Inhibition of Gliadin-α2 Protein Using Novel Peptide Synthesis

Grant Proposal

Derek Desrosiers

Massachusetts Academy of Math and Science at WPI

Worcester, MA, USA

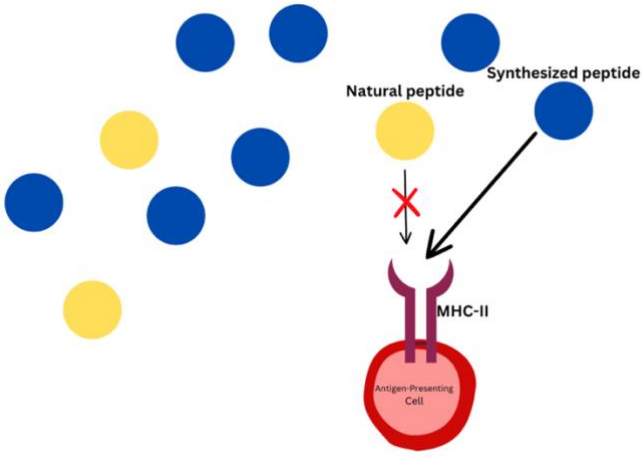
Abstract

Caused from an immune response to ingesting gluten proteins, celiac disease (CeD) most commonly relies on peptides found in wheat (Wei et al., 2020). When people with CeD eat gluten, their immune response attacks the small intestine, causing negative symptoms like abdominal pain, iron-deficiency anemia, and headaches. The only reliable treatment for CeD is a strict, lifelong gluten-free diet, however, little research has been done into alternative treatments. ~~In order to design a more accessible treatment to CeD,~~ the question was proposed of determining the effect of altered sequences of gluten proteins on binding strength and probability between the gluten-derived peptides and MHC-II molecules.

To test this question, novel peptide sequences based on gluten proteins will be made and tested in silico using the web-based software NetMHCIIpan-4.1. Testing will be done to determine the binding strength and binding likelihood between the peptide and MHC-II molecule. Due to unique properties of the peptide, such as its ability to competitively inhibit natural gliadin-based peptides, this drug can be the basis for individual-based peptide therapies if extended to test an individual's specific MHC-II complexes.

Keywords: peptide synthesis, CeD, TCR, gliadin-α2, MHC-II, NetMHCIIpan-4.1

Graphical Abstract



Inhibition of Gliadin-α2 Protein Using Novel Peptide Synthesis

Celiac Disease (CeD) is the most common food-induced heritable and life-long inflammatory disease in humans. Approximately 1% of the world's population is affected by CeD, however, the majority of patients with CeD remain undiagnosed due to large bias errors in testing (Singh et al., 2018; Hujoel et al., 2021). CeD is triggered upon ingesting wheat gluten or similar proteins found in cereals such as barley and rye (Wei et al., 2020). Gluten contains hundreds of proteins, each differing slightly in their amino acid sequences (Dahal-Koirala et al., 2019). In general, when people with celiac disease eat gluten, their body enacts an immune response that recognizes specific motifs in the gluten protein as foreign and attacks the small intestine. These attacks lead to damage on the villi - small fingerlike projections that line the small intestine that promote nutrient absorption (Celiac Disease Foundation, 2023b). Patients suffering from CeD have signs and symptoms of malabsorption, including diarrhea and weight loss or growth failure in children (Celiac Disease Foundation, 2023a).

Treatments and Pharmacological Research

Currently, the only treatment is a gluten-free diet. Eating gluten, even in small amounts, can damage the intestine and cause further negative effects. Past research has suggested that enzymatic treatments can be useful in degrading some peptides, however, these enzymes denature in the stomach and do not have enough time to fully digest gluten peptides (Wei et al., 2020). These enzymes also can only reduce the level of gluten in the immune system and not prevent certain proteins found in gluten from causing effects.

Gliadin-α2 Peptide

CeD relies on T-cell recognition (TCR) of a gluten-derived peptide bound to HLA-DQ2 or HLA-DQ8 - the genes most involved in CeD - which help the immune system identify agents that may be harmful (Anderson, 2022). These complexes are also called the MHC-II complexes. Little pharmacological research has been done into the role of peptides and their effect on T-cell recognition, especially regarding degradation of peptides because they are notoriously known for being non-degradable by most intestinal enzymes (Wei et al., 2020). When the gluten protein enters the body, it breaks down into gliadins and glutenins. Specifically, gliadins get transported into the stomach where they get digested. The problem with this sequence is that the entire gliadin protein does not get digested, and some sequences are left over. Digested and undigested peptide sequences get absorbed into the small intestine, and after modification by tissue transglutaminase (tTG), become a substance that is attacked by the immune system after binding to a MHC-II molecule and a T-cell binding to that molecule. However, the gliadin-α2 peptide is of utmost importance because it has the most reported binding in CeD patients and cannot be naturally broken down. Table 1, to the right, depicts a study conducted by Arentz-Hansen et al. (2000) where various tTG-treated peptides were measured for binding affinity. The table demonstrates that gliadin-α2 had the highest average binding affinity of all gliadin-alpha peptides across the seven T-cell clones (TCCs) and highlights its importance in creating a way to inhibit this reaction more than other gliadin-based peptides. While peptides are supposed to bind to these MHC-II molecules, T-cells recognize these peptide sequences as foreign and therefore create an autoimmune response damaging the patient's body.

TCC	+	α-1	α-2	α-3	α-4	α-5	α-6	α-7	α-8	α-9	α-10	α-11
412 R3	13.8	0.9	32.3	1.0	1.0	1.2	1.1	1.5	1.4	1.3	2.3	2.4
412 R5.32	15.7	1.2	19.0	0.9	1.1	1.2	1.2	1.1	1.2	0.9	1.6	0.9
370 R2.3	17.7	1.1	5.0	0.9	1.1	0.3	1.2	1.0	1.4	1.3	4.3	4.6
387 E9	52.6	1.0	25.0	ND	0.8	ND	1.2	1.2	6.1	25.3	23.7	19.4
387 E34	25.1	0.9	25.1	ND	1.0	ND	1.3	1.1	5.0	15.2	15.3	16.2
380 E3	2.4	1.0	2.2	ND	0.9	ND	1.3	1.1	1.1	1.0	0.6	1.0
389 E37	6.5	1.1	3.7	1.1	1.2	1.6	1.4	1.0	2.8	2.1	ND	ND

Table 1. Recognition of a Panel of tTG-treated Recombinant a-Gliadin Antigens by Seven TCCs. (Arentz-Hansen et al., 2000)

Hypotheses

In order to inhibit the binding of gliadin-α2 to the MHC-II molecule, the question was developed of determining the effect of different sequences of amino acids on binding strength and probability. The following hypotheses were formed to answer this question:

Hyp. 1a: If a peptide is modified by tTG, it will have a higher binding affinity to HLA-DQ2 because it will have an increased number of negatively-charged amino acids.

Hyp. 1b: If positions 4, 6, 8, and 9 are changed in a core peptide sequence then its binding affinity will decrease because these are the locations for binding to HLA-DQ2.

Hyp. 2: If a peptide sequence has a higher binding affinity and likelihood then it will competitively inhibit its natural peptide sequence because it will both bind stronger and be more likely to bind.

Methodology

To test this question, multiple mutations and novel amino acid sequences based on gliadin-α2 will be made and tested in silico using NetMHCIIpan-4.1. This web-based software predicts binding between peptides and the MHC-II complexes using tailored machine learning strategies (Reynisson et al., 2020). Testing will be done to determine binding strength and binding likelihood between each peptide and its associated MHC-II molecule. After determining baseline binding strength and likelihood scores, novel peptide sequences will be designed based on gliadin-derived peptides that will competitively inhibit the native peptide sequence by having both a stronger binding strength and likelihood.

Inhibition As a Treatment

Inhibition of the gliadin-α2 peptide can be a powerful step in creating a novel drug for CeD. Even though this project only inhibits one specific peptide, this process can easily be repeated for similar peptides. The research presented in this study will also show specific amino acid positions that are important for binding and sequences that repeat in other HLA-DQ2-binding peptides. In addition, this drug can be the basis for individual-based peptide therapies if extended to test an individual's specific MHC-II complexes. This treatment does not rely on the use of enzymes, therefore opening the possibility to inhibit a larger range of peptides as they are not restricted to being degraded before entering the stomach and small intestine.

Section II: Specific Aims

There are two specific aims of this study:

Specific Aim 1: determine which amino acid positions have the strongest effect on binding affinity.

Specific Aim 2: determine a sequence of amino acids that will most effectively inhibit the tTG-modified gliadin-α2 peptide.

The expected outcome of this work is the production of a peptide that will competitively inhibit gliadin-α2 by binding to the corresponding MHC-II molecule. Specific aim 1 will be used to develop specific aim 2, and as such, this project is reliant on both aims being successful. This work can be applied to various peptides with similar function to gliadin-α2 and will provide the basis for a novel drug to treat CeD.

Section III: Project Goals and Methodology

Relevance/Significance

CeD is one of the most common diseases caused by eating food, affecting over 70,000,000 people, including over 3,000,000 Americans, where 60-70% of Americans who have the disease are undiagnosed and needlessly suffering (Celiac Disease Foundation, 2016). Little research has been done into CeD treatment because it has long been considered that a gluten-free diet is the only option. However, with newly developed biotechnologies and knowledge, it is now possible to develop a drug that targets the pathology itself. The specific aims and methodology described below provides the basis for an individualized peptide-based therapy for the disease and does not affect any CeD blood tests or biopsies.

Innovation

Many current treatments of CeD rely on the use of enzymes to digest gluten before it enters the small intestine, but the problem with this is that these enzymes do not fully break down gluten and often degrade before entering the stomach. The proposed treatment does not rely on the use of enzymes and opens the possibility to inhibit a larger range of peptides as they are not restricted to being degraded before entering the stomach and small intestine.

While the methods and software used for this study are not novel, their combination for this specific purpose is an innovation for the field. Utilizing the vast resources available for protein and peptide sequences and their identification, it is now possible to perform an experiment on a peptide and make it more effective than what is naturally found in the body. In addition, many past studies have focused on the inhibition of the T-cell itself, which presents many additional problems, so inhibiting TCR by changing the peptide binding cleft is a more efficient and novel way to do so.

Methodology

Specific Aim #1: Determine which amino acid positions have the strongest effect on binding affinity.

Justification and Feasibility. Amino acid sequences in peptides can entirely alter a peptide's function. If one amino acid is changed in one specific position, it can either cause an increase in activation of the immune response or can stop any effect from happening at all. Therefore, analyzing which amino acid positions in a peptide sequence have the most effect on MHC-II binding affinity is of the utmost importance. To do this, the MHC-II molecule bound to a gliadin-α2 peptide will be searched in the Protein Data Bank database (RCSB PDB) and then a visual analysis will be conducted to determine which amino acid positions share a hydrogen bond with the MHC-II molecule. Any amino acids that share this bond must not be altered when making a novel peptide sequence, as it would prevent its binding to the same MHC-II molecule.

Figure 1, shown to the right, contains data from an experiment by Arentz-Hansen et al. (2000) to show how a difference in amino acid sequence effects binding affinity. IC_{50} represents the quantity of peptide needed to cause a 50% recognition by T-cells, where a lower number indicates a stronger binding affinity. This figure demonstrates how a tTG-treated peptide variant of gliadin-α2 has an increased binding affinity efficiency by over 50%, and provides support for the methods provided above and shows how important amino acid sequences are on binding affinity.

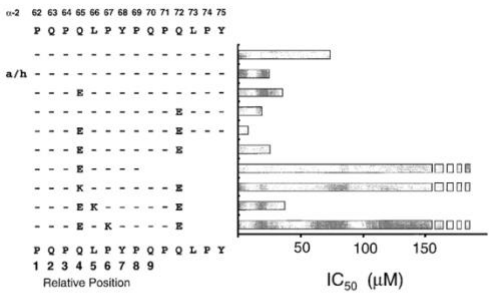


Figure 1: Analysis of peptide a-2(62–75) binding to DQ2. (Arentz-Hansen et al., 2000)

Expected Outcomes. This methodology will provide a specific set of amino acid positions that have a significant effect on binding affinity, meaning an alteration of that specific amino acid will impact the overall binding affinity. These findings will be used to further advance the study and provide support for specific aim 2.

Potential Pitfalls and Alternative Strategies. Another methodology to solve the aim identified above is to do an in vitro analysis of peptide sequences. While this method may be more accurate, it is much more expensive in cost and time. For the purposes of this study, an in silico analysis works more efficiently due to the sheer number of amino acid sequences being tested with various different MHC-II molecules.

Specific Aim 2: Determine a sequence of amino acids that will most effectively inhibit the tTG-modified gliadin-α2 peptide.

Justification and Feasibility. To produce a peptide that will competitively inhibit the gliadin-α2 peptide from binding to the MHC-II molecule, it is essential to make a peptide that binds to the same cleft but has a stronger binding affinity. The results from specific aim 1 come into play in this section; they are used as a basis to make a peptide that has the amino acids with the strongest binding affinity, as well as additional amino acid sequences on the tails of the peptide to cause a greater binding affinity. This part of the study is more of a trial and error, where it is about trying many different possible amino acid combinations until one is found to have the highest binding affinity. This peptide will then be synthesized and tested in vitro for T-cell proliferation and binding affinity to its MHC-II molecule.

Figure 2 (to the right) is based on a study by Xia et al. (2006) that depicts T-cell proliferation with gliadin-α2 and four modified peptides based on gliadin-α2. Specifically, three out of the four

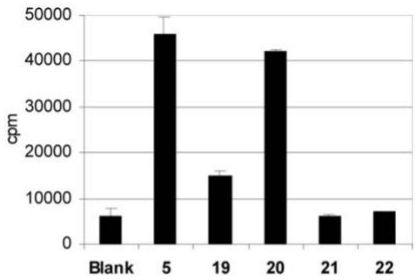


Figure 2: Comparison of T cell proliferation in the presence of modified peptides 19, 20, 21, and 22 against original peptide 5. (Xia et al., 2006)

peptides tested have a statistically significant difference in T-cell proliferation, providing evidence that modified peptides based on a natural sequence have a strong possibility of decreasing proliferation and the negative effects in a CeD patient.

Expected Outcomes. This method will produce a peptide with the possibility to become a novel drug for CeD that competitively inhibits the gliadin-α2 peptide by a stronger binding affinity. For this to work in real life, there would also have to be a large number of novel peptides in comparison to the natural peptide to truly block any possible TCR.

Potential Pitfalls and Alternative Strategies. An alternative method for this aim is to synthesize and test a novel chemical that uses the same binding cleft as its peptide counterpart but has a different peptide backbone, preventing T-cell recognition. However, this method has a much higher chance of either getting degraded before reaching the small intestine or to bind to another protein/molecule and can cause unwanted effects in patients.

Section III: Resources/Equipment

In Silico Analysis

By definition, this portion of the experiment exclusively uses computer-based software to run the experiment. A laptop with WiFi access and access to Microsoft Excel will be needed to access the software needed to run the tests and do to some basic data analysis and manipulation.

In Vitro Analysis

All of the work done for the in vitro analysis portion of this study will be done in a controlled lab environment. Directions will be followed based on a paper published by Yin & Stern (2014), to measure peptide binding affinity.

Section V: Ethical Considerations

Previous CeD studies focusing on creating a novel drug often utilized the inhibition of the T-cell instead of its binding to the MHC-II molecule. However, the main issue with those studies was that the drugs proposed either had no effect or a strong negative effect in patients as they caused overstimulation of the T-cells. Giving a patient a drug that has known issues is a large ethical issue, but the purpose of this study is to propose a theoretical solution that could turn into a drug if more testing were to be done. This study inhibits one specific peptide, and if a drug was to be made, more would likely have to be tested and added to the substance. If this project was continued, this drug would further have to meet rigorous testing in patients, and before that, ensure that the peptides presented in the drug do not interfere with any other functions in the intestines or body as a whole.

Section VI: Timeline

Early-mid November: Determining amino acid sequences and binding sites of MHC II molecules

Late November-late December: analysis of amino acid positions and effect of tTG on binding affinity and likelihood

Early-late January: novel design of peptides, starting with core sequence and adding terminal ends after

Section VII: Appendix

Appendix 1: HLA-DQ2.5-Bound Peptide Core Sequences

Peptide sequence	Epitope	Reference
PFPQPELPY	DQ2.5-glia-a1a	Arentz-Hansen et al. 2000
PYPQPELPY	DQ2.5-glia-a1b	Arentz-Hansen et al. 2002
PQPELPYPQ	DQ2.5-glia-a2	Arentz-Hansen et al. 2000
FRPEQPYPQ	DQ2.5-glia-a3	Vader et al. 2002
PQQSFPEQQ	DQ2.5-glia-y1	Sjostrom et al. 1998
IQPEQPAQL	DQ2.5-glia-y2	Qiao et al. 2005; Vader et al. 2002
QQPEQPYPQ	DQ2.5-glia-y3	Arentz-Hansen et al. 2002
SQPEQEFPQ	DQ2.5-glia-y4a	Arentz-Hansen et al. 2002
PQPEQEFPQ	DQ2.5-glia-y4b	Qiao et al. 2005
QQPEQPFPQ	DQ2.5-glia-y4c	Arentz-Hansen et al. 2002
PQPEQPFCQ	DQ2.5-glia-y4d	Qiao and Sollid 2019
LQPEQPFPQ	DQ2.5-glia-y4e	Qiao and Sollid 2019
QQPFPEQPQ	DQ2.5-glia-y5	Arentz-Hansen et al. 2002
PFPQPEQPF	DQ2.5-glia-w1	Tye-Din et al. 2010
PQPEQPFPW	DQ2.5-glia-w2	Tye-Din et al. 2010
PFSEQEQPV	DQ2.5-glut-L1	Vader et al. 2002
FSQQQESPF	DQ2.5-glut-L2	Stepniak et al. 2005; Vader et al. 2003
PFPQPEQPF	DQ2.5-hor-1	Tye-Din et al. 2010; Vader et al. 2003
PQPEQPFPQ	DQ2.5-hor-2	Vader et al. 2003
PIPEQPQPY	DQ2.5-hor-3a	Tye-Din et al. 2010
PYPEQPQPY	DQ2.5-hor-3b	Hardy et al. 2015b
PFPQPEQPF	DQ2.5-sec-1	Tye-Din et al. 2010, Vader et al. 2003
PQPEQPFPQ	DQ2.5-sec-2	Vader et al. 2003
PFPEQPEQI	DQ2.5-sec-3	Hardy et al. 2015b
PYPEQEEPF	DQ2.5-ave-1a	Arentz-Hansen et al. 2004; Vader et al. 2003
PYPEQEQPF	DQ2.5-ave-1b	Arentz-Hansen et al. 2004; Vader et al. 2003
PYPEQEQPI	DQ2.5-ave-1c	Hardy et al. 2015b

Section VIII: References

Anderson, J. (2022). *HLA-DQ2: The Primary Celiac Disease Gene*. Verywell Health. https://www.verywellhealth.com/hla-dq2-the-primary-celiac-disease-gene-562569

Arentz-Hansen, H., Körner, R., Øyvind Molberg, Hanne Quarsten, Vader, W., Kooy, Y., Knut E.A. Lundin, Koning, F., Roepstorff, P., Sollid, L. M., & McAdam, S. N. (2000). The Intestinal T Cell Response to α-Gliadin in Adult Celiac Disease Is Focused on a Single Deamidated Glutamine Targeted by Tissue Transglutaminase. *Journal of Experimental Medicine*, *191*(4), 603–612. https://doi.org/10.1084/jem.191.4.603

Celiac Disease Foundation. (2016). *20 Things You Might Not Know About Celiac Disease | Celiac Disease Foundation*. Celiac Disease Foundation; Celiac. https://celiac.org/about-the-foundation/featured-news/2016/08/20-things-you-might-not-know-about-celiac-disease/#:~:text=1%20in%20100%20people%20worldwide,diagnosed%20and%20are%20needlessly%20suffering.

Celiac Disease Foundation. (2023a). *Symptoms of Celiac Disease*. Celiac Disease Foundation. https://celiac.org/about-celiac-disease/symptoms-of-celiac-disease/

Celiac Disease Foundation. (2023b). *What is Celiac Disease?* Celiac Disease Foundation. https://celiac.org/about-celiac-disease/what-is-celiac-disease/

Dahal-Koirala, S., Ciacchi, L., Petersen, J., Louise Fremgaard Risnes, Ralf Stefan Neumann, Christophersen, A., Knut E.A. Lundin, Hugh Harrington Reid, Qiao, S., Rossjohn, J., & Sollid, L. M. (2019). Discriminative T-cell receptor recognition of highly homologous HLA-DQ2–bound gluten epitopes. *Journal of Biological Chemistry*, *294*(3), 941–952. https://doi.org/10.1074/jbc.ra118.005736

Hujoel, I. A., Jansson-Knodell, C., Hujoel, P. P., Margaux L.A. Hujoel, Rok Seon Choung, Murray, J. A., & Rubio–Tapia, A. (2020). Estimating the Impact of Verification Bias on Celiac Disease Testing. *Journal of Clinical Gastroenterology*, *55*(4), 327–334. https://doi.org/10.1097/mcg.0000000000001361

Reynisson, B., Alvarez, B., Paul, S., Peters, B., & Nielsen, M. (2020). NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Research*, *48*(W1), W449–W454. https://doi.org/10.1093/nar/gkaa379

Singh, P., Arora, A., Strand, T. A., Leffler, D. A., Catassi, C., Green, P. H., Kelly, C. P., Ahuja, V., & Makharia, G. (2018). Global Prevalence of Celiac Disease: Systematic Review and Meta-analysis.

Clinical Gastroenterology and Hepatology, *16*(6), 823-836.e2. https://doi.org/10.1016/j.cgh.2017.06.037

Wei, G., Helmerhorst, E. J., Darwish, G., Blumenkranz, G., & Schuppan, D. (2020). Gluten Degrading Enzymes for Treatment of Celiac Disease. *Nutrients*, *12*(7), 2095–2095. https://doi.org/10.3390/nu12072095

Xia, J., Siegel, M., Bergseng, E., Sollid, L. M., & Khosla, C. (2006). Inhibition of HLA-DQ2-Mediated Antigen Presentation by Analogues of a High Affinity 33-Residue Peptide from α2-Gliadin. *Journal of the American Chemical Society*, *128*(6), 1859–1867. https://doi.org/10.1021/ja056423o

Yin, L., & Stern, L. J. (2014). Measurement of Peptide Binding to MHC Class II Molecules by Fluorescence Polarization. *Current Protocols in Immunology*, *106*(1). https://doi.org/10.1002/0471142735.im0510s106