# Towards peptide based therapeutics-applications in celiac disease and infectious diseases

Kapoerchan, V.V.

# Chapter 1

# General Introduction

## 1.1 Introduction

Celiac disease (CD) is one of the most prevalent food-related disorders, with a occurrence of up to 1% in the Western population.[1] Originally thought to be primarily a problem in the Western world, cases of CD are now also emerging in Asia and Africa.[2] Symptoms include fatigue, chronic diarrhea, weight loss, anemia, psychiatric syndromes and neurological symptoms. CD patients also suffer from an elevated risk of infertility, bone disorders and intestinal malignancies.[3] CD can appear in early childhood, but many patients do not show symptoms until later in life. Furthermore, the severity of the symptoms varies strongly.[4] Typically, CD is indicated by the presence of tissue transglutaminase specific antibodies in the blood of patients,[5] and confirmed by studying a biopsy of the small intestine. In CD patients, the small intestinal morphology is affected which can culminate in the complete flattening of the epithelial barrier. The extent of the damage is scored with the use of the Marsh classification: Marsh 0 to Marsh III.[6] Marsh 0 stands for normal small intestine histology, while Marsh II and III are typical for CD patients. Complete remission of symptoms and normal histology after a gluten-free diet (GFD) definitively establishes the diagnosis. However, there has been much debate concerning the recovery of the small intestine as a consequence of the GFD, as complete recovery is not always

observed.[7] Moreover, symptoms may appear before the development of small intestinal damage and therefore new diagnostic methods are currently under development.[8] Although symptoms have been observed already in the second century A.D.,[9] it was not until the 1950s that the link between wheat gluten and CD was discovered by the Dutch pediatrician W.K. Dicke, who observed that the shortage of bread during World War II actually benefited children suffering from the abovementioned symptoms.[10] Today, it is known that the symptoms are caused by a misdirected immune response against gluten in genetically predisposed individuals. In healthy individuals, the immune system of the upper gastrointestinal tract is characterized by a so-called hyporesponsive state; the active inhibition of intestinal immune responses against harmless food-derived antigens such as gluten. For currently unknown reasons, in CD patients, this hyporesponsive state is either not established or broken, leading to a gluten-specific immune response.[11] This response leads to small intestinal damage, and due to this damage the surface of the small intestine flattens (also known as 'villous atrophy'), with deficient nutrient uptake (such as iron, folic acid, calcium and fat-soluble vitamins[12]) and the abovementioned symptoms as a result.

It is well established that, with a few exceptions, CD only develops in Human Leukocyte Antigen (HLA)-DQ2 and/or HLA-DQ8 positive individuals. These HLA-molecules bind gluten-derived peptides and induce inflammatory T cell responses which play a crucial role in the disease process. Native gluten-derived peptides, however, do not bind to HLA-DQ2 or HLA-DQ8 as these peptides lack the negatively charged amino acids required for binding to HLA-DQ2 or HLA-DQ8. These negative charges are introduced in the small intestine by the activity of the enzyme tissue transglutaminase which converts selected glutamine residues in gluten peptides into the negatively charged glutamate. It is known that CD development is not only the result of gluten-induced inflammation and genetical predisposition, but that environmental factors also play a role, such as breastfeeding and the moment of introduction of gluten in the diet.[13] Three key factors play an essential role in the pathogenesis of CD: gluten, the HLA-DQ2 protein and the enzyme tissue transglutaminase (tTG or TG2). Each will be discussed in the next sections.

### 1.1.1 The gluten proteins

Gluten proteins are storage proteins occurring in wheat and these proteins were identified as toxic for CD patients. On the basis of their ethanol solubility, gluten can be divided in gliadins and glutenins. The gliadins are further divided in the α/β-, γ-and ω-gliadins, while the glutenins can be divided into low- and high molecular weight glutenins (LMW and HMW, respectively), making gluten a complex mixture of

comparable but distinctly different proteins.[14] Typical for these proteins is the unusually high glutamine (Q, 35%) and proline (P, 15%) content, as compared to other proteins.[15] The glutamine residues serve as nitrogen source for germinating seeds, but the significance of the proline residues in this matter is unknown. However, the high amount of proline makes the gluten proteins resistant to gastropancreatic and brush border enzymes, which usually cleave proteins from food to tripeptides, dipeptides and single amino acids.[16] As a consequence, peptide fragments are formed which may elicit an immune response. Indeed, various toxic peptides have been identified, mainly from α- and γ-gliadins (for examples, see Table 1). Longer gluten peptides containing these nine-amino acid motifs are toxic to CD patients and elicit an immune response. Interestingly, young CD patients respond to a larger repertoire of toxic gluten peptides than adult CD patients.[17]

**Table 1** Examples of toxic gluten epitopes identified.[18]

| *Peptide* | *Sequence* |
|---|---|
| Glia-α2 | PQPQLPYPQ |
| Glia-α9 | PFPQPQLPY |
| Glia-α20 | FRPQQPYPQ |
| Glu-5 | QXPQQPQQF |
| Glia-γ1 | PQQSFPQQQ |
| Glia-γ2 | PYPQQPQQP |
| Glia-γ30 | IIQPQQPAQ |
| Glt-17 | PFSQQQQPV |

Proteins similar to gluten are also found in barley (hordeins), rye (secalins) and oats (avenins). While barley and rye cause symptoms similar to those caused by wheat, oats are considered safe for CD patients.[19] This is due to the lower number of toxic sequences present in the avenins, a feature that is linked to the lower amount of proline (10%) in the avenins[20] and the lower amount of total storage protein compared to wheat, barley and rye (10% avenins in oats as compared to 40-50% gluten, hordeins or secalins).[21]

### 1.1.2 The HLA-DQ2 molecule

The HLA-DQ2 protein is expressed by 95% of the CD patients, while the remaining 5% expresses the analogous HLA-DQ8 molecule.[17] However, HLA-DQ2 is also often found (~30%) in the population not suffering from CD, which suggests that additional factors play a role in the development of CD.[22] The 45 kD HLA-DQ2 heterodimer is formed by two protein subunits, termed the α and β chain (See

Appendix, Figure 1). There are many different α and β chains which can combine to form different heterodimers of HLA-DQ2.[23] HLA molecules play an essential role in immunity against pathogens. Their function is to bind peptides derived from proteins of pathogens and to display these in order to alert the immune system to the presence of pathogens. Upon specific recognition of these HLA-peptide complexes by white blood cells an immune response is triggered. In the case of CD, HLA-DQ2 binds to gluten peptides and presents these to gluten specific $CD4^+$ T cells.

Formation of a complex between a peptide and a MHC Class II molecule is governed by two sources of binding energy: (1) the positioning of specific peptide side chains, also known as 'anchors', in the binding pocket of HLA-DQ2 and (2) a hydrogen bond network formed between conserved residues of the HLA molecule and carbonyl oxygens and amide nitrogens in the main chain of the peptide epitope.[24] The p1, p4, p6, p7 and p9 positions of the HLA-DQ2 binding pocket harbor the anchor residues for peptide binding and the amino acids preferred at these positions have been identified (Table 2).[23] For the p1 and p9 positions, and to a lesser extent the p4 position, mainly bulky hydrophobic residues are preferred, while for the p4, p6 and p7 positions negatively charged residues are favored. For the p4 and p6 positions, this is the result of the presence of Lys71 in the β-chain of HLA-DQ2, which is a unique feature for this protein. This lysine can interact with the p4 and p6 negative charges in the peptide by hydrogen bonding and salt-bridge formation.[25]

**Table 2** Amino acid preference of HLA-DQ2. Anchor positions are indicated in bold.

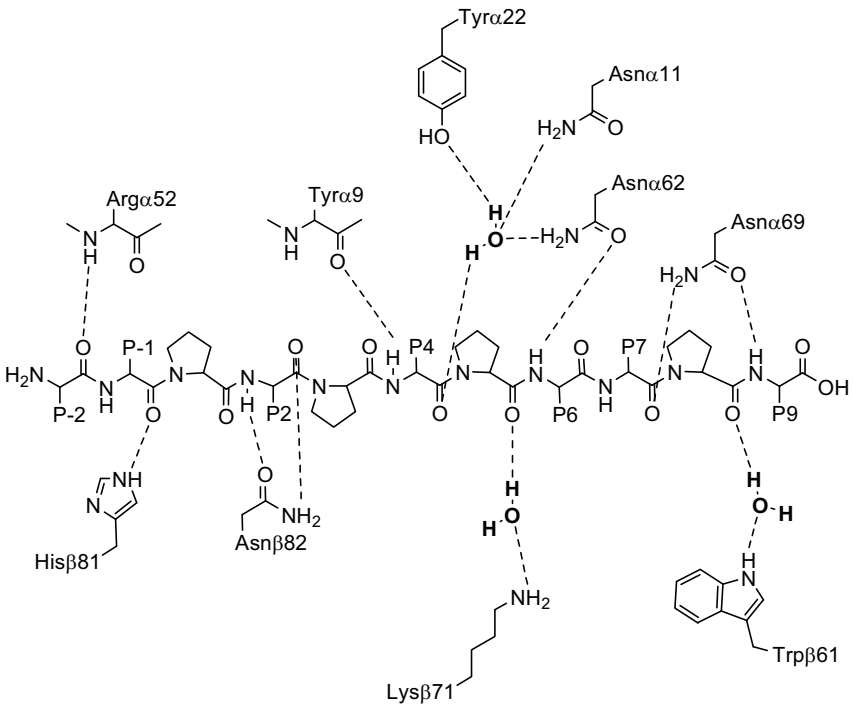
| *Position in the epitope* | *Amino acid preference*[23,26] |
|---|---|
| **1** | **F, I, L, P, W, V, Y** |
| 2 | E |
| 3 | E |
| **4** | **D, E, I, L, V** |
| 5 | E |
| **6** | **A, E, P** |
| **7** | **D, E** |
| 8 | E |
| **9** | **F, I, L, P, W, V, Y** |

In addition, it was observed that a negatively charged glutamate is accepted by HLA-DQ2 at every position in the binding pocket, including non-anchor positions, and that a proline is favored at p1 and p8.[26] The polyproline helix type II-like (PPII) conformation of the gluten peptides is the preferred conformation of MHC Class II

bound molecules and is essential in positioning the anchor side chains in the binding pocket of HLA-DQ2.

The interactions between the HLA-DQ2 protein and the gliadin peptide LQ**PFPQPELPY** (epitope in bold) upon binding have been investigated in detail.[27] Four hydrogen bonds are formed, involving the amide bonds of the p2, p4, p6 and p9 residues (Figure 1). The hydrogen bonds at p2 and p4 are important for binding, whereas those at p6 and p9 contribute less to the binding energy. A hydrogen bond between the NH of amide bond connecting the p1 and p-1 residues of the peptide and HLA-DQ2 was not observed in other peptides without proline at this position. Since peptides with a proline at the p1 position form the majority of the HLA-DQ2 binding peptides, it was concluded that a proline can be accommodated at this position without any energetic penalty.

**Figure 1** The hydrogen bond network between HLA-DQ2 and the peptide LQPFPEPELPY. Figure based on ref. 27. Hydrogen bonds are indicated by dashed lines, and the four hydrogen bonds between the amide bonds of p2, p4, p6 and p9 and HLA-DQ2 are indicated by bold dashed lines. Water molecules participating in the hydrogen bonding network are indicated in bold.




The 'anchor-pocket' model for peptide/HLA-DQ2 binding, comparable to the 'lock-and-key' model known in biology, assumes that both the peptide and the MHC Class II molecule each independently contribute to the binding energy and bind without

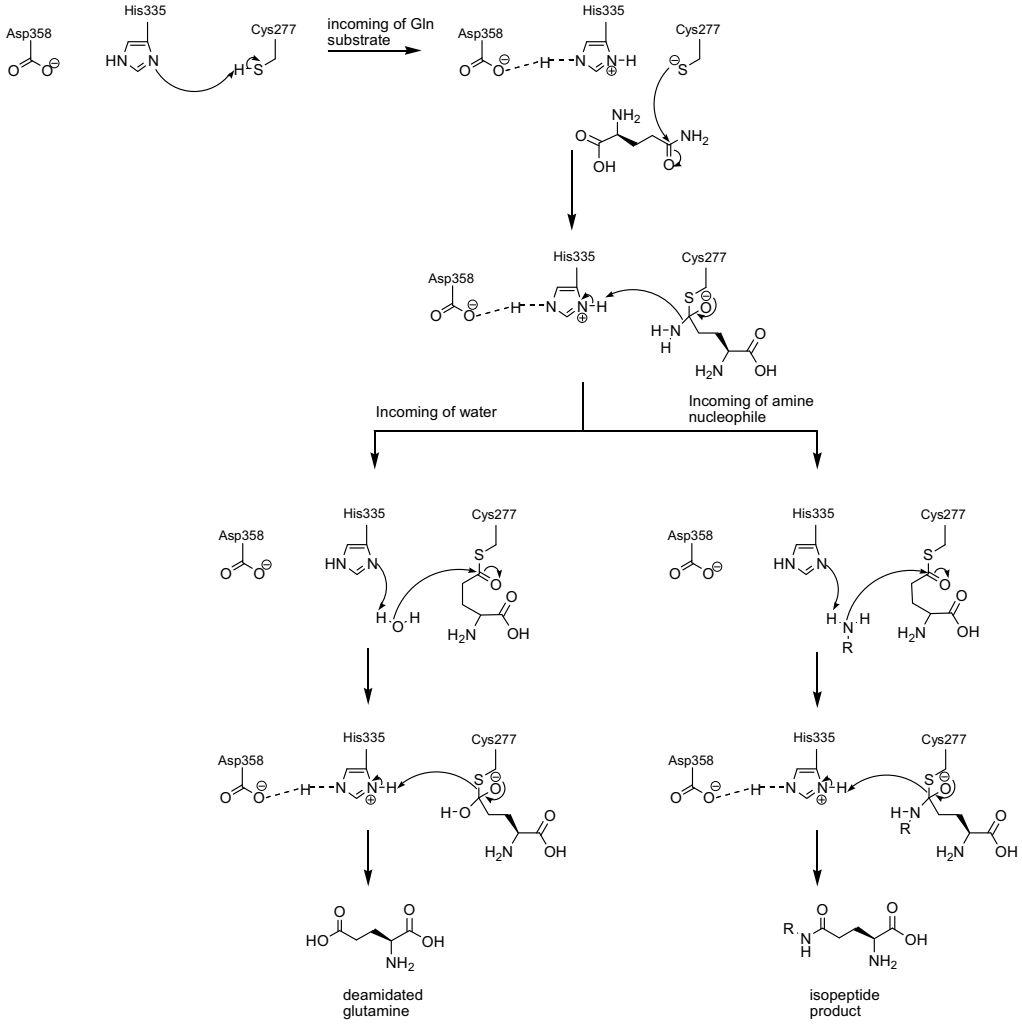
conformational changes. However, recent findings suggest that peptide binding to these proteins is a process in which both molecules undergo conformational changes in order to reach a lower energy state of the complex. This is supported by the observation that peptide/MHC Class II complexes adopt at least two conformational states *in vitro*: a fast dissociating 'open' or 'active' conformer and a kinetically stable 'closed' or 'compact' conformer. [24,28] It is expected that the same binding process will occur *in vivo* within the antigen presenting cells (APCs) where the MHC-Class II molecules bind to antigenic peptides.

### 1.1.3 Tissue transglutaminase

The human tissue transglutaminase (also known as TG2, tTG or transglutaminase 2) is a monomeric 76 kD protein and contains 686 amino acid residues.[29] It belongs to the vast family of the $Ca^{2+}$- and GTP-dependent transglutaminases, which mainly catalyze the crosslinking of the ε-amino group of a lysine of one protein to the γ-carboxamide group of a glutamine of another protein, forming an isopeptide bond. Instead of a lysine, other low molecular weight amine-functionalized compounds can participate in tTG-catalyzed amide-bond formation reactions.[30] However, when lysine/amine concentrations are low,[31] and at low pH,[32] deamidation of glutamine to glutamic acid is favored. The TG enzymes are expressed throughout all mammalian cells and tissues and play a role in various biological processes, such as wound healing and tissue repair, apoptosis and inflammation.[33] Furthermore, the transglutaminases are associated with various diseases, with tTG involved in Huntington's, Alzheimer's and Parkinson's disease, besides celiac disease.[34] tTG is mainly expressed in the lamina propria, the reservoir of the gut in which the majority of HLA-DQ2/DQ8-reactive T cells are present, and stimuli such as mechanical stress or a bacterial/viral infection increase its expression.[32] One remarkable feature of CD is the development of auto-immunity against tTG. The recognition of this autoantigen by highly specific antibodies serves as a screening method for CD.[35] The other important role of tTG in CD is the deamidation of selected glutamine residues of the gluten peptides. The glutamine-rich gluten peptides are good substrates for tTG and are also favored because of their PPII-like conformation.[25,36] The amino acid residues flanking the glutamine residue determine whether the glutamine is a good substrate for deamidation. It was found that the glutamines in one of the sequences QXP, QXX(F,Y,W,I,M,V or L) and QXP(F,Y,W,I,M,V or L) (X = any amino acid) are good residues for deamidation, while deamidation in QP and QXXP motifs is inhibited.[37] Recently, it was observed that tTG is also capable of hydrolyzing crosslinked products and thereby generating T cell stimulatory gluten peptides.[38]

The mechanism of acylation and deamidation involves a catalytic triad of Cys277, His335 and Asp358.[39] Upon hydrolysis of GTP and binding of the released GDP, this active site is blocked by two loops, with Cys277 forming a hydrogen bond with a Tyr residue. Upon binding of $Ca^{2+}$, the binding to the nucleotide is weakened, a conformational change takes place and as a consequence the active site becomes accessible.[29] In the first step of the acylation mechanism, a thio-ester is formed between the γ-carboxamide of the glutamine substrate and the active site cysteine. Then, the acyl moiety is transferred to a nucleophilic substrate, releasing the free enzyme and the isopeptide product. The nucleophile is an amine moiety, or in the absence of amines, a water molecule. This is schematically represented in Figure 2.

**Figure 2** The acylation mechanism of tTG. Picture based on ref 39.



### 1.1.4 Pathogenesis of CD

With all the key factors identified, the pathogenesis of CD can be summarized as follows (See appendix, Figure 2). Upon ingestion of gluten by CD patients, the proteins are hydrolyzed to small Gln- and Pro-rich peptide fragments. These fragments are taken up by the APC and, in some cases, further processed. Inside the cell, some of these fragments bind weakly to HLA-DQ2 or HLA-DQ8. These HLA-gluten peptide complexes are recognized by specific $CD4^+$ T cells[40] and this leads to the activation and proliferation of these T cells. Subsequently, interferon-γ and other inflammatory cytokines are released. This leads to inflammation of the small intestine through the release of metalloproteinases by fibroblasts and inflammatory cells.[41,9]

Not much is known about the interaction of the gluten peptide/HLA-DQ2 complex with the T cell receptor (TCR) on the T cell, as no crystal structures are available. However, binding studies have shown that T cell recognition is strongly dependent on the amino acid at p-1 (the amino acid preceding the epitope). Molecular modeling suggests that a proline at p-1 points towards the TCR and thus interacts with the TCR.[42] Following the immune response, tissue damage of the small intestine occurs, which results in the release of tissue transglutaminase. Confronted with a high concentration of glutamine-rich gluten peptides, tissue transglutaminase will take on specific glutamine residues and transform these into the corresponding glutamates, instead of catalyzing the crosslinking reaction due to the low abundance of Lys residues in gluten.[43,44] As HLA-DQ2 and HLA-DQ8 are known to preferentially bind peptides that harbor negatively charged amino acids, some of these glutamate containing peptides bind to HLA-DQ2 or HLA-DQ8 with a much enhanced affinity compared to their glutamine counterparts,[45,46,47] with an amplified immune response as a result,[17,31,37,43,44,48,49,50] leading to severe intestinal damage resulting in the clinical symptoms typically observed. Although the site-specific deamidation of selected glutamine residues has been considered as the most important role of tTG in CD pathogenesis, there is an increasing amount of evidence that tTG antibodies, which are used to diagnose CD, also have an important function. In a recent study, it was suggested that these antibodies are involved in the development of the small-intestinal damage typical for CD,[51] instead of being the consequence of tTG release.

## 1.2 The development of therapies for the treatment of CD

Today, the only therapy available for CD patients is strict adherence to a gluten-free diet. Upon following this diet, the small intestine recovers to its original state and symptoms disappear. However, adhering to a GFD is quite burdensome as wheat, barley and rye are very frequently used in the food industry. Additionally, the presence of gluten in food is not always obvious and so-called gluten-free products may contain

traces of gluten.[52] Finally, the GFD has a large impact on the quality of life.[53] Therefore, alternative therapies are needed. Various approaches targeted at the three key factors already mentioned have been explored and will be discussed hereafter.

### 1.2.1 Safe alternatives for gluten in foods

Some effort has been taken towards the expansion of the gluten-free diet. Whereas wheat, barley and rye are 'forbidden' for CD patients, oats is considered safe. Although oats does contain T cell stimulatory epitopes, a study in which CD patients consumed oat products for 5 years showed no immunological response in any of the test subjects.[54] The Ethiopian cereal Tef (Eragrostis Tef) was investigated for the presence of stimulatory epitopes, but no such peptides were found.[55] Therefore, this cereal can be considered as 'gluten-free' and safe for use in a GFD.

### 1.2.2 Enzymatic degradation of toxic peptides

Due to the presence of multiple prolines, gluten is relatively stable against gastropancreatic and brush border enzymes, which usually hydrolyze peptides and proteins from food to single amino acids, dipeptides and tripeptides. Inefficient breakdown by these enzymes generates toxic gluten peptides. When these peptides are degraded completely before they reach the small intestine, no gluten peptides can bind to HLA-DQ2 and therefore, no immune response takes place.

In 2002, Shan and co-workers suggested that detoxification of gluten with a prolyl endopeptidase (PEP) from *Flavobacterium meningosepticum*, a serine protease that cleaves peptides after Pro residues, could potentially be an effective therapy for CD patients.[56] The enzyme would be administered to the patient before or during the consumption of gluten-containing foods. Structural analysis of a porcine PEP revealed that the enzyme consists of a peptidase domain, built up from the N- and C-terminal region, with a catalytic triad of Ser554, His680 and Asp641 at the interface of the two domains. Attached to the catalytic residues is a seven blade β-propeller domain, based on a sevenfold repeat of four stranded antiparallel β-sheets.[57,58] This secondary structure is representative for all PEPs. Various PEPs from different bacterial sources (*Myxococcus xanthus, Sphingomonas capsulate*) or recombinant PEPs have been evaluated for their ability to degrade gluten peptides. All of these enzymes were found to be effective, however, in some cases the activity was found to be dependent on substrate length.[59] Moreover, practical application is limited as the PEP enzyme is inactivated by pepsin and at low pH (below pH 5). To overcome this problem, an enzyme combination approach was proposed. [60] EP-B2 (a barley cysteine endoprotease) was the enzyme of choice and is complementary to the PEPs in its sequence specificity, chain length specificity and pH activity range. The EP-B2 itself was shown to

effectively degrade gluten in the rat stomach,[61] and in combination with a PEP from *Flavobacterium meningosepticum* rapid and efficient degradation of gluten peptides was observed under simulated gastric conditions.[60,62] Similarly, a combination of aspergillopepsin (ASP) from *Aspergillus Niger* and dipeptidyl peptidase IV from *Aspergillus oryzae*, two enzymes commonly used in the food industry, was found to degrade gluten as well under simulated gastric conditions.[63] Another approach is the modification of the PEP to obtain variants which are more stable and suitable for therapeutic use.[64] Alternatively, it was shown that a single PEP, originating from *Aspergillus Niger*, was able to degrade gluten completely within a few hours under conditions present in the gastrointestinal tract.[65,66] Moreover, this enzyme was found to be stable against pepsin treatment, and was active in a pH range from 2-8. The difference in stability probably arises from the sequence homology with other peptidases rather than with PEPs. The enzyme therapy discussed above seems a promising treatment for CD patients and the EP-B2/PEP formulation[62] is currently being tested in clinical trials.

### 1.2.3 Inhibitors of the tissue transglutaminase enzyme

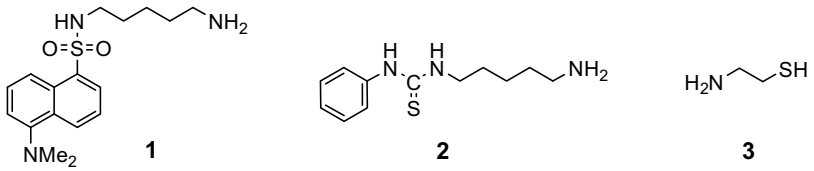
The natural gluten peptides are marginally toxic for CD patients, but after tTG-induced deamidation negative charges are introduced in the peptide, which enhance the affinity for the HLA-DQ2 protein, with amplified immune responses as a result. Therefore, a considerable amount of research has gone into the development of tTG inhibitors, either competitive amine inhibitors, reversible inhibitors or irreversible inhibitors. Some competitive inhibitors are amine substrates which, as enzyme substrate, can compete with the natural lysine substrate in the crosslinking reaction catalyzed by tTG. The parameter used to assess the inhibition potency of competitive inhibitors is the dissociation constant $K_I$ (given in μM) for the inhibitor/enzyme complex. This constant represents the affinity of the substrate for the enzyme. The reversible inhibitors block substrate access to the active site, but without covalently binding to the enzyme. The potency of these inhibitors is expressed by $K_I$ (given in μM), but $IC_{50}$ values, which stand for the concentration of inhibitor needed to reduce enzyme activity to half as compared to the activity in the absence of inhibitor, are used as well. Finally, the irreversible inhibitors form a covalent bond with the active site of the enzyme and as a consequence, the natural substrate can not be converted by the enzyme. The chemical parameter used to assess the potency of irreversible inhibitors is $k_i/K_I$, where $k_i$ (in $min^{-1}$) is the kinetic parameter describing the reaction rate for formation of the inhibitor-enzyme covalent bond, and thus represents the rate of inactivation of the enzyme.[67] The quotient of these two entities, $k_i/K_I$ (in $\mu M^{-1}$ $min^{-1}$) is a measure of the effectivity of the inhibitor: a high value indicates that the inhibitor

both shows high affinity for the enzyme and that the enzyme is quickly inactivated. $IC_{50}$ values are also often used to compare the potency of irreversible inhibitors. Many inhibitors were assayed using guinea pig liver transglutaminase as a model system, because of its resemblance to the human transglutaminases and its availability.[68] However, because of the various methods used to calculate the constants and the many different assays used to probe enzyme activity, results from different studies cannot always be compared directly.

### 1.2.3.1 Competitive amine inhibitors

The competitive amine inhibitors described so far were based on simple molecules, such as dansylcadaverine[69] (**1**, Figure 3), and phenylthiourea derivative **2**.[70] Both compounds displayed moderate efficiency as inhibitors with $K_I$ values between 20 and 50 μM, but were also shown to be non-specific for tTG. β-mercaptoethanolamine (**3**) was assessed for its inhibitory activity for tTG[71] and was found to be less effective ($K_I$ = 150 μM) than compounds **1** and **2**. Moreover, specificity for tTG was not evaluated.

**Figure 3** Competitive amine inhibitors.

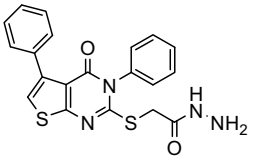



### 1.2.3.2 Reversible inhibitors

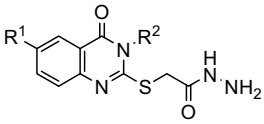
The simplest of reversible inhibitors for tTG was found to be the $Zn^{2+}$ ion. The enzyme needs $Ca^{2+}$ for activation, but $Zn^{2+}$ ions are able to compete with $Ca^{2+}$ for binding to the enzyme and thus prevent the activation of the enzyme. The $Zn^{2+}$ ion shows a $K_I$ of 0.1 μM, and is thus fairly effective.[72] Interestingly, it was later found that the GTP cofactor of tTG inhibited crosslinking activity when the nucleotide was complexed with $Mg^{2+}$ and the same was the case for Mg-GDP and Mg-GMP ($IC_{50}$ values 9, 9, and 400 μM, respectively). It was found that these complexes were allosteric inhibitors, having a distinct binding site other than the active site.[73]

Thieno[2,3-*d*]pyrimidin-4-one acylhydrazide **4** (Figure 4) was found in a high-throughput screen for tTG to be a reasonably effective inhibitor.[74] The structure-activity relationship of **4** ($IC_{50}$ = 0.8 μM) was extensively investigated by Duval *et al.*[75] who prepared a wide range of analogs which were tested for their activity on tTG.

**Figure 4** Structure-activity relationships investigated for **4**.

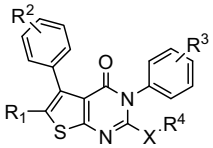


**4**



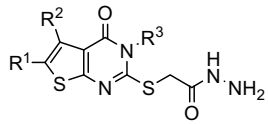
**5**

$R^1$ = H, Ph
$R^2$ = Ph, 3-F-Ph



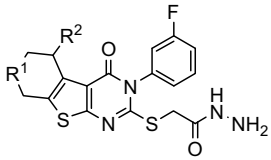
**6**

$R^1$ = H, Me
$R^2$ = H, 2-F, 4-F
$R^3$ = H, 3-F, 4-F
$R^4$ = $CH_2C(O)NHNH_2$, $CH_2CH_2C(O)NHNH_2$, $CH_2(Me)C(O)NHNH_2$, $CH_2C(O)NMeNH_2$, $CH_2C(O)NHNHMe$, $CH_2CO_2H$, $CH_2CO_2Et$, $CH_2C(O)NHMe$, $CH_2C(O)NH_2$, $CH_2C(O)NHOH$, $CH_2C(O)CH_2NH_2$,
X = S, O, NH, NMe



**7**

$R^1$ = H, Cl, Me, *i*Pr, Ph
$R^2$ = H, Me, *i*Pr, Ph, 2-OMe-Ph, 3-OMe-Ph, 4-OMe-Ph, 2-($OC_3H_6$-$NEt_2$)-Ph, 2-F-Ph, 3-F-Ph, 4-F-Ph
$R^3$ = Me, Ph, $CH_2Ph$, 3-Py, Cy, 2-OMe-Ph, 3-OMe-Ph, 4-OMe-Ph, 2-Cl-Ph, 3-Cl-Ph, 4-Cl-Ph, 2-F-Ph, 3-F-Ph, 4-F-Ph



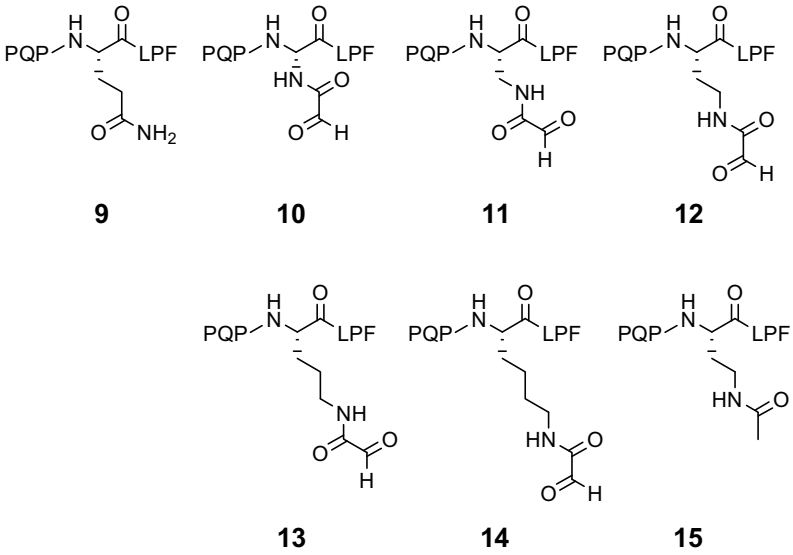
**8**

$R^1$ = $CH_2$, NH, NMe, N-Pr, N-Boc
$R^2$ = H, Me, Ph, $CH_2Ph$

First, the replacement of the thiophene ring with a benzene ring (**5**) was studied, but this was found to be detrimental for tTG inhibition. Next, the thioether and acylhydrazide moieties were systematically replaced (**6**). Replacement of the sulfur atom with nitrogen or oxygen was again not beneficial for tTG activity ($IC_{50}$ = 3 μM and higher). Elongation of the acylhydrazide, removal of the acylhydrazide or replacement with amines, also performed in combination with substitution of the aromatic rings by 3- or 4-fluorobenzene, did only result in diminished activity ($IC_{50}$ = 5 μM and higher). Subsequently, the substitution pattern on the aromatic rings was investigated (**7**). Replacement of the phenyl group at $R^3$ with a methyl, cyclohexyl, benzyl or 3-pyridyl group was detrimental, as well as the introduction of a 2- or 4-methoxyphenyl or chlorophenyl group ($IC_{50}$ = 1.2-8 μM). A 3-methoxyphenyl substituent at $R^3$, however, showed equal activity as **4** ($IC_{50}$ = 0.82 μM). A fluorophenyl substitution both at $R^2$ and $R^3$ increased activity, just as a 2- or 3-methoxyphenyl at $R^2$ ($IC_{50}$ = 0.14-0.32 μM). At $R^1$, a chloro- or methyl substituent in combination with fluorophenyl substituent at $R^2$ and $R^3$ increased activity ($IC_{50}$ = 0.13-0.16 μM), but a larger group such as isopropyl was not beneficial ($IC_{50}$ = 0.90 μM). Finally, a cyclohexyl ($R_1$ = $CH_2$) or piperidyl ($R_1$ = NH or NR) moiety was fused

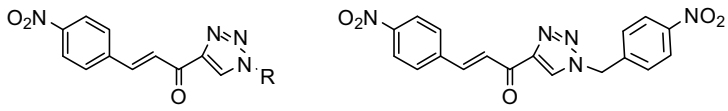
to the thiophene ring (**8**). However, no significant increase in activity was noted. It was concluded that while the acylhydrazide moiety and the thiophene ring were necessary for inhibitory properties, the other substituents could be slightly modified.
In a later study by the same group, the kinetics of inhibition of tTG by **4** was investigated in detail.[76] It was found that **4** is a slow-binding and reversible inhibitor of tTG with an $IC_{50}$ value of 0.6 μM. It is an allosteric binder, as it does not bind in the active site, but probably at the GTP binding site. Furthermore, compound **4** was also evaluated for inhibitory activity against coagulation factor XIIIa and caspase 3. The first enzyme is a member of the TG family, the second has an acyl-transfer mechanism identical to that of tTG. It was found that **4** did not have any effect on these two enzymes. Moreover, this compound did not show any cytotoxic activity.
Based on the observation that aldehyde-containing gluten peptide analogs show reasonable binding affinity for DQ2, a series of short aldehyde-containing peptides was synthesized by the group of Khosla (Figure 5).[77] It was postulated that the active site cysteine would react with the aldehyde and form a hemithioacetal. The peptides were evaluated for their inhibitory activity, and it was observed that peptides **10** and **14** did not show any inhibition ($K_I$ > 500 μM), probably because of steric factors. The optimal side chain length for inhibition was found to be longer than the glutamine substrate **9**, as peptide **12** was the best inhibitor of this series ($K_I$ = 40 μM). Peptides **11** and **13** showed poor inhibition ($K_I$ > 200 μM). The fact that peptide **15**, lacking the aldehyde component, also displayed poor inhibition ($K_I$ > 100 μM), proved that the aldehyde was indeed involved in inhibition.

**Figure 5** Aldehyde-containing peptides as reversible tTG inhibitors.



The small peptide Cbz-Gln-Gly is a known TG substrate.[78] The fact that the cinnamoyl group is a more rigid analog of the Cbz group led to the development of cinnamoyl-containing inhibitors (Figure 6), which were found to be reversible inhibitors of the tTG enzyme.[79] By systematic screening of the substituents on the aromatic ring, and the variation in the nature of the cyclic moiety, the lead structure **16** was identified. In a follow-up study[80] inhibitor **17** was synthesized and found to be a very effective inhibitor of tTG ($IC_{50}$ = 2.1 μM). Moreover, some of the inhibitors evaluated in this study were selective for tTG and did not exhibit any activity towards related enzymes.

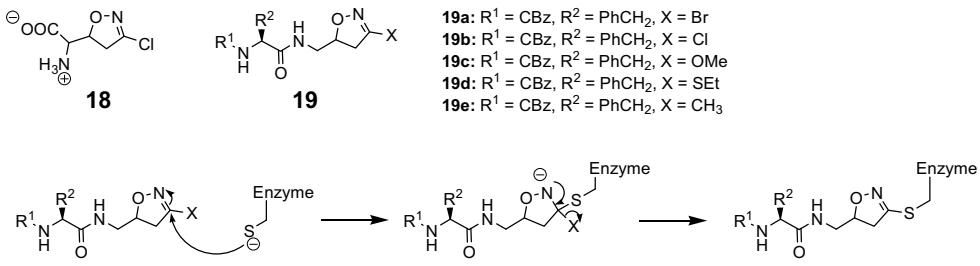
**Figure 6** Representative examples of cinnamoyl-containing inhibitors.




### 1.2.3.3 Irreversible inhibitors

The natural product acivicin (**18**, Figure 7) is known to inhibit the enzyme anthranilate synthase by reacting with the active site cysteine. Since the active site of tTG also contains a Cys residue, Krantz and co-workers reasoned that analogous peptidyl 3-halo-4,5-dihydroisoxazoles (such as **19a** and **19b**) might inhibit TG inhibitory activity.[81] The Cbz-group was incorporated because the small peptide Cbz-Gln-Gly is a known TG substrate.[78]

**Figure 7** Acivicin derived 3-substituted-4,5-dihydroisoxazoles as irreversible inhibitors of TG and inactivation of the enzyme by these compounds.

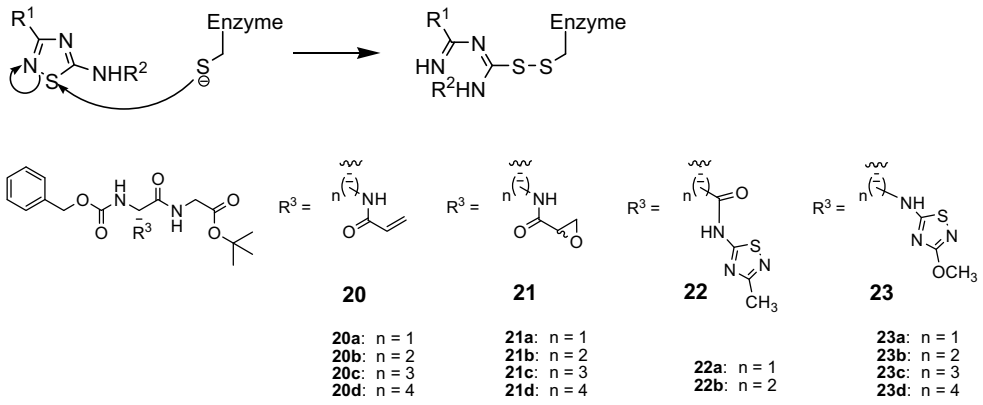



Indeed, it was found that these compounds inhibited TG in a irreversible manner by nucleophilic displacement of the ring substituent X by the active site cysteine thiol of the enzyme (Figure 7), with the potency being dependent of the nature of the 3-substituent on the ring. Compounds **19a** and **19b** were the most potent ($k_i/K_I$ = 0.054

and 0.013 $\mu M^{-1}$ $min^{-1}$, respectively). The importance of the 3-substituent was confirmed by the lack of activity observed for **19e**.

Alternative irreversible inhibitors for the guinea pig liver TG were synthesized and evaluated for their activity by the group of Keillor.[82,83] These compounds consisted of α,β-unsaturated amides, epoxides and 1,2,4-thiadiazoles as electrophilic traps attached to either a mono- or dipeptide which is recognized by the enzyme. Again, the enzyme is inactivated by attack of the active site cysteine on the inhibitor (Figure 8). Cbz-X-OH, Cbz-X-Gly-O*t*Bu (shown in Figure 8) and Cbz-X-Gly-OH were chosen as scaffolds, in which X is the amino acid containing the electrophilic trap.

**Figure 8** Inactivation of TG by thiadiazoles and examples of inhibitors generated by Keillor and co-workers.[82,83]

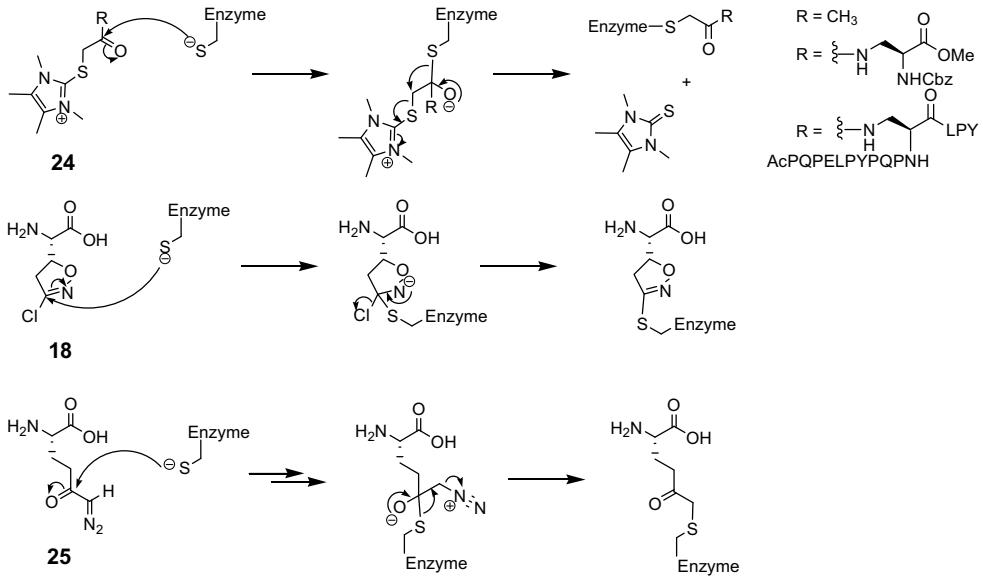



It was found that the α,β-unsaturated amide dipeptide substrates **20** showed higher affinity for the enzyme ($K_I$ = 0.23-1.48 μM) than the same inhibitors based on the Cbz-X-OH scaffold ($K_I$ = 0.52-2.32 μM) and that removal of the *t*-butyl ester further improved affinity ($K_I$ = 0.15-0.48 μM). The optimal side chain length was observed to be four methylene units, and thus the Cbx-X-Gly-OH compound with n = 4 was the most efficient inhibitor in this series, with the highest affinity and the highest rate of inactivation: $K_I$ = 0.15 μM and $k_i/K_I$ = 3.00 $\mu M^{-1}$ $min^{-1}$. For the epoxide compounds, which were tested as diasteromeric mixtures, the monopeptides were found to have the highest affinity for the enzyme, namely $K_I$ = 0.28-2.13 μM. The protected dipeptide **21** displayed a decreased affinity of $K_I$ = 0.95-5.91 μM and the removal of the *t*-butyl ester partially restored the affinity. Again, the side chain length of four methylene units was found to be optimal and therefore the Cbz-X-OH peptide with n = 4 was the most efficient inhibitor in the epoxide series: $K_I$ = 0.28 μM and $k_i/K_I$ = 2.03 $\mu M^{-1}$ $min^{-1}$. The difference between these two classes of compounds could be

explained by the steric bulk of the epoxide reactive group as compared to the double bond. Next, the 1,2,4-thiadiazole compounds were evaluated for their inhibition of guinea pig liver TG. For these compounds, it was observed that the substrates with one methylene unit in the side chain did not have any inhibitory effect on the enzyme. As for the epoxide series described above, the attachment of the glycine residue was not beneficial for inhibition, and removal of the *t*-butyl group did not restore the activity. Compound Cbz-X-OH with n = 2 was found to be the most effective inhibitor with high affinity and efficient inactivation: $K_I$ = 0.77 μM and $k_i/K_I$ = 0.89 $\mu M^{-1}$ $min^{-1}$. For the second series of thiadiazoles, compound Cbx-X-Gly-OH with n = 2 was found to be the most effective with $K_I$ = 1.3 μM and $k_i/K_I$ = 0.71 $\mu M^{-1}$ $min^{-1}$, again showing the importance of side chain length and a free carboxylic acid moiety in the inhibitor.

Shortly thereafter, Khosla and co-workers synthesized glutamine isosteres, either as single amino acid or as part of either a Cbz-X-OMe (with X being the glutamine isostere) or high affinity peptide tTG substrates, and tested them for inhibition of tTG.[84] The isosteres contain either a tetramethyl thioimidazole (**24**, Figure 9), a 3-chloro-4,5-dihydroisoxazole as in acivicin **18**, or a 6-diazo-5-oxo-norleucine (DON) moiety (**25**).

**Figure 9** Inactivation of tTG by glutamine isosteres.

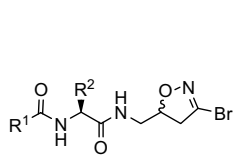



As controls, protected and unprotected glutamine and a glutamine-containing peptide were used. The Cbz-X-OMe motif is a 'low-affinity scaffold', while the peptide-based

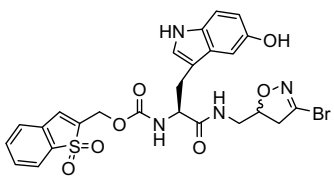
inhibitors are 'high affinity scaffolds'. The inactivation by these compounds is depicted in Figure 9. The tetramethyl thioimidazoles inactivates the enzyme by transfer of the acyl portion of the inhibitor to the enzyme and release of the tetramethyl imidazolethione.[85] The inactivation of active-site cysteine containing enzymes by DON is believed to proceed via attack of the active site thiol on the ketone moiety, after which a thiohemiketal is formed. Collapse of this intermediate and exclusion of nitrogen leads to the product.[86] All compounds were evaluated for their tTG inhibitory properties. Compounds derived from **24** (See Figure 9) showed significantly decreased affinity for the enzyme as compared to the glutamine controls. Compounds **18** and **25** showed low activity against tTG ($k_i/K_I = 0.17 \cdot 10^{-6}$ and $0.2 \cdot 10^{-6}$ $\mu M^{-1}$ $min^{-1}$, respectively), but incorporating the glutamine isostere in a peptide markedly increased its affinity for tTG and its inhibition efficiency (PQ-**18**-LPY: $k_i/K_I = 1.8 \cdot 10^{-5}$ $\mu M^{-1}$ $min^{-1}$, Cbz-**25**-OMe: $k_i/K_I = 8.9 \cdot 10^{-4}$ $\mu M^{-1}$ $min^{-1}$, Ac-PQP-**25**-LPF-$NH_2$: $k_i/K_I = 2.9$ $\mu M^{-1}$ $min^{-1}$). The DON containing peptide Ac-PQP-**25**-LPF-$NH_2$ showed the best inhibitory activity. Moreover, no cytotoxic effects were observed towards epithelial cells and fibroblasts, and it was found to be selective for tTG over other related TG enzymes.

The group of Khosla also prepared and evaluated a series of acivicin analogs (Figure 10) based on the core structure **26**.[87]

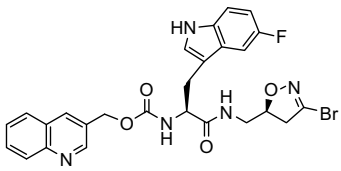
**Figure 10** Representative acivicin analogs evaluated as tTG inhibitors.




**26**




**27**

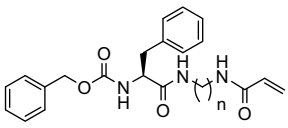



**28**

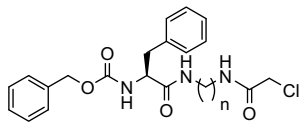
Varying the $R^1$ and $R^2$ positions finally led to compound **27** which was found to exhibit both high affinity and rapid inactivation of tTG, with $k_i/K_I = 6.8 \cdot 10^{-3}$ $\mu M^{-1}$ $min^{-1}$. No activity was found towards naturally occurring thiols such as glutathione (δ-Glu-Cys-Gly). A similar compound (**26** with $R^1 = C_6H_5\text{-}CH_2O\text{-}$ and $R^2 = \text{-}CH_2C_6H_4\text{-}$*p*-OH) was evaluated for its toxicity *in vivo* by administering the compound to mice. These assays demonstrated that this compound possessed good oral bioavailability, was active and not toxic in mice. Further elaboration on this motif and the enzymatic resolution of the racemic dihydroisoxazoles used previously to obtain enantiomerically pure compounds revealed the 5-fluorotryptophane **28**, which was found to be a

potent inhibitor ($k_i/K_I$ = 0.057 $\mu M^{-1}$ $min^{-1}$) that showed a eight-fold increase in potency compared to compound **27** ($k_i/K_I = 6.8 \cdot 10^{-3}$ $\mu M^{-1}$ $min^{-1}$).
Keillor and co-workers prepared a series of eight compounds, combining the Cbz-Phe scaffold[87] with α,β-unsaturated amides, chloroacetamides and maleimides as cysteine reactive groups and with spacers of different length between the scaffold and the reactive group (Figure 11).[88] The α,β-unsaturated amides had been applied before, as part of a Cbz-X-Gly scaffold and were shown to be reasonably effective irreversible inhibitors of guinea pig liver TG.[82]

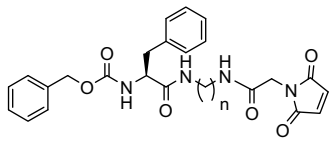
**Figure 11** Inhibitors based on the Cbz-Phe scaffold.



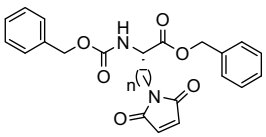
**29** **a**: n = 2
**b**: n = 4
**c**: n = 6
**d**: n = 8



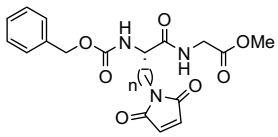
**30** **a**: n = 2
**b**: n = 4
**c**: n = 6



**31** **a**: n = 4



**32** **a**: n = 2
**b**: n = 4
**c**: n = 6



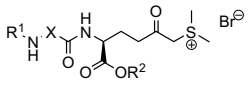
**33** **a**: n = 2
**b**: n = 4
**c**: n = 6

It was found that the affinity of compounds **29a**-**d** (Figure 11) was about 60-fold higher ($K_I$ = 3.5-13 μM) than similar compounds evaluated in the previous study[87] ($K_I$ = 0.26-0.74 mM, compounds not shown), suggesting that the large dihydroisoxazole moiety in the latter is not accommodated well in the TG substrate pocket. The inhibition efficiency of the compounds was found to be slightly variable with the spacer length; the longer spacers were more effective than the shorter ones. This might be explained by the increasing flexibility of the longer spacers, due to which a better positioning of the reactive group in the active site could be achieved. Compounds **30b**-**c** showed the same correlation between linker length and activity as seen for **29a**-**d**, but were less effective as inhibitors ($k_i/K_I$ = 0.010-0.053 $\mu M^{-1}$ $min^{-1}$ for **29a**-**d** vs. $k_i/K_I$ = 0.069-0.109 $\mu M^{-1}$ $min^{-1}$ for **30b**-**c**), while **30a** did not show any inhibition at all. Peptide **31** also did not show inhibition, probably due to the large maleimido-moiety. Compounds **29c** and **30c** were found to react only very slowly with glutathione, indicating their selectivity for tTG. In a later study the effect of the

incorporation of the maleimido-group as amino acid side chain (Figure 11) and the effect of the distance of the maleimide from the peptide backbone were investigated.[89] It was found that the dipeptide inhibitors were two- to twenty-fold less effective than the benzyl esters (**32a**-**c**: $k_i/K_I$ = 0.67-17.08 · $10^{-3}$ $\mu M^{-1}$ $min^{-1}$, **33a**-**c**: $k_i/K_I$ = 0.28-0.83 · $10^{-3}$ $\mu M^{-1}$ $min^{-1}$). Probably, because of the aromatic ring of the benzyl esters, these inhibitors are better positioned in the active site, as observed before.[87] When considering the chain length, it was found that the efficiency increases with side chain length, which was again observed previously.[82,83,88] However, these inhibitors were $10^3$-$10^4$ fold less effective than the α,β-unsaturated amides identified from the previous study,[82] again showing that the large maleimide ring does not fit well into the active site of tTG.

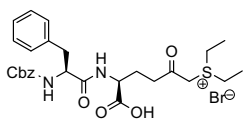
Based on the fact that sulfonium peptidylmethylketones have been shown to irreversibly inhibit proteases by transfer of the peptidyl portion to the active site cysteine sulphur atom, a series of water-soluble dipeptide-based sulfonium peptidylmethylketones (Figure 12) was synthesized and evaluated for their inhibitory activity.[90]

**Figure 12** Sulfonium peptidylmethylketone-based inhibitors of tTG.

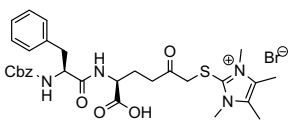


**34**

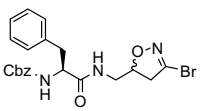
**a**: $R^1$ = Cbz, X = Pro, $R^2$ = H
**b**: $R^1$ = Cbz, X = Gly, $R^2$ = H
**c**: $R^1$ = $PhCH_2CH_2OC$=O, X = Phe, $R^2$ = H
**d**: $R^1$ = Cbz, X = Phe, $R^2$ = H
**e**: $R^1$ = Cbz, X = Trp, $R^2$ = H
**f**: $R^1$ = Cbz, X = Hyp, $R^2$ = H
**g**: $R^1$ = Cbz, X = Ala, $R^2$ = H
**h**: $R^1$ = Cbz, X = Asp, $R^2$ = H
**i**: $R^1$ = Fmoc, X = Phe, $R^2$ = H
**j**: $R^1$ = $PhCH_2C$=O, X = Phe, $R^2$ = H
**k**: $R^1$ = Cbz, X = Lys, $R^2$ = H
**l**: $R^1$ = Boc, X = Phe, $R^2$ = H
**m**: $R^1$ = Cbz, X = Ile, $R^2$ = H
**n**: $R^1$ = Cbz, X = Phe, $R^2$ = Me



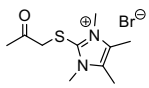
**35**



**36**



**37**



**38**

These compounds were hypothesized to possess reduced toxicity. Due to their water-solubility and permanent charge, they should be more likely to remain outside cells and target extracellular tTG, instead of crossing the cell membrane and react with other TG enzymes and proteases. Compounds **37** and **38** were taken as reference compounds. It was found that the Cbz group was the optimal substituent for the N-terminus, in accordance with the results shown above. When the phenylalanine was replaced with bulky amino acids, aliphatic amino acids or charged amino acids, a decrease in activity was seen (**34e**-**k**, **m**, $IC_{50}$ values from 12-305 μM) as compared to **34d** ($IC_{50}$ = 10 μM) but the introduction of a proline or glycine slightly increased activity to $IC_{50}$ = 5-8 μM (**34a**, **34b**). Protection of the acid with a methyl ester (**34n**)

also slightly improved the activity ($IC_{50}$ = 5 μM), as compared to **34d**. Changing the dimethylsulfonium to the diethyl moiety reduced the activity by 50% ($IC_{50}$ of **35** = 20 μM), while incorporation of the tetramethylimidazolium group resulted in the most potent compound in this study, with an $IC_{50}$ value of 3 μM, being as active as reference compound **38**. Although peptide **37** proved to be effective in a previous study,[91] here it did not show any activity. Compound **34d** was also applied to an aminal model and did not display any toxicity.

A wide variety of irreversible, reversible and competitive tTG inhibitors have been synthesized and evaluated, leading to the establishment of valuable structure-activity relationships. It has been shown that selective inhibitors for tTG can be prepared. Although tTG is implicated in a wide variety of biological processes and irreversible inhibition of tTG can result in severe side-effects, it has been shown that irreversible tTG inhibitors can be administered in animal models without showing any toxicity. However, more research concerning this issue is necessary before such agents can be used in the treatment of CD patients.

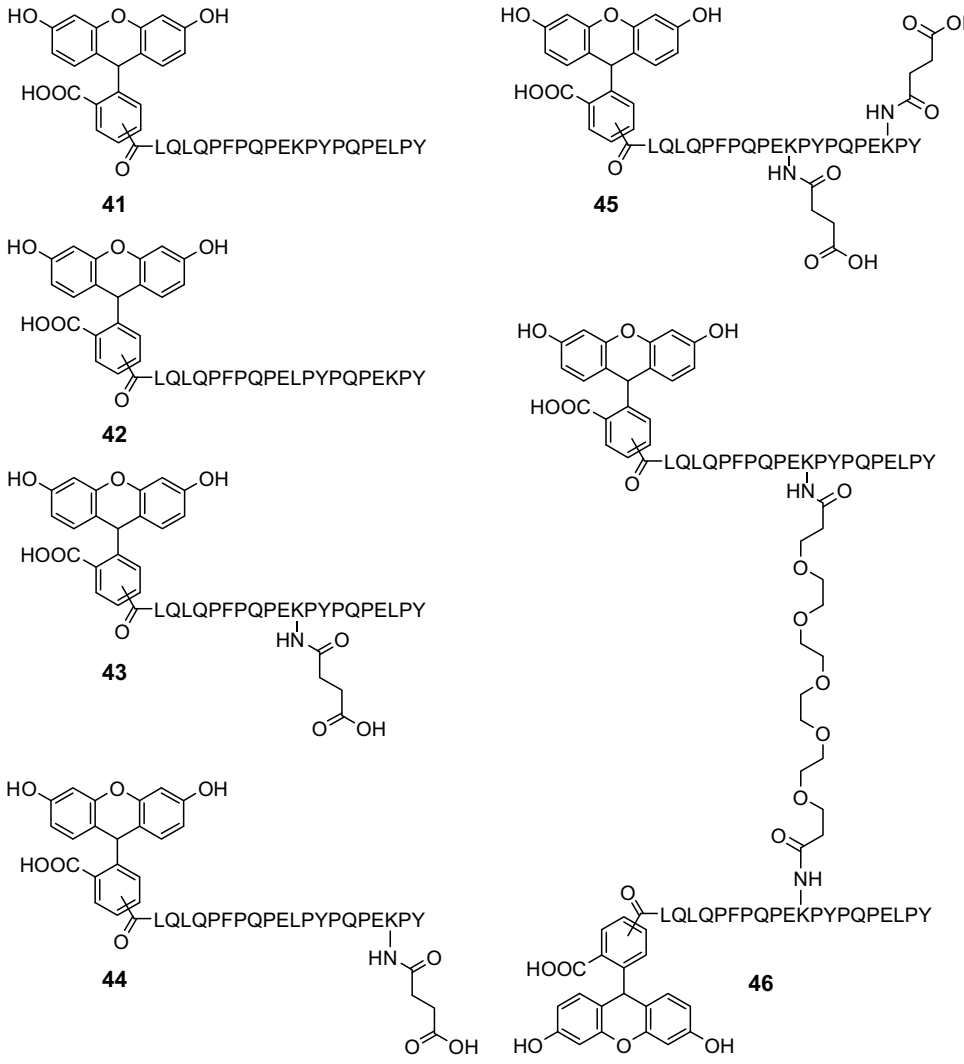
### 1.2.4 HLA-DQ2 blocking peptides

Much attention has been paid to development of inhibitors for the tTG enzyme. However, potential CD therapies involving the blocking of the HLA-DQ2 protein have, until recently, not been investigated. In 2004, a crystal structure of HLA-DQ2 in complex with the gluten peptide QLQ**PFPQPELPY** (epitope in bold) was published by Kim *et al.*[25] In this crystal structure the authors observed that the prolines at the p5 and p8 positions had limited interaction with the HLA-DQ2 molecule and are solvent-exposed. Therefore, it was suggested that these residues could be modified with additional side chains to disrupt T cell recognition of the peptide/HLA-DQ2 complex. Such peptides, the so-called HLA-DQ2 blockers, should bind better than the toxic gluten peptides and when in complex with HLA-DQ2, should not be recognized by gluten-specific $CD4^+$ T cells, so that the immune response to gluten is completely inhibited.

Khosla and co-workers developed a series of potent HLA-DQ2 blockers, starting from a T cell stimulatory 33-mer peptide,[92] with the sequence LQLQPFPQPELPYPQPELPYPQPELPYPQPQPF (**39**). This peptide consists of overlapping copies of the epitopes PFPQPELPY, PQPELPYPQ and PYPQPELPY and has a much higher affinity for HLA-DQ2 than the individual epitopes. Multiple truncated versions of this peptide were synthesized and evaluated for their HLA-DQ2 affinity. It was found that peptide LQLQPFPQPELPYPQPELPY (**40**) showed comparable activity with respect to **39**. This peptide was subsequently modified at L1 and L18, which corresponds to the p5 position of epitope PQPELPYPQ, of which

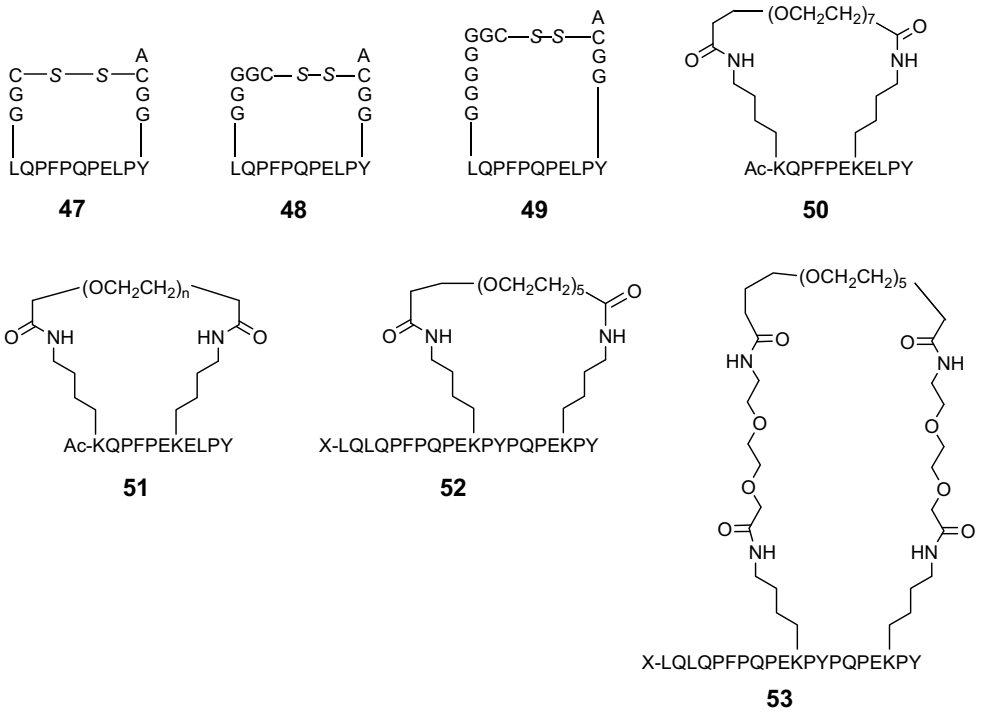
two copies are present in **40**. Thus, 6 peptides with large side chains at these positions were prepared (Figure 13).

**Figure 13** Blockers for the HLA-DQ2 protein based on a T cell stimulatory 33-mer. The fluorescein moiety was attached for the assessment of the binding properties.



It was found that peptides **41** and **42** did not bind well to HLA-DQ2, probably due to salt bridge formation between the side chains of the Lys and the preceding Glu residue. When a succinyl moiety was attached (**43**, **44** and **45**), the binding affinity improved considerably. Dimer **46** showed even more affinity together with faster binding. The blocking properties of **45** and **46** were assessed and these peptides were shown to be able to compete with the unmodified peptide and effectively inhibit T cell proliferation, based on steric hindrance on the side of the solvent-exposed face of the peptide/HLA-DQ2 complex. As a consequence, the TCR failed to recognize the complex. Similarly, dimeric and cyclic gluten peptide analogs, designed to inhibit T cell recognition based on steric hindrance, were evaluated for their ability to inhibit T cell proliferation.[93] These cyclic peptides were made by either incorporating cysteine residues to the N- and C-terminus of a known epitope and linking the cysteines together via disulfide bridges, or by replacing leucine by lysine residues and connecting the side chain amines of these lysines with polyethyleneglycol linkers (Figure 14).

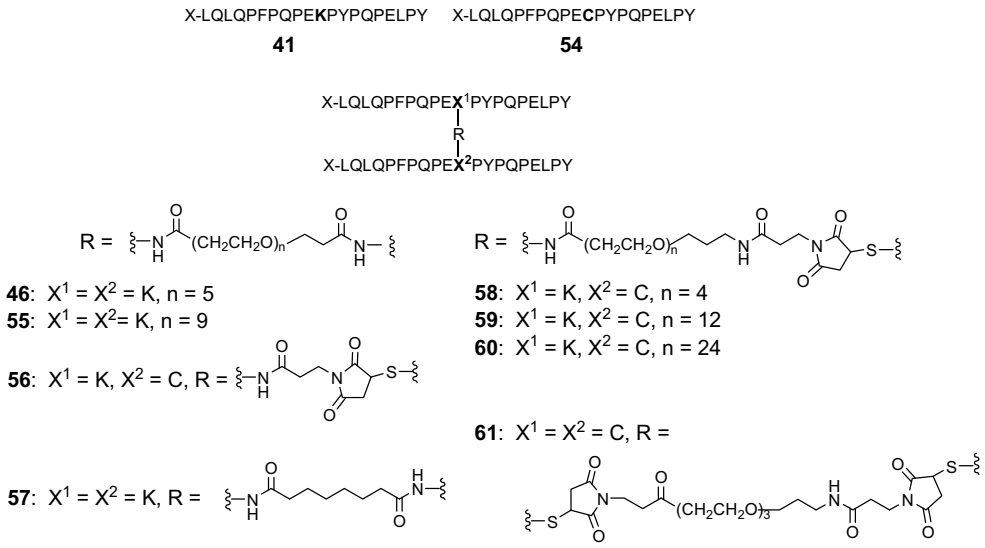
**Figure 14** Cyclic gluten-based peptides as HLA-DQ2-blockers. X = carboxyfluorescein, as in Figure 13. Peptide **51** is a heterogeneous mixture with n = ~11. Disulfide bridges are indicated in italic.




These peptides were tested for their HLA-DQ2 binding affinity and ability to induce T cell proliferation. All peptides were found to have a good affinity ($IC_{50}$ between 0.2

and 4.2 μM, as compared to epitope LQPFPQPELPY which displayed an $IC_{50}$ value of 2.5 μM), with **52** and **53** being the best binders with $IC_{50}$ values of 0.2 and 0.3 μM, respectively. Compounds **47**-**49** were recognized by specific T cells, probably resulting from the instability of the disulfide bridges under the assay conditions. Peptides **50**-**53** did inhibit T cell proliferation, with peptides **52** and **53** again being the most effective inhibitors. The peptide dimers were based on dimer **46** (Figure 13), but the linkers connecting the two peptide sequences were varied. As controls, modified monomers **41** and **54** were used, in which a leucine residue was replaced with either a lysine or cysteine (Figure 15), which showed equal affinity for HLA-DQ2 as the unmodified peptide.[92]

**Figure 15** Dimeric peptides as HLA-DQ2 blockers. X = carboxyfluorescein, as in Figure 13. For peptides **41** and **54**, the bold residue indicates the amino acid substitution.

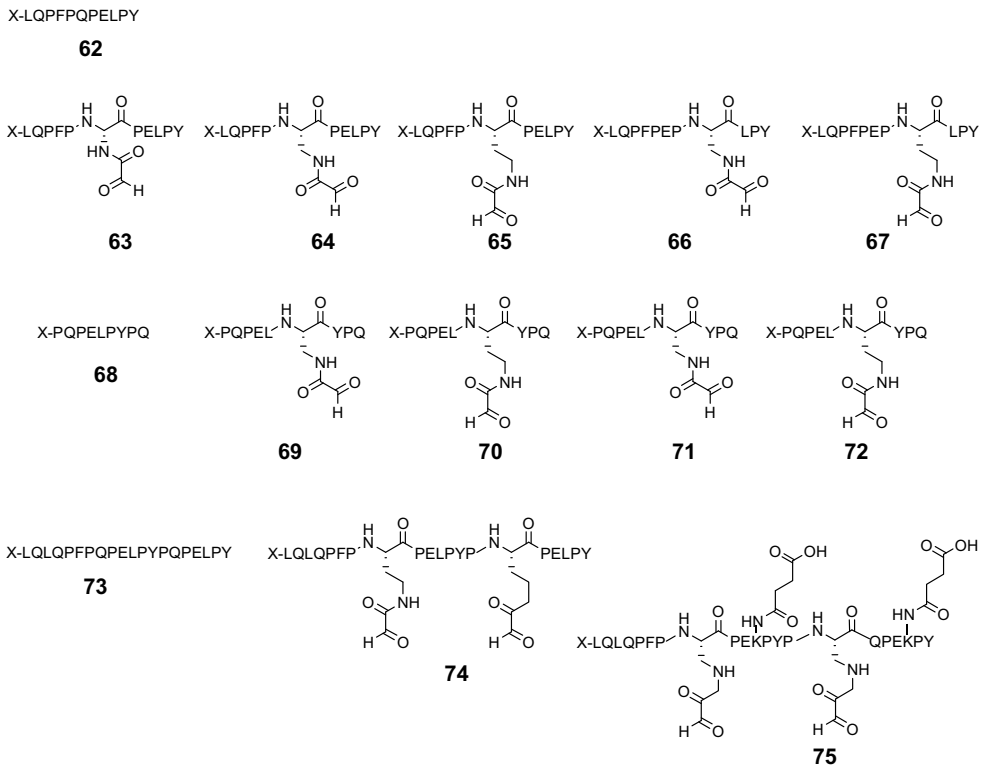



All peptides were able to form 1:2 complexes with HLA-DQ2. Peptides **56** and **58**-**61**, all containing thioether linkages between the cysteine residue and the linker, did stimulate T cell proliferation, again suggesting that a thioether bond is not stable under the conditions of the assay. Compounds **46**, **55** and **57** were able to inhibit T cell proliferation with approximately the same effectivity.

Another approach was based on covalent modification of the HLA-DQ2 molecule with, as a consequence, that the peptide-binding groove is unavailable for gluten peptides.[77] For this purpose, a series of aldehyde-containing gluten peptide analogs was prepared. The aldehyde reacts with a nearby amine of a lysine of HLA-DQ2, forming an imine. This might stabilize the peptide-HLA-DQ2 complex. It has been

known that Lys71 of the β-chain is in close proximity to the p4 and p6 positions in the HLA-DQ2 binding pockets.[25] Therefore, the peptides were designed in such a way that the aldehydes were present at positions p4 (**63**-**65**, Figure 16), p6 (**66**, **67**, **69**-**72**) or both (**74** and **75**).

**Figure 16** Aldehyde-containing gluten peptide analogs. X = carboxyfluorescein, as in Figure 13.




Of the short peptides **62**-**72**, peptide **64** showed the highest affinity, probably due to its resemblance to the glutamine residue in the unmodified peptide **62**. Also, peptide **64** was shown to enhance T cell proliferation as compared to **62** ($IC_{50}$ = 3.3 μM for **62**, $IC_{50}$ = 1.0 μM for **64**). Peptide **73**, a high affinity HLA-DQ2 binder also used in the previous study,[92] was modified with two aldehydes, and indeed, significantly higher binding affinity was seen. Peptide **74** showed a decrease in T cell activation ($IC_{50}$ = 0.039 μM) as compared to **73** ($IC_{50}$ = 0.016 μM) and thus may represent a good blocker. Based on peptide **45** from the previous study,[92] dialdehyde-containing peptide **75** was synthesized. However, it was found to block T cell proliferation less effectively than peptides **45** and **46**.

Anderson *et al.* investigated the effect of amino acid substitutions in the known toxic gluten peptide QLQPFPQPELPYPQPQS, originating from α-gliadin.[94] It was found that substitutions in the central PELP sequence with lysine, arginine or glutamine rendered the peptide inactive towards recognition by T cells. The parts flanking the central sequence were found to be less sensitive to amino acid substitutions. The binding properties of the generated peptides were not evaluated.

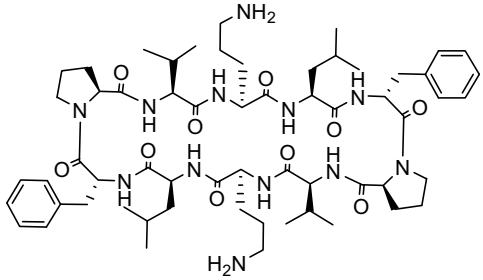
In summary, there are several possible therapies to treat the symptoms of CD, aimed at different key factors in CD pathogenesis. However, except for the enzyme therapy described in section 1.2.2, none of these therapies are in a stage where they can be assayed in clinical trials. Several tTG inhibitors discussed above show affinity for other enzymes next to tTG. Moreover, although it has been proven that some irreversible tTG inhibitors can be administered in animals without any toxic effects, irreversibly inhibiting this enzyme in humans may have an adverse effect as tTG is involved in many biological processes. The majority of the reversible inhibitors have been shown to be selective for tTG, combined with low cytotoxicity. The most promising approach in this matter thus seems to be reversible inhibition of tTG. The HLA-DQ2 blocker approach is not as thoroughly investigated as the tTG inhibitor approach, but shows promise in that HLA-DQ2 selectivity can be obtained by modifying natural HLA-DQ2 ligands, as has been shown in section 1.2.4. Moreover, a crystal structure of HLA-DQ2 in complex with a gluten peptide exists,[25] making design of high affinity ligands for HLA-DQ2 possible. In this Chapter only the most frequently investigated approaches were described. Several new strategies are emerging,[95] such as modulation of the T cell response. For example, De Vincenzi and co-workers identified a peptide from durum wheat, which is considered toxic for CD patients, that turned out to be inhibitory for some processes occurring downstream of T cell recognition, such as the inflammatory response that causes small intestinal damage.[96] It was found that the T cell response was shifted from type Th1 to type Th2 with the release of different cytokines as a result. Therefore, the inflammatory process was inhibited. Another example is the use of regulatory T cells to inhibit the action of the $CD4^+$ T cells involved in CD pathogenesis.[97] Such new therapies may prove to be as promising as the therapies described here.

### 1.3 Aim and outline of this Thesis

The first part of this Thesis aims at elaborating on the HLA-DQ2 blocker approach. When the crystal structure of HLA-DQ2 with QLQ**PFPQPELPY** (epitope in bold) was studied, it was observed that besides the p5 and p8 positions,[25] the residue on the p3 position of the epitope does not have interaction with the HLA-DQ2 binding pocket as well. The p8 position, in contrast, shows this exposure to a lesser extent.

Therefore, high affinity HLA-DQ2 binding peptides which are modified at the p3 position may also be suitable as HLA-DQ2 blockers. **Chapter 2** describes the modification at the p3 and p5 positions of a known T cell stimulatory gluten peptide (QLQ**PFPQPELPY**PQ, epitope in bold). The prolines at these positions were replaced with azidoprolines, to which a side chain was attached *via* 'click chemistry'. These peptides were tested for their binding affinity for HLA-DQ2 and their T cell response inhibitory activity. In **Chapter 3** attempts to improve the binding affinity of HLA-DQ2-binding peptides are described. From earlier studies, high affinity HLA-DQ2 peptides have been identified and a lead sequence is further optimized using systematic substitutions with both natural and unnatural amino acids. In **Chapter 4** it is described how three high affinity gluten peptides identified from the studies in Chapter 3 are modified. The same approach was applied as in Chapter 2; the amino acid at the p3 position was replaced with the corresponding azido amino acid and various side chains were attached. These compounds were evaluated for their binding properties.
The second part of this Thesis deals with Gramicidin S (GS) (**76**, Figure 17), a cationic cyclic peptide antibiotic isolated from the soil bacterium *Bacillus Brevis*.[98] This peptide is described in more detail in the introduction of Chapter 5.

**Figure 17** The structural formula of Gramicidin S.



**76**

The occurrence of multiple-resistant bacterial strains poses a great problem to healthcare institutions, as infections with these bacteria are difficult to cure. Therefore, the development of new antibiotics is urgently needed to combat these bacterial strains. Gramicidin S has been shown to be effective against a number of bacterial strains, both Gram-positive and Gram-negative,[99] but also lyses human red blood cells, limiting its clinical use.[100] Modification of GS to tune its antibacterial and hemolytic activity appears to be an attractive strategy to develop new efficient antibiotics. In **Chapter 5**, the modification of GS with adamantyl amino acids is

described. The newly generated peptides are evaluated for their structural and biological properties. In **Chapter 6**, the modification of GS with morpholine amino acids (MAAs) is described, and their structural and biological properties are presented. **Chapter 7** summarizes Chapters 2-6 and discusses future directions for the research described in this Thesis.

## References


1. Rewers, M. *Gastroenterology* **2005**, *128*, S47-S51.
2. Barton, S. H.; Kelly, D. G.; Murray, J. A. *Gastroenterol. Clin. N. Am.* **2007**, *36*, 93-108.
3. Shan, L.; Qiao, S. W.; Arentz-Hansen, H.; Gray, G. M.; Sollid, L. M.; Khosla, C. *J. Proteome Res.* **2005**, *4*, 1732-1741.
4. Barker, J. M.; Liu, E. *Adv. Pediatr.* **2008**, *55*, 349-365.
5. Hill, I. D.; Dirks, M. H.; Liptak, G. S.; Colletti, R. B.; Fasano, A.; Guandalini, S.; Hoffenberg, E. J.; Horvath, K.; Murray, J. A.; Pivor, M.; Seidman, E. G. *J. Pediatr. Gastroenterol. Nutr.* **2005**, *40*, 1-19.
6. Marsh, M. N. *Immunol. Invest.* **1989**, *18*, 509-531.
7. Lanzini, A.; Lanzarotto, F.; Villanacci, V.; Mora, A.; Bertolazzi, S.; Turini, D.; Carella, G.; Malagoli, A.; Ferrantes, G.; Cesana, B. M.; Ricci, C. *Aliment. Pharmacol. Ther.* **2009**, *29*, 1299-1308.
8. Salmi, T. T.; Collin, P.; Reunala, T.; Mäki, M.; Kaukinen, K. *Dig. Liv. Dis.* **2009**, doi: 10.1016/j.dld.2009.04.004.
9. Briani, C.; Samaroo, D.; Alaedini, A. *Autoimmun. Rev.* **2008**, *7*, 644-650.
10. Van de Kamer, J. H.; Weijers, H. A.; Dicke, W. K. *Acta Paediatr.* **1953**, *42*, 223-231.
11. Holtmann, M. H.; Neurath, M. F. *Gut* **2004**, *53*, 1065-1067.
12. Feighery, C. *British Med. J.* **1999**, *319*, 236-239.
13. DuBois, P. C.; Van Heel, D. A. *Clin. Exp. Immunol.* **2008**, *153*, 162-173.
14. McAdam, S. N.; Sollid, L. M. *Gut* **2000**, *47*, 743-745.
15. Dewar, D.; Pereira, S. P.; Ciclitira, P. J. *Int. J. Biochem. Cell. Biol.* **2004**, *36*, 17-24.
16. Kagnoff, M. F. *Gastroenterology* **2005**, *128*, S10-S18.
17. Vader, W.; Kooy, Y.; Van Veelen, P.; De Ru, A.; Harris, D.; Benckhuijsen, W.; Peña, S.; Mearin, L.; Drijfhout, J. W.; Koning, F. *Gastroenterology* **2002**, *122*, 1729-1737.
18. Vader, W.; Stepniak, D.; Kooy, Y.; Mearin, L.; Thompson, A.; Van Rood, J. J.; Spaneij, L.; Koning, F. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12390-12395.

19. Kemppainen, T.; Janatuinen, E.; Holm, K.; Kosma, V. M.; Heikkinen, M.; Mäki, M.; Laurila, K.; Uusitupa, M.; Julkunen, R. *Scand. J. Gastroenterol.* **2007**, *42*, 54-59.
20. Wieser, H.; Seilmeier, W.; Eggert, M.; Beliz, H. D. *Z. Lebensm. Unters. Forshc.* **1983**, *177*, 457-460.
21. Hoffenberg, E. J.; Haas, J.; Drescher, A.; Barnhurst, R.; Osberg, I.; Bao, F.; Eisenbarth, G. *J. Pediatr.* **2000**, *137*, 361-366.
22. Hunt, K. A.; Monsuur, A. J.; McArdle, W. L.; Kumar, P. J.; Travis, S. P. L.; Walters, J. R. F.; Jewell, D. P.; Strachan, D. P.; Playford, R. J.; Wijmenga, C.; Van Heel, D. A. *Gut* **2006**, *55*, 969-972.
23. Koning, F. *Gastroenterology* **2005**, *129*, 1294-1301.
24. Anderson, M. W.; Gorski, J. *Biochemistry* **2005**, *44*, 5617-5624.
25. Kim, C. Y.; Quarsten, H.; Bergseng, E.; Khosla, C.; Sollid, L. M.; *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4175-4179.
26. Stepniak, D.; Wiesner, M.; De Ru, A. H.; Moustakas, A. K.; Drijfhout, J. W.; Papadopulos, G. K.; Van Veelen, P. A.; Koning, F. *J. Immunol.* **2008**, *180*, 3268-3278.
27. Bergseng, E.; Xia, J.; Kim, C. Y.; Khosla, C.; Sollid, L. M. *J. Biol. Chem.* **2005**, *280*, 21791-21796.
28. Ferrante, A.; Gorski, J. *J. Immunol.* **2007**, *178*, 7181-7189.
29. Fesus, L.; Piancentini, M. *Trends Biochem. Sci.* **2002**, *27*, 534-539.
30. Qiao, S. W.; Piper, J.; Haraldsen, G.; Oynebraten, I.; Fleckenstein, B.; Molberg, Ø.; Khosla, C.; Sollid, L. M. *J. Immunol.* **2005**, *174*, 1657-1663.
31. Arentz-Hansen, H.; Korner, R.; Molberg, Ø.; Quarsten, H.; Vader, W.; Kooy, Y. M. C.; Lundin, K. E. A.; Koning, F.; Roepstorff, P.; Sollid, L. M.; McAdam, S. N. *J. Exp. Med.* **2000**, *191*, 603-612.
32. Schuppan, D.; Dieterich, W. *Chem. Biol.* **2003**, *10*, 199-201.
33. Elli, L.; Bergamini, C. M.; Bardella, M. T.; Schuppan, D. *Dig. Liv. Dis.* **2009**, *41*, 541-550.
34. Lai, T. S.; Liu, Y. S.; Tucker, T.; Daniel, K. R.; Sane, D. C.; Toone, E.; Burke, J. R.; Strittmatter, W. J.; Greenberg, C. S. *Chem. Biol.* **2008**, *15*, 969-978.
35. Freitag, T.; Schulze-Koops, H.; Niedobitek, G.; Melino, G.; Schuppan, D. *Autoimmun. Rev.* **2004**, *3*, 103-110.
36. Mutter, M.; Wöhr, T.; Gioria, S.; Keller, M. *Biopolymers* **1999**, *51*, 121-128.
37. Vader, L. W.; De Ru, A.; Van der Wal, Y.; Kooy, Y. M. C.; Benckhuijsen, W.; Mearin, M. L.; Drijfhout, J. W.; Van Veelen, P.; Koning, F. *J. Exp. Med.* **2002**, *195*, 643-649.

38. Stamnaes, J.; Fleckenstein, B.; Sollid, L. M. *Biochim. Biophys. Acta* **2008**, *1784*, 1804-1811.
39. Esposito, C.; Caputo, I.; Troncone, R. *Curr. Med. Chem.* **2007**, *14*, 2572-2580.
40. Indeed, it has been found that nondeamidated peptides can induce immune reponses, especially in children suffering from CD. See ref. 17.
41. a) Gianfrani, C.; Troncone, R.; La Cava, A. *Mini Rev. Med. Chem.* **2008**, *8*, 129-134.
42. Stepniak, D.; Vader, L. W.; Kooy, Y.; Van Veelen, P.; Moustakas, A.; Papandreou, N. A.; Eliopoulos, E.; Drijfhout, J. W.; Papadopoulos, G. K.; Koning, F. *Immunogenetics* **2005**, *57*, 8-15.
43. Molberg, Ø.; McAdam, S.; Körner, R.; Quarsten, H.; Kristiansen, C.; Madsen, L.; Fugger, L.; Scott, H.; Norén, O.; Roepstorff, P.; Lundin.; K. E. A.; Sjöström, H.; Sollid, L. M. *Nature Med.* **1998**, *4,* 713-717.
44. Van de Wal, Y.; Kooy, Y. M. C.; Van Veelen, P.; Peña, A. S.; Mearin, M. L.; Papadopoulos, G. K.; Koning, F. *J. Immunol.* **1998**, *161*, 1585-1588.
45. Van de Wal, Y.; Kooy, Y. M. C.; Drijfhout, J. W.; Amons, R.; Koning, F. *Immunogenetics* **1996**, *44*, 246-253.
46. Johansen, B. H.; Vartdal, F.; Eriksen, J. A.; Thorsby, E.; Sollid, L. M. *Int. Immunol.* **1996**, *8*, 177-182.
47. Kwok, W. W.; Domeier, M. E.; Raymond, F. C.; Byers, P.; Nepom, G. T. *J. Immunol.* **1996**, *156*, 2171-2177.
48. Van de Wal, Y.; Kooy, Y.; Van Veelen, P.; Peña, S.; Mearin, L.; Molberg, Ø.; Lundin, L.; Mutis. T.; Benckhuijsen, W.; Drijfhout, J. W.; Koning, F. *Proc. Natl. Acad. Sci. USA,* **1998**, *95*, 10050-10054.
49. Anderson, R. P.; Degano, P.; Godkin, A. J.; Jewell, D. P.; Hill, A. V. *Nature Med.* **2000**, *6*, 337-342.
50. Vader, W.; Stepniak, D.; Bunnik, E. M.; Kooy, Y.; De Haan, W.; Drijfhout, J. W.; van Veelen, P. A.; Koning, F. *Gastroenterology* **2003**, *125*, 1105-1113.
51. Caputo, I.; Barone, M. V.; Martuciello, S.; Lepretti, M.; Esposito, C. *Amino Acids* **2009**, *36*, 693-699.
52. Collin, P.; Thorell, L.; Kaukinen, K.; Mäki, M. *Aliment. Pharmacol. Ther.* **2004**, *19*, 1277-1283.
53. a) Zarkadas, M.; Cranney, A.; Case, S.; Molloy, M.; Switzer, C.; Graham, I. D.; Butzner, J. D.; Rashid, M.; Warren, R. E.; Burrows, V. *J. Hum. Nutr. Diet.* **2006**, *19*, 41-49. b) Lee, A.; Newman, J. M. *J. Am. Diet. Assoc.* **2003**, *103*, 1533-1535.
54. Kemppainen, T.; Janatuinen, E.; Holm, K.; Kosma, V. M.; Heikkinen, M.; Mäki, M.; Laurila, K.; Uusitupa, M.; Julkunen, R. *Scand. J. Gastroenterol.* **2007**, *42*, 54-59.

55. Spaneij-Dekking, L.; Kooy-Winkelaar, Y.; Koning, F. *N. Engl. J. Med.* **2005**, *353*, 1748-1749.
56. Shan, L.; Molberg, Ø.; Parrot, I.; Hausch, F.; Filiz, F.; Gray, G. M.; Sollid, L. M.; Khosla, C. *Science* **2002**, *297*, 2275-2279.
57. Polgar, L. *Cell. Mol. Life Sci.* **2002**, *59*, 349-362.
58. Fulop, V.; Bocskei, Z.; Polgar, L. *Cell* **1998**, *94*, 161-170.
59. a) Shan, L.; Martin, T.; Sollid, L. M.; Gray, G. M.; Khosla, C. *Biochem. J.* **2004**, *383*, 311-318. b) Gass, J.; Ehren, J.; Strohmeier, G.; Isaacs, I.; Khosla, C. *Biotech. Bioeng.* **2005**, *92*, 674-684. c) Piper, J. L.; Gray, G. M.; Khosla, C. *J. Pharm. Exp. Ther.* **2004**, *311*, 213-219. d) Shan, L.; Qiao, S. W.; Arentz-Hansen, H.; Molberg, Ø.; Gray, G. M.; Sollid, L. M.; Khosla, C. *J. Proteome Res.* **2005**, *4*, 1732-1741. e) Matysiak-Budnik, T.; Candalh, C.; Cellier, C.; Dugave, C.; Namane, A.; Vidal-Martinez, T.; Cerf-Benussan, N.; Heyman, M. *Gastroenterology* **2005**, *129*, 786-796.
60. Siegel, M.; Bethune, M. T.; Gass, J.; Ehren, J.; Xia, J.; Johannsen, A.; Stuge, T. B.; Gray, G. M.; Lee, P. P.; Khosla, C. *Chem. Biol.* **2006**, *13*, 649-658.
61. Gass, J.; Vora, H.; Bethune, M. T.; Gray, G. M.; Khosla, C. *J. Pharm. Exp. Ther.* **2006**, *318*, 1178-1186.
62. Gass, J.; Bethune, M. T.; Siegel, M.; Spencer, A.; Khosla, C. *Gastroenterology* **2007**, *133*, 472-480.
63. Ehren, J.; Morón, B.; Martin, E.; Bethune, M. T.; Gray, G. M.; Khosla, C. *PLoS One* **2009**, *4*, e6313.
64. Ehren, J.; Govindarajan, S.; Moron, B.; Minshull, J.; Khosla, C. *Prot. Eng. Des. Sel.* **2008**, *21*, 699-707.
65. Stepniak, D.; Spaneij-Dekking, L.; Mitea, C.; Moester, M.; De Ru, A.; Baak-Pablo, R.; Van Veelen, P. A.; Edens, L.; Koning, F. *Am. J. Physiol. Gastrointest. Liver. Phys.* **2006**, *291*, G621-G629.
66. Mitea, C.; Havenaar, R.; Drijfhout, J. W.; Edens, L.; Dekking, L.; Koning, F. *Gut* **2008**, *57*, 25-32.
67. Siegel, M.; Khosla, C. *Pharmacol. Ther.* **2007**, *115*, 232-245.
68. Lorand, L.; Parameswaran, K. N.; Stenberg, P.; Tong, Y. S.; Velasco, P. T.; Jonsson, N. A.; Mikiver, L.; Moses, P. *Biochemistry* **1979**, *18*, 1756-1765.
69. Lorand, L.; Rule, N. G.; Ong, H. H.; Furlanetto, R.; Jacobsen, A.; Downey, J.; Oner, N.; Bruner-Lorand, J. *Biochemistry* **1968**, *7*, 1214-1223.
70. Lee, N. K.; Fesus, L.; Yancey, S. T.; Girard, J. E.; Chung, S. I. *J. Biol. Chem.* **1985**, *260*, 14689-14694.
71. Jeitner, T. M.; Delikatny, E. J.; Ahlqvist, J.; Capper, H.; Cooper, A. J. L. *Biochem. Pharm.* **2005**, *69*, 961-970.

72. Lorand, L.; Conrad, S. M. *Mol. Cell. Biochem.* **1984**, *58*, 9-35.
73. Lai, T. S.; Slaughter, T. F.; Peoples, K. A.; Hettasch, J. M.; Greenberg, C. S. *J. Biol. Chem.* **1998**, *273*, 1776-1781.
74. Case, A.; Ni, J.; Yeh, L. A.; Stein, R. L. *Anal. Biochem.* **2005**, *338*, 237-244.
75. Duval, E.; Case, A.; Stein, R. L.; Cuny, G. D. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1885-1889.
76. Case, A.; Stein, L. *Biochemistry* **2007**, *46*, 1106-1115.
77. Siegel, M.; Xia, J.; Khosla, C. *Bioorg. Med. Chem.* **2007**, *15*, 6253-6261.
78. Folk, J. E.; Finlayson, J. S.; *Adv. Prot. Chem.* **1977**, *31*, 1-133.
79. Pardin, C.; Pelletier, J, N.; Dubell, W. D.; Keillor, J. W. *J. Org. Chem.* **2008**, *73*, 5766-5775.
80. Pardin, C.; Roy, I.; Lubell, W. D.; Keillor, J. W. *Chem. Biol. Drug Des.* **2008**, *72*, 189-196.
81. Castelhano, A. L.; Billedeau, R.; Pliura, D. H.; Bonaventura, B. J.; Krantz, A. *Bioorg. Chem.* **1988**, *16*, 335-340.
82. Marrano, C.; de Macedo, P.; Keillor, J. W. *Bioorg. Med. Chem.* **2001**, *9*, 1923-1928.
83. Marrano, C.; de Macedo, P.; Gagnon, P.; Lapierre, D.; Gravel, C.; Keillor, J. W. *Bioorg. Med. Chem.* **2001**, *9*, 3231-3241.
84. Hausch, F.; Halttunen, T.; Mäki, M.; Khosla, C. *Chem. Biol.* **2003**, *10*, 225-231.
85. Freund, K. F.; Doshi, K. P.; Gaul, S. L.; Claremon, D. A.; Remy, D. C.; Baldwin, J. J.; Pitzenberger, S. M.; Stern, A. M. *Biochemistry* **1994**, *33*, 10109-10119.
86. Brocklehurst, K.; Malthouse, P. G. *Biochem. J.* **1978**, *175*, 761-764.
87. Choi, K.; Siegel, M.; Piper, J. L.; Yuan, L.; Cho, E.; Strnad, P.; Omary, B.; Rich, K. M.; Khosla, C. *Chem. Biol.* **2005**, *12*, 469-475.
88. Pardin, C.; Gillet, S. M. F. G.; Keillor, J. W. *Bioorg. Med. Chem.* **2006**, *14*, 8379-8385.
89. Halim, D.; Caron, C.; Keillor, J. W. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 305-308.
90. Griffin, M.; Mongeot, A.; Collighan, R.; Saint, A. E.; Jons, R. A.; Coutts, I. G. C.; Rathbone, D. L. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5559-5562.
91. Watts, R. E.; Siegel, M.; Khosla, C. *J. Med. Chem.* **2006**, *49*, 7493-7501.
92. Xia, J.; Siegel, M.; Bergseng, E.; Sollid, L. M.; Khosla, C. *J. Am. Chem. Soc.* **2006**, *128*, 1859-1867.
93. Xia, J.; Bergseng, E.; Fleckenstein, B.; Siegel, M.; Kim, C. Y.; Khosla, C.; Sollid, L. M. *Bioorg. Med. Chem.* **2007**, *15*, 6565-6573.
94. Anderson, R. P.; Van Heel, D. A.; Tye-Din, J. A.; Jewell, D. P.; Hill, A. V. S. *Gut* **2006**, *55*, 485-491.
95. Lerner, A.; *Autoimmunity Rev.* **2009**, doi: 10.1016/j.autrev.2009.05.002.

96. a) Silano, M.; Di Benedetto, R.; Trecca, A.; Arrabito, G.; Leonardi, F.; De Vincenzi, M. *Pediatr. Res.* **2007**, *61*, 67-71. b) Silano, M.; Leonardi, F.; Trecca, A.; Mancini, E.; Di Benedetto, R.; De Vincenzi, M. *Scand. J. Gastroenterol.* **2007**, *42*, 786-787.
97. Gianfrani, C.; Levings, M. K.; Sartirana, C.; Mazzarella, G.; Barba, G.; Zanzi, D.; Camarca, A.; Iaquinto, G.; Giardullo, N.; Auricchio, S.; Troncone, R.; Roncarolo, M. G. *J. Immunol.* **2006**, *177*, 4178-4186.
98. Gause, G. F.; Brazhnikova, M. G. *Nature* **1944**, *154*, 703.
99. Kondejewski, L. H.; Farmer, S. W.; Wishart, D. S.; Hancock, R. E. W.; Hodges, R. S. *Int. J. Peptide Protein Res.* **1996**, *47*, 460-466.
100. O' Grady, F.; Greenwood, D. in *Antibiotic and chemotherapy: anti-infective agents and their use in therapy*; ed. O' Grady, F. (Churchill Livingstone, New York); **1997**, pp 337-338.