Application of Peptide Modeling in Celiac Disease Risk Assessment

Distribution of HLA-DQ2.5 Q/EX$_1$PX$_2$ Motif

From Whole Genome Translation

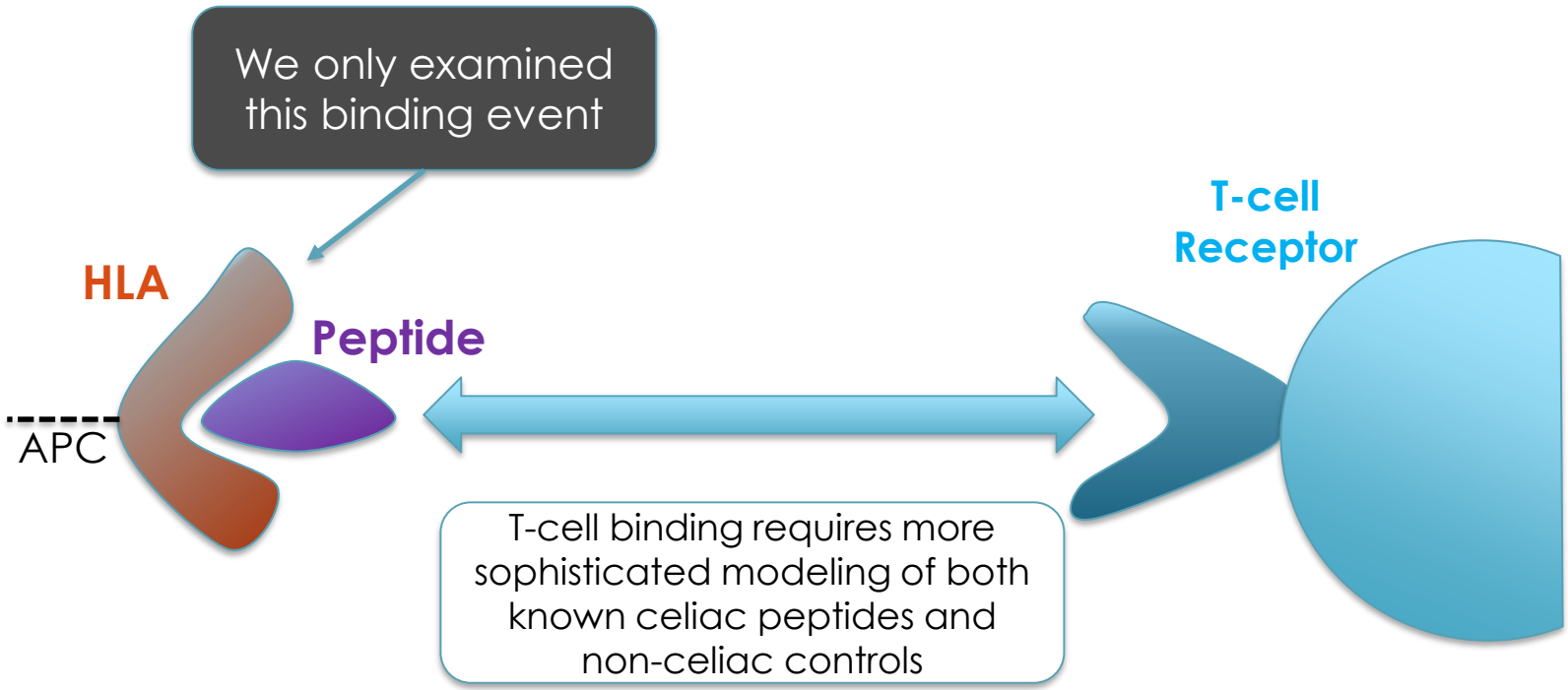
Organism	Total number of 9mers containing Q/EX$_1$PX$_2$ motif starting at the 4th or 6 th amino acid	Percent of 9mers possibly subject to peptide modeling
Triticum Urartu	6682	57.80
Wheat	46248	57.11
Maize	19080	53.97
Rice	13748	56.27
Soybean	21563	57.77
Sunflower	21244	55.39
Tomato	12120	56.54
Apple	17775	56.91
Potato	13060	57.37
Banana	16347	55.57
Cattle	19758	57.05
Swine	49178	55.89

Song et al., 2018

From UniProtKB Download

Organism	Total number of 9mers containing Q/EX$_1$PX$_2$ motif starting at the 4th or 6 th amino acid	Percent of 9mers possibly subject to peptide modeling
Barley	1037	78.01
Rye	2021	93.77
Wheat	76465	68.74
Oat	253	90.51
Maize	42945	56.56
Rice	43231	56.42
Potato	11058	56.19
Peanut	253	54.94
Sunflower	11629	57.19
Apple	471	59.45
Banana	8423	55.56
Chicken	17195	54.17
Cattle	15108	57.07

Modeling Concept

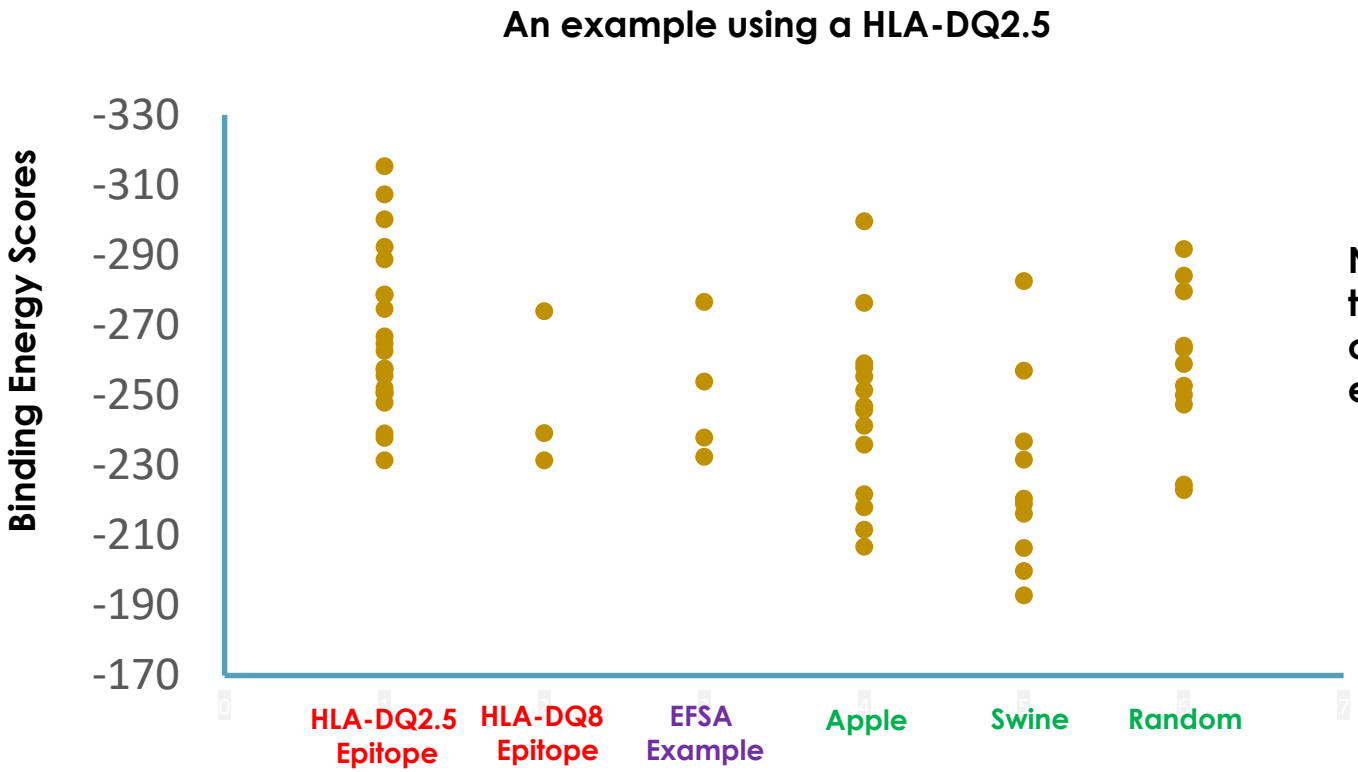


HLA-DQ Peptide Modeling Test

9mer peptides used for HLA peptide modeling:

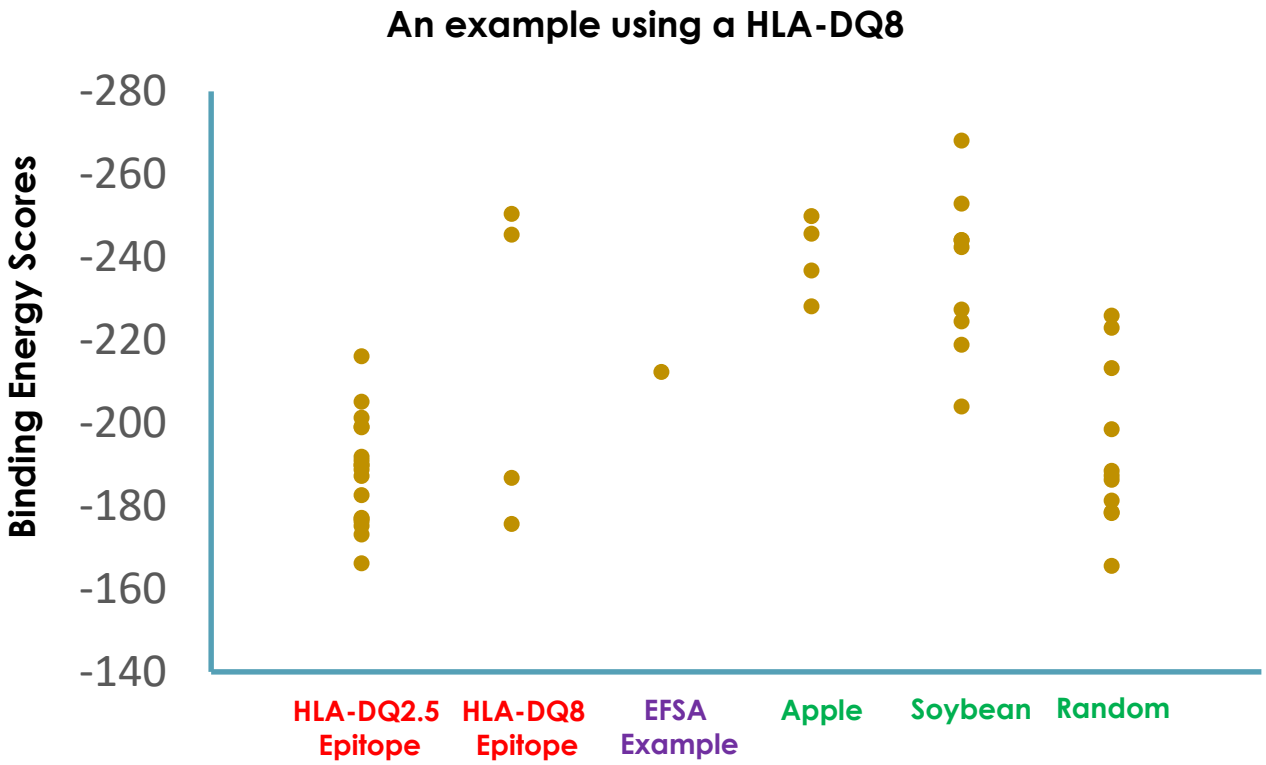
- Known HLA-DQ2.5/DQ8 binders (epitopes).
- Random 9mer sequences.
- 9mers from apple, soybean, and swine – not considered to cause celiac disease but with QX_1PX_2 motif.
- 9mer sequences that contained the HLA-DQ2.5 motif, but with parameters making binding unlikely – not subject to modeling based on EFSA guideline.
- 9mers containing DQ2.5 motif, but no parameters to dismiss modeling requirement – subject to modeling based on EFSA guideline.

HLA-DQ Peptide Modeling – Distribution of Binding Energy Scores



No clear threshold based on the binding energy scores

HLA-DQ Peptide Modeling – Distribution of Binding Energy Scores



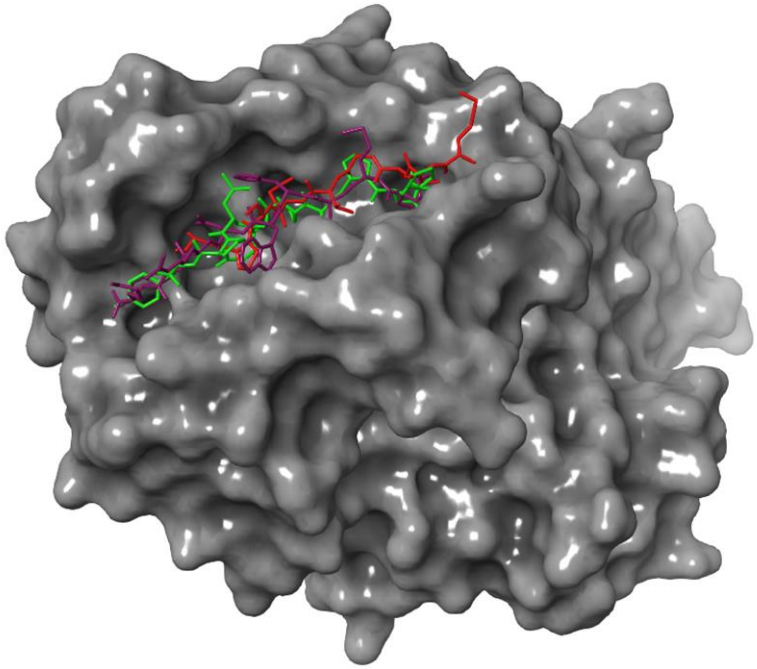
No clear threshold based on the binding energy scores

HLA-DQ Peptide Modeling – an Example Using a HLA-DQ2.5

Peptides:
DQ2.5-glia-a1a: PFPQP**ELPY**
Energy scores: -273.538

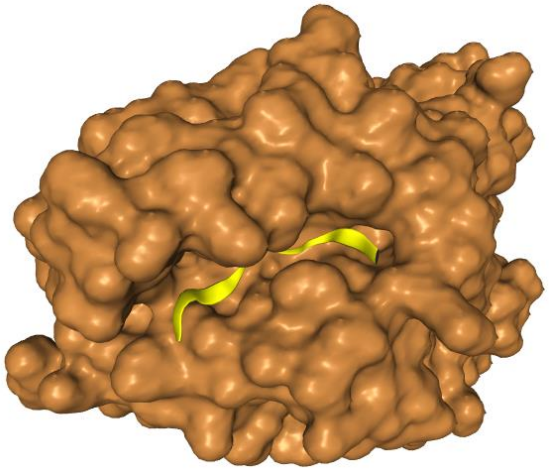
With QX_1PX_2motif but not subject to modeling due to presence of a positively charged amino acid that makes binding unlikely:
KARGV**ESPA**
Energy scores: -189.756

Random 9mer: WMHHWDRYK
Energy scores: -299.810

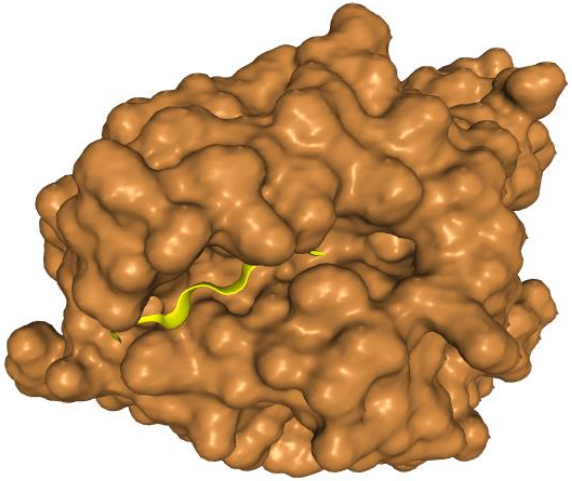


HLA-DQ Peptide Modeling/Docking – an Example Using a HLA-DQ2.5

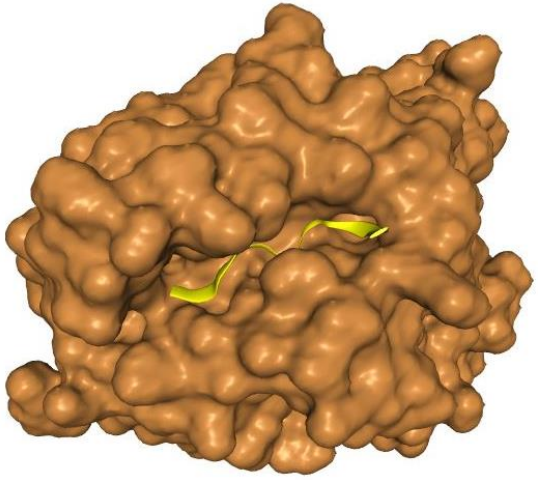
DQ2.5-glia-ω1: PFPQPEQPF
Energy scores:-272.759



Apple: QSQQQEQPF
Energy scores: -193.414

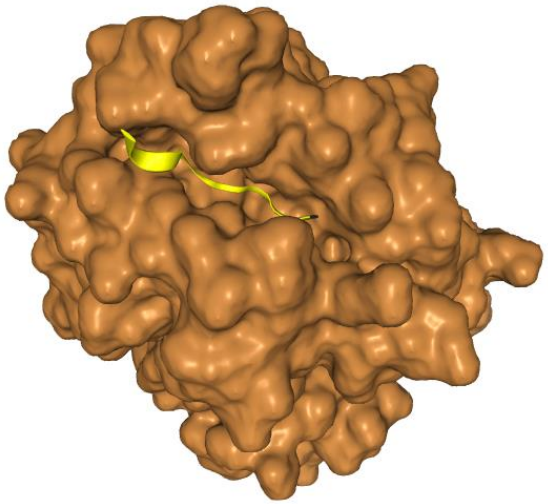


Random: NTPYAVFGL
Energy scores: -266.791

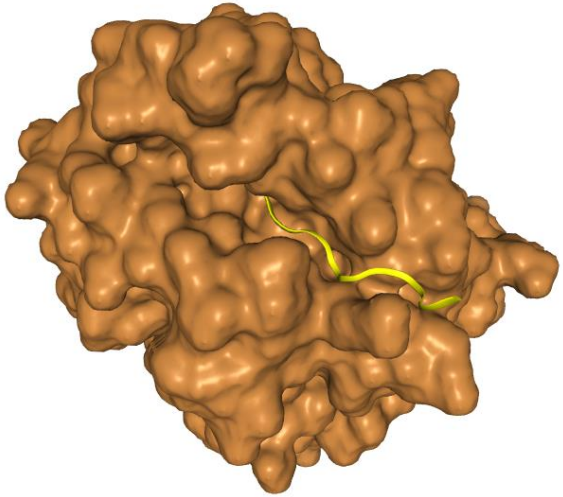


HLA-DQ Peptide Modeling/Docking – an Example Using a HLA-DQ8

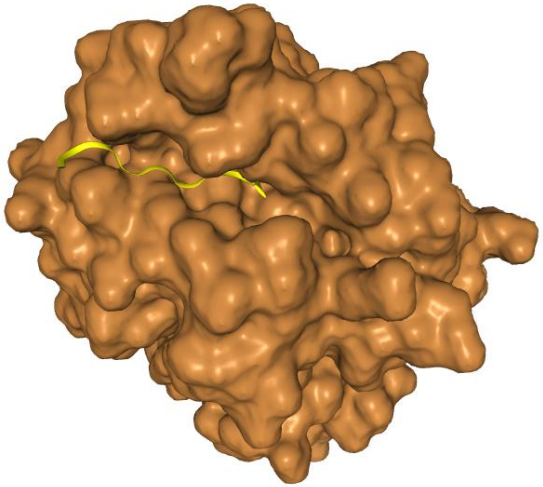
DQ8-glia-γ1a: EQPQQPFPE
Energy scores: -220.187



Soybean: PQQQQPQQE
Energy scores: -232.73



Random: YRQTDPHWE
Energy scores: -247.076



Take-Home Message

- **Peptide modeling is only appropriate when other HLA-DQ peptide binding exclusion criteria are absent**
- **Several software packages were tested for HLA-DQ peptide modelling and one was presented based on its ability to estimate energy scores for binding**
  - Binding energy score was unable to differentiate the 9mer HLA-DQ2.5 or -DQ8 core epitopes from 9-mers not associated with celiac disease
  - Candidate criteria and tools for peptide modeling need a thorough validation for their ability to differentiate celiac peptides from random peptides before being adopted for risk assessment
  - *In silico* modelling criteria and thresholds have yet to be identified for distinguishing celiac peptides from non-celiac peptides
- **Is the 9mer core epitope sufficient to quantify HLA-DQ peptide binding *in silico*?**
  - or are other software packages and/or criteria needed to distinguish true celiac-disease risk ?
  - or is flanking sequence also critical to distinguish the binding of 9mers associated with true celiac-disease risk ?

Thank you!