

## **Specificity of Tissue Transglutaminase Explains Cereal Toxicity in Celiac Disease**

L. Willemijn Vader,<sup>1</sup> Arnoud de Ru,<sup>1</sup> Yvonne van der Wal,<sup>1</sup>  
Yvonne M.C. Kooy,<sup>1</sup> Willemien Benckhuijsen,<sup>1</sup> M. Luisa Mearin,<sup>2</sup>  
Jan Wouter Drijfhout,<sup>1</sup> Peter van Veelen,<sup>1</sup> and Frits Koning<sup>1</sup>

<sup>1</sup>Department of Immunohematology and Blood Transfusion, and the <sup>2</sup>Department of Paediatrics, Leiden University Medical Centre, 2300 RC Leiden, Netherlands

### **Abstract**

Celiac disease is caused by a selective lack of T cell tolerance for gluten. It is known that the enzyme tissue transglutaminase (tTG) is involved in the generation of T cell stimulatory gluten peptides through deamidation of glutamine, the most abundant amino acid in gluten. Only particular glutamine residues, however, are modified by tTG. Here we provide evidence that the spacing between glutamine and proline, the second most abundant amino acid in gluten, plays an essential role in the specificity of deamidation. On the basis of this, algorithms were designed and used to successfully predict novel T cell stimulatory peptides in gluten. Strikingly, these algorithms identified many similar peptides in the gluten-like hordeins from barley and secalins from rye but not in the avenins from oats. The avenins contain significantly lower percentages of proline residues, which offers a likely explanation for the lack of toxicity of oats. Thus, the unique amino acid composition of gluten and related proteins in barley and rye favors the generation of toxic T cell stimulatory gluten peptides by tTG. This provides a rationale for the observation that celiac disease patients are intolerant to these cereal proteins but not to other common food proteins.

Key words: HLA-DQ2 • deamidation • T cell • cereals • algorithm

### **Introduction**

Celiac disease (CD) is a permanent intolerance to gluten (1). Gluten is a complex mixture of storage proteins found in wheat. Similar proteins are present in other grains like barley, rye, and oats. Typical disease symptoms include chronic diarrhea, fatigue, and failure to thrive. These symptoms are associated with a lesion in the upper small intestine, characterized by a (sub) total villous atrophy, an increased number of intraepithelial lymphocytes, and a chronic inflammatory response of the lamina propria lymphocytes (1). A strong indication for CD is the presence of antibodies specific for the enzyme tissue transglutaminase (tTG) in patients (2). Moreover, recent studies demonstrated the involvement of this enzyme in the generation of T cell stimulatory peptides (3, 4). CD is strongly associated with HLA-DQ2 (A1\*0501, B1\*0201) and HLA-DQ8 (A1\*0301, B1\*0302) (5). In addition, HLA-DQ2 and/or HLA-DQ8 restricted, gluten-specific CD4<sup>+</sup> T lymphocytes can be isolated from small intestinal biopsies of CD

patients, and these are thought to cause disease (6, 7). The peptide-binding motif of HLA-DQ2 and -DQ8 predicts a preference for negative charges at anchor positions in the bound peptides (8–10). However, gluten molecules contain few negative charges. This discrepancy was solved by the finding that tTG can convert glutamine residues into glutamic acid (a process termed deamidation), and thus introduces negative charges in gluten peptides. Indeed it was found that tTG is required for, or enhances the gluten-specific response of T cell clones and T cell lines derived from the majority of CD patients (3, 4, 11, 12). Considering the abundance of glutamine residues in gluten (~30–40%) tTG has many potential target sites in gluten. Mass spectral analysis of the deamidated T cell stimulatory gluten peptides, however, showed a highly restricted pattern of deamidation (3, 4) that often coincides with the positions where negative charges are preferred in the HLA-binding motif, thus favoring the binding of gluten peptides to HLA-DQ2 and/or -DQ8. The specificity of gluten deamidation by tTG, therefore, is a crucial factor in the generation of toxic gluten peptides but it is not known why only particular glutamine residues are targeted by tTG. This has been investigated in this study.

Address correspondence to F. Koning, Dept. of Immunohematology and Blood Transfusion, Leiden University Medical Centre, PO Box 9600, 2300 RC Leiden, The Netherlands. Phone: 31-0-71-526-6673; Fax: 31-0-71-521-6751; E-mail: F.Koning@lumc.nl

## Materials and Methods

**Peptides and tTG Treatment.** Peptides were synthesized by standard Fmoc chemistry on a multiple peptide synthesizer (SyrtoII, MultiSynTech GmbH). Integrity of synthetic peptides was checked by rpHPLC and mass spectrometry. tTG treatment was performed by incubating the peptides (500 µg/ml) with the enzyme (100 µg/ml, T-5398; Sigma-Aldrich) at 37°C for 4 h minimum, in 50 mM triethylamine-acetate, pH 6.5, 2 mM CaCl<sub>2</sub>. Throughout the manuscript the one letter code for amino acids is used.

**Mass Spectrometry.** Electrospray ionization mass spectrometry was performed on the synthetic gluten peptides before and after tTG treatment, using a Q-TOF hybrid mass spectrometer (Micromass). Precursor ions were selected with the quadrupole window set to 3 Da and fragments were collected with high efficiency with the orthogonal time of flight mass spectrometer. The collision gas applied was argon (pressure  $4 \times 10^{-5}$  mbar) and the collision voltage  $\sim 30$  V. Deamidation of glutamine residues in synthetic gluten peptides was determined by mass spectrometric analysis as described previously (4). Deamidation of a glutamine residue results in an increase of 1 Da. These conversions were assigned to particular glutamine residues by comparison of the fragmentation spectra of tTG treated and nontreated peptides (4).

**Database Searching.** The program PeptideSearch was used for pattern searches (13). The patterns were composed on the basis of the newly identified specificity of tTG in combination with the previously identified HLA-DQ2 peptide-binding motif (8, 9). For database searching selected subsets of gluten, hordein, secalin, and avenin proteins were compiled from the Swiss Prot databank, SPTREMBL and PIR.

**T Cell Clones and T Cell Lines.** Gluten-specific T cell lines were generated from small intestinal biopsies of CD patients as described previously (4, 7, 14). All patients gave informed consent to the study, which was approved by the hospital ethics committee.

**T Cell Proliferation Assays.** Proliferation assays were performed in duplicate in 150 µl RPMI 1640 (GIBCO BRL) supplemented with 10% human serum in 96-well, flat-bottomed plates (Falcon) using  $10^4$  T cells stimulated with  $10^5$  irradiated PBMCs (3,000 rad) in the presence or absence of antigen (1–10 µg/ml). After 48 h at 37°C, cultures were pulsed with 0.5 µCi of <sup>3</sup>[H]thymidine and harvested 18 h thereafter as described previously (7).

## Results

**The Specificity of tTG.** Previously we have characterized an HLA-DQ8-restricted, T cell stimulatory gliadin peptide (4, 7). In the core of this gliadin peptide (sequence PSGQ<sub>208</sub>GSFQPSQQ<sub>216</sub>NPQA) the Q<sub>208</sub> and Q<sub>216</sub> are deamidated by tTG, whereas positions Q<sub>212</sub>, Q<sub>215</sub>, and Q<sub>219</sub> are not (4). Preliminary studies indicated that the nature of COOH-terminal flanking amino acids influenced deamidation of glutamine residues (data not shown). We have now investigated this systematically by analyzing the deamidation in amino acid substitution analogs of a shorter version of the gliadin peptide (sequence: PSGQ<sub>208</sub>GSFQPSQQ<sub>216</sub>N). In this peptide only the Q<sub>208</sub> is deamidated by tTG treatment (Fig. 1 A). Mass spectrometric analysis demonstrated that at only two positions in the peptide substitutions had effect on deamidation of Q<sub>208</sub> (Fig. 1 A, and data not shown). First, the replacement of Q<sub>209</sub> by a P abolished deamidation

of Q<sub>208</sub>. Second, several amino acid substitutions at F<sub>211</sub> had a strong negative effect on deamidation of Q<sub>208</sub>, in particular replacements of F<sub>211</sub> for amino acids with significantly smaller or charged side chains or replacement by proline (Fig. 1 A).

Next, we performed a similar analysis to determine the influence of amino acid substitutions on the deamidation of Q<sub>216</sub> in a longer version of the peptide (Fig. 1 A). The results demonstrate that the substitution of N<sub>217</sub> with a P abolishes deamidation of Q<sub>216</sub>, a finding that underscores the observation that in a QP sequence the Q is not a target for tTG. Moreover, the substitution of Q<sub>219</sub> with a P had a negative effect on deamidation of Q<sub>216</sub>, confirming the negative influence of the P in the sequence QXXP (X stands for any amino acid). Other amino acid substitutions at Q<sub>216</sub>+3, however, had no effect on deamidation of Q<sub>216</sub>, a result that is in contrast to the strong influence of amino acid substitutions at position Q<sub>208</sub>+3 on deamidation of Q<sub>208</sub>. A marked difference between the sequence Q<sub>208</sub>GSF and Q<sub>216</sub>NPQ is the presence of a P in the latter sequence at Q+2 suggesting that in the sequence QXPX the Q is a good target for deamidation and the nature of the amino acid at Q+3 is much less important (see below).

Together the results indicate that in the sequences QP and QXXP the Q is not a target for tTG. In contrast the sequences QXP, QXXF(Y, W, M, L, I, or V) and QXPF(Y, W, M, L, I, or V) favor deamidation of the Q (Fig. 1 B). To verify these “patterns” for deamidation by tTG, a large panel of gluten peptides was synthesized, treated with tTG and analyzed by mass spectrometry (Table I). Without exception we found that a Q preceding a P was not deamidated by tTG while in the majority of cases a P at Q+3 also inhibited deamidation. In contrast, a bulky/hydrophobic residue at Q+3, and P at Q+2 allowed good deamidation. In the case that two opposite rules coincide, as is found in peptide 20 (LQPPQQ<sub>6</sub>PQFQPQQQF), the negative effect of QP↓ overruled the positive effect of QXXF↑ at position Q<sub>6</sub>. Together, these results confirmed the defined patterns. Proline thus plays a crucial role in the deamidation of gluten peptides.

**Prediction of Novel Gluten Epitopes.** The peptide-binding motifs of the disease-associated molecules HLA-DQ2 and HLA-DQ8 display preferential binding of negative charges at anchor positions (8–10). HLA-DQ2 has a strong preference for a negative charge at p4 and p7 (8, 9). Strikingly, at p6 a proline is one of the preferred amino acids in the HLA-DQ2 peptide binding motif that could thus target specific deamidation at p4 (Table II). Therefore, we combined the HLA-DQ2 peptide-binding motif with the specificity of tTG to construct a search algorithm that could be used to predict T cell stimulatory gluten peptides from the gluten database (Table II). In this algorithm glutamine residues were introduced at the putative DQ2 binding positions p4 and p7. These serve as potential target sites for tTG to allow introduction of a negative charge at these positions. Next, proline was selected for p6 to

A

Short	P	S	G	Q <sub>208</sub>	G	S	F	Q	P	S	Q	Q <sub>216</sub>	N			
Long	P	S	G	Q <sub>208</sub>	G	S	F	Q	P	S	Q	Q <sub>216</sub>	N	P	Q	A
G <sup>a</sup>				Q <sub>208</sub>	-	-	↓↓					Q <sub>216</sub>	-	-	-	-
A				Q <sub>208</sub>	-	-	↓↓					Q <sub>216</sub>	-	-	-	-
V				Q <sub>208</sub>	-	-	-					Q <sub>216</sub>	-	-	-	-
L				Q <sub>208</sub>	-	-	↓					Q <sub>216</sub>	-	-	-	-
I				Q <sub>208</sub>	-	-	-					Q <sub>216</sub>	-	-	-	-
P				Q <sub>208</sub>	↓↓	-	↓↓					Q <sub>216</sub>	↓↓		↓↓	
F				Q <sub>208</sub>	-	-	-					Q <sub>216</sub>	-	-	-	-
W				Q <sub>208</sub>	-	-	-					Q <sub>216</sub>	-	-	-	-
M				Q <sub>208</sub>	-	-	-					Q <sub>216</sub>	-	-	-	-
S				Q <sub>208</sub>	-	-	↓↓					Q <sub>216</sub>	-	-	-	-
T				Q <sub>208</sub>	-	-	↓					Q <sub>216</sub>	-	-	-	-
Y				Q <sub>208</sub>	-	-	-					Q <sub>216</sub>	-	-	-	-
N				Q <sub>208</sub>	-	-	↓↓					Q <sub>216</sub>	-	-	-	-
Q				Q <sub>208</sub>	-	-	↓					Q <sub>216</sub>	-	-	-	-
D				Q <sub>208</sub>	-	-	↓↓					Q <sub>216</sub>	↓	-	-	-
E				Q <sub>208</sub>	-	-	↓↓					Q <sub>216</sub>	-	-	-	-
R				Q <sub>208</sub>	-	-	↓↓					Q <sub>216</sub>	-	-	-	-
H				Q <sub>208</sub>	-	-	↓↓					Q <sub>216</sub>	-	-	-	-

B

Pattern	Effect on deamidation
QP ↓	no deamidation
QXXP ↓	no deamidation
QXP ↑	good deamidation
QXX(F,Y,W,M,L,I,V) ↑	good deamidation
QXP(F,Y,W,M,L,I,V) ↑	good deamidation

promote deamidation at Q<sub>4</sub>, which is also consistent with the requirements for peptide binding at this position (Table II). For the p<sub>9</sub> position two options were designed. In the first option bulky/hydrophobic residues were selected for p<sub>9</sub>, according to the peptide-binding motif. In the second option, a proline residue was placed at p<sub>9</sub>, followed by bulky/hydrophobic at p<sub>10</sub>. This option does not agree with the peptide-binding motif but should allow optimal deamidation at Q<sub>7</sub> according to our QXPF pattern. The remaining positions p<sub>1</sub>, p<sub>2</sub>, p<sub>3</sub>, p<sub>5</sub>, and p<sub>8</sub> were not defined (Table II). These two algorithms were used to search in the gluten database. A total of 98 matches were obtained, which corresponded to 27 unique gluten sequences. 13 of these were predicted by algorithm 1, the others by algorithm 2. These 27 peptides were synthesized and deamidated by tTG. Subsequent mass spectral analysis confirmed deamidation of the predicted Q<sub>4</sub> and Q<sub>7</sub> residues (data not shown). The native and deamidated peptides were pooled in pools of five peptides each and tested in T cell proliferation assays with three gluten-specific T cell lines isolated from small intestinal biopsies of CD patients. Two of these T cell lines responded to stimulation by several of the deamidated peptide pools (Fig. 2 A, results shown for one of the T cell lines). Strong T cell re-

**Figure 1.** The influence of amino acid substitutions on deamidation of Q<sub>208</sub> and Q<sub>216</sub> in an HLA-DQ8-restricted gliadin peptide. (A) Overview of the influence of COOH-terminal amino acid substitutions on deamidation of Q<sub>208</sub> in the short version of the gliadin peptide and Q<sub>216</sub> in the long version of the peptide. <sup>a</sup>The amino acids G<sub>209</sub>, S<sub>210</sub>, F<sub>211</sub>, N<sub>217</sub>, and Q<sub>219</sub> were substituted for all amino acids except lysine and cysteine, because these residues could affect deamidation by formation of disulphide bridges and tTG driven cross-linking, respectively. The effect on deamidation was determined by mass spectrometry. Substitutions that resulted in major differences in deamidation are designated with arrows: ↓↓ decrease of 70–100%; ↓ decrease of 30–69%. Dash indicates no effect of the substitution; substitution at other positions in the peptide had only minor effects on deamidation of the Q<sub>208</sub> and Q<sub>216</sub> residues (data not shown). (B) Patterns for deamidation in the HLA-DQ8-restricted gliadin peptide. X stands for any amino acid.

sponses were only found against pools that contained peptides predicted by the second algorithm, X X X Q<sub>4</sub> X P<sub>6</sub> Q<sub>7</sub> X P<sub>9</sub> (Y, F, W, I, L)<sub>10</sub> (Fig. 2 A and B). Subsequently, the T cell line was tested against the individual peptides of the positive pools, and found to respond to eight out of 13 peptides tested (Fig. 2 B).

In addition we performed searches with a less strict variant of algorithm 2, X X X X<sub>4</sub> X P<sub>6</sub> Q<sub>7</sub> X P<sub>9</sub> (Y,F,W,I,L)<sub>10</sub>, which predicts the deamidation of only one Q residue. This search yielded 261 matches in the gluten database (Table III). 18 of these matches corresponded to a known T cell stimulatory  $\alpha$ -gliadin peptide (sequence QLQPF-PQPQLPY) (11, 12).

*Presence of T Cell Stimulatory Sequences in Rye, Barley, and Oats.* Next we used the three algorithms to perform database searches in the gluten-like hordeins from barley, secalins from rye, and avenins from oats. While the first two cereals are toxic for CD patients, the latter is not (15, 16). Strikingly, all three search algorithms yielded many matches in the available hordein and secalin sequences but none in the avenin sequences (Table III). Consequently, the predicted T cell stimulatory gluten peptide QQPF-PQQPQQPFPQ (Fig. 2) is present in hordeins and secalins but not in avenins (Table III). Additional database searches

**Table I.** Validation of the Deamidation Patterns

Peptides	Pattern			
	QP ↓	QXXF ↑	QXXP ↓	QXP ↑
1 VPVPQLQPQNPS <b>Q</b> Q <b>Q</b> P <b>Q</b> E <b>Q</b> VPL	100%	100%	100%	66%
2 PLV <b>Q</b> <b>Q</b> <b>Q</b> QFLG <b>Q</b> <b>Q</b> Q <b>Q</b> PFPPQ	100%	100%	100%	100%
3 YYPTSP <b>Q</b> QSGQG	-	-	-	-
4 <b>Q</b> <b>Q</b> <b>Q</b> GGYYPTSPQ	-	50%	-	-
5 SS <b>Q</b> VSFQPSQLN	100%	100%	-	-
6 FPQT <b>Q</b> QPQQLFPQSQ	100%	0%	100%	100%
7 <b>Q</b> <b>Q</b> <b>Q</b> GGYYPTSV <b>Q</b> <b>Q</b> SGQ	-	50%	-	-
8 GSVQP <b>Q</b> <b>Q</b> QLPQFEIR	100%	100%	100%	100%
9 FPQHHCNY <b>Q</b> <b>Q</b> QPQTFPQP	100%	-	100%	100%
10 QQPIQP <b>Q</b> <b>Q</b> FP <b>Q</b> <b>Q</b> QFF	100%	50%	100%	50%
11 QSG <b>Q</b> Y <b>Q</b> <b>Q</b> QPQMOTT	100%	-	100%	100%
12 ES <b>Q</b> <b>Q</b> <b>Q</b> QDEPQPF	100%	-	100%	-
13 FPQPED <b>Q</b> <b>Q</b> SQQSE	100%	-	-	-
14 LQQV <b>Q</b> <b>Q</b> GGP <b>Q</b> <b>Q</b> QPFPPQPQPF	100%	100%	66%	100%
15 QVQWP <b>Q</b> <b>Q</b> QPFPPQP <b>Q</b> QPF	100%	50%	100%	33%
16 PLLP <b>Q</b> <b>Q</b> PFPSQQEQPQF	100%	50% <sup>a</sup>	100%	100%
17 W <b>Q</b> <b>Q</b> QPPFSQEQPIL	100%	0% <sup>a</sup>	33%	100%
18 QSNLPQPA <b>Q</b> <b>Q</b> PFQPQVPQQP	100%	100%	-	66%
19 WFQPSQLNP <b>Q</b> <b>Q</b> QDQPQ	100%	-	100%	-
20 LQQP <b>Q</b> <b>Q</b> QPFQP <b>Q</b> <b>Q</b> QF	100%	50% <sup>a</sup>	100%	50%
21 V <b>Q</b> <b>Q</b> <b>Q</b> IPVVQPSIL	100%	33% <sup>a</sup>	100%	100%

The observed deamidation patterns in a set of 21 gluten peptides were compared with those predicted according to the deamidation patterns given in Fig. 1. The deamidation patterns are represented in the 4 columns. The percentages shown indicate whether the pattern is fully or partially applicable to the specific deamidation of Q residues in the listed gluten peptides. A bar (-) indicates the absence of that sequence in the peptide. Deamidated glutamine residues are shown in bold and underlined.

<sup>a</sup>QP inhibition of deamidation overrules the positive influence of QXXF.

with less strict algorithms revealed a general lack of matches in the avenin database as compared with the gluten, hordein, and secalin databases (Table III). Notably, no matches were found in the avenins with three variants of the predictive algorithm 2.

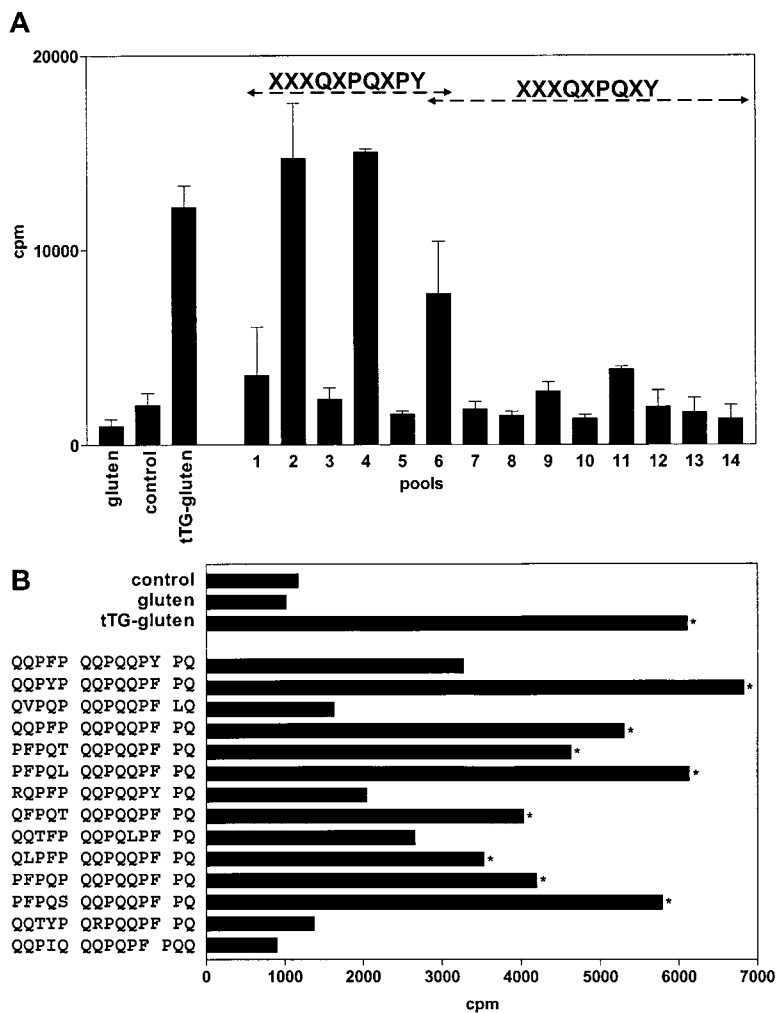
## Discussion

We have investigated the deamidation patterns of gluten peptides by tTG. The results demonstrate clear and selective deamidation patterns that to a large extent can be explained by strong effects of the spacing between the target

**Table II.** Design of Search Algorithms Based on the HLA-DQ2 Peptide-binding Motif and the tTG Specificity

DQ2 peptide-binding motif										
Residue	1	2	3	4	5	6	7	8	9	10
Preferred	F, W, Y, I, L, V			D, E, V, L, I		P, A, E	D, E		F, W, Y, I, L, V, M	-
DQ2 epitope search algorithms										
Algorithm 1	X	X	X	Q	X	P	Q	X	F, Y, W, I, L	X
Algorithm 2	X	X	X	Q	X	P	Q	X	P	F, Y, W, I, L

X stands for any amino acid.



**Figure 2.** Stimulation of gluten-specific T cell lines by predicted peptides. (A) A gluten-specific T cell line isolated from a small intestinal biopsy of a CD patient was tested against gluten, tTG-treated gluten, and against the predicted peptides that were present in pools of five peptides each (pools 1–14). The odd numbers represent pools that contain the native peptides, whereas the even numbers represent corresponding pools that were treated with tTG. Strong reactivity of the T cell line was only observed with tTG-treated gluten and tTG-treated pools of peptides predicted by the XXXQXPQXPY algorithm. Cpm indicates  $^3\text{H}$ thymidine incorporation. (B) Reactivity of the T cell line against the individual peptides from the T cell stimulatory pools. Bars indicated with an asterisk are considered positive (Stimulation Index > 3).

**Table III.** Database Searches with the Search Algorithms

Search algorithm	Number of matches in databases				
	Gluten	Hordein	Secalin	Avenin	
Algorithm 1	Q <sub>4</sub> XPQ <sub>7</sub> X(YFWIL)	52	48	8	-
Algorithm 1 – Q <sub>4</sub>	XXPQ <sub>7</sub> X(YFWIL)	334	91	14	6
Algorithm 1 – Q <sub>7</sub>	Q <sub>4</sub> XPXX(YFWIL)	286	142	60	11
Algorithm 2	Q <sub>4</sub> XPQ <sub>7</sub> XP(YFWIL)	46	60	33	-
Algorithm 2 – Q <sub>4</sub> <sup>a</sup>	XXPQ <sub>7</sub> XP(YFWIL)	262	196	89	-
Algorithm 2 – Q <sub>7</sub>	Q <sub>4</sub> XPXXP(YFWIL)	51	68	33	-
Short algorithm	QXP(YFWIL) X ≠ P <sup>b</sup>	>500	276	105	7
Predicted epitope	QQPFPQQPQQPFPQ	6	12	2	-

The search algorithms reveal many matches in the gluten, hordeins, and secalin sequences but very few in the avenin sequences. The predictive algorithm 2 (QXPQXP[YFWIL]) and the less strict variants of that algorithm have no matches in the avenin database. The novel T cells stimulatory gluten peptide (QQPFPQQPQQPFPQ) is also found in the hordeins and secalin but not in the avenins. X stands for any amino acid.

<sup>a</sup>Algorithm 2 without the specification of a Q at position 4 predicts the known gliadin  $\alpha$ 9 epitope.

<sup>b</sup>A proline at p2 in this algorithm inhibits deamidation of the glutamine.

glutamine residue and COOH-terminal proline residues. It has been described previously that a glutamine between proline residues is not modified by tTG (17). An important general role for proline in guiding tTG specificity, however, has not been appreciated. Proline, after glutamine, is the second most abundant amino acid in gluten (18) and thus the dominant factor in the selective deamidation of gluten peptides observed. In addition, the presence of a large hydrophobic amino acid three positions COOH terminally from the target glutamine has a strong positive influence on deamidation. This is highly significant since the sequence QXPF(Y) is one of the most frequently found sequences in gluten molecules and is predicted to be invariably deamidated by tTG. Moreover, cleavage by the enzyme pepsin is known to occur after phenylalanine and to a lesser extent after tyrosine, leucine, and isoleucine, and will thus generate many gluten peptides with the tTG substrate sequence QXPF(Y) at the COOH terminus. The sequential activity of pepsin and tTG, therefore, favors the generation of gluten peptides with an appropriate p7 anchor for binding to HLA-DQ2. We also demonstrate that by combining the HLA-DQ2 peptide-binding motif with two deamidation patterns an algorithm is obtained that predicts novel T cell stimulatory peptides in gluten. Strikingly, in this predictive algorithm (X X X Q<sub>4</sub> X P<sub>6</sub> Q<sub>7</sub> X P<sub>9</sub> [Y,F,W,I,L]<sub>10</sub>) we had incorporated a deamidation pattern that results in peptides that do not optimally fit the HLA-DQ2 peptide-binding motif, e.g., they lack an amino acid with a large hydrophobic side chain at position 9 in the peptide. The lack of this anchor is probably compensated for by the introduction of two negative charged anchor residues in the peptide at p4 and p7 as the result of deamidation. Thus, the T cell stimulatory activity of these peptides appears to depend more on optimal deamidation than on strict adherence to the exact HLA-DQ2 peptide-binding motif.

A less stringent search algorithm, combining only one deamidation rule with the HLA-DQ2 peptide-binding motif, proved equally successful since it predicted a previously identified T cell stimulatory  $\alpha$ -gliadin peptide (11, 12). Importantly, the first predictive algorithm identified 13 peptides, eight out of which stimulated T cells. The second, less stringent algorithm predicted 261 peptides. 18 of those (1 in 15) matched a known T cell stimulatory gliadin peptide. These algorithms, therefore, have a high predictive value. Therefore, it is striking that these algorithms identified very similar and sometimes identical peptides in the barley- and rye-derived hordeins and secalins but not in the oats-derived avenins. Oats is considered safe for CD patients (15, 16). While gliadins, hordeins, and secalins contain ~36% glutamine and 20% proline, avenins contain a similar percentage of glutamine (34%) but half the amount of proline (10%) (18). Deamidation of avenins is thus likely to occur in a much more random fashion. Indeed, the lack of proline residues in Q-rich regions leads to nonselective deamidation of the glutamine residues in such regions (data not shown). Collectively, our results indicate that due to nonselective deamidation no T cell stimulatory

neoepitopes can be generated from the avenins and thus offer an explanation for the clinical observation that oats is not toxic for CD patients. Moreover, our results indicate that the replacement of particular proline residues in gluten by other amino acids would yield gluten molecules with considerably less toxicity for CD patients but that would retain many of the chemical and physical properties desired in gluten.

tTG is a ubiquitous enzyme and could thus be involved in the modification of proteins at other sites in the body. Therefore, it will be of interest to investigate if the defined deamidation patterns and search algorithms can be used to identify novel T cell stimulatory peptides that are relevant for autoimmune diseases.

We thank Drs. F.H.J. Claas, R.R.P. de Vries, and R. Offringa for critical reading of the manuscript.

This study was supported by grants from the European Community (BHM4-CT98-3087 and QLK1-2000-00657) and the Dutch Digestive Disease Foundation (WS 98-24).

Submitted: 6 December 2001

Accepted: 14 January 2002

## References

1. Marsh, M.N. 1992. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology*. 102:330–354.
2. Dieterich, W., T. Ehnis, M. Bauer, P. Donner, U. Volta, E.O. Riecken, and D. Schuppan. 1997. Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nat. Med.* 3:797–801.
3. Molberg, O., S.N. McAdam, R. Korner, H. Quarsten, C. Kristiansen, L. Madsen, L. Fugger, H. Scott, O. Noren, P. Roepstorff, K.E. Lundin, H. Sjostrom, and L.M. Sollid. 1998. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat. Med.* 4:713–717.
4. van de Wal, Y., Y. Kooy, P. van Veelen, S. Pena, L. Mearin, G. Papadopoulos, and F. Koning. 1998. Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *J. Immunol.* 161:1585–1588.
5. Spurkland, A., G. Ingvarsson, E.S. Falk, I. Knutsen, L.M. Sollid, and E. Thorsby. 1997. Dermatitis herpetiformis and celiac disease are both primarily associated with the HLA-DQ ( $\alpha$ 1\*0501,  $\beta$ 1\*02) or the HLA-DQ ( $\alpha$ 1\*03,  $\beta$ 1\*0302) heterodimers. *Tiss. Antigenes.* 49:29–34.
6. Lundin, K.E., H. Scott, T. Hansen, G. Paulsen, T.S. Halstensen, O. Fausa, E. Thorsby, and L.M. Sollid. 1993. Gliadin-specific, HLA-DQ( $\alpha$ 1\*0501,  $\beta$ 1\*0201) restricted T cells isolated from the small intestinal mucosa of celiac disease patients. *J. Exp. Med.* 178:187–196.
7. van de Wal, Y., Y.M. Kooy, P.A. van Veelen, S.A. Pena, L.M. Mearin, O. Molberg, K.E. Lundin, L.M. Sollid, T. Mutis, W.E. Benckhuijsen, J.W. Drijfhout, and F. Koning. 1998. Small intestinal T cells of celiac disease patients recognize a natural pepsin fragment of gliadin. *Proc. Natl. Acad. Sci. USA.* 95:10050–10054.
8. Johansen, B.H., T. Jensen, C.J. Thorpe, F. Vartdal, E.

- Thorsby, and L.M. Sollid. 1996. Both  $\alpha$  and  $\beta$  chain polymorphisms determine the specificity of the disease-associated HLA-DQ2 molecules, with  $\beta$  chain residues being most influential. *Immunogenetics*. 45:142–150.
9. van de Wal, Y., Y.M.C. Kooy, J.W. Drijfhout, R. Amons, and F. Koning. 1996. Peptide binding characteristics of the coeliac disease-associated DQ( $\alpha$ 1\*0501,  $\beta$ 1\*0201) molecule. *Immunogenetics*. 44:246–253.
  10. Kwok, W.W., M.L. Domeier, F.C. Raymond, P. Byers, and G.T. Nepom. 1996. Allele-specific motifs characterize HLA-DQ interactions with a diabetes-associated peptide derived from glutamic acid decarboxylase. *J. Immunol.* 156:2171–2177.
  11. Arentz-Hansen, H., R. Korner, O. Molberg, H. Quarsten, W. Vader, Y.M. Kooy, K.E. Lundin, F. Koning, P. Roepstorff, L.M. Sollid, and S.N. McAdam. 2000. The intestinal T cell response to  $\alpha$ -gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J. Exp. Med.* 191:603–612.
  12. Anderson, R.P., P. Degano, A.J. Godkin, D.P. Jewell, and A.V. Hill. 2000. In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat. Med.* 6:337–342.
  13. Mann, M., and M. Wilm. 1994. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal.Chem.* 66:4390–4399.
  14. van de Wal, Y., Y.M. Kooy, P. van Veelen, W. Vader, S.A. August, J.W. Drijfhout, S.A. Pena, and F. Koning. 1999. Glutenin is involved in the gluten-driven mucosal T cell response. *Eur. J. Immunol.* 29:3133–3139.
  15. Janatuinen, E.K., P.H. Pikkarainen, T.A. Kempainen, V.M. Kosma, R.M. Jarvinen, M.I. Uusitupa, and R.J. Julkunen. 1995. A comparison of diets with and without oats in adults with celiac disease. *N. Engl. J. Med.* 333:1033–1037.
  16. Janatuinen, E.K., T.A. Kempainen, P.H. Pikkarainen, K.H. Holm, V.M. Kosma, M.I. Uusitupa, M. Maki, and R.J. Julkunen. 2000. Lack of cellular and humoral immunological responses to oats in adults with coeliac disease. *Gut.* 46:327–331.
  17. Pastor, M.T., A. Diez, E. Perez-Paya, and C. Abad. 1999. Addressing substrate glutamine requirements for tissue transglutaminase using substance P analogues. *FEBS Lett.* 451: 231–234.
  18. Wieser, H., W. Seilmeier, M. Eggert, and H.D. Belitz. 1983. Tryptophan content of cereal proteins. *Z. Lebensm. Unters. Forsch.* 177:457–460.