos taurus</i>)	-	-	-	-
5	P00709	Alfa-lactalbumina (<i>Homo sapiens</i>)	0,00	68,31	23	4
6	P00698	Lisozima C (<i>Gallus gallus</i>)	1,65	100,00	29	3
7	P00125	Citocromo C1 (<i>Bos taurus</i>)	nS	nS	nS	nS

proteínas com mesma numeração e cor indicam isoformas da biomolécula em questão. *Total de peptídeos identificados (403), nS - não sequenciado.

Tabela 6 - Dados da segunda melhor condição experimental da **Figura 13B** para uma mistura proteolítica de 13 espécies digeridas (entre proteínas, fosfoproteínas e isoformas) submetidas ao método SCX×RP-Q-ToF-MS/MS

#	Acesso Uniprot	Descrição	Score	Cobertura	Peptídeos*	Peptídeos únicos
1	P02769	Albumina de Soro Bovino (<i>Bos taurus</i>)	118,65	99,34	141	22
2	P68871	β-Hemoglobina (<i>Homo sapiens</i>)	82,29	100,00	21	6
2	P02042	δ-Hemoglobina (<i>Homo sapiens</i>)	72,80	100,00	23	3
2	P69892	γ-Hemoglobina (<i>Homo sapiens</i>)	50,85	100,00	25	1
2	P69905	α-Hemoglobina (<i>Homo sapiens</i>)	10,40	80,28	19	5
3	P68082	Mioglobina (<i>Equus caballus</i>)	45,20	88,96	35	8
4	P02662	Alfa-S1-caseína (<i>Bos taurus</i>)	33,00	80,37	26	6
4	P02663	Alfa-S2-caseína (<i>Bos taurus</i>)	3,91	82,88	47	6
4	P02668	Kappa-caseína (<i>Bos taurus</i>)	3,91	52,63	12	2
4	P02666	Beta-caseína (<i>Bos taurus</i>)	0,00	52,23	17	1
5	P00711	Alfa-lactalbumina (<i>Bos taurus</i>)	0,00	70,42	19	3
6	P00698	Lisozima C (<i>Gallus gallus</i>)	8,70	100,00	35	6
7	P00125	Citocromo C1 (<i>Bos taurus</i>)	nS	nS	nS	nS

proteínas com mesma numeração e cor indicam isoformas da biomolécula em questão. *Total de peptídeos identificados (420), nS - não sequenciado

Figura 14 - Gráfico 2D com tempo de retenção normalizado do método SCX×RP-Q-ToF-MS/MS. A) para BSA digerida. B) 13 espécies digeridas.

Fonte: Autoria própria.

Figura 15 - Cromatograma de íon total da mistura de proteínas digeridas da melhor condição instrumental da **Figura 13** e proveniente do resultado da **Tabela 5**.

Fonte: Autoria própria.

A avaliação da ortogonalidade da modalidade SCX×RP foi realizada por meio de gráficos cujos eixos são os tempos de retenção dos peptídeos de cada dimensão cromatográfica, conforme

a estratégia reportada por Gilar *et al.* [26] e denominada como ortogonalidade geométrica. Devido a modalidade do método ser *on-line*, realizaram-se algumas adaptações, em conformidade com os estudos de Motoyama, *et al.* [38] Assim a identificação dos peptídeos e os dados do tempo de retenção de ambas as dimensões cromatográficas foram obtidas no programa *Sequest*. Em especial, na 1D foi considerado a retenção de cada fração de eluição dos plugues de solução salina (por esta razão há espaços simétricos em branco no eixo das ordenadas do gráfico 2D). Desta forma, os tempos de retenção foram normalizados, $RT_{i(norm)}$, de acordo com a **Equação 9**, onde RT_i é o tempo de retenção do peptídeo questionado, RT_{max} e RT_{min} são o tempo de retenção do peptídeo de maior e menor grandeza, respectivamente.

$$RT_{i(norm)} = \frac{RT_i - RT_{min}}{RT_{max} - RT_{min}} \quad \text{Equação 9}$$

A **Figura 14** mostra a alta ortogonalidade do nosso método, devido a baixa correlação entre os tempos de retenção dos diferentes peptídeos e, portanto, alto grau de separação na área 2D. Estes resultados estão de comum acordo com trabalhos anteriores. [24-26]

A elevada ortogonalidade e seletividade da modalidade SCX×RP tem importante influência sobre os resultados em proteômica, devido à diminuição da complexidade amostral na separação, proporcionando melhor detecção e caracterização biomolecular no MS/MS. Como consequência, diminui-se a eluição dos analitos de mesma massa e tempo de retenção, minimizando a fragmentação de íons/peptídeos de mesma m/z para a etapa de isolamento do MS/MS. Este efeito é denominado como mistura co-fragmentada ou espectro “*chimera*” sendo um assunto que tem recebido muita atenção na proteômica *shotgun*, uma vez que o fenômeno gera resultados tendenciosos no sequenciamento e identificação das proteínas. [39]

Cientistas em bioinformática têm desenvolvido alternativas para diminuir o efeito *chimera*, por meio do tratamento estatístico dos espectros de massas. Adicionalmente, torna-se importante o desenvolvimento de espectrômetros de massas de alta resolução ou estratégias complementares como a mobilidade de íons (*ion mobility-MS/MS*), para aumento da taxa de isolamento dos íons e consequente melhora na detecção do MS.

É interessante mencionar que os nossos resultados, no quesito aumento do poder de separação, reforçam as evidências previstas por **Kocher *et al.***, onde o grupo demonstrou, de

maneira muito clara, a correlação positiva linear entre a capacidade de pico em LC e o número de peptídeos identificados. O grupo em questão utilizou a modalidade cromatografia líquida unidimensional de coluna de fase reversa (C₁₈) e detecção em espectrometria de massas sequencial no equipamento LTQ-Orbitrap Velos. Podemos concluir que métodos oriundos com elevada eficiência de separação, são essenciais para a caracterização total do proteoma. [40]

Conforme mencionado no início do tópico, outro ponto importante do nosso estudo foi a elevada qualidade dos cromatogramas de íons obtidos em MS/MS e consequentemente os ótimos resultados no sequenciamento dos peptídeos. Portanto, é de extrema importância que estudos futuros voltados para proteômica evidenciem os cromatogramas de íons de modo a atribuir a qualidade da análise química, permitindo assim maiores clarezas nos resultados e interpretações para a biologia molecular.

2.4.3. Estudo do sequenciamento dos peptídeos em proteínas

O sequenciamento das proteínas referente às **Tabela 5** e **Tabela 6** foi realizado pelo *software* *Sequest*, o qual utilizou o banco de dado *on-line* proveniente da plataforma *Uniprot*. Portanto, o primeiro passo é realizar o *download* dos espectros de massas sequenciais (MS/MS) deste banco de dados. Desta forma, tornam-se necessárias algumas considerações de extrema importância na identificação correta do arquivo para que o sequenciamento seja adequado: *i*) é necessário realizar o *download* das proteínas ou proteoma investigado e estes devem ser provenientes do organismo em questão, ou seja, não é razoável sequenciar mioglobina humana se no banco de dados contém mioglobina de *Equus caballus*; *ii*) deve-se atentar a informação genética dos organismos, portanto o analista deve possuir informação do(s) gene(s) de sua amostra e correlacionar com o arquivo do banco de dados; *iii*) a proteína a ser sequenciada deve possuir o banco de dados proveniente do mesmo órgão biológico, ou seja, uma proteína mitocondrial de coração bovino, não poderá ser sequenciada por um arquivo de MS/MS proveniente de uma proteína mitocondrial da rede respiratória.

Conforme mencionado no tópico anterior, os nossos resultados foram baseados no sequenciamento de 13 espécies, entre proteínas, fosfoproteínas e isoformas, sendo que somente uma delas não foi recuperada, trata-se do citocromo C, conforme reportado na **Tabela 5** e **Tabela**

6. Provavelmente, o sequenciamento não foi possível conforme explicitado no item *iii*) do parágrafo anterior. Realizamos o sequenciamento com citocromo C mitocondrial bovino da rede celular de mesmo gene, contudo não obtemos sucesso nesta estratégia. Desta forma, foi enviado um e-mail a empresa relatando o fato, contudo não se obteve resposta. Outra justificativa, para o não sequenciamento do citocromo C, pode ser advinda da característica hidrossolúvel da biomolécula em questão. Possivelmente, os peptídeos gerados possuem elevada solubilidade em água. Desta forma, estes podem ter sido descartados na etapa de dessalinização do método.

Além disso, geralmente a digestão trípica desta pequena proteína gera somente 12 peptídeos e estes podem ter sido correlacionadas de forma errônea pelo algoritmo do *Sequest*, computando para a cobertura das outras proteínas. Aqui podemos evidenciar uma das limitações da proteômica *shotgun*.

Com as considerações acima, torna-se inviável realizar o *download* de um proteoma, uma vez que a amostra deste estudo possui somente 13 espécies e estas são provenientes de 4 diferentes organismos, tais como (nome científico): *Bos taurus*, *Gallus gallus*, *Equus caballus* e *Homo sapiens*. A plataforma *Uniprot* possui uma estratégia interessante para o *download* de proteínas selecionadas, cujo intuito é montar um arquivo de espectros de MS/MS personalizado. Para esta finalidade, a cada proteína de interesse em <http://www.uniprot.org/uniprot/> é possível selecionar e adicionar a mesma em uma cesta (botão *Add to basket*), após selecionar todas as proteínas necessárias, basta clicar em “*Basket*” e realizar o *download*, assim todos os espectros de MS/MS das proteínas selecionadas ficarão restrita em somente um arquivo (banco de dados), ou seja, o analista possui um arquivo personalizado conforme a necessidade de seu sequenciamento.

[41]

Ao realizar o *download* é possível escolher várias formas de arquivo, uma delas é somente a “canônica” e outra é “canônica e suas isoformas”. A opção canônica trata-se de uma estratégia para reduzir a redundância nos resultados das proteínas sequenciadas. Desta forma, ela descreve todos os produtos das proteínas codificadas de um dado gene e uma sequência canônica baseada em alguns critérios para a elucidação da(s) proteína(s) investigada(s). Tais critérios podem ser: *i*) proteínas predominantes; *ii*) maior semelhança na sequência peptídica de ortólogos encontrado em outras espécies; *iii*) em virtude do comprimento ou quantidade de aminoácidos que permitam a descrição mais clara dos domínios, isoformas, polimorfismo e modificações pós-traducionais;

iv) na ausência de qualquer informação é escolhido a sequência mais longa. Portanto, tratam-se de diretrizes o qual o algoritmo do *software Sequest* irá se basear para melhor interpretação dos resultados.

2.4.4. Desenvolvimento do modelo-químico matemático

Como a estratégia *m-DoE* permite avaliar o sistema otimizado com maior abrangência, devido os vários planejamentos experimentais realizados (**Capítulo 1**), elaborou-se a ideia de transformar este resultado em uma superfície de resposta com maior significado na otimização, assim como melhor previsão nos resultados, ou seja, um modelo matemático abrangente, que evidencie os pontos fortes e fracos da técnica SCX×RP-MS/MS. E ainda, por meio de tal estratégia, compreender em termos químicos a perda de peptídeos do método e desta forma prever a maximização da cobertura da sequência peptídica para amostras análogas em proteômica *shotgun*. Aplicou-se esta estratégia por meio da validação de um segundo e terceiro modelo químico-matemático, o qual será discutido adiante.

Assim, o primeiro passo para gerar uma superfície de resposta de maior abrangência é encontrar uma equação matemática que evidencie o comportamento real da otimização do sistema avaliado, para tal finalidade utiliza-se os fatores significativos, os quais são as variáveis independentes da equação, de modo que a resposta é a variável dependente. Normalmente, tal equação é encontrada por meio da matriz do planejamento experimental, em conjunto com ferramentas estatísticas (Tabela de Análise de Variância – ANOVA), as quais avaliam se os valores experimentais estão de acordo com os valores previstos na superfície de resposta, de modo que haja coerência estatística entre os resultados experimentais e o modelado. [42-44]

As modelagens partem de uma superfície tridimensional, uma vez que se tem normalmente dois fatores significativos na resposta de otimização. Portanto um polinômio de segunda ordem é a função matemática que representa de forma adequada esta superfície. Neste caso, temos a seleção de dois fatores significativos, de modo que o eixo *z* é resultado da resposta dos eixos *x* e *y*, abscissas e ordenadas, respectivamente da superfície tridimensional. Caso a superfície em questão não esteja bem ajustada (valores experimentais coerentes com valores experimentais), um ou mais coeficientes do polinômio podem ser retirados de modo à melhorar a

previsão da equação matemática e, portanto, satisfazer uma modelagem próxima do comportamento real. Esta consideração somente é válida por meio da supervisão de testes estatísticos. É importante ressaltar que a constatação de uma superfície de resposta ajustada é realizada pelos valores obtidos da Tabela ANOVA. [42-44]

A maneira mais fácil de encontrar a equação de um modelo matemático (superfície de resposta) é por meio de programas estatísticos, tais como: *Statistica*, *Design Expert* e outros. A desvantagem é que tais programas são plataformas fechadas e, portanto, não permitem que a equação matemática extrapole além da função quadrática padrão ou linear já estabelecida, assim como, o número de linhas e colunas da matriz matemática (experimentos $\times$ fatores) devem possuir quantidade exata conforme o modelo matemático padrão do *software*. Somente desta maneira o programa permite o seu processamento de forma adequada. Como a proposta do modelo matemático deste estudo é inédita e tais programas não permitem atingir o nosso objetivo, ou seja: atingir um modelo matemático com maior abrangência, em outras palavras, uma matriz com valores extrapolados e, conseqüentemente, uma equação que envolva todos os experimentos de *m-DoE*. A estratégia de se utilizar “*softwares* fechados” foi descartada.

Em contrapartida, os programas com plataformas abertas, tais como: MATLAB e Octave, permitem a geração da equação por meio da programação de funções matemáticas as quais proporcionam encontrar a equação do modelo. Contudo, para tal finalidade é necessário que o usuário tenha conhecimento da matemática assim como da estatística em conjunto com a programação/linguagem computacional. Podemos extrapolar estas ideias iniciais para a multidisciplinaridade com o pessoal da informática. [45]

Além disso, a modelagem de uma superfície de resposta de maior abrangência poderia exigir novos experimentos para a elaboração da matriz. Uma vez que as matrizes realizadas em cada *DoE* do *m-DoE* são diferentes, em termos de número de linhas e colunas ($m \times n$), impossibilitando a utilização desses dados em uma nova matriz, a qual seria utilizada na matriz da superfície de resposta abrangente. Aspecto que iria exigir muito tempo seja na afirmação do parágrafo anterior (elaboração da equação matemática no MATLAB), acrescida da elaboração dos experimentos, de modo a tornar a matriz simétrica a modelagem (superfície de resposta). Portanto, propomos uma alternativa de solucionar este desafio, o qual será evidenciado no parágrafo seguinte, cuja estratégia foi aproveitar os dados já obtidos. [42]

Assim, elaboramos uma equação genérica, cujo modelo matemático tenha dois máximos. Desta forma, realizamos a adaptação da equação por meio da alteração (tentativa e erro) dos valores numéricos dos coeficientes de modo a respeitar coerência com os dados empíricos obtidos em *m-DoE* e pelo composto central (modelos estes submetidos à ANOVA), assim atingimos a **Equação 10** e definiu-se a matriz $x = [-3:3]$ e $y = x$. É importante mencionar que tais pressupostos são permitidos, se e somente se a modelagem representa o comportamento real do fenômeno a ser modelado. Tais consideração são embasadas por meio da literatura especializada em modelagem matemática. [45]

$$Z = \frac{1,7}{(1.75 + (x + 1.5)^2 + y^2)} + \frac{1,4}{(1.75 + (x - 1.5)^2 + (y - 1)^2)} \quad \text{Equação 10}$$

Desta forma, os valores das abscissas e ordenadas foram alteradas conforme a **Tabela 7**, cujos eixos (x e y), são os fatores significativos. Portanto, a superfície de resposta abrangente é resultado da matriz $x = [-3:3]$ e $y = x$, de modo que, após o processamento do modelo tridimensional, foram colocados os valores experimentais nos eixos referentes aos dois fatores significativos, conforme constatados na **Figura 16**. É interessante mencionar que o valor de x , expande-se da matriz usual do composto central, cuja extremidade é $+1,414$ e $-1,414$, fato que permite que a superfície de resposta seja explicada com maior abrangência, uma vez que sua matriz matemática possui esta característica. A **Figura 17** lista os comandos utilizados no programa MATLAB para a modelagem do modelo químico-matemático.

Tabela 7 - Valores do modelo químico matemático

Valores da matriz	-3	-2	-1	0	1	2	3
Tampão na 1D % ACN (eixo x)	4	7	10	13	16	19	22
Pressão do gás nebulizador (eixo y)	0	0,4	0,8	1,2	1,6	2,0	2,4

Como resultado, para o digerido proteico da BSA, obtivemos a superfície de resposta da **Figura 16** a qual faz referência a **Equação 10**. O eixo z da **Figura 16** faz referência aos valores da resposta. Assim a região vermelha são as condições instrumentais que fornecem os maiores

valores do *score*, cobertura da proteína e número de peptídeos. Por meio destes resultados, levantamos a hipótese que a condição ótima do método analítico, para maior cobertura das proteínas, poderia estar relacionada com a hidrofobicidade dos peptídeos. A primeira evidência para esta afirmação foi por meio da interpretação do eixo da abscissa (*x*) da **Figura 16**. Sabe-se que as colunas SCX, além do grupo sulfônico de troca catiônica, possuem também grupos hidrofóbicos em sua fase estacionária. Por esta razão, adiciona-se um pouco de solvente orgânico no tampão, o qual se constitui como parte de 5 a 25% de ACN no tampão da fase móvel da 1D. Assim, diminui-se a interação hidrofóbica do analito com a fase estacionária, tornando a interação eletrostática mais “pura”. Desta forma, a interpretação da **Figura 16** permite-nos levantar a hipótese que peptídeos hidrofílicos necessitam condições de ionização brandas, enquanto peptídeos hidrofóbicos necessitam de condições de ionização mais intensas, referente ao uso do gás nebulizador na plataforma espectrometria de massas sequencial.

Figura 16 - Modelo químico-matemático BSA, aplicado ao método SCX×RP-MS/MS: superfície de resposta como função de dois fatores para a BSA digerida. Região hidrofílica: peptídeos com baixo grau de interações hidrofóbicas não-específicas (hidrofílicos). Região hidrofóbica: peptídeos com alto grau de interações hidrofóbicas não-específicas (hidrofóbicos).

Fonte: Autoria própria.

2.4.5. Interpretação do modelo químico-matemático para o aumento da cobertura de proteínas em proteômica *shotgun* e *bottom-up*

As discussões iniciais a respeito da interpretação química da perda de peptídeos do método SCX×RP-MS/MS foram previamente discutidas em Batiston (2015). [36] Contudo, naquele momento não havia uma explicação plausível de como as propriedades físico-químicas dos peptídeos poderiam interferir na perda dos peptídeos do método em questão.

Figura 17 - Comandos do MATLAB para a geração do modelo químico-matemático, referente à **Figura 16**.

```
>> x=[-3:3];  
>> y=x;  
>> x=linspace(-3,3,100); y=x;  
>> [x,y]=meshgrid(x,y);  
>> z=1.7./(1.75+(x+1.5).^2+y.^2)+1.4./(1.75+(x-1.5).^2+(y-1).^2);  
>> mesh(x,y,z);  
>> surf(x,y,z);
```

Fonte: Autoria própria.

As informações da **Figura 16** permite entender a razão da perda de peptídeos ou porque estes não foram detectados no nosso método (LC×LC-MS/MS). Assim, nós assumimos que a pressão do gás nebulizador (N₂) entre 1,2 bar para os peptídeos com baixa hidrofobicidade e 1,6 bar para os peptídeos com alta hidrofobicidade determina a eletronebulização eficiente na escala capilar, referente a interface entre a cromatografia líquida e o espectrômetro de massas. [29] Ressaltamos, que a coluna cromatográfica utilizada neste estudo possui diâmetro interno de 300 µm com vazão de 8 µL min⁻¹. Nestas condições, torna-se necessário a utilização de um gás secante proveniente da pressão do nebulizador, referente ao processo de ionização do equipamento microQTOF-QII da Bruker.

Conforme mencionado, a primeira consideração a respeito do grau de hidrofobicidade das biomoléculas foram interpretadas por meio da quantidade de ACN no tampão na etapa de eluição da coluna SCX (eixo da abscissa da **Figura 16**), portanto quanto maior a percentagem de solvente orgânico maior o caráter hidrofóbico do peptídeo.

Verifica-se a possibilidade de interpretação conjunta dos dois fatores significativos (% ACN no tampão na 1D com a pressão do gás nebulizador na autoMS/MS), uma vez que por meio

da seletividade em termos da hidrofobicidade da coluna SCX podemos correlacionar estes resultados com a melhor condição da eletronebulização no espectrômetro de massas (observe as condições das duas regiões de máximo da superfície de resposta da **Figura 16**). Aspecto este que seria extremamente difícil de verificar por meio da otimização univariada, e ainda, uma correlação até agora inédita na literatura.

Os valores menores a 1,2 bar da pressão do gás nebulizador, provavelmente estão relacionados à incompleta evaporação da gota no processo de eletronebulização, fato que não permite sua dispersão por meio da força coulômbica (repulsão intermolecular eletrostática) em gotículas menores, com cargas de mesmo sinal. Desta forma, alguns peptídeos não são direcionados no campo elétrico entre o capilar da LC-2D e o contra eletrodo localizado na interface do Q-ToF. Por outro lado, valores superiores a 1,6 bar da pressão do gás nebulizador são capazes de vencer o campo elétrico, de forma que os peptídeos são direcionados para fora da interface Q-ToF. A **Figura 18** esquematiza o processo descrito. [29]

Figura 18 - Câmara da eletronebulização e ionização do Q-ToF. A) conexão com cromatógrafo e entrada da amostra. B) entrada do gás nebulizador (N_2). C) Agulha para pulverização com gás nebulizador com fluxo coaxial de N_2 . 1C) ionização efetiva com gás nebulizador (1,2 ou 1,6 bar). 2C) perda de peptídeos e eliminação de espécies neutras D) escudo de pulverização (-3.6 kV). E) capa do capilar (-4 kV). F) descarte G) câmara de pulverização. H) N_2 de secagem (fluxo e temperatura controlada). I) capilar de vidro. J) estágio de vácuo.

Fonte: Adaptação de Batiston (2019) a partir de BRUKER Daltonics micrOTOF II User Manual, Version 1.2 (Dezembro 2008), Alemanha.

Alguns estudos descreveram a importância da pressão do gás nebulizador para melhorar a sensibilidade, assim como o limite de detecção em análises químicas para moléculas pequenas. [46-48] Apesar do aumento da sensibilidade estar relacionado com os nossos resultados,

verificamos que não há estudos sobre o relato da importância da pressão do gás nebulizador (N₂) para o aumento de coberturas na sequência peptídica das biomoléculas. E ainda, não foi verificado nenhum estudo que relacione a eletronebulização eficiente com a hidrofobicidade dos peptídeos. Desta forma, as conclusões demonstradas em nossos trabalhos possuem uma colaboração adicional ao trabalho do grupo de **Matthias Mann**, onde o grupo constatou que de 100.000 peptídeos eluídos na cromatografia líquida unidimensional, a grande maioria dos peptídeos não foram identificados no MS, especificamente na etapa de fragmentação do MS/MS. [17]

A primeira discussão a respeito do comportamento da hidrofobicidade dos peptídeos na etapa de eluição na coluna SCX teve início há 25 anos, e a elucidação deste fenômeno ainda continua em aberto. [49-51] Isso porque o comportamento químico e o equilíbrio das espécies na solução são extremamente diversificadas. Em suma, a adição de solvente orgânico no tampão de eluição da coluna SCX, não é somente para diminuir a interação hidrofóbica e desta forma maximizar a interação eletrostática das moléculas com a fase estacionária da trocadora catiônica. Mas, além da carga líquida total dos peptídeos, outros fatores podem afetar o tempo de retenção dos peptídeos sobre a fase estacionária da coluna SCX. Portanto tratam-se de propriedades relativas à hidrofobicidade, seja na contribuição referente aos peptídeos com diferentes tamanhos de cadeia molecular, a conformação estrutural, a distribuição da carga, a densidade da carga, as regiões hidrofóbicas na molécula, e as interações inter- e intramoleculares. O conjunto destes comportamentos possui denominação na literatura de interações hidrofóbicas não-específicas. [49-51]

Conforme evidenciado em Batiston (2015), [36] investigamos o conjunto de peptídeos dos experimentos que fazem referência para as duas regiões de máximo do modelo químico-matemático referente aos peptídeos de BSA e verificamos que alguns peptídeos são detectados somente com a condição de 8% de ACN no tampão de eluição da primeira dimensão, assim como 1,2 bar da pressão do gás nebulizador e outros peptídeos são somente detectados na condição de 18% de ACN no tampão de eluição da 1D com 1,6 bar da pressão do gás nebulizador em autoMS/MS.

Constatamos que quando a primeira dimensão cromatográfica possui baixa percentagem de solvente orgânico no tampão de eluição, os peptídeos com elevado grau de hidrofobicidade

ficam retidos de forma irreversível na fase estacionária na etapa de análise, de modo que são removidos após a etapa de limpeza da coluna SCX, esta etapa é necessária para evitar o efeito de memória. Por outro lado, quando a primeira dimensão cromatográfica possui elevada percentagem de ACN no tampão de eluição, os peptídeos com baixo grau de hidrofobicidade são perdidos na etapa de limpeza e dessalinização na coluna de armadilha.

2.4.6. Comprovação da hidrofobicidade dos peptídeos nos modelos químico-matemáticos

Após o levantamento das hipóteses e descrição sobre a perda de peptídeos do método em questão, torna-se necessário provar numericamente tais evidências. Para tal, utilizamos a ferramenta *ProtParam*, proveniente da *ExPASy Bioinformatics Resource Portal* (<https://web.expasy.org/protparam/>). [52] Esta ferramenta permite a análise físico-química de um dado peptídeo ou proteína, por meio de algumas métricas cientificamente comprovadas. Dentre estas, nós utilizamos a métrica *Gravy* (do inglês *Grand average of hydropathicity*). Esta medida, foi aprimorada por J. Kyte e R.F. Doolittle como uma escala de hidrofobicidade para peptídeos e proteínas. [53] Vários *softwares* de bioinformática têm utilizado tal escala em sua plataforma. A escala funciona da seguinte maneira, cada aminoácido possui um valor de hidrofobicidade, dado uma sequência de aminoácidos o somatório destes valores, dividido pelo número de aminoácidos, trata-se do valor numérico de *Gravy*. Portanto, quanto mais negativo o valor numérico mais hidrofílico é o peptídeo ou proteína avaliada. Assim, o análogo é verdadeiro, onde quanto mais positivo o valor mais hidrofóbico é a biomolécula.

Os valores médios de *Gravy* obtidos para todos os peptídeos detectados nas regiões de máxima resposta, **A** e **B** da **Figura 16** (modelo BSA) foram $-0,421$ e $-0,350$, respectivamente. Portanto, há mais peptídeos hidrofóbicos na região **B** em relação aos peptídeos identificados na região **A** do modelo BSA. Após eliminar os peptídeos idênticos das duas regiões **A** e **B** do modelo BSA, ou seja, quando se considera somente os diferentes peptídeos de cada região de máxima resposta, a média de *Gravy* foi $-0,530$ para região **A** e $0,083$ para a região **B**. Nesta condição, os peptídeos da região **A** são 6,37 vezes mais hidrofílicos do que a região **B**.

Conforme mencionado no **tópico 2.4.1.**, verifica-se que ambas as superfícies de resposta da **Figura 13** possuem semelhanças. Contudo a **Figura 13B** possui somente um máximo de

resposta. Adicionalmente, a interpretação da **Tabela 5** e **Tabela 6**, evidenciam que algumas proteínas não possuem 100% de cobertura, ou seja, alguns peptídeos foram perdidos no processo de análise ou não foram identificados na plataforma autoMS/MS. Assim, nós levantamos a hipótese de uma segunda região de máxima resposta para a **Figura 13B**. Uma vez que as espécies não identificadas nas **Tabela 5** e **Tabela 6** poderiam fazer referência ao grupo dos peptídeos hidrofóbicos, perfazendo a segunda região de máxima resposta. Aspecto este que estaria de acordo com o modelo BSA, mas que infelizmente não foi detectada no modelo Mix da **Figura 13B**. Provavelmente, este comportamento se deve pela escolha dos intervalos dos fatores significativos ou pela natureza química de maior complexidade dos peptídeos na amostra, em relação ao modelo BSA.

Uma maneira razoável de conhecer todos os peptídeos provenientes da digestão proteolítica de um conjunto de proteínas é por meio do *Software* da *ExPASy Peptide Cutter*. Assim realizamos a digestão virtual triptica das proteínas amostrais (13 no total) no *Software Peptide Cutter*. Comparamos os nossos resultados experimentais (peptídeos identificados no método) com o resultado do *Software Peptide Cutter* (peptídeos virtuais) e selecionamos os peptídeos não identificados no nosso método. Desta forma, nós dividimos em dois diferentes grupos de peptídeos: *i*) peptídeos de cadeia muito curta, cuja média de *Gravy* foi de $-2,00$ (baixa hidrofobicidade) e massa molecular $< 445,00$ e *ii*) peptídeos de cadeia longa com média de *Gravy* de $-0,284$ (alta hidrofobicidade) e massa molecular $> 4.000,00$ (**Anexo I**).

Os peptídeos de cadeia curta (1 a 3 aminoácidos na sequência) possuem menor relevância no aumento da cobertura da sequência das proteínas e normalmente são muito fáceis de serem perdidos no método. Na maioria das vezes, tais peptídeos não são computados em análises proteômicas. Em contrapartida, é evidente que os peptídeos de cadeia longa (34 a 56 aminoácidos na sequência) contribuem significativamente com o aumento da cobertura das proteínas. Por meio destas considerações, nós desenvolvemos o segundo modelo químico-matemático para a mistura de proteínas (**Figura 19**), na qual a região hipotética foi visualizada por meio das interpretações de hidrofobicidade e tamanho das biomoléculas.

Adicionalmente, assumindo que os peptídeos do modelo da **Figura 19 região B** possuem maior grau de hidrofobicidade em relação ao modelo da **Figura 16 região B** (em termos gerais isto se deve aos peptídeos de cadeia longa). Desta forma, para a detecção destas espécies, espera-

se que a pressão do gás nebulizador seja relativamente elevada, uma vez que as propriedades hidrofóbicas da molécula induzem a uma condição ótima na eletronebulização, provavelmente o valor ótimo será por volta de 1,8 bar. A **Equação 11** trata-se da função matemática para a geração do modelo químico-matemático da **Figura 19**, enquanto que a **Figura 20** trata-se dos comandos no MATLAB

Por meio da literatura científica, comprovou-se os pressupostos elucidados em termos das interações hidrofóbicas não específicas em relação à coluna SCX do modelo químico-matemático da **Figura 16**, assim como da **Figura 19**. Com por exemplo, Mant, C.T. e Hodges, R.S demonstraram que o uso de 10% de ACN no tampão de eluição em duas colunas SCX diferentes melhorou significativamente a forma do pico cromatográfico. Provavelmente, o solvente orgânico torna a interação eletrostática mais pura, de modo que a interação hidrofóbica não específica é diminuída. [51] Peter Stanton menciona que o uso de 10 a 20% de ACN é necessário para limitar as interações hidrofóbicas da fase estacionária da trocadora catiônica. [54] Alpert, J.A., especialista em colunas SCX e inventor da fase estacionária poli(2-sulfoetil aspartamida)-sílica, observou uma boa separação com o uso de 25% de ACN no tampão da fase móvel. [55]

Figura 19 - Modelo químico-matemático para a mistura proteolítica de sete proteínas para o método SCX×RP-MS/MS. Região Hidrofílica: peptídeos com baixo grau de interações hidrofóbicas não específicas. Região Hidrofóbica: área hipotética para os peptídeos com alto grau de interações hidrofóbicas não específicas e elevada massa molecular.

Fonte: Autoria própria.

$$Z = \frac{1,8}{(1,75 + (x + 1,5)^2 + y^2)} + \frac{1,1}{(1,75 + (x - 2,5)^2 + (y - 1,5)^2)} \quad \text{Equação 11}$$

Figura 20 - Comandos do MATLAB para a geração do modelo químico-matemático, referente à Figura 19.

```
>> x=[-3:3];
>> y=x;
>> x=linspace(-3,3,100); y=x;
>> [x,y]=meshgrid(x,y);
>> z=1.8./(1.75+(x+1.5).^2+y.^2)+1.1./(1.75+(x-2.5).^2+(y-1.5).^2);
>> mesh(x,y,z);
>> surf(x,y,z);
```

Fonte: Autoria própria.

É importante reconhecer que a pressão do gás nebulizador composta por N₂ irá depender da escala cromatográfica da segunda dimensão. Assim escalas cromatográficas maiores devem ser avaliadas. De forma análoga, na cromatografia de nanofluxo não é necessário a utilização de N₂ para a eletronebulização. Pois a dimensão e vazão da fase móvel reduzida (colunas com diâmetro interno de 75 µm), favorece a repulsão eletrostática dos peptídeos para a formação do aerossol.

2.4.7. Robustez sobre a hidrofobicidade dos peptídeos nos modelos químico-matemáticos

Com o intuito de verificar a robustez dos modelos químico-matemáticos, nós comparamos a média de hidrofobicidade dos peptídeos identificados na região de máxima resposta da **Figura 16** (modelo BSA) com a região de máxima resposta da **Figura 19** (modelo Mix). Ambas as regiões se referem a 8% de ACN na fase móvel do tampão de eluição da 1D e 1,2 bar para o gás nebulizador no ESI-MS. Verificou-se que a média de hidrofobicidade das regiões foram próximas. Assim nos modelos BSA e Mix a média de *Gravy* foi de -0,421 e -0,405, respectivamente (**Anexo II**).

Adicionalmente, nós plotamos gráficos de distribuição gaussiana para ambas as regiões de máxima resposta. A **Figura 21** mostra que a distribuição gaussiana de ambas é deslocada para a região de caráter hidrofílico. Portanto, a previsão teórica em termos da média de hidrofobicidade dos peptídeos mantém uma ligação estreita com o comportamento de interação hidrofóbica não-específica da coluna SCX, em termos de % ACN no tampão para eluição da 1D.

A comparação da distribuição gaussiana entre os modelos BSA e Mix da região hidrofóbica torna-se limitada e verificamos que não gerou informação relevante, por dois motivos: *i*) tal região é hipotética no modelo Mix e *ii*) esta região de máximo contém somente 6 peptídeos (estes de cadeia muito longa > 33 resíduos de aminoácidos). Contudo, conforme anteriormente descrito temos a média de *Gravy* (hidrofobicidade) destas regiões e ambas se mostraram que tais peptídeos possuem elevada hidrofobicidade. Desta forma, um modelo químico-matemático proveniente de um proteoma real poderia comprovar as hipóteses até aqui constatadas.

Figura 21 - Distribuição gaussiana dos peptídeos identificados na região hidrofílica do método SCX×RP-MS/MS, o qual os parâmetros instrumentais significativos possuem valores de 8% ACN no tampão eluente da 1D e 1,2 bar da pressão do nebulizador. A) Peptídeos provenientes do Modelo BSA; B) Peptídeos provenientes do Modelo Mix. Quanto mais distante a gaussiana é do eixo zero, mais hidrofílico é o conjunto dos peptídeos.

Fonte: Autoria própria.

2.4.8. Superfície de resposta para um proteoma de levedura (*Saccharomyces cerevisiae*)

A teoria do modelo químico-matemático foi avaliada por meio da aplicação de uma amostra de maior complexidade, o digerido trípico do proteoma de *Saccharomyces cerevisiae*. Este foi escolhido, pois trata-se da amostra ouro para otimização de métodos analíticos em

proteômica, e adicionalmente, permite comparação com a literatura. Assim, os peptídeos foram separados na *Multidimensional Protein Identification Technology* (MudPIT) acoplado ao cromatógrafo líquido de nanofluxo *EASY-nLC*, seguido pela separação e detecção por espectrometria de massas sequencial *LTQ-Orbitrap Velos*. A interpretação dos espectros de massas, foi realizada pelo algoritmo computacional *ProLuCID*, a qual é o principal componente do software *Integrated Proteomics Pipeline* (IP2). IP2 é uma plataforma operacional desenvolvida no *Yates Lab* o qual permite estudos abrangentes dos proteomas. [14]

Como MudPIT possui uma estratégia de separação semelhante ao SCX×RP-Q-ToF-MS/MS (método dos modelos matemáticos anteriores). Nós assumimos que a mesma condição da 1D cromatográfica seria um dos fatores significativos para controlar a hidrofobicidade dos peptídeos e permitir o aumento da cobertura do proteoma. Em suma, a percentagem de acetonitrila no tampão eluente da 1D. Portanto, seria necessário encontrar o parâmetro significativo na etapa de ionização, uma vez que nós utilizamos um espectrômetro de massas diferente (*LTQ-Orbitrap Velos*). Desta forma, planejou-se teste rápido, simples e eficiente para a verificação dos possíveis fatores e intervalos dos parâmetros instrumentais na etapa de ionização do espectrômetro de massas, uma vez que nos modelos anteriores a etapa de ionização foi um fator significativo. Para atingir este pressuposto, nós aplicamos um planejamento experimental (*DoE*) em método unidimensional: coluna com partículas de C_{18} , *Thermo Scientific PepMap®* (75 μm d.i. $\times$ 15 cm, C_{18} , 3 μm , 100 Å), acoplado ao cromatógrafo líquido de nanofluxo hifenizado ao espectrômetro de massas sequencial *LTQ-Orbitrap Velos*. A **Tabela 8** mostra os parâmetros avaliados e o raciocínio da investigação.

Tabela 8 - Fatores e intervalos independentes para otimização da nanoLC(coluna C_{18})-MS/MS

#	Dimensão	Fatores	Níveis		Raciocínio
			Baixo (-1)	Alto (+1)	
1	LC	Distância da ponta da coluna no MS (mm)	2,0	4,0	Pode influenciar a formação do spray e a entrada do número de íons em fase gasosa no orifício do MS
2	MS	Tensão elétrica da eletronebulização (kV)	2,0	2,5	Melhora a formação do aerossol
3	LC	Vazão na coluna (nL min^{-1})	300	400	Pode influenciar a formação do spray, sensibilidade e tamanho da gota para ionização
4	MS	Temperatura do capilar ($^{\circ}\text{C}$)	225	325	Condições aprimoradas para o analito obter uma dessolvatação eficaz

Utilizou-se do planejamento fatorial completo 2^4 com 16 experimentos e 1 réplica. Foram avaliadas duas respostas, número de proteínas e peptídeos identificados (**Tabela 9**).

Tabela 9 - Níveis de fatores codificados e respostas de saída da matriz, para encontrar os fatores significativos na etapa de ionização do espectrômetro de massas.

Fatores		Níveis	
		(-1)	(+1)
1	Distância da ponta da coluna do MS (mm)	2,0	4,0
2	Tensão elétrica da eletronebulização (kV)	2,0	2,5
3	Vazão na coluna C ₁₈ (nL min ⁻¹)	300	400
4	Temperatura do capilar (°C)	225	325

	1	2	3	4	Valores Absolutos		<i>d</i>		
	Distância	ESI	Vazão	Capilar	Proteínas IDs	Peptídeos IDs	Proteínas IDs	Peptídeos IDs	<i>D</i>
1	1	1	1	1	714	7927	0,85	0,87	0,86
2	1	1	1	-1	661	6410	0,31	0,15	0,22
3	1	1	-1	1	697	7161	0,68	0,51	0,58
4	1	1	-1	-1	661	6349	0,31	0,12	0,19
4B	1	1	-1	-1	650	6286	0,20	0,09	0,13
5	1	-1	1	1	631	6723	0,01	0,30	0,05
6	1	-1	1	-1	662	6751	0,32	0,31	0,32
7	1	-1	-1	1	729	8201	1,00	1,00	1,00
8	1	-1	-1	-1	630	6099	0,00	0,00	0,00
9	-1	1	1	1	668	7105	0,38	0,48	0,43
10	-1	1	1	-1	649	6482	0,19	0,18	0,19
11	-1	1	-1	1	711	7329	0,82	0,59	0,69
12	-1	1	-1	-1	656	6459	0,26	0,17	0,21
13	-1	-1	1	1	723	7410	0,94	0,62	0,77
14	-1	-1	1	-1	668	6408	0,38	0,15	0,24
15	-1	-1	-1	1	691	7001	0,62	0,43	0,51
16	-1	-1	-1	-1	680	6742	0,51	0,31	0,39
<i>Máximo (A)</i>					729	8201			
<i>Mínimo (L)</i>					630	6099			

IDs = Número de identificações

Como o sistema foi avaliado por múltiplas respostas, foi necessário aplicar a função de desejabilidade descrita por *Derringer R. e Suich, R.* (**Equação 2**) seguida pela média geométrica, indicada pela **Equação 4 (Capítulo 1)**. Acreditamos que neste experimento ambas as respostas possuem mesmo peso para o modelo, devido a elevada possibilidade de co-eluição dos peptídeos (cromatografia unidimensional) e a identificação relativa das proteínas. Desta forma, para o valor de t nós assumimos como 1.

Notou-se uma boa distribuição do sinal m/z ao longo do cromatograma de íons unidimensional, o qual faz referência a maior desejabilidade global ($D = 0,65$) da **Tabela 9** e ensaio 7. Esta característica permite maiores capacidades de pico e consequente maior número de identificação de peptídeos por espectrometria de massas. Adicionalmente, a réplica 4B da **Tabela 9** mostrou boa reprodutibilidade nas respostas.

O gráfico de Pareto revela o peso da contribuição de cada parâmetro instrumental (fatores) sobre a resposta, assim como, os efeitos das interações entre estas variáveis. Dentre os fatores avaliados da **Tabela 9**, apenas a temperatura do capilar mostrou-se como efeito significativo e positivo. O gráfico de Pareto, **Figura 22** indica que a melhor condição é aumentar a temperatura do capilar no MS. Nós verificamos que todas as interações não geraram confusão. Portanto, todas as contribuições das interações podem ser consideradas neste *DoE*, uma vez que não há redundância (mesmo valores na resposta da matriz *DoE*, medida não mostrada). Portanto, a temperatura do capilar pode ser considerada como um fator significativo no aumento de identificações de espécies em proteômica *shotgun*, com 95% do intervalo de confiança. [56, 57]

O grupo do Professor Muddiman possui resultados que suportam nossas conclusões. O qual faz referência ao aumento da cobertura de um dado proteoma, utilizando a temperatura do capilar como o fator mais significativo em toda a plataforma espectrometria de massas. Eles avaliaram 2 *DoE*'s na otimização de 9 parâmetros instrumentais no *LTQ-Orbitrap XL*. O segundo *DoE*, o qual considerou a temperatura do capilar na otimização, permitiu maior cobertura do proteoma trípico de *S. cerevisiae*. Enquanto que o primeiro *DoE*, o qual não considerou a temperatura do capilar, a cobertura do proteoma foi menor. Assim, o melhor resultado na otimização mostrou um total de 570 proteínas e 3.932 peptídeos únicos, proporcionando cerca de 20% de melhora na sequência dos peptídeos e identificações das proteínas. [58]

Em um segundo estudo, o grupo avaliou 2 *DoE*'s na otimização de 14 parâmetros instrumentais, cujo a plataforma foi um espectrômetro de massas híbrido tipo quadrupolo acoplado ao Orbitrap. A resposta avaliada foi o número de identificação de peptídeos únicos. Os dois *DoE*'s foram realizados sob os mesmos 14 fatores. No entanto, o valor dos intervalos dos fatores foi diferente entre os *DoE*'s. Assim, dentre os 14 fatores avaliados, apenas a temperatura do capilar teve um efeito significativo para aumento da cobertura proteômica em ambos os *DoE*'s. Os resultados mostram que a escolha dos valores dos intervalos influencia, se os mesmos serão estatisticamente significativos ou não. [59]

Figura 22 - Avaliação da contribuição dos efeitos do *DoE*. Os efeitos principais padronizados do gráfico de Pareto mostram somente um fator significativo: temperatura do capilar. A linha vertical vermelha do gráfico define o nível de confiança de 95%. Os números das coordenada y referem-se aos fatores codificados da **Tabela 98** e **Tabela 109**.

Fonte: Autoria própria.

Por meio dos dois fatores significativos é possível plotar um modelo para prever o aumento da cobertura proteômica. Assim, o planejamento composto central (CCD) foi o modelo matemático escolhido, pois permite maior área de superfície com menor número de experimentos. A **Tabela 10** indica os fatores e níveis avaliados no CCD por meio do *software*

Statística. Os valores axiais ou ponto estrela selecionados como $\alpha = \pm 2$, em vez de $\alpha = \pm 1,41$, foram escolhidos para proporcionar a maior área da superfície de resposta, de modo a identificar adequadamente as duas regiões de máxima resposta do modelo de sela (o qual foi hipotetizado a partir das **Figura 16** e **Figura 19**). Esta estratégia pode permitir a obtenção de valores menores da variância dos coeficientes da equação polinomial do modelo, sendo esta a propriedade de rotatibilidade descrita por Box e Hunt. [60,61] É importante mencionar que a escolha de $\pm\alpha$ pode ser realizada por sua conveniência e viabilidade, conforme reportado pelos autores e revisada por Ferreira *et al.* [60, 61, 62]

Tabela 10 - Fatores avaliados no planejamento composto central de um projeto com 2^2 e $\alpha = \pm 2$.

Cód.	Dimensão	Fatores	$-\alpha$	-1	0	$+1$	$+\alpha$
1	SCX	%ACN no tampão	5	9	13	17	21
2	MS/MS	Temperatura do capilar (°C)	275	300	325	350	375
Valores absolutos					di (desejabilidade)		
#	% ACN no tampão	Temperatura do capilar (°C)	Proteínas ids	Peptídeos ids	Proteínas ids	Peptídeos ids	D
1	-1	-1	1228	14707	0,22	0,02	0,07
2	1	-1	1029	10763	0,07	0,00	0,01
3	-1	1	2229	42532	1,00	1,00	1,00
4	1	1	1380	16459	0,34	0,04	0,11
5C	-2	0	1471	14586	0,41	0,02	0,09
6B	2	0	938	10109	0,00	0,00	0,00
7	0	-2	2181	27000	0,96	0,27	0,51
8	0	2	2138	37585	0,93	0,72	0,82
9st	0	0	2170	37528	0,95	0,72	0,83
10	0	0	1731	35380	0,61	0,61	0,61
12D	0	0	2136	28142	0,93	0,31	0,54
13D	0	0	2182	30737	0,96	0,40	0,62
Máximo (A)			2229	42532			
Mínimo (L)			938	10109			

IDs = Número de identificações

As respostas avaliadas foram o número de identificações de proteínas e peptídeos. Utilizamos novamente a função desejabilidade **Equação 2** e **Equação 4**. Neste caso, $t = 1$ para os

Figura 23 - Cromatogramas de íons provenientes do método MudPIT-LTQ-Orbitrap-MS/MS. A) Ortogonalidade de separação para a modalidade SCX-RP(C₁₈). Cada cor dos cromatogramas de íons indica a separação na coluna C₁₈, após a etapa de eluição do tampão salino da fase estacionária SCX. Os valores percentuais indicam a quantidade de carregamento de acetato de amônio (tampão salino) na fase estacionária SCX, primeira dimensão cromatográfica. B) Gráfico 3D a partir da separação da coluna RP-C₁₈ versus a análise por espectrometria de massas. O intervalo de massa selecionado foi de 300-2000 *m/z*. Os cromatogramas de íons referem-se a uma etapa de eluição da fase estacionária SCX (30% de acetato de amônio). Condição: coluna C₁₈ 15 cm × 100 µm d.i.; 5 µm; 100 Å; fase estacionária SCX 2,5 cm × 250 µm d. i. 5 µm; temperatura ambiente; gradiente linear de 15 para 30% acetonitrila em 0,1% ácido fórmico aquoso em 93 minutos; vazão de 250 nL min⁻¹; concentração de amostra 20 µg na coluna.

Fonte: Autoria própria.

ids de proteínas e $t = 2$ para os *ids* de peptídeos. Como o método em questão possui poder de separação maior, espera-se que os *ids* de peptídeos sejam um valor numérico mais preciso em relação aos *ids* de proteínas. Além disso, as identificações de proteínas na proteômica *shotgun* possuem precisão relativamente baixa, pois trata-se de uma técnica relativa de identificação. Verificamos que tal consideração não altera significativamente o resultado final, porém torna o modelo mais exato. A **Figura 23** mostra o poder da ortogonalidade de separação entre dimensões do MudPIT-MS/MS. Sabe-se que o elevado grau de ortogonalidade aumenta a cobertura proteômica. Por isso, sua avaliação é muito importante. A **Figura 23** mostra a superfície de resposta em função dos dois fatores significativos.

O gráfico de Pareto da **Figura 24A** indica que ambos os fatores, assim como a função linear (L), quadrática (Q) e a interação de segunda ordem (1L*2L), são significantes para o aumento da cobertura das proteínas. Verifica-se dois máximos de resposta na superfície de resposta da **Figura 24B** e **Figura 24C**. Portanto, temos a primeira evidência que a teoria do *modelo químico-matemático* funciona para proteomas. Uma vez que os nossos estudos prévios mostram dois máximos de resposta. Para tornar nossos resultados mais plausíveis, nós estendemos a superfície de resposta do CCD para dois pontos adicionais, os quais podem representar os melhores valores instrumentais para o aumento de cobertura do proteoma referente as **Figura 24B** e **Figura 24C**, conforme reportado na **Tabela 11**.

A **Figura 25** e **Equação 12**, indica a superfície de resposta com os experimentos adicionais para a mais abrangente representação da perda de peptídeos do método, ou as condições de máxima resposta. A **Figura 26** descreve a linguagem computacional para plotar o modelo químico-matemático (MQM) do proteoma de *S. cerevisiae*. Como o CCD evidencia duas máximas respostas, com os experimentos adicionais da **Tabela 11**, a superfície de resposta estendida mostra que houve um decréscimo das identificações de peptídeos e proteínas. Assim, nós podemos inferir que o intervalo dos fatores em CCD estão adequados e MQM possui ótima predição da cobertura máxima do proteoma.

Figura 24 - Planejamento composto central para o método MudPIT-LTQ-Orbitrap-MS/MS A) Efeitos principais padronizados para o gráfico de Pareto do digerido trípico de levedura. Linha vertical do gráfico define 95% do nível de confiança. B) Superfície de resposta como função de dois fatores significativos para o aumento da cobertura de proteínas do digerido trípico de levedura. Os números da **Figura A** referem-se aos fatores codificados da **Tabela 10**. As letras L e Q da coordenada y significam função linear e quadrática da superfície de resposta, respectivamente. C) Gráfico 2D para a superfície de resposta da **Figura B**. D) Equação para plotar a superfície de resposta.

Fonte: Autoria própria.

Tabela 11 - Ensaios adicionais para avaliação da região máxima de detecção dos peptídeos no planejamento composto central.

#	% ACN no tampão	Temperatura do capilar (°C)	Valores absolutos		di		
			Proteínas ids	Peptídeos ids	Proteínas ids	Peptídeos ids	D
MX1	-2 (5)	+2 (375)	1724	21322	0,61	0,35	0,46
MX2	+1 (17)	-3 (250)	1840	24056	0,70	0,43	0,55

Nós sequenciamos em conjunto no *software* IP2 os dois experimentos que se referem as duas máximas regiões do *modelo químico-matemático* da **Figura 25**. Os valores obtidos foram 2.711 proteínas e 65.511 peptídeos, para 20µg de amostra injetada na coluna. Webb, K. J. *et al.* identificaram 4.488 proteínas e 44.154 peptídeos únicos de 40µg de amostra injetada na coluna, por meio do somatório de três métodos cromatográfico de um gradiente modificado; cujo o método foi o MudPIT acoplado ao espectrômetro de massas *Velos Pro Orbitrap Elite*. [63] Uma maior quantidade de amostra injetada na coluna, permite maior identificação, se não ocorrer saturação na detectabilidade do MS. Richards, A.L. *et al.* identificaram 4.002 proteínas e 34.255 peptídeos únicos de 1,4µg de amostra, utilizando coluna capilar de 30cm de comprimento, com partículas de diâmetro de 1,7µm acoplado ao espectrômetro de massas de maior sensibilidade o *LTQ Orbitrap Fusion*. [37] Para ambos os estudos, a amostra padrão foi o digerido trípico de levedura (*Sacchromyces cerevisiae*) e utilização de um FDR (*false discovery rate*) de 1%. As quais, em suma, são as mesmas condições do nosso método bioanalítico.

Figura 25 - Modelo químico-matemático para maximizar a cobertura do proteoma de *S. cerevisiae*. A) Gráfico 3D e B) Gráfico 2D.

Fonte: Autoria própria.

$$Z = \frac{1.7}{(1.75 + (x + 1.0)^2 + (y - 1)^2)} + \frac{1.3}{(1.75 + (x - 0.1)^2 + (y + 1.41)^2)} \quad \text{Equação 12}$$

Figura 26 - Linguagem do MATLAB para plotar o modelo químico-matemático da **Figura 25**.

```
>> x=[-2:2];  
  
>> y=x;  
  
>> x=linspace(-2,2,100); y=x;  
  
>> [x,y]=meshgrid(x,y);  
  
>> z=1.7./(1.75+(x+1.0).^2+(y-1).^2)+1.3./(1.75+(x-0.1).^2+(y+1.41).^2);  
  
>> mesh(x,y,z);  
  
>> surf(x,y,z);
```

Fonte: Autoria própria.

2.5. CONCLUSÕES

Com os resultados e interpretações obtidas até o momento, pode-se afirmar que o *modelo químico-matemático* para maximizar a cobertura da sequência das proteínas, para as técnicas proteômicas *shotgun e bottom-up*, funciona, uma vez que foram obtidas duas máximas respostas do modelo de proteoma de leveduras. Este fato está em comum acordo com as hipóteses e dados experimentais dos modelos de padrões de proteínas digeridas. Assim, cada máximo de resposta, pode ser relacionado com o grau de hidrofobicidade dos peptídeos. Para tal finalidade, ajustes no algoritmo do *software* IP2 estão em processo de desenvolvimento. O modelo pode permitir, de maneira simples, uma maior compreensão dos proteomas, rápida otimização no desenvolvimento de métodos analíticos, padronização de métodos entre laboratoriais e melhor caracterização das expressões de isoformas de proteínas, contribuindo assim, com os avanços nas figuras analíticas de mérito em proteômica e, sobretudo, à teoria proteômica multidimensional. É interessante mencionar que a possibilidade de simplificar a otimização de técnica analítica complexa, em simples modelos com informações amostrais tridimensionais (químicas, físicas e instrumentais) reflete no fenômeno matemático observado pelo princípio de Pareto, ou seja, 80% das consequências advém de 20% das causas.

A aplicação de maior concentração de amostra no método poderia permitir a identificação de todo o proteoma, entretanto a possibilidade de entupimento da coluna capilar aumentaria. Perante o nosso melhor conhecimento, o modelo-químico matemático permitiu a identificação recorde de números de peptídeos (> 65k) por meio de quantidade de amostra reduzida (20µg), proveniente do proteoma de *Sacchromyces cerevisiae*.

Adicionalmente, sabe-se que uma das limitações da proteômica *top-down* é a baixa capacidade de identificação de espécies nos métodos. Desta forma, acreditamos que a lógica do *m-DoE* (**Capítulo 1**) em conjunto com o *modelo químico-matemático* poderiam permitir maior caracterização do proteoma provenientes da técnica *top-down*.

2.6. REFERÊNCIAS

- 1 NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger**. Porto Alegre: artmed. 7. ed., 2018, p. 89-182.
- 2 PAULING, L. How my interest in proteins developed. **Protein Science**, Hoboken, v. 2, n. 6, p. 1060-1063, 1993
- 3 PANKOW, S.; BAMBERGER, C.; YATES, J. R. A posttranslational modification code for CFTR maturation is altered in cystic fibrosis. **Science Signaling**, Washington, v. 12, n. 562, 2019.
- 4 SALERMO, M.S.; MATSUMOTO, C.; FERRAZ, M. S. Biofármacos no Brasil: características, importância e delineamento de políticas públicas para seu desenvolvimento. 2018 Brasília: Rio de Janeiro: **Ipea Instituto de Pesquisa Econômica Aplicada. Ministério do Planejamento, Desenvolvimento e Gestão**. Disponível em: <<http://www.ipea.gov.br/portal/publicacoes>> Acesso em: 13 abril 2020.
- 5 LARANCE, M.; LAMOND, A. I. Multidimensional proteomics for cell biology. **Nature Reviews Molecular Cell Biology**, v. 16, n. 5, p. 269-280, 2015.
- 6 ZHANG, Y.; FONSLow, B. R.; SHAN, B.; BAEK, M. C.; YATES, J. R. Protein analysis by shotgun/bottom-up proteomics. **Chemical Reviews**, Washington, v. 113, n. 4, p. 2343-2394, 2013.
- 7 YATES, J. R.; RUSE, C. I.; NAKORCHEVSKY, A. Proteomics by mass spectrometry: approaches, advances, and applications. **Annual Reviews of Biomedical Engineering**, Danbury, v. 11, p. 49-79, 2009.
- 8 BREUKER, K.; MCLAFFERTY, F. W. Stepwise evolution of protein native structure with electrospray into the gas phase, 10^{-12} to 10^2 s. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 105, n. 47, p. 18145-18152, 2008.
- 9 VAN DEN HEUVEL, R. H. H.; HECK, A. J. R. Native protein mass spectrometry: from intact oligomers to functional machineries. **Current Opinion in Chemical Biology**, Oxford, v. 8, n. 5, p. 519-526, 2004.
- 10 TOBY, T. K.; FORNELLI, L.; KELLEHER, N. L. Progress in top-down proteomics and the analysis of proteoforms. **Annual Review of Analytical Chemistry**, Palo Alto, v. 9, n. 1, p. 499-519, 2016.
- 11 CHAIT, B. T. Mass Spectrometry: bottom-up or top-down? **Science**, Washington, v. 314, n.

5796, p. 65-66, 2006.

12 HAN, X.; ASLANIAN, A.; YATES, J. R. Mass spectrometry for proteomics. **Current Opinion in Chemical Biology**, Oxford, v. 12, n.5, p. 483-490, 2008.

13 FONSLow, B. R.; CARVALHO, P. C.; ACADEMIA, K.; FREEBY, S.; XU, T.; NAKORCHEVSKY, A.; PAULUS, A.; YATES, J. R. Improvements in proteomic metrics of low abundance proteins through proteome equalization using ProteoMiner prior to MudPIT. **Journal of Proteome Research**, Washington, v. 10, n. 8, p. 3690-3700, 2011.

14 XU, T.; PARK, S. K.; VENABLE, J. D.; WOHLSCHLEGEL, J. A.; DIEDRICH, J. K.; COCIORVA, D.; LU, B.; LIAO, L.; HEWEL, J.; HAN, X.; WONG, C. C. L. FONSLow, B.; DELAHUNTY, C.; GAO, Y.; SHAH; YATES, J. R. ProLuCID: An improved SEQUEST-Like algorithm with enhanced sensitivity and specificity. **Journal of Proteomics**, Amsterdam, v. 129, n. 3, p. 16-24, 2015.

15 ZHANG, Z.; WU, S.; STENOIEN, D. L.; PAŠA-TOLIĆ, L. High-throughput proteomics. **Annual Review of Analytical Chemistry**, Palo Alto, v. 7, n. 1, p. 427-454, 2014.

16 SHEN, Y.; ZHAO, R.; BERGER, S. J.; ANDERSON, G. A.; RODRIGUEZ, N.; SMITH, R. D. High-efficiency nanoscale liquid chromatography coupled on-line with mass spectrometry using nanoelectrospray ionization for proteomics. **Analytical Chemistry**, Washington, v. 74, n. 16, p. 4235-4249, 2002.

17 MICHALSKI, A.; COX, J.; MANN, M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. **Journal of Proteome Research**, Washington, v. 10, n. 4, p. 1785-1793, 2011.

18 MCLAFFERTY, F. W.; LORY, E. R. Mass spectrometry as a separation technique analysis of involatile samples. **Journal of Chromatography A**, v. 203, n. 9. p. 109-116, 1981.

19 PERRY, R. H.; COOKS, R. G.; NOLL, R. J. Orbitrap mass spectrometry: instrumentation, ion motion and applications. **Mass Spectrometry Reviews**, Hoboken, v. 27, p. 661-699, 2015.

20 VALEJA, S. G.; XIU, L.; GREGORICH, Z. R.; GUNER, H.; JIN, S.; GE, Y. Three-dimensional liquid chromatography coupling ion exchange chromatography/hydrophobic interaction chromatography/reverse phase chromatography for effective protein separation in top-down proteomics. **Analytical Chemistry**, Washington, v. 87, n. 10, p. 5363-5371, 2015.

21 STOLL, D. R.; CARR, P. W. Two-dimensional liquid chromatography: a state-of-the-art tutorial. **Analytical Chemistry**, Washington, v. 89, n. 1, p. 519-531, 2017.

22 STOLL, D. R.; LI, X.; WANG, X.; CARR, P. W.; PORTER, S. E. G.; RUTAN, S. C. Fast, comprehensive two-dimensional liquid chromatography. **Journal of Chromatography A**,

Amsterdam, v. 1168, n. 1-2, p. 3-43, 2007.

23 CASSIANO, N. M.; BARREIRO, J. C.; OLIVEIRA, R. V.; CASS, Q. B. Direct bioanalytical sample injection with 2D LC-MS. **Bioanalysis**, London, v. 4, n. 22, p. 2737-2756, 2012.

24 FRANÇOIS, I.; SANDRA, K.; SANDRA, P. Comprehensive liquid chromatography: fundamental aspects and practical considerations - A review. **Analytica Chimica Acta**, Amsterdam, v. 641, n. 1-2, p. 14-31, 2009.

25 DOLAN, J. W.; SNYDER, L. R.; DJORDJEVIC, N. M.; HILL, D. W.; WAEGHE, T. J. Reversed-phase liquid chromatographic separation of complex samples by optimizing temperature and gradient time: I. peak capacity limitations. **Journal of Chromatography A**, Amsterdam, v. 857, n. 1-2, p. 1-20, 1999.

26 GILAR, M.; OLIVOVA, P.; DALY, A. E.; GEBLER, J. C. Orthogonality of separation in two-dimensional liquid chromatography. **Analytical Chemistry**, Washington, v. 77, n. 19, p. 6426-6434, 2005.

27 GIDDINGS, J. C. Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation. **Journal of Chromatography A**, Amsterdam, v. 703, n. 1-2, p. 3-15, 1995.

28 MOGOLLÓN, N. G. S.; DE LIMA, P. F.; GAMA, M. R.; FURLAN, M. F.; BRAGA, S. C. G. N.; PRATA, P. S.; JARDIM, I. C. S. F.; AUGUSTO, F. O estado da arte da cromatografia líquida bidimensional: conceitos fundamentais, instrumentação e aplicações. **Química Nova**, Sao Paulo, v. 37, n. 10, p. 1680-1691, 2014.

29 FENN, J. B.; MANN, M.; MENG, C. K.; WONG, S. F.; WHITEHOUSE, C. M.; Electrospray ionization of large for mass spectrometry biomolecules. **Science**, Washington, v. 246, n. 4926, p. 64-71, 1989.

30 TANAKA, K.; The origin of macromolecule ionization by laser irradiation (Nobel Lecture). **Angewandte Chemie International Edition**, Weinheim, v. 42, n. 33, p. 3860-3870, 2003.

31 TAKÁTS, Z.; WISEMAN, J. M.; GOLOGAN, B.; COOKS, R. G. Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. **Science**, Washington, v. 306, n. 5695, p. 471-473, 2004.

32 WILM, M.; MANN, M. Analytical properties of the nanoelectrospray ion source. **Analytical Chemistry**, Washington, v. 68, n. 1, p. 1-8, 1996.

33 SCHWARTZ, J. C.; SENKO, M. W.; SYKA, J. E. P. A two-dimensional quadrupole ion trap mass spectrometer. **Journal American Society Mass Spectrometry**, Washington, v. 13, p. 659-669, 2002.

- 34 MAKAROV, A. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. **Analytical Chemistry**, Washington, v. 72, n. 6, p. 1156-1162, 2000.
- 35 BRODBELT, J. S. Photodissociation mass spectrometry: new tools for characterization of biological molecules. **Chemical Society Reviews**, Cambridge, v. 43, n. 8, p. 2757-2783, 2014.
- 36 BATISTON, WELITON PEDRO. **Quimiometria aplicada à cromatografia líquida multidimensional capilar hifenizada a espectrometria de massas sequencial para proteômica shotgun**. 2015. Dissertação (Mestrado em Química Analítica) – Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2015.
- 37 RICHARDS, A. L.; HEBERT, A. S.; ULBRICH, A.; BAILEY, D. J.; COUGHLIN, E. E.; WESTPHALL, M. S.; COON, J. J. One-hour proteome analysis in yeast. **Nature Protocols**, London, v. 10, n. 5, p. 701-714, 2015.
- 38 MOTOYAMA, A.; XU, T.; RUSE, C. I.; WOHLSCHEGEL, J. A.; YATES, J. R. Anion and cation mixed-bed ion exchange for enhanced multidimensional separations of peptides and phosphopeptides. **Analytical Chemistry**, Washington, v. 79, n. 10, p. 3623-3634, 2007.
- 39 HOUEL, S.; ABERNATHY, R.; RENGANATHAN, K.; MEYER-ARENDET, K.; AHN, N. G.; OLD, W. M. Quantifying the impact of chimera MS/MS spectra on peptide identification in large-scale proteomics studies. **Journal Proteome Research**, Washington, v. 9, p. 4152-4160, 2010.
- 40 THOMAS, K.; SWART, R.; MECHTLER, K.; Ultra-high-Pressure RPLC hyphenated reveals a linear relation between peak capacity and number of identified peptides. **Analytical Chemistry**, Washington, v. 83, n. 7, p. 2699-2704, 2011.
- 41 UNIVERSAL PROTEIN RESOURCE. **UniProt**. Hinxton: European Bioinformatics Institute. Disponível em: <<http://www.uniprot.org/uniprot/>> Acesso em: 20 jan. 2015.
- 42 FERREIRA, M. M. C. **Quimiometria: conceitos, métodos e aplicações**. Campinas: Unicamp, 2015. 493. p.
- 43 BARROS NETO, B; SCARMINIO, I. S.; BRUNS, R. E. **Como fazer experimentos: pesquisa e desenvolvimento na ciência e na indústria**. Porto Alegre: Bookman, 2010. 401 p.
- 44 PEREIRA FILHO, E. R. **Planejamento fatorial em química: maximizando a obtenção de resultados**. São Carlos: EdUFSCar, 2015. 88. p.
- 45 POHJOLAINEN, S. **Mathematical Modeling**. Suíça: springer. 2016, p. 1-237.
- 46 KRUBE, A.; HERODES, K.; LEITO, I. Optimization of electrospray interface and quadrupole ion trap mass spectrometer parameters in pesticide liquid

chromatography/electrospray ionization mass spectrometry analysis. **Rapid Communications in Mass Spectrometry**, Oxford, v. 24, n.7, p. 919-926, 2010.

47 KRUBE, A.; LEITO, I.; HERODES, K.; LAANISTE, A.; LÖHMUS, R. Enhanced nebulization efficiency of electrospray mass spectrometry: improved sensitivity and detection limit. **Journal of the American Society for Mass Spectrometry**, Washington, v. 23, n. 12, p. 2051-2054, 2012.

48 SCHMID, S.; JECKLIN, M. C.; ZENOBI, R. Electrosonic spray ionization — an ideal interface for high-flow liquid chromatography applications. **Journal of Chromatography A**, Amsterdam, v. 1218, n. 23, p. 3704-3710, 2011.

49 ISOBE, T.; TAKAYASU, T.; TAKAI, N.; OKUYAMA, T. High-performance liquid chromatography of peptides on a macroreticular resin: application to peptide mapping of Bence-Jones proteins. **Analytical Biochemistry**, Philadelphia, v. 425, n. 122, p. 417-425, 1982.

50 CACHIA, P. J.; VAN EYK, J.; CHONG, P. C. S.; TANEJA, A.; HODGES, R. S. Separation of basic peptides by cation-exchange high-performance liquid chromatography. **Journal of Chromatography**, Amsterdam, v. 266, p. 651-659, 1983.

51 BURKE, T. W.; MANT, C. T.; BLACK, J. A.; HODGES, R. S. Strong cation-exchange high-performance liquid chromatography of peptides effect of non-specific hydrophobic interactions and linearization of peptides retention behavior. **Journal of Chromatography**, Amsterdam, v. 416, n. 4, p. 377-389, 1989.

52 EXPASY BIOINFORMATICS RESOURCES PORTAL. **Protparam**. Lausanne: Swiss Institute of Bioinformatics. 2015. Disponível em: <<https://web.expasy.org/protparam/>>. Acesso em: 20 jan. 2015.

53 KYTE, J.; DOOLITTLE, R. F. A simple method for displaying the hydropathic character of a protein. **Journal of Molecular Biology**, London, v. 157, n. 1, p. 105-132, 1982.

54 AGUILAR, M. I. **HPLC of peptides and proteins methods and protocols**. Totowa: Human Press, 2004. 395 p. v. 251

55 ALPERT, A. J.; ANDREWS, P. C. Cation-exchange chromatography of peptides on poly(2-sulfoethyl aspartamide)-silica. **Journal of Chromatography**, Amsterdam, v. 443, p. 85-96, 1988.

56 BOX, G. E. P.; MEYER, R. D. Dispersion effects from fractional designs. **Technometrics**, Milwaukee, v. 28, n. 1, p. 19-27, 1986.

57 BOX, G. E. P.; MEYER, R. D. An analysis for unreplicated fractional factorials. **Technometrics**, Milwaukee, v. 28, n.1, p. 11-18, 1986.

- 58 ANDREWS, G. L.; DEAN, R. A.; HAWKRIDGE, A. M.; MUDDIMAN, D. C. Improving proteome coverage on a LTQ-Orbitrap using design of experiments. **Journal of the American Society for Mass Spectrometry**, Washington, v. 22, n. 4, p. 773-783, 2011.
- 59 RANDALL, S. M.; CARDASIS, H. L.; MUDDIMAN, D. C. Factorial experimental designs elucidate significant variables affecting data acquisition on a quadrupole Orbitrap mass spectrometer. **Journal of the American Society for Mass Spectrometry**, Washington, v. 24, n. 10, p. 1501-1512, 2013.
- 60 BOX, G. E. P.; WILSON, K. B. On the experimental attainment of optimum conditions. **Journal of the royal statistical society. Series B (Methodological)**, London, v. 12, n. 1, p. 1-45, 1951.
- 61 BOX, G. E. P.; HUNTER, J. S. Multi-factor experimental designs for exploring response surfaces. **Annals of mathematical statistics**, Shaker Heights, v. 28, n. 1, p. 195-241, 1957.
- 62 FERREIRA, S. L. C.; BRUNS, R. E.; DA SILVA, E. G. P.; DOS SANTOS, W. N. L.; QUINTELLA, C. M.; DAVID, J. M.; DE ANDRADE, J. B.; BREITKREITZ, M. C.; JARDIM, I. C. S. F.; NETO, B. B. Statistical designs and response surface techniques for the optimization of chromatographic systems. **Journal of Chromatography A**, Amsterdam, v. 1158, p. 2-14, 2007.
- 63 WEBB, K. J.; XU, T.; PARK, S. K.; YATES, J. R. Modified MuDPIT separation identified 4488 proteins in a system-wide analysis of quiescence in yeast. **Journal of Proteome Research**, Washington, v. 12, n. 5, p. 2177-2184, 2013.

Capítulo 3

Microfluídica

Ramo da química analítica que estuda a dinâmica dos fluidos em sistemas analíticos miniaturizados. Faço a minha homenagem ao maior inventor multidisciplinar de todos os tempos:

“The natural desire of good men is knowledge.”

Leonardo da Vinci

3.1. INTRODUÇÃO

3.1.1. Uma breve história da evolução da cromatografia líquida

O raciocínio matemático é sem dúvida uma das atividades mais instigadoras dentre as ciências. Perante ela podemos prever fenômenos, muitas são as equações que permitiram esta afirmação. Em 1956, J.J. *van Deemter*; F.J. *Zuiderweg* e A. *Klinkenbertg* propuseram uma equação que direcionou o futuro da ciência de separação, a famosa equação de *van Deemter*, a qual prevê que partículas pequenas e uniformes em colunas cromatográficas reduzem significativamente dois termos da equação, a difusão de *Eddy* ou múltiplos caminhos (termo A_v) e resistência à transferência de massa (termo C), aumentando assim a eficiência de separação cromatográfica. [1,2] Em 1956, a tecnologia disponível para a confecção do tamanho das partículas em colunas cromatográficas era de até 100-200 μm , alcançando assim eficiências de separação de apenas 200 pratos (para colunas com 15 cm de comprimento). Contudo, inúmeros grupos de pesquisa se inspiraram nas previsões do “trio de 1956” até que, em 1996 foram introduzidas as colunas com partículas esféricas não porosas com tamanho de 1,5 μm , obtendo-se assim eficiência de até 30.000 pratos/15 cm de coluna, um aumento de 150 vezes em eficiência cromatográfica. [3]

Entretanto, colunas preenchidas com partículas minúsculas com comprimento usual de 10 a 15 cm geram pressões elevadas (aproximadamente 7.000 bar), aspecto que limitou a utilização de bombas convencionais na HPLC, as quais resistem à pressão de bombeamento em até 400 bar. Nesta perspectiva, os primeiros esforços com resultados significativos para análises cromatográficas em pressões elevadas, acima de 4.100 bar, foram desenvolvidas pelos grupos dos Professores J.W. *Jorgenson* e M. *Lee*, por volta de 1997. [3]

Como consequência, a cromatografia líquida de elevada pressão e eficiência (*ultra high pressure liquid chromatography* – UHPLC) tornou-se uma instrumentação sofisticada, pois grande parte do sistema deve suportar pressões elevadas, fato que aumentou o custo das análises, assim como os cuidados. Além disso, a sofisticação da síntese das partículas porosas de até 1 μm no recheio de colunas capilares indica que a equação de *van Deemter*, provavelmente tenha

chegado ao seu ápice de previsão, para a eficiência de separação, em termos de comportamentos microscópios de separações cromatográficas. [2,3]

Uma alternativa promissora aos problemas relatados da UHPLC são as colunas com fases estacionárias monolíticas. O monolito é um meio contínuo de fase estacionária para a separação, ou seja, pode-se mencionar que se trata de um material único ao longo da coluna cromatográfica. A morfologia deste material é composta por canais e domínios, os quais podem conter microporos ou mesoporos (estes às vezes nomeados como nanoporos, por possuir poros com tamanhos na escala nanométrica $< 1 \mu\text{m}$), assim como a mistura de microporos e mesoporos ou não conter poros. O material polimérico (monolito) possui uma estrutura sólida altamente porosa, garantindo alta permeabilidade da fase móvel sem comprometer a eficiência de separação cromatográfica. Como resultado, é possível análise cromatográfica rápida com elevada eficiência de separação, devido ao elevado grau de transferência de massa, em menor pressão e, portanto, sem a utilização de bombas cromatográficas sofisticadas. A permeabilidade do monolito é justificada pela presença dos canais, os quais podem atingir valores de até $2 \mu\text{m}$ de diâmetro. Desta forma, este conjunto de estrutura e geometria, garante menor resistência à vazão da fase móvel em relação às colunas empacotadas com partículas. [4]

3.1.2. Entendendo a química dos materiais porosos de fases monolíticas

Os monolitos são formados a partir de reações de polimerização, e há três tipos de materiais monolíticos atribuídos para separações cromatográficas: *i*) o monolito polimérico orgânico; *ii*) o monolito polimérico de sílica e *iii*) o monolito polimérico híbrido organo-sílica. Os monolitos orgânicos são geralmente polimerizados por reações radicalares e os monolitos à base de sílica, assim como os monolitos híbridos (organo-sílica) são geralmente polimerizados pelo processo sol-gel. Tais reações poliméricas podem ser iniciadas pelos processos térmico, foto, micro-ondas, feixe de elétrons ou γ -radiação. [4-7]

Apesar dos três monolitos apresentarem elevada permeabilidade, tais materiais possuem diferenças significativas na morfologia, conforme explicitado na **Figura 27**. Estudos relacionaram que H depende principalmente da estrutura morfológica (poros e macroporos) e a natureza da difusão dos analitos no meio cromatográfico. [5,8]

Figura 27 - Microscopia de varredura eletrônica evidenciando as diferenças da morfologia de monolitos poliméricos tipos orgânico e inorgânico (sílica). A) estrutura bimodal dos poros em monolito polimérico de sílica. B) estrutura reticulada microglobular de monolito polimérico orgânico.

Fonte: Adaptação de Batiston (2019) a partir de: NISCHANG, I. Porous polymer monoliths: Morphology, porous properties, polymer nanoscale gel structure and their impact on chromatography performance – A review. *Journal of Chromatography A*, Holanda, v.1287, 2013, p. 41.

3.1.3. Monolitos poliméricos de sílica ou inorgânico

Os monolitos à base de sílica possuem estrutura de poro bimodal (**Figura 27A**), com presença significativa de mesoporos (~13%), os quais possuem em média 72 nm de diâmetro. Alguns pesquisadores acreditam que tais características morfológicas, assim como a elevada área superficial, são as condições que justificam a elevada eficiência de separação para moléculas pequenas. Entretanto, tais monolitos, infelizmente não possuem boa eficiência de separação para macromoléculas. [5-8]

Os monolitos poliméricos de sílica demonstram boa resistência de expansão ou contração frente a diferentes solventes, elevada estabilidade mecânica, elevada área superficial (150 - 400

m² g⁻¹), podem ser modificados por ligações químicas a partir de vários grupos funcionais, permitindo elevada eficiência e seletividade de separação. Os principais inconvenientes são: maiores dificuldades na preparação, adsorção não específicas, causadas pelos grupos Si-OH (principalmente pelas interações hidrofílicas e eletrostáticas das proteínas), limitação à temperaturas superiores a 60 °C e possui instabilidade em pH básicos (< 8). [5-8]

Estes monolitos são preparados pelo processo de polimerização sol-gel, o qual envolve hidrólise e policondensação de alcoxissilano na presença de um catalisador e aditivos. Os alcoxissilanos empregados são geralmente o tetrametoxissilano e tetraetoxissilano e como aditivos são utilizados o polietilenoglicol (PEG) ou seus compostos análogos. [4,8]

Estudos reportam que os monolitos de sílica permitem separações rápidas com eficiência de separação superior a 100.000 pratos por metro, aspecto que justifica a atração das indústrias do ramo. Atualmente, esta tecnologia tem sido comercializada pela Merck (Darmstadt, Germany) como *Chromolith*® e pela Phenomenex (Torrance, CA, Estados Unidos) como Onyx™ de modo que a superfície do monolito é modificada com grupos funcionais apolares, tais como C₁₈ e C₈. [5]

Recentemente, os monolitos de sílica demonstraram aplicações bem sucedidas na separação de peptídeos devido a elevada eficiência de separação, por meio da modalidade de cromatografia em fase reversa. [5] Vaast *et al.* compararam a eficiência de separação de uma coluna capilar com monolito polimérico de sílica, em relação à duas colunas capilares cromatográficas comercialmente disponíveis, com partículas porosas de 3 µm e partículas de sílica *fused-core*. O desempenho da coluna monolítica de sílica foi superior em relação às duas colunas empacotadas com partículas. Os autores justificam os resultados devido a melhor transferência de massa e permeabilidade da coluna monolítica. [9]

3.1.4. Monolitos poliméricos orgânicos

Os monolitos poliméricos orgânicos possuem uma estrutura heterogênea, com morfologia reticulada de microglóbulos (semelhante a uma couve-flor) com poros de diâmetro menores em relação aos monolitos inorgânicos (**Figura 27B**). Os monolitos poliméricos são amplamente utilizados como fase estacionária em cromatografia líquida, possui fácil preparação na etapa de

polimerização dos monômeros, normalmente dissolvido em solvente orgânico (agente porogênico), o qual permite a formação dos poros, elevada estabilidade química em amplo intervalo de pH e facilidade na modificação, para a incorporação de grupos funcionais. Os principais inconvenientes são a baixa área superficial (10 a $150 \text{ m}^2 \text{ g}^{-1}$), inchaço e encolhimento em alguns solventes orgânicos e deficiência de estabilidade mecânica. [5,6,8]

Pesquisadores acreditam que tais características, como o elevado tamanho do domínio morfológico, em conjunto com a baixa área superficial, são os aspectos favoráveis para a separação de moléculas maiores, como proteínas e polímeros. Além disso, os monolitos orgânicos são biocompatíveis com as biomoléculas. [6,10] Assim, a falta de mesoporos e a presença de uma estrutura heterogênea, conduzida pela pequena área superficial, são as condições que justificam a menor eficiência de separação de moléculas pequenas, inclusive para peptídeos, quando comparado com os monolitos baseados em sílica. [6,11]

Estes monolitos são polimerizados por reações radiculares, geralmente pelo processo térmico, mas há trabalhos que reportam a polimerização pelo processo foto-iniciado. [5] Tradicionalmente, polímeros preparados por radicais livres enfrentam, na maioria dos casos, desafios como a baixa área superficial e estrutura descontrolada, aspectos que são desfavoráveis para a eficiência de separação cromatográfica, assim como a reprodutibilidade da característica morfológica dentre as colunas. [6] Entretanto, alguns grupos de pesquisa têm conseguido controlar de forma eficiente a reação radicalar, de modo a obter estruturas bem definidas na morfologia do monolito. [12]

A porosidade de um monolito orgânico depende do ajuste do agente porogênico e da quantidade do monômero de entrecruzamento, na mistura de reação, assim como da temperatura e tempo de polimerização. Arrua, R.D. *et al.* mencionam que os processos físico-químicos de partição e adsorção na separação cromatográfica irão depender do monômero utilizado na reação polimérica. [13]

Eeltink, S. *et al.* mencionaram que a eficiência de separação das fases monolíticas orgânicas depende do tamanho do domínio, representado pela média da distribuição de macroporos e microglóbulos. Outros estudos mostram que a composição da mistura de polimerização e condições de reação afetam fortemente a morfologia das fases monolíticas. [14] Hasegawa, G. *et al.* retrataram que a efetiva separação de macromoléculas está atribuída a

ausência de micro- e mesoporos nos microglóbulos do monolito, formados pelas reações radiculares, uma vez que, macromoléculas possuem baixa difusão na transferência de massa em microglóbulos porosos. [12]

Em especial, por meio dos resultados de Eeltink, S. *et al.*, aplicados para a separação de peptídeos, é possível constatar que a permeabilidade do monolito será resultado do conjunto entre: *i*) o tipo de monômeros, *ii*) o agente porogênico e *iii*) a composição de ambos no meio reacional. É interessante mencionar que o grupo não encontrou nenhuma tendência entre o aumento de solvente porogênico na etapa de polimerização e consequente aumento da permeabilidade, uma vez que o aumento do volume do solvente porogênico poderia fornecer o aumento da quantidade de poros no monolito e, desta forma, aumentar a permeabilidade do material. Contudo, diminuindo a quantidade de solvente porogênico, houve diminuição do esqueleto microglobular, fato que diminui a permeabilidade. Entretanto, eles encontraram resultados satisfatórios ao produzir um monolito com menores tamanhos de domínios e com elevada distribuição de microglóbulos, aumentando assim a área superficial, os quais foram justificados pelos termos A e C da equação de *van Deemter*. Eeltink, S. *et al.* comentam que para monolitos orgânicos, os monômeros derivados de metacrilato são bons para a separação de proteínas intactas, de modo que para peptídeos há problemas devido à fraca retenção e baixa resolução cromatográfica. [14]

Hasegawa, G. *et al.* encontraram uma condição ideal para a formação de macroporos bem definidos, com o controle da reação radicalar por meia da razão de uma mistura composta por estireno (S) e divinilbenzeno (DVB) iniciada com o radical livre 2,2'-azobis-(isobutironitrila) (AIBN). Eles deduziram que para obter macroporos bem definidos na separação de alquilbenzenos, deve-se utilizar uma razão S/DVB de 3 ou 4 e um pequeno decréscimo da quantidade de etil-2-metil-2-butilopropionato de etila. [12]

3.1.5. Monolitos poliméricos híbridos organo-sílica

Os monolitos híbridos apresentam as vantagens dos monolitos tipo orgânico e de sílica. Desta forma, esta classe de monolito possui fácil e rápida preparação, boa estabilidade mecânica, estabilidade de pH em elevado intervalo e área superficial elevada. Portanto, tais monolitos têm

atraído atenção de pesquisadores nos últimos anos e são considerados como o estado da arte em tecnologia de monolitos. [11,15]

Assim como o monolito de sílica, a polimerização ocorre pelo processo sol-gel, contudo utilizam-se monômeros que possuem tanto a função alcoxissilano quanto a função orgânica, ou ainda, na etapa de policondensação é adicionado dois precursores, geralmente o trialcóxissilano e tetraalcóxissilano, os quais sofrem ligação química em etapa única. Uma vantagem excepcional dos monolitos híbridos organo-sílica é a dupla funcionalidade (orgânica e silanol) presentes na superfície do poro, as quais permitem modificação adicional, seja com a utilização de solvente organossilano ou monômeros orgânicos. [11]

3.1.6. Entendendo os padrões morfológicos de monolitos híbridos para as separações cromatográficas

Neste item foi feita a revisão da literatura, a respeito de monolitos híbridos (organo-sílica) a base de alcóxidos de metacrilato, ([3-(metacriloiloxi)-propil]-trimetoxissilano]) (MPTMS), os quais fazem referência ao protocolo desta tese. Portanto, esta revisão teve como finalidade buscar evidências que ajudem a entender a química do monolito em termos da morfologia aliado a separações cromatográficas. Este aspecto é de suma importância para uma pesquisa efetiva e certa, principalmente com o auxílio de ferramentas quimiométricas, resultando assim o desenvolvimento de um produto promissor.

Weed, A-M.K. *et al.* verificaram que a variação da quantidade de solvente porogênico, resulta na formação de monolitos híbridos com permeabilidade e área superficial diferente. Os pesquisadores também compararam a polimerização iniciada por radiação térmica e fótons. Os resultados mostraram que processos de polimerização térmica em menor tempo reduz a eficiência da coluna cromatográfica. Por outro lado, radiação superior a 2 h não providencia mudança significativa na eficiência de separação. Além disso, verificou-se que o processo térmico gera melhores eficiências de separação para moléculas pequenas, tais como tiourea e benzeno. [11]

Geiser *et al.* reportaram boa reprodutibilidade dos poros de monolitos híbridos à base de metacrilato em ambos os processos, seja foto e termicamente iniciado. [16] Em contrapartida, Weed, A-M.K. *et al.*, verificaram que a reprodutibilidade coluna-coluna foi melhor pelo processo

de iniciação térmica. A justificativa parte do pressuposto que o processo térmico de polimerização ocorre de maneira lenta, de modo que o calor é transferido do capilar até o interior, permitindo melhor fixação do monolito na parede do capilar. Por outro lado, a fotopolimerização ocorre de forma rápida, de modo que o monolito tende a se encolher, se não for bem fixado na parede do capilar. Curiosamente, neste artigo, o grupo menciona um tratamento químico na superfície interna do capilar, o qual produz monolitos foto-polimerizados de maneira muito homogênea. [11] Entretanto, em outro artigo mais recente, o mesmo grupo não tem reportado tal tratamento químico, antes da etapa de fotopolimerização. Além disso, o grupo verificou que o aumento da percentagem de MPTMS na mistura de polimerização, melhora significativamente o número de pratos. Contudo, uma concentração do monômero com excesso de 25% v/v diminui a permeabilidade do monolito. [15]

Weed, A-M.K. *et al.*, concluíram que o monolito foto-iniciado possui dois tipos de poros: *i*) poros grandes e altamente permeáveis, os quais facilitam a passagem da fase móvel; *ii*) presença de mesoporos os quais aumentam a área superficial e, conseqüentemente, contém maiores quantidades de sítios ativos para adsorção. [11] Segundo Jandera, as macromoléculas são retidas em poros com tamanho de pelo menos 15-100 nm (grandes). Contudo, verifica-se que as melhores separações ocorrem em monolitos com área superficial baixa. Entretanto, os feitos das aplicações de macromoléculas em poros grandes de elevada área superficial, do meu melhor conhecimento, não foram relatados na literatura. [5]

Kato, M. *et al.*, evidenciaram algumas conclusões interessantes a respeito dos monolitos híbridos a base de metacrilatos de alcóxidos, como a concentração de ácido na etapa de polimerização está relacionada com a rigidez do monolito, uma vez que o ácido afeta a reação de hidrólise, de modo que baixa concentração de HCl não produz uma quantidade suficiente de grupo silanol, aspecto importante para a formação da rede polimérica. É mencionado também, que ao aumentar a quantidade do monômero em relação ao agente porogênico, a densidade do esqueleto do monolito também aumenta e, portanto, a permeabilidade diminui, uma vez que a quantidade de poros também diminui. O grupo menciona que a percentagem de monômero pode variar devido o tempo de radiação e comprimento de onda utilizado no processo de foto-iniciação. [17]

Gharbharan, D. *et al.* realizaram uma pesquisa a respeito das condições para a otimização de monolitos híbridos, comparando o processo termicamente e foto-iniciado, na separação de alquilbenzeno. Os pesquisadores relatam alguns pontos importantes como: *i)* polimerizações rápidas, com o aumento da concentração do catalisador ácido e temperatura, geram insuficiente formação de poros; *ii)* a redução de solvente porogênico na mistura de polimerização e prolongado tempo de reação permitiram colunas capilares monolíticas com elevada retenção e eficiência de separação (111.000 pratos/m) para alquilbenzenos de cadeia curta; *iii)* os monolitos termicamente iniciados possuem superfícies ligeiramente mais hidrofóbicas, com redução do número de resíduos de silanol; *iv)* a seletividade estérica foram a mesma para os dois tipos de coluna; *v)* a reprodutibilidade coluna-coluna foi melhor para o processo termicamente iniciado; *vi)* monolito fotopolimerizado proporcionou separações mais rápidas em relação ao termicamente iniciado. Gharbharan, D. *et al.*, mencionam que os monolitos a base de alcóxidos de metacrilato constituem de materiais promissores para a separação de biomoléculas, tais como proteínas. [15]

Algumas observações mais específicas tornam os estudos de Gharbharan, D. *et al.*, importantes para o desenvolvimento desta tese. Assim, os pesquisadores concluíram que o melhor resultado, em termos do fator de retenção da molécula benzeno, para processo foto-iniciado, foi utilizar 65% de tolueno e 35% de MPTMS na mistura de reação. Eles observaram também, que ao aumentar de 0,12 mol L⁻¹ de HCl para 0,20 mol L⁻¹, não resultou em significativa mudança na eficiência de separação cromatográficas. Entretanto, foi verificado maior permeabilidade para colunas capilares monolíticas com 0,12 mol L⁻¹ de HCl. Além disso, foi evidenciado, que o monolito foto-iniciado proporcionou separações mais rápidas em relação ao monolito termicamente iniciado. Esta conclusão foi avaliada por meio do uso de 50% ACN (modo isocrático) como fase móvel, a este fato o grupo atribuiu a menor quantidade de hidrocarbonetos funcionalizados na superfície do poro do monolito em questão. Finalmente, por meio da curva de *van Deemter* conclui-se que ambos monolitos (foto e termicamente iniciados) possuem elevada eficiência de separação, com o mínimo de altura do prato de 9 µm. [15]

3.2. OBJETIVOS

Desenvolver colunas cromatográficas de fase estacionária monolítica híbrida organo-sílica para separações de biomoléculas aplicadas à proteômica *shotgun*, *bottom-up* e *top-down*.

3.2.1. Objetivos específicos

1. Preparação e estudo da reação de fotopolimerização dos monolitos híbridos em capilares e canais microfluídicos.
2. Desenvolvimento de métodos analíticos para análise de peptídeos, aminoácidos e proteínas intactas: LC-MS/MS e LC-UV-Vis.
3. Otimização quimiométrica da reação de fotopolimerização em capilares de sílica e μ chips de vidro.
4. Avaliações da eficiência cromatográfica das colunas monolíticas desenvolvidas.

3.3. RESULTADOS E DISCUSSÃO

3.3.1. Testes iniciais das fases monolíticas em capilares e μ chip

As investigações basearam-se no desenvolvimento de monolitos do tipo híbrido (organo-sílica) a base de metacrilato de alcóxidos pelo processo sol-gel de foto-iniciação, inspirados principalmente nos resultados de Weed, A-M.K., [11] Dulay, M.T., [18] Gharbharan, D. [15] Kato, M., [17] Vaz, F.A.S. [7] Morichina, K. [19] Frente a estes trabalhos, procurou-se inovar a reação de fotopolimerização com a utilização de um novo foto-iniciador, cuja aplicação ao tema é incipiente. Pode-se destacar que os monolitos híbridos possuem poucos estudos, principalmente quando se referem à separação de biomoléculas.

A **Figura 28** mostra as estruturas morfológicas dos monolitos por meio da análise de microscopia eletrônica de varredura (MEV). Verificou-se que os monolitos, tanto nos capilares avaliados quanto no μ chip de vidro, foram polimerizados. Adicionalmente, verifica-se a formação da cadeia reticulada típica da síntese de monolitos, ou seja, a formação de estruturas globulares do macroporos e possivelmente a presença dos mesoporos (nanoporos).

3.3.2. Otimização da síntese dos monolitos

Realizou-se a pesquisa na literatura a respeito dos fatores/variáveis que poderiam proporcionar significância para o planejamento experimental, o objetivo nesta etapa foi otimizar a reação polimérica *in-situ*, em termos da homogeneidade e permeabilidade do material. Desta forma, elaborou-se a **Tabela 12**, a qual apresenta os fatores e intervalos dos valores avaliados tanto para o monolito fotopolimerizado no capilar com revestimento de TSU da *Polymicro*, quanto para o capilar com revestimento de acrilato da *Microtube*. Em ambos os casos, utilizou-se do planejamento fatorial fracionário 2^{4-1} com total de 8 experimentos sem réplicas. Os dados empíricos de cada capilar encontram-se nas **Tabela 13** e **Tabela 14**, respectivamente. Foram avaliadas duas respostas: *i*) a homogeneidade do monolito no capilar, representado por *Ho*; e *ii*) a permeabilidade do monolito em acetonitrila, representado pela letra *P*, conforme as respectivas

Tabela 13 e Tabela 14. Portando, o desejado neste experimento é obter elevados valores de Ho e P .

Figura 28 - Imagens de MEV dos monolitos em capilares e μ chip. A) Monolito fotopolimerizado em 254 nm no capilar de sílica fundida com revestimento de TSU da *Polymicro Technologies*. B) Monolito fotopolimerizado com luz negra no capilar de sílica fundida com revestimento de Acrilato da *Microtube* (Araraquara, Brasil). C) Monolito fotopolimerizado com luz negra no μ chip de vidro.

Fonte: Autoria própria.

Tabela 12 - Fatores e intervalos para a otimização quimiométrica do monolito em capilares de sílica fundida da *Polymicro* e *Microtube*.

Fatores			Níveis	
			-1	+1
Fatores:	1	Concentração de HCl (mol L ⁻¹)	0,12	0,50
	2	MPTMS/Tolueno (%)	25	40
	3	Massa do foto-iniciador (mg)	30,0	60,0
	4	Tempo de radiação (minutos)	10	20

Tabela 13 - Dados dos experimentos referentes ao monolito híbrido organo-sílica em capilar de sílica fundida com revestimento de TSU (Polymicro).

Estudo de FSC (Polymicro):									
Fatores					Respostas		Desejabilidade		Média Geométrica
Ensaio	1	2	3	4	Ho	P(cm ²)	dH	dP	D
1	-1	-1	-1	-1	0	1,96E-04	0,00	0,00	0,00
2	1	-1	-1	1	6	4,24E-05	0,60	0,78	0,69
3	-1	1	-1	1	4	5,36E-05	0,40	0,73	0,54
4	1	1	-1	-1	5	5,18E-05	0,56	0,74	0,64
5	-1	-1	1	1	0	1,42E-04	0,00	0,28	0,00
6	1	-1	1	-1	3	6,81E-05	0,30	0,65	0,44
7	-1	1	1	-1	9	3,53E-05	0,90	0,82	0,86
8	1	1	1	1	9	1,37E-07	0,90	1,00	0,95
LI					0	1,96E-04			
A					10	7,31E-10			
t					1	1			
Contrastes: C1 0,33; C2 0,46; C3 0,10; C4 0,06; C12 -0,23; C13 -0,06; C14 0,22; C23 0,22; C24 -0,06; C34 -0,23									
Ho – Homogeneidade do monolito; P – Permeabilidade do monolito; dH - resultado da função de desejabilidade para H; dP – resultado da função de desejabilidade para P; D média geométrica.									

A resposta da homogeneidade foi realizada por microscópio ótico. Desta forma, atribui-se por meio da observação do material monolítico contido no capilar, notas de 0 (menor homogeneidade) à 10 (maior homogeneidade), ou seja, foi verificado visualmente (e de certa forma, subjetivamente) a quantidade de luz que o material permite transmitir, assim quanto maior a transmissão menor a nota. Para cada capilar foi verificado 12 imagens de microscopia ótica. Em especial, a **Figura 29A** mostra a elevada homogeneidade ao longo do capilar obtida do **ensaio 8** da **Tabela 13**. Esta característica do material é um dos parâmetros para a condição ideal de uma boa eficiência de separação cromatográfica, uma vez que está correlacionado com baixa difusão de *Eddy* e elevada transferência de massa do analito com a fase estacionária. Em contrapartida, a **Figura 29B** e **Figura 29C** ilustram materiais monolíticos com homogeneidade inferior ao longo da coluna cromatográfica.

Tabela 14 - Dados dos experimentos referentes ao monolito híbrido organo-sílica em capilar de sílica fundida com revestimento de acrilato (Microtube)

Fatores					Respostas		Desejabilidade		Média Geométrica
Ensaio	1	2	3	4	Ho	P(cm²)	dH	dP	D
1	-1	-1	-1	-1	6	1,28E-05	0,6	0,73	0,66
2	1	-1	-1	1	8	2,85E-05	0,8	0,39	0,56
3	-1	1	-1	1	8	9,53E-06	0,8	0,80	0,80
4	1	1	-1	-1	6	1,44E-05	0,6	0,69	0,64
5	-1	-1	1	1	5	1,45E-05	0,5	0,69	0,59
6	1	-1	1	-1	7	1,54E-05	0,7	0,67	0,69
7	-1	1	1	-1	8	1,69E-05	0,8	0,64	0,72
8	1	1	1	1	10	2,73E-06	1,0	0,94	0,97
LI					0	4,69E-05			
A					10	7,31E-10			
t					1	1,0			
Contrastes: C1 -0,08; C14 -0,06; C23 -0,06; C4 -0,02; C12 0,17; C34 0,17; C3 0,18 C2 0,29; C13 0,30; C24 0,30									
Ho - Homogeneidade do monolito; P - Permeabilidade do monolito; dH - resultado da função de desejabilidade para H; dP - resultado da função de desejabilidade para P; D média geométrica.									

Em relação à resposta de permeabilidade do monolito, esta foi calculada por meio da lei de Darcy (**Equação 13**):

$$P = (F_m \times \eta \times L) / (\Delta p \times \pi \times r^2) \quad \text{Equação 13}$$

Onde F_m é o fluxo da fase móvel, η é a viscosidade da fase móvel (acetonitrila), L é o comprimento da coluna, Δp é a variação da pressão no capilar e r é o raio da coluna. As medidas de pressão da coluna capilar monolítica foram efetuadas por meio de bombas de HPLC da Shimadzu Prominence 9A.

Por se tratar de duas respostas no planejamento experimental, torna-se necessário aplicar a função de desejabilidade descrita por Derringer *R.* e Suich, *R.* e média geométrica, conforme as **Equações 2 e 4.** [20]

Figura 29 - Imagens de microscopia ótica para colunas capilares monolíticas híbridas organo-sílica em diferentes concentrações dos reagentes de partida da reação de polimerização. A) Elevada homogeneidade do material ao longo da coluna (material de alta densidade e baixa permeabilidade). Condições: concentração de HCl 0,50 mol L⁻¹ e MPTMS/Tolueno de 40% – ensaio 8 da **Tabela 13**. B) Homogeneidade média do material ao longo da coluna (densidade média e baixa permeabilidade) Condição: concentração de HCl de 0,12 mol L⁻¹ e MPTMS/tolueno de 40% – ensaio 3 da **Tabela 13**. C) Baixa homogeneidade do material ao longo da coluna (material de baixa densidade e alta permeabilidade). Condição: concentração de HCl de 0,12 mol L⁻¹ e MPTMS/tolueno de 25% – ensaio 1 da **Tabela 13**. Em cada figura são indicadas 5 imagens de áreas diferentes da mesma coluna capilar monolítica. Os valores de homogeneidade e permeabilidade são encontrados na **Tabela 13**.

Fonte: Autoria própria.

A **Figura 30** sugere que para atingir os melhores resultados em termos de *Ho* e *P*, é necessário aumentar a razão da % MPTMS/Tolueno, assim como a concentração de HCl. Estes dois fatores são responsáveis respectivamente por 39,20% e 19,73%, na otimização. Tais resultados referem-se a otimização quimiométrica para o monolito híbrido organo-sílica polimerizado no capilar da *Polymicro*. Já os resultados da otimização quimiométrica, referente ao monolito híbrido organo-sílica polimerizado no capilar da *Microtube*, representado pela **Figura 31**, sugere que as interações dos contrastes C24 (24,5%) e C13 (24,5%), assim como o contraste C2 (22,2%) são os resultados mais significativos na otimização. Estas evidências mostram que se torna necessário considerar todos os fatores na otimização do monolito em questão. No entanto, o fator MPTMS/Tolueno, por ser o contraste de primeira ordem, de maior contribuição na resposta, talvez seja o fator mais importante na otimização. Como os resultados não foram estatisticamente significativos, com 95% do intervalo de confiança, pode-se inferir que o intervalo dos fatores 1 e 2 da **Tabela 12** podem não ter sido adequados para a avaliação dos monolitos polimerizados em ambos os capilares. Isso estaria de acordo com estudos anteriores. [11,15] Assim, para a aplicação da superfície de resposta (planejamento composto central) o ajuste dos intervalos dos fatores em questão foram reformulados.

Figura 30 - Gráfico normal dos contrastes e tabela da percentagem de contribuição dos contrastes na resposta ótima, referente à otimização da **Tabela 13** (monolito fotopolimerizado no capilar da *Polymicro*).

Fonte: Autoria própria.

Figura 31 - Gráfico normal dos contrastes e tabela da percentagem de contribuição dos contrastes na resposta ótima, referente a otimização da **Tabela 14** (monolito fotopolimerizado no capilar da *Microtube*).

Fonte: Autoria própria.

3.3.3. Validação da instrumentação para avaliação da eficiência das colunas capilares monolíticas

Com o objetivo de desenvolver um método cromatográfico simplificado para a avaliação das colunas capilares monolíticas, em termos da eficiência de separação, antes da análise no espectrômetro de massas, empregou-se a cromatografia líquida de nano-fluxo(vazão) (*nano-LC*) com detecção capilar *UV-Vis* (*LabAlliance*), conforme representado pela **Figura 32**. Desta forma,

a aquisição dos dados foi realizada pelo programa *Labview* escrito por ex-integrantes do grupo *BioMicS*. [21,22]

Figura 32 - Fotografia do sistema cromatográfico nano-LC com coluna monolítica acoplada ao detector UV-Vis capilar.

Fonte: Autoria própria.

Por se tratar de um sistema de detecção capilar, cujo caminho ótico é de 75 μm e volume da cela de 2,78 pL, torna-se necessário avaliar algumas figuras de mérito, de modo a conhecer o sistema de detecção. Para este fim, selecionou-se a histidina (His) como padrão de referência, devida esta possuir absorção na região UV em 211 nm, próxima às regiões de absorbância dos analitos de interesse (proteínas e peptídeos), cuja absorção ocorre preferencialmente em 214 nm ou 280 nm. Adicionalmente, o padrão em questão possui baixo custo, fácil preparação, solubilidade em água e boa estabilidade.

A **Figura 33** apresenta a curva analítica do sistema de detecção capilar acoplado ao nano-LC, cuja sensibilidade foi de 0,089 mAU. Adicionalmente, verifica-se boa faixa linear de 1-5 mmol L⁻¹ de His. A partir deste intervalo há desvio da linearidade e aumento do desvio padrão. Sabe-se que o aumento da concentração do analito na cela de detecção, está correlacionado com a resposta não ideal do detector ou devido a efeitos químicos. Além disso, calculou-se o limite de detecção (LD) de 92,26 nmol L⁻¹ cujo nível de confiança foi de 95%.

Figura 33 - Curva analítica para quantificação de histidina utilizando-se detector capilar UV-Vis (Lab Alliance).

Fonte: Autoria própria.

Apesar do nano-LC da empresa *Thermo* possuir um sistema de desgaseificação próprio, a detecção de bolhas no UV-Vis capilar na maioria das vezes é inevitável. Este fenômeno é observado devido a presença de sinais característicos no UV-Vis, tais como: sinais sequenciais e de máxima amplitude (saturado). Desta forma, tomou-se os cuidados necessários com as conexões que contenham volume morto, uma vez que a quantidade de ar inserida ao conectar o sistema cromatográfico pode perdurar por muito tempo durante a análise, principalmente em condições cromatográficas com vazão de nL min^{-1} . Por meio da observação, verificou-se que quando o sistema cromatográfico está pressurizado acima de 200 bar, verificou-se menor formação de bolhas. Provavelmente, a pressurização aumenta a pressão de vapor da fase móvel, fato que diminui a formação de bolhas. Além disso, oscilações da temperatura na sala onde se situa o cromatógrafo podem favorecer a formação de bolhas. Uma hipótese adicional é que a própria estrutura do monolito interfira e de alguma forma propicie a formação de bolhas.

3.3.4. Obtenção da curva de *van Deemter*

Para que as medidas da eficiência de separação cromatográfica (N) ou altura do prato (H) sejam adequadas, torna-se necessário conhecer a velocidade ótima da fase móvel do analito em

relação a fase estacionária. Na prática, trata-se de obter formato de pico gaussiano com largura de base mínima e com tempo de retenção elevado. Para esta finalidade, a curva de *van Deemter* é a ferramenta ouro para as primeiras evidências de uma otimização cromatográfica da coluna. Sabe-se que tanto N quanto H são afetados pelo *i*) comprimento da coluna; *ii*) temperatura; *iii*) soluto; *iv*) solvente; *v*) vazão; *vi*) tamanho da amostra e *vii*) técnica de injeção. Desta forma, selecionamos as 4 melhores colunas monolíticas, após a otimização realizada pela aplicação da superfície de resposta, em termos da permeabilidade e homogeneidade do material polimérico ao longo do capilar, perfazendo assim, a curva de *van Deemter*, conforme representado pela **Figura 34**.

Verificou-se que a menor altura do prato (H) para as 4 colunas monolíticas (*DoE 1*, *DoE 6*, *DoE 9* e *DoE 10*) fazem referência a vazão de 250 nL min^{-1} (velocidade linear de aproximadamente $0,24 \text{ mm s}^{-1}$), cuja H variou de $596\text{--}885 \text{ }\mu\text{m}$. Não encontramos a aplicação da curva de *van Deemter* para biomoléculas na literatura. Contudo, Liu, Z. *et al.* relatou uma vazão ótima para alquilbenzenos em colunas monolíticas cuja as dimensões são semelhantes ao nosso estudo ($20,0 \text{ cm} \times 75 \text{ }\mu\text{m d.i.}$), onde o regime ideal da fase móvel foi de 220 nL min^{-1} . [23] Entretanto, o referido grupo obteve melhor eficiência cromatográfica com valores de H em torno de $10\text{--}12 \text{ }\mu\text{m}$. Isso se deve a maior interação intermolecular hidrofóbica de alquilbenzenos em fase reversa de separação cromatográfica, o qual promove elevado tempo de retenção, consequentemente têm-se melhoras nos valores de N e H . Contudo, acredita-se que o monolito desenvolvido neste trabalho não atingiu a condição de síntese ótima e possivelmente possui vacâncias em seu interior, fato que contribui com o alargamento de banda e diminuição na eficiência cromatográfica. A **Figura 35** mostra os cromatogramas obtidos, cujo o alargamento de banda foi de aproximadamente $4,0 \text{ min.}$ e tempo de retenção de $11,5 \text{ min.}$ para o triptofano. Em contrapartida, por meio dos estudos quimiométricos, conseguiu-se encontrar a região que possivelmente torne a eficiência cromatográfica melhor e equiparável aos estudos de Liu, Z. *et al.*, conforme será discutido no próximo tópico. [23]

Figura 34 - Curva de *van Deemter*, cujo o analito foi o aminoácido triptofano. Cada cor indica as diferentes colunas monolíticas avaliadas, conforme o planejamento experimental proveniente da superfície de resposta (DoE 01, DoE 06, DoE 09 e DoE 10), a qual será apresentada no próximo tópico. As condições foram: concentração do analito de 1 mg mL⁻¹; volume de injeção 0,25 µL; coluna capilar monolítica híbrida orgânico-sílica com dimensões de 20 cm × 75 µm d.i.; fase móvel 50% ACN_(aq) e detecção no comprimento de onda em 280 nm.

Fonte: Autoria própria.

Figura 35 - Cromatogramas obtidos para eluição de triptofano com vazão de 250 nL min⁻¹. A) Cromatograma referente a coluna monolítica DoE 01, cuja altura do prato teórico foi de 596 µm. B) Cromatogramas referente as diferentes colunas monolíticas avaliadas. As condições foram: concentração do analito de 1 mg mL⁻¹; volume de injeção 0,25 µL; coluna capilar monolítica híbrida organo-sílica com dimensões de 20 cm × 75 µm d.i.; fase móvel 50% ACN_(aq), vazão de 250 nL min⁻¹ e detecção no comprimento de onda em 280 nm.

Fonte: Autoria própria

Construiu-se a curva de *van Deemter* para a proteína mioglobina, entretanto verificamos baixa reprodutibilidade e curva assimétrica para as colunas monolíticas referentes ao *DoE 09* e *DoE 10*.

Para as colunas monolíticas referente ao *DoE 01* e *DoE 02*, obteve-se uma curva de *van Deemter* simétrica (**Figura 36**). Para a coluna monolítica *DoE 01* a vazão ótima foi 250 nL min⁻¹. Em especial a coluna monolítica *DoE 02* atingiu melhores valores de *H* (98 µm), contudo a vazão da fase móvel foi realizada em regime lento (50 µL min⁻¹), uma vez que verificou-se baixa permeabilidade do material ($2,48 \times 10^{-9}$ cm²). Fato que não permitiu conduzir as análises para vazões superiores a 200 nL min⁻¹, uma vez que gerou pressão > 300 bar, de modo que o sistema nano-cromatográfico resiste pressão de até 300 bar. Sabe-se que a baixa vazão da fase móvel e longo tempo de análise contribuem com o efeito de dispersão do analito e consequente alargamento de banda, conforme evidenciado na **Figura 37**.

Figura 36 - Curva de *van Deemter* simétricas para as colunas monolíticas referente ao *DoE 01* e *DoE 02*, cujo analito foi a mioglobina. As condições foram: concentração do analito de 5 mg mL⁻¹; volume de injeção 0,25 µL; coluna capilar com dimensões de 20 cm × 75 µm d.i.; fase móvel 50% ACN_(aq) e detecção com comprimento de onda em 280 nm.

Fonte: Autoria própria.

A **Figura 37** mostra o elevado tempo de retenção da proteína de mioglobina na coluna monolítica *DoE 02*, o que justifica a melhor eficiência cromatográfica desta. É importante mencionar que a baixa permeabilidade e o entupimento desta coluna não nos permitiu avaliar a mesma para o triptofano, o qual poderia apresentar resultados melhores. Contudo, podemos inferir que as condições de polimerização da coluna monolítica *DoE 02* podem estar próxima da região de síntese ótima.

3.3.5. Avaliação da polimerização do material monolítico nas colunas otimizadas por meio da superfície de resposta

Tendo em mente a melhor condição da vazão da fase móvel, aplicamos alguns cálculos para a caracterização física de todas as fases monolíticas sintetizadas, na otimização quimiométrica dirigida pelo composto central, tais como: *i*) eficiência cromatográfica (N); *ii*) altura de pratos (H) e *iii*) permeabilidade do monolito (P), conforme mencionado pelas **Equação 14 e Equação 15**. Onde, t_R é o tempo de retenção do analito e W_b é a largura da base do pico, ambas as medidas foram efetuadas em unidade de tempo (min.) e L refere-se ao comprimento da coluna.

Figura 37 - Cromatograma da proteína mioglobina referente a coluna monolítica DoE 02. As condições foram: concentração do analito de 5 mg mL⁻¹; volume de injeção 0,25 µL; coluna capilar com dimensões de 20 cm × 75 µm d.i.; fase móvel 50% ACN(aq), vazão de 50 nL min⁻¹, velocidade linear de 0,05 mm s⁻¹ e detecção com comprimento de onda de 280 nm.

Fonte: Autoria própria.

$$N = 16 \left(\frac{t_R}{W_b} \right)^2 \quad \text{Equação 14}$$

$$H = \frac{L}{N} \quad \text{Equação 15}$$

Aplicou-se o planejamento composto central (CCD) para a otimização das colunas monolíticas referente aos dois fatores significativos na síntese dos monolitos: catalisador (Cat.), ou seja, concentração de HCl e monômero (Mon.) (razão %MPTMS/tolueno). Desta forma as

respostas foram *N*, *H* e *P* para triptofano (**Tabela 15**) e mioglobina (**Tabela 16**). O triptofano foi escolhido devido apresentar absorção na região UV, semelhança com peptídeos de massa mole-

Tabela 15 - Valores obtidos referente a aplicação do composto central na otimização das colunas monolíticas híbridas organo-sílica para análise cromatográfica e física do triptofano.

<i>Desejabilidade</i>								
	<i>Cat.</i>	<i>Mon.</i>	<i>N</i>	<i>H (μm)</i>	<i>P (cm²)</i>	<i>dH</i>	<i>dP</i>	<i>D</i>
1	0,06	25,00	337,47	595,61	8,57E-07	0,97	0,82	0,89
2	0,18	25,00	2197,16	98,31	2,48E-09	1,00	1,00	1,00
3	0,06	30,00	9,00	22000,00	5,62E-09	0,00	1,00	0,01
4	0,18	30,00	182,16	1207,89	3,43E-07	0,95	0,93	0,94
5	0,01	27,50	9,00	22000,00	8,57E-06	0,00	0,00	0,00
6	0,23	27,50	272,03	720,51	7,41E-07	0,97	0,84	0,90
7	0,12	22,50	18,71	10580,52	7,21E-07	0,52	0,85	0,66
8	0,12	32,50	205,20	887,61	7,59E-07	0,96	0,84	0,90
9	0,12	27,50	201,35	885,32	7,10E-07	0,96	0,85	0,90
10	0,12	27,50	270,32	654,78	7,03E-07	0,97	0,85	0,91
11	0,12	27,50	500,13	411,89	8,33E-07	0,98	0,82	0,90
			<i>LS</i>	22001	9,00E-06			
			<i>A</i>	9,00	7,31E-10			

Tabela 16 – Valores obtidos referente a aplicação do composto central na otimização das colunas monolíticas híbridas organo-sílica para análise cromatográfica e física da mioglobina.

<i>Desejabilidade</i>								
	<i>Cat.</i>	<i>Mon.</i>	<i>N</i>	<i>H (μm)</i>	<i>P (cm²)</i>	<i>dH</i>	<i>dP</i>	<i>D</i>
1	0,06	25,00	100,00	1800	8,57E-07	0,92	0,90	0,91
2	0,18	25,00	2197,16	98,31	2,48E-09	1,00	1,00	1,00
3	0,06	30,00	9,00	22000	5,47E-09	0,00	1,00	0,01
4	0,18	30,00	12,98	16976,27	3,43E-07	0,23	0,96	0,47
5	0,01	27,50	9,00	22000	8,57E-06	0,00	0,05	0,00
6	0,23	27,50	15,39	1192,55	7,41E-07	0,95	0,92	0,93
7	0,12	22,50	28,50	689,80	7,21E-07	0,97	0,92	0,94
8	0,12	32,50	18,17	1012,60	7,59E-07	0,95	0,92	0,93
9	0,12	27,50	95,43	188,62	7,10E-07	0,99	0,92	0,96
10	0,12	27,50	69,96	271,78	7,03E-07	0,99	0,92	0,95
11	0,12	27,50	47,58	403,46	8,33E-07	0,98	0,91	0,94
			<i>LS</i>	22001	9,00E-06			
			<i>A</i>	9	7,31E-10			

cular baixa assim como polaridade e hidrofobicidade, além da estabilidade, alta pureza e baixo custo. A mioglobina foi escolhida devido ser uma proteína menor, que em tese, diminui a complexidade de análise cromatográfica, ajudando a compreender melhor os sistemas nesta etapa inicial de refinamento da otimização do material polimérico (monolito).

Os valores em negrito da **Tabela 15** e **Tabela 16** não foram calculados. Pois não foi possível realizar a medida devido ao entupimento da coluna monolítica ou pela má formação de banda devido a elevada dispersão do analitos no meio cromatográfico durante as análises.

Como se avaliou duas respostas na aplicação do composto central, tornou-se necessário aplicar a função de desejabilidade (d) descrita por *Derringer* e *Suich*, e média geométrica, conforme as **Equações 2, 3, 4 e 5 (Capítulo 1)**. [20]

As **Figura 38** e **Figura 39** indicam a superfície de resposta e região ótima em termos de eficiência cromatográfica e permeabilidade do material para triptofano e mioglobina, respectivamente. Ao comparar as duas superfícies de respostas, podemos inferir que há uma região de ótimo para melhor eficiência cromatográfica para molécula pequena (triptofano) e outra região de ótimo para molécula grande (mioglobina). Entretanto, é factível que a falta de reprodutibilidade e consistência nas análises da mioglobina pode ter ocasionado esse fenômeno. Portanto, mais experimentos são necessários para maiores conclusões.

3.3.6. Refinamento da otimização das colunas capilares monolíticas para nanofluxo

Com os resultados obtidos, verificou-se que algumas condições de síntese não formaram o monolito ou proporcionaram entupimento da coluna monolítica. Além do mais, uma condição de síntese refinada se torna necessária de modo a diminuir a formação de vacâncias, os quais, em tese, justificam o alargamento de banda e baixa eficiência cromatográfica. Com estas considerações, reformulou-se o planejamento quimiométrico composto central, com intervalo de parâmetros os quais acreditamos que sejam mais adequados para a efetiva otimização da fase monolítica.

Figura 38 - Superfície de resposta para a otimização das colunas monolíticas referentes as medidas para o triptofano.

Fonte: Autoria própria.

Figura 39 - Superfície de resposta para a otimização das colunas monolíticas referentes as medidas para a mioglobina.

Fonte: Autoria própria.

A **Figura 40** mostra a curva de *van Deemter* para a melhor coluna monolítica obtida. Verifica-se que a velocidade ótima da fase móvel com maior número de partições entre a fase estacionária é de fato 250 nL min^{-1} , tanto para o triptofano quanto para BSA. Estes resultados indicam $N = 2.680,14$ e $H = 82 \text{ } \mu\text{m}$ e $N = 1.228,20$ e $H = 179 \text{ } \mu\text{m}$ para triptofano e BSA, respectivamente. Estes valores são considerados como adequados na literatura ($H < 600 \text{ } \mu\text{m}$).

Contudo, colunas empacotadas com partículas com 10 μm de diâmetro possuem altura de pratos (H) de aproximadamente 2,0 μm e eficiência que podem atingir até 20.000, em contrapartida, o sistema cromatográfico deverá suportar elevada pressão (7000 bar).[2] A vantagem da coluna monolítica é que N e H podem ser melhorados pelo simples fato de aumentar o comprimento da coluna, sem aumento significativo da pressão no sistema cromatográfico, uma vez que o monolito possui elevada permeabilidade em relação as colunas empacotadas. [24]

Figura 40 - Curva de *van Deemter* para as melhores coluna monolíticas obtidas. A) Curva de *van Deemter* para triptofano. B) curva de *van Deemter* para BSA. As condições foram: volume de injeção 0,25 μL ; coluna capilar com dimensões de 22 $\text{cm} \times 75 \mu\text{m}$ d.i.; fase móvel 50% $\text{ACN}_{(\text{aq})}$ e detecção com comprimento de onda de 280 nm.

Fonte: Autoria própria.

A **Figura 41** mostra os cromatogramas obtidos referente ao tempo de retenção do triptofano e do BSA em diversas vazões. Verifica-se a formação de bandas estreitas onde os valores da largura da meia altura do pico cromatográfico (W_h), para o triptofano e BSA foram 0,67 min. e 0,99 min., respectivamente. Os tempos de retenção (t_R) foram 14,73 min. para o triptofano e 14,85 min. para BSA. Um parâmetro importante para avaliar se de fato as moléculas hidrofílicas possuem retenção na coluna é o cálculo do tempo de retenção ajustado ($t_{R'}$), o qual considera o tempo transcorrido do analito desde o momento de injeção da amostra, ou seja, desconsidera-se o tempo morto (t_M) ou volume morto (V_M) do sistema cromatográfico, conforme representado pela **Equação 16**.

$$t_{R'} = t_R - t_M \quad \text{Equação 16}$$

Figura 41 - Cromatogramas das colunas capilares monolíticas DoE 16. A) análise de triptofano em diversas vazões; B) análise cromatográfica de albumina de soro bovino (BSA) em diversas vazões; C) Tempo de retenção do triptofano em vazão de 250 nL min⁻¹; D) Tempo de retenção da BSA em vazão de 250 nL min⁻¹. As condições foram: volume de injeção 0,25 µL; coluna capilar com dimensões de 22 cm × 75 µm d.i.; fase móvel 50% ACN_(aq) e detecção com comprimento de onda em 280 nm.

Fonte: Autoria própria.

Para obter o valor de t_M é necessário injetar um analito que não tenha interação com a fase estacionária da coluna cromatográfica. Geralmente, o analito de escolha para colunas cromatográficas é a tioureia. Contudo, verificamos que este analito possui interação com o nosso monolito, impedindo o cálculo de t_M . Desta forma, injetamos uma concentração elevada de tioureia [1 mol L⁻¹] na coluna monolítica de modo a saturar a fase estacionária da mesma. Por meio de um gráfico é possível verificar o ponto de saturação do detector UV-Vis capilar, o qual indica o ponto de equivalência do analito não interagido com a fase estacionária, sendo assim é possível estimar o valor de t_M com precisão (**Figura 42**). Como a primeira derivada de uma constante é zero, é possível aplicar esta propriedade matemática de modo a obter o ponto de equivalência de saturação no detector, o qual faz referência ao t_M do sistema cromatográfico (**Figura 42B e Figura 42C**).

Figura 42 - Gráficos para a obtenção do tempo morto (t_M) do sistema nano-cromatográfico com vazão de 250 nL min⁻¹ ou velocidade linear de 0,24 mm s⁻¹. A) Curvas de equivalências de saturação no detector UV-Vis capilar. B) Primeira derivada da injeção 1. C) Primeira derivada da injeção 2.

Fonte: Autoria própria.

Por meio das **Figura 42B** e **Figura 42C**, nota-se que t_M possui o valor de 8,5 minutos. Sendo assim, o tempo de retenção ajustado ($t_{R'}$) é de 6,23 min. para triptofano e 6,35 min. para BSA. Podemos inferir que a coluna monolítica híbrida possui adequado tempo de retenção para tais moléculas hidrofílicas.

Ao analisar outra coluna monolítica (DoE 14) verificamos que o triptofano e BSA possuem tempos de retenção distintos, conforme representado pela **Figura 43**. Uma hipótese, é

que o triptofano, por ser menor, interage com os mesoporos (nanoporos), ocasionando maior tempo de retenção em relação ao BSA. Em contrapartida a molécula de BSA percola entre os macroporos. Por meio desta observação podemos inferir que esta coluna monolítica possui seletividades em termos de tamanho e hidrofobicidade.

Figura 43 - Cromatogramas indicando o tempo de retenção distinto das moléculas na coluna monolítica híbrida organo-sílica DoE 14. As condições foram: volume de injeção 0,25 μ L; coluna capilar com dimensões de 22 cm $\times$ 75 μ m d.i.; fase móvel 50% ACN_(aq) e detecção com comprimento de onda de 280 nm.

Fonte: Autoria própria.

3.4. Avaliação das colunas capilares monolíticas para proteômica *bottom-up* ou *shotgun*

Para melhor entendimento da fase estacionária monolítica, investigou-se sua seletividade por meio da separação de 14 peptídeos com massa molar de 843 a 1600. Realizou-se a análise em diferentes condições cromatográficas (isocrático e gradiente) utilizando como fase móvel acetonitrila e água.

A **Figura 44** ilustra a separação de alguns peptídeos, mas com elevado grau de alargamento de banda (< 5 minutos). A **Figura 45** mostra a separação dos mesmos 14 peptídeos em coluna de C₁₈ com partículas de 3 μ m de diâmetro. Embora nossas colunas monolíticas apresentem as ótimas propriedades de um material híbrido (organo-sílica), *i.e.* *i*) alta permeabilidade; *ii*) estabilidade em elevado intervalo de pH; *iii*) estabilidade térmica e *iv*) boa resistência mecânica em elevada vazão (testes não mostrados). As evidências mostram que

atualmente, nossa coluna capilar monolítica híbrida, não possui eficiência cromatográfica equiparável as colunas capilares empacotadas com partículas de C₁₈ de 3 µm. Entretanto, as colunas monolíticas desenvolvidas podem ser equiparadas com a eficiência de separação de colunas com baixa eficiência cromatográficas, tais como SCX e SEC.

Figura 44 - Cromatogramas de íons para mistura de 14 padrões de peptídeos obtidos na coluna monolítica híbrida organo-sílica utilizando um nLC-LTQ-Orbitrap XL com intervalo de massas observado de 400 a 1000 m/z . As letras destacadas em vermelho no cromatograma de íons totais referem-se a uma mistura de m/z $z=+2$: A) 488,76, 609,31; B) 488,77, 609,31, 447,28, 492,28; C) 492,28, 609,31; D) 492,28, 609,31, 447,28; E) 418,73, 691,83; F) 418,73, 691,83; G) 418,73, 691,83; H) 581,80, 418,73, 691,73; I) 581,80, 769,89, 418,73; J) 741,39, 493,80, 796,41, 769,89, 581,80; L) 493,80, 553,32, 741,39, 796,41; M) 568,30; 676,37. As condições do método foram: coluna 60 cm $\times$ 75 µm i.d.; gradiente linear com aumento a uma taxa de 5%ACN min^{-1} em 0.1% ácido fórmico aquoso em 70 minutos; vazão 300 nL min^{-1} ; injeção de amostra de 1 pmol na coluna de cada peptídeo. Os números nos cromatogramas de íons indicam as réplicas.

Fonte: Autoria própria.

3.5. Avaliação de colunas capilares monolíticas híbridas organo-sílica para proteômica *top-down*

Avaliou-se a seletividade da coluna capilar monolítica híbrida organo-sílica para uma

mistura de 3 macromoléculas com as respectivas massas moleculares: *i*) citocromo C 12.327 Da; *ii*) lisozima 14.388 Da e *iii*) ribonuclease A 13.700 Da. Foi realizado vários testes cromatográficos para promover uma boa separação em termos de picos gaussianos e capacidade de pico. A **Figura 46** mostra a separação de macromoléculas e o espectro de massas típico para proteínas intactas.

Figura 45 - Cromatogramas de íons para 14 peptídeos obtidos na coluna C₁₈ utilizando nLC-LTQ-Orbitrap XL com intervalo de massas observado de 400 a 1000 com $z = +2$. As letras em vermelho no cromatograma de íons referem-se a uma mistura de m/z : A) 488.76; B) 609.31; C) 492.28; D) 447.28; E) 418.73; F) 691.82; G) 581.80; H) 769.89; I) 553.32; J) 741.38; L) 796.41; M) 493.80; N) 568.30; O) 676.37. As condições cromatográficas foram: coluna 12 cm $\times$ 75 μ m d.i.; gradiente linear com proporção crescente de 5%ACN min^{-1} em 0.1% ácido fórmico aquoso em 0.1% em 40 minutos; vazão 300 nL min^{-1} ; concentração de amostra na coluna 1 nmol, para cada peptídeo.

Fonte: Autoria própria.

No topo da **Figura 46**, o cromatograma de íon mostra que a forma do pico não possui distribuição gaussiana simétrica. Entretanto, trata-se de um comportamento físico comum em proteômica *top-down*. Adicionalmente, houve uma ligeira coeluição entre o citocromo C e a ribonuclease A. Contudo, houve uma considerável diferença no tempo de retenção entre a lisozima para citocromo C e ribonuclease A, uma boa separação e boa reprodutibilidade entre as réplicas.

Assim, pode-se inferir que nossas colunas capilares monolíticas híbridas são promissoras para proteômica *top-down* e ela poderia ser mais efetiva para outros grupos de proteínas intactas, uma vez que a estrutura da molécula, assim como, o comprimento da coluna, influenciam na eficiência cromatográfica. Curiosamente, dependendo da condição cromatográfica as proteínas

apresentam perfil de espectro e estado de carga diferentes, os quais podem estar relacionados com a protonação do solvente nas proteínas.

Figura 46 - Cromatograma de íons e espectro de massas para uma mistura de 3 padrões de biomoléculas para duas condições cromatográficas diferentes A) Condição 1: coluna monolítica com 35 cm × 75 µm d.i.; gradiente linear de 6% ACN min⁻¹ em 0,1% ácido fórmico aquoso; vazão 500 nL min⁻¹; temperatura 70 °C; volume de injeção de 5 µL para 2 mg mL⁻¹ de amostra e B) Condição 2: coluna 35 cm × 75 µm i.d.; gradiente linear de 3% ACN min⁻¹ em 0,1% ácido fórmico aquoso; vazão de 1500 nL min⁻¹; temperatura 70 °C; volume de injeção de 5 µL de 1 mg mL⁻¹ de amostra.

Fonte: Autoria própria.

3.6. Alternativas para as colunas monolíticas híbridas organo-sílica desenvolvidas

Verificou-se que as colunas monolíticas híbridas organo-sílica possuem elevado poder de retenção de peptídeos, em modo cromatográfico isocrático, utilizando como fase móvel uma mistura de 95% de acetonitrila e 5% de água na vazão de 1 µL min⁻¹ (**Figura 47**). Portanto, o meio monolítico desenvolvido pode funcionar como uma armadilha de peptídeos no método; cromatografia líquida multidimensional HILIC×RP, cuja a primeira dimensão trata-se da cromatografia líquida de interação hidrofílica, onde a eluição dos analitos é realizada por acréscimo de água em relação a acetonitrila, seguido pela coluna armadilha monolítica para elevada retenção dos peptídeos em até 95% ACN, e consequente cromatografia líquida de fase

reversa, onde a eluição dos analitos é realizada por acréscimo de acetonitrila em relação a água (oposto da primeira dimensão).

Figura 47 - Eluição direta de 14 peptídeos em nLC-MS, para quatro réplicas de colunas/traps cromatográficas monolíticas, denominadas como A, B, C e D. Para cada coluna/trap (letra) são representadas 8 réplicas. As condições cromatográficas e espectrométricas foram: condição isocrática com 95% acetonitrila aquoso e 0,1% de ácido fórmico; injeção de amostra de 1 pmol; vazão de 1 $\mu\text{L min}^{-1}$; tensão elétrica da eletronebulização de 2.5 kV e temperatura do capilar de 325 °C.

A) Coluna/trap 1 - 30 cm

B) Coluna/trap 2 - 30 cm

C) Coluna/trap 3 - 30 cm

D) Coluna/trap 4 - 40 cm

Fonte: Autoria própria.

Este tipo de modalidade, possui compatibilidade entre as dimensões cromatográficas, em termos da fase móvel; rápido equilíbrio para as mudanças de gradientes entre as dimensões, pois o monolito permite elevada vazão; além de possuir uma das mais elevadas ortogonalidades cromatográficas, conforme reportado por Gilar, *et. al.*; [25]. Segundo nosso conhecimento, este

tipo de aplicação não foi reportado. Adicionalmente, tais propriedades do material monolítico, permite a realização da limpeza, dessalinização e pré-concentração dos analitos. E ainda, a empresa [glysci](#) comercializa colunas de armadilha no valor de U\$ 225,00 dólares por unidade. Portanto, há uma oportunidade de ganhos de capital nesta tecnologia.

Para uma segunda possibilidade de aplicação, verificou-se por meio de um método cromatográfico, cuja configuração possui ponta de vidro para eletronebulização no MS, empacotada no máximo com 5 mm de fase estacionária com partículas de C₁₈ de 3 µm, conectado a coluna monolítica híbrida organo-sílica (**Figura 48**). Esta configuração permitiu a separação de 7 peptídeos em menos de 12 minutos, conforme indicado pela **Figura 49**.

Figura 48. Esquema do método cromatográfico multidimensional para a separação rápida de sete peptídeos.

Figura 49 - Cromatogramas de íons referente a separação de sete peptídeos com m/z ($z = +2$): 488,76, 609,31, 418,73, 691,82, 769,89 741,38 e 568,30. Provenientes de um sistema cromatográfico multidimensional com a ponta para eletronebulização preenchida, no máximo com 5 mm, com partículas de C₁₈ conectada a coluna monolítica com 30 cm de comprimento. Esta ponta de eletronebulização trata-se de um acessório extra acoplado a coluna monolítica. A) Gradiente orgânico com taxa de 3% ACN min⁻¹ e nanofluxo de 900 nL min⁻¹. B) Gradiente orgânico com taxa de 5% ACN min⁻¹ e nano-fluxo de 900 nL min⁻¹.

Fonte: Autoria própria.

3.6.1. CONCLUSÕES

As colunas capilares monolíticas híbridas organo-sílica possuem excelentes propriedades físico-químicas para superar os desafios da proteômica. Por décadas, as partículas C_{18} têm sido a fase estacionária de escolha para a separação de peptídeos em cromatografia líquida. No momento, acredito que dificilmente outro tipo de fase estacionária poderá superar a elevada capacidade de separação dos peptídeos das colunas empacotadas com partículas de C_{18} . Contudo, as colunas monolíticas desenvolvidas podem ser equiparáveis ou até mesmo melhores, em termos de eficiência cromatográfica, em relação às colunas de troca iônica ou por exclusão de tamanho.

No mercado atual há colunas capilares monolíticas de sílica com grupo de C_{18} imobilizado na superfície do monolito. Embora essas colunas tenham excelente poder de separação, acrescidas das vantagens do material monolítico, quando comparadas com o monolito híbrido, elas possuem alguns inconvenientes, tais como: baixa estabilidade química, instabilidade em pH básico, baixa estabilidade térmica, elevado efeito de memória e dificuldades na fabricação.

Futuros estudos sobre a imobilização de C_{18} na superfície dos nanoporos poderiam ser uma estratégia para melhorar o meio cromatográfico para proteômica *shotgun*. Atualmente, os estudos sobre nanoporos são uma tendência em química e ciência geral.

Adicionalmente, os resultados sugerem que as colunas capilares monolíticas híbridas organo-sílica funcionam para macromoléculas. Contudo, para um produto eficiente e inovador, os resultados exigem melhorias para atingir os requerimentos do mercado. Desta forma, menciono duas sugestões. A primeira e mais fácil, seria aplicar diferentes classes de proteínas intactas para o atual material monolítico, uma vez que as estruturas das moléculas escolhidas, como os analitos padrão, influenciam na eficiência cromatográfica. É interessante mencionar que alguns estudos indicam que o monolito deste estudo possui potencial para resolução cromatográfica de proteínas. A segunda e mais difícil, seria desenvolver um novo desenho de reação polimérica. Atualmente, o grupo C_4 é o melhor meio de separação para proteínas intactas em cromatografia líquida. A fase estacionária monolítica deste estudo possui grupos metil C_1 na superfície. Por meio da experiência adquirida, uma possibilidade seria desenvolver uma polimerização com grupos C_4 ,

como produto final da reação, na superfície do monolito, contudo não encontramos os reagentes que poderiam proporcionar tal polimerização.

A elevada capacidade de retenção do material monolítico, em 95% de ACN e vazão da fase móvel de $1\mu\text{L min}^{-1}$, permite que este funcione como um *trap* para peptídeos, cujo sua aplicação se torna viável para métodos cromatográficos multidimensionais, tais como HILIC×RP. Adicionalmente, permite a limpeza, dessalinização ou pré-concentração dos peptídeos. Uma segunda abordagem de aplicação do monolito trata-se de um sistema cromatográfico multidimensional, cuja a ponta de vidro para eletronebulização no MS foi empacotada com partículas de C_{18} , esta quando conectada a coluna monolítica, permitiram rápida separação de 7 peptídeos em menos de 12 minutos.

3.4. REFERÊNCIAS

- 1 VAN DEEMTER, J. J.; ZUIDERWEG, F. J.; KLINKENBERG, A. Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. **Chemical Engineering Science**, Oxford, v. 5, p. 271-289, 1956.
- 2 JORGENSEN, J. W. Capillary liquid chromatography at ultrahigh pressures. **Annual Review of Analytical Chemistry**, Palo Alto, v. 3, p. 129-150, 2010.
- 3 MALDANER, L.; SALES, C.; JARDIM, F. O estado da arte da cromatografia líquida de ultra eficiência. **Química Nova**, São Paulo, v. 32, n. 1, p. 214-222, 2009.
- 4 FARIA, A. M.; BOTTOLI, C. B. G.; JARDIM, I. C. S. F.; COLLINS, C. H. Fases estacionárias monolíticas para separações cromatográficas. **Química Nova**, São Paulo, v. 29, n. 2, p. 300-309, 2006.
- 5 JANDERA, P. Advances in the development of organic polymer monolithic columns and their applications in food analysis — A review. **Journal of Chromatography A**, Amsterdam, v. 1313, p. 37-53, 2013.
- 6 LIANG, Y.; ZHANG, L.; ZHANG, Y. Recent advances in monolithic columns for protein and peptide separation by capillary liquid chromatography. **Analytical and Bioanalytical Chemistry**, Heidelberg, v. 405, p. 2095-2106, 2013.
- 7 VAZ, F. A. S.; MOUTINHO, A. D.; MENDONÇA, J. P. R. F.; DE ARAÚJO, R. T.; RIBEIRO, S. J. L.; POLACHINI, F. C.; MESSADDEQ, Y.; DE OLIVEIRA, M. A. L. Optimization of photo-polymerized sol-gel monolithic stationary phases prepared in polyacrylate-coated fused-silica capillaries for capillary electrochromatography. **Microchemical Journal**, Philadelphia, v. 100, n. 1, p. 21-26, 2012.
- 8 NISCHANG, I. Porous polymer monoliths: morphology, porous properties, polymer nanoscale gel structure and their impact on chromatographic performance. **Journal of Chromatography A**, Amsterdam, v. 1287, p. 39-58, 2013.
- 9 VAAST, A.; BROECKHOVEN, K.; DOLMAN, S.; DESMET, G.; EELTINK, S. Comparison of the gradient kinetic performance of silica monolithic capillary columns with columns packed with 3 μm porous and 2.7 μm fused-core silica particles. **Journal of Chromatography A**, Amsterdam, v. 1228, p. 270-275, 2012.
- 10 LI, Y.; LEE, M. L. Biocompatible polymeric monoliths for protein and peptide separations. **Journal of Separation Science**, Weinheim, v. 32, n. 20, p. 3369-3378, 2009.

- 11 WEED, A. K.; DVORNIK, J.; STEFANCIN, J. J.; GYAPONG, A. A.; SVEC, F.; ZAJICKOVA, Z. Photopolymerized organo-silica hybrid monolithic columns: characterization of their performance in capillary liquid chromatography **Journal of Separation Science**, Weinheim, v. 36, n. 2, p. 270-278, 2013.
- 12 HASEGAWA, G.; KANAMORI, K.; ISHIZUKA, N.; NAKANISHI, K. New monolithic capillary columns with well-defined macropores based on poly(styrene-co-divinylbenzene). **ACS Applied Materials and Interfaces**, Washington, v. 4, p. 2343-2347, 2012.
- 13 ARRUA, R. D.; TALEBI, M.; CAUSON, T. J.; HILDER, E. F. Review of recent advances in the preparation of organic polymer monoliths for liquid chromatography of large molecules. **Analytica Chimica Acta**, Amsterdam, v. 738, p. 1-12, 2012.
- 14 EELTINK, S.; GEISER, I.; SVEC, F.; FRÉCHET, J. M. Optimization of the porous structure and polarity of polymethacrylate-based monolithic capillary columns for the LC-MS separation of enzymatic digests. **Journal of Separation Science**, Weinheim, v. 30, n. 17, p. 2814-2820, 2007.
- 15 GHARBHARAN, D.; BRITSCH, D.; SOTO, G.; WEED, A. K.; SVEC, F.; ZAJICKOVA, Z. Tuning preparation conditions towards optimized separation performance of thermally polymerized organo-silica monolithic columns in capillary liquid chromatography. **Journal of Chromatography A**, Amsterdam, v. 1408, p. 101-107, 2015.
- 16 GEISER, L.; EELTINK, S.; SVEC, F.; FRECHÉT, J. M. Stability and repeatability of capillary columns based on porous monoliths of poly(butyl methacrylate-co-ethylene dimethacrylate). **Journal of Chromatography A**, Amsterdam, v. 1140, n. 1-2, p. 140-146, 2007.
- 17 KATO, M.; SAKAI-KATO, K.; TOYO, T.; DULAY, M. T.; QUIRINO, J. P.; BENNETT, B. D.; ZARE, R. N. Effect of preparatory conditions on the performance of photopolymerized sol – gel monoliths for capillary electrochromatography. **Journal of Chromatography A**, Amsterdam, v. 961, n. 1, 45-51, 2002.
- 18 DULAY, M. T.; QUIRINO, J. P.; BENNETT, B. D.; KATO, M.; ZARE, R. N. Photopolymerized sol-gel monoliths for capillary electrochromatography. **Analytical Chemistry**, Washington, v. 73, n. 16, p. 3921-3926, 2001.
- 19 MORISHIMA, K.; BENNETT, B. D.; DULAY, M. T.; QUIRINO, J. P.; ZARE, R. N. Toward sol-gel electrochromatographic separations on a chip. **Journal of Separation Science**, Weinheim, v. 25, n. 15-17, p. 1226-1230, 2002.
- 20 DERRINGER, G.; SUICH, R. Simultaneous-optimization of several response variables. **Journal of Quality Technology**, Milwaukee, v. 12, n. 4, p. 214-219, 1980.

- 21 COLTRO, WENDELL KARLOS TOMAZELLI. **Detecção condutométrica sem contato: uma nova ferramenta para monitoramento de interações biomoleculares em microssistemas analíticos**. 2008. Tese (Doutorado em Química Analítica) – Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2008.
- 22 SEGATO, THIAGO PINOTTI. **Avanços no processo de fabricação de microdispositivos analíticos e em seu acoplamento com a detecção condutométrica sem contato**. 2011. Tese (Doutorado em Química Analítica) – Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2011.
- 23 LIU, Z.; OU, J.; LIN, H.; WANG, H.; LIU, Z.; DONG, J.; ZOU, H. Preparation of monolithic polymer columns with homogeneous structure via photoinitiated thiol-yne click polymerization and their application in separation of small molecules. **Analytical Chemistry**, Washington, v. 86, n. 24, p. 12334-12340, 2014.
- 24 SVEC, F. Porous polymer monoliths: amazingly wide variety of techniques enabling their preparation. **Journal of Chromatography A**, Amsterdam, v. 1217, n. 6, p. 902-924, 2010.
- 25 GILAR, M.; OLIVOVA, P.; DALY, A. E.; GEBLER, J. C. Orthogonality of separation in two-dimensional liquid chromatography. **Analytical Chemistry**, Washington, v. 77, n. 19, p. 6426-6434, 2005.
- 26 GLYGEN CORPORATION COMPANY. **Glysci**. Columbia: Glygen Corporation, 2019. Disponível em: <<http://www.glysci.com/products/Trap-Column.html>> Acesso em: 1 nov. 2019.

Capítulo 4

Lab-on-a-chip

Os conceitos desta ciência permitem a criação de laboratórios miniaturizados, ou seja, transposição da química analítica tradicional de bancada para microdispositivos que cabem na palma da sua mão. Este capítulo faz homenagem ao reverente e entusiasta cientista (dentre os maiores) o qual, visionou a nanotecnologia:

*“Study hard what interests you the most in the
most undisciplined, irreverent and original
manner possible.”*

Richard Feynmann

4.1. INTRODUÇÃO

Desde 1992, verifica-se avanço considerável dos μ chips com princípios eletroforéticos de separação (μ chip-CE). Muitos são os atrativos científicos da técnicas que justificam o seu progresso, tais como: *i*) facilidade em aplicar o potencial no μ chip; *ii*) simplicidade na instrumentação; *iii*) custo reduzido; *iv*) pouco consumo de amostra; *v*) exige menor força de adesão na etapa de selagem na plataforma microfluídica; *vi*) sistema de injeção facilitado; *viii*) análises rápidas e *viii*) ótima eficiência de separação. Infelizmente, em muitos casos, a reprodutibilidade das réplicas perdura como uma limitação da técnica. Adicionalmente, Ramsey J.D. *et al.* tem reportado que canais em forma de serpentina para μ chip-CE não proporcionam campo elétrico contínuo e, portanto, reduzem significativamente a eficiência de separação no dispositivo. [1-6]

Por outro lado, apesar do seu surgimento em 1990 com plataformas de silício, os μ chips regidos por princípios de separação, provenientes da cromatografia líquida (μ chip-LC), se configuram como um progresso científico lento e tardio. Verifica-se, desde o seu surgimento até o ano de 2000, uma total estagnação da técnica. Provavelmente os maiores desafios são: atingir a velocidade ótima da fase móvel, uma vez que a plataforma microfluídica deverá resistir a uma pressão elevada (4.000 bar), devido o empacotamento de partículas (< 3.0 μ m i.d.) no microcanal. Além disso, exige da necessidade de *frit*, para restrição das partículas a qual se configura como habilidade dispendiosa e árdua, além da difícil habilidade de se empacotar partículas em microcanais nos chips. [1,3,5,6]

Entretanto, o μ chip-LC apresenta-se como um grande atrativo a ciência de separações, devido a teoria consolidada e abrangente da LC, muito bem fundamentada e simples, a elevada reprodutibilidade análise após análise, robustez, precisão, exatidão e facilidade no acoplamento ao espectrômetro de massas, uma vez que não há necessidade de isolar ou controlar as magnitudes dos potenciais, como ocorre no CE-MS.

Na busca de plataformas microfluídicas robustas, em termos da elevada resistência a pressão interna do microcanal, pesquisadores têm investigado o uso de titânio e aço inoxidável, confeccionados por meio de impressora 3D de fusão a laser seletiva. Aspecto que certamente encarece e dificulta a microfabricação. [7]

Na última década o μ chip-LC tem se tornado comum nos laboratórios de análise e crescente número de publicações têm sido relatado devido o surgimento do *HPLC-chip/MS* da *Agilent Technologies* no mercado. [8] O qual é estruturado por uma microválvula e uma dimensão cromatográfica com empacotamento de partículas de 3 μ m diâmetro interno (d.i.), exigindo assim de uma plataforma altamente robusta a pressão. Recentemente, a *Waters Corporation* iniciou a comercialização do *ionKey/MS* cuja a dimensão do canal microfluídico é de 150 μ m de largura com partículas de 2 μ m de d.i. e canal microfluídico adicional para melhorar a ionização ao ESI-MS por meio do auxílio de solvente apropriado. [9]

Uma alternativa interessante às limitações do μ chip-LC, principalmente em termos da resistência a elevadas pressões, são os monolitos produzidos *in-situ* no canal microfluídico. Pois, os monolitos respeitam todas as vantagens da cromatografia líquida, acrescidas da elevada permeabilidade do material, permitindo sua condução em plataformas microfluídicas menos robustas, ou seja, menor resistência à pressão. Adicionalmente, a eficiência cromatográfica não é prejudicada, devido à elevada transferência de massa do analito com a fase estacionária. Além disso, não é necessário a utilização de *frit*, uma vez que o material é ligado quimicamente a parede do canal. Contudo, a grande limitação do monolito é sua reprodutibilidade, uma vez que o monolito é inerente a variações da síntese. [5]

Recentemente, Grinias e Kennedy revisaram os avanços do μ chip-LC. [5] O que chamou a atenção foi que ao considerar a **Tabela 1** do artigo em questão, verifica-se que das técnicas disponíveis para a confecção do μ chip-LC, a qual apresenta maior número de vantagens e menor número de desvantagens é o μ chip-LC com meio monolítico de separação.

Com estas considerações, e na busca de simplicidade e rapidez na microfabricação, nós planejamos o μ chip multidimensional, conforme representado pela **Figura 50**. Nesta configuração, a separação é realizada essencialmente por cromatografia líquida, cujo monolito estará preenchido ao longo de todo o canal microfluídico. A ortogonalidade do sistema será provida pela alteração do pH durante a separação, ou seja, os analitos serão eluídos em pH básico na 1D numa distância de canal/monolito de 14,26 cm e conseguinte alteração para pH ácido na 2D e eluição numa distância de canal/monolito de 10,63 cm. A alteração de pH mudará a conformação da proteína/peptídeo permitindo separações independentes em cada dimensão cromatográfica, garantido assim a ortogonalidade do sistema.

Figura 50 - Ideia do μ chip com separação cromatográfica multidimensional. A) Injeção de amostra pelo nano-LC e eluição dos analitos em meio básico. B) Auxílio de uma segunda bomba cromatográfica para eluição dos analitos em meio ácido. C) Eletronebulização para o espectrômetro de massas, por meio de emissor de capilar de sílica.

Fonte: Autoria própria.

Algumas das vantagens e simplicidades na microfabricação deste sistema microfluídico são: *i*) as dimensões cromatográficas possuem compatibilidade, uma vez que será utilizado o mesmo solvente na 1D e na 2D; *ii*) consequentemente não será necessário a utilização de válvulas ou divisão do fluxo (*split*) da fase móvel entre as dimensões cromatográficas; *iii*) não será necessário a aplicação do potencial para a separação eletroforética na 2D, assim como isolamento deste potencial para o acoplamento ao espectrômetro de massas, uma vez que este interfere na etapa de eletronebulização e *iv*) segundo a teoria fundamentada por Gilar *et al.* a mudança de pH entre as dimensões cromatográficas se configura como separação altamente ortogonal para biomoléculas. [10] Contudo, um sistema miniaturizado para cromatografia líquida proverá de maior resistência a pressões elevadas. Desta forma, um dos desafios será obter uma plataforma microfluídica e conexões que resistam pelo menos 200 bar. [5]

É interessante mencionar que a integração da fonte ESI ao μ chip de vidro se caracteriza como a etapa mais desafiadora na microfabricação do μ chip-MS, fato que tem despertado o

interesse de inúmeros pesquisadores. [11] Uma vez que, vários parâmetros devem ser considerados para a eletronebulização eficiente, de modo que os fatores mais críticos são o diâmetro do emissor do *spray* e a vazão. [4] Não obstante, o formato e a propriedade físico-química do material a proporcionar a emissão do *spray*, devem ser levados em consideração na microfabricação. [12] Além disso, por se tratar de *nanofluxo*, o desempenho da eletronebulização é sensível a pequenas variações. Neste âmbito, os *μchip* de vidro têm se consolidado como a melhor escolha devido a sua inércia e transparência ótica. Contudo a projeção de microcanais pontiagudos com baixa área de emissão do *spray* é o maior desafio. [2]

Os primeiros *μchip* de vidro com ESI reportados na literatura mostraram-se ineficientes na etapa de ionização devido ao elevado grau de espalhamento do fluído sobre a superfície do vidro, gerando baixa área de emissão do *spray*. Consequente houve a formação de gotas com elevado diâmetro e eletronebulização prejudicada. De modo que emissores de *spray* hidrofóbicos não permitem o espalhamento da gota sobre a superfície do vidro e podem ser conduzidos por um processo chamado de molhamento eletrolítico (*electrowetting*). Contudo, o volume associado a estas gotas proporcionam alargamento de banda ao cromatograma de íons e requer aplicação de potencial elevado. A estratégia de junção líquida também foi muito investigada, com acoplamento em emissores de *nanospray* para *μchip*. De modo que os resultados se demonstraram ótimos em termos de sensibilidade. Contudo, os problemas foram o volume morto na ponta da eletronebulização e a perda de junção líquida devido a aplicação do potencial requerido ao ESI-MS. [1,2,4]

Várias estratégias engenhosas, por meio da manipulação controlada de materiais, têm sido publicadas nos últimos anos. Como a microfabricação de bicos em superfícies de pastilhas de silício, com diâmetros menores a 15 μm . [11] A microfabricação de um emissor por meio do esmagamento de um cone dentro do canal (com o auxílio de uma máquina CNC). Sendo assim, o cone foi alinhado ao centro e na extremidade do microcanal, este foi puxado para a formação de uma ponta, utilizando um puxador de platina aquecido (*home-made*). [12] Outra estratégia foi a microfabricação da ponta de emissão realizada em um dos cantos pontiagudos do *μchip* e afiada até atingir 300 μm , formando assim uma ponta de vidro integrada para eletronebulização. [2] Além disso, diferentes polímeros têm sido investigados como plataformas de emissão, tais como:

polidimetilsiloxano (PDMS), polimetilmetacrilato (PMMA), poliimida (PI), policarbonato (PC) e copolímero de olefina cíclica (COC). [11]

Em termos de ótima sensibilidade e estabilidade do *spray* os emissores com capilares de sílica fundida acoplados ao μ chip mostram-se muito adequado devido a geração de microgotículas e ótima sensibilidade. Contudo o principal problema é o volume morto associado ao acoplamento μ chip-capilar e, conseqüentemente, a microfabricação minuciosa e precisa. [2,11] Diante de todas estas considerações, optamos conduzir a eletronebulização por meio da utilização de capilares de sílica fundida, uma vez que este possui fácil adaptação ao *nanospray FLEXTM Ion Sources* da *Thermo Scientific*.

4.2. OBJETIVOS

Desenvolver um *μchip* na modalidade cromatografia líquida multidimensional acoplada ao espectrômetro de massas sequencial para análises proteômicas.

4.2.1. Objetivos específicos

- 1) Planejar e microfabricar canais microfluídicos para separações multidimensionais.
- 2) Promover a selagem vidro/vidro das plataformas microfluídicas.
- 3) Realizar a caracterização dos microcanais, por meio do perfilômetro e microscopia ótica.
- 4) Elaborar e aplicar estratégias para a conexão capilar do *μchip* aos equipamentos: nano-LC e espectrômetro de massas.
- 5) Sintetizar o monolito híbrido no canal microfluídico para separação em cromatografia líquida multidimensional.

4.3. PARTE EXPERIMENTAL

4.3.1. Substratos e reagentes

As lâminas de vidro para microscopia (26 × 76 mm de tamanho e 1,0 a 1,2 mm de espessura) foram adquiridas da empresa Perfecta, São Paulo – SP. O fotolito foi fabricado pela empresa DGM Design, Curitiba – PR. O resiste positivo AZ 4620; a solução reveladora de potássio de borato aquoso (AZ[®] 400K); acetona; ácido clorídrico (HCl) e ácido fluorídrico (HF) foram adquiridas da *Clariant Corporation*, Somerville, NJ, USA. O hexametildisilazano (HMDS); xilol; fluoreto de amônio (NH₄F) foram adquiridos da Sigma-Aldrich, São Paulo-SP. A solução reveladora de cromo (Chrome ETCH N°1), foi adquirida da Technic, França. As brocas diamantadas para uso odontológico foram adquiridas no comércio local de São Carlos-SP. O ácido sulfúrico (H₂SO₄), assim como o peróxido de hidrogênio (H₂O₂) foram adquiridos da Synth.

4.3.2. Equipamentos

- Chapa aquecedora TE-038/2-MP da Tecnal
- Agitador Magnético da Fisatom
- Fotoalinhadora da Karl Suss America Inc. MJB 3 UV 200 com potência de 10 mW cm⁻²
- Microscópio ótico da Micronal
- Microrretífica Dremel[®] MultiPro 395JU
- Mufla EDG modelo F3000
- Millipore[®] (Billerica, MA, USA)
- Veeco modelo Dektak 150

4.3.3. Fabricação dos microcraniais em vidro

Os experimentos deste tópico foram conduzidos no Laboratório de Microfabricação do LNNano (Laboratório Nacional de Nanotecnologia Brasileiro) do CNPEM (Centro Nacional de Pesquisa em Energias e Materiais), em Campinas – SP.

1. As lâminas de vidro foram lavadas com solução piranha (H_2SO_4 e H_2O_2 ; 2:1, v/v) e secadas em chapa aquecedora.
2. O desenho dos microcanais foi gerado por meio de uma máscara de campo escuro (fotolito).
3. Por meio da técnica de deposição metálica por *sputtering*, realizou-se o depósito de uma fina camada de cobre sobre as lâminas de vidro.
4. Depositou-se sobre o cobre uma fina camada da solução de xilol contendo 20% de hexametildisilazano (HMDS).
5. Em seguida, depositou-se o fotorresiste positivo (AZ 4620) por *spinning* a 3.000 rpm durante 30 s, sobre a camada de cobre contida na plataforma de vidro.
6. A cura do fotorresiste foi realizada a 95 °C por 10 min em chapa térmica.
7. A máscara do fotolito foi sobreposta na lâmina de vidro contendo cobre e fotorresiste, este conjunto foi exposto à radiação UV por 90 s por meio de uma fotoalinhadora.
8. Removeu-se as partes solúveis do fotorresiste AZ 4620, por meio da imersão das lâminas de vidros tratadas, em solução reveladora de potássio de borato aquoso (AZ[®] 400K), durante 40 s sob agitação.
9. A etapa de pós cura foi realizada a 120 °C por 5 min em chapa aquecedora.
10. O cromo contido no microcanal, após gravação, foi retirado por meio da imersão das plataformas de vidro em solução reveladora de cromo (Chrome ETCH N°1, TECHNIC) à 50°C.
11. Verificou-se a revelação do cromo foi eficiente nos microcanais por meio de microscopia ótica.
12. Em seguida, secou-se as plataformas de vidros tratadas, em chapa aquecedora em 76 °C.
13. Posteriormente, imergiu as plataformas de vidro tratadas, em solução contendo: *i*) $\text{HF}:\text{NH}_4\text{F}:\text{HCl}$ na proporção de 1:41:1 (HF 49%, NH_4F 1,38 mol L⁻¹ e HCl 38%) e *ii*) HF e H_2O 1:1, durante 5 min.
14. Lavou-se as plataformas de vidro em acetona.

15. Os orifícios das plataformas de vidro, as quais se referem ao topo do μ chip, foram realizadas com auxílio de uma microrretífica composta de broca diamantadas de uso odontológico.

4.3.4. Selagem térmica vidro/vidro do μ chip

Este protocolo foi desenvolvido pela ex-integrante do grupo BioMicS Dra. Alessandra Maffei Monteiro.

4.3.4.1. Instrumentação e materiais

O processo de selagem térmica das plataformas de vidro foi realizado com o auxílio de uma barra de ferro (1,6 kg), adquirida na Antari Comércio de Metais, e duas placas de alumina 1 (9,0 x 9,0 x 2,0 cm) produzidas sob encomenda pela Engecer (São Carlos, SP, Brasil). Estas foram polidas em rebolo de diamante para que irregularidades não fossem impressas na superfície externa dos μ chips, durante a etapa de selagem térmica. Empregou-se mufla e água ultrapura obtida de um sistema Millipore®, conforme ilustrado pela **Figura 51**.

4.3.4.2. Etapa prévia da selagem

- 1) Todas as plataformas de vidro foram lavadas em solução piranha (H_2SO_4 e H_2O_2 ; 2:1, v/v).
- 2) Duas plataformas de vidro contendo *i*) microcanais e *ii*) orifícios foram lavadas com etanol.
- 3) Com o auxílio de uma bacia contendo água ultrapura, ambas as plataformas de vidro (topo e microcanal) foram submersas em água para alinhamento das partes. Este procedimento garante maior aderência e evita que partículas de poeira penetrem na região de contato entre as lâminas de vidro.
- 4) Após colocados em contato, as plataformas de vidro foram dispostas entre duas placas polidas de alumina e encaminhados a uma mufla. Para tal, pressionou-se as plataformas de vidro contra uma das placas de alumina, promovendo a aderência da plataforma à alumina. Colocou-se a segunda placa de alumina de maneira alinhada sobre a plataformas de vidro e levou-se cuidadosamente para a mufla. Adicionou-se posteriormente o bloco de ferro sobre a

estrutura, conforme representado pela **Figura 51B**. Quando houve o deslocamento dos substratos vítreos na estrutura, reiniciou-se todo o procedimento, de modo a garantir selagem simétrica entre as plataformas vítreas.

Figura 51 - Processo de selagem térmica vidro/vidro para confecção dos μ chips multidimensionais. A) Instrumentação necessária para a execução do processo de selagem térmica. B) Disposição dos blocos de alumina e ferro no interior da mufla.

Fonte: Autoria própria.

4.3.4.3. Programação de temperatura

A mufla foi programada para taxa de aquecimento inicial de 10°C/min até atingir 550°C, durante 60 min. Posteriormente, a temperatura foi elevada a 590°C a uma taxa de aquecimento de 2°C/min, durante 900 min. Terminado o processo, a mufla foi mantida fechada até que se atingisse a temperatura ambiente. Como resultado obteve-se μ chip com a arquitetura dos microcanais.

4.4. Resultados e Discussões

4.4.1. Microfabricação do μ chip multidimensional

Os microcanais em plataformas de vidro foram fabricadas por fotolitografia e corrosão química por via úmida. O desenho destes foi gerada a partir de uma máscara de campo escuro (**Figura 52**), a qual tem a função de permitir radiação de luz e conseguinte revelação do fotorresiste na região do vidro a qual será corroída pelo ácido fluorídrico (HF), desta forma tem-se a criação do canal microfluídico. É interessante mencionar que as dimensões desejadas do canal microfluídico não dependem somente do desenho elaborado na máscara, mas deve-se ter em mente o processo de corrosão isotrópico por via úmida do HF em vidro. Motivo pelo qual, aplica-se uma fina camada de cobre de modo a amenizar a corrosão do HF nas bordas dos microcanais. Com estas considerações e com o objetivo de criar microcanais com dimensões próximas ao diâmetro interno das colunas capilares, ou seja, secção de área de $4.417,86 \mu\text{m}^2$, foram consultados os Pesquisadores Renato Sousa Lima, Maria Helena e Angelo Gobbi do Laboratório Nacional de Nanotecnologia – CNPEM, onde atingiu-se o consenso que a espessura do canal na máscara deveria ser de $80 \mu\text{m}$ com tempo de corrosão de 5 min.

Figura 52 - Projeto do software Autocad para a confecção da máscara de fotolito para a fotogração dos canais microfluídicos em plataformas de vidro, referente a microfabricação do μ chip-LC×LC.

Fonte: Autoria própria

As dimensões de secção de área $4.417,86 \mu\text{m}^2$, assim como, comprimento total de aproximadamente 25 cm do canal, sendo 14,26 cm para 1D e 10,63 cm para a 2D foram

planejados de modo a adaptar as otimizações realizadas em capilares para os sistemas microfluídicos, com atenção especial aos quesitos: *i*) diminuir as chances de entupimento, seja na polimerização do monolito ou amostral; *ii*) diminuir a ação de fenômenos difusivos dos analitos para sistema multidimensionais de separação com longo canal microfluídico, *iii*) pressão adequada ao sistema com dimensões de canal adequado e *iv*) análises rápidas na 2D de modo a não perder a resolução obtida na 1D.

Após a criação dos microcanais estes foram medidos no perfilômetro (*Veeco* modelo *Dektak 150*) de modo a verificar a homogeneidade de corrosão ao longo do canal, assim como sua dimensão final. Desta forma, selecionou-se três regiões distintas do canal microfluídico de cada plataforma, com total de 6 plataformas de vidro ($n = 18$). Como o processo de corrosão é isotrópico, decidiu-se medir três regiões aleatórias dentro do canal (topo, centro e fundo). Em seguida realizou-se os cálculos da secção de área, cujo valor foi obtido pelo produto múltiplo da profundidade com a média da largura do canal referentes aos valores do topo, meio e fundo, conforme reportados na **Tabela 17**. Verifica-se que a secção de área do canal microfluídico foi próximo ao valor desejado ($4.417,9 \mu\text{m}^2$) com média e desvio padrão (DP) de $4.352,0 \mu\text{m}^2 \pm 358,1$, respectivamente. A **Figura 53** mostra o imageamento de uma parte do microcanal processado no perfilômetro. Verifica-se que os microcanais possuem profundidade homogênea ao longo do seguimento.

Em seguida as plataformas de vidro foram submetidas a selagem térmica vidro/vidro de modo a criar de fato o canal microfluídico. Nesta etapa é importante verificar se houve a formação de anéis de Newton (bolhas de ar inseridas entre as placas de vidro) e se a selagem formou de fato um canal com segmento retilíneo nas superfícies seladas. A **Figura 54** mostra as plataformas microfluídicas de vidro seladas e submetidas com vazão de um corante azul solúvel em água. Verifica-se que a selagem foi eficiente pois não houve a formação de anéis de Newton, assim como, o corante azul indica que não há vazamentos ao longo do canal. Entretanto, verifica-se algumas saliências nas paredes dos canais. Estas fazem referência, principalmente a qualidade da máscara (fotolito). Uma vez que ao observar a máscara de fotolito, verifica tais imperfeições e talvez esta seja uma limitação do próprio processo de fabricação da mesma ou do material (fotolito). Outra possibilidade seria um excesso na retirada do cromo, o qual é depositado na plataforma de vidro na etapa de microfabricação de modo a amenizar a corrosão isotrópica do HF

nas extremidades laterais do microcanal, contudo, foi tomado muita cautela para este quesito. Em suma, acreditamos que estas pequenas imperfeições não sejam um agravante considerável para o sucesso de uma separação LC×LC.

Tabela 17 - Medidas e dimensões do canal microfluídico multidimensional obtido no perfilômetro

	Profundidade (μm)	Dimensões da largura do canal microfluídico				Secção de área (μm^2)
		Topo (μm)	Centro (μm)	Fundo (μm)	Média (μm)	
1a	36,10	147,9	-	91,3	119,6	4.454,5
1b	36,68	139,7	-	96,8	118,3	
1c	37,54	153,4	-	97,7	125,6	
Média	36,8	147,0	-	95,3	121,1	
DP	$\pm 0,7$	$\pm 6,9$	-	$\pm 3,5$	$\pm 3,5$	
2a	38,10	144,3	113,12	105,0	120,8	4.316,0
2b	37,07	150,0	117,8	105,9	124,6	
2c	32,68	129,7	104,10	110,5	114,8	
Média	36,0	141,3	111,7	107,1	120,1	
DP	$\pm 2,9$	$\pm 10,5$	$\pm 7,0$	$\pm 3,0$	$\pm 4,8$	
3a	36,10	137,90	102,3	92,2	110,8	3.992,6
3b	36,10	132,40	105,0	91,3	109,6	
3c	36,11	134,20	105,0	95,0	111,4	
Média	36,1	134,8	104,1	92,8	110,6	
DP	$\pm 0,0$	$\pm 2,8$	$\pm 1,6$	$\pm 1,9$	$\pm 0,9$	
4a	41,03	147,0	122,4	106,8	125,4	4.882,9
4b	39,88	146,1	116,9	95,0	119,3	
4c	39,81	139,7	118,7	99,50	119,3	
Média	40,2	144,3	119,3	100,4	121,3	
DP	$\pm 0,7$	$\pm 4,0$	$\pm 2,8$	$\pm 6,0$	$\pm 3,5$	
5a	37,80	149,80	120,50	97,70	122,7	4.541,4
5b	37,17	176,30	116,90	98,60	130,6	
5c	36,46	143,4	106,8	90,4	113,5	
Média	37,1	156,5	114,7	95,6	122,3	
DP	$\pm 0,7$	$\pm 17,4$	$\pm 7,1$	$\pm 4,5$	$\pm 8,5$	
6a	39,30	139,70	112,3	89,5	113,8	3.924,6
6b	36,67	133,3	106,8	94,10	111,4	
6c	31,27	125,10	100,50	86,80	104,1	
Média	35,7	132,7	106,5	90,1	109,8	
DP	$\pm 4,1$	$\pm 7,3$	$\pm 5,9$	$\pm 3,7$	$\pm 5,0$	
					Média DP	4.352,0 $\pm 358,1$

Figura 53 - Imagens dos canais (azul) obtidas do perfilômetro. A) Imagem 2D dos microcanais e B) Imagem 3D com a profundidade dos microcanais, quanto maior a tonalidade de azul maior a profundidade.

Fonte: Autoria própria.

Figura 54 - Diferentes imagens de microscopia óptica (4× de magnitude) ao longo dos microcanais microfabricados e selados. Estes foram submetidos a preenchimento de um corante azul em água (50:50) à vazão de $0,20 \mu\text{L min}^{-1}$ com auxílio de uma bomba seringa.

Fonte: Autoria própria.

Por se tratar de um sistema microfluídico multidimensional com princípios de separação cromatográficos, torna-se necessário o acoplamento do *μchip* ao *nano-LC* assim como ao espectrômetro de massas para a eletronebulização. Desta forma, desenvolvemos junto a oficina

mecânica suportes de aço-inox e conectores rosqueáveis ao suporte, conforme mostrados na **Figura 55**.

Figura 55 - Visualização em diferentes ângulos do μ chip multidimensional com suporte de aço-inox. 1) Acoplamento para o nano-LC; 2) Acoplamento em bomba cromatográfica para alterar o pH do meio microfluídico; 3) Canal microfluídico para separação multidimensional e 4) acoplamento para eletronebulização ao espectrômetro de massas.

Fonte: Autoria própria.

O suporte de aço-inox é reutilizável e possui compatibilidade para acoplamento ao espectrômetro de massas. Além do mais, estes suportes foram úteis para o preenchimento do microcanal com a solução sol-gel para a polimerização do monolito (sem formação de bolhas).

Em seguida, realizou-se a polimerização dos monolitos *in-situ* nos microcanais. Por meio da experiência obtida na polimerização dos monolitos em capilares, adotou-se que os fatores significativos na otimização seriam a concentração de %MPTMS/Tolueno e HCl. Desta forma, aplicou-se um planejamento fatorial completo $2^4 = 16$ experimentos. Contudo, verificamos que após a polimerização, na etapa de limpeza do microcanal, os microcanais da **Figura** ocasionaram entupimento em todos os 16 μ chips. Portanto, trata-se de um fenômeno que deve ser levado em consideração para posteriores etapas de microfabricação. Perante a este desafio, o μ chip multidimensional para proteômica não obteve sucesso até o momento.

4.5. CONCLUSÕES

As estratégias de planejamento e caracterização do *μchip* 2D LC mostraram-se promissoras para as aplicações da cromatografia líquida multidimensional. Foram desenvolvidos conectores para o preenchimento da solução sol-gel no canal, assim como para o acoplamento ao espectrômetro de massas. Apesar destes conectores mostrarem-se resistentes a uma pressão de 200 bar, apresentando-se compatíveis com a permeabilidade do monolito, os mesmos possuem a desvantagem de conter volume morto e devem ser considerados em futuros estudos. A caracterização dos microcanais com 25 cm de comprimento mostraram-se adequados devido a homogeneidade na corrosão realizada pelo HF e selagem sem a presença de anéis de Newton. As dimensões e estratégias planejadas para o acoplamento ao nano-LC e ao ESI-MS se mostraram adequados. Contudo, o entupimento inevitável nos canais após a polimerização do monolito não permitiu o sucesso da aplicação até o momento.

4.6. REFERÊNCIAS

- 1 RAMSEY, J. D.; JACOBSON, S. C.; CULBERTSON, C. T.; RAMSEY, J. M. High-efficiency, two-dimensional separations of protein digests on microfluidic devices. **Analytical Chemistry**, Washington, v. 75, n. 15, p. 3758-3764, 2003.
- 2 MELLORS, J. S.; GORBOUNOV, V.; RAMSEY, R. S.; RAMSEY, J. M. Fully integrated glass microfluidic device for performing high-efficiency capillary electrophoresis and electrospray ionization mass spectrometry. **Analytical Chemistry**, Washington, v. 80, n. 18, p. 6881-6887, 2008.
- 3 LIU, J.; CHEN, C.; TSAO, C.; CHANG, C.; CHU, C.; DEVOE, D. L. Polymer microchips integrating solid-phase extraction and high-performance liquid chromatography using reversed-phase polymethacrylate monoliths. **Analytical Chemistry**, Washington, v. 81, n. 7, p. 2545-2554, 2009.
- 4 KOSTER, S.; VERPOORTE, E. A decade of microfluidic analysis coupled with electrospray mass spectrometry: an overview. **Lab on a Chip**, Cambridge, v. 7, p. 1394-1412, 2007.
- 5 GRINIAS, J. P.; KENNEDY, R. T. Advances in and prospects of microchip liquid chromatography. **Trends in Analytical Chemistry**, Amsterdam, v. 81, p. 110-117, 2016.
- 6 COLTRO, W. K. T.; DE JESUS, D. P.; DA SILVA, J. A. F.; DO LAGO, C. L.; CARRILHO, E. Toner and paper-based fabrication techniques for microfluidic applications. **Electrophoresis**, Weinheim, v. 31, n. 15, p. 2487-2498, 2010.
- 7 SANDRON, S.; HEERY, B.; GUPTA, V.; COLLINS, D. A.; NESTERENKO, E. P.; NESTERENKO, P. N.; TALEBI, M.; BEIRNE, S.; THOMPSON, F.; WALLACE, G. G.; BRABAZON, D.; REGAN, F.; PAULL, B. 3D printed metal columns for capillary liquid chromatography. **Analyst**, Cambridge, v. 139, n. 24, p. 6343-6347, 2014.
- 8 AGILENT TECHNOLOGIES. **Infinitely better for microfluidic nanospray LC/MS**. 2010. Santa Clara: Agilent Technologies. Disponível em:
<<https://www.agilent.com/cs/library/brochures/5990-6221EN.pdf>> Acesso em: 13 ago. 2017.
- 9 WATERS CORPORATION. **Enhanced MS with the turn of a key experience the future of MS, experience ionKey/MS**. 2014. Milford: Waters Corporation. Disponível em:
<<http://www.waters.com/webassets/cms/library/docs/720004662en.pdf>> Acesso em: 13 ago. 2017.

10 GILAR, M.; OLIVOVA, P.; DALY, A. E.; GEBLER, J. C. Orthogonality of separation in two-dimensional liquid chromatography. **Analytical Chemistry**, Washington, v. 77, n. 19, p. 6426-6434, 2005.

11 FENG, X.; LIU, B. F; LI, J.; LIU, X. Advances in coupling microfluidic chips to mass spectrometry. **Mass Spectrometry Reviews**, Hoboken, v. 34, n. 5, p. 535-557, 2015.

12 HOFFMANN, P.; HÄUSIG, U.; SCHULZE, P.; BELDER, D. Microfluidic glass chips with an integrated nanospray emitter for coupling to a mass spectrometer. **Angewandte Chemie**, Weinheim, v. 46, p. 4913-4916, 2007.

Capítulo 5

Uma Nova Equação Para a Ciência de Separações

Apesar de ser uma ciência majotariamente experimental, os maiores cientista do campo eram téóricos (Dr. Archer J. P. Martin, Dr. John Calvin Giddings, Dr. Georges Guiochon, Dr. Csaba Horváth e outros). Este capítulo possui dedicatória especial, será para o brasileiro que descobriu experimentalmente o méson pi, como consequência, ele quase permitiu o nosso primeiro Nobel.

“Tudo o que eu fiz não vale uma sinfonia de Mozart”

César Lattes

5.1. INTRODUÇÃO

5.1.1. A cromatografia líquida multidimensional com parada de fluxo na 1D ocasiona alargamentos de bandas?

Há uma forte ideia na comunidade de ciência de separação que os métodos cromatográficos multidimensionais com parada de fluxo na 1D podem ocasionar o alargamento de banda. Esta afirmação é fortalecida pelo equacionamento da obtenção da melhor eficiência de separação, em termos da menor altura do prato (H), prevista pela equação de *van Deemter*, mais especificamente pelo termo B da equação, a qual trata do fenômeno de difusão longitudinal das moléculas (analitos) ao longo da coluna cromatográfica. Segundo a referida equação, tempo longo ou lento de separação ou interrupção da vazão no processo de eluição gera a difusão das moléculas. Fato que é muito compreensível, devido às regras da segunda lei da termodinâmica, cujo aumento da desordem (entropia) é favorável até o nível máximo do equilíbrio. [1-3]

Contudo, a equação de *van Deemter*, em termos entrópicos, não se restringe somente a difusão longitudinal. A contribuição do termo C da equação, ou seja, a transferência de massa dos analitos seja com a fase estacionária (C_s) ou com a fase móvel (C_m) tem a sua magnitude na forma do pico cromatográfico e, portanto, em termos práticos, com a diminuição ou aumento do alargamento de banda. Assim como o termo A (múltiplos caminhos ou difusão de Eddy), neste caso espera-se que as partículas sejam uniformes e pequenas para que não ocorra a difusão do(s) analito(s) no meio cromatográfico.

O grupo do Professor Janssen, H-G. provavelmente foi o pioneiro em declarar de forma convincente que o alargamento de banda na modalidade cromatografia líquida multidimensional com parada de fluxo pode estar incorreto. [4] Entretanto, os estudos foram baseados em uma coluna de exclusão por tamanho (SEC) sendo esta a 1D e coluna hidrofóbica como 2D. É interessante mencionar que apesar da SEC permitir separações provenientes de grandezas físicas devido ao tamanho diferenciado dos poros na fase estacionária, os quais analogamente “peneiram” os analitos (obs: este termo tem sido usado por *J.C. Giddings*), esta coluna também possui mecanismos de separação físico-químicos devido às interações do analito com a fase estacionária. [5] Assim, os experimentos foram conduzidos por meio da comparação do fluxo

contínuo e com parada de fluxo da molécula (peptídeo) Ala-Val na 1D, este foi escolhido devido ao seu elevado coeficiente de difusão. Os cálculos foram derivados a partir da equação de *van Deemter*. [1]

Desta forma, verificou-se que após 8 horas de análise, o total de variância do desvio padrão (largura do pico) do método com parada de fluxo (σ^2_{tpf}) aumentou somente 0,10 vezes em relação ao total de variância do desvio padrão do método com fluxo contínuo (σ^2_{tfc}). Como conclusão, os autores afirmam que o método com parada de fluxo não ocasiona o alargamento de banda. Esta afirmação é extremamente importante devido um conceito um tanto quanto estabelecido, cujos métodos cromatográficos multidimensionais abrangentes devam possuir uma coluna menor na 2D para que a análise seja rápida de modo a não comprometer com a resolução obtida na 1D, em termos do alargamento de banda. Fato que diminui drasticamente a capacidade de pico da 2D e, portanto, do método LC×LC por completo. No entanto, é óbvio que um método com parada de fluxo irá ter um tempo de análise drasticamente superior em relação com o método 2D de fluxo contínuo. Neste âmbito, o pesquisador deverá pensar qual aspecto é mais importante, a resolução ou o tempo de análise.

Em contrapartida, o grupo do Professor T. Greibrokk publicou no mesmo ano o artigo a respeito da influência do alargamento de banda de peptídeos em cromatografia multidimensional com parada de fluxo, [6] na qual a 1D foi uma coluna em escala micro de fase estacionária hidrofóbica (C₁₈). O grupo menciona que a largura de banda depende do tamanho das moléculas analisadas, assim como a natureza da interação química do analito com a fase estacionária. Em especial foi mencionado que o peptídeo bradicinina (Brad) por possuir maior massa molecular e interação eletrostática mais forte com a fase estacionária em relação ao peptídeo leucine-encefalina (Leu-en) faz com que Brad tenha menor alargamento de banda. Fato que é muito coerente com os princípios das interações químicas, assim como a baixa dispersão de objetos com maior massa.

Em suma, as conclusões referentes a este artigo mostraram que ao comparar o método sem parada de fluxo com a modalidade com parada de fluxo de 6 horas na 1D, houve aumento na largura de banda do peptídeo Brad de 26% em temperatura de 20 °C e aumento de 49% em temperatura de 40 °C, respectivamente. Em relação ao peptídeo Leu-en houve aumento de 2,2 para 2,5 vezes maiores em relação a Brad, para as respectivas temperaturas. Conforme os

próprios autores mencionaram é óbvio que o aumento de temperatura ocasiona aumento da largura de banda, se mantida uma vazão constante e se o processo de transferência de massa for ineficiente, uma vez que o aumento de temperatura proporciona aumento da energia cinética das moléculas e, portanto aumento dos processos de transferência de massa, se tal processo for eficiente provavelmente não acarretará em drástico alargamento de banda. Consequentemente, não iremos expandir tais pressupostos os quais não são o objetivo deste estudo. O grupo em questão também concluiu que a utilização do gradiente em vez da condição isocrática diminui o efeito de alargamento de banda para a coluna em modo de fase reversa com partículas de C₁₈.

Em contrapartida, A. P. Köhne, U. Dornberger e T. Welsch realizaram a separação de amostras complexas de amino- e nitro-orgânicos por meio da modalidade 2D com parada de fluxo, cuja a primeira dimensão foi uma coluna hidrofóbica de C₈ e segunda dimensão com coluna hidrofóbica C₁₈. O estudo mostrou que após a parada do fluxo de 30 minutos na primeira dimensão cromatográfica utilizando 60% de metanol e 40% de água não houve alargamento de banda para os analitos 2,4-dinitrotolueno e 2,4,6-trinitrotolueno. [7]

Com as informações descritas nos três trabalhos mencionados nos parágrafos anteriores, podemos deduzir teoricamente que o alargamento de banda na modalidade 2D com parada de fluxo depende de três pontos principais: *i*) da magnitude da interação química dos analitos com a fase estacionária, ou seja, quanto mais forte a interação, menor é o alargamento de banda, o qual está diretamente relacionado com; *ii*) o coeficiente de difusão das moléculas no meio cromatográfico e *iii*) o tempo de análise.

Diante de tais evidências e ainda diante da leitura do artigo intitulado: “*Theory for Electrostatic Interaction Chromatography of Proteins*” de autoria de Jan Ståhlberg, Bengt Jonsson e Csaba Horváth. Surgiu uma ótima oportunidade para discutir tais fenômenos termodinâmicos em colunas de troca catiônica, os quais podem contribuir com a teoria da ciência de separação, principalmente com as modalidades cromatográficas multidimensionais com parada de fluxo na primeira dimensão. [8]

5.2. OBJETIVO

Elaborar uma equação para previsão do alargamento de bandas provenientes de colunas cromatográficas de troca iônica.

5.3. RESULTADOS

5.3.1. Elaboração de uma equação termodinâmica para a previsão do alargamento de banda em separações cromatográficas provenientes de colunas de troca catiônica

O grupo do Professor Csaba Horváth mostrou a linearidade do fator de retenção (k) de proteínas intactas com o aumento da concentração de sal no processo de eluição da coluna SCX. [8] Desta forma, foi obtido correlações positivas de $\ln k$ vs $1/\sqrt{I}$, onde I é a força iônica do meio. Com a proposta de entender o fenômeno, em termos termodinâmicos, do baixo alargamento de pico na 1D de métodos cromatográficos com parada de fluxo, realizamos correlações matemáticas por meio de interpretações provenientes de outros estudos reportados na literatura e da consolidada teoria da ciência de separação. [5,9] Assim podemos introduzir a **Equação 18**, onde t_R é o tempo de eluição de uma dada banda e t_0 é o tempo de retenção de um componente não retido. Portanto, podemos utilizar a **Equação 17** com a teoria baseada na relação de k com a energia livre de retenção ΔG_{ret} (**Equação 18**), onde ϕ é a razão de fases da coluna, R é a constante universal dos gases e T é a temperatura absoluta.

$$k = \frac{(t_R - t_0)}{t_0} \quad \text{ou} \quad t_R = (1 + k)t_0 \quad \text{Equação 17}$$

$$\ln k = -\frac{\Delta G_{ret}}{RT} + \ln \phi \quad \text{Equação 18}$$

Nós podemos combinar a **Equação 17** com a **Equação 19** para obter uma função dependente da variância de banda em unidades de tempo ou desvio padrão (σ^2), conforme a **Equação 20**.

$$N = \left(\frac{t_R}{\sigma} \right)^2 \quad \text{Equação 19}$$

$$(k + 1)^2 = \sigma^2 \left(\frac{N}{t_0^2} \right) \quad \text{Equação 20}$$

Podemos assumir uma mesma condição cromatográfica da coluna em termos do número de pratos (N) e t_0 , assim podemos descrever o termo da direita em parêntese da **Equação 20** como uma constante (ϖ), e então simplificar para a **Equação 21**.

$$(k + 1)^2 = \sigma^2 \varpi \quad \text{Equação 21}$$

A fim de relacionar o alargamento de banda com o efeito entrópico em função do tempo assim como as interações eletrostáticas (I) dos analitos com a fase estacionária de troca catiônica, permita-me aplicar a equação de Einstein sobre o efeito *Browniano*, onde $\sigma^2 = 2Dt$, aqui D é o coeficiente de difusão das moléculas e t é o tempo de difusão, o qual pode ser considerado em nosso caso como o tempo da parada do fluxo na 1D, conforme previamente discutido por Janssen, H-G. [4] e Greibrokk T. [6] Então, após o rearranjo da **Equação 21**, é conclusivo uma aproximação dos cálculos na linearidade positiva de $\ln(\sqrt{2Dt\varpi}) - 1$ vs $1/\sqrt{I}$. Assim, quando se aumenta a magnitude de I na fase móvel $\ln(\sqrt{2Dt\varpi}) - 1$ é diminuído, essa observação foi sugerida do mesmo resultado de $\ln k$ vs $1/\sqrt{I}$. De modo que o tempo de retenção das proteínas é vencido com o aumento da concentração de sal na etapa de eluição, assim $\ln k$ é diminuído. Fato muito condizente com a teoria de equilíbrios químicos iônicos em suportes de troca iônica.

Como nossas deduções matemáticas atingem a seguinte igualdade: $\ln k = \ln(\sqrt{2Dt\varpi}) - 1$, podemos aplicar estes resultados na **Equação 21**, de modo a obter ΔG_{ret} em função do coeficiente de difusão e tempo da parada do fluxo, assim obtemos a **Equação 22**, de modo que D_1 é o coeficiente de difusão dos analitos na 1D e t_1 é o tempo da parada do fluxo na 1D. Assim, para que a **Equação 22** seja válida é necessário que o termo da raiz seja positivo e maior que 1, uma vez que não há solução de *raiz quadrada* e *logaritmo natural* para número negativo.

$$\ln(\sqrt{2D_1t_1\varpi} - 1) = -\frac{\Delta G_{ret}}{RT} + \ln \varnothing \quad \text{Equação 22}$$

A **Equação 22** mostra a relação da *Variação da Energia Livre de Gibbs de Retenção* (ΔG_{ret}) com o coeficiente de difusão das moléculas (D_1), cuja grandeza está relacionada com a magnitude de interação química e, portanto quanto mais forte a interação química da fase estacionária junto ao analito menor será o valor de D_1 , portanto é conclusivo mencionar menor

difusão das moléculas e consequentemente menor alargamento de banda, assim como o oposto pode ser verificado, caso D_I possua grandeza elevada e consequentemente interações químicas fracas com a fase estacionária. Por fim, a referida equação elucida o tempo da parada de fluxo por meio do multiplicador t . Assim, a **Equação 22** está condizente com os três pontos experimentais os quais interpretamos referentes aos trabalhos de Janssen, H-G., T. [4] Greibrokk [6] e T. Welsch, [7] conforme descrito anteriormente.

Por meio da **Equação 22** é possível realizar algumas aproximações matemáticas de modo a prever quando ocorrerá o alargamento de banda nos picos da 1D. Para facilitar o nosso raciocínio, iremos considerar o valor de $\ln \phi$ insignificante. Assim, com o rearranjo obtemos a **Equação 23**, como esta equação trata-se da *Variação da Energia Livre de Gibbs de Retenção*, podemos considerar a constante dos gases ideais como $R = 8,31 \text{ J mol}^{-1} \text{ K}^{-1}$ e a temperatura a 25 °C em *Kelvin* como $T = 298,15 \text{ K}$, portanto o produto de $RT = 2477,6265 \text{ J mol}^{-1}$. Assim obtemos a unidade adequada para a *Variação da Energia livre de Gibbs de Retenção* (J mol^{-1}). Agora, estamos aptos a realizar as previsões matemáticas.

$$-\Delta G_{ret} = \ln(\sqrt{2D_1 t_1 \varpi} - 1)RT \quad \text{Equação 23}$$

Primeiramente, para que a **Equação 23** tenha solução, é necessário que o produto múltiplo de $D_1 t_1 \varpi$ seja $> 0,5$ caso contrário a solução de $(\sqrt{2D_1 t_1 \varpi} - 1)$ será negativo, uma vez que não há solução de $\ln(x)$ para números negativos. Contudo, se $D_1 t_1 \varpi > 0,5$ e $< 2,0$ a **Equação 23** possui solução e ΔG_{ret} será **positivo**. Por outro lado, se $D_1 t_1 \varpi > 2$ a equação 12 faz com que ΔG_{ret} seja **negativo**. A **Tabela 18** exemplifica o raciocínio matemático.

Tabela 18 - Estimativas dos possíveis valores da Variação da Energia Livre de Gibbs de Retenção (ΔG_{ret})

Valor de $D_1 t_1 \varpi$	Solução de $\ln(\sqrt{2D_1 t_1 \varpi} - 1)$	Valor de $-\Delta G_{ret} (\text{kJ mol}^{-1})$	Sinal de ΔG_{ret}
0,50	Não há	Não há	Não há
0,60	-2,35	-5,82	Positivo
1,00	-0,88	-2,18	Positivo
1,90	-0,05	-0,13	Positivo
2,00	0,00	0,00	Equilíbrio
2,10	0,05	0,11	Negativo
3,00	0,37	0,91	Negativo

Por meio da interpretação da **Tabela 18**, concluímos que há um “divisor de águas”, que prevê quando o sistema não será espontâneo $+\Delta G_{ret}$ (**positivo** para valores de $D_I t_I \varpi$ entre 0,5 e 1,9) e quando o sistema será espontâneo $-\Delta G_{ret}$ (**negativo** para valores de $D_I t_I \varpi$ acima de 2). Portanto, o alargamento de banda na 1D em métodos cromatográficos multidimensionais com parada de fluxo ocorrerá para valores de $D_I t_I \varpi > 2,00$. Provavelmente, os resultados aqui elucidados e com maiores estudos empíricos que comprovem a **Equação 23** serão um atrativo para a utilização de colunas com forte interação química entre o analito e a fase estacionária, em modos práticos a utilização de colunas iônicas na 1D.

Os resultados discutidos até aqui podem ser utilizados com o seguinte argumento: como as moléculas estão aprisionadas na coluna SCX por interações eletrostáticas, a difusão (D_I) é baixa. Assim, podemos assumir que o longo tempo de análise (t_I) que permite a difusão das moléculas na coluna SCX possui magnitude insuficiente para proporcionar difusão significativa na contribuição do alargamento de banda. Portanto, se $D_I \ll t_I$ a contribuição do tempo de parada do fluxo possui efeito insignificante nos resultados, devido ao baixo valor de D_I , que comanda o resultado final da equação. De forma análoga, quando $D_I > t_I$ a contribuição do tempo na parada do fluxo possui efeito significativo no alargamento de banda.

A **Equação 22** e **Equação 23** possuem alguns pontos de retratação de suma importância: i) o valor de D_I é referente ao coeficiente de difusão das moléculas no meio cromatográfico, portanto além do efeito *Browniano*, deve-se considerar o efeito das interações químicas que ocorrem entre o analito a fase móvel e a fase estacionária; ii) deve-se observar a magnitude do valor de D_I , t_I e ϖ , pois por meio do produto ($D_I t_I \varpi$) é possível prever o alargamento de banda de um método com parada de fluxo na 1D e iii) interações fortes são interessantes, contudo espera-se que ocorra a eluição de todos analitos na 1D, ou seja, que não ocorra adsorção irreversível dos analitos, caso contrário pode haver o efeito de memória e comprometer com a repetibilidade do método analítico.

Em suma, desenvolveu-se uma equação que prevê, de maneira muito simples, o alargamento de banda na 1D em métodos cromatográficos multidimensionais com parada de fluxo, ou seja, para valores de $D_I t_I \varpi > 2,00$. Tais resultados foram baseados por meio da literatura (ciência de separações) e por meio das observações de nossos resultados das análises proteômicas. Provavelmente, a hipótese aqui elucidada e com maiores estudos empíricos que

comproven a **Equação 23** serão um atrativo para a utilização de colunas com forte interação química entre o analito e a fase estacionária, em modos práticos a utilização de colunas iônicas na 1D.

5.4. CONCLUSÕES

Elaborou-se uma equação para a previsão do alargamento de banda em cromatografia líquida multidimensional com parada de fluxo na 1D. A equação é explicada por meio da *Variação da Energia de Gibbs*, e, portanto, possui uma função diretamente relacionada em termos termodinâmicos. Esta previsão pode ajudar a técnica de modo a evidenciar a importância da utilização de colunas cromatográficas com elevado poder de retenção (interação e/ou adsorção) dos analitos, provavelmente, evidenciando o uso de colunas de troca iônica ou análogas em sistemas multidimensionais com parada de fluxo, além disso, ajudar no tempo de amostragem na 1D. É interessante mencionar que em proteômica, onde exige eficiência de separação elevada, métodos cromatográficos multidimensionais com trocadora catiônica na 1D, são as preferidas pela ortogonalidade e solução de compostos para o espectrômetro de massas, qualidades estas da técnica que podem ser explicadas em parte, pela equação aqui retratada.

5.5. REFERÊNCIAS

- 1 VAN DEEMTER, J. J.; ZUIDERWEG, F. J.; KLINKENBERG, A. Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. **Chemical Engineering Science**, Oxford, v. 5, p. 271-289, 1956.
- 2 JORGENSEN, J. W. Capillary liquid chromatography at ultrahigh pressures. **Annual Review of Analytical Chemistry**, Palo Alto, v. 3, p. 129-150, 2010.
- 3 KNOX, J. H.; SCOTT, H. P. B and C terms in the van Deemter equation for liquid chromatography. **Journal of Chromatography A**, Amsterdam, v. 282, p. 297-313, 1983.
- 4 BEDANI, F.; KOK, W.T.; JANSSEN, H. G. A theoretical basis for parameter selection and instrument design in comprehensive size-exclusion chromatography×liquid chromatography. **Journal of Chromatography A**, Amsterdam, v. 1133, n. 1-2, p. 126-134, 2006.
- 5 GIDDINGS, J. C. **Dynamics of Chromatography**. Part I: principles and theory. New York: Marcel Dekker, Inc., 1965. p. 1-118.
- 6 TRAN, B. Q.; LUNDANES, E.; GREIBROKK, T. Influence of stop-flow on band broadening of peptides in micro-liquid chromatography. **Chromatographia**, Heidelberg, v. 64, n. 1-2, p. 1-5, 2006.
- 7 KOHNE, A. P.; DORNBERGER, U.; WELSCH, T. Two-dimensional high-performance liquid chromatography for the separation of complex mixtures of explosives and their by-products. **Chromatographia**, Heidelberg, v. 48, n. 1-2, p. 9-16, 1998.
- 8 STAHLBERG, J.; JONSSON, B.; HORVÁTH, C. Theory for electrostatic interaction chromatography of proteins. **Analytical Chemistry**, Washington, v. 63, n. 17, p. 1867-1874, 1991.
- 9 KARGER, B. L.; SNYDER, L. R.; HORVATH, C. **An introduction to separation science**. Hoboken: Wiley, 1973. p. 1-102.

ANEXOS

Relação de peptídeos não identificados (Modelo Mix) do método SCX×RP-Q-

ToF(MS/MS)

Relação dos peptídeos de baixa massa molecular não identificados no método			Peptídeos de elevada massa molecular não identificados no método (Região hidrofóbica hipotética da Figura 19)		
Peptídeos	Gravy	Massa molecular (Da)	Peptídeos	Gravy	Massa molecular (Da)
MK	-1,00	295,40	QFYQLDAYPSGAWYYVPLGTQYTDAPSFSDIPNPIGSENSEK	-0,693	4718,08
INK	-0,97	409,48	YPSYGLNYYQKQPVALINNQLPYPYYAKPAAVR	-0,509	4011,59
K	-3,90	146,19	YPVEPFTEQSQTLDVENLHLPLLLQSWMHQPHQLPPTVMFPPQSVLSLSQSK	-0,257	6362,32
IEK	-0,97	424,49	DIDGYGGIALPELICTMFHTSGYDTQAIVENNESTEYGLFQISNK	-0,224	4556,44
VK	0,15	263,34	TEIPTINTIASGEPTSTPTTEAVESTVATLEDSPEVIESPPEINTVQVTSTAV	-0,123	5455,92
HK	-3,55	301,35	IHPFAQTQSLVYPFGPIPNSLPQNIPPLTQTPVVVPPFLQPEVMGVSK	0,104	5319,25
MMK	-0,03	444,61	Média	-0,284	5070,6
CEK	-1,63	414,48			
DER	-3,83	426,42			
IAK	0,93	366,45			
K	-3,90	146,19			
K	-3,90	146,19			
LR	-0,35	305,37			
YR	-2,90	355,39			
MR	-1,30	323,41			
K	-3,90	146,19			
L	3,80	131,17			
MK	-1,00	295,40			
QEK	-3,63	439,47			
R	-4,50	174,20			
K	-3,90	146,19			
K	-3,90	146,19			
TK	-2,30	265,31			
NR	-4,00	306,32			
YQK	-2,90	411,46			
TK	-2,40	265,31			
YL	-2,60	312,36			
MK	-1,00	295,40			
EK	-3,70	293,32			
LK	-0,05	277,36			
K	-3,90	146,19			
YK	-2,60	327,38			
FDK	-1,53	444,48			

FK	-0,55	311,38
HLK	-1,10	146,19
K	-3,90	146,19
K	-3,90	146,19
K	-3,00	146,19
HK	-3,55	301,35
YK	-2,60	327,38
Média	-2,24	277,93

Média de *Gravy* dos peptídeos para as regiões hidrofílicas dos Modelos BSA e MIX

Região hidrofílica (Modelo BSA)		Região hidrofóbica (Modelo Mix)	
Peptídeos	Grand average of hydropathicity (GRAVY):	Peptídeos	Grand average of hydropathicity (GRAVY):
GVFR	0,525	KNQDK	-3,660
DTHK	-2,825	QEPER	-3,320
QNcDQFEK	-2,013	DTHK	-2,825
QEPERNEcFLSHKDDSPDLPK	-1,762	YTRK	-2,600
ADEKKFWGK	-1,711	LTEEEKNR	-2,413
DTHKSEIAHR	-1,700	TKLTEEEKNR	-2,390
LKPDPNTLcDEFKADEKK	-1,433	NRcK	-2,350
ATEEQLK	-1,357	KTCLTEEEK	-2,200
LKPDPNTLcDEFKADEK	-1,288	QRLR	-2,175
KFWGK	-1,260	RHGLDNYR	-2,138
FKDLGEEHFK	-1,250	KISQRYQK	-2,112
HLVDEPQNLIKQNcDQFEK	-1,184	EQLSTSEENSKK	-2,050
LSQKFPK	-1,014	QNcDQFEK	-2,013
LKPDPNTLcDEFK	-0,985	HKPKATEEQLK	-2,009
YNGVFQEccQAEDK	-0,914	TKLTEEEK	-1,988
RHPYFYAPELLYYANK	-0,812	RDTHKSEIAHR	-1,955
NYQEAKDAFLGSFLYEYSR	-0,789	AHGKK	-1,920
YlcDNQDTISSKLK	-0,721	LTEEEK	-1,883
DDPHAcYSTVFDKLLK	-0,720	EQLSTSEENSK	-1,882
LYYEIARR	-0,625	ISQRYQK	-1,857
LKEccDKPLLEK	-0,617	LKKYK	-1,840
FWGK	-0,600	DDSPDLPK	-1,825
QNcDQFEKLGEYGFQNALIVR	-0,586	DERFFSDK	-1,762
HLVDEPQNLIK	-0,582	QEPERNEcFLSHKDDSPDLPK	-1,762
HPYFYAPELLYYANK	-0,567	HIQKEDVPSEK	-1,755
RPcFSALTPDETYVPK	-0,537	cEKDERFFSDK	-1,727
VHKEccHGDLLcEADDRADLAK	-0,527	QEPERNEcFLSHK	-1,723
QTALVELLKHKPK	-0,523	ADEKKFWGK	-1,711
LKHLVDEPQNLIK	-0,500	DTHKSEIAHR	-1,700
LRcASIQKFGER	-0,475	TVYQHQQ	-1,700
EccHGDLLcEADDRADLAK	-0,458	ETYGDmADccEKQEPER	-1,659
VASLRETYGDmADccEK	-0,418	TPGSR	-1,600
YTRKVPQVSTPTLVEVSR	-0,417	mGHFTEEDK	-1,566
DAFLGSFLYEYSRR	-0,400	GVFRRDTHK	-1,522
LAKEYEATLEEccAK	-0,387	DTHKSEIAHRFK	-1,508
LSQKFPKAEFVEVTK	-0,380	ccTKPESER	-1,500

AFDEKLFTFHADl _c TLPDTEKQIK	-0,354	ITVDDKHYQK	-1,480
YNGVFQEccQAEDKGAcLLPK	-0,324	KGHHEAELKPLAQSHATKHK	-1,470
AFDEKLFTFHADl _c TLPDTEK	-0,267	EGIHAAQQK	-1,462
LcVLHEKTPVSEKVTk	-0,188	DVcKNYQEAk	-1,460
RHPEYAVSVLLR	-0,133	QEKNmAINPSK	-1,455
ccAADDKEAcFAVEGPK	-0,124	HYQKALNEINQFYQK	-1,447
FPKAEFVEVTK	-0,118	LKPDPNTLcDEFKADEKK	-1,433
LFTFHADl _c TLPDTEKQIK	-0,116	ASEDLKK	-1,429
DAFLGSFLYEYSR	-0,085	SSQVPQSR	-1,413
YLYEIAR	-0,071	FESNFNTQATNR	-1,400
KVPQVSTPTLVEVSR	-0,067	FGER	-1,400
LFTFHADl _c TLPDTEK	0,044	IETmREK	-1,386
SLHTLFGDELcK	0,058	KKGHHEAELKPLAQSHATK	-1,379
mPcTEDYLSLILNR	0,064	TEAEmKASEDLKK	-1,377
LVNELTEFAK	0,130	ATEEQLK	-1,357
AEFVEVTK	0,175	GHHEAELKPLAQSHATKHK	-1,342
KQTALVELLK	0,190	VKAHGKK	-1,329
GLVLIAFSQYLQcPFDEHVK	0,262	QFTK	-1,325
HPEYAVSVLLR	0,264	TEAEmK	-1,317
AWSVAR	0,267	LKPDPNTLcDEFKADEK	-1,288
LGEYGFQNALIVR	0,292	NEcFLSHKDDSPDLPK	-1,275
AEFVEVTKLVTDLTk	0,293	HPHPHLSFmAIPPKKNQDK	-1,274
SLHTLFGDELcKVASLR	0,306	AFDEK	-1,260
ALKAWSVAR	0,367	KFWGK	-1,260
LVTDLTK	0,429	FKDLGEEHFK	-1,250
GVFR	0,525	HLKTEAEmK	-1,244
LcVLHEK	0,529	KGHHEAELKPLAQSHATK	-1,239
TVmENFVAFVDK	0,567	EKVNELSKDIGSESTEDQAmEDIK	-1,229
QTALVELLK	0,644	TVYQHQA _m KPWIQPK	-1,194
LVVSTQTALA	1,390	HLVDEPQNLIKQ _{Nc} DQFEK	-1,184
	-0,42	TEAEmKASEDLK	-1,167
	0,73	HLKTEAEmKASEDLK	-1,153
		DIGSESTEDQAmEDIK	-1,150
		NAVPTPTLNREQLSTSEENSKK	-1,148
		VGTRccTKPESER	-1,146
		FDKFK	-1,140
		NFGKEFTPQ _m QAAYQK	-1,137
		SEIAHRFKDLGEEHFK	-1,137
		FDKFKHLK	-1,125
		FKHLKTEAEmK	-1,118
		NRLNFLKK	-1,113
		NTDGSTDYGILQINSRW _{Wc} NDGR	-1,109
		EFTPQ _m QAAYQK	-1,100
		EFTPEVQASWQK	-1,092

HLDDLK	-1,083
LKHLVDEPQNLIKQNcDQFEK	-1,076
ISQR	-1,075
AmKPWIQPKTK	-1,073
LHSmKEGIHAQQK	-1,069
DERFFSDKIAK	-1,064
TPVSEK	-1,050
YQKFALPQYLKTVYQHQQ	-1,039
ASEDLK	-1,017
LSQKFPK	-1,014
LHVDPENFR	-1,000
NDIAAKYK	-1,000
LKPDPNTLcDEFK	-0,985
KYKVPQLEIVPNSAEER	-0,976
NYQEAKDAFLGSFLYEYSRR	-0,975
TNVK	-0,975
ETYGDMADccEK	-0,967
NRcKGTDVQAWIR	-0,954
SEIAHR	-0,950
LHVDPENFK	-0,933
TYFPHFDLSHGSAQVKGHGKK	-0,919
YNGVFQEccQAEDK	-0,914
EKVLASSARQR	-0,900
QmEAEISSSEEIVPNSVEQKHIQK	-0,900
NTDGSTDYGILQINSR	-0,894
ALNEINQFYQK	-0,891
HIQKEDVPSERYLGYLEQLLR	-0,886
EFTPEVQAAWQK	-0,875
KTVDMESTEVFTKK	-0,857
SEIAHRFK	-0,850
TVDmESTEVFTKKTK	-0,847
LFTGHPETLEKFDK	-0,836
YIcDNQDTISSK	-0,833
DDPHAcYSTVFDK	-0,823
LWcKSSQVPQSR	-0,817
HPIKHQGLPQEVLENLLR	-0,816
QTALVELLKHKPKATEEQLK	-0,815
RHPYFYAPELLYYANK	-0,812
QRLRcASIQK	-0,810
AmKPWIQPK	-0,800
LFTGHPETLEKFDKFK	-0,800
YKELGFQG	-0,800
YKVPQLEIVPNSAEER	-0,794
NYQEAKDAFLGSFLYEYSR	-0,789

LSELHcDKLHVDPENFK	-0,776
TYFPHFDLSHGSAQVKGHGK	-0,770
KQTALVELLKHKPK	-0,764
HQGLPQEVLENENLLR	-0,753
EFTPPVQAAYQK	-0,750
TPVSEKVTK	-0,744
ISQRYQKFALPQYLK	-0,740
EDVPSERYLGYLEQLLR	-0,735
SSQVPQSRNiDIScDK	-0,735
HPGDFGADAQGAmTK	-0,733
EccDKPLLEK	-0,730
NEcFLSHK	-0,725
YIcDNQDTISSKLKEccDKPLLEK	-0,725
mVHLTPEEK	-0,722
YIcDNQDTISSKLK	-0,721
DDPHAcYSTVFDKLLK	-0,720
NRLNFLK	-0,714
RHGLDNYRGYSLGNWVcAAK	-0,710
RPcFSALTPDETYVPKAFDEK	-0,710
YKVPQLEIVPNSAEERLHSmK	-0,710
NmDNLKPAFAK	-0,709
VHKEccHGDLLcEADDR	-0,682
FALPQYLKTVYQHQQ	-0,667
ScQAQPTTmAR	-0,664
NmDNLKPAFAKLSELHcDK	-0,663
ScQAQPTTmARHPHPLSFmAIPPKK	-0,658
NDIAAKYKELGFQG	-0,657
HPHPLSFmAIPPKK	-0,653
LEQWLcEK	-0,650
LFTGHPETLEK	-0,645
NmAInPSK	-0,637
EYEATLEccAK	-0,625
YLYEIARR	-0,625
KTVDmESTEVFTK	-0,623
TVDmESTEVFTKK	-0,623
EccHGDLLcEADDR	-0,621
FWGKYLYEIARR	-0,617
LKEccDKPLLEK	-0,617
VNELSK	-0,617
TNVKAAWGK	-0,611
FWGK	-0,600
HKIPIK	-0,600
LSELHcDK	-0,600
ADLAKYIcDNQDTISSK	-0,588

QNcDQFEKLGEYGFQNALIVR	-0,586
HLVDEPQNLIK	-0,582
GTFSQLSELHcDK	-0,569
HPYFYAPELLYYANK	-0,567
KFWGKYLYEIAR	-0,567
DAIPENLPPLTADFAEDK	-0,561
FGERALK	-0,557
RPcFSALTPDETYVPK	-0,537
VPQLEIVPNSAEER	-0,536
ADLAKYIcDNQDTISSKLK	-0,532
EGIHAAQQKEPmIGVNQELAYFYPELFR	-0,530
VHKEccHGDLLcEADDRADLAK	-0,527
QTALVELLKHKPK	-0,523
LNFLKKISQR	-0,520
TcVADESHAGcEK	-0,515
HGLDNYRGYSLGNWVcAAK	-0,511
VPQLEIVPNSAEERLHSmK	-0,511
cASIQKFGER	-0,500
GTFATLSELHcDKLHVDPENFR	-0,500
LKHLVDEPQNLIK	-0,500
IVSDGNGmNAWVAWRNRcK	-0,495
VNVEDAGGETLGR	-0,492
DAIPENLPPLTADFAEDKDVcK	-0,491
cKGTDVQAWIRGcR	-0,486
ITVDDK	-0,483
GVFRR	-0,480
mNVEEAGGEALGR	-0,477
LRcASIQKFGER	-0,475
IVSDGNGmNAWVAWRNR	-0,471
TYFPHFDLSHGSAQVK	-0,469
IETmR	-0,460
EccHGDLLcEADDRADLAK	-0,458
TNVKAAWGKVGAHAGEYGAEALER	-0,458
mVHFTAEEK	-0,456
GTDVQAWIRGcR	-0,450
LKEccDKPLLEKSHcIAEVEK	-0,443
AWSVARLSQKFPK	-0,423
HPHPHLSFmAIPPK	-0,421
VASLRETYGDmADccEK	-0,418
YTRKVPQVSTPTLVEVSR	-0,417
cKGTDVQAWIR	-0,400
DAFLGSFLYEYSRR	-0,400
LAKEYEATLEEccAK	-0,387
FFESFGDLSSPDAmGNPKVKAHGK	-0,384

LSQKFPKAEFVEVTK	-0,380
TPVSEKVTKccTESLVNR	-0,372
ALcTEKLEQWLcEK	-0,371
GTFAQLSELHcDK	-0,369
VGAHAGEYGAEALER	-0,367
FPQYLQYLYQGPIVLNPWDQVK	-0,355
AFDEKLFTFHADlcTLPDTEKQIK	-0,354
TVDmESTEVFTK	-0,350
NmAINPSKENLcSTFcK	-0,347
ccTESLVNRRPcFSALTPDETYVPK	-0,344
GIDYWLAKH	-0,344
SLGKVGTR	-0,338
GTDVQAWIR	-0,333
ccAADDK	-0,329
HPGDFGADAQGAmTKALELFR	-0,324
YNGVFQEccQAEDKGAcLLPK	-0,324
AAWGK	-0,320
IETmREKVLASSAR	-0,300
VLASSARQRLR	-0,291
AWSVARLSQK	-0,280
FLDDITDDImcAKK	-0,280
VLASSARQR	-0,278
ALELFRNDIAAKYK	-0,271
AFDEKLFTFHADlcTLPDTEK	-0,267
FWGKYLYEIAR	-0,265
GAcLLPKIETmREK	-0,264
NmAINPSKENLcSTFcKEVVR	-0,262
QFTKcELSQLLK	-0,258
cASIQKFGERALK	-0,254
KIVSDGNGmNAWVAWR	-0,244
TcVADESHAGcEKSLHTLFGDELcK	-0,240
mVHLTPEEKTAVNALWGK	-0,239
LRVDPVNFK	-0,222
FFESFGDLSPDAVmGNPK	-0,221
FFESFGDLSTPDAVmGNPK	-0,216
mVLSPADKTNVKAAGWK	-0,212
EKVLASSAR	-0,211
SHcIAEVEK	-0,211
GIDYWLAKHcALcTEK	-0,207
ELGFQG	-0,200
LcVLHEKTPVSEK	-0,200
FFDSFGNLSSPSAILGNPKVKAHGK	-0,188
LcVLHEKTPVSEKVTK	-0,188
FFESFGDLSPDAVmGNPKVK	-0,186

VDPVNFK	-0,186
VEADIAGHGQEV LIRLFTGHPETLEK	-0,185
FFESFGDLSTPD A V mGNPKVK	-0,181
KVLTSFGDAIK N mDNLKPAFAK	-0,177
DVcK	-0,175
LGEYGFQNALIVRYTR	-0,169
ATITSLWGKVNVEDAGGETLGR	-0,168
mVLSPADKTNVK	-0,167
NAVPIPTLNR	-0,164
GTFATLSELHcDK	-0,154
LNFLKK	-0,150
ATEEQLKTV mENFVAFVDK	-0,142
EPmIGVNQELAYFYPELFR	-0,137
NlcDIScDK	-0,133
RHPEYAVSVLLR	-0,133
ccAADDKEAcFAVEGPK	-0,124
GTDVQAWIRGcRL	-0,123
AHGKKVLTSFGDAIK	-0,120
FPKAEFVEVTK	-0,118
mNVEEAGGEALGRLLVVYPWTQR	-0,117
LFTFHADl cTLPDTEKQIK	-0,116
VNVDEVGGEALGR	-0,115
KILDIKGIDYWLAHK	-0,107
KVLTSLGDAIKHLDDLK	-0,094
ALcTEKLEQWLcEKL	-0,093
ENLcSTFcK	-0,089
mVHLTPEEKSAVTALWGK	-0,089
ALKAWSVARLSQK	-0,085
DAFLGSFLYEYSR	-0,085
VGAHAGEYGAEALER mFLSFPTTK	-0,079
QIKKQTALVELLK	-0,077
YLYEIAR	-0,071
AAVTSLWSK mNVEEAGGEALGR	-0,070
KVPQVSTPTLVEVSR	-0,067
AHGKKVLTSLGDAIK	-0,053
VTKccTESLVNR	-0,033
ENLcSTFcKEVVR	-0,031
FLDDDITDDImcAK	-0,021
SPAQILQWQVLSNTVPAK	-0,011
ALcTEK	0,000
ccTESLVNR	0,000
IVSDGNGmNAWVAWR	0,000
RHPEYAVSVLLRLAK	0,007
LVTDLTKVHK	0,010

EAcFAVEGPK	0,020
FFDSFGNLSSPSAILGNPK	0,037
LFTFHADlcTLPDTEK	0,044
YLGYLEQLRLK	0,050
SLHTLFGDELcK	0,058
FLDDITDDImcAKKILDIK	0,060
mPcTEDYLSLILNR	0,064
YLGYLEQLLR	0,070
AEFVEVTKLVTDLTkVHK	0,083
VPQVSTPTLVEVSRSLGK	0,089
FPKAEFVEVTKLVTDLTk	0,094
cASIQK	0,100
TKVIPYVR	0,113
ILDIKGIDYWLAHKALcTEK	0,115
FFDSFGNLSSASAIImGNPK	0,116
ALELFRNDIAAK	0,117
LVNELTEFAK	0,130
VLTSLGDAIKHLDDLK	0,144
VEADIAGHGQEVLR	0,153
SAVTALWGKVNVDVGGGEALGR	0,155
KVLGAFSDGLAHLNLK	0,159
ILDIKGIDYWLAHK	0,164
AEFVEVTK	0,175
KQTALVELLK	0,190
VPQVSTPTLVEVSR	0,207
FFSDKIAKYIPIQYVLSR	0,211
FFVAPFPEVFGKEK	0,214
FALPQYLK	0,237
mVLSPADK	0,237
TVmENFVAFVDKccAADDK	0,237
GYSLGWNWcAAK	0,242
TAVNALWGK	0,244
KILDIK	0,250
GLVLIAFSQYLQcPFDEHVK	0,262
HPEYAVSVLLR	0,264
AWSVAR	0,267
TAVNALWGKVNVDVGGGEALGR	0,273
HKIPIKYLEFISDAIHVLHSK	0,277
LGEYGFQNALIVR	0,292
VFGRcELAAAmKR	0,292
VNVDVGGGEALGR	0,292
AEFVEVTKLVTDLTk	0,293
TTmPLW	0,300
ASEDLKKHGTVVLTALGGILK	0,305

SLHTLFGDELcKVASLR	0,306
GAcLLPKIETmR	0,308
VNVDAVGGEALGRLLVVYPWTQR	0,317
HPEYAVSVLLRLAK	0,329
TKVIPYVRYL	0,340
LLVVYPWTQR	0,350
KVLTSFGDAIK	0,355
ALKAWSVAR	0,367
mVTGVASALSSRYH	0,371
LWcK	0,375
mFLSFPTTK	0,400
YIPIQYVLSR	0,400
VLGAFSDGLAHLNKL	0,412
NLcNIPcSALLSSDITASVNCaKK	0,425
LVTDLTK	0,429
KVLTSLGDAIK	0,445
IAKYIPIQYVLSR	0,492
AAVTSLWSK	0,500
FLASVSTVLTskyr	0,507
GVFR	0,525
VFGR	0,525
LcVLHEK	0,529
mVTAVASALSSRYH	0,529
HGTVVLTALGGILKKK	0,537
SAVTALWGK	0,544
TVmENFVAFVDK	0,567
VVAGVANALAHKYH	0,579
LHVDPENFKLLGNV _m VIIILATHFGK	0,592
LNFLK	0,600
YLEFISDAIIHVLHsk	0,606
NLcNIPcSALLSSDITASVNCaK	0,613
QTALVELLK	0,644
VFGRcELAAAmK	0,692
ALELFR	0,700
cELAAAmK	0,755
VLTSFGDAIK	0,780
mVTGVASALSSR	0,808
HGTVVLTALGGILKK	0,833
KHGTVVLTALGGILK	0,833
GAcLLPK	0,857
FFVAPFPEVFGK	0,867
VLTSLGDAIK	0,880
VIPYVR	0,917
mKWVTFISLLLLFSSAYSr	0,984

mVTAVASALSSR	0,992
VIPYVRYL	1,000
VVAGVANALAHK	1,050
FLASVSTVLTSK	1,075
WVTFISLLLLFSSAYSRGVFR	1,086
LVSVAIALAHKYH	1,093
LLGNVLVcVLARNFGK	1,169
HGTVVLTALGGILK	1,171
WVTFISLLLLFSSAYSR	1,218
LLGNVLVcVLAHHFGK	1,269
mRFFVPLFLVGILFPAILAKQFTK	1,333
LVVSTQTALA	1,390
LLGNVmVILATHFGK	1,450
LLGNVLVTVLAIHFGK	1,550
mRSLILVLcFLPLAALGKVFGK	1,609
LVSVAIALAHK	1,650
LLILTcLVAVALARPK	1,819
mRSLILVLcFLPLAALGK	1,837
mKFFIFTcLLAVALAK	1,863
mRFFVPLFLVGILFPAILAK	1,865
SLLILVLcFLPLAALGKVFGK	1,886
LLGNVLVcVLAR	1,975
SLLILVLcFLPLAALGK	2,206
FFVPLFLVGILFPAILAK	2,217
FFIFTcLLAVALAK	2,271
<hr/>	
Média	-0,405
DP	0,870

m-DoE – Multi-Design of Experiments with Multi-Responses for Optimization of on-line 2D LC-MS/MS in Proteomic Analysis

Weliton Pedro Batiston,^{†,ϕ} Emanuel Carrilho,^{†,ϕ} Álvaro José dos Santos-Neto^{†,*}

[†] *Instituto de Química de São Carlos, Universidade de São Paulo
Av. Trabalhador São-carlense, 400, 13566-590 São Carlos, SP, Brazil*

^ϕ *Instituto Nacional de Ciência e Tecnologia de Bioanalítica – INCTBio, 13083-970, Campinas, SP, Brazil*

ABSTRACT: The optimization of all platforms in hyphenated analytical techniques, such as in 2D LC-MS/MS, is a hard and necessary work, which has a high impact on successful data acquisition and analysis. However, the literature is incipient when addressing multivariate statistical evaluation of a considerable number of parameters and responses. Here, we propose a chemometrics alternative for saturated design, named m-DoE, applied for full optimization of on-line comprehensive multidimensional liquid chromatography coupled to tandem mass spectrometry for shotgun proteomics (LC×LC-MS/MS). This investigation overcomes the limitations of the mathematical pattern of Plackett-Burman Design (PBD) for example, not considering the confounding patterns, or even the high tendency of PBD to select inadequate range levels of significant factors. These limitations can generate biased results on chemometrics performance. The m-DoE allows efficient optimization of the probed system by a combination of multiple “interconnected” responses using Derringer’s desirability function. Here, we introduce a response surface evaluated by ANOVA on a LC×LC-MS/MS method of high orthogonality applied to shotgun proteomics. The m-DoE approach yielded unseen protein identification data as *i*) a score of 3816.29, *ii*) the number of peptides 73, *iii*) the number of unique peptides 32, and *iv*) protein coverage of 92.92% when BSA was a model protein or *v*) average coverage of 83.07% for 11 species identified (between proteins, phosphoprotein, and isoforms), reaching up to 100% coverage for 45.45% of analyzed species. However, this strategy can lead to an additional number of experiments concerning PBD. Here, we discuss when each approach is most appropriate and how to make decisions.

Keywords: design of experiments; multidimensional liquid chromatography; tandem mass spectrometry; bottom-up proteomics

1. INTRODUCTION

During the past twenty years, much progress has been made in the elucidation of relevant biological problems and the understanding of biological phenomena in terms of the identification and quantification of molecules and their interconnected properties of cells.[1,2] The first great scientific impact in this field was the sequencing of the human genome, whereas now, the next step is the determination of complex dynamic functions of proteins in the proteome.[3,4] To overcome this challenge the evolution of hyphenated analytical techniques such as multidimensional ultrahigh-pressure liquid chromatography coupled to tandem high-resolution mass spectrometry (LC×LC-MS/MS) has been primordial to advancing in the area.[5–17]

Although there was a great evolution in instrumental analytical chemistry for analysis of complex mixtures in large scale, the optimization strategies of all LC×LC-MS/MS platforms are incipient in the literature.[6,16,18] Probably, the high complexity and number of parameters that can be significant in the optimization is a great barrier.[17,19] Moreover, it is known that proteomics has great challenges to obtain reproducibility on analysis, affecting the logical optimization of instrumental parameters. Thus, many scientific organizations have focused on improving the analytical strategies to proteomics, as Human Proteome Organization (HUPO); Proteome Standard Initiative (PSI) and Clinical Proteomic Technology Assessment for Cancer (CPTAC).[20]

Currently, an elegant strategy to optimization has gained prominence in some fields of chemistry, named as design of experiments (*DoE*). This is one of the approaches of largest application in chemometrics. Chemometrics studies are based on statistical principles with maximum acquisition of useful information of a system evaluated with the minimum of experiments, as result it is possible to have the most satisfactory response and considerably less laboratory work.[21–25] Consequently, statistical science has the power to make the interpretation of a mathematical tendency to create rules, which can be applied to explain phenomena as well as the problem solving. On the other hand, chemistry science has the power of discernment; *i.e.*, choose the best conditions to different types of reactions, analysis, separation, identification, and for the development of new technologies.[26] This combination of sciences is called chemometrics.[26,27] Muddiman's group was the pioneer in the application of *DoE* for large-scale investigation on the MS platform to increase proteome coverage.[20,28] Thus, the *DoE* application in proteomic analysis shows as a challenge to achieve the *Pareto Principle*. This principle relies on the fact that the model response is influenced mainly by just a few factors (usually two or three), which permits to plot a tridimensional surface to visualize the better result. So, a response surface on the instrumental optimization has not been described for proteomic analysis. We can suggest that the lack of LC platform optimization, analyzing a complex proteome and investigating a large number of instrumental parameters, leads obtaining multiples

significant factors (>5) on the optimization of the hyphenated techniques. The range level of such factors, their impact on a single response, and the chemometrics strategy could define the outcome of such an approach.

To date, there are not studies discussing an integrated optimization of instrumental parameters in LC×LC-MS/MS aiming the output or response towards improved coverage of the sequencing of proteins and the increasing of species identifications. This could mean an important advance in multi-dimensional proteomics for cell biology in terms of a better understanding of proteome dynamic nature.[1] Here, we introduce the concepts and strategies for the development of a new method and optimization of *on-line* LC×LC-MS/MS with stop-flow in the first dimension (1D) for a greater peptides sequence coverage. Currently, LC×LC-MS/MS is one of the most complex hyphenated analytical techniques and when this platform is operated in *on-line* mode, adequate analysis without losing sample is a great challenge.[6,19,29,30] In this paper, we showed an alternative chemometrics strategy focused on the mathematical limitations of the saturated model proposed by *Plackett-Burman Design* (PBD).[31] Moreover, we propose a response surface for two significant parameters in LC×LC-MS/MS, the evaluation by ANOVA table shows that model fit was effective. For this new approach, we named multi-DoE with multi-Responses (*m-DoE*).[30,32] For a condensed revision on the fundamentals on chemometrics, the limitations of PBD, and the rationale behind our design of experiments (pros and cons), please, refer to the **Background Section, Supporting Information**.

2. EXPERIMENTAL SECTION

2.1. Chemicals and reagents

The standard of the protein and phosphoprotein mix constituted of bovine serum albumin (BSA) - A7906, myoglobin - M0630, lysozyme - L6876, casein and its isoforms (α S1, α S2, β , κ) - C3400, hemoglobin human and its isoforms (HBB, HbD, HBG2, HbA) - H7379, lactalbumin - L7252, was purchased from Sigma-Aldrich (São Paulo, SP, Brazil). The reagents used in sample preparation were: urea (Sigma-Aldrich), ammonium bicarbonate (Sigma-Aldrich), dithiothreitol (DTT) (Fluka Analytical), calcium chloride (CaCl_2) (Mallinckrodt GenAR), iodoacetamide (IAA) (GE Healthcare), and trypsin (Sigma-Aldrich). The solvents, salts, and acids of HPLC grade, used in the composition of the mobile phases for chromatography were: acetonitrile (ACN) (Tedia), sodium phosphate monobasic monohydrate ($\text{Na}_2\text{HPO}_4 \cdot \text{H}_2\text{O}$) (Tedia), formic acid 98% (FA) (Fluka Analytical), trifluoroacetic acid 99% (TFA) (Sigma-Aldrich), methanol (Tedia) and sodium chloride (NaCl) (Mallinckrodt).

2.2. Sample preparation

Proteins aqueous solution were diluted in ratio 1:1 in 8 mol L⁻¹ of urea, followed by reduced with 5 mmol L⁻¹ dithiolthreitol (DTT) incubated per 25 min at 56 °C and alkylated with 14 mmol L⁻¹ iodoacetamide (IAA) incubated per 30 min at room temperature. After that, solution was adjusted in terms of final concentration for 5 mmol L⁻¹ IAA to quench IAA free on solution, incubated per 15 min at room temperature. Sample was diluted at 1:50 ratio in 50 mmol L⁻¹ ammonium bicarbonate to reduce the urea concentration from 0.8 mol L⁻¹, followed by CaCl₂ addition at a final concentration of 1.0 mmol L⁻¹. Then, the digestion was carried out with the addition of trypsin solution (20 ng μL⁻¹) in the proportion of enzyme/substrate of 1:50, in the cold stock solution of ammonium bicarbonate 50 mmol L⁻¹.

2.3. LC×LC-MS/MS Instrumentation

The on-line comprehensive SCX×RP-MS/MS was performed on an ultrahigh-pressure liquid chromatograph (UHPLC) micro/capillary system LC×LC directly hyphenated with the MS/MS instrument. Our setup uses a Dionex Ultimate 3000 RSL nanoLC system, in which the reverse phase column outlet is coupled to ESI source of a quadrupole time-of-flight (Q-TOF) micrOTOF-QII from Bruker. For the first dimension (1D) LC, we used a 500 μm × 50 mm, 54 μm, 72 Å SCX column (*homemade*) followed by a 500 μm × 50 mm, 33 μm, 91 Å Strata X polymeric trap column (*homemade*), and the second dimension (2D) LC was a 300 μm × 15 cm, 2 μm, 100 Å Acclaim PepMap C₁₈ RP analytical column (Dionex). The UHPLC was equipped with a thermostated autosampler, one ternary pump, two binary pumps, switching six- and ten-port valves, a photodiode-array and a UV-vis detector, and a thermostated column compartment. This method is described with more details in **Supporting Information**.

2.4. Bioinformatics (SEQUEST Algorithm)

Tandem mass spectra were extracted from *mzXML files using DataAnalysis Version 4.0, Bruker Daltonics and searched with the Discoverer Proteome (Thermo) with the SEQUEST algorithm against Uniprot/Swiss-Prot database for BSA, myoglobin, lysozyme, casein, hemoglobin, lactalbumin protein. The workflow of software was as follows: filter for 1-4 positive charge state, precursor mass adjusted between *m/z* 100-5000, total intensity threshold of 100, signal noise ratio threshold of 3, precursor ion mass tolerance of 2 Da, fragment ions mass was searched with a 0.8 Da mass window, two missed cleavage sites for trypsin were allowed, carboxymethyl at cysteine and oxidation of methionine were set as fixed modifications, and the data was selected on a maximum 0.01 protein false discovery rate (FDR).

2.5. Chemometrics for *m-DoE*

In this study, *m-DoE* consists of three Fractional Factorial Designs - FracFD (*DoE* 1, 3, and 5) two Full Factorial Designs - FullFD (*DoE* 2 and 4), and the response surface by central composite design (CCD). Key experiments as well as their responses were planned for each dimension evaluated with the following assumptions: if possible, fast experiments; multiple responses evaluated in same magnitude and all experiments with 95% of the confidence interval in the selection of significant factors. All experiments in *DoE*'s were analyzed and selected at random. The software packages of Excel, Statistica, and Matlab were used for tables, response surfaces, variance analysis, and equations. More details are described in **Supporting Information**.

3. RESULTS AND DISCUSSION

3.1. Multi-DoE with Multi-Responses (*m-DoE*)

The *m-DoE* consists of evaluating the system by individual modules. Therefore, *m-DoE* is recommended for hyphenated analytical techniques, where the method can be compartmentalized. The optimization through *m-DoE* can be carried out using FullFD or FracFD to each module. Thus, the main effects and their cross-interactions of each module are evaluated. **Figure 1A** illustrates the approach made in this study.[19] In addition, different types of responses can be used in each module through *Derringer's* desirability function (**Equation 1**) followed by the geometric average of the values (**Equation 2**). Where, A is target value, L is the lower acceptable experimental value from trials, γ is questioned value (obtained in the response), and t is magnitude for a specific response, which we assumed as 1. For values of γ out of range L was used $d = 0$ (**Equation 3**). So the global desirability (D) is function of each individual desirability (d_1, d_2, d_m).[33,34] Thus, for multi-responses evaluated are converted to a unique response in the range of 0 to 1.

It is important to mention that, as *m-DoE* uses various responses in the design, the response of each module can be different, but they must have an analytical consensus to improve the other modules. For example, the response of module 1 (blue box) in **Figure 1A** has an analytical sense to improve the response in module 2 (green box). This concept is summarized in **Figure 1B**.

We can infer that *m-DoE* strategy has more information on the evaluated system. Consequently, *m-DoE* is more precise in the selection of factors or levels than PBD. However, in *m-DoE*, it is necessary to increase the number of trials, because the interactions of factors are considered. On the other hand, *m-DoE* becomes advantageous over the One Factor At a Time (OFAT) approach, in terms of number of trials.[35,36] Moreover, when the planning of PBD is inappropriate owing to the

unfamiliarity of the significant factors and their ranges, *m-DoE* outperforms PBD. In addition, the *m-DoE* strategy has a robust way to evaluate output responses through the desirability function.

Figure 1. Illustration of chemometrics strategy of multi-design of experiment with multi-responses (*m-DoE*) for LC×LC-MS/MS. **1A.** Each cube represents one technique used in this study. Thus, the blue cube represents the SCX column, the green cube represents the reversed-phase column (RP column) and the purple cube represents the tandem mass spectrometry platform (MS/MS). The number of responses in each cube indicates the number of different responses evaluated for each dimension. **1B.** The strategy of *m-DoE* to calculate multiple responses in the analytical hyphenated methods in *shotgun* proteomics. The dimensions mean different techniques used. Thus, the optimization of the 1st Dimension for increase sample recovery, it means to optimize indirectly the 2nd Dimension, where the goal is increasing the number of peaks (peptides), and consequently, it means to optimize indirectly the 3rd Dimension, once the goal is to increase coverage of proteins. The optimization of each module increases the knowledge on the chemical information on the evaluated system, so it is more precise to select the exact range levels of factors investigated than PBD.

$$d = \left(\frac{\gamma - L}{A - L} \right)^t \quad \text{when } L \leq \gamma \leq A; \quad (1)$$

$$D = \sqrt[m]{d_1 d_2 \dots d_m} \quad (2)$$

$$d = 0 \quad \text{when } \gamma \text{ out range of } L \quad (3)$$

3.2. Application of *m-DoE* in the LC×LC-MS/MS method

3.2.1. Evaluation of the significant factors for 1stD-LC

We performed a fast, simple, and efficient test to verify the relevant factors, their cross-effects, and adequate levels of SCX column that could represent significance in the outcome (system response). It was the first module for screening the instrumental parameters of **Figure 1A** (*DoE 1*). Therefore, we used the detection by UV-Vis spectrophotometry at 214 nm to get real sample recovery. Integration of the chromatographic peak for each injected plug salt, or the final column cleaning, provided the input parameter for the response factor. **Figure S3, Supporting Information** illustrates the results and **Table 1** indicates the factors and levels evaluated in this factorial design.

Table 1. Independent factor levels for *DoE 1*, 1st D-LC evaluation

Code	Factors	Level		Reasoning
		Low (-1)	High (+1)	
1	Buffer concentration (mmol L ⁻¹)	5	15	Ionic strength and buffer capacity
2	%ACN in buffer	5	15	To decrease non-specific hydrophobic interaction on the SCX stationary phase
3	pH	2.5	3.0	
4	Flow rate (μL min ⁻¹)	50	100	Influences on the partition equilibrium in terms of linear velocity in the van Deemter equation
5	Sample concentration (mg mL ⁻¹)	0.63	1.00	Sensitivity and peak shape

DoE 1 was performed with a FracFD 2^{v5-1} with 16 trials and 4 replicates (**Table S3, Supporting Information**).^[25] The response (peak area) is the sum of the eight peaks areas that refer to all fractions of salt plug separated on SCX column (for example sum of peaks areas of **Figure S3, Supporting Information**). Thus, large peak areas indicate less peptide lost in the 1st D system.

The Pareto chart, normal plot and contribution percentage (**Figure S11A, S11B and S11C Supporting Information**) reveal the weight of the contribution of each factor to the response and possible significant interaction effects among these variables. Among the factors evaluated in *DoE 1*, only the acetonitrile percentage at the loading buffer showed a significant and positive effect. The Pareto Chart, normal plot and contribution percentage (**Figures S11A, S11B and S11C Supporting Information**) indicate that the best condition is to increase %ACN of mobile phase in 1stD. We verified that the third-order interactions generate confusion. These are linear combinations of other effects and cannot be estimated. Because of that, third-order interactions are redundant effects, in other words, they are nonsignificant in the design and can be neglected to resolve ambiguities. This behavior follows sparsity-of-effects principle.^[37–39] To improve the performance of the other steps of the optimization process, we maintained the following values of factors: 15% ACN on buffer due to better sample recovery; phosphate buffer concentration of 15 mmol L⁻¹ to increase buffer capacity; pH 3 according to Palma et al.,^[16] flow rate of 125 μL min⁻¹ for fast equilibration and short analysis time; the sample concentration of 1 mg mL⁻¹ to keep good detectability.

3.2.2. Evaluation of the significant factors for 2ndD-LC

The capillary hydrophobic column, the second module of **Figure 1A** (*DoE 2*), was performed with a similar strategy to that of *DoE 1*, however, the 2ndD separation was carried out (**Figure S4, Supporting Information**). At this time, we used an organic gradient with acetonitrile for separative elution of the compounds from the C₁₈ column, followed by UV/vis

detection at 214 nm. With this approach, *DoE* 2 was evaluated for two responses: number of peaks (pn_2) and peak capacity (pc_2), both values were taken only from the second dimension, by means of **Equation S2, Supporting Information**, based on Snyder *et al.* with slight modification.[40] As the system was evaluated by multiple responses, it was necessary to apply a desirability function as described by Derringer and Suich,[33] towards maximum response (**Equation 1**), followed by the geometric average of the values (**Equation 2 or 3**). **Table 2** indicates the factors and levels evaluated in module 2 from **Figure 1**. *DoE* 2 was performed with a FullFD 2^3 with 8 assays and 2 additional replicates (**Table S4, Supporting Information**).

Table 2. Independent factors levels for 2ndD-LC evaluation

Factor	Level		Reasoning
	Low (-1)	High (+1)	
1 %ACN in the gradient	40.00	65.00	Peak capacity and analysis time
2 Flow rate in 2D-LC ($\mu\text{L min}^{-1}$)	5.00	11.00	Influences on the partition equilibrium in terms of linear velocity in the van Deemter equation
3 Temperature ($^{\circ}\text{C}$)	35.00	42.00	Solubility of peptides and partition equilibrium with stationary phase

The Pareto Chart, normal plot and contribution percentage (**Figure S12A, S12B and S12C Supporting Information**) show that all main factors were significant as well as some second-order interactions (C13 and C23). It is interesting to note that this behavior would not be observed in PBD, as mentioned in **Background Design of Experiments Topic 1.2 in SI**, because PBD does not evaluate the interaction effects of second-order or higher. In addition, Riter *et al.* have reported that it is important to elucidate interaction effects on hyphenated analytical techniques, because it is quite common to occur significant interactions of variables (factors) in the optimization of instrumental parameters.[25] Our study indicated that, high levels of organic mobile phase gradient have positive effect on peaks numbers and peak capacity. Temperature and flow rate have a negative effect to increase the responses.

3.2.3. Evaluation of the significant factors in the whole LC×LC setup

In the sequence, we evaluated the factors that could be significant in the whole LC×LC platform. With this assumption, we performed in the *DoE* 3, a FracFD $2\sqrt{5-1}$ with 16 assays and 4 replicates (**Table 3 and Table S5, Supporting Information**).[25] We evaluated the peak number (pn_2) and the peak capacity (pc_2) only of the second separation dimension by UV/vis detection. The better pc_2 value was 257 that is a great value for a capillary chromatography column scale compare to the microbore column scale ($pc = 280$), where usually it easier to obtain the highest values of pc . [41] The

values were converted through the global desirability (D), equations 1, 2 and 3.[34] The total time of each analysis was 6 hours.

Table 3. Independent factors and levels for LC×LC evaluation

Code	Dimension	Factors	Level	
			Low (−1)	High (+1)
1	1 st D-SCX	% ACN in buffer	15	25
2	1 st D-SCX	Flow rate (μL min ^{−1})	100	150
3	2 nd D-C ₁₈	% ACN in the gradient	45	70
4	2 nd D-C ₁₈	Temperature (°C)	30	38
5	2 nd D-C ₁₈	Flow rate (μL min ^{−1})	5	11

We observed that there were similar values of cross-effect of second-order, such as C14 and C23, as well as, C25 and C35 contrasts (**Table S5, Supporting Information**). Such values are redundant, and they do not generate a relevant response to the model. Therefore, it becomes necessary to carry out calculations only with the main factors.[38,39] With 95% of the confidence interval, the Pareto Chart, normal plot of main effects and contribution percentage were evaluated (**Figure S13A, S13B and S13C Supporting Information**). We confirmed that %ACN on sodium phosphate buffer is really the most significant factor in the whole LC×LC system. Thus, the 1D-SCX-LC showed more relevance than 2D-C₁₈-LC module. A positive effect to flow rate on 2nd D was noted; probably, higher velocity of mobile phase keeps the entropic effects (dispersion) smaller, according to the *van Deemter* equation and therefore peaks width is smaller.[19,42]

3.2.4. Evaluation of the factors on the autoMS/MS platform

The third module of *m-DoE* of **Figure 1A**, the evaluation of factors on the autoMS/MS platform (*DoE 4*), was investigated by an initial screening with fast trials. The workflow consisted of two main stages: *i*) sample injection and elution with phosphate buffer on 1stD, followed by desalting on trap-column and organic gradient separation on 2ndD, and detection by autoMS/MS; *ii*) elution with 3.5 M of NaCl, followed by desalting on the trap-column, separation with organic gradient on 2ndD, and detection by autoMS/MS. Therefore, the aliquots of salt plug were not performed. The time of each analysis was 2 hours.

Table 4 indicates the factors analyzed and their levels on the *DoE 4*. In this design, we evaluated a FullFD 2⁴ with 16 assays without replicates. The responses evaluated were intensity (I) and ion profile chromatogram (PC) by assigning a score 1 to 10 (**Table S6, Supporting Information**).

Table 4. Independent factors and levels for LC×LC-MS/MS evaluation

Code	Dimension		Levels		Reasoning
			Low (−1)	High (+1)	
1	MS/MS	Capillary voltage (kV)	1.6	4.5	Aerosol formation
2	1D-SCX	Sample concentration (mg mL ^{−1})	0.50	1.00	Sensitivity and peak shape
3	2D-C ₁₈	Flow rate C ₁₈ (μL min ^{−1})	5	10	Influences on the partition equilibrium in terms of linear velocity in the van Deemter equation to LC and aerosol formation to MS
4	MS/MS	Transfer time (μs)	70	120	Increases resolution and decreases noise

Figures S14A, S14B and S14C Supporting Information shows that increasing the capillary voltage and sample concentration are the significant factors in *DoE 4*. However, by analysis of the results and ion chromatogram interpretation, we verified that the range levels of the effects of the 1st order were not adequate in the chemometrics study. The capillary voltage value of 1.6 kV did not produce good shape of peaks in the ion chromatograms. Thereby, there was a misleading tendency to maximum values of response. In other words, the statistical analysis indicated the capillary voltage as a significant factor. However, the resulting chemical analysis was not adequate for optimization, because of the inadequate selection of range levels. Here, we can introduce the PBD limitation: if its matrix has inadequate range(s) level(s) the results will be seriously biased.[32,43] Thus, the analyst has two options: *i*) accept the results as true, or *ii*) build another PBD matrix. Certainly, it will lead to more laboratory work, which breaks the concept of PBD to conduct fewer experiments.[32,37] In relation to the factor 2 (sample concentration), it is expected that high concentration injected in LC×LC-MS/MS method should increase the signal intensity. However, this *DoE* shows that inadequate values of parameters have a stronger influence on results since factor 1 (capillary voltage) is about two times more significant than factor 2 (Figure S14A, Supporting Information).

3.2.5. Evaluation of the significant factors for the whole LC×LC-MS/MS method

The analysis of the fourth *DoE* allowed us to obtain a great deal of knowledge about LC×LC-MS/MS platform in terms of the significant factors and their range levels. However, a true optimization through *DoE* must be evaluated in whole LC×LC-MS/MS method, since the factors of one platform can be more significant than in other platforms evaluated in a hyphenated method. In addition, we need to have in mind that usually, the optimization by *DoE* follows the *Pareto Principle* and the effect of sparsity. Thus, the whole method must be evaluated in conjunction. Here the aim is to avoid biased results on chemometrics performance. At this point, we learned that %ACN on loading buffer on 1stD is one significant factor in optimization, and the level of capillary voltage should have a more thorough investigation in its range levels for a better sense of its significance. Table S9, Supporting Information shows the values factors maintained and

the reasoning for better response on the whole LC×LC-MS/MS method. **Table 5** shows the factors analyzed, their levels, and reasoning in the *DoE 5*.

Table 5. Factors evaluated on the whole method LC×LC-MS/MS

Code	Module	Factor	Level		Reasoning
			Low (-1)	High (+1)	
1	SCX	%ACN in buffer	10	20	Significant factor in others DoE's
2	C ₁₈	%ACN on gradient	40	65	Can influence the spray formation due to less superficial tension of drop and separation of peptides
3	ESI-MS	Capillary voltage (kV)	2.8	4.5	Improved aerosol formation
4	MS/MS	Dry temperature (°C)	180	250	Improved conditions for the analyte to go to the gaseous phase via the analyte desolvation
5	MS/MS	Nebulizer pressure (bar)	0.4	0.8	Can influence the spray formation and size of drop to ionization

Table S7, Supporting Information indicates the factors and the range investigated in *DoE 5*. FracFD 2_{III}⁵⁻² (resolution of 3) with 8 assays and 2 replicates. In *DoE 5* we evaluated three responses through bioinformatics (software *Sequest*): *i*) score, *ii*) coverage (C) and *iii*) number of peptides (NP). **Equations 1, 2, and 3** were applied in function of global desirability (D). Values of maximum (A) in **Equation 1** were obtained from *Software ExPASy Peptide Cutter*, indicating the virtual number of peptides by tryptic BSA protein digested was 78. The time of each analysis in LC×LC-MS/MS method was 6 hours and 20 minutes.

There were confounding interactions, *i.e.*, contrasts of same value (**Table S7, Supporting Information**). Therefore, it becomes necessary to carry out calculations only with the main factors.[38,39] The Pareto Chart, normal plot and contribution percentage are products only of main factors (**Figure S15, Supporting Information**). Significant factors were %ACN on the buffer (**Code 1 from Table 5**) and gas nebulizer (**Code 5 from Table 5**). *DoE 5* showed interesting results in terms of the elucidation of the properties of method and on the limitations of PBD, such as: *i*) capillary voltage is not a significant factor, *ii*) the significant factors of *DoE 5* overlap the significant factors of *DoE 4*, *iii*) the nebulizer gas value is not equivalent to the optimal conditions of the manufacturer's protocol, probably the sample properties can influence this parameter, and *iv*) in *DoE 5*, as well as *DoE 3*, the %ACN on buffer had negative effect. Nevertheless, in *DoE 1* the effect was positive. Thus, we can infer that we are close to the maximum response. These assumptions strengthen our mention of the increasing of the chemical knowledge on the evaluated system in **Figure 1B**.

3.2.6. Central composite design

A central composite design (CCD) was chosen as the mathematical model since it is the best strategy using two factors with fewer numbers of trials. Furthermore, CCD analyzes a larger area of response surface than Box-Behnken and Doehlert Designs.[22,27,34] **Table 6** indicates the factors and levels evaluated in the CCD. **Table S8, Supporting Information** indicates the range level of the experiments and their responses evaluated: score, protein coverage, and number of peptides. We used again the desirability function (**Equations 1, 2 and 3**). **Figure 2** shows the response surface as a function of two significant factors. The time of each analysis was 6 hours and 30 minutes.

Table 6. Factors evaluated in the Central Composite Design of a polynomial for a (hyper)ellipsoid with runs of the 2^2 and $\alpha_1 = \pm 1.67$ and $\alpha_2 = 1.5$

Code	Module	Factor	$-\alpha$	-1	0	$+1$	$+\alpha$
1	SCX	%ACN on Buffer	8	10	13	16	18
2	MS/MS	Nebulizer gas (bar)	0.6	0.8	1.2	1.6	1.8

Figure 2. Central composite design for the LCxLC-MS/MS method **A**) Standardized main effect Pareto Chart for BSA sequence coverage. Vertical line in the chart defines 95% confidence level. **B**) Response surface (saddle surfaces) as a function of two factors for BSA sequence coverage. The numbers in Figure **A** refers to the factors coded in **Table 6**. The letters L and Q in y-coordinate means linear function and quadratic function of response surface, respectively.

Figure 2 suggests that the selection of factor levels was suitable. Imagine that we are walking on response surface to maximum zone, which is linked with assay 5 of **Table S8, Supporting Information**, if we go straight there will a decline to fewer response values. Because we verified in *DoE 1* where the amount of 5% of ACN on buffer causes a decrease of sample recovery. The same fact is observed in another maximum response, which is linked with experiment 6 of the **Table S8, Supporting Information** and **Figure 2B** because the *DoE 3* used 25% ACN on the buffer and there was a strong decrease of response. Here, we can introduce the main idea of the power of *m-DoE*; the chemical discernment is increased (**Figure 1B**). Thus, we can affirm that with *m-DoE* we can achieve true and exact optimization. Moreover, this particular

study was a challenging situation because presents two maximum responses; probably it could be harder to PBD detect the behavior of saddle surface from model (**Figure S2 C, Supporting Information**).

In addition, the design showed the rotatability behavior proposed by Box and Hunter, because the selection of $\alpha_1 = \pm 1.67$ and $\alpha_2 = 1.5$ from **Table 6** infer the design of polynomial with a (hyper)ellipsoid instead of a (hyper)sphere $\alpha_1 = \alpha_2 \pm 1.41$. A (hyper)ellipsoid design is a more suitable representation for a saddle (hyper)surface (two maxima response) instead a (hyper)surface with one maximum response. Rotatability behavior was validated because of the low range variance obtained (0.05-0.09) of the coefficients from **Equation 5**. [44–46] It is important to mention that choosing α can be done for its convenience and feasibility as described by Box and Hunter and reviewed by Ferreira, S.L.C. *et al.* [45,46] Here we can introduce the symmetrical fusion between chemical experimental condition to math prediction which means a perfect chemometric synergism achieved by the *m-DoE* strategy.

Equation 5, that includes coefficient and confidence interval, were measured manually in Excel because it is more precise than Software. Through it, we can predict the maximal response. Let us replace the values $x = +\alpha$ and $y = 0$ in **Equation 5**, where these values present better conditions of the response surface of **Figure 2B**. Thus, it is possible to obtain the Global Desirability value of 0.75. This value is equivalent to the best condition of **Table S8, Supporting Information** (trial 5).

$$Z = 0.35(\pm 0.09) - 0.12x(\pm 0.05) + 0.10y(\pm 0.05) + 0.12x^2(\pm 0.06) - 0.08y^2(\pm 0.06) \quad (5)$$

The most reliable way to evaluate the quality of the model fitted is by variance analysis (ANOVA) of the quadratic model (**Table 7**). Here we can compare the variation due to the change in the combination of variables levels with the variation due to random error inherent to the measurements of the generated responses. [47–49] As the calculated Fisher distribution (Test F_{cal}), the square sum for the regression (R) is higher than tabulated value of Fisher distribution (F_{value}), we can affirm that the model is robust and has a good fit. Thus, Mean Square of regression is not confounded with Mean Square of residuals because of these variances are different ($MSR > MSr$) with 95% of confidence. Furthermore, 93% of the total variance around of means is explained by model regression, in other words, the mathematical model is able to model 93% of results, so 7% is residual. Therefore, 7% of the model of our study cannot be modeled. The determination coefficient R^2_{max} indicates that it is possible to model until 99% of the response. For R to have a significant value, the value of pure error must be low. Since F_{value} is higher than F_{cal} of pure error, the model is adequate. [48,50] Moreover, the three central points of CCD, trials 10; 11 and 12 of **Table 7**, had good reproducibility on the replicates (SD = 0.02 with relative standard deviations of 5.71% for global desirability, **Excel Spreadsheet Tables S10-S14**). We can affirm that the model has a good fit. Then, the surface response is well aligned with the experimental results.

Table 7. Variance Analyses (ANOVA) of the quadratic model calculated in Excel.

	Sum Square (SS)	Degree of Freedom (d.f.)	Mean Square (MS)	Test F _{cal}	F _{value} (95%)
Regression (R)	0.35	4	0.09	21.81	4.12
Residuals (r)	0.03	7	0.00		
Total (R_{ss} + r_{ss})	0.38	11	0.03		
Pure error	0.00	3	0.00	5.31	6.59
Lack of fit	0.02	4	0.01		
R²	0.93	R	0.96		
R²_{max}	0.99	R_{max}	1.00		

Let us mention that the studies of *DoE* in liquid chromatography and mass spectrometry platforms to proteomics are incipient. Andrews, *et al.* used 43 different experiments and 2 *DoE* to optimize 9 instrumental parameters of LTQ-Orbitrap (MS/MS) platform. They concluded that: *i*) monoisotopic precursor selection, *ii*) ion trap maximum ionization time, *iii*) tube lens voltage and *iv*) capillary temperature were preferred for increasing the proteins coverage.[20] Randall, *et al.* did 54 different combinations of 14 instruments parameters with 3 *DoE*, where the factors responsible for increasing proteome coverage were: *i*) normalized collision energy, *ii*) capillary temperature, *iii*) isolation window ion accumulation time on duty cycle, *iv*) number of MS/MS acquired, and *v*) dynamic range.[28] The two studies did not propose a response surface and they did not evaluate the different responses using desirability function. However, detailed descriptions of instrumentation and MS parameters to increase proteome coverage were very well elucidated. Novotná, K. *et al.*, optimized the one-dimensional liquid chromatography platform by 32 experiments to 3 factors with 2 CCD, where the resolution and number of peaks were investigated. They considered that gradient conditions were significant factors in this platform.[51] We used in *m-DoE* the amount of 64 experiments and 13 replicates for evaluation of 12 instrumentals parameters. If we consider a PBD model for 11 factors for 12 experiments, *m-DoE* had 5.33 times more experiments than PBD. On the other hand, we can affirm the good optimization of *m-DoE* since the LC×LC-MS/MS method obtained great coverage of proteins (92.92%, score 3816.29, 73 peptides and 32 unique peptides - **Figure S16 and Excel spreadsheet Table S15, Supporting Information**) as well as good results in ANOVA table. In addition, we can infer that *m-DoE* uses a consolidated math of *DoE*, which makes it easy application for the chemometrics community.

Another important observation on *m-DoE* is that the multiple responses must have a connection among them. So, the better response obtained in a platform must improve the response in the other platform (**Figure S17, Supporting Information**). For example, the sample recovery evaluated in *DoE 1* at 1D-LC is linked with Number of Peak (*DoE 2* and 3) at 2D-LC as well as Number of Peptides and Coverage (*DoE 5* and CCD) at MS/MS. In special, the great value of score is linked with intensity and high quality of ion chromatogram (*DoE 4*) at MS/MS as well as peak capacity (*DoE 2* and 4) at LC×LC. Thus, we may consider that chemometrics evaluation separated from each platform reduces the complexity of optimization,

in terms of selection of factors and the range of the levels studied. Therefore, this observation must be noted in future studies. **Excel Spreadsheet Table S16, Supporting Information** shows the values of the Pearson correlation coefficient and covariance from multiples responses evaluated of each DoE or cube from **Figure 1A**. The two responses of the green cube have a moderate correlation (0.5 – 0.7) between them, while the purple cube has moderate, high (0.7 – 0.9) and a very high correlation (0.9 – 1.0) between them. The covariance showed a large range between the response of dimensions 718 – 659,393.75. It is important to mention that the evaluating of multiples responses decreases significantly random errors, mainly with a low correlation between responses from the same platform, consequently increase the accuracy and precision of results that infer a true optimization.[52]

We applied our model to a mixture of six digested proteins and their isoforms on the SCX×RP-MS/MS method. For that, it was evaluated the identical conditions of **Table 6** and **Figure 2**. **Figure 3** shows a close mathematical correlation between **Figure 2**, which is driven by the **equation S2, Supporting Information**. It is noted that nebulizer pressure in 1.2 bars holds as the ideal condition of method autoMS/MS. Furthermore, the use of 8% ACN in buffer mobile phase in the first-dimension chromatography is the best condition. Moreover, we obtained the identification of isoforms for casein (α S1, α S2, κ) and hemoglobin human (HBB, HbD, HBG2, HbA). Considering 11 species identified between proteins, phosphoprotein and, isoforms, the results showed an average coverage of 83.07% overall species analyzed, reaching up to 100% coverage for 45.45% of the sample mix (**Excel Spreadsheet Table S17, Supporting Information**). The behavior of

Figure 3. Central composite design for the LC×LC-MS/MS method **A)** Standardized main effect Pareto Chart for six proteins sequence coverage. Vertical line in the chart defines 95% confidence level. **B)** Response surface as a function of two factors for digested of six proteins. The numbers in Figure **A** refers to the factors coded in **Table 6**. The letters L and Q in y-coordinate means linear function and quadratic function of response surface, respectively.

the mathematical surface curves to a single maximum region. The reasoning for the bottom plateau in <0.3 (dark green region) is due to loss of some peptides, because of a high percentage of organic solvent, and the use of an inadequate condition of nebulizer pressure.

Figure 4 and Excel Spreadsheet Tables S18 and S19, Supporting Information, show the high orthogonality of the SCX×RP-MS/MS method, which were evaluated following the studies of Gilar, M. *et al.* and Motoyama, A. *et al.*[53,54] These figures clearly show the separation space is quasi-fully occupied mainly in a more complex sample. They are expressive results compared to contour plots from the most influential literature of the strategies for analytical scale 2D separations.[55] Additionally, **Figure 5** shows the great quality of the ion chromatogram of the method in terms of intensity and multidimensional separation. It is very known such performance indicates an efficient optimization, consequently the highest protein coverage. It is essential to mention, that the prominent advances in terms of sampling time for 2D LC faster analysis using univariate optimization strategies must be substantially considered for further studies. Probably, in this case, opposite our method, the first dimension must be a column of high peak capacity, such as RP.[56–58]

Figure 4. Normalized retention time plots for selected LC×LC-MS/MS method. A) digested serum albumin bovine protein. B) digested mix of proteins. The points in area space are unique peptides detected from MS/MS. The high coverage of peaks in all area space justifies high orthogonality and the use of multiplication in LC×LC.

It is interesting to mention that 100% coverage in *shotgun* proteomics is one of the main challenges of this field and *m-DoE* was a great strategy to evolve in on-line hyphenated techniques optimization. Moreover, to our best knowledge, this was the first study that evaluated the whole platform LC×LC-MS/MS by chemometrics tools for *shotgun* proteomics and additionally showed two surfaces of response with adequate results demonstrated by ANOVA Table, which follows

completely the *Pareto Principle*, once the two significant factors represent 16.17% out of the 12 factors evaluated. In addition, high coverage of proteins 92.92% was obtained, (when the assays 5 and 9 in **Figure S16 and Excel Spreadsheet Table S15, Supporting Information** are sequenced together). Let us mention that the description of 100% of BSA sequence coverage in the literature is absent. Recent studies on the development of new unidimensional methods, where the tendency to lose peptides is lower, showed a range of 67 to 90% for BSA coverage.[59–61] As the tendency of losing peptides in multidimensional methods is larger than unidimensional methods, m-DoE showed an efficient strategy to increase the proteins coverage, which it was validated with a complex mixture of six digested proteins and their isoforms. We expect m-DoE approach could improve the understanding of dynamics in multi-dimensional proteomics analysis in terms of a higher number of identifications for proteins, isoforms and proteoforms coming from biological systems.[1]

Figure 5. Optimized total ion chromatogram from m-DoE of digested proteins as sample applied to LC×LC-MS/MS method. A) Better method condition from the response surface of digested serum albumin bovine protein. B) Better method condition from the response surface of digested six mix protein.

4. CONCLUSION

We propose a new chemometrics strategy to design of experiments applied to hyphenated analytical techniques of greater complexity to optimization, such as 2D LC-MS/MS. We called this strategy Multi-Design of Experiment with Multi-Responses (*m-DoE*). *m-DoE* is an alternative to the limitations of the saturated design proposed by *Plackett-Burman design*. Therefore, *m-DoE* attacks two principles where PBD is limited: *i*) effective detection of significant interactions of factors analyzed on model and *ii*) better selection of factors and its range levels. Thus, *m-DoE* can have a larger number of experiments than PBD. However, the chemical information on the system is increased and the discernment to true optimization has more chances of success. We can infer that PBD is the best strategy when the ranges of levels factors are

well known, in other words, the range of chemical conditions is clear. When this fact is limited, we recommend the use of *m-DoE*. Therefore, *m-DoE* is useful in optimizations with ≥ 7 factors for a system with little chemical information on the factors or range levels that are relevant to the system. It is important to mention that in *m-DoE* is necessary that the responses of modules have a strategy of connection to improve the response of another module. Thus, when a module is being optimized, this must have an indirect optimization to improve other modules. The surface response was evaluated by Analysis of Variance (ANOVA) that showed the great statistical significance of the model where it indicated two maximum responses to optimization. When the two maxima regions in **Figure 2** were sequenced in the set, the results were 92.92% of coverage, score of 3816.29, 73 peptides and 32 unique peptides to BSA protein with good reproducibility of the method. In addition, the choosing of the values of axial or star points $\pm\alpha$ for a (hyper)ellipsoid drove to low values of variance of coefficients of the model, which indicates a rotatability behavior suitable. Moreover, a further and more challenging application of *m-DoE* yielded an average coverage of 83.07% for 11 species identified (between proteins, phosphoprotein, and isoforms), reaching up to 100% coverage for 45.45% of analyzed species. In addition, we showed that the chemometrics performance of the LC×LC-MS/MS method followed the *Pareto Principle*, which is a normal tendency that occurs in math. Lastly, the high orthogonality was shown of the LC×LC-MS/MS method in the 2D plot evidencing the optimization efficiency of *m-DoE*. This study contributes to reaching better analysis of biological samples and performance intrinsic of instrumental parameters for the most different categories of hyphenated analytical systems. A scenery that *m-DoE* approach could be useful it would be in Multi-dimensional proteomics for more comprehensive cell biology analysis in terms of increasing molecular detection.

ACKNOWLEDGEMENTS

Financial support from Council for Scientific and Technological Development (CNPq), grant no. [131311/2013-1](#) to WPB) and São Paulo Research Foundation (FAPESP), grant no. [2015/16025-8](#) to WPB, grant no. [2010/19910-9](#) to AJSN) are gratefully acknowledged. Moreover, INCTBio is kindly acknowledged for the grants from FAPESP (grant no. [2014/50867-3](#)) and CNPq (grant no. [465389/2014-7](#)). We thank Professor Edenir Pereira Filho (UFSCar) for his pleasant classes and great discussions about DoE and chemometrics in general.

REFERENCES

- [1] M. Larance, A.I. Lamond, Multidimensional proteomics for cell biology, *Nat. Rev. Mol. Cell Biol.* 16 (2015) 269–280. doi:10.1038/nrm3970.
- [2] E. Stokstad, NOBEL PRIZES. DNA's repair tricks win chemistry's top prize, *Science*. 350 (2015) 266.

doi:10.1126/science.350.6258.266.

- [3] Y. Zhang, B.R. Fonslow, B. Shan, M. Baek, J.R. Yates, Protein analysis by shotgun/bottom-up proteomics, *Chem. Rev.* 113 (2013) 2343–2394. doi:10.1021/cr3003533.
- [4] M. Tyers, M. Mann, From genomics to proteomics, *Nature*. 422 (2003) 193–7. doi:10.1038/nature01510.
- [5] J.C. Giddings, Two-dimensional separations: concept and promise, *Anal. Chem.* 56 (1984) 1258A-1270A. doi:10.1021/ac00276a003.
- [6] R.J. Vonk, A.F.G. Gargano, E. Davydova, H.L. Dekker, S. Eeltink, L.J. de Koning, P.J. Schoenmakers, Comprehensive two-dimensional liquid chromatography with stationary-phase-assisted modulation coupled to high-resolution mass spectrometry applied to proteome analysis of *Saccharomyces cerevisiae*, *Anal. Chem.* 87 (2015) 5387–5394. doi:10.1021/acs.analchem.5b00708.
- [7] K.M. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, Review of chemometric analysis techniques for comprehensive two dimensional separations data, *J. Chromatogr. A.* 1255 (2012) 3–11. doi:10.1016/j.chroma.2012.05.050.
- [8] J.S. Mellors, W.A. Black, A.G. Chambers, J.A. Starkey, N.A. Lacher, J.M. Ramsey, Fully integrated glass microfluidic device for performing high-efficiency capillary electrophoresis and electrospray ionization mass spectrometry, *Anal. Chem.* 85 (2013) 4100–4106. doi:10.1021/ac400205a.
- [9] A.C. Johnson, M.T. Bowser, High-speed, comprehensive, two dimensional separations of peptides and small molecule biological amines using capillary electrophoresis coupled with micro free flow electrophoresis, *Anal. Chem.* 89 (2017) 1665–1673. doi:10.1021/acs.analchem.6b03768.
- [10] N.M. Cassiano, Q.B. Cass, Direct bioanalytical sample injection with 2D LC – MS, *Bioanalysis.* 4 (2012) 2737–2756. doi:10.4155/bio.12.226.
- [11] B.L. Karger, L.R. Snyder, C. Horvath, *An Introduction to Separation Science*, 1973.
- [12] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates, Direct analysis of protein complexes using mass spectrometry, *Nat. Biotechnol.* 17 (1999) 676–682. doi:10.1038/10890.
- [13] D.R. Stoll, X. Li, X. Wang, P.W. Carr, S.E.G. Porter, S.C. Rutan, Fast, comprehensive two-dimensional liquid chromatography, *J. Chromatogr. A.* 1168 (2007) 3–43. doi:10.1016/j.chroma.2007.08.054.
- [14] D.A. Wolters, M.P. Washburn, J.R. Yates, An automated multidimensional protein identification technology for shotgun proteomics, *Anal. Chem.* 73 (2001) 5683–5690. doi:10.1021/ac010617e.
- [15] J.W. Jorgenson, Capillary liquid chromatography at ultrahigh pressures, *Annu. Rev. Anal. Chem.* 3 (2010) 129–150. doi:10.1146/annurev.anchem.1.031207.113014.
- [16] S. Di Palma, M.L. Hennrich, A.J.R. Heck, S. Mohammed, Recent advances in peptide separation by multidimensional liquid chromatography for proteome analysis, *J. Proteomics.* 75 (2012) 3791–3813. doi:10.1016/j.jprot.2012.04.033.
- [17] L. Hu, M. Ye, X. Jiang, S. Feng, H. Zou, Advances in hyphenated analytical techniques for shotgun proteome and peptidome analysis-A review, *Anal. Chim. Acta.* 598 (2007) 193–204. doi:10.1016/j.aca.2007.07.046.
- [18] B.W.J. Pirok, D.R. Stoll, P.J. Schoenmakers, Recent developments in two-dimensional liquid chromatography: fundamental improvements for practical applications, *Anal. Chem.* 91 (2018) 240–263. doi:10.1021/acs.analchem.8b04841.

- [19] W.P. Batiston, Chemometric approach to a capillary multidimensional liquid chromatography coupled to tandem mass spectrometry for shotgun proteomics, University of São Paulo, 2015.
- [20] G.L. Andrews, R.A. Dean, A.M. Hawkrige, D.C. Muddiman, Improving proteome coverage on a LTQ-orbitrap using design of experiments, *J. Am. Soc. Mass Spectrom.* 22 (2011) 773–783. doi:10.1007/s13361-011-0075-2.
- [21] R.E.B. Neto, Benicio de Barros, Ieda Spacino Scarminio, Como fazer experimentos. Pesquisa e desenvolvimento na ciência e na indústria, 4th ed., Campinas, 2006.
- [22] D.B. Hibbert, Experimental design in chromatography: A tutorial review, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 910 (2012) 2–13. doi:10.1016/j.jchromb.2012.01.020.
- [23] R. Leardi, Experimental design in chemistry: A tutorial, *Anal. Chim. Acta.* 652 (2009) 161–172. doi:10.1016/j.aca.2009.06.015.
- [24] A.L. Pomerantsev, O.Y. Rodionova, Process analytical technology: A critical view of the chemometricians, *J. Chemom.* 26 (2012) 299–310. doi:10.1002/cem.2445.
- [25] L.S. Riter, O. Vitek, K.M. Gooding, B.D. Hodge, R.K. Julian, Statistical design of experiments as a tool in mass spectrometry, *J. Mass Spectrom.* 20 (2005) 565–579. doi:10.1002/jms.871.
- [26] G.M. Whitesides, Reinventing chemistry, *Angew. Chemie - Int. Ed.* 54 (2015) 3196–3209. doi:10.1002/anie.201410884.
- [27] S.L.C. Ferreira, R.E. Bruns, H.S. Ferreira, G.D. Matos, J.M. David, G.C. Brand, E.G.P. Silva, L.A. Portugal, P.S. Reis, A.S. Souza, W.N.L. Santos, Box-Behnken design: An alternative for the optimization of analytical methods, *Anal. Chim. Acta.* 597 (2007) 179–186. doi:10.1016/j.aca.2007.07.011.
- [28] S.M. Randall, H.L. Cardasis, D.C. Muddiman, Factorial experimental designs elucidate significant variables affecting data acquisition on a quadrupole orbitrap mass spectrometer, *J. Am. Soc. Mass Spectrom.* 24 (2013) 1501–1512. doi:10.1007/s13361-013-0693-y.
- [29] J.R. Yates, C.I. Ruse, A. Nakorchevsky, Proteomics by mass spectrometry: approaches, advances, and applications, *Annu. Rev. Biomed. Eng.* 11 (2009) 49–79. doi:10.1146/annurev-bioeng-061008-124934.
- [30] H. Malerod, E. Lundanes, T. Greibrokk, Recent advances in on-line multidimensional liquid chromatography, *Anal. Methods.* 2 (2010) 110–122. doi:10.1039/b9ay00194h.
- [31] L. Plackett, R. P. Burman, J, The design of optimum multifactorial experiments, *Biometrika Trust.* 33 (1946) 305–325. doi:10.1093/biomet/33.4.305.
- [32] H. Ebrahimi-Najafabadi, R. Leardi, M. Jalali-Heravi, Experimental design in analytical chemistry - Part I: theory, *J. AOAC Int.* 97 (2014) 3–11. doi:10.5740/jaoacint.SGEEbrahimi1.
- [33] G. Derringer, R. Suich, Simultaneous-optimization of several response variables, *J. Qual. Technol.* 12 (1980) 214–219. doi:10.1017/CBO9781107415324.004.
- [34] L. Vera Candioti, M.M. De Zan, M.S. Cámara, H.C. Goicoechea, Experimental design and multiple response optimization. Using the desirability function in analytical methods development, *Talanta.* 124 (2014) 123–138. doi:10.1016/j.talanta.2014.01.034.
- [35] L. Moruz, P. Pichler, T. Stranzl, K. Mechtler, L. Käll, Optimized nonlinear gradients for reversed-phase liquid chromatography in shotgun proteomics, *Anal. Chem.* 85 (2013) 7777–7785. doi:10.1021/ac401145q.
- [36] N. Marchetti, J.N. Fairchild, G. Guiochon, Comprehensive off-line, two-dimensional liquid chromatography.

- Application to the separation of peptide digests, *Anal. Chem.* 80 (2008) 2756–2767. doi:10.1021/ac7022662.
- [37] K.R. Quinlan, D.K.J. Lin, Run order considerations for Plackett and Burman designs, *J. Stat. Plan. Inference.* 165 (2015) 56–62. doi:10.1016/j.jspi.2015.04.001.
- [38] G.E.P. Box, R.D. Meyer, An analysis for unreplicated fractional factorials, *Technometrics.* 28 (1986) 11–18. <http://www.jstor.org/stable/1269599>.
- [39] G.E.P. Box, R.D. Meyer, Dispersion effects from fractional designs, *Technometrics.* 28 (2012) 19–27. doi:10.1080/00401706.1986.10488094.
- [40] J.W. Dolan, L.R. Snyder, N.M. Djordjevic, D.W. Hill, T.J. Waeghe, Reversed-phase liquid chromatographic separation of complex samples by optimizing temperature and gradient time I. Peak capacity limitations, *J. Chromatogr. A.* 857 (1999) 1–20. doi:10.1016/s0021-9673(99)00767-0.
- [41] M. Vajrychova, J. Klimentova, Conventional-flow liquid chromatography – mass spectrometry for exploratory bottom-up proteomic analyses, *Anal. Chem.* 90 (2018) 5381–5389. doi:10.1021/acs.analchem.8b00525.
- [42] J.J. van Deemter, F.J. Zuiderweg, A. Klinkenberg, Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography., *Chem. Eng. Sci.* 5 (1956) 271.
- [43] L.B. Barrentine, Illustration of confounding in Plackett-Burman designs, *Qual. Eng.* 9 (1996) 11–20. doi:10.1080/08982119608919013.
- [44] G.E.P. Box, K.B. Wilson, On the experimental attainment of optimum conditions, *J. R. Stat. Soc. Ser. B.* 13 (1951) 1–45.
- [45] G.E.P. Box, J.S. Hunter, Multi-factor experimental designs for exploring response surface, *Ann. Math. Stat.* 28 (1957) 195–241.
- [46] S.L.C. Ferreira, R.E. Bruns, E.G.P. da Silva, W.N.L. dos Santos, C.M. Quintella, J.M. David, J.B. de Andrade, M.C. Breikreitz, I.C.S.F. Jardim, B.B. Neto, Statistical designs and response surface techniques for the optimization of chromatographic systems, *Anal. Chim. Acta.* 1158 (2007) 2–14. doi:10.1016/j.chroma.2007.03.051.
- [47] M. Mulholland, J. Waterhouse, Investigation of the limitations of saturated fractional factorial experimental designs, with confounding effects for an HPLC ruggedness test, *Chromatographia.* 25 (1988) 769–774. doi:10.1007/BF02262082.
- [48] E.R. Pereira, Planejamento fatorial em química: maximizando a obtenção de resultados, Primeira, São Carlos, 2015.
- [49] B. de B. Neto, I.S. Scarminio, R.E. Bruns, Como fazer experimentos pesquisa e desenvolvimento na ciência e na indústria, 2010.
- [50] M.A. Bezerra, R.E. Santelli, E.P. Oliveira, L.S. Villar, L.A. Escaleira, Response surface methodology (RSM) as a tool for optimization in analytical chemistry, *Talanta.* 76 (2008) 965–977. doi:10.1016/j.talanta.2008.05.019.
- [51] K. Novotná, J. Havliš, J. Havel, Optimisation of high performance liquid chromatography separation of neuroprotective peptides: Fractional experimental designs combined with artificial neural networks, *J. Chromatogr. A.* 1096 (2005) 50–57. doi:10.1016/j.chroma.2005.06.048.
- [52] M.M. Mukaka, Statistics Corner : A guide to appropriate use of correlation coefficient in medical research, *Malawi Med. J.* 24 (2012) 69–71.

- [53] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography, 77 (2005) 6426–6434.
- [54] A. Motoyama, T. Xu, C.I. Ruse, J.A. Wohlschlegel, J.R. Yates, Anion and cation mixed-bed ion exchange for enhanced multidimensional separations of peptides and phosphopeptides, *Anal. Chem.* 79 (2007) 3623–3634. doi:10.1021/ac062292d.
- [55] M. Sarrut, A.D. Attoma, S. Heinisch, Optimization of conditions in on-line comprehensive two-dimensional reversed phase liquid chromatography . Experimental comparison with one-dimensional reversed phase liquid chromatography for the separation of peptides, *J. Chromatogr. A.* 1421 (2015) 48–59. doi:10.1016/j.chroma.2015.08.052.
- [56] R.E. Murphy, M.R. Schure, J.P. Foley, Effect of sampling rate on resolution in comprehensive two-dimensional liquid chromatography, *Anal. Chem.* 70 (1998) 1585–1594. doi:10.1021/ac971184b.
- [57] Y. Huang, H. Gu, M. Filgueira, P.W. Carr, An experimental study of sampling time effects on the resolving power of on-line two-dimensional high performance liquid chromatography, *J. Chromatogr. A.* 1218 (2011) 2984–2994. doi:10.1016/j.chroma.2011.03.032.
- [58] D.R. Stoll, X. Wang, P.W. Carr, Comparison of the practical resolving power of liquid chromatography analysis of metabolomic samples, *Anal. Chem.* 80 (2008) 268–278. doi:10.1021/ac701676b.
- [59] X. Zhao, X. Xie, S. Sharma, L.T. Tolley, A. Plistil, H.E. Barnett, M.P. Brisbin, A.C. Swensen, J.C. Price, P.B. Farnsworth, H.D. Tolley, S.D. Stearns, M.L. Lee, Compact ultrahigh-pressure nanoflow capillary liquid chromatograph, *Anal. Chem.* 89 (2017) 807–812. doi:10.1021/acs.analchem.6b03575.
- [60] C. Young, A. V. Podtelejnikov, M.L. Nielsen, Improved reversed phase chromatography of hydrophilic peptides from spatial and temporal changes in column temperature, *J. Proteome Res.* 16 (2017) 2307–2317. doi:10.1021/acs.jproteome.6b01055.
- [61] T.T.T.N. Nguyen, N.J. Petersen, K.D. Rand, A simple sheathless CE-MS interface with a sub-micrometer electrical contact fracture for sensitive analysis of peptide and protein samples, *Anal. Chim. Acta.* 936 (2016) 157–167. doi:10.1016/j.aca.2016.07.002.