



Article **Impacts of Sourdough Technology on the Availability of Celiac Peptides from Wheat** α - and γ -Gliadins: In Silico Approach

Annick Barre 🗅, Hervé Benoist 🕩 and Pierre Rougé *🕩

UMR 152 PharmaDev, Faculté de Pharmacie, Institut de Recherche et Développement, Université Paul Sabatier, 35 Chemin des Maraîchers, 31062 Toulouse, France

* Correspondence: pierre.rouge.perso@gmail.com; Tel.: +33-069-552-0851

Abstract: Celiac peptide-generating α - and γ -gliadins consist of a disordered N-terminal domain extended by an α-helical-folded C-terminal domain. Celiac peptides, primarily located along the disordered part of α - and γ -gliadin molecules, are nicely exposed and directly accessible to proteolytic enzymes occurring in the gastric (pepsin) and intestinal (trypsin, chymotrypsin) fluids. More than half of the potential celiac peptides identified so far in gliadins exhibit cleavage sites for pepsin. However, celiac peptides proteolytically truncated by one or two amino acid residues could apparently retain some activity toward HLA-DQ2 and HLA-DQ8 receptors in docking experiments. Together with the uncleaved peptides, these still active partially degraded CD peptides account for the incapacity of the digestion process to inactivate CD peptides from gluten proteins. In contrast, sourdough fermentation processes involve other proteolytic enzymes susceptible to the deep degradation of celiac peptides. In particular, sourdough supplemented by fungal prolyl endoproteases enhances the degrading capacities of the sourdough fermentation process toward celiac peptides. Nevertheless, since tiny amounts of celiac peptides sufficient to trigger deleterious effects on CD people can persist in sourdough-treated bread and food products, it is advisable to avoid consumption of sourdoughtreated food products for people suffering from celiac disease. As an alternative, applying the supplemented sourdough process to genetically modified low gluten or celiac-safe wheat lines should result in food products that are safer for susceptible and CD people.

Keywords: gliadin; celiac disease; celiac peptide; HLA-DQ2; HLA-DQ8; pepsin; trypsin; prolyl oligopeptidase; prolyl endoprotease; bioinformatics

1. Introduction

Celiac disease (CD), an immune reaction to gluten proteins resulting in deleterious effects on genetically predisposed individuals, requires strict avoidance of gluten-containing foods [1]. In spite of the prescription of a gluten-free diet to people suffering from severe CD, the repeated unintended/accidental consumption of gluten proteins, e.g., as hidden gliadins and glutenins occurring in food products, can result in life-threatening complications in CD subjects. Therefore, the consumption of gluten-free foods is of paramount importance to ensure a strict avoidance of gluten proteins, especially gliadins. According to WHO guidelines on "Standard for Foods for Special Dietary Use for Persons Intolerant to Gluten" from the Codex alimentarius (http://www.fao.org/fao-who-codexalimentarius, accessed on 8 November 2022), foods labeled as gluten-free must contain $\leq 20 \text{ mg/kg}$ (20 ppm) of gluten protein [2].

Gluten corresponds to a complex mixture of proteins, predominantly represented by two groups of proteins: gliadins and glutenins. Gliadins from wheat (Triticum aestivum) and other gluten-containing cereals, such as spelt (Triticum spelta), kamut (Triticum turanicum), triticale (Triticum \times Secale hybrid), rye (Secale cereale) and barley (Hordeum vulgare), consist of single polypeptide chains of 30–75 kDa MW, characterized by a high content in glutamine (Q) and proline (P). Wheat gliadins have classically been divided into



Citation: Barre, A.; Benoist, H.; Rougé, P. Impacts of Sourdough Technology on the Availability of Celiac Peptides from Wheat α - and γ -Gliadins: In Silico Approach. *Allergies* **2023**, *3*, 39–57. https:// doi.org/10.3390/allergies3010004

Academic Editors: Carmen Cuadrado and Rosario Linacero

Received: 9 November 2022 Revised: 29 December 2022 Accepted: 1 February 2023 Published: 3 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). four groups of (α) alpha-, (β) beta-, (δ) delta-, (γ) gamma- and (ω) omega-gliadins, which differ in their amino acid sequences and their electrophoretic behavior [3]. Alpha-gliadins and γ -giadins have been recognized for a long time as the most important providers of HLA-DQ2 and HLA-DQ8 immunodominant epitopes responsible for CD and other severe intolerances to gluten proteins.

Alpha-gliadins (α -gliadin) and gamma-gliadins (γ -gliadin), consist of proteins rich in glutamine and proline (P), composed of a disordered *N*-terminal part associated with an ordered C-terminal part, which contains α -helices, as predicted from the HCA plots of the proteins (Figure 1), and conformational analyses performed on partially hydrolyzed ggliadins [4]. Omega-gliadins (ω -gliadin) readily differ from other gliadins by a completely disordered structure. All these gliadins contain a variable number of either identical or different potential celiac peptides. These immunotoxic peptides are systematically located in the disordered N-terminal part of α -, γ - and ω -gliadins (Figure 1). However, although ω -gliadins exhibit a high content of Q and P residues and contain immunodominant epitopes for celiac disease, they do not apparently activate CD but play a prominent role in IgE-mediated wheat-dependent exercise anaphylaxis [5].

Although the conformational features of both gliadins and glutenins have been poorly investigated, essentially due to the lack of consistent and reproducible crystals allowing the structure of gluten proteins to be solved by X-ray crystallographic approaches, some physicochemical and modeling approaches have suggested an extended rod-shaped conformation, especially for gliadins [4]. In addition, the occurrence of an ordered a-helical C-terminal end in both molecules should allow homology modeling approaches in order to obtain some insight into their structural organization and, particularly, into the availability of their celiac peptides.

The availability of celiac peptides from the a- and g-gliadins depends on the proteolytic cleavage of their polypeptide chain into smaller fragments that contain the celiac peptide core of 9 amino acids. Gastric (pepsin) and intestinal (trypsin, chymotrypsin) proteases are currently involved in the digestive cleavage of gliadins. In addition, cleavage sites also occur inside many celiac peptide cores, which makes celiac peptides potentially sensitive to proteolytic inactivation. However, it is noteworthy that the digestive breakdown of gluten proteins is unable to prevent the formation of CD peptides from the gliadins and glutenins. Moreover, it is noteworthy that the occurrence of proline residues in CD peptides makes these peptides more resistant to attack by proteases in the upper digestive tract.

The sourdough technology, and especially sourdough bread making, implies a high number of proteolytic enzymes that are secreted by various microorganisms, especially *Lactobacillus* species [6]. Some of these proteases, especially the so-called prolyl peptidases, readily differ from the digestive proteases, pepsin, trypsin and chymotrypsin, by their ability to cleave the short peptides in even shorter peptides depending on the occurrence of proline-containing cleavage sites, which are more or less abundantly represented in CD peptides [7]. In this regard, sourdough fermentation of bread is known for a long time as a process susceptible to degrading gluten proteins and CD peptides into harmless fragments [8]. However, the degradation of gluten proteins during sourdough fermentation is often incomplete [9,10], and tiny amounts of residual CD peptides are sufficient to trigger deleterious effects on people suffering from severe CD [11]. Accordingly, the regular consumption of sourdough bread must be avoided as a precaution in people with severe celiac disease.

Here, we focus on in silico approaches suggesting that potential CD peptides are well exposed on structural models built for wheat α - and γ -gliadins, and are thus readily available for a proteolytic attack by the digestive enzymes pepsin and trypsin. However, since digestive proteases fail to inactivate CD peptides, we further focused on the in silico activity of sourdough technology associated with fungal proteases as a possible way to eliminate harmful CD peptides from bread and other gluten-containing food products. The goal is to render food products safer for susceptible people and people suffering from CD.

<u>alpha gliadin (α-gliadin)</u>:

VRVPVPQLQPQNPSQQQPQEQVPLVQQQQFPGQQQPFPPQQPYPQLQPFPSQQPYMQLQ<mark>PFP QPQLPYPQPQLPYPQ</mark>PQPFRPQQSYPQPQVSQPQQPISQQQQQQQQQQQQQQQLQQILQQILQQ LIPCRDVVLQQHSIAHGSSQVLQQSTYQLVQQLCCQQLWQIPEQSRCQAIHNVVHAIILHQQ QQQQQQQQQQQQPLSQVCFQQSQQQYPSGQGSFQPSQQNPQAQGSVQPQQLPQFEEIRNLAL ETLPAMCNVYIPPYCTIAPVGIFGTN



<u>gamma3 gliadin (γ3-gliadin):</u>



<u>omega gliadin (ω-gliadin)</u>:

ARELNPSNKELQSPQQSFSHQQQPFPQQPYPQQPYPSQQPYPSQQPFPTPQPQFPQQSQQPF TQPQQPTPLQPQQPFPQQPQQPQQ<mark>PFPQPQQPFPW</mark>QPQQPFPQTQQSFPLQPQQPFPQQPQQ PFPQPQLQFPQQPEQIIPQQPQQPFLLESQQPFPQQPFPQPQQLIPMQPQQPFPQQSQQ SQQPFPGPQQLFPELQQPIPQQPRQPFPLQPQQPFPQQSQQPFPQQPHQPQQPYPQQQPYGS SLTSIGGQ



Figure 1. Amino acid sequences and HCA plots of representatives of the α -gliadin, γ -gliadin and ω -gliadin groups. CD peptides are highlighted in red and blue in the sequences, and are boxed in red and blue in the HCA plots.

2. Materials and Methods

The amino acid sequences of α -, δ - and ω -gliadins were taken from the non-redundant NCBI protein database (http://www.ncbi.nlm.nih.gov/pubmed), accessed on 18 May 2022. Multiple amino acid sequence alignment of gliadins carried out with CLUSTAL-X [12]. Amino acid sequences of immunotoxic peptides generated from α -gliadins were retrieved from the most recent list reported by Sollid et al. [13].

Hydrophobic Cluster Analysis (HCA) was performed to delineate the conserved secondary structure features along the amino acid sequences of gliadins [14]. Regions containing clusters of hydrophobic amino acid residues correspond to the ordered regions of gliadins, whereas no clusters occur along the disordered regions of gliadins. HCA plots were generated using the drawhca program (http://osbornite.impmc.upmc.fr/hca/hca-form.html) and the RPBS web portal (http://www.mobyle.rpbs.univ-paris-diderot.fr) [15], as accessed on 20 May 2022.

The program PeptideCutter (https://web.expasy.org/peptide_cutter/), accessed on 24 May 2022, was used to predict the potential cleavage of the modeled gliadins and α -gliadin-derived immunotoxic peptides by pepsin (pH 1.3) and trypsin (pH 8.5).

Homology modeling of the α - and γ 3-gliadin was performed using the YASARA Structure program [16]. The X-ray coordinates of different proteins from the protein databank (PDB) that provide the best partial alignments and display partial HCA profile similarities were used as templates for building the three-dimensional models of α - and γ -gliadins. They include the α -amylase/trypsin inhibitor from *Eleusine oracana* (PDB codes 1B1U and 1TMQ) [17,18], the α -amylase/trypsin inhibitor from *Triticum aestivum* (PDB code 1HSS) [19], the limit dextrinase inhibitor from *Hordeum vulgare* (PDB code 4CVW) [20], and the maltose/maltodextrin-binding periplasmic protein from Escherichia coli (PDB code 3OB4) [21]. Finally, a single hybrid model was built for each gliadin from the different previous models. PROCHECK, K [22], ANOLEA [23], and the calculated QMEAN scores [24,25] were used to assess the geometric and thermodynamic qualities of the ordered C-terminal parts of the three-dimensional models. For example, a single K136 residue of the C-terminal ordered region of the modelled δ -gliadin occurred in the non-allowed region of the Ramachandran plot, and using ANOLEA to evaluate the model, only 11 residues (out of 139) exhibited energy over the threshold value. Both residues are mainly located in the loop regions connecting the α -helices in the model. The calculated QMEAN score of the model yielded a rather acceptable value of 0.39. In addition, the disordered N-terminal part of the γ 3-gliadin has been tentatively modeled using the atomic coordinates of the periodic tryptophan protein 2-like protein (Z) associated with the 35S rRNA in the ribosome of Chaetomium thermophilum (PDB code 5OQL) [26,27], which contains disordered regions particularly rich in proline (P), acidic residues (D,E) and their amides (N,Q).

Electrostatic potentials were calculated with YASARA, using the classical parameter values for the water dielectric contant (4.0, 80.0) and the AMBER96 forcefield, and displayed on the molecular surface as red (electronegatively charged) and blue (electropositively charged) patches. The surfaces occupied by the potential CD toxic peptide stretches on the modeled gliadins were differently colored and displayed on the molecular surface of allergens with Chimera-X [28].

Active gliadin peptides were built with the editing program of Chimera [29], and energy minimization of the built peptides was performed using 1000 steps of steepest descent followed by 100 steps of conjugated gradient. Docking experiments of the gliadin peptides to the CMH-II basket of HLA-DQ2 (PDB code 1S9V) [30] and HLA-DQ8 (PDB code 4GG6) [31] were performed with the HDOCK server (http://www.hdock.phys.hust.edu.cn, accessed on 15 July 2022) [32,33]. Docking results exhibiting the best docking scores were retained and analyzed with Chimera-X to decipher the H-bonding network anchoring the peptides to the HLA-DQ2 and HLA-DQ8 basket and the complete or incomplete occupancy of the basket sub-sites by the amino acids of the gliadin peptides.

The 33 mer peptide LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF derived from the proteolytic attack of α -gliadin was built in its extended conformation with the editing program of Chimera [29] and minimized by two runs of 1000 steps of steepest descent followed by 100 steps of conjugate gradient.

A three-dimensional model for γ3-gliadin was retrieved from the AlphaFold Protein Structure Database (https://www.alphafold.com, accessed on 22 August 2022) [34,35].

Except for ω -gliadins which consist exclusively of a single disordered domain, α - and γ -gliadins are composed of a highly disordered *N*-terminal domain, extended by a *C*-terminal domain exhibiting the canonical α -helical fold of the cereal α -amylase inhibitors (α -AI). As shown in Figure 2, the HCA plot of the *C*-terminal domain of the γ 3-gliadin from wheat (*Triticum aestivum*) exhibits striking similarities with the corresponding HCA plots of α -AIs from ragi (*Eleusine coracana*), wheat (*T. aestivum*) and barley (*Hordeum vulgare*) (Figure 2).



Figure 2. Comparison of the HCA plot for γ 3-gliadin from wheat with the HCA plots from α -amylase inhibitors from ragi, wheat and barley, showing the structural homologies of the inhibitors with the ordered C-terminal part of the gliadin. Homologous cysteine and hydrophobic cluster residues are colored yellow and red or blue, respectively.

The three-dimensional models built up for the α - and γ 3-gliadin exhibit a very similar α -helical *C*-terminal domain associated with a more (α -gliadin) or less (γ 3-gliadin) incomplete disordered *N*-terminal domain (Figure 3A–D). Although largely hypothetical, the three-dimensional models built for both gliadins suggest that the potential celiac peptides are well exposed and thus accessible for proteolytic attack by digestive proteases (Figure 3E–H).

A very similar three-dimensional model of γ 3-gliadin, was retrieved from the AlphaFold Protein Structure Database, comprising a highly disordered *N*-terminal domain linked to a C-terminal domain possessing the canonical α -helical fold of the α -amylase inhibitor from wheat (*Triticum aestivum*) (Figure 4). The potential celiac peptides were similarly exposed, as compared with our modelled γ 3-gliadin.



Figure 3. (**A**,**B**). Ribbon diagram of the three-dimensional models built for α -gliadin (**A**) and γ -gliadin (**B**). The disordered N-terminal part and the ordered α -helical C-terminal part of the molecules are colored green and violet, respectively. Immunodominant epitopes P1 and P2 of α -gliadin are colored red and magenta, respectively. The four epitopes P1–P4 present in the disordered *N*-terminal part of γ 3-gliadin (**D**). (**E**,**F**). Mapping of immunodominant epitopes on the molecular surface of α -gliadin (**E**) and γ 3-gliadin (**F**). (**G**,**H**). Distribution of surface electrostatic potentials on the surfaces of α -gliadin (**G**) and γ 3-gliadin (**H**). Electropositively and electronegatively charged surfaces are colored blue and red, respectively. Neutral surfaces are colored white.



Figure 4. Ribbon diagram (**A**) of the three-dimensional model retrieved from the AlphaFold Protein Structure Databank (https://www.alphafold.com, accessed on 22 August 2022). The signal peptide, the disordered *N*-terminal domain and the α -helical *C*-terminal domain are colored yellow, green and violet purple, respectively. The four potential celiac peptides (1–4) are colored red. (**B**). Molecular surface of the three-dimensional model (colored sand), showing the exposure of the four potential celiac peptides 1–4 (colored red).

A large number of potential cleavage sites for pepsin and fewer cleavage sites for trypsin/chymotrypsin, are predicted to occur along the amino acid sequences of α - and γ 3-gliadin (Figure 5). Most sites are located in the ordered α -helical *C*-terminal domain of the polypeptide chains. However, the putative cleavage sites for pepsin concern the CD peptide of α -gliadin, whereas they locate outside the potential CD peptides in γ 3-gliadin.

<u>Alpha-gliadin (α-gliadin)</u>:



Figure 5. Localization of the putative cleavage sites for pepsin (red arrow) and trypsin (blue arrow) along the amino acid sequences of α -gliadin and γ 3-gliadin predicted with PeptideCutter. The disordered *N*- and *C*-terminal regions are colored green, and the C-terminal α -helical core is colored cyan. Potential immunodominant epitopes are in red and magenta letters highlighted in grey.

More than half, 20 vs. 17, out of the 37 potential celiac peptides identified in Sollid's list of gliadins [13], contain putative sites of cleavage for pepsin and a single DQ2.5- α 3-gliadin exhibits a putative cleavage site for trypsin/chymotrypsin, as predicted from the PeptideCutter program (Figure 6).

Depending on their shortening, the truncated 9 mer celiac peptides lacking up to 2 amino acids display a more or less reduced capacity to properly interact with the basket of HLA-DQ2 and HLA-DQ8 compared to intact peptides. Docking experiments of the DQ2-specific peptide PFPQPELPY (peptide no. 9 of Sollid's list) and its truncated form PFPQPEL to HLA-DQ2 (PDB code 1S9V) revealed striking differences in the number of hydrogen bonds connecting the peptide (10) and its truncated form (5) to the HLA-basket and in the corresponding docking scores (230 vs. 207) (Figure 7A–F, Table 1). In contrast, the promiscuous binding of this DQ2-specific peptide to HLA-DQ8 (PDB code 4GG6) revealed no difference between the intact (11 H-bonds, docking score of 245) and truncated forms (11 H-bonds, docking score of 252) (Figure 7G–L, Table 1).

02 DQ2.2-glia-alpha1: QGSVQPQQ 02_de1 DQ2.2-glia-alpha1_de1: EGSVQPQQ 02_de2 DQ2.2-glia-alpha1_de2: OGSVOPOE 02_de3 DQ2.2-glia-alpha1_de3: EGSVQPQE 03 DQ2.2-glia-alpha2: QYSQPEQP: 07 DQ2.5-glia-alpha1a DQ2-a-I, a9: PFPOPELP 08 DQ2.5-glia-alpha1b DQ2-a-III: PYPQPELP 09 DQ2.5-glia-alpha2 DQ2-a-II, a2: POPELPYP 10 DQ2.5-glia-alpha3 glia-a20: FRPEQPYPQ 11 DQ2.5-glia-gamma1 DQ2-y-I: POOSFPEOC 11_de DQ2.5-glia-gamma1_de: PQQSFPEQE 12 DQ2.5-glia-gamma2 DQ2-y-II, y30: IQPEQPAQ 13 DQ2.5-glia-gamma3 DQ2-y-III: QQPEQPYPC 13_de1 DQ2.5-glia-gamma3_de1: EQPEQPYPQ 13_de2 DQ2.5-glia-gamma3_de2: QQPEQPYPE 13_de3 DQ2.5-glia-gamma3_de3: TOPEOPYPE 14 DQ2.5-glia-gamma4a DQ2-y-IV: SOPEOEFPC 15 DQ2.5-glia-gamma4b DQ2-y-VIIC: PQPEQEFPQ 16 DQ2.5-glia-gamma4c DQ2-y-VIIa: QQPEQPFPQ 16_de DQ2.5-glia-gamma4c_de: EQPEQPFPC 17 DQ2.5-glia-gamma4d DQ2-y-VIIb: PQPEQPFCQ 17_de DQ2.5-glia-gamma4d_de: PQPEQPFCE 18 DQ2.5-glia-gamma4e: LQPEQPFPQ 18_de DQ2.5-glia-gamma4e_de: LQPEQPFPE 19 DQ2.5-glia-gamma5 DQ2-y-V: QQPFPEQPQ 19_de DQ2.5-glia-gamma5_de: EQPFPEQPQ 20 DQ2.5-glia-omega1 DQ2-w-I: PFPQPEQPF PQPEQPFPW 21 DQ2.5-glia-omega2 DQ2-w-II: 31 DQ8-glia-alpha1 DQ8-a-I: EGS<mark>FQPSQE</mark> 32 DQ8-glia-gamma1a DQ8-y-Ia: EQPQQPFPQ 32_de DQ8-glia-gamma1a_de: EQPEQPFPC 33 DQ8-glia-gamma1b DQ8-y-Ib: EOPOOPYPE 33_de DQ8-glia-gamma1b_de: COPEOPYPE 34 DQ8-glia-gamma2: PQQSFPEQE 36 DQ8.5-glia-alpha1 DQ8-a-I: FOPSOE 37 DQ8.5-glia-gamma1: POOSFPEOF

Figure 6. List of immunodominant peptides from α - and γ -gliadins determined by Sollid et al. [13], showing the putative pepsin (red exclamation point) and trypsin (blue exclamation point) cleavage sites predicted by PeptideCutter. Excised parts from the peptides are in red-boxed white letters. Peptides uncleaved by pepsin or trypsin are highlighted in green.



Figure 7. (A-C): (A). Docking of the DQ2-specific PFPQPELPY peptide (colored green) to the basket of HLA-DQ2 (PDB code 1S9V). The A and B domains forming the HLA-DQ2 basket are colored violet and pink, respectively. (B). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (C). Network of hydrogen bonds (dashed black lines) connecting the peptide to the amino acid residues of the A and B domains, labeled and colored violet and pink, respectively. (D-F): (D). Docking of the truncated PFPQPEL peptide (colored green) to the basket of HLA-DQ2 (PDB code 1S9V). The A and B domains forming the HLA-DQ2 basket are colored violet and pink, respectively. (E). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (F). Network of hydrogen bonds (dashed black lines) connecting the truncated peptide to the amino acid residues of the A and B domains, labeled and colored violet and pink, respectively. (G-I): (G). Docking of the DQ2-specific PFPQPELPY peptide (colored green) to the basket of HLA-DQ8 (PDB code 4GG6). The A and B domains forming the HLA-DQ8 basket are colored violet and pink, respectively. (B). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (C). Network of hydrogen bonds (dashed black lines) connecting the peptide to the amino acid residues of the A and B domains, labeled and colored violet and pink, respectively. (J-L): (J). Docking of the DQ2-specific truncated PFPQPEL peptide (colored green) to the basket of HLA-DQ8 (PDB code 4GG6). The A and B domains forming the HLA-DQ8 basket are colored violet and pink, respectively. (B). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (C). Network of hydrogen bonds (dashed black lines) connecting the truncated peptide to the amino acid residues of the A and B domains, labeled and colored violet and pink, respectively.

Docking experiments to HLA-DQ8 performed with the DQ8-specific peptide QGSVQ-PQQL and its truncated form QGSVQPQQ revealed no difference between both peptides: a similar number of hydrogen bonds (16 vs. 17) and similar docking scores (243 vs. 241) were obtained for the intact peptide (Figure 8G–I) and its truncated form (Figure 8J–L) (Table 1). Similarly, the promiscuous binding of this DQ8-specific peptide to HLA-DQ2 (PDB code 1S9V) revealed no difference in the number of hydrogen bonds (11 vs. 11) and docking score values (221 vs. 217) between the intact peptide (Figure 8A–C) and its truncated form (Figure 8D–F) (Table 1).



Figure 8. (**A–C**): (**A**). Docking of the DQ8-specific QGSVQPQQL peptide (colored green) to the basket of HLA-DQ2 (PDB code 1S9V). The A and B domains forming the HLA-DQ2 basket are colored violet and pink, respectively. (**B**). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (**C**). Network of hydrogen bonds (dashed black lines) connecting the peptide to the amino acid residues of the A and B domains, labeled and colored violet and pink, respectively. (**D–F**): (**D**). Docking of the truncated QGSVQPQQ peptide (colored green) to the basket of HLA-DQ2 (PDB code 1S9V). The A and B domains forming the HLA-DQ2 basket are colored violet and pink, respectively. (**D–F**): (**D**). Docking of the domains forming the HLA-DQ2 basket are colored violet and pink, respectively. (**E**). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (**F**). Network of hydrogen bonds (dashed black lines) connecting the A and B domains, labeled and colored violet and pink, respectively. (**B**). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (**F**). Network of hydrogen bonds (dashed black lines) connecting the truncated peptide to the amino acid residues of the A and B domains, labeled and

colored violet and pink, respectively. (**G**–**I**): (**G**). Docking of the DQ8-specific QGSVQPQQ peptide (colored green) to the basket of HLA-DQ8 (PDB code 4GG6). The A and B domains forming the HLA-DQ8 basket are colored violet and pink, respectively. (**B**). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (**C**). Network of hydrogen bonds (dashed black lines) connecting the peptide to the amino acid residues of the A and B domains, labeled and colored violet and pink, respectively. (**J**–L): (**J**). Docking of the DQ8-specific truncated QGSVQPQQ peptide (colored green) to the basket of HLA-DQ8 (PDB code 4GG6). The A and B domains forming the HLA-DQ8 basket are colored violet and pink, respectively. (**B**). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (**C**). Network of hydrogen bonds (dashed black lines) connecting the truncated peptide to the amino acid residues of the A and B domains, labeled and colored violet and pink, respectively. (**B**). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (**C**). Network of hydrogen bonds (dashed black lines) connecting the truncated peptide to the amino acid residues of the A and B domains, labeled and colored violet and pink, respectively. (**B**). Coulombic surfaces (acidic in red, basic in blue, neutral in white) of the basket. (**C**). Network of hydrogen bonds (dashed black lines) connecting the truncated peptide to the amino acid residues of the A and B domains, labeled and colored violet and pink, respectively.

Conversely, docking experiments performed with truncated peptides lacking 3 or 4 amino acid residues (see Figure 6), were constantly unsuccessful, most probably due to an exceeding number of unoccupied subsites in the HLA-DQ2 and HLA-DQ8 basket (results not shown).

Table 1. Hydrogen bonds and docking score values calculated for the DQ2-specific peptide PF-PQPELPY and its truncated form PFPQPEL and the DQ8-specific peptide QGSVQPQQL and its truncated form QGSVQPQQ upon docking to the basket of HLA-DQ2 (PDB code 1S9V) and HLA-DQ8 (PDB code 4GG6).

Peptide	HLA-DQ2/DQ8	Hydrogen * Bonds	Docking Score
PFPQPELPY	HLA-DQ2	10	230
PFPQPEL	HLA-DQ2	5	207
PFPQPELPY	HLA-DQ8	11	245
PYPQPEL	HLA-DQ8	11	252
QGSVQPQQL	HLA-DQ2	11	221
QGSVQPQQ	HLA-DQ2	11	217
QGSVQPQQL	HLA-DQ8	16	243
QGSVQPQQ	HLA-DQ8	17	241

* Some H-bondings correspond to multiple H-bonds that could be less energetically favorable than single H-bonds, depending on the alignment of the H donor and receptor, and the distance between both partners. This may explain why docking scores calculated for QGSVQPQQL (243 for 16 bonds) and QGSVQPQQ (241 for 17 bonds) are lower than those calculated for PFPQPELPY (245 for 11 bonds).

Results from the docking experiments suggest that celiac peptides lacking a limited number of amino acid residues and, especially, residues located at the *C*-terminal end of the peptide, could maintain some residual capacity to accommodate the basket of HLA-DQ2 and HLA-DQ8 and perhaps retain decreased activity toward people suffering from CD.

In addition, upon proteolysis by pepsin, a longer peptide of 33 amino acids has been identified, even though it contains putative cleavage sites for pepsin and is readily exposed, as shown in the three-dimensional model built up by homology modeling for the 33mer-containing α -gliadin (acess. AFX69628.1). This points out that the exposure character of the potential celiac peptides on the surface of gliadin is not the sole factor responsible for cleavage by pepsin. Most probably, other factors are also involved in this process, depending on the physiological conditions of gastric digestion, such as the intragastric pH evolution, the intermittent secretion of pepsin, the mixing of gastric contents during the digestive process, etc.

The 33 mer peptide derived from the a-gliadin exhibits potential cleavage sites for pepsin, as predicted with the PeptideCutter program of Expasy (Figure 9A). These cleavage sites are readily exposed on a three-dimensional model built up for the 33 mer peptide in the extended conformation and thus should be readily accessible for a proteolytic attack by pepsin (Figure 9B). In addition, potential cleavage sites for other proteases of the lactic bacteria and yeasts involved in sourdough fermentation, such as the prolyl-peptidase, also occur along the amino acid sequence of the 33 mer peptide, which could achieve the proteolytic breakdown of its celiac peptides and prevent the release of peptides with deleterious effects on people suffering from CD (Figure 9C).



Figure 9. (**A**). Potential cleavage sites for pepsin (numbered red arrows) predicted along the amino acid sequence of the 26 mer peptide derived from α -gliadin. (**B**). Three-dimensional model in the extended conformation built for the 33 mer peptide, showing the exposure of the potential cleavage sites for pepsin (numbered red lines). (**C**). Potential cleavage sites polyl peptidases predicted along the amino acid sequence of the 33 mer peptide. The two potential cleiac peptides of the 33 mer peptide are highlighted in yellow.

A similar situation exists for an unattacked 26 mer peptide resulting from the incomplete degradation of α -gliadin (Figure 10A). This peptide is well exposed on the molecular surface of α -gliadin but resists digestive proteolysis since it contains a single cleavage site for pepsin at the beginning of the sequence and no cleavage site for trypsin/chymotrypsin (Figure 10A). However, additional cleavage sites for prolyl-peptidases occur along the amino acid sequence of the peptide, which suggests possible degradation by the corresponding proteases during the sourdough fermentation process (Figure 10B,C).

Prolyl oligopeptidases, which comprise members of the S28 serine protease family, were used as efficient gluten degradation proteases to reduce gluten intake in CD subjects [36,37]. However, prolyl oligopeptidases have a limited impact on gluten peptides because they require an optimum pH between 7.0 and 8.0 to develop their proteolytic activity and are therefore almost inoperative during the gastric digestion of gluten proteins [38]. Other prolyl endoproteases isolated from different strains of *Aspergillus niger* have been identified as very efficient proteases for mitigating the gluten peptide content. Their cleavage sites are well distributed on the *N*-terminal disordered domains of the three-dimensional models built for the γ 3- and γ 5-gliadins but are very scarce in their *C*-terminal ordered domains (Figure 11). Accordingly, the CD peptide-containing domains of gliadins should be readily degraded into short harmless fragments in the presence of sourdough supplemented with fungal prolyl endoproteases (Figure 12).



Figure 10. (**A**). Three-dimensional model built by homology modeling for the 26 mer-containing a-gliadin. The disordered *N*-terminal domain and the α -helical *C*-terminal domain are colored green and violet purple, respectively. The well-exposed 26 mer peptide on the molecular surface of α -gliadin is colored red. (**B**). Potential cleavage sites for polyl-peptidase (numbered green arrows) predicted along the amino acid sequence of the 26 mer peptide. The 5 potential overlapping celiac peptides occurring along the 26 mer peptide are highlighted in yellow. (**C**). Three-dimensional model in the extended conformation built for the 26 mer peptide, showing the exposure of the potential cleavage sites for pepsin (numbered red line), and prolyl-peptidase (green lines).



Figure 11. (**A**). Localization of cleavage sites for prolyl endoprotease (red spots) exposed on the molecular surface of the *N*-disordered (colored tan) and *C*-ordered (colored light blue) domains of the three-dimensional model built for the γ 5-gliadin. (**B**). Localization of cleavage sites for prolyl endoprotease (red spots) exposed on the molecular surface of the *N*-disordered (colored tan) and *C*-ordered (colored light blue) three-dimensional model built by AlphaFold for g3-gliadin. The signal peptide is colored yellow.

02 DQ2.2-glia-alpha1: QGSVQP<mark>QQ</mark>L 02_de1 DQ2.2-glia-alpha1_de1: EGSVQPQQL 02_de2 DQ2.2-glia-alpha1_de2: QGSVQP<mark>QE</mark>L 02_de3 DQ2.2-glia-alpha1_de3: EGSVQPQEL 03 DQ2.2-glia-alpha2: QYSQPEQPI <mark>PFP</mark>QPEL<mark>PY</mark> 07 DQ2.5-glia-alpha1a DQ2-a-I, a9: 08 DQ2.5-glia-alpha1b DQ2-a-III: <mark>PYP</mark>QPEL<mark>PY</mark> 09 DQ2.5-glia-alpha2 DQ2-a-II, a2: PQPELPYPQ <mark>FR</mark>PEQP<mark>YP</mark>Q 10 DQ2.5-glia-alpha3 glia-a20: 11 DQ2.5-glia-gamma1 DQ2-y-I: PQQSF<mark>PEQQ</mark> 11_de DQ2.5-glia-gamma1_de: PQQSFPEQ 12 DQ2.5-glia-gamma2 DQ2-y-II, y30: IQPEQPAQ 13 DQ2.5-glia-gamma3 DQ2-y-III: OOPEOP<mark>YP</mark>C 13_de1 DQ2.5-glia-gamma3_de1: EQPEQP<mark>YP</mark>Q 13_de2 DQ2.5-glia-gamma3_de2: **OOPEOP**YPE 13_de3 DQ2.5-glia-gamma3_de3: QPEQP<mark>YPE</mark> 14 DQ2.5-glia-gamma4a DQ2-y-IV: SQPEQEFPQ 15 DQ2.5-glia-gamma4b DQ2-y-VIIC: <mark>P</mark>QPEQEF<mark>PQ</mark> 16 DQ2.5-glia-gamma4c DQ2-y-VIIa: QQPEQP<mark>FP</mark>C 16_de DQ2.5-glia-gamma4c_de: EQPEQP<mark>FP</mark>Q 17 DQ2.5-glia-gamma4d DQ2-y-VIIb: PQPEQPFCC 17_de DQ2.5-glia-gamma4d_de: PQPEQPFC 18 DQ2.5-glia-gamma4e: LOPEOP<mark>FP</mark>C 18_de DQ2.5-glia-gamma4e_de: LQPEQP<mark>FP</mark>E QQPFPEQPQ 19 DQ2.5-glia-gamma5 DQ2-y-V: 19_de DQ2.5-glia-gamma5_de: EQP<mark>FPEQP</mark>Q 20 DQ2.5-glia-omega1 DQ2-w-I: PEOPEOPE 21 DQ2.5-glia-omega2 DQ2-w-II: QPEQP<mark>FPW</mark> <mark>EGS</mark>FQPSQE 31 DQ8-glia-alpha1 DQ8-a-I: 32 DQ8-glia-gamma1a DQ8-y-Ia: EQP<mark>QQP</mark>FP<mark>Q</mark> 32_de DQ8-glia-gamma1a_de: EQPEQP<mark>FP</mark>Q EQP<mark>QQP</mark>YPE 33 DQ8-glia-gamma1b DQ8-y-Ib: 33_de DQ8-glia-gamma1b_de: QPEQP<mark>YPE</mark> 34 DQ8-glia-gamma2: PQQSF<mark>PEQE</mark> EGSFQP<mark>SQE</mark> 36 DQ8.5-glia-alpha1 DQ8-a-I: 37 DQ8.5-glia-gamma1: PQQSF<mark>PEQE</mark>

Figure 12. List of immunodominant peptides from α - and γ -gliadins, showing the peptide fragments excised by pepsin (highlighted red), trypsin (highlighted blue) and prolyl-endopeptidase (highlighted yellow). Peptides uncleaved by pepsin or trypsin are highlighted in green.

3. Discussion

The three-dimensional models built by homology modeling for α - and γ -gliadins are largely hypothetical, essentially for the *N*-terminal domain, which consists of a highly disordered and flexible coiled structure. Nevertheless, both models suggest that the potential CD peptides are located in regions of this disordered *N*-terminus that are well exposed and readily accessible to proteases, especially gastric (pepsin) and intestinal (trypsin, chymotrypsin) proteases. In this respect, the proteolytic cleavage of gliadins by pepsin and trypsin will allow the release of either long peptide fragments containing the CD peptides or shorter fragments that roughly correspond to the CD peptide cores, susceptible to being accommodated by the HLA-DQ2 and HLA-DQ8 antigens, thus triggering deleterious effects in people suffering from CD. In addition to this extra-CD peptide cleavage by pepsin, which releases fully functional CD peptides, an intra-CD peptide cleavage by pepsin can occur since a high proportion of potential CD peptides contains a readily accessible pepsin cleavage site. The resulting truncated peptides should be inactivated and thus unable to be properly accommodated by the CMH-II basket from HLA-DQ2 and HLA-DQ8 antigens. However, docking experiments performed with various truncated CD peptides suggest that peptides truncated by no more than two N- or C-terminal amino acids could retain some binding activity toward the HLA-DQ2 and/or HLA-DQ8 basket. Conversely, potential CD peptides truncated by more than two N- or C-terminal residues apparently lose their ability to properly accommodate the HLA-DQ2 and/or HLA-DQ8 basket. Moreover, a significant proportion of potential CD peptides from gliadins lack cleavage sites for pepsin and/or trypsin, and thus escape a possible cleavage by digestive proteases. Taken together, these predictions are in agreement with the well-known incapacity of the digestive attack of gliadins by pepsin and trypsin/chymotrypsin to prevent the release of CD peptides from gluten-containing foods and food products. In addition, discrepancies occur among individuals concerning the efficacy of pepsin to cleave some cleavage sites exposed on the surface of gliadins, depending most probably on other physiological factors affecting pepsin digestion, such as the pH value of the gastric content, the transient release of pepsin into the gastric fluid, the duration of gastric digestion, etc. Accordingly, an uncleaved 33 mer peptide resistant to pepsin proteolysis is systematically identified even though it contains well-exposed cleavage sites that seem readily accessible to pepsin [8]. Other long peptides resulting from incomplete digestion of gliadins by pepsin have been identified [9].

Sourdough technology has been known for a long time as a process capable of mitigating the harmful effects of CD of gluten-containing food products for people suffering from CD due to the occurrence in sourdough of diverse proteolytic enzymes, especially prolyl oligopeptidases, excreted by different strains of *Lactobacilli* [7,36–39]. Compared to digestive proteolysis by pepsin and trypsin, the proteolytic processes involved in sourdough fermentation are more diverse and effective in causing inactivation of the potential celiac peptides issued from gliadins. In particular, the occurrence of many proline residues in CD peptides allows their cleavage by prolyl-oligopeptidase, which generates shorter harmless peptides. However, the inactivation of gluten by prolyl oligopeptidase remains limited to intestinal digestion by trypsin because this protease requires a pH optimum between 7.0 and 8.0 to develop its proteolytic activity.

Hopefully, other prolyl endoproteases isolated from different strains of Aspergillus *niger* have been identified as very efficient proteases for mitigating the CD peptide content, active in the range between pH 3.0 and pH 7.0, compared to prolyl oligopeptidases [40–46]. The MALDI-TOF-MS experiments performed with various gluten peptides identified the predominant post-proline cleavage sites $P \downarrow F$, $P \downarrow Q$, $P \downarrow Y$, $P \downarrow L$, $P \downarrow T$ and $P \downarrow S$ for the *A. niger* prolyl endoprotease [42]. The CD peptide-harboring sequences readily exposed on the molecular surface of gliadins, which are insufficiently degraded by digestive proteases or in the presence of sourdough supplemented with prolyl oligopeptidases, become prime targets for prolyl endoproteases. Nevertheless, complete degradation requires very prolonged sourdough fermentation [41], which remains incompatible with conventional sourdough bread making. Moreover, the degradation of gluten proteins by Aspergillus niger prolyl endoprotease is highly dependent on meal composition [47]. Accordingly, even sourdough bread supplemented with fungal extracts usually contains less CD peptides and is well tolerated in many CD subjects [48–50]. Its low levels of CD peptides can be sufficient to trigger deleterious effects on people suffering from severe CD [1]. Thus, regular consumption by people suffering from severe CD should be avoided. Nevertheless, sourdough bread and other food products issued from sourdough fermentation in the presence of additional proteases from fungal extracts remain very interesting for people suffering from CD because their accidental consumption should not result in any deleterious effect. In addition, sourdough food products have been recognized as potential enhancers for the

recovery of intestinal inflammation in celiac patients at the early stage of the gluten-free diet [50].

In our in silico approaches, the sensitivity of gliadins to proteases has been predicted using three-dimensional models built by homology modeling, which are largely hypothetical because of the lack of X-ray or NMR-solved three-dimensional structures usable as templates to build the models. In addition, the ability of gliadins to become embedded in polymeric structures resulting from the sourdough bread-making process could mask some exposed epitopic regions that are usually exposed to monomeric gliadins. According to these limitations, our in silico predictions have to be interpreted with caution, even though they agree with experimental approaches and will have to be ascertained by proper experiments.

In parallel with sourdough technology, another promising approach is developing, based on the use of RNA interference silencing and CRISPR/Cas9 gene editing, to reduce the gluten content in wheat and wheat products [51–53]. RNA interference (RNAi) silencing has been used to silent γ -gliadins [54], α -gliadins [55], ω -gliadins [56], and downregulate the gliadin gene families [57]. The downregulation of three gliadin gene families resulted in a marked decrease (>90%) in the reactivity toward the R5-specific monoclonal antibody (moAb) and a 10–100-fold reduction of the HLA-DQ2 and HLA-DQ8 epitopes in T-cell stimulation tests [57]. The CRISPR/Cas9 gene editing is in development to produce wheat lines with fewer gluten genes. This technology has been used to edit α -gliadin and γ gliadin genes [58–61]. However, although they represent beneficial progress for subjects suffering from CD, the public acceptance of these transgenic foods and food products still remains a puzzling problem. In addition, low-gluten wheat lines may present some difficulties in being properly bread-integrated or processed. In this regard, new mutation breading strategies have been developed that focus on the removal of CD peptides from the gluten proteins while maintaining their food-processing properties. The selective targeting of potential CD peptides with CRISPR/Cas9 allows them to be switched into harmless peptides [51]. These forthcoming celiac-safe wheat lines should be beneficial for avoiding genetically susceptible people from developing celiac disease. Moreover, applying sourdough technology to food products prepared from these celiac-safe wheat lines should allow the discarding of trace amounts of CD peptides, which could remain, to make these food products even safer for people suffering from CD.

4. Conclusions

Despite the presence of cleavage sites for digestive proteases in the potential CD peptides of gliadins, the digestive proteolysis of these gluten proteins by pepsin and trypsin is unable to prevent the release of peptides with deleterious effects on subjects suffering from severe CD. By introducing an additional source of proteases, the sourdough bread-making process improves the cleavage of the gliadins and their potential CD peptides, thus making sourdough bread safer for patients with CD. However, the enhanced lowering of CD peptides in sourdough bread remains insufficient to fully protect people suffering from severe CD from deleterious effects since the repeated presence of tiny amounts of CD peptides in the diet is sufficient to deteriorate the intestinal villi. Accordingly, the consumption of foods and food products issued from sourdough technology should be avoided by people suffering from severe celiac sprue, but their accidental intake should not have serious consequences for these people.

Author Contributions: Conceptualization, P.R.; methodology, P.R.; software, P.R.; validation, A.B. and H.B.; resources, A.B.; data curation, H.B.; writing—original draft preparation, P.R.; writing—review and editing, P.R.; project administration, A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Raiteri, A.; Granito, A.; Giamperoli, A.; Catenaro, T.; Negrini, G.; Tovoli, F. Current guidelines for the management of celiac disease: A systematic review with comparative analysis. *World J. Gastroenterol.* 2022, 28, 154–176. [CrossRef] [PubMed]
- 2. Standard for Foods for Special Dietary Use for Persons Intolerant to Gluten. CXS 118-1979, Adopted in 1979. Available online: http://www.fao.org/fao-who-codexalimentarius (accessed on 20 July 2022).
- Barak, S.; Mudgil, D.; Khatkar, B.S. Biochemical and Functional Properties of Wheat Gliadins: A Review. *Crit. Rev. Food Sci. Nutr.* 2015, 55, 357–368. [CrossRef] [PubMed]
- Sahli, L.; Boire, A.; Solé-Jamault, V.; Rogniaux, H.; Giuliani, A.; Roblin, P.; Renard, D. New exploration of the γ-gliadin structure through its partial hydrolysis. *Int. J. Biol. Macromol.* 2020, 165, 654–664. [CrossRef] [PubMed]
- Morita, E.; Matsuo, H.; Chinuki, Y.; Takahashi, H.; Dahlström, J.; Tanaka, A. Food-dependent exercise-induced anaphylaxis— Importance of omega-5 gliadin and HMW-glutenin as causative antigens for wheat-dependent exercise-induced anaphylaxis. *Allergol. Int.* 2009, *58*, 493–498. [CrossRef] [PubMed]
- Thiele, C.; Grassl, S.; Gänzle, M. Gluten Hydrolysis and Depolymerization during Sourdough Fermentation. J. Agric. Food Chem. 2004, 52, 1307–1314. [CrossRef]
- Vermeulen, N.; Pavlovic, M.; Ehrmann, M.A.; Gänzle, M.G.; Vogel, R.F. Functional Characterization of the Proteolytic System of Lactobacillus sanfranciscensis DSM 20451 T during Growth in Sourdough. *Appl. Environ. Microbiol.* 2005, 71, 6260–6266. [CrossRef]
- Mandile, R.; Picascia, S.; Parrella, C.; Camarca, A.; Gobbetti, M.; Greco, L.; Troncone, R.; Gianfrani, C.; Auricchio, R. Lack of immunogenicity of hydrolysed wheat flour in patients with coeliac disease after a short-term oral challenge. *Aliment. Pharmacol. Ther.* 2017, 46, 440–446. [CrossRef]
- Shan, L.; Molberg, Ø.; Parrot, I.; Hausch, F.; Filiz, F.; Gray, G.M.; Sollid, L.M.; Khosla, C. Structural Basis for Gluten Intolerance in Celiac Sprue. Science 2002, 297, 2275–2279. [CrossRef]
- Shan, L.; Qiao, S.-W.; Arentz-Hansen, H.; Molberg, Ø.; Gray, G.M.; Sollid, L.M.; Khosla, C. Identification and Analysis of Multivalent Proteolytically Resistant Peptides from Gluten: Implications for Celiac Sprue. J. Proteome Res. 2005, 4, 1732–1741. [CrossRef]
- 11. Coto, L.; Mendia, I.; Sousa, C.; Bai, J.C.; Cebolla, A. Determination of gluten immunogenic peptides for the management of the treatment adherence of celiac disease: A systematic review. *World J. Gastroenterol.* **2021**, *27*, 6306–6321. [CrossRef]
- 12. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tool. *Nucleic Acids Res.* **1997**, *25*, 4876–4882. [CrossRef] [PubMed]
- 13. Sollid, L.M.; Tye-Din, J.A.; Qiao, S.-W.; Anderson, R.P.; Gianfrani, C.; Koning, F. Update 2020: Nomenclature and listing of celiac disease-related gluten epitopes recognized by CD4+ T cells. *Immunogenetics* 2020, 72, 85–88. [CrossRef] [PubMed]
- Gaboriaud, C.; Bissery, V.; Benchetrit, T.; Mornon, J.P. Hydrophobic cluster analysis: An efficient new way to compare and analyse amino acid sequences. *FEBS Lett.* 1987, 224, 149–155. [CrossRef] [PubMed]
- Callebaut, I.; Labesse, G.; Durand, P.; Poupon, A.; Canard, L.; Chomilier, J.; Henrissat, B.; Mornon, J.P. Deciphering protein sequence information through hydrophobic cluster analysis (HCA): Current status and perspectives. *Cell. Mol. Life Sci.* 1997, 53, 621–645. [CrossRef] [PubMed]
- 16. Krieger, E.; Koraimann, G.; Vriend, G. Increasing the precision of comparative models with YASARA NOVA-a self-parameterizing force field. *Proteins* **2002**, *47*, 393–402. [CrossRef]
- 17. Gourinath, S.; Alam, N.; Srinivasan, A.; Betzel, C.; Singh, T.P. Structure of the bifunctional inhibitor of trypsin and α-amylase from ragi seeds at 2.2 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* **2000**, *56*, 287–293. [CrossRef]
- Strobl, S.; Maskos, K.; Wiegand, G.; Huber, R.; Gomis-Rüth, F.X.; Glockshuber, R. A novel strategy for inhibition of α-amylases: Yellow meal worm α-amylase in complex with the Ragi bifunctional inhibitor at 2.5 Å resolution. *Structure* **1998**, *6*, 911–921. [CrossRef]
- 19. Oda, Y.; Matsunaga, T.; Fukuyama, K.; Miyazaki, T.; Morimoto, T. Tertiary and quaternary structures of 0.19 α-amylase inhibitor from wheat kernel determined by X-ray analysis at 2.06 Å resolution. *Biochemistry* **1997**, *36*, 13503–13511. [CrossRef]
- Møller, M.S.; Vester-Christensen, M.B.; Jensen, J.M.; Hachem, M.A.; Henriksen, A.; Svensson, B. Crystal Structure of Barley Limit Dextrinase-Limit Dextrinase Inhibitor (LD-LDI) Complex Reveals Insights into Mechanism and Diversity of Cereal Type Inhibitors. J. Biol. Chem. 2015, 290, 12614–12629. [CrossRef]
- 21. Mueller, G.A.; Gosavi, R.A.; Pomés, A.; Wünschmann, S.; Moon, A.F.; London, R.E.; Pedersen, L.C. Ara h 2: Crystal structure and IgE binding distinguish two subpopulations of peanut allergic patients by epitope diversity. *Allergy* **2011**, *66*, 878–885. [CrossRef]
- 22. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: A program to check the stereochemistry of protein structures. *J. Appl. Cryst.* **1993**, *26*, 283–291. [CrossRef]
- 23. Melo, F.; Feytmans, E. Assessing protein structures with a non-local atomic interaction energy. J. Mol. Biol. 1998, 277, 1141–1152. [CrossRef] [PubMed]

- 24. Benkert, P.; Biasini, M.; Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **2011**, *27*, 343–350. [CrossRef] [PubMed]
- Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 2006, 22, 195–201. [CrossRef] [PubMed]
- Cheng, J.; Kellner, N.; Berninghausen, O.; Hurt, E.; Beckmann, R. 3.2-Å-resolution structure of the 90S preribosome before A1 pre-rRNA cleavage. *Nat. Struct. Mol. Biol.* 2017, 24, 954–964. [CrossRef]
- 27. Cheng, J.; Kellner, N.; Berninghausen, O.; Hurt, E.; Beckmann, R. Author Correction: 3.2-Å-resolution structure of the 90S preribosome before A1 pre-rRNA cleavage. *Nat. Struct. Mol. Biol.* **2020**, *27*, 683. [CrossRef]
- 28. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Meng, E.C.; Couch, G.S.; Croll, T.I.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **2021**, *30*, 70–82. [CrossRef]
- 29. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera?A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [CrossRef]
- Kim, C.-Y.; Quarsten, H.; Bergseng, E.; Khosla, C.; Sollid, L.M. Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease. *Proc. Natl. Acad. Sci. USA* 2004, 101, 4175–4179. [CrossRef]
- Broughton, S.E.; Petersen, J.; Theodossis, A.; Scally, S.W.; Loh, K.L.; Thompson, A.; van Bergen, J.; Kooy-Winkelaar, Y.; Henderson, K.N.; Beddoe, T.; et al. Biased T Cell Receptor Usage Directed against Human Leukocyte Antigen DQ8-Restricted Gliadin Peptides Is Associated with Celiac Disease. *Immunity* 2012, *37*, 611–621. [CrossRef]
- 32. Yan, Y.; Zhang, D.; Zhou, P.; Li, B.; Huang, S.-Y. HDOCK: A web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res.* 2017, 45, W365–W373. [CrossRef] [PubMed]
- Yan, Y.; Tao, H.; He, J.; Huang, S.-Y. The HDOCK server for integrated protein–protein docking. *Nat. Protoc.* 2020, 15, 1829–1852. [CrossRef] [PubMed]
- 34. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef] [PubMed]
- Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 2022, 50, D439–D444. [CrossRef]
- 36. Gobbetti, M.; Rizzello, C.G.; Di Cagno, R.; De Angelis, M. How the sourdough may affect the functional features of leavened baked goods. *Food Microbiol.* **2014**, *37*, 30–40. [CrossRef]
- Ogilvie, O.J.; Gerrard, J.A.; Roberts, S.; Sutton, K.H.; Larsen, N.; Domigan, L.J. A Case Study of the Response of Immunogenic Gluten Peptides to Sourdough Proteolysis. *Nutrients* 2021, 13, 1906. [CrossRef]
- Shan, L.; Marti, T.; Sollid, L.M.; Gray, G.M.; Khosla, C. Comparative biochemical analysis of three bacterial prolyl endopeptidases: Implications for celiac sprue. *Biochem. J.* 2004, 383, 311–318. [CrossRef]
- Rizzello, C.G.; De Angelis, M.; Di Cagno, R.; Camarca, A.; Silano, M.; Losito, I.; De Vincenzi, M.; De Bari, M.D.; Palmisano, F.; Maurano, F.; et al. Highly Efficient Gluten Degradation by Lactobacilli and Fungal Proteases during Food Processing: New Perspectives for Celiac Disease. *Appl. Environ. Microbiol.* 2007, *73*, 4499–4507. [CrossRef]
- Stepniak, D.; Spaenij-Dekking, L.; Mitea, C.; Moester, M.; De Ru, A.; Baak-Pablo, R.; van Veelen, P.; Edens, L.; Koning, F. Highly efficient gluten degradation with a newly identified prolyl endoprotease: Implications for celiac disease. *Am. J. Physiol. Gastrointest. Liver Physiol.* 2006, 291, G621–G629. [CrossRef]
- De Angelis, M.; Cassone, A.; Rizzello, C.G.; Gagliardi, F.; Minervini, F.; Calasso, M.; Di Cagno, R.; Francavilla, R.; Gobbetti, M. Mecanism of degradation of immunogenic gluten epitopes from *Triticum turgidum* L. var. durum by sourdough lactobacilli and fungal proteases. *Appl. Environ. Microbiol.* 2010, 76, 508–518. [CrossRef]
- 42. Heredia-Sandoval, N.G.; de la Barca, A.M.C.; Islas-Rubio, A.R. Gluten degradation in wheat flour with Aspergillus niger prolyl-endopeptidase to prepare a gluten-reduced bread supplemented with an amaranth blend. *J. Cereal Sci.* 2016, 71, 73–77. [CrossRef]
- Shetty, R.; Vestergaard, M.; Jessen, F.; Hägglund, P.; Knorr, V.; Koehler, P.; Prakash, H.S.; Hobley, T.J. Discovery, cloning and characterization of proline specific prolyl endopeptidase, a gluten degrading thermos-stable enzyme from Sphaerobacter thermophiles. *Enzym. Microb. Technol.* 2017, 107, 57–63. [CrossRef] [PubMed]
- Sakandar, H.A.; Usman, K.; Imran, M. Isolation and characterization of gluten-degrading Enterococcus mundtii and Wickerhamomyces anomalus, potential probiotic strains from indigenously fermented sourdough (Khamir). LWT 2018, 91, 271–277. [CrossRef]
- 45. Yu, J.; Wu, J.; Xie, D.; Du, L.; Tang, Y.-J.; Xie, J.; Wei, D. Characterization and rational design for substrate specificity of a prolyl endopeptidase from Stenotrophomonas maltophilia. *Enzym. Microb. Technol.* **2020**, *138*, 109548. [CrossRef] [PubMed]
- 46. Garcia, V.S.; Bersanetti, P.A.; de Araújo Morandim-Giannetti, A. Peptidases production by fungi obtained from Manihot esculenta Crantz waste and its application in gluten hydrolysis. *Biocat. Agric. Biotechnol.* **2021**, *37*, 102184. [CrossRef]
- Montserrat, V.; Bruins, M.J.; Edens, L.; Koning, F. Influence of dietary components on Aspergillus niger prolyl endoprotease mediated gluten degradation. *Food Chem.* 2015, 174, 440–445. [CrossRef]
- Di Cagno, R.; De Angelis, M.; Auricchio, S.; Greco, L.; Clarke, C.; De Vincenzi, M.; Giovannini, C.; D'Archivio, M.; Landolfo, F.; Parrilli, G.; et al. Sourdough Bread Made from Wheat and Nontoxic Flours and Started with Selected Lactobacilli Is Tolerated in Celiac Sprue Patients. *Appl. Environ. Microbiol.* 2004, 70, 1088–1096. [CrossRef]

- Di Cagno, R.; Barbato, M.; Di Camillo, C.; Rizzello, C.G.; De Angelis, M.; Giuliani, G.; De Vincenzi, M.; Gobbetti, M.; Cucchiara, S. Gluten-free Sourdough Wheat Baked Goods Appear Safe for Young Celiac Patients: A Pilot Study. *J. Pediatr. Gastroenterol. Nutr.* 2010, *51*, 777–783. [CrossRef]
- 50. Calasso, M.; Vincentini, O.; Valitutti, F.; Felli, C.; Gobbetti, M.; Di Cagno, R. The sourdough fermentation may enhance the recovery from intestinal inflammation of coeliac patients at the early stage of the gluten-free diet. *Eur. J. Nutr.* **2012**, *51*, 507–512. [CrossRef]
- Jouanin, A.; Gilissen, L.J.W.J.; Boyd, L.A.; Cockram, J.; Leigh, F.J.; Wallington, E.J.; van den Broeck, H.C.; van der Meer, I.M.; Schaart, J.G.; Visser, R.G.F.; et al. Food processing and breeding strategies for coeliac-safe and healthy wheat products. *Food Res. Int.* 2018, 110, 11–21. [CrossRef]
- 52. García-Molina, M.D.; Giménez, M.J.; Sánchez-León, S.; Barro, F. Gluten free wheat: Are we there? *Nutrients* 2019, 11, 487. [CrossRef] [PubMed]
- 53. Jouanin, A.; Gilissen, L.J.; Schaart, J.G.; Leigh, F.J.; Cockram, J.; Wallington, E.J.; Boyd, L.A.; van den Broeck, H.C.; van der Meer, I.M.; America, A.H.P.; et al. CRISPR/Cas9 gene editing of gluten in wheat to reduce gluten content and exposure—Reviewing methods to screen for coeliac safety. *Front. Nutr.* **2020**, *7*, 51. [CrossRef] [PubMed]
- Gil Humanes, J.; Piston, F.; Hernando, A.; Alvarez, J.B.; Shewry, P.R.; Barro, F. Silencing of γ-gliadins by RNA interference (RNAi) in bread wheat. J. Cereal Sci. 2008, 48, 565–568. [CrossRef]
- 55. Becker, D.; Wieser, H.; Koehler, P.; Folck, A.; Mühling, K.H.; Zörb, C. Protein composition and techno-functional properties of transgenic wheat with reduced a-gliadin content obtained by RNA interference. *J. Appl. Bot. Food Qual.* **2012**, *85*, 23.
- Altenbach, S.B.; Chang, H.C.; Yu, X.B.; Seabourn, B.W.; Green, P.H.; Alaedini, A. Elimination of omega-1,2 gliadins from breade wheat (*Triticum aestivum*) flour: Effects of immunogenic potential and end-use quality. *Front. Plant Sci.* 2019, 10, 580. [CrossRef]
- 57. Gil-Humanes, J.; Pistón, F.; Tollefsen, S.; Sollid, L.M.; Barro, F. Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 17023–17028. [CrossRef] [PubMed]
- Sánchez-León, S.; Gil Humanes, J.; Ozuna, C.V.; Gimenez, M.J.; Sousa, C.; Voytas, D.; Barro, F. Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant Biotechnol. J.* 2018, 16, 902–910. [CrossRef]
- 59. Jouanin, A.; Schaart, J.G.; Boyd, L.A.; Cockram, J.; Leigh, F.J.; Bates, R.; Wallington, E.J.; Visser, R.G.F.; Smulders, M.J.M. Ourlook for coeliac disease patients: Towards bread wheat with hypoimmunogenic gluten by gene editing of α- and γ-gliadin gene families. *BMC Plant Biol.* 2019, 19, 333. [CrossRef]
- Jouanin, A.; Born, T.; Boyd, L.A.; Cockram, J.; Leigh, F.; Santos, B.A.C.M.; Visser, R.G.F.; Smulders, M.J.M. Development of the GlutEnSeq capture system for sequencing gluten gene families in hexaploidy bread wheat with deletions and mutations induced by γ-irradiation of CRISPR/Cas9. J. Cereal Sci. 2019, 88, 157–166. [CrossRef]
- Jouanin, A.; Tenorio-Berrio, R.; Schaart, J.G.; Leigh, F.; Visser, R.G.; Smulders, M.J. Optimization of droplet digital PCR for determining copy number variation of α-gliadin genes in mutant and gene-dited polyploid bread wheat. J. Cereal Sci. 2020, 92, 102903. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.