Predictive tools in the risk assessment of new proteins in GMOs: the case of Celiac Disease

Frits Koning
LUMC, Leiden, The Netherlands
Celiac Disease Consortium
Gluten proteins in wheat

HLA-DQ2/8

T-cells
Gluten specific T cell response in the small intestine

Gluten peptide → HLA-DQ2(8) → APC

→ Tissue damage
Gluten peptides generated by pepsin

It takes three to tango!!!!

Tissue damage: release of intracellular tTG

T Cells

- Enhanced binding
- Amplification of the immune response

tTG

deamidation

Q → E

tissue repair

gluten peptides

HLA-DQ2(8)

APC

T Cells
The specificity of tTG is determined by proline, the 2\textsuperscript{nd} most abundant aa in gluten.

Characteristic gluten sequences:

- QP: no modification
- QXP: yes
- QXXP: no
- QXPY or QXPF: yes

```
LGQQQPFPPQQPYQPQPQPFPFPQLYPYLQLQPFPQQL
LGQEQPFPPPEQPYQPQPQPFPFPSELQPYLQLQPFPQQL
```
Predict toxic gluten sequences?

<table>
<thead>
<tr>
<th>Gluten</th>
<th>Hordein</th>
<th>Secalin</th>
<th>Avenin</th>
<th>Tcells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td>Barley</td>
<td>Rye</td>
<td>Oats</td>
<td></td>
</tr>
</tbody>
</table>

Specificity of tissue transglutaminase explains cereal toxicity in celiac disease.

Identification of T cell stimulatory peptides in cereals

**Gliadin (wheat):**

QLQPFPQPQLPYPQPQ

PFPQPQLPY

PQPQLPYPQ

**Secalin (rye):**

PQQPFQPQQPFPQSQ

PFPQPQQPFPQ

PQPQPQFPQF

PQPQFPQFPQ
DQ2-glia-α2 recognition

TRAV26-01
TRBV7-02+

Conserved β-chain footprint

T cell receptor
Peptide
DQ2

Petersen et al, NMSB 2014
DQ2-glia-a2 recognition: PQPQLPYPQ

Broughton Immunity 2012; Petersen et al, NMSB 2014; Petersen JI 2015
Bona fide toxicity of gluten for patients with celiac disease

• Well defined
• Mechanism underlying toxicity clear
RA of (novel) proteins: celiac disease

**Fig 2. Search for sequence identity**

- **100% match with T-cell stimulatory epitope**
  - **Hazard identified**

- **Concerns raised from a partial match* with T-cell stimulatory epitope**
  - **Further investigations are necessary**

- **No concerns raised from a partial match* with T-cell stimulatory epitope**
  - **No hazard identified**

* A partial match with a known T cell-stimulatory peptide raises concern because of the position and nature of the identical amino acids.
## Celiac disease — DQ2 T-cell epitopes

<table>
<thead>
<tr>
<th>Epitope</th>
<th>Motif</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DQ2.5-glia-α1a</td>
<td>PFPQPQQLPY</td>
<td>Arentz-Hansen et al. (2000)</td>
</tr>
<tr>
<td>DQ2.5-glia-α1b</td>
<td>PYQPQQLPY</td>
<td>Arentz-Hansen et al. (2002)</td>
</tr>
<tr>
<td>DQ2.5-glia-α2</td>
<td>PQPQLPPQ</td>
<td>Arentz-Hansen et al. (2000)</td>
</tr>
<tr>
<td>DQ2.5-glia-α3</td>
<td>ELPY</td>
<td>Vader et al. (2002b)</td>
</tr>
<tr>
<td>DQ2.5-glia-γ1</td>
<td>QQFQ</td>
<td>Sjöström et al. (1998)</td>
</tr>
<tr>
<td>DQ2.5-glia-γ2</td>
<td>QQFA</td>
<td>Qiao et al. (2005), Vader et al. (2002b)</td>
</tr>
<tr>
<td>DQ2.5-glia-γ3</td>
<td>SVEQ</td>
<td>Arentz-Hansen et al. (2002)</td>
</tr>
<tr>
<td>DQ2.5-glia-γ4a</td>
<td>QQFA</td>
<td>Qiao et al. (2005)</td>
</tr>
<tr>
<td>DQ2.5-glia-γ4b</td>
<td>QQFA</td>
<td>Arentz-Hansen et al. (2002)</td>
</tr>
<tr>
<td>DQ2.5-glia-γ4c</td>
<td>QFAQ</td>
<td>Qiao (unpublished)</td>
</tr>
<tr>
<td>DQ2.5-glia-γ4d</td>
<td>EQFQ</td>
<td>Arentz-Hansen et al. (2002)</td>
</tr>
<tr>
<td>DQ2.5-glia-γ5</td>
<td>FSVEQ</td>
<td>Tye-Din et al. (2010)</td>
</tr>
<tr>
<td>DQ2.5-glia-ω1</td>
<td>FSVEQ</td>
<td>Arentz-Hansen et al. (2002)</td>
</tr>
<tr>
<td>DQ2.5-glia-ω2</td>
<td>FSVEQ</td>
<td>Tye-Din et al. (2010)</td>
</tr>
<tr>
<td>DQ2.2-glut-L1</td>
<td>EPFPFP</td>
<td>Vader et al. (2002b)</td>
</tr>
<tr>
<td>DQ2.2-glut-L2</td>
<td>PFPFPF</td>
<td>Stepniak et al. (2005), Vader et al. (2002b)</td>
</tr>
<tr>
<td>DQ2.5-hor-1</td>
<td>EFPFPF</td>
<td>Tye-Din et al. (2010), Vader et al. (2003)</td>
</tr>
<tr>
<td>DQ2.5-hor-2</td>
<td>EPFPFP</td>
<td>Vader et al. (2003)</td>
</tr>
<tr>
<td>DQ2.5-sec-1</td>
<td>PQPQQPQPQ</td>
<td>Tye-Din et al. (2010), Vader et al. (2003)</td>
</tr>
<tr>
<td>DQ2.5-sec-2</td>
<td>PQPQQPQPQ</td>
<td>Vader et al. (2003)</td>
</tr>
<tr>
<td>DQ2.5-ave-1</td>
<td>PYPEQFQQPF</td>
<td>Arentz-Hansen et al. (2004), Vader et al. (2003)</td>
</tr>
<tr>
<td>DQ2.5-ave-1b</td>
<td>PYPEQFQQPF</td>
<td>Arentz-Hansen et al. (2004), Vader et al. (2003)</td>
</tr>
</tbody>
</table>

Sollid et al., 2012. Immunogenetics, 64, 455-460
Q-X-P-X

- PFPQPQLPY
- PQPQLPYPQ

- PXP in addition to QXPX is associated with the most immunogenic epitopes
- In contrast: PP is never found in T cell epitopes
- Positively charged amino acids in general diminish likelihood of DQ-binding and T cell recognition. Positive charge at p1, p4, p6, p7 and p9 bad for DQ-binding.
# Celiac disease — DQ8 T-cell epitopes

*Sollid et al., 2012. Immunogenetics, 64, 455-460*

## DQ8 restricted epitopes

<table>
<thead>
<tr>
<th>Epitope</th>
<th>Motif</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DQ8-glia-(\alpha)1</td>
<td>QGSFQPSQQ</td>
<td>van de Wal et al. (1998b)</td>
</tr>
<tr>
<td>DQ8-glia-(\gamma)1a</td>
<td>QQPQQFPQP</td>
<td>Tollefsen et al. (2006)</td>
</tr>
<tr>
<td>DQ8-glia-(\gamma)1b</td>
<td>QQPQQYPQP</td>
<td>Tollefsen et al. (2006)</td>
</tr>
<tr>
<td>DQ8-glut-H1</td>
<td>QGYYPTSQP</td>
<td>van de Wal et al. (1999)</td>
</tr>
</tbody>
</table>

Partial matches without the Q/E-X1-P-X2 to be investigated
Partial matches: Q/E-X1-P-X2 motif is present

\[
\text{PFQPQLPY and ALPLTQLPA}
\]

4 identical, two invisible, one conservative: POTENTIAL HAZARD

\[
\text{PQPQLPYPQ and PLTQLPASR}
\]

4 identical, one conservative BUT Y > A, P > S and Q > R prohibit recognition: NO HAZARD
Partial matches: Q/E-X1-P-X2 motif is NOT present

QGSFQPSQQ and
EGSIQAGQQ

5 identical, one conservative, one enhances binding:
POTENTIAL HAZARD

QGSFQPSQQ and
QGLFSPSAQ

6 identical BUT
Critical T cell receptor contact residues differ:
NO HAZARD
Peptide binding and Modelling
PFPQP ELPY
PLLMQ ALPM
Molecular mimicry?
Cross-reactivity between microbial antigens and gluten epitopes?

<table>
<thead>
<tr>
<th>Protein</th>
<th>Peptide Sequence</th>
<th>Matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glia-α1</td>
<td>PFPQPELPY</td>
<td>6</td>
</tr>
<tr>
<td>Bacterial</td>
<td>peptide 1</td>
<td>5</td>
</tr>
<tr>
<td>Bacterial</td>
<td>peptide 2</td>
<td>5</td>
</tr>
<tr>
<td>Glia-α2</td>
<td>PQPELPYPQ</td>
<td>7</td>
</tr>
<tr>
<td>Bacterial</td>
<td>peptide 3</td>
<td>5</td>
</tr>
<tr>
<td>Bacterial</td>
<td>peptide 4</td>
<td>5</td>
</tr>
</tbody>
</table>

All have the **Q/E-X1-P-X2** motif
Conclusion

Potential antigenicity can be predicted

AND YES

There are bacterial peptides that trigger gluten-specific T cells