Amino Acids (2001) 21: 243–253

Amino
Acids

Structure-activity relationships in coeliac-toxic gliadin peptides

H. J. Cornell and **G. Wills-Johnson**

Department of Applied Chemistry, R.M.I.T. University, Melbourne, Victoria, Australia

Accepted December 11, 2000

Summary. Computer modelling studies of two groups of biologically-active peptides derived from A-gliadin indicated that the most likely structures were α-helical ones, in the case of serine-containing peptides, and random peptides coil types featuring β-turns, in the case of proline-rich, tyrosine-containing peptides.

The serine-containing group of peptides appear to be essentially cytotoxic in animal models of coeliac disease, whilst the tyrosine-containing group have the capacity to initiate damaging immunological reactions in patients with coeliac disease.

Both types of activity in coeliac disease are only possible if there is defective digestion of the active peptides, as mucosal digestion studies indicate. In the case of the serine-containing peptides, activity of the peptides is linked to the presence of PSQQ and also probably QQQP motifs. With the tyrosine-containing peptides, sequences such as QQPY and/or QPYP are associated with immunological activity and hence toxicity.

Keywords: Amino acid motifs – Coeliac disease – Wheat gliadin – Toxic peptides – Computer modelling

Introduction

Elucidation of the structures of the peptides in enzymic digests of wheat gliadin is vital to the understanding of coeliac disease (C.D.). Small peptides produced by digestion of wheat proteins have been shown to be toxic to individuals with (C.D.) (Cornell and Townley, 1974). Those derived from wheat gliadin appear to be the most toxic and since the elucidation of the sequence of an α-gliadin (A-gliadin) by Kasarda et al. (1984), studies by De Ritis et al. (1988) have shown the 4-mer amino acid motifs PSQQ and QQQP were present in toxic peptides from A-gliadin, but absent in non-toxic peptides. Since then, Kocna et al. (1991) have confirmed these observations using synthetic peptides based on the A-gliadin structure. They tested

dodecapeptides corresponding to residues 8–19, 45–56 and 208–219 in the foetal chick assay of Mothes et al. (1985) and found them to be active, insofar as they inhibited the biosynthesis of sucrase. The motif PSQQ, present in all of these active peptides, was shown to overlap with β-turns predicted by computer modelling.

Work with a fraction (Fraction 9) of an enzymic digest of wheat gliadin has pointed to the presence of two groups of active peptides: a) tyrosine-containing peptides, such as 75–86 of A-gliadin; b) serine-containing peptides, such as 7–20 of A-gliadin (Cornell et al., 1992). Peptides within these sequences have been synthesised and evaluated for activity using the foetal chick assay (Cornell and Mothes, 1993, 1995). Fraction 9, shown to be toxic "in vitro" and "in vivo" to patients with C.D., was used as the positive control in the assay.

Peptide 75–86 was shown to be active in the assay, but alone did not account for the high activity of Fraction 9. However, its activity was comparable to that of Fraction 9 in increasing production of γ-interferon in peripheral blood cultures, not only in blood from patients with C.D. but also in blood from normal individuals (Cornell et al., 1994). The results indicated that peptides such as 75–86 of A-gliadin were probably involved in cell-mediated reactions which have the potential to damage the enterocyte. It was suggested that a tyrosine-containing motif such as QQPY could be a feature of these peptides associated with toxicity, just as QQQP and PSQQ were shown to be associated.

Peptides containing the latter motifs such as 8–19 of A-gliadin have been shown to be almost as active as Fraction 9 in the foetal chick assay and when deletions of certain amino acids were made, it was shown that activity increased (Cornell and Mothes, 1995). Peptides 9–19 and 11–19 were observed to be considerably more active than 8–19. However, when a proline residue formed the N-terminus (peptide 10–19) or when the peptide became too small e.g. peptide 13–18 (PSQQQP) much of the activity was lost, despite the fact that this latter peptide contains both PSQQ and QQQP motifs. It is very important to distinguish between aetiology and pathogenesis in C.D. Aetiology, the basic cause of the disease, seems to be linked to defective mucosal digestion. Pathogenesis, the manner in which the disease is manifested through damage to small intestinal tissue , may be due to different mechanisms. The T-cell response is one of the cell-mediated mechanisms thought to be prevalent. This response is associated with HLA-DQ2 (histocompatibility type) although direct cytotoxic action also occurs and is predominant in the foetal chick (Mothes et al., 1985) and foetal rat assays (De Ritis et al., 1979).

It was therefore important to determine if these observations could be interpreted in terms of differences in secondary structure and an explanation given for the difference in the mechanism of toxicity of the tyrosine – as against the serine – containing peptides.

The present study seeks to bring together the evidence obtained from "in vitro" toxicity studies with structural studies in an attempt to explain why certain gliadin-derived peptides are much more active than others and why the mechanisms are different.

Materials and methods

Structures of the peptides were simulated using version 5.01 of the molecular modelling program HyperChem (Hypercube Inc., Gainesville, Florida). Various conformations in a molecular mechanics force field (MM+) were studied and the semi-empirical NDDO methods AM1, MNDO and PM3 were used for geometry optimisation.

Two groups of peptides were studied viz:

1) Those within the 8–19 sequence of A-gliadin (the serine-containing peptides).

and 2) Those within the 75–86 sequence of A-gliadin (containing tyrosine).

Elimination of amino acids from each of these peptides was carried out with studies of the changes in secondary structure and an assessment of how this relates to the known toxicity of each peptide. Boltzman statistics were applied in order to obtain restricted forms of each structure.

The parameters studied were:

- Energy figures calculated on both the bond-dipole and the electrostatic interaction models in the MM+ modification of the Alinger force field as implemented by HyperChem. The energy minimisation search used the Polak-Ribiere protocol on conformations obtained by simulated annealing. When exploring the potential energy surface, the minimum energy figure will have a zero gradient.
- Hydrogen bonding (H-bonding) in regard to the number of H-bonds able to be formed by the approach of the relevant groups. HyperChem indicates a H-bond if the distance between the donor hydrogen and acceptor atom is less than 3.2 Ångström units(Å) and the angle made by covalent bonds to the donor and acceptor atoms is greater than 150 degrees. Apparent H-bonds are stretched as the conformation of the molecule changes during optimisation (HyperChem, Computational Chemistry, 1996). The introduction of different amino acid residues into glycine helices was used to study the effect on stabilisation of the helix by H-bonding parallel to the axis, independently of H-bonding among groups in side chains.
- The distances between the nitrogen atom of the terminal amino group and the carbon atom of the terminal carboxylic acid group (N-C distances) determined for stable conformations of peptides starting from various cyclic structures.

Results

Serine-containing peptides

The program was able to create cyclic forms of peptides from the inactive hexapeptide 13–18 to the active nonapeptide 11–19 using the protocols mentioned. Convergence of all these peptides was obtained and many structures had low energy (−10 to −87 kcal/mol by bond dipole and −22 to −154 kcal/mol by electrostatic approaches). Bond dipole results covered a greater range of values and were thought to be more meaningful. N-C distances obtained were as low as 6 Å, but convergence was also seen with structures that had N-C distances about 35 Å. However, these structures depended upon the notion that they could be stabilised by an electrovalent bond between the N-terminal NH3+ and the C-terminal COO−. This is very unlikely, as when these groups draw close together, proton transfer was observed.

When COO− and NH3+ were brought to interact, the ring was twisted out of its plane and when the backbone was folded to obtain rings, unstable structures were produced. When optimization was begun, the first movement observed takes the COO− perpendicular to the plane of the ring.

Graphical plots of energy per atom (E/n) against toxicity of the peptides showed some degree of correlation. There was a tendency for E/n values by

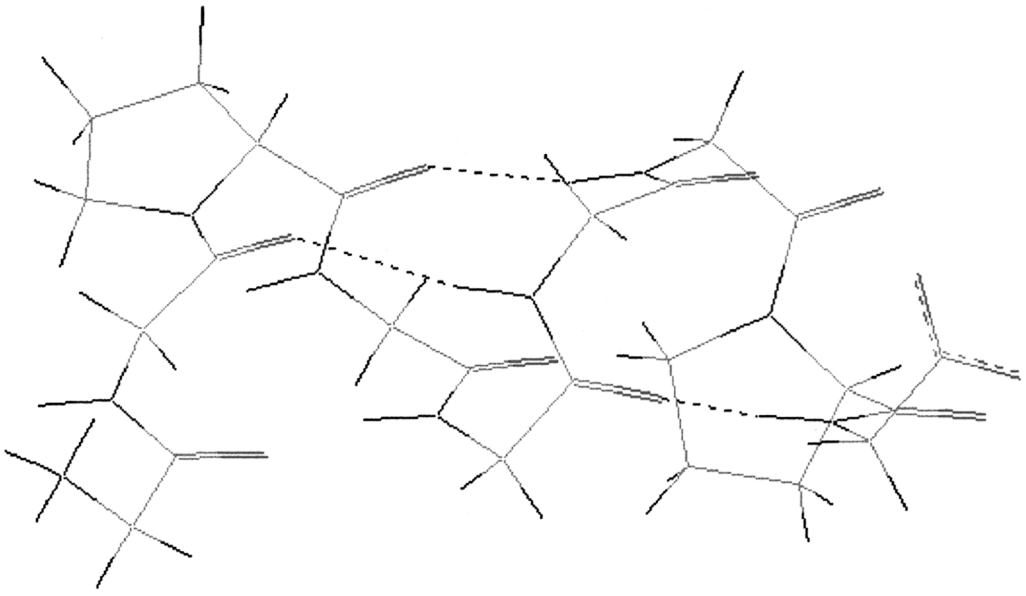

Fig. 1. A-gliadin fragment 11–19 showing three hydrogen bonds (dotted lines). For clarity, only the backbone is shown

both bond dipole and electrostatic energy calculations to be lower with the most toxic fractions and higher with the least toxic fractions.

When structures generated with the serine-containing peptides were subjected to annealing at simulation temperature 300 K, the energy figures were lowered to less than 100 k cal/mol (bond dipole approach) for the most toxic fractions. However, during this procedure, N-C distances generally increased from about 3–5 Å to 6–9 Å. Lowest energy figures (after annealing) generally appeared to be associated with the most toxic fractions. Two of the least toxic peptides (13–18 and 11–17) were notable for having significantly higher energy values than the other peptides. Some reliance can be placed upon these structures, but there is no absolute significance in the results as it depends on the severity of the criteria imposed on the structures.

It was seen that peptide 11–19 was also able to form an α-helical type structure which was stabilised by three H-bonds (ref. Fig. 1). If the same number (9) of glutamine residues was used, the structure was stabilised by five H-bonds. If the glutamine in the middle of this latter peptide was replaced with a proline, the number of hydrogen bonds was once again reduced to three and if the nonapeptide was formed entirely from proline, no hydrogen bonds were observed.

Table 1 shows that four H-bonds were present in the active peptides 8–19 and 9–19, although the most toxic peptide, 11–19, could form only three H-bonds. Toxicity was lost when the peptides became smaller with a corresponding decrease in the number of H-bonds, or when the peptide had proline at the N-terminus (e.g. 10–19). When glycine was substituted for the amino acids other than proline in peptides 11–19 and 12–19, three H-bonds were again observed. When glycine was substituted for all amino acids in an

Table 1. Number of proline residues, number of H-bonds and toxicity of various peptides derived from A-gliadin

Peptide	Number of proline residues	Number of H-bonds (backbone)	Toxicity ± SEM (100 = nil toxicity)
8–19*	3	4	60 ± 12
9–19	3	4	−6 ± 14
10–19	3 (incl.N-term.)	4	89 ± 17
11–19	2	3	−11 ± 7 (highest)
12–19	2	3	68 ± 11
13–19	2	2	not tested
11–18	2	2	95 ± 10
12–18	2	2	not tested
13–18	2 (incl.N-term.)	1	107 ± 17 (lowest)
11–17	2	2	91 ± 11

* The sequence of this peptide is LQPQNPSQQQPQ.

undecapeptide, seven hydrogen bonds can be formed, but when a proline residue is substituted for the glycine in the middle of the structure, only five H-bonds are observed (Table 2).

Tyrosine-containing peptides

The dodecapeptide 75–86 again showed convergence with energy values of about −40 kcal/mol by bond-dipole calculations. Many cyclic structures were thus able to be generated, but in general, convergence was more difficult to obtain, particularly with the low-toxicity peptides within this sequence, and N-C distances obtained (15–20 Å) were generally higher than those of the serine-containing peptides.

Annealing of the tyrosine-containing peptides at a simulation temperature of 300 K lowered the energy figures usually from positive figures to negative figures with corresponding increases in the N-C distance from 3–4 Å to 6–9 Å. No correlation was observed for the energy figures against toxicity of the peptides. Toxicity of peptides 75–86 and 75–85 are the highest in the series studied but activity is lost with peptide 76–85 which has an N-terminal proline residue (ref. Table 3). Much of the activity is regained in peptide 77–84 (N-terminal glutamine) but is again lost with the smaller peptide 77–82. Peptide 75–86 was stabilised by only one H-bond due to it being proline-rich. Because it contains five proline residues, it apears to have a random coil type structure featuring β-turns, quite different from that of peptide 11–19 (refer Fig. 2). No H-bonds were observed in any of the peptides smaller than peptide 76–85 or even with peptide 77–86 of the same size (refer Table 3). However, peptide 77–84 still retains activity.

Glycine substitution for amino acids other than proline also showed peptide 75–86 with only one H-bond and measurable activity in 77–84 without any H-bonds (refer Table 4).

Table 2. Hydrogen bonding in peptides within the sequence 11–19 of A-gliadin, but with glycine substituted for all amino acids except proline. Positions of N- and C-terminal amino acids in A-gliadin shown as superscript. The effect of proline on the number of H-bonds is reinforced with the example of an all glycine peptide. Toxicity figures are ones before substitution

Peptide	Number of H-bonds	Toxicity ± SEM (100 = nil toxicity)
^{11}GGPGGGGPG^{19}	3	−11 ± 7
^{12}GPGGGGPG^{19}	3	68 ± 11
^{12}GPGGGGP^{18}	2	not tested
^{13}PGGGGPG^{19}	2	not tested
^{13}PGGGGP^{18}	1	107 ± 17
All glycine 9–19	7	−6 ± 14
9–19 with P at 14	5	−6 ± 14

Table 3. Number of proline residues, number of H-bonds and toxicity of various peptides within the 75–86 sequence of A-gliadin

Peptide	Proline residues	H-bonds (backbone)	Toxicity ± SEM (100 = nil toxicity)
75–86*	5	1	57 ± 12
75–85	5	1	63 ± 9
76–86	5 (1 N-term.)	1	not tested
76–85	5 (1 N-term.)	1	114 ± 28 (lowest)
77–86	4	0	not tested
77–85	4	0	not tested
77–84	3	0	69 ± 13
77–82	2	0	96 ± 12

*The sequence of this peptide is RPQQPYPQPQPQ.

Discussion

Mucosal digestion aspects

Mucosal digestion experiments indicate that the active serine-containing peptides like 11–19 and the active tyrosine-containing peptides like 75–86 are incompletely digested by remission coeliac mucosa (Cornell and Rivett, 1995; Cornell, 1998). The residual peptides, such as peptides 11–18 and 77–84, are still toxic, which suggests that the aetiology of C.D. has to do with defective mucosal digestion and that the pathogenesis of the disease is the result of the action of the undigested peptides on the mucosa.

Table 4. Hydrogen bonding in peptides within the sequence 75–86 of A-gliadin, but with glycine substituted for all amino acids except proline. Positions of N- and C-terminal amino acids in A-gliadin shown as superscript. Toxicity figures are ones before substitution

Peptide	Number of H-bonds	Toxicity ± SEM (100 = nil toxicity)
^{75}GPGGPGPGPGPG^{86}	1	57 ± 12
^{75}GPGGPGPGPGP^{85}	1	63 ± 9
^{76}PGGPGPGPGPG^{86}	1	not tested
^{76}PGGPGPGPGP^{85}	1	114 ± 28
^{77}GGPGPGPGPG^{86}	0	not tested
^{77}GGPGPGPGP^{85}	0	not tested
^{77}GGPGPGPG^{84}	0	69 ± 13
^{77}GGPGPGP^{83}	0	not tested
^{77}GGPGPG^{82}	0	96 ± 12

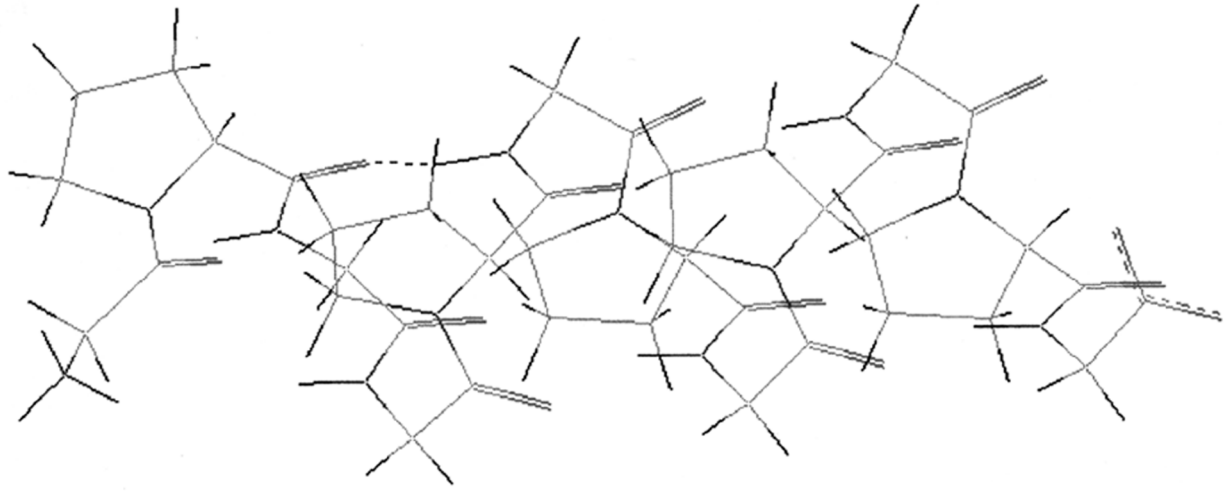

Fig. 2. A-gliadin fragment 75–86 showing one hydrogen bond (dotted line). For clarity, only the backbone is shown

It is important to note that the amino acid motifs associated with toxicity are present in the undigested (residual) peptides. In the case of peptide 11–19, the active residual peptide is NPSQQQPQ (12–19) and for 75–86 it is QQPYPQPQ (77–84). For the former, the hydrophilic nature of asparagine

and serine side chains, coupled with the β-turn caused by the proline at position 13, are likely to be key factors. For the latter, three proline residues at positions 79, 81 and 83 could distort the structure and minimise protease attack. However, this is not likely to be the only factor, as proline-rich regions are present in other parts of the A-gliadin structure which are digestible by coeliac mucosa. Coeliac mucosal digestion does not proceed as far as the inactive peptide 77–82. The latter contains only two proline residues, further supporting the notion that the 4-mer motif by itself is not bioactive. The enzyme deficiency in C.D. is likely to be due to only one enzyme but at least two different types of peptide residues build up and initiate damage to tissue.

In peptide 11–19, the key motifs PSQQ and QQQP overlap. However, peptide 13–18 is not active and has established that these sequences are not solely responsible for the high activity of peptides such as 9–19 and 11–19.

Hydrogen bonding

The serine-containing peptides from peptide 8–19 seem able to form as many as four H-bonds, which help to stabilise structures of the partial α-helical type. As amino acids are removed from 8–19, there comes a point at peptide 11–19 where the activity is maximal and three H-bonds are still able to be formed. Changes in the conformation of the peptide 8–19 would occur during digestion, which again, does not proceed past the hepta- and octa-peptide stage. Whilst these latter peptides are still toxic, they are not as active as peptides 9–19 and 11–19.

These observations are in agreement with those of Javadpour et al. (1996) who showed that the propensity to α-helical conformation of lysine-rich peptides in amphipathic media was proportional to their cytotoxicity to mouse fibroblasts. Their 14-mer and 21-mer peptides showed modest levels of helicity by circular dichroism spectroscopy but the 7-mer peptides were devoid of biological activity and of secondary structure. Thus, a more significant secondary structure also appears to be important for the activity of the gliadin peptides in C.D.

Hydrogen bonding thus needs to be considered, as well as peptide size, but the preservation of key motifs (PSQQ, QQQP and QQPY) and the types of amino acids which flank these motifs are important factors for retention of activity. These principles also hold in the case of peptides within the 75–86 section of A-gliadin. Even large sequences in the 75–86 peptide do not appear to be stabilised to any great extent by H-bonding. Their biological activity may instead relate to a random coil-type of structure with β-turns. Furthermore, these peptides resemble the more opioid structures and the evidence from Graf et al. (1987) supports the view that tyrosine-containing peptides with sequences such as YPQPQ form part of opioid receptors on peripheral blood lymphocytes. This sequence is incomplete in the inactive peptide 77–82 (QQPYPQ). The activity of peptide 75–86 in peripheral blood lymphocytes from coeliacs (and some normals) was detected by its ability to increase

production of γ-interferon, which may initiate damage to tissue. This is a typical example of a cell-mediated immune reaction.

The data obtained from amino acid substitution experiments confirm that proline disrupts hydrogen bonding and hence α-helical conformation, and once again the loss of activity is seen in a peptide (76–85) in which this amino acid is at the N-terminus. No peptides in this group had the extremely high activity in the foetal chick assay of peptides within the 8–19 section of A-gliadin, but in the γ-interferon assay, their activity was far in excess of any of the serine-containing peptides tested (Cornell et al. 1994).

In vivo, enhancement of this type of activity may be related to deamidation of specific glutamine residues. For example, the partial deamidation of gliadin peptides such as A-gliadin peptide 45–56 (which contains the QQPY motif and four proline residues) resulting from acidic environment in the stomach has been shown to lead to increased binding to the HLA-DQ2 molecule (Terreaux et al., 1998). This could be a mechanism which enhances the immunological reactions of pathogenesis in C.D. Similarly, the more specific deamidation of A-gliadin peptide 57–73 brought about by a tissue transglutaminase was shown to enhance the T-cell responses to this peptide (Anderson et al., 2000). The peptide studied contains a tyrosine and six proline residues.

β-Turns due to proline

According to Tatham et al. (1990), β-turns are a predominant structural feature of gliadin-derived peptides and may be involved in the pathogenesis of C.D. These are a predominant structural feature of peptide 75–86 and may create a spatial arrangement which becomes a recognition site for immunological reactions, such as stimulation of lymphocytes. With this particular site, hydrophobic reactions and electrostatic reactions also are likely to contribute. Folding of the chains caused by the proline residues interrupting H-bonding may be especially significant in the orientation of the tyrosine residue, which is essential to its opioid activity and ability to produce γ-interferon in cell-mediated reactions. Alfonso et al. (1998) have reported a role for β-turn motifs in recognition of anti-gliadin antibodies in C.D. A 20-amino acid synthetic peptide corresponding to the sequence from the N-terminal region of an oats peptide showed the highest β-turn content using circular dichroism and reactivity with the antibodies was comparable to that of a gliadin extract.

Cyclic forms of the peptides

The possibility that the peptides are present in cyclic forms, as are some antibiotics, is most unlikely as proton transfer occurred with both types of peptides when the N-terminal and C-terminal amino acids in the zwitter ion form came into close proximity. However, in the case of peptide 75–86, the H-bond (from the guanidinium group) seems to be able to hold the peptide in a type of cyclised structure, although it was shown to have "slipped" from the

original oxygen at the end of the backbone to an oxygen on the side chain of the last amino acid (Q).

We can conclude that the activity of the serine-containing and the tyrosine-containing toxic peptides appears to be due to different mechanisms; the former is due chiefly to direct cytotoxicity on the enterocyte and is enhanced by its partly α-helical conformation; the latter is due to cell-mediated reactions initiated by its random coil conformation.

References

Alfonso P, Soto C, Albar J, Camafeita E, Escobar H, Suarez L, Rico M, Bruix M, Mendez E (1998) β-Structure motif recognition by anti-gliadin antibodies in coeliac disease. FEBS Lett 427: 36–40

Anderson R, Degano P, Godkin A, Jewell D, Hill V (2000) In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. Nature Med 6: 337–342

Cornell HJ (1998) Partial in vitro digestion of active gliadin-related peptides in celiac disease. J Protein Chem 17: 739–744

Cornell HJ, Mothes T (1993) The activity of wheat gliadin peptides in "in vitro" assays for coeliac disease. Biochim Biophys Acta 1181: 169–173

Cornell HJ, Mothes T (1995) Further studies of the "in vitro" activity of synthetic gliadin peptides in coeliac disease. Biochim Biophys Acta 1270: 168–172

Cornell HJ, Rivett DE (1995) "In vitro" mucosal digestion of synthetic gliadin-derived peptides in celiac disease. J Protein Chem 14: 335–339

Cornell HJ, Townley RRW (1974) The toxicity of certain cereal proteins in coeliac disease. Gut 15: 862–869

Cornell HJ, Wieser H, Belitz H-D (1992) Characterisation of the gliadin-derived peptides which are biologically active in coeliac disease. Clin Chim Acta 213: 37–50

Cornell HJ, Skerritt JH, Puy R, Javadpour M (1994) Studies of "in vitro" gamma-interferon production in coeliac disease as a response to gliadin peptides. Biochim Biophys Acta 1226: 126–130

De Ritis G, Occorsio P, Auricchio S, Gramenzi F, Morosi G, Silano V (1979) Toxicity of wheat flour proteins and protein-derived peptides for "in vitro" developing intestine from rat fetus. Pediatr Res 13: 1255–1261

De Ritis G, Auricchio S, Jones HW, Lew EJ-L, Bernardin JE, Kasarda DD (1988) "In vitro" (organ culture) studies of the toxicity of specific A-gliadin peptides in celiac disease. Gastroenterology 94: 41–49

Graf L, Horvath K, Walcz E, Berzetei I, Burnier J (1987) Effect of two synthetic alpha-gliadin peptides on lymphocytes in celiac disease: identification of a novel class of opioid receptors. Neuropeptides 9: 113–122

Javadpour M, Juban M, Lo W-CJ, Bishop SM, Alberty JB, Cowell SM, Becker CL, McLaughlin ML (1996) "De novo" antimicrobial peptides with low mammalian cell toxicity. J Med Chem 39: 3107–3113

Kasarda DD, Okita TW, Bernardin JE, Baeker PA, Nimmo CC, Lew EJ-L, Dietler MD, Greene FC (1984) Nucleic acid (c-DNA) and amino acid sequences of alpha-type gliadins from wheat (Triticum aestivum). Proc Natl Acad Sci USA 81: 4712–4716

Kocna P, Mothes T, Krchnák V, Frič P (1991) Relationship between gliadin peptide structure and their effects on the fetal chick duodenum. Zeitschrift für Lebensmit-teluntersuchung und-Forschung 192: 116–119

Mothes T, Muhle W, Mäller F, Hekkens WTJM (1985) Influence of gliadin on fetal chick intestine in tissue culture. Biol Neonate 48: 59–64

Tatham AS, Marsh MN, Wieser H, Schewry PR (1990) Conformational studies of peptides corresponding to the coeliac-activating regions of wheat α-gliadin. Biochem J 270: 313–318

Terreaux C, Walk T, van de Wal Y, Koning F, Jung G, Fleckenstein B (1998) Increased HLA-DQ2-affinity of a sythetic gliadin peptide by acid-induced deamidation of glutamine residues. Bioorg Med Chem Lett 8: 2039–2044

Authors' address: Prof. Hugh J. Cornell, Department of Applied Chemistry, R.M.I.T. University, GPO Box 2476V, Melbourne, Victoria, 3001, Australia,
E-mail: hugh.cornell@rmit.edu.au

Received November 8, 2000