

# **Targeted und untargeted Proteomics von Weizenmehlproteinen in Zusammenhang zur Backqualität**

Zur Erlangung des akademischen Grades einer

DOKTORIN DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und Biowissenschaften

des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

staatlich geprüfte Lebensmittelchemikerin

Christine Kämper

1. Referentin: Prof. Dr. Katharina Scherf

2. Referentin: Prof. Dr. Andrea Hartwig

Tag der mündlichen Prüfung: 18.07.2025







## Publikationsliste

Diese Arbeit wurde im Rahmen des Projekts „BigBaking“ (FKZ: 2818404B18) angefertigt. Die Förderung des Vorhabens erfolgte aus Mitteln des Bundesministeriums für Landwirtschaft, Ernährung und Heimat (BMLEH) aufgrund eines Beschlusses des Deutschen Bundestags. Die Projektträgerschaft erfolgte über die Bundesanstalt für Landwirtschaft und Ernährung (BLE). Während der Projektlaufzeit wurden Teile dieser Arbeit in jährlichen Projekttreffen und Zwischenberichten präsentiert.

### Publikationen in Fachzeitschriften (Peer-Review)

**Kaemper C.,** Mossburger J., Geyer, M., Hartl L., Geisslitz S., Scherf K. A. (2025). Comparative shotgun proteomics analysis of wheat gluten proteins digested by various peptidases. *Current Research in Food Science*, 11:101095.

Beitrag zur Publikation: Datenerfassung, Auswertung, Recherche, Methodik, Visualisierung, Schreiben - ursprünglicher Entwurf, Überprüfung und Bearbeitung.

**Kaemper C.,** Geyer, M., Hartl L., Geisslitz S., Scherf K. A. (2025). Absolute quantification of gluten protein groups and their relation to wheat baking quality. *Food Research International*, 227:118228.

Beitrag zur Publikation: Datenerfassung, Auswertung, Recherche, Methodik, Visualisierung, Schreiben - ursprünglicher Entwurf, Überprüfung und Bearbeitung.

### Sonstige Publikationen

Stemler C. D., Kaemper C., Hammann S., Börner A., Scherf K. A. (2024). Lipidomic profiling of common wheat flours from 1891–2010. *Journal of Agricultural and Food Chemistry*, 72:25997–26005.

Beitrag zu Publikation: Methodik, Schreiben, Überprüfung und Bearbeitung.

### Vorträge

Muller O., Geyer M., Kämper C., Knopf O., Hartl L., Geisslitz S., Scherf K. A. (2024). Project BigBaking – Ensuring high baking quality and nitrogen-use efficiency through optimization of storage protein composition and enzymatic activity. *proWeizen-Conference*. Quedlinburg, Deutschland, 16.-17.04.2024.

Knopf O., Geyer M., Kämper C., Muller O., Hartl L., Geisslitz S., Scherf K. A. (2023). Project BigBaking – Ensuring high baking quality and nitrogen-use efficiency through optimization of storage protein composition and enzymatic activity. *proWeizen-Conference*. Quedlinburg, Deutschland, 18.-19.04.2023.

Kämper C., Geißlitz S., Scherf K. A. (2022). Weizen-Proteomics zur Bestimmung der Backqualität (Wheat proteomics to determine baking quality). *Arbeitstagung des Regionalverbandes Südwest der Lebensmittelchemischen Gesellschaft (Workshop of the Regional Association Southwest within the Society of Food Chemistry)*. Online, 29.-30.03.2022.

Knopf O., Geyer M., Kämper C., Muller O., Hartl L., Geißlitz S., Scherf K. A. (2022). Project BigBaking – Ensuring high baking quality and nitrogen-use efficiency through optimization of storage protein composition and enzymatic activity. *proWeizen-Conference*. Online, 23.-24.03.2022.

Knopf O., Kaemper C., Muller O., Geisslitz S., Scherf K. A. (2021). BigBaking – Ensuring high baking quality and nitrogen-use efficiency through optimization of storage protein composition and enzymatic activity. *1st Virtual Wheat Quality Meeting, WheatInitiative*. Online, 24.-25.11.2021.

Geyer M., Kämper C., Knopf O., Muller O., Hartl L., Geißlitz S., Scherf K. A. (2021). BigBaking Project – Quantitative wheat protein analysis using LC-MS/MS. *GFPi Annual Meeting*. Online, 11.11.2021.

## Posterbeiträge

Kaemper C., Geisslitz S., Scherf K. A. (2024). Wheat proteomics for analysing baking quality. *51. European Bioinformatics Community for Mass Spectrometry (EuBIC-MS) Winter School 2024*. Winterberg, Deutschland, 15.-19.01.2024.

Kämper C., Geißlitz S., Scherf K. A. (2023). Weizen-Proteomics für die Untersuchung der Backqualität. *51. Deutsche Lebensmittelchemietage (51. German Food Chemistry Days)*. Bonn, Deutschland, 21.-23.08.2023.

Knopf O., Castro A., Lenzen D., Quiros J., Markwitz S., Kleist E., Quarten M., Steier A., Geyer M., Hartl L., Kaemper C., Scherf K. A., Pude R., Poorter H., Rascher U., Muller O. (2022). Effects of Free-Air Carbon Dioxide Enrichment (FACE) on photosynthesis, phenology and yield of winter wheat. *7th International Plant Phenotyping Symposium 2022*. Wageningen, Niederlande, 27.-30.09.2022.

Kaemper C., Geisslitz S., Scherf K. A. (2022). Untargeted and targeted proteomic analysis of wheat related to baking quality. *The International Mass Spectrometry Conference (IMSC) 2022*. Maastricht, Niederlande, 27.08.-02.09.2022.

# Inhaltsverzeichnis

Abkürzungsverzeichnis.....	VIII
Ein-Buchstaben-Code für Aminosäuren.....	X
Zusammenfassung.....	XI
Summary.....	XII
1 Einleitung.....	1
1.1 Weizen.....	1
1.1.1 Weizenkornproteine.....	2
1.1.2 Weizengenom.....	4
1.1.3 Backqualität.....	9
1.2 Proteomics.....	14
1.2.1 Bottom-Up Proteomics.....	15
1.2.2 Quantitative Proteomics.....	17
1.2.3 Proteomics von Weizenproteinen.....	18
2 Zielsetzung.....	21
3 Ergebnisse und Diskussion.....	22
3.1 Untargeted Proteomics von Weizenmehl: Methodenoptimierung.....	22
3.1.1 Proteinextraktion.....	22
3.1.2 Proteinverdau.....	27
3.1.3 Proteinaufreinigung.....	28
3.1.4 Weitere Optimierungsversuche.....	29
3.1.5 Optimierung der MS-Parameter.....	32
3.1.6 Zusammenfassung.....	35
3.2 Untargeted Proteomics von Weizenmehl: Variation der Peptidasen.....	36
3.2.1 Anzahl identifizierter Peptide.....	37
3.2.2 Anzahl identifizierter Proteingruppen.....	40
3.2.3 Vergleich der Glutenproteinzusammensetzung.....	44
3.3 Targeted Proteomics von Weizenmehl.....	54
3.3.1 Methodenentwicklung PRM-Methode.....	54

3.3.2	Vergleich mit untargeted Ergebnissen.....	61
3.4	Weizenproteomics in Verbindung mit Backqualität .....	68
3.4.1	Untargeted Messung der BMWpop.....	68
3.4.2	Targeted Messung der BMWpop.....	90
4	Fazit und Ausblick.....	101
5	Experimenteller Teil .....	103
5.1	Probensortiment.....	103
5.2	Chemikalien und Enzyme .....	104
5.3	Quantifizierung der Glutenfraktionen mittels RP-UHPLC .....	105
5.4	Methodenoptimierung untargeted Proteomics .....	107
5.4.1	Untargeted Proteomics: Finale LC-MS/MS-Probenaufarbeitung .....	107
5.4.2	Proteinextraktion Variation.....	109
5.4.3	Proteinverdau: Variation Enzym-Substrat-Verhältnis.....	110
5.4.4	Proteinaufreinigung .....	110
5.4.5	Weitere Optimierungsversuche.....	111
5.4.6	Optimierung der MS-Parameter.....	111
5.5	Untargeted Proteomics mit verschiedenen Peptidasen .....	113
5.5.1	Proteinaufarbeitung.....	113
5.5.2	LC-MS/MS-Messung.....	113
5.5.3	Proteinidentifizierung und -quantifizierung.....	114
5.6	Methodenentwicklung targeted Proteomics .....	115
5.6.1	Peptidstandards.....	115
5.6.2	Identifizierung von Markerpeptiden.....	115
5.6.3	Probenaufarbeitung für SIDA und targeted LC-MS/MS.....	116
5.6.4	Responsegerade .....	116
5.6.5	Targeted LC-MS/MS Messung .....	117
5.6.6	Proteinquantifizierung .....	117
5.6.7	Präzision und Wiederfindung .....	118
5.6.8	Bestimmungs- und Nachweisgrenze.....	118
5.6.9	Targeted Proteomics: Finale LC-MS/MS Aufarbeitung .....	118
5.7	Instrumentelle Angaben.....	119
5.8	Statistische Analysen.....	120

6	Literaturverzeichnis.....	121
7	Anhang.....	138
7.1	Untargeted Proteomics von Weizen: Variation der Peptidasen.....	138
7.2	Targeted Proteomics von Weizen.....	140
7.3	Weizenproteomics in Verbindung mit Backqualität .....	147

## Abkürzungsverzeichnis

ACN	Acetonitril
AGC	<i>Automatic Gain Control</i>
ATI	Amylase/Trypsin-Inhibitoren
BMWpop	<i>Bavarian MAGIC Wheat Population</i>
C	Chymotrypsin
CAA	2-Chloracetamid
CT	Chymotrypsin + Trypsin
DAD	Dioden-Array-Detektor
DDA	<i>Data Dependent Acquisition</i>
DIA	<i>Data Independent Acquisition</i>
DNA	Desoxyribonukleinsäure ( <i>Deoxyribonucleic Acid</i> )
DTT	1,4-Dithiothreitol
ESI	Elektrospray-Ionisation
EtOH	Ethanol
FA	Ameisensäure ( <i>Formic Acid</i> )
FASP	<i>Filter Aided Sample Preparation</i>
HCD	<i>Higher-energy Collisional Dissociation</i>
HorRat	<i>Horwitz Ratio</i>
HMW/LMS-GS	Hoch-/niedermolekulare Gluteninuntereinheiten ( <i>High/Low-Molecular-Weight Glutenin Subunits</i> )
ICC	<i>International Association for Cereal Science and Technology</i>
IS 1-11	Interner Standard 1-11
IT	Injektionszeit (Injection Time)
KI	Künstliche Intelligenz
LC-MS/MS	Flüssigkeitschromatographie – Tandem-Massenspektrometrie
LFQ	Labelfreie Quantifizierung ( <i>Label-free Quantification</i> )
LOD	Nachweisgrenze ( <i>Limit of Detection</i> )

LOQ	Bestimmungsgrenze ( <i>Limit of Quantification</i> )
MAGIC	<i>Multiparent Advanced Generation Intercross</i>
MS	Massenspektrometrie
MW	Molekulargewicht
<i>m/z</i>	Masse-zu-Ladungs-Verhältnis
NCE	Normalisierte Kollisionsenergie ( <i>Normalized Collision Energy</i> )
P	Pepsin
P1-11	Peptid 1-11
PAGE	Polyacrylamidgelelektrophorese
PCA	Hauptkomponentenanalyse ( <i>Principal Component Analysis</i> )
PWG	<i>Prolamin Working Group</i>
PRM	<i>Parallel Reaction Monitoring</i>
QTL	<i>Quantitative Trait Loci</i>
R	Replikat
RIL	Rekombinante Inzuchtslinien ( <i>recombinant inbred lines</i> )
RP-HPLC	Umkehrphasen-Hochleistungsflüssigkeitschromatographie ( <i>Reversed Phase – High Performance Liquid Chromatography</i> )
RSD	Relative Standardabweichung
SDS	Natriumdodecylsulfat ( <i>Sodium Dodecyl Sulfate</i> )
SIDA	Stabilisotopen-Verdünnungsanalyse ( <i>Stable Isotope Dilution Assay</i> )
SPE	Festphasenextraktion ( <i>Solid Phase Extraction</i> )
T	Trypsin
TFA	Trifluoressigsäure ( <i>Trifluoroacetic Acid</i> )
TCEP	Tris(2-carboxyethyl)phosphin
TLY	Thermolysin
TRIS	Tris(hydroxymethyl)aminomethan
UV	Ultraviolettes Licht
US	Ultraschall

## Ein-Buchstaben-Code für Aminosäuren

Alanin	A
Arginin	R
Asparagin	N
Asparaginsäure	D
Cystein	C
Glutaminsäure	E
Glutamin	Q
Glycin	G
Histidin	H
Isoleucin	I
Leucin	L
Lysin	K
Methionin	M
Phenylalanin	F
Prolin	P
Serin	S
Threonin	T
Tryptophan	W
Tyrosin	Y
Valin	V

# Zusammenfassung

Weizen zählt weltweit zu den wichtigsten Grundnahrungsmitteln. Angesichts von Bevölkerungswachstum, Klimawandel und begrenzten Ressourcen bleibt die Züchtung qualitativ hochwertiger Sorten von zentraler Bedeutung. Proteomics bildet dabei ein zentrales Bindeglied zwischen Genotyp und Phänotyp und ist entscheidend für die gezielte Optimierung der Proteinzusammensetzung mit dem Ziel einer guten Backqualität. In komplexen Populationen wie der Bavarian Multiparent Advanced Intercross Weizenpopulation (BMWpop) ermöglicht sie die effektive Verknüpfung genetischer und phänotypischer Daten für züchterische Zwecke. Im Rahmen dieser Arbeit wurden eine untargeted und eine targeted Bottom-Up Proteomics-Methode entwickelt, optimiert und auf die BMWpop angewendet, um proteomische Merkmale mit Backqualitätsparametern zu verknüpfen.

Die targeted Methode baute auf der optimierten Probenaufarbeitung der untargeted Methode auf, bei der jeder Schritt hinsichtlich Peptidausbeute und Robustheit optimiert wurde. Bei der enzymatischen Hydrolyse wurden vier Peptidasen und eine Kombination getestet. Die Peptidasen lieferten unterschiedliche Einblicke in das Weizenglutenproteom. Mit Trypsin wurden die meisten Peptide identifiziert und die Spaltung war am selektivsten, wobei jedoch keine  $\omega$ -Gliadine detektiert wurden. Die kombinierte Sequenzabdeckung betrug 61,3 %, wobei Thermolysin allein 51,1 % erreichte. Die Identifikation geeigneter Glutenmarkerpeptide war durch geringe Sequenzspezifität und ungünstige Peptidlängen anspruchsvoll. Bei der targeted Proteomics-Methodenentwicklung gelang die erfolgreiche Identifikation, Verifikation und Validierung von Markerpeptiden für elf definierte Glutenproteingruppen.

Beide Methoden ermöglichten eine umfassende Charakterisierung der Proteinzusammensetzung der BMWpop - sowohl relativ als auch absolut - und lieferten konsistente Ergebnisse. Glutenproteine zeigten Korrelationen mit zentralen Backqualitätsparametern wie Proteingehalt und Brotvolumen. Den größten Einfluss auf die Backqualität hatten hochmolekulare Gluteninuntereinheiten (HMW-GS), gefolgt von niedermolekulare Gluteninuntereinheiten und  $\gamma$ -Gliadinen, während  $\alpha$ -Gliadine nur geringe Effekte zeigten. Innerhalb der Fraktionen variierten die Beiträge einzelner Proteine deutlich. Eine genetische Rangfolge ließ sich nur für die HMW-GS ableiten, wobei Dy10 den stärksten Einfluss zeigte. Insgesamt wirkten HMW-GS vom  $\gamma$ -Typ stärker in Bezug auf die Backqualität als solche vom  $\alpha$ -Typ. Die Ergebnisse betonen die Bedeutung der Glutenproteinzusammensetzung für die Weizenzüchtung. Verbesserte Datenbanken, fortschrittlichere Analyseverfahren und durch künstliche Intelligenz gestützte Methoden könnten in Kombination mit Quantitative Trait Locus-Analysen künftig sowohl die Identifikation relevanter Proteinmarker als auch die Vorhersage der Backqualität wesentlich verbessern.

## Summary

Wheat is one of the most important staple foods worldwide. In light of population growth, climate change, and limited resources, the breeding of high-quality varieties remains of central importance. Proteomics serves as a crucial link between genotype and phenotype and is essential for the targeted optimization of protein composition with the aim of achieving good baking quality. In complex populations such as the Bavarian Multiparent Advanced Intercross Wheat Population (BMWpop), proteomics enables the effective integration of genetic and phenotypic data for breeding purposes. In this study, both untargeted and targeted bottom-up proteomics methods were developed, optimized, and applied to the BMWpop to link proteomic features with baking quality parameters.

The targeted method was based on the optimized sample preparation established for the untargeted approach, with each step refined for peptide yield and robustness. During enzymatic hydrolysis, four peptidases and one combination were tested. These enzymes produced distinct representations of the wheat gluten proteome. Trypsin yielded the highest number of identified peptides and exhibited the most selective cleavage, although no  $\omega$ -gliadins were detected. The combined sequence coverage across all enzymes was 61.3%, with thermolysin alone achieving 51.1%. The identification of suitable gluten marker peptides proved challenging due to low sequence specificity and unfavorable peptide lengths. Nevertheless, the development of the targeted proteomics method enabled the successful identification, verification, and validation of marker peptides for eleven defined gluten protein groups.

Both methods enabled a comprehensive characterization of the protein composition of the BMWpop - both in relative and absolute terms - and yielded consistent results. Gluten proteins showed correlations with key baking quality parameters such as protein content and loaf volume. High-molecular-weight glutenin subunits (HMW-GS) had the greatest influence on baking quality, followed by low-molecular-weight glutenin subunits (LMW-GS) and  $\gamma$ -gliadins, while  $\alpha$ -gliadins had only minor effects. Within each fraction, the contributions of individual proteins varied considerably. A genetic ranking could only be established for the HMW-GS, with Dy10 showing the strongest effect. Overall, HMW-GS of the  $\gamma$ -type exhibited a stronger impact on baking quality than those of the  $\alpha$ -type. These findings highlight the importance of specifically considering gluten protein composition in wheat breeding. Improved databases, more advanced analytical techniques, and methods supported by artificial intelligence could, in combination with quantitative trait locus analyses, significantly enhance both the identification of relevant protein markers and the prediction of baking quality in the future.

# 1 Einleitung

Weichweizen stellt eine der global bedeutendsten Kulturpflanzen dar und dient als essenzielles Grundnahrungsmittel für einen großen Teil der Weltbevölkerung. Schätzungsweise 788,7 Millionen Tonnen wurden weltweit im Jahr 2023/24 angebaut, wobei der weltweite pro-Kopf-Verbrauch 67,7 kg/Jahr betrug [1]. Dies macht etwa 20 % der menschlichen Kalorienaufnahme aus [2]. Die Backeigenschaften von Weizen sind einzigartig und haben wesentlich zur Entwicklung einer breiten Palette an Lebensmitteln beigetragen. Insbesondere seine Eignung zur Herstellung vielfältiger Brotspezialitäten hat dem Weichweizen die Bezeichnung „Brotweizen“ eingebracht [3]. Vor diesem Hintergrund ist eine hohe Backqualität der angebauten Weichweizensorten von zentraler Bedeutung.

Ein weiterer Erfolgsfaktor des Weizens ist seine Fähigkeit, sich an unterschiedlichste klimatische Bedingungen anzupassen [4]. Dennoch ist die Weizenproduktion aufgrund klimatischer Entwicklungen von 2000 bis 2008 rückläufig gewesen [5]. Weiterhin stellt der Einsatz von stickstoffbasierten Düngemitteln zur Ertragssteigerung eine Belastung für die Umwelt dar [6]. Aus Umwelt- und Klimaschutzgründen unterliegt die Düngung in Deutschland deshalb agrarpolitischen Einschränkungen, welche in der Düngeverordnung (DüV) gesetzlich festgelegt sind [7]. Das dadurch reduzierte Stickstoffangebot führt zu einer Verringerung des Kornproteingehalts, einem entscheidenden Parameter für die Backqualität. Um Sorten zu identifizieren und zu züchten, die auch bei reduziertem Proteingehalt gute Backeigenschaften aufweisen, ist eine detaillierte Analyse jener Proteine erforderlich, die maßgeblich zur Ausprägung der Backqualität beitragen.

## 1.1 Weizen

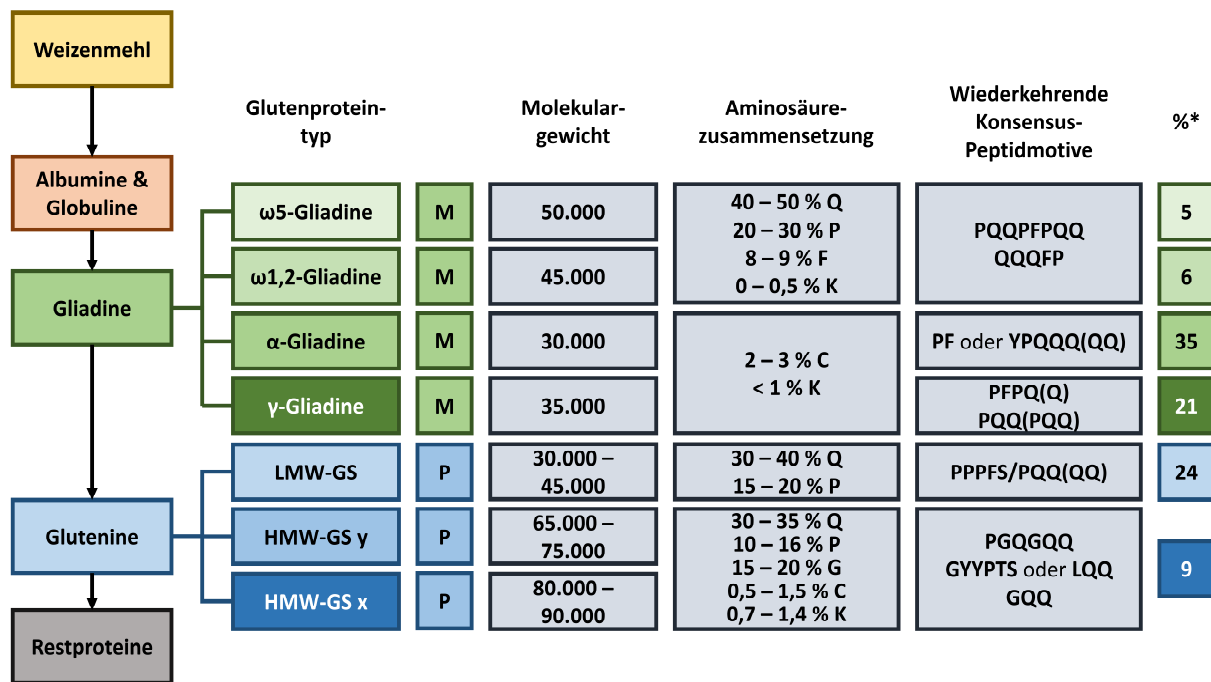
Der heutige Weichweizen (*Triticum aestivum*) gehört zur Familie der Süßgräser (*Poaceae*). Die Gattung *Triticum* umfasst weitere Weizenarten wie Hartweizen (*Triticum durum*), Einkorn (*Triticum monococcum*) und Emmer (*Triticum dicoccon*), wobei *Triticum aestivum* weltweit am bedeutendsten ist. Beim Vermahlen des reifen Weizenkorns wird das nährstoffreiche Endosperm von Kleie und Keimling getrennt, um Mehl zu gewinnen. Wenn Mehl mit Wasser gemischt wird, bildet sich ein viskoelastisches Glutenproteinnetzwerk aus, welches die Grundlage für die Herstellung von Brot und anderen Backwaren darstellt [8]. Die Proteine im Weizenkorn sind daher maßgeblich für die charakteristischen Backeigenschaften von Weizen verantwortlich und stehen im Fokus zahlreicher wissenschaftlicher Untersuchungen.

### 1.1.1 Weizenkornproteine

Während das Weizenkorn für den Menschen vor allem als Nahrungsmittel von Bedeutung ist, erfüllt es für die Pflanze eine andere Hauptfunktion: Es dient als Nährstoffspeicher für den Keimling. Zur Vorbereitung darauf lagert das heranreifende Korn im Endosperm Proteine, Stärke, Lipide und Zucker ein [9]. Der Gesamtproteingehalt von Weizenkörnern variiert je nach Sorte in einem weiten Bereich von etwa 8 % bis 16 % [10, 11]. Die Proteine teilen sich auf in etwa 20–30 % Stoffwechselproteine und 70–80 % Speicherproteine, die sogenannten Glutenproteine. Diese befinden sich ausschließlich im stärkehaltigen Endosperm, wo sie zunächst in separaten Proteinkörpern gelagert sind [12]. Während der Kornreifung bilden sie eine zusammenhängende Proteinmatrix um die Stärkekörner [8, 12]. Schließlich werden sie bei der Keimung mobilisiert, um das Wachstum des Keimlings zu fördern, indem sie ihn mit Stickstoff, Aminosäuren und Energie versorgen [13]. Neben ihrer biologischen Funktion bestimmen die Glutenproteine die viskoelastischen Eigenschaften des Teiges, indem sie dort das Glutennetzwerk ausbilden [14].

#### 1.1.1.1 Proteinfraktionen

Analytisch können Weizenkornproteine nach ihrer Löslichkeit in verschiedene Proteinfraktionen (Osborne-Fraktionen) eingeteilt werden (**Abbildung 1**). Dabei handelt es sich um die Fraktionen der Albumine (wasserlöslich), Globuline (salzwasserlöslich), Gliadine (löslich in wässrigen Alkoholen) und Glutenine (unlöslich) [15]. Die monomeren Gliadine können nicht mit Wasser oder Salzlösungen extrahiert werden, dafür mit 60 % Ethanol oder auch 50 % Propanol. Die polymeren Glutenine können erst extrahiert werden, nachdem sie zu monomeren Gluteninuntereinheiten reduziert wurden. Das ist möglich mit einem Reduktionsmittel wie Dithiothreitol (DTT) in 50 % Propanol. Ein kleiner, unlöslicher Teil der Proteine verbleibt zusammen mit Stärke im Rückstand. Dabei handelt es sich unter anderem um Membranproteine und Lipoproteine [16].



**Abbildung 1:** Klassifikation und Charakterisierung von Weizenproteinen in verschiedene Fraktionen und Typen. Inhalt basiert auf Wieser et al. [13] und Shewry & Belton [17]. \*Die angegebenen Prozentsätze stellen typische Werte für die jeweiligen Glutenproteintypen dar, bezogen auf die Gesamtmenge an Gluten, die durch modifizierte Osborne-Fraktionierung und Umkehrphasen-Hochleistungsflüssigkeitschromatographie bestimmt wurde [18]; HMW-GS, hochmolekulare Gluteninuntereinheit; LMW-GS, niedermolekulare Gluteninuntereinheit; **M**, Monomer; **P**, Polymer; C, Cystein; F, Phenylalanin; G, Glycin; K, Lysin; L, Leucin; P, Prolin; Q, Glutamin; S, Serin; T, Threonin; Y, Tyrosin.

Die Albumin/Globulinfraktion enthält hauptsächlich Stoffwechselproteine und Enzyme. Die Gliadine und die Glutenine bilden zusammen die Gruppe der Glutenproteine, welche weiter in verschiedene Glutenproteintypen unterteilt werden. Gliadine werden anhand ihrer unterschiedlichen Mobilität bei der sauren Polyacrylamidgelelektrophorese (polyacrylamide gel electrophoresis, PAGE) weiter unterteilt in α-, γ-, ω1,2- und ω5-Gliadine. Getrennt durch Natriumdodecylsulfat(sodium dodecyl sulfate, SDS)-PAGE werden Glutenine in niedermolekulare (low-molecular-weight, LMW-) und hochmolekulare Gluteninuntereinheiten (high-molecular-weight glutenin subunit, HMW-GS) unterschieden [19].

### 1.1.1.2 Glutenproteine

Die α- und γ-Gliadine haben ein Molekulargewicht (molecular weight, MW) von 30.000 bis 35.000, was einer Länge von etwa 260–330 Aminosäureresten entspricht (**Abbildung 1**). Die ω-Gliadine sind mit einem MW von 45.000 bis 50.000 größer, mit Aminosäurelängen von 370–420 Resten. Die LMW-GS sind mit einem MW von 30.000 bis 45.000 eine diversere Gruppe. Die HMW-GS werden weiter unterteilt in den x-Typ (MW von 80.000–90.000, entspricht etwa 800 Aminosäureresten) und den γ-Typ (MW von 65.000–75.000, entspricht etwa 600 Aminosäureresten) [13, 17].

Glutenproteine haben insgesamt eine unausgewogene Aminosäurezusammensetzung mit 11–29 % Prolin und 32–53 % Glutamin. Durch die sekundäre Aminosäuregruppe des Prolins entstehen Knicke in der Proteinfaltung, sodass sich die Proteine dichter in das stärkehaltige Endosperm anordnen lassen. Glutamin enthält eine zweite Aminogruppe, die als wichtiger Nährstofflieferant bei der Keimung dient [12]. Die hohen Anteile von Glutamin und Prolin sind folglich der biologischen Funktion von Glutenproteinen in dem Weizenkorn geschuldet. Daneben enthalten Glutenproteine viele hydrophobe Aminosäurereste wie Phenylalanin, Leucin und Isoleucin (in der Summe bis zu 28 %) und sind dadurch eher hydrophob. Dies sorgt bei der Keimung einerseits dafür, dass die Proteine nicht aus dem Korn austreten können. Andererseits ermöglicht es beim Backprozess unter anderem die Bildung eines Teigs, da sich die Proteine in dem zugesetzten Wasser nicht lösen. Eine weitere für die spätere Teigbildung sehr wichtige Aminosäure ist Cystein, welche zu rund 2 % in Gluten enthalten ist. Die Formation von Disulfidbrücken durch Cysteinreste (intra- und intermolekular) bestimmt wesentlich die Funktionalität des Glutens im Teig [20]. Des Weiteren enthalten Glutenproteine nur wenige Aminosäuren mit geladenen Seitenketten (2–7 %), welche sich vorwiegend in den nicht-repetitiven Regionen (siehe nächster Abschnitt) befinden. Die essentiellen Aminosäuren Lysin (0,3–1,1 %), Methionin (0,0–1,8 %) und Tryptophan (0,0–1,0 %) sind ebenfalls nur in geringen Mengen enthalten [13, 17].

Zusätzlich sind die Aminosäuren in repetitiven Sequenzen mit hoher Sequenzhomologie angeordnet [13, 17]. Ihre Aminosäuresequenzen lassen sich in mehrere sogenannte „Domänen“ gliedern, von denen eine durch Tandem- oder verstreute Wiederholungen kurzer Peptidmotive (mit einer Länge von 3 bis 12 Aminosäuren) charakterisiert ist (**Abbildung 1**). Die Länge dieser repetitiven Abschnitte variiert sowohl zwischen den verschiedenen Glutenproteintypen als auch innerhalb dieser. Im Allgemeinen sind die Wiederholungsdomänen bei den HMW-GS und den  $\omega$ -Gliadinen länger als bei den schwefelreichen  $\alpha$ - und  $\gamma$ -Gliadinen sowie bei den LMW-GS. Unabhängig von ihrer Gruppe enthalten alle Wiederholungsdomänen der Glutenproteine hohe Anteile an Glutamin und Prolin und sind somit verantwortlich für den insgesamt hohen Anteil dieser Aminosäuren. Zudem treten in bestimmten Proteinen weitere Aminosäuren vermehrt auf, wie etwa Phenylalanin in  $\omega$ -Gliadinen oder Glycin in HMW-GS [17]. Der repetitive Charakter und die hohe Ähnlichkeit innerhalb der Weizenglutenproteine ist dem komplexen, hexaploiden Weichweizengenom geschuldet.

### 1.1.2 Weizengenom

Weizenpflanzen wie wilder Emmer und Einkorn gehörten zu den ersten Pflanzen, die vor etwa 10.000 Jahren im westlichen Teil des fruchtbaren Halbmonds (Winterregengebiet am nördlichen Rand der

Syrischen Wüste) kultiviert wurden [21]. Der heutige Weizen ist aus natürlichen Kreuzungen mehrerer Wildarten hervorgegangen. Anhand des Genomvergleichs von Weizen mit anderen Sorten konnte die evolutionäre Verwandtschaft und damit die stammesgeschichtliche Entwicklung des Weizens weitgehend nachvollzogen werden. So ist sein hexaploides Genom aus zwei Polyploidisierungsereignissen (Verdopplung des gesamten Chromosomensatzes) aus diploiden Pflanzen entstanden [4]. Das erste war die Hybridisierung der Genome von rotem, wildem Einkorn (*Triticum urartu*;  $2n = 2x = 14$ ; AA;  $2n$  ist die Anzahl der Chromosomen in jeder somatischen Zelle und  $2x$  ist die Basischromosomennummer) und wahrscheinlich einer Spezies, die mit dem Wildgras *Aegilops speltoides* ( $2n = 2x = 14$ ; SS) verwandt ist [22]. Es entstand das tetraploide *Triticum turgidum* ( $2n = 4x = 28$ ; AABB), ein Vorfahre von wildem Emmer. Das zweite Ereignis war die natürliche Hybridisierung mit dem Wildgras *Aegilops tauschii* (DD) [22, 23]. Somit enthält das Weizengenom sieben Chromosomen pro Subgenom A, B und D, die als Duplikatsätze vorliegen ( $2n = 6x = 42$ ; AABBDD). Daraus ergeben sich 7 Chromosomengruppen, die mit 1A, 1B, 1D, 2A, 2B etc. bezeichnet werden, je nach Chromosomengruppe und zugehörigem Subgenom.

Aufgrund der Polyploidie ist das Weizengenom mit knapp 16 Gb sehr groß (fünffach größer als das menschliche Genom) [24]. Des Weiteren enthält es mehr als 85 % repetitive DNA. Die Größe und der repetitive Charakter des Weizengenoms haben die Erstellung eines Referenzgenoms lange verzögert, sodass es erst in den letzten Jahren vollständig sequenziert worden ist. Dies wurde mit dem Weizenkultivar *Chinese Spring* durchgeführt, um die IWGSC RefSeq v1.0 zu erstellen. [25]. Es wurden insgesamt 105.200 und 154.779 Gene identifiziert, die mit hoher beziehungsweise geringerer Zuverlässigkeit proteincodierend sind. Insgesamt wurden 290.756 Pseudogene festgestellt, die damit einen höheren Anteil ausmachen als proteincodierende Gene. In diesem Zusammenhang konnten weiterhin Genfamilien zugeordnet werden, die mit biotischen oder abiotischen Stress, dem RNA-Stoffwechsel in Organellen und der männlichen Fruchtbarkeit oder der Endproduktqualität in Verbindung stehen. Für die Endproduktqualität sind unter anderem jene Genfamilien relevant, die für Albumine, Globuline, Gliadine und Glutenine codieren [25].

#### 1.1.2.1 Kartierung von Glutenproteingenen

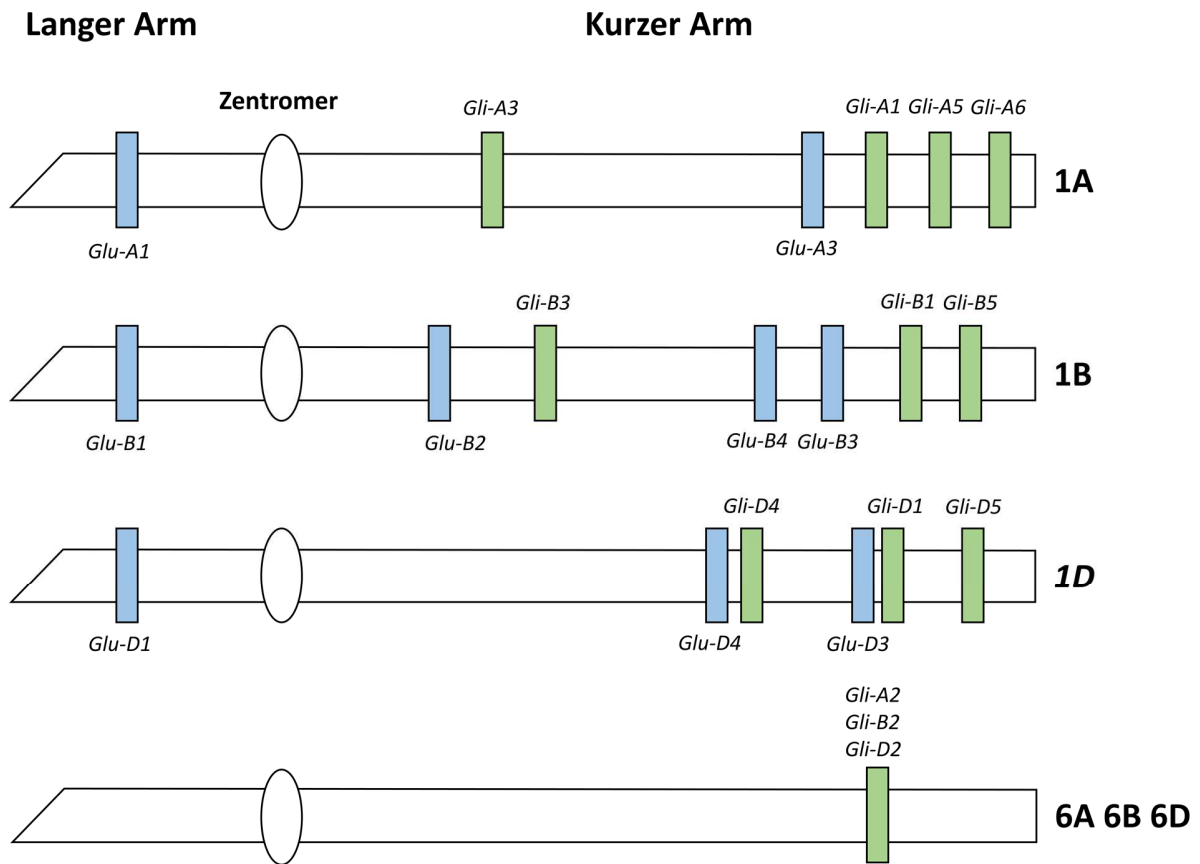
Ein fundiertes Verständnis der genetischen Organisation sowie eine einheitliche Nomenklatur sind entscheidend, um Gliadin- und Glutenin-Proteine ihren spezifischen Genorten zuzuordnen und deren Expression sowie funktionelle Rolle im Weizenkorn gezielt untersuchen zu können. Die genetische Organisation der Gliadine und Glutenine im Weizen wird primär durch Gene auf den Chromosomen 1 und 6 bestimmt. Die Gliadine werden hauptsächlich durch Genloci der Familien *Gli-1* und *Gli-2* kontrolliert. Die Expression der  $\alpha$ -Gliadine unterliegt der Kontrolle der *Gli-2* Loci (*Gli-A2*, *Gli-B2*, *Gli-D2*), die sich auf den

kurzen Armen der Chromosomen 6A, 6B, 6D befinden [26]. Die  $\omega$ - und  $\gamma$ -Gliadine werden durch Gruppen eng benachbarter Gene (Gencluster) an den *Gli-1* Loci (*Gli-A1*, *Gli-B1*, *Gli-D1*) reguliert, welche sich auf den kurzen Armen der Chromosomen 1A, 1B, 1D befinden. Zusätzlich wurden einige Minorgene innerhalb der *Gli-1* Region identifiziert, die ebenfalls zur Expression dieser Gliadinfraktionen beitragen. Einige  $\gamma$ -Gliadin-Gene wurden zudem auf den langen Armen der Chromosomen 3B und 3D nachgewiesen [27]. Schätzungen zufolge werden in einer Weizensorte rund 25 bis 35  $\alpha$ -Gliadin- und 10 bis 40  $\gamma$ -Gliadin-Gene exprimiert [28]. Eine Übersicht über die verschiedenen Glutenproteintypen und ihre jeweiligen Genorte zeigt

**Abbildung 2.**

Die Gene für LMW-GS an den *Glu-3* Loci (*Glu-A3*, *Glu-B3*, *Glu-D3*) sind auf den kurzen Armen der Chromosomen 1A, 1B, 1D lokalisiert. Es wurde beobachtet, dass es in diesen Regionen zu Rekombinationen zwischen Genen, die für  $\omega$ -Gliadine und bestimmte LMW-GS kodieren, kommen kann. Die Gene für weitere LMW-GS konnten den Chromosomengruppen 1 und 6 zugeordnet werden und liegen dicht gekoppelt an oder innerhalb der *Gli-1* bzw. *Gli-2* Loci [29]. Die LMW-GS stellen eine sehr heterogene Gruppe dar, und in einer Weizensorte sind typischerweise etwa 30 bis 40 LMW-GS-Gene vorhanden [30]. Basierend auf ihrer Mobilität in der SDS-PAGE können die LMW-GS weiter unterteilt werden in B-, C- und D-LMW-GS [31]. Die B-LMW-GS können wiederum in LMW-s, LMW-m und LMW-i Typen differenziert werden, je nachdem, ob ihre N-terminale Aminosäure Serin, Methionin oder Isoleucin ist [32]. Die LMW-s und LMW-m kommen in vergleichbaren Anteilen vor, während die LMW-i deutlich weniger vorhanden sind [33]. Bei den C- und D-LMW-GS handelt es sich um modifizierte  $\alpha$ -,  $\gamma$ - und  $\omega$ -Gliadine [34].

Die HMW-GS werden an komplexen Loci kodiert, die als *Glu-A1*, *Glu-B1*, *Glu-D1* bezeichnet werden und sich auf dem langen Arm der Chromosomengruppe 1A, 1B, 1D befinden. Jeder Locus enthält zwei eng gekoppelte Gene, die für den x- und den y-Typ codieren. Der *Glu-D1* Locus exprimiert typischerweise zwei HMW-GS, während der *Glu-B1* Locus eine oder zwei und der *Glu-A1* Locus eins oder keins (Null-Allel) exprimiert [29]. Folglich werden pro Weizensorte drei bis fünf HMW-GS exprimiert [35]. Um individuelle HMW-GS unterscheiden zu können, werden sie oft nach ihrem Genort (1A, 1B, 1D) und Typ (x- oder y-typ) benannt. Ergänzend wird ihre Mobilität in der SDS-PAGE häufig durch numerische Bezeichnungen angegeben; ein bekanntes Beispiel für eine HMW-GS-Allelkombination ist *Dx2+Dy10*. Dieses System muss jedoch für jedes verwendete Gel und je nach Puffer neu verifiziert werden. Außerdem muss beachtet werden, dass es bei manchen HMW-GS zu Unregelmäßigkeiten und Anomalien in ihrer Mobilität in der SDS-PAGE kommt [36].



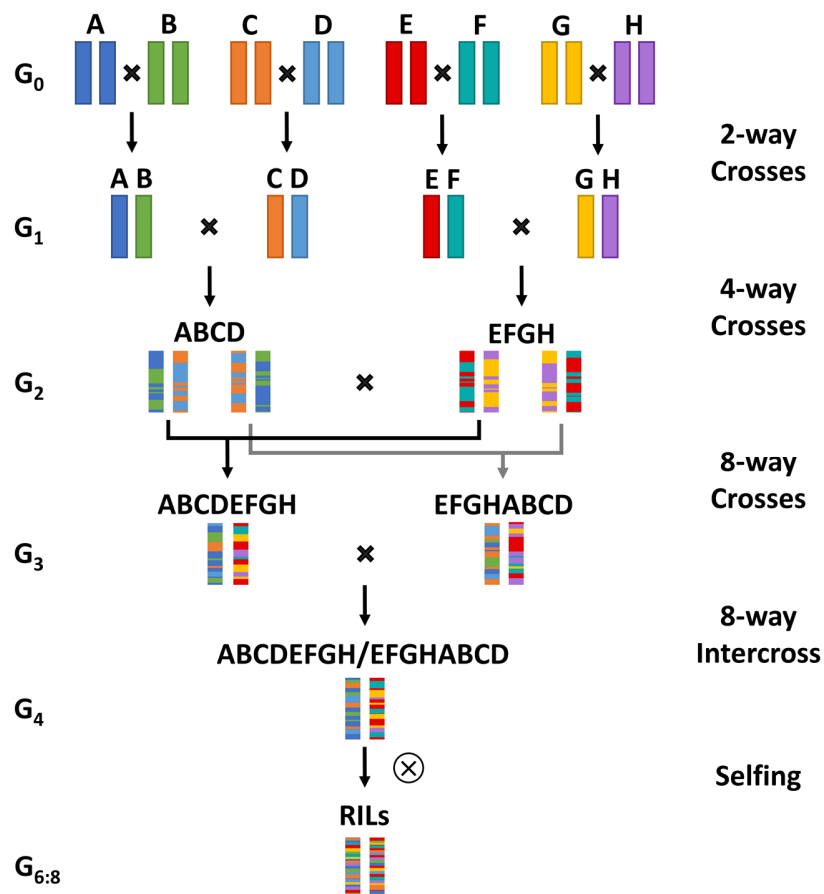
**Abbildung 2:** Lage der Genorte von Gliadinen und Gluteninen auf den Chromosomen, adaptiert und modifiziert nach Shewry et al. [29].

In den letzten zehn Jahren wurden viele neue genetische Informationen über Weizen gesammelt, darunter vollständige Genomdaten. Diese helfen dabei, Gene zu finden und zu untersuchen, die wichtige Eigenschaften, wie zum Beispiel die Qualität, beeinflussen. Deshalb gibt es Ansätze, die Regeln für die Benennung von Weizengenen systematisch entwickeln, um die Nomenklatur zu vereinheitlichen und den Austausch von Daten zu erleichtern [37].

#### 1.1.2.2 MAGIC-Populationen

Mithilfe von Multiparent Advanced Intercross (MAGIC) Populationen können komplexe Merkmale genetisch aufgeschlüsselt werden (**Abbildung 3**). Dazu werden mehrere (vier, acht oder mehr) Gründerlinien über viele Generationen nach einem festgelegten Kreuzungsschema („intercrossing“) miteinander gekreuzt. Die Nachkommen werden zudem untereinander gekreuzt und anschließend durch Selbstbefruchtung (Selfing) stabilisiert, um rekombinante Inzuchtlinien (recombinant inbred lines, RILs) zu erzeugen. Dies resultiert in einer Vielzahl genetisch differenzierter Genotypen, von denen jeder ein einzigartiges Mosaik der Gründerallele repräsentiert [38]. Die vielen neuen Kombinationen der Allele

helfen dabei, komplexe genetische Merkmale genetisch zu kartieren. Außerdem ist die Kartierung von Quantitative Trait Loci (QTL) möglich. Dabei handelt es sich um eine genetische Methode, die die Position und den Effekt von Genomregionen bestimmt, die quantitative Merkmale beeinflussen. Denn im Gegensatz zu qualitativen Merkmalen werden quantitative Merkmale (z. B. Ertrag) durch viele Gene (Loci) und Umwelteinflüsse bestimmt. Mit Weizen wurden bereits MAGIC Populationen mit vier, acht und 16 Elternlinien entwickelt. Stadlmeier et al. entwickelten die Bavarian MAGIC wheat population (BMWpop) aus acht Elternlinien. Sie umfasst 394 RILs, welche 71,7% der Allelvielfalt in der deutschen Weizenzüchtung repräsentieren. Außerdem konnten sie anhand der Kartierung von quantitativen Merkmalen von Sämlingsresistenz gegen echten Mehltau bereits nachweisen, dass die Population für QTL-Analysen geeignet ist [38]. Solche Ressourcen ermöglichen es, genetische Zusammenhänge und potenzielle Wechselwirkungen zwischen weiteren bedeutenden Merkmalen – etwa der qualitativen und quantitativen Proteinexpression – aufzuklären.



**Abbildung 3:** Kreuzungsschema der Bavarian MAGIC wheat population nach Stadlmeier et al. [38]. A: Event; B: Format; C: BAYP4535; D: Potenzial; E: Ambition; G: FIRL3565; H: Julius; G<sub>6:8</sub>: sechs Kreuzungsgenerationen, acht Selbstungsgenerationen. MAGIC, Multiparent Advanced Generation Intercross.

### 1.1.3 Backqualität

Die Backqualität zählt zu den wichtigsten Eigenschaften des Weizens. Dabei handelt es sich um eine komplexe Eigenschaft, die von vielen Faktoren abhängt (**Abbildung 4**). Die Weizensorte bzw. ihre genetische Ausstattung bildet hierfür die Basis. Ein grundlegender Faktor für das Wachstum der Weizenpflanze ist die Umwelt: Wetter, Klima, Wasserversorgung und Bodenbeschaffenheit haben einen entscheidenden Einfluss auf die Entwicklung der Pflanze. Eng damit verbunden sind Zeitpunkt und Art der Düngung, die darauf abzielen, der Pflanze zusätzliche Nährstoffe bereitzustellen [39].



**Abbildung 4:** Einflussfaktoren auf die Backqualität von Weizen, adaptiert und modifiziert nach Langenkämper & Zörb [39].

Diese äußeren Bedingungen beeinflussen maßgeblich die spätere qualitative und quantitative Zusammensetzung des Weizenkorns in Bezug auf Stärke, Fette und Proteine, aber auch Korngröße und -form. Letztere haben außerdem Einfluss auf die Vermahleigenschaften der Körner, womit noch ein weiterer Einflussfaktor hinzukommt, die Mahl- und Backtechnik. Für die Industrie sind unter anderem die Vermahlungseigenschaften der Weizenkörner, sowie die Verarbeitungseigenschaften des Teiges und die Teigstabilität wichtig. Bei den Endprodukten zählt vor allem, dass sie sensorisch ansprechend sind [40].

Der Stärkegehalt des Weizenkorns (70–75 %) stellt einen wichtigen Parameter für die Verarbeitung, Konsistenz und Gärung des Teiges dar. Der Lipidgehalt des Weizenkorns ist mit 2–3 % recht gering. Die Lipide teilen sich auf in Stärke-(30–40 %) und Nichtstärke-(60–70 %)lipide. Vor allem die Nichtstärkelipide haben einen Einfluss auf die Backqualität, was sich molekular mit ihrer Wirkung auf die Gaszellenstabilität

des Brotteiges erklären lässt [41]. Die Feuchtigkeit des Korns (frisch geerntet: 18–22 %, getrocknet: 12–14 %) ist wichtig für die Haltbarkeit und Verarbeitungseigenschaften der Weizenkörner.

Obwohl Stärke der Hauptbestandteil des Weizenkorns ist, haben der Gehalt und die Zusammensetzung des Proteins im Korn einen erheblichen Einfluss auf die Teigeigenschaften und sind der wichtigste Bestandteil zur Beurteilung der Backqualität. Gliadine und Glutenine bilden ein zusammenhängendes Glutennetzwerk, welches für die viskoelastischen Eigenschaften des Teiges sorgt [42]. Während der Teigfermentation wird das entstehende Gas durch dieses Netzwerk zurückgehalten, sodass beim Backen die charakteristische fluffige Brotkrume entsteht [43]. Bezogen auf Teig- und Backeigenschaften haben die Glutenproteintypen verschiedene Funktionen. Gliadine tragen hauptsächlich zu Teigviskosität und -dehnbarkeit bei, während die Glutenine Teigstärke und -elastizität bestimmen [16]. In der Folge bestimmt auch das Verhältnis von Gliadinen zu Gluteninen wesentlich die Teigeigenschaften. Auf molekularer Ebene kann die Funktionalität damit erklärt werden, dass die Gluteninpolymere durch Disulfidbrücken zwischen Cysteinresten stabilisiert werden (**Abbildung 5**). Proteomanalysen zeigen, dass die meisten dieser Bindungen HMW-GS betreffen, z. B. Verknüpfungen zwischen  $\alpha$ - und  $\gamma$ -Typ-Untereinheiten. Bindungen mit LMW-GS wurden ebenfalls nachgewiesen [13]. HMW-GS gelten daher als strukturelles „Rückgrat“ des Glutens, welches ein Gerüst für weitere Interaktionen mit LMW-GS, Gliadinen und anderen Teigkomponenten bildet. Daher beeinflussen die Qualität und Quantität der Glutenproteine maßgeblich die Teigeigenschaften und somit die Backqualität.



**Abbildung 5:** Schematische Darstellung der Disulfidstruktur von Weizenglutenproteinen anhand eines Beispiels, adaptiert und modifiziert nach Wieser et al. [13]. Braune Segmente repräsentieren die N- und C-terminalen Enden; Disulfidbrücken sind durch schwarze Linien zwischen den Gluteninen dargestellt; HMW-GS hochmolekulare Gluteninuntereinheit; LMW-GS, niedermolekulare Gluteninuntereinheit.

### 1.1.3.1 Bestimmung der Backqualität

Die Beurteilung der Backqualität von Weizenmehl erfolgt am zuverlässigsten mittels standardisierter Backversuche, bei denen das Mehl den vollständigen Produktionsprozess durchläuft – von der Teigbereitung über die Gärung bis zur Lagerung des Endprodukts. Zu diesen Versuchen zählen z. B. der standardisierte Weizen-Backversuch [44] oder das Rapid-Mix-Testverfahren [45]. Dabei ist das in Backversuchen bestimmte Brotvolumen die wichtigste Messgröße für die Beurteilung der Backqualität. In

wissenschaftlichen Studien wird das Brotvolumen unter anderem als Referenzgröße herangezogen, um die Korrelation mit anderen Backqualitätsparametern zu untersuchen und deren Eignung zur Beurteilung der Backqualität zu evaluieren [46, 47].

Aufgrund des hohen Zeit- und Ressourcenaufwands sind Backversuche für Routineanalysen nur eingeschränkt geeignet. Deshalb wurden andere, indirekte Qualitätsparameter definiert, um die Backqualität einfacher und schneller vorherzusagen. Dazu zählen z. B. Rohproteingehalt, Feuchtklebergehalt, Glutenindex, Sedimentationswert, Fallzahl, Kornhärte, oder Wasserabsorption. Als indirekter Indikator für die Backqualität wird häufig der Rohproteingehalt verwendet, da es sich um eine einfach zu bestimmende Größe handelt und der Proteingehalt bereits sehr früh mit einer höheren Backqualität in Zusammenhang gebracht wurde [48]. Üblicherweise wird der Proteingehalt indirekt über die Bestimmung des Stickstoffgehalts ermittelt; beispielsweise mittels der Dumas-Methode [49]. Anschließend erfolgt die Umrechnung in den Proteingehalt durch Multiplikation mit dem Faktor 5,7, der sich aus dem durchschnittlichen Stickstoffanteil in Weizenproteinen von etwa 17,5 % ergibt. Der Feuchtklebergehalt ist die Menge an Gluten, die aus Weizenmehl extrahiert werden kann, wobei die Stärke mit Wasser ausgewaschen wurde. Nach diesem Vorgang wird der Feuchtkleber getrocknet und gewogen [50]. Zur Bestimmung des Glutenindex wird das Gluten aus Weizenmehl zentrifugiert, um den Feuchtkleber unter genormten Bedingungen durch ein speziell entwickeltes Sieb zu pressen. Der Anteil an Feuchtkleber, der nach der Zentrifugation auf dem Sieb verbleibt, wird als Glutenindex definiert. Bei sehr schwachem Gluten kann der gesamte Feuchtkleber durch das Sieb gehen, was zu einem Glutenindex von 0 führt. Wenn kein Gluten durch das Sieb geht, beträgt der Index 100 [51]. Eine weitere Möglichkeit zur Bewertung der Glutenqualität und -stärke ist der Sedimentationswert (auch Zeleny-Sedimentationstest). Der Test basiert auf der Fähigkeit der Glutenproteine, in saurer Lösung Wasser zu absorbieren und zu quellen [52]. Die Fallzahl liefert einen indirekten Hinweis auf die  $\alpha$ -Amylase-Aktivität, indem sie die Viskosität einer erhitzten Mehl-Wasser-Suspension über die Fallzeit eines standardisierten Rührstabs bestimmt [53]. Die Kornhärte bezieht sich auf die Festigkeit der Weizenkörner, welche sich auf die Mahl- und Teigeigenschaften auswirkt. Zur Ermittlung der Kornhärte wird das Bruchverhalten der Weizenkörner unter standardisierten Bedingungen bestimmt. Das Mehl-Wasser-Verhältnis spielt eine entscheidende Rolle für das Ergebnis vieler Brot-, Gebäck- und Keksrezepte. Um einen guten Geschmack, eine gute Textur und eine gute Ausbeute bei Backwaren zu erzielen, ist es wichtig, einen standardisierten, verarbeitungsfähigen Teig mit einer definierten Konsistenz herzustellen. Dazu wird die optimale Wassermenge für den Teig ermittelt [54].

Manche der genannten Parameter bewerten zwar die Glutenqualität, allerdings nicht anhand der inhaltlichen Zusammensetzung, sondern aufgrund von indirekten, technologischen Eigenschaften des Mehls bzw. des Teiges. Da die Glutenproteine wesentlich die Teigeigenschaften bestimmen, die Glutenproteintypen aber auch verschiedene Funktionen im Teigsystem einnehmen, werden diese gezielter im Hinblick auf ihre Bedeutung für die Backqualität untersucht.

#### 1.1.3.2 Weizenproteine in Zusammenhang mit der Backqualität

Die Forschung zu der Bedeutung von Proteinen für die Backqualität von Weizen befasst sich mit unterschiedlichen Aspekten. Uthayakumaran & Wrigley untersuchten unter anderem die Gesamtmenge an Protein, die Menge an Gliadin und Glutenin, sowie deren Verhältnis untereinander, die Ausbildung von Disulfidbrücken, vor allem innerhalb der Glutenine, sowie die extrahierbare Menge des SDS-unlöslichen Glutenins, dem sogenannten Glutenin-Makropolymer. Darüber hinaus wurde der Einfluss verschiedener Glutenine und Gliadine in unterschiedlichen Weizensorten analysiert. Neben Glutenproteinen spielen Puroindoline zudem eine zentrale Rolle bei der Bestimmung der Kornhärte und beeinflussen maßgeblich die Verarbeitungseigenschaften [55].

Der Proteingehalt lässt sich zuverlässig und unkompliziert bestimmen und kann bereits im Stadium des unverarbeiteten Weizenkorns (z. B. mit Nahinfrarotspektroskopie) analysiert werden. Finney et al. entdeckten bereits 1948 einen Zusammenhang zwischen Proteingehalt und Brotvolumen, den sie als linear postulierten [48]. Auch 60 Jahre später wurde der Proteingehalt als wichtigster Parameter für die Vorhersage des Brotvolumens und damit der Backqualität festgestellt [56]. Allerdings wurden auch Nachteile bei der Verwendung des Proteingehalts als einzigem Indikator bemerkt. So wurden stark schwankende Korrelationskoeffizienten zwischen 0,01 und 0,86 bei der Korrelation von Proteingehalt und Brotvolumen erhalten, was auf Wetter- und Umweltbedingungen, Düngung sowie Weizensorte zurückgeführt wurde [56]. Die lineare Beziehung zwischen Proteingehalt und Backqualität wurde zudem kritisch hinterfragt. Es wurde vermutet, dass ab einem bestimmten Proteingehalt ein Sättigungseffekt eintritt, wodurch eine weitere Verbesserung der Backqualität ausbleibt [57]. Gabriel et al. wiesen nach, dass sich bei gleichem Proteingehalt die Backqualität zwischen verschiedenen Sorten unterscheiden kann [47]. Zum einen stellt der Proteingehalt als alleiniger Indikator damit nicht die beste Möglichkeit dar, Unterschiede in Mehl- bzw. des Teigeigenschaften zu erklären oder die Backqualität zuverlässig vorherzusagen [47, 58]. Zum anderen ergibt sich daraus, dass eine hohe Backqualität nicht allein durch den Proteingehalt bestimmt wird, sondern wesentlich von der qualitativen Zusammensetzung der Proteine abhängt.

Der Einfluss einzelner Weizenproteinfraktionen (Albumine/Globuline, Gliadine, Glutenine, unlöslicher Rest) auf die Backqualität wurde bereits in den Siebzigern untersucht (Orth, 1972). Damals konnte ein positiver Einfluss der Albumine und des unlöslichen Rests festgestellt werden, während die Glutenine negativ mit Brotvolumen korreliert wurden. Thanhaeuser et al. untersuchten den Einfluss der Menge an Gliadinen und Gluteninen, sowie des Glutenin-Makropolymers auf die Backqualität [46]. Es zeigte sich, dass Gliadin- und Gluteningehalte eine stärkere Korrelation mit dem Brotvolumen aufwiesen als der Rohproteingehalt und somit als verlässlichere Prädiktoren für die Backqualität gelten könnten. In einer weiteren Studie wurde der Einfluss von Proteingehalt und –zusammensetzung von 82 Weizenlinien auf eine Reihe rheologischer und Backqualitätsparameter untersucht. Unter anderem wurden die Gehalte der einzelnen Glutenproteintypen mit dem Brotvolumen korreliert [59]. Es stellte sich heraus, dass keine signifikante Beziehung zwischen einzelnen Glutenproteintypen und dem Brotvolumen bestand. Im gesamten Datensatz konnte kein anderer Parameter ermittelt werden, der einen messbaren Einfluss auf das Brotvolumen hatte. Dies und die widersprüchlichen Ergebnisse vorhergehender Arbeiten weisen darauf hin, dass es sich bei der Backqualität um eine komplexe Eigenschaft handelt, welche sich nicht durch einen einzelnen oder wenige Parameter zuverlässig erfassen lässt, was auch von Schuster et al. geschlussfolgert wurde [59]. Andererseits ergeben sich hierdurch immer noch keine Hinweise darauf, ob individuelle Glutenproteine einen Einfluss haben.

Im Jahr 1979 isolierten Payne et al. einzelne HMW-GS aus verschiedenen Weizenlinien und identifizierten eine spezifische Untereinheit, deren Anwesenheit stark mit einer verbesserten Backqualität korrelierte [60]. Zusammen mit der Aufklärung der essentiellen funktionellen Eigenschaften der HMW-GS im Teig bildete dies die Basis für die Klassifizierung von HMW-GS-Genorten bzw. Allelen und deren individueller Bedeutung für die Backeigenschaften von Weizenmehl. Es wurde z. B. festgestellt, dass die HMW-GS-Kombination Dx5+Dy10 mit einer höheren Backqualität assoziiert ist, während Dx2+Dy12 mit einer geringeren Backqualität in Verbindung gebracht wird [26]. Generell wurde beobachtet, dass die Anreicherung von Teig mit HMW-GS des x-Typs einen größeren Einfluss auf Teigparameter ausübte als die Zugabe von y-Typ HMW-GS [61]. Allgemein gibt es eine Reihe von Studien, die unterschiedliche Weizenmaterialien und Ansätze zur Bewertung der Auswirkungen von Gluteninen auf Qualitätsmerkmale verwendet. Meistens untersuchen sie jedoch nur einen kleinen Datensatz oder nur bestimmte Weizensorten bestimmter Regionen. Guzmán et al. fassten den Effekt von HMW-GS und LMW-GS auf Backqualität anhand von 2550 Weichweizenlinien zusammen. Sie stellten fest, dass die HMW-GS 1, 2\*, 7+8, 17+18, 13+16, 5+10 sowie die LMW-GS-Allele *Glu-A3*, *Glu-A3d*, *Glu-A3f*, *Glu-B3c* und *Glu-B3d* mit höherem Brotvolumen assoziiert sind. Im Gegensatz dazu waren die HMW-GS Null, 7, 6+8, 2+12 und die

LMW-GS-Allele *Glu-A3e* und *Glu-B3j* mit einem schlechten Backqualitätsprofil verbunden [62]. Um den Einfluss einzelner Glutenproteine beurteilen zu können, ist zunächst deren qualitative und quantitative Erfassung in der Probe erforderlich. Aufgrund der komplexen Zusammensetzung der Proteine in organischen Proben stellt die Identifizierung mittels LC-MS/MS eine besonders geeignete Methode dar.

## 1.2 Proteomics

Bei Proteomics wird die Gesamtheit der Proteine, die in einer Zelle oder einem Organismus zu einem bestimmten Zeitpunkt (unter definierten Bedingungen) vorhanden sind, qualitativ und quantitativ analysiert [63]. Analog zum Genom kann so ein Proteom erstellt werden, welches im Gegensatz zum Genom allerdings hochdynamisch ist, da es unter anderem abhängig von Zustand und Umwelt ist. Typische Analysemethoden zur Untersuchung des Proteoms sind die 2D-Gelelektrophorese und die Flüssigkeitschromatographie gekoppelt mit Tandem-Massenspektrometrie (LC-MS/MS).

Prinzipiell gibt es zwei verschiedene Varianten: Top-Down Proteomics und Bottom-Up Proteomics (auch Shotgun Proteomics oder Discovery Proteomics). Bei Top-Down Proteomics werden intakte Proteine analysiert, während bei Bottom-Up Proteomics die Proteine durch Enzyme in Peptide gespalten und diese dann identifiziert werden [64]. Grundsätzlich ermöglicht Top-Down Proteomics eine präzisere Identifizierung von Proteinen, da sie auf einzigartigen Fragmenten basiert, aus denen zusätzlich noch detailliertere Informationen über Proteinstrukturen gewonnen werden können. Allerdings stellt dies hohe Anforderungen an die Massenspektrometrie (MS). Zudem lassen sich komplexe Proteingemische mit mehr als zehn Komponenten nur schwer auflösen. Insbesondere bei stark schwankenden Proteinmengen und großen Proteinen (MW größer als 50.000), die aufgrund einer ineffizienten Ionisation und Fragmentierung nur eine geringe Anzahl aussagekräftiger Fragmente erzeugen, ist die Analyse erschwert [65]. Die Bottom-Up Strategie ermöglicht eine effiziente Analyse komplexer Proteingemische mit hoher Empfindlichkeit und hohem Durchsatz. Außerdem ist sie technisch robuster, kostengünstiger und besser skalierbar. Durch die mittlerweile breite Anwendung profitiert sie außerdem von je nach Organismus mehr oder weniger ausgereiften Datenbanken, standardisierten Protokollen und einer breiten Kompatibilität mit etablierten analytischen Plattformen [66].

Zur akkuraten Identifikation der Peptide sind hochauflösende Massenanalytoren wie Orbitrap-Systeme notwendig. Der Orbitrap-Massenanalysator besteht aus drei speziell geformten Elektroden: zwei becherförmigen äußeren Elektroden und einer zentralen, spindelförmigen Mittelelektrode. Diese Anordnung erzeugt ein elektrisches Feld, das Ionen in einer stabilen Bahn hält. Die Ionen bewegen sich dabei harmonisch entlang der Achse der Mittelelektrode. Diese Schwingungen erzeugen einen Bildstrom

an den äußeren Elektroden. Das Signal wird verstärkt, digitalisiert und per Fourier-Transformation in ein Massenspektrum umgewandelt [67].

### 1.2.1 Bottom-Up Proteomics

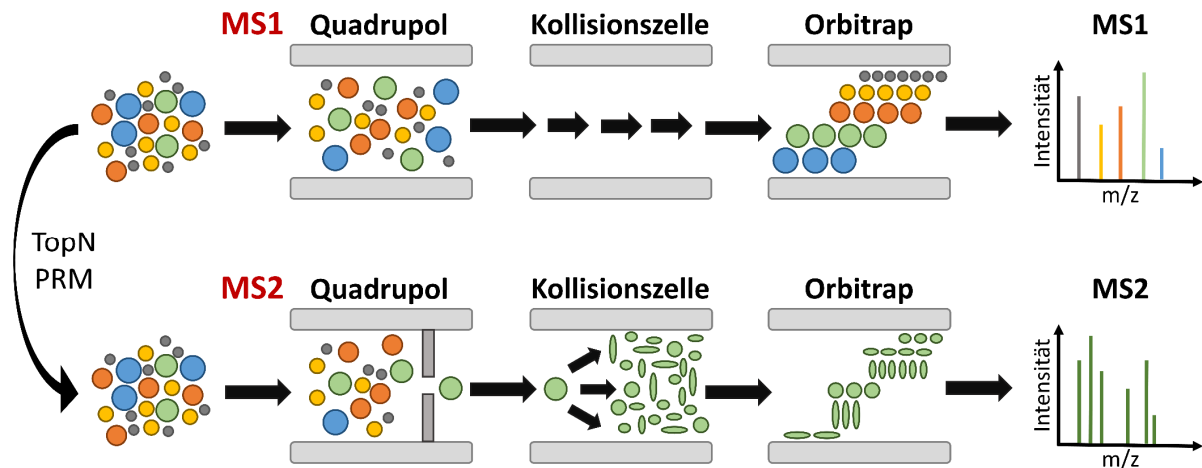
Die Durchführung eines Bottom-Up Proteomics-Experiments umfasst mehrere Schritte (siehe auch **Abbildung 7**). Zunächst müssen die Proteine aus der Probe isoliert werden. Danach kann eine Quantifizierung der Proteinkonzentrationen (z. B. Bradford Assay) erfolgen. Ebenfalls möglich ist eine vorangestellte Fraktionierung der Proteine etwa mittels SDS-PAGE. Im nächsten Schritt und zur Vorbereitung auf den enzymatischen Verdau werden die Proteine oftmals reduziert und alkyliert. Dabei werden die Disulfidbrücken zwischen Cysteinresten gespalten (z. B. mit DTT). Die entstehenden freien Sulfhydrylgruppen werden anschließend alkyliert (z. B. mit Chloracetamid, CAA), um so eine erneute Bildung von Disulfidbrücken zu verhindern. Dieser Schritt fördert die Proteindenaturierung und führt dadurch zu einer besseren Zugänglichkeit für proteolytische Enzyme. Der nächste essentielle Schritt ist die proteolytische Spaltung der Proteine durch Enzyme (i. d. R. Trypsin). Je nachdem sind noch weitere Aufreinigungs- bzw. Aufkonzentrierungsschritte notwendig bevor das Peptidgemisch massenspektrometrisch analysiert werden kann [66].

Je nach Komplexität des Peptidgemischs erfolgt zuerst eine Auftrennung mittels Hochleistungsflüssigkeitschromatographie (HPLC). Für die Messung mittels MS müssen die Peptide ionisiert werden. Dies kann z. B. durch Elektrospray-Ionisation (ESI) erfolgen [68]. Danach erfolgt die Erfassung der Peptidionen durch das Massenspektrometer. Die Analyse kann entweder untargeted erfolgen, wobei sämtliche Peptide/Proteine einer Probe umfassend untersucht werden, oder targeted, bei der gezielt bestimmte Peptide/Proteine erfasst werden.

#### 1.2.1.1 Untargeted Proteomics

Bei untargeted Proteomics werden sämtliche detektierbaren Peptide analysiert. Außerdem sind die mittels MS gemessenen Peptidsignale proportional zu den Peptidkonzentrationen. Somit können diese sowohl qualitativ als auch quantitativ erfasst werden. Dabei gibt es drei MS-Datenerfassungsmodi: Full-Scan, datenabhängige Erfassung (Data Dependent Acquisition, DDA) und datenunabhängige Erfassung (Data Independent Acquisition, DIA). Bei dem Full Scan (MS1) wird das Massenspektrum aller ionisierten Moleküle innerhalb eines vorgegebenen  $m/z$  Fensters (z. B. 200–1000  $m/z$ ) aufgenommen. Diese Methode liefert einen ersten Überblick über alle detektierten Ionen in der Probe und bildet die Grundlage für DDA. Bei DDA wird anhand des Full Scans eine vordefinierte Anzahl (TopN) der häufigsten Peptide gewählt

(**Abbildung 6**). Diese werden in der Kollisionszelle fragmentiert und anschließend analysiert. Anhand der Fragmente können die Peptide exakt identifiziert werden [69]. Bei DIA werden alle Peptide/Ionen innerhalb eines bestimmten  $m/z$  Fensters gesammelt, fragmentiert und anschließend per MS/MS identifiziert. Es werden dadurch komplexere Massenspektren als bei DDA erhalten, was die spätere Auswertung aufwendiger macht.



**Abbildung 6:** Schematische Darstellung des DDA und PRM Messprinzips, adaptiert und modifiziert nach Ludwig & Aebersold [70]. DDA, Data Dependent Acquisition; PRM, Parallel Reaction Monitoring.

Unabhängig von der Erfassungsmethode ist im Anschluss spezialisierte Software für die Auswertung der generierten Rohdaten notwendig, z. B. MaxQuant [71]. Diese nutzen Suchmaschinen (Andromeda bei MaxQuant) um die experimentellen Daten mit Daten aus einer Proteindatenbank (z. B. UniProtKB) zu vergleichen. Für DIA-Daten ist zusätzlich oft noch eine spektrale Bibliothek erforderlich. Für eine erfolgreiche Proteinidentifizierung ist es weiterhin notwendig, dass die entsprechende Proteinsequenz in der Datenbank hinterlegt ist. Falls dies nicht der Fall ist, werden die Peptide dem am stärksten homologen Eintrag zugeordnet; häufig einem Protein aus einer nah verwandten Art. Kann ein Peptid oder Protein nicht identifiziert werden, erfolgt eine *de novo* Sequenzierung, bei der die Aminosäuresequenz direkt aus den Massenspektren abgeleitet wird. Die so gewonnenen Sequenzdaten können anschließend manuell mit bestehenden Datenbankeinträgen abgeglichen werden.

#### 1.2.1.2 Targeted Proteomics

Targeted Proteomics wird dann angewendet, wenn gezielt bestimmte Peptide analysiert werden sollen. Als besonders geeignete MS-Messmethode hat sich das Selected Reaction Monitoring (SRM, auch bekannt als Multiple Reaction Monitoring, MRM) etabliert, da es eine präzise, akkurate, empfindliche sowie

reproduzierbare Detektion der Analyten ermöglicht. Eine Weiterentwicklung dieser Technik ist das Parallel Reaction Monitoring (PRM), bei dem ein Orbitrapanalysator mit einem vorgeschalteten Quadrupol genutzt wird (**Abbildung 6**). Das Messprinzip beider Methoden basiert darauf, dass im ersten Schritt das Präkursorion (des interessierenden Peptids) durch den ersten Quadrupol selektiert wird. Anschließend erfolgt dessen Fragmentierung in einer Kollisionszelle. Die entstehenden Produkt-Ionen werden durch den zweiten Masseanalysator detektiert. Beim Einsatz eines Orbitrap-Analysators können alle Produkt-Ionen parallel und mit hoher Auflösung erfasst werden. Anschließend wird daraus ein MS2-Spektrum generiert [70].

## 1.2.2 Quantitative Proteomics

### 1.2.2.1 Relative Proteinquantifizierung

Grundsätzlich wird zwischen markierungsbasierten (label-based) und labelfreien (label-free) Quantifizierungsansätzen unterschieden. Die labelfreie Quantifizierung (label-free quantification, LFQ) erfordert keine zusätzlichen Schritte bei der Proteinaufarbeitung, sondern die Proben werden einzeln gemessen und jeweils separate Rohdateien erzeugt. Softwareprogramme berechnen die LFQ, indem die Intensität eines Peptids über die Fläche des extrahierten Ionenchromatogramms ermittelt wird. Bei DDA werden dazu die Signale der Präkursorionen verwendet [72]. Die relative Quantifizierung bestimmt somit Unterschiede in der Proteinmenge zwischen zwei oder mehr Proben, d. h. es findet ein Vergleich der relativen Häufigkeit statt. Obwohl sich damit keine absoluten Proteinmengen ermitteln lassen, ermöglicht die Methode dennoch die Erkennung von Proteinanreicherungen oder -abnahmen bei verschiedenen Behandlungen, Zeitpunkten oder experimentellen Bedingungen. Dadurch, dass keine Markierung notwendig ist, handelt es sich um eine kostengünstige Methode, die für alle Probenarten geeignet ist. Dafür weist sie eine geringere Genauigkeit auf als markierungsbasierte Methoden wie z. B. die Tandem Mass Tags. Bei dieser Methode werden die Analyten mit isobaren Markierungen versehen, die im MS1-Spektrum nicht unterscheidbar sind. Erst nach Fragmentierung im MS2 entstehen unterschiedliche Produkt-Ionen, sogenannte Reporterionen, die eine quantitative Unterscheidung der Proben ermöglichen. Im Gegensatz dazu werden bei der Isotope-Coded Affinity Tag-Methode cysteinhaltige Peptidsequenzen markiert, wodurch diese bereits im MS1-Spektrum von unmarkierten Peptiden differenzierbar sind und eine relative Quantifizierung erlauben [73].

### 1.2.2.2 Absolute Proteinquantifizierung

Bei Bottom-Up Proteomics wird das Protein von Interesse indirekt durch ein oder mehrere Peptide identifiziert, die bei dem Verdau des Proteins entstehen. Die absolute Quantifizierung dieser Peptide ist mit der Stabilisotopenverdünnungsanalyse (Stable Isotope Dilution Assay, SIDA) möglich. Dabei wird die Probe mit einer definierten Menge stabilisotopenmarkierter Peptidstandards als Referenz "verdünnt". Die Peptidstandards sind chemisch identisch mit den Zielpeptiden in der Probe, wobei bestimmte Aminosäuren mit schweren Isotopen markiert wurden. Dabei handelt es sich in der Regel um  $^{13}\text{C}$  oder  $^{15}\text{N}$ . Folglich weisen die markierten Peptide dieselben physikochemischen Eigenschaften wie die Analyten auf, unterscheiden sich aber in der molekularen Masse, sodass sie durch MS unterscheidbar sind. Nach der PRM-Messung kann anhand des Vergleichs der Signale von Analyt und Standard absolut quantifiziert werden, da die Konzentration des internen Standards bekannt ist [70]. Anschließend lässt sich aus dem Peptidgehalt auf die Menge des jeweiligen Proteins rückschließen.

Ein Vorteil dieser Methode besteht darin, dass die stabilisotopenmarkierten Peptide denselben Bedingungen unterliegen wie die Probenpeptide, da sie als interner Standard möglichst früh bei der Probenaufarbeitung zugesetzt werden. Folglich werden Analytverluste bei der Aufarbeitung bereits berücksichtigt. Auch während der LC-MS/MS-Analyse verhalten sich Standard und Analyt nahezu identisch: Sie werden gleichzeitig aus dem LC-System eluiert, werden anschließend unter denselben Bedingungen ionisiert und sind dabei denselben Matrix- oder Hintergrundeffekten ausgesetzt. Auf diese Weise können sie Störungen während der Aufarbeitung oder Messung kompensieren, was die Präzision und Reproduzierbarkeit der Ergebnisse deutlich verbessert [70].

### 1.2.3 Proteomics von Weizenproteinen

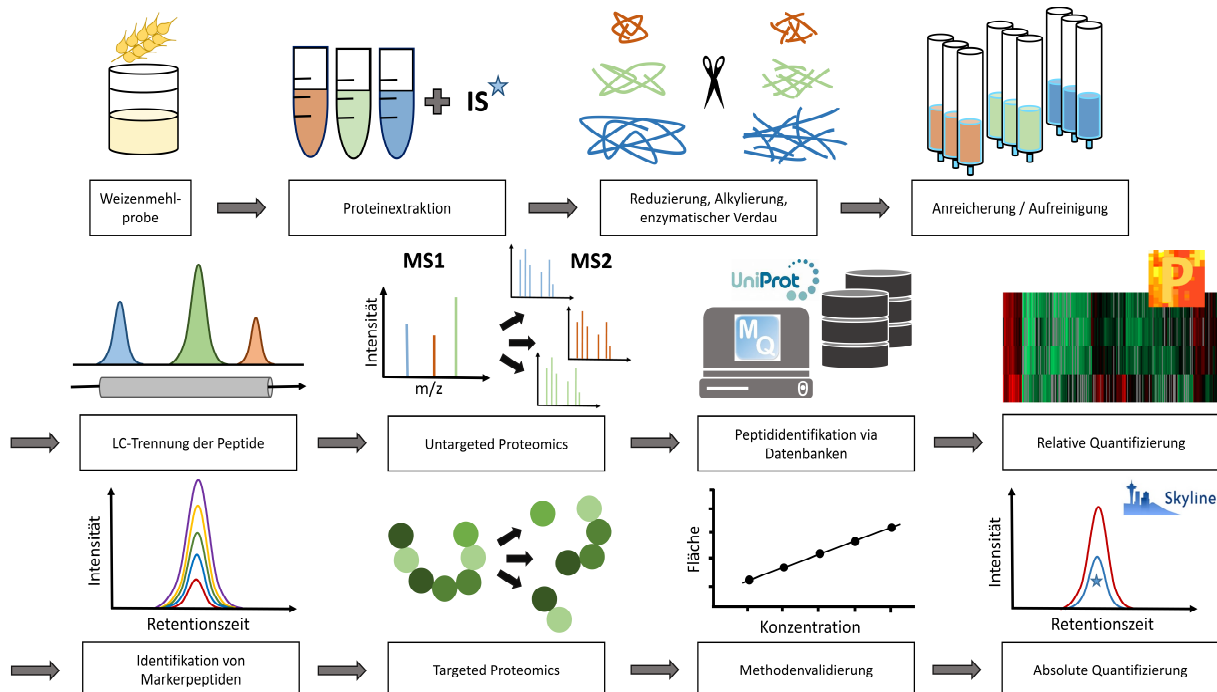
Aufgrund der hohen Selektivität, Empfindlichkeit und Vielseitigkeit sind LC-MS/MS-Methoden ein vielversprechender Ansatz für den Nachweis und die Quantifizierung von Weizenproteinen [74]. Mehrere Studien in der Literatur widmen sich der Charakterisierung des Weizenproteoms und dokumentieren den Nachweis von individuellen Glutenproteinen. Dupont et al. waren die ersten, die erfolgreich Weizenmehlproteine spezifischen Gensequenzen zuordneten und relativ quantifizierten. Sie identifizierten insgesamt 157 Weizenproteinsequenzen, darunter 20,4 %  $\alpha$ -Gliadine, 17,1 %  $\gamma$ -Gliadine, 12,2 %  $\omega$ -Gliadine, 18,0 % LMW-GS und 17,1 % HMW-GS in der Weizensorte Butte 86 mittels SDS-PAGE und anschließender MS-Messung [75]. Bromilow et al. identifizierten mit einer Kombination aus DIA und DDA insgesamt 26 Gliadine (4  $\omega$ -, 14  $\alpha$ - und 8  $\gamma$ -Gliadine) und 37 Glutenine (29 LMW-GS und 8 HMW-GS) in der Weizensorte Hereward [76]. Mittels DIA wurden 19 Glutenine und 23 Gliadine in zehn Weizensorten

identifiziert. Dabei lag der Fokus auf genetisch stabilen Proteinen, die in mehr als 50 % der Sorten sowie in mindestens zwei von drei Umwelten in über 80 % der Sorten nachweisbar waren [77].

**Abbildung 7** veranschaulicht den Ablauf einer Probenaufbereitung von Weizenmehl mit anschließender Proteomics-Analyse. Der erste Schritt ist die Extraktion der Proteine. Bei Weizen oftmals angewendet wird die Osborne-Fraktionierung, bei der die Proteinfractionen der Albumine/Globuline, Gliadine und Glutenine mit drei verschiedenen Extraktionsmitteln getrennt gewonnen werden. Die Fraktion der Albumine/Globuline wird im ersten Schritt mit phosphatgepufferter Kochsalzlösung extrahiert. Danach erfolgt die Extraktion der Gliadine mit wässrigem Ethanol. Die Glutenine werden mit einem gepufferten Gemisch aus 1-Propanol, DTT und Urea bei 60 °C extrahiert. DTT dient als Reduktionsmittel, um die intra- und intermolekularen Disulfidbindungen der Gluteninproteine aufzubrechen. Urea besitzt disaggregierende Eigenschaften. Es wirkt als chaotropes Agens, in dem es die Wasserstoffbrückenbindungen zwischen Proteinen aufbricht und dadurch deren Denaturierung und Auffaltung erleichtert [19]. Im nächsten Schritt werden die Proteine reduziert und alkyliert, um die Disulfidbindungen dauerhaft zu spalten und eine erneute Ausbildung von Disulfidbrücken zu verhindern. Als Reduktionsmittel wird Tris(2-carboxyethyl)phosphin (TCEP) verwendet. Im Gegensatz zu DTT ist es oxidationsstabil und wirkt über einen breiteren pH-Bereich, wodurch eine stabile und vollständige Reduktion erfolgt. Zum Alkylieren wird z. B. CAA verwendet, welches die freien Thiolgruppen der Cysteine kovalent blockiert. So wird das erneute Ausbilden von Disulfidbindungen dauerhaft verhindert. Im Anschluss erfolgt die hydrolytische Spaltung der Proteine durch Peptidasen. Das Enzym Trypsin wird dazu besonders häufig in Proteomics verwendet, da es sich um eine sehr effiziente und spezifische Peptidase handelt [78]. Den Abschluss der Probenaufarbeitung bildet eine Festphasenextraktion (Solid Phase Extraction, SPE), die der Entfernung von Salzen sowie der Aufkonzentrierung der Probe dient.

Für die LC-MS/MS-Messung werden die Proben gelöst und mittels HPLC aufgetrennt. Nach der Ionisation erfolgt die untargeted Proteomics-Analyse. Die anschließende Zuordnung von Proteinen und Ergebnisauswertung erfolgt mithilfe von Datenbanken wie UniProtKB und Auswertesoftware wie MaxQuant und Perseus. MaxQuant ist eine frei verfügbare Softwareplattform zur automatisierten Verarbeitung und quantitativen Auswertung (LFQ) von massenspektrometrie-basierten Proteomics-Daten. Die Identifikation von Peptiden und Proteinen erfolgt dabei mithilfe der integrierten Suchmaschine Andromeda. Perseus ist eine Begleitsoftware zu MaxQuant für die bioinformatische Nachbearbeitung und statistische Analyse von quantitativen Proteomics-Daten. Neben grundlegender Datenfilterung und Normalisierung ermöglicht Perseus auch die Durchführung von Korrelationsanalysen, Clusterbildungen

sowie die visuelle Aufbereitung komplexer Datensätze, wodurch biologische Zusammenhänge und relevante Proteinmuster effizient identifiziert werden können [71, 79].



**Abbildung 7:** Schematische Darstellung der Aufarbeitung von Weizenmehlproben für untargeted und targeted Proteomics. IS, interner Standard; LC, Flüssigkeitschromatographie.

Im Falle der targeted Proteomics-Messung müssen zunächst geeignete Markerpeptide identifiziert, verifiziert und validiert werden. Ein zentrales Kriterium bei der Auswahl eines Markerpeptids ist seine eindeutige Zuordenbarkeit, das heißt, es muss spezifisch einem einzelnen Protein bzw. einer Proteingruppe zugeordnet werden können. Zudem sind gute massenspektrometrische Detektionseigenschaften erforderlich, damit das Peptid zuverlässig, reproduzierbar und in quantifizierbarer Weise nachgewiesen werden kann. Die stabilisotopenmarkierten Markerpeptide werden als interner Standardmix direkt nach der Proteinextraktion zu jeder Probe gegeben (**Abbildung 7**). Nach der restlichen Probenaufarbeitung und LC-Trennung werden die Proben mittels targeted Proteomics gemessen. Die absolute Quantifizierung erfolgt anhand des Verhältnisses von Standard- zu Analytpeptid in Kombination mit einer externen Kalibrierkurve (Responsegerade). Die Auswertung erfolgt mittels Software wie Skyline. Skyline ist eine frei verfügbare, plattformübergreifende Software zur Entwicklung, Visualisierung und quantitativen Auswertung von targeted Proteomics-Daten, insbesondere aus SRM/MRM, PRM, DIA und DDA-Experimenten. Sie ermöglicht die gezielte Auswahl von Peptiden und Transitionen, die Integration von MS-Signalen sowie statistische und visuelle Qualitätskontrollen [80].

## 2 Zielsetzung

Weichweizen (*Triticum aestivum* L.) gehört zu den bedeutendsten Kulturpflanzen weltweit und dient als zentrales Grundnahrungsmittel für einen großen Teil der Weltbevölkerung. Seine besondere Eignung für die Herstellung backfähiger Teige und Brotprodukte beruht maßgeblich auf der einzigartigen Zusammensetzung seiner Glutenproteine, die für die viskoelastischen Eigenschaften des Teigs verantwortlich sind. Der fortschreitende Klimawandel gefährdet jedoch zunehmend den globalen Weizenanbau. Zudem führen ökologische und politische Rahmenbedingungen zu Einschränkungen beim Einsatz von Stickstoffdünger. Beide Faktoren können zu einer Reduktion des Proteingehalts im Weizenkorn beitragen, was potenziell mit einer Verschlechterung der Backqualität einhergeht.

Deshalb ist die Entwicklung von Weizensorten mit einer hochwertigen Proteinzusammensetzung erforderlich, die trotz eines geringeren Gesamtproteingehalts eine hohe Backqualität gewährleisten. Hierfür ist ein präzises Verständnis darüber notwendig, welche spezifischen Weizenproteine die Backqualität positiv beeinflussen. Diese könnten gezielt in der Züchtung neuer Sorten gefördert und selektiert werden.

Ziel dieser Arbeit war die umfassende proteomische Charakterisierung der BMWpop, die aufgrund ihrer breiten genetischen Basis eine ideale Grundlage für Assoziationsanalysen bietet. Im Mittelpunkt standen die Entwicklung geeigneter Methoden sowie die quantitative Erfassung und funktionelle Interpretation verschiedener Glutenproteingruppen, darunter die  $\alpha$ -,  $\gamma$ - und  $\omega$ -Gliadine sowie die niedermolekularen und hochmolekularen Gluteninuntereinheiten.

Zu diesem Zweck wurden sowohl eine untargeted als auch eine targeted Bottom-Up Proteomics-Methode zur Analyse von Weizenmehl entwickelt, optimiert und auf die BMWpop angewendet. Durch die Verknüpfung der gewonnenen proteomischen Daten mit backqualitätsrelevanten Merkmalen sollen systematische Korrelationen zwischen spezifischen Glutenproteinen und funktionellen Backqualitätsparametern ermittelt werden. Ein besonderes Augenmerk lag dabei auf der Identifikation von Glutenproteinen, die mit einer verbesserten Backqualität assoziiert sind. Durch diese integrative Herangehensweise leistet die Arbeit einen Beitrag zum besseren Verständnis des funktionellen Einflusses einzelner Glutenproteine auf die Endprodukteigenschaften von Weizenmehl und bietet eine Grundlage für die gezielte Verbesserung der Backqualität durch molekulargestützte Selektion.

## 3 Ergebnisse und Diskussion

### 3.1 Untargeted Proteomics von Weizenmehl: Methodenoptimierung

Basierend auf der dreistufigen Osborne-Fraktionierung zur RP-UHPLC-UV-Analyse nach Schalk et al. [81], sowie der Bottom-Up Proteomics-Methode zur LC-MS/MS-basierten Untersuchung von Amylase-Trypsin-Inhibitoren in Weizen nach Geisslitz et al. [82], wurde eine optimierte Methode für die Aufarbeitung und LC-MS/MS-Messung von Weizenmehlproteinen entwickelt. Die in Kapitel 1.2.3 dargestellte Proteomics-Methode für Weizenproteine (**Abbildung 7**) veranschaulicht die einzelnen Arbeitsschritte, die nacheinander optimiert wurden.

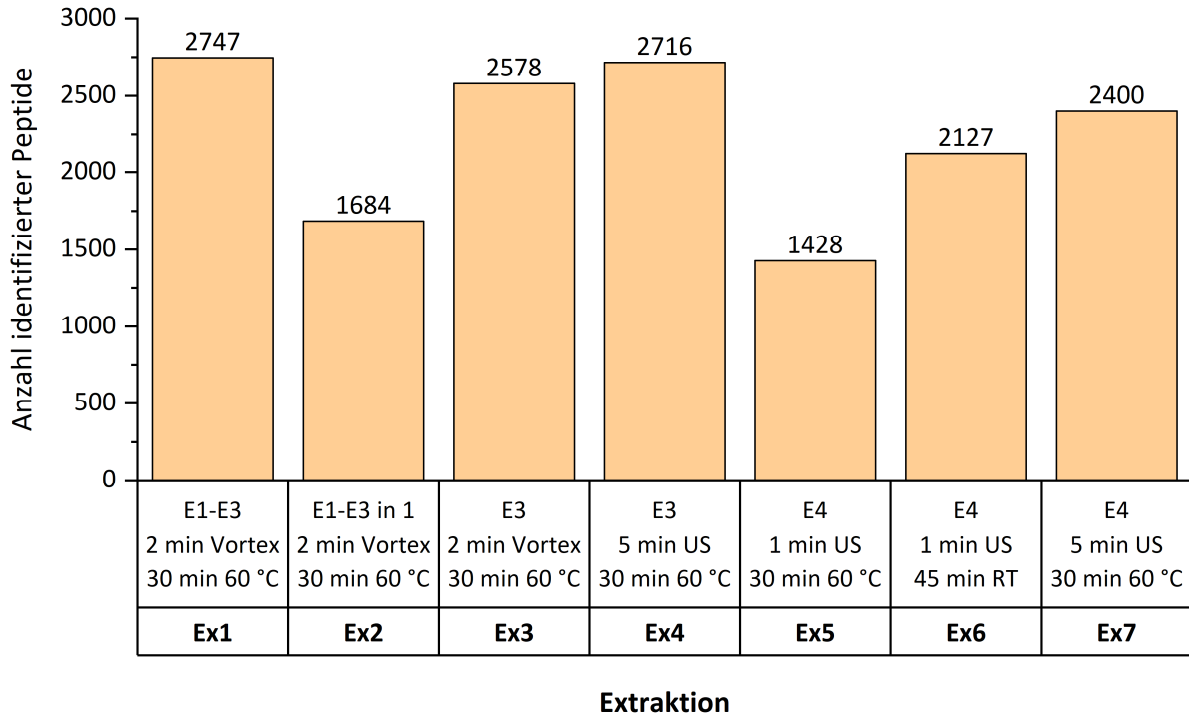
Hauptziel der Optimierung war es, die Anzahl an identifizierten Peptiden zu maximieren. Da letztlich ein umfangreicher Datensatz von 403 Weizenproben (die BMWpop) in Dreifachbestimmung analysiert werden sollte, wurde die Methode auch im Hinblick auf Probenaufarbeitung und Zeitaufwand entsprechend angepasst.

#### 3.1.1 Proteinextraktion

Die Proteinextraktion stellt den ersten kritischen Schritt dar. Von ihr hängt ab, welche Proteine überhaupt extrahiert und somit später detektiert werden können. Bei Weizenproteinen galt lange Zeit die fraktionierte Osborne-Extraktion als Maßstab, da sie unterschiedliche Proteinfractionen getrennt nach ihrer Löslichkeit extrahieren soll. Aufgrund ihres dreistufigen Aufbaus ist diese Methode jedoch sehr zeitintensiv, insbesondere bei der Bearbeitung umfangreicher Probensätze. Aus diesem Grund wurden sieben alternative Extraktionsansätze erprobt und hinsichtlich der Anzahl der jeweils identifizierten Peptide miteinander verglichen.

Zunächst wurde die etablierte Osborne-Fraktionierung angewendet, wobei die drei Extrakte während der restlichen Aufarbeitung weiterhin getrennt aufgearbeitet wurden (Ex1). In einem weiteren Ansatz wurden die drei Fraktionen nach Entnahme in einem Extrakt vereinigt und dieser weiter aufgearbeitet. Es folgte eine Extraktion nur mit der dritten Extraktionslösung E3 (Ex3). Die vierte Extraktion (Ex4) erfolgte wie Ex3, jedoch wurde zusätzlich noch Ultraschall (US) während der Extraktion eingesetzt. Bei Extraktion 5 (Ex5) wurde ACN/Wasser mit 0.1 % Trifluoressigsäure (TFA) als alternatives Extraktionsmittel eingesetzt. Nach dessen Zugabe wurde für 1 Minute mit Ultraschall behandelt und anschließend 45 Minuten bei Raumtemperatur gerührt. Bei Ex6 und Ex7 wurde analog zu Ex5 extrahiert. Es wurde nach der fünfminütigen Ultraschallbehandlung für 45 Minuten bei Raumtemperatur (Ex6) bzw. 30 Minuten bei

60 °C (Ex7) gerührt. Die einzelnen Extraktionsverfahren sind detailliert in 5.4.1.1 beschrieben. Eine Übersicht der Anzahl identifizierter Peptide ist in **Abbildung 8** dargestellt.

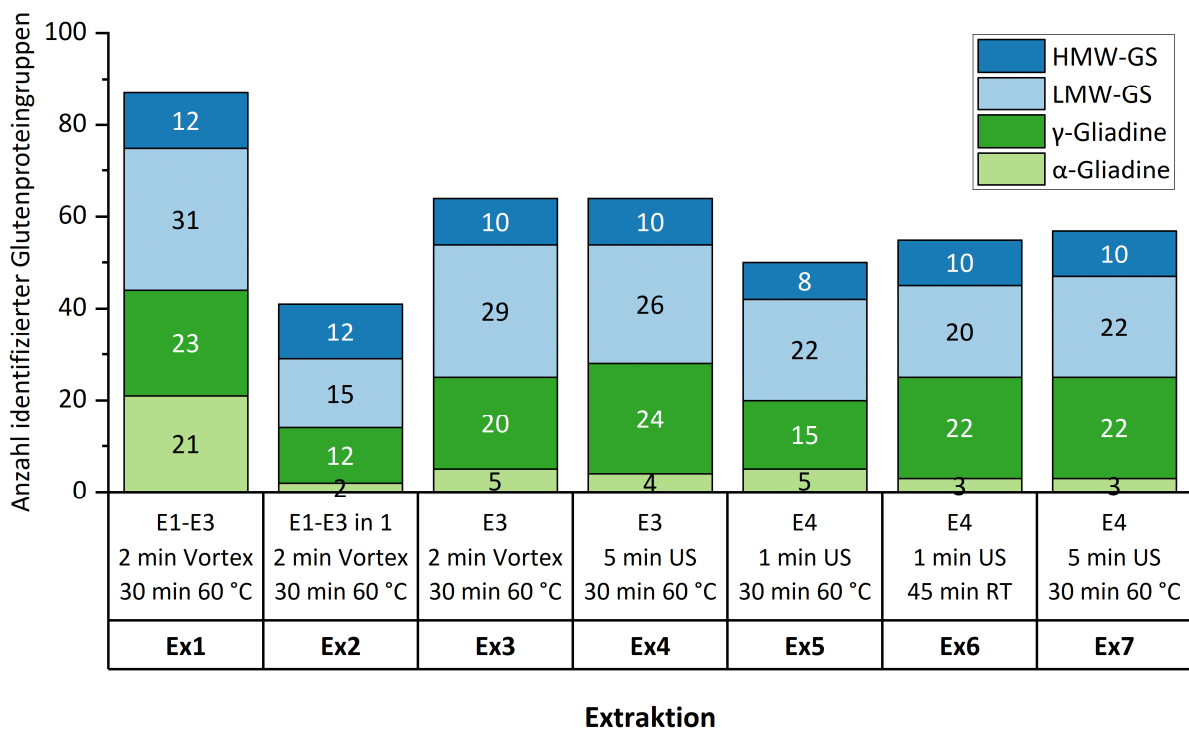


**Abbildung 8:** Anzahl der identifizierten Peptide nach verschiedenen Proteinextraktionen aus Weizenmehl der Sorte RGT Reform. E1-4, Extraktionsmittel 1-4; E1: 0,07 mol/L Na<sub>2</sub>HPO<sub>4</sub> × H<sub>2</sub>O und 0,4 mol/L NaCl (pH = 7,6); E2: 60 % Ethanol in Wasser; E3: 50 % (v/v) 1-Propanol in 2 mol/L Harnstoff und 0,1 mol/L TRIS-HCl (pH 7,5) plus 1 % (w/v) Dithiothreitol in Wasser; E4: 50 % Acetonitril, 0,1 % Trifluoressigsäure in Wasser; Ex1-Ex7, Extraktionsversuch 1-7; US, Ultraschall.

In Ex1 wurde mit 2747 identifizierten Peptiden die höchste Ausbeute erzielt. Ex4 lieferte mit 2716 Peptiden ein nahezu vergleichbares Ergebnis (knapp 99 % im Vergleich zu Ex1). Wurde bei der Einfachextraktion kein Ultraschall eingesetzt (Ex3), lag die Ausbeute bei 93,8 %. Die Zusammenführung der Fraktionen E1, E2 und E3 in einem Extrakt (Ex2) führte zu einer deutlich geringeren Peptidanzahl von 61,3 %. Die geringste Ausbeute wurde bei der kurzen Ultraschallbehandlung mit E4 (Ex5) erzielt, mit nur 1428 identifizierten Peptiden, entsprechend 52 % von Ex1. Durch Verlängerung der Ultraschalldauer (Ex6) bzw. Erhöhung der Rührtemperatur auf 60 °C (Ex7) konnte die Peptidanzahl auf 77,4 % bzw. 87,4 % gesteigert werden.

Insgesamt zeigte der Vergleich der Extraktionsmethoden, dass die Einfachextraktion mit Ultraschall (Ex4) hinsichtlich der Anzahl identifizierter Peptide eine Alternative zu der ausführlichen Extraktion Ex1 darstellt, denn es wurden insgesamt nur geringfügig weniger Peptide identifiziert. Die Extraktion mit E4 und Ultraschall- und Hitzebehandlung (Ex7) lieferte ebenfalls ein recht gutes Ergebnis, das an die Extraktion mit komplexeren Extraktionsmitteln wie E3 heranreicht. Die Daten aus Ex2 stammen aus einem

unvollständigen Datensatz, da es infolge der hohen Salzbelastung zur Verblockung der SPE-Kartuschen kam. Diese Extraktionsmethode erwies sich daher als nicht geeignet. Für eine zukünftige Anwendung wäre eine Anpassung der Salzkonzentration bzw. des Bettgewichts der SPE-Kartuschen erforderlich. In Anbetracht der guten Ergebnisse der zeitlich deutlich weniger aufwendigen Einfachextraktion (Ex3 und 4) stellt sich die Frage, ob der höhere Aufwand der dreistufigen Extraktion gerechtfertigt ist. Eine mögliche Rechtfertigung könnte darin liegen, dass durch die dreistufige Extraktion auch solche Proteingruppen identifiziert werden können, die mit der einfachen Extraktion nicht erfasst werden und die im Fokus der wissenschaftlichen Fragestellung stehen. Diese Möglichkeit ist auch bei Ex5-7 zu berücksichtigen. Aufgrund der unterschiedlichen Zusammensetzung der Extraktionsmittel ist es deshalb naheliegend, dass unterschiedliche Proteinmengen bzw. andere Proteine extrahiert werden, die zu veränderten Proteinzuordnungen führen. Deshalb wurde in einem nächsten Schritt die Anzahl der zugeordneten Glutenproteingruppen analysiert (**Abbildung 9**).



**Abbildung 9:** Anzahl der identifizierten Glutenproteingruppen nach verschiedenen Proteinextraktionen aus Weizenmehl der Sorte RGT Reform. Ex1-Ex7, Extraktionsversuch 1-7; HMW-GS, hochmolekulare Gluteninuntereinheit; LMW-GS, niedermolekulare Gluteninuntereinheit; US, Ultraschall.

Die Ergebnisse zeigen, dass mit der umfassendsten Extraktion, Ex1, auch die meisten Glutenproteine (90) identifiziert wurden, gefolgt von Ex3/4 (jeweils 66) und Ex7 (59). Mit Ex6, Ex5 und Ex2 wurden jeweils 59, 57 und 43 Glutenproteine identifiziert. Aufgrund von Schwierigkeiten bei der Aufarbeitung wird Ex2 bei

der weiteren Auswertung nicht mehr berücksichtigt. Mit Extraktionsmittel E3 wurden geringfügig mehr Glutenproteingruppen identifiziert als mit E4. Dies äußert sich in erster Linie bei den Gluteninen. Dort wurden mit Ex1 und Ex3/4 36-44 Glutenine identifiziert, während bei Ex5-7 30-33 festgestellt wurden. Dies lag vor allem an den LMW-GS. 26-31 wurden nach Ex1-4 detektiert, während nach Ex5-7 20-22 festgestellt wurden. Bei den HMW-GS wurden nach allen Extraktionen 8-12 individuelle HMW-GS analysiert. Bei den Gliadinen wurden mit Ex3-7 ähnlich viele Gliadine (22-27) zugeordnet. Den größten Unterschied, der sich auch in der Gesamtanzahl widerspiegelt, wurde bei Ex1 gefunden. 46 Gliadine konnten identifiziert werden, was nahezu doppelt so viel ist, wie bei den anderen Extraktionen. Nähere Betrachtung zeigte weiterhin, dass dies an der identifizierten Anzahl an  $\alpha$ -Gliadinen lag, die bei Ex1 21 betrug, während nach allen anderen Extraktionen lediglich 2-5  $\alpha$ -Gliadine analysiert wurden. Bei den  $\gamma$ -Gliadinen wurde nach allen Extraktionen eine ähnliche Anzahl festgestellt (15-23). Weiterhin fällt auf, dass die  $\omega$ -Gliadine nach keiner der Extraktionen identifiziert wurden. Es wurde bereits festgestellt, dass einigen Glutenproteinen mögliche Spaltstellen für Trypsin fehlen [75, 83]. Daraus ergibt sich, dass die Ursache vermutlich nicht in der Extraktion, sondern vielmehr im eingesetzten Verdauungsenzym liegt.

Der bedeutendste Unterschied innerhalb der getesteten Extraktionsansätze bestand in den verwendeten Extraktionsmitteln E3 und E4, die sich sowohl in ihrer chemischen Zusammensetzung als auch in ihrer Wirkung auf die Proteinextraktion unterschieden. E3 besteht aus 50 % (v/v) 1-Propanol und 50 % 0,1 mol/L TRIS-HCl (pH 7,5), ergänzt mit 2 mol/L Harnstoff und 1 % (w/v) DTT. Diese Komponenten entfalten eine starke denaturierende Wirkung: Harnstoff löst die native Proteinstruktur auf, wodurch DTT effektiv Disulfidbrücken reduzieren kann. In der Folge werden polymere Glutenine in ihre löslichen monomeren Untereinheiten überführt und in der alkoholischen Lösung extrahierbar gemacht. Damit eignet sich E3 besonders gut zur gezielten Extraktion der Gluteninfraktion.

Im Gegensatz dazu ist E4 einfacher aufgebaut und besteht lediglich aus Acetonitril (ACN) und Wasser im Verhältnis 1:1, ergänzt mit 0,1 % TFA. Dieses Extraktionsmittel erfordert keine pH-Einstellung, ist schnell herzustellen und enthält keine ionisationshemmenden Substanzen wie Harnstoff oder DTT, die bei der ESI-Ionisation zur Ionensuppression führen können. Daher ist bei der Anwendung von E4 keine Aufreinigung durch SPE erforderlich. Dies stellt einen deutlichen praktischen Vorteil gegenüber E3 dar. Gleichzeitig erklärt sich daraus aber auch, warum mit E4 weniger Glutenine extrahiert wurden. Ohne Denaturierung und Reduktion bleiben diese Proteine in ihrer polymeren Form unlöslich bzw. reicht die US-Behandlung nicht aus. Generell zeigte die Verwendung von US lediglich einen moderaten Anstieg der Peptidanzahl, hatte jedoch keinen Einfluss auf die Gesamtanzahl identifizierter Proteine: In beiden Fällen (Ex3 und Ex4) wurden jeweils 66 Proteine detektiert. Die erhöhte Peptididentifikation bei Ex4 deutet jedoch auf eine

verbesserte Proteinsequenzabdeckung hin, was die Qualität der Analyse trotz konstanter Proteinanzahl verbessern kann. Ein Vergleich der Extraktionsansätze Ex1 bis Ex4 verdeutlicht zudem einen qualitativen Unterschied in der Extraktion der  $\alpha$ -Gliadine: Mit Ex1, das die vollständige dreistufige Osborne-Extraktion umfasst, wurden deutlich mehr  $\alpha$ -Gliadine identifiziert als mit den einfacheren Ansätzen Ex3 und Ex4. Dies weist darauf hin, dass diese Proteine effizienter mit E2 (60 % Ethanol) extrahiert werden, welches nur im Rahmen der vollständigen Fraktionierung eingesetzt wird.

Die Osborne-Fraktionierung hat mehrere Nachteile, auch wenn sie die meisten Peptid- und Proteinidentifizierungen erbrachte. Erstens gibt es viele Variationen der Durchführung in der Literatur, die zu einem breiten Ergebnisspektrum führen. Zweitens verbleibt ein Teil der Proteine im Rückstand und wird somit nicht erfasst. Drittens kann es zu einer Überlappung der extrahierten Fraktionen kommen [84]. Das heißt, es befinden sich nicht nur die erwarteten Proteine in der Fraktion, sondern auch Proteine aus anderen Fraktionen. Dies kann mit RP-HPLC nicht überprüft werden, da dort Chromatogrammflächen der Glutenproteintypen verglichen werden und keine Bestimmung individueller Glutenproteine möglich ist [81]. Ein Vorteil der LC-MS/MS-Messung ist, dass mit dieser eine exakte Identifizierung individueller Proteine bzw. Proteingruppen erfolgt. So konnten in der ersten Fraktion von Ex1, welche die Albumine und Globuline extrahieren soll, bereits 14 Gliadin- und 10 Gluteninproteingruppen identifiziert werden. Bei der zweiten Fraktion, die die Gliadine extrahieren soll, wurden neben 30 verschiedenen Gliadinen 19 Gluteninproteingruppen gefunden. Die letzte Fraktion enthielt neben 43 Gluteninen 25 Gliadinproteingruppen. Auf Grundlage dieser Ergebnisse ist von einer Überlappung der Fraktionen auszugehen. Der Umfang der Fraktionsüberlappung wurde im Rahmen dieser Arbeit nicht weiter untersucht, da für die Analyse der BMWpop die Einfachextraktion gewählt wurde und eine detaillierte Abgrenzung der Fraktionen somit nicht im Fokus stand. Eine weiterführende Untersuchung wäre unter anderem im Hinblick auf den Vergleich mit UHPLC-Messungen von Interesse. Hierfür bietet sich ein systematischer Vergleich der LFQ-Intensitäten ausgewählter Proteine an. Ergänzend könnten Reproduzierbarkeitsprüfungen Aufschluss darüber geben, ob bestimmte Proteine verlässlich und in vergleichbarer Menge extrahiert werden.

Bose et al. analysierten die Gesamtheit der aus Weizenmehl extrahierbaren Proteine unter Verwendung von drei verschiedenen Extraktionsmethoden. Dabei kamen Pufferlösungen auf Basis von TRIS-HCl, TRIS-HCl-Harnstoff bzw. Isopropanolalkohol (alle drei mit 2 % DTT) zum Einsatz. Aus den Ergebnissen zweier Weizensorten konnten so 1761, 1574 bzw. 408 Proteine nach der nanoLC-MS/MS-Messung der Peptide zugeordnet werden [85]. Nach Ex1, Ex4 und Ex7 konnten insgesamt 716, 661 und 754 Proteingruppen identifiziert werden (Daten nicht gezeigt). Die großen Unterschiede in der Gesamtanzahl identifizierter

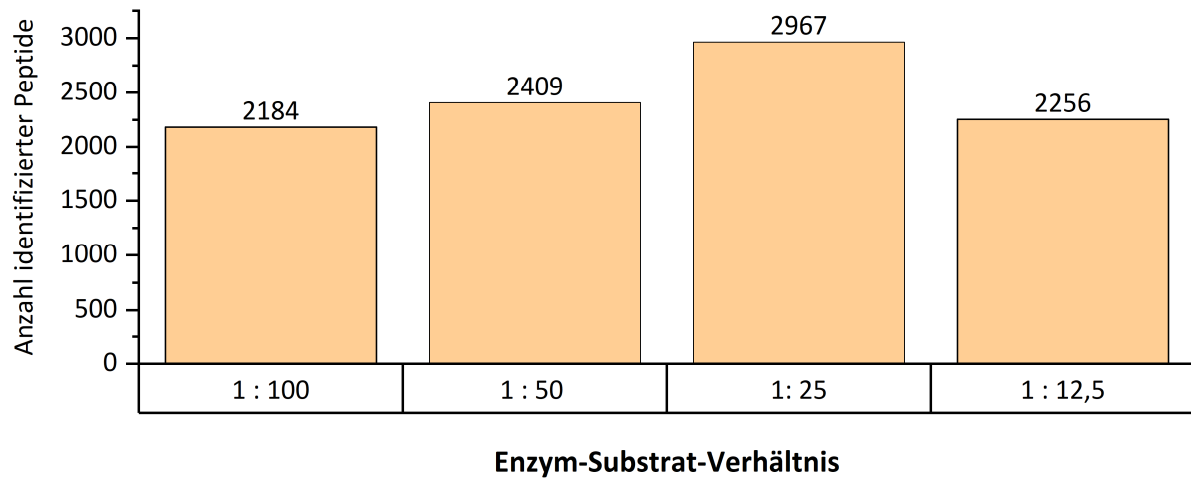
Proteine nach TRIS-HCl bzw. TRIS-HCl-Harnstoff-Extraktion zu den Ergebnissen hier können damit erklärt werden, dass hier nur eine Weizensorte analysiert wurde und ein UHPLC-System zur Trennung der Peptide verwendet wurde, welches nicht so sensitiv ist wie ein nanoLC-System. Dennoch zeigte sich bei Bose et al., dass TRIS-HCl- bzw. TRIS-HCl-Harnstoff-Puffer wesentlich besser geeignet waren, als der Isopropanolalkohol. Das ist hier nicht der Fall, da nach Ex7 die meisten Proteingruppen identifiziert wurden. Dies zeigt erneut, dass E4 bei der Extraktion von Weizenproteinen eine Alternative zu den herkömmlichen TRIS-HCl-Harnstoff-Puffern darstellt.

Ex1 ist zwar die umfangreichste Extraktion, aber auch die zeitaufwendigste. Dies ist ein Nachteil bei der Aufarbeitung eines großen Probensatzes. Weiterhin ist die fraktionierte LC-MS/MS-Messung auch bei der späteren Auswertung mittels MaxQuant problematisch. Dadurch, dass sich manche Glutenproteine in allen drei Extrakten identifizieren lassen, müssen diese Ergebnisse bei der Auswertung wieder zusammengeführt werden. Dies kann vor allem bei der quantitativen Auswertung zu einer Verzerrung der Ergebnisse führen. Ex7 ist zwar eine vielversprechende Alternative zur Extraktion mit TRIS-HCl-Harnstoff-DTT-Puffern, jedoch war diese für die Extraktion der Glutenine weniger gut geeignet. Da diese Fraktion die bedeutendste Gruppe im Hinblick auf die Backqualität darstellt, ist eine ausführliche Extraktion erstrebenswert. Um sowohl eine zuverlässige Extraktion als auch ein Zeitersparnis zu gewährleisten, wurde die Methode mit Ex4 weiterverwendet, da sie einen guten Kompromiss zwischen Extraktionsdauer und Ausbeute an Glutenproteinen bot. Im Folgenden dient sie zudem als Referenz für alle weiteren Optimierungsansätze.

### 3.1.2 Proteinverdau

Bei Bottom-Up Proteomics ist das Enzym-Substrat-Verhältnis (E/S) ein entscheidender Parameter für die Effizienz und Spezifität des Proteinverdauprozesses, insbesondere bei der Verwendung von Trypsin. Typischerweise wird das E/S-Verhältnis als Massenverhältnis (w/w) angegeben. Die Verhältnisse 1:50 und 1:100 (w/w) sind Standard in Proteomics und bieten einen guten Kompromiss zwischen Verdau-Effizienz und Vermeidung von Autolyse. Sie werden häufig für Übernacht-Verdauungen bei 37 °C eingesetzt [86].

Die optimalen Bedingungen können je nach Proteinprobe variieren, weshalb bei dem Verdauungsschritt mit der Hydrolase Trypsin vier verschiedene E/S-Verhältnisse (1:100, 1:50, 1:25 und 1:12,5) nach Ex4 getestet wurden (5.4.3). Nach Auswertung mit MaxQuant wurden unterschiedliche Mengen an Peptiden identifiziert (**Abbildung 10**).



**Abbildung 10:** Anzahl der identifizierten Peptide nach enzymatischem Verdau von Weizenmehl der Sorte RGT Reform mit unterschiedlichen Enzym-Substrat-Verhältnissen.

Die Anzahl der Peptididentifizierungen variierte in Abhängigkeit vom Enzym-Substrat-Verhältnis. Es wurde festgestellt, dass mit einem Enzym-Substrat-Verhältnis von 1 zu 25 die meisten Peptide bestimmt wurden. Zur Optimierung der Bottom-Up Proteomics-Methode ist eine möglichst hohe Zahl an Peptididentifizierungen wünschenswert, da sie die Präzision der Zuordnung und Sequenzabdeckung von Proteinen erhöht. Folglich wurde im Hinblick auf die Methodenoptimierung der Bottom-Up Proteomics-Methode von Weizenmehl ein Enzym-Substrat-Verhältnis von 1:25 gewählt.

### 3.1.3 Proteinaufreinigung

Zur Entfernung der Salze wurden die Proben vor der Messung durch LC-MS mittels SPE aufgereinigt. Dies erfolgte zunächst in einzelnen SPE-Röhrchen 5.4.4.1. Dabei wurden insgesamt 2349 Peptide identifiziert.

Dies ist ein sehr zeitaufwendiger Prozess, wenn viele Proben aufgereinigt werden müssen, da nur 24 Proben parallel bearbeitet werden können. Deshalb wurde die SPE in 96-Well-Mikroplatten durchgeführt (**Ex1**, 3.1.1). Dabei besteht die Gefahr des Trockenlaufens der einzelnen Wells, was zu einem Verlust an Analyt führen kann. Da allerdings insgesamt 2747 Peptide identifiziert wurden, wurde diese Form der SPE als gleichwertig angesehen. Dies stellte eine große Zeitersparnis dar, da mit weniger Aufwand die vierfache Menge an Proben gleichzeitig aufgereinigt werden kann. Folglich wurden in allen weiteren Versuche nur noch 96-Well-Mikro-SPE-Platten verwendet.

### 3.1.4 Weitere Optimierungsversuche

#### 3.1.4.1 Verwendung von 96-Well-Mikroplatten

Die Verwendung von 96-Well-Mikroplatten erlaubt eine parallele Probenverarbeitung, wodurch Einzelbearbeitungen entfallen und sich die Aufarbeitungszeit deutlich reduziert. Die gleichzeitige Bearbeitung von bis zu 96 Proben erhöht nicht nur die Effizienz, sondern sorgt auch für konsistente Bedingungen, was die Probenvariabilität minimiert. Diese Vorteile sind besonders bei der Verarbeitung großer Probensätze relevant, weshalb in dieser Arbeit angestrebt wurde, die Plattenformate gezielt in den Workflow zu integrieren.

Die optimierte Aufarbeitung wurde komplett durchgeführt. Im letzten Schritt der SPE wurde direkt in eine 96-Well-Mikroplatte eluiert. Nach der Entfernung des Lösungsmittels wurden die Proben für die Messung ebenfalls direkt in der Platte gelöst und diese zur Messung in die UHPLC gestellt (5.4.5.1). Nach der Messung konnten insgesamt 1796 Peptide in der Dreifachbestimmung von RGT Reform identifiziert werden. Das sind knapp 1000 Peptide weniger als bei der Einfachextraktion Ex4 aus Abschnitt 3.1.1 (2716 Peptididentifikationen) und kann deshalb auch nicht mehr als normale Standardabweichung angesehen werden. Es wurde noch ein Versuch durchgeführt, bei dem komplett in 96-Well-Mikroplatten gearbeitet wurde (Protokoll nicht im Experimentellen Teil hinterlegt). Danach wurden noch weniger Peptide (insgesamt 1044) identifiziert.

96-Well-Mikroplatten finden in der aktuellen Literatur bereits breite Anwendung in der automatisierten Hochdurchsatz-Proteomics und haben sich dort als erfolgreiches Format etabliert [87-90]. In der vorliegenden Arbeit konnte dieser Vorteil jedoch nicht bestätigt werden, da keine vergleichbaren Ergebnisse erzielt wurden. Ein zentrales Problem zeigte sich beim Abdampfen des Lösungsmittels (ACN/Wasser, 40:60, v/v) in der Vakuumzentrifuge: Die Verdunstung verlief in den Wells der Platte uneinheitlich, sodass einige Proben bereits nach drei Stunden vollständig getrocknet waren, während andere bis zu zwölf Stunden benötigten. Dies führte möglicherweise dazu, dass sich bestimmte Peptide beim anschließenden Lösen nicht mehr vollständig rekonstituieren ließen und somit nicht detektiert wurden. Bei dem durchgeführten Lösungsprozess wurde das Lösungsmittel zunächst auf die getrockneten Proben pipettiert und 30 Minuten ruhen gelassen. Anschließend wurde die Platte zehn Minuten vortexgemischt und für fünf Minuten in ein Ultraschallbad gestellt. Eine Verlängerung der Misch- und Ultraschalldauer könnte hier potenziell die Wiederlösbarkeit verbessern, würde sich je nachdem negativ auf die Analyten auswirken.

Viele in der Literatur beschriebenen Protokolle führen keine Trocknungsschritte durch. Seo et al. verzichteten in ihrem Workflow vollständig auf Entsalzungs- und Trocknungsschritte und konnten dennoch erfolgreich die gesuchten Zielanalyten in Serumproben menschlicher Patienten detektieren – trotz des Vorhandenseins von Harnstoff, der bekanntlich die LC-MS/MS-Messung beeinträchtigen kann [87]. Problematisch ist dabei nicht nur die mögliche Signalunterdrückung durch Salze, sondern auch die Kontamination des Massenspektrometers sowie die damit verbundene Beeinträchtigung der Datenqualität, insbesondere bei der vollständigen Erfassung des Proteoms.

Andere Protokolle nutzen die Filter Aided Sample Preparation (FASP) in Kombination mit 96-Well-Mikroplatten [90] oder magnetische Beads, bei denen nur recht kurz getrocknet wird [89]. Der Einsatz von FASP war in dieser Arbeit jedoch problematisch, da die in Weizenmehl enthaltene Stärke die Filtermembranen verstopfte. Magnetische Beads stellen daher eine potenziell geeignete Alternative dar, deren Einsatz in zukünftigen Arbeiten getestet werden sollte. Alternativ wurden weitere Versuchsansätze durchgeführt, bei denen gezielt auf die Trocknungsschritte verzichtet wurde. Diese werden im Folgenden näher erläutert.

#### 3.1.4.2 Aufarbeitung ohne Zwischentrocknung

Da die Proben bei der Standardaufarbeitung sowohl nach der Extraktion als auch nach der Reduktion und Alkylierung jeweils über mehrere Stunden in der Vakuumzentrifuge eingedampft werden, erstreckt sich die gesamte Aufarbeitung über drei Arbeitstage. Durch den Verzicht auf die Trocknungsschritte könnte die Gesamtzeit auf zwei Tage reduziert werden. Zudem würde dies die Durchführung im 96-Well-Format ermöglichen, wie bereits zuvor diskutiert (3.1.4.1). Ein weiterer Vorteil dieses Ansatzes besteht darin, dass lediglich das Extraktionsmittel vorbereitet werden muss. In der Standardmethode wird 1-Propanol verwendet, das jedoch die Trypsinaktivität hemmt. Daher wurde es in diesem Ansatz durch 50 % ACN ersetzt. Zusätzlich wurde der pH-Wert des TRIS-HCl-Puffers auf 8,5 eingestellt, um optimale Bedingungen für den enzymatischen Verdau zu gewährleisten. Anschließend wurde ein Versuch durchgeführt, bei dem die einzelnen Aufarbeitungsschritte ohne Trocknung direkt hintereinander erfolgten (5.4.5.2).

Nach der anschließenden Auswertung konnten insgesamt nur 838 Peptide identifiziert werden im Gegensatz zu 2716 Peptiden mit Ex4. Dies deutet darauf hin, dass 50 % ACN den Verdau durch Trypsin ebenfalls hemmt. Es gibt in der Literatur Versuchsprotokolle, die ACN während des Trypsinverdau nutzen und damit gute Erfahrung gemacht haben. So erzielten Hervey et al. die meisten Peptididentifizierungen nach dem Verdau mit Trypsin in 80 % ACN-Lösung [91]. In einer weiteren Studie konnte die Zugabe von 10 % ACN zum Reaktionspuffer die Effizienz des tryptischen Verdau soweit steigern, dass sich die

erforderliche Inkubationszeit von über Nacht auf nur vier Stunden verkürzte [92]. Eine Reduktion des ACN-Anteils im Extraktionspuffer auf 10 % könnte möglicherweise zu einem besseren tryptischen Verdau führen. Allerdings ist dann unklar, wie effizient die Extraktion funktioniert. Dies müsste erst überprüft werden, indem die restliche Aufarbeitung nach dem Standardversuch durchgeführt wird. Dies gilt auch für den bereits durchgeführten Versuch mit 50 % ACN-Anteil im Extraktionspuffer. Dadurch, dass nur so wenige Peptide identifiziert werden konnten, stellte der durchgeführte Versuch keine Alternative zu dem Standardversuch dar. Dennoch bestehen Optimierungspotenziale wie die Reduktion des ACN-Anteils in der Extraktionslösung, die in zukünftigen Untersuchungen berücksichtigt werden könnten.

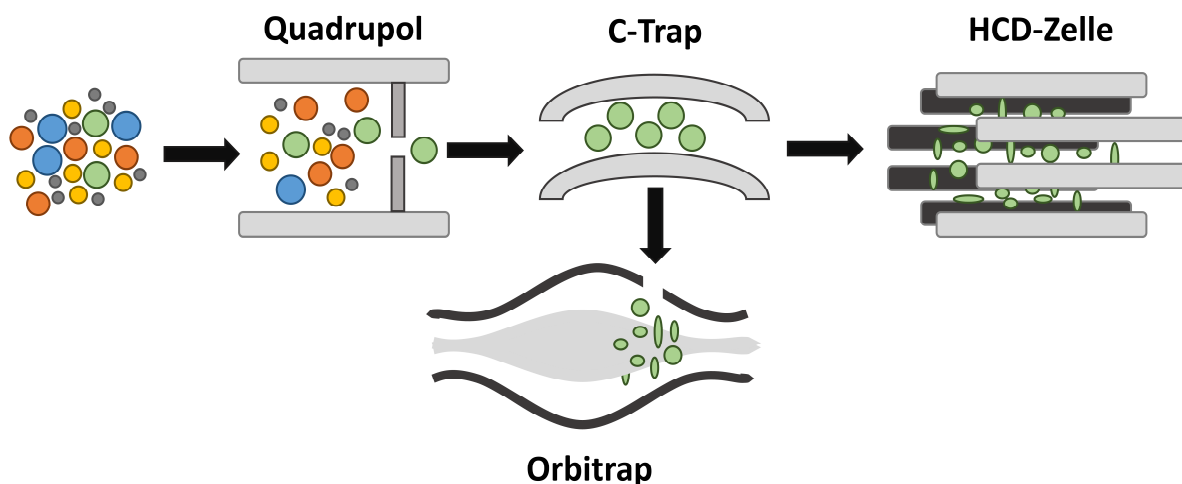
#### 3.1.4.3 Aufarbeitung ohne Zwischentrocknung nach Alkylierung

Um zumindest einen Trocknungsschritt zeitlich einzusparen, wurden die Proben nach der Alkylierung nicht in der Vakuumzentrifuge getrocknet, sondern direkt im Anschluss enzymatisch verdaut (5.4.5.3). Es konnten insgesamt 2056 Peptide identifiziert werden, was 76 % von Ex4 entspricht.

Die Proben wurden für die Reduzierung und Alkylierung in 100 % TRIS-HCl-Harnstoff gelöst (pH 7,8). Bei der Standardmethode werden sie in 50 % TRIS-HCl (pH 8,5) und 50 % 1-Propanol gelöst. Das 1-Propanol wurde in diesem Versuch weggelassen, da es bei dem Verdau mit Trypsin stört. Eventuell waren die Proben für die Reduzierung deshalb nicht optimal gelöst. Allerdings war auch das Puffersystem der TRIS-HCl-Lösung mit pH 7,8 nicht so stark und alkalisch angesetzt, wie bei der Standardmethode. Sowohl TCEP, welches zur Reduktion der Proteine eingesetzt wird, als auch CAA, welches zur Alkylierung der Cysteingruppen eingesetzt wird, haben eine leicht saure Wirkung in wässriger Lösung. Deshalb wird in Proteomics-Protokollen ein neutraler Puffer verwendet, um den pH-Wert konstant zu halten. Dies ist notwendig, um eine effiziente Alkylierung mit CAA zu gewährleisten. Dieses reagiert mit den freien Thiolgruppen (-SH) von Cystein. Dazu müssen diese deprotoniert als nucleophiles Thiolat-Ion (-S<sup>-</sup>) vorliegen, was nur bei neutralem bis leicht basischem pH der Fall ist. Bei saurem pH liegt Cystein vollständig protoniert vor und ist damit reaktionsträge gegenüber der Alkylierung. Weiterhin hat die Serinprotease Trypsin eine stabile Aktivität zwischen pH 7 und 9, sein pH-Optimum liegt bei pH 7,5–8,5. Wenn der pH-Wert unter sechs fällt, verliert Trypsin schnell seine Struktur und damit seine Enzymaktivität. Ein vollständiger Verdau ist damit nicht mehr gewährleistet. Die Wirkung des in diesem Versuch verwendeten Puffers könnte nicht ausgereicht haben, um die saure Wirkung von TCEP und CAA auszugleichen, weshalb der pH ins Saure umgeschlagen ist. In der Folge wären die Alkylierung und der Verdau beeinträchtigt gewesen. Bei Folgeversuchen sollte deshalb ein stärkerer Puffer getestet werden bzw. der pH-Wert während der Versuchsdurchführung beobachtet werden.

### 3.1.5 Optimierung der MS-Parameter

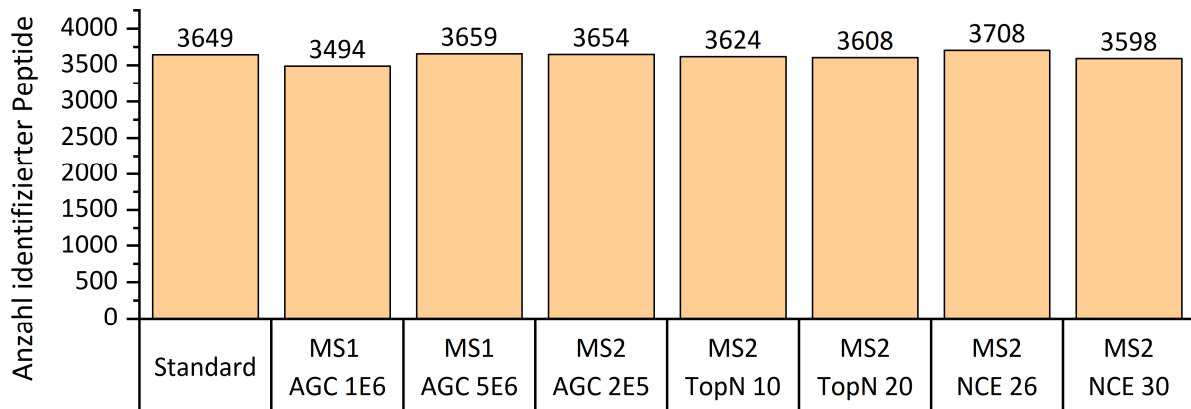
Zunächst wurden die ESI-Parameter optimiert (**Tabelle 7** in 5.4.6). Bei der ESI werden gelöste Ionen in die Gasphase überführt. Dazu wird eine Sprühspannung an die ESI-Nadel angelegt. Zur Unterstützung des Ionisationsprozesses kommen verschiedene Gase zum Einsatz. Das Mantelgas ist ein Gasstrom, der die Probenlösung in der ESI-Quelle umgibt. Seine Hauptfunktion besteht darin, die Bildung eines feinen Aerosols zu unterstützen, indem es die Probenlösung stabilisiert und den Verdampfungsprozess fördert. Das Hilfsgas wird zusätzlich zum Mantelgas eingesetzt. Es beschleunigt den Trocknungsprozess und verbessert so die Ionisationseffizienz. Das Sweep-Gas entfernt unerwünschte Dämpfe aus der Ionenquelle, reduziert Hintergrundrauschen und trägt zur Stabilität und Sauberkeit der Messung bei [93, 94]. Es stellte sich heraus, dass die angegebenen Standardbedingungen für einen LC-Fluss von 0,2 mL/min am besten geeignet waren (Daten nicht gezeigt).



**Abbildung 11:** Schematischer Aufbau eines Orbitrap-Massenspektrometers, adaptiert und modifiziert nach Michalski [95]. Die Ionen werden zunächst im Quadrupol massenselektiv vorselektiert und anschließend in der C-Trap gesammelt, um das AGC-Target zu erreichen. Für die Erzeugung von MS<sup>2</sup>-Spektren erfolgt die Fragmentierung in der HCD-Zelle. Die entstehenden Produkt-Ionen werden anschließend in der Orbitrap detektiert. AGC, Automatic Gain Control; HCD, Higher-energy Collisional Dissociation.

Weiterhin wurden ausgehend von einer Standardmethode verschiedene Einstellungen der DDA-Methode getestet (**Tabelle 8** in 5.4.6). Bei den variierten Parametern handelte es sich zunächst um das Automatic Gain Control Target (AGC Target), TopN und die Kollisionsenergie. Das AGC Target steuert die Anzahl der Ionen, welche in die Orbitrap gelangen, indem die vorgeschaltete C-Trap für eine berechnete Zeitspanne gefüllt wird. Innerhalb dieser Zeit werden die Ionen in der C-Trap gesammelt, zu einer engen Wolke komprimiert und in die Orbitrap weitergeleitet (**Abbildung 11**). Auf diese Weise gelangen die Ionen nur zu bestimmten Zeiten in den Massendetektor [96]. Wenn das AGC Target niedrig angesetzt wird, erfolgt ein

schnellerer Scan, aber die Signalintensität fällt geringer aus und es besteht das Risiko, dass es unterhalb der Detektionsgrenze ist. Ein hohes AGC Target benötigt längere Zykluszeiten, dafür ist die Signalqualität besser. Die TopN-Methode steuert die Anzahl der intensitätsstärksten Präkursorionen, die aus einem MS1-Scan für eine anschließende Fragmentierung (MS2-Scan) ausgewählt werden (siehe auch **Abbildung 6**). Je höher das TopN gewählt wird, desto mehr MS2-Scans gibt es pro Zyklus, was die Zykluszeit verlängert. Eine zu lange Zykluszeit kann allerdings dazu führen, dass schnell eluierende Peptide verpasst werden. Damit hängt TopN auch von den gewählten LC-Parametern ab. In komplexeren Proben kann ein höheres TopN mehr Ionen abdecken, aber es kann gegebenenfalls zu einer Verschlechterung der Fragmentqualität aufgrund kürzerer Injektionszeiten kommen. Ein kleiner gewählter TopN-Wert in Kombination mit längerer Injektionszeit führt zu qualitativ hochwertigeren MS2-Spektren, kann jedoch die Anzahl identifizierter Peptide einschränken. Die Wahl des passenden TopN stellt daher einen Kompromiss zwischen Spektrenqualität und Proteomabdeckung dar. Je nach Komplexität der Proben, wird TopN gewöhnlich zwischen drei und 20 gewählt [95]. Die Kollisionsenergie bestimmt, mit welcher Energie Ionen in der HCD-Zelle (Higher-energy Collisional Dissociation) fragmentiert werden. Sie wird als dimensionslose, normalisierte Kollisionsenergie (normalized collision energy, NCE) angegeben. Eine optimal gewählte NCE sorgt für eine effektive Fragmentierung und ermöglicht dadurch eine zuverlässige Identifikation der Peptide. Typische NCE-Werte für Peptide liegen üblicherweise in einem Bereich zwischen 25 und 30 [96]. Die Ergebnisse der verschiedenen DDA-Messungen sind in **Abbildung 12** dargestellt.



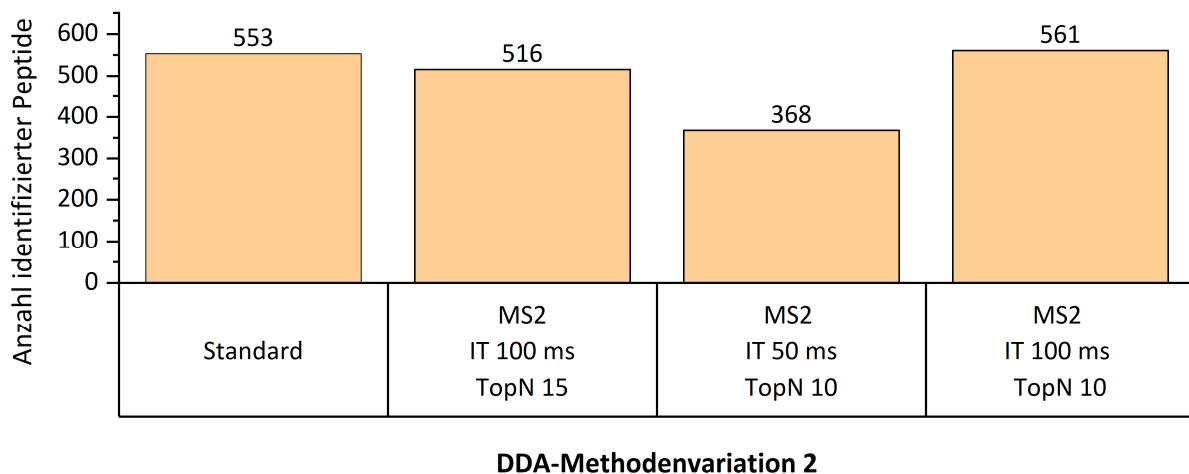
**DDA-Methodenvariation 1**

**Abbildung 12:** Variation der DDA-Methode und detektierte Anzahl an Peptiden. AGC, Automatic Gain Control; DDA, Data Dependent Acquisition; MS1, Full Scan; MS2, MS/MS Scan; NCE, Normalized Collision Energy.

Es wurden kaum Unterschiede zwischen den einzelnen DDA-Methoden beobachtet. Alle Werte befanden sich zwischen 3500 und 3700 identifizierten Peptiden. Deshalb wurden bereits die Standardeinstellungen als gut geeignet eingeschätzt. Weiterhin zeigte sich, dass genügend Analyt vorhanden war, da sich die

Variation des AGC Targets kaum ausgewirkt hat. Auch die Variation von TopN zwischen 10, 15 und 20 ergab keine großen Unterschiede. Weiterhin wurden noch zwei Kollisionsenergie-Einstellungen getestet. In diesem Versuch schien die etwas niedrigere Kollisionsenergie (26) leicht besser geeignet zu sein.

TopN und die maximale Injektionszeit sind eng miteinander verknüpft. Ein hoher TopN-Wert in Kombination mit langer Injektionszeit verlängert die Zykluszeit, was das Risiko erhöht, kurz eluierende Peptide nicht zu erfassen. Umgekehrt kann ein hoher TopN-Wert bei kurzer Injektionszeit zu einer unzureichenden Ionenspeicherung in der C-Trap führen, was die Qualität der MS2-Spektren und somit die Identifizierungsrate negativ beeinflusst. Niedrigere TopN decken insgesamt weniger Peptide und damit Proteine ab (vor allem bei komplexen Proben). Allerdings erlaubt ein niedriger TopN-Wert eine längere Injektionszeit pro MS2-Scan, wodurch mehr Ionen akkumuliert und qualitativ hochwertigere Fragmentierungsspektren erzeugt werden können. Dies kann die Identifikationswahrscheinlichkeit für einzelne Präkursorionen erhöhen und die Quantifizierungsgenauigkeit verbessern. Gleichzeitig wird die Gesamtzyklusdauer nicht zu lang. Es wurde ein Folgeversuch durchgeführt, bei dem TopN und maximale Injektionszeit (IT) variiert wurden (**Abbildung 13**).



**Abbildung 13:** Variation der Parameter Injektionszeit und TopN innerhalb der DDA-Methode. DDA, Data Dependent Acquisition; IT, Injection Time; MS2, Einstellungen des MS/MS-Scans.

Dabei stellten sich die Einstellungen IT 50 ms / TopN 15 und IT 100 ms / TopN 10 als gleichwertig heraus (553 und 561 identifizierte Peptide), während die Einstellungen IT 100 ms / TopN 15 und IT 50 ms / TopN 10 als weniger geeignet erwiesen (516 bzw. 368 identifizierte Peptide). Die festgestellte Anzahl identifizierter Peptide war so niedrig, da es sich um eine Gliadinprobe handelte, in der nicht so viele verschiedene detektierbare Peptide vorhanden sind. Um den TopN-Messzyklus zu verkürzen, wurde für die Standardmethode die DDA-Methode mit TopN = 10 und Maximum IT = 100 ms gewählt. Dies sollte

zum einen gewährleisten, dass genügend Messpunkte für jeden Analyten erfasst werden. Zum anderen sind akkuratere Fragmentierungsspektren von Vorteil bei der Entwicklung der targeted Proteomics-Methode. Alle anderen Parameter wurden von der Standardmethode übernommen.

### 3.1.6 Zusammenfassung

Auf der Grundlage bestehender Aufarbeitungsprotokolle sollte eine robuste und effiziente Bottom-Up Proteomics-Methode für die Analyse von Weizenmehlproteinen entwickelt und optimiert werden. Die Herausforderung bestand darin, einen Kompromiss zwischen einer möglichst hohen Zahl identifizierter Peptide und Zeitaufwand zu finden.

Im ersten Aufarbeitungsschritt wurde die Proteinextraktion systematisch untersucht. Die klassische dreistufige Osborne-Fraktionierung (Ex1) zeigte zwar die höchste Anzahl an identifizierten Peptiden und Glutenproteinen, ist jedoch sehr zeitintensiv und aufwändig. Eine alternative Extraktion mit einem weniger komplexen Extraktionsmittel (Ex7) erwies sich als nahezu gleichwertig hinsichtlich der Peptidanzahl, bei reduziertem Arbeitsaufwand. Allerdings wurde die wichtige Glutenfraktion der Glutenine nicht in dem Umfang analysiert wie bei den anderen Extraktionen. Dies verdeutlichte die Notwendigkeit, die Extraktionsstrategie auf die Zielproteine abzustimmen. Die Einfachextraktion Ex4 stellte den besten Kompromiss aus Effizienz und Analysequalität dar und wurde daher in der finalen Methode angewendet. Anschließend wurden weitere Parameter entlang der Probenaufarbeitung optimiert. Der Proteinverdau zeigte, dass ein Enzym-Substrat-Verhältnis von 1:25 die höchste Peptidausbeute erzielte. In Bezug auf die Proteinaufreinigung konnte durch den Wechsel von Röhrchen- auf Mikroplatten-SPE eine signifikante Zeitersparnis erzielt werden, ohne Verluste bei der Peptididentifikation. Versuche, die Proben vollständig in 96-Well-Mikroplatten aufzuarbeiten oder Trocknungsschritte wegzulassen, führten hingegen zu erheblichen Einbußen in der Peptidanzahl und wurden daher nicht weiterverfolgt. Innerhalb der DDA-Methode wurden die gewählten Standard-LC-MS/MS-Parameter überprüft. Es zeigten sich nur geringe Unterschiede zwischen den verschiedenen DDA-Einstellungen, sodass nur eine leichte Anpassung von TopN (10) und Injektionszeit (100 ms) notwendig war.

Zusammenfassend konnte eine effizientere Methode entwickelt werden, welche für große Probensätze besser geeignet ist. Sie basiert auf einer einstufigen Extraktion, einem optimierten Trypsinverhältnis (1:25), der Verwendung von 96-Well-Mikroplatten SPE sowie angepassten MS-Parametern (5.4.1). Die Methode bildete die Grundlage für die Proteomics-Analyse der BMWpop.

### 3.2 Untargeted Proteomics von Weizenmehl: Variation der Peptidasen

Teile dieses Kapitels wurden bereits bei der Fachzeitschrift *Current Research in Food Science* veröffentlicht [97].

Das Enzym Trypsin wird standardmäßig für den Verdau von Proteinen in Bottom-Up Proteomics-Aufarbeitungen verwendet. So wurde es bereits zur Untersuchung des Proteoms verschiedener Weizenarten eingesetzt [77]. Es handelt sich um eine sehr effiziente und spezifische Serin-Endopeptidase, die nach den basischen Aminosäuren Lysin und Arginin spaltet. Tryptische Peptide enthalten daher mindestens eine basische Aminosäure und sind für die ESI im positiven Modus gut geeignet, da sie leicht protoniert werden können [98].

In den Extraktionsexperimenten (3.1.1) wurde der Glutenproteintyp der  $\omega$ -Gliadine in keiner der getesteten Extraktionen identifiziert (**Abbildung 9**).  $\omega$ -Gliadine enthalten 20–30 % Prolin [99]. Prolin behindert den tryptischen Verdau, da es normalerweise nicht gespalten wird, wenn auf Lysin oder Arginin Prolin folgt. Zusätzlich wird eine vollständige Abdeckung der Proteinsequenz durch eine einzige Peptidase oftmals nicht erreicht. Dies wirft die Frage auf, ob die resultierenden tryptischen Peptide für die Analyse durch LC-MS/MS geeignet sind und wie umfangreich die Proteinabdeckung ist. Der Einsatz nur eines Verdauungsenzyms führt häufig zu einer unvollständigen und potenziell verzerrten Darstellung des Proteoms [78]. Daher besteht die Notwendigkeit, neben Trypsin auch andere Peptidasen oder Peptidasekombinationen zu testen.

Neben Trypsin stehen zahlreiche weitere proteolytische Enzyme als potenzielle Alternativen zur Verfügung, die auch schon bei Bottom-Up Proteomics von Weizenmehl eingesetzt wurden. Chymotrypsin wurde bereits häufiger im Zusammenhang mit Glutenproteinen verwendet [100-104]. Es spaltet nach Phenylalanin, Tryptophan und Tyrosin, und weniger spezifisch nach Leucin und Methionin. Wie Trypsin verhindert jedoch Prolin an der P1'-Position die Spaltung durch Chymotrypsin. Aufgrund seiner Spezifität bildet Chymotrypsin komplementäre Peptide zu Trypsin, was es sowohl für den alleinigen als auch für den kombinierten Verdau mit Trypsin interessant macht. Pepsin spaltet je nach pH-Wert an verschiedenen Stellen. Bei pH 1,3 spaltet es vor und nach Phenylalanin und Leucin. Bei pH-Werten >2 spaltet es zusätzlich vor und nach Tryptophan und Tyrosin. Da die Proteinextraktion in der Regel bei einem pH-Wert von 6–8 durchgeführt wird, stellt eine anschließende Ansäuerung einen zusätzlichen Verarbeitungsschritt dar. Pepsin wurde zur Untersuchung von Gersten- [100] und Hartweizenproteinen [105] verwendet und wird zusammen mit Trypsin und Chymotrypsin in Studien eingesetzt, die die menschliche Magen-Darm-Verdauung nachahmen wollen [106, 107]. Die Peptidase Thermolysin wurde auch schon zur Untersuchung

des Weizenproteoms [75, 108] und zur Aufklärung der intermolekularen Disulfidbindungen zwischen Glutenproteinen eingesetzt [109]. Thermolysin spaltet vor Alanin, Phenylalanin, Isoleucin, Lysin, Methionin und Valin, aber nicht nach Asparaginsäure und Glutaminsäure. Das bedeutet, dass es im Vergleich zu allen vier genannten Enzymen die höchste Anzahl potenzieller Spaltstellen aufweist, was ein Vorteil für Glutenproteine sein könnte.

Zur Untersuchung des Einflusses verschiedener Peptidasen auf das Weizenproteom wurden die acht Elternlinien der BMWpop sowie der Vergleichsstandard RGT Reform mit vier verschiedenen Peptidasen und einer Kombination verdaut: Trypsin, Chymotrypsin, Pepsin, Thermolysin und die Kombination aus Trypsin und Chymotrypsin (TC) (5.5). Da diese Versuchsreihe mit einem kleinen Probensatz durchgeführt wurde, wurde die Dreifachextraktion (modifizierte Osborne-Fraktionierung) angewendet, da sie die umfassendste Proteinextraktion ermöglichte (3.1.1). Ziel war es, einen Überblick über die Verdaulichkeit von Weizenproteinen, insbesondere von Glutenproteinen, mit verschiedenen Peptidasen zu geben. Außerdem sollte festgestellt werden, welche Peptidase für die Entwicklung einer targeted Proteomics-Methode am besten geeignet ist.

### 3.2.1 Anzahl identifizierter Peptide

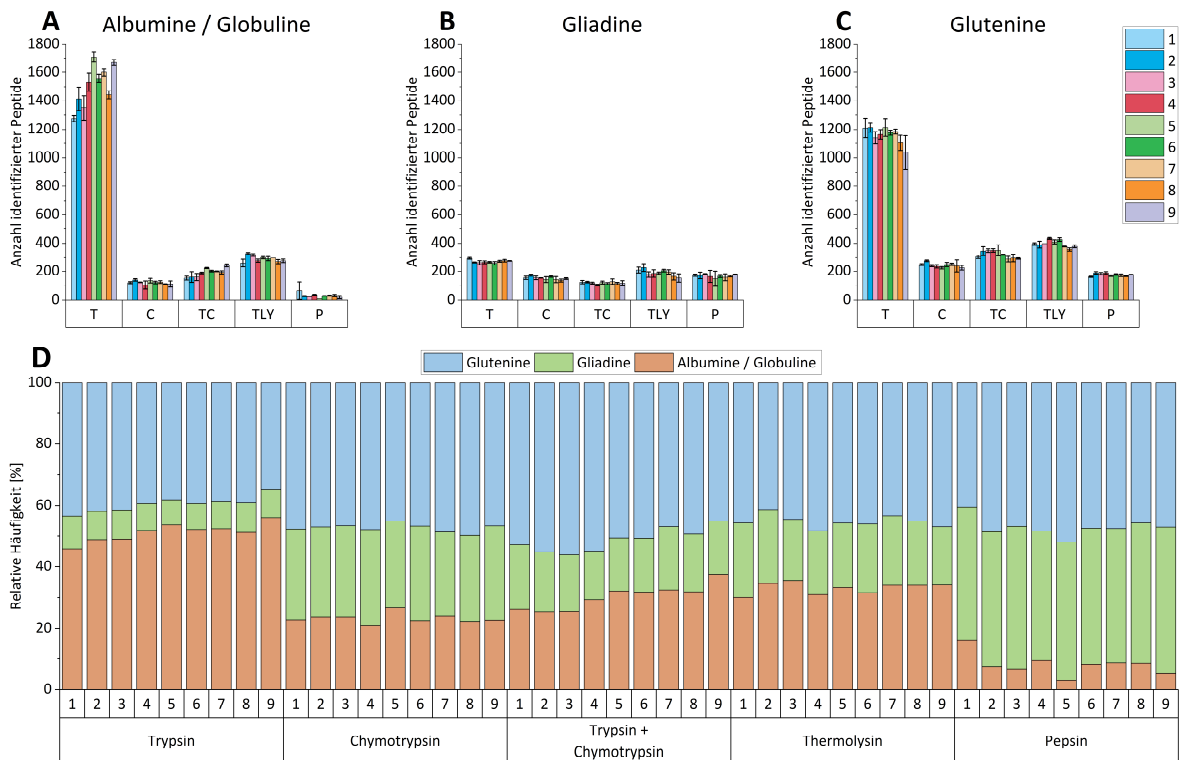
Die höchste Anzahl Peptide wurde nach dem Trypsinverdau in allen drei Extrakten E1-E3 identifiziert (**Abbildung 14 A-C**). Insbesondere in der Albumin/Globulin- und Gluteninfraktion gab es einen großen Unterschied zu den anderen Verdauungsenzymen. Nur in der Gliadinfraktion lieferten alle Enzyme ähnliche Ergebnisse.

Bei näherer Betrachtung der Albumin/Globulinfraktion (**Abbildung 14 A**) wurden nach dem Trypsinverdau 1250 bis 1700 Peptide in den Weizensorten identifiziert. Diese Fraktion zeigte zudem die größte Variation zwischen den neun untersuchten Weizensorten. Nach dem Thermolysinverdau wurden in den Weizensorten 250 bis 350 Peptide identifiziert, was weit weniger ist als nach dem Trypsinverdau. Der TC-Verdau ergab 150 bis 250 Peptide, gefolgt vom Chymotrypsinverdau mit 100 bis 130 Peptiden. Der Pepsinverdau führte nur zu 30 bis 60 Peptiden. Die Ergebnisse belegen eindeutig, dass Trypsin die effizienteste Spaltung der Proteine in identifizierbare Peptide ermöglicht und im Vergleich zu den anderen getesteten Enzymen etwa 1000 zusätzliche Peptididentifizierungen liefert. Pepsin schien für die Analyse der Albumin/Globulinfraktion nicht geeignet zu sein, da nur wenige Peptide (60) identifiziert wurden.

Bei der Gliadinfraktion (**Abbildung 14 B**) wurde mit allen Enzymen eine ähnliche Peptidanzahl erzielt. Insgesamt wurden jedoch nur 100 bis 300 Peptide identifiziert. Die meisten Peptide wurden mit Trypsin

nachgewiesen (250 bis fast 300). Thermolysin folgte mit 150 bis 200 Peptiden. Außerdem wurden 150 bis 180 Peptide mit Pepsin nachgewiesen. Etwa 130 bis 160 Peptide wurden mit Chymotrypsin identifiziert, während der TC-Verdau 100 bis 130 Peptide ergab. Da in der Gliadinfraktion mit jedem Enzym eine eher geringe Anzahl von Peptiden festgestellt wurde, scheint diese Fraktion „schwer verdaulich“ zu sein. Keines der Enzyme scheint für die Analyse der Gliadinfraktion gut geeignet zu sein. Dies könnte aber auch darauf hindeuten, dass nicht viele verschiedene Proteine in dieser Fraktion vorhanden sind.

Die Gluteninfraktion (**Abbildung 14 C**) zeigt ein ähnliches Muster wie die Albumin/Globulinfraktion. Nach dem Trypsinverdau wurden 1100 bis 1250 Peptide identifiziert. Nach dem Thermolysinverdau wurden 350 bis 440 Peptide nachgewiesen. Nach dem kombinierten TC-Verdau 290 bis 370 Peptide und nach dem Chymotrypsinverdau waren es etwa 190 bis 280. Pepsin wies auch hier mit 150 bis 200 Peptiden die geringste Anzahl von Identifizierungen auf. Wie bei der Albumin/Globulinfraktion scheint Trypsin für die Analyse der Gluteninfraktion am besten geeignet zu sein, da die Zahl der identifizierten Peptide wesentlich höher ist.



**Abbildung 14:** Gesamtzahl der identifizierten Peptide pro Weizensorte eins bis neun und verwendeter Peptidase in der Albumin-/Globulinfraktion (A), in der Gliadinfraktion (B) und in der Gluteninfraktion (C). Die relative Häufigkeit der identifizierten Peptide zwischen den drei Fraktionen ist in Feld D dargestellt. 1: Ambition; 2: FIRL3565; 3: Bussard; 4: Event; 5: Format; 6: Julius; 7: BAYP4535; 8: Potenzial; 9: RGT Reform; T, Trypsin; C, Chymotrypsin; TLY, Thermolysin; P, Pepsin.

Über alle neun Weizensorten und alle drei Fraktionen hinweg wurden mit Trypsin insgesamt 4115 verschiedene Peptide identifiziert. Mit Thermolysin waren es insgesamt 1421 Peptide, gefolgt von TC mit 1141, Chymotrypsin mit 910 und Pepsin mit 606 Peptiden. Zusammenfassend lässt sich sagen, dass Trypsin am besten geeignet war, um Weizenpeptide zu erhalten, die mittels LC-MS/MS identifiziert werden können. Alle anderen verwendeten Peptidasen haben eine höhere Anzahl möglicher Spaltstellen. Dies könnte zur Bildung kürzerer Peptide führen, die für die LC-MS/MS-Identifikation ungeeignet sind, da ihnen aufgrund ihrer Unspezifität eine eindeutige Zuordnung zum Ursprungprotein fehlt. Als optimal für die Analyse gilt daher eine Peptidlänge von etwa sieben bis 25 Aminosäuren. Zu kurze Peptide wurden bereits in chymotryptischen Verdauungen von Glutenproteinen beobachtet [100]. Selbst nach dem Trypsinverdau wurde berichtet, dass mehr als die Hälfte der Peptide weniger als sechs Aminosäuren lang und daher schwer zu identifizieren waren [78, 110]. Diese Erkenntnisse basierten jedoch auf Hefepeptiden und gelten nicht unbedingt für Weizenpeptide. Insbesondere die Peptide von Glutenproteinen enthalten oft repetitive Sequenzen, die hauptsächlich aus Prolin und Glutamin und nur wenigen Lysin- und Argininresten bestehen. Folglich gibt es nicht so viele Spaltungsmöglichkeiten für Trypsin. Die hier erzielten Ergebnisse bestätigen dies: Nach dem Trypsinverdau betrug die durchschnittliche Gluten-Peptidlänge 21 Aminosäuren. Nach dem Verdau mit Pepsin und Thermolysin waren die Peptide im Durchschnitt 14 und 15 Aminosäuren lang. Bei Chymotrypsin und TC lag die durchschnittliche Peptidlänge bei jeweils 16 Aminosäuren. Darüber hinaus wurden mit Trypsin die meisten Peptide mit einer Länge von mehr als 31 Aminosäuren gemessen: 59 im Gegensatz zu maximal 21 bei den anderen Peptidasen (Anhang, **Abbildung 32**). Die meisten Glutenpeptide wurden mit Thermolysin identifiziert, gefolgt von Chymotrypsin, TC, Trypsin und dann Pepsin. Dies steht im Einklang mit der Literatur, in der die meisten Peptide ebenfalls mit Thermolysin, gefolgt von Chymotrypsin und dann Trypsin nachgewiesen wurden [104]. Durch die sequentielle Verdauung mit Chymotrypsin und Trypsin sowie die Kombination dieser Ergebnisse mit denen einer sequentiellen Lys-C/Trypsin-Verdauung wurden insgesamt 434 Glutenpeptide aus Weizen identifiziert [111]. Dieses Resultat liegt in einem ähnlichen Bereich wie das des TC-Verdau, obwohl kein Lys-C eingesetzt wurde. Bemerkenswert ist, dass sich die optimale Reihenfolge der Peptidasen verändert, wenn ausschließlich die Anzahl der identifizierten Glutenpeptide betrachtet wird und nicht die Gesamtheit aller detektierten Peptide im Extrakt.

Ein weiterer wichtiger Punkt, auch für targeted Proteomics, ist die vollständige und reproduzierbare Spaltung der Proteine in Peptide, d. h. dass es keine missed cleavages (verpassten Spaltstellen) gibt. Von den tryptischen Peptiden wiesen 1210 missed cleavages auf, was einem Anteil von 29 % entspricht. Dies zeigt, dass Trypsin nicht nur zu den weitaus meisten Peptididentifikationen führt, sondern im Vergleich zu

den anderen Peptidasen unter den gewählten Bedingungen auch am vollständigsten zu spalten scheint. Dies bestätigt, dass Trypsin eine sehr effiziente und spezifische Peptidase ist [78]. Nach dem Thermolysinverdau zeigten 83 % der Peptide missed cleavages. Mit 47 % wies fast die Hälfte aller Peptide nach dem kombinierten TC-Verdau missed cleavages auf und 39 % der Peptide wurden beim Verdau mit Chymotrypsin nicht vollständig gespalten. Dass Chymotrypsin nicht so spezifisch spaltet wie Trypsin, wurde ebenfalls in der Literatur festgestellt [100]. Nach dem Pepsinverdau enthielten 86 % der Peptide missed cleavages, was ebenfalls auf die vielen Spaltstellen von Pepsin und die Tatsache zurückzuführen sein könnte, dass es stark pH-abhängig spaltet. Je unspezifischer das Enzym ist, desto mehr Spaltstellen gibt es, aber desto größer ist auch die Wahrscheinlichkeit von missed cleavages [98]. Die hier erzielten Ergebnisse bestätigen dies.

Wie in **Abbildung 14 D** dargestellt, variiert die relative Häufigkeit der Peptididentifizierungen zwischen den Fraktionen je nach Peptidase. Nach dem Trypsinverdau wurden 46–56 % der Peptide in der Albumin/Globulinfraktion nachgewiesen, während 35–44 % in der Gluteninfraktion identifiziert wurden. Auffallend ist, dass nur 9–11 % der Peptide in der Gliadinfraktion enthalten waren. Wie bereits erwähnt, ist dies wahrscheinlich auf das Vorhandensein ähnlicher Proteine in dieser Fraktion und den Mangel an geeigneten Spaltstellen für Trypsin zurückzuführen. Der Verdau mit Chymotrypsin, TC und Thermolysin zeigte eine ähnliche Verteilung: Die meisten Peptide (41–46 %) wurden in der Gluteninfraktion identifiziert. Ein ähnlicher Anteil an Peptiden wurde sowohl in der Albumin/Globulin- (21–37 %) als auch in der Gliadinfraktion (16–31 %) nachgewiesen, wobei es leichte Unterschiede zwischen den Peptidasen gab. Im Gegensatz dazu wurde in den Gliadin- (42–47 %) und Gluteninfraktionen (40–48 %) nach Pepsinverdau eine ähnliche Anzahl von Peptiden identifiziert, aber weit weniger in der salzlöslichen Fraktion (3–16 %). Daher ist Pepsin unter den gewählten Bedingungen nicht gut geeignet, die Proteine in der Albumin/Globulinfraktion zu verdauen.

### 3.2.2 Anzahl identifizierter Proteingruppen

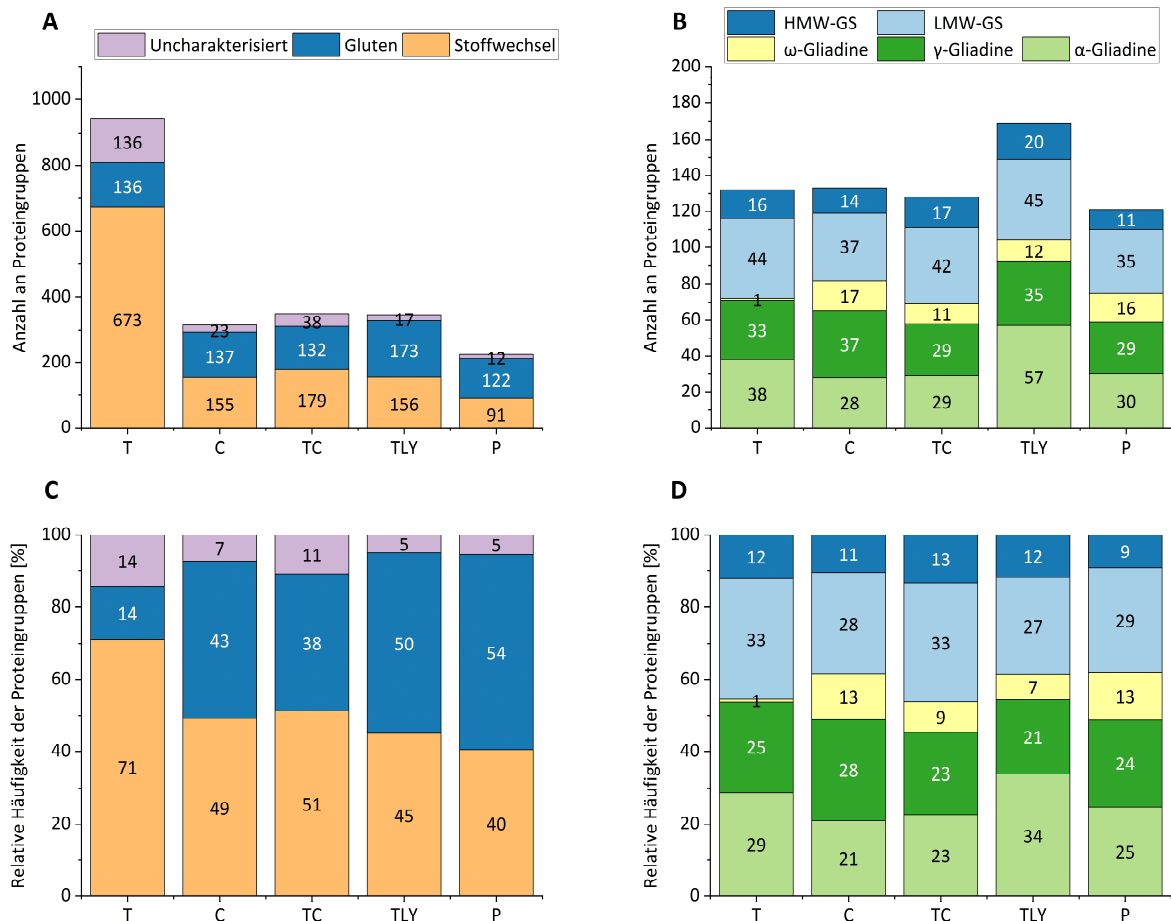
Die dreistufige Extraktion wurde durchgeführt, um eine umfassendere Abdeckung des Proteoms zu erreichen als bei der Einfachextraktion. Daher sollten in jeder Fraktion unterschiedliche Proteine angereichert werden. Eine detaillierte Analyse der Albumin-/Globulinfraktion nach Trypsinverdau zeigte jedoch, dass 31 der 701 identifizierten Proteingruppen den Gliadinen und 18 den Gluteninen zugeordnet werden konnten. Dies deutet darauf hin, dass Gliadine und Glutenine bereits während des ersten Extraktionsschritts extrahiert werden. Wie zu erwarten war, enthielt die Gliadinfraktion überwiegend monomere Gliadine. Darüber hinaus konnten nach Trypsinverdau auch 30 Glutenine in dieser Fraktion

identifiziert werden. Wenn man bedenkt, dass insgesamt 60 Gluteninproteingruppen identifiziert wurden, ist die Hälfte davon bereits in dieser Fraktion enthalten. Es wurde schon berichtet, dass monomere Glutenine auch mit wässrigen Alkohollösungen aus dem Mehl extrahiert werden können [112]. Die Gluteninfraktion ist primär auf die Extraktion großer, polymerer Glutenine ausgerichtet. Da hierbei ein starkes Extraktionsmittel zum Einsatz kommt und es sich um den dritten Schritt der Fraktionierung handelt, enthält diese Fraktion auch zahlreiche weitere Proteinarten, die in den vorherigen Extraktionsschritten nicht gelöst wurden. Alle drei Fraktionen überschneiden sich also in ihren Proteingruppen, was mit früheren Berichten übereinstimmt [113]. Darüber hinaus führte selbst eine weitere Trennung durch präparative HPLC immer noch nicht zu reinen Proteinfractionen, die nur bestimmte Proteingruppen enthielten [114].

Daher wurden alle drei Fraktionen der mehrstufigen Extraktion gemeinsam ausgewertet, insbesondere im Hinblick auf die quantitative Auswertung (5.5.3). Insgesamt wurden nach Trypsinverdau 945 Proteingruppen in den neun Weizensorten nachgewiesen (**Abbildung 15 A**). Erwartungsgemäß war dies wiederum der mit Abstand höchste Wert im Vergleich zu den anderen enzymatischen Hydrolysaten. Nach dem kombinierten TC-Verdau wurden 349 Proteingruppen identifiziert, nach Thermolysin- und Chymotrypsinverdau 346 bzw. 315 Proteingruppen. Mit Pepsin konnten nur 225 Proteingruppen zugeordnet werden, was sich durch die geringe Anzahl der identifizierten Peptide erklären lässt. Wie schon bei den Peptididentifizierungen zeigt sich auch hier der große Einfluss der Peptidase.

Da die Proteingruppen nicht mehr in die drei Fraktionen nach Löslichkeit unterteilt wurden, wurden sie nach ihren Funktionen in Stoffwechsel-, Gluten- und uncharakterisierte Proteingruppen unterteilt (**Abbildung 15 A**). Die deutlichsten Unterschiede zwischen den Peptidasen zeigten sich bei den Stoffwechselproteingruppen: Nach Trypsinverdau konnten 673 dieser Proteingruppen identifiziert werden, während die anderen Enzyme bzw. die Enzymkombination lediglich 91 bis 179 Gruppen lieferten. Bei den am häufigsten vorkommenden Proteinen im Weizen, den Glutenproteinen, wurden mit jedem Enzym etwa 130 Glutenproteingruppen identifiziert. Nur Thermolysin sticht mit 169 identifizierten Glutenproteingruppen hervor. Zwischen 50 und 100 Glutenproteine kommen in einer einzigen Weizensorte vor [17]. Vensel et al. identifizierten in der Weizensorte Butte 86 insgesamt 52 Glutenproteine durch die kombinierte Auswertung von Verdauungen mit Trypsin, Chymotrypsin und Thermolysin [104]. Da es sich um eine gemeinsame Auswertung von neun Weizensorten handelt, stimmen diese Ergebnisse mit der Literatur überein. Es ist zu beachten, dass einige Proteingruppen aufgrund fehlender Informationen in der Weizenproteindatenbank uncharakterisiert bleiben. Bei der Trypsin-Verdauung waren 136 Proteingruppen betroffen (14 % der Gesamtmenge, **Abbildung 15 C**). Weitere 14 %

aller tryptischen Proteingruppen gehören zu den Glutenproteingruppen, während die Mehrheit (71 %) Stoffwechselproteingruppen sind. Bei allen anderen Peptidasen ist die relative Verteilung sehr ähnlich: bei 40–51 % und 43–54 % handelt es sich um Stoffwechselproteingruppen bzw. Glutenproteingruppen. Die übrigen Proteingruppen (5–11 %) waren nicht charakterisiert. Im Vergleich zu Glutenproteinen verfügen Stoffwechselproteine über ein günstigeres Aminosäureprofil, welches eine größere Anzahl potenzieller Spaltstellen aufweist und sie dadurch besser für die Analyse mittels Bottom-up-Proteomics geeignet macht [115]. Interessanterweise zeigt sich, dass Trypsin allein bereits eine hohe Effizienz bei der Analyse von Stoffwechselproteinen aufweist. In der Literatur wurde ebenfalls berichtet, dass Trypsin diese Proteine weitgehend abdeckt und ein zusätzlicher Tandemverdau mit Glu-C nur marginale Verbesserungen erzielte [116].



**Abbildung 15:** Absolute (A) und relative (C) Häufigkeit von Stoffwechsel-, Gluten- und uncharakterisierten Proteingruppen. Absolute (B) und relative (D) Häufigkeit spezifischer Glutenproteingruppen. Die Ergebnisse für alle neun Weizenlinien wurden für jedes verwendete Verdauungsenzym zusammengefasst. Die Proteingruppen wurden mit einer False Discovery Rate von 1 % identifiziert. T, Trypsin; C, Chymotrypsin; TLY, Thermolysin; P, Pepsin; HMW-GS, hochmolekulare Gluteninuntereinheit; LMW GS, niedermolekulare Gluteninuntereinheit.

Im weiteren Verlauf wurden die zugehörigen Proteintypen innerhalb der Glutenproteingruppen genauer betrachtet (**Abbildung 15 B & D**). Jeder Glutenproteintyp wurde mit jedem Enzym identifiziert. Nach dem Trypsinverdau wurden insgesamt 72 verschiedene Gliadinproteingruppen identifiziert. Nach dem Chymotrypsinverdau waren es 82 und nach dem kombinierten Verdau 69. 104 Gliadinproteingruppen wurden nach Thermolysin- und 75 nach Pepsinverdauung gefunden. Ausgehend von der Gesamtzahl der identifizierten Gliadinproteingruppen scheint es, dass nach einem Chymotrypsin- oder Thermolysinverdau mehr identifiziert oder differenziert (und somit in mehr verschiedene Gruppen eingeteilt) werden können. Dies könnte ein Vorteil sein, wenn diese Fraktion eingehender untersucht werden soll.

Innerhalb der Gliadinfraktion schwankte die Zahl der den  $\alpha$ -Gliadinen zugeordneten Proteingruppen zwischen 28 und 57. Insbesondere Thermolysin scheint für die Untersuchung von  $\alpha$ -Gliadinen gut geeignet zu sein, da die meisten Proteingruppen (57) differenziert werden konnten. Die Zahl der  $\gamma$ -Gliadinproteingruppen schwankte zwischen 29 und 37 und lag damit für alle Peptidasen in einem ähnlichen Bereich. Die meisten Proteingruppen (37) wurden nach dem Chymotrypsinverdau nachgewiesen.  $\alpha$ - und  $\gamma$ -Gliadine sind in etwa gleichen Anteilen am Gesamtglutengehalt vorhanden [13]. Auch ihre relative Häufigkeit zeigte eine ähnliche Anzahl an Proteingruppen (**Abbildung 15 D**). Die Identifizierung von  $\omega$ -Gliadinproteingruppen stellte den größten Unterschied zwischen den Peptidasen dar. Mit Trypsin konnte nur eine Proteingruppe den  $\omega$ -Gliadinen zugeordnet werden.  $\omega$ -Gliadine bestehen bis zu 80 % aus den Aminosäuren Glutamin, Prolin und Phenylalanin [16]. Da Trypsin selektiv nach Lysin und Arginin spaltet, gibt es möglicherweise nicht genügend Spaltstellen innerhalb der  $\omega$ -Gliadine. Infolgedessen sind die Peptide zu lang, um mit LC-MS/MS identifiziert zu werden. Alle anderen Peptidasen, die verwendet wurden, spalten ebenfalls vor oder nach Phenylalanin, was zu geeigneten  $\omega$ -Gliadin-Peptiden führt, die mit LC-MS/MS nachgewiesen werden können. Die  $\omega$ -Gliadine machen nur etwa 10 % der Glutenproteine aus [10, 11], was der Grund dafür sein könnte, dass insgesamt nicht so viele verschiedene  $\omega$ -Gliadine gefunden werden.  $\omega$ -Gliadine machen folglich nur einen kleinen Anteil der Glutenproteine aus und wurden bislang nicht mit Eigenschaften wie der Backqualität von Weizen in Verbindung gebracht. Dennoch erscheint für die Analyse der  $\omega$ -Gliadine im Bottom-Up Proteomics der Einsatz alternativer Verdauungsenzyme wie Chymotrypsin geeigneter als Trypsin.

Innerhalb der Gluteninproteine wiesen die LMW-GS eine größere Variation auf als die HMW-GS [117]. Hier konnten mit jeder der eingesetzten Peptidasen zwischen 35 und 45 individuelle Proteingruppen den LMW-GS zugeordnet werden. Die höchste Anzahl (45) wurde nach dem Verdau mit Thermolysin identifiziert,

wobei die Unterschiede zwischen den Peptidasen insgesamt gering ausfielen. Im Gegensatz zu den LMW-GS (7–16) sind nur 3–5 einzelne HMW-GS pro Weizensorte vorhanden [117]. Es wurden nicht so viele verschiedene HMW-GS nachgewiesen, sondern insgesamt 11-20 innerhalb der neun Weizenlinien. Wiederum wurden die meisten Proteingruppen (20) nach dem Thermolysinverdau nachgewiesen. HMW-GS enthalten einen relativ hohen Anteil an Glycin in den repetitiven Regionen [16]. Dies vereinfacht grundsätzlich die Zuordnung von Peptiden zu bestimmten Proteinen. Außerdem kommen polare Aminosäuren wie Lysin häufiger in den nicht-repetitiven Regionen vor [17]. Dies ist ein Vorteil bei der Verwendung von Trypsin, da es selektiv nach Lysin spaltet. Dies schlug sich jedoch nicht in der Anzahl identifizierter Proteingruppen nieder, da mit allen eingesetzten Enzymen vergleichbare Ergebnisse erzielt wurden. Insgesamt stehen die hier gewonnenen Erkenntnisse im Einklang mit Literaturangaben, in denen etwa 20 verschiedene HMW-GS und über 40 unterschiedliche LMW-GS in Weizen beschrieben sind [118].

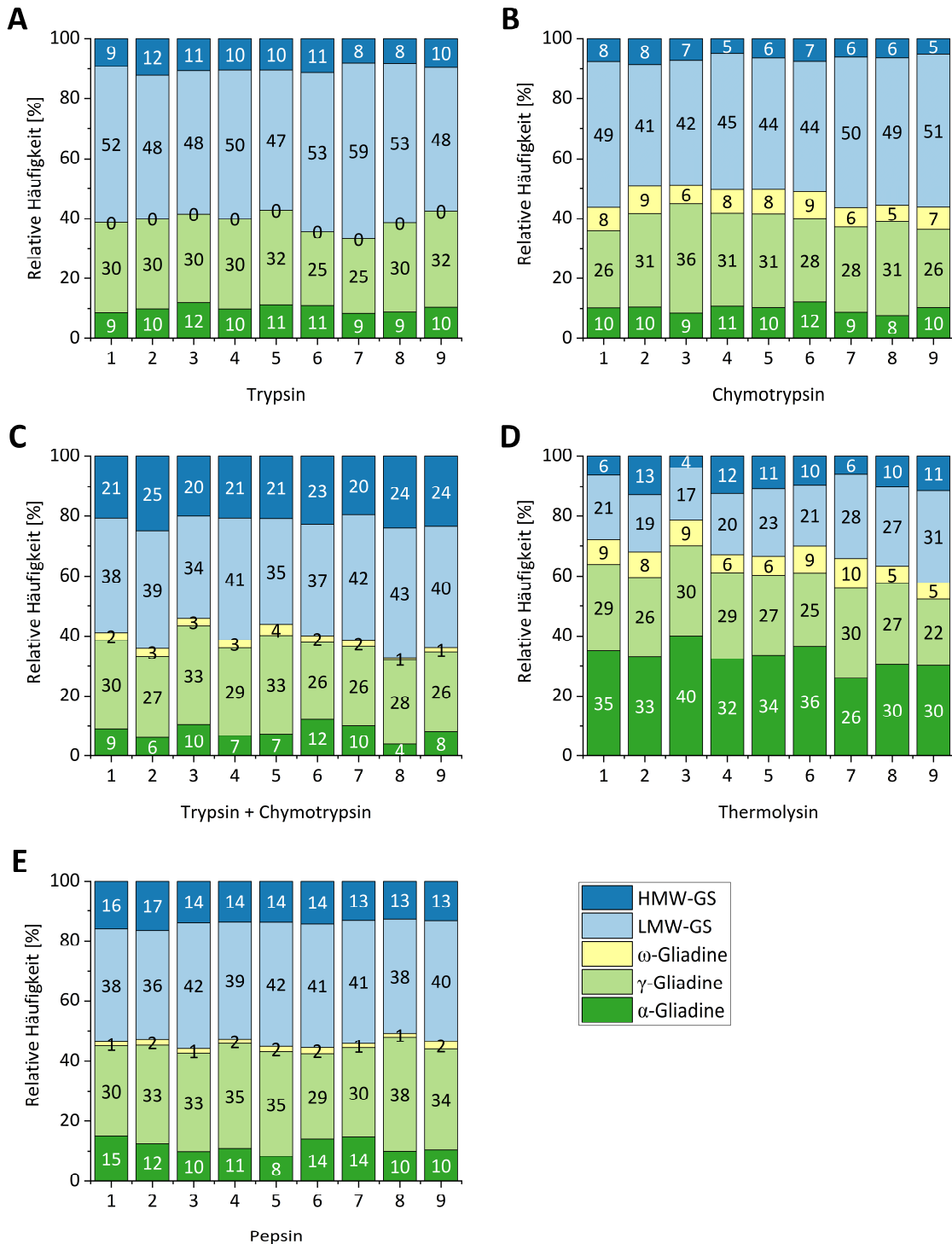
### 3.2.3 Vergleich der Glutenproteinzusammensetzung

Die Gesamtmenge der Glutenproteine in einem Mehl kann mit HPLC und einem externen Standard bestimmt werden [19, 119]. Daher wurden die LFQ-Intensitäten aller Proteingruppen innerhalb des jeweiligen Glutenproteintyps summiert und die jeweiligen prozentualen Anteile bestimmt (**Abbildung 16**). Das Muster variierte in Abhängigkeit von der Peptidase. Das Verteilungsmuster mit Trypsin, Chymotrypsin, TC und auch Pepsin sah sehr ähnlich aus: ca. 10 %  $\alpha$ -Gliadine und 30–35 %  $\gamma$ -Gliadine waren vorhanden, gefolgt von einem kleineren Anteil (0–10 %)  $\omega$ -Gliadine. Nur ein  $\omega$ -Gliadin wurde nach dem Trypsinverdau identifiziert. Allerdings hatte es eine LFQ-Intensität von Null, so dass es nicht berücksichtigt wurde. In der Literatur wurde ebenfalls berichtet, dass nach dem Trypsinverdau keine  $\omega$ -Gliadin-Peptide nachgewiesen wurden [104]. Die LMW-GS und HMW-GS waren zu 40–50 % und 10–20 % vorhanden. Nur der Thermolysinverdau zeigte ein abweichendes Muster, das sich vor allem in proportional mehr  $\alpha$ -Gliadinen (26–40 %) und weniger LMW-GS (17–31 %) widerspiegelt. Nach dem Thermolysinverdau wurden geringfügig weniger  $\gamma$ -Gliadine bestimmt (22–30 %). Die  $\omega$ -Gliadine lagen zwischen 5 % und 10 % und die HMW-GS zwischen 4 % und 11 %. Dies zeigt, dass die Häufigkeit stark von der Peptidase abhängt. Daraus folgt auch, dass der Vergleich von LFQ-Daten verschiedener Peptidasen nicht möglich ist. Stattdessen können nur Datensätze verglichen werden, die mit der gleichen Peptidase gewonnen wurden. Zu diesem Schluss kamen auch Giansanti et al., die das menschliche Phosphoproteom mit mehreren Proteasen untersuchten [120]. Demnach zeigt Trypsin jedoch ein verzerrtes Bild der Glutenproteine, da z. B. die  $\omega$ -Gliadine nach Trypsinverdau nicht gefunden wurden. Auf diese Proteomverzerrung wurde bereits in anderen Studien hingewiesen [78, 121, 122]. Selbst bei absoluten Quantifizierungsansätzen weisen dieselben Peptide, die von verschiedenen Peptidasen produziert werden, eine Verzerrung auf. Dies wurde

auf Unterschiede bei den Verdauungsbedingungen, der Proteinstruktur und der Peptidase selbst zurückgeführt [123].

Nach der RP-HPLC-UV-Analyse (Quantifizierung der Glutenfraktionen mittels RP-UHPLC) lagen die  $\alpha$ -,  $\gamma$ - und  $\omega$ -Gliadine zwischen 25–32 %, 23–29 % bzw. 7–11 % des gesamten Glutens. LMW-GS und HMW-GS bewegten sich zwischen 22–30 % und 8–12 % (Anhang, **Abbildung 33**). Dies stimmt mit der Literatur überein, wo  $\alpha$ -,  $\gamma$ - und  $\omega$ -Gliadine 35 %, 21 % bzw. 11 % des Gesamtglutens ausmachen [99]. Nur die Ergebnisse von Thermolysin entsprechen in etwa dieser Verteilung (**Abbildung 16 D**). Insbesondere bei den  $\alpha$ -Gliadinen gab es einen großen Unterschied zwischen Thermolysin und den anderen Peptidasen. Nach dem Thermolysinverdau konnten mehr  $\alpha$ -Gliadinproteingruppen unterschieden werden, was sich hier in einem höheren Anteil an  $\alpha$ -Gliadinen niederschlägt.

Nach Scherf [99] haben die HMW-GS einen Anteil von 9 % und die LMW-GS einen Anteil von 24 %, was auch mit den hier erzielten RP-HPLC-UV-Ergebnissen übereinstimmt. Die Ergebnisse der LC-MS/MS von Thermolysin stimmten damit überein. Für die HMW-GS lagen die Ergebnisse der meisten Peptidasen in einem ähnlichen oder leicht höheren Bereich. Die LMW-GS wiesen die größten Unterschiede auf, da der Anteil mit 40–50 % mehr als doppelt so hoch war (die einzige Ausnahme war Thermolysin mit LMW-GS-Anteilen von 20–30 %). Da einige LMW-GS den  $\gamma$ - und  $\omega$ -Gliadinen ähnlich sind, ist es möglich, dass Peptide, die ursprünglich zu den  $\gamma$ - und  $\omega$ -Gliadinen gehörten, den LMW-GS zugeordnet wurden. Der Grund dafür könnte in den kurzen N- und C-terminalen Resten der  $\omega$ -Gliadine liegen, die nur 10 bis 20 Aminosäuren enthalten [17]. Wenn ein Enzym ein charakteristisches terminales Peptid nicht spaltet, wird das entsprechende Protein nicht gefunden. Dies gilt insbesondere für tryptische Peptide. Dies spiegelt sich auch in der relativen Menge der LMW-GS wider, die bei Trypsin am höchsten war (**Abbildung 16 A**). Allerdings werden die verschiedenen Glutenproteintypen bei HPLC-Analysen nur über ihre Retentionszeit zugeordnet und nicht mit einer genauen Identifizierung wie bei LC-MS/MS. Folglich kann die HPLC-Analyse nicht zwischen  $\alpha$ -/ $\gamma$ -Gliadinen und LMW-GS unterscheiden. Wie zuvor erläutert, wurden Glutenine bereits in der Gliadinfraktion nachgewiesen, was darauf hinweist, dass sie dort in gewissem Umfang mitextrahiert werden. In der HPLC-Analyse lassen sich diese nicht eindeutig von Gliadinen unterscheiden und werden daher dieser Fraktion zugeordnet. Entsprechende Überschneidungen traten auch in den beiden anderen Fraktionen auf, in denen mittels LC-MS/MS ebenfalls verschiedene Glutenproteine identifiziert wurden.



**Abbildung 16:** Relative Häufigkeit der Glutenproteine in den untersuchten Weizensorten auf der Grundlage der LFQ-Intensitäten der zugehörigen Proteingruppen. Die Abundanz variiert je nach Peptidase, darunter Trypsin (A), Chymotrypsin (B), Trypsin + Chymotrypsin (C), Thermolysin (D) und Pepsin (E). LFQ, labelfreie Quantifizierung; 1: Ambition; 2: FIRL3565; 3: Bussard; 4: Event; 5: Format; 6: Julius; 7: BAYP4535; 8: Potenzial; 9: RGT Reform; HMW-GS, hochmolekulare Gluteninuntereinheit; LMW GS, niedermolekulare Gluteninuntereinheit.

Mit Ausnahme von Thermolysin stellte die Gluteninfraktion bei allen anderen Peptidasen mit jeweils rund 60 % den größten Anteil der identifizierten Proteine dar. Dies widerspricht den Angaben von Wieser et al., die für Winterweizen einen Anteil von lediglich 35 % berichten. [124]. In einer anderen Studie, in der die Proteine mittels Osborne-Fraktionierung extrahiert und mittels HPLC bestimmt wurden, machen die Glutenine fast 30 % der Glutenproteine in Weichweizen aus [11]. Das ist die Hälfte des hier berechneten Gluteninanteils. Das Verhältnis von Gliadin zu Glutenin lag bei allen Peptidasen zwischen 0,5 und 1,0, außer bei Thermolysin, wo es zwischen 1,4 und 3,6 lag. Dies liegt im selben Bereich wie die HPLC-Ergebnisse in der Literatur, die z. B. 2,0-3,2 oder 2,1-4,1 betragen [10, 11]. Folglich bleibt von Interesse, welche der ermittelten Ergebnisse die tatsächliche Proteinverteilung im Weizenkorn – sowohl qualitativ als auch quantitativ – am besten widerspiegeln. Insgesamt zeigten sich auch deutliche Unterschiede zwischen den untersuchten Weizenlinien, sowohl in der Zusammensetzung der Proteingruppen als auch in deren relativen Anteilen. Diese natürlichen Variationen bieten ein wertvolles Potenzial für den gezielten Einsatz in züchterischen Programmen [85, 125].

Zusammengenommen unterstreichen die Ergebnisse die Herausforderungen bei der Bottom-Up Proteomicsanalyse von Weizenglutenproteinen. Sie enthalten stark repetitive Sequenzabschnitte, die entweder nicht genügend Spaltstellen aufweisen und/oder zu ähnlichen Peptiden führen, die sich nicht eindeutig einzelnen Proteinen zuordnen lassen. Dies lässt vermuten, dass einige der Peptide nicht ihren ursprünglichen Proteinen, sondern Proteinen anderer Glutenproteinfraktionen zugeordnet wurden. Dies würde zu falsch positiven Ergebnissen führen. Aber auch die HPLC-Analysen liefern keine exakten Ergebnisse, da die Glutenproteine in allen drei Fraktionen mit LC-MS/MS identifiziert wurden.

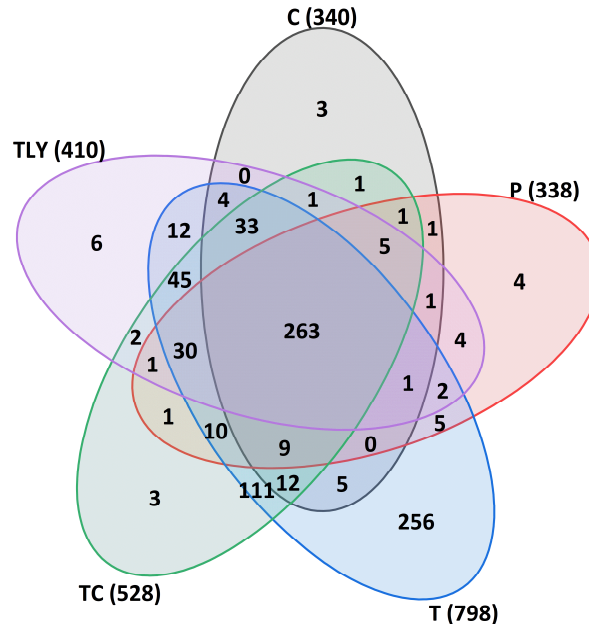
Weiterhin führte die LC-MS/MS-Analyse zu sehr unterschiedlichen Ergebnissen zwischen den Peptidasen. Im Allgemeinen blieben die Werte für  $\omega$ -Gliadine und HMW-GS über alle Peptidasen und Auswertungen hinweg am konstantesten. Dies ist möglicherweise auf ihren charakteristischen Phenylalanin- ( $\omega$ -Gliadine) bzw. Glycingehalt (HMW-GS) zurückzuführen, der ihre eindeutige Zuordnung erleichtert [17]. Der prinzipielle Vorteil von Proteomics – die präzise Identifizierung einzelner Moleküle – lässt sich bei Glutenproteinen aufgrund ihrer hohen Sequenzhomologie nur eingeschränkt nutzen. Aus diesem Grund wird im folgenden Abschnitt exemplarisch eine Weizensorte detailliert betrachtet, um zu untersuchen, inwieweit sich die eingesetzten Peptidasen hinsichtlich ihrer Ergebnisse ergänzen oder unterscheiden.

### 3.2.3.1 RGT Reform: Proteinidentifizierungen und Sequenzabdeckung

In diesem Abschnitt wird nur die Weizensorte RGT Reform als Vergleichsstandard für die Elternlinien der BMWpop näher untersucht [38]. Zu diesem Zweck wurden die Messungen aller Peptidasen in einem MaxQuant-Lauf ausgewertet. Zunächst wurde verglichen, welche Proteingruppen den einzelnen Peptidasen zugeordnet sind und inwieweit sie sich überschneiden (**Abbildung 17**). Insgesamt wurden 798 Proteingruppen mit Trypsin identifiziert, während es 528, 410, 340 und 338 Proteingruppen mit TC, Thermolysin, Chymotrypsin bzw. Pepsin gab. Von diesen Proteingruppen wurden 263 mit allen Peptidasen identifiziert. 256 Proteingruppen wurden nur mit Trypsin identifiziert. Bei allen anderen wurden nur drei bis sechs Proteingruppen mit nur einer Peptidase nachgewiesen.

Bei den Überschneidungen zwischen den Enzymen stachen Trypsin und TC mit 111 gemeinsamen Proteingruppen besonders hervor. Dies war zu erwarten, da Trypsin auch in der Kombination verwendet wird. Die Tatsache, dass Chymotrypsin und TC nur eine gemeinsame Proteingruppe aufweisen, zeigt, dass Trypsin der wichtigere Akteur in der Kombination ist. Trypsin, TC und Thermolysin hatten 45 Proteingruppen gemeinsam. Vergleicht man Chymotrypsin mit Trypsin, TC und Thermolysin, so gab es 33 gemeinsame Proteingruppen, während es bei Pepsin 30 waren. Alle anderen Überschneidungen waren sehr gering (0–12). Bei den Proteingruppen, die nur mit Trypsin gefunden wurden, handelte es sich fast ausschließlich um Stoffwechselproteine, aber auch um einige nicht charakterisierte Proteine. Das Gleiche gilt für die 111 Proteingruppen, die mit Trypsin und TC nachgewiesen wurden. Die Ergebnisse bestätigen erneut, dass Trypsin das Glutenproteom qualitativ am umfassendsten abbildet. Die übrigen Peptidasen lieferten kaum zusätzliche Identifizierungen von Proteingruppen. Eine Ausnahme stellten die  $\omega$ -Gliadine dar, die mit allen Enzymen außer Trypsin detektiert wurden. Alle anderen Glutenproteintypen konnten unabhängig vom eingesetzten Enzym nachgewiesen werden. Dies spiegelt sich auch in den zehn am häufigsten identifizierten Proteingruppen wider, die jeweils mit den unterschiedlichen Peptidasen erhalten wurden (Anhang, **Tabelle 12**). Dabei handelte es sich überwiegend um Glutenproteine mit Anteilen zwischen 2,3 % und 24,1 %. Im Fall von Trypsin gehörten vier der zehn häufigsten Proteine zu den Stoffwechselproteinen, hauptsächlich Amylase/Trypsin-Inhibitoren (ATIs), mit Anteilen von 1,9–3,4 %. Auch bei Thermolysin und Pepsin fanden sich jeweils drei Stoffwechselproteine unter den zehn häufigsten, mit Anteilen von 2,0–7,3 %. Bei Chymotrypsin und TC hingegen wurde jeweils nur ein Stoffwechselprotein häufiger identifiziert; mit Anteilen von 13,8 % bzw. 2,9 %. Auffällig ist, dass  $\omega$ -Gliadine mit Anteilen zwischen 10,1 % und 15,1 % bei allen Peptidasen – außer Trypsin – zu den zehn häufigsten Proteinen zählten. Dies unterstreicht erneut die bessere Eignung dieser Peptidasen für die Identifizierung dieses

spezifischen Glutenproteintyps. Im Gegensatz dazu scheint Trypsin besonders gut für die Identifizierung von ATIs geeignet zu sein, wo es bereits erfolgreich in der gezielten Proteomik eingesetzt wurde [126, 127].



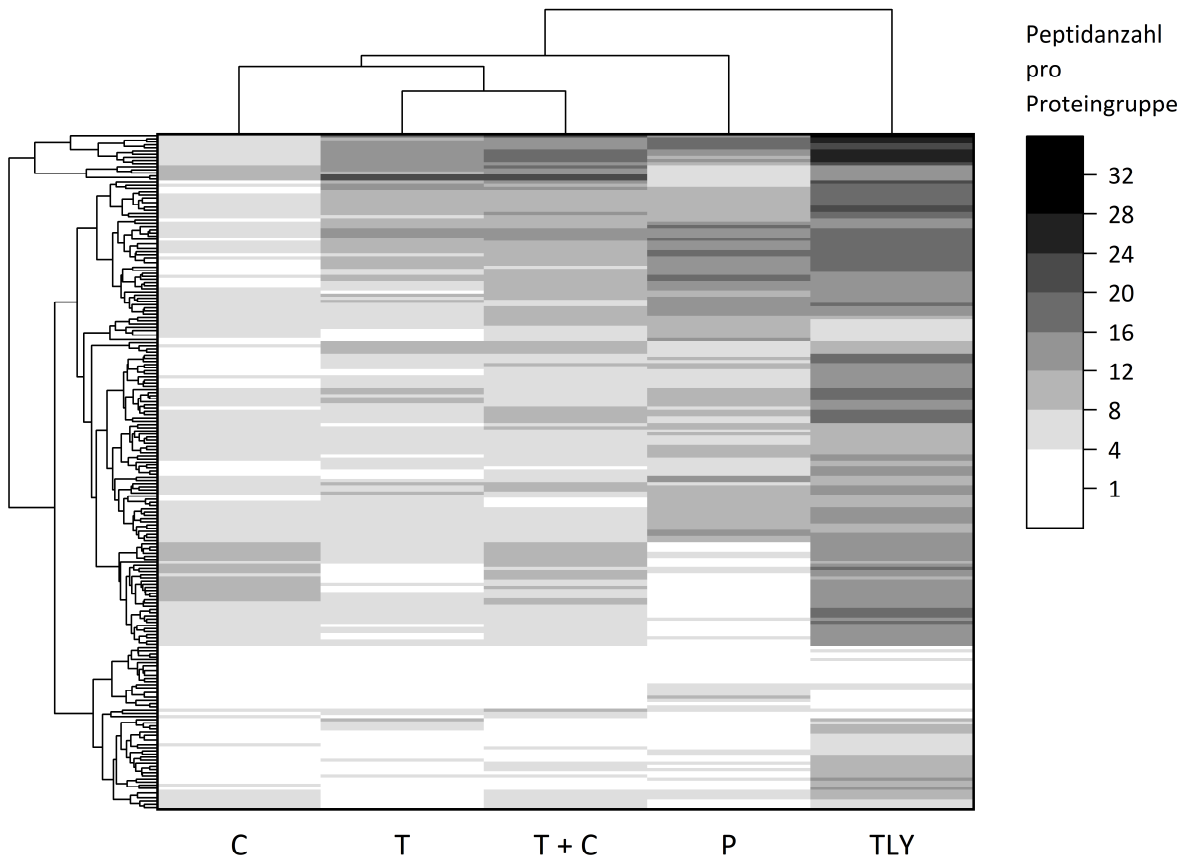
**Abbildung 17:** Anzahl der Proteingruppen in der Weizensorte RGT Reform, die bei den fünf verschiedenen Peptidasen einzigartig oder gemeinsam sind. Die Zahl in Klammern gibt die Gesamtzahl der bei der jeweiligen Peptidase gefundenen Proteingruppen an. T, Trypsin; C, Chymotrypsin; TLY, Thermolysin; P, Pepsin.

Eine hohe Spezifität der Peptidase führt zu mehr Proteinidentifizierungen. Peptidasen, die nicht so spezifisch sind, wie z. B. Chymotrypsin, können von einem Vorverdau mit Trypsin profitieren, was zu mehr Proteinidentifizierungen führt [128]. Dies konnte hier mit dem TC-Verdau nicht bestätigt werden. Ein Grund dafür könnte sein, dass der TC-Verdau nicht nacheinander, sondern gleichzeitig stattfand.

Die durchschnittliche Sequenzabdeckung der Glutenproteingruppen betrug 61,3 %. Bei den Versuchen, bei denen nur ein Enzym bewertet wurde, betrug die durchschnittliche Sequenzabdeckung für Thermolysin 51,1 %. Für TC, Chymotrypsin und Pepsin lag sie bei 37,2 %, 35,1 % bzw. 30,0 %. Trypsin ergab nur eine Abdeckung von 27,3 %. In der Literatur wurde die höchste Sequenzabdeckung ebenfalls für Thermolysin gefunden (zwischen 31–49 %), gefolgt von Chymotrypsin (16–40 %) und Trypsin (0–28 %). Nur bei den  $\gamma$ -Gliadinen schnitt Chymotrypsin mit 40 % etwas besser ab als Thermolysin mit 35 % [104]. Dies bestätigt, dass Trypsin zwar effizient und spezifisch schneidet, aber möglicherweise nicht alle interessierenden Regionen abdeckt [129]. Der TC-Verdau führte zwar nicht zu mehr Proteinidentifikationen, aber zu einer besseren Abdeckung der Glutenproteinsequenz. Guo et al. berichteten ebenfalls, dass die Hinzufügung von Doppel- und Dreifach-Enzymverdauungen nicht zu einem wesentlichen Anstieg der

Proteinidentifizierungen führte, aber zu einer verbesserten mittleren Sequenzabdeckung von 32,5 % bis 42,0 % [129]. Im Vergleich zur ausschließlichen Trypsinverdauung ermöglichte die kombinierte Anwendung mehrerer Peptidasen eine um 18 % höhere Proteinidentifikation und eine doppelt so hohe Sequenzabdeckung [110]. Die Verwendung von mehr als einer Peptidase könnte auch bei der Suche nach bestimmten Proteinen oder Sequenzbereichen von Vorteil sein. So haben beispielsweise Giansanti et al. gezeigt, dass Phosphorylierungsstellen mit spezifischen Peptidasen verknüpft werden können, was ihre Nachweisbarkeit um den Faktor 1000 erhöhte [120].

Thermolysin erreichte eine vergleichsweise hohe Sequenzabdeckung für die Glutenproteine. Um dies zu veranschaulichen, wurde eine Heatmap erstellt, die zeigt, wie viele Peptide pro Glutenproteingruppe identifiziert wurden (**Abbildung 18**). Die meisten Peptide pro Proteingruppe wurden mit Thermolysin identifiziert. Im Durchschnitt gab es zwölf Peptide pro Gruppe, was sich auch in der Heatmap sehr gleichmäßig widerspiegelt. Die Gruppen mit der höchsten Anzahl an Peptididentifikationen waren überwiegend HMW-GS und LMW-GS. Mit TC und Pepsin wurden im Durchschnitt sieben Peptide pro Proteingruppe identifiziert. Unter den Peptiden, die mit TC nachgewiesen wurden, waren besonders viele HMW-GS in den Proteingruppen mit den meisten Peptiden zu finden. Bei Pepsin waren es vor allem LMW-GS und  $\gamma$ -Gliadine. Pro Proteingruppe wurden durchschnittlich sechs tryptische Peptide gefunden, wobei die Gruppen mit den meisten Peptiden HMW-GS und LMW-GS waren. Mit Chymotrypsin wurden durchschnittlich vier Peptide pro Gruppe nachgewiesen. Unter den Gruppen mit den meisten Peptididentifikationen befanden sich viele  $\alpha$ - und  $\gamma$ -Gliadine. Obwohl mit Trypsin bei weitem die meisten Peptide gefunden wurden (**Abbildung 14**), gilt dies nicht für Glutenpeptide. Insgesamt weist eine höhere Proteinabdeckung darauf hin, dass die identifizierten Glutenproteine mit höherer Wahrscheinlichkeit richtig zugeordnet wurden.



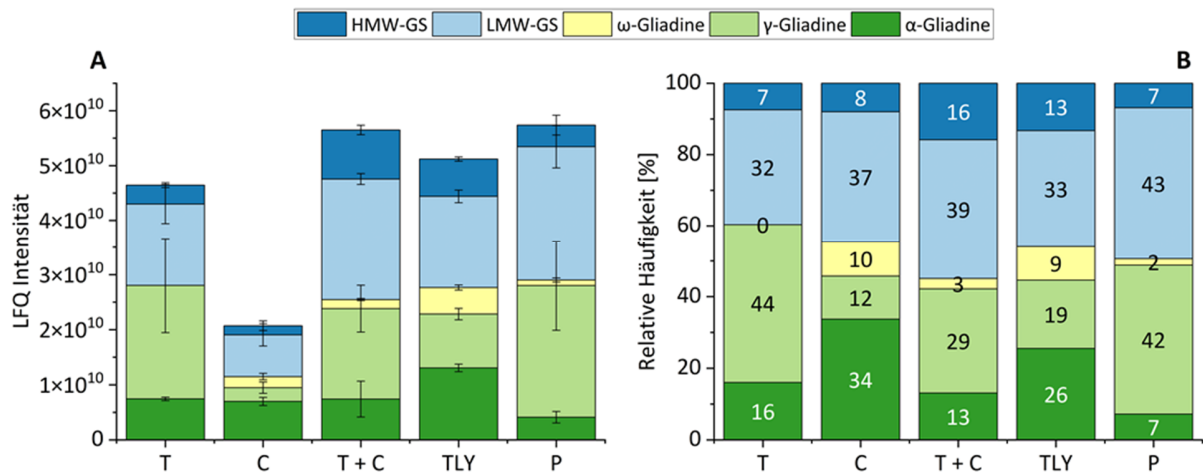
**Abbildung 18:** Heatmap basierend auf der Anzahl der identifizierten Peptide pro Glutenproteingruppe (insgesamt 215). Dargestellt sind die Bottom-Up Proteomics-Ergebnisse für die Weizensorte RGT Reform, die mit fünf verschiedenen Peptidasen verdaut wurde. Weiße Farbe bedeutet, dass für diese Proteingruppe keine Peptide nachgewiesen wurden. T, Trypsin; C, Chymotrypsin; TLY, Thermolysin; P, Pepsin.

### 3.2.3.2 RGT Reform: Glutenproteinzusammensetzung

Der Vergleich der LFQ-Intensitäten der Glutenproteintypen (**Abbildung 19**) verteilt sich anders als in **Abbildung 15**. Bei den absoluten Intensitätswerten liegen alle Enzyme in einem ähnlichen Bereich, außer Chymotrypsin. Dies bedeutet, dass nur halb so viele Proteine identifiziert wurden. Während die  $\alpha$ -Gliadine immer noch auf einem ähnlichen Niveau wie die anderen Enzyme liegen, wurde für alle anderen Glutenproteintypen weniger gefunden. Dies könnte darauf hindeuten, dass Chymotrypsin nicht so umfangreich spaltet wie die anderen Enzyme.

Im relativen Vergleich unterschieden sich die Enzyme am meisten innerhalb der Gliadine, während die Glutenine in einem ähnlichen Bereich lagen. Die  $\alpha$ -Gliadine variierten von 7 % (Pepsin) bis 34 % (Chymotrypsin). Der Unterschied bei den  $\gamma$ -Gliadinen war noch größer und reichte von 12 %

(Chymotrypsin) bis 44 % (Trypsin).  $\omega$ -Gliadine wurden mit Trypsin nicht identifiziert, sondern reichten von 2 % (Pepsin) bis 10 % (Chymotrypsin). Bei den LMW-GS wurden 32 % (Trypsin) bis 43 % (Pepsin) und bei den HMW-GS 7 % (Trypsin und Pepsin) bis 16 % (TC) detektiert.



**Abbildung 19:** Absolute (A) und relative (B) Häufigkeit spezifischer Glutenproteingruppen in RGT Reform nach gemeinsamer Auswertung aller Daten in MaxQuant. Die Proteingruppen wurden mit einer False Discovery Rate von 1 % identifiziert. T, Trypsin; C, Chymotrypsin; TLY, Thermolysin; P, Pepsin; HMW-GS, hochmolekulare Gluteninuntereinheit; LMW-GS, niedermolekulare Gluteninuntereinheit.

Das relative Muster veränderte sich im Vergleich zu **Abbildung 15** ebenfalls. Mit Trypsin wurden deutlich mehr  $\alpha$ - und  $\gamma$ -Gliadine identifiziert: 16 % und 44 % im Vergleich zu 10 % und 32 % zuvor. Andererseits sank der Anteil der LMW-GS auf 32 % (vorher 48 %). Das Verhältnis von Gliadinen zu Gluteninen war hier also genau umgekehrt und betrug 60 % Gliadine zu 40 % Glutenine. Bei Chymotrypsin war der Anstieg der  $\alpha$ -Gliadine am auffälligsten, von 10 % auf 34 %, während die  $\gamma$ -Gliadine von 26 % auf 12 % zurückgingen. Die Verteilung für TC und Thermolysin blieb in einem ähnlichen Bereich. Bei Pepsin war ein Rückgang der  $\alpha$ -Gliadine von 10 % auf 7 % und der HMW-GS (13 % auf 7 %) zu verzeichnen, während die  $\gamma$ -Gliadine und LMW-GS von 34 % auf 42 % bzw. 40 % auf 43 % zunahmen.

Die Auswertung bestätigt erneut die starke Abhängigkeit der Proteinverteilung vom eingesetzten Verdauungsenzym. Auffällig ist jedoch, wie groß die Unterschiede zwischen den beiden Auswertungsansätzen ausfallen. Die kombinierte Enzymanalyse in MaxQuant scheint die Zuordnung der Peptide zu den jeweiligen Proteingruppen anders zu beeinflussen. Dies deutet erneut darauf hin, dass die hohe Sequenzähnlichkeit der Glutenpeptide eine eindeutige Identifikation erschwert. Möglicherweise ist MaxQuant für die Analyse komplexer, hochredunder Proteinfamilien wie der Glutenproteine nur eingeschränkt geeignet. Zukünftig könnte eine spezifisch angepasste oder trainierte Auswertungssoftware erforderlich sein, um dieser Problematik besser gerecht zu werden.

Zusätzlich zur Weiterentwicklung von Datenbanken und Auswertungssoftware sollte künftig auch der Einsatz weiterer Peptidasen in Betracht gezogen werden. Neben Trypsin stehen zahlreiche andere proteolytische Enzyme als potenzielle Alternativen zur Verfügung, darunter Lys-C, Lys-N, Asp-N, Glu-C und Arg-C. Liu et al. setzten eine Tandem-Verdauungsstrategie mit Trypsin und der Glutamyl-Endopeptidase Glu-C ein, da Letztere über komplementäre Spaltspezifitäten zu Trypsin verfügt. Sie beobachteten, dass Glu-C die langen Peptidfragmente, die durch Trypsin entstehen, weiter in kürzere Stücke zerlegt, was deren Kompatibilität mit der anschließenden MS-Analyse deutlich erhöht. Mithilfe dieses Tandem-Ansatzes konnten sie insgesamt 121 Glutenproteine identifizieren, während bei ausschließlicher Trypsin-Verdau lediglich 89 detektiert wurden. Zudem ermöglichte die Methode die Identifikation von  $\omega$ -Gliadinen. Die Sequenzabdeckung der Glutenproteine konnte durch die Tandemverdauung von 19,4 % auf 25,3 % ebenfalls gesteigert werden [116]. Dies liegt im gleichen Bereich wie die Ergebnisse der vorliegenden Arbeit und reicht nicht aus, um die Zusammensetzung der Glutenproteine präzise abzubilden. Daher sollten weitere Alternativen geprüft werden – sowohl der Einsatz einzelner zusätzlicher Peptidasen als auch Tandem-Verdauungsstrategien.

Mit dem Fortschritt in der Entwicklung hochauflösender Massenspektrometer könnte es zukünftig möglich sein, eine Top-Down Proteomis-Analyse der Glutenproteine durchzuführen. Dies würde letztlich eine eindeutige Identifizierung und Zuordnung ermöglichen. Derzeit ist eine Umsetzung jedoch aufgrund der hohen molekularen Masse (>50 kDa), der ausgeprägten Sequenzhomologie sowie der strukturellen Heterogenität der Glutenproteine nicht realisierbar. Zusätzlich neigen native Glutenproteine zur Aggregation, was gezielte Optimierungen in der Probenvorbereitung erforderlich macht. Ein Kompromiss zwischen Bottom-Up und Top-Down Proteomics ist Middle-Down Proteomics, bei der größere Peptidfragmente (3–15 kDa) analysiert werden [130]. Die Methode ermöglicht eine verbesserte Sequenzabdeckung sowie den Erhalt struktureller Informationen (etwa zu Disulfidbrücken) bei gleichzeitig moderater Fragmentgröße, die massenspektrometrisch zuverlässig analysiert werden kann. Insbesondere im Fall von Glutenproteinen stellt dieser Ansatz eine vielversprechende Alternative dar, da er der hohen Sequenzhomologie und Isoformvielfalt besser gerecht wird als Bottom-Up Proteomics-Strategien, ohne die technischen Anforderungen der vollständigen Top-Down-Analyse mit sich zu bringen.

### 3.3 Targeted Proteomics von Weizenmehl

Teile dieses Kapitels wurden bereits bei der Fachzeitschrift *Food Research International* veröffentlicht [131].

Bei targeted Proteomics werden ausgewählte Zielpeptide selektiv untersucht. Eine konsistente und möglichst vollständige Spaltung der Proteine ist dabei von zentraler Bedeutung, da nur unter diesen Voraussetzungen eine zuverlässige und reproduzierbare Quantifizierung der Zielpeptide möglich ist. Der Vergleich der eingesetzten Peptidasen ergab, dass Trypsin mit deutlichem Abstand die höchste Spaltspezifität aufwies, was sich in der geringsten Anzahl an missed cleavages widerspiegelte. Diese ausgeprägte Selektivität war der entscheidende Faktor für die Auswahl von Trypsin als bevorzugte Peptidase für die Entwicklung der targeted Proteomics-Methode.

#### 3.3.1 Methodenentwicklung PRM-Methode

Ausgehend von der Entwicklung einer targeted LC-MS/MS Methode für Amylase/Trypsin-Inhibitoren (ATI) in Weizen [82, 126] wurde eine targeted Methode für die Aufarbeitung und Messung von Glutenpeptiden entwickelt (5.6). Die einzelnen Schritte werden im Folgenden detailliert erläutert.

##### 3.3.1.1 Identifizierung potenzieller Markerpeptide

Mehlproteinextrakte aus neun Weizengenotypen wurden mit dem im Vergleich zur UHPLC sensitiveren nanoLC-MS/MS-System im DDA-Modus gemessen, um potenzielle Markerpeptide für Gluten nach tryptischem Verdau zu identifizieren. Bei der anschließenden MaxQuant-Auswertung wurden ausschließlich Glutenpeptide ohne missed cleavages ausgewählt. Insgesamt konnten 262 solcher Peptide detektiert werden (Anhang, **Tabelle 13 & Tabelle 14**), darunter 72  $\alpha$ -Gliadin-, 61  $\gamma$ -Gliadin-, 76 LMW-GS- und 53 HMW-GS-Peptide. Diese wurden anschließend mittels PRM gemessen (5.6.2) um potenzielle Gluten-Markerpeptide zu identifizieren.

Die Auswahl der Gluten-Markerpeptide für die targeted Proteomics-Analyse erfolgte ausschließlich anhand ihrer MS-Nachweisbarkeit. Das bedeutete, dass diese Peptide (**Tabelle 1**) eine reproduzierbare tryptische Spaltung und ein gutes MS-Verhalten mit hohen Signalen ( $>1E6$ ) und stabilen Übergängen (mindestens fünf) zeigten. Darüber hinaus wurden sie bei wiederholter Aufarbeitung und Messung der Proben stets mit ähnlicher Intensität nachgewiesen, was auf ihre Stabilität während der gesamten Durchführung hindeutet. Die entsprechenden synthetischen Peptide (P1–P11) wurden zunächst zur

Verifizierung eingesetzt und anschließend zusammen mit den stabilisotopenmarkierte Standards (IS1–IS11) für die Quantifizierung verwendet (**Tabelle 1**).

**Tabelle 1:** LC-MS/MS-Parameter für die Quantifizierung von 11 Glutenpeptiden.

Peptid/ Interner Standard	Aminosäure- sequenz <sup>a</sup>	Einzig- artig	Gluten- proteintyp	Führendes (Razor) Protein UniProtKB ID	Präkursorion <i>m/z</i>		Produkt-Ionen	NCE	RT [min]
					P	IS			
P1/IS1	QCAIHNVVHAIILHQQQQ*R	Gruppe	α-Gliadin 1t	R9XW75	606,1 <sup>4+</sup>	608,6 <sup>4+</sup>	y9/y8/y7/y6/y5/y4/y3/b3	27	8,6
P2/IS2	NLALETLPAMC NVYIPPYCTIAPV GI*FGTN	Nein	α-Gliadin 2t	R9XUM8	1103,9 <sup>3+</sup>	1107,2 <sup>3+</sup>	y15/y9/y4/y3/b3/b4/b5/b6/b7	25	18,1
P3/IS3	SLVLQTLPSMCN VYVPECSIM*R	Nein	γ-Gliadin 1t	H9BFB6	898,8 <sup>3+</sup>	902,1 <sup>3+</sup>	y10/y9/y8/y7/y6/y5	25	14,7
P4/IS4	SVNVPLY*R	Protein	LMW-GS 1t	A0A0S2GJQ0	474,3	479,3	y6/y5/y4/y3/b3/b4	25	8,2
P5/IS5	TLPTMCSVNVPL YSSITSAPLGVGS *R	Gruppe	LMW-GS 2t	P93790	1354,2	1359,2	y17/y16/y14/y13/y10/y7/y3/b8	25	14,7
P6/IS6	VFLQQQCIPVAMQ*R	Nein	LMW-GS 3t	R4JAQ1	859,5	864,5	y12/y11/y10/y9/y8/y7/y6/y4/b3	25	10,8
P7/IS7	GQQGYPTSLQ QPGQQQGGYY PTSLQHTGQ*R	Gruppe	HMW-GS 1t	Q0Q5D3	1151,9 <sup>3+</sup>	1155,2 <sup>3+</sup>	y13/y11/y10/y8/y6/y5/y4/b6	29	9,4
P8/IS8	EGEASELQCE*R	Protein	HMW-GS 2t	A0A060MZP1	718,3	723,3	y10/y9/y8/y7/y6/y5/y4/y3/b3/b4	25	5,1
P9/IS9	AQQPATQLPTV C*R	Nein	HMW-GS 3t	Q6RX92	735,4	740,4	y11/y10/y9/y8/y7/y6/y5/y4/b3	25	7,3
P10/IS10	ELQESSLEAC*R	Nein	HMW-GS 4t	Q6RX92	661,3	666,3	y9/y8/y7/y6/y4/y3/b3/b4	25	6,5
P11/IS11	GGSFYPGETTTPP QQLQQ*R	Nein	HMW-GS 5t	B1B520	996,0	1001,0	y13/y12/y10/y9/y8/y7/y3/b4	27	8,7

<sup>a</sup>C: S-Carboxyamidomethylcystein; \*R: Arginin (<sup>13</sup>C<sub>6</sub>, <sup>15</sup>N<sub>4</sub>); \*F: Phenylalanin (<sup>13</sup>C<sub>9</sub>, <sup>15</sup>N); NCE: Normalized Collision Energy; RT, Retention Time; <sup>3+</sup>Präkursorionen waren 3<sup>+</sup>, <sup>4+</sup>Präkursorionen waren 4<sup>+</sup>, alle anderen 2<sup>+</sup>; Führendes (Razor) Protein: In MaxQuant wird bei der Protein-Gruppierung das "Leading razor protein" als das Protein bezeichnet, dem Peptide mit mehrfacher Zuordnung bevorzugt zugewiesen werden, wenn keine eindeutige Zuordnung möglich ist; LMW-GS, niedermolekulare Gluteninuntereinheiten; HMW-GS, hochmolekulare Gluteninuntereinheiten; t, targeted.

Dieses Verfahren ist ungewöhnlich, weil die Peptide nicht, wie in der Literatur empfohlen, nach Einzigartigkeit, Länge, biologischen und chemischen Modifikationen etc. sortiert wurden [70]. Die eindeutige Proteinzuordnung eines Peptids – im Sinne einer exklusiven Zuordnung zu genau einem Protein – gilt als eines der entscheidenden Kriterien. Nur mit diesen können einzelne Proteine später absolut quantifiziert werden. Dies stellt jedoch eine große Herausforderung bei der Bottom-up Proteomics von Weizenglutenproteinen dar. Glutenproteine bestehen aus Aminosäuresequenzen mit sich

wiederholenden Teilen desselben Peptidmotivs (**Abbildung 1**). Darüber hinaus enthalten die Peptidmotive einen hohen Anteil an Glutamin und Prolin [17, 99]. Diese sich wiederholenden Sequenzen erschweren die spätere Zuordnung der Peptide zu einzelnen Proteinen, wie schon mehrfach diskutiert wurde. Bereits das Auffinden weniger, eindeutig zuordenbarer Peptide (im Folgenden als "eindeutige Peptide" bezeichnet) pro Proteingruppe gestaltete sich daher als schwierig. In der untargeted Analyse der neun Weizenlinien wurden insgesamt 85 verschiedene Gliadinproteingruppen identifiziert. Allerdings enthielten lediglich 57 dieser Gruppen eindeutige Peptide, wobei in 47 Fällen jeweils nur ein einziges solches Peptid detektiert wurde. Es gab nur eine Proteingruppe, für die zehn eindeutige Peptide gemessen wurden. Die übrigen neun Proteingruppen enthielten zwischen zwei und sechs eindeutige Peptide. Bei den Gluteninproteingruppen konnten insgesamt 72 innerhalb der neun Weizenlinien differenziert werden. Davon wiesen 50 Gruppen eindeutige Peptide auf. In 39 dieser Fälle konnte jeweils nur ein einziges eindeutiges Peptid identifiziert werden. Die verbleibenden elf Proteingruppen enthielten zwischen zwei und vier solcher Peptide. Folglich gab es keine große Auswahl an eindeutig zuordenbaren Glutenpeptiden für eine targeted Analyse. Die Auswahl mehrerer eindeutiger Peptide, die eine Proteingruppe optimal repräsentieren können, war für die meisten Glutenproteingruppen gar nicht möglich. Dementsprechend gab es fast keine Peptide, die nur einem einzigen Protein zugeordnet werden konnten. Aufgrund eines Mangels an eindeutig zuordenbaren Peptiden wurden deshalb auch Peptide ausgewählt, die mehreren Proteingruppen zugeordnet wurden. Diese wurden in Tabelle 1 mit dem Eintrag „Nein“ für den Parameter eindeutig zuordenbar gekennzeichnet. Der Eintrag „Protein“ bedeutet, dass das Peptid eindeutig für ein einzelnes Protein ist, „Gruppe“ bedeutet, dass es eindeutig für eine Proteingruppe ist. Alle mit den ausgewählten Peptiden korrespondierenden Proteine sind im Anhang aufgeführt (Anhang, **Tabelle 15**).

Weiterhin behindert der hohe Anteil an Glutamin und Prolin die enzymatische Hydrolyse durch Trypsin. Es wurde bereits gezeigt, dass Glutenproteine nur wenige Spaltstellen für Trypsin aufweisen [75, 83]. Infolgedessen werden Peptide gebildet, die länger als 25 Aminosäuren sind, wie auch der Peptidlängenvergleich der verschiedenen Hydrolysate bestätigte (Anhang, **Abbildung 32**). Für den MS-Nachweis werden jedoch Peptide mit einer Länge zwischen acht und 25 Aminosäuren bevorzugt [70]. Im Falle der  $\alpha$ -Gliadine hatten 37 der 72 identifizierten Peptide eine Länge von 25 Aminosäuren oder mehr. Das längste identifizierte Peptid hatte eine Sequenzlänge von 38 Aminosäuren. Bei den  $\gamma$ -Gliadinen waren es 13 von 61 Peptiden mit mehr als 25 Aminosäuren, wobei das längste Peptid 39 Aminosäuren betrug. Dieser Trend ändert sich nicht bei den Gluteninpeptiden. 29 von 77 identifizierten Peptiden in den LMW-GS waren länger als 25 Aminosäuren; das längste Peptid war 40 Aminosäuren lang. Innerhalb der HMW-GS waren 14 von 53 Peptiden länger als 25 Aminosäuren. Hier zählte das längste Peptid 42 Aminosäuren.

Insgesamt waren 35 % der identifizierten Glutenpeptide länger als 26 Aminosäuren; von diesen ließ sich knapp die Hälfte eindeutig einem Protein zuordnen.

Darüber hinaus gibt es chemische Veränderungen, die sich im Allgemeinen nur schwer vermeiden lassen, z. B. die Oxidation von Methionin und Tryptophan. Die unvollständige Alkylierung von Cysteinen ist ebenfalls problematisch. Wären jedoch alle Glutenpeptide mit einem (alkylierten) Cystein herausgefiltert worden, wären nur 102 (39 %) übrig geblieben. Eine weitere chemische Modifikation ist die Möglichkeit der Zyklisierung, wenn sich Glutamin, Glutaminsäure oder carbamidomethyliertes Cystein am N-terminalen Ende der Peptide befindet [132]. Aufgrund des hohen Glutamingehalts in Glutenproteinen wurden Peptide mit N-terminalem Glutamin recht häufig identifiziert. Bei den  $\alpha$ - und  $\gamma$ -Gliadinen waren es 10 bzw. 12 Peptide. Bei den LMW-GS-Peptiden waren es nur 2, aber 20 aller HMW-GS-Peptide hatten ein N-terminales Glutamin. Das bedeutet, dass 16 % aller Glutenpeptide N-terminale Glutamine aufwiesen. Im Falle des carbamidomethylierten Cysteins waren nur  $\alpha$ - und  $\gamma$ -Gliadine betroffen. 10 bzw. 2 Peptide enthielten carbamidomethyliertes Cystein am N-terminalen Ende.

Obendrein gibt es auch einige biologische Modifikationen, die zu vermeiden sind, z. B. N- und C-terminale Peptide. Da sie sich an den Enden befinden, sind sie besonders anfällig für Proteolyse oder Abbau [70]. Die Aminosäuresequenz von Glutenproteinen unterscheidet sich jedoch am stärksten an ihren Enden [17, 99]. Es ist daher wahrscheinlich, dass sich an den Enden jene Peptide befinden, die für das betreffende Protein charakteristisch sind. Bei den ausgewählten Peptiden trifft dies auf das Peptid P8 zu. Da es sich um ein eindeutig einem einzelnen Protein zuordenbares Peptid handelt, wurde es dennoch in die Auswahl aufgenommen.

Es ist anzumerken, dass P9 und P10 das gleiche führende Razor Protein haben. Allerdings wurde P9 in 23 Proteinen gefunden, während P10 aus 51 Proteinen bestand (Anhang, **Tabelle 15**). Von diesen Proteinen hatten sie 10 Proteine gemeinsam, darunter das führende Razor Protein. Da die beiden Peptide jedoch in 13 bzw. 41 anderen Proteinen vorkamen, könnte die gezielte Analyse zu unterschiedlichen Mengen führen. Dies würde Rückschlüsse auf die Mengen der nicht-gemeinsamen Proteine zulassen.

Abschließend betrachtet stellten tryptische Glutenpeptide eine große Herausforderung bei der Suche nach geeigneten Peptiden für die targeted Analyse dar. Wenn jedes Peptid, das die Kriterien nicht erfüllt, aussortiert würde, blieben nicht viele übrig. Dennoch wurde bei der Entwicklung der targeted Methode Trypsin gewählt, weil aus der Literatur und dem Vergleich der Peptidasen (3.2) bekannt war, dass es am vollständigsten und selektivsten spaltet.

### 3.3.1.2 Targeted LC-MS/MS und SIDA

Die Entwicklung der targeted LC-MS/MS- und SIDA-Methode erfolgte in Anlehnung an Geißlitz et al. [82, 126]. Zunächst wurden geeignete Präkursorionen auf Basis der gemessenen Intensitäten ihrer jeweiligen Ladungszustände ausgewählt. In den meisten Fällen zeigte der Ladungszustand 2<sup>+</sup> die höchste Signalintensität und wurde daher bevorzugt verwendet. Bei längeren Peptiden eignen sich jedoch höhere Ladungszustände besser, da sie häufig stärkere Signale erzeugen. Aus diesem Grund wurde für die Peptide P2, P3 und P7 der Ladungszustand 3<sup>+</sup> und für P1 der Ladungszustand 4<sup>+</sup> gewählt. Dies galt auch für die entsprechenden stabilisotopenmarkierten Standards. Im nächsten Schritt wurden die Präkursorionen fragmentiert, um mindestens fünf bis acht stabile Übergänge zu definieren. Dabei wurde zusätzlich vorausgesetzt, dass mindestens ein Produkt-Ion einen höheren *m/z*-Wert als das Präkursorion aufweist. Abschließend wurden die Kollisionsenergien experimentell optimiert, um eine maximale Signalintensität zu erzielen.

Die Responsegeraden wurden zusammen mit den Proben analysiert. Um die Stabilität der synthetischen Peptide (markiert und unmarkiert) und der Peptide in der Probe zu bewerten, wurden die Responsegeraden und der Weizengenotyp RGT Reform mehrmals über einen Zeitraum von 15 Wochen gemessen. Insgesamt wurden die Mischungen für die Responsegeraden für jede P/IS-Kombination viermal neu hergestellt und Aliquots bei –80 °C gelagert. Ein frisch aufgetautes Aliquot wurde für eine Serie von 30 Proben (einmal pro Woche) verwendet. Insgesamt wurden die Responsegeraden mindestens 15 Mal in Doppelbestimmung gemessen. RGT Reform wurde wiederholt aufgearbeitet und 31 Mal in Dreifachbestimmung gemessen.

Für die Peptide P1/IS1, P4/IS4, P5/IS5, P6/IS6 und P8/IS8 war die Responsegerade im Laufe der Zeit nicht stabil. Die Response von P1/IS1 und P8/IS8 führte in kurzer Zeit zu größeren Verhältnissen. Dies bedeutete, dass wiederholte Messungen der gleichen Probe diesen Effekt bereits zeigten. Bei P4/IS4, P5/IS5 und P6/IS6 ergab die Response über einen Zeitraum von 15 Wochen größere Verhältnisse. Im Allgemeinen bedeutet die Verschiebung zu höheren Verhältnissen, dass das isotopenmarkierte Peptid abgebaut wird und daher eine geringere Signalintensität im Vergleich zum unmarkierten Peptid aufweist. P1/IS1 hat einen alkylierten Cysteinrest und P8/IS8 hat Glutaminsäure am N-terminalen Ende. Wahrscheinlich fand bereits beschriebene Zyklisierungsreaktion statt [132]. Das Peptid P10/IS10 hatte ebenfalls eine N-terminale Glutaminsäure. Seine Response blieb jedoch über die 15 Wochen hinweg stabil mit einer Standardabweichung von weniger als 5 %. Dennoch war auch hier ein leichter Trend zu höheren Verhältnissen zu erkennen. Es gibt keinen offensichtlichen Grund für die Degradation von IS4, IS5 und IS6.

IS5 ist mit 26 Aminosäuren recht lang, aber IS6 hat 16 Aminosäuren und IS4 nur 8 Aminosäuren. Auffallend ist, dass diese drei Peptide alle zu den LMW-GS gehörten. Bei allen genannten Peptiden schienen jedoch hauptsächlich die markierten Standards vom Abbau betroffen zu sein, während die nicht markierten Standards stabil blieben.

Um vergleichbare Ergebnisse für alle Proben zu erhalten, wurden Durchschnittswerte für die Responsegeraden aller Peptide benötigt. Aus diesem Grund wurden alle Responsewerte auf Ausreißer überprüft und diese entfernt. Anschließend wurde der Mittelwert für alle Punkte auf der Responsegeraden berechnet. Die Berechnungen der Peptidmengen in den Proben wurden dann anhand dieser gemittelten Responsegerade durchgeführt. Obwohl dies für die instabilen Peptide eine höhere Abweichung der gemittelten Response von der tatsächlich gemessenen Response bedeutete, war dies die einzige Möglichkeit, einheitliche Ergebnisse für die Peptide in den Proben zu erzielen. Alle anderen Peptide zeigten stabile Responsegeraden, weshalb der Mittelwert recht gut mit den jeweils gemessenen Werten übereinstimmt. Die Steigungen aller Responsegeraden lag zwischen 0,595 und 1,213, während die Schnittpunkte nahe bei 0,0 lagen. Dies ist vergleichbar mit den Ergebnissen, die in anderen Studien für Weizenproteine erzielt wurden [82, 133].  $R^2$  reichte von 0,997 bis 1.

Die 31 Messungen der Weizenlinie RGT Reform wurden verwendet, um festzustellen, ob die Peptide in den Proben stabil waren (Anhang, **Abbildung 34** & **Abbildung 35**). Die Ergebnisse zeigten, dass die meisten Peptide, mit Ausnahme von zwei, eine Standardabweichung von unter 5 % über alle Messungen hinweg aufwiesen. P4 und P11 wiesen jedoch Standardabweichungen von 19 % bzw. 12 % auf. P4 war in der RGT Reform nur in einer geringen Konzentration vorhanden, was die höhere Abweichung erklären könnte. Bei P11 war dies nicht der Fall: Trotz höherer Konzentration wies es stärkere Abweichungen auf, die sich nicht durch Parameter wie die Peptidlänge (18 Aminosäuren) erklären lassen. Insgesamt lagen alle Abweichungen in einem akzeptablen Bereich (<20 %).

Beim Vergleich der Stabilität der Peptide in der Response und in den Proben könnte die Matrix einen Einfluss gehabt haben. Die Responsegeraden wurden in einer wässrigen Matrix hergestellt, die frei von Analyten war. Da einige Peptide nur in Dimethylsulfoxid (DMSO) löslich waren, befand sich auch eine geringe Menge DMSO in der Verdünnung. DMSO ist dafür bekannt, dass es oxidierende Eigenschaften hat. Die Aminosäuren Cystein, Methionin und Tryptophan sind anfällig für Oxidation und sollten daher sogar unter anaeroben Bedingungen gelagert werden. Allerdings waren die Peptide in der IS1/IS11-Mischung, die den Proben als interner Standard zugesetzt wurde, ebenfalls in Wasser und DMSO gelöst und waren in den Messungen stabil. Neben der unterschiedlichen Matrix gab es einen wesentlichen Unterschied in der

Aufarbeitung der Responsegeraden im Vergleich zu den Proben. Nach der Reduktion und Alkylierung bei pH 8,5 wurden die Response-Proben nicht mehr mit SPE aufgereinigt, sondern direkt für die Messung weiter verdünnt. Bei einem pH-Wert von acht oder höher verschlechtert sich die Stabilität der Peptidlösung. Die Hersteller von synthetischen Peptiden empfehlen daher, die Peptidlösung bei einem pH-Wert von vier bis sechs zu lagern. Die Weizenproben mit internem Peptidstandard wurden nach der Alkylierung auf einer SPE gereinigt und dann in Wasser mit 2 % ACN und 0,1 % FA gelöst. Sie befanden sich demnach in einer sauren Umgebung und blieben auch während der Messungen weitgehend stabil.

### 3.3.1.3 Bestimmung von LOD/LOQ, Präzision und Wiederholbarkeit

Mit der endgültigen Methode wurden Wiederholbarkeit, Präzision, Nachweisgrenze (LOD) und Bestimmungsgrenze (LOQ) anhand einer Mischung aus neun Weizenlinien berechnet (**Tabelle 2**). Die Werte für die Wiederholbarkeit lagen für neun Peptide zwischen 1,4 % und 3,9 %. Dies liegt in einem guten Bereich, was sich auch in der entsprechenden HorRat<sub>r</sub> widerspiegelt, die stets zwischen 0,5 und 2,0 lag. Die mittlere Präzision lag mit 1,8-4,6 % ebenfalls knapp darüber und damit in einem sehr guten Bereich. P4 und P11 zeigten eine Wiederholbarkeit von 15,4 % bzw. 7,5 %. Die entsprechenden HorRat<sub>r</sub> lagen bei 4,9 und 3,0 und damit über dem Maximalwert von 2,0 für eine gute Wiederholbarkeit. Da P4 in allen Proben und damit auch in der Mischung nur in geringen Konzentrationen vorhanden ist, könnte dies der Grund für die hohe Abweichung sein. Im Falle von P11 könnte dies an der bereits erwähnten begrenzten Stabilität in Lösungen liegen. Es enthält viermal Glutamin, das anfällig für Oxidation ist. Allerdings enthalten viele der anderen ausgewählten Peptide ebenfalls hohe Mengen dieser Aminosäure und sind dennoch nicht betroffen. P11 ist das einzige Peptid, bei dem der isotopenmarkierte Standard in einer höheren Menge zugesetzt wurde als das ursprüngliche Peptid in der Probe, was zu Verhältnissen unter 1 führte. Dies könnte sich auf die Empfindlichkeit ausgewirkt haben. Bei allen anderen Proben wurde darauf geachtet, dass das Verhältnis größer als eins ist, idealerweise zwischen eins und zwei.

Die LOD für die entsprechenden Proteine der Markerpeptide lag zwischen 0,2 und 5,5 µg/g, die LOQ entsprechend zwischen 0,8 und 18,4 µg/g (**Tabelle 2**). Die erhaltenen Werte zeigen, dass die Peptide mit einer hohen Empfindlichkeit gemessen wurden. Die Ergebnisse stehen auch im Einklang mit früheren Studien für Glutenpeptide [133] und Weizen-ATI-Peptide [126, 127].

Im Allgemeinen wurden für die meisten Glutenproteine Wiederfindungswerte zwischen 98,9 und 119,2 % erzielt. Nur P4 zeigte aufgrund seiner geringen Konzentration in den Proben eine hohe Wiederfindung von 177,8 %. Für alle anderen Peptide konnte gezeigt werden, dass die Extraktion von verdünntem und

unverdünntem Weizenmehl ähnliche Ergebnisse in Bezug auf den Glutenproteingehalt erzielte und somit funktionierte.

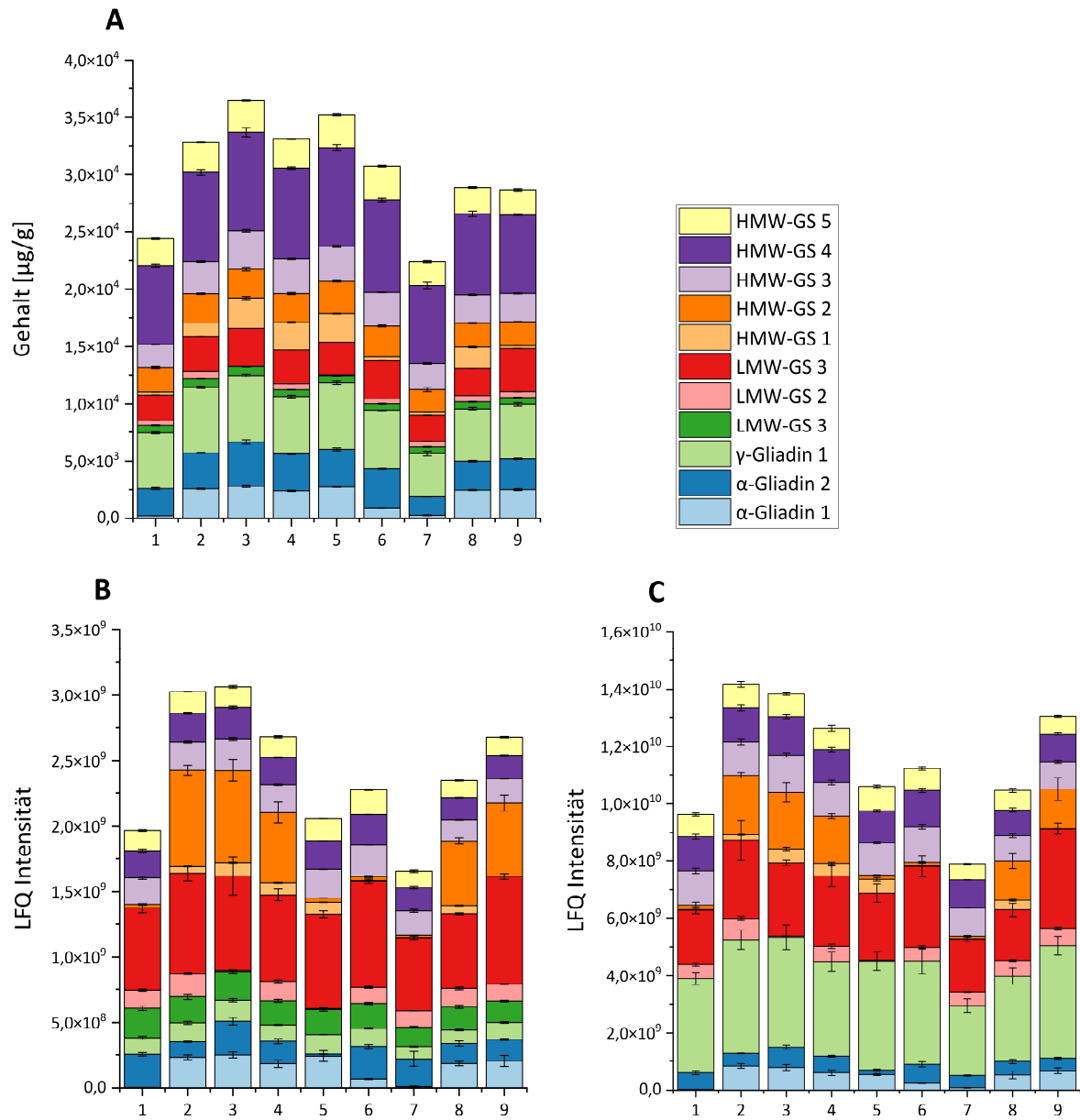
**Tabelle 2:** Ergebnisse der Methodenvalidierung der targeted LC-MS/MS-Analyse für elf Glutenpeptide.

Nr.	Glutenproteintyp	Präzision <sup>a</sup>			Sensitivität (Protein) [µg/g] <sup>b</sup>		Wiederfindung <sup>c</sup> %
		Wiederholbarkeit %	HorRat,	Durchschnitt %	LOD	LOQ	
P1	α-gliadin 1t	2,8	1,1	2,9	1,5	5,0	109,5
P2	α-gliadin 2t	3,0	1,2	2,9	5,5	18,4	111,7
P3	γ-gliadin 1t	2,1	0,9	2,2	1,1	3,6	98,93
P4	LMW-GS 1t	15,4	4,9	18,6	0,2	0,8	177,8
P5	LMW-GS 2t	2,8	0,8	3,4	1,8	6,1	122,5
P6	LMW-GS 3t	2,0	0,8	2,1	2,4	8,1	110,9
P7	HMW-GS 1t	3,9	1,4	4,6	3,8	12,8	119,2
P8	HMW-GS 2t	2,1	0,9	2,4	4,5	15,1	104,8
P9	HMW-GS 3t	1,4	0,6	1,8	4,1	13,7	102,2
P10	HMW-GS 4t	1,9	0,9	1,9	1,8	5,9	100,3
P11	HMW-GS 5t	7,5	3,0	8,4	4,5	15,1	117,8

<sup>a</sup>Präzision in Weichweizenmehl: Wiederholbarkeit: sechs biologische Replikate an einem Tag (n = 6), HorRat: Berechnetes Verhältnis des gemessenen und des berechneten Wertes für RSD<sub>r</sub>, annehmbare Werte liegen zwischen 0,5 und 2; Durchschnitt: sechs biologische Wiederholungen an drei Tagen (n = 18); <sup>b</sup>Nachweisgrenzen von Peptiden in glutenfreier Weizenstärke: LOD, Nachweisgrenze; LOQ, Bestimmungsgrenze; <sup>c</sup>Wiederfindung in Weizenmehl, das mit glutenfreier Weizenstärke verdünnt wurde (1 + 1); RSD<sub>r</sub>, relative Standardabweichung der Wiederholbarkeit; LMW-GS, niedermolekulare Gluteninuntereinheit; HMW-GS, hochmolekulare Gluteninuntereinheit.

### 3.3.2 Vergleich mit untargeted Ergebnissen

Die acht Elternlinien der BMWpop und der Vergleichsstandard RGT Reform wurden mit der finalen SIDA-Methode untersucht (**Abbildung 20**). Zusätzlich wurden die relativen Proteinmengen mithilfe von untargeted Proteomics im DDA-Modus bestimmt, wobei sowohl mit UHPLC als auch nanoLC in Kopplung mit einem Orbitrap-Massenspektrometer gemessen wurde. Für die Quantifizierung wurden die LFQ-Intensitäten jener Proteingruppen herangezogen, die das führende Razor Protein enthielten. Peptide, die auch in Proteinen anderer Gruppen vorkamen, wurden dabei nicht berücksichtigt. Dies entspricht der Vorgehensweise in MaxQuant, da zur Quantifizierung ausschließlich das führende Razor-Protein innerhalb einer Proteingruppe herangezogen wird.



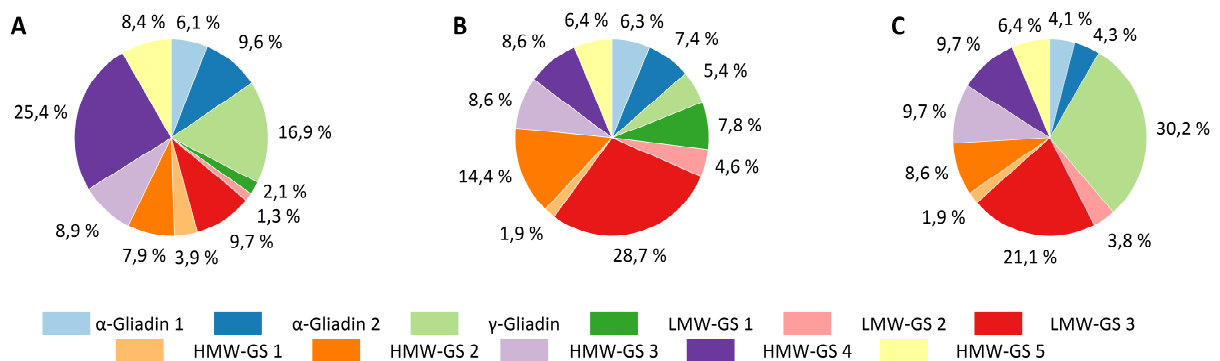
**Abbildung 20:** Vergleich der Verteilung der Glutenproteintypen in neun Weizensorten, die mittels SIDA (A), UHPLC-MS/MS (B) und nanoLC-MS/MS (C) analysiert wurden. 1: Ambition; 2: FIRL3565; 3: Bussard; 4: Event; 5: Format; 6: Julius; 7: BAYP4535; 8: Potenzial; 9: RGT Reform; LFQ, labelfreie Quantifizierung; SIDA, Stabilisotopenverdünnungsanalyse.

Betrachtet man zunächst die summierte Gesamtmenge an Weizenglutentypen in den Linien 1–4 und 6–9, so zeigt sich ein recht ähnliches Muster mit nur geringen Unterschieden. Das Gesamtmuster war lediglich bei der Linie 5 unterschiedlich. Während die relativen Daten ähnlich aussehen, zeigte die absolute Analyse der Weizenlinie 5 eine höhere Intensität im Vergleich zu den anderen Linien. Bei der Betrachtung der

einzelnen Proteinmengen/Intensitäten fällt auf, dass es deutliche Unterschiede zwischen relativen und absoluten Daten sowie zwischen den LFQ-Daten gab.

Zusammenfassend lässt sich sagen, dass die analysierten relativen Proteinmengen zwischen den drei Methoden teilweise übereinstimmen (**Abbildung 21**). Betrachtet man die Durchschnittswerte über alle Weizenlinien, so gilt dies für  $\alpha$ -Gliadin 1 (SIDA: 6,1 %, UHPLC-MS/MS: 6,3 %, nanoLC-MS/MS: 4,1 %) und 2 (SIDA: 9,6 %, UHPLC-MS/MS: 7,4 %, nanoLC-MS/MS: 4,3 %), LMW-GS 2 (SIDA: 1,3 %, UHPLC-MS/MS: 4,6 %, nanoLC-MS/MS: 3,8 %), HMW-GS 1 (SIDA: 3,9 %, UHPLC-MS/MS: 1,9 %, nanoLC-MS/MS: 1,9 %), HMW-GS 3 (SIDA: 8,9 %, UHPLC-MS/MS: 8,6 %, nanoLC-MS/MS: 9,7 %) und HMW-GS 5 (SIDA: 8,4 %, UHPLC-MS/MS: 6,4 %, nanoLC-MS/MS: 6,4 %), wobei es auch hier kleine Abweichungen gab. Dies steht im Einklang mit einer Studie, in der eine SIDA-Methode für fünf ATIs mit der intensitätsbasierten absoluten Quantifizierung (iBAQ) derselben Proben verglichen wurde [82].

Bei den anderen Proteingruppen wurden unterschiedliche Mengen festgestellt. Über alle Weizenlinien hinweg waren die Durchschnittswerte für LMW-GS 1 (SIDA: 16,9%, UHPLC-MS/MS: 5,4%, nanoLC-MS/MS: 30,2 %), LMW-GS 3 (SIDA: 9,7 %, UHPLC-MS/MS: 28,7 %, nanoLC-MS/MS: 21,1 %), HMW-GS 2 (SIDA: 7,9 %, UHPLC-MS/MS: 14,4 %, nanoLC-MS/MS: 8,6 %) und HMW-GS 4 (SIDA: 25,4 %, UHPLC-MS/MS: 8,6 %, nanoLC-MS/MS: 9,7 %).



**Abbildung 21:** Vergleich der durchschnittlichen relativen Konzentrationen der Glutenproteintypen in den neun Weizenlinien, bestimmt mittels SIDA (A), UHPLC-MS/MS (B) und nanoLC-MS/MS (C). LFQ, labelfreie Quantifizierung; SIDA: Stabilisotopenverdünnungsanalyse.

Die Daten für die  $\alpha$ -Gliadine unterschieden sich nicht sehr. Die absoluten Daten zeigten höhere Mengen im Vergleich zu den anderen Proteinen, aber die relative Quantifizierung geringere Mengen. Die einzige Ausnahme bildeten die relativen Daten für die Weizenlinie 5, wo nur geringe Mengen von  $\alpha$ -Gliadin 2 nachgewiesen wurden. Dennoch stimmten die Ergebnisse der verschiedenen Messmethoden recht gut überein.

Deutliche Unterschiede gab es bei  $\gamma$ -Gliadin 1. Die geringste Menge wurde hier bei der relativen Quantifizierung mittels UHPLC-MS/MS nachgewiesen. Eine deutlich größere Menge wurde mit SIDA nachgewiesen und noch mehr wurde mit nanoLC-MS/MS analysiert. Es ist widersprüchlich, dass die LFQ-Ergebnisse so inkonsistent sind und die SIDA-Ergebnisse genau dazwischen lagen.

Dies zeigt sich auch in der Gruppe der LMW-GS, da es auch hier erhebliche Unterschiede zwischen den verschiedenen Untersuchungsmethoden gab. Zunächst wurde LMW-GS 1 mit nanoLC-MS/MS nicht in allen Proben nachgewiesen. Mit SIDA war es in einer geringen Menge vorhanden, während es mit UHPLC-MS/MS in einer größeren, aber immer noch geringen Menge nachgewiesen wurde. LMW-GS 2 wurde in beiden LFQ-Daten in ähnlicher Menge analysiert, während es bei SIDA generell in einer geringeren Menge gefunden wurde. Interessanterweise wurde LMW-GS 2 in den Weizenlinien 3 und 5 mit keiner der Methoden nachgewiesen, was darauf hindeutet, dass es in diesen Weizenlinien nicht vorkommt. Dies könnte zu interessanten Schlussfolgerungen führen, da LMW-GS 2 mit dem Peptid P5 quantifiziert wurde, das für diese Gruppe einzigartig ist. Außerdem ist P5 das einzige Peptid unter allen elf Peptiden, das in einigen Linien nicht nachgewiesen werden konnte. LMW-GS 3 wies die größten Unterschiede auf. Es wurde in den UHPLC-MS/MS-Messungen in sehr großen Mengen nachgewiesen, gefolgt von der nanoLC-MS/MS, wo es ebenfalls in recht großen Mengen analysiert wurde. Im Gegensatz dazu wurde mittels SIDA deutlich weniger gefunden.

Interessanterweise war es bei den HMW-GS-Proteingruppen genau umgekehrt wie bei den LMW-GS. Im Allgemeinen wurde mit SIDA deutlich mehr analysiert als mit LFQ. HMW-GS 1 wurde bei LFQ nur in geringen Mengen nachgewiesen. In der SIDA wurde es in größeren Mengen nachgewiesen, allerdings in starker Abhängigkeit von der Weizenlinie. Im Gegensatz dazu wurde HMW-GS 2 bei den SIDA-Messungen in allen Weizenlinien in gleichen Mengen gefunden. Darüber hinaus wurde es bei beiden LFQ-Messungen in sehr unterschiedlichen Mengen innerhalb der Weizenlinien nachgewiesen. Beide Messungen lieferten das gleiche Ergebnis: In den Weizenlinien 1, 5, 6 und 7 wurde nur eine geringe Menge nachgewiesen, während die Weizenlinien 2, 3, 4, 8 und 9 relativ große Mengen davon enthielten. HMW-GS 3 wurde bei allen Methoden in ähnlichen Mengen nachgewiesen; die größten Unterschiede gab es jedoch bei der Proteingruppe HMW-GS 4. In den LFQ-Daten wies HMW-GS 4 die gleichen Mengen auf wie die HMW-GS 3-Proteingruppe, da sie das gleiche führende Razor Protein aufweisen. Dies gilt jedoch nicht für SIDA, so dass bei SIDA deutlich größere Mengen für HMW-GS 4 gefunden wurden. P10 (HMW-GS 4) kommt in neun verschiedenen Proteingruppen mit insgesamt 51 Proteinen vor und P9 (HMW-GS 3) nur in zwei verschiedenen Proteingruppen mit insgesamt 23 Proteinen. Um die unterschiedliche Anzahl von Proteingruppen auszugleichen, wurden alle zu einem Peptid gehörenden Proteingruppen berücksichtigt.

Wenn ein Peptid in mehreren Proteingruppen vorkam, wie z. B. P10, wurden die LFQ-Intensitäten aller Proteingruppen summiert, um eine „Gesamtintensität“ zu ermitteln. Diese Gesamtintensität unterschied sich nicht wesentlich von den Ergebnissen, wenn nur das führende Razor Protein betrachtet wurde (Anhang, **Abbildung 36**). Die letzte Proteingruppe, HMW-GS 5, wurde in allen Weizenlinien bei allen Methoden in ähnlichen Mengen analysiert.

Die Unterschiede zwischen den Messmethoden lassen sich dadurch erklären, dass mittels SIDA nur ein Peptid zur Quantifizierung eines Proteins/einer Proteingruppe/mehrerer Proteingruppen verwendet wurde. Dies ist nicht ausreichend, insbesondere wenn es andere Peptide gibt, die zur gleichen Proteingruppe gehören und bessere Messeigenschaften haben als das ausgewählte Markerpeptid. Diese werden dann in der untargeted Analyse erkannt und entsprechend in die LFQ-Berechnung einbezogen. Darüber hinaus besteht auch die Möglichkeit eines Peptidverlustes, z. B. durch posttranslationale Modifikationen (PTMs), unvollständigen Verdau oder Instabilität des Peptids [134, 135]. In diesen Fällen wurde bei SIDA weniger gefunden wie bei LMW-GS 2 und 3. Dies ist ein weiteres Problem, wenn nur ein Peptid pro Protein für die absolute Quantifizierung gewählt wird. Um die SIDA-Daten zu überprüfen, müssten mehr Peptide pro Proteingruppe analysiert werden, um zu vergleichen, ob die ermittelten Mengen in einem ähnlichen Bereich liegen. So wurde beispielsweise bereits erfolgreich eine SIDA-Methode für ATIs in verschiedenen Weizenarten etabliert, bei der 9 von 13 untersuchten ATIs mit mindestens zwei Peptiden qualifiziert und quantifiziert wurden [126]. Allerdings wurden auch dort für zwei Peptide, die von demselben Protein stammen, unterschiedliche Ergebnisse erzielt.

Ein weiterer möglicher Grund für die Unterschiede in der Proteinverteilung zwischen den Methoden könnten fehlerhafte Peptidzuweisungen bei der Auswertung der untargeted Daten sein, die durch die hohe Sequenzähnlichkeit mancher Glutenproteine bedingt sind. So unterscheidet sich beispielsweise ein Proteinpaar mit insgesamt 658 Aminosäuren lediglich an einer einzigen Position (vgl. UniProtKB-IDs Q6RX92 und Q6RX93). Ein weiteres Beispiel ist, dass  $\gamma$ -Gliadine und bestimmte LMW-GS in ihrer Struktur sehr ähnlich sind. Es ist daher möglich, dass Peptide, die eigentlich zu den LMW-GS gehören, bei der Berechnung dem  $\gamma$ -Gliadin zugeordnet wurden. Dies wäre eine Erklärung für den großen Anteil an LMW-GS 3 in den UHPLC- und nanoLC-MS/MS-Daten bzw.  $\gamma$ -Gliadin-Anteil in den nanoLC-MS/MS-Daten.

Diese Beobachtungen verdeutlichen erneut ein grundsätzliches Problem bei der Analyse komplexer Weizenproteine mittels Bottom-up Proteomics: Die hohe Redundanz und Sequenzähnlichkeit, insbesondere aufgrund repetitiver Motive, erschwert eine eindeutige Zuordnung der Peptide zu individuellen Proteinen. Hinzu kommt eine weitere Einschränkung durch die LFQ-Quantifizierung über das

sogenannte „führende Razor-Protein“, wie sie beispielsweise in MaxQuant implementiert ist. Dabei wird zur Quantifizierung nur das dominante Protein innerhalb einer Proteingruppe berücksichtigt. Dies kann zu einer Unterschätzung der tatsächlichen Proteinverteilung führen, insbesondere wenn ein Peptid in mehreren, funktionell relevanten Proteinen vorkommt. Daher sollten untargeted Ergebnisse hinsichtlich ihrer biologischen Aussagekraft mit Vorsicht betrachtet und durch targeted Analysen ergänzt werden. Zusätzlich lag die durchschnittliche Sequenzabdeckung der Glutenproteine nach tryptischer Hydrolyse bei weniger als 30 %. Vor diesem Hintergrund stellt sich die Frage, ob eine korrekte Zuordnung tryptischer Peptide in der untargeted Analyse in allen Fällen gewährleistet werden kann. Methodisch lässt sich dies jedoch nur schwer eindeutig nachweisen.

Auffällig war auch, dass mit den LFQ-Daten verschiedener HPLC-Systeme teilweise sehr unterschiedliche Ergebnisse ermittelt wurden, insbesondere für LMW-GS 1, das mit nanoLC-MS/MS nicht nachgewiesen wurde. Die Tatsache, dass LMW-GS 1 mit UHPLC-MS/MS und SIDA, die ebenfalls mit derselben Säule über UHPLC durchgeführt wurde, nachgewiesen wurde, könnte darauf hinweisen, dass das nanoLC-System das entsprechende Peptid P4 nicht gut trennen konnte. Darüber hinaus verdeutlicht auch das Beispiel von  $\gamma$ -Gliadin 1, dessen gemessene Menge im SIDA-Experiment zwischen den Werten der untargeted Experimente lag, dass die HPLC-Trennung der Peptide einen erheblichen Einfluss auf die LFQ-Auswertung hatte. Dies lässt darauf schließen, dass Peptide derselben Proteingruppe unterschiedlich chromatographisch getrennt und infolgedessen mit variierenden Intensitäten detektiert wurden. Dies führte letztlich bei einigen Proteingruppen zu deutlichen Unterschieden in den LFQ-Intensitäten zwischen den beiden HPLC-Messungen. Grundsätzlich ist die relative Quantifizierung von Weizenglutenproteinen mittels Bottom-Up Proteomics möglich. Allerdings können bereits Messungen derselben Probe auf unterschiedlichen Systemen zu stark abweichenden Ergebnissen führen. Abgesehen von den glutenbezogenen Herausforderungen mangelt es labelfreien Ansätzen nach wie vor an ausreichender Robustheit und Reproduzierbarkeit [134].

SIDA bietet im Vergleich zu relativen Quantifizierungsmethoden eine höhere Präzision, Genauigkeit und Empfindlichkeit [70]. Demnach sollte diese Methode präzisere Ergebnisse in Bezug auf die untersuchten Peptide liefern. Folglich gilt dies auch für die entsprechenden Proteine, in denen dieses Peptid vorkommt. Interessant ist daher, dass HMW-GS 2 in allen Weizenlinien der SIDA-Messungen in ähnlichen Mengen nachgewiesen wurde, während in den beiden ungezielten Messungen signifikante Unterschiede zwischen den Linien analysiert wurden (was wiederum ein Hinweis auf die schwierige Zuordnung der Weizenpeptide in LFQ-Auswertungen sein könnte). Besonders auffällig waren die großen Mengen an HMW-GS 4, die im SIDA-Versuch gefunden wurden, was auch darauf hindeutet, dass das zugehörige Peptid P10 in der PRM-

Messung viel besser gefiltert und folglich erkannt wurde. Dies gilt auch für alle anderen Proteingruppen, für die mit SIDA größere Mengen gefunden wurden.

Geisslitz et al. verglichen verschiedene absolute und relative Quantifizierungsmethoden für die ATI-Peptidauswahl in Weizenlinien. Es zeigte sich, dass die verschiedenen Methoden bei mengenmäßig dominanten ATIs sehr gut zur Peptididentifizierung passten, während bei weniger dominanten ATIs teilweise unterschiedliche Peptide entdeckt wurden. Die Auswahl der Peptide für die Quantifizierung zeigte jedoch eine größere Variation [136]. Außer P4 und P5 wurden alle für SIDA ausgewählten Markerpeptide in großen Mengen nachgewiesen. Es ist daher unwahrscheinlich, dass die manchmal recht großen Unterschiede zwischen den Methoden auf die detektierte Menge zurückzuführen sind.

Es gibt verschiedene Studien, die das Weizenproteom untersucht haben, die jedoch keinen targeted Ansatz verfolgten [125, 137, 138]. Einige Studien verfolgten auch einen targeted Ansatz für die Suche nach Weizenpeptiden [111, 139]. Diese Ansätze zielten jedoch in erster Linie darauf ab, kleine Mengen von Weizenproteinen als Kontamination in anderen Proben nachzuweisen. Eine Bedingung für die Aminosäuresequenz der ausgewählten Weizenpeptide war daher, dass sie in keiner anderen Getreidesorte vorkommen. Ein weiterer targeted Ansatz bestand darin, den Gesamtgehalt an Weizengluten mit Peptidmarkern zu bestimmen [133]. Die Menge der chymotryptischen Peptide wurde mit einem internen isotopenmarkierten Standard über SIDA quantifiziert. Mit Hilfe der quantifizierten Daten aus den HPLC-Messungen war es möglich, Rückschlüsse auf den Gesamtglutengehalt zu ziehen. Im Unterschied dazu bestand ein Ziel dieser Arbeit in der absoluten Quantifizierung spezifischer Glutenproteingruppen mittels der SIDA-Methode. Dazu wurden für jede ausgewählte Gluten-Proteingruppe stabilisotopenmarkierte Referenzpeptide eingesetzt. Es ist anzumerken, dass sowohl die methodische Validierung als auch der Einsatz synthetischer, stabilisotopenmarkierter Peptide als interne Standards mit erheblichem Zeit- und Kostenaufwand verbunden sind. Dennoch ermöglicht deren Verwendung eine deutlich robustere und präzisere Quantifizierung der ausgewählten Glutenproteingruppen. Nach derzeitigem Kenntnisstand handelt es sich hierbei um die erste Arbeit, in der eine SIDA-basierte LC-MS/MS-Methode zur Quantifizierung spezifischer Glutenpeptide bzw. -proteine entwickelt wurde.

### 3.4 Weizenproteomics in Verbindung mit Backqualität

Teile dieses Kapitels wurden bereits bei der Fachzeitschrift *Food Research International* veröffentlicht [131].

Anhand der targeted und untargeted Proteomics-Daten sowie der Daten zur Verarbeitungsqualität der BMWpop 2018 wurde der Zusammenhang zwischen Proteom und Backqualität untersucht. Die Backversuche wurden an der Bayerischen Landesanstalt für Landwirtschaft (LfL) durchgeführt und die Daten zur Verfügung gestellt.

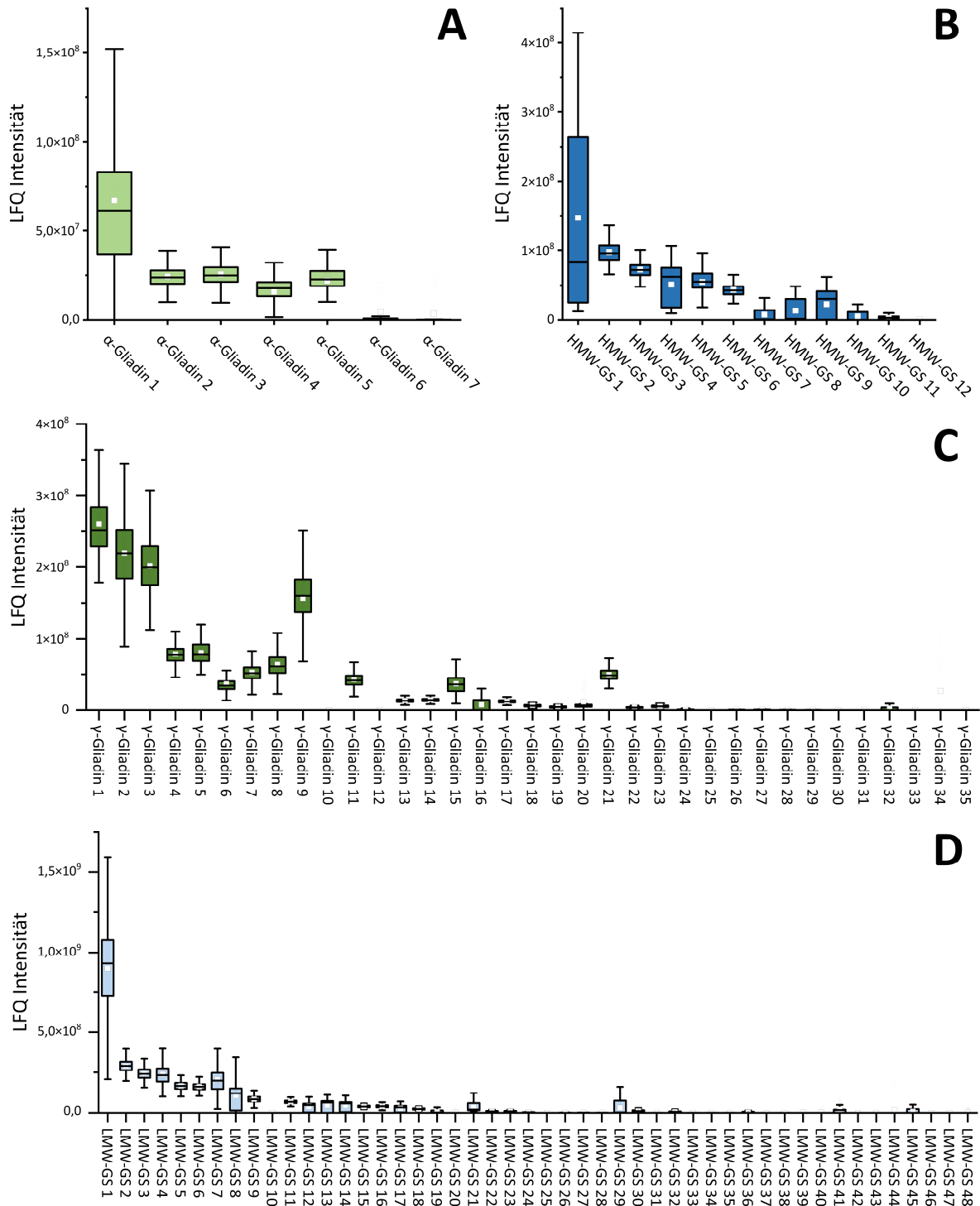
#### 3.4.1 Untargeted Messung der BMWpop

Die BMWpop aus der Ernte 2018 – bestehend aus acht Elternlinien, 394 rekombinanten Tochterlinien und RGT Reform als Vergleichsstandard – wurde vollständig in Dreifachbestimmung aufgearbeitet, mittels untargeted Proteomics gemessen und die Rohdaten mit MaxQuant und der UniProtKB ausgewertet (5.4.1).

##### 3.4.1.1 Allgemeine Übersicht

Insgesamt wurden 816 verschiedene Proteingruppen identifiziert. Darunter befanden sich 101 Glutenproteingruppen aufgeteilt in sieben  $\alpha$ -Gliadine, 35  $\gamma$ -Gliadine, 47 LMW-GS und 12 HMW-GS (Anhang, **Tabelle 16 - Tabelle 19**). Es wurden keine  $\omega$ -Gliadine identifiziert, da mit Trypsin verdaut wurde. In der Literatur werden  $\omega$ -Gliadine nicht mit einer positiven Beeinflussung der Backqualität in Verbindung gebracht. Im Gegenteil, es wurde festgestellt, dass sowohl das Entfernen von  $\omega$ 5-, als auch von  $\omega$ 1,2-Gliadinen zu besseren Teigeigenschaften und erhöhter Proteinstabilität führten [140-142]. Vor diesem Hintergrund wurde die eingeschränkte Nachweisbarkeit der  $\omega$ -Gliadine durch Trypsin in Kauf genommen. Die geringe Anzahl an identifizierten  $\alpha$ -Gliadinen ist der Extraktionsmethode geschuldet, wie bereits diskutiert wurde (3.1.1).

Die LFQ-Intensitätsverteilung in den 403 Weizenlinien zeigte, dass unabhängig vom Glutenproteintyp ein paar wenige Proteingruppen in sehr hoher Menge und Variabilität identifiziert wurden, während andere kaum in den Weizenlinien vorhanden waren (**Abbildung 22**). Die Proteingruppe mit der höchsten gemessenen Intensität ( $>1E9$ ) und Variabilität war LMW-GS 1 gefolgt von HMW-GS 1 (4E8) und  $\gamma$ -Gliadin 1–3 (3E8). Die  $\alpha$ -Gliadine wurden generell mit geringerer Intensität (zwischen 5E7 und 1E8) detektiert.



**Abbildung 22:** Verteilung der LfQ-Intensitäten der  $\alpha$ -Gliadine (A), HMW-GS (B),  $\gamma$ -Gliadine (C) und LMW-GS (D) in den 403 Weizenlinien der BMWpop 2018. Boxplots zeigen den Interquartilbereich (Box), den Mittelwert (Quadrat), den Median (horizontale Linie) sowie Minima und Maxima (Whisker). HMW-GS, hochmolekulare Gluteninuntereinheit; LfQ, labelfreie Quantifizierung; LMW-GS, niedermolekulare Gluteninuntereinheit.

Die genetische Zuordnung der  $\alpha$ - und  $\gamma$ -Gliadine sowie der LMW-GS mithilfe der UniProtKB gestaltete sich als herausfordernd, da entweder keine oder zahlreiche redundante und uneinheitliche Informationen verfügbar waren. Zudem erwies sich die manuelle Recherche als sehr zeitaufwendig. Dennoch konnten die LMW-GS in die Typen LMW-s, -m und -i unterteilt werden (Anhang, **Tabelle 18**). Dabei zeigte sich, dass 30 der 47 LMW-GS Gruppen dem m-Typ zugeordnet werden konnten. Lediglich fünf Gruppen (LMW-GS 8, 26, 31, 33 und 39) fielen in die Kategorie des s-Typs. Sowohl LMW-GS des m- als auch des s-Typs fanden sich in den Gruppen 3, 6, 9, 15, 16 und 46, während LMW-GS 2, 32 und 41 dem selteneren i-Typ zugeordnet werden konnten. Die Gruppe LMW-GS 34 konnte keinem der Typen eindeutig zugewiesen werden; sie gehört zur LMW-GS-Gruppe D, welche strukturelle Ähnlichkeiten zu den  $\omega$ -Gliadinen aufweist und in der Literatur deshalb auch als  $\omega$ b-Gliadine bezeichnet werden [143].

Bei den HMW-GS sind mehr Informationen in der Datenbank hinterlegt, die in einer Tabelle zusammengefasst wurden (Anhang, **Tabelle 19**). Zudem wurde mittels gekoppelten DNA-Markern von der LfL bestimmt, welche HMW-GS-Allele in den Weizenlinien der BMWpop vertreten sind. Dabei wurden die Allele an den *Glu-A1* Loci (Ax1, Ax2\*, Null), am *Glu-B1* Locus (Bx7+By8, Bx7+By9, Bx6+By8) sowie am *Glu-D1* Locus (Dx2+Dy12, Dx5+Dy10) erfasst. Auf Basis dieser Daten konnte eine plausible genetische Zuordnung vorgenommen werden (**Tabelle 3**). Es ist darauf hinzuweisen, dass keine Unterscheidung zwischen Dx2 und Dx5 (beide HMW-GS 3) möglich war. Ungewöhnlich ist weiterhin die Detektion von 1Ay (1Ay21\*, HMW-GS 11). Das Null-Allel ist eine Variante von 1Ay, bei der keine Expression erfolgt. Dieses ist in den meisten modernen Weizensorten vorhanden. In der BMWpop scheint es dennoch Weizenlinien zu geben, die aktive 1Ay-Varianten exprimieren. Im Vergleich mit **Abbildung 22** fällt weiterhin auf, dass die gekoppelten HMW-GS x- und y-Typen unterschiedlich stark exprimiert werden, da sie in variierenden Mengen nachgewiesen wurden. Die unterschiedliche Expression der x- und y-Typen könnte auf verschiedene regulatorische Mechanismen zurückzuführen sein. Studien zeigen, dass *Glu-1*-Gene beider Typen durch unterschiedliche Transkriptionsfaktoren gesteuert werden: Während die x-Typ-Gene überwiegend von Myb-Faktoren reguliert werden, erfolgt die Regulation der y-Typ-Gene in der späten Kornfüllungsphase durch andere Transkriptionsfaktoren [144]. Weitere Forschungsarbeiten beschäftigen sich mit der Analyse zusätzlicher regulatorischer Faktoren und Promotorregionen, um die Steuerung der HMW-GS-Expression umfassender zu verstehen [145, 146].

**Tabelle 3:** Mutmaßliche Zuordnung der identifizierten HMW-GS zu spezifischen Untereinheiten.

Nr.	Untereinheit
HMW-GS 1	Ax1, Ax2*
HMW-GS 2	Dy12 (By8)
HMW-GS 3	Dx2, Dx5
HMW-GS 4	Bx7
HMW-GS 5	By9
HMW-GS 6	By8
HMW-GS 7	Ax1, Ax2*
HMW-GS 8	Bx6
HMW-GS 9	Dy10
HMW-GS 10	Bx6, Bx7
HMW-GS 11	Ay
HMW-GS 12	Ax2* (Fragment)

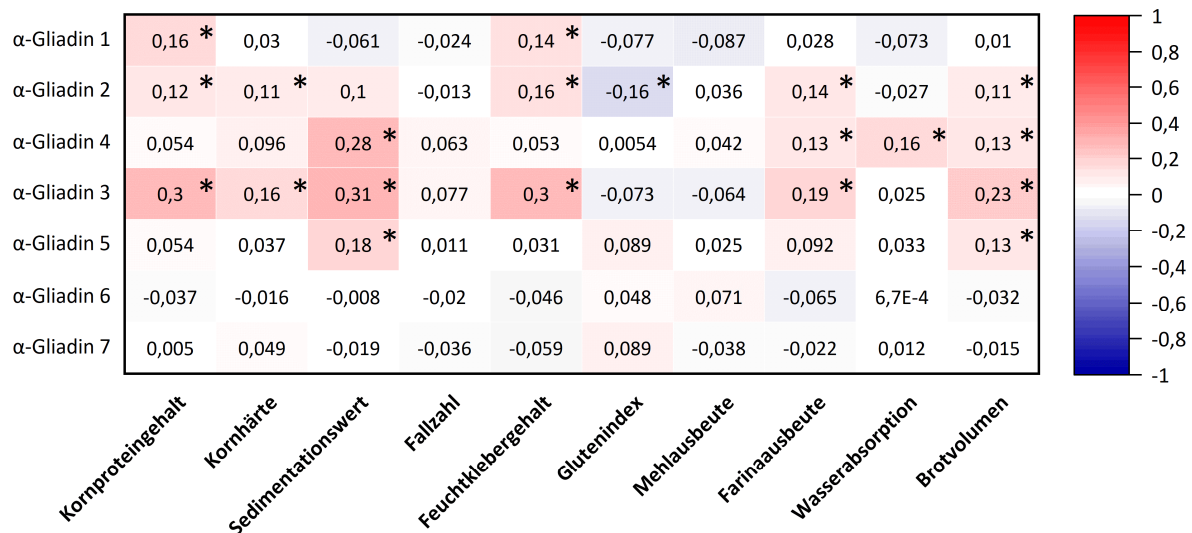
HMW-GS, hochmolekulare Gluteninuntereinheit.

#### 3.4.1.2 Korrelation mit Backqualitätsdaten

Aus den LFQ-Intensitätswerten der MaxQuant Auswertung aller Proteingruppen der BMWpop 2018 wurden die einzelnen Glutenproteintypen gefiltert. Diese wurden jeweils typweise mit Backqualitätsdaten der BMWpop 2018 korreliert, um festzustellen, ob einzelne Proteingruppen enger mit der Backqualität zusammenhängen. Die Korrelationen waren mehrheitlich positiv, insgesamt jedoch schwach ausgeprägt. In der Mehrzahl der Fälle bewegten sich die Korrelationskoeffizienten im sehr schwachen ( $r < 0,2$ ) oder schwachen Bereich ( $0,2 \leq r < 0,4$ ). Moderate Korrelationen ( $0,4 \leq r < 0,6$ ) traten nur vereinzelt auf, während starke oder sehr starke Zusammenhänge ( $r \geq 0,6$ ) nicht festgestellt wurden. Korrelationen ab  $r = 0,4$  gelten als potenziell relevant, da sie auf einen messbaren Einfluss einzelner Proteine auf technologische Eigenschaften hinweisen können [147]. Darüber hinaus wurden vereinzelt auch sehr schwache bis schwache negative Korrelationen beobachtet. Trotz der insgesamt geringen Ausprägung werden im Folgenden alle signifikanten Korrelationen ( $p \leq 0,05$ ) analysiert, da sie Gruppen mit statistisch relevanten Zusammenhängen von solchen ohne signifikante Ergebnisse abgrenzen.

Insgesamt konnten nur sieben verschiedene  $\alpha$ -Gliadine relativ quantifiziert werden. Innerhalb der  $\alpha$ -Gliadine zeigten insbesondere die Gruppen eins bis fünf mehrere signifikante Korrelationen mit Qualitätsparametern (**Abbildung 23**). Für  $\alpha$ -Gliadin 6 und 7 konnten hingegen keine relevanten Zusammenhänge festgestellt werden, was wahrscheinlich daran lag, dass diese nur in geringen Mengen in den Weizenlinien vorhanden waren (siehe auch **Abbildung 22**). Mit den Parametern Fallzahl und Mehlausbeute ergaben sich keine signifikanten Korrelationen zu einem der  $\alpha$ -Gliadine. Die deutlichsten Zusammenhänge zeigten sich bei  $\alpha$ -Gliadin 3, das signifikant mit Kornproteingehalt (0,30),

Sedimentationswert (0,31), Feuchtklebergehalt (0,30) und Brotvolumen (0,23) positiv korrelierte. Weitere  $\alpha$ -Gliadine zeigten meist nur sehr schwache Korrelationen mit einzelnen Qualitätsparametern. Eine Ausnahme bildete  $\alpha$ -Gliadin 4, das schwach mit Sedimentationswert (0,28) assoziiert war. Insgesamt blieben die  $\alpha$ -Gliadine hinsichtlich ihrer Korrelationen mit den Backqualitätsmerkmalen vergleichsweise unauffällig.



**Abbildung 23:** Korrelationsmatrix der relativen  $\alpha$ -Gliadin-LFQ-Intensitäten und Backqualitätsdaten der BMWpop 2018. \* $p \leq 0,05$ ; LFQ, labelfreie Quantifizierung.

Bei  $\alpha$ -Gliadin 3 handelt es sich um eine Proteingruppe, die vier Proteine enthält, welche nach UniProtKB alle dem *Gli-2* Locus zugeordnet werden. Ein Vergleich des ersten Hauptproteins dieser Gruppe mit dem erstgenannten Hauptprotein der anderen  $\alpha$ -Gliadine zeigte, dass es sich mit 330 Aminosäuren um das längste  $\alpha$ -Gliadin handelte, während die anderen Aminosäurelängen von 221 bis 318 Aminosäuren aufwiesen. Auffällig war, dass das Hauptprotein von  $\alpha$ -Gliadin 3 (UniProtKB-ID A0A2R2Y404) besonders viele Abschnitte mit langen Glutaminwiederholungen aufwies; das längste Segment bestand aus 45 aufeinanderfolgenden Glutaminresten. Diese Segmente sind typisch für  $\alpha$ -Gliadine und kommen in keinem der anderen Glutenproteintypen vor [13]. Bei  $\alpha$ -Gliadin 6 und 7, welche keinerlei Korrelationen aufwiesen, handelte es sich mit 277 und 221 Aminosäuren um die kürzesten  $\alpha$ -Gliadine. Die hier verglichenen  $\alpha$ -Gliadine enthielten alle sechs Cysteinreste. Dies stimmt mit der Literatur überein, wo ebenfalls sechs Cysteinreste in den nicht-repetitiven Regionen festgestellt wurden, welche intramolekulare Disulfidbrücken formen [148]. Deshalb bilden sie keine Disulfidbrücken zu anderen Proteinen aus und tragen folglich weniger zur Stabilisierung des Glutennetzwerks bei, sondern eher zur Viskosität des Teiges. Es wurden allerdings auch Gene für  $\alpha$ -Gliadine beschrieben, die sieben Cysteinreste enthalten [149].

Gliadine mit einer ungeraden Anzahl an Cysteinen gelten jedoch auch als Terminatoren für die Bildung des Glutennetzwerks, da sie die Polymerisation der HMW-GS unterbrechen [16]. Deshalb sollten auch sie sich negativ auf die Teigstabilität auswirken. Bei den  $\alpha$ -Gliadinen lässt sich dies nicht überprüfen, da keines mit einer ungeraden Anzahl an Cysteinen detektiert wurde. Um dies zu ermöglichen, müssten mehr  $\alpha$ -Gliadine mithilfe einer alternativen Extraktionsmethode, wie der Dreifachextraktion, aus den Proben isoliert werden.

Da viele T-Zell-stimulierende Epitope aus  $\alpha$ -Gliadinen stammen, wird diese Proteinfraction häufig mit weizenbedingten Erkrankungen wie Zöliakie in Verbindung gebracht. Aus diesem Grund stehen  $\alpha$ -Gliadin-Allele im Fokus der Forschung zur Krankheitsentstehung. Allerdings ist auch bekannt, dass sie in der Lage sind, die Backqualität zu verbessern [150]. Van den Broeck et al. zeigten, dass die Deletion des  $\alpha$ -Gliadin Locus auf dem kurzen Arm von Chromosom 6D zu einer deutlichen Verschlechterung der technologischen Eigenschaften führte. Im Gegensatz dazu blieben diese Eigenschaften unbeeinträchtigt, wenn Deletionen auf dem kurzen Arm von Chromosom 1D vorgenommen wurden, der die  $\omega$ -Gliadin-,  $\gamma$ -Gliadin- und LMW-GS-Loci umfasst [151]. In einer späteren Studie berichteten Van den Broeck et al. jedoch, dass die gezielte Deletion des *Gli-D2* Locus auf dem kurzen Arm des 6D Chromosoms, welcher  $\alpha$ -Gliadin-Gene enthält, die Backqualität positiv beeinflussen kann. Durch die gezielte Stilllegung von  $\alpha$ -Gliadin-Genen in Weichweizen wurde der  $\alpha$ -Gliadinegehalt um 63 % reduziert. Gleichzeitig wurden die  $\alpha$ -Gliadine durch Avenine ersetzt. Dies führte zu einer verbesserten Teigstärke, obwohl das Brotvolumen leicht abnahm [152].

Bei den  $\gamma$ -Gliadinen zeigten sich ebenfalls signifikante, allerdings nur schwache Korrelationen mit den Backqualitätsparametern (**Abbildung 24**). Besonders für Kornproteingehalt, Sedimentationswert, Feuchtklebergehalt und Brotvolumen wurden überwiegend positive Zusammenhänge beobachtet. Im Gegensatz dazu standen der Glutenindex und die Wasseraufnahme, die häufiger negativ mit den  $\gamma$ -Gliadinen korrelierten. Innerhalb der  $\gamma$ -Gliadin-Gruppe traten die Korrelationen hauptsächlich bei den  $\gamma$ -Gliadinen 1 bis 28 auf, während für die  $\gamma$ -Gliadine 29 bis 35 keine signifikanten Zusammenhänge festgestellt wurden. Wie bei den  $\alpha$ -Gliadinen lag dies wahrscheinlich daran, dass diese nur in wenigen Weizenlinien vorhanden waren. Die stärksten positiven Korrelationen der  $\gamma$ -Gliadine mit Qualitätsparametern lagen zwischen 0,31 und 0,38, insbesondere zwischen dem Kornproteingehalt und  $\gamma$ -Gliadin 11, 14 und 21.  $\gamma$ -Gliadin 11 zeigte zudem signifikante Zusammenhänge mit dem Feuchtklebergehalt (0,36) und dem Brotvolumen (0,35). Weitere moderate Korrelationen traten zwischen  $\gamma$ -Gliadinen 1, 3, 4, 9, 13, 15, 17 und 23 mit den Parametern Sedimentationswert und Feuchtklebergehalt auf. Negative Korrelationen, im Bereich von  $-0,10$  bis  $-0,20$ , wurden hauptsächlich für den Glutenindex und die Wasserabsorption festgestellt, insbesondere bei  $\gamma$ -Gliadin 3, 4, 11, 14, 17 und 21. Diese Ergebnisse deuten

darauf hin, dass bestimmte  $\gamma$ -Gliadine, insbesondere  $\gamma$ -Gliadin 11 und 21, einen stärkeren Einfluss auf die Backqualität haben als andere.

Insgesamt fällt auf, dass die  $\gamma$ -Gliadine mehr und stärkere Korrelationen mit den Backqualitätsparametern zeigten als die  $\alpha$ -Gliadine. Zudem traten innerhalb der  $\gamma$ -Gliadine deutlichere Unterschiede zwischen den einzelnen Untergruppen auf, was darauf hindeutet, dass bestimmte  $\gamma$ -Gliadine stärker mit der Backqualität assoziiert sind als andere. So scheinen einige Gruppen generell in einem positiven Zusammenhang zu Qualitätsparametern zu stehen, da sie mehrere positive Korrelationen aufwiesen. Besonders hervorzuheben war in dieser Gruppe das  $\gamma$ -Gliadin 11, welches mehrere der höchsten Korrelationen zu Qualitätsmerkmalen aufwies. Diese Gruppe besteht aus sieben Proteinen, die vermutlich vom Chromosom 1B stammen. Sechs dieser Proteine besitzen die für  $\gamma$ -Gliadine typischen acht Cysteinreste, während ein Protein dieser Gruppe (UniProtKB-ID: B6UKQ2) neun Cysteinreste aufweist. Diese zusätzliche Thiolgruppe ermöglicht eine weitere Disulfidbrücke, die jedoch aufgrund der ungeraden Anzahl nur intermolekular gebildet werden kann. Eine solche Vernetzung mit anderen Proteinen könnte zur verstärkten Beteiligung am Glutennetzwerk beitragen und damit die hohe Korrelation dieser Gruppe mit Backqualitätsparametern erklären. Auch  $\gamma$ -Gliadin 21 zeigte eine starke Assoziation mit dem Kornproteingehalt (0,38). Interessanterweise enthält auch diese Gruppe ein Protein mit neun Cysteinresten (UniProtKB-ID: D0ES81), was ebenfalls auf ein potenziell erhöhtes Vernetzungspotenzial hinweist. Allerdings steht diese Theorie im Gegensatz zur Literatur, in der Gliadine mit einer ungeraden Anzahl an Cysteinresten als Terminatoren des Glutennetzwerks angesehen werden, wie es bereits bei den  $\alpha$ -Gliadinen beschrieben wurde [16].

Liu et al. entdeckten ein Allel auf *Gli-y1-1D*, welches mit einer verbesserten Backqualität assoziiert war [153]. Nach Zhou et al. verbesserte die Zugabe des  $\gamma$ -Gliadinfragments TaGli-y-2.1 die Teigeigenschaften von Gluten [154]. Sherman et al. berichteten, dass ein bestimmtes Allel auf dem *Gli-B1* Locus die Backqualität von Hartweizen verbessern kann, um daraus die Herstellung von Vollkorn-Sauerteigbrot zu ermöglichen [155].  $\gamma$ -Gliadin 11 stammt höchstwahrscheinlich von Chromosom B und wies die höchsten positiven Korrelationen zur Backqualität auf. Deshalb ist es plausibel, dass die Inkorporation bestimmter  $\gamma$ -Gliadine auch bei anderen Weizenarten deren Backeigenschaften verbessern kann. Da auch andere Gliadine positive Korrelationen mit Qualitätsparametern zeigten, ist eine standardisierte genetische Nomenklatur sowie eine kuratierte Erfassung in öffentlich zugänglichen Datenbanken dringend erforderlich, um einen zuverlässigen Vergleich mit bestehenden Literaturdaten zu ermöglichen.

ERGEBNISSE UND DISKUSSION

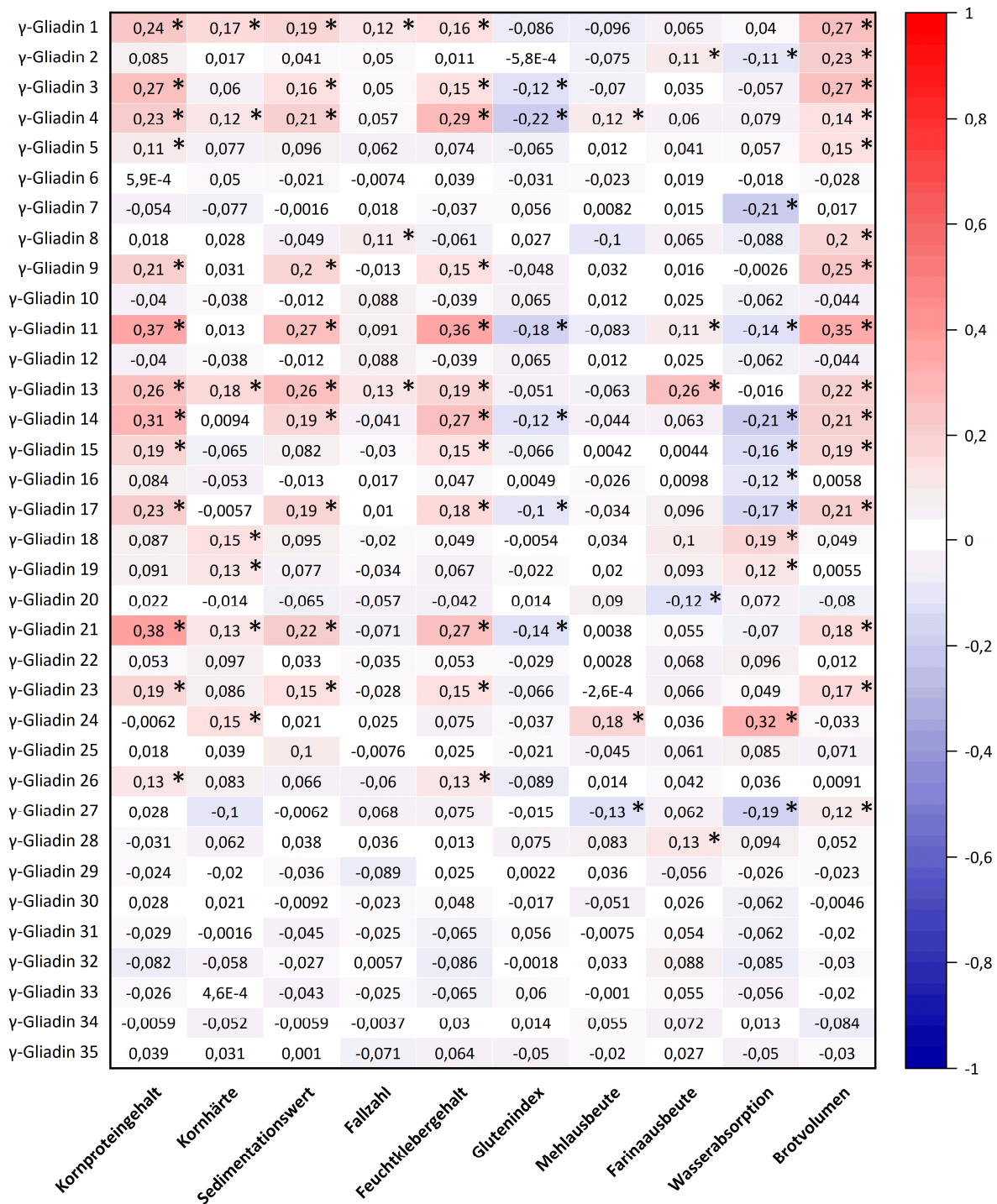


Abbildung 24: Korrelationsmatrix der relativen γ-Gliadin-LFQ-Intensitäten und Backqualitätsdaten der BMWpop 2018. \*p ≤ 0,05; LFQ, labelfreie Quantifizierung.

Ähnlich wie bei den  $\gamma$ -Gliadinen zeigten auch die LMW-GS die meisten signifikanten Korrelationen mit den Backqualitätsmerkmalen Kornproteingehalt, Sedimentationswert, Feuchtklebergehalt und Brotvolumen (**Abbildung 25**). Diese Zusammenhänge waren überwiegend schwach positiv ausgeprägt bzw. generell schwächer ausgeprägt als bei den  $\gamma$ -Gliadinen. Beim Brotvolumen traten jedoch auch einzelne schwach negative Korrelationen auf, ebenso wie beim Glutenindex und der Wasserabsorption. Es gab kaum Korrelationen zwischen LMW-GS 32–47 und den Qualitätsparametern. Auffällig sind die Proteingruppen LMW-GS 2, 7, 9, 11, 15, 16, 18 und 27, die alle schwach positiv mit den Parametern Kornproteingehalt, Sedimentationswert, Brotvolumen und teilweise noch mit dem Feuchtklebergehalt korrelierten. Sie scheinen sich damit am stärksten positiv auf die Qualitätsparameter auszuwirken. Im Gegensatz dazu stehen die LMW-GS Gruppen 21, 25, 26, 29 und 31, welche alle einen leicht negativen Zusammenhang zum Brotvolumen zeigten. Die deutlichsten positiven Korrelationen zeigten sich beim Kornproteingehalt mit den LMW-GS 2 (0,30), 11 (0,30) und 18 (0,34). Darüber hinaus wurde eine Korrelation von 0,31 zwischen LMW-GS 43 und der Wasserabsorption festgestellt. Bei dem Parameter Brotvolumen wurden die stärksten Zusammenhänge mit LMW-GS 1 (0,26), 11 (0,28) und 15 (0,28) beobachtet.

Wie bei den  $\gamma$ -Gliadinen deuten diese Ergebnisse darauf hin, dass bestimmte LMW-GS-Proteine stärker mit der Ausprägung von Qualitätsparametern assoziiert sind als andere. Dies spricht für eine funktionelle Differenzierung innerhalb der LMW-GS-Familie. Besonders Proteine wie LMW-GS 11, LMW-GS 15 und LMW-GS 18 scheinen aufgrund ihrer positiven Korrelationen mit mehreren Parametern in positivem Zusammenhang mit der Backqualität zu stehen. Sie gehören alle dem m-Typ an, wobei LMW-GS 15 auch Proteine des s-Typs enthält. Der m-Typ beginnt mit einem Methionin und zeigt typische N-terminale Sequenzen wie METSHIPGL-, METSRIPGL- oder METSCIPGL- [156]. Hier wurde in einigen Fällen auch die N-terminale Sequenz METSCIS- gefunden. Zur Gruppe LMW-GS 11 zählen drei Proteine, von denen zwei acht und eines sechs Cysteinreste aufweisen. Die Gruppe LMW-GS 15 ist mit insgesamt 42 Proteinen vergleichsweise groß; eine manuelle Analyse der ersten sechs Sequenzen ergab, dass alle acht Cysteinreste enthalten. LMW-GS 18 umfasst ebenfalls zwei Proteine mit jeweils acht Cysteinresten. LMW-GS 1 zeigte im Vergleich eine relativ hohe Korrelation mit dem Brotvolumen (0,26) und gehört ebenfalls zum m-Typ der LMW-GS mit acht Cysteinresten. Dies ist typisches Strukturmerkmal für funktionelle LMW-GS. Zwei dieser Cysteinreste befinden sich in flexiblen Bereichen der Polypeptidkette und fördern die Ausbildung intermolekularer Disulfidbrücken, was zur Verlängerung der Gluteninpolymere beiträgt [157]. Deshalb besteht die generelle Annahme, dass LMW-GS einen positiven Einfluss auf Backqualität haben können. Stichprobenartige Überprüfungen zeigten jedoch, dass auch andere LMW-Gruppen des m-Typs ohne

ERGEBNISSE UND DISKUSSION

signifikante positive Korrelation zur Backqualität häufig Proteine mit acht Cysteinresten enthielten, sodass ein genereller Zusammenhang nicht angenommen werden kann.

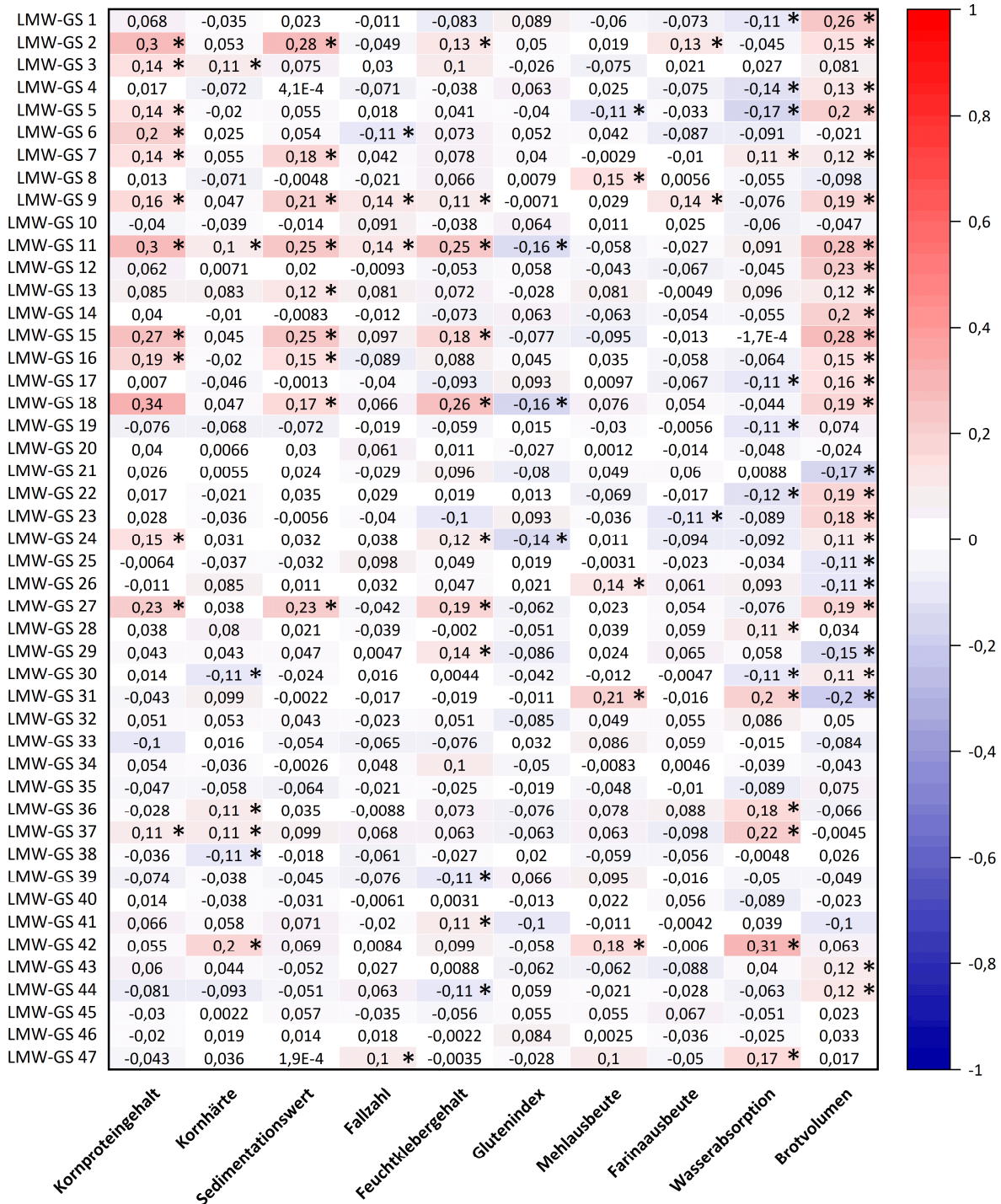


Abbildung 25: Korrelationsmatrix der relativen LMW-GS-LFQ-Intensitäten und Backqualitätsdaten der BMWpop 2018. \*p ≤ 0,05; LMW-GS, niedermolekulare Gluteninuntereinheit; LFQ, labelfreie Quantifizierung.

Fünf Gruppen (LMW-GS 8, 26, 31, 33 und 39) konnten ausschließlich den LMW-GS des s-Typs zugeordnet werden. Sechs Gruppen konnten sowohl s- als auch m-Typ zugeordnet werden (LMW-GS 3, 6, 9, 15, 16 und 46). Nur die gemischten Gruppen wiesen geringe Korrelationen zur Backqualität auf. Daher lässt sich zwar nicht eindeutig feststellen, welcher Typ den beobachteten Einfluss auf die Backqualität ausübt. Es liegt jedoch nahe, dass der m-Typ dafür verantwortlich ist, da auch Gruppen, die ausschließlich diesen Typ enthalten, höhere Korrelationen mit Backqualitätsparametern zeigten. Der s-Typ ist durch ein Serin an der N-terminalen Position gekennzeichnet, zeigt typischerweise Sequenzmotive wie SHIPGL- und stellt die häufigste Untergruppe innerhalb der B-Gruppe der LMW-GS dar, gefolgt von dem m-Typ [156]. Eine Bestätigung ist in diesem Fall nicht möglich, da nach dieser Einteilung lediglich elf LMW des s-Typs, jedoch 35 LMW des m-Typs identifiziert wurden – eventuell bedingt durch die Extraktionsmethode. Bei dem s-Typ wird aufgrund der acht Cysteinreste ebenfalls ein positiver Zusammenhang zur Backqualität gesehen [157]. Weiterhin wurde ein LMW-GS s-Typ-Gen in Zusammenhang mit der Backqualität gebracht [158]. Daraus ergibt sich die Vermutung, dass auch innerhalb dieser Gruppe unterschiedliche Effekte auf die Backqualität bestehen. Da jedoch lediglich fünf LMW-GS-Gruppen identifiziert wurden, die ausschließlich den s-Typ enthielten und keinen Zusammenhang mit der Backqualität zeigten, lässt sich diese Annahme bislang nicht eindeutig belegen. Andere Befunde deuten zudem darauf hin, dass eine eindeutige Differenzierung zwischen m- und s-Typen auf Grundlage der Nukleotidsequenzen nicht möglich ist [34].

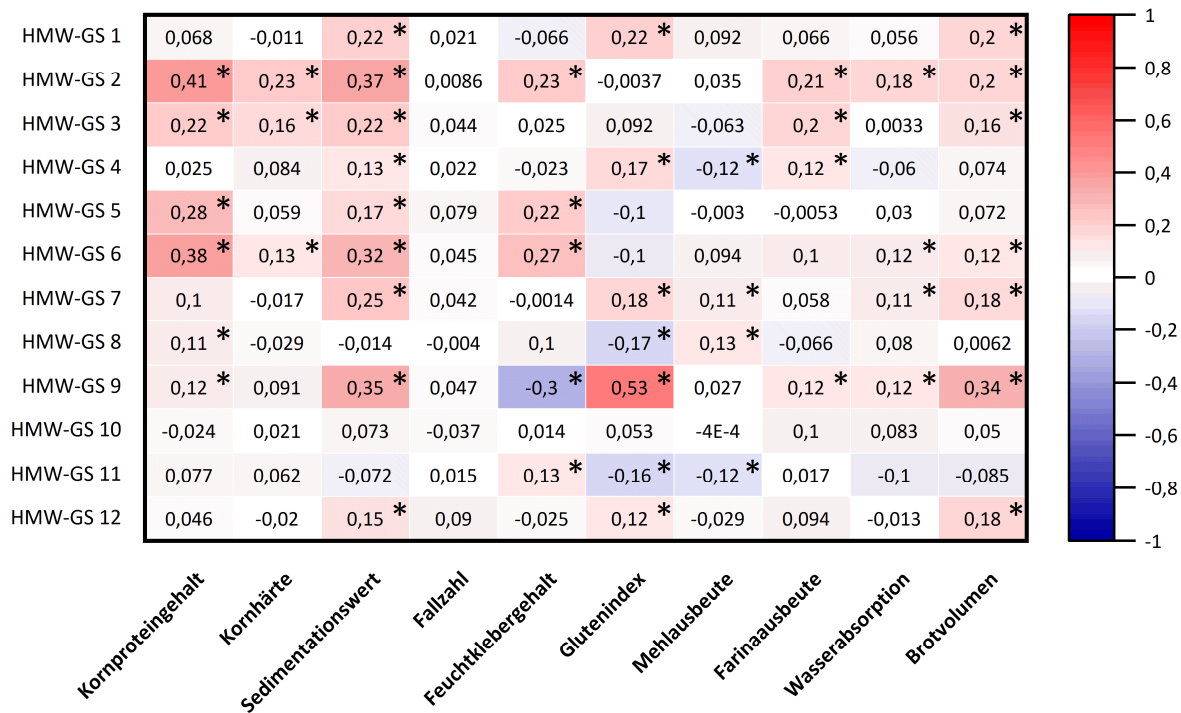
Drei LMW-GS-Gruppen (2, 32 und 41) wurden dem i-Typ zugeordnet, wobei lediglich LMW-GS 2 eine Korrelation mit Backqualitätsparametern zeigte, während für die anderen beiden Gruppen keine signifikanten Zusammenhänge festgestellt wurden. Alle Proteine der LMW-GS-2-Gruppe wiesen jeweils acht Cysteinreste auf. In der LMW-GS 32-Gruppe enthielten einige der insgesamt 45 Proteine nur sieben Cysteinreste, während die Proteine der LMW-GS 41-Gruppe neun Cysteinreste aufwiesen. Ferrante et al. integrierten gezielt den i-Typ in Teige aus Brot- und Hartweizen. Während dies in Hartweizenteig zu einer Abnahme der Teigstärke führte, zeigte der Brotweizenteig hingegen kaum messbaren Effekte [159]. Huang et al. isolierten 14 LMW-GS-Gene des i-Typs aus chinesischen Landweizensorten. Neun dieser Gene codierten für i-Typen mit neun Cysteinresten, die übrigen für Varianten mit acht Cysteinresten. Zudem stellten sie die Hypothese auf, dass die i-Typen mit neun Cysteinresten mit einer reduzierten Backqualität assoziiert sind [160]. Dies deckt sich mit den Ergebnissen der vorliegenden Arbeit, in der ebenfalls gezeigt wurde, dass LMW-i mit ungerader Cysteinanzahl zumindest keinen erkennbaren Einfluss auf die Backqualität haben. Wie bei den Gliadinen scheinen sie durch den zusätzlichen Cysteinrest eher als Terminator zu fungieren [159].

LMW-GS 34 wurde als ein LMW-GS des D-Typs identifiziert, wobei weitere Proteine innerhalb dieser Gruppe den  $\omega$ -Gliadinen zugeordnet wurden. Die D-Gruppe der LMW-GS besteht aus modifizierten  $\omega$ -Gliadinen, welche einen Cysteinrest erworben haben. Aufgrund dieses zusätzlichen Cysteinrests und ihrer abweichenden Elektromobilität im Vergleich zu Gliadinen werden sie den LMW-GS zugeordnet [161]. Wie in der Literatur bereits beschrieben, haben  $\omega$ -Gliadine kaum Einfluss auf die Backqualität. Da auch für LMW-GS 34 nur geringe Korrelationen festgestellt wurden, liegt die Vermutung nahe, dass die gesamte Gruppe der D-LMW-GS ebenfalls keine relevante Wirkung auf die Backqualität hat. Dabei dürfte ebenfalls Ähnliches gelten wie für die meisten Gliadine mit ungerader Anzahl an Cysteinresten, die (wie bereits mehrfach erwähnt) als potenzielle Terminatoren der Gluteninpolymerbildung betrachtet werden [34, 159].

Wang et al. untersuchten die funktionellen Eigenschaften des *Glu-B3h* Allels, welches unter anderen für LMW-GS codiert, und stellten fest, dass das Fehlen zu einem Rückgang von Backqualitätsmerkmalen wie Brotvolumen führte [162]. Weiterhin wurden die *Glu-A3b* und *Glu-B3b* Allele mit stärkerem Gluten und besseren Backeigenschaften assoziiert [163]. Guzman et al. untersuchten den Effekt verschiedener LMW-GS-Allelkombinationen. Die Kombination mit dem höchsten Effekt auf Glutenstärke war *Glu-A3b/Glu-B3i/Glu-D3c*. Dies führte jedoch zu sehr zähem Gluten, während die Kombination *Glu-A3c/Glu-B3d/Glu-D3b* in einem hohen Brotvolumen resultierte. Die häufigste LMW-GS-Kombination *Glu-A3c/Glu-B3h/Glu-D3b* war mit mittlerer Glutenstärke und Brotvolumen assoziiert [62]. Hier wurden ebenfalls spezifische Effekte bei manchen LMW-GS beobachtet, die genetisch noch nicht zugeordnet werden können, weshalb eine Überprüfung mit anderen Ergebnissen aus der Literatur nicht möglich ist.

Die HMW-GS zeigten signifikante Korrelationen mit nahezu allen untersuchten Backqualitätsparametern; mit Ausnahme der Fallzahl, zu der keine relevanten Zusammenhänge festgestellt wurden (**Abbildung 26**). Besonders ausgeprägt waren die Korrelationen zu Kornproteingehalt, Sedimentationswert, Glutenindex und Brotvolumen, während für Kornhärte, Mehl- und Farinausbeute sowie Wasserabsorption überwiegend geringe oder keine Korrelationen beobachtet wurden. Die meisten Zusammenhänge waren positiv und bewegten sich im Bereich von 0,20 bis 0,40, vereinzelt traten jedoch auch negative Korrelationen, insbesondere mit Feuchtklebergehalt, Glutenindex und Mehlausbeute, auf. Während die Mehrheit der HMW-GS-Gruppen klare Zusammenhänge aufwies, blieb HMW-GS 10 (Bx6, Bx7) vollständig unauffällig, und auch HMW-GS 11 (Ay) und 12 (Ax2\* Fragment) zeigten nur sehr schwache Korrelationen. Besonders auffällig war die Gruppe HMW-GS 9, die mit einem Korrelationskoeffizienten von 0,53 zum Glutenindex die stärkste Einzelkorrelation der gesamten Untersuchung zeigte. Mit dem Brotvolumen

(0,34) und Sedimentationswert (0,35) zeigten sich ebenfalls deutliche positive Zusammenhänge. Dabei ist zu vermuten, dass die Gruppe von dem *Dy10* Allel stammt. Wang et al. analysierten gezielt den Einfluss der *Dy10*-Untereinheit auf die Teigeigenschaften von Weizen, indem sie einen Mutanten mit einem *Dy10* Nullallel erzeugten, bei dem keine *Dy10*-Expression erfolgte. Das Fehlen von *Dy10* reduzierte das Glutenin zu Gliadinverhältnis und wirkte sich negativ auf die Teigeigenschaften aus, welche sie unter anderem mit Sedimentationswert und Glutenindex bestimmten [164]. Die hier erhaltenen Ergebnisse stimmen mit dieser Beobachtung überein und stützen die ausgeprägten Effekte von *Dy10* auf Glutenindex und Sedimentationswert. Sie zeigen, dass die *Dy10*-Untereinheit wesentlich für die Aufrechterhaltung von Teigstärke ist, was sich auch durch den stärksten Zusammenhang mit dem Brotvolumen (0,34) äußerte. Einen vermeintlichen Widerspruch hierzu bildet jedoch die negative Korrelation von HMW-GS 9 mit dem Feuchtklebergehalt (-0,3).



**Abbildung 26:** Korrelationsmatrix der relativen HMW-GS-LFQ-Intensitäten und Backqualitätsdaten der BMWpop 2018. \* $p \leq 0,05$ ; HMW-GS, hochmolekulare Gluteninuntereinheit; LFQ, labelfreie Quantifizierung.

Moderate Korrelationen fanden sich zudem bei HMW-GS 2, insbesondere mit Kornproteingehalt (0,41) und Sedimentationswert (0,37), was auf eine funktionelle Relevanz der Untereinheit *Dy12* hinweist. Weiterhin wurde für HMW-GS 2 die höchste Korrelation mit der Kornhärte festgestellt (0,23) und die zweithöchste mit dem Feuchtklebergehalt (0,2). Wang et al. (2024) zeigten, dass die Kombination

Dx2+Dy12 den Kornproteingehalt, den Härteindex sowie den Feuchtklebergehalt signifikant erhöht [165]. Dennoch wird diese Kombination in anderen Arbeiten mit einer schwachen Backqualität assoziiert [62, 166]. Farahani et al. untersuchten den Einfluss von HMW-GS aus 28 Brotweizensorten auf verschiedene Backqualitätsparameter. Dabei zeigte sich, dass insbesondere Ax2\* sowie Dx5+Dy10 einen deutlich stärkeren positiven Effekt auf die Backeigenschaften hatten als andere Untereinheiten, während das Null-Allel und Dx2+Dy12 mit einer geringeren Backqualität in Verbindung standen [167]. Ergänzend dazu weisen Ergebnisse aus der Analyse von HMW-GS 2 darauf hin, dass Dy12 zwar einen positiven Einfluss auf die Backqualität haben kann, dieser jedoch nicht so stark ausgeprägt ist wie bei Dy10 (HMW-GS 9). Diese Befunde stützen die häufig in der Literatur zu findende Einordnung der Kombination Dx2+Dy12 als eher ungünstig im Hinblick auf die Backqualität.

HMW-GS 5 und HMW-GS 6, die zu By9 und By8 gehören, zeigten schwache, aber konsistente Zusammenhänge mit mehreren Qualitätsparametern. Karaduman et al. beschrieben, dass die Kombination Bx7+By8 einen verbessernden Effekt auf die Backqualität ausübte als Bx7+By9, wobei die Untereinheit Bx7 mit schwacher Teigstärke assoziiert wurde [168]. Dies passt zu den Ergebnissen dieser Arbeit, da By8 (HMW-GS 6) stärkere Korrelationen zur Backqualität zeigte als By9 (HMW-GS 5). Im Gegensatz dazu brachte eine weitere Studie die Kombination Bx6 und By8 mit reduziertem Kornprotein- und Feuchtklebergehalt in Verbindung, während Bx7+By9 mit höherer Weizenqualität assoziiert wurde [62, 169].

Weitere Gruppen wie HMW-GS 3 (Dx2, Dx5), HMW-GS 1 und HMW-GS 7 (beide Ax1, Ax2\*) wiesen schwache Korrelationen insbesondere mit dem Sedimentationswert, dem Glutenindex und dem Brotvolumen auf. Obwohl Ax1 und Ax2 in früheren Studien mit guten Qualitätseigenschaften assoziiert wurden [62, 167, 168], konnte dieser Zusammenhang in der vorliegenden Analyse nicht eindeutig bestätigt werden.

Einzelne Gruppen wie HMW-GS 12 (Ax) und HMW-GS 10 (Bx6/Bx7) zeigten entweder nur sehr schwache oder gar keine signifikanten Korrelationen mit Qualitätsparametern. Besonders Bx7 war bereits in früheren Studien mit einer geringen Teigstärke assoziiert worden [168]. Da es sich bei HMW-GS 12 um ein Fragment handelt, das nur in sehr niedrigen Mengen detektiert wurde (**Abbildung 22**), wird dieses als Artefakt angesehen und in der weiteren Auswertung nicht mehr berücksichtigt. Die einzige Gruppe mit ausschließlich schwach negativen Korrelationen war HMW-GS 11, welche dem Ay-Typ zugeordnet wird. Möglicherweise lag das an der generell sehr geringen LFQ-Intensität (**Abbildung 22**). Dies deutet darauf

hin, dass es kaum oder nur von wenigen Weizenlinien exprimiert wurde und wahrscheinlich auch deshalb keinen Effekt ausübte. Möglicherweise wurde sie jedoch auch fehlerhaft identifiziert oder annotiert.

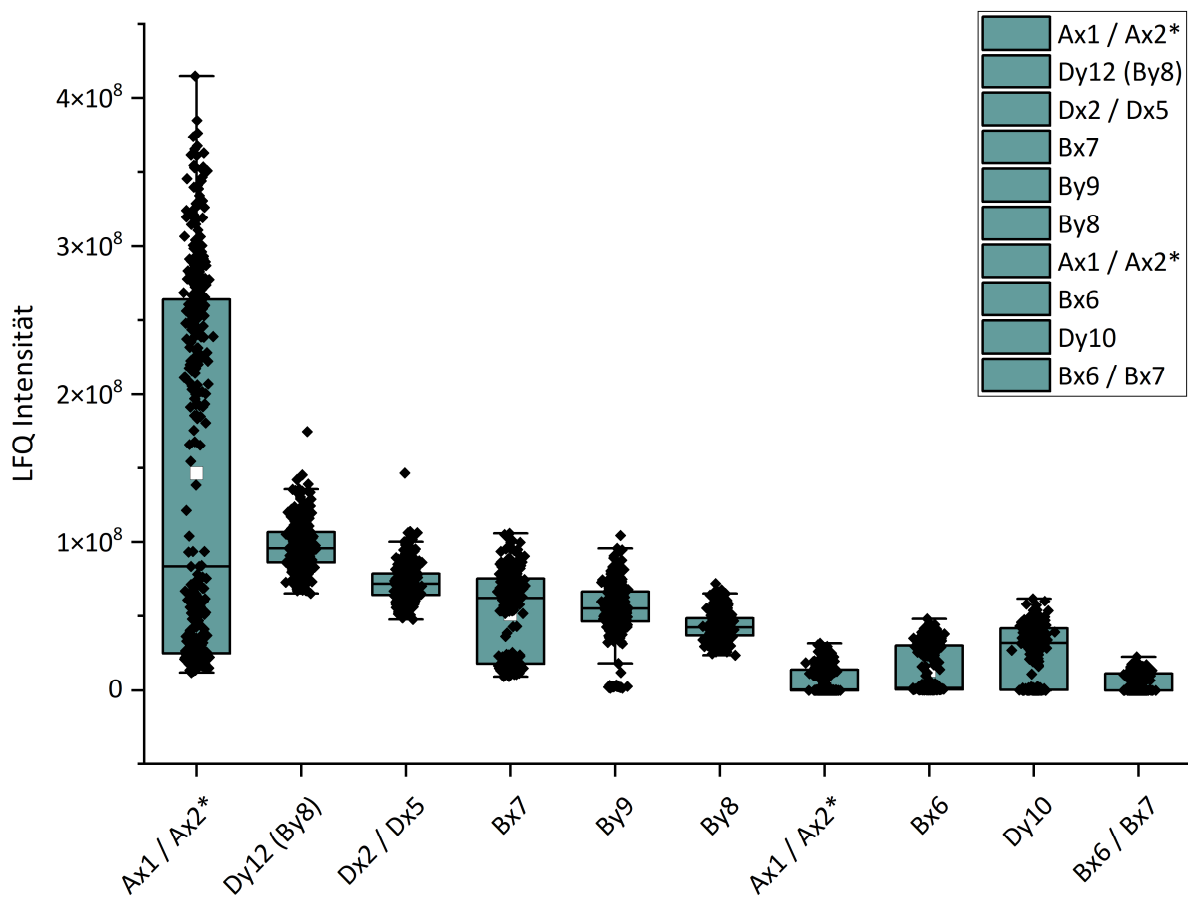
Auf Grundlage der Ergebnisse lässt sich eine Rangfolge der HMW-GS hinsichtlich ihres Einflusses auf die Backqualität ableiten: Den stärksten Effekt zeigte Dy10, gefolgt von Dy12, By8 und By9. Danach folgen die x-Typen in der Reihenfolge Dx2/Dx5, Ax1/2\*, Bx7, Ax1/2\* sowie Bx6. Den geringsten Einfluss wies die Ay-Untereinheit auf. Dies deutet darauf hin, dass die y-Typen einen stärkeren Einfluss auf die Backqualität ausüben als die x-Typen.

In der Literatur wird der Einfluss der HMW-GS meist auf Ebene der Allel- bzw. Untereinheitenkombinationen untersucht. Besonders häufig wurden dabei die Kombinationen Dx5+Dy10 und Dx2+Dy12 hinsichtlich ihrer Wirkung auf die Backqualität von Weizen vergleichend analysiert. Die Grundlage hierfür bildete eine Studie von Payne et al. (1987) die zeigte, dass Dx5+Dy10 signifikant bessere Backqualitätsmerkmale, insbesondere hinsichtlich Brotvolumen und Teigstärke, aufweist als Dx2+Dy12 [60]. Spätere Studien bestätigten diese Ergebnisse weitgehend. So zeigten Horvat et al. und Liang et al. ebenfalls die zentrale Bedeutung von Dx5+Dy10 für gute Teigeigenschaften [170, 171].

Eine Begründung für die hohe Glutenstärke der Kombination Dx5+Dy10 gegenüber Dx2+Dy12 ist, dass Dx5 mehr Cystein enthält als Dx2 [29]. Während Dx2 vier Cysteinreste enthält, weist Dx5 ein weiteres Cystein direkt am Übergang zur repetitiven Domäne auf [172]. Dies resultiert in mehr intermolekularen Disulfidbindungen,  $\beta$ -Schleifen und intermolekularen  $\beta$ -Faltblättern und damit einem stärkeren Glutennetzwerk [172, 173]. Interessanterweise zeigte die vorliegende Analyse, dass Dy10 und Dy12 einen insgesamt stärkeren Einfluss auf die Backqualität hatten als Dx2 und Dx5. Dies steht teilweise im Widerspruch zu früheren Studien, die den x-Typ-Untereinheiten eine dominantere Rolle zuschreiben. Anderson et al. stellten beispielsweise fest, dass die Zugabe von Dx-Untereinheiten einen stärkeren Einfluss auf Teigeigenschaften ausübte, als die Zugabe von Dy-Untereinheiten [61]. Im Gegensatz zu den x-Typen enthalten y-Typ-Untereinheiten in der Regel mehr Cysteinreste – fünf im N-terminalen, einen im repetitiven und einen im C-terminalen Bereich. Dadurch sind sie besser in der Lage, intra- und intermolekulare Disulfidbrücken zu bilden und könnten daher eine wichtigere Rolle bei der Verbesserung der Backqualität spielen [174, 175]. Diese Annahme wird durch die vorliegenden Ergebnisse gestützt.

Da die x- und y-Typen unterschiedlich stark exprimiert werden, könnte es ebenfalls sein, dass die y-Typen in der BMWpop stärker exprimiert wurden und dadurch einen größeren Effekt ausübten. **Abbildung 27** zeigt die LFQ-Intensitätsverteilung der einzelnen HMW-GS im Detail. Dabei ist erkennbar, dass Dx2 und Dx5 in höherem Umfang vorliegen als Dy10, welches vergleichsweise geringe Intensitäten aufwies und

damit zu den am schwächsten detektierten Untereinheiten zählt. Dy12 hingegen zeigte die zweithöchsten Intensitäten. Dies könnte teilweise darauf zurückzuführen sein, dass auch By8-Untereinheiten dieser Gruppe zugeordnet wurden. Insgesamt lässt sich aus diesen Daten ableiten, dass die Unterschiede in der Wirkung der HMW-GS auf die Backqualität nicht primär durch unterschiedliche Proteinmengen erklärt werden können. Im Gegenteil: Dy10 zeigte trotz seiner vergleichsweise geringen LFQ-Intensität den stärksten positiven Einfluss auf die Backqualität.



**Abbildung 27:** Verteilung der LFQ-Intensitäten der HMW-GS in den 403 Weizenlinien der BMWpop. Boxplots zeigen den Interquartilbereich (Box), den Mittelwert (Quadrat), den Median (horizontale Linie) sowie Minima und Maxima (Whisker). LFQ, labelfreie Quantifizierung.

Um potenzielle funktionelle Unterschiede zwischen den y-Typen besser zu verstehen, wurde ihre Aminosäuresequenz verglichen. Dy10 und Dy12 weisen dabei nur minimale strukturelle Abweichungen auf (**Abbildung 28**), sondern sind in weiten Teilen nahezu identisch; Dy12 ist lediglich um 12 Aminosäuren länger, was auf zusätzliche Wiederholungen in der zentralen repetitiven Domäne zurückzuführen ist. Abgesehen von einem Cysteinrest in der repetitiven Region befinden sich alle übrigen Cysteinreste an

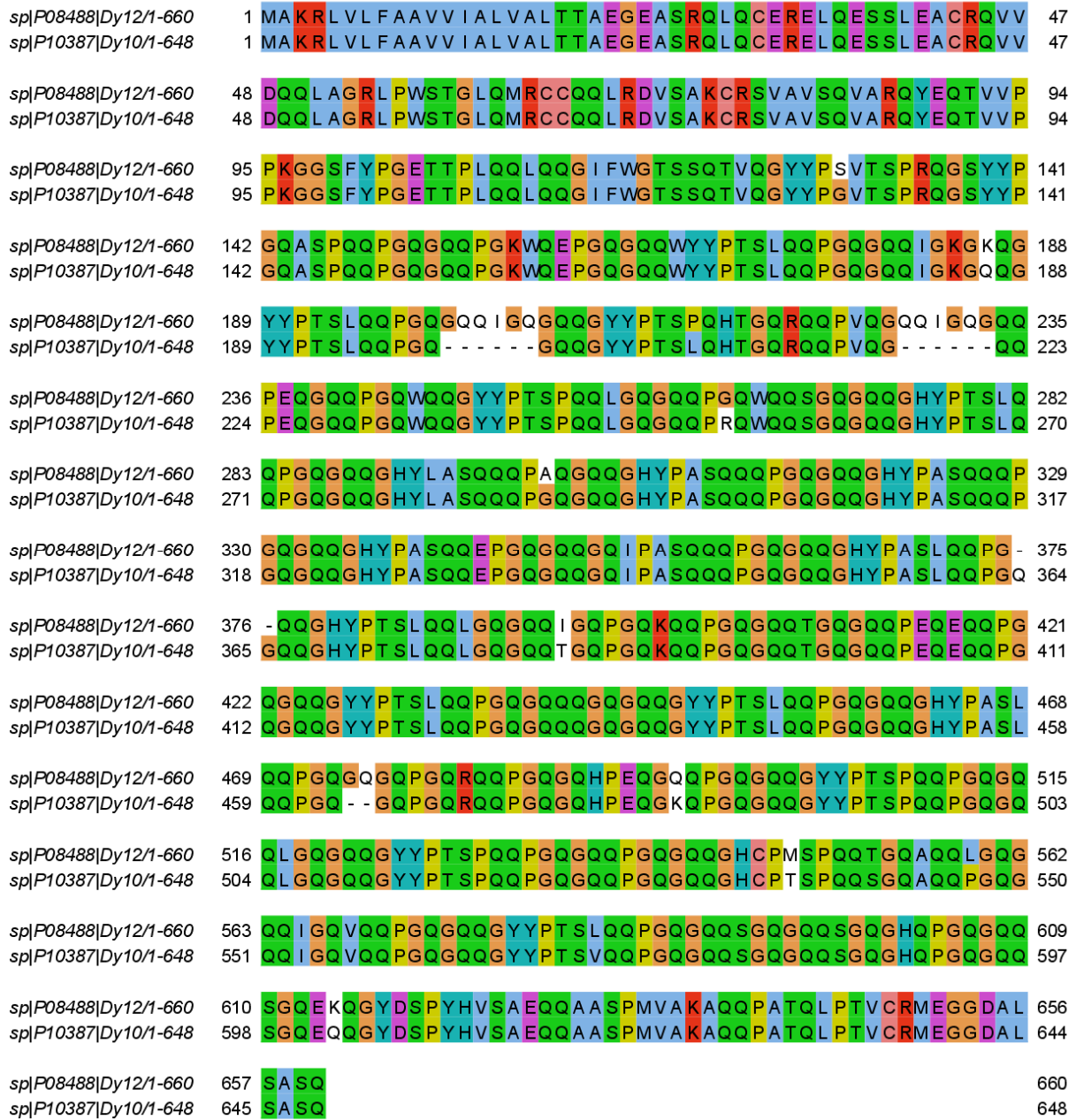
vergleichbaren Positionen. Die Hauptunterschiede betreffen somit die Länge und die feine Sequenzstruktur der repetitiven Domäne. Diese Verlängerung scheint jedoch keinen funktionellen Vorteil zu bringen, sondern könnte vielmehr die räumliche Anordnung der Cysteinreste so verändern, dass vermehrt intramolekulare statt intermolekulare Disulfidbrücken gebildet werden. Dies würde die Vernetzungsfähigkeit des Glutenpolymers und damit die Teigstabilität einschränken.

Es ist möglich, dass die Anwesenheit von By8 die Korrelationen von Dy12 beeinflusst und deren Stärke abgeschwächt hat, obwohl Dy10 und Dy12 aufgrund ihrer strukturellen Ähnlichkeit grundsätzlich vergleichbare Beiträge zur Vernetzung leisten könnten. Um verlässliche Aussagen über die tatsächliche Wirkung von Dy12 zu treffen, müsste dieser Einfluss zunächst ausgeschlossen werden. Die Überlegenheit der Kombination Dx5+Dy10 lässt sich vermutlich durch den zusätzlichen, fünften Cysteinrest in Dx5 erklären, der eine weitere intermolekulare Vernetzung ermöglicht. Besonders starke Effekte auf die Teigeigenschaften wurden zudem beobachtet, wenn x- und y-Typ-Untereinheiten im Verhältnis 1:1 eingesetzt wurden, was eine ausgewogene Polymerbildung begünstigt [61].

Der Vergleich der Aminosäuresequenzen von Dy10, Dy12, By8 und By9 (Anhang, **Abbildung 37**) zeigte, dass auch By8 und By9 eine hohe Sequenzähnlichkeit zu Dy10 und Dy12 aufweisen. Der auffälligste Unterschied lag erneut in der Gesamtlänge der Sequenzen: By8 besteht aus 720, By9 aus 705 Aminosäuren – verursacht durch zusätzliche Wiederholungen in der repetitiven Domäne. Diese Unterschiede könnten strukturelle Auswirkungen haben. Allerdings lässt sich daraus nicht ableiten, dass eine längere Sequenz generell mit einer geringeren Wirkung auf die Backqualität einhergeht: By8 zeigte trotz der größten Sequenzlänge stärkere Effekte als By9. Daher müssen weitere feinstrukturelle Unterschiede zwischen den y-Typ-Untereinheiten für die unterschiedlichen Beiträge zur Backqualität verantwortlich sein.

1Ay übte den geringsten Effekt aus und steht damit im Gegensatz zu den anderen y-Typ Varianten. Wang et al. integrierten das *Glu-1 Ay* Allel aus wildem Emmer in Weichweizen und stellten eine Zunahme des Kornproteingehalts, des Sedimentationswerts und des Feuchtklebergehalts fest [176]. Die allgemein positiven Effekte der y-Typ-Untereinheiten lassen vermuten, dass auch der 1Ay-Typ bei vollständiger Genaktivität einen positiven Einfluss auf die Backqualität ausüben würde.

ERGEBNISSE UND DISKUSSION



**Abbildung 28:** Darstellung des Sequenzvergleichs der HMW-GS Dy12 (P08488) und Dy10 (P10387) mit Jalview [177]. Farbcodierung der Aminosäuren: blau, Alanin, Isoleucin, Leucin, Methionin, Phenylalanin, Tryptophan, Valin (hydrophob); gelb, Prolin (Strukturbrecher); grün, Glutamin, Serin, Threonin (polar, ungeladene Seitenketten); orange, Glycin; rot, Arginin, Lysin (basisch); türkis, Histidin, Tyrosin (polare Seitenkette, Wasserstoffbrücke); violett, Asparaginsäure, Glutaminsäure (sauer); HMW-GS, hochmolekulare Gluteninuntereinheit.

Bei den x-Typen gab es ebenfalls eine absteigende Reihenfolge im Effekt auf die Backqualität:  $Dx2/Dx5 > Ax1/2^* > Bx7 > Bx6$ . In der HMW-GS 3 Gruppe sind Gluteninuntereinheiten vorhanden, die sowohl vom  $Dx2$  als auch vom  $Dx5$  Allel stammen. Somit war eine eindeutige Differenzierung und Zuordnung der

beiden Untereinheiten zu separaten Gruppen nicht möglich. Offenbar wurden keine Peptide nachgewiesen, die eindeutig nur einem der beiden Subtypen zuzuordnen sind. Allerdings zeigte der Sequenzvergleich der in der HMW-GS 3-Gruppe enthaltenen Proteine, dass sieben von 33 Sequenzen einen zusätzlichen Cysteinrest unmittelbar vor Beginn der repetitiven Domäne aufweisen und somit dem HMW-GS Dx5 zugeordnet werden können (Anhang, **Tabelle 19**). Die Effekte dieser Gruppe auf die Backqualitätsparameter waren im Vergleich zu den übrigen x-Typen am stärksten ausgeprägt, was – wie bereits erläutert – vermutlich auf die Präsenz der Dx5-Untereinheit mit ihrem zusätzlichen Cysteinrest zurückzuführen ist. Eine abschließende Bestätigung ist erst dann möglich, wenn die beiden Untereinheiten eindeutig differenzierbar sind.

Die Untereinheiten Bx7 und Bx6 zeigten die schwächsten Zusammenhänge mit der Backqualität, wobei insbesondere Bx7 bereits zuvor mit geringer Teigstärke in Verbindung gebracht wurde [168]. Der Vergleich der Aminosäuresequenzen (Anhang, **Abbildung 38**) offenbarte auch hier eine hohe strukturelle Ähnlichkeit zwischen Bx- und Dx-Untereinheiten. Allerdings waren die Unterschiede zwischen den Dx- und Bx-Typen ausgeprägter als bei den jeweiligen  $\gamma$ -Pendants. Diese lagen vor allem in der Länge und Zusammensetzung der repetitiven Domäne: Bx-Typen wiesen mit 824 Aminosäuren eine etwas kürzere Gesamtlänge auf als Dx-Typen (838 bzw. 848 Aminosäuren). Die Unterschiede in der Backqualität sind daher vermutlich auch hier auf feinstrukturelle Merkmale zurückzuführen. Besonders hervorzuheben ist der nachgewiesene positive Effekt des zusätzlichen Cysteinrests bei Dx5. Möglicherweise ließe sich dieses Merkmal künftig gezielt auf andere x-Typ-Untereinheiten übertragen, etwa im Rahmen genetischer Modifikation.

Ax1 und Ax2\* konnten anhand der Daten nicht eindeutig innerhalb der HMW-GS getrennt werden und gehören zu den HMW-GS 1 und 7. Sie hatten allerdings beide einen ähnlichen Effekt auf Backqualitätsdaten, wobei HMW-GS 1 durch die größte Variabilität und die höchsten detektierten Mengen auffiel, was auf eine starke Expression dieser Untereinheiten hindeutet (**Abbildung 27**). Ihr positiver Einfluss auf die Backqualität könnte folglich mit ihrer hohen Proteinmenge zusammenhängen. Sie zeigten ebenfalls strukturelle Ähnlichkeiten in ihrer AS-Sequenz zu anderen x-Typ-Untereinheiten, insbesondere zu den Bx-Typen (Anhang, **Abbildung 38**). Mit Aminosäurelängen von 815 bzw. 830 Aminosäuren liegen sie größtenteils zwischen den Dx- und Bx-Typen.

### 3.4.1.3 Zusammenfassende Diskussion

Auf Grundlage der relativen LFQ-Intensitäten aus der MaxQuant-Auswertung wurden verschiedene Glutenproteinfraktionen hinsichtlich ihrer Korrelation zu Backqualitätsparametern analysiert. Die Ergebnisse ermöglichen eine differenzierte Bewertung der einzelnen Glutenproteintypen im Hinblick auf ihre potenzielle Bedeutung für die Backqualität.

Auf alle Backparameter bezogen korrelierten die Glutenproteingruppen am stärksten mit den Parametern Kornproteingehalt, Sedimentationswert, Feuchtklebergehalt, Glutenindex und Brotvolumen. Mit Kornhärte, Fallzahl, Mehl- und Farinaausbeute sowie Wasserabsorption gab es nur sehr schwach signifikante oder keine Korrelationen. Der Einfluss der Glutenproteine auf diese Parameter scheint also nur sehr gering zu sein. Das Endosperm besteht hauptsächlich aus Stärkekörnern und Glutenproteinen. In diesem Zusammenhang wurde jedoch ein Protein (Friabilin) entdeckt. In Weichweizen ist dieses Protein mit der Oberfläche der Stärkekörner verbunden. Die gleiche Menge wurde in Hartweizen gefunden, wo man annimmt, dass es direkt mit den Proteinen assoziiert ist [178, 179]. Hogg et al. zeigten zudem, dass die Speicherproteine Puroindolin A (PINA) und Puroindolin B (PINB) die Kornhärte beeinflussen. Sie stellten die Hypothese auf, dass beide Proteine miteinander interagieren, um Friabilin zu bilden, und dadurch gemeinsam die Textur des Weizenkorns mitbestimmen [180]. Der exakte Mechanismus ist noch nicht aufgeklärt, allerdings vermuten Geneix et al. eine Wechselwirkung zwischen PINA und den Gliadinen [181]. Im Fall von Kornhärte haben somit wahrscheinlich andere Proteine wie Friabilin bzw. Puroindoline mehr Einfluss. Sowohl PINA als auch PINB konnten in der untargeted Messung der BMWpop festgestellt werden.

Die Fallzahl ist ein indirektes Maß für die Amylaseaktivität. Amylasen sind Enzyme, die Stärke zu Zuckern abbauen. Eine hohe Fallzahl weist auf geringe Amylaseaktivität hin, eine niedrige auf hohe Aktivität. Wenn zu wenig Stärke abgebaut wird, kann dies zu schlechter Gärung führen. Dagegen führt ein starker Stärkeabbau auch zu Teigabbau und damit schlechterer Backqualität. Folglich ist eine ausgewogene Amylaseaktivität ideal [182]. Daraus ergibt sich, dass diese Enzyme vermutlich einen stärkeren Einfluss als die Glutenproteine auf die Fallzahl ausüben. Spezifische Glutenproteine scheinen hingegen keinen zusätzlichen Einfluss darauf zu haben. In der untargeted Analyse der BMWpop konnten vereinzelt  $\beta$ -Amylasen identifiziert werden.

Mehl- und Farinaausbeute sind im Gegensatz zum Kornproteingehalt vermutlich zu unspezifisch, da die anderen Mehlbestandteile (vor allem Stärke) den Parameter zu stark mitbeeinflussen. Bei der Wasserabsorption ist das gesamte Glutennetzwerk für die Wasserbindung oder die Hydrophilie der beteiligten Glutenproteine verantwortlich. Bei diesem Parameter scheint es allerdings wichtiger zu sein,

die zugegebene Wassermenge für jeden Teig individuell zu bestimmen, um das bestmögliche Ergebnis zu erzielen [183]. Insgesamt ist davon auszugehen, dass neben den Glutenproteinen weitere Einflussfaktoren – insbesondere andere Proteine wie Amylasen und Puroindoline – eine bedeutendere Rolle als Glutenproteine für bestimmte Qualitätsparameter spielen könnten. Diese Proteinarten standen jedoch nicht im Fokus der vorliegenden Untersuchung. Eine vertiefte Analyse ihres Einflusses auf spezifische Qualitätsmerkmale erscheint im Rahmen zukünftiger Studien jedoch vielversprechend.

Bei den Backqualitätsparametern mit vorhandenen Korrelationen (Kornproteingehalt, Feuchtklebergehalt, Sedimentationswert, Glutenindex und Brotvolumen) zeigte der Vergleich der vier Glutenproteinfraktionen folgendes Muster: Die  $\alpha$ -Gliadine zeigten insgesamt die schwächsten Korrelationen, während die  $\gamma$ -Gliadine moderat positive Zusammenhänge aufwiesen, insbesondere die Gruppen mit neun Cysteinresten. Die LMW-GS waren ebenfalls positiv mit verschiedenen Qualitätsparametern korreliert, wobei sich bestimmte Gruppen wie LMW-GS 15 und 18 besonders hervortaten; beide gehörten zu den LMW-GS des m-Typs. Generell konnten jedoch kein typabhängiger Zusammenhang zwischen den LMW-GS und Backqualitätsparametern identifiziert werden. Ein eindeutiger Zusammenhang zwischen der Anzahl an Cysteinresten und der Backqualität konnte nicht festgestellt werden, da die Ergebnisse – etwa bei den  $\gamma$ -Gliadinen – teils widersprüchlich waren und selbst zwischen Proteingruppen mit identischer Cysteinanzahl erhebliche Unterschiede beobachtet wurden. Bei den LMW-i konnte hingegen bestätigt werden, dass Proteine mit einer ungeraden Zahl an Cysteinresten keinen Einfluss auf die Backqualität hatten. Dies deutet darauf hin, dass neben der Anzahl an Cysteinresten weitere strukturelle Merkmale auf molekularer Ebene zur funktionellen Differenzierung innerhalb der Glutenproteintypen beitragen.

Weiterhin war es nicht möglich, die Gliadine und LMW-GS spezifischen Genorten zuzuordnen, was einen direkten Vergleich mit anderen Ergebnissen aus der Literatur nicht ermöglichte. Zum einen besteht ein dringender Bedarf an Datenbanken mit hinterlegten genetischen Informationen und standardisierter Nomenklatur. Darüber hinaus ist eine umfassende Analyse der Gliadin- und LMW-GS-Gene in der gesamten BMWpop-Population erforderlich. Dies ist aufgrund der hohen Komplexität der zugrunde liegenden Genfamilien mit erheblichem Aufwand verbunden.

Die stärksten Effekte wurden für die HMW-GS beobachtet; insbesondere für Dy10 und Dy12, die mit einem Großteil der Backqualitätsparameter in signifikantem Zusammenhang standen. Dies bestätigt die etablierte Annahme, dass HMW-GS die Hauptträger der strukturellen Teigeigenschaften sind. Die hier beschriebenen Befunde stehen in engem Einklang mit zahlreichen Literaturquellen. So wurde bereits von

Payne et al. (1987) die überlegene Wirkung der Kombination Dx5+Dy10 auf die Backqualität beschrieben [26]. Die besonders hohe Korrelation von Dy10 mit Glutenindex und Brotvolumen unterstreicht dessen zentrale Rolle bei der Netzwerkbildung im Teig. Dy12 zeigte ebenfalls moderate positive Effekte, was auf eine gewisse Relevanz auch dieser HMW-GS hinweist, wenngleich sie schwächer ausfallen als bei Dy10. Insgesamt zeigten die  $\gamma$ -Typ-HMW-GS einen stärkeren Einfluss auf die Backqualität als die  $\alpha$ -Typen. Die Verteilung der LFQ-Intensitäten liefert hierfür keine überzeugende Erklärung, da die  $\gamma$ -Typen überwiegend in vergleichbaren oder geringeren Mengen vorlagen als ihre korrespondierenden  $\alpha$ -Typen. Dies dürfte darauf zurückzuführen sein, dass sie mit sieben Cysteinresten mehr Vernetzungsmöglichkeiten bieten als die  $\alpha$ -Typen, die in der Regel nur vier Cysteinreste aufweisen. Die Unterschiede innerhalb der  $\alpha$ - und  $\gamma$ -Typen könnten auf feinstrukturelle Merkmale einzelner Untereinheiten zurückzuführen sein – etwa auf eine vorteilhaftere Positionierung von Cysteinresten oder eine effizientere Integration in das Glutenin-Polymernetzwerk. Wechselwirkungen mit bestimmten LMW-GS oder teigsystembedingte Matrixeffekte könnten ebenfalls zur beobachteten Wirkung beitragen. Insgesamt zeigen die Ergebnisse, dass nicht die Proteinmenge allein, sondern insbesondere die spezifische Sequenzstruktur und Cysteinverteilung der HMW-GS entscheidend für deren Einfluss auf die Backqualität sind. Ein vertieftes Verständnis des molekularen Aufbaus des Glutennetzwerks ist daher unerlässlich [184].

Keine der untersuchten Glutenproteingruppen zeigte eine starke oder sehr starke Korrelation mit der Backqualität, sodass keine belastbaren Aussagen über einen dominierenden Einfluss einzelner Gruppen getroffen werden können. Dies weist darauf hin, dass die Ausprägung der Backqualität nicht primär durch eine einzelne Proteingruppe gesteuert wird und somit keine Gruppe als verlässlicher Indikator dienen kann – ein Ergebnis, das angesichts der komplexen, multigenetischen Natur dieses Merkmals nicht überraschend ist. Die Mehrheit der beobachteten Korrelationen lag im sehr schwachen bis schwachen Bereich ( $r < 0,4$ ). Bei den HMW-GS konnten jedoch vereinzelt moderate Zusammenhänge ( $0,4 \leq r < 0,6$ ) festgestellt werden, die auf potenziell funktionale Beziehungen zwischen spezifischen Proteinen und den Backqualitätseigenschaften hinweisen, sowie die besondere Bedeutung der HMW-GS in diesem Kontext unterstreichen.

Dennoch weisen die Ergebnisse insgesamt darauf hin, dass bestimmte Glutenproteingruppen aller Glutenproteintypen stärker zur Ausprägung der Backqualität beitragen. Dieser Zusammenhang sollte in künftigen molekularbiologischen Studien weiter untersucht und im Rahmen züchterischer Strategien berücksichtigt werden. Die Züchtung wurde bisher neben einer generell hohen Kornausbeute auch auf möglichst hohe Proteinausbeute ausgerichtet. Die Herausforderung besteht hierbei darin, dass diese beiden Parameter (Korn- und Proteinausbeute) normalerweise negativ miteinander korreliert sind. Der

Anbau von Winterweizensorten, die in Deutschland zwischen 1983 und 2014 angebaut wurden, brachte deutliche Fortschritte zutage: Der Kornertrag erhöhte sich um 31,6 %, begleitet von einem moderaten Anstieg des Proteingehalts um 1,5 %. Gleichzeitig verbesserten sich auch wichtige Backqualitätsmerkmale: der Sedimentationswert stieg um 45,4 % und das Brotvolumen nahm um 8,3 % zu. Dies war nur möglich aufgrund neu gezüchteter Weizenkultivare in Kombination mit modernen landwirtschaftlichen Praktiken [185]. Die hier gewonnenen Erkenntnisse können weiter dazu beitragen, dass die Züchtung noch gezielter in Richtung Proteinqualität ausgerichtet wird. So können Sorten entwickelt werden, die trotz geringerem Proteingehalt dennoch gute Backeigenschaften besitzen.

### 3.4.2 Targeted Messung der BMWpop

Die BMWpop aus der Ernte 2018 wurde vollständig aufgearbeitet, mittels targeted Proteomics gemessen und mit Skyline ausgewertet (5.6.9). Die Methodenentwicklung wurde ausführlich in 3.3 “Targeted Proteomics von Weizenmehl” beschrieben. Die Nummerierung der targeted Daten erfolgte unabhängig von der der untargeted Daten. Die entsprechende Zuordnung ist in **Tabelle 4** dargestellt.

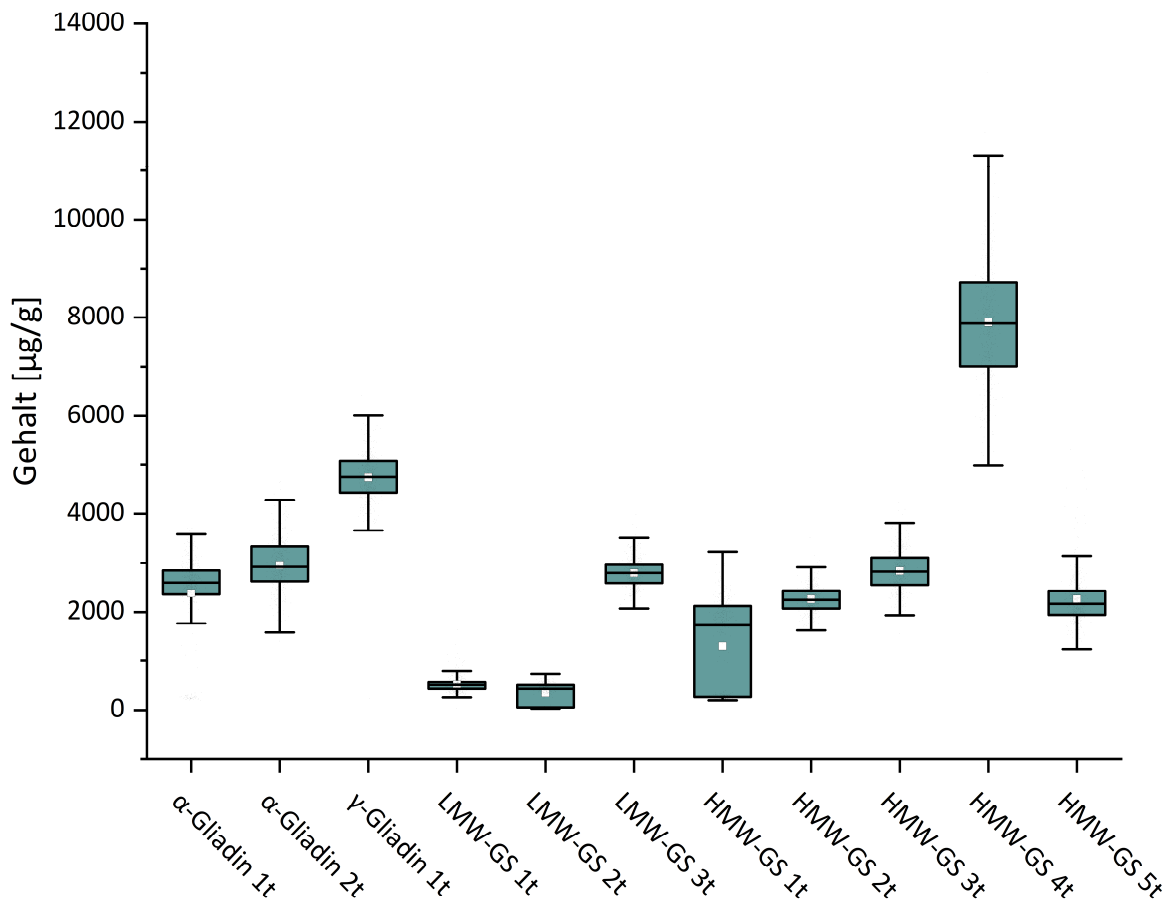
**Tabelle 4:** Zuordnung der targeted Glutenproteingruppen zu den entsprechenden untargeted Glutenproteingruppen. Zur Unterscheidung von der Nummerierung der untargeted Daten wurde ein ‚t‘ ergänzt.

Nummerierung targeted Glutenproteingruppen	Zuordnung untargeted Glutenproteingruppen
α-Gliadin 1t	-
α-Gliadin 2t	α-Gliadin 2
γ-Gliadin 1t	γ-Gliadin 7 & 19
LMW-GS 1t	-
LMW-GS 2t	LMW-GS 12
LMW-GS 3t	LMW-GS 2, 25, 32, 41
HMW-GS 1t	HMW-GS 9
HMW-GS 2t	HMW-GS 1
HMW-GS 3t	HMW-GS 2 & 9
HMW-GS 4t	HMW-GS 2, 5, 6, 9, 11, 12
HMW-GS 5t	HMW-GS 3

HMW-GS, hochmolekulare Gluteninuntereinheiten; LMW-GS, niedermolekulare Gluteninuntereinheiten; t, targeted.

#### 3.4.2.1 Allgemeine Übersicht

Insgesamt lagen die meisten Proteine in einem ähnlichen Gehaltsbereich (2000–4000 µg/g, **Abbildung 29**). Auffällig waren LMW-GS 1t und 2t, welche in vergleichsweise geringen Mengen (<5000 µg/g) gemessen wurden. Im Kontrast steht HMW-GS 4t, welches in sehr hohen Mengen (5000–11000 µg/g) und mit hoher Variabilität in den Weizenlinien festgestellt wurde.



**Abbildung 29:** Gehaltsverteilung der absolut untersuchten Proteine in den 403 Weizenlinien der BMWpop 2018. Boxplots zeigen den Interquartilbereich (Box), den Mittelwert (Quadrat), den Median (horizontale Linie) sowie Minima und Maxima (Whisker). HMW-GS, hochmolekulare Gluteninuntereinheit; LMW, niedermolekulare Gluteninuntereinheit.

Die genetische Zuordnung der  $\alpha$ - und  $\gamma$ -Gliadine sowie der LMW-GS anhand der UniProtKB gestaltete sich analog zu den untargeted Daten schwierig, da entweder keine oder sehr viel redundante und nicht einheitliche Information in der UniProtKB verfügbar war. Zudem ist die manuelle Suche sehr zeitaufwendig. Bei den HMW-GS sind mehr Informationen in der Datenbank hinterlegt. Alle verfügbaren genetischen Informationen zu den Peptiden wurden zusammengefasst (Anhang, **Tabelle 20**).

Die Proteingruppe  $\alpha$ -Gliadin 1t umfasst drei Proteine, die mit hoher Wahrscheinlichkeit vom *Gli-D2* Locus auf dem Chromosom 6D stammen. Bei  $\alpha$ -Gliadin 2t handelt es sich um eine größere Gruppe mit insgesamt 56 Proteinen. Die zugehörigen Genannotationen verweisen überwiegend auf die *Gli-2* Loci auf Chromosom 6, ohne spezifische Lokalisierung. Bei  $\gamma$ -Gliadin 1t, das 34 Proteine umfasst, liegen nur vereinzelt Angaben zu Genloci vor. Vier dieser Proteine lassen sich jedoch explizit dem *Gli-1* Locus auf Chromosom 1D zuordnen. Die Gruppe LMW-GS 1t gehört zu einem Cluster, das mit LMW-GS-Genen auf

Chromosom 1D (*Glu-D3*) assoziiert ist. Von den 13 Proteinen der Gruppe LMW-GS 2t weist lediglich eines die Annotation „D3-4“ auf, was einen Hinweis auf den *Glu-D3* Locus darstellen könnte. Die größte Gruppe bildet LMW-GS 3t mit 72 Proteinen, von denen zahlreiche eine Zuordnung zum *Glu-A3* Locus aufweisen, was darauf hindeutet, dass sie überwiegend diesem Locus entstammen. Für die HMW-GS konnten erneut genetische Zuordnungen vorgenommen werden (**Tabelle 5**). Dabei ist zu beachten, dass die Gruppen HMW-GS 1t, 3t, 4t und 5t sich in Teilen überlappen. Diese Überschneidungen resultieren aus der Auswahl der Peptide für die targeted Analyse, die primär aufgrund ihrer guten MS-Detektierbarkeit gewählt wurden (3.3.1.1). Aufgrund der hohen Sequenzähnlichkeit zwischen verschiedenen Gluteninuntereinheiten enthalten viele dieser Proteine identische Peptidabschnitte, was zu einer Mehrfachzuordnung führen kann. Die Gruppe HMW-GS 4t beispielsweise umfasst sämtliche  $\gamma$ -Typ-Glutene, denen bereits im vorangegangenen Abschnitt zur untargeted Analyse der BMWpop (3.4.1) der größte Einfluss auf die Backqualität zugeschrieben wurde.

**Tabelle 5:** Zuordnung der targeted HMW-GS-Proteingruppen zu spezifischen Untereinheiten.

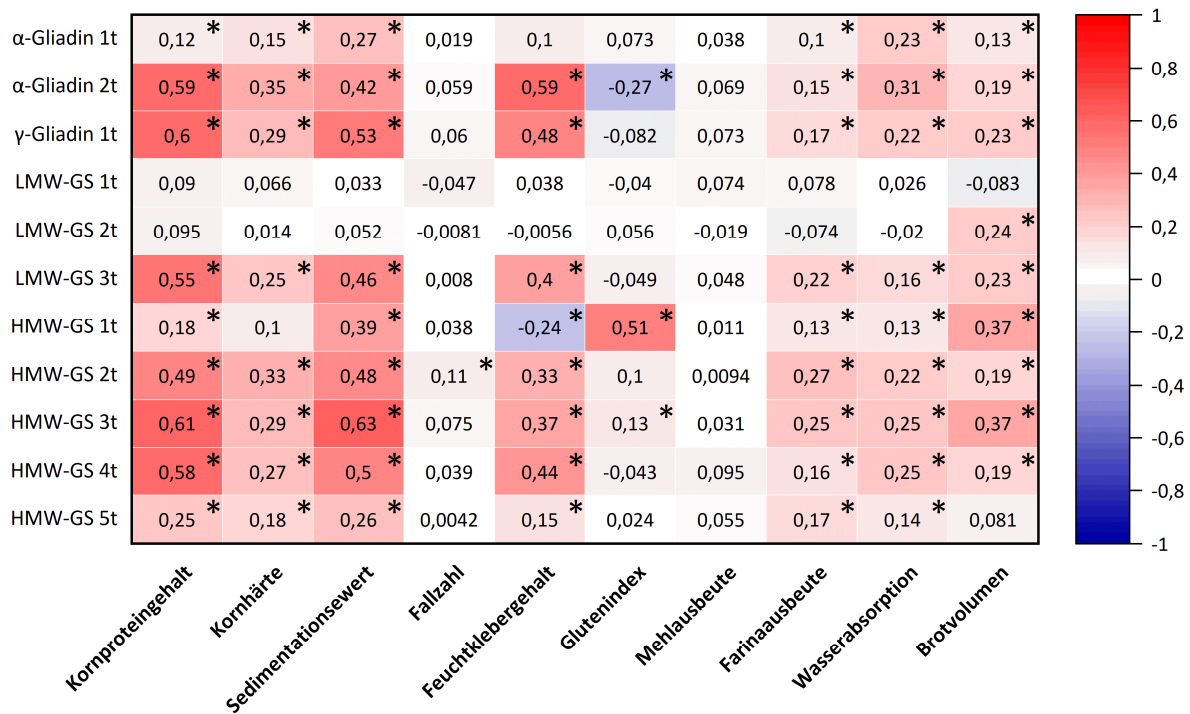
Nr.	Untereinheit
HMW-GS 1t	Dy10
HMW-GS 2t	Ax2*
HMW-GS 3t	Dy10, Dy12, By8
HMW-GS 4t	Dy10, Dy12, By8, By9, 1Ay
HMW-GS 5t	Dx2, Dx5

HMW-GS, hochmolekulare Gluteninuntereinheit; t, targeted.

#### 3.4.2.2 Korrelation mit Backqualitätsdaten und Vergleich mit untargeted Korrelationen

Die targeted Ergebnisse der BMWpop 2018 wurden analog zu den untargeted Ergebnissen mit den Backqualitätsparametern korreliert (**Abbildung 30**). In der Mehrzahl der Fälle bewegten sich die Korrelationskoeffizienten im sehr schwachen ( $r < 0,2$ ) oder schwachen Bereich ( $0,2 \leq r < 0,4$ ). Moderate Korrelationen ( $0,4 \leq r < 0,6$ ) traten etwas öfter auf als bei den untargeted Ergebnissen, während starke Zusammenhänge ( $0,6 \leq r < 0,8$ ) vereinzelt festgestellt wurden. Sehr starke Korrelationen ( $r \geq 0,8$ ) traten nicht auf. Korrelationen ab  $r = 0,4$  gelten als potenziell relevant, da sie auf einen messbaren Einfluss einzelner Proteine auf technologische Eigenschaften hinweisen können [144]. Analog zu den untargeted Ergebnissen wurden bei den targeted Daten alle signifikanten Korrelationen ( $p \leq 0,05$ ) analysiert, da sie Gruppen mit statistisch relevanten Zusammenhängen von solchen ohne signifikante Ergebnisse abgrenzen.

ERGEBNISSE UND DISKUSSION



**Abbildung 30:** Korrelationsmatrix der absoluten Glutenproteingehalte und Backqualitätsdaten der BMWpop 2018. \*  $p \leq 0,05$ ; LMW-GS: niedermolekulare Gluteninuntereinheit; HMW-GS: hochmolekulare Gluteninuntereinheit; t, targeted.

Insgesamt zeigten die meisten Glutenproteinfraktionen schwache bis mäßige Korrelationen mit Backqualitätsparametern. Ausnahmen bildeten LMW-GS 1t und 2t, die weitgehend keine Korrelationen zeigten, was möglicherweise auf ihre geringe Konzentration zurückzuführen ist (**Abbildung 29**). Feuchtklebergehalt, Sedimentationswert und Kornproteingehalt wiesen die stärksten Zusammenhänge mit einzelnen Proteinfraktionen auf. Der Feuchtklebergehalt zeigte moderate positive Korrelationen mit  $\gamma$ -Gliadin 1t (0,48),  $\alpha$ -Gliadin 2t (0,59), LMW-GS 3t (0,40) sowie HMW-GS 4t (0,44). Eine negative Korrelation bestand hingegen mit HMW-GS 1t (-0,24). Der Sedimentationswert korrelierte stark mit HMW-GS 3t (0,63) und wies darüber hinaus mäßige positive Zusammenhänge mit  $\alpha$ -Gliadin 2t,  $\gamma$ -Gliadin 1t, LMW-GS 3t sowie den HMW-GS-Gruppen 1t, 2t und 4t auf. Bezüglich des Kornproteingehalts zeigten sich starke positive Korrelationen mit  $\gamma$ -Gliadin 1t (0,60) und HMW-GS 3t (0,61). Zudem konnten mäßige Korrelationen mit  $\alpha$ -Gliadin 2t, LMW-GS 3t und HMW-GS 4t festgestellt werden.

Für den Glutenindex zeigte HMW-GS 1t die höchste Korrelation (0,51),  $\alpha$ -Gliadin 2t hingegen eine signifikante negative (-0,27). Das Brotvolumen korrelierte meist nur schwach (max. 0,37 mit HMW-GS 1t und 3t). Wasseraufnahme und Mehlausbeute zeigten insgesamt nur schwache oder keine signifikanten Korrelationen. Die Fallzahl war nur mit HMW-GS 2t minimal korreliert (0,11). Die Kornhärte zeigte sehr

schwache bis schwache Zusammenhänge (max. 0,35), vor allem mit  $\alpha$ -Gliadin 2t,  $\gamma$ -Gliadin 1t, LMW-GS 3t und HMW-GS 2t–5t. Insgesamt traten nur zwei signifikant negative Zusammenhänge auf (Feuchtklebergehalt mit HMW-GS 1t und Glutenindex mit  $\alpha$ -Gliadin 2t). Im Vergleich zu den untargeted Daten fielen die Korrelationen tendenziell etwas stärker aus. Nachfolgend werden die einzelnen Glutenproteingruppen einzeln betrachtet und, wenn möglich, mit den Ergebnissen der untargeted Proteingruppen verglichen.

Bei den  $\alpha$ -Gliadinen fällt besonders  $\alpha$ -Gliadin 2t auf, da es moderat (0,59) mit Kornproteingehalt und Feuchtklebergehalt korrelierte. Der Zusammenhang mit dem Sedimentationswert war mit 0,42 auch recht hoch. Dagegen gab es bei  $\alpha$ -Gliadin 1t nur ein paar schwach signifikante Korrelationen. Dies zeigt auch bei den targeted Ergebnissen, dass sich innerhalb der  $\alpha$ -Gliadine funktionelle Unterschiede ergeben. Zu  $\alpha$ -Gliadin 1t wurden drei Proteine zugeordnet. Eines davon (UniProtKB ID J7HT14) enthält sieben Cysteinreste. Dies könnte die Polymerisation der Glutenine beeinträchtigen und somit die Backqualität negativ beeinflussen, was die Theorie von Wieser stützen würde [16]. Die  $\alpha$ -Gliadin 2t Gruppe ist mit 56 Proteinen sehr groß. Eine spezifische Zuordnung zu einzelnen Genen kann nicht erfolgen. Es kann lediglich festgestellt werden, dass  $\alpha$ -Gliadine durchaus einen Effekt auf Backqualitätsparameter ausüben können. Bei der untargeted Analyse zeigte zuvor die  $\alpha$ -Gliadin 3 Gruppe den stärksten Einfluss auf die Backeigenschaften (vgl. **Abbildung 23**). Die  $\alpha$ -Gliadin 2t stimmt jedoch nicht mit  $\alpha$ -Gliadin 3 überein, sondern enthält teilweise Proteine der  $\alpha$ -Gliadin 2-Gruppe aus der untargeted Analyse. Diese ergaben allerdings nur schwache Korrelationen mit den Backqualitätsparametern. Daraus lässt sich schließen, dass die übrigen Proteine innerhalb der  $\alpha$ -Gliadin 2t-Gruppe einen stärkeren Einfluss auf die Backqualität haben müssen. Dennoch kann auch durch die Korrelation der targeted Proteomics-Daten bestätigt werden, dass bestimmte  $\alpha$ -Gliadine zur Backqualität beitragen, während andere keinen signifikanten Einfluss zeigten.

Mit  $\gamma$ -Gliadin 1t ergaben sich ebenfalls recht starke Korrelationen mit Kornproteingehalt (0,6), Sedimentationswert (0,53) und Feuchtklebergehalt (0,48). Es handelt sich um eine Gruppe von 34  $\gamma$ -Gliadinen, wobei keins der Gliadinproteine enthalten ist, die bei der untargeted Messung einen höheren Einfluss auf die Backqualität hatten ( $\gamma$ -Gliadin 11, 13, 14 und 21). Stattdessen entsprechen sie den Proteinen von  $\gamma$ -Gliadin 7 und 19 (vgl. **Abbildung 24**), die in der untargeted Analyse kaum Korrelationen zu Backqualitätsparametern aufwiesen. Die hier beobachteten höheren Korrelationen stehen somit im Widerspruch zu den Ergebnissen der untargeted Analyse. Dennoch weisen sie ebenfalls darauf hin, dass bestimmte  $\gamma$ -Gliadine einen Einfluss auf Backqualität haben. Qi et al. untersuchten die  $\gamma$ -Gliadinfamilie und stellten fest, dass  $\gamma$ -Gliadin-Gene genetisch stark variieren. Die Autoren unterteilten die  $\gamma$ -Gliadine zusätzlich in zwei Typen, basierend auf der Länge ihrer repetitiven Domäne. Dabei zeigten sich auch

Unterschiede in ihren ernährungsphysiologischen Eigenschaften:  $\gamma$ -Gliadine mit kürzerer repetitiver Domäne wiesen ein günstigeres Verhältnis essenzieller Aminosäuren auf und gelten daher als ernährungsphysiologisch hochwertiger [186]. Auch diese unterschiedlichen Gliadintypen könnten spezifische Effekte auf die Backqualität haben, die in zukünftigen Studien noch zu untersuchen wären.

Mit LMW-GS 1t und 2t ergaben sich kaum signifikante Korrelationen. LMW-GS 1t beinhaltet nur ein Protein, welches in der untargeted Analyse nicht identifiziert wurde. Dies deutet darauf hin, dass es in so geringen Mengen vorkommt, dass das entsprechende Peptid nur mittels PRM bei der targeted Messung spezifisch gefiltert und detektiert wurde. Die untargeted Messung hat allerdings auch gezeigt, dass Proteingruppen, die in geringer Intensität gemessen wurden, oftmals keine hohen Korrelationen zeigten. LMW-GS 2t entspricht der LMW-GS 12 Gruppe der untargeted Daten (**Abbildung 25**). Weder bei untargeted noch bei targeted Analysen wurden Korrelationen festgestellt; somit entsprechen die Messungen einander. Die LMW-GS 3t Gruppe umfasst 72 Proteine, welche den untargeted Gruppen LMW-GS 2, 25, 32 und 41 entsprechen. Davon korrelierte nur die LMW-GS 2 Gruppe stärker mit Backqualität und könnte folglich auch bei der targeted Messung von LMW-GS 3t hauptverantwortlich für die moderaten Korrelationen mit Glutenproteingehalt (0,55) und Sedimentationswert (0,46) sein. Sowohl LMW-GS 1t als auch LMW-GS 2t gehörten dem m-Typ der LMW-GS an. Die LMW-GS der 3t-Gruppe sowie die LMW-GS 2 Gruppe der untargeted Analyse gehörten sämtlich dem seltenen i-Typ an, der durch eine N-terminale Aminosäuresequenz mit Isoleucin als Startaminosäure charakterisiert ist [187]. Alle Proteine der LMW-GS-3t-Gruppe wiesen durchgehend acht Cysteinreste auf. Diese Ergebnisse stützen die Befunde der untargeted Analyse und bekräftigen, dass LMW-i mit acht Cysteinresten einen positiven Einfluss auf die Backqualität ausüben.

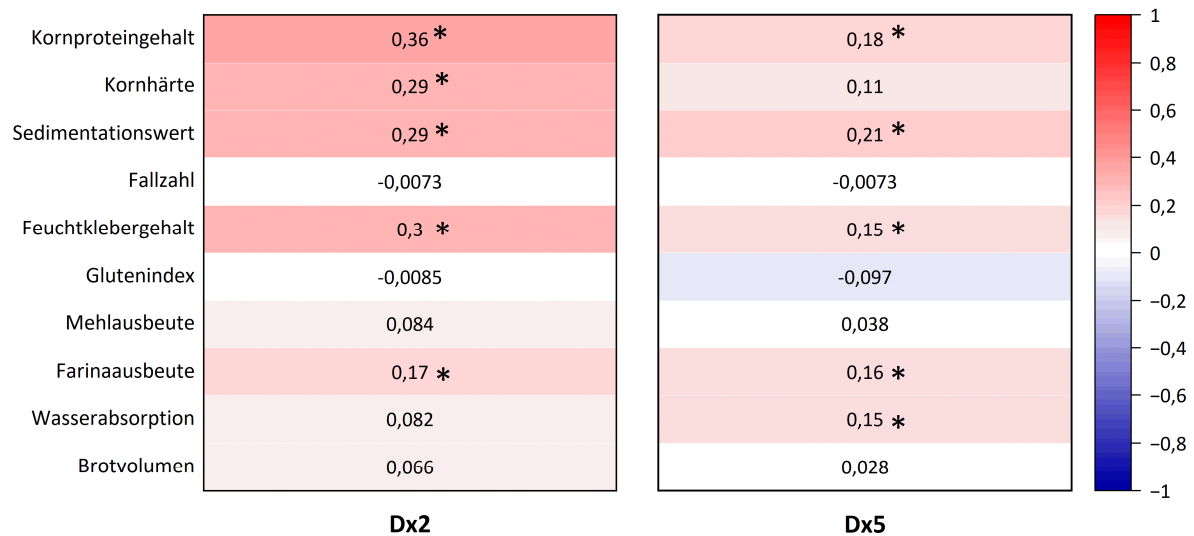
Die HMW-GS 1t Gruppe entspricht Dy10 sowie HMW-GS 9 (**Abbildung 26**). Die beiden Messungen zeigten ein analoges Korrelationsmuster. So betrug die Korrelation für HMW-GS 1t mit Glutenindex 0,51 (HMW-GS 9: 0,53), mit Sedimentationswert 0,39 (HMW-GS 9: 0,35) und mit Brotvolumen 0,37 (HMW-GS 9: 0,34). Beide Gruppen wiesen zudem eine negative Korrelation zum Feuchtklebergehalt auf (HMW-GS 1t: -0,24; HMW-GS 9: -0,30). Diese Übereinstimmungen untermauern den positiven Einfluss von Dy10 auf die Backqualität. Die HMW-GS 3t Gruppe umfasste neben Dy10 noch Dy12 und By8. Sie zeigte starke Korrelationen mit Sedimentationswert (0,63) und Kornproteingehalt (0,61), was auf eine zusätzliche positive Wirkung durch Dy12 und By8 hindeutet. Zudem wirkten sie sich positiv auf den Feuchtklebergehalt aus (0,37), welcher mit HMW-GS 1t negativ korrelierte (-0,24). Die Korrelation zum Brotvolumen blieb mit 0,37 dagegen unverändert. Deshalb liegt die Vermutung nahe, dass die Gruppen Dy12 und By8 keinen zusätzlich positiven Effekt auf diesen zentralen Backparameter hatten. Die HMW-GS 4t-Gruppe, die alle  $\gamma$ -

Typen umfasst, zeigte teils gegensätzliche Effekte zu HMW-GS 1t und 3t. Die Korrelation mit dem Brotvolumen war gering (0,19), während ein mittlerer Zusammenhang mit dem Sedimentationswert (0,50) und ein hoher mit dem Kornproteingehalt (0,58) bestand. Dies deutet auf eine mögliche funktionelle Diversität innerhalb der  $\gamma$ -Typen mit unterschiedlich starkem Einfluss auf die technologischen Eigenschaften hin.

HMW-GS 2t entspricht Ax2\* und bei den untargeted Messungen den Gruppen HMW-GS 1, 7 und 12. Es ergaben sich insgesamt stärkere Zusammenhänge, vor allem bei Kornproteingehalt (0,49, HMW-GS 2: 0,41) und Sedimentationswert (0,48, HMW-GS 2: 0,37). Die Korrelation mit dem Parameter Brotvolumen blieb allerdings nahezu unverändert bei 0,19 vs. 0,2 bei HMW-GS 2. Die generellen Tendenzen stimmen zwischen beiden Messansätzen überein, jedoch deuten die targeted Daten auf einen etwas stärkeren Einfluss von HMW-GS Ax2\* auf bestimmte Backqualitätsparameter hin.

HMW-GS 5t entspricht den Untereinheiten Dx2 und Dx5 bzw. HMW-GS 3 bei den untargeted Messungen. Es zeigten sich leicht stärkere Korrelationen für Kornproteingehalt (0,25, HMW-GS 3: 0,22), Sedimentationswert (0,26, HMW-GS 3 0,22) und Kornhärte (0,16, HMW-GS 3: 0,18). Allerdings ergab sich bei dieser Auswertung keine signifikante Korrelation zum Brotvolumen während mit HMW-GS 3 ein leichter Zusammenhang von 0,16 festgestellt wurde. Dafür ergab sich bei HMW-GS 5t ein leichter Zusammenhang mit dem Feuchtklebergehalt (0,15), der bei HMW-GS 3 nicht signifikant war. Insgesamt weisen sowohl die targeted als auch die untargeted Analysen lediglich auf einen geringfügig positiven Einfluss dieser Gruppe auf die Backqualität hin. Damit zeigen die  $\gamma$ -Typen in dieser Analyse ebenfalls einen stärkeren Einfluss auf die Backqualität als die  $\alpha$ -Typen und stützen somit die Ergebnisse der untargeted Analytik.

Da Dx2 und Dx5 weder in der untargeted noch in der targeted Analyse eindeutig differenziert werden konnten, wurde auf das genetische Hintergrundwissen zu den BMWpop-Linien zurückgegriffen. Anhand der bekannten Allelzusammensetzung der einzelnen Weizenlinien wurden die Datensätze entsprechend gruppiert und jeweils mit den Backqualitätsdaten der BMWpop 2018 in Beziehung gesetzt (**Abbildung 31**). Die Korrelationen wiesen auf einen schwachen positiven Zusammenhang zwischen Dx2 und dem Kornproteingehalt (0,36), der Kornhärte (0,29), dem Sedimentationswert (0,29) sowie dem Feuchtklebergehalt (0,30) hin. Für Dx5 zeigte sich ein schwacher Zusammenhang mit dem Sedimentationswert (0,21), während die Korrelationen mit dem Kornproteingehalt (0,18), dem Feuchtklebergehalt, der Farinausbeute (0,16) und der Wasserabsorptionsfähigkeit (0,15) lediglich sehr schwach ausfielen.



**Abbildung 31:** Korrelationsmatrix der HMW-GS Dx2 bzw. Dx5 und Backqualitätsdaten der BMWpop 2018. \* $p \leq 0,05$ ; HMW-GS, hochmolekulare Gluteninuntereinheit.

Insgesamt waren die Korrelationen zwischen Dx2 und den Backqualitätsparametern etwas ausgeprägter als bei Dx5. Dies ist im Widerspruch zur Literatur, die der Dx5-Untereinheit mit ihrem zusätzlichen Cysteinrest einen stärkeren Einfluss auf Backqualität zuschreibt [29, 172]. Dabei ist zu hinterfragen, ob der zusätzliche Cysteinrest möglicherweise auch die Polymerisation der Glutene beeinträchtigt [16]. Allerdings zeigten sowohl Dx5/Dx2 bzw. HMW-GS 5t keinen Zusammenhang mit dem wichtigsten Backqualitätsparameter, dem Brotvolumen. Dies deutet erneut darauf hin, dass diese Untereinheiten (und damit die x-Typen insgesamt) lediglich einen geringen Einfluss auf die Backqualität haben.

#### 3.4.2.3 Zusammenfassende Diskussion

Auffällig waren insbesondere die moderaten Korrelationen zwischen der targeted Proteomics-Analyse und Kornproteingehalt, Feuchtklebergehalt sowie Sedimentationswert, während die Zusammenhänge mit Kornhärte und Brotvolumen schwächer ausfielen. Diese Ergebnisse unterstreichen den praktischen Nutzen dieser technologischen Parameter als verlässliche Indikatoren für das Vorhandensein von Glutenproteinen und die Ausbildung eines funktionellen Glutennetzwerks. Auch in den targeted Daten zeigten sich überwiegend nur moderate bis schwache Korrelationen mit den Backqualitätsparametern, sodass – analog zu den Ergebnissen der untargeted Analyse – keine belastbaren Aussagen über einen dominierenden Einfluss einzelner Glutenproteingruppen möglich sind. Die höchsten Korrelationen, teilweise über 0,6, traten mit dem Kornproteingehalt und Sedimentationswert auf. Der Parameter Kornproteingehalt ist jedoch nicht unbedingt ein verlässlicher Indikator für die Backqualität, da bereits gezeigt wurde, dass ein hoher Proteingehalt nicht zwangsläufig mit guten Backeigenschaften korreliert

[47]. Zudem ergaben sich für das Brotvolumen als zentraler Backqualitätsparameter höchstens schwache Korrelationen, selbst mit dem als besonders einflussreich geltenden HMW-GS Dy10. Dies unterstreicht, dass das Brotvolumen nicht durch einzelne Glutenproteingruppen bestimmt wird, sondern vielmehr von der Gesamtzusammensetzung der Glutenproteine abhängt. So wurde bereits nachgewiesen, dass der Effekt auf die Backqualität am stärksten war, wenn HMW-GS der  $\alpha$ - und  $\gamma$ -Typen in gleichen Anteilen dem Teig zugesetzt wurden; im Gegensatz zur einseitigen Anreicherung mit nur einem Typ [61].

Die bei der targeted Analyse erhaltenen Ergebnisse bestätigten die Tendenz der untargeted Daten größtenteils. Die Targeted-Proteomics-Daten zeigten insgesamt stärkere Zusammenhänge mit der Backqualität, insbesondere auch bei den  $\alpha$ - und  $\gamma$ -Gliadinen, den LMW-i sowie bestimmten HMW-GS-Gruppen, was deren Einfluss auf die Backeigenschaften zusätzlich bestätigt. Im Gegensatz zur untargeted Analyse ermöglicht die targeted Proteomics-Strategie eine hochspezifische Detektion ausgewählter Peptide. Nach der eindeutigen Identifizierung erfolgt eine automatisierte und konsistente Integration der Peptidsignale. Der Einsatz eines stabilisotopenmarkierten internen Standards erhöht dabei die Validität der Messergebnisse, da sowohl Probenverluste während der Probenaufarbeitung als auch matrixbedingte Effekte, z. B. bei der Ionisation, kompensiert werden. Dieses Vorgehen gilt als Goldstandard in der absoluten Proteinquantifizierung [70]. Die methodischen Vorteile führen zu einer höheren quantitativen Genauigkeit und geringeren Messvariabilität im Vergleich zu LFQ-basierten Verfahren, bei denen größere Abweichungen auftreten können, sodass zunächst Normalisierungsschritte erforderlich sind, um vergleichbare Ergebnisse zu erhalten. Die Ergebnisse der targeted Analytik sollten bestehende Korrelationen zwischen Proteingehalt und funktionellen Parametern deshalb zuverlässiger darstellen können.

Bei dem targeted Ansatz wäre es jedoch notwendig, für individuelle Glutenproteine eindeutige Peptidmarker zu finden, die eine gezielte Quantifizierung derer ermöglichen. So konnte bei der  $\gamma$ -Gliadin-1t Gruppe ein Einfluss auf die Backqualität festgestellt werden. Es bleibt jedoch unklar, welche spezifischen Gliadin-Proteine innerhalb dieser Gruppe dafür verantwortlich sind. Da die Entwicklung einer targeted Proteomics-Methode mit erheblichem Zeit- und Kostenaufwand verbunden ist, ist zudem eine priorisierte Auswahl potenziell relevanter Proteine im Vorfeld empfehlenswert. Bei den  $\alpha$ - und  $\gamma$ -Gliadinen sowie bei den LMW-GS könnten dies z. B. jene Proteingruppen sein, die einen höheren Einfluss auf die Backqualität hatten. Im Fall der HMW-GS wäre insbesondere Dy12 von Interesse, um einen gezielteren Vergleich mit Dy10 zu ermöglichen. Die Unterscheidung zwischen Dx2 und Dx5 ist erstrebenswert, um den nachgewiesenen starken Einfluss von Dx5 auf die Backqualität klarer belegen zu können. Die Identifikation eindeutig zuordenbarer Peptide mit geeigneten MS-Detektionseigenschaften stellte jedoch eine

Herausforderung dar (3.3.1.1). Eine Unterscheidung zwischen Dx2 und Dx5 war letztlich nur auf Basis der bekannten genetischen Daten der BMWpop möglich. Die Identifikation eindeutig zuordenbarer Peptide würde hingegen eine Unabhängigkeit von genetischen Informationen ermöglichen.

Beide Proteomics-Ansätze erlaubten eine Unterscheidung von Proteingruppen innerhalb der  $\alpha$ - und  $\gamma$ -Gliadine sowie der LMW-GS, wobei einige einen deutlichen Zusammenhang mit der Backqualität zeigten, während andere keinen signifikanten Einfluss aufwiesen. Diese Ergebnisse konnten nicht auf einzelne Allele zurückgeführt werden. In der Literatur gibt es oftmals nur vereinzelte Hinweise auf spezifische Allele, die mit guten Backeigenschaften in Verbindung gebracht werden [153-155]. Wie bereits bei den untargeted Ergebnissen diskutiert, müsste eine umfassende Aufklärung über die vollständige Identifikation und genaue Zuordnung sämtlicher Gliadine und LMW-GS zu ihren jeweiligen Allelen erfolgen, um die hier gewonnenen Proteomics-Ergebnisse mit anderen genetischen oder funktionellen Ergebnissen in Verbindung zu setzen. Darüber hinaus müssten diese Informationen verfügbar sein. Zwar wurde das Weizengenom in den letzten Jahren vollständig sequenziert [25], doch besteht nach wie vor eine Lücke in der Verknüpfung zu den entsprechenden Proteinen. Dies liegt vor allem an der Komplexität des hexaploiden, multigenomischen Aufbaus von Weizen, der eine eindeutige Zuordnung von Genen zu Proteinen erheblich erschwert. Dies hat unter anderem zur Folge, dass Glutenproteine häufig nicht nur von einzelnen Genen codiert werden. Die detaillierte Charakterisierung von Gliadin-Allelvarianten gestaltet sich z. B. herausfordernd, da diese Proteinfraktionen von hochkomplexen und zum Teil redundanten Genfamilien codiert werden. So wurden 47  $\alpha$ -Gliadine identifiziert, aus denen lediglich 26 Proteine resultierten [188]. Wang et al. identifizierten 52 Gliadine, die für 25  $\alpha$ -,  $\gamma$ - und  $\omega$ -Gliadine kodierten [189]. Im Gegensatz dazu wurden die LMW-GS-Allele bereits ausführlicher in Bezug auf ihre Relevanz für die Backqualität untersucht und klassifiziert [62]. Bereits verfügbare genetische Informationen sollten systematisch in öffentlich zugänglichen Proteindatenbanken hinterlegt werden, um künftig eine zuverlässige Zuordnung sowohl der in dieser Studie generierten als auch anderer Proteomics-Daten zu gewährleisten.

$\gamma$ -Gliadin 1t war die einzige Proteingruppe, bei der es gegenteilige Korrelationen zwischen den Messmethoden gab. Dies könnte an der LFQ-Quantifizierung liegen, welche die jeweiligen Proteingruppen in der Regel durch mehrere Peptide (soweit vorhanden) quantifiziert. Dadurch werden die Intensitätsverhältnisse eventuell anders zugeordnet. Dadurch, dass die Glutenpeptide eine hohe Ähnlichkeit aufweisen, ist eine eindeutige Zuordnung ebenfalls erschwert und kann je nach Messmethode variieren, wie bereits in Kapitel 3.2.3 und 3.3.2 mehrfach dargestellt. Dies verdeutlicht einmal mehr die bereits mehrfach diskutierten Herausforderungen bei der Analyse und eindeutigen Zuordnung von

Glutenproteinen. Generell besteht ein erheblicher Bedarf an kuratierten, nicht-redundanten Datenbanken, die eine präzise Proteinannotation ermöglichen. Die durch gemeinsam vorkommende Peptide verursachten Unsicherheiten bei der Proteinidentifikation und -quantifizierung bleiben eine zentrale Herausforderung der Shotgun-Proteomics. Zukünftige Entwicklungen bioinformatischer Software sollten gezielt darauf ausgerichtet werden, diese Problematik zu adressieren. Die Weiterentwicklung von Middle-Down-Proteomics, bei der größere Peptidfragmente mit höherer Sequenzspezifität analysiert werden, bietet ebenfalls einen vielversprechenden Ansatz, um dieser auch in anderen Proteinen auftretenden Herausforderung gezielter zu begegnen [190].

Die derzeit noch bestehenden Unterschiede in der Proteinerkennung und -quantifizierung zwischen beiden Methoden verdeutlichen die nach wie vor bestehende Notwendigkeit komplementärer analytischer Ansätze, um die funktionelle Relevanz einzelner Proteinfractionen im Hinblick auf die Backqualität differenziert beurteilen zu können. Eine vielversprechende Erweiterung wäre die Analyse der BMWpop mittels DIA, da sie ein noch detaillierteres Proteomprofil liefern könnte; auch wenn die Auswertung von DIA-Daten deutlich umfangreicher und komplexer ist. Da in dieser Arbeit ebenfalls eine targeted Analyse durchgeführt wurde, diente DDA zunächst der gezielten Auswahl geeigneter Peptide. Zudem hat sich für eine breit angelegte und gleichzeitig targeted Proteinquantifizierung die Kombination aus targeted- und DDA-basierter LC-MS/MS-Analyse bewährt [191].

## 4 Fazit und Ausblick

In dieser Arbeit wurde der Zusammenhang zwischen Glutenproteingruppen und technologischen Backqualitätsparametern in der BMWpop untersucht. Zum Einsatz kamen eine optimierte untargeted Proteomics-Methode (LFQ-basiert) sowie eine neu entwickelte targeted Strategie (SIDA-basiert), die jeweils konsistente Ergebnisse lieferten. Auf dieser Grundlage konnten spezifische Beziehungen zwischen Einzelproteinen und Backqualitätsmerkmalen detailliert erfasst werden. Starke Korrelationen wurden nicht festgestellt, was auf den fehlenden dominanten Einfluss einzelner Proteingruppen hinweist und die komplexe Natur des Merkmals Backqualität unterstreicht. Signifikante, meist sehr schwache bis moderate Zusammenhänge zeigten sich mit Kornproteingehalt, Sedimentationswert, Feuchtklebergehalt, Glutenindex und spezifischem Brotvolumen. Dagegen korrelierten Parameter wie Kornhärte, Fallzahl, Mehl- und Farinaausbeute sowie Wasseraufnahme kaum oder gar nicht mit den Proteinmengen – ein Hinweis auf ihre weitgehende Unabhängigkeit von der Glutenproteinzusammensetzung.

Die HMW-GS zeigten den stärksten Einfluss auf die Backqualität, gefolgt von LMW-GS und  $\gamma$ -Gliadinen mit ähnlicher Wirkung, während  $\alpha$ -Gliadine nur geringe Effekte aufwiesen. Diese Ergebnisse stimmen mit der Literatur überein, die den HMW-GS eine Schlüsselrolle zuschreibt [28]. Innerhalb der  $\alpha$ - und  $\gamma$ -Gliadine sowie LMW-GS variierten die Beiträge einzelner Proteingruppen deutlich, was auf die besondere Bedeutung spezifischer Einzelproteine hinweist. Der Einfluss struktureller Merkmale auf die beobachteten Unterschiede erfordert weitere Untersuchungen, da sich diese nicht allein durch die Anzahl oder Position der Cysteinreste erklären lassen.

Aufgrund unzureichender Datenbankeinträge konnten viele Gliadine und LMW-GS nicht eindeutig Genorten zugewiesen werden. Für die HMW-GS hingegen ließ sich eine genetische Rangfolge bezüglich ihres Einflusses auf die Backqualität ableiten:  $Dy10 > Dy12 > By8 \geq By9 > Dx2/Dx5 > Ax1/2^* > Bx7 > Ax1/2^* > Bx6 > Ay$ . Die Ergebnisse standen im Einklang mit der Literatur und unterstreichen den vielfach beschriebenen starken Einfluss von Dy10 auf die Backqualität [62]. In dieser Arbeit erwiesen sich die  $\gamma$ -Typ-Untereinheiten als einflussreicher als die  $\alpha$ -Typen – entgegen Teilen der Literatur, die insbesondere Dx5 hervorhebt [61], jedoch im Einklang mit Studien, die  $\gamma$ -Typen wegen zusätzlicher Cysteinreste als besonders relevant betrachten [174]. Unterschiedliche Auswirkungen auf die Backqualität innerhalb der  $\alpha$ - und  $\gamma$ -Typen deuten auf feinstrukturelle Ursachen hin, deren Einfluss auf die Glutennetzwerkbildung weiter untersucht werden sollte. Insgesamt bestätigen die Ergebnisse die zentrale Bedeutung der individuellen Glutenproteinzusammensetzung für die Backqualität.

Die vorliegende Arbeit schafft eine Grundlage für ein vertieftes Verständnis des Einflusses einzelner Glutenproteine auf die Backqualität von Weizen. Die gewonnenen Erkenntnisse eröffnen die Möglichkeit, spezifische proteomische Marker zu identifizieren, die künftig in züchterischen Selektionsprozessen eingesetzt werden könnten, um Qualitätsmerkmale gezielter und effizienter vorherzusagen. Ein nächster Schritt wäre die Durchführung von QTL-Analysen unter Einbeziehung der BMWpop-Genetik, um Genorte zu identifizieren, die mit Unterschieden in der Proteinexpression assoziiert sind – etwa solche, die die Bildung von Proteinen wie Dy10 fördern. Zur Berücksichtigung von Umwelteinflüssen wäre zudem eine weitere proteomische Messung der BMWpop in einem zusätzlichen Erntejahr sinnvoll. Dies würde ermöglichen, stabil exprimierte Marker über verschiedene Umweltbedingungen hinweg zu validieren. Darauf aufbauend könnten gezielte Eingriffe mittels *Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated proteins* (CRISPR/Cas) genutzt werden, um relevante Loci präzise zu modifizieren und so die Backqualität gezielt über Genomeditierung zu verbessern.

Eine zentrale Limitation bestand in der oftmals fehlenden eindeutigen Zuordnung einzelner Proteine zu spezifischen Genorten, was ihre klare Differenzierung und genetische Einordnung erschwerte. Um eine fundierte genetische Interpretation der proteomischen Daten zu ermöglichen, besteht ein dringender Bedarf an umfassend kuratierten Referenzdatenbanken mit konsistenter genetischer Annotation und standardisierter Nomenklatur. Solche Ressourcen würden nicht nur die Vergleichbarkeit der Ergebnisse mit anderen Studien erheblich verbessern, sondern auch deren Einordnung in den bestehenden Forschungskontext erleichtern.

Bottom-Up Proteomics bei Weizen ist nach wie vor herausfordernd, da die hohe Sequenzhomologie der Glutenproteine zu zahlreichen redundanten Peptiden führt und die Proteinzuzuordnung erschwert. Insbesondere der Einsatz verschiedener Peptidasen führte zu deutlichen Unterschieden in der Proteinidentifikation. Für eine präzisere Zuordnung und höhere Sequenzabdeckung sollten alternative Hydrolysestrategien systematisch geprüft werden. In Kombination mit kuratierten, redundanzfreien Referenzdatenbanken und spezialisierter Auswertesoftware, die auf die Komplexität des Weizenproteoms zugeschnitten ist, könnten die Differenzierung und die Identifikation einzelner Glutenproteine deutlich verbessert werden. Langfristig eröffnet die Weiterentwicklung hochauflösender Massenspektrometer die Möglichkeit, intakte Glutenproteine im Rahmen von Top-down Proteomics oder größere Peptidfragmente mittels Middle-down Proteomics direkt zu analysieren, um damit eine eindeutige Identifizierung zu ermöglichen. KI-gestützte Methoden wie Machine Learning und neuronale Netze ermöglichen die Integration und Analyse komplexer omischer sowie weiterer heterogener Daten – mit dem Ziel, die Produktqualität präzise vorherzusagen [192].

## 5 Experimenteller Teil

### 5.1 Probensortiment

Für die untargeted und targeted Proteomics-Methodenentwicklung und den Peptidasenvergleich wurden Vollkornmehle von neun Weizenlinien, die 2018 in Söllingen, Deutschland, geerntet wurden, verwendet: Ambition, FIRL3565, Bussard, Event, Format, Julius, BAYP4535, Potenzial und RGT Reform. Sie wurden von der Bayerischen Landesanstalt für Landwirtschaft (LfL), Deutschland, bezogen. Mit Ausnahme von RGT Reform handelt es sich bei diesen Sorten um die Elternlinien der bayerischen Multiparent Advanced Generation Intercross Weizenpopulation (BMWpop), die an der LfL entwickelt wurde [38]. RGT Reform ist derzeit eine der am häufigsten angebauten Sorten in Deutschland und diente als Vergleichsstandard.

Für die untargeted und targeted Proteomics-Gesamtanalyse wurden alle Vollkornmehle (8 Elternlinien, 394 Tochterlinien, RGT Reform als Vergleichsstandard) der BMWpop 2018 verwendet.

## 5.2 Chemikalien und Enzyme

Chemikalie	Hersteller	Reinheit
1-Propanol	VWR (Radnor, PA, USA)	100 %
1,4-Dithiothreitol (DTT)	Appllichem (Darmstadt, Deutschland)	Min. 99,5 %
2-Chloracetamid (CAA)	Acros Organics (Geel, Belgien)	98 %
Acetonitril (ACN)	VWR Fisher Scientific (Waltham, MA, USA)	99,9 % 100 %
Ameisensäure (FA)	VWR	≥99,9 %
Ammoniumbicarbonat (NH <sub>4</sub> HCO <sub>3</sub> )	Carl Roth (Karlsruhe, Deutschland)	99 %
Calciumchlorid (CaCl <sub>2</sub> )	Carl Roth	≥98 %
α-Chymotrypsin aus Rinderpankreas	Sigma-Aldrich (St. Louis, MO, USA)	-
Dinatriumhydrogenphosphat-Dihydrat (Na <sub>2</sub> HPO <sub>4</sub> x 2 H <sub>2</sub> O)	VWR	99,9 %
Ethanol (EtOH) abs.	VWR	100 %
Harnstoff	Sigma-Aldrich	>99,5 %
Kaliumhydrogenphosphat (KH <sub>2</sub> PO <sub>4</sub> )	VWR	99,9 %
Natriumchlorid (NaCl)	VWR	99,9 %
Pepsin aus porziner Magenschleimhaut	Sigma-Aldrich	-
Prolamin Working Group-Gliadin (PWG)	Arbeitsgemeinschaft Getreideforschung e.V. (Detmold, Deutschland)	93,1 %
Salzsäure (HCl)	VWR	p.A.
Thermolysin von <i>Geobacillus</i> <i>Stearothermophilus</i>	Sigma-Aldrich	-
Trifluoressigsäure (TFA)	Alfa Aesar (Ward Hill, MA, USA)	99,5 %
Tris(2-carboxyethyl)phosphin (TCEP) Hydrochlorid	Alfa Aesar	98 %
Tris(hydroxymethyl)aminomethan (TRIS)	VWR	100 %
TRIS-HCl	Carl Roth	≥99 %, p.A.
Trypsin aus Rinderpankreas	Sigma-Aldrich	-

### 5.3 Quantifizierung der Glutenfraktionen mittels RP-UHPLC

Die Extraktion der Albumine und Globuline, Gliadine und Glutenine wurde nach der modifizierten Osborne-Fraktionierung durchgeführt [19]. Dazu wurde Vollkornmehl (100 mg) der acht Elternlinien der BMWpop plus RGT Reform als Vergleichsstandard in Dreifachbestimmung eingewogen.

Die Extraktion der Albumine und Globuline, Gliadine und Glutenine erfolgte schrittweise aus 100 mg Mehl. Dabei wurden folgende Lösungen verwendet:

- Extraktionslösung E1: 0,4 mol/L NaCl + 0,067 mol/L Na<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub> (pH 7,6)
- Extraktionslösung E2: 60 % (v/v) Ethanol
- Extraktionslösung E3: 50 % (v/v) 1-Propanol plus 2 mol/L Harnstoff plus 0,05 mol/L Tris(hydroxymethyl)aminomethan Hydrochlorid (TRIS-HCl) (pH 7,5) plus 1 % (w/v) DTT

Zunächst wurden die Albumine und Globuline jeweils zweimal mit je 1 mL der Extraktionslösung E1 extrahiert. Anschließend erfolgte die dreifache Extraktion der Prolamine mit jeweils 0,5 mL der Lösung E2. Zum Abschluss wurden die Gluteline zweimal mit je 1 mL des Puffers E3 extrahiert. Jeder Extraktionsschritt begann mit 2 min Vortexen, gefolgt von einer Durchmischung bei 22 °C für 10 min (bei Albuminen, Globulinen und Gliadinen) bzw. bei 60 °C für 30 min (bei Gluteninen). Im Anschluss wurden die Proben für 30 min bei 25 °C und 3550 rcf (relative centrifugal force, relative Zentrifugalbeschleunigung) zentrifugiert. Die gewonnenen Überstände wurden abgenommen, mit der jeweiligen Extraktionslösung auf 2 mL aufgefüllt und anschließend durch einen 0,45 µm-Spritzenfilter mit regenerierter Zellulosemembran (WICOM, Heppenheim, Deutschland) filtriert. Die gefilterten Extrakte wurden für die anschließende Analyse mittels RP-HPLC verwendet.

Die Messung der Extrakte erfolgte mittels RP-HPLC. Die Bedingungen sind in **Tabelle 6** dargestellt. Die externe Kalibrierung erfolgte mit PWG-Gliadin (Arbeitsgemeinschaft Getreideforschung e.V., Detmold, Deutschland) gemäß van Eckert et al. [193]. Das Referenzmaterial wurde in Lösung E2 auf eine Konzentration von 2,5 mg/mL gelöst, dabei im Ultraschallbad behandelt und anschließend filtriert. Die Bestimmung der Gehalte an Albuminen, Globulinen, Gliadinen und Gluteninen erfolgte anhand der jeweiligen Gesamtpeakfläche im RP-HPLC-Chromatogramm. Die Unterfraktionen der Gliadine ( $\omega$ 5-,  $\omega$ 1,2-,  $\alpha$ - und  $\gamma$ -Gliadine) wurden entsprechend ihrem Anteil an der Gesamtgliadinfläche quantifiziert. Die HMW- und LMW-Glutenine sowie die  $\omega$ b-Gliadine wurden analog auf Basis ihres prozentualen Anteils an der Gesamtpeakfläche der Glutenine berechnet. Alle Analysen erfolgten unter Berücksichtigung einer Basislinienkorrektur mit dem jeweils verwendeten Lösungsmittel.

EXPERIMENTELLER TEIL

**Tabelle 6:** Bedingungen für die Messung der Albumine/Globuline, Gliadine, Glutenine für die Messung mittels Umkehrphasen-Hochleistungsflüssigkeitschromatographie.

Vorsäule	Phenomenex (Torrance, CA, USA) Security Guard, Widedpore C18		
Säule	YMC (Dinslaken, Deutschland) Triart Bio C18, 150 mm × 2,1 mm, 3 μm		
Injektionsvolumen Albumine/Globuline	20–30 μL		
Injektionsvolumen Gliadine	5–30 μL		
Injektionsvolumen Glutenine	15–60 μL		
Injektionsvolumen PWG-Kalibrierung	5, 10, 15 und 20 μL		
Säulentemperatur	60 °C		
Wellenlänge	210 nm		
Fluss	0,5 mL/min		
Eluenten	Eluent A: Wasser/TFA (99,9/0,1; v/v) Eluent B: ACN/TFA (99,9/0,1; v/v)		
Gradient Albumine/Globuline	<b>Zeit [min]</b>	<b>A [%]</b>	<b>B [%]</b>
	0	100	0
	0,4	100	0
	0,5	80	20
	8,0	40	60
	8,1	0	100
	13,0	0	100
	27,0	100	0
Gradient Gliadine, Glutenine und PWG-Standard	<b>Zeit [min]</b>	<b>A [%]</b>	<b>B [%]</b>
	0	95	5
	0,4	95	5
	0,5	70	30
	16,0	40	60
	16,1	0	100
	22,1	0	100
	36,0	95	5

ACN: Acetonitril; PWG: Prolamin Working Group; TFA: Trifluoressigsäure.

## 5.4 Methodenoptimierung untargeted Proteomics

### 5.4.1 Untargeted Proteomics: Finale LC-MS/MS-Probenaufarbeitung

#### 5.4.1.1 Proteinextraktion

Die Einfachextraktion erfolgte mit einer einzigen Extraktionslösung (E3 für nicht lösliche Proteine) bestehend aus 50 % (v/v) 1-Propanol in 2 mol/L Harnstoff und 0,05 mol/L TRIS-HCl (pH 7,5) plus 1 % (w/v) DTT. Vollkornmehl (50 mg) wurde in Dreifachbestimmung in ein 2 mL Mikroröhrchen eingewogen. E3 (1 mL) wurde zugegeben und es wurde 2 min mit dem Vortexer gemischt. Die Röhrchen wurden 5 min bei 22 °C mit Ultraschall (US) behandelt, anschließend in einem Thermomixer inkubiert (1500 U/min, 10 min, 22 °C) und 15 min lang bei 21.380 rcf und 22 °C zentrifugiert. 800 µL des Überstands wurden in einem 2 mL Mikroröhrchen überführt und bis zur Trockene eingedampft (8 mbar, 4–6 h, 40 °C).

#### 5.4.1.2 Proteinreduzierung und –alkylierung

Die extrahierten Proteine (jeweils n = 3) wurden in 300 µL 0,5 mol/L TRIS-HCl (pH 8,5) und 300 µL 1-Propanol gelöst. Für die Reduktion wurden 100 µL 0,05 mol/L TCEP in 0,5 mol/L TRIS-HCl (pH 8,5) zugegeben. Die Röhrchen wurden in einem Thermomixer (1000 U/min, 30 min und 60 °C) inkubiert. Für die Alkylierung wurden 100 µL 0,05 mol/L CAA in 0,5 mol/L TRIS-HCl (pH 8,5) zugegeben. Die Röhrchen wurden im Thermomixer im Dunkeln inkubiert (1000 U/min, 45 min, 37 °C) und das Lösungsmittel bis zur Trockene eingedampft (8 mbar, 4–6 h, 40 °C).

#### 5.4.1.3 Proteinverdau

Die reduzierten und alkylierten Proteinextrakte wurden in 800 µL 0,1 mol/L TRIS-HCl (pH 7,8) und 0,04 mol/L Harnstoff gelöst. Trypsin wurde in einer Konzentration von 1 mg/mL gelöst und den Extrakten in einem Enzym-Substrat-Verhältnis von 1:25 zugesetzt. Die Röhrchen wurden in einem Thermomixer im Dunkeln (200 U/min, 37 °C, 18 h) inkubiert. Die Aktivität des Enzyms wurde gestoppt, indem die Röhrchen 5 min lang auf 95 °C erhitzt wurden. Die Extrakte wurden anschließend direkt mit SPE gereinigt.

#### 5.4.1.4 Proteinaufreinigung

Die SPE wurde unter Verwendung von Discovery DSC-18 SPE 96-Well-Mikroplatten mit einem Bettgewicht von 100 mg pro Well (Supelco, Sigma-Aldrich) gemäß dem Protokoll des Herstellers durchgeführt. Die Wells wurden mit 2 mL Methanol aktiviert und mit 2 mL 80 % (v/v) ACN und 0,1 % (v/v) FA in Wasser äquilibriert. Anschließend wurde ein Konditionierungsschritt mit 3 mL von 2 % (v/v) ACN und 0,1 % (v/v)

FA in Wasser durchgeführt. Die verdauten Proben wurden in die Wells gegeben und ohne Vakuum durchtropfen gelassen. Anschließend wurden sie mit 5 mL 2 % (v/v) ACN und 0,1 % (v/v) FA in Wasser gewaschen. Die Peptide wurden ohne Vakuum mit 2 × 0,5 mL von 40 % (v/v) ACN und 0,1 % (v/v) FA in Wasser eluiert. Die Eluate wurden in 2 mL-Mikroröhrchen überführt und das Lösungsmittel bis zur Trockene eingedampft (8 mbar, 4–6 h, 40 °C).

#### 5.4.1.5 LC-MS/MS Messung

Die gereinigten Proben wurden in 1 mL 2% ACN und 0,1% FA rekonstituiert. Die LC-MS/MS-Messungen wurden mit einem Vanquish UHPLC-System (Thermo Fisher Scientific, Waltham, MA, USA) durchgeführt, das mit einem Orbitrap Q Exactive plus MS/MS-System (Thermo Fisher Scientific) gekoppelt war.

Die Peptide (Injektionsvolumen von 10 µL) wurden auf einer Aeris PEPTIDE XB-C18 (1,7 µm, 10 nm, 150 mm x 2,1 mm) LC-Säule (Phenomenex, Torrance, CA, USA) bei einer Flussrate von 0,2 mL/min getrennt. Die verwendeten Lösungsmittel waren (A) 0,1 % FA in Wasser und (B) 0,1 % FA in ACN. Der lineare Gradient wurde wie folgt eingestellt: 2 % B bei 0 min, 10 % B bei 1,5 min, 30 % B bei 23 min, 40 % B bei 28 min, 60 % B bei 31 min, 80 % B für 34–37 min und 2 % B für 38–45 min. Der Eluent der UHPLC wurde direkt an die ESI-Quelle der Q Exactive plus MS/MS-Systeme gekoppelt. Die Sprühspannung wurde auf 3,0 kV, das Mantelgas auf 35 und das Hilfsgas auf 10 eingestellt. Es wurde kein Sweep-Gas verwendet. Die Kapillartemperatur wurde auf 350 °C und die Radiofrequenz der S-Linse auf 60 eingestellt.

Die Daten wurden via DDA im positiven ESI-Modus aufgenommen. Die vollständigen MS-Parameter wurden wie folgt eingestellt: Resolution: 70.000, AGC Target: 3e6, Maximum Injection Time (IT): 50 ms, Scan Range: 360-1300 *m/z*. Die Parameter für dd-MS2-Scans wurden wie folgt festgelegt: Resolution: 17.500; AGC Target: 1e5; Maximum IT: 100 ms; TopN: 10; Isolation Window: 2,4 *m/z*; Fixed First Mass: 140,0 *m/z*; Normalized Collision Energy: 28; Dynamic Ion Exclusion: 20 s. Die Laufzeit der Methode betrug 45 min.

#### 5.4.1.6 Proteinidentifizierung

Zur Proteinidentifizierung wurde die Software MaxQuant (Version 2.6.7.0) mit der integrierten Suchmaschine Andromeda verwendet [6]. Die MS/MS-Rohdaten wurden mit einer Weizen-Proteindatenbank aus UniProtKB abgeglichen (FASTA-Datei für Organismus-ID [4565], heruntergeladen am 11. Februar 2025, enthielt 112.753 Proteineinträge). Die Standardeinstellungen von MaxQuant wurden beibehalten, mit Ausnahme der folgenden Änderungen: Die Oxidation von Methionin sowie die N-terminale Acetylierung von Proteinen wurden als variable Modifikationen festgelegt, die

Carbamidomethylierung von Cysteinen als fixe Modifikation. Die LFQ wurde ausgewählt und mit den Voreinstellungen durchgeführt. Als proteolytisches Enzym wurde Trypsin ausgewählt. Die Funktion ‚Match between runs‘ war aktiviert.

## 5.4.2 Proteinextraktion - Variation

### 5.4.2.1 Dreifachextraktion

Die Fraktionierung erfolgte schrittweise mit drei Extraktionslösungen nach Wieser et al. [19].

- E1 für salzlösliche Proteine (Albumine und Globuline): 0,07 mol/L  $\text{Na}_2\text{HPO}_4 \times \text{H}_2\text{O}$  und 0,4 mol/L NaCl, die mit einer Lösung, die 0,07 mol/L  $\text{KH}_2\text{PO}_4$  und 0,4 mol/L NaCl auf einen pH-Wert von 7,6 eingestellt wurde
- E2 für alkohollösliche Proteine (Gliadine): 60 % (v/v) Ethanol in Wasser
- E3 für nicht lösliche Proteine (Glutenine): 50 % (v/v) 1-Propanol in 2 mol/L Harnstoff und 0,05 mol/L TRIS-HCl (pH 7,5) plus 1 % (w/v) DTT

Vollkornmehl der Weizensorte RGT Reform (50 mg) wurde in Dreifachbestimmung in ein 2 mL Mikroröhrchen eingewogen. E1 (1 mL) wurde zugegeben und es wurde 2 min mit dem Vortexer gemischt. Die Röhrchen wurden 5 min bei 22 °C mit US behandelt, anschließend in einem Thermomixer inkubiert (1500 U/min, 10 min, 22 °C) und 15 min lang bei 21.380 rcf und 22 °C zentrifugiert. Der Überstand wurde in einem 2 mL Mikroröhrchen gesammelt und bis zur Trockene eingedampft (8 mbar, 4–6 h, 40 °C). Dem verbleibenden Pellet wurde 1 mL E2 zugegeben und es wurde 2 min gevortext. Die Röhrchen wurden 5 min bei 22 °C mit US behandelt, anschließend in einem Thermomixer inkubiert (1500 U/min, 10 min, 22 °C) und 15 min lang bei 21.380 rcf und 22 °C zentrifugiert. Der Überstand wurde in einem 2 mL Mikroröhrchen gesammelt und bis zur Trockene eingedampft (8 mbar, 4–6 h, 40 °C). Dem verbleibenden Pellet wurde 1 mL E3 zugegeben und 2 min lang gevortext. Die Röhrchen wurden 5 min bei 22 °C mit US behandelt, anschließend in einem Thermomixer inkubiert (1500 U/min, 10 min, 22 °C) und 15 min lang bei 21.380 rcf und 22 °C zentrifugiert. Der Überstand wurde in einem 2 mL Mikroröhrchen gesammelt und bis zur Trockene eingedampft (8 mbar, 4–6 h, 40 °C).

Alle Proben wurden nach der Extraktion analog zu 5.4.1 weiterverarbeitet.

#### 5.4.2.2 Dreifachextraktion vereint

Wie in 5.4.1.1, jedoch wurden die Überstände aus E1–E3 nach der Extraktion zusammengeführt und gemeinsam weiter verarbeitet.

#### 5.4.2.3 Einfachextraktion ohne/mit Ultraschall

Die Extraktion erfolgte analog zu 5.4.1, jedoch entweder gänzlich ohne oder mit 5 min US-Behandlung. Der Überstand wurde vollständig abgenommen und eingedampft.

#### 5.4.2.4 Extraktion mit ACN/Wasser

Die Einfachextraktion erfolgte mit einer Extraktionslösung E4 bestehend aus 50 % ACN in Wasser plus 0,1 % Trifluoressigsäure (TFA). Vollkornmehl (50 mg) wurde in Dreifachbestimmung in ein 2 mL Mikroröhrchen eingewogen. E4 (1 mL) wurde zugegeben und es wurde 2 min mit dem Vortexer gemischt. Die Röhrchen wurden 1 min bei 22 °C mit US behandelt, anschließend in einem Thermomixer inkubiert (1500 U/min, 45 min, 22 °C) und 15 min lang bei 21.380 rcf und 22 °C zentrifugiert. Der Überstand wurde in einem 2 mL Mikroröhrchen gesammelt und bis zur Trockene eingedampft (8 mbar, 4–6 h, 40 °C).

Variationen:

- Die US-Dauer wurde auf 5 min erhöht.
- Die US-Dauer betrug 5 min. Die Inkubation im Thermomixer erfolgte bei 60 °C und 1500 U/min über einen Zeitraum von 30 min.

#### 5.4.3 Proteinverdau: Variation Enzym-Substrat-Verhältnis

Die Aufarbeitung erfolgte analog zu 5.4.1. Bei dem Schritt Proteinverdau wurden vier verschiedene Enzym-Substrat-Verhältnisse angewendet: 1:100, 1:50, 1:25 und 1:12,5.

#### 5.4.4 Proteinaufreinigung

##### 5.4.4.1 SPE in Kartuschen

Die Aufarbeitung erfolgte analog zu 5.4.1. Die SPE wurde mit Discovery DSC-18 Röhrchen mit einem Bettgewicht von 100 mg/Well (Supelco, Sigma-Aldrich) gemäß dem Herstellerprotokoll durchgeführt (5.4.1.4).

## 5.4.5 Weitere Optimierungsversuche

### 5.4.5.1 Elution in 96-Well-Mikroplatten

Die Aufarbeitung erfolgte analog zu 5.4.1. Nach der SPE wurden die Proben in eine 96-Well-Mikroplatte eluiert und das Lösungsmittel bis zur Trockene eingedampft (8 mbar, 4–6 h, 40 °C). Die gereinigten Proben wurden in 1 mL 2 % ACN und 0,1 % FA rekonstituiert und die LC-MS/MS-Messung erfolgte direkt aus der 96-Well-Mikroplatte.

### 5.4.5.2 Aufarbeitung ohne Zwischentrocknung

Bei dem Extraktionsmittel E3 wurde 1-Propanol durch 50 % ACN ersetzt. Außerdem wurde der pH-Wert des TRIS-HCl Puffers auf 8,5 eingestellt. Die restliche Aufarbeitung erfolgte analog zu 5.4.1, wobei alle Zwischentrocknungsschritte ausgelassen wurden.

### 5.4.5.3 Aufarbeitung ohne Trocknung nach Alkylierung

Die Aufarbeitung erfolgte analog zu 5.4.1. Die Reduzierung und Alkylierung erfolgte in 100 % TRIS-HCl-Harnstoff (pH 7,8). Anschließend wurden die Proben nicht bis zur Trockene eingedampft, sondern direkt mit Trypsin verdaut.

## 5.4.6 Optimierung der MS-Parameter

Die Aufarbeitung erfolgte analog zu 5.4.1. Bei der LC-MS/MS-Messung wurden zunächst die ESI-Einstellungen variiert (**Tabelle 7**). Dabei wurde bei jeder Messung immer nur ein Parameter variiert, um dessen Einfluss zu testen.

**Tabelle 7:** Standardeinstellungen und Variationen für die ESI-Quelle bei einem LC-Fluss von 0,2 mL/min.

ESI-Quelle	Standard-einstellungen	Variation
Sprayspannung	3,0 kV	2,5-4,0
Mantelgas	35 psi	30-45
Hilfsgas	10 psi	5-10
Sweep-Gas	0 psi	0-2

ESI, Elektrospray-Ionisation; LC, Flüssigkeitschromatographie.

Es wurde ausgehend von einer adaptierten DDA-Methode nach Geisslitz et al. [82] eine Standard-DDA-Methode geschrieben (**Tabelle 8**).

**Tabelle 8:** Standardeinstellungen der DDA-Methode adaptiert von Geisslitz et al. [82].

MS1 (Full-MS)		MS2 (MS/MS)		dd Settings	
Auflösung	70.000	Auflösung	17.500	Minimum AGC Target	5.00e3
AGC Target	3e6	AGC Target	1e5	Intensity threshold	1.0e5
Maximum IT	50 ms	Maximum IT	50 ms	Apex trigger	3 to 5 s
Scan Range	360–1300 <i>m/z</i>	TopN	15	Charge exclusion	unassigned, 5-8, >8
		Isolation Window	2,4 <i>m/z</i>	Peptide match	preferred
		Fixed first mass	140,0 <i>m/z</i>	Exclude isotopes	on
		NCE	28	Dynamic exclusion	20.0 s

AGC, Automatic Gain Control; DDA, Data Dependent Acquisition; IT, Injection Time; MS1, Einstellungen des MS1-Full Scans; MS2, Einstellungen des MS/MS-Scans; NCE, Normalized Collision Energy.

Anschließend wurden verschiedene Parameter der DDA-Einstellungen variiert, wobei bei jeder Messung immer nur eine Einstellung verändert wurde (**Tabelle 9**).

**Tabelle 9:** Variation der Parameter ausgehend von der DDA-Standardmethode.

Ebene	Parameter	Wert	Kommentar
MS1	AGC Target	1e6	Standardwert für ausreichende Ionenladung
MS1	AGC Target	5e6	Höhere Ionenzahl, kann Zykluszeit verlängern
MS2	AGC Target	2e5	Typischer Wert für MS/MS
MS2	TopN	10	Geringere Sequenzierungstiefe, bessere Spektrenqualität
MS2	TopN	15	Kompromiss zwischen Tiefe und Qualität
MS2	TopN	20	Höhere Abdeckung, potenziell schlechtere MS2-Qualität
MS2	NCE	26	Moderat, geeignet für viele Peptide
MS2	NCE	30	Etwas energiereicher, z. B. für modifizierte Peptide
MS2	Maximum IT	100 ms (TopN 15)	Längere IT für bessere Spektren
MS2	Maximum IT	50 ms (TopN 10)	Kürzere IT bei niedrigem TopN
MS2	Maximum IT	100 ms (TopN 10)	Maximale Spektrenqualität bei niedrigem TopN

AGC, Automatic Gain Control; DDA, Data Dependent Acquisition; IT, Injection Time; MS1, Einstellungen des MS1-Full Scans; MS2, Einstellungen des MS/MS-Scans; NCE, Normalized Collision Energy.

## 5.5 Untargeted Proteomics mit verschiedenen Peptidasen

### 5.5.1 Proteinaufarbeitung

Die gesamte Aufarbeitung wurde bis auf den Proteinverdau analog zur Dreifachextraktion durchgeführt (5.4.2.1). Für den enzymatischen Verdau wurden die reduzierten und alkylierten Proteinextrakte in Puffer gelöst (TRIS-HCl/Urea oder HCl, je nach Enzym). Verwendet wurden Trypsin, Chymotrypsin, Pepsin, Thermolysin oder die Kombination Trypsin + Chymotrypsin (TC) in dem Enzym-Substrat-Verhältnis 1:50. Die Inkubation erfolgte im Thermomixer (200 U/min, 37 °C, 18 h bzw. 4 h für Pepsin). Für Thermolysin wurde zusätzlich 0,5 mmol/L CaCl<sub>2</sub> zugesetzt; Pepsinverdau erfolgte in 0,15 mol/L HCl (pH<2). Die Aktivität aller Enzyme wurde gestoppt, indem die Röhrchen 5 min bei 95 °C erhitzt wurden. Die Extrakte wurden anschließend direkt mit SPE gereinigt.

### 5.5.2 LC-MS/MS-Messung

Die gereinigten Proben wurden in 1 mL 2 % ACN und 0,1% FA (E1–2) bzw. 0,5 mL 2 % ACN und 0,1 % FA (E3) rekonstituiert. Die LC-MS/MS-Messungen wurden mit einem Vanquish UHPLC-System (Thermo Fisher Scientific, Waltham, MA, USA), das an ein Orbitrap Q Exactive plus MS/MS-System (Thermo Fisher Scientific) gekoppelt war, durchgeführt.

Die Peptide (Injektionsvolumen von 20 µL) wurden auf einer Aeris PEPTIDE XB-C18 (1,7 µm, 10 nm, 150 mm x 2,1 mm) LC-Säule (Phenomenex, Torrance, CA, USA) bei einer Flussrate von 0,2 mL/min getrennt. Die verwendeten Lösungsmittel waren (A) 0,1 % FA in Wasser und (B) 0,1 % FA in ACN. Der lineare Gradient wurde wie folgt eingestellt: 2 % B bei 0 min, 10 % B bei 1 min, 30 % B bei 18 min, 40 % B bei 21 min, 60 % B bei 23 min, 80 % B für 25–27 min und 2 % B für 28–35 min. Der Eluent der UHPLC wurde direkt an die ESI-Quelle des Q Exactive plus MS/MS-Systems gekoppelt. Die Sprühspannung wurde auf 3,0 kV, das Mantelgas auf 35 und das Hilfgas auf 10 eingestellt. Es wurde kein Sweep-Gas verwendet. Die Kapillartemperatur wurde auf 350 °C und die Radiofrequenz der S-Linse auf 60 eingestellt.

Die Daten wurden im DDA-Modus im positiven ESI-Modus aufgenommen. Die vollständigen MS-Parameter wurden wie folgt eingestellt: Resolution: 70.000, AGC Target: 3e6, Maximum Injection Time (IT): 50 ms, Scan Range: 360-1300 *m/z*. Die Parameter für dd-MS<sup>2</sup>-Scans wurden wie folgt festgelegt: Resolution: 17.500; AGC Target: 1e5; Maximum IT: 100 ms; TopN: 10; Isolation Window: 2,0 *m/z*; Fixed First Mass: 120,0 *m/z*; Normalized Collision Energy: 28; Dynamic Ion Exclusion: 20 s. Die Laufzeit der Methode betrug 35 min.

### 5.5.3 Proteinidentifizierung und -quantifizierung

Für die Proteinidentifizierung wurde die Software MaxQuant (Version 2.4.9.0) mit der integrierten Suchmaschine Andromeda verwendet [71]. Zur Identifikation wurden die MS/MS-Rohdaten mit der Weizenproteindatenbank (UniProtKB) verglichen (Fasta-Datei für Organismus-ID [4565], heruntergeladen am 28. September 2023 mit 151.978 Proteineinträgen). Die Standardeinstellungen von MaxQuant wurden bis auf die folgenden Änderungen beibehalten. Die Oxidation von Methionin und die N-terminale Proteinacetylierung wurden als variable Modifikationen und die Carbamidomethylierung an Cysteinen als fixe Modifikationen festgelegt. Die LFQ wurde ausgewählt und die restlichen Standardeinstellungen beibehalten. Trypsin, Chymotrypsin, Thermolysin, Pepsin (pH<2) und die Kombination TC (beide Peptidasen wurden im MaxQuant-Lauf ausgewählt) wurden ausgewählt. Der Punkt „Match between runs“ wurde aktiviert. Die MaxQuant-Auswertungen wurden für jedes Enzym einzeln durchgeführt. Für eine detailliertere Untersuchung wurden die Ergebnisse aller Enzyme in einem MaxQuant-Lauf für die Sorte RGT Reform ausgewertet.

## 5.6 Methodenentwicklung targeted Proteomics

### 5.6.1 Peptidstandards

Die Peptide (**Tabelle 10**) wurden von GenScript (Piscataway, NJ, USA) als unmarkierte Peptide (P1–P11) und als mit stabilen Isotopen markierte interne Standards (IS1–IS11) synthetisiert. Für die Stammlösungen wurden die Peptide in Wasser (1 mg/mL) gelöst und vor der Verwendung bei -80 °C gelagert.

**Tabelle 10:** Unmarkierte (P) und markierte Peptide (IS), welche als Standard zur Quantifizierung der Glutenproteine verwendet wurden.

Peptide	Glutenproteintyp	Aminosäuresequenz	Reinheit [%]
P1/IS1	α-Gliadin 1t	<u>C</u> QAIHNVVHAILHQQQQ* <u>R</u>	90,2/93,6
P2/IS2	α-Gliadin 2t	NLALETLPAM <u>C</u> NVYIPPY <u>C</u> TIAPVGI*FGTN	90,6/93,0
P3/IS3	γ-Gliadin 1t	SLVLQTLPSM <u>C</u> NVYVPPE <u>C</u> SIM* <u>R</u>	91,9/93,8
P4/IS4	LMW-GS 1t	SVNVPLY* <u>R</u>	94,8/96,7
P5/IS5	LMW-GS 2t	TLPTM <u>C</u> SVNVPLYSSITSAPLGVGS* <u>R</u>	90,9/90,3
P6/IS6	LMW-GS 3t	VFLQQQ <u>C</u> IPVAMQ* <u>R</u>	91,3/92,0
P7	HMW-GS 1t	GQQGYPTSLQQPGQQGQQGYPTSLQHTGQ* <u>R</u>	91,1/95,2
P8/IS8	HMW-GS 2t	EGEASEQLQ <u>C</u> E* <u>R</u>	95,5/94,2
P9/IS9	HMW-GS 3t	AQQPATQLPTV <u>C</u> * <u>R</u>	92,3/95,3
P10/IS10	HMW-GS 4t	ELQESSLEAC* <u>R</u>	93,1/96,4
P11/IS11	HMW-GS 5t	GGSFYPGETTPPQQLQ* <u>R</u>	93,5/92,7

C, S-Carboxyamidomethylcystein; \*K, Lysin (<sup>13</sup>C<sub>6</sub>, <sup>15</sup>N<sub>2</sub>); \*R, Arginin (<sup>13</sup>C<sub>6</sub>, <sup>15</sup>N<sub>4</sub>).

### 5.6.2 Identifizierung von Markerpeptiden

Die Proben wurden wie in Abschnitt 2.4 beschrieben vorbereitet, jedoch ohne Zugabe des internen Standardmixes (IS-Mix). Nach dem Lösen in 1 mL einer Lösung aus 2 % ACN und 0,1 % FA wurden sie mit einem U3000 nanoLC-System (Thermo Fisher Scientific, Waltham, MA, USA) gemessen, das mit einer Orbitrap Q Exactive Plus (Thermo Fisher Scientific, Waltham, MA, USA) gekoppelt war.

Die Peptidtrennung erfolgte auf einer bioZen™ XB-C18 nanoLC-Säule (2,6 µm, 250 mm × 0,075 mm; Phenomenex, Torrance, CA, USA) mit einer Flussrate von 0,3 µL/min. Als Laufmittel wurden (A) 0,1 % FA in Wasser und (B) 0,1 % FA in ACN verwendet. Der lineare Gradient der NC-Pumpe gestaltete sich wie folgt: 0–5 min 2 % B, 5–6 min 2–5 % B, 6–45 min 5–20 % B, 45–60 min 20–33 % B, 60–62 min 33–100 % B, 62–65 min 100 % B, 65–66 min 100–2 % B, 66–80 min 2 % B. Das Injektionsvolumen betrug 10 µL, die Säulentemperatur wurde auf 40 °C eingestellt. Der isokratische Gradient der Loading-Pumpe war 100 % A mit einer Flussrate von 8,0 µL/min. Die Sprühspannung der nanoFlex-Quelle lag bei 2,8 kV.

Die Datenaufnahme erfolgte im DDA-Modus. Die Parameter für die MS1-Scans waren: Resolution: 140.000; AGC Target: 3e6; Maximum IT: 50 ms; Scan Range: 300–1300  $m/z$ . Die Parameter für die dd-MS2-Scans waren: Resolution 17.500; AGC Target: 1e5; Maximum IT: 50 ms; TopN: 10; Isolation Window: 2,4  $m/z$ ; Fixed First Mass: 120,0  $m/z$ ; Normalized Collision Energy: 28; Dynamic Ion Exclusion: 45 s. Die Gesamtlaufzeit der Methode betrug 82 min.

### 5.6.3 Probenaufarbeitung für SIDA und targeted LC-MS/MS

Die Probenaufarbeitung erfolgte analog zu 5.4.1. Vor dem Reduktionsschritt wurde der interne Standard (IS) als Mix zugegeben. Der IS-Mix wurde aus einzelnen Peptiden (1 mg/mL-Stammlösungen) in folgenden Volumina zusammengestellt: 70  $\mu\text{L}$  ( $\alpha$ -Gliadin 1t), 120  $\mu\text{L}$  ( $\alpha$ -Gliadin 2t), 90  $\mu\text{L}$  ( $\gamma$ -Gliadin 1t), 10  $\mu\text{L}$  (LMW-GS 1), 20  $\mu\text{L}$  (LMW-GS 2), 50  $\mu\text{L}$  (LMW-GS 3), 70  $\mu\text{L}$  (HMW-GS 1), jeweils 20  $\mu\text{L}$  (HMW-GS 2 und 3), 30  $\mu\text{L}$  (HMW-GS 4) sowie 40  $\mu\text{L}$  (HMW-GS 5). Die Mischung wurde anschließend mit Wasser auf ein Endvolumen von 1 mL (insgesamt 460  $\mu\text{L}$  Wasser) aufgefüllt.

### 5.6.4 Responsegerade

Die Responsegerade wurde hergestellt basierend auf Geisslitz et al. [82, 126]. Zwei Lösungen mit jeweils 50  $\mu\text{L}$  jedes Peptids wurden aus den Stammlösungen hergestellt: Lösung 1 enthielt die Peptide P1–P11, Lösung 2 die internen Standards IS1–IS11. Jeweils ein Aliquot der Lösungen 1 und 2 wurde analog zu den Weizenmehlproben mit TCEP reduziert und mit CAA alkyliert (5.4.1.2). Die alkylierten Lösungen 1 und 2 (jeweils 20  $\mu\text{g/mL}$  jedes Peptids) wurden zur Kalibrierung in molaren Verhältnissen von  $n(\text{P})/n(\text{IS})$  zwischen 9,1 und 0,1 gemischt (9 + 1, 4 + 1, 3 + 1, 1 + 1, 1 + 3, 1 + 4, 1 + 9).

Die Response wurde insgesamt 31 Mal gemessen. Um vergleichbare Ergebnisse für die Proben zu erzielen, mussten Mittelwerte über alle Response-Kurven aller Peptide gebildet werden. Daher wurden alle Response-Werte auf Ausreißer überprüft und diese entfernt. Anschließend wurde für jeden Punkt der Response-Kurven der Mittelwert berechnet. Die Berechnung der Peptidmengen in den Proben erfolgte auf Basis dieser gemittelten Responses. Obwohl dies für instabile Peptide (P4 und P11) zu einer höheren Abweichung vom tatsächlich gemessenen Response führen kann, war dies die einzige Möglichkeit, einheitliche Ergebnisse für alle Peptide in den Proben zu erzielen. Alle anderen Peptide (P1, P2, P3, P5, P6, P7, P8, P9 und P10) zeigten stabile Response-Verläufe, sodass der Mittelwert den jeweils gemessenen Werten entsprach. Die Steigungen aller Response-Kurven lagen zwischen 0,6 und 1,3, während die Achsenabschnitte nahe 0,0 lagen. Diese Werte stimmen mit Ergebnissen aus anderen Studien zu

Weizenproteinen überein [82, 194]. Die Reinheiten der nicht alkylierten Standards wurden vom Hersteller bestimmt und in die Berechnungen einbezogen (5.6.1, **Tabelle 10**).

### 5.6.5 Targeted LC-MS/MS Messung

Die Proben wurden in 1 mL einer Lösung aus 2 % ACN und 0,1 % FA rekonstituiert. Die Messung erfolgte mit einem Vanquish UHPLC-System (Thermo Fisher Scientific, Waltham, MA, USA), gekoppelt an eine Orbitrap Q Exactive Plus (Thermo Fisher Scientific, Waltham, MA, USA).

Die Peptide wurden auf einer Aeris™ PEPTIDE XB-C18-Säule (1,7  $\mu\text{m}$ , 10 nm, 150 mm  $\times$  2,1 mm; Phenomenex, Torrance, CA, USA) bei einer Flussrate von 0,2 mL/min getrennt. Als Laufmittel dienten (A) 0,1 % FA in Wasser und (B) 0,1 % FA in ACN. Der Gradient war wie folgt: 0–2 min 2–10 % B, 2–28 min 10–30 % B, 28–34 min 30–40 % B, 34–38 min 40–60 % B, 38–42 min 60–80 % B, 42–46 min 80 % B, 46–48 min 80–2 % B, 48–55 min 2 % B. Das Injektionsvolumen betrug 10  $\mu\text{L}$ ; die Säulentemperatur lag bei 30 °C. Der Eluent der UHPLC wurde direkt an die ESI-Quelle der Q Exactive plus MS/MS-Systeme gekoppelt. Die Sprühspannung wurde auf 3,0 kV, das Mantelgas auf 35 und das Hilfsgas auf 10 eingestellt. Es wurde kein Sweep-Gas verwendet. Die Kapillartemperatur wurde auf 350 °C und die Radiofrequenz der S-Linse auf 60 eingestellt.

Die Datenerfassung erfolgte im PRM-Modus. Die MS2-Parameter wurden wie folgt eingestellt: Resolution: 17.500; AGC Target: 5e5; Maximum IT: 50 ms; Isolation Window: 2,4  $m/z$ . Eine Isolationsliste mit den  $m/z$ -Werten der internen Standards wurde hinterlegt. Die Kollisionsenergie für jedes Peptid wurde in Vorversuchen individuell optimiert. Die Gesamtlauzeit der Methode betrug 30 min.

### 5.6.6 Proteinquantifizierung

Die PRM-Peakflächenintegration wurde mit Skyline (64-bit) Version 24.1.0.199 (6a0775ef83) (MacCoss Lab Software, University of Washington, Seattle, WA, USA) durchgeführt [195]. Die Mittelwerte der Peakflächenverhältnisse zwischen den 5-8 Übergängen (Präkursor–Produkt-Ionenpaare) wurden für die Quantifizierung genommen. Diese konstanten Übergangsverhältnisse dienten zusätzlich als Identifikationskriterium für die Markersignale der Peptide. Die Response-Kurven wurden durch lineare Regression der Peakflächenverhältnisse  $A(P1-P11)/A(IS1-IS11)$  gegen die molaren Verhältnisse  $n(P1-P11)/n(IS1-IS11)$  dargestellt. Alle Quantifizierungen basierten auf drei biologischen Replikaten mit jeweils einer Injektion. Die Gehalte an Glutenproteingruppen wurden berechnet, indem die Peptidmengen mit dem jeweiligen Umrechnungsfaktor ( $M_{\text{Protein}}/M_{\text{Peptid}}$ ) multipliziert wurden. Dabei wurde die Molekülmasse des Proteins ohne Signalpeptid verwendet.

### 5.6.7 Präzision und Wiederfindung

Die Wiederholpräzision wurde anhand von sechs biologischen Replikaten ( $n = 6$ ) einer Weizenmehlmischung (acht Elternlinien der BMWpop und Vergleichsstandard RGT Reform) bewertet. Die durchschnittliche Präzision wurde durch Analyse von insgesamt 18 biologischen Replikaten an drei unterschiedlichen Tagen bestimmt (jeweils  $n = 6$ ; gleicher Analytiker, gleiches Gerät) [196]. Die Wiederfindung wurde nach Geisslitz et al. [126] als Verhältnis des in der Weizenmehlmischung gemessenen Peptidgehalts zur entsprechenden Probe definiert, bei der die Mehlmischung im Verhältnis 1 : 1 mit glutenfreier Weizenstärke verdünnt wurde.

### 5.6.8 Bestimmungs- und Nachweisgrenze

Die Ermittlung der LOD und der LOQ erfolgte nach Geisslitz et al. [126]. Dafür wurden die synthetischen Zielseptide (P1–P11) sowie die entsprechenden internen Standards (IS1–IS11) in acht abgestuften Konzentrationen (2, 1,5, 1,25, 1,00, 0,75, 0,5, 0,25 und 0,1  $\mu\text{g}/\text{mL}$ ) auf 50 mg glutenfreie Weizenstärke als analytenfreie Matrix aufgestockt ( $n = 3$ ). Die Probenaufarbeitung und LC-MS/MS-Analyse erfolgten analog zu 5.4.1. Zur Berechnung von LOD und LOQ wurde der niedrigste Konzentrationspunkt herangezogen, bei dem alle Identifikationskriterien erfüllt waren. Die LOD wurde dabei als das Dreifache, die LOQ als das Zehnfache der Standardabweichung dieses letzten gültigen Konzentrationspunkts definiert [197].

### 5.6.9 Targeted Proteomics: Finale LC-MS/MS Aufarbeitung

Die Aufarbeitung erfolgte analog zu 5.4.1. Vor der Reduktion wurden jeweils 50  $\mu\text{L}$  des IS-Mix hinzugefügt.

**Tabelle 11:** Zusammensetzung der stabilisotopenmarkierten Peptide in dem internen Standardmix (IS-Mix).

Label	Bemerkung	Volumen [ $\mu\text{L}$ ]
IS1	Aus 1 mg/mL	70
IS2	Aus 1 mg/mL	120
IS3	Aus 1 mg/mL	90
IS4	Aus 1 mg/mL	10
IS5	Aus 1 mg/mL	20
IS6	Aus 1 mg/mL	50
IS7	Aus 1 mg/mL	70
IS8	Aus 1 mg/mL	20
IS9	Aus 1 mg/mL	20
IS10	Aus 1 mg/mL	30
IS11	Aus 1 mg/mL	40
H <sub>2</sub> O		460

IS, interner Standard.

## 5.7 Instrumentelle Angaben

<b>RP-HPLC</b>	
Hersteller	Shimadzu (Kyoto, Japan)
Modell	Prominence UFLC
Degasser	DGU-20A5R
Pumpe	LC-20AD
Autosampler	SIL-20AC
Säulenofen	CTO-20AC
Detektor	Dioden-Array-Detektor, SPD000-M20A
Kommunikationsmodul	CBM-20A
Software	LabSolutions V5.93
<b>UHPLC-MS/MS</b>	
Hersteller	Thermo Fisher Scientific
Modell	Vanquish
Pumpe	VF-P10-A
Autosampler	VF-A10-A
Säulenofen	VH-C10-A
Detektor	DAD, VF-D11-A
Massenanalysator	Orbitrap Q Exactive Plus
Software	Xcalibur V 4.2.47
<b>nanoLC-MS/MS</b>	
Hersteller	Thermo Fisher Scientific
Modell	Dionex UltiMate 3000 RSLCnano System
Pumpe	NCS-3500RS Nano ProFlow
Autosampler	WPS-3000TPL RS
Massenanalysator	Orbitrap Q Exactive Plus
Software	Xcalibur V 4.2.47

## 5.8 Statistische Analysen

Die Prozessierung der Rohdaten erfolgte mit der Software MaxQuant [71]. Zur Kontrolle falsch positiver Identifizierungen wurde eine False Discovery Rate von 1 % auf Peptid- und Proteinebene angewendet (Voreinstellung von MaxQuant). Die jeweiligen methodenspezifischen Einstellungen sind in den entsprechenden Abschnitten detailliert beschrieben.

Die bereits durch MaxQuant normierten LFQ-Daten wurden vor der statistischen Auswertung mithilfe der Software Perseus (Version 2.0.3.0) einer standardisierten Vorverarbeitung unterzogen [198]: Proteingruppen, die als potenzielle Kontaminationen (Only identified by site, Reverse, Potential contaminant) identifiziert wurden, wurden aus dem Datensatz entfernt. Für die weitere Analyse der relativen Häufigkeit von Glutenproteinen wurden ausschließlich Proteingruppen berücksichtigt, die in mindestens zwei der drei technischen Replikate zuverlässig identifiziert worden waren. Identifikationen, die nur in einem Replikat vorlagen, wurden systematisch ausgeschlossen. Auf dieser Grundlage wurden für jede Analyse drei biologische Replikate berücksichtigt.

Zur Quantifizierung wurde der Mittelwert der LFQ-Intensitäten aus den jeweiligen Replikaten berechnet. Die Berechnung von Mittelwerten und Standardabweichungen erfolgte mithilfe von Microsoft Office Excel 2016 (Microsoft Corporation, Seattle, WA, USA).

Säulen- und Boxplot-Diagramme wurden mit OriginPro 2023 (OriginLab, Northampton, MA, USA) erstellt. Für weiterführende statistische Auswertungen und Visualisierungen wurden die integrierten Module HeatMap Dendrogram (Version 2.00), Correlation Plot (Version 1.31) und Venn Diagram (Version 1.10) eingesetzt.

Zur Beurteilung der Stärke der Zusammenhänge wurden Pearson-Korrelationskoeffizienten ( $r$ ) verwendet. Die Klassifikation basiert auf dem Betrag des Korrelationskoeffizienten ( $|r|$ ) und wurde unabhängig vom Vorzeichen in gleich große Wertebereiche vorgenommen: sehr schwach ( $|r| < 0,2$ ), schwach ( $0,2 \leq |r| < 0,4$ ), moderat ( $0,4 \leq |r| < 0,6$ ), stark ( $0,6 \leq |r| < 0,8$ ), sehr stark ( $|r| \geq 0,8$ ). Zusätzlich wurde für alle Korrelationen die statistische Signifikanz anhand eines Signifikanzniveaus von  $p \leq 0,05$  berücksichtigt.

## 6 Literaturverzeichnis

- [1] Food and Agriculture Organization of the United Nations (2024). *Food outlook – biannual report on global food markets*. FAO, Rome, ITA. <https://doi.org/10.4060/cd1158en>.
- [2] M. Pfeifer, K. G. Kugler, S. R. Sandve, B. Zhan, H. Rudi, et al. (2014). Genome interplay in the grain transcriptome of hexaploid bread wheat. *Science*, 345(6194):1250091-1 - 1250091-7. <https://doi.org/10.1126/science.1250091>.
- [3] C. F. Morris (2016). Cereals: overview of uses: accent on wheat grain. In C. Wrigley, K. Seetharaman, H. Corke & J. Faubion (Hrsg.), *Encyclopedia of Food Grains*, Seiten 1-7. Elsevier, Kidlington, GBR. <https://doi.org/10.1016/b978-0-12-394437-5.00142-x>.
- [4] The International Wheat Genome Sequencing Consortium (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, 345(6194):1251788-1 - 1251788-11. <https://doi.org/10.1126/science.1251788>.
- [5] D. B. Lobell, W. Schlenker & J. Costa-Roberts (2011). Climate trends and global crop production since 1980. *Science*, 333(6042):616-620. <https://doi.org/10.1126/science.1204531>.
- [6] K. Butterbach-Bahl, E. M. Baggs, M. Dannenmann, R. Kiese & S. Zechmeister-Boltenstern (2013). Nitrous oxide emissions from soils: how well do we understand the processes and their controls? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1621):20130122. <https://doi.org/10.1098/rstb.2013.0122>.
- [7] Bundesministerium für Ernährung und Landwirtschaft (2017). Verordnung über die Anwendung von Düngemitteln, Bodenhilfsstoffen, Kultursubstraten und Pflanzenhilfsmitteln nach den Grundsätzen der guten fachlichen Praxis beim Düngen (Düngeverordnung – DüV). Abrufbar unter: <https://www.ble-medien-service.de/1756-4-duengeverordnung-2020.html>.
- [8] P. R. Shewry, N. G. Halford, P. S. Belton & A. S. Tatham (2002). The structure and properties of gluten: an elastic protein from wheat grain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 357(1418):133-142. <https://doi.org/10.1098/rstb.2001.1024>.
- [9] P. R. Shewry & P. Tosi (2016). Protein synthesis and deposition. In C. Wrigley, K. Seetharaman, H. Corke & J. Faubion (Hrsg.), *Encyclopedia of Food Grains*, Seiten 115-122. Elsevier, Kidlington, Oxford, GBR. <https://doi.org/10.1016/b978-0-12-394437-5.00106-6>.
- [10] L. Hajas, K. A. Scherf, K. Torok, Z. Bugyi, E. Schall, et al. (2018). Variation in protein composition among wheat (*Triticum aestivum* L.) cultivars to identify cultivars suitable as reference material for wheat gluten analysis. *Food Chemistry*, 267:387-394. <https://doi.org/10.1016/j.foodchem.2017.05.005>.
- [11] S. Geisslitz, H. Wieser, K. A. Scherf & P. Koehler (2018). Gluten protein composition and aggregation properties as predictors for bread volume of common wheat, spelt, durum wheat, emmer and einkorn. *Journal of Cereal Science*, 83:204-212. <https://doi.org/10.1016/j.jcs.2018.08.012>.

- [12] P. Tosi (2012). Trafficking and deposition of prolamins in wheat. *Journal of Cereal Science*, 56:81-90. <https://doi.org/10.1016/j.jcs.2012.02.004>.
- [13] H. Wieser, P. Koehler & K. A. Scherf (2022). Chemistry of wheat gluten proteins: qualitative composition. *Cereal Chemistry*, 100(1):23-35. <https://doi.org/10.1002/cche.10572>.
- [14] J. D. Schofield (1994). Wheat proteins: structure and functionality in milling and breadmaking. In W. Bushuk & V. F. Rasper (Hrsg.), *Wheat - Production, Properties and Quality*, Seiten 73-106. Springer, Boston, MA, USA. [https://doi.org/10.1007/978-1-4615-2672-8\\_7](https://doi.org/10.1007/978-1-4615-2672-8_7).
- [15] T. B. Osborne (1908). Our present knowledge of plant proteins. *Science*, 28(718):417-427.
- [16] H. Wieser (2007). Chemistry of gluten proteins. *Food Microbiology*, 24(2):115-119. <https://doi.org/10.1016/j.fm.2006.07.004>.
- [17] P. R. Shewry & P. S. Belton (2024). What do we really understand about wheat gluten structure and functionality? *Journal of Cereal Science*, 117:103895. <https://doi.org/10.1016/j.jcs.2024.103895>.
- [18] B. Lexhaller, C. Tompos & K. A. Scherf (2017). Fundamental study on reactivities of gluten protein types from wheat, rye and barley with five sandwich ELISA test kits. *Food Chemistry*, 237:320-330. <https://doi.org/10.1016/j.foodchem.2017.05.121>.
- [19] H. Wieser, S. Antes & W. Seilmeier (1998). Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography. *Cereal Chemistry*, 75(5):644-650. <https://doi.org/10.1094/Cchem.1998.75.5.644>.
- [20] P. Köhler, B. Keck-Gassenmeier, H. Wieser & D. D. Kasarda (1997). Molecular modeling of the N-terminal regions of high molecular weight glutenin subunits 7 and 5 in relation to intramolecular disulfide bond formation. *Cereal Chemistry*, 74(2):154-158. <https://doi.org/10.1094/cchem.1997.74.2.154>.
- [21] M. Feldman & A. A. Levy (2023). Evolution of wheat under cultivation. In M. Feldman & A. A. Levy (Hrsg.), *Wheat Evolution and Domestication*, Seiten 605-663. Springer, Basel, CHE. [https://doi.org/10.1007/978-3-031-30175-9\\_13](https://doi.org/10.1007/978-3-031-30175-9_13).
- [22] G. Petersen, O. Seberg, M. Yde & K. Berthelsen (2006). Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). *Molecular Phylogenetics and Evolution*, 39:70-82. <https://doi.org/10.1016/j.ympev.2006.01.023>.
- [23] E. Cavalet-Giorsa, A. Gonzalez-Munoz, N. Athiyannan, S. Holden, A. Salhi, et al. (2024). Origin and evolution of the bread wheat D genome. *Nature*, 633:848-855. <https://doi.org/10.1038/s41586-024-07808-z>.
- [24] K. Arumuganathan & E. D. Earle (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter*, 9(3):208-218. <https://doi.org/10.1007/bf02672069>.

- [25] The International Wheat Genome Sequencing Consortium (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361(6403):ear7191. <https://doi.org/10.1126/science.aar7191>.
- [26] P. I. Payne (1987). Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Annual Review of Plant Biology*, 38:141-153. <https://doi.org/10.1146/annurev.pp.38.060187.001041>.
- [27] A. Juhasz, T. Belova, C. G. Florides, C. Maulis, I. Fischer, et al. (2018). Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Science Advances*, 4(8):ear8602. <https://doi.org/10.1126/sciadv.aar8602>.
- [28] K. Pourmohammadi, E. Abedi & S. M. B. Hashemi (2023). Gliadin and glutenin genomes and their effects on the technological aspect of wheat-based products. *Current Research in Food Science*, 7:100622. <https://doi.org/10.1016/j.crfs.2023.100622>.
- [29] P. R. Shewry, N. G. Halford & D. Lafiandra (2003). Genetics of wheat gluten proteins. *Advances in Genetics*, 49:111-184. [https://doi.org/10.1016/s0065-2660\(03\)01003-4](https://doi.org/10.1016/s0065-2660(03)01003-4).
- [30] B. G. Cassidy, J. Dvorak & O. D. Anderson (1998). The wheat low-molecular-weight glutenin genes: characterization of six new genes and progress in understanding gene family structure. *Theoretical and Applied Genetics*, 96:743-750. <https://doi.org/10.1007/s001220050797>.
- [31] P. I. Payne & K. G. Corfield (1979). Subunit composition of wheat glutenin proteins, isolated by gel filtration in a dissociating medium. *Planta*, 145:83-88. <https://doi.org/10.1007/bf00379931>.
- [32] S. Masci, J. L. E. Lew, D. Lafiandra, E. Porceddu & D. D. Kasarda (1995). Characterization of low molecular weight glutenin subunits in durum wheat by reversed-phase high-performance liquid chromatography and N-terminal sequencing. *Cereal Chemistry*, 72(1):100-104.
- [33] T. M. Ikeda, T. Nagamine & H. Yano (2004). Characterisation of LMW-GS genes and the corresponding proteins in common wheats. In D. Lafiandra, S. Masci & R. D'Ovidio (Hrsg.), *The gluten proteins*, Seiten 101-104. Royal Society of Chemistry, London, GBR. <https://doi.org/10.1039/9781847552099-00101>.
- [34] R. D'Ovidio & S. Masci (2004). The low-molecular-weight glutenin subunits of wheat gluten. *Journal of Cereal Science*, 39(3):321-339. <https://doi.org/10.1016/j.jcs.2003.12.002>.
- [35] H. Wieser, P. Koehler & K. A. Scherf (2022). Chemistry of wheat gluten proteins: quantitative composition. *Cereal Chemistry*, 100(1):36-55. <https://doi.org/10.1002/cche.10553>.
- [36] B. Lagrain, I. Rombouts, H. Wieser, J. Delcour & P. Koehler (2012). A reassessment of the electrophoretic mobility of high molecular weight glutenin subunits of wheat. *Journal of Cereal Science*, 56(3):726-732. <https://doi.org/10.1016/j.jcs.2012.08.003>.
- [37] S. A. Boden, R. A. McIntosh, C. Uauy, S. G. Krattinger, J. Dubcovsky, et al. (2023). Updated guidelines for gene nomenclature in wheat. *Theoretical and Applied Genetics*, 136(4):72. <https://doi.org/10.1007/s00122-023-04253-w>.

- [38] M. Stadlmeier, L. Hartl & V. Mohler (2018). Usefulness of a multiparent advanced generation intercross population with a greatly reduced mating design for genetic studies in winter wheat. *Frontiers in Plant Science*, 9:1825. <https://doi.org/10.3389/fpls.2018.01825>.
- [39] G. Langenkämper & C. Zörb (2019). Modern aspects of wheat grain proteins. *Journal of Applied Botany and Food Quality*, 92:240-245. <https://doi.org/10.5073/JABFQ.2019.092.033>.
- [40] J. Huen, J. Börsmann, I. Matullat, L. Böhm, F. Stukenborg, et al. (2017). Wheat flour quality evaluation from the baker's perspective: comparative assessment of 18 analytical methods. *European Food Research and Technology*, 244(3):535-545. <https://doi.org/10.1007/s00217-017-2974-3>.
- [41] S. Melis & J. A. Delcour (2020). Impact of wheat endogenous lipids on the quality of fresh bread: Key terms, concepts, and underlying mechanisms. *Comprehensive Reviews in Food Science and Food Safety*, 19(6):3715-3754. <https://doi.org/10.1111/1541-4337.12616>.
- [42] T. Amend & H. D. Belitz (1991). Microstructural studies of gluten and a hypothesis on dough formation. *Food Structure*, 10(4):277-288.
- [43] J. A. Delcour, I. J. Joye, B. Pareyt, E. Wilderjans, K. Brijs, et al. (2012). Wheat gluten functionality as a quality determinant in cereal-based food products. *Annual Review of Food Science and Technology*, 3:469-492. <https://doi.org/10.1146/annurev-food-022811-101303>.
- [44] ICC Standard No 131 (1980). *Method for test baking of wheat flours*. International Association for Cereal Science and Technology.
- [45] R. Kieffer, H.-D. Belitz, M. Zweier, R. Ipfelkofer & G. Fischbeck (1993). Der Rapid-Mix-Test als 10-g-Mikrobackversuch. *Zeitschrift für Lebensmittel-Untersuchung und -Forschung*, 197(2):134-136. <https://doi.org/10.1007/bf01260309>.
- [46] S. M. Thanhaeuser, H. Wieser & P. Koehler (2014). Correlation of quality parameters with the baking performance of wheat flours. *Cereal Chemistry*, 91(4):333-341. <https://doi.org/10.1094/Cchem-09-13-0194-Cesi>.
- [47] D. Gabriel, C. Pfitzner, N. U. Haase, A. Hüsken, H. Prüfer, et al. (2017). New strategies for a reliable assessment of baking quality of wheat – rethinking the current indicator protein content. *Journal of Cereal Science*, 77:126-134. <https://doi.org/10.1016/j.jcs.2017.08.002>.
- [48] K. F. Finney & M. A. Barmore (1948). Loaf volume and protein content of hard winter and spring wheats. *Cereal Chemistry*, 25:291-312.
- [49] ICC Standard No 167 (2000). *Determination of crude protein in grain and grain products for food and feed by the Dumas combustion principle*. International Association for Cereal Science and Technology.
- [50] ICC Standard No 137/1 (1994). *Mechanical determination of the wet gluten content of wheat flour*. International Association for Cereal Science and Technology.

- [51] ICC Standard No 155 (1980). *Determination of wet gluten quantity and quality (gluten index according to Perten) of whole wheat meal and wheat flour*. International Association for Cereal Science and Technology.
- [52] ICC Standard No 116/1 (1994). *Determination of sedimentation value (according to Zeleny) as an approximate measure of baking quality*. International Association for Cereal Science and Technology.
- [53] ICC Standard No 107/1 (1995). *Determination of the 'falling number' according to Hagberg-Perten as a measure of the degree of alpha-amylase activity in grain and flour*. International Association for Cereal Science and Technology.
- [54] ICC Standard No 115/1 (1992). *Method for using the Brabender farinograph*. International Association for Cereal Science and Technology.
- [55] S. Uthayakumaran & C. Wrigley (2017). Wheat: grain-quality characteristics and management of quality requirements. In I. Batey, D. Miskelly & C. Wrigley (Hrsg.), *Cereal Grains - Assessing and Managing Quality*, Seiten 91-134. Woodhead Publishing, Cambridge, GBR.  
<https://doi.org/10.1016/b978-0-08-100719-8.00005-x>.
- [56] F. E. Dowell, E. B. Maghirang, R. O. Pierce, G. L. Lookhart, S. R. Bean, et al. (2008). Relationship of bread quality to kernel, flour, and dough properties. *Cereal Chemistry*, 85(1):82-91.  
<https://doi.org/10.1094/cchem-85-1-0082>.
- [57] F. Békés (2012). New aspects in quality related wheat research: II. New methodologies for better quality wheat. *Cereal Research Communications*, 40(3):307-333.  
<https://doi.org/10.1556/crc.2012.0003>.
- [58] H. Wieser & R. Kieffer (2001). Correlations of the amount of gluten protein types to the technological properties of wheat flours determined on a micro-scale. *Journal of Cereal Science*, 34(1):19-27. <https://doi.org/10.1006/jcrs.2000.0385>.
- [59] C. Schuster, J. Huen & K. A. Scherf (2022). Comprehensive study on gluten composition and baking quality of winter wheat. *Cereal Chemistry*, 100(1):142-155.  
<https://doi.org/10.1002/cche.10606>.
- [60] P. I. Payne, M. A. Nightingale, A. F. Krattiger & L. M. Holt (1987). The relationship between HMW glutenin subunit composition and the bread-making quality of British-grown wheat varieties. *Journal of the Science of Food and Agriculture*, 40(1):51-65.  
<https://doi.org/10.1002/jsfa.2740400108>.
- [61] O. D. Anderson & F. Bekes (2011). Incorporation of high-molecular-weight glutenin subunits into doughs using 2 gram mixograph and extensigraphs. *Journal of Cereal Science*, 54(3):288-295.  
<https://doi.org/10.1016/j.jcs.2011.06.001>.
- [62] C. Guzmán, J. Crossa, S. Mondal, V. Govindan, J. Huerta, et al. (2022). Effects of glutenins (Glu-1 and Glu-3) allelic variation on dough properties and bread-making quality of CIMMYT bread wheat breeding lines. *Field Crops Research*, 284:108585.  
<https://doi.org/10.1016/j.fcr.2022.108585>.

- [63] J. V. Jorin Novo (2021). Proteomics and plant biology: contributions to date and a look towards the next decade. *Expert Review of Proteomics*, 18(2):93-103. <https://doi.org/10.1080/14789450.2021.1910028>.
- [64] S. Geisslitz & K. A. Scherf (2021). Proteomics in food quality. In A. Cifuentes (Hrsg.), *Comprehensive Foodomics*, Seiten 699-717. Elsevier, Amsterdam, NLD. <https://doi.org/10.1016/b978-0-08-100596-5.22762-5>.
- [65] F. W. McLafferty, K. Breuker, M. Jin, X. Han, G. Infusini, et al. (2007). Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS Journal*, 274(24):6256-6268. <https://doi.org/10.1111/j.1742-4658.2007.06147.x>.
- [66] E. J. Dupree, M. Jayathirtha, H. Yorkey, M. Mihasan, B. A. Petre, et al. (2020). A critical review of bottom-up proteomics: the good, the bad, and the future of this field. *Proteomes*, 8(3):14. <https://doi.org/10.3390/proteomes8030014>.
- [67] E. S. Hecht, M. Scigelova, S. Eliuk & A. Makarov (2019). Fundamentals and advances of orbitrap mass spectrometry. In R. A. Meyers (Hrsg.), *Encyclopedia of Analytical Chemistry*, Seiten 1-40. Wiley, Hoboken, NJ, USA. <https://doi.org/10.1002/9780470027318.a9309.pub2>.
- [68] M. L. Alexandrov, L. N. Gall, N. V. Krasnov, V. I. Nikolaev, V. A. Pavlenko, et al. (2008). Extraction of ions from solutions under atmospheric pressure as a method for mass spectrometric analysis of bioorganic compounds. *Rapid Communications in Mass Spectrometry*, 22(3):267-270. <https://doi.org/10.1002/rcm.3113>.
- [69] N. W. Bateman, S. P. Goulding, N. J. Shulman, A. K. Gadok, K. K. Szumlinski, et al. (2014). Maximizing peptide identification events in proteomic workflows using data-dependent acquisition (DDA). *Molecular & Cellular Proteomics*, 13(1):329-338. <https://doi.org/10.1074/mcp.M112.026500>.
- [70] C. Ludwig & R. Aebersold (2015). Getting absolute: determining absolute protein quantities via selected reaction monitoring mass spectrometry. In S. Gaskell & C. Eyers (Hrsg.), *Quantitative Proteomics - New Developments in Mass Spectrometry*, Seiten 80-109. Royal Society of Chemistry, London, GBR. <https://doi.org/10.1039/9781782626985-00080>.
- [71] S. Tyanova, T. Temu & J. Cox (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols*, 11(12):2301-2319. <https://doi.org/10.1038/nprot.2016.136>.
- [72] Y. Jiang, D. A. B. Rex, D. Schuster, B. A. Neely, G. L. Rosano, et al. (2024). Comprehensive overview of bottom-up proteomics using mass spectrometry. *ACS Measurement Science Au*, 4(4):338-417. <https://doi.org/10.1021/acsmeasuresciau.3c00068>.
- [73] P. Sinitcyn, J. D. Rudolph & J. Cox (2018). Computational methods for understanding mass spectrometry-based shotgun proteomics data. *Annual Review of Biomedical Data Science*, 1:207-234. <https://doi.org/10.1146/annurev-biodatasci-080917-013516>.
- [74] K. A. Scherf & R. E. Poms (2016). Recent developments in analytical methods for tracing gluten. *Journal of Cereal Science*, 67:112-122. <https://doi.org/10.1016/j.jcs.2015.08.006>.

- [75] F. M. Dupont, W. H. Vensel, C. K. Tanaka, W. J. Hurkman & S. B. Altenbach (2011). Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Science*, 9:10. <https://doi.org/10.1186/1477-5956-9-10>.
- [76] S. N. Bromilow, L. A. Gethings, J. I. Langridge, P. R. Shewry, M. Buckley, et al. (2016). Comprehensive proteomic profiling of wheat gluten using a combination of data-independent and data-dependent acquisition. *Frontiers in Plant Science*, 7:2020. <https://doi.org/10.3389/fpls.2016.02020>.
- [77] M. Afzal, M. Sielaff, U. Distler, D. Schuppan, S. Tenzer, et al. (2023). Reference proteomes of five wheat species as starting point for future design of cultivars with lower allergenic potential. *npj Science of Food*, 7:9. <https://doi.org/10.1038/s41538-023-00188-0>.
- [78] L. Tsiatsiani & A. J. R. Heck (2015). Proteomics beyond trypsin. *FEBS Journal*, 282(14):2612-2626. <https://doi.org/10.1111/febs.13287>.
- [79] S. Tyanova & J. Cox (2018). Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. *Methods in Molecular Biology*, 1711:133-148. [https://doi.org/10.1007/978-1-4939-7493-1\\_7](https://doi.org/10.1007/978-1-4939-7493-1_7).
- [80] L. K. Pino, B. C. Searle, J. G. Bollinger, B. Nunn, B. MacLean, et al. (2020). The Skyline ecosystem: Informatics for quantitative mass spectrometry proteomics. *Mass Spectrometry Reviews*, 39(3):229-244. <https://doi.org/10.1002/mas.21540>.
- [81] K. Schalk, P. Koehler & K. A. Scherf (2018). 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. <https://doi.org/10.1021/acs.jafc.7b05286>.
- [82] S. Geisslitz, C. Ludwig, K. A. Scherf & P. Koehler (2018). Targeted LC-MS/MS reveals similar contents of alpha-amylase/trypsin-inhibitors as putative triggers of nonceliac gluten sensitivity in all wheat species except Einkorn. *Journal of Agricultural and Food Chemistry*, 66(46):12395-12403. <https://doi.org/10.1021/acs.jafc.8b04411>.
- [83] P. Ferranti, G. Mamone, G. Picariello & F. Addeo (2007). Mass spectrometry analysis of gliadins in celiac disease. *Journal of Mass Spectrometry*, 42(12):1531-1548. <https://doi.org/10.1002/jms.1361>.
- [84] P. R. Shewry (2009). Wheat Grain Proteins. In K. Khan & P. R. Shewry (Hrsg.), *WHEAT: Chemistry and Technology*, Seiten 223-298. Woodhead Publishing and AACC International Press, Cambridge, GBR. <https://doi.org/10.1094/9781891127557.fm>.
- [85] U. Bose, J. A. Broadbent, K. Byrne, S. Hasan, C. A. Howitt, et al. (2019). Optimisation of protein extraction for in-depth profiling of the cereal grain proteome. *Journal of Proteomics*, 197:23-33. <https://doi.org/10.1016/j.jprot.2019.02.009>.
- [86] J. Nickerson & A. Doucette (2022). Maximizing cumulative trypsin activity with calcium at elevated temperature for enhanced bottom-up proteome analysis. *Biology*, 11(10):1444. <https://doi.org/10.3390/biology11101444>.

- [87] Y. Seo, I. Kang, H. J. Lee, J. Hwang, S. H. Kwak, et al. (2024). Simple and robust high-throughput serum proteomics workflow with low-microflow LC-MS/MS. *Analytical and Bioanalytical Chemistry*, 416(29):7007-7018. <https://doi.org/10.1007/s00216-024-05603-3>.
- [88] S. Seyedmohammad, M. Zhgamadze, A. Haghani, S. Vasantdadkar, E. Daviso, et al. (2024). *Full automation of a Covaris-integrated high throughput cell-based assay system: A 96-well plate sonication workflow for cell-culture based proteomic drug and perturbation screening*. US HUPO 2024 Conference, Portland, OR, USA, March 2024.
- [89] A. H. Kverneland, F. Harking, J. M. Vej-Nielsen, M. Huusfeldt, D. B. Bekker-Jensen, et al. (2024). Fully automated workflow for integrated sample digestion and Evotip loading enabling high-throughput clinical proteomics. *Molecular & Cellular Proteomics*, 23(7):100790. <https://doi.org/10.1016/j.mcpro.2024.100790>.
- [90] S. Loroch, D. Kopczynski, A. C. Schneider, C. Schumbrutzki, I. Feldmann, et al. (2022). Toward zero variance in proteomics sample preparation: positive-pressure FASP in 96-well format (PF96) enables highly reproducible, time- and cost-efficient analysis of sample cohorts. *Journal of Proteome Research*, 21(4):1181-1188. <https://doi.org/10.1021/acs.jproteome.1c00706>.
- [91] W. J. Hervey, M. B. Strader & G. B. Hurst (2007). Comparison of digestion protocols for microgram quantities of enriched protein samples. *Journal of Proteome Research*, 6(8):3054-3061. <https://doi.org/10.1021/pr070159b>.
- [92] C. C. Liu, L. H. Liang, Y. Yang, H. L. Yu, L. Yan, et al. (2021). Direct acetonitrile-assisted trypsin digestion method combined with LC-MS/MS-targeted peptide analysis for unambiguous identification of intact ricin. *Journal of Proteome Research*, 20(1):369-380. <https://doi.org/10.1021/acs.jproteome.0c00458>.
- [93] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong & C. M. Whitehouse (1990). Electrospray ionization—principles and practice. *Mass Spectrometry Reviews*, 9(1):37-70. <https://doi.org/10.1002/mas.1280090103>.
- [94] Thermo Fisher Scientific (2009). *Ion Max and Ion Max-S API Source Hardware Manual*. Revision E, Bremen, GER.
- [95] A. Michalski, E. Damoc, J.-P. Hauschild, O. Lange, A. Wieghaus, et al. (2011). Mass Spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole orbitrap mass spectrometer. *Molecular & Cellular Proteomics*, 10(9):M111.011015. <https://doi.org/10.1074/mcp.m111.011015>.
- [96] Thermo Fisher Scientific (2018). *Exactive Series - Operating Manual*. Revision A, Bremen, GER.
- [97] C. Kaemper, J. Mossburger, M. Geyer, L. Hartl, S. Geisslitz, et al. (2025). Comparative shotgun proteomics analysis of wheat gluten proteins digested by various peptidases. *Current Research in Food Science*, 11:101095. <https://doi.org/10.1016/j.crf.2025.101095>.
- [98] T. W. Jaskolla, D. G. Papatotiriou & M. Karas (2009). Comparison between the matrices  $\alpha$ -Cyano-4-hydroxycinnamic acid and 4-Chloro- $\alpha$ -cyanocinnamic acid for trypsin, chymotrypsin, and pepsin digestions by MALDI-TOF mass spectrometry. *Journal of Proteome Research*, 8(7):3588-3597. <https://doi.org/10.1021/pr900274s>.

- [99] K. A. Scherf (2023). Gluten proteins. In P. R. Shewry, H. Koxsel & J. R. N. Taylor (Hrsg.), *ICC - Handbook of 21st Cereal Science Technology*, Seiten 71-78. Elsevier Academic Press, London, GBR. <https://doi.org/10.1016/B978-0-323-95295-8.00041-1>.
- [100] M. L. Colgrave, K. Byrne & C. A. Howitt (2017). Food for thought: Selecting the right enzyme for the digestion of gluten. *Food Chemistry*, 234:389-397. <https://doi.org/10.1016/j.foodchem.2017.05.008>.
- [101] A. Manfredi, M. Mattarozzi, M. Giannetto & M. Careri (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. <https://doi.org/10.1016/j.aca.2015.09.008>.
- [102] K. L. Fiedler, S. C. McGrath, J. H. Callahan & M. M. Ross (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. <https://doi.org/10.1021/jf500997j>.
- [103] M. C. Norwig, S. Geisslitz & K. A. Scherf (2024). Comparative label-free proteomics study on celiac disease-active epitopes in common wheat, spelt, durum wheat, emmer, and einkorn. *Journal of Agricultural and Food Chemistry*, 72(26):15040-15052. <https://doi.org/10.1021/acs.jafc.4c02657>.
- [104] W. H. Vensel, F. M. Dupont, S. Sloane & S. B. Altenbach (2011). Effect of cleavage enzyme, search algorithm and decoy database on mass spectrometric identification of wheat gluten proteins. *Phytochemistry*, 72(10):1154-1161. <https://doi.org/10.1016/j.phytochem.2011.01.002>.
- [105] B. Prandi, M. Bencivenni, T. Tedeschi, R. Marchelli, A. Dossena, et al. (2012). Common wheat determination in durum wheat samples through LC/MS analysis of gluten peptides. *Analytical and Bioanalytical Chemistry*, 403(10):2909-2914. <https://doi.org/10.1007/s00216-012-5731-2>.
- [106] O. Ogilvie, S. Roberts, K. Sutton, J. Gerrard, N. Larsen, et al. (2021). The effect of baking time and temperature on gluten protein structure and celiac peptide digestibility. *Food Research International*, 140:109988. <https://doi.org/10.1016/j.foodres.2020.109988>.
- [107] B. Lexhaller, C. Ludwig & K. A. Scherf (2020). Identification of isopeptides between human tissue transglutaminase and wheat, rye, and barley gluten peptides. *Scientific Reports*, 10:7426. <https://doi.org/10.1038/s41598-020-64143-9>.
- [108] W. H. Vensel, C. K. Tanaka & S. B. Altenbach (2014). Protein composition of wheat gluten polymer fractions determined by quantitative two-dimensional gel electrophoresis and tandem mass spectrometry. *Proteome Science*, 12(1):8. <https://doi.org/10.1186/1477-5956-12-8>.
- [109] E. Lutz, H. Wieser & P. Koehler (2012). Identification of disulfide bonds in wheat gluten proteins by means of mass spectrometry/electron transfer dissociation. *Journal of Agricultural and Food Chemistry*, 60(14):3708-3716. <https://doi.org/10.1021/jf204973u>.
- [110] D. L. Swaney, C. D. Wenger & J. J. Coon (2010). Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *Journal of Proteome Research*, 9(3):1323-1329. <https://doi.org/10.1021/pr900863u>.

- [111] M. J. Martinez-Esteso, J. Norgaard, M. Brohee, R. Haraszi, A. Maquet, et al. (2016). Defining the wheat gluten peptide fingerprint via a discovery and targeted proteomics approach. *Journal of Proteomics*, 147:156-168. <https://doi.org/10.1016/j.jprot.2016.03.015>.
- [112] I. Rombouts, B. Lagrain & J. A. Delcour (2012). Heat-induced cross-linking and degradation of wheat gluten, serum albumin, and mixtures thereof. *Journal of Agricultural and Food Chemistry*, 60(40):10133-10140. <https://doi.org/10.1021/jf3024672>.
- [113] A. Fallahbaghery, W. Zou, K. Byrne, C. A. Howitt & M. L. Colgrave (2017). Comparison of gluten extraction protocols assessed by LC-MS/MS analysis. *Journal of Agricultural and Food Chemistry*, 65(13):2857-2866. <https://doi.org/10.1021/acs.jafc.7b00063>.
- [114] B. Lexhaller, M. L. Colgrave & K. A. Scherf (2019). Characterization and relative quantitation of wheat, rye, and barley gluten protein types by liquid chromatography-tandem mass spectrometry. *Frontiers in Plant Science*, 10:1530. <https://doi.org/10.3389/fpls.2019.01530>.
- [115] A. Khalid, A. Hameed & M. F. Tahir (2023). Wheat quality: a review on chemical composition, nutritional attributes, grain anatomy, types, classification, and function of seed storage proteins in bread making quality. *Frontiers in Nutrition*, 10:1053196. <https://doi.org/10.3389/fnut.2023.1053196>.
- [116] Z. Liu, Z. Sun, J. Mao, Y. Zhang, L. Zhou, et al. (2025). Accelerating wheat (*Triticum aestivum*) proteomics using a tandem digestion approach. *Journal of Cereal Science*, 123:104133. <https://doi.org/10.1016/j.jcs.2025.104133>.
- [117] P. R. Shewry & D. Lafiandra (2022). Wheat glutenin polymers 1. structure, assembly and properties. *Journal of Cereal Science*, 106:103486. <https://doi.org/10.1016/j.jcs.2022.103486>.
- [118] X. Hu, L. Cheng, Y. Hong, Z. Li, C. Li, et al. (2023). An extensive review: How starch and gluten impact dough machinability and resultant bread qualities. *Critical Reviews in Food Science and Nutrition*, 63(13):1930-1941. <https://doi.org/10.1080/10408398.2021.1969535>.
- [119] K. Schalk, B. Lexhaller, P. Koehler & K. A. Scherf (2017). Isolation and characterization of gluten protein types from wheat, rye, barley and oats for use as reference materials. *PLoS One*, 12(2):e0172819. <https://doi.org/10.1371/journal.pone.0172819>.
- [120] P. Giansanti, T. T. Aye, H. van den Toorn, M. Peng, B. van Breukelen, et al. (2015). An Augmented Multiple-Protease-Based Human Phosphopeptide Atlas. *Cell Reports*, 11(11):1834-1843. <https://doi.org/10.1016/j.celrep.2015.05.029>.
- [121] M. Peng, N. Taouatas, S. Cappadona, B. van Breukelen, S. Mohammed, et al. (2012). Protease bias in absolute protein quantitation. *Nature Methods*, 9(6):524-525. <https://doi.org/10.1038/nmeth.2031>.
- [122] P. Giansanti, L. Tsiatsiani, T. Y. Low & A. J. R. Heck (2016). Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nature Protocols*, 11(5):993-1006. <https://doi.org/10.1038/nprot.2016.057>.

- [123] J. Woessmann, D. Kotol, A. Hober, M. Uhlén & F. Edfors (2022). Addressing the protease bias in quantitative proteomics. *Journal of Proteome Research*, 21(10):2526-2534. <https://doi.org/10.1021/acs.jproteome.2c00491>.
- [124] H. Wieser (2000). Comparative investigations of gluten proteins from different wheat species. *European Food Research and Technology*, 211(4):262-268. <https://doi.org/10.1007/s002170000165>.
- [125] M. Afzal, M. Sielaff, V. Curella, M. Neerukonda, K. El Hassouni, et al. (2021). Characterization of 150 wheat cultivars by LC-MS-based label-free quantitative proteomics unravels possibilities to design wheat better for baking quality and human health. *Plants*, 10(3):424. <https://doi.org/10.3390/plants10030424>.
- [126] S. Geisslitz, C. F. H. Longin, P. Koehler & K. A. Scherf (2020). Comparative quantitative LC-MS/MS analysis of 13 amylase/trypsin inhibitors in ancient and modern *Triticum* species. *Scientific Reports*, 10:14570. <https://doi.org/10.1038/s41598-020-71413-z>.
- [127] N. Jahn, S. Geisslitz, U. Konradl, K. Fleissner & K. A. Scherf (2025). Amylase/trypsin-inhibitor content and inhibitory activity of German common wheat landraces and modern varieties do not differ. *npj Science of Food*, 9:24. <https://doi.org/10.1038/s41538-025-00385-z>.
- [128] T. Dau, G. Bartolomucci & J. Rappsilber (2020). Proteomics using protease alternatives to trypsin benefits from sequential digestion with trypsin. *Analytical Chemistry*, 92(14):9523-9527. <https://doi.org/10.1021/acs.analchem.0c00478>.
- [129] X. Guo, D. C. Trudgian, A. Lemoff, S. Yadavalli & H. Mirzaei (2014). Confetti: A multiprotease map of the *HeLa* proteome for comprehensive proteomics. *Molecular & Cellular Proteomics*, 13(6):1573-1584. <https://doi.org/10.1074/mcp.m113.035170>.
- [130] A. Cristobal, F. Marino, H. Post, H. W. van den Toorn, S. Mohammed, et al. (2017). Toward an optimized workflow for middle-down proteomics. *Analytical Chemistry*, 89(6):3318-3325. <https://doi.org/10.1021/acs.analchem.6b03756>.
- [131] C. Kaemper, M. Geyer, L. Hartl, S. Geisslitz & K. A. Scherf (2025). Absolute quantification of gluten protein groups and their relation to wheat baking quality. *Food Research International*, 227:118228. <https://doi.org/10.1016/j.foodres.2025.118228>.
- [132] J. Reimer, D. Shamshurin, M. Harder, A. Yamchuk, V. Spicer, et al. (2011). Effect of cyclization of N-terminal glutamine and carbamidomethyl-cysteine (residues) on the chromatographic behavior of peptides in reversed-phase chromatography. *Journal of Chromatography A*, 1218(31):5101-5107. <https://doi.org/10.1016/j.chroma.2011.05.079>.
- [133] K. Schalk, P. Koehler & K. A. Scherf (2018). Targeted liquid chromatography tandem mass spectrometry to quantitate wheat gluten using well-defined reference proteins. *PLoS One*, 13(2):e0192804. <https://doi.org/10.1371/journal.pone.0192804>.
- [134] F. Calderon-Celis, J. R. Encinar & A. Sanz-Medel (2018). Standardization approaches in absolute quantitative proteomics with mass spectrometry. *Mass Spectrometry Reviews*, 37(6):715-737. <https://doi.org/10.1002/mas.21542>.

- [135] A. N. Hoofnagle, J. R. Whiteaker, S. A. Carr, E. Kuhn, T. Liu, et al. (2016). Recommendations for the generation, quantification, storage, and handling of peptides used for mass spectrometry-based assays. *Clinical Chemistry*, 62(1):48-69. <https://doi.org/10.1373/clinchem.2015.250563>.
- [136] S. Geisslitz, S. Islam, L. Buck, C. Grunwald-Gruber, F. Sestili, et al. (2022). Absolute and relative quantitation of amylase/trypsin-inhibitors by LC-MS/MS from wheat lines obtained by CRISPR-Cas9 and RNAi. *Frontiers in Plant Science*, 13:974881. <https://doi.org/10.3389/fpls.2022.974881>.
- [137] O. Duncan, J. Trosch, R. Fenske, N. L. Taylor & A. H. Millar (2017). Resource: Mapping the *Triticum aestivum* proteome. *The Plant Journal*, 89(3):601-616. <https://doi.org/10.1111/tpj.13402>.
- [138] D. Vincent, A. Bui, V. Ezernieks, S. Shahinfar, T. Luke, et al. (2022). A community resource to mass explore the wheat grain proteome and its application to the late-maturity alpha-amylase (LMA) problem. *Gigascience*, 12:1-25. <https://doi.org/10.1093/gigascience/giad084>.
- [139] M. L. Colgrave, H. Goswami, K. Byrne, M. Blundell, C. A. Howitt, et al. (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. <https://doi.org/10.1021/acs.jproteome.5b00187>.
- [140] S. B. Altenbach, H. C. Chang, M. H. Rowe, X. B. Yu, A. Simon-Buss, et al. (2020). Reducing the immunogenic potential of wheat flour: silencing of alpha gliadin genes in a U.S. wheat cultivar. *Frontiers in Plant Science*, 11:20. <https://doi.org/10.3389/fpls.2020.00020>.
- [141] S. B. Altenbach, H. C. Chang, X. B. Yu, B. W. Seabourn, P. H. Green, et al. (2019). Elimination of omega-1,2 gliadins from bread wheat (*Triticum aestivum*) flour: effects on immunogenic potential and end-use quality. *Frontiers in Plant Science*, 10:580. <https://doi.org/10.3389/fpls.2019.00580>.
- [142] J. Waga & A. Skoczowski (2013). Development and characteristics of  $\omega$ -gliadin-free wheat genotypes. *Euphytica*, 195:105-116. <https://doi.org/10.1007/s10681-013-0984-1>.
- [143] H. Wieser & W. Seilmeier (1998). The influence of nitrogen fertilisation on quantities and proportions of different protein types in wheat flour. *Journal of the Science of Food and Agriculture*, 76(1):49-55.
- [144] S. Makai, C. Eva, L. Tamas & A. Juhasz (2015). Multiple elements controlling the expression of wheat high molecular weight glutenin paralogs. *Functional & Integrative Genomics*, 15(6):661-672. <https://doi.org/10.1007/s10142-015-0441-4>.
- [145] C. Éva, B. Moncsek, K. Szőke-Pázsai, V. Kunos, K. Mészáros, et al. (2023). bZIP transcription factors repress the expression of wheat (*Triticum aestivum* L.) high molecular weight glutenin subunit genes in vegetative tissues. *Acta Physiologiae Plantarum*, 45:29. <https://doi.org/10.1007/s11738-022-03503-6>.
- [146] W. Guo, H. Yang, Y. Liu, Y. Gao, Z. Ni, et al. (2015). The wheat transcription factor TaGAMYb recruits histone acetyltransferase and activates the expression of a high-molecular-weight glutenin subunit gene. *The Plant Journal*, 84(2):347-359. <https://doi.org/10.1111/tpj.13003>.

- [147] J. Cohen (2013). *Statistical Power Analysis for the Behavioral Sciences*. 2. Edition, Routledge, New York, NY, USA. <https://doi.org/10.4324/9780203771587>.
- [148] O. D. Anderson & F. C. Greene (1997). The  $\alpha$ -gliadin gene family. II. DNA and protein sequence variation, subfamily structure, and origins of pseudogenes. *Theoretical and Applied Genetics*, 95:59-65. <https://doi.org/10.1007/s001220050532>.
- [149] S. Noma, K. Kawaura, K. Hayakawa, C. Abe, N. Tsuge, et al. (2016). Comprehensive molecular characterization of the alpha/beta-gliadin multigene family in hexaploid wheat. *Molecular Genetics and Genomics*, 291:65-77. <https://doi.org/10.1007/s00438-015-1086-7>.
- [150] S. Noma, M. Yamagishi, Y. Ogihara & K. Kawaura (2023). Characterization of  $\alpha$ -gliadin alleles of Japanese wheat cultivars in relation to flour dough extensibility and celiac disease epitopes. *Journal of Cereal Science*, 109:103591. <https://doi.org/10.1016/j.jcs.2022.103591>.
- [151] H. C. van den Broeck, T. W. van Herpen, C. Schuit, E. M. Salentijn, L. Dekking, et al. (2009). Removing celiac disease-related gluten proteins from bread wheat while retaining technological properties: a study with Chinese Spring deletion lines. *BMC Plant Biology*, 9:41. <https://doi.org/10.1186/1471-2229-9-41>.
- [152] H. C. van den Broeck, L. J. W. J. Gilissen, M. J. M. Smulders, I. M. van der Meer & R. J. Hamer (2011). Dough quality of bread wheat lacking  $\alpha$ -gliadins with celiac disease epitopes and addition of celiac-safe avenins to improve dough quality. *Journal of Cereal Science*, 53(2):206-216. <https://doi.org/10.1016/j.jcs.2010.12.004>.
- [153] D. Liu, H. Yang, Z. Zhang, Q. Chen, W. Guo, et al. (2023). An elite  $\gamma$ -gliadin allele improves end-use quality in wheat. *New Phytologist*, 239(1):87-101. <https://doi.org/10.1111/nph.18722>.
- [154] Z. Zhou, C. Liu, M. Qin, W. Li, J. Hou, et al. (2022). Promoter DNA hypermethylation of *TaGli- $\gamma$ -2.1* positively regulates gluten strength in bread wheat. *Journal of Advanced Research*, 36:163-173. <https://doi.org/10.1016/j.jare.2021.06.021>.
- [155] J. D. Sherman, A. C. Varella, S. P. Lanning, J. M. Martin, H. Y. Heo, et al. (2018). Effect of a gene for high dough strength on whole wheat baking parameters of hard white spring wheat. *Cereal Chemistry*, 95(3):411-417. <https://doi.org/10.1002/cche.10042>.
- [156] E. Lew, D. D. Kuzmicky & D. D. Kasarda (1992). Characterization of low molecular weight glutenin subunits by reversed-phase high-performance liquid chromatography, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and N-terminal amino acid sequencing. *Cereal Chemistry*, 69(5):508-515.
- [157] S. Masci, R. D'Ovidio, D. Lafiandra & D. D. Kasarda (1998). Characterization of a low-molecular-weight glutenin subunit gene from bread wheat and the corresponding protein that represents a major subunit of the glutenin polymer. *Plant Physiology*, 118(4):1147-1158. <https://doi.org/10.1104/pp.118.4.1147>.
- [158] W. Maruyama-Funatsuki, K. Takata, H. Funatsuki, T. Tabiki, M. Ito, et al. (2005). Identification and characterization of a novel LMW-s glutenin gene of a Canadian Western Extra-Strong wheat. *Journal of Cereal Science*, 41(1):47-57. <https://doi.org/10.1016/j.jcs.2004.07.003>.

- [159] P. Ferrante, M. C. Gianibelli, O. Larroque, C. Volpi, R. D'Ovidio, et al. (2006). Effect of incorporation of an i-type low-molecular-weight glutenin subunit and a modified  $\gamma$ -gliadin in durum and in bread wheat doughs as measured by micro-mixographic analyses. *Journal of Cereal Science*, 44(2):194-202. <https://doi.org/10.1016/j.jcs.2006.06.004>.
- [160] J. Huang, S. Dai, Z. Song, L. Liu, L. Liang, et al. (2021). Characterization of novel LMW-i genes with nine cysteine residues from Chinese wheat landraces (*Triticum aestivum* L.) and analysis of their functional properties on dough mixing. *3 Biotech*, 11:499. <https://doi.org/10.1007/s13205-021-03044-8>.
- [161] S. Masci, D. Lafiandra, E. Porceddu, E. J. L. Lew, H. P. Tao, et al. (1993). D-Glutenin subunits: N-terminal sequences and evidence for the presence of cysteine<sup>1</sup>. *Cereal Chemistry*, 70(5):581-585.
- [162] Y. Wang, S. Zhen, N. Luo, C. Han, X. Lu, et al. (2016). Low molecular weight glutenin subunit gene *Glu-B3h* confers superior dough strength and breadmaking quality in wheat (*Triticum aestivum* L.). *Scientific Reports*, 6:27182. <https://doi.org/10.1038/srep27182>.
- [163] N. D. Rathan, A. Mahendru-Singh, V. Govindan & M. I. Ibba (2020). Impact of high and low-molecular-weight glutenins on the processing quality of a set of biofortified common wheat (*Triticum aestivum* L.) lines. *Frontiers in Sustainable Food Systems*, 4:583367. <https://doi.org/10.3389/fsufs.2020.583367>.
- [164] Y. Wang, Z.-r. Guo, Q. Chen, Y. Li, K. Zhao, et al. (2023). Effect of high-molecular-weight glutenin subunit Dy10 on wheat dough properties and end-use quality. *Journal of Integrative Agriculture*, 22(6):1609-1617. <https://doi.org/10.1016/j.jia.2022.08.041>.
- [165] X. Wang, Y. An, J. Chen, M. Wang, C. Wang, et al. (2024). Genetic diversity of HMW-GS and the correlation of grain quality traits in bread wheat (*Triticum aestivum* L.) in Hubei province, China. *Agronomy*, 14(6):1158. <https://doi.org/10.3390/agronomy14061158>.
- [166] V. Brönneke, G. Zimmermann & B. Killermann (2000). Effect of high molecular weight glutenins and D-zone gliadins on bread-making quality in German wheat varieties. *Cereal Research Communications*, 28:187-194. <https://doi.org/10.1007/bf03543592>.
- [167] H. Farahani, M. E. Moghaddam, A. Mohammadi, K. Z. Nezhad & F. Naghipour (2023). Effect of high molecular weight glutenin subunits on wheat quality properties, across a wide range of climates and environments in Iran. *Physiology and Molecular Biology of Plants* 29(6):889-902. <https://doi.org/10.1007/s12298-023-01324-6>.
- [168] Y. Karaduman, Z. Sirel Yeşildağ & A. Akın (2022). Evaluating selection efficacy of high molecular weight glutenin subunits (HMWGs) by relating gluten quality parameters. *LWT – Food Science and Technology*, 155:112949. <https://doi.org/10.1016/j.lwt.2021.112949>.
- [169] J. Wang, N. Liu, S. Yang, G. Qiu, H. Tian, et al. (2024). Research progress in the synthesis of stable isotopes of food flavour compounds. *Food Chemistry*, 435:137635. <https://doi.org/10.1016/j.foodchem.2023.137635>.
- [170] D. Horvat, G. Drezner, Z. Jurkovic, G. Šimić, D. Magdić, et al. (2006). The importance of high-molecular-weight glutenin subunits for wheat quality evaluation. *Poljoprivreda (Osijek)*, 12(1):53-57.

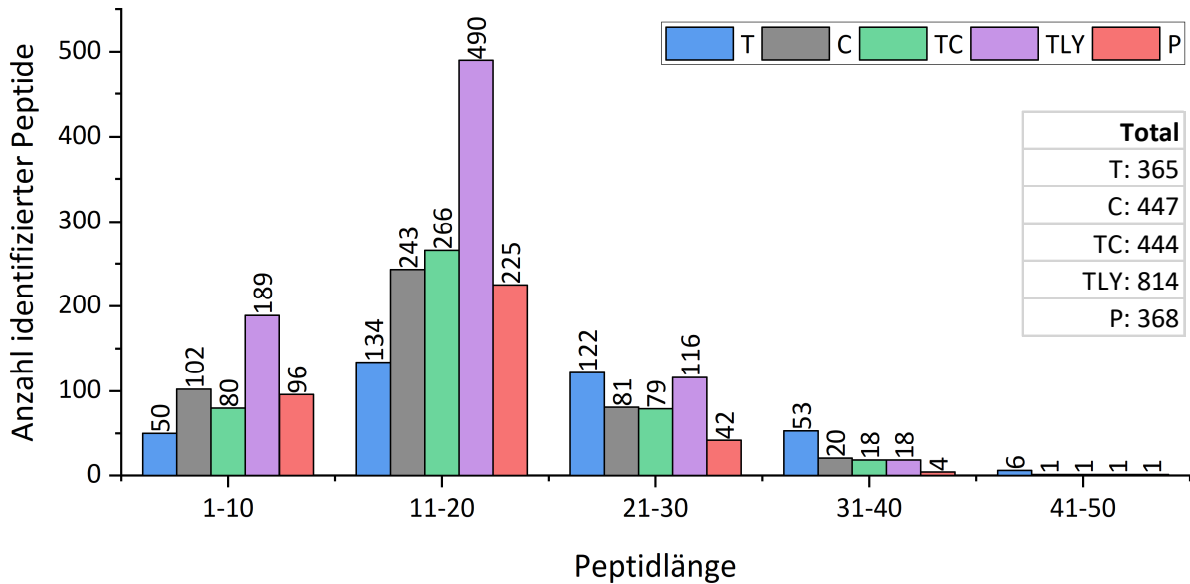
- [171] D. Liang, J. Tang, R. J. Peña, R. Singh, X. He, et al. (2009). Characterization of CIMMYT bread wheats for high- and low-molecular weight glutenin subunits and other quality-related genes with SDS-PAGE, RP-HPLC and molecular markers. *Euphytica*, 172(2):235-250. <https://doi.org/10.1007/s10681-009-0054-x>.
- [172] Y. Li, J. Fu, Q. Shen & D. Yang (2020). High-molecular-weight glutenin subunits: genetics, structures, and relation to end use qualities. *International Journal of Molecular Sciences*, 22(1):184. <https://doi.org/10.3390/ijms22010184>.
- [173] X. Wang, Y. Zhang, B. Zhang, C. Florides, Z. Gao, et al. (2018). Comparison of quality properties between high-molecular-weight glutenin subunits 5+10 and 2+12 near-isogenic lines under three common wheat genetic backgrounds. *Cereal Chemistry*, 95(4):575-583. <https://doi.org/10.1002/cche.10061>.
- [174] Y. Li, X. An, R. Yang, X. Guo, G. Yue, et al. (2015). Dissecting and enhancing the contributions of high-molecular-weight glutenin subunits to dough functionality and bread quality. *Molecular Plant*, 8(2):332-334. <https://doi.org/10.1016/j.molp.2014.10.002>.
- [175] Y. Peng, K. Yu, Y. Zhang, S. Islam, D. Sun, et al. (2015). Two novel Y-type high molecular weight glutenin genes in Chinese wheat landraces of the Yangtze-River region. *PLoS One*, 10(11):e0142348. <https://doi.org/10.1371/journal.pone.0142348>.
- [176] Z. Wang, L. Huang, B. Wu, J. Hu, Z. Jiang, et al. (2018). Characterization of an integrated active *Glu-1Ay* allele in common wheat from wild emmer and its potential role in flour improvement. *International Journal of Molecular Sciences*, 19(4):923. <https://doi.org/10.3390/ijms19040923>.
- [177] A. M. Waterhouse, J. B. Procter, D. M. Martin, M. Clamp & G. J. Barton (2009). Jalview Version 2-- a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9):1189-1191. <https://doi.org/10.1093/bioinformatics/btp033>.
- [178] H. F. Darlington, L. Tecsi, N. Harris, D. L. Griggs, I. C. Cantrell, et al. (2000). Starch granule associated proteins in barley and wheat. *Journal of Cereal Science*, 32(1):21-29. <https://doi.org/10.1006/jcrs.2000.0312>.
- [179] P. Greenwell & J. Schofield (1986). A starch granule protein associated with endosperm softness. *Cereal Chemistry*, 63(4):379-380.
- [180] A. C. Hogg, T. Sripo, B. Beecher, J. M. Martin & M. J. Giroux (2004). Wheat puroindolines interact to form friabilin and control wheat grain hardness. *Theoretical and Applied Genetics*, 108(6):1089-1097. <https://doi.org/10.1007/s00122-003-1518-3>.
- [181] N. Geneix, M. Dalgalarroondo, C. Tassy, I. Nadaud, P. Barret, et al. (2020). Relationships between puroindoline A-prolamin interactions and wheat grain hardness. *PLoS One*, 15(9):e0225293. <https://doi.org/10.1371/journal.pone.0225293>.
- [182] A. C. Eliasson (2003). Starch structure and bread quality. In S. Cauvain (Hrsg.), *Bread Making - Improving Quality*, Seiten 145-167. Woodhead Publishing, Cambridge, GBR. <https://doi.org/10.1533/9781855737129.1.145>.

- [183] M. Schopf & K. A. Scherf (2021). Water absorption capacity determines the functionality of vital gluten related to specific bread volume. *Foods*, 10(2):228. <https://doi.org/10.3390/foods10020228>.
- [184] D. Lafiandra & P. R. Shewry (2022). Wheat Glutenin polymers 2. The role of wheat glutenin subunits in polymer formation and dough quality. *Journal of Cereal Science*, 106:103487. <https://doi.org/10.1016/j.jcs.2022.103487>.
- [185] F. Laidig, H. P. Piepho, D. Rentel, T. Drobek, U. Meyer, et al. (2017). Breeding progress, environmental variation and correlation of winter wheat yield and quality traits in German official variety trials and on-farm during 1983-2014. *Theoretical and Applied Genetics*, 130:223-245. <https://doi.org/10.1007/s00122-016-2810-3>.
- [186] P. F. Qi, Y. M. Wei, T. Ouellet, Q. Chen, X. Tan, et al. (2009). The  $\gamma$ -gliadin multigene family in common wheat (*Triticum aestivum*) and its closely related species. *BMC Genomics*, 10:168. <https://doi.org/10.1186/1471-2164-10-168>.
- [187] P. Ferrante, C. Patacchini, S. Masci, R. D'Ovidio & D. Lafiandra (2007). LMW-i type subunits are expressed in wheat endosperm and belong to the glutenin fraction. In R. D'Ovidio, S. Masci & D. Lafiandra (Hrsg.), *The Gluten Proteins*, Seiten 136-139. Royal Society of Chemistry, London, GBR. <https://doi.org/10.1039/9781847552099-00136>.
- [188] N. Huo, T. Zhu, S. Altenbach, L. Dong, Y. Wang, et al. (2018). Dynamic evolution of  $\alpha$ -gliadin prolamin gene family in homeologous genomes of hexaploid wheat. *Scientific Reports*, 8:5181. <https://doi.org/10.1038/s41598-018-23570-5>.
- [189] D. W. Wang, D. Li, J. Wang, Y. Zhao, Z. Wang, et al. (2017). Genome-wide analysis of complex wheat gliadins, the dominant carriers of celiac disease epitopes. *Scientific Reports*, 7:44609. <https://doi.org/10.1038/srep44609>.
- [190] Y. Zhang, B. R. Fonslow, B. Shan, M. C. Baek & J. R. Yates (2013). Protein analysis by shotgun/bottom-up proteomics. *Chemical Reviews*, 113(4):2343-2394. <https://doi.org/10.1021/cr3003533>.
- [191] G. Hart-Smith, R. S. Reis, P. M. Waterhouse & M. R. Wilkins (2017). Improved quantitative plant proteomics via the combination of targeted and untargeted data acquisition. *Frontiers in Plant Science*, 8:1669. <https://doi.org/10.3389/fpls.2017.01669>.
- [192] V. Lullien-Pellerin (2024). How can we evaluate and predict wheat quality? *Journal of Cereal Science*, 119:104001. <https://doi.org/10.1016/j.jcs.2024.104001>.
- [193] R. van Eckert, E. Berghofer, P. J. Ciclitira, F. Chirido, S. Denery-Papini, et al. (2006). Towards a new gliadin reference material— isolation and characterisation. *Journal of Cereal Science*, 43(3):331-341. <https://doi.org/10.1016/j.jcs.2005.12.009>.
- [194] K. Schalk, C. Lang, H. Wieser, P. Koehler & K. A. Scherf (2017). Quantitation of the immunodominant 33-mer peptide from  $\alpha$ -gliadin in wheat flours by liquid chromatography tandem mass spectrometry. *Scientific Reports*, 7:45092. <https://doi.org/10.1038/srep45092>.

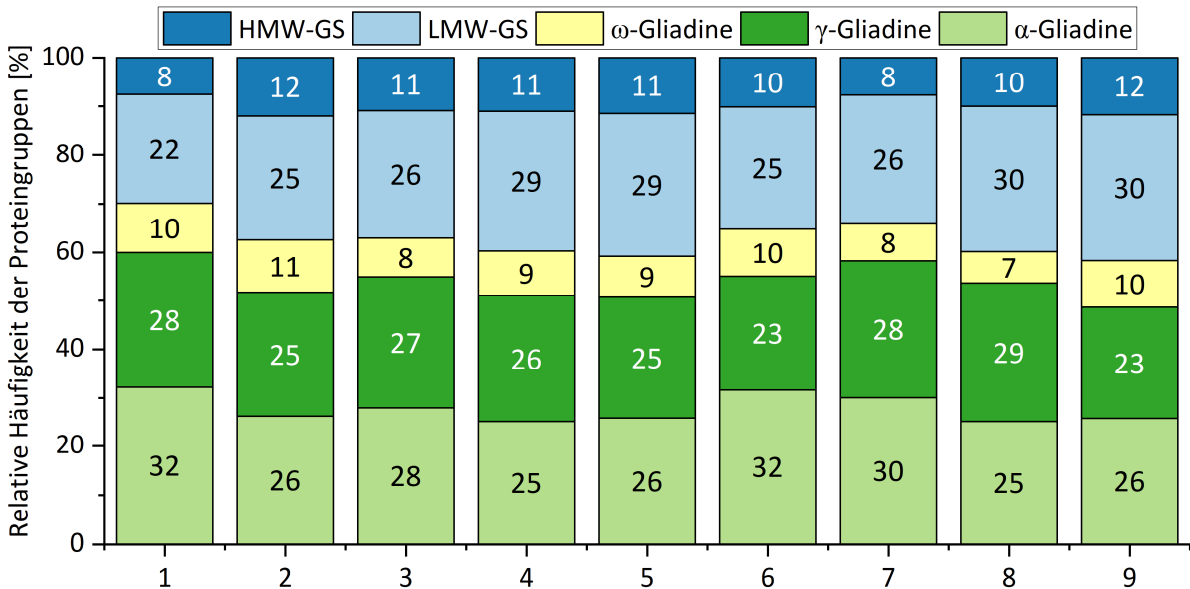
- [195] B. MacLean, D. M. Tomazela, N. Shulman, M. Chambers, G. L. Finney, et al. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, 26:966-968. <https://doi.org/10.1093/bioinformatics/btq054>.
- [196] AOAC International (2013). Appendix K: Guidelines for Dietary Supplements and Botanicals. In G. W. Latimer (Hrsg.), *Official Methods of Analysis of AOAC INTERNATIONAL*, Seiten 1-32. AOAC International, Gaithersburg, MD, USA.
- [197] B. Magnusson & U. Ornemark (2014). *Eurachem Guide: The fitness for purpose of analytical methods - a laboratory guide to method validation and related topics* 2. Edition, ISBN: 978-91-87461-59-0, abrufbar unter: <https://www.eurachem.org>.
- [198] S. Tyanova, T. Temu, P. Sinitcyn, A. Carlson, M. Y. Hein, et al. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, 13(9):731-740. <https://doi.org/10.1038/nmeth.3901>.

## 7 Anhang

### 7.1 Untargeted Proteomics von Weizen: Variation der Peptidasen



**Abbildung 32:** Verteilung der Glutenpeptidlängen in Abhängigkeit von den eingesetzten Enzymen. Die Tabelle auf der rechten Seite zeigt die Gesamtanzahl der identifizierten Glutenpeptide pro Peptidase. T: Trypsin; C: Chymotrypsin; TLY: Thermolysin; P: Pepsin.



**Abbildung 33:** Relative Häufigkeit der Glutenproteintypen in den untersuchten Weizensorten basierend auf den RP-UHPLC-UV-Ergebnissen der intakten Proteine. 1: Ambition; 2: FIRL3565; 3: Bussard; 4: Event; 5: Format; 6: Julius; 7: BAYP4535; 8: Potenzial; 9: RGT Reform. HMW-GS, hochmolekulare Gluteninuntereinheit; LMW-GS, niedermolekulare Gluteninuntereinheit.

**Tabelle 12:** Liste der zehn am häufigsten vorkommenden Proteintypen im Weizenmehl der Sorte RGT Reform nach Verdau mit verschiedenen Peptidasen. Unter der Haupt-Protein-ID sind maximal drei Protein-IDs aufgeführt; der relative Anteil ist in Klammern hinter dem jeweiligen Proteintyp angegeben.

Trypsin		Chymotrypsin		Trypsin + Chymotrypsin		Thermolysin		Pepsin	
Major Protein ID	Protein-typ	Major Protein ID	Protein-typ	Major Protein ID	Protein-typ	Major Protein ID	Protein-typ	Major Protein ID	Protein-typ
B6UKN7	γ-Gliadin (7.8%)	A9YSK5;Q42451;M4M8L5	HMW-GS (14.4%)	A0A7D5DPJ7	ω-Gliadin (10.1%)	A0A7D5DPJ7	ω-Gliadin (11.1%)	B6UKN7	LMW-GS (24.1%)
A0A2U8JCZ2; B2Y2Q1;R4JBL6	LMW-GS (4.8%)	A0A3B6GZ31; A0A077RQJ1	Calcyclin-binding protein (13.8%)	A0A2U8JCZ2; B2Y2Q1;R4JBL6	LMW-GS (6.2%)	A9YSK5;Q42451;M4M8L5	HMW-GS (8.0%)	A0A2U8JD23;M9TK56	ω-Gliadin (15.1%)
Q5MD68;P01084Q4U1A4	ATI (3.4%)	R4JFL5;R4JBH8;R4JB19	LMW-GS (5.4%)	A9YSK5;Q42451;M4M8L5	HMW-GS (6.0%)	R9XV26;A0A1K0J4W8	α-Gliadin (7.6%)	Q5MD68;P01084;Q4U1A4	HMW-GS (11.6%)
A0A3B6IJ76	Cupin (3.4%)	A0A290XZ34	γ-Gliadin (4.7%)	Q3S4V7	α-Gliadin (5.7%)	X2KYP9;A0A7D5RE50;A0A7D5R7Y7	ATI (7.3%)	A0A2U8JCZ2;B2Y2Q1;R4JBL6	γ-Gliadin (9.4%)
A0A1S6KXP9	ATI (2.7%)	P21292;R9XWA1;K7XEF1	γ-Gliadin (4.4%)	R4JFL5;R4JBH8;R4JB19	LMW-GS (5.4%)	A0A290XZ34	γ-Gliadin (6.4%)	A0A3B6IJ76	Pyruvat kinase (4.0%)
A0A060N0C7; V9P6D7;Q7X9J9	LMW-GS (2.6%)	V9P6N2;Q5MFP8	LMW-GS (4.2%)	A0A290XZ34	γ-Gliadin (3.7%)	W5CQ97;A0A3B6GQ71;Q1XHC6	Cysteine proteinase (4.9%)	A0A1S6KXP9	γ-Gliadin (3.6%)
A0A2U8JD23; M9TK56;Q30DX3	γ-Gliadin (2.5%)	Q3S4V7	α-Gliadin (4.0%)	R9XV26;A0A1K0J4W8	α-Gliadin (2.9%)	A0A3B6I5M0	Bi-functional inhibitor (4.2%)	P21292;R9XWA1K7XEF1	α-Gliadin (2.8%)
R9XSX7	α-Gliadin (2.0)	R9XV26;A0A1K0J4W8	α-Gliadin (3.8%)	A0A3B6T7X8;Q7X9M4	Bifunctional Inhibitor (2.9%)	Q5UHH8;Q5UHH6;Q4U1A1	ATI (3.3%)	A0A060N0C7;V9P6D7;Q7X9J9	α-Gliadin (2.7%)
P18573;J7HWD3;Q1WA39	α-Gliadin (1.9)	A0A2U8JD23;M9TK56	ω-Gliadin (3.7%)	Q41543;Q94G92;Q9M4L5	γ-Gliadin (2.5%)	A0A0K2QJX0	α-Gliadin (3.2%)	R9XSX7	Grain softness protein (2.5%)
A4ZIZ6;A4ZIW9;A4ZIX1	ATI (1.9%)	A0A7D5DPJ7	ω-Gliadin (3.4%)	A0A060MZP1	HMW-GS (2.3%)	A0A3B6T7X8;Q7X9M4	Bifunctional inhibitor (2.9%)	P18573;J7HWD3;Q1WA39	Serpin domain protein (2.0%)

ATI, Amylase/Trypsin-Inhibitor; HMW-GS, hochmolekulare Gluteninuntereinheit; LMW-GS, niedermolekulare Gluteninuntereinheit.

## 7.2 Targeted Proteomics von Weizen

**Tabelle 13:** Liste der mittels nanoLC-MS/MS im Parallel Reaction Monitoring analysierten Gliadinpeptide.

Nr.	$\alpha$ -Gliadine	$\gamma$ -Gliadine
1	CQAIHNVVHAILHQQQQQQQQQPLSQVFSFQQPQQQ	QGVQILVPLSQQQQVGGTTLVQGGGIIQPQQPAQLEVIR
2	DVIVLQQHNIHESQVLLQSSYQVLQQLCCQQLR	GFGQPQPQQLGQEMPMQPQHQLGQHSILPQQLAQYK
3	CLAINVAHAILHEQQQHQQQQQYSWGVTGFK	QPQQPFQPQQPQQSFPQQPSLIQQSLQQQLNPCK
4	ELCCQHLWQIPEQSQCAIHNVVHAILHQQQK	QQCCQQLAQIPQQLQCAAHSIVHSIIMQQEQQEQR
5	DVVLQQHNIHASSQVLLQSSYQLLQQLCCQR	QPFPPQQPQPYPQQPQQPFPQTQQPQQPFPQSK
6	NLALETLPAMCNVYIPPYCSTTIAPFGIFGTN	TLPTMCNVYVPPDCSTINVPYANIDAGIGGQ
7	NLALQTLPAMCNVYIPPHCSTTIAPFGIFGTN	TLPTMCNVYVSPDCSTINAPFANIVVIGGGQ
8	NLALQTLPAMCNVYIPPHCSTTIAPFGISGTN	TLPTMCSVNVPLYSATTSVFPFVGTGVGAY
9	NLALQTLPAMCNVYIPPYCSTTIAPFGIFGTN	QLQCAAHNVVHSIIMQQEQQQGIIQLR
10	NLALQTLPAMCNVYIPPYCSTTIAPFGIMSTN	QLQCAAHSVVHSIIMQQEQQQGIIQLR
11	NLALQTLPAMCNVYIPPYCSTTIAPSGIFGTN	QLQCAAHSVVHSIVMQQEQQQGIIQLR
12	NLALQTLPAVCNVYIPPYCSTTIAPFGIFGTN	QPQQPFYQQPQQTFPQPQQAFPHQPK
13	NLALETLPAMCNVYIPPYCTIAPFGIFGTN	PLFQLAQGLGIIQPQQPAQLEGIR
14	NLALETLPAMCNVYIPPYCTIAPVGGIFGTN	PLFQLVQGGGIIQPQQPAQLEVIR
15	NLALETLPAMCNVYIPPYCTIAPVGFIFGTN	PLFQLVQGGGIIQPQQPAQYEVIR
16	NLALETLPAMCNVYIPPYCTIAQVGFIFGTN	NFLQQCNHVSLSVSSLVSIILPR
17	NLALETLPAMCNVYIPPYCTVAPVGFIFGTN	NFLQQCNPMSLVSSLVSIILPR
18	NLALQMLPAMCNVYIPPYCTIAPFGIFGTN	NFLQQCNPVSLVSSLISMILPR
19	NLALQTLPAICNVYIPPYCTIAPFGIFGTN	NFLQQCNPVSLVSSLVSIILPR
20	NLALQTLPAMCNVYIPLYCTIAPFGIFGTN	NFLQQCNVSVLSSIVSLILPR
21	NLALQTLPAMCNVYIPPYCTIAPFGIFGTN	NYLLQQCDPVSLVSSLVSMILPR
22	NLALQTLPAMCNVYIPPYCTIAPFGIFGTN	NYLLQQCNPVSLVSSLVSMILPR
23	NLALQTLPAMCNVYIPPYCTIVPFGIFGTN	NYLLQQCNPMVSSLVSMILPR
24	NLALQTLPAMCNVYIPPYCTMAPFGIFGTN	NYLLQQCSPVSLVSSLVSMILPR
25	NLALQTLPSMCNVYIPPYCTIAPFGIFGTN	NYPLQQCNPVSLVSSLVSMILPR
26	QILQQILQQQLIPCMDVVLQQHNIHAGR	PLFQIVQGQSIIQQPQAQLEVIR
27	QTLPAMCNVYIPPYCSTTIAPSGIFGTN	SLVLGTLPTMCNVFVPPPECSTTK
28	SAWEPQHPSPEHQPTPQPEHPVPHQK	SLVLQTLPSMCNVYVPECAIMR
29	SQVLQQSTYQLLQELCCQQLWQIPEQSR	SLVLQTLPSMCNVYVPECSIMR
30	SQVLQQSTYQLVQQLCCQQLWQIPEQSR	SLVLQTLPTMCNVYVPECSIIR
31	SQVLQQSTYQLLQELCCQQLWQIPEQSR	SLVLQTLPTMCNVYVPPYCSTIR
32	CQAIHNVVHAILHQQQQQQQQQQQK	SSVLQTLATMCNVYVPPYCSTIR
33	CQAIHNVVHAILHQQQQQQQQQQQK	PASLVSSLWSIIPQSDCQVMR
34	PQQQPSCQVSFQQPLQYPLGQGSFR	PVSLVSSLWSIILPPSDCQVMR
35	QQQPSSQVSFQQPLQYPLGQGSFR	AAIHSIVHSIIMQQEQQEQR
36	QQQPSSQVSFQQPQQYPLGQGSFR	CPAIIHNVHAIVMQQQHVDR
37	SQVLQQSTYQLLQELCCQHLWQIPEK	LCCQAIHNVVHAILHQQQK
38	QQILQQQLIPCMDVVLQQHNIHAGR	GDCQVMQQCCQQLAQIPR
39	VSQSYQLLQQLCCLQLWQTPEQSR	NDCQVMQQCCQQLAQIPR
40	VSQSYQLLQQLCCLQLWQTPEQSR	NDSQVMQQCCQQLAQIPR
41	CQAIHNVVAHAIIMHQQQQQQQQEQR	SDCQVMQQCCQQLAQIPR
42	PSQQNPQAQGSVQPQQLPQFEIR	SDCQVMQQQYCCQQLAQIPR
43	MCNVYIPPYCSTTIAPFGIFGTN	NPVSLVSSLVSMILPR
44	QQQQQQQQQLLQQLLQQLLIPCR	PFIQPSLQQQLNPCK
45	QQQQQQQQQLLQQLLQQLLIPCR	PVSLVSSIVSLILPR
46	SQVVQHSSCLVMWEQCCQQLK	APFASIVADIGGQ
47	CQAIHNVVHAILHQQQQQR	APFASIVAGIGGQ
48	CQAIQNVVHAILHQQQQQR	APFASIVASIGGQ
49	NVYIPPYCTIAPVGFIFGTN	APFASIVTGIGGQ
50	CQAIHNVVHAILHQQQR	APFSSVAGIGGQ
51	VPVPQLQPQNPSQQQPQK	TPFASTIAGIGGQ
52	DALLQQCSPVADMSFLR	
53	QAIHNVVHAILHQQQK	

54	CQAIHNVIAIILHHR	TPFASTVAGIGGQ
55	QQILQQILQQQLPCR	TPFASTVDGIGGQ
56	DVLLMQCNILVPSFK	EFLQQCNPEEK
57	DVVLQQPNIAHASSK	PFIQPSLQQR
58	GSVQPQQLPQFEEIR	QLAHISEPSR
59	QILQQILQQQLPCR	CQQLAQIPR
60	SDQPQQSFPPQPPQK	QQCCQQLAR
61	MDVVLQQHNIVHGR	QQLNLCK
62	TVQSFFEQLISCR	
63	DVVLQQHNIAHAR	
64	DVVLQQHNIAHGR	
65	SQILQENVCVVMR	
66	SQVLQQSTYQLLR	
67	CQQLWQIPEQSR	
68	VVHAILHQQK	
69	LWQIPEQSR	
70	NLALQTLPR	
71	ECCEQFR	
72	LIPEQSR	

**Tabelle 14:** Liste der mittels nanoLC-MS/MS im Parallel Reaction Monitoring-Modus analysierten Gluteninpeptide.

Nr.	Niedermolekulare Gluteninuntereinheiten	Hochmolekulare Gluteninuntereinheiten
1	AIISILVQEQQHGQGLNQPQQQQPQQSVQGVSQPQQQQK	QQSGQGQQLGQGQQSGQGQEQQPQGQQAYYPASSQQSGQR
2	AIISIVLQEQQHGQGFNQPHQQQPQQSVQGVSQPQQQQK	QQSGQGQQPQGQGGQGGQGGQGGQGGQGGQAYYPTSSQQSR
3	AIISIVLQEQQHGQGFNQPQQQQPQQSVQGVSQPQQQQK	GGSFYPGETTPLQQLQQGIFWGTSSQTVQGYYPGVTSR
4	ETFPQQPSSQQQPPFPQQPFLQQQPSFSQQPLFSQK	GGSFYPGETTPLQQLQQGIFWGTSSQTVQGYYPVTSR
5	TLPMMCVSNVPLYSSSTTSVPFGVGTGVGAY	QQPVQGGQPEQGQPGQWQQGYPTSPQQLGQGQQPR
6	TLPMMCVSNVPLYSSSTTSVPFGVGTGVGAY	QGYPTSLQQPGQGGQIGQGQQGYPTSPQHTGQR
7	TLPTMCGVNVPLYSSSTIMPFSIGTGVGGY	QQGYPTSPQPGQGGQLGQGQPGYPTSPQPGQK
8	TLPTMCGVNVPLYSSSTIMPFSIGTGVGGY	QQPGQGGQPEQGQPGQGGQGGYPTSPQPGQK
9	TLPTMCGVNVPLYSSSTIMPFSIGTGVGGY	GQQGYPTSLQQPGQGGQGYPTSLQHTGQR
10	TLPTMCGVNVPLYSSSAPFGVGTGVGAY	AGSFYPSETTSPQQLQGMIFWGIPALLR
11	TLPTMCGVNVPLYSSSAPFGVGTGVGAY	GGSFYPGETTTPQQLQGSILWGIPALLR
12	PQMLQQSSCHVMQQQCCQQLPQIPQQSR	DISPECHPVVSPVAGQYEQQIVVPPK
13	SQMLQQSSCHVMQQQCCQQLPQIPQQSR	WQEPGQGGQWYPTSLQQPGQGGQIGK
14	SQMLQQSSCHVMQQQCCQQLPQIPQQSR	YYPTSSQPGQLQQLAQGGQGGQPER
15	SQMLQQSSCHVMQQQCCQQLPQIPQQSR	GQQGQGGQGGQQLGQGQGGQGGQPGQK
16	SQMLQQSSCHVMQQQCCQQLPQIPQQSR	QLQQPEQGQGGQPEQGQGGQGGQQR
17	SQMLQQSSCHVMQQQCCQQLPQIPQQSR	QPGYPTSPWQPEQLQPTQGGQR
18	SQMLQQSSCHVMQQQCCQQLPQIPQQSR	QGSYYPGQASPPQPGQGGQPGK
19	SQMLQQSSCHVMQQQCCQQLPQIPQQSR	QGYDSPYHVSAEQQAASPMVAK
20	SQMLQQSSCHVMQQQCCQQLPQIPQQSR	QQPGQGGQIGQGQQLGQGR
21	SQMLWQSSCHVMQQQCCQQLPQIPEQSR	GGSFYPGETTTPQQLQQR
22	SQMWPQSSCHVMQQQCCQQLPQIPEQSR	QGQQLGQGGQPGQGGQTR
23	SQMWPQSSCHVMQQQCCQQLPQIPEQSR	QPGYPTSPQQLGQGQPR
24	SQMWPQSSCHVMQQQCCQQLPQIPEQSR	DVSPECQPVGGGPVAR
25	SQMWPQSSCHVMQQQCCQQLPQIPEQSR	YYPTSPQPGQEQQPR
26	SQMWPQSSCHVMQQQCCQQLPQIPEQSR	ELEACQVVDQQLR
27	TLPTMCGVNVPLYSSSAPLGVGSR	AQQLAAQLPAMCR
28	TLPTMCGVNVPLYSSSAPLGVGSR	AQQPATQLPTVCR
29	TLPTMCGVNVPLYSSSAPLGVGSR	AQQPVAQLPAMCR
30	SQMLQQSSCHVMQQQCCQQLPQR	VQQPATQLPIMCR
31	SQMLWQSSCHVMQQQCCQQLPQR	EGEASEQLQCR
32	SQMWPQSSCHVMQQQCCQQLPQR	ACQQVMDQQLR

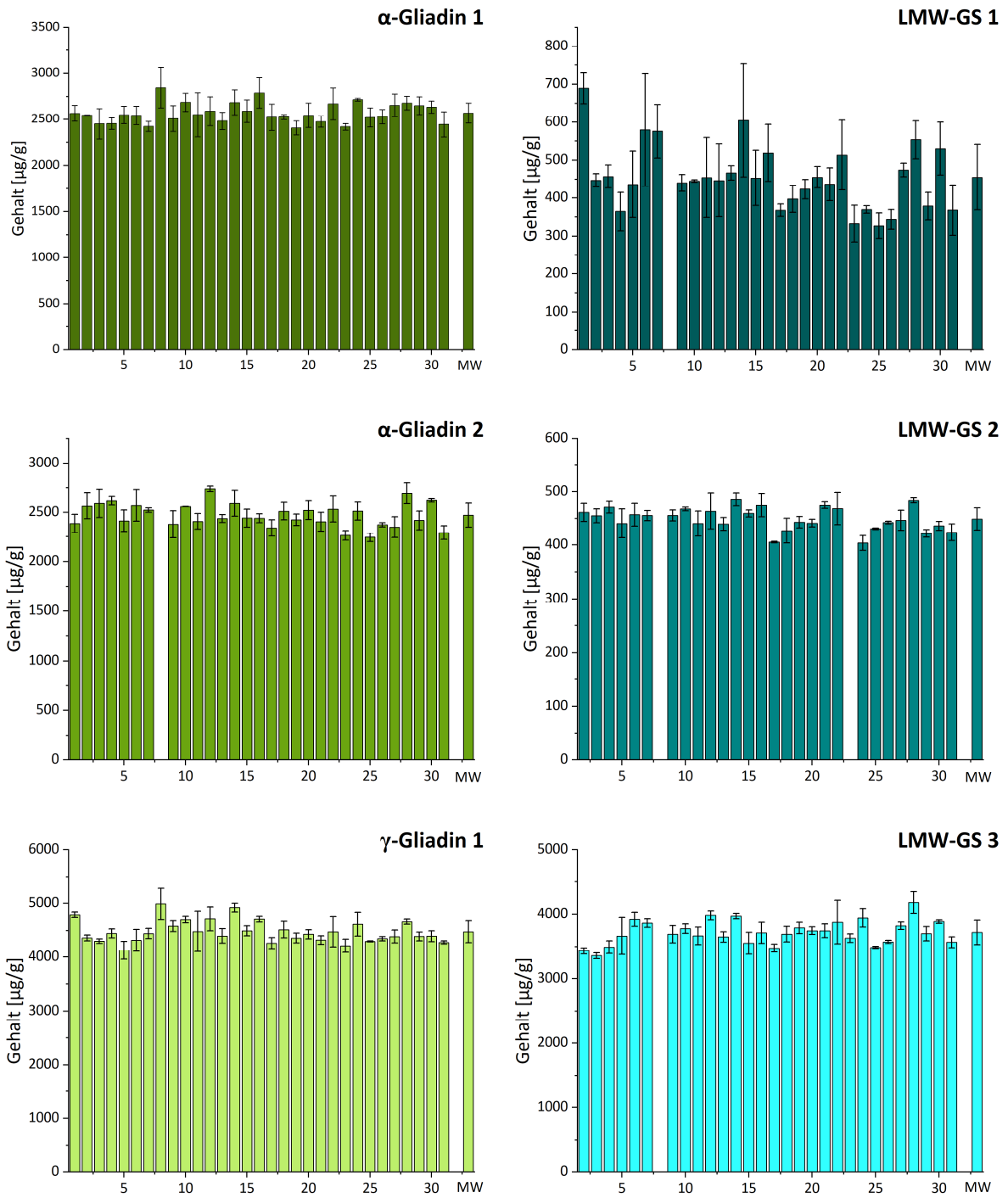
ANHANG

---

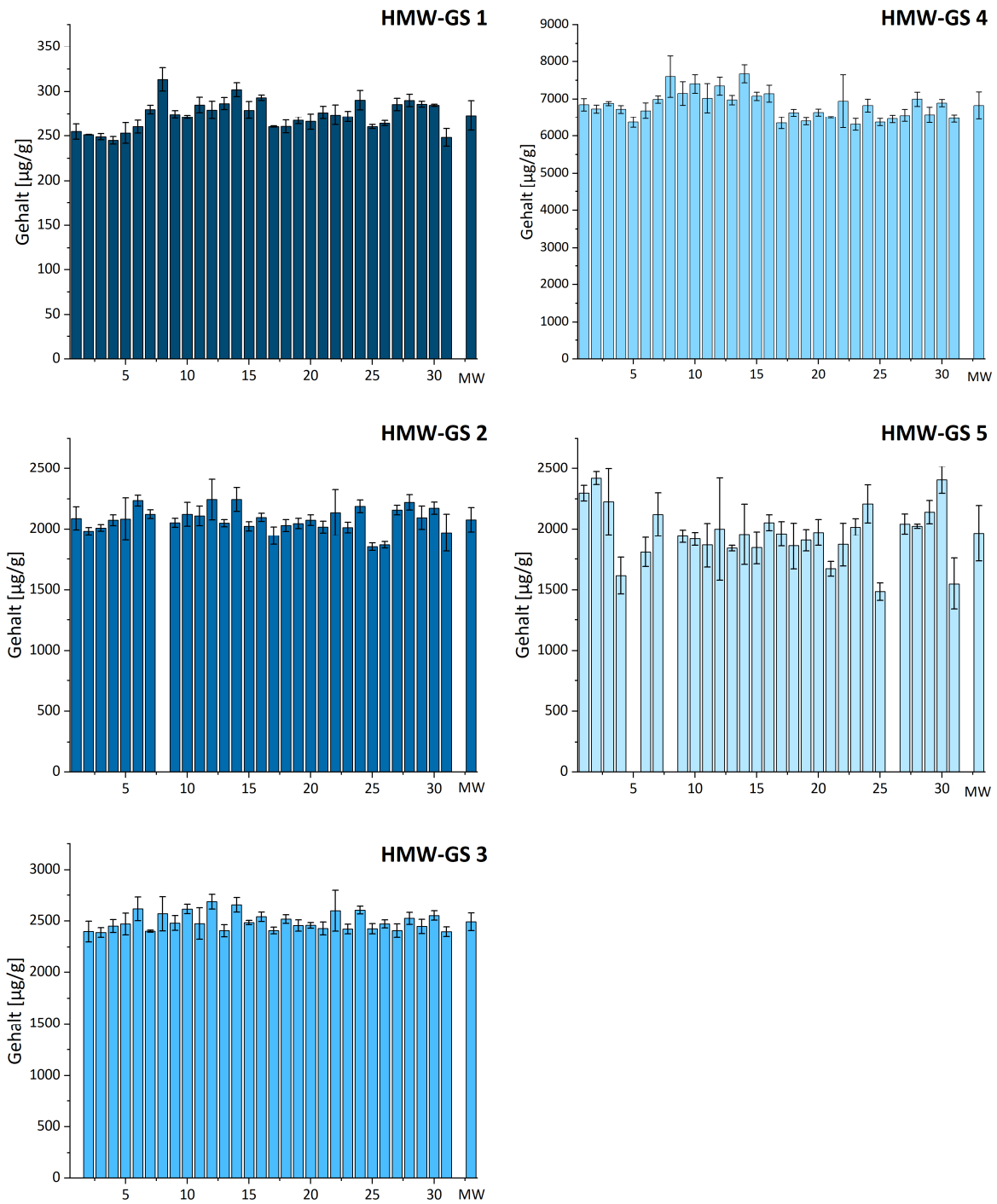
33	SQMWWQSSCHVVQQCCQQLPR	ELQESSLEACR
34	SQMWWQSSCNVMQQCCQQLPR	LEGGDALLASQ
35	AIVYSIILQQQQQQQQQQQR	LEGGDALSASQ
36	GIFLQPHQISQLEVMTSIALR	MEGGDALSASQ
37	PQMLQQSICHVMQQCCQQLR	QYEQQPVVPSK
38	SCHVMQQCCQQLPQIPQQSR	QYEQQVVVPPK
39	SQMLQQSICHVMQQCCQQLR	IFWGIPALLK
40	QQQIPVIHPSVLQQLNPCK	LEGSDALSAR
41	LFLQQQCSPVAMPQSLAR	LEGSDALSTR
42	SQTLWQSSCHVMQQCCCR	LPWSTGLQMR
43	VFLQQQCNPVAMPQSLAR	LVVDQQLAGR
44	VFLQQQCSPVAIPQSLAR	QQPGQGGQQLR
45	VFLQQQCSPVAMPQGLAR	QVVDQQLAGR
46	VFLQQQCSPVAMPQHLLAR	QYEQTVVPPK
47	VFLQQQCSPVAMPQSLAR	YYLSVTSPQR
48	SQMLWQSSCHVMQQCCQ	LVAVSQVVR
49	TTTNVPGVGTGVGSY	PITVSPGTR
50	TTTSVPFDVGTGVGAY	SVAVSQVAR
51	TTTSVPGVAGVAGAY	ELQEHSK
52	TTTSVPGVGTGVGAY	ELQELQER
53	TTTSVPGVGTGVGSY	QVVDQQLR
54	TTTSVPGVSAGVAGAY	
55	AIISIVLQEQQQVR	
56	IFLQQCNHVAMSQR	
57	LFLQQCSPVAMPQR	
58	LFPQQCSPVAMPQR	
59	VFLQQCSHVAMSQR	
60	VFLQQCSPVAMPQR	
61	VFLQQCSPVMPQR	
62	ICHVMQQCCQQLR	
63	ILPTMCSNVNPLYR	
64	TLPTMCNVNPLYR	
65	VFLQQCIPVAMQR	
66	TTTSVPGVGTG	
67	GTFLQPHQIAR	
68	MDTSCIPGLER	
69	PSVLQQLNPCK	
70	LEVMTSIALR	
71	QLPQIPEQSR	
72	SHIPGLER	
73	SVNPLYR	
74	TLPMCR	
75	TLPTMCR	
76	VNPLYR	

**Tabelle 15:** Markerpeptide und zugehörige Protein-IDs der UniProtKB-Datenbank.

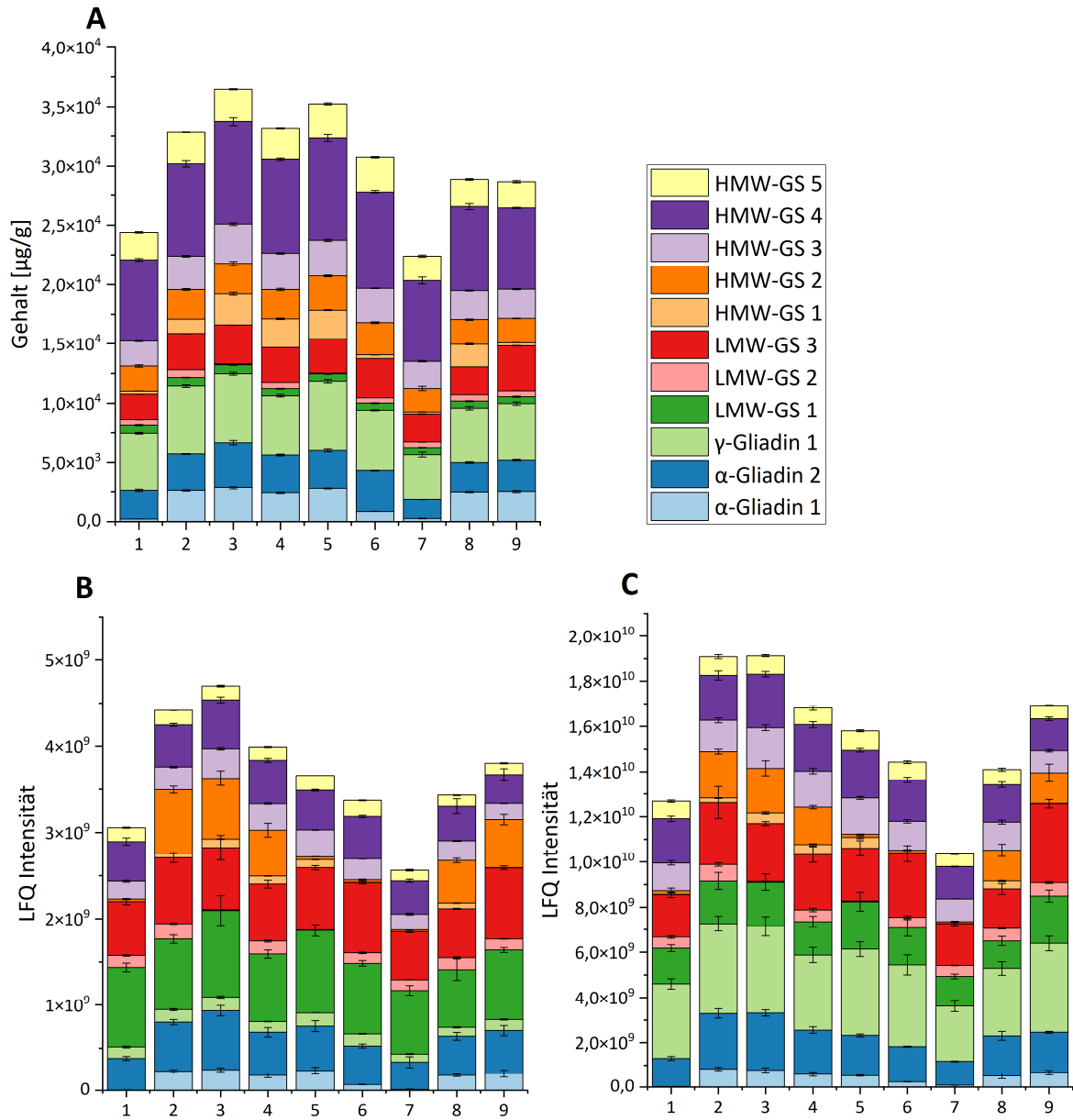
Nr.	UniProtKB ID des führenden (Razor-) Proteins	Protein- anzahl	Anzahl an Protein- gruppen	Protein IDs
P1	R9XW75	3	1	R9XW75;A0A1K0J4W5;J7HT14
P2	R9XUM8	56	6	R9XSX7;A7LHB1;A0A3B6UAZ5;R9XV34;A7LHB6;K7X0I0;K7X1K5;K7WV53;K7X1F5;K7XRC1;K7XRA8;K7XE59;A0A9R0G5C7;J7I021;R9XV48;D0UTY3;K7XR81;A0A0E3UR68;A0A0E3UQU4;X2KS61;R9XW94;R9XW80;K7XR49;K7WUZ4;R9XW73;K7WV57;R9XSW3;K7X0I8;K7XR61;X2KVI4;A0A0E3Z527;K7XR67;A0A1K0JS80;A0A2R2Y407;K7X1G2;K7XR55;K7X1F1;K7WV06;A0A1K0ITH3;A5JSB5;A0A2R2Y406;K7X1J5;R9XW75;A0A1K0J4W5;J7HT14;A5JSA6;A0A9R0GKV9;A0A1K0JNG6;A0A0K2QJ9;K7XRC7;A0A1K0IT13;A0A1K0JSA4;R9XUM8;A0A2P1H695;A0A0E3Z7F7;R9XW87
P3	H9BFB6	34	4	R9XUB9;K7X1S8;Q9M6P7;Q94G91;R9XUB0;U5UA54;R9XW98;P08453;Q6EEW9;R9XT02;R9XUT4;R9XV87;B6DQB3;F4YT72;H9BFB6;DOES83;A0A9R1DJ49;R9XWD0;A0A4D6Q653;A0A4D6Q658;A0A4D6Q7Q1;A0A5B9Y2N1;A0A5B9Y3R8;R9XU95;A0A4D6Q614;R9XSY7;F2X323;F2XAR4;M9TGF7;B6UKN8;B6DQB6;B6DQB5;B6DQB8;B6UKN0
P4	A0A0S2GJQ0	1	1	A0A0S2GJQ0
P5	P93790	13	1	P93790;Q5MFN7;Q670Q4;Q8W3V4;D3U321;K9K7W7;D0EVP3;A0A7D4XMB8;A0A0S2GJR9;A0A2R4Q9V2;A0A0S2GJR2;A0A0S2GJS4;Q0PW10
P6	R4JQA1	72	6	A0A0S2GJQ1;R4JF25;Q19MN3;R9XUD0;Q5MFH2;Q5MFH0;M1GL80;B3EY91;Q19N55;Q19MN2;M1GP58;Q6SPZ1;R9XVB6;R9XT45;A0A3G3C4G6;R9XT25;Q6SPZ3;A0A1U9XLB3;R4JBI2;D3UAL5;R4JAR7;A0A1U9XLA9;P93794;Q6SPZ2;B8R6A0;A0A3G3C4Q1;R9XWEO;R9XUF6;R9XV98;R9XWD5;Q0PW11;Q8W3U9;D2DII7;M1GMA2;R4V1P5;A0A0S2GJQ4;R4JFL5;Q6SPZ0;X2JAE7;R4JBE8;R4JB50;R4JD42;R9XU00;R4JB27;R4JB18;R9XUF2;R9XWV9;X2J8E3;R4JB19;R4JBH8;Q8W3V0;C3VN76;C3VN75;R9XV91;D3U317;Q8W3V1;R4JB34;A0A1U9XL81;D2DI11;C3VN77;C3VN74;Q6QGWO;Q5MFH3;Q5MFG8;Q6SPY7;A0A1U9XLC8;R4JQA1;P10385;D2DII3;Q6SPY8;F8SGQ4;A0A1U9XLB4
P7	Q0Q5D3	8	1	Q0Q5D3;P10387;A9ZMG8;A9YSK3;A0A165B8S1;A0A1W6C2J9;V9TRL3;A0A2L1K3K4
P8	A0A060MZP1	1	1	A0A060MZP1
P9	Q6RX92	23	2	Q0Q5D3;P10387;A9ZMG8;A9YSK3;A0A165B8S1;A0A1W6C2J9;V9TRL3;Q8HOL0;Q6RX92;Q52JL3;P08488;Q6RX93;E4W506;A4URY8;A0A0H4SR12;A9QUS3;D2CPI7;A0A1W6C2K2;A0A0H4SMF2;A0A1G4P1V2;Q670Q5;A0A1G4P1T5;Q8HOK9
P10	Q6RX92	51	9	A5HMG2;Q9SDM2;Q94IJ8;Q9S8V9;Q03871;W6AX70;U5YQV1;Q52JL2;Q84UY5;B8PSA6;A0A4Y6GR70;Q4JHY1;W0C8U3;A0A060AER5;Q0Q5D8;Q94IJ6;A0A9R1DFL1;Q0Q5D3;P10387;A9ZMG8;A9YSK3;A0A165B8S1;A0A1W6C2J9;V9TRL3;A0A2L1K3K4;B8YPU3;Q8HOL3;Q6RX92;Q52JL3;P08488;Q6RX93;E4W506;A4URY8;A0A0H4SR12;D2CPI7;A0A1W6C2K2;A0A0H4SMF2;A0A1G4P1V2;Q670Q5;A0A9R1IXL7;A0A1G4P1T5;Q8HOL6;W6AWK6;W6AW92;A0A023GQG9;X5CHT6;Q7XZA7;A0A1G4P1T3;A0A1G4P1S9;Q7XZB4;Q7X8H1
P11	B1B520	25	1	B1B520;Q599I0;G3FLC7;J7G6L4;P10388;Q0Q5D2;C6L669;A9YSK4;E2CT66;X2JUA0;D7REK2;D0IQ07;A0A0H4PMT1;Q8GV12;P08489;Q6R2V1;A0A0K0KCV6;C0SUC3;D0IQ05;Q38LF5;Q6R507;Q6R508;Q6R509;G1E6K7



**Abbildung 34:** Mittelwerte und Standardabweichungen der jeweiligen Peptide basierend auf 31 RGT Reform-Messungen. LMW-GS, niedermolekulare Gluteninuntereinheit; MW, Mittelwert.



**Abbildung 35:** Mittelwerte und Standardabweichungen der jeweiligen Peptide basierend auf 31 RGT Reform-Messungen. HMW-GS, hochmolekulare Gluteninuntereinheit; MW, Mittelwert.



**Abbildung 36:** Vergleich der Verteilung der Glutenproteintypen in neun Weizensorten, die mittels SIDA (A), UHPLC-MS/MS (B) und nanoLC-MS/MS (C) analysiert wurden. Berücksichtigt wurde nicht nur die Proteingruppe mit dem führenden Razor-Protein, sondern es wurden die LRFQ-Intensitäten aller Proteingruppen summiert, die das jeweilige Peptid enthielten. 1: Ambition; 2: FIRL3565; 3: Bussard; 4: Event; 5: Format; 6: Julius; 7: BAYP4535; 8: Potenzial; 9: RGT Reform; UHPLC, Ultrahochleistungsflüssigkeitschromatographie; LC-MS/MS, Flüssigkeitschromatographie-Tandemmassenspektrometrie; LRFQ, labelfreie Quantifizierung; SIDA, Stabilisotopenverdünnungsanalyse.

### 7.3 Weizenproteomics in Verbindung mit Backqualität

**Tabelle 16:** Übersicht der in der untargeted Analyse der BMWpop 2018 identifizierten  $\alpha$ -Gliadine.

Name	UniProtKB ID
$\alpha$ -Gliadin 1	K7X0N8;Q3S4V7;Q41531;A0A1K0IT72;P02863;A0A1K0J4U0;A0A1K0JS64;B2ZRD0;R9XUM3;Q41509;K7XEB5;B2ZRD3;Q53WW7;A0A060N3B9;A0A1K0ISW8;A5JSA5;A0A1K0J4S9;A0A1K0ITF2;A0A1K0ISY1;Q9M4M1;Q9M4M2;Q9M4M5;Q9M4L7;P04721
$\alpha$ -Gliadin 2	A5JSA6;A0A9R0GKV9;A0A1K0JNG6;A0A0K2QJJ9;K7XRC7;A0A1K0IT13;A0A1K0JSA4;A0A455QSW4
$\alpha$ -Gliadin 3	A0A2R2Y404;I0IT62;A0A0E3Z6L3;A0A0K2QJT3
$\alpha$ -Gliadin 4	A0A9R1I5A8;I0IT63;R9XUP2
$\alpha$ -Gliadin 5	Q41545;P04725
$\alpha$ -Gliadin 6	P04726;F4YT75;A0A1W6C2K1;A0A023WGF7
$\alpha$ -Gliadin 7	A0A023WHE7;A0A455QTH1
$\alpha$ -Gliadin 8	A0A455QTG1;A0A9R1IFK0;A0A0K2QJY6;A0A0K2QJ82;A0A0K2QJ8

**Tabelle 17:** Übersicht der in der untargeted Analyse der BMWpop 2018 identifizierten  $\gamma$ -Gliadine.

Name	Uniprot ID
$\gamma$ -Gliadin 1	Q41543;D0ES84
$\gamma$ -Gliadin 2	R9XUS2
$\gamma$ -Gliadin 3	D0ES80
$\gamma$ -Gliadin 4	A0A290XZ11;A0A290XYW2;A0A290XYV6
$\gamma$ -Gliadin 5	R9XV62
$\gamma$ -Gliadin 6	B6UKM7;A0A4D6JZ15;Q94G93;M9TE79;B6UKN5;B6UKM5;A0A4D6K3Q4;A0A9R1INI1;Q94G95;A0A4D6K2W9;A0A290XZ23;A0A290XZ17;B6UKP8;B6UKP7;B6UKN3;A0A4D6JZB7;Q94G90;A0A290XZ14
$\gamma$ -Gliadin 7	H9BFB6;D0ES83;A0A9R1DJ49;R9XWD0;A0A4D6Q653;A0A4D6Q658;A0A4D6Q7Q1;A0A5B9Y2N1;A0A5B9Y3R8;R9XU95;A0A4D6Q614;R9XSY7;F2X323;F2XAR4;M9TGF7;B6UKN8;B6DQB6;B6DQB5
$\gamma$ -Gliadin 8	A0A5B9Y471
$\gamma$ -Gliadin 9	K7WVC4;R9XV71;R9XV82;B6UKP3;B6UKM4;B6UKP4;B6UKN2;U5UA50;K7XEG8;U5U7C3;A0A4D6Q642;A0A4D6QE74;U5U9Q4;B6UKM9;U5U6N8;B6UKN6;K7XRJ1;Q94G94;R9XT13;R9XT17;R9XT08;R9XWA9;U5U9Q1;A0A4D6Q653;A0A4D6Q702;R9XUB4;R9XUT2;R9XV74;B6UKQ1;A0A4D6K428;A0A4D6K4V7;A7XDG7;R9XSZ9;R9XUC3;B6UKQ6;A0A4D6K1Y6;A0A290XZ51;A0A4D6K1B1
$\gamma$ -Gliadin 10	U5U7C7;R9XUR5;R9XUR0;R9XSY2;K7XRH9;K7XEG0;K7XOY6;B5ANT0;H9A6C2;H9A6C0;H9A6B9;R9XUQ5;K7XRI7;K7WVB5;B5ANT1
$\gamma$ -Gliadin 11	P06659;M9TE83;B6UKQ4;F2X0K8;B6UKQ2;B6UKP2;B6UKN4
$\gamma$ -Gliadin 12	A0A4D6Q720;A0A4D6Q669;A0A5B9Y2H5;A0A4D6Q640;B6UKQ8;B6UKQ5;A7XDG6;A0A4D6Q692;U5U6M7;K7XEG5;A0A4D6Q625;U5U6P1;U5UA51;U5U7C0
$\gamma$ -Gliadin 13	I7KM78;A0A9R1DIU5;A0A2U8JD37;A0A9R1IN17;A0A1D6RERO
$\gamma$ -Gliadin 14	A0A4D6QBN0;A0A9R1DIX6;B6UKN9;Q9SYX8;A0A4D6Q623;A0A4D6Q633;A0A4D6Q9I7;A0A4D6QBP5;A0A5B9Y1L3;A1EHE7;A0A060N479;R9XV78;B6UKM6;Q8L6B5;Q8L6B3;Q8L6B2;Q6EEW8;Q8L6B4;Q9FS76;Q9FS75;Q9FS77;P08079
$\gamma$ -Gliadin 15	K7X1R6
$\gamma$ -Gliadin 16	F6KV47
$\gamma$ -Gliadin 17	A0A290XY3;A0A290XZ34;A0A290XYX6;A0A290XZ50;A0A290XYZ4;Q6EEW7;A0A290XZ27;Q9M4L5;A0A290XZ28;A0A290XZ26;A0A290XZ01;A0A290XY9;A0A290XZ16;A0A290XYX5;A0A290XYX2;A0A290XYW9;A0A290XYW5;A0A290XYV1;A0A290XYU0;A0A290XYX4;A0A290XZ21;A0A290XYX7;Q6EEW0;A0A290XYZ2;A0A290XY5;A0A290XZ08;A0A290XYX1;A0A290XYV3;B6DQB7;A0A290XYV4;A0A290XYZ6
$\gamma$ -Gliadin 18	R9XUS6

ANHANG

γ-Gliadin 19	R9XUB9;K7X158;Q9M6P7;Q94G91;R9XUB0;U5UA54;R9XW98;P08453;Q6EEW9;R9XWB3;R9XT02;R9XUT4;R9XV87;A0A290XZ33;A0A290XYU1;B6DQB3;A0A290XYW4;A0A290XZ20;F4YT72
γ-Gliadin 20	A0A4D6Q668;F2XAR5
γ-Gliadin 21	Q94G97;Q94G96;D0ES81;M9TG60;A0A3B5XT99;Q94G98
γ-Gliadin 22	B6DQB2;A0A4D6QE67;B6DQB4
γ-Gliadin 23	A0A4D6Q9J3;H9A6C1;H9A6B8;A0A290XZ18
γ-Gliadin 24	B6DQB9
γ-Gliadin 25	B6UKP5;D0ES79
γ-Gliadin 26	A0A4D6Q708
γ-Gliadin 27	F2X322
γ-Gliadin 28	R9XWB8
γ-Gliadin 29	A0A5B9Y1T4
γ-Gliadin 30	A0A290XYZ9
γ-Gliadin 31	A0A4D6QBN8
γ-Gliadin 32	B6UKQ3
γ-Gliadin 33	L7R666;L7R5A3;L7R599
γ-Gliadin 34	P04730
γ-Gliadin 35	U5U6M3

**Tabelle 18:** Übersicht der in der untargeted Analyse der BMWpop 2018 identifizierten niedermolekularen Gluteninuntereinheiten (LMW-GS) einschließlich zugehöriger Subtypen.

Name	UniProtKB ID	Typ
LMW-GS 1	B2BZC7	m
LMW-GS 2	R4JJAQ1;P10385;D2DII3;Q6SPY8	i
LMW-GS 3	D0EVDN9;A0A165R811;V9P748;A0A3G3C4J5;A0A3G3C4I5;A0A3G3C4H7;A0A3G3C4H3;A0A3G3C4H2;A0A3G3C4E8;A0A0S2GJV0;F8SGP0;F8SGP2;Q00M59;R9YUB3;A0A3G3C4Q9;R9XUV6;R9XUE1;F8SGP4;P93791;P10386;I3XHQ3;I3XHQ1;R4JFK6;I3XHP9;I3XHP8;F8SGP1;A0A3G3C4H4;Q0PW08;A0A3G3C4I0;Q0GQX0;Q52NZ3;Q30DX5;H6VLQ2;A0A3G3C4W4;A0A0S2GJP9;P94021;H6VLQ5;A0A3G3C4J8;F4YT78;D0EVP5;C0KK82;Q7DM83;C0KK81;A0A7D4XM12;C0KK80;A0A7D5BXD7;A0A7D4XYC5;Q30DX7;H6VLQ1;X2JC84;R9YQY9;H6VLQ0;F8SGP3;X2JBS3;A0A3G3C4G5;Q8H0J5;A0A3G3C4G7;A0A3G3C4W3;P93793;A0A3G3C4W5;H9XH01;A0A3G3C4P3;Q30DX4;Q30DX6	m & s
LMW-GS 4	A0A7D5CN18;A0A2R3C1P4;A0A7D5BxB5;Q3LGB7;V9P7D5;V9P6Q0;V9P7D0;V9P778	m
LMW-GS 5	A0A7D5BJ88	m
LMW-GS 6	A0A060N0C7;Q7X9J9;V9P6D7;F5A655;B5ANT2	m & s
LMW-GS 7	A0A7D5BVX1;A0A7D4Y114;A0A7D5C0C0;A0A7D5BIP1;A0A7D5BVU8	m
LMW-GS 8	Q8W3W5	s
LMW-GS 9	Q5MFAQ0;Q5MFN4;Q0GNF9;R9XWG9;R9XVC9;I1XB60;I1XB41;R9XT50;B2Y2Q7;B2Y2Q6;R9XVA5	m & s
LMW-GS 10	R4JDK8;Q8W3V6;Q8W3V7;Q8W3V8;Q8W3V9;D3U315;D2DII5;A0A7D4XNV2;R4JB46;A0A1B1RV25;A0A1U9XLA1;V9P766	m
LMW-GS 11	B2BZC4;R4JAN5;Q0ZCB0	m

ANHANG

LMW-GS 12	P93790;Q5MFN7;Q670Q4;Q8W3V4;D3U321;K9K7W7;A0A7D4XMB8;Q5MFN8;A0A2R4Q9V2;Q8W3V3;D0EVP3;A0A0S2GJR9;A0A0S2GJR2;A0A0S2GJS4;B2BZC5;Q0PW10;V9P781	m
LMW-GS 13	I3XHQ0	m
LMW-GS 14	Q1ZZT4	m
LMW-GS 15	A0A0S2GJT4;D3U318;R4JB53;F8SGL7;R4JFQ3;B2Y2S2;A0A0S2GJQ7;F8SGL4;C3VN79;R4JFP9;B2Y2S1;A0A0S2GJR7;B2Y2S6;F8SGM7;F8SGM8;F8SGM9;R4JDM5;F8SGL5;F8SGL6;F8SGL9;F8SGL8;R4JB62;F8SGM5;R4JAN9;C3VN78;F8SGL3;A0A0S2GJR1;V6BPA5;C5IXL3;Q41551;D3UAL6;V9P785;C8KIL6;A0A7D4XR9;Q5TLY8;F8SGL2;Q3LGB3;Q0ZCA8;Q5TLY9;C3VN80;F8SGN0;I3QPI2	m & s
LMW-GS 16	Q8H737;Q6J161;R4JBK0;Q8W3X6;F8SGQ3;B2BZD1;Q41552;D3U326;R4JAP5	m & s
LMW-GS 17	A4KZ73	m
LMW-GS 18	H6VLP8;D7RFI3	m
LMW-GS 19	B5ANT3	m
LMW-GS 20	F8SGN9	m
LMW-GS 21	A0A7D4XM26;A0A7D4XY91;Q52NZ6;F4YT77;I3QPG6;H9XH00;P93792;A0A7D5C1I4;A0A7D5BLN1;A0A7D4XMD4;R4JAT4;V9P7D6;R9YQY5;X2JC78;A0A7D5C1J7;V9P6P2;V9P6Q3;V9P754;D0PX24	m
LMW-GS 22	Q571Q5	-
LMW-GS 23	V9P7D3	m
LMW-GS 24	B2BZD2	m
LMW-GS 25	C3VN75;R9XV91;D3U317;Q8W3V1;R4JB34;A0A1U9XL81;D2DII1;C3VN77;C3VN74;Q6QGW0;Q5MFH3;Q5MFG8;Q6SPY7;A0A1U9XLC8;R4JAS6;Q6SPY9;O22116;I1XB50;D2DII2	i
LMW-GS 26	Q8W3X1	s
LMW-GS 27	Q8W3X5	m
LMW-GS 28	U6BEQ9	m
LMW-GS 29	F6M7E1;I3QPH4;R9YUB0	m
LMW-GS 30	A0A0S2GJS0;Q0QBR3;Q0GNG1	m
LMW-GS 31	Q6J162	s
LMW-GS 32	M1GP58;Q6SPZ1;R9XVB6;R9XT45;A0A3G3C4G6;R9XT25;Q6SPZ3;A0A1U9XLB3;R4JBI2;D3UAL5;R4JAR7;A0A1U9XLA9;P93794;Q6SPZ2;B8R6A0;A0A3G3C4Q1;R9XWE0;R9XUF6;R9XV98;R9XWD5;Q0PW11;Q8W3U9;D2DII7;M1GMA2;R4V1P5;A0A0S2GJQ4;R4JFL5;R9XWF4;Q6SPZ0;X2JAE7;R4JBE8;R4JB50;R4JD42;R9XUU0;R4JB27;R4JB8;R9XUF2;R9XWF9;X2J8E3;R4JB19;R4JBH8;Q8W3V0;C3VN76;R4JB44;A0A1G4P206	i
LMW-GS 33	B2Y2R3	s
LMW-GS 34	B6ETR9;A0A9R1DIZ0;A0A290XYT7;U5UA46;A0A290XYZ5;A0A290XYW6;R9XUY1;U5U9P7;A0A290XYS8;U5U6L3;A0A7D5DKX6;R9XWH8;A0A2U8JD15;A0A7D5DPJ7	D
LMW-GS 35	R4JD66;R4JBL2;A0A7D5BM85;O22108;D2DII6;I1XB44;D2DII8;A0A7D4XMC6;R4JFK3	m
LMW-GS 36	A0A0S2GJQ3;A0A0S2GJP7;V9P6P6;F8SGQ2;Q5MFP7;R9XUH5;R9XVC3;Q5MFAQ6;Q8W3W1;A0A0S2GJQ8;B3EY88;D0EVP6;Q8W3W2;R9XT30;R9XT35;R9XUG8;Q00M61;F8SGP6;F8SGP8;R9XVD4;A0A3G3C4G2;D0EVP2;Q5MFAQ2;R9XUE5;R9XT62;R9XUW9;Q5MFAQ1;R9XT57;R9XUW0;Q5MFAQ5;R9XUX7;Q8W3W0;X2JFZ5;V9P6D0;V9P767;D0EVP4;R9XUX4;D0EVP0;A0A2R3C1T0;B5TWK4;R9YU07;V9P6C5;X2JC81;V9P6N8;A0A0S2GJQ6;V9P741;K7X1P3;B3EY90;Q5MFP8;R4JAT6	m

ANHANG

	;V9P769;G3F3Z2;G3F3Z1;I3RTU0;V9P776;V9P757;A0A2R3C1P3;G3F3Z3;D0EVP1;Q5MFP4;R9XWH3;V9P7D8;B3EY89;B5TWK5;R9YSZ0;V9P7D1;I3RTU2;A0A7D4Y136;R4JFH1;Q3LGB6	
LMW-GS 37	A0A0S2GJR0	m
LMW-GS 38	A0A3G3C4I6	m
LMW-GS 39	B2Y2S3;K0I5W1;K0HZJ2	s
LMW-GS 40	F8SGQ4	i
LMW-GS 41	H9XGZ9	m-
LMW-GS 42	I3XHQ2;Q5MFP0;R4JB49;Q0QBR5;P04729;R9XWG5;Q8W3V5;Q5MFP3;Q5MFP1;Q0PW09;K7XED9;A0A7D5BHT4;A0A7D5BHS3;A0A7D4XLZ4;A0A0S2GKA3;K7XRG7;D2K763;F8SGM3;D3U323;Q41547;F8SGM2;A0A7D5BHQ3;A0A7D5BIQ7	m
LMW-GS 43	P16315	m
LMW-GS 44	Q52NZ5	m
LMW-GS 45	Q84U20;R9YTM4	m
LMW-GS 46	Q8W3W7;R9XWE6;R9XVB0	m & s
LMW-GS 47	V9P760	m

**Tabelle 19:** Übersicht der in der untargeted Analyse der BMWpop 2018 identifizierten hochmolekularen Gluteninuntereinheiten (HMW-GS) einschließlich Protein- und Genzuordnung anhand der UniProtKB-Datenbank.

	UniProtKB			
	Protein ID	Protein	Gen	Sequenzlänge
HMW-GS 1	A0A060MZP1	High molecular weight glutenin subunit	glu	794
	Q8H0L2	Glutenin high molecular weight subunit	HMW-GS(Ax)	238
	Q8H0L5	Glutenin high molecular weight subunit	HMW-GS(Ax)	191
	P02862	Glutenin, high molecular weight subunit PC237	-	39
	X5CQ58	HMW glutenin	DAL1Slx	231
	W6AW98	High molecular weight glutenin subunit x	Glu	742
	Q94IJ7	High molecular weight glutenin subunit x	glu-1R	811
HMW-GS 2	Q6RX92	HMW glutenin subunit	Glu	658
	Q52JL3	HMW glutenin subunit	-	658
	P08488	Glutenin, high molecular weight subunit 12	-	660
	Q6RX93	HMW glutenin subunit	-	658
	E4W506	HMW glutenin subunit	1Dy12.2*	658
	A4URY8	Dy-type high molecular weight subunit protein	1Dy12.3	652
	A0A0H4SR12	High molecular weight glutenin subunit	Glu-1Dy12.6	673
	A9QUS3	High molecular weight glutenin subunit	1Dy12*	673
	D2CPI7	HMW glutenin subunit	1Dy11	637
	A0A1W6C2K2	High molecular weight glutenin	HMW-G	658
	A0A0H4SMF2	High molecular weight glutenin subunit	Glu-1Dy12.7	658
	A0A1G4P1V2	HMW glutenin y-type subunit By8	Glu-1By	717

ANHANG

	Q670Q5	High molecular weight glutenin subunit 1Dy10.1	-	655
	A0A9R1IXL7	High-molecular-weight glutenin subunit	-	220
	A0A1G4P1T5	HMW glutenin $\gamma$ -type subunit 1By7	Glu-1By	702
	Q8H0L6	Glutenin high molecular weight subunit	HMW-GS(Dy type)	196
	Q8H0K9	HMW glutenin	-	201
HMW-GS 3	B1B520	High-molecular-weight glutenin subunit 2.6	Glu-D1	1025
	Q599I0	HMW subunit of glutenin 1Dx2.2*	glu-D1x	1024
	G3FLC7	High-molecular-weight glutenin subunit	Glu	971
	J7G6L4	High-molecular-weight glutenin 1Dx2.2	-	971
	P10388 <sup>a</sup>	Glutenin, high molecular weight subunit DX5	GLU-1D-1D	848
	Q0Q5D2 <sup>a</sup>	X-type HMW glutenin	-	848
	X2JUA0	High molecular weight glutenin subunit Dx5	-	839
	C6L669	High molecular weight glutenin subunit 2.8	Glu-D1	752
	E2CT66	High molecular weight glutenin subunit	1Dx1.5*	827
	A9YSK4	High molecular weight glutenin subunit	Glu-D1-1	839
	D7REK2	HMW glutenin subunit 1Dx5'	-	833
	D0IQ07	High-molecular-weight glutenin Dx2	Glu-D1	835
	A0A0H4PMT1	High molecular weight glutenin subunit	1Dx2.1t	836
	P08489	Glutenin, high molecular weight subunit PW212	-	838
	Q8GV12	HMW glutenin subunit	1Dx2.2	971
	Q1KL96	1Dx high molecular weight glutenin subunit	-	839
	Q6R2V1	High-molecular-weight glutenin subunit 1Dx2.1	-	836
	Q7M1M7 <sup>a</sup>	High-molecular-weight glutenin	-	161
	A0A0K0KCV6	High molecular weight glutenin subunit 1Dx2	-	831
	C0SUC3 <sup>a</sup>	High-molecular-weight glutenin subunit x5	Glu1Dx5	848
	D0IQ05 <sup>a</sup>	High-molecular-weight glutenin Dx5	-	848
	Q38LF5 <sup>a</sup>	HMW glutenin protein Dx5	glu	628
	A0A9R1IXH0	HMW glutenin subunit	-	221
	Q6R509	HMW glutenin subunit	HMW-GS	188
	Q6R508	HMW glutenin subunit	HMW-GS	188
	Q6R507	HMW glutenin subunit	HMW-GS	188
	G1E6K7 <sup>a</sup>	High molecular weight glutenin subunit Dx5	HMW-GS	118
	A0A249Y2Y2	High molecular weight glutenin subunit 1Slx2.3	1Slx2.3*	941
	A0A19RW78	HMW glutenin 1Mgx	-	820
	Q6R510	HMW glutenin subunit	HMW-GS	188
	A0A2D0UFQ0	X-type HMW-GS 1Slx2.2	1Slx2.2	974
	A0A19RW80	HMW glutenin 1Ugx	-	995
	X5D5A6	HMW glutenin	DAL1Slx	233
HMW-GS 4	Q18MZ6	High-molecular-weight glutenin subunit Bx17	Glu-B1	747
	A9YSK5	High molecular weight glutenin subunit	Glu-B1-1	783
	Q42451	Glu-B1-1b HMW glutenin subunit	-	789

ANHANG

	Q45R38	HMW glutenin x-type subunit Bx7	Glu-B1a1	795
	G4Y3Y1	High-molecular-weight glutenin subunit Bx7.2	Glu-B1br	795
	G3FLC5	High-molecular-weight glutenin subunit	Glu	795
	G4Y3Y0	High-molecular-weight glutenin subunit Bx7.1	Glu-B1br	795
	M4M8L5	High-molecular-weight glutenin subunit Bx17	HMWBx17	754
	G4Y3Y2	High-molecular-weight glutenin subunit Bx7.3	Glu-B1bs	795
	Q1KL95	1Bx high molecular weight glutenin subunit	-	789
HMW-GS 5	Q03871	HMW glutenin subunit 1By9	Glu-1By9	705
	W6AX70	High molecular weight glutenin subunit	Glu	720
	U5YQV1	High molecular weight glutenin subunit By18	-	720
	Q52JL2	HMW glutenin subunit	-	705
	Q84UY5	HMW glutenin	-	712
	B8PSA6	High molecular weight glutenin subunit 1By15	-	717
	A0A4Y6GR70	High molecular weight glutenin subunit 1By20	Glu-B1-2	717
	Q4JHY1	High molecular weight glutenin subunit 1By15	-	723
	W0C8U3	High-molecular-weight glutenin subunit 1By15	Glu-B1	717
	A0A060AER5	High-molecular-weight glutenin subunit 1By15	HMW-GS 1By15*	702
	X5CN16	HMW glutenin	DAL1Sly	164
HMW-GS 6	Q0Q5D8	High-molecular-weight glutenin By8	Glu	720
	Q94IJ6	High molecular weight glutenin subunit y	glu-1R	720
	A0A9R1DFL1	High-molecular-weight glutenin subunit	-	258
HMW-GS 7	Q41553	HMW glutenin subunit Ax2	HMWG	815
	Q03872	High molecular weight glutenin subunit 1Ax1	Glu-1AX1	830
	A0A059UHD1	High molecular weight glutenin subunit 1Ax1	Glu-1Ax1	830
	G3K725	High molecular weight glutenin subunit	Glu-1Ax1.1	866
	A0A0X9ARV2	High molecular weight glutenin subunit	Glu-Ax	830
	A0A0X9BSF8	High molecular weight glutenin subunit	Glu-Ax	836
	A0MZ38	High molecular weight glutenin subunit	Glu-1Ax	815
	T2HRF3	High-molecular-weight glutenin subunit	Glu-A1	824
	Q19AE4	High molecular weight glutenin subunit	GluA1	824
	A0A0X9BZV4	High molecular weight glutenin subunit	Glu-Ax	830
	A0A1G4P1V3	HMW glutenin x-type subunit 1Ax21	Glu-1Ax	814
	P02861	Glutenin, high molecular weight subunit PC256	-	101
	Q8H0L1	HMW glutenin	-	220
Q9SDM3	Glutenin, high molecular weight subunit type x	Glu-R1	766	
HMW-GS 8	A0A1B0Z3C8	Glutenin	Glu1Bx6	824
	Q7XAJ1	High-molecular-weight glutenin subunit	Glu-B1-1	119
	Q7XAH3	High-molecular-weight glutenin subunit	Glu-B1-1	114
	Q7XAH2	High-molecular-weight glutenin subunit	Glu-B1-1	114
	Q7XAG7	High-molecular-weight glutenin subunit	Glu-B1-1	114
	Q7XAH1	High-molecular-weight glutenin subunit	Glu-B1-1	114
	Q7XAH0	High-molecular-weight glutenin subunit	Glu-B1-1	114
	Q7XAG6	High-molecular-weight glutenin subunit	Glu-B1-1	114

ANHANG

	Q7XAI4	High-molecular-weight glutenin subunit	Glu-B1-1	119
	Q7XAH7	High-molecular-weight glutenin subunit	Glu-B1-1	114
HMW-GS 9	Q0Q5D3	Y-type HMW glutenin	-	648
	P10387	Glutenin, high molecular weight subunit DY10	GLU-D1-2B	648
	A9ZMG8	High-molecular-weight glutenin subunit y10	Glu1Dy10	648
	A9YSK3	High molecular weight glutenin subunit	Glu-D1-2	648
	A0A165B8S1	High molecular weight glutenin subunit Dy10-m328SF	-	648
	A0A1W6C2J9	High molecular weight glutenin	HMW-G	603
	V9TRL3	High molecular weight glutenin subunit 1Dy protein	Glu-D1-2	646
	A0A2L1K3K4	High molecular weight glutenin	Glu-B1	629
	B8YPU3	High molecular weight glutenin y-type	-	158
	Q8H0L3	Glutenin high molecular weight subunit	HMW-GS(Dy)	181
	Q8H0L0	HMW glutenin	-	211
	Q38LF4	HMW glutenin protein 1Dy10	glu	444
HMW-GS 10	A0A1G4P1V7	HMW glutenin x-type subunit Bx7	Glu-1Bx	824
	A0A1G4P1W4	HMW glutenin x-type subunit 1Bx6	Glu-1Bx	824
	A0A060AAZ6	High-molecular-weight glutenin subunit 1Bx14	HMW-GS 1Bx14*	789
HMW-GS 11	A0A1G4P1T3	HMW glutenin y-type subunit 1Ay21	Glu-1Ay	587
	A0A1G4P1S9	HMW glutenin y-type subunit 1Ay21	Glu-1Ay	588
HMW-GS 12	Q7XZB4	High-molecular-weight glutenin subunit	Glu-A1-2	137
	Q7X8H1	High-molecular-weight glutenin subunit	Glu-A1-2	137

<sup>a</sup>Stammen von dem Dx5 Allel (enthalten zusätzlichen Cysteinrest unmittelbar vor Beginn der repetitiven Domäne).

**Tabelle 20:** Übersicht der in der targeted Analyse der BMWpop 2018 verwendeten Peptide einschließlich Protein- und Genzuordnung anhand der UniProtKB-Datenbank.

Gluten-proteintyp	UniProtKB		
	Protein ID	Protein	Gene
α-Gliadin 1t	R9XW75	Prolamin	<i>Gli-2</i>
	A0A1K0J4W5	Prolamin	Gli-D2-11
	J7HT14	Prolamin	Gli-SB-4
α-Gliadin 2t	R9XSX7	Prolamin	<i>Gli-2</i>
	A7LHB1	Prolamin	-
	A0A3B6UAZ5	Uncharacterized protein	-
	R9XV34	Prolamin	<i>Gli-2</i>
	A7LHB6	Prolamin	-
	K7X0I0	Prolamin	<i>Gli-2</i>
	K7X1K5	Prolamin	<i>Gli-2</i>
	K7WV53	Prolamin	<i>Gli-2</i>
	K7X1F5	Prolamin	<i>Gli-2</i>
	K7XRC1	Prolamin	<i>Gli-2</i>
	K7XRA8	Prolamin	<i>Gli-2</i>
	K7XE59	Prolamin	<i>Gli-2</i>
	A0A9R0G5C7	Uncharacterized protein	-
	J7I021	Prolamin	Gli-CS-3
	R9XV48	Prolamin	<i>Gli-2</i>
	D0UTY3	Prolamin	-
	K7XR81	Prolamin	<i>Gli-2</i>
	A0A0E3UR68	Prolamin	-
	A0A0E3UQU4	Prolamin	-
	X2KS61	Prolamin	-
	R9XW94	Prolamin	<i>Gli-2</i>
	R9XW80	Prolamin	<i>Gli-2</i>
	K7XR49	Prolamin	<i>Gli-2</i>
	K7WUZ4	Prolamin	<i>Gli-2</i>
	R9XW73	Prolamin	<i>Gli-2</i>
	K7WV57	Prolamin	<i>Gli-2</i>
	R9XSW3	Prolamin	<i>Gli-2</i>
	K7X0I8	Prolamin	<i>Gli-2</i>
	K7XR61	Prolamin	<i>Gli-2</i>
	X2KVI4	Prolamin	-
	A0A0E3Z527	Prolamin	-
	K7XR67	Prolamin	<i>Gli-2</i>
	A0A1K0JS80	Prolamin	Gli-B2-1
A0A2R2Y407	Prolamin	<i>Gli-2</i>	

ANHANG

	K7X1G2	Prolamin	<i>Gli-2</i>
	K7XR55	Prolamin	<i>Gli-2</i>
	K7X1F1	Prolamin	<i>Gli-2</i>
	K7WV06	Prolamin	<i>Gli-2</i>
	A0A1K0ITH3	Prolamin	Gli-D2-8
	A5JSB5	Prolamin	Gli-Z5
	A0A2R2Y406	Prolamin	<i>Gli-2</i>
	K7X1J5	Prolamin	<i>Gli-2</i>
	R9XW75	Prolamin	<i>Gli-2</i>
	A0A1K0J4W5	Prolamin	Gli-D2-11
	J7HT14	Prolamin	Gli-SB-4
	A5JSA6	Prolamin	Gli-G3
	A0A9R0GKV9	Alpha-gliadin	Gli-D2-15
	A0A1K0JNG6	Prolamin	<i>Gli-2</i>
	A0A0K2QJ9	Prolamin	-
	K7XRC7	Prolamin	<i>Gli-2</i>
	A0A1K0IT13	Prolamin	Gli-D2-14
	A0A1K0JSA4	Prolamin	Gli-D2-13
	R9XUM8	Prolamin	<i>Gli-2</i>
	A0A2P1H695	Prolamin	Gli-D2
	A0A0E3Z7F7	Prolamin	-
	R9XW87	Prolamin	<i>Gli-2</i>
$\gamma$ -Gliadin 1t	R9XUB9	Gamma-gliadin	-
	K7X1S8	Gamma-gliadin	-
	Q9M6P7	Gamma-gliadin	-
	Q94G91	Gamma-gliadin	Gli-Yn1
	R9XUB0	Gamma-gliadin	-
	U5UA54	Gamma-gliadin	-
	R9XW98	Gamma-gliadin	-
	P08453	Gamma-gliadin	-
	Q6EEW9	Gamma-gliadin	-
	R9XT02	Gamma-gliadin	-
	R9XUT4	Gamma-gliadin	-
	R9XV87	Gamma-gliadin	-
	B6DQB3	Gamma-gliadin	-
	F4YT72	Gamma-gliadin	-
	H9BFB6	Gamma-gliadin	Gligamma
	D0ES83	Gamma-gliadin	-
	A0A9R1DJ49	Gamma-gliadin 2	-
	R9XWD0	Gamma-gliadin	-
	A0A4D6Q653	Gamma-gliadin	-
	A0A4D6Q6S8	Gamma-gliadin	-
A0A4D6Q7Q1	Gamma-gliadin	-	

ANHANG

	A0A5B9Y2N1	Gamma-gliadin	-
	A0A5B9Y3R8	Gamma-gliadin	-
	R9XU95	Gamma-gliadin	-
	A0A4D6Q614	Gamma-gliadin	-
	R9XSY7	Gamma-gliadin	-
	F2X323	Gamma-gliadin	Gli-Rz1
	F2XAR4	Gamma-gliadin	gli
	M9TGF7	Gamma gliadin-D2	<i>Gli-1</i>
	B6UKN8	Gamma-gliadin	II923
	B6DQB6	Gamma-gliadin	-
	B6DQB5	Gamma-gliadin	-
	B6DQB8	Gamma-gliadin	-
	B6UKN0	Gamma-gliadin	II910
LMW-GS 1t	A0A0S2GJQ0	Prolamin	LMW-GS
LMW-GS 2t	P93790	Low-molecular-weight glutenin storage protein	-
	Q5MFN7	Low molecular weight glutenin	-
	Q670Q4	Low molecular weight glutenin subunit XJLMW-1347	-
	Q8W3V4	LMW-D3	lmw-gs
	D3U321	Low molecular weight glutenin subunit D3-4	LMW-GS
	K9K7W7	Low molecular weight glutenin subunit	-
	D0EVP3	Prolamin	-
	A0A7D4XMB8	Low-molecular-weight glutenin subunit	LMW-GS
	A0A0S2GJR9	Low-molecular-weight glutenin subunit	LMW-GS
	A0A2R4Q9V2	Prolamin	LMW-GS
	A0A0S2GJR2	Low-molecular-weight glutenin subunit	LMW-GS
	A0A0S2GJS4	Low-molecular-weight glutenin subunit	LMW-GS
	Q0PW10	Low molecular weight glutenin subunit GF-2	-
LMW-GS 3t	A0A0S2GJQ1	Low-molecular-weight glutenin subunit	LMW-GS
	R4JF25	Low-molecular-weight glutenin subunit	LMW-GS
	Q19MN3	Low-molecular-weight glutenin subunit Glu-A3	-
	R9XUD0	LMW-GS	glu-3
	Q5MFH2	Low molecular weight glutenin	-
	Q5MFH0	Low molecular weight glutenin	-
	M1GL80	Low-molecular-weight glutenin subunit	Glu-A3
	B3EY91	LMWGS1	SYGluA3
	Q19N55	Low molecular weight glutenin subunit Glu-A3	-
	Q19MN2	Low-molecular-weight glutenin subunit Glu-A3	-
	M1GP58	Low-molecular-weight glutenin subunit	Glu-A3
	Q6SPZ1	LMW-GS	Glu-A3
	R9XVB6	LMW-GS	glu-3
	R9XT45	LMW-GS	glu-3
	A0A3G3C4G6	Low molecular weight glutenin subunit A3-3	-
	R9XT25	LMW-GS	glu-3

ANHANG

Q6SPZ3	LMW-A2	Glu-A3
A0A1U9XLB3	Low-molecular-weight glutenin subunit	GluA3-19
R4JBI2	Low-molecular-weight glutenin subunit	LMW-GS
D3UAL5	Low molecular weight glutenin subunit A3-2	LMW-GS
R4JAR7	Low-molecular-weight glutenin subunit	LMW-GS
A0A1U9XLA9	Low-molecular-weight glutenin subunit	GluA3-111
P93794	Low-molecular-weight glutenin storage protein	-
Q6SPZ2	Low molecular weight glutenin	Glu-A3
B8R6A0	Low molecular weight glutenin	Glu-A3
A0A3G3C4Q1	Low molecular weight glutenin subunit A3-6	-
R9XWE0	LMW-GS	glu-3
R9XUF6	LMW-GS	glu-3
R9XV98	LMW-GS	glu-3
R9XWD5	LMW-GS	glu-3
Q0PW11	Low molecular weight glutenin subunit GF-1	-
Q8W3U9	Prolamin	lmw-gs
D2DII7	Low molecular weight glutenin subunit A3-4	GluA3-4
M1GMA2	Low-molecular-weight glutenin subunit	Glu-A3
R4V1P5	Low molecular weight glutenin subunit	LMW-GS
A0A0S2GJQ4	Low-molecular-weight glutenin subunit	LMW-GS
R4JFL5	Low-molecular-weight glutenin subunit	LMW-GS
Q6SPZ0	Low molecular weight glutenin	Glu-A3
X2JAE7	Low-molecular-weight glutenin subunit Glu-A3	-
R4JBE8	Low-molecular-weight glutenin subunit	LMW-GS
R4JB50	Low-molecular-weight glutenin subunit	LMW-GS
R4JD42	Low-molecular-weight glutenin subunit	LMW-GS
R9XUU0	LMW-GS	glu-3
R4JB27	Low-molecular-weight glutenin subunit	LMW-GS
R4JBJ8	Low-molecular-weight glutenin subunit	LMW-GS
R9XUF2	LMW-GS	glu-3
R9XWF9	LMW-GS	glu-3
X2J8E3	Low-molecular-weight glutenin subunit Glu-A3	-
R4JB19	Low-molecular-weight glutenin subunit	LMW-GS
R4JBH8	Low-molecular-weight glutenin subunit	LMW-GS
Q8W3V0	Low-molecular-weight glutenin subunit group 12 type VI	lmw-gs
C3VN76	Low-molecular-weight glutenin subunit	Glu-A3
C3VN75	Low molecular weight glutenin	Glu-A3
R9XV91	LMW-GS	glu-3
D3U317	Low molecular weight glutenin subunit A3-2	LMW-GS
Q8W3V1	Low-molecular-weight glutenin subunit group 11 type VI	lmw-gs
R4JB34	Low-molecular-weight glutenin subunit	LMW-GS
A0A1U9XL81	Low-molecular-weight glutenin subunit	GluA3-112
D2DII1	Low-molecular-weight glutenin subunit	GluA3-14

ANHANG

	C3VN77	Prolamin	Glu-A3
	C3VN74	Low molecular weight glutenin	Glu-A3
	Q6QGW0	LMW-i glutenin pGH3.1	-
	Q5MFH3	Low molecular weight glutenin	-
	Q5MFG8	Low molecular weight glutenin	-
	Q6SPY7	LMW-GS	Glu-A3
	A0A1U9XLC8	Low-molecular-weight glutenin subunit	GluA3-18
	R4JAQ1	Low-molecular-weight glutenin subunit	LMW-GS
	P10385	Glutenin, low molecular weight subunit	-
	D2DII3	Low-molecular-weight glutenin subunit	GluA3-16
	Q6SPY8	Low molecular weight glutenin	Glu-A3
	F8SGQ4	Low-molecular-weight glutenin subunit	-
	A0A1U9XLB4	Prolamin	GluA3-312
HMW-GS 1t	Q0Q5D3	Y-type HMW glutenin	-
	P10387	Glutenin, high molecular weight subunit DY10	GLU-D1-2B
	A9ZMG8	High-molecular-weight glutenin subunit y10	Glu1Dy10
	A9YSK3	High molecular weight glutenin subunit	Glu-D1-2
	A0A165B8S1	High molecular weight glutenin subunit Dy10-m328SF	-
	A0A1W6C2J9	High molecular weight glutenin	HMW-G
	V9TRL3	High molecular weight glutenin subunit 1Dy protein	Glu-D1-2
	A0A2L1K3K4	High molecular weight glutenin	Glu-B1
HMW-GS 2t	A0A060MZP1	High molecular weight glutenin subunit	glu
HMW-GS 3t	Q0Q5D3	Y-type HMW glutenin	-
	P10387	Glutenin, high molecular weight subunit DY10	GLU-D1-2B
	A9ZMG8	High-molecular-weight glutenin subunit y10	Glu1Dy10
	A9YSK3	High molecular weight glutenin subunit	Glu-D1-2
	A0A165B8S1	High molecular weight glutenin subunit Dy10-m328SF	-
	A0A1W6C2J9	High molecular weight glutenin	HMW-G
	V9TRL3	High molecular weight glutenin subunit 1Dy protein	Glu-D1-2
	Q8H0L0	HMW glutenin	-
	Q6RX92	HMW glutenin subunit	Glu
	Q52JL3	HMW glutenin subunit	-
	P08488	Glutenin, high molecular weight subunit 12	-
	Q6RX93	HMW glutenin subunit	-
	E4W506	HMW glutenin subunit	1Dy12.2*
	A4URY8	Dy-type high molecular weight subunit protein	1Dy12.3
	A0A0H4SR12	High molecular weight glutenin subunit	Glu-1Dy12.6
	A9QUS3	High molecular weight glutenin subunit	1Dy12*
	D2CPI7	HMW glutenin subunit	1Dy11
	A0A1W6C2K2	High molecular weight glutenin	HMW-G
	A0A0H4SMF2	High molecular weight glutenin subunit	Glu-1Dy12.7
	A0A1G4P1V2	HMW glutenin y-type subunit By8	Glu-1By
Q670Q5	High molecular weight glutenin subunit 1Dy10.1	-	

ANHANG

	A0A1G4P1T5	HMW glutenin $\gamma$ -type subunit 1By7	Glu-1By
	Q8H0K9	HMW glutenin	
HMW-GS 4t	A5HMG2	HMW glutenin subunit 1By16	
	Q9SDM2	Glutenin, high molecular weight subunit type $\gamma$	Glu-R1
	Q94IJ8	High molecular weight glutenin subunit $\gamma$	glu-1R
	Q9S8V9	High-molecular-weight glutenin SUBUNI	
	Q03871	HMW glutenin subunit 1By9	Glu-1By9
	W6AX70	High molecular weight glutenin subunit	Glu
	U5YQV1	High molecular weight glutenin subunit By18	
	Q52JL2	HMW glutenin subunit	
	Q84UY5	HMW glutenin	
	B8PSA6	High molecular weight glutenin subunit 1By15	
	A0A4Y6GR70	High molecular weight glutenin subunit 1By20	Glu-B1-2
	Q4JHY1	High molecular weight glutenin subunit 1By15	
	W0C8U3	High-molecular-weight glutenin subunit 1By15	Glu-B1
	A0A060AER5	High-molecular-weight glutenin subunit 1By15	HMW-GS 1By15*
	Q0Q5D8	High-molecular-weight glutenin By8	Glu
	Q94IJ6	High molecular weight glutenin subunit $\gamma$	glu-1R
	A0A9R1DFL1	High-molecular-weight glutenin subunit	
	Q0Q5D3	$\gamma$ -type HMW glutenin	-
	P10387	Glutenin, high molecular weight subunit DY10	GLU-D1-2B
	A9ZMG8	High-molecular-weight glutenin subunit $\gamma$ 10	Glu1Dy10
	A9YSK3	High molecular weight glutenin subunit	Glu-D1-2
	A0A165B8S1	High molecular weight glutenin subunit Dy10-m328SF	-
	A0A1W6C2J9	High molecular weight glutenin	HMW-G
	V9TRL3	High molecular weight glutenin subunit 1Dy protein	Glu-D1-2
	A0A2L1K3K4	High molecular weight glutenin	Glu-B1
	B8YPU3	High molecular weight glutenin $\gamma$ -type	
	Q8H0L3	Glutenin high molecular weight subunit	HMW-GS(Dy type)
	Q6RX92	HMW glutenin subunit	Glu
	Q52JL3	HMW glutenin subunit	
	P08488	Glutenin, high molecular weight subunit 12	
	Q6RX93	HMW glutenin subunit	
	E4W506	HMW glutenin subunit	1Dy12.2*
	A4URY8	Dy-type high molecular weight subunit protein	1Dy12.3
	A0A0H4SR12	High molecular weight glutenin subunit	Glu-1Dy12.6
	D2CPI7	HMW glutenin subunit	1Dy11
	A0A1W6C2K2	High molecular weight glutenin	HMW-G
	A0A0H4SMF2	High molecular weight glutenin subunit	Glu-1Dy12.7
	A0A1G4P1V2	HMW glutenin $\gamma$ -type subunit By8	Glu-1By
	Q670Q5	High molecular weight glutenin subunit 1Dy10.1	
	A0A9R1IXL7	High-molecular-weight glutenin subunit	
	A0A1G4P1T5	HMW glutenin $\gamma$ -type subunit 1By7	Glu-1By

ANHANG

	Q8H0L6	Glutenin high molecular weight subunit	HMW-GS(Dy type)
	W6AWK6	High molecular weight glutenin subunit	Glu
	W6AW92	High molecular weight glutenin subunit y	Glu
	A0A023GQG9	Y-type high molecular weight glutenin subunit	
	X5CHT6	HMW glutenin	DAL1Sly
	Q7XZA7	High-molecular-weight glutenin subunit	Glu-A1-2
	A0A1G4P1T3	HMW glutenin y-type subunit 1Ay21	Glu-1Ay
	A0A1G4P1S9	HMW glutenin y-type subunit 1Ay21	Glu-1Ay
	Q7XZB4	High-molecular-weight glutenin subunit	Glu-A1-2
	Q7X8H1	High-molecular-weight glutenin subunit	Glu-A1-2
HMW-GS 5t	B1B520	High-molecular-weight glutenin subunit 2.6	Glu-D1
	Q59910	HMW subunit of glutenin 1Dx2.2*	glu-D1x
	G3FLC7	High-molecular-weight glutenin subunit	Glu
	J7G6L4	High-molecular-weight glutenin 1Dx2.2	
	P10388	Glutenin, high molecular weight subunit DX5	GLU-1D-1D
	Q0Q5D2	X-type HMW glutenin	
	C6L669	High molecular weight glutenin subunit 2.8	Glu-D1
	A9YSK4	High molecular weight glutenin subunit	Glu-D1-1
	E2CT66	High molecular weight glutenin subunit	1Dx1.5*
	X2JUA0	High molecular weight glutenin subunit Dx5	
	D7REK2	HMW glutenin subunit 1Dx5'	
	D0IQ07	High-molecular-weight glutenin Dx2	Glu-D1
	A0A0H4PMT1	High molecular weight glutenin subunit	1Dx2.1t
	Q8GV12	HMW glutenin subunit	1Dx2.2
	P08489	Glutenin, high molecular weight subunit PW212	
	Q6R2V1	High-molecular-weight glutenin subunit 1Dx2.1	
	A0A0K0KCV6	High molecular weight glutenin subunit 1Dx2	
	C0SUC3	High-molecular-weight glutenin subunit x5	Glu1Dx5
	D0IQ05	High-molecular-weight glutenin Dx5	
	Q38LF5	HMW glutenin protein Dx5	glu
	Q6R507	HMW glutenin subunit	HMW-GS
	Q6R508	HMW glutenin subunit	HMW-GS
	Q6R509	HMW glutenin subunit	HMW-GS
	G1E6K7	High molecular weight glutenin subunit Dx5	HMW-GS

HMW-GS, hochmolekulare Gluteninuntereinheit; LMW-GS, niedermolekulare Gluteninuntereinheit.

ANHANG

<i>sp P10387 Dy10/1-648</i>	1	MAKRLVLF AAVVIALVALTTAEGEASRQLQCERELQESSLEACRQVV	47
<i>sp P08488 Dy12/1-660</i>	1	MAKRLVLF AAVVIALVALTTAEGEASRQLQCERELQESSLEACRQVV	47
<i>tr Q0Q5D8 By8/1-720</i>	1	MAKRLVLFATVVITLVALTTAEGEASRQLQCERELQESSLEACRQVV	47
<i>tr Q03871 By9/1-705</i>	1	MAKRLVLFATVVITLVALTTAEGEASRQLQCERELQESSLEACRQVV	47
<i>sp P10387 Dy10/1-648</i>	48	DQQLAGRLPWSTGLQMRCCQQLRDVSAKCRSVAVSQVARQYEQT VVP	94
<i>sp P08488 Dy12/1-660</i>	48	DQQLAGRLPWSTGLQMRCCQQLRDVSAKCRSVAVSQVARQYEQT VVP	94
<i>tr Q0Q5D8 By8/1-720</i>	48	DQQLAGRLPWSTGLQMRCCQQLRDVSAKCRSVAVSQVVRQYEQT VVP	94
<i>tr Q03871 By9/1-705</i>	48	DQQLAGRLPWSTGLQMRCCQQLRDVSAKCRPVAVSQVVRQYEQT VVP	94
<i>sp P10387 Dy10/1-648</i>	95	PKGGSFYPGETTPLQQLQQGIFWGTSSQTVQGYYPVTS PRQGS S YYP	141
<i>sp P08488 Dy12/1-660</i>	95	PKGGSFYPGETTPLQQLQQGIFWGTSSQTVQGYYPVTS PRQGS S YYP	141
<i>tr Q0Q5D8 By8/1-720</i>	95	PKGGSFYPGETTPLQQLQQVIFWGTSSQTVQGYYPVSSPQQG P YYP	141
<i>tr Q03871 By9/1-705</i>	95	PKGGSFYPGETTPLQQLQQVIFWGTSSQTVQGYYPVSSPQQG P YYP	141
<i>sp P10387 Dy10/1-648</i>	142	GQAS PQQPGQGQQPGKWE PGQGQQWY - - - - - YPTSL	173
<i>sp P08488 Dy12/1-660</i>	142	GQAS PQQPGQGQQPGKWE PGQGQQWY - - - - - YPTSL	173
<i>tr Q0Q5D8 By8/1-720</i>	142	GQAS PQQPGQGQQPGKWE LGQGQQGYPTSLHQSGQGQQGY P S S L	188
<i>tr Q03871 By9/1-705</i>	142	GQAS PQQPGQGQQPGKWE LGQGQQGYPTSLHQSGQGQQGY P S S L	188
<i>sp P10387 Dy10/1-648</i>	174	QQPGQGQQIGKKGQQGYPTSLQQPGQ - - - - - GQQGYPTSLQHTGQ	214
<i>sp P08488 Dy12/1-660</i>	174	QQPGQGQQIGKKGKQQGYPTSLQQPGQGQQIGQGQQGYPTSPQH T G Q	220
<i>tr Q0Q5D8 By8/1-720</i>	189	QQPGQGQQIGQGQQGYPTSLQQPGQGQQIGQGQQGYPTSPQH P G Q	235
<i>tr Q03871 By9/1-705</i>	189	QQPGQGQQIGQGQQGYPTSLQQPGQGQQIGQGQQGYPTSPQH P G Q	235
<i>sp P10387 Dy10/1-648</i>	215	RQQPVQG - - - - - QQP - - - - - EQGQPGQWQQGYPTSPQQ LGQGQ	249
<i>sp P08488 Dy12/1-660</i>	221	RQQPVQGQQIGQGQQP - - - - - EQGQPGQWQQGYPTSPQQ LGQGQ	261
<i>tr Q0Q5D8 By8/1-720</i>	236	RQQPGQGQQIGQGQQLGQGRQIGQGQQSGQGQQGYPTSPQQ LGQGQ	282
<i>tr Q03871 By9/1-705</i>	236	RQQPGQGQQIGQGQQLGQGRQIGQGQQSGQGQQGYPTSPQQ LGQGQ	282
<i>sp P10387 Dy10/1-648</i>	250	QPRQWQSGQGQQGHYPTSLQQPGQGQQGHYLA SQQQPGQGQQGHYP	296
<i>sp P08488 Dy12/1-660</i>	262	QPGQWQSGQGQQGHYPTSLQQPGQGQQGHYLA SQQQPAQGQQGHYP	308
<i>tr Q0Q5D8 By8/1-720</i>	283	QPGQWQSGQGQQGYPTSLQQPGQGQQGYPASQQQPGQGQQGQYP	329
<i>tr Q03871 By9/1-705</i>	283	QPGQWQSGQGQQGYPTSLQQPGQGQQGYPASQQQPGQGQQGQYP	329
<i>sp P10387 Dy10/1-648</i>	297	ASQQQPGQGQQGHYPA SQQQPGQGQ - - - - - QGHYPAS	328
<i>sp P08488 Dy12/1-660</i>	309	ASQQQPGQGQQGHYPA SQQQPGQGQ - - - - - QGHYPAS	340
<i>tr Q0Q5D8 By8/1-720</i>	330	ASQQQPGQGQQGYPASQQQPAQGQQGYYPASQQQPGQGQQGHYLA S	376
<i>tr Q03871 By9/1-705</i>	330	ASQQQPG - - - - - QGQGGQYPASQQQPGQGQQGHYLA S	361
<i>sp P10387 Dy10/1-648</i>	329	QQEPGQGGQQGIPASQQQPGQGQQGHYPASLQQPGQGQQGHYP T S L Q	375
<i>sp P08488 Dy12/1-660</i>	341	QQEPGQGGQQGIPASQQQPGQGQQGHYPASLQQPG - - QGHYP T S L Q	385
<i>tr Q0Q5D8 By8/1-720</i>	377	QQQPGQGGQRHYPASLQQPGQGQQGHYTASLQQPGQGQQGHYPAS L Q	423
<i>tr Q03871 By9/1-705</i>	362	QQQPGQGGQRHYPASLQQPGQGQQGHYTASLQQPGQGQQGHYPAS L Q	408
<i>sp P10387 Dy10/1-648</i>	376	QLGQGQQTGQPGQKQQPGQGQQTGQG - - - - - QQP	404
<i>sp P08488 Dy12/1-660</i>	386	QLGQGQQIGQPGQKQQPGQGQQTGQG - - - - - QQP	414
<i>tr Q0Q5D8 By8/1-720</i>	424	QVGQGQQIGQLGQRQQPGQGQQTRQGQQLLEQGQQPGQGQQTRRQ G Q Q L	470
<i>tr Q03871 By9/1-705</i>	409	QVGQGQQIGQLGQRQQPGQGQQTRQGQQLLEQGQQPGQGQQTRRQ G Q Q L	455
<i>sp P10387 Dy10/1-648</i>	405	EEQQPGQGQQGYPTSLQQPGQG - - - - - QQQGQQGGYPTSLQQ	445
<i>sp P08488 Dy12/1-660</i>	415	EEQQPGQGQQGYPTSLQQPGQG - - - - - QQQGQQGGYPTSLQQ	455
<i>tr Q0Q5D8 By8/1-720</i>	471	EQGQQPGQGQQGYPTSPQQSGQGQQPGQSQQPGQGQQGYSS S L Q Q	517
<i>tr Q03871 By9/1-705</i>	456	EQGQQPGQGQQGYPTSPQQSGQGQQPGQSQQPGQGQQGYSS S L Q Q	502
<i>sp P10387 Dy10/1-648</i>	446	PGQGQQGHYPASLQQPGQ - - GQPGQRQQPGQGQHPEQGKQPGQGQQG	490
<i>sp P08488 Dy12/1-660</i>	456	PGQGQQGHYPASLQQPGQGQGQPGQRQQPGQGQHPEQGKQPGQGQQG	502
<i>tr Q0Q5D8 By8/1-720</i>	518	PGQGLQGHYPASLQQPGQ - - GHFPGRQQPGQGQPEQGQQPGQGQQG	562
<i>tr Q03871 By9/1-705</i>	503	PGQGLQGHYPASLQQPGQ - - GHFPGRQQPGQGQPEQGQQPGQGQQG	547

ANHANG

<i>sp P10387 Dy10/1-648</i>	491	YYPTSPQQPGQGGQLGQQGQQGYYPTSPQQPGQGGQGGQGGHCP	537
<i>sp P08488 Dy12/1-660</i>	503	YYPTSPQQPGQGGQLGQQGQQGYYPTSPQQPGQGGQGGQGGHCP	549
<i>tr Q0Q5D8 By8/1-720</i>	563	YYPTSPQQPGQGGQLGQQGQQGYYPTSPQQPGQGGQGGQGGHCP	609
<i>tr Q03871 By9/1-705</i>	548	YYPTSPQQPGQGGQLGQQGQQGYYPTSPQQPGQGGQGGQGGHCP	594
<i>sp P10387 Dy10/1-648</i>	538	PQQSGQAQQPGQGGQIGQVQQPGQGGQGGYYPTSVQQPGQGGQGGSGQG	584
<i>sp P08488 Dy12/1-660</i>	550	PQQTGQAQQLGQGGQIGQVQQPGQGGQGGYYPTSLQQPGQGGQGGSGQG	596
<i>tr Q0Q5D8 By8/1-720</i>	610	PQQTGQAQQPGQGGQIGQVQQPGQGGQGGYYPTISLQQSGQGGQGGSGQG	656
<i>tr Q03871 By9/1-705</i>	595	PQQTGQAQQPGQGGQIGQVQQPGQGGQGGYYPTISLQQSGQGGQGGSGQG	641
<i>sp P10387 Dy10/1-648</i>	585	QSGQGHQPGQGGQSGQEQQGYDSPYHVSAAEQQAASPMVAKAQQPATQ	631
<i>sp P08488 Dy12/1-660</i>	597	QSGQGHQPGQGGQSGQEQQGYDSPYHVSAAEQQAASPMVAKAQQPATQ	643
<i>tr Q0Q5D8 By8/1-720</i>	657	QSGQGHQLGQGGQSGQEQQGYDNPYHVNTAQKVAKVQQPATQ	703
<i>tr Q03871 By9/1-705</i>	642	QSGQGHQLGQGGQSGQEQQGYDNPYHVNTAQKVAKVQQPATQ	688
<i>sp P10387 Dy10/1-648</i>	632	LPTVCRMEGGDALSA	648
<i>sp P08488 Dy12/1-660</i>	644	LPTVCRMEGGDALSA	660
<i>tr Q0Q5D8 By8/1-720</i>	704	LPI MCRMEGGDALSA	720
<i>tr Q03871 By9/1-705</i>	689	LPI MCRMEGGDALSA	705

**Abbildung 37:** Darstellung des Sequenzvergleichs der HMW-GS Dy12 (P08488), Dy10 (P10387), By8 (Q0Q5D8) und By9 (Q03871) mit Jalview [177]. Farbcodierung der Aminosäuren: blau, Alanin, Isoleucin, Leucin, Methionin, Phenylalanin, Tryptophan, Valin (hydrophob); gelb, Prolin (Strukturbrecher); grün, Glutamin, Serin, Threonin (polar, ungeladene Seitenketten); orange, Glycin; rot, Arginin, Lysin (basisch); türkis, Histidin, Tyrosin (polare Seitenkette, Wasserstoffbrücke); violett, Asparaginsäure, Glutaminsäure (sauer); HMW-GS, hochmolekulare Gluteninuntereinheit.





ANHANG

<i>tr Q03872 Ax1/1-830</i>	639	QQPGQ - - - - - WQQPGQGLPGYYPTSSLQPEQQQGYYPTSQQQP	677
<i>tr Q41553 Ax2/1-815</i>	624	QQPGQ - - - - - WQQPGQGGQPGYYPTSSLQPEQQQGYYPTSQQQP	662
<i>tr A0A1B0Z3C8 Bx6/1-824</i>	625	QQGYY - - - T T S L Q Q S G Q G Q Q G Y Y L T S P Q Q S G Q G Q Q G Y Y P T S P Q Q S	666
<i>tr A0A1G4P1V7 Bx7/1-824</i>	625	QQGYY - - - T T S L Q Q S G Q G Q Q G Y Y L T S P Q Q S G Q G Q Q G Y Y P T S P Q Q S	666
<i>sp P10388 Dx5/1-848</i>	673	QQPGQ - - - - - WQQPGQGGQPGYYLTSPLQLGGQQGYYP TSLQQP	711
<i>sp P08489 Dx2/1-838</i>	657	QQPGQWQQPGQWQQPGQGGQPGYYLTSPLQLGGQQGYYP TSLQQP	701
<i>tr Q03872 Ax1/1-830</i>	678	GQGPGPGQWQQSGQGQQGYYPTSPQQSGQGQQPGQWLPQPGQWLPQS	722
<i>tr Q41553 Ax2/1-815</i>	663	GQGPGPGQWQQSGQGQQGYYPTSPQQSGQGQQPGQWLPQPGQWLPQS	707
<i>tr A0A1B0Z3C8 Bx6/1-824</i>	667	GQGQQPGQGQQPRQGQQGYYPI SPQQSGQGQQPGQG - - - - - QQ	704
<i>tr A0A1G4P1V7 Bx7/1-824</i>	667	GQGQQPGQGQQPRQGQQGYYPI SPQQSGQGQQPGQG - - - - - QQ	704
<i>sp P10388 Dx5/1-848</i>	712	GQGQQPGQWQQSGQGQHWYYPTSPQLSGQGQR - - - - - 743	743
<i>sp P08489 Dx2/1-838</i>	702	GQGQQPGQWQQSGQGQHGYYPTSPQLSGQGQR - - - - - 733	733
<i>tr Q03872 Ax1/1-830</i>	723	GYYL TSPQQ LGQQG - - - - - QPRQWLQPRQGQQGYYP TSPQQSGQ	761
<i>tr Q41553 Ax2/1-815</i>	708	GYYL TSPQQ LGQQG - - - - - QPRQWLQPRQGQQGYYP TSPQQSGQ	746
<i>tr A0A1B0Z3C8 Bx6/1-824</i>	705	GYYPTSPQQSGQGQQPGHEQQPGQWLQPGQGQQGYYP TSSQQSGQ	749
<i>tr A0A1G4P1V7 Bx7/1-824</i>	705	GYYPTSPQQSGQGQQPGHEQQPGQWLQPGQGQQGYYP TSSQQSGQ	749
<i>sp P10388 Dx5/1-848</i>	744	- - - - - PGQWLQPGQGQQGYYP TSPQQPGQ	767
<i>sp P08489 Dx2/1-838</i>	734	- - - - - PGQWLQPGQGQQGYYP TSPQQSGQ	757
<i>tr Q03872 Ax1/1-830</i>	762	GQQL - - - - - GQGQQGYYP TSPQQ - - - - - SGQGQQGYDSSPYHVS	794
<i>tr Q41553 Ax2/1-815</i>	747	GQQL - - - - - GQGQQGYYP TSPQQ - - - - - SGQGQQGYDSSPYHVS	779
<i>tr A0A1B0Z3C8 Bx6/1-824</i>	750	GQQ - - - - - SGQGQQGYYP TSLWQPGQGQQPGQGQQGYDSSPYHVS	788
<i>tr A0A1G4P1V7 Bx7/1-824</i>	750	GQQ - - - - - SGQGQQGYYP TSLWQPGQGQQPGQGQQGYDSSPYHVS	788
<i>sp P10388 Dx5/1-848</i>	768	GQQLGQWLQPGQGQQGYYP TSLQQTGGQGQQSGQGQQGYSSYHVS	812
<i>sp P08489 Dx2/1-838</i>	758	GQQLGQWLQPGQGQQGYYP TSLQQTGGQGQQSGQGQQGYSSYHVS	802
<i>tr Q03872 Ax1/1-830</i>	795	A E H Q A A S L K V A K A Q Q L A A Q L P A M C R L E G G D A L L A S Q	830
<i>tr Q41553 Ax2/1-815</i>	780	A E H Q A A S L K V A K A Q Q L A A Q L P A M C R L E G G D A L L A S Q	815
<i>tr A0A1B0Z3C8 Bx6/1-824</i>	789	A E Y Q A A R L K V A K A Q Q L A A Q L P A M C R L E G S D T L S A R Q	824
<i>tr A0A1G4P1V7 Bx7/1-824</i>	789	A E Y Q A A R L K V A K A Q Q L A A Q L P A M C R L E G S D A L S A R Q	824
<i>sp P10388 Dx5/1-848</i>	813	V E H Q A A S L K V A K A Q Q L A A Q L P A M C R L E G G D A L S A S Q	848
<i>sp P08489 Dx2/1-838</i>	803	V E H Q A A S L K V A K A Q Q L A A Q L P A M C R L E G G D A L S A S Q	838

**Abbildung 38:** Darstellung des Sequenzvergleichs der HMW-GS Ax1 (Q03872), Ax2 (Q41553), Bx6 (A0A1B0Z3C8), Bx7 (A0A1G4P1V7), Dx5 (P10388) und Dx2 (P08489) mit Jalview [177]. Farbcodierung der Aminosäuren: blau, Alanin, Isoleucin, Leucin, Methionin, Phenylalanin, Tryptophan, Valin (hydrophob); gelb, Prolin (Strukturbrecher); grün, Glutamin, Serin, Threonin (polar, ungeladene Seitenketten); orange, Glycin; rot, Arginin, Lysin (basisch); türkis, Histidin, Tyrosin (polare Seitenkette, Wasserstoffbrücke); violett, Asparaginsäure, Glutaminsäure (sauer); HMW-GS, hochmolekulare Gluteninuntereinheit.