PROGRAMA DE PÓS-GRADUAÇÃO EM ALIMENTOS E NUTRIÇÃO

CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE

UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO

Thais de Oliveira Alves

OBSERVATIONS ON THE IMPACT OF FOOD PROCESSING ON THE PROTEIN AND ALLERGENIC PROFILES OF WHEAT-BASED PRODUCTS: A PROTEOMIC STUDY

IMPLICAÇÕES DO PROCESSAMENTO DE ALIMENTOS NOS PERFIS PROTEICOS E ALERGÊNICOS DE PRODUTOS À BASE DE TRIGO: UM ESTUDO PROTEÔMICO

Rio de Janeiro

2023

Thais de Oliveira Alves

Observations on the impact of food processing on the protein and allergenic profiles of wheat-based products: a proteomic study

Doctoral Thesis submitted to the Food and Nutrition Graduate Program, Federal University of the State of Rio de Janeiro as partial requirement for obtaining the Doctoral Degree and PhD title in Food and Nutrition

Supervisor: Prof. Dr. Mariana S. L. Ferreira

Co-supervisor: Dr. Cristina Yoshie Takeiti

Rio de Janeiro

2023

Catalogação informatizada pelo(a) autor(a)

A474

Alves, Thais de Oliveira
OBSERVATIONS ON THE IMPACT OF FOOD PROCESSING ON THE PROTEIN AND ALLERGENIC PROFILES OF WHEAT-BASED PRODUCTS: A PROTEOMIC STUDY / Thais de Oliveira Alves. -- Rio de Janeiro, 2023.
132 f

Orientadora: Mariana Simões Larraz Ferreira.
Coorientadora: Cristina Yoshie Takeiti.
Tese (Doutorado) - Universidade Federal do Estado do Rio de Janeiro, Programa de Pós-Graduação em Alimentos e Nutrição, 2023.

1. foodomics. 2. glutenomics. 3. wheat processing. 4. technological quality. 5. immunogenic profile. I. Ferreira, Mariana Simões Larraz, orient. II. Takeiti, Cristina Yoshie, coorient. III. Título.

Thais de Oliveira Alves

Doctoral Thesis submitted to the Food and Nutrition Graduate Program, Federal University of the State of Rio de Janeiro.

Approved on: 16 / 06 / 2023

JUDGING BOARD

_____

President: Prof. Dr. Mariana Simões Larraz Ferreira – PPGAN/UNIRIO

_____

Prof. Dr. Maria Gabriela Bello Koblitz – PPGAN/UNIRIO

_____

Prof. Dr. Carlos Wanderlei Piler de Carvalho – EMBRAPA CTAA

_____

Dr. Maíra Fasciotti Pinto Lima – INMETRO

_____

Prof. Dr. Talita Pimenta do Nascimento – UNIRIO

ASTEN ASSINATURA

Datas e horários baseados em Brasília, Brasil
Sincronizado com o NTP.br e Observatório Nacional (ON) em
19/07/2023 às 11:01:25

ntp.br

Folha_Assinatura_Tese_Thais Alves

Data e Hora de Criação: 17/07/2023 às 14:52:07

Doocumentos que originaram esse envelope:
- ThaisOAlves_Assinaturas.pdf (Arquivo PDF) - 1 página(s)

Hashs únicas referente à esse envelope de documentos

[SHA256]: f6abe8f2e04fb6d9db96299ef2b4283540fc26eeaadcfe6e7fc957fbfdaf10e4
[SHA512]: d40569835981b212c1888138589b4c11a7a6a0cce0dac1ceb507e0242f3ac1db4bfb81535da54199f8710adbc9212c5ed68638d97a3c9d4b24191297db2dd77b

Lista de assinaturas solicitadas e associadas à esse envelope

ASSINADO - Carlos Wanderley Pller de Carvalho (carlos.pller@embrapa.br)
Data/Hora: 18/07/2023 - 07:37:12, IP: 200.143.246.132
[SHA256]: e8a7eb8ac2670a77059335cdb9719a1fbc994ec2eb6d7d725ed09fcd1f0d7781

ASSINADO - Maria Gabriela Bello Koblitz (maria.koblitz@unirio.br)
Data/Hora: 19/07/2023 - 11:01:25, IP: 200.156.27.160
[SHA256]: 8682eab621fce2730ebeb97415d9c96822c47e55300e92f356d7c44447c0cc54

ASSINADO - Mariana Simões Larraz Ferreira (mariana.ferreira@unirio.br)
Data/Hora: 18/07/2023 - 06:49:45, IP: 147.100.172.65, Geolocalização: [48.8046592, 2.0873216]
[SHA256]: bcd9341c507015c6b6939c3d22255226b07ed4c3210e352381993b77713c7499

ASSINADO - Maira Fasciotti Pinto Lima (mfasciotti@inmetro.gov.br)
Data/Hora: 17/07/2023 - 15:58:28, IP: 200.20.211.4, Geolocalização: [-22.596812, -43.279974]
[SHA256]: 2eded4e82b5717994439ea283d62a880bf178edd59fdb3d62b268d9d7b51bede

ASSINADO - Talita Pimenta do Nascimento (talitapiment@gmail.com)
Data/Hora: 17/07/2023 - 15:21:40, IP: 200.202.168.254, Geolocalização: [-15.731943, -47.900339]
[SHA256]: 9259bf8e39f35d7f776bf25702237b72420200a1a151443a75ed3f2d2e100362

Histórico de eventos registrados neste envelope

19/07/2023 11:01:25 - Envelope finalizado por maria.koblitz@unirio.br, IP 200.156.27.160
19/07/2023 11:01:25 - Assinatura realizada por maria.koblitz@unirio.br, IP 200.156.27.160
19/07/2023 11:01:13 - Envelope visualizado por maria.koblitz@unirio.br, IP 200.156.27.160
18/07/2023 07:37:12 - Assinatura realizada por carlos.pller@embrapa.br, IP 200.143.246.132
18/07/2023 06:49:45 - Assinatura realizada por mariana.ferreira@unirio.br, IP 147.100.172.65
18/07/2023 06:49:10 - Envelope visualizado por mariana.ferreira@unirio.br, IP 147.100.172.65
17/07/2023 15:58:28 - Assinatura realizada por mfasciotti@inmetro.gov.br, IP 200.20.211.4
17/07/2023 15:58:13 - Envelope visualizado por mfasciotti@inmetro.gov.br, IP 200.20.211.4
17/07/2023 15:21:40 - Assinatura realizada por talitapiment@gmail.com, IP 200.202.168.254
17/07/2023 15:21:35 - Envelope visualizado por talitapiment@gmail.com, IP 200.202.168.254
17/07/2023 14:59:21 - Envelope registrado na Blockchain por ppgan.secretaria@unirio.br, IP 200.156.27.158
17/07/2023 14:59:20 - Envelope encaminhado para assinaturas por ppgan.secretaria@unirio.br, IP 200.156.27.158
17/07/2023 14:52:08 - Envelope criado por ppgan.secretaria@unirio.br, IP 200.156.27.158

Dedico este trabalho a mim. Só eu sei o que me custou chegar até aqui.

ACKNOWLEDGMENTS

Agradeço à CAPES pelas bolsas concedidas durante o período de Doutorado, à. FAPERJ e ao CNPq por financiarem a pesquisa, os eventos científicos e as viagens para apresentação dos trabalhos oriundos dessa Tese.

À UNIRIO, por ser minha casa há quase 13 anos, e a todos os que foram professores do curso de Nutrição entre 2010 e 2015, pelos sorrisos quando nos cruzamos pelos corredores. Me emociona ver que vocês ainda lembram daquela moleca-representante de turma-faladeira-cheia de sonhos que era a Thais da graduação.

Ao PPGAN, pelos 7 anos de tamanho aprendizado e evolução, como pessoa e como profissional.

Aos membros do Colegiado do PPGAN, pela oportunidade de participar do mesmo como representante discente, por me darem voz e permitirem aprender tanto na vivência e discussões das nossas longas reuniões.

Ao prof. Dr. Luiz Carlos Gutkoski e à empresa OR Melhoramento de Sementes Ltda. pelo fornecimento das amostras, até nos "45 do segundo tempo"!

À Profa. Dra. Mariana Simões Larraz Ferreira, minha "Ori". Eu não consigo expressar em palavras tamanha gratidão que eu sinto por você e espero um dia conseguir fazê-lo. Ainda que às vezes eu pareça distante e fechada, e, como você já colocou mais de uma vez, "não te dou abertura para tentar entender o que se passa comigo", você nunca cogitou desistir de mim (diferente dessa que vos fala). Obrigada pela paciência, pelos conselhos e por acreditar que eu cumpriria mais essa etapa.

À Dra. Cristina Takeiti, por aceitar me coorientar quando decidi trabalhar com processamento do trigo, e não só com a farinha! Obrigada por me ensinar tudo sobre extrusão, panificação, pastificação e por correr a Embrapa pra cima e pra baixo junto comigo, pra fazer isso dar certo.

Zu Prof. Dr. Katharina Scherf, die die Herausforderung angenommen hat, mich in ihrer neuen Forschungsgruppe an einer neuen Universität aufzunehmen, wo sie selbst eine neue Phase ihrer Karriere beginnen würde. Sie sind eine große Inspiration, wenn ich über Lebensmittelwissenschaft, Getreide und Weizen nachdenke. Mögen wir uns auf diesem langen Weg der Erforschung allergener Peptide in Getreide wiedersehen und zusammenarbeiten.

A todos os membros da banca examinadora: Profa. Dra. Maria Gabriela Koblitz, Prof. Dr. Carlos Piler, Dra. Maíra Fasciotti e Profa. Dra. Talita Pimenta, pelo aceite em julgar e contribuir com melhorias no presente trabalho.

Ao Prof. Dr. Paulo Sergio Marcellini, que me ensinou a amar a ciência e a sala de aula, por ser meu espelho e maior inspiração quando eu penso em na palavra PROFESSOR. Não digo que quero ser igual a você, afinal, cada um tem as suas características e é isso que nos diferencia uns dos outros, mas minha meta é ter a sua capacidade de ensinar aliada ao meu jeitinho de fazê-lo, me tornando a melhor professora que eu puder ser, dentro da minha área.

Ao Prof. Marcelo Costa, por confiar na minha capacidade e me dar a oportunidade de alcançar o maior objetivo de toda a minha jornada acadêmica: lecionar. A UCAM e toda a Escola de Saúde são um marco na minha carreira!

A todos os que passaram, em algum momento, pelo LABBIO 1: Ana, Amanda, Andrés, Beatriz Barroso, Beatriz Moura, Carolina, Cátia, Diego, Fabiana, Fernanda, Ingrid, Louise, Luciana, Mariana Moraes, Mariana Pimentel, Millena, Moara, Nathália, Oscar, Pamela, Patrícia, Renata, Sónia, Talita Braga, Talita Pimenta, Tatiana e Verônica (espero não ter esquecido ninguém...). Vocês são o melhor time que eu jamais sonharia trabalhar. Nosso espírito de união e cooperação é algo raríssimo de se encontrar em qualquer equipe por esse mundo afora. Que cada um de nós possa encontrar um *Team Gluten* (eu nunca vou abandonar esse nome) nos próximos passos das suas jornadas. Menção honrosa pras minhas Cocotinhas! O que seria de mim sem as nossas fofocas e ligações de madrugada “pra trabalhar junto”?

Zu AK Scherfchen: Alena, Annika, Charlotte, Christine, Lukas, Majlinda, Nina, Nora, Ruben, Sabrina und Sarah, die 17 Monate lang wegen mir gezwungen war, bei der Arbeit Englisch zu sprechen. Vielen Dank für all die gemeinsamen Momente, auch wenn wir aufgrund des Coronavirus eine gewisse soziale Distanz wahren; und dafür, dass du mir beigebracht hast, wie man Glühwein trinkt, und Brezel und Currywurst isst! Ich vermisse euch alle und unser GuMos, Mittagessen in der Sommersonne, Apache im Labor, Kaffee und Kuchen, Feierabende mit Bayerisches Bier und Pizza und Harry Potter-Filmsitzungen. Wie die Deutschen sagen, “Man sieht sich immer zweimal im Leben”. Ich hoffe, dass ich jeden von Ihnen wiedersehen kann.

Aos meus pais, Francisco José e Jane, por absolutamente tudo! Por todo o esforço e sacrifício feitos por e para mim. Se hoje eu sou quem sou, devo isso única e exclusivamente a vocês. O tamanho do meu amor e gratidão é imensurável!

Ao Raniel, minha Capivarinha, e talvez uma das pouquíssimas boas decisões que eu tomei nos últimos anos. Tá pra nascer namorado que sinta mais orgulho da mulher que tem como companheira do que ele! Obrigada por vibrar a cada conquista minha como se ela também fosse sua, por me incentivar a crescer cada vez mais e por torcer tanto por mim e por nós. Obrigada pelas noites me esperando no laboratório para me dar carona até em casa, sem nem ideia de que horas eu conseguiria me liberar. Obrigada pelos "corres" com amostra, catalisador, produtos químicos e outras coisas que você nem sabia o que eram. Obrigada também por me ensinar que o amor não tem barreiras e me permitir sentir um amor que eu não acreditava ser possível pela nossa Ana Laura.

Aos meus afilhados João Eduardo e Pedro Jaime, por serem o amor mais puro e profundo que uma pessoa é capaz de sentir.

À minha dinda Carla Maria, por ceder sua casa e se privar da sua liberdade nos dias em que ir e voltar da UNIRIO era praticamente impossível.

Às minhas primas Dayana, Mayara e Paula Carla, por serem porto seguro, meus lugares de desabafo e as pessoas que eu sei que posso confiar os meus segredos mais obscuros.

Às minhas cadelinhas Alice e Petit, minha maior saudade durante o intercâmbio e melhor companhia da vida.

Zu meine Deutsche Tante Vera, que, desde que me conheceu, aos 4 anos, dizia que eu um dia iria morar e estudar na Alemanha, por se desdobrar e dirigir por mais de 500 km para me receber no aeroporto e me levar para minha nova casa, numa cidade e país até então desconhecidos para mim, por livre e espontânea vontade. Ich werde trotzdem Deutsch lernen, damit wir uns in deiner Sprache unterhalten können!

Às minhas amigas da vida, Andressa, Maria Clara e Mariana, com quem, independente do tempo ou distância, eu sei que posso contar pra tudo.

Aos meus pedacinhos do Brasil na Alemanha: Danielle, minha eterna RuSis, pelas ligações, pelos furos no lockdown para beber vinho e comer besteira, pelos jantares e por me apresentar RuPaul's Drag Race; e Rhael, a melhor companhia pra ir ao clube tomar banho de piscina no verão, pela comida mexicana, pelo feijão em caixinha vencido, pelas sessões de stand-up comedy na Netflix, pela varinha do Dumbledore e por ter perguntado se tinha sobremesa!

Zu Patrick, meine Deutsche gute Fee, dass Sie mir so oft das Leben gerettet haben, dass man es nicht an den Fingern abzählen kann!

Zu Jonas, eine unglaubliche Begleiterin, die zufällig in mein Leben trat, hat meine letzten sechs Monate in Deutschland geprägt und verändert.

À psiquiatra Adriana Dallier e à psicóloga Amanda Bengozi por me auxiliarem de forma profissional a lidar com os obstáculos e diagnósticos recebidos nos últimos anos.

À Ciência, à Pesquisa, ao conhecimento.

Ao desbravamento de novas experiências e oportunidades.

Por último, mas mais importante, a Deus. Tudo que aconteceu até aqui, foi pela Sua vontade. Ainda que eu não entenda Seus planos ou Seus caminhos, confio que reservas o melhor para mim.

O futuro ainda é uma incógnita, mas eu definitivamente não paro por aqui.

Meu mais sincero “muito obrigada”.

T.

“Das wahre Zeichen von Intelligenz ist nicht das Wissen, sondern die Vorstellungskraft. […] Vorstellungskraft ist wichtiger als Wissen. Das Wissen ist begrenzt. Imagination umkreist die Welt.”

“The true sign of intelligence is not knowledge, but imagination. [...] Imagination is more important than knowledge. Knowledge is limited. Imagination circles the world.”

(Albert Einstein)

ABSTRACT

This thesis aimed to better understand the impact of food processing on the protein and allergen profiles of wheat-based products obtained from contrasting wheat flours by applying modern proteomic techniques. Thermoplastic/cold extrusion, baking and fermentation processes were performed to assess how they affect protein behavior and profile, also considering two contrasting wheat flours (improver/bread and biscuit wheat genotypes) (*Triticum aestivum*). The higher gluten strength in ORS Agile (bread type) impaired protein extraction compared to the ORS Vintecinco genotype (biscuit type), especially for extrudate samples, probably due to the intense glutenin polymerization. Structural changes caused by kneading and yeast action in the dough result in significant changes in the protein profile. The increase in glutenin content was driven by the incorporation of gliadin subunits and soluble proteins into the polymers. Proteomic analysis showed higher expression of prolamins and glutenins with the increase in screw rotation speed during thermoplastic extrusion. In pasta, the greatest influence was on prolamins, induced by heat during cooking. In bread samples, α-, γ- and ω-gliadins were the most affected proteins. Mechanical force and fermentation cause a reduction in gliadin expression when breads were compared to flour and unfermented breads. Samples produced from bread wheat type were more impacted by processing than those produced with biscuit bread type. Different protein composition and expression was found between them: products from ORS Vintecinco showed a lower amount of gluten proteins compared to products from ORS Agile. Globally, around 12% of the proteins were classified as allergenic, comprising also soluble and metabolic wheat proteins. The distinct allergen content between the two flours was remarkable, bread type flour showed an expression of allergens more than three times higher than biscuit wheat flours. This result can be explained by the higher gluten content in the bread wheat flour, as most allergens are gluten proteins. After processing, wheat-based samples from bread type genotype showed a more pronounced reduction in the intensity of allergen expression. Although interesting observations were pointed out, suggesting some effect of processing, further studies including a larger number of genotypes are needed to determine candidate proteins for biomarkers of quality/technological suitability and the role of wheat processing in causing changes in the proteomic profile. Further studies, also involving the molecular size distribution of proteins can help to understand the structural modifications of gluten in the different wheat-based products.

Keywords: Foodomics, glutenomics, wheat processing, technological quality, immunogenic profile.

RESUMO

Esta tese teve como objetivo compreender melhor o impacto do processamento de alimentos nos perfis de proteínas e alérgenos de produtos à base de trigo obtidos a partir de farinhas de trigo contrastantes, aplicando técnicas proteômicas modernas. Foram realizados processos de extrusão termoplástica/a frio, panificação e fermentação para avaliar como eles afetam o comportamento e perfil das proteínas, considerando também dois tipos contrastantes de farinhas de trigo (genótipos de trigo melhorador/pão e biscoito) (*Triticum aestivum*). A maior força do glúten em ORS Agile (tipo pão) dificultou a extração de proteínas em comparação com o genótipo ORS Vintecinco (tipo biscoito), especialmente para as amostras extrudadas, provavelmente devido à intensa polimerização das gluteninas. As mudanças estruturais causadas pelo amassamento e ação do fermento na massa resultam em alterações significativas no perfil de proteínas. O aumento no conteúdo de gluteninas foi impulsionado pela incorporação de subunidades de gliadina e proteínas solúveis nos polímeros. A análise proteômica mostrou maior expressão de prolaminas e gluteninas com o aumento da velocidade de rotação do parafuso durante a extrusão termoplástica. Na massa, a maior influência foi sobre as prolaminas, induzida pelo calor durante o cozimento. Nas amostras de pão, as proteínas mais afetadas foram α-, γ- e ω-gliadinas. A força mecânica e a fermentação causam uma redução na expressão de gliadinas quando comparadas com as farinhas e pães não-fermentados. As amostras produzidas a partir de trigo tipo pão foram mais afetadas pelo processamento do que aquelas produzidas com trigo tipo biscoito. Foram encontradas diferentes composições e expressões de proteínas entre elas: os produtos de ORS Vintecinco apresentaram menor quantidade de proteínas de glúten em comparação com os produtos de ORS Agile. De forma geral, cerca de 12% das proteínas foram classificadas como alergênicas, incluindo proteínas solúveis e metabólicas do trigo. A distinção no teor de alérgenos entre as duas farinhas foi notável, a farinha tipo pão apresentou uma expressão de alérgenos mais de três vezes maior do que as farinhas de trigo tipo biscoito. Esse resultado pode ser explicado pelo maior teor de glúten na farinha de trigo tipo pão, já que a maioria dos alérgenos são proteínas do glúten. Após o processamento, as amostras à base de trigo do tipo pão mostraram uma redução mais pronunciada na intensidade da expressão de alérgenos. Embora tenham sido apontadas observações interessantes, sugerindo algum efeito do processamento, estudos adicionais envolvendo um maior número de genótipos são necessários para determinar proteínas candidatas a biomarcadores de qualidade/adequação tecnológica e o papel do processamento do trigo em causar alterações no perfil proteômico. Estudos adicionais,

envolvendo também a distribuição do tamanho molecular das proteínas, podem ajudar a compreender as modificações estruturais do glúten nos diferentes produtos à base de trigo.

Palavras-chave: Foodomics, glutenomics, processamento de trigo, qualidade tecnológica, perfil imunogênico.

LIST OF ACRONYMS AND ABBREVIATIONS

AG – Albumins and Globulins

ALP – Alkaline Phosphatase

APCI – Atmospheric Pressure Chemical Ionization

APPI – Atmospheric Pressure Photoionization

BLAST – Basic Local Alignment Search Tool

CCS – Collision Cross-Section

CD – Celiac Disease

DAP – Different Abundant Proteins

DIA – Data Independent Analysis

DRP-5 – Dihydropyrimidinase-Related Protein 5

EBI – European Bioinformatics Institute

ELISA – Enzyme-Linked Immunosorbent Assay

ESI – Electrospray Ionization

FT-ICR – Fourier Transform-Ion Cyclotron Resonance

Gli – Gliadins

Glu - Glutenins

GMO – Genetic Modified Organisms

GO – Gene Ontology

GP-HPLC-FLD – High Performance Liquid Chromatography with Gel Permeation and Fluorescence Detection

GRD – Gluten-Related Disorders

GSP-1 – Grain Softness Proteins

$HDMS^E$ – High-Definition Mass Spectrometry

HMW-GS – High Molecular Weight-Glutenin Subunit

IMS – Ion Mobility Spectrometry

IT – Ion Trap

LC – Liquid Chromatography

LC-MS – Liquid Chromatography coupled to Mass Spectrometry

LMW-GS – Low Molecular Weight-Glutenin Subunit

LTPs – Lipid Transfer Proteins

MALDI – Matrix-Assisted Laser Desorption Ionization

MRM – Multiple Reaction Monitoring

MS – Mass Spectrometry

MS/MS – *Tandem* Mass Spectrometry

NMR – Nuclear Magnetic Resonance

MS^E – Multiplex data-independent acquisition

PCA – Principal Component Analysis

Pin – Puroindolines

PIR – Protein Information Resource

PLS-DA – Partial Least Square-Discriminant Analysis

PRM – Parallel Reaction Monitoring

QCM – Quartz Crystal Microbalance

Q – Quadrupole

QqQ – Triple Quadrupole

QToF – Quadrupole coupled to Time of Flight

RP-HPLC – Reverse Phase-High Performance Liquid Chromatography

RRM – RNA Recognition Motif protein

SDS-PAGE – Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

sHsps – small-Heat shock proteins

SIB – Swiss Institute of Bioinformatics

SMP – Synaptotagmin-like Mitochondrial-lipid-binding

SPR – Surface Plasmon Resonance

SRM – Single or Selected Reaction Monitoring

ToF – Time of Flight

TQS – Triple Quadrupole Spectrometer

UDMS^{E} – Ultra Definition Mass Spectrometry

UHPLC – Ultra High-Performance Liquid Chromatography

WDEIA – Wheat-Dependent Exercise-Induced Anaphylaxis

WSCI – Wheat Subtilisin/Chymotrypsin Inhibitor

SUMMARY

INTRODUCTION

Wheat (*Triticum aestivum*) is a cereal from the Poaceae family and is one of the oldest and most important crops in the world. Wheat is the second most consumed cereal (as food) in the world, with an average annual production of 764 million tons (FAO, 2023) in the last 5 years. However, the Brazilian production of wheat is lower than demand, and Brazil imports about 50% to maintain the country's supply and ensure food security for the population. The average annual Brazilian production in the last 5 years was 5.8 million tons, but in 2022 it reached a record of 9.5 million tons (Conab, 2023; FAO, 2023) In this scenario, it is possible to reach 76% of national demand, but still being necessary to import 6 million tons to guarantee national consumption.

Wheat and its derivatives, such as wheat flour, are fundamental components of human nutrition in many cultures around the world. It is believed to have first been cultivated in the Middle East about 10,000 years ago. Since then, its production and consumption have spread across the globe. This cereal is considered a vital source of nutrients and energy, and its importance in human nutrition is related to its nutritional composition. Whole wheat is an excellent source of complex carbohydrates, providing energy gradually. In addition, it contains proteins, fibers, B vitamins, and minerals such phosphorus and magnesium, as well as iron and zinc (Šramková et al., 2009). Although Fe and Zn are not the main minerals in wheat, their amount in this cereal is significant due to its enrichment in wheat crops, which contribute for the population to reach the daily intake recommendation of these minerals, which are the most deficient in the human diet (Hussain et al., 2010).

Wheat flour is the main product derived from this cereal, and the most widely consumed, for its versatility in being used in the production of different processed foods. During the milling process, the wheat grains are crushed through a series of corrugated and smooth metal rollers, occurring the separation of the inner layer – endosperm – from the bran and germ, followed by gradual size reduction of endosperm, where the ground wheat is sifted between stages. The final result is a fine, white powder used in the preparation of various foods such as breads, cakes, pastas, and cookies. Several types of wheat flour can be produced, such as whole wheat flour, common wheat flour and enriched wheat flour. The whole meal is obtained from the milling of the whole wheat grain, including bran and germ. It is richer in fiber, vitamins, minerals, and bioactive compounds, being a healthier option when

compared to refined flours. Since common wheat flour goes through a refining process, this material loses the nutrients located within the germ and the bran. Enriched wheat flour, on the other hand, is a refined version that goes through a fortification process, in which essential vitamins and minerals are added to compensate for the loss during refining (Jones & Engleson, 2010).

Despite its high versatility, the destination of wheat flour is defined by a classification system based on the technological aptitude or quality. Wheat flour presents different physicochemical properties or characteristics that define technological qualities and then directly impact the quality of the end products. Protein content and composition, gluten strength, water absorption, dough stability and gas holding capacity are important characteristics to be considered during flour selection and use. The protein content of wheat flour is a key determinant of its technological qualities. Wheat storage proteins, known as gluten, are responsible for the formation of the bread structure and provide elasticity and extensibility to the dough.

Wheat proteins are essential components of wheat and play a key role in the structure and properties of wheat flour products. The main proteins found in wheat are glutenins and gliadins, which, with the addition of water and mechanical support, form gluten. Gluten is a network of proteins formed by the interaction of glutenins and gliadins present in wheat. Glutenins are responsible for the elasticity and strength of gluten, while gliadins confer extensibility and viscosity. These characteristics are crucial for the formation and retention of gases during the fermentation process, resulting in the soft and elastic texture of wheat products. Glutenins and gliadins are subdivided into different groups based on their structural properties and functional characteristics. There are two main subunits of glutenins: high molecular weight (HMW) and low molecular weight (LMW). Gliadins, on the other hand, are divided into different subunits: alpha, beta, gamma, and omega. Each subunit contributes to the functional characteristics of gluten, such as viscosity and extensibility. In addition to these characteristics, wheat proteins confer dough stability and tolerance to processing, helping to maintain structure and prevent excessive collapse during cooking (Wieser, 2007; Wieser et al., 2023).

Flours with high protein content, such as “strong” wheat flour, are suitable to produce breads with good expansion and volume. On the other hand, flours with low protein content, such as “weak” wheat flour, are more suitable for confectionery products like cakes and

cookies. Although, A high protein content can also impede bread crumb formation by constricting air expansion, thus compromising its structural integrity (Janssen et al., 1996). However, one of the most important features appears to be the degree of protein polymerization, characterized by the molecular weight distribution (MWD) of polymeric gluten proteins. Hence, the formation of gluten polymers has been since the 1990–2000s and still is the subject of numerous studies because of the preponderant role they seem to play in the definition of the technological properties of flours and/or bread wheat doughs (Aussenac et al., 2020).

The composition of glutenin subunits is also an important parameter and the amounts of x-type HMW-GS subunits showed positive correlations with dough development time, maximum resistance of dough, amount of gluten and bread volume (as reviewed by Lafiandra and Shewry, 2022). Some specific subunits are associated with superior dough properties, such as 1Dx5, which has an extra cysteine residue within the repetitive domain, and 1Bx7 present in greater amount, showed stronger effects on breadmaking than other x-type subunits.

Gluten strength is related to the ability to retain gases during bread fermentation, which directly affects the volume and texture of the final product. Wheat flours with strong and elastic gluten tend to produce breads with greater volume and softer crumb. On the other hand, flours with weak gluten can result in breads with lower volume and denser crumb (Delcour et al., 2012). The gas retention capacity is directly associated with the previous characteristic and is an important property of the wheat flour during bread fermentation. The gluten present in the flour can retain the carbon dioxide produced by the yeast, contributing to the growth and formation of alveoli in the dough. Flours with high gas retention capacity produce breads with lighter, softer crumb and well distributed alveoli. Water absorption by wheat flour is also a relevant technological characteristic. The amount of water required to achieve the proper dough consistency varies among different flours. Flours with high water absorption require more liquid to obtain the desired dough consistency. This characteristic can influence the final texture of baked goods such as bread and pasta. Finally, dough stability refers to the wheat flour ability to maintain its structure during processing and fermentation. A stable dough makes it easier to handle and shape products. Flours with good dough stability help maintain the shape and structure of baked goods, resulting in products with better sensory quality (Thanhaeuser et al., 2014).

Despite in many countries, wheat is commercially classified according to the end-use product destination, such as in Brazil: improver, bread, domestic or basic types accordingly to gluten strength, farinography and Falling number (Brasil, 2010). There is a lack of correlation between this kind of classification and the practical observations of industrial tests, for instance between protein composition and loaf volume (Schuster et al., 2023). Usually, millers blend wheat cultivars of related technological qualities to meet the desired end-product specifications. Baking quality is considered as the result of a complex combination of different parameters and little knowledge is available on how different processing techniques, from milling to baking or extrusion, influence protein behavior.

Wheat-based products play a significant role in modern diets, offering convenience and diversity of choice. However, it is important to understand the science related to these products, their nutritional characteristics, and potential impacts on human health. As already mentioned, wheat flour is a versatile ingredient that is widely used in the food industry to produce a variety of products. Wheat flour is a key ingredient in the production of a wide variety of products, from breads and pastas to cookies and snacks (Mengli Zhang et al., 2022). From flat breads to baguettes, wheat flour is the basis for obtaining the desired dough structure. In addition, other bakery products, such as cakes, muffins, and cookies, also rely on wheat flour to obtain soft textures and proper structures. Pasta, such as noodles, are often made with wheat flour. The flour is mixed with water and, in some cases, eggs, to form an elastic, pliable dough that can be stretched and shaped into different shapes. The quality of the wheat flour used directly affects the texture and taste of these products. Cookies, crackers, and various snacks also use wheat flour as the main ingredient. The flour is combined with other ingredients such as sugar, fats, and leavening agents to create the dough or rolled dough required to produce these products. The wheat flour contributes to the crunchy texture and desired structure.

The processing of these products involves techniques such as mixing, extrusion, fermentation, and baking to achieve desired textures, structures, and flavors. Understanding the techniques and processes involved in processing wheat flour products is essential for producing quality and consistent foods (Islam et al., 2019). The mixing and kneading technique is commonly used in the processing of wheat flour products. This step involves combining the dry ingredients (flour, sugar, salt, etc.) and liquid ingredients (water, eggs, oils, etc.) to form a uniform dough. Proper kneading helps develop gluten, giving the dough

elasticity and structure (Redl et al., 2003). Extrusion is a widely used technique in the food industry, including wheat flour processing. Extrusion of wheat flour involves the application of heat, pressure, and mechanical force to transform the flour into products with specific shapes and textures (Min Zhang et al., 2011). Fermentation is a crucial step in the processing of bread and bakery products. During this process, the yeast in the dough consumes sugars and releases carbon dioxide, resulting in an increase in the volume of the dough. Fermentation is essential for obtaining the proper texture and softness in wheat flour products (Verheyen et al., 2014). Cooking is the final step in processing wheat flour products. The proper baking temperature and time ensure that the dough will fully develop to the desired texture and color. The wheat flour, along with other ingredients, undergoes chemical reactions such as caramelization and starch gelatinization, which contribute to the final taste and appearance of the products (Bakke & Vickers, 2007).

Although wheat is a widely consumed food and an important source of nutrients, some people may develop diseases related to its consumption. Wheat-related diseases, such as celiac disease, non-celiac gluten sensitivity, and wheat allergy, are distinct conditions that require specific approaches to diagnosis and treatment (Sapone et al., 2012). Celiac disease is an autoimmune condition in which the consumption of gluten triggers an immune response in the small intestine. The gluten present in wheat, as well as in other cereals such as rye and barley, contains components that are toxic to people with celiac disease. This immune response results in damage to the intestinal mucosa, leading to a variety of gastrointestinal and nutritional symptoms. Celiac disease is a chronic condition that requires the total elimination of gluten from the diet. Strict adherence to a gluten-free diet is the only way to manage celiac disease and prevent long-term complications (Presutti et al., 2007). Non-celiac gluten sensitivity is a condition in which people experience symptoms like celiac disease, but do not have the immunological markers or damage to the intestinal mucosa characteristic of celiac disease. Symptoms may include abdominal discomfort, bloating, fatigue, and gastrointestinal disturbances (Catassi et al., 2013). Wheat allergy is an allergic reaction to the proteins in wheat. Unlike celiac disease and non-celiac gluten sensitivity, which are triggered by gluten, wheat allergy involves an immune response to specific wheat proteins. Symptoms can range from mild, such as hives and itching, to severe, such as difficulty breathing and anaphylactic shock. Wheat allergy is more common in children and in many cases is overcome with time. Diagnosis is made by skin tests, blood tests, and a history of allergic reactions. Prevention

involves eliminating wheat and wheat-containing foods from the diet, and appropriate treatment in case of allergic reactions (Inomata, 2009).

In addition to the diseases discussed above, exposure to wheat can also trigger other specific conditions. Two such wheat-related diseases are wheat-associated exercise-induced anaphylaxis (WDEIA) and baker's asthma. WDEIA is a rare form of food allergy that occurs because of wheat consumption followed by exercise. Symptoms usually manifest during or after intense physical activity and may include hives, swelling, difficulty breathing, drop in blood pressure, and in severe cases, anaphylaxis. WDEIA is thought to be triggered by the combination of wheat intake and increased body temperature during exercise, leading to an exacerbated allergic reaction. Its treatment involves avoiding wheat intake before exercise and being prepared for emergencies in case of severe allergic reactions (K. A. Scherf et al., 2016). Baker's asthma, also known as occupational asthma, is a respiratory condition that occurs due to exposure to allergens present in the work environment in bakeries and baking industries. Inhalation of wheat flour, baking enzymes, and other components in the air can trigger an allergic response in the lungs, resulting in asthma symptoms such as shortness of breath, wheezing, coughing, and chest tightness (Larre et al., 2011).

The etiology of Wheat-related diseases, especially CD, has been linked to the amino acid composition and chain length of peptides produced during the gastrointestinal digestion of gluten proteins (Kucek et al., 2015). The primary trigger of the immune response in CD is the resistance of gluten peptides to proteolysis. This resistance is due to the scarcity of lysine and arginine and the compact structure of the proteins, which impedes the action of proteases such as trypsin, pepsin, and chymotrypsin. Upon reaching the lamina propria of the small intestine, these proline- and glutamine-rich polypeptides act as immune mediators by binding to DQ2 or DQ8 antigens on intestinal epithelial cells in genetically susceptible individuals (Katharina Anne Scherf et al., 2016; Scherf & Poms, 2016). Several gluten peptides capable of eliciting an immune response by being recognized by T lymphocytes have been identified in gliadins, glutenins, hordeins, and secalins (Kucek et al., 2015). The difference in amino acid composition of the different prolamins may be responsible for the different reactivities associated with CD. Grains belonging to the subtribe Triticeae (wheat, barley and rye) contain significantly high levels of glutamine and proline, being primarily responsible for triggering the immune response in celiacs (Michelle L. Colgrave et al., 2015). Gluten allergenicity is a problem not only in foods that contain this protein through the presence of wheat, barley or

rye, but in foods that are considered to be hidden sources of gluten due to improper labeling or cross-contamination in manufacturing or transportation. There is also a growing concern about the presence of gluten due to the trend of its incorporation in foods that traditionally do not contain wheat protein (Hlywiak, 2008). Accurate diagnosis and proper treatment are essential for the management of any of these conditions. People who experience symptoms related to wheat consumption or exposure require medical advice for a correct diagnosis and an appropriate treatment plan.

Proteomics involves the study of proteins on a large scale, including their identification, characterization, quantification, and interactions. This approach enables a more comprehensive and detailed understanding of the proteins in a biological system. Proteomics uses advanced techniques, such as mass spectrometry, to analyze the proteins present in a sample. Proteomics has applications in many areas of research such as medicine, biotechnology, agriculture, and food safety (Graves & Haystead, 2002). Through proteomic analysis, it is possible to identify disease biomarkers, elucidate metabolic pathways, understand the response of organisms to external stimuli, and assess the quality and authenticity of foods. The presence of traces of wheat and wheat allergens in food may pose risks for people with different Wheat-related diseases. Proteomics plays an important role in the detection and identification of these allergens, providing sensitive and specific methods for food analysis (Alves et al., 2017; Alves et al., 2019; Xhaferaj et al., 2020).

Proteomic analysis in wheat trace and allergen detection involves several steps. First, proteins are extracted from the food sample. Next, the proteins are separated and purified using techniques such as gel electrophoresis and chromatography. Subsequently, the proteins are identified by mass spectrometry by comparing their mass spectra with proteomic databases. This approach allows the detection of wheat-specific proteins and related allergens, even in small quantities (Schalk et al., 2018a, 2018b). Proteomics also plays a crucial role in the analysis of processed food ingredients such as wheat flour to ensure regulatory compliance and absence of cross-contamination. Proteomic techniques can identify and quantify wheat proteins in processed products, even when wheat is not declared as the main ingredient (Scherf & Poms, 2016). In addition, proteomics can assist in the development of food processing methods that minimize the presence of wheat allergens to produce safe foods for people with dietary restrictions. In summary, proteomics is a powerful tool in detecting traces of wheat and wheat allergens in food. Through proteomic analysis, wheat-specific

proteins can be identified and quantified, providing important information to ensure food safety and meet the needs of people with wheat-related diseases. The continued evolution of proteomics will contribute to improved methods for detecting and monitoring food allergens, promoting safer diets tailored to individual needs (Abril et al., 2023).

The term "Foodomics" is used to describe a new approach in food analysis that combines several analytical platforms and data processing for transcriptomics, proteomics, and metabolomics studies. "Omics" stand for the comprehensive study of a particular class of biological molecules or components within a biological system, which allows for the comprehensive evaluation of the health benefits of food ingredients at the molecular level (Cifuentes, 2009; Ibáñez et al., 2012). A recent literature review published by Ortea (2022) explores the study of food and nutrition using omics approaches to investigate the effect of diets, foods or food components on health/disease status. Bioactive food components can alter gene expression and protein levels and functions, leading to various beneficial effects on human health. By analyzing and integrating gene variants and expression, epigenetic regulation, protein levels and post-translational modifications, as well as metabolites globally, functional information of a diverse food range can be obtained. This allows the investigation of cellular processes, functional mechanisms and molecules involved, as well as the definition of targets for bioactive compounds useful for the development of nutritional intervention strategies and the discovery of biomarkers linking nutrition, health, and disease.

Although the application of Foodomics in food analysis proves to be a promising strategy, it is important to emphasize that challenges associated to food composition evaluation and applying Omics technology will be faced (Ahmed et al., 2022). The wide dynamic range of protein concentration used to hinder the detection and quantification of low abundance proteins. With the development of more sensitive proteomic methods, such as targeted analysis like the multiple reaction monitoring (MRM), this issue has been defeated, after the development of a quantitative MS-based method to detect wheat contamination in foods (Michelle L. Colgrave et al., 2015). In food allergens identification and quantification, for example, interferents present in the food matrix or generated during processing may affect and impair allergen detection (García-Cañas et al., 2012). Furthermore, the extensive amount of data provided makes challenging to analyze and interpret the results. The application of bioinformatics, such as tools for data processing, statistics, and functional interpretation, imply difficulties because results vary depending on the statistical thresholds used. The wide

variety of bioinformatics tools used also lead to limitations in reproducibility and comparability of results and an inadequacy of validation. Integration and comprehension of different omics results is the most challenging stage of Foodomics application (Ortea, 2022).The present thesis aims to understand, using proteomic techniques, the changes caused by processing in the expression of wheat proteins in wheat flour samples of contrasting technological qualities (bread/improver and biscuit wheat genotypes), as well as the influence of this processing on the presence of proteins and peptides already established as wheat allergens. For this purpose, this manuscript was organized in 5 chapters and structured in the format of scientific papers:

Chapter 1 provides a published literature review of the most modern mass spectrometry techniques used for the identification and quantification of immunogenic peptides in cereals. The bibliographic review is complemented by the published review coauthored and presented in Appendix 1 “Recent progress in analytical method development to ensure the safety of gluten-free foods for celiac disease patients”.

Chapters 2, 3 and 4 portray the results obtained from the proteomics experimental analyses conducted during the PhD. Chapter 2 addresses the impact of different extrusion processes - thermoplastic extrusion and cold extrusion - and cooking on the protein profile and assembly in wheat flours of contrasting technological end-use qualities. In chapter 3, the same proteomics techniques were applied to understand how fermentation by *S. cerevisiae* impacts protein proteome in breads manufactured with two wheat flours of contrasting technological end-use qualities. Chapter 4 presents results obtained by the proteomic analysis to evaluate the allergenic profile of wheat proteins in all these different wheat-based processed products evaluated in chapters 2 and 3.

Chapter 5, which closes the manuscript, was written to present the final considerations and perspectives of the thesis.

This doctoral thesis was carried out in the Graduate Program in Food and Nutrition (PPGAN), within the institutional research project "Food Proteomics and Metabolomics" which is part of the research line "Processing, quality, recovery of foods, coproducts and wastes" that aims to meet the challenges of the agricultural and industrial sectors, considering the need to increase the production of safe and healthy foods in the context of the Bioeconomy. Moreover, this doctoral thesis was conducted in the framework of the research group of CNPq “Proteomics and metabolomics of bioactive compounds - Omics sciences

applied to organisms of economic and biotechnological interest". The experimental steps occurred in three different institutions at UNIRIO in the Laboratory of Bioactives-PPGAN (Rio de Janeiro, Brazil), at Embrapa Food Technology in the Pilot Plant 4 (Rio de Janeiro, Brazil) and at Karlsruhe Institute of Technology (KIT) in the Institute of Applied Biosciences (Germany) where a period of 17 months of sandwich doctorate was performed granted by CAPES-PDSE, despite the pandemic situation.

CHAPTER 1 – MODERN APPROACHES IN THE IDENTIFICATION AND QUANTIFICATION OF IMMUNOGENIC PEPTIDES IN CEREALS BY LC-MS/MS

Thais O. Alves[1,2], Carolina T. S. D'Almeida[1,2], Katharina A. Scherf[3], Mariana S. L. Ferreira[1,2]

[1] *Food and Nutrition Graduate Program (PPGAN), Laboratory of Bioactives, Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro, Brazil.*

[2] *Laboratory of Protein Biochemistry - Center of Innovation in Mass Spectrometry (LBP-IMasS), UNIRIO, Rio de Janeiro, Brazil.*

[3] *Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.*

Published in Frontiers in Plant Science, 2019. DOI: 10.3389/fpls.2019.01470

Abstract

Celiac disease (CD) is an immunogenic disorder that affects the small intestine. It is caused by the ingestion of gluten, a protein network formed by prolamins and glutelins from cereals such as wheat, barley, rye and, possibly, oats. For predisposed people, gluten presents epitopes able to stimulate T-cells causing symptoms like nausea, vomiting, diarrhea, amongst others unrelated to the gastrointestinal system. The only treatment for CD is to maintain a gluten-free diet, not exceeding 20 mg/kg of gluten, what is generally considered the safe amount for celiacs. Due to this context, it is very important to identify and quantify the gluten content of food products. ELISA is the most used method to detect gluten traces in food. However, by detecting only prolamins, the results of ELISA tests may be underestimated. For this reason, more reliable and sensitive assays are needed to improve gluten quantification. Because of high sensitivity and the ability to detect even trace amounts of peptides in complex matrices, the most promising approaches to verify the presence of gluten peptides in food are non-immunological techniques, like liquid chromatography coupled to mass spectrometry. Different methodologies using this approach have been developed and described in the last years, ranging from non-targeted and exploratory analysis to targeted and specific methods depending on the purpose of interest. Non-targeted analyses aim to define the proteomic profile of the sample, while targeted analyses allow the search for specific peptides, making it

possible to quantify them. This review aims to gather and summarize the main proteomic techniques used in the identification and quantitation of gluten peptides related to CD-activity and gluten-related allergies.

Keywords: allergenic peptides, cereals, gluten, LC-MS, MRM, prolamins, proteomics.

1. Introduction

Cereals are one of the main food sources in the world. The nutrients provided by this group represent about 50% of the recommended daily intake (RDI) of carbohydrates and one third of the RDI for proteins. Cereal grains are also considered a good source of minerals and vitamins, especially complex B vitamins (Belitz et al., 2009). According to updated FAO data (2018), the cereal production, including non-food uses specially for maize, in the last year exceeded 2,600 million tons, with a slight decrease in production expected for 2019.

Wheat is one of the most important cereals in the world for human consumption and is considered the most suitable raw material for bread and pasta making. Its production has remained constant over the years, currently only behind maize and followed by rice (FAOSTAT, 2018). In recent data reported by USDA (2018), world wheat production reached 733 million tons, whereas the estimated consumption is about 745 million tons. Barley, rye, and oats also have large production and consumption, but not so expressive as wheat, their production corresponds to about 25% of that of wheat. Rye is mostly applied for baking, while barley is applied in beer production and oats essentially commercialized as flour, bran and other products for immediate consumption (Owusu-Apenten, 2002).

The search for practical ways in the preparation and consumption of meals combined with the promotion of healthier eating habits, sparked an increase in research for new processes for products (Pfeifer et al., 2014). Grain processing involves techniques that can alter protein structure, causing changes in solubility, viscoelastic properties, spatial conformation of proteins, and other changes (Hayta & Alpaslan, 2001). Amongst the main treatments used in cereal processing, extrusion and cooking can be highlighted, as well as baking and pasta production. However, there is a lack of studies to elucidate how processing may alter not only technological characteristics, but also nutritional and health implications, since cereal proteins, especially wheat, have a high allergenic potential in susceptible individuals.

The allergenic potential of cereals has been mainly related to gluten, a complex mixture of storage proteins found in cereals that is composed mainly of prolamins (responsible for the cohesiveness and extensibility of the gluten) and glutelins (maintenance of the elasticity and strength of the gluten). Gluten proteins have common structural characteristics. Their primary structure is subdivided into distinct domains that may exhibit repetitive sequences rich in the amino acids proline (P) and glutamine (Q) (Shewry & Halford, 2002), but low in amino acids with charged side groups. Different compositions in amino acids can be responsible for different reactivity associated with celiac disease (CD) (Belitz et al., 2009; Michelle L. Colgrave et al., 2015). Grains belonging to the *Triticeae* subtribe (wheat, barley and rye) contain significantly higher levels of Q and P, being the main cereal grains responsible for triggering the immune response in celiacs (Michelle L. Colgrave et al., 2015). Cysteines represent only 2% of the amino acids of gluten proteins, but are extremely important for their structure and functionality, since they allow the formation of disulphide bonds, responsible for gluten polymerization (Wieser, 2007).

The disorders associated to gluten consumption are known as GRD (gluten-related diseases) and are classified into three types according to the response triggered in the body: autoimmune, allergic and neither autoimmune nor allergic (Sapone et al., 2012). Examples of autoimmune diseases are dermatitis herpetiformis, gluten-induced ataxia, and CD. Among IgE antibody-mediated allergies, WDEIA (wheat-dependent exercise-induced anaphylaxis), contact urticaria, food allergy and respiratory allergies are prominent. The respiratory allergies are related to the proteins of the albumin and globulin fractions, and are known as "baker's asthma" (Weiss et al., 1997). There are also disorders of non-allergic and non-autoimmune origin known as non-celiac gluten sensitivity or intolerance (Sapone et al., 2012).

In all cases of GRD, diagnosed patients cannot consume foods containing gluten or its traces, since even minimal amounts can trigger the reaction, causing variable symptoms, ranging from abdominal pain, bloating and diarrhea, to osteoporosis and long-term infertility. The severity of the reaction is due to the degree of intolerance of each individual (Banerjee, 2010; Pietzak & Fasano, 2005). Therefore, it is extremely important to correctly identify the presence of immunogenic proteins in cereal products, to guarantee the safety of their consumption by the patients. One major problem for patients is the “hidden sources of gluten” that may be present in foods due to inadequate labelling or cross contamination during

manufacturing or transportation. There is also concern about the presence of gluten due to the tendency of its incorporation into foods that traditionally do not contain wheat in its composition (e. g. sausages, nuggets, meatballs) (Day et al., 2006).

Some authors indicated the natural genetic variability as a strategy to be further exploited for the development of wheat varieties with lower levels of immunogenic epitopes (Spaenij–Dekking et al., 2005). By using the R5-based quantitation of immunodominant toxic epitopes as the trait of interest, Ribeiro et al. (2016) demonstrated that tetraploid varieties had a lower amount of toxic epitopes than hexaploid varieties, especially when compared to *Triticum aestivum* landraces, which were not subjected to breeding practices. Despite the advances in the study of genetic variability of wheat toxicity, at present there is no common hexaploid wheat that might be safe for CD patients. Furthermore, considering the wide range of in vivo immunoresponse between celiac patients and the limitation of the immunological techniques for quantifying gluten proteins, the quantification and identification of cereal reactive proteins and peptides has been a complex task requiring constant analytical improvements.

Currently, the gold standard method to detect and quantify gluten in foods is the R5 ELISA and it is recommended by the Codex Alimentarius Commission (2008). More recently, the G12 ELISA was accepted by AOAC International as an official method of analysis, first action (Halbmayr-Jech et al., 2015) . ELISAs are based on the immune reaction between specific antibodies that have been raised to detect the antigen to be determined, such as gluten. Due to their sensitivity, adequate recovery, repeatability, and reproducibility as demonstrated by collaborative studies, ELISAs are most used to check for the presence of gluten in gluten-free raw materials and products. However, in some cases, ELISAs may give false negative results, because the monoclonal antibodies have been raised against prolamins (R5: raised against a rye extract and G12: raised against the α-gliadin 33-mer peptide) and are not suitable for all gluten protein types. As a consequence, the quantification can be compromised since the result is converted to gluten amount by multiplying the prolamin content by two, assuming the prolamin/glutelin ratio to be constant (Thompson & Méndez, 2008; Wieser & Koehler, 2009). ELISA methods currently cannot distinguish between the different gluten-containing cereals and are affected by the cross-reactivity of antibodies (Diaz-Amigo & Popping, 2013; Martínez-Esteso et al., 2017; Wieser & Koehler, 2009).

In this context, proteomic approaches appear to be more sensitive and reliable techniques than the currently used assays to identify gluten proteins, which present high amino acid sequence similarity and are difficult to distinguish. Especially when applying modern in tandem tools, proteomics can undoubtedly provide additional information to ELISA results, such as the confirmation of specific proteins by unravelling the peptide sequences (Martínez-Esteso et al., 2017).

A general workflow for cereal proteomics, as shown in Figure 1, should first consider the appropriate extraction taking into account the solubility of gluten proteins (Osborne, 1907) that usually requires the use of reducing (e.g. DTT-dithiothreitol, DTE-dithioerythritol, and TCEP-Tris2-Carboxyethyl phosphine hydrochloride) and denaturing agents (e.g. SDS or urea) (Schalk et al. 2018 a, b). Enzymatic digestion is the crucial step in bottom-up proteomics. This high-throughput analysis is based on the detection of peptides to assign the proteins. Digestion is important, because the sensitivity of methods depends on the optimal size of peptides, considering the ability to be ionized and fragmented. Trypsin is the most used enzyme due to its specific cleavage on the C-terminal side of lysine and arginine residues. However, due to the small number of these proteolytic cleavage sites in gluten proteins, multiple enzymatic digestion or less specific enzymes have been used for cereal proteomics (Vensel et al., 2011; Fiedler et al., 2014). After that, the peptides can be separated by electrophoresis or liquid chromatography (LC).

LC coupled to mass spectrometry (LC-MS) is the most important tool for the identification and quantification of immunoreactive cereal proteins (Alves et al., 2017). One of the major contributions of proteomics in the study of CD has been the identification of the immunogenic epitope sequences of gluten peptides. The application of LC-MS methods makes it possible to identify the cereal species, the protein subunit and to quantify thousands of peptides and proteins in the same experiment. Having a well-curated database that includes all possible proteins present in that organism is a great advantage for the identification of the sequences. However, peptide sequences may also be identified by de novo sequencing (Ferreira et al., 2014).

Other aspects, such as ionization source and type of MS analyzer, also influence the analysis and consequently the identification and quantification of the proteins. All of these topics will be briefly covered in this review. With the use of this information, significant advances in the understanding of GRD mechanisms, such as aspects related to resistance to

proteolysis of these proteins and influence of cereal processing can be clarified, contributing to various aspects from the development of peptide detection and quantification methods to the selection of less reactive genotypes for better tolerability of these cereals.

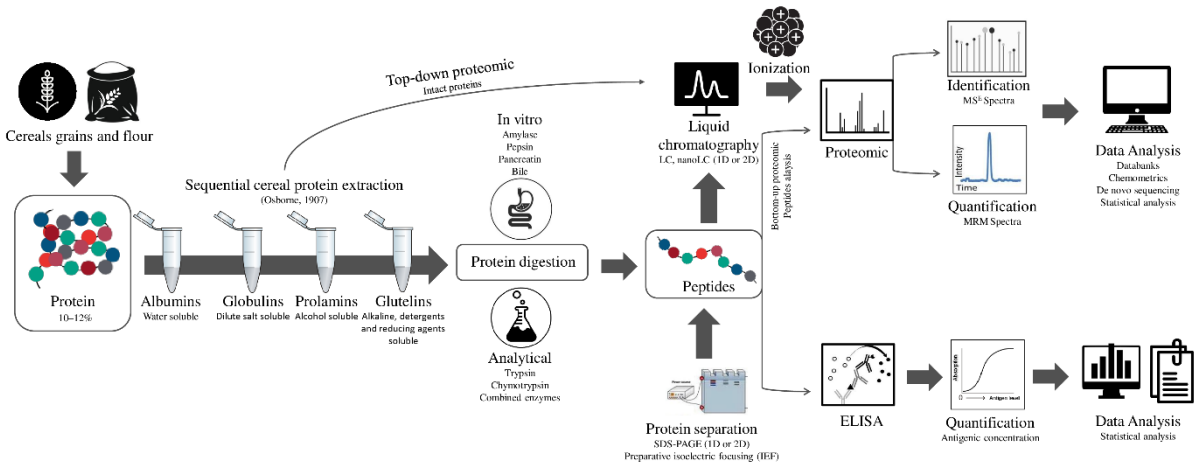

Figure 1. General workflow for cereal proteomic analysis.

2. Available gluten protein and customized databases

For LC-MS/MS analysis is important to define and use a well-curated gluten protein sequence database to improve the identification of immunogenic peptides. For this, it may be necessary to build a custom database based on an existing general database.

To provide the scientific community with a high-quality protein knowledge base, the Swiss Institute of Bioinformatics (SIB), the European Bioinformatics Institute (EBI) and the Protein Information Resource (PIR) group have joined forces and created the UniProt consortium in 2002 (https://www.uniprot.org/). The UniProt Knowledgebase (UniProtKB), the main product of this consortium, combines UniProtKB/Swiss-Prot (contains over 560,823 sequences that have been created by experimental information extracted from the literature, organized and summarized, 379 belonging to *Triticum aestivum* – accessed Oct. 2019) and UniProtKB/TrEMBL (171,501,488 sequences that have been largely derived from high throughput DNA sequencing, 142,558 belong to wheat) (The UniProt Consortium, 2017). Besides this, the UniProt consortium also produces and maintains UniRef (which consists of clusters of sequences sharing 100%, 90% or 50% of identity), UniParc (a highly redundant archive that contains original protein sequences retrieved from several different sources) or

UniMES (a collection of metagenomic and environmental sequences) (Schneider et al., 2009). All known sequences can be BLAST searched against the entire database or a part of it and the resulting sequence of high homology can be downloaded from UniProt in FASTA format.

To customize a database, other softwares should be applied. Clustal Omega (Goujon et al., 2010) and Jalview (Waterhouse et al., 2009) are used in multiple sequence alignments. Clustal Omega is an online software tool that allows protein sequences to be entered in a text file format, with optional output formats (msf output format). Jalview is a desktop program or online software for editing, visualizing, and analyzing multiple sequence alignments using Clustal Omega. Lastly, it is necessary to count the number of sequences within the file and remove redundant sequences with DBtoolkit software (Martens et al., 2005). A custom database (GluPro V1.0) of wheat gluten proteins containing 630 unique protein sequences was created to be used in LC-MS/MS data analysis to identify the presence of immunoreactive gluten peptides in foods (Bromilow et al., 2017). All software tools mentioned above were used to create this database and it provides more reliable protein IDs compared to the general database (*Viridiplantae*).

Juhász et al. (2015) also collected datasets from various public databases (UniprotKB, IEDB, NCBI GenBank) to create a specific database addressed to cereal prolamin protein families. The ProPepper database contains 2,484 unique and complete prolamin sequences, but also their peptides obtained with single- and multi-enzyme in silico digestions and specific epitopes that are responsible for wheat-related food disorders. Accordingly, is provided 667,402 unique digestion events, but also including redundant protein-peptide connections due to the simultaneous presence of some protein sequences in many genotypes and the frequency of the same peptide within a protein. Besides to be highly specific in the identification of protein sequences, this database provides specific information, such as the possible disease associated with the sequence.

Developed in 2005, Allergen Online database provides an updated peer reviewed allergen list and sequence searchable dataset to offer a risk assessment tool for evaluating the potential allergenicity of new food proteins produced by genetically modified organisms (GMO) and novel protein ingredients in processed foods (Goodman et al., 2016). The main goal is to identify proteins that may present a potential risk of allergenic cross-reactivity. This database currently presents a list of 72 proteins known to induce CD together with a downloadable list containing more than 1,000 CD-active peptide sequences. However, this

function cannot be used to search mass spectrometry (MS) data directly due to the restrictive size and not adapted format of the database (e.g., not available in FASTA format).

3. Proteomics as a tool for the screening for immunogenic peptides

The "omics" suffix means collectively considering all constituents. Proteomics consists of the analysis of the set of proteins encoded by the genome and its component molecules responsible for the control of almost all biological processes (Graves & Haystead, 2002). The use of proteomics in food analysis has become a key technological tool for the characterization and quantification of proteins and peptides, especially when it comes to the evaluation of biological markers (Carr & Anderson, 2008; Herrero et al., 2012). The coupling of the chromatographic separation and mass spectrometer detection techniques (LC-MS) increases the speed of the analyzes, allowing a large number of samples to be analyzed in a short period of time (Alves et al., 2017). In these studies, the amount of data generated is enormous and requires an important computational analytical effort to process data in a systemic and comparative way in order to deliver a practical conclusion and application (Victorio et al., 2018).

MS analyses can be divided into two types: untargeted and targeted approaches. While untargeted approaches aim to establish a comprehensive profile of the proteome of the sample, the targeted analysis allows the selection of specific molecules to be screened and studied in the sample (Saghatelian & Cravatt, 2005). Both types follow a standard workflow, where the sample is ionized via an ion source; the ions are separated according to their mass-to-charge ratios (m/z) and monitored by a mass analyzer prior to detection. In tandem MS (MS/MS) these precursor ions are then introduced into a collision cell where they undergo specific fragmentation through collision-induced dissociation (CID) by an inert gas, usually nitrogen or argon, resulting in the formation of product ions (Everett, 2011). MS/MS is usually applied for complex samples, where identified peptides are selected and subjected to fragmentation to decipher the amino acid sequence, allowing the identification of sequences that differ from each other by a single amino acid (Graves & Haystead, 2002).

The ionization source significantly impacts MS analysis as there are many ionization techniques and each has its advantages and ideal applications. The selection of the ideal ionization technique should be made based on the structure of the analyte of interest as well as the desired application (Buse et al., 2014). Various ionization techniques have been used with MS, including Electrospray Ionization (ESI), Atmospheric Pressure Chemical Ionization

(APCI), Atmospheric Pressure Photoionization (APPI) and Matrix-Assisted Laser Desorption Ionization (MALDI). For the ion source, it is important to be efficient, but at the same time sensitive and "soft", to avoid the destruction of the analyte by unwanted fragmentation in-source (Everett, 2011). Of these, the techniques most commonly used for having this feature are ESI and MALDI (El-Aneed et al., 2009). MALDI ionization essentially generates monocharged ions and thus does not require any deconvolution step. This technique emerged as an alternative to characterize wheat storage proteins due to its robustness and ability to ionize intact proteins and tolerate the presence of contaminants, such as detergents (SDS) commonly used for gluten extraction (Ferreira et al., 2014). However, this technique cannot be hyphenated directly to LC.

Conversely, ESI is powerful technique for the analysis of complex protein and peptide mixtures that benefit from the additional separation. Jira and Münch (2019) used LC-ESI-MS/MS for the simultaneous MS detection of the six most important grain species (barley, maize, oats, rice, rye and wheat) in meat products based on marker peptides. ESI was also suitable to detect traces of immunogenic gluten marker peptides in a variety of foods (Sealey-Voyksner et al., 2010) and gluten marker peptides (e.g., Colgrave et al., 2016, Manfredi et al., 2015, Schalk et al., 2018a,b).

A miniaturized version of ESI, termed nanospray, has become the preferred method of introducing large peptides into the mass spectrometer in case peptide contents are suspected to be low to very low (Hopper, 2019; Nadler et al., 2017). nanoLC-ESI-MS/MS was efficient to identify 29 immunogenic peptides from wheat flour carrying a high number of epitopes (Alves et al., 2018). Droplets produced from nanoESI are smaller than in conventional ESI (of the order of a few hundred nanometers), greatly improving the sensitivity and explaining the predominance of this technique in quantitative large-scale proteomics. The use of nanoLC to analyze complex peptide mixtures, especially when combined to orthogonal separation such as 2D RP/RP separation prior to MS/MS analysis, improves the resolution facilitating the identification and quantification of peptides containing CD immunogenic epitopes even at low femtomolar levels of detection (van den Broeck et al., 2015). When sample amounts are limited, nanoLC remains the best option due to the increased analytical sensitivity, otherwise UPLC or even HPLC separation is also useful for gluten detection.

Quadrupole is one of the most common types of mass analyzer, which four parallel metal rods are opposite connected electrically, and voltage is applied to the diagonally placed

pair of rods, resulting in an electrical field that causes the ions to travel forward. Nonetheless, a set of mass analyzers can be used for this purpose, such as ToF (Time of Flight), IT (ion trap), Orbitrap® or FT-ICR (Fourier transform ion cyclotron resonance), they can also be combined to improve the sensitivity of the method (Herrero et al., 2012).

3.1. MS-based identification of immunogenic peptides

The variability of cereal protein composition caused by the different species and varieties (genetic variability) and by growing conditions (environmental variability) leads to methodological difficulties for the analysis of immunoreactive peptides and also for the selection of genotypes (Juhász et al., 2015). In addition, the high amount of repetitive units and the similarity of the amino acid sequences of the different prolamins, with limitations in the available methodologies, make it difficult to accurately identify peptides that cause diseases related to cereal consumption, as well as their genotype frequency, variability and stability (Juhász et al., 2015).

As mentioned, MS is considered to be the golden standard for the analysis of biomolecules in complex samples, such as food matrices, because it presents high levels of sensitivity and specificity, and has been increasingly used in food analysis (Michelle L Colgrave, 2017). In cereal proteins, multiple acquisition methods or DIA (data independent acquisition), such as MSE allow minimizing data loss (e.g.: non-fragmented precursors) (Victorio et al., 2018). In MSE methods, all ions generated in the source are transmitted to the collision chamber, which alternates between low and high energy, sending precursors and fragments quasi-simultaneously to the TOF (Time of Flight) analyzer (Egertson et al., 2015). In DIA methods there is no previous selection of precursors or a threshold of ion intensity to undergo fragmentation, while for DDA typically the three most intensive single or multiple charged ions eluting from the column are selected for fragmentation (van den Broeck et al., 2015).

The use of label-free acquisition methods, such as the multiplex MSE method, takes advantage of a data collection approach that focuses on maximizing peptide fragmentation and then improving the identification and proteome coverage (Victorio et al., 2018). MSE methods have been applied to gluten protein identification and quantitation (Bromilow et al., 2016; van den Broeck et al., 2015; Uvackova et al., 2013). Label-free absolute quantification is based on the relationship between MS signal response and proteinpeptide concentration: the average MS signal response for the three most intense tryptic peptides per mole of protein

(top 3) is constant (CV<10%) and this relationship is used to calculate a universal signal response factor given an internal standard (Silva et al., 2006). However, due to data complexity many steps of data processing are required in DIA such as peak alignment, ion detection, clustering and normalization prior to peptide matching by search algorithms from a database of protein sequences.

In general, there are two possible approaches when applying LC-MS/MS for gluten detection, both of which are valid, but depend on the question to be answered. The first option is to specifically detect known CD-immunogenic peptides in order to estimate the immunogenicity of gluten. This has been reported for a selection of α- and γ-gliadin peptides (Sealey-Voyksner et al., 2010), α-gliadin peptides (van den Broeck et al., 2015), the 33-mer peptide (Schalk et al., 2017) and various gluten-derived peptides (Malalgoda et al., 2018, Alves et al., 2018). In contrast, the second option is to look for the presence of gluten, but not necessarily for CD-immunogenic peptides. Due to their length of at least nine amino acids, the poor enzymatic digestibility of the corresponding repetitive sequences and their high contents of glutamine and proline, CD-immunogenic peptides often have properties unfavourable for MS detection, whereas other gluten peptides might be more abundant. With the overall aim to detect gluten, this approach was also used to identify marker peptides in wheat, rye, barley, and oats (Manfredi et al., 2015, Schalk et al., 2018 a, b).

Recent examples demonstrating the successful application of proteomics in the evaluation of the presence of gluten marker peptides, include the detection of the presence of gluten in beers (Allred et al., 2014; Tanner et al., 2013). Tanner et al. (2013) also made a comparison between two different gluten detection methods, reinforcing the superiority of LC-MS/MS to detect gluten peptides in relation to the ELISA due to its higher sensibility and the ability to detect both, glutelin and prolamins, and not only prolamins as ELISA. This fact can be corroborated by Michelle L. Colgrave et al. (2014), where MS was used to detect and confirm the presence of hydrolyzed gluten proteins in beers which had been previously estimated as gluten-free by ELISA. A set of barley-specific peptide markers was also proposed to evaluate the contamination of processed food, ensuring the food safety for CD patients (Michelle L. Colgrave et al., 2016).

In fact, MS has been effectively applied to define a set of specific analytical targets, such as signature peptides specific to prolamins or cereal-containing gluten proteins. The main interest of these works is to apply new methodologies that can overcome food

adulteration and mislabelling or to check authenticity of cereal based-products (Table 1). Bönick et al. (2017) reported an analytical strategy, based on in silico steps and LC-MS/MS, to check the authenticity of wheat, spelt and rye addition in bread products. MS has been reported as a promising alternative to ELISA, in particular for the detection but also quantification of proteins in contaminated food, as it can target multiple and very specific analytes (Martínez-Esteso et al., 2016).

Fiedler et al. (2014) identified a list of specific grain peptides of wheat, barley, rye, and oats for the detection of gluten contamination in several types of commercial flours. Specifically, targeted MS/MS method enabled the detection of two wheat peptide markers at a level of 10 ppm of wheat flour spiked into gluten-free oat flour. Martínez-Esteso et al. (2016) identified a set of unique wheat gluten peptides and proposed their use as markers of the presence of gluten related to the manifestation of CD symptoms. The authors reinforce the idea that this strategy can be applied to other allergens and that this is the first step towards the standardization of a new methodology, using LC-MS techniques, to evaluate the immunogenicity of different food matrices but also to produce reference materials, since the establishment of a set of markers is the first step to infer the presence of gluten and that enable the quantity of gluten present to be determined.

In the last decade, ion mobility spectrometry (IMS) has appeared as an analytical separation technique, especially important to the analysis of primary structures with a high degree of homology, such as gluten proteins. The IMS consists of an orthogonal separation technique, where for each value of m/z a spectrum of drift time is added. The drift time corresponds to the time the ion takes to cross the ionic mobility cell where an inert gas is inserted, allowing the determination of shock sections, or collision cross-section (Michaelevski et al., 2010). Thus, the ions can be further differentiated by size, shape, and charge, which allow separating by the three-dimensional conformation even peptides that present the same m/z or reverse peptides. In this way, the IMS can be applied to improve LC-MS and GC-MS workflows, since it increases method sensitivity by isolating the compounds of interest from background noise, improving confidence of identification, either in targeted or non-targeted approaches (Hernández-Mesa et al., 2019).

Wheat allergens from the non-gluten soluble protein fraction (albumins and globulins) have also been reported and identified by MS (Larre et al., 2011). Samples of diploid and hexaploid wheat were used to incite immunological reaction with human sera and then were

subsequently analyzed and identified by MS. The analysis of 2D spots revealed by immunoblotting leads to the MS-based identification of 39 IgE-binding proteins, some of them unknown thus far as wheat allergens. A recent study evaluated albumins and globulins from different genotypes of Brazilian wheat flour through the application of MSE and IMS, called UDMSE (Ultra Definition Mass Spectrometry). Collectively, about 5,900 proteins and 45,000 peptides (Victorio et al., 2018) were identified in the dataset and relatively quantified with 8 peptides/protein. Alves et al. (2018) reported that some of these proteins found have been previously described and associated with the development of respiratory allergies such as baker's asthma. Serpins, purinins, α-amylase/protease inhibitors, globulins, and farinins have also been associated with the humoral response to celiac disease (Huebener et al., 2014).

Following the same approach, Alves et al. (2018) evaluated the allergenic potential of nine wheat flours of different technological qualities by assessment of their immunogenic profiles. Peptides responsible for the manifestation of CD and other wheat-related allergies were identified in both gluten and soluble protein fractions. This work points to a relation between the variability in the expression of allergens and the technological quality of wheat flour, showing a distinct proteomic profile in flours of inferior technological quality, concluding that they can be more immunoreactive than the other qualities, especially due to the highest expression of two isoforms of serpins.

It is important to highlight that, to reach the identification of the peptide sequences by proteomic tools, the peptides must be present in the databases, so that the results obtained in the analyses can be cross-checked with those already consolidated (Altenbach et al., 2010). One of the major limitations to conducting proteomic studies in wheat was the lack of complete sequencing of the wheat genome (Bromilow et al., 2017). It is important to note that a high percentage of non-annotated proteins makes difficult the functional classification based on Gene Ontology. From the 414 soluble proteins found differentially expressed in common wheat flours, 85% proteins were not yet described, according to their biological function (Victorio et al., 2018). An alternative to reduce the misidentification of sequences is the use of de-novo sequencing to assemble wheat gluten gene sequences (W. Zhang et al., 2014). However, recently, the complete wheat genome was released, making it possible to improve the identifications of the proteins present in this cereal, since more peptides will be annotated in the proteomic databases (Ramírez-González et al., 2018).

Table 1. Overview of studies using LC-MS to detect gluten in foods.

Title	Food matrix	Techniques/methods	Reference
Novel aspects of quantitation of immunogenic wheat gluten peptides by liquid chromatography–mass spectrometry/mass spectrometry	Quinoa flour; whole grain corn flour; whole grain soy flour; vital wheat gluten flour; whole wheat flour; rye flour; barley flour; rice flour; oat flour; powdered iced tea mix; pasta; orzo; cheerios; hot sauce; bread; goldfish crackers; white vinegar; toothpaste; body lotion; body wash; beer; gin; vodka; rum; red wine; white wine and GF product	HPLC-ESI-TQS-MS/MS	Sealey-Voyksner et al. (2010)
Assessment of allergenicity of diploid and hexaploid wheat genotypes: identification of allergens in the albumin/globulin fraction.	Wheat; Human sera	ELISA; SDS-PAGE; immunoblotting; LC-MS/MS	Larre et al. (2011)
Measuring hordein (gluten) in beer – a comparison of ELISA and mass spectrometry.	Beer	Western blot; ELISA sandwich; MRM-MS	Tanner et al. (2013)
MS^E based multiplex protein analysis quantified important allergenic proteins and detected relevant peptides carrying known epitopes in wheat grain extracts	Wheat	nanoUPLC-QTOF-MS/MS	Uvackova et al. (2013)
The MS^E-proteomic analysis of gliadins and glutenins in wheat grain identifies and quantifies proteins associated with celiac disease and baker's asthma	Wheat	nanoUPLC-QTOF-MS/MS	Uvackova et al., 2013
Evaluation of qualitative and quantitative immunoassays to detect barley contamination in gluten-free beer with confirmation using LC-MS/MS.	Barley; GF beer	EZ Gluten assay; AllerTek Gluten ELISA; LC-QTof-MS/MS	Allred et al. (2014)

Characterization of grain-specific peptide markers for the detection of gluten by mass spectrometry.	Gluten; wheat flour; barley flour; rye flour; oat flour	nanoHPLC-ESI-pSMR; MS/MS	Fiedler et al. (2014)
Assessment of the allergenicity of soluble fractions from GM and commercial genotypes of wheats.	Wheat; GM wheat (*T. aestivum* and *T. durum*); Human sera	SDS-PAGE; western blot; immunoblotting; nanoLC-QTof-MS/MS	Lupi et al. (2014)
Specific nongluten proteins of wheat are novel target antigens in celiac disease humoral response	Wheat; Human sera	ELISA; SDS-PAGE; immunoblotting; MS/MS	Huebener et al. (2014)
Using mass spectrometry to detect hydrolyzed gluten in beer that is responsible for false negatives by ELISA	Beer	ELISA; nanoHPLC-ESI-MRM-MS	Michelle L. Colgrave et al. (2014)
Qualitative and quantitative determination of peptides related to celiac disease in mixtures derived from different methods of simulated gastrointestinal digestion of wheat products	Durum wheat (ground kernels; semolina; dough; extruded pasta; dried pasta and cooked pasta)	LC-ESI-MS	Prandi et al. (2014)
Label free targeted detection and quantification of celiac disease immunogenic epitopes by mass spectrometry	Wheat	On-line 2D nanoLC–MS/MS; UPLC-MRM-MS/MS	van den Broeck et al. (2015)
Allergen relative abundance in several wheat varieties as revealed via a targeted quantitative approach using MS	Wheat (*T. aestivum, T. durum, T. monococcum*)	LC-MS/MS	Rogniaux et al. (2015)
Proteomic profiling of 16 cereal grains and the application of targeted proteomics to detect wheat contamination	Barley; wheat; rye; oats; green wheat; amaranth; chia; quinoa; sorghum; tef; buckwheat; soy; millet; maize	SDS-PAGE; western blot; nanoUPLC-ESI-MRM-MS	Michelle L. Colgrave et al. (2015)
Multiplex liquid chromatography-tandem mass spectrometry for the detection of wheat, oat, barley, and rye prolamins towards the assessment of gluten-free product safety	Flour; seeds; pasta; biscuits; cookies; crackers; beverages; breads; breakfast cereals; snacks	HPLC-IonTrap-MS/MS	Manfredi et al. (2015)
Defining the wheat gluten peptide fingerprint via a discovery and targeted proteomics approach	Wheat gluten; GluVital®	ELISA; nanoUPLC-ESI-MS/MS	Martínez-Esteso et al. (2016)

Identification of barley-specific peptide markers that persist in processed foods and are capable of detecting barley contamination by LC-MS/MS	Barley; wheat; rye; oats; green wheat; amaranth; chia; quinoa; sorghum; tef; buckwheat; soy; millet; maize; breakfast cereals	nanoUPLC-ESI-MRM-MS	Michelle L. Colgrave et al. (2016)
Quantitation of the immunodominant 33-mer peptide from α-gliadin in wheat flours by liquid chromatography tandem mass spectrometry	Wheat flour	RP-HPLC; 1H qNMR; untargeted MS/MC; ESI-MRM-MS/MS	Schalk et al. (2017)
Determination of wheat, rye, and spelt authenticity in bread by targeted peptide biomarkers	Wheat; spelt; emmel wheat; einkorn wheat; barley; maize; oat; rye	UPLC-ESI-MRM-MS/MS	Bönick et al. (2017)
Peptides from gluten digestion: A comparison between old and modern wheat varieties	Wheat (*T. aestivum, T. durum, T. monococcum, T. dicoccum, T. spelta*)	UPLC-ESI-MS; HPLC-ESI-MS/MS	Prandi et al. (2017)
Development and validation of the detection method for wheat and barley glutens using mass spectrometry in processed foods	Seeds; flour; beers; cookies; beverages; GF products (GF flour; corn flour; apple wine; rice wine)	ELISA; LC-ESI-MRM-MS	Liao et al. (2017)
Using LC-MS to examine the fermented food products vinegar and soy sauce for the presence of gluten	Vinegar; malt vinegar; soy sauce	ELISA; UHPLC-MRM-MS/MS	Li et al. (2018)
Differential expression of albumins and globulins of wheat flours of different technological qualities revealed by nanoUPLC-UDMS^{E}	Wheat flour	nanoUPLC-HDMSE; nanoUPLC-UDMSE	Victorio et al. (2018)
Immunogenic and allergenic profile of wheat flours from different technological qualities revealed by ion mobility mass spectrometry	Wheat flour	nanoUPLC-MSE; nanoUPLC-UDMSE	Alves et al. (2018)
Detection and quantitation of immunogenic epitopes related to celiac disease in historical and modern hard red spring wheat cultivars	Wheat	RP-HPLC; SDS-PAGE; SRM-MS	Malalgoda et al. (2018)
Targeted liquid chromatography tandem mass spectrometry to quantitate wheat gluten using well-defined reference proteins	Wheat	RP-HPLC; untargeted MS/MS; MRM-MS	Schalk et al. (2018b)

Quantitation of specific barley, rye, and oat marker peptides by targeted liquid chromatography–mass spectrometry to determine gluten concentrations	Barley; Rye; Oat	RP-HPLC; untargeted MS/MS; MRM-MS; competitive R5-ELISA; SDS-PAGE	Schalk et al. (2018a)
A complete mass spectrometry (MS)-based peptidomic description of gluten peptides generated during in vitro gastrointestinal digestion of durum wheat: implication for celiac disease	Durum wheat	SDS-PAGE; UHPLC-ESI-MS/MS; UPLC-ESI-MS	Boukid et al. (2019)

3.2. MS-based quantification of immunogenic peptides

MS can also be applied for the selection and quantification of specific peptides by methods called MRM (multiple reaction monitoring) (Anderson & Hunter, 2006) or also called SRM (selected reaction monitoring) or PRM (parallel reaction monitoring) (Peterson et al., 2012), depending on the instrument manufacturer, which allow a targeted analysis of these peptides and their quantification even at minimum or trace concentrations. A set of strategies has been developed to measure the allergenic potential of various cereal species and LC-MRM/MS technology has been useful for the identification and quantification of peptides containing immunogenic epitopes at low levels of detection, such as femtomolar (van den Broeck et al., 2015). Different approaches can be used to quantify these peptides, like label-free quantification combined with external calibration.

This methodology was used by van den Broeck et al. (2015) to quantify CD immunogenic epitopes in three varieties of wheat (two hexaploid and one tetraploid). A list of nine peptides was proposed to create the calibration curves that quantified the amount of glia-α2 and glia-α20 in gluten extracts from the samples (Table 2). The reliability of the results depends on optimal digestion conditions and limit of detection and/or ionization properties of the peptides. Malalgoda et al. (2018) used the same approach to quantify immunogenic peptides from old and modern hard red spring wheat cultivars. Even though, it was not possible to associate the year of harvesting with the amounts of immunogenic epitopes and α-gliadin since it was randomly detected in all samples analyzed.

Table 2. List of gluten peptides selected for the creation of calibration curves (van den Broeck et al., 2015).

Peptide sequence
LQLQ**PFPQPQLPY**
LQLQ**PFPQPQLPYPQ**PQPF
LQLQ**PFPQPQLPYPQ**P*H*LPYPQPQPF
LQLQ**PFPQPQLPYPQPQLPYPQ**PQPF
LQLQ**PFPQPQLPYPQPQLPYPQPQLPYPQ**PQPF
RPQQPYPQPQPQY
RPQQPYPQ*S*QPQY
QQQLIPCRDVVL
QQILQQQLIPCRDVVL

CD-epitope sequences within the peptides are shown in bold.

Schalk et al. (2017) developed a targeted LC-MS/MS method to quantify the immunodominant gluten peptide called 33-mer (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF), which contains three different overlapping T-cell epitopes (PFPQPQLPY; PYPQPQLPY; PQPQLPYPQ) that initiate a strong immunological response (Shan et al., 2002). In this study, the quantitative data on contents of 33-mer peptide in different wheat cultivars was carried out by combining a stable isotope dilution assay with LC-MS/MS, as first reported for peptides by Stöcklin et al. (1997). The authors detected the presence of this peptide in 23 common wheat flours and in two spelt flours (T. spelta), but it was absent in tetraploid and diploid wheat flours. No obvious cluster formation between modern and old wheat cultivars and no correlations between contents of 33-mer and those of α-gliadins, gliadins, gluten, or crude protein were observed. Indeed, the harvest year had a higher influence on 33-mer contents than the cultivar. It is important to highlight that this was the first study that accurately quantitated the 33-mer peptide in wheat flours.

Recent studies use the combination of untargeted and targeted methods as a strategy to quantify gluten marker peptides in cereals and determine gluten concentrations in different types of samples (Schalk et al., 2018a, 2018b). Schalk et al. (2018b) developed a methodology that allowed the simultaneous determination of 33 marker peptides, 16 for wheat, seven for rye, seven for barley and three for oats using LC-MS/MS in MRM mode, using a labelled peptide as internal standard. Furthermore, they compared the LC-MS/MS results with those of R5 ELISA RP-HPLC and GP-HPLC-FLD (gel-permeation high-performance liquid chromatography with fluorescence detection) and found a strong correlation between LC-MS/MS and the other methods. When analyzing wheat starch samples, the LC-MS/MS and ELISA results agreed well in 4 out of 7 cases, but there were two samples where LC-MS/MS found substantially higher and one with lower gluten contents than ELISA. The lower values obtained by LC-MS/MS may be explained by the presence of other gluten peptides that were not monitored with the targeted method, whereas the higher values may be due to variable gliadin/glutenin ratios in wheat starches that may lead to an underestimation of gluten contents by ELISA (Schalk et al., 2018a).

One of the most important considerations when using targeted LC-MS/MS is the careful selection of gluten marker peptides, because only these pre-defined peptides will be monitored. Even a single amino acid substitution, deletion, insertion, or post-translational

modification will result in that marker peptide not being detected anymore, even if the sample may still contain other gluten-derived and possibly immunogenic peptides. While it is possible to use stable isotope labelled peptides or concatamers as internal standards to precisely quantify the selected peptides, the conversion of gluten peptide contents to gluten contents is far from being trivial. Legislation requires the result to be expressed as mg gluten/kg of the food, so that the correspondence between the amount of gluten and the resulting peptides needs to be established by careful calibration, also considering the whole sample preparation procedure. One of the most important points to verify is the extent of enzymatic hydrolysis. Matrix-matched calibration has been applied in many cases (Fiedler et al., 2014; Manfredi et al., 2015), but the use of well-defined gluten reference materials revealed the complexity of converting marker peptide contents to gluten contents (Schalk et al. 2018a,b). Further pros and cons of using ELISA and LC-MS/MS for gluten detection are given in Figure 2.

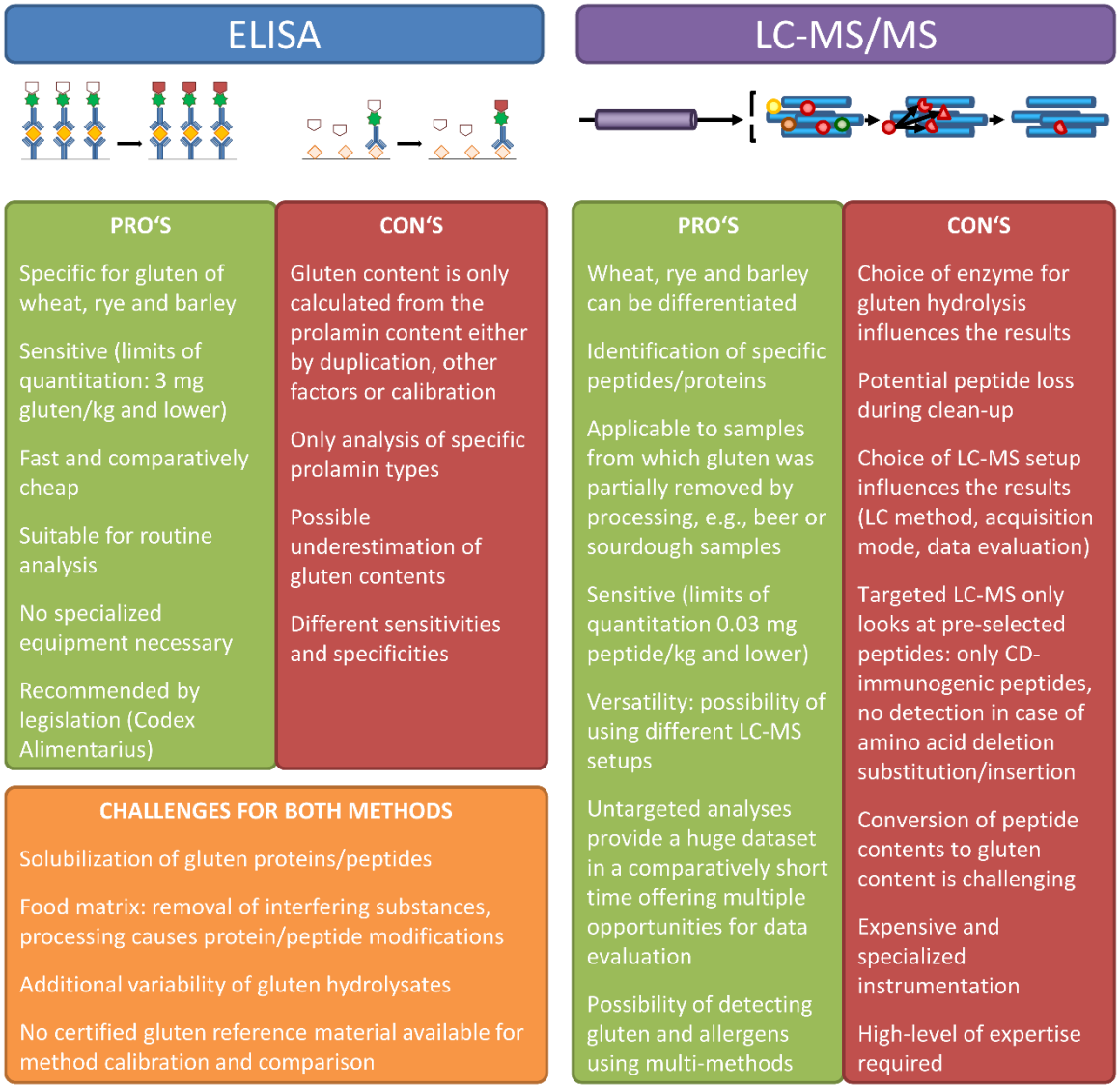

Figure 2. Overview of advantages and disadvantages of ELISA and LC-MS/MS for gluten detection.

A quantitative approach was also used to compare the relative abundance of 12 allergens in the albumins/globulins fraction in seven wheat varieties (Rogniaux et al., 2015). Allergens were monitored by targeted investigation of one to two proteotypic peptides (single protein peptides), and the abundance of some allergens was found to be quite stable among genotypes, while others, such as α-amylase inhibitors, showed clear differences depending on the wheat species, revealing themselves as possible markers of allergenicity in wheat. The content of allergenic polypeptides from these fractions was also investigated in common and genetically modified wheat (Lupi et al., 2014) revealing a large variation in the amounts of these allergens. The lack of information on the peptide sequences and epitopes responsible for

the allergies triggered by albumins/globulins render targeted studies in this protein fraction even more complicated.

4. Other strategies to unravel and to detect gluten peptides

Even with all the benefits of LC-MS/MS, such as the identification and quantification of specific proteins and peptides, new techniques have also been highlighted, such as the use of biosensors. Soler et al. (2016) used Surface Plasmon Resonance (SPR), a biosensor able to detect and quantify chemical and biological analytes quickly, sensitively, and specifically in complex field samples. SPR was able to detect gluten toxic peptides in the urine of CD patients and directly quantify the small digestive peptides without the need for prior extraction or purification procedures, so that the assay can be performed in 20 min. White et al. (2018) developed a floating gate transistor biosensor with longer analysis time (1.5 h), but it was still able to quantify wheat proteins faster than ELISA.

In addition to the shortest analysis time, biosensors also have high sensitivity at low detection limits and low cost. Chu et al. (2012) used a quartz crystal microbalance (QCM) immunosensor to detect gliadin in foods and had high sensitivity, being able to detect 8 ppb of this protein. In addition, the cost of materials for biosensor analyzes is estimated to be approximately 3-fold less than the cost of a single ELISA kit (Soler et al., 2016). In the future, immunosensors may be promising alternatives for existing immunochemical tests, such as ELISAs, because of their specificity and sensitivity (Scherf et al. 2016). However, this method does not allow the characterization of proteins and their respective peptides, as in LC-MS/MS. In addition, the type of sensor that is the best candidate to replace the ELISA still needs to be evaluated.

LC-MRM/MS analysis can also be linked to genomics to improve our understanding of the genes responsible for expressing allergenic proteins, culminating in the development of wheat varieties with a lower allergenic potential (Salentijn et al., 2013), increasing the variety of food options that can be consumed by GRD patients by ensuring food safety. Moreover, the studies about authenticity requires also an approach towards a well-defined "proteogenomic annotation" looking carefully at the specific peptide candidates from an enzymatic digest (Bönick et al., 2017).

5. Concluding remarks and perspectives

The use of LC-MS/MS strategies is the most useful and promising path to improve the identification and quantification of immunogenic peptides. Despite the methodological difficulties, it proves to be a fast, sensitive, and reproducible method. In addition, it can be extended to several other allergenic food matrices, like dairy, nuts, and seafood. Thus, knowing the profile of allergenic proteins of cereals is necessary as a basis, not only for future applications of MS in the quantification of gluten in food, but also to ensure the safety of consumers regarding food labelled cereal- or gluten-free.

Although the declaration of gluten-containing cereals on products labelled gluten-free is mandatory worldwide, there is no certified reference material available for gluten. The available reference material contains only gliadins that underestimate the gluten content, besides the problem of reproducing a new batch with similar properties and composition. The majority of MS-based studies have been conducted with the final objective to establish a reference material for gluten analysis starting from the study of specific grain peptide markers. Therefore, targeted high-resolution MS/MS methods allowed the quantification of low levels of specific marker peptides from different species and protein types.

When comparing LC-MS/MS methods to ELISA for gluten detection, ELISA still remains the method of choice in most cases, because it is fast, comparatively cheap, suitable for routine analyses and does not require highly specialized equipment. However, several studies have shown that ELISA may underestimate gluten contents especially in processed foods that have been extensively heat-treated or hydrolyzed. Untargeted LC-MS/MS is recommended to screen for the presence of gluten-derived peptides in products such as beer, malt vinegar and fermented sauces. However, there are some points that will equally all analytical methods because gluten extractability has been shown to be reduced substantially in heat-treated foods and processing-induced post-translational protein modifications will lead to reduced gluten detectability irrespective of the analytical method used.

The use of modern MS-based techniques, combining orthogonal separations with high sensitivity and reliable certified references materials will hopefully help to better comprehend the effect of food processing or plant breeding on gluten immunogenicity. Continued efforts in this area will also help to solve the questions about the selection of relevant target epitopes and even antibodies, taking account the high protein polymorphism and the fact that patients react individually to different proteins and present variable sensitivities.

6. Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7. Author Contributions Statement

T.O.A. organized and wrote the manuscript and the summarized table, C.T.S.D. complemented the writing and designed the figure, K.A.S. reviewed the manuscript and M.S.L.F. supervised and reviewed the manuscript.

8. Funding

This work was funded by Federal University of State of Rio de Janeiro (UNIRIO), FAPERJ (E-26/202.709/2018), CNPq (427116/2018-0) and in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

CHAPTER 2 – A PROTEOMIC APPROACH TO ASSESS THE IMPACT OF DIFFERENT EXTRUSION PROCESSES ON WHEAT-BASED PRODUCTS

Thais de O. Alves[1,2], Sabrina M. Geisslitz[3], Talita P. do Nascimento[1,2], Carlos W. P. de Carvalho[4], Katharina A. Scherf[3], Cristina Y. Takeiti[1,4], Mariana S. L. Ferreira[1,2]

[1] *Food and Nutrition Graduate Program (PPGAN), Laboratory of Bioactives, Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro, Brazil.*

[2] *Laboratory of Protein Biochemistry - Center of Innovation in Mass Spectrometry (LBP-IMasS), UNIRIO, Rio de Janeiro, Brazil.*

[3] *Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.*

[4] *Embrapa Agroindústria de Alimentos, Rio de Janeiro, Brazil.*

In preparation for submission.

Abstract

The study aimed to access the changes in the protein profile of wheat flours (*Triticum aestivum*) presenting contrasting technological end-use qualities after thermoplastic and cold extrusion processes. Two wheat genotypes (ORS Agile, bread/improver wheat; ORS Vintecinco, biscuit wheat) were used to obtain the extrudates, applying two different screw rotation speeds (150 and 300 rpm), and pasta by cold-extrusion process with further cooking. Protein were sequentially extracted – salt-soluble (AG), ethanol-soluble (Gli), urea-soluble with reduction (Glu) and characterized by RP-HPLC. Peptides were analyzed by nanoLC-MS/MS in a Q Exactive Plus Orbitrap by using data-dependent acquisition (DDA) in positive ion mode. MS data was processed in MaxQuant using a Triticum database from UniprotKB. Total protein content ranged from 8.6% to 13.8%. Independent of the processing applied, wheat genotypes showed different protein recovery yield: averaging 59% for ORS Agile, considerably lower than ORS Vintecinco (94%). Extruded samples shown the same protein profile for both genotypes, with a decrease in AG and Gli and increase of Glu content, indicating the protein insolubilization. The highest gluten polymerization occurred with the increasing of the screw rotation speed. The protein profile of raw and cooked pasta was also similar for both genotypes. AG and Gli contents were decreased when sample was submitted

to the cooking process, with the increasing of Glu. The Gli/Glu ratios were significantly lower than the respective wheat flours for all samples, except for raw pasta, which follows the same Osborne distribution as wheat flour. ω-5- and ω-1,2-gliadin contents were the least affected by processing in both genotypes, while α- and γ-gliadins, and LMW-GS, were mainly responsible by promoting interactions contributing to improve the gluten fraction. After processing proteomic data, a total of 339 proteins were evaluated, from which 250 were present in ORS Agile and 162 in ORS Vintecinco. PCA allowed the separation of samples in three distinct clusters, with the same pattern for both samples. Flours and expanded extrudates (300 rpm) presented a more distinct profile from the samples in third cluster. Heatmaps and cluster analysis confirmed flours as the most hierarchically distant from products. ORS Agile produces a stronger gluten network than biscuit wheat genotype ORS Vintecinco, the highest gluten force rendered more difficult the protein extraction after processing. Thermoplastic extrusion decreased protein extractability more than cold extrusion, but cooking pasta intensified the protein cross-linkages.

Keywords: cold extrusion, thermoplastic extrusion, foodomics, LC-MS/MS, RP-HPLC.

CHAPTER 3 – THE ROLE OF FERMENTATION IN WHEAT BREADS UNRAVELED BY PROTEOMICS-BASED ANALYSIS

Thais de O. Alves[1,2], Sabrina M. Geisslitz[3], Katharina A. Scherf[3], Cristina Y. Takeiti[1,4], Mariana S. L. Ferreira[1,2]

[1] *Food and Nutrition Graduate Program (PPGAN), Laboratory of Bioactives, Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro, Brazil.*

[2] *Laboratory of Protein Biochemistry - Center of Innovation in Mass Spectrometry (LBP-IMasS), UNIRIO, Rio de Janeiro, Brazil.*

[3] *Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.*

[4] *Embrapa Agroindústria de Alimentos, Rio de Janeiro, Brazil.*

In preparation for submission.

Abstract

Common or bread wheat (*Triticum aestivum*) is by far the most prevalent wheat species and represents around 95% of cultivated wheat. Gluten proteins are the main responsible to the unique rheological features allowing bread making and several other applications. Gluten viscoelasticity is essential in the manufacture of breads because it helps the dough rising and maintain its shape. Modern proteomic techniques such as LC-MS/MS help to understand the contribution of these proteins to the quality of the final product and even to identify the wheat cultivars the most suited for bread making. This study aimed to understand how fermentation by *S. cerevisiae* impacts protein proteome in breads manufactured with two wheat flours of contrasting technological end-use qualities (ORS Agile, bread/improver wheat; ORS Vintecinco, biscuit wheat). Proteins were sequentially extracted based on solubility (AG, Gli and reduced Glu) and characterized/quantified by RP-HPLC and Kjeldahl. Peptides were analysed by nanoLC-MS/MS in a Q Exactive Plus Orbitrap and data processed in MaxQuant using a *Triticum* database from UniprotKB. Total protein content ranged between 8.2%-12.9%. Extractability was lower in ORS Agile (62%) than ORS Vintecinco samples (98%).

AG and Gli amounts decreased in breads for both genotypes, while Glu contents increased. This increase was higher in ORS Vintecinco breads (~200%) than ORS Agile breads (54%). α-gliadin and γ-gliadin were mostly affected by the breadmaking process. LMW-glutenin was the main glutenin subunit affected. HMW-glutenins presented different behavior depending on the genotype. From proteomic data, after filter application, 250 proteins were evaluated (187 proteins for ORS Agile, 122 for ORS Vintecinco and 59 shared by both genotypes). ORS Agile samples presented 18 commonly expressed proteins, while ORS Vintecinco samples had 16. Some exclusive proteins could be selected as possible biomarkers to characterize the composition of commercial wheat flours as well as to have information about which flours were used to produce commercial breads. Proteomic analysis reveals differences in protein composition among samples produced from different flours. Differences in the molecular structure and composition of gluten proteins between genotypes may affect the susceptibility of gluten proteins to chemical and enzymatic modifications during flour processing. ORS Vintecinco is most affected by processing and the hypothesis to explain this finding is that weak wheat flour may be more inclined to suffer changes in gluten content during processing due to its lower gluten protein content when compared to strong wheat flour.

Keywords: bakery, wheat quality, foodomics, RP-HPLC, LC-MS/MS.

CHAPTER 4 – SHOTGUN PROTEOMIC ANALYSIS TO ASSESS ALLERGEN PROFILE OF GLUTEN AND NON-GLUTEN PROTEINS OF WHEAT-BASED PRODUCTS

Thais de O. Alves[1,2], Sabrina M. Geisslitz[3], Katharina A. Scherf[3], Cristina Y. Takeiti[1,4], Mariana S. L. Ferreira[1,2]

[1] *Food and Nutrition Graduate Program (PPGAN), Laboratory of Bioactives, Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro, Brazil.*

[2] *Laboratory of Protein Biochemistry - Center of Innovation in Mass Spectrometry (LBP-IMasS), UNIRIO, Rio de Janeiro, Brazil.*

[3] *Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.*

[4] *Embrapa Agroindústria de Alimentos, Rio de Janeiro, Brazil.*

Short communication in preparation to be submitted.

Abstract

Although wheat is widely produced and consumed worldwide, predisposed people can present sensitivities or inflammatory reactions to the wheat proteins. These reactions are caused by the stimulation of the immune system of individuals presenting celiac disease, non-celiac gluten sensitivity, wheat allergy, dermatitis herpetiformis, amongst other disorders. This study aimed to understand how wheat flours of contrasting technological end-use qualities submitted to different processes behave concerning the expression of allergenic proteins. Wheat-based products were produced from two different genotypes (*Triticum aestivum*) (ORS Agile, classified as bread/improver wheat; ORS Vintecinco, biscuit wheat): refined flours, thermoplastic extrudates (with different screw rotation speeds 150 and 300 rpm), raw and cooked pasta and breads with different formulations. Proteins were sequentially extracted: salt-soluble (AG), ethanol-soluble (Gli), urea-soluble with reduction (Glu) and then quantified by RP-HPLC. Peptides were analyzed by nanoLC-MS/MS in a Q Exactive Plus Orbitrap by using data-dependent acquisition (DDA). MS data was processed in MaxQuant using a Triticum database from UniprotKB. The identified proteins were searched against a customized databank of wheat allergenic proteins. Preliminary results concerning the

expression of allergenic proteins showed that from the set of identified proteins, only 3% were stablished as wheat allergens, but presenting proteins of all fractions (AG, Gli, Glu). From those, five were shared by samples from ORS Agile and ORS Vintecinco. ORS Agile flour had the highest expression intensity of allergenic proteins, while ORS Vintecinco expanded extrudate has the lowest. Although interesting and pointing to some effect of processing, further studies are necessary to determine the role of technological quality and processing in causing changes to the allergenic potential of wheat products.

CHAPTER 5 – CONCLUDING REMARKS

This thesis was dedicated to applying modern proteomic techniques to better understand the impact of food processing on the protein and allergenic profiles of wheat-based products obtained from wheat flour samples of contrasting technological qualities (bread/improver and biscuit wheat genotypes).

First, the bibliographic study developed reinforces and point out the utilization of the UHPLC-MS/MS strategy as a promising way to enhance the identification and quantification of immunogenic peptides in wheat products. Despite the methodological challenges, it proves to be a rapid, sensitive, and consistent method. Additionally, it can be extended to several other allergenic food matrices, such as dairy, nuts, and seafood. Therefore, understanding the profile of allergenic proteins of cereals is essential as a foundation for future applications of MS in gluten quantification in food and to ensure consumer safety regarding food labeled cereal- or gluten-free. Significant achievement has been done in the establishment of certified reference material (RM) for gluten, e.g. PWG-gliadin (Prolamin Working Group) (Van Eckert et al., 2006), wheat RM based on gluten and gliadin isolates (Schall et al., 2020), but until this moment, no consensus has been reported about this topic (Bugyi, 2022).

The first part of experiments conducted during this thesis aimed to apport better knowledge on how different processing techniques, such as thermoplastic and cold extrusion, cooking and fermentation affect protein behavior and profile, considering also two contrasting technological classes of wheat flour. Taken all these results together, wheat genotype, here presented by two different commercial classes of wheat flour, played a strong influence on protein profile and composition, because it significantly impacted the protein assembly when flour was submitted to different processing.

Comparing thermoplastic and cold extrudate samples, higher gluten force of ORS Agile (bread/improver wheat type) impaired protein extractability in relation to ORS Vintecinco. Due to the more intense polymerization of glutenins, proteins of high molecular weight were not efficiently extracted. Thermoplastic extrusion decreased protein extractability more than cold extrusion due to a higher protein insolubilization caused by the intensification of cross-linkages and polymerization of glutenins. When it comes to baking, structural changes caused by kneading and yeast action on the dough result in significant alterations in the protein profile in both technological quality flours and breads samples. The increase in

glutenin content was driven by the incorporation of gliadin subunits and soluble proteins to the polymers, followed by a decrease in the content of these fractions. ORS Vintecinco samples were more impacted by processing than ORS Agile.

Proteomic analysis allowed the identification of mostly affected protein subunits. In extrudates, prolamins and glutenins had their expression intensities elevated with the increase in thermomechanical forces caused by the increase in screw rotation speed. Evaluating pasta, the biggest influence was on prolamins, heat-induced during cooking. Proteomic evaluation of baked samples demonstrated the different protein expression among samples produced from different flours and allowed the understanding that α-, γ-, and ω-gliadins are most affected by processing than other proteins. Mechanical force and fermentation provokes a reduction in their expression when flour and non-fermented breads were compared to breads.

ORS Vintecinco samples showed a lower amount of gluten proteins compared to ORS Agile samples. The composition of gluten proteins was also proven to be different between these genotypes. Differences in gluten composition may influence the ability of gluten proteins to go through both chemical and enzymatic modifications during flour treatment (Abedi & Pourmohammadi, 2021). Thus, it is suggested that ORS Vintecinco, due to its lower amount of gluten proteins and possible differences in molecular composition and structure, may be more susceptible to modifications in gluten content during processing compared to ORS Agile.

Succeeding the understanding of the entire proteome of the samples, the use of this data for enlightening the allergenicity of wheat and wheat-based products could be applied. Approximately 12% were classified as allergens based on available online wheat allergen databases and the group of proteins was represented by soluble and metabolic wheat proteins. The gap in initial allergen content between the two flours was notable, with F1 expressing allergens more than three times as intensely as F2. This result can be explained by the higher gluten content in improver wheat flour, as most allergens are gluten proteins. The reduction in allergen expression intensity was more pronounced in samples from genotype 1.

Although interesting observations could be made pointing to some effect of processing, further studies including a higher number of genotypes are necessary to determine candidate proteins for biomarkers of quality or technological aptitude and also the role of processing in causing changes to the proteomic profile of wheat products. Additional studies, involving for instance the molecular size distribution of these protein fractions, are needed to

confirm the results found in this thesis and to understand the underlying mechanisms involved in the modifications of gluten content in different types of wheat flours.

Concerning scientific production associated with this Thesis, one literature review, and presented in Chapter 1, and three original papers (Chapters 2, 3 and 4), were composed. Literature review entitled "Modern Approaches in the Identification and Quantification of Immunogenic Peptides in Cereals by LC-MS/MS" is already published in a high-impact scientific Journal (Frontiers in Plant Science, IF = 4.298). The other three papers are currently under preparation for publication in high-impact scientific Journals. "A Proteomic Approach to Assess the Impact of Different Extrusion Processes on Wheat-Based Products", shown in Chapter 2, will be submitted to Food Chemistry (IF = 9.231), "The Role of Fermentation in Wheat Breads Unraveled by Proteomics-Based Analysis", in Chapter 3, will be submitted to Food Research International (IF = 7.425) and "Shotgun Proteomic Analysis To Assess Allergen Profile Of Gluten And Non-Gluten Proteins Of Wheat-Based Products" will be submitted as a short-communication to The Journal of Nutritional Biochemistry (IF = 6.117).

It is notorious to underline that two of the original research papers appear to be the first-of-its-kind, using proteomic approaches to understand how different types of extrusion and fermentation by *S. cerevisiae* promote changes in the protein structure and expression of different technological qualities of wheat flours. Another important contribution to gluten detection field is the most recent compiled list of wheat proteins and peptides with associated immunogenic action: 99 non-redundant proteins and 1,233 epitopes were grouped from different databases, repetitive contents were removed, modified or non-wheat sequences were also depleted to facilitate access to the proteins of interest. This list will be published as Supplementary material in paper 4.

In conclusion, the proteomic comprehension of wheat and gluten proteins is crucial for understanding the factors that affect the quality and functionality of wheat flour and its derived products. The proteomic analysis of wheat proteins requires advanced methods that can cope with the complexity and diversity of the gluten protein mixture. A combination of data-independent and data-dependent acquisition strategies using mass spectrometry has been proposed as a powerful approach to achieve comprehensive proteomic profiling of wheat gluten. Many challenges and gaps in the proteomic comprehension of wheat and gluten proteins are constantly faced in science, such as the identification of cultivar-specific sequences, the quantification of gluten protein fractions and subunits, the characterization of

gluten protein interactions and networks, and the elucidation of the molecular mechanisms of gluten protein synthesis and folding.

New perspectives and studies are proposed to address these issues and unravel the proteomics-based understanding of wheat and gluten proteins. The development of novel bioinformatics tools to integrate omics data and improve Foodomics interpretations and the application of label-free quantitative proteomics to compare the gluten protein profiles of different wheat cultivars, varieties and genotypes with different technological qualities and processing performances are strategies that can be developed in the future, in a possible continuation of the project.

Regarding allergenic and immunogenic potential of wheat flours and wheat-based products, the use of high-resolution proteomics techniques, such as nano LC-ESI-MS/MS, to compare the proteomes of different wheat flours, has been proved to be able revealing differences in the abundance and expression of proteins that may be involved in wheat hypersensitivities. The replication of analysis used in this study using different varieties of Brazilian wheat flours could lead to the identification of unique proteins for each wheat species or variety, which may serve as potential targets for research on wheat sensitivities.

The crop of Brazilian wheat varieties in warm regions, like the Brazilian Cerrado and Northeast, has emerged in the last decade. The nutritional and technological quality of these varieties have been described and Brazilian wheat cultivars are proving to be a potential substitute for wheat imported from Argentina, reducing the need to import this cereal in our country (Oliveira et al., 2022). However, there are no studies characterizing the presence of allergens and the reactivity of these cultivars. Expanding the current project to characterize and quantify the immunogenic potential in these samples is a promising field of research for the coming years.

Another perspective is to use targeted proteomics techniques, such as LC-MRM-MS, to detect and quantify specific allergenic or immunogenic proteins in wheat flours, as well as to monitor possible contamination of wheat proteins in other cereal grains or products. For this purpose, the definition of biomarkers that are related to the specific range of samples analyzed is essential. This approach can provide accurate and sensitive measurements of proteins that are relevant for wheat allergy or intolerance diagnosis. Furthermore, this approach can also ensure the safety and quality of gluten-free products by detecting traces of wheat contamination in grains or flours that are supposed to be gluten-free, such as sorghum,

buckwheat, oats, and soy, as well as to confirm the absence of gluten in commercial products labeled as gluten-free.

REFERENCES

AACC International. Approved Methods of Analysis. 11th ed. Method 10-09.01. – Basic Straight-Dough Bread-Baking Method – Long Fermentation. St Paul, MN: AACC International.

AACC International. Approved Methods of Analysis. 11th ed. Method 54-30.02. – Alveograph Method for Soft and Hard Wheat Flour. St Paul, MN: AACC International.

AACC International. Approved Methods of Analysis. 11th ed. Method 54-21.02. – Rheological Behavior of Flour by Farinograph: Constant Flour Weight Procedure. St Paul, MN: AACC International.

AACC International. Approved Methods of Analysis. 11th ed. Method 56-81.03. – Determination of Falling Number. St Paul, MN: AACC International.

Abedi, E., & Pourmohammadi, K. (2021). Chemical modifications and their effects on gluten protein: An extensive review. *Food chemistry, 343*, 128398. doi:https://doi.org/10.1016/j.foodchem.2020.128398

Abril, A. G., Carrera, M., Sánchez-Pérez, Á., & Villa, T. G. (2023). Gut Microbiome Proteomics in Food Allergies. *International Journal of Molecular Sciences, 24*(3), 2234.

Ahmed, S., de la Parra, J., Elouafi, I., German, B., Jarvis, A., Lal, V., . . . Watkins, S. (2022). Foodomics: A Data-Driven Approach to Revolutionize Nutrition and Sustainable Diets. *Frontiers in nutrition, 9*. doi:10.3389/fnut.2022.874312

Allred, L. K., Sealey Voyksner, J. A., & Voyksner, R. D. (2014). Evaluation of qualitative and quantitative immunoassays to detect barley contamination in gluten-free beer with confirmation using LC-MS/MS. *Journal of AOAC International, 97*(6), 1615-1625.

Altenbach, S. B., Vensel, W. H., & DuPont, F. M. (2010). Integration of transcriptomic and proteomic data from a single wheat cultivar provides new tools for understanding the roles of individual alpha gliadin proteins in flour quality and celiac disease. *Journal of Cereal Science, 52*(2), 143-151.

Alves, T. O., D'Almeida, C. T. S., & Ferreira, M. S. L. (2017). Determination of Gluten Peptides Associated with Celiac Disease by Mass Spectrometry *Celiac Disease and Non-Celiac Gluten Sensitivity*: InTech.

Alves, T. O., D'Almeida, C. T. S., Scherf, K. A., & Ferreira, M. S. L. (2019). Modern approaches in the identification and quantification of immunogenic peptides in cereals by LC-MS/MS. *Frontiers in Plant Science, 10*.

Alves, T. O., D'Almeida, C. T. S., Victorio, V. C. M., Souza, G. H. M. F., Cameron, L. C., & Ferreira, M. S. L. (2018). Immunogenic and allergenic profile of wheat flours

from different technological qualities revealed by ion mobility mass spectrometry. *Journal of Food Composition and Analysis, 73*, 67-75.

Anderson, L., & Hunter, C. L. (2006). Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Molecular & Cellular Proteomics, 5*(4), 573-588.

Aussenac, T., Rhazi, L., & Branlard, G. (2020). Molecular Weight Distribution of Polymeric Proteins in Wheat Grains: The Rheologically Active Polymers. *Foods, 9*(11), 1675.

Bakke, A., & Vickers, Z. (2007). Consumer liking of refined and whole wheat breads. *Journal of food science, 72*(7), S473-S480.

Banerjee, B. (2010). *Nutritional management of digestive disorders*: CRC Press.

Belitz, H. D., Grosch, W., & Schieberle, P. (2009). *Food chemistry, 4th revised and extended edn* (4 ed.).

Bönick, J., Huschek, G., & Rawel, H. M. (2017). Determination of wheat, rye and spelt authenticity in bread by targeted peptide biomarkers. *Journal of Food Composition and Analysis, 58*, 82-91. doi:https://doi.org/10.1016/j.jfca.2017.01.019

Boukid, F., Prandi, B., Faccini, A., & Sforza, S. (2019). A Complete Mass Spectrometry (MS)-Based Peptidomic Description of Gluten Peptides Generated During In Vitro Gastrointestinal Digestion of Durum Wheat: Implication for Celiac Disease. *Journal of The American Society for Mass Spectrometry*, 1-10.

Instrução Normativa nº 38, de 30 de novembro de 2010 - Regulamento técnico do trigo, (2010).

Bromilow, S., Gethings, L. A., Buckley, M., Bromley, M., Shewry, P. R., Langridge, J. I., & Mills, E. N. C. (2017). A curated gluten protein sequence database to support development of proteomics methods for determination of gluten in gluten-free foods. *Journal of proteomics*.

Bugyi, Z., Muskovics, G., Schall, E., Török, K., Hajas, L., Scherf, K., ... & Tömösközi. (2022). Classics in a new perspective: gluten as a special food safety and analytical challenge. . *Journal of Food Investigation, 68*(4).

Buse, J., Purves, R. W., Verrall, R. E., Badea, I., Zhang, H., Mulligan, C. C., . . . El-Aneed, A. (2014). The development and assessment of high-throughput mass spectrometry-based methods for the quantification of a nanoparticle drug delivery agent in cellular lysate. *J Mass Spectrom, 49*(11), 1171-1180. doi:10.1002/jms.3444

Carr, S. A., & Anderson, L. (2008). Protein Quantitation Through Targeted Mass Spectrometry: the Way Out of Biomarker Purgatory? *Clin Chem, 54*(11), 1749-1752.

Catassi, C., Bai, J. C., Bonaz, B., Bouma, G., Calabrò, A., Carroccio, A., . . . Dolinsek, J. (2013). Non-celiac gluten sensitivity: the new frontier of gluten related disorders. *Nutrients, 5*(10), 3839-3853.

Chu, P. T., Lin, C. S., Chen, W. J., Chen, C. F., & Wen, H. W. (2012). Detection of gliadin in foods using a quartz crystal microbalance biosensor that incorporates gold nanoparticles. *J Agric Food Chem, 60*(26), 6483-6492. doi:10.1021/jf2047866

Cifuentes, A. (2009). Food analysis and foodomics.

Colgrave, M. L. (2017). *Proteomics in Food Science - From Farm to Fork* (M. L. Colgrave Ed. 1 ed. Vol. 1).

Colgrave, M. L., Byrne, K., Blundell, M., & Howitt, C. A. (2016). Identification of barley-specific peptide markers that persist in processed foods and are capable of detecting barley contamination by LC-MS/MS. *Journal of proteomics, 147*, 169-176. doi:http://dx.doi.org/10.1016/j.jprot.2016.03.045

Colgrave, M. L., Goswami, H., Blundell, M., Howitt, C. A., & Tanner, G. J. (2014). Using mass spectrometry to detect hydrolysed gluten in beer that is responsible for false negatives by ELISA. *Journal of Chromatography A, 1370*, 105-114.

Colgrave, M. L., Goswami, H., Byrne, K., Blundell, M., Howitt, C. A., & Tanner, G. J. (2015). Proteomic profiling of 16 cereal grains and the application of targeted proteomics to detect wheat contamination. *Journal of Proteome Research, 14*(6), 2659-2668.

Conab. (2023). *Acompanhamento da safra brasileira grãos safra 2022/2023.* Retrieved from

Day, L., Augustin, M. A., Batey, I. L., & Wrigley, C. W. (2006). Wheat-gluten uses and industry needs. *Trends in Food Science & Technology, 17*(2), 82-90.

Delcour, J. A., Joye, I. J., Pareyt, B., Wilderjans, E., Brijs, K., & Lagrain, B. (2012). Wheat Gluten Functionality as a Quality Determinant in Cereal-Based Food Products. *Annual review of food science and technology, 3*(1), 469-492. doi:10.1146/annurev-food-022811-101303

Diaz-Amigo, C., & Popping, B. (2013). Accuracy of ELISA detection methods for gluten and reference materials: a realistic assessment. *Journal of agricultural and food chemistry, 61*(24), 5681-5688.

Egertson, J. D., MacLean, B., Johnson, R., Xuan, Y., & MacCoss, M. J. (2015). Multiplexed peptide analysis using data-independent acquisition and Skyline. *Nature Protocols, 10*(6), 887-903.

El-Aneed, A., Cohen, A., & Banoub, J. (2009). *Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers* (Vol. 44).

Everett, A. D. (2011). Introducing Proteomics: From Concepts to Sample Separation, Mass Spectrometry and Data Analysis Josip Lovric Wiley-VCH, 2011, p. 296 ISBN: 978-0-470-03523-8. *PROTEOMICS, 11*(22), 4439-4439. doi:10.1002/pmic.201190118

FAO. (2018). World Food Situation - FAO Cereal Supply and Demand Brief. Retrieved from http://www.fao.org/worldfoodsituation/csdb/en/

FAO. (2023). FAO Cereal Supply and Demand Brief. Retrieved from https://www.fao.org/worldfoodsituation/csdb/en/

FAOSTAT. (2018). *Food and agricultural organization of the United Nations*. Retrieved from http://www.fao.org/worldfoodsituation/csdb/en/

Ferreira, M. S. L., Mangavel, C., Rogniaux, H., Bonicel, J., Samson, M.-F., & Morel, M.-H. (2014). A MALDI-TOF based study of the in-vivo assembly of glutenin polymers of durum wheat. *Food research international, 63*, 89-99.

Fiedler, K. L., McGrath, S. C., Callahan, J. H., & Ross, M. M. (2014). Characterization of grain-specific peptide markers for the detection of gluten by mass spectrometry. *Journal of agricultural and food chemistry, 62*(25), 5835-5844.

García-Cañas, V., Simó, C., Herrero, M., Ibáñez, E., & Cifuentes, A. (2012). Present and future challenges in food analysis: foodomics: ACS Publications.

Goodman, R. E., Ebisawa, M., Ferreira, F., Sampson, H. A., van Ree, R., Vieths, S., . . . Taylor, S. L. (2016). AllergenOnline: A peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity. *Mol Nutr Food Res, 60*(5), 1183-1198. doi:10.1002/mnfr.201500769

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., & Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res, 38*(Web Server issue), W695-699. doi:10.1093/nar/gkq313

Graves, P. R., & Haystead, T. A. J. (2002). Molecular biologist's guide to proteomics. *Microbiology and Molecular Biology Reviews, 66*(1), 39-63.

Halbmayr-Jech, E., Rogers, A., Don, C., Prinster, M., & Collaborators: Augustin, G. B. C. B. Z. C. D. C. P. F. A. G. J. H. M. H. R. J. J. K. P. K. P. (2015). Gluten in rice flour and baked rice products by G12 sandwich ELISA: First Action 2014.03. *Journal of AOAC International, 98*(1), 103-111.

Hayta, M., & Alpaslan, M. (2001). Effects of processing on biochemical and rheological properties of wheat gluten proteins. *Nahrung/Food, 45*(5), 304.

Hernández-Mesa, M., Ropartz, D., García-Campaña, A. M., Rogniaux, H., Dervilly-Pinel, G., & Le Bizec, B. (2019). Ion Mobility Spectrometry in Food Analysis: Principles, Current Applications and Future Trends. *Molecules, 24*(15), 2706.

Herrero, M., Simó, C., García-Cañas, V., Ibáñez, E., & Cifuentes, A. (2012). Foodomics: MS-based strategies in modern food science and nutrition. *Mass spectrometry reviews, 31*(1), 49-69.

Hlywiak, K. H. (2008). Hidden sources of gluten. *Pract Gastroenterol, 32*, 27-39.

Hopper, J. T. R., C. V. (2019). Mass Spectrometry of Intact Protein Complexes *Proteomics for Biological Discovery*.

Huebener, S., Tanaka, C. K., Uhde, M., Zone, J. J., Vensel, W. H., Kasarda, D. D., . . . Altenbach, S. B. (2014). Specific nongluten proteins of wheat are novel target antigens in celiac disease humoral response. *J Proteome Res, 14*(1), 503-511.

Hussain, A., Larsson, H., Kuktaite, R., & Johansson, E. (2010). Mineral composition of organically grown wheat genotypes: contribution to daily minerals intake. *International journal of environmental research and public health, 7*(9), 3442-3456.

Ibáñez, C., Valdés, A., García-Cañas, V., Simó, C., Celebier, M., Rocamora-Reverte, L., . . . Segura-Carretero, A. (2012). Global Foodomics strategy to investigate the health benefits of dietary constituents. *Journal of chromatography A, 1248*, 139-153.

Inomata, N. (2009). Wheat allergy. *Current opinion in allergy and clinical immunology, 9*(3), 238-243.

Islam, S., Yu, Z., She, M., Zhao, Y., & Ma, W. (2019). Wheat gluten protein and its impacts on wheat processing quality. *Frontiers of Agricultural Science and Engineering, 6*(3), 279-287.

Janssen, A. M., Van Vliet, T., & Vereijken, J. M. (1996). Fundamental and empirical rheological behaviour of wheat flour doughs and comparison with bread making performance. *Journal of Cereal Science, 23*(1), 43-54.

Jira, W., & Münch, S. (2019). A sensitive HPLC-MS/MS screening method for the simultaneous detection of barley, maize, oats, rice, rye and wheat proteins in meat products. *Food chemistry, 275*, 214-223. doi:https://doi.org/10.1016/j.foodchem.2018.09.041

Jones, J. M., & Engleson, J. (2010). Whole grains: benefits and challenges. *Annual review of food science and technology, 1*, 19-40.

Juhász, A., Haraszi, R., & Maulis, C. (2015). ProPepper: a curated database for identification and analysis of peptide and immune-responsive epitope composition of cereal grain protein families. *Database, 2015*, bav100.

Kucek, L. K., Veenstra, L. D., Amnuaycheewa, P., & Sorrells, M. E. (2015). A grounded guide to gluten: how modern genotypes and processing impact wheat sensitivity. *Comprehensive Reviews in Food Science and Food Safety, 14*(3), 285-302.

Lafiandra, D., & Shewry, P. R. (2022). Wheat Glutenin polymers 2, the role of wheat glutenin subunits in polymer formation and dough quality. *Journal of Cereal Science*, 103487.

Larre, C., Lupi, R., Gombaud, G., Brossard, C., Branlard, G., Moneret-Vautrin, D. A., . . . Denery-Papini, S. (2011). Assessment of allergenicity of diploid and hexaploid wheat genotypes: identification of allergens in the albumin/globulin fraction. *Journal of proteomics, 74*(8), 1279-1289.

Li, H., Byrne, K., Galiamov, R., Mendoza-Porras, O., Bose, U., Howitt, C. A., & Colgrave, M. L. (2018). Using LC-MS to examine the fermented food products vinegar and soy sauce for the presence of gluten. *Food chemistry, 254*, 302-308.

Liao, Y.-S., Kuo, J.-H., Chen, B.-L., Tsuei, H.-W., Lin, C.-Y., Lin, H.-Y., & Cheng, H.-F. (2017). Development and validation of the detection method for wheat and barley glutens using mass spectrometry in processed foods. *Food Analytical Methods, 10*(8), 2839-2847.

Lupi, R., Masci, S., Rogniaux, H., Tranquet, O., Brossard, C., Lafiandra, D., . . . Larré, C. (2014). Assessment of the allergenicity of soluble fractions from GM and commercial genotypes of wheats. *Journal of Cereal Science, 60*(1), 179-186.

Malalgoda, M., Meinhardt, S. W., & Simsek, S. (2018). Detection and quantitation of immunogenic epitopes related to celiac disease in historical and modern hard red spring wheat cultivars. *Food chemistry, 264*, 101-107.

Manfredi, A., Mattarozzi, M., Giannetto, M., & Careri, M. (2015). Multiplex liquid chromatography-tandem mass spectrometry for the detection of wheat, oat, barley and rye prolamins towards the assessment of gluten-free product safety. *Analytica Chimica Acta, 895*, 62-70.

Martens, L., Vandekerckhove, J., & Gevaert, K. (2005). DBToolkit: processing protein databases for peptide-centric proteomics. *Bioinformatics, 21*(17), 3584-3585. doi:10.1093/bioinformatics/bti588

Martínez-Esteso, M. J., Nørgaard, J., Brohée, M., Haraszi, R., Maquet, A., & O'Connor, G. (2016). Defining the wheat gluten peptide fingerprint via a discovery and targeted proteomics approach. *Journal of proteomics*.

Martínez-Esteso, M. J., Brohée, M., Nørgaard, J., & O'Connor, G. (2017). Label-Free Proteomic Analysis of Wheat Gluten Proteins and Their Immunoreactivity to ELISA Antibodies. *Cereal chemistry, 94*(5), 820-826.

Michaelevski, I., Kirshenbaum, N., & Sharon, M. (2010). T-wave ion mobility-mass spectrometry: basic experimental procedures for protein complex analysis. *Journal of Visualized Experiments*(41), e1985-e1985.

Nadler, W. M., Waidelich, D., Kerner, A., Hanke, S., Berg, R., Trumpp, A., & Rosli, C. (2017). MALDI versus ESI: The Impact of the Ion Source on Peptide Identification. *J Proteome Res, 16*(3), 1207-1215. doi:10.1021/acs.jproteome.6b00805

Oliveira, M. E. A. S., Alves, T. d. O., Gutkoski, L. C., de Miranda, M. Z., Larraz Ferreira, M. S., & Takeiti, C. Y. (2022). Brazilian Cerrado wheat: Technological quality of genotypes grown in tropical locations. *Journal of Food Processing and Preservation, 46*(2), e16228. doi:https://doi.org/10.1111/jfpp.16228

Ortea, I. (2022). Foodomics in health: Advanced techniques for studying the bioactive role of foods. *TrAC Trends in Analytical Chemistry*, 116589.

Osborne, T. B. (1907). *The proteins of the wheat kernel*: Carnegie Institution of Washington.

Owusu-Apenten, R. (2002). *Food protein analysis: quantitative effects on processing* (Vol. 118): CRC Press.

Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., & Coon, J. J. (2012). Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Molecular & Cellular Proteomics, 11*(11), 1475-1488.

Pfeifer, M., Kugler, K. G., Sandve, S. R., Zhan, B., Rudi, H., Hvidsten, T. R., . . . International Wheat Genome Sequencing, C. (2014). Genome interplay in the grain transcriptome of hexaploid bread wheat. *Science, 345*(6194), 1250091.

Pietzak, M. M., & Fasano, A. (2005). Celiac disease: a new paradigm of an immune-mediated disorder due to dietary gluten *Reviews in Food and Nutrition Toxicity* (Vol. 3, pp. 243): Citeseer.

Prandi, B., Faccini, A., Tedeschi, T., Cammerata, A., Sgrulletta, D., D'Egidio, M. G., . . . Sforza, S. (2014). Qualitative and quantitative determination of peptides related to celiac disease in mixtures derived from different methods of simulated gastrointestinal digestion of wheat products. *Analytical and bioanalytical chemistry, 406*(19), 4765-4775.

Prandi, B., Tedeschi, T., Folloni, S., Galaverna, G., & Sforza, S. (2017). Peptides from gluten digestion: A comparison between old and modern wheat varieties. *Food research international, 91*, 92-102.

Presutti, R. J., Cangemi, J. R., Cassidy, H. D., & Hill, D. A. (2007). Celiac disease. *American family physician, 76*(12), 1795-1802.

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., . . . Pasha, A. (2018). The transcriptional landscape of polyploid wheat. *Science, 361*(6403), eaar6089.

Redl, A., Guilbert, S., & Morel, M. H. (2003). Heat and shear mediated polymerisation of plasticized wheat gluten protein upon mixing. *Journal of Cereal Science, 38*(1), 105-114.

Ribeiro, M., Rodriguez-Quijano, M., Nunes, F. M., Carrillo, J. M., Branlard, G., & Igrejas, G. (2016). New insights into wheat toxicity: Breeding did not seem to contribute to a prevalence of potential celiac disease's immunostimulatory epitopes. *Food chemistry, 213*, 8-18.

Rogniaux, H., Pavlovic, M., Lupi, R., Lollier, V., Joint, M., Mameri, H., . . . Larré, C. (2015). Allergen relative abundance in several wheat varieties as revealed via a targeted quantitative approach using MS. *PROTEOMICS, 15*(10), 1736-1745.

Saghatelian, A., & Cravatt, B. F. (2005). Global strategies to integrate the proteome and metabolome. *Current opinion in chemical biology, 9*(1), 62-68.

Salentijn, E. M. J., Esselink, D. G., Goryunova, S. V., van der Meer, I. M., Gilissen, L. J. W. J., & Smulders, M. J. M. (2013). Quantitative and qualitative differences in celiac disease epitopes among durum wheat varieties identified through deep RNA-amplicon sequencing. *BMC genomics, 14*(1), 905.

Sapone, A., Bai, J. C., Ciacci, C., Dolinsek, J., Green, P. H. R., Hadjivassiliou, M., . . . Schumann, M. (2012). Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC medicine, 10*(1), 13.

Schalk, K., Koehler, P., & Scherf, K. A. (2018a). Quantitation of Specific Barley, Rye, and Oat Marker Peptides by Targeted Liquid Chromatography–Mass Spectrometry To Determine Gluten Concentrations. *Journal of agricultural and food chemistry, 66*(13), 3581-3592.

Schalk, K., Koehler, P., & Scherf, K. A. (2018b). Targeted liquid chromatography tandem mass spectrometry to quantitate wheat gluten using well-defined reference proteins. *PloS one, 13*(2), e0192804.

Schalk, K., Lang, C., Wieser, H., Koehler, P., & Scherf, K. A. (2017). Quantitation of the immunodominant 33-mer peptide from α-gliadin in wheat flours by liquid chromatography tandem mass spectrometry. *Scientific Reports, 7*, 45092.

Schall, E., Scherf, K. A., Bugyi, Z., Török, K., Koehler, P., Schoenlechner, R., & Tömösközi, S. (2020). Further Steps Toward the Development of Gluten Reference Materials – Wheat Flours or Protein Isolates? *Frontiers in Plant Science, 11*. doi:10.3389/fpls.2020.00906

Scherf, K. A., Brockow, K., Biedermann, T., Koehler, P., & Wieser, H. (2016). Wheat-dependent exercise-induced anaphylaxis. *Clinical & Experimental Allergy, 46*(1), 10-20.

Scherf, K. A., Ciccocioppo, R., Pohanka, M., Rimarova, K., Opatrilova, R., Rodrigo, L., & Kruzliak, P. (2016). Biosensors for the diagnosis of celiac disease: current status and future perspectives. *Molecular biotechnology, 58*, 381-392.

Scherf, K. A., & Poms, R. E. (2016). Recent developments in analytical methods for tracing gluten. *Journal of Cereal Science, 67*, 112-122.

Schneider, M., Lane, L., Boutet, E., Lieberherr, D., Tognolli, M., Bougueleret, L., & Bairoch, A. (2009). The UniProtKB/Swiss-Prot knowledgebase and its Plant Proteome Annotation Program. *Journal of proteomics, 72*(3), 567-573. doi:10.1016/j.jprot.2008.11.010

Schuster, C., Huen, J., & Scherf, K. A. (2023). Comprehensive study on gluten composition and baking quality of winter wheat. *Cereal chemistry, 100*(1), 142-155.

Sealey-Voyksner, J. A., Khosla, C., Voyksner, R. D., & Jorgenson, J. W. (2010). Novel aspects of quantitation of immunogenic wheat gluten peptides by liquid chromatography–mass spectrometry/mass spectrometry. *Journal of Chromatography A, 1217*(25), 4167-4183.

Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., . . . Khosla, C. (2002). Structural basis for gluten intolerance in celiac sprue. *Science, 297*(5590), 2275-2279.

Shewry, P. R., & Halford, N. G. (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. *Journal of Experimental Botany, 53*(370), 947-958.

Silva, J. C., Gorenstein, M. V., Li, G.-Z., Vissers, J. P. C., & Geromanos, S. J. (2006). Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Molecular & Cellular Proteomics, 5*(1), 144-156.

Soler, M., Estevez, M. C., Moreno Mde, L., Cebolla, A., & Lechuga, L. M. (2016). Label-free SPR detection of gluten peptides in urine for non-invasive celiac disease follow-up. *Biosens Bioelectron, 79*, 158-164. doi:10.1016/j.bios.2015.11.097

Spaenij–Dekking, L., Kooy–Winkelaar, Y., van Veelen, P., Drijfhout, J. W., Jonker, H., van Soest, L., . . . Koning, F. (2005). Natural variation in toxicity of wheat: potential for selection of nontoxic varieties for celiac disease patients. *Gastroenterology, 129*(3), 797-806.

Šramková, Z., Gregová, E., & Šturdík, E. (2009). Chemical composition and nutritional quality of wheat grain. *Acta chimica slovaca, 2*(1), 115-138.

Stöcklin, R., Vu, L., Vadas, L., Cerini, F., Kippen, A. D., Offord, R. E., & Rose, K. (1997). A stable isotope dilution assay for the in vivo determination of insulin levels in humans by mass spectrometry. *Diabetes, 46*(1), 44-50.

Tanner, G. J., Colgrave, M. L., Blundell, M. J., Goswami, H. P., & Howitt, C. A. (2013). Measuring hordein (gluten) in beer – a comparison of ELISA and mass spectrometry. *PloS one, 8*(2), e56452.

Thanhaeuser, S. M., Wieser, H., & Koehler, P. (2014). Correlation of Quality Parameters with the Baking Performance of Wheat Flours. *Cereal chemistry, 91*(4), 333-341. doi:https://doi.org/10.1094/CCHEM-09-13-0194-CESI

The UniProt Consortium. (2017). UniProt: the universal protein knowledgebase. *Nucleic acids research, 45*(D1), D158-D169. doi:10.1093/nar/gkw1099

Thompson, T., & Méndez, E. (2008). Commercial assays to assess gluten content of gluten-free foods: why they are not created equal. *Journal of the american dietetic association, 108*(10), 1682-1687.

USDA. (2018, December 12, 2018). Wheat Data. Retrieved from https://www.ers.usda.gov/data-products/wheat-data/

Uvackova, L., Skultety, L., Bekesova, S., McClain, S., & Hajduch, M. (2013). MSE based multiplex protein analysis quantified important allergenic proteins and detected relevant peptides carrying known epitopes in wheat grain extracts. *J Proteome Res, 12*(11), 4862-4869.

van den Broeck, H. C., Cordewener, J. H. G., Nessen, M. A., America, A. H. P., & van der Meer, I. M. (2015). Label free targeted detection and quantification of celiac disease immunogenic epitopes by mass spectrometry. *Journal of Chromatography A, 1391*, 60-71.

Van Eckert, R., Berghofer, E., Ciclitira, P. J., Chirdo, F., Denery-Papini, S., Ellis, H. J., . . . Mamone, G. (2006). Towards a new gliadin reference material–isolation and characterisation. *Journal of Cereal Science, 43*(3), 331-341.

Verheyen, C., Jekle, M., & Becker, T. (2014). Effects of Saccharomyces cerevisiae on the structural kinetics of wheat dough during fermentation. *LWT-Food Science and Technology, 58*(1), 194-202.

Victorio, V. C. M., Souza, G. H. M. F., Santos, M. C. B., Vega, A. R., Cameron, L. C., & Ferreira, M. S. L. (2018). Differential expression of albumins and globulins of wheat flours of different technological qualities revealed by nanoUPLC-UDMSE. *Food chemistry, 239*, 1027-1036.

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics, 25*(9), 1189-1191. doi:10.1093/bioinformatics/btp033

Weiss, W., Huber, G., Engel, K. H., Pethran, A., Dunn, M. J., Gooley, A. A., & Görg, A. (1997). Identification and characterization of wheat grain albumin/globulin allergens. *Electrophoresis, 18*(5), 826-833.

White, S. P., Frisbie, C. D., & Dorfman, K. D. (2018). Detection and Sourcing of Gluten in Grain with Multiple Floating-Gate Transistor Biosensors. *ACS Sensors, 3*(2), 395-402. doi:10.1021/acssensors.7b00810

Wieser, H. (2007). Chemistry of gluten proteins. *Food Microbiology, 24*(2), 115-119.

Wieser, H., & Koehler, P. (2009). Is the calculation of the gluten content by multiplying the prolamin content by a factor of 2 valid? *European Food Research and Technology, 229*(1), 9-13.

Wieser, H., Koehler, P., & Scherf, K. A. (2023). Chemistry of wheat gluten proteins: Quantitative composition. *Cereal chemistry, 100*(1), 36-55.

Xhaferaj, M., Alves, T. O., Ferreira, M. S. L., & Scherf, K. A. (2020). Recent progress in analytical method development to ensure the safety of gluten-free foods for celiac disease patients. *Journal of Cereal Science, 96*, 103114.

Zhang, M., Bai, X., & Zhang, Z. (2011). Extrusion process improves the functionality of soluble dietary fiber in oat bran. *Journal of Cereal Science, 54*(1), 98-103.

Zhang, M., Jia, R., Ma, M., Yang, T., Sun, Q., & Li, M. (2022). Versatile wheat gluten: functional properties and application in the food-related industry. *Critical Reviews in Food Science and Nutrition*, 1-17.

Zhang, W., Ciclitira, P., & Messing, J. (2014). PacBio sequencing of gene families—A case study with wheat gluten genes. *Gene, 533*(2), 541-546.

APPENDIX - RECENT PROGRESS IN ANALYTICAL METHOD DEVELOPMENT TO ENSURE THE SAFETY OF GLUTEN-FREE FOODS FOR CELIAC DISEASE PATIENTS

Majlinda Xhaferaj[1], Thais O. Alves[2,3], Mariana S. L. Ferreira[2,3], Katharina A. Scherf[1]

[1]*Department of Bioactive and Functional Food Chemistry, Institute of Applied Biosciences, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.*

[2]*Food and Nutrition Graduate Program (PPGAN), Laboratory of Bioactives, Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro, Brazil.*

[3]*Laboratory of Protein Biochemistry - Center of Innovation in Mass Spectrometry (LBP-IMasS), UNIRIO, Rio de Janeiro, Brazil.*

Published in Journal of Cereal Science. DOI: 10.1016/j.jcs.2020.103114.

Abstract

As laid down by the Codex Alimentarius, products bearing a gluten-free label must not contain gluten levels above 20 mg/kg to be safe for consumption by celiac disease patients. Analytical methods to detect gluten from wheat, rye and barley need to be sufficiently sensitive, specific, suitable for routine analyses and validated by collaborative studies. With continuous progress in the field of gluten analysis, the aim of this paper is to provide an up-to-date overview of legislation regarding gluten-free products worldwide, as well as immunochemical, proteomics-based, genomics-based and other methods designed to analyze gluten traces. While ELISA test kits and PCR are still most widely used in quality control, liquid chromatography tandem mass spectrometry (LC-MS/MS) is gaining more and more importance by providing unprecedented insights into gluten. Several other methods such as immunosensors, other sensors and microarrays are being developed. The pros and cons of the different methods are discussed as well as the remaining challenges, including the need for improved extraction procedures, comprehensive reference materials and independent reference methods.

Keywords: enzyme-linked immunosorbent assay (ELISA), gluten, gluten-free, liquid chromatography, mass spectrometry, proteomics.