



Contract NP/EFSA/NIF/2023/02

preDQ – Update & Upgrade

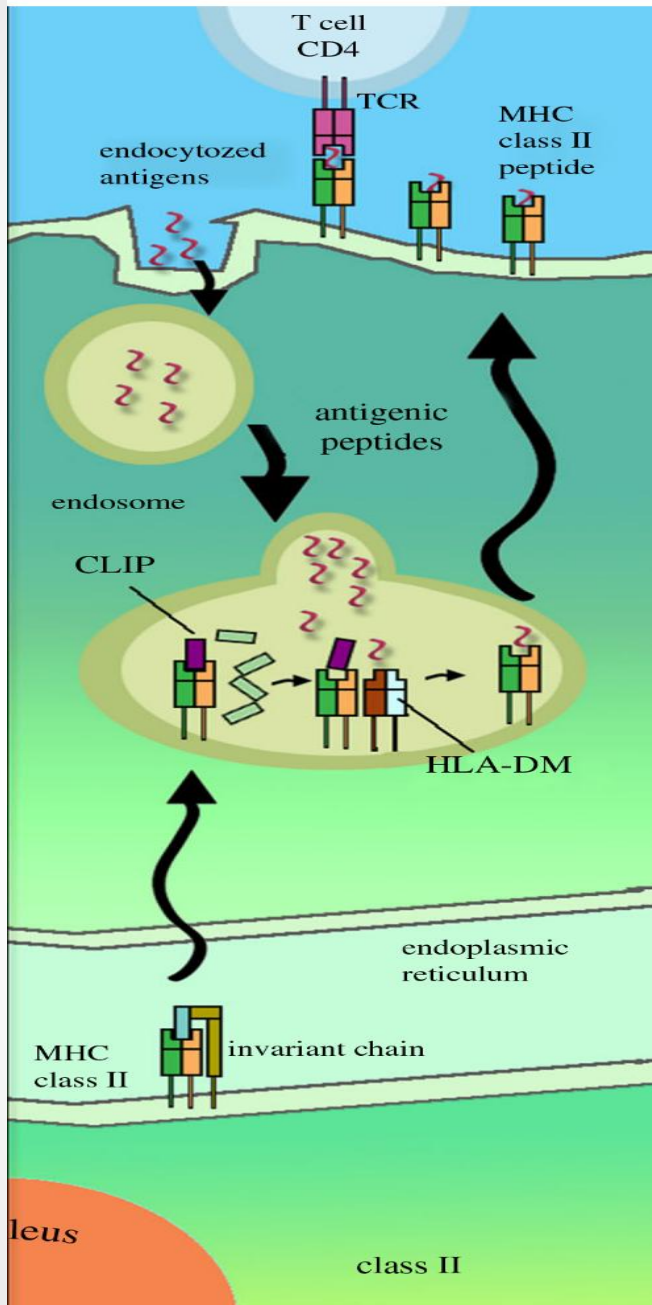
Update and maintenance of the **preDQ** –
software tool for peptide binding prediction
to HLA-DQ2 and HLA-DQ8

Irini Doytchinova, Ivan Dimitrov, Mariyana Atanasova

Medical University of Sofia, Bulgaria

Ad-hoc Meeting with GMO Industry Representatives

27 May 2026



Not all peptide-HLA complexes are recognized by T-cell receptors.

All T-cell epitopes are HLA binders, but not all HLA binders are T-cell epitopes.

preDQ is a tool for **peptide binding** prediction to HLA-DQ2.5 and DQ8.1.

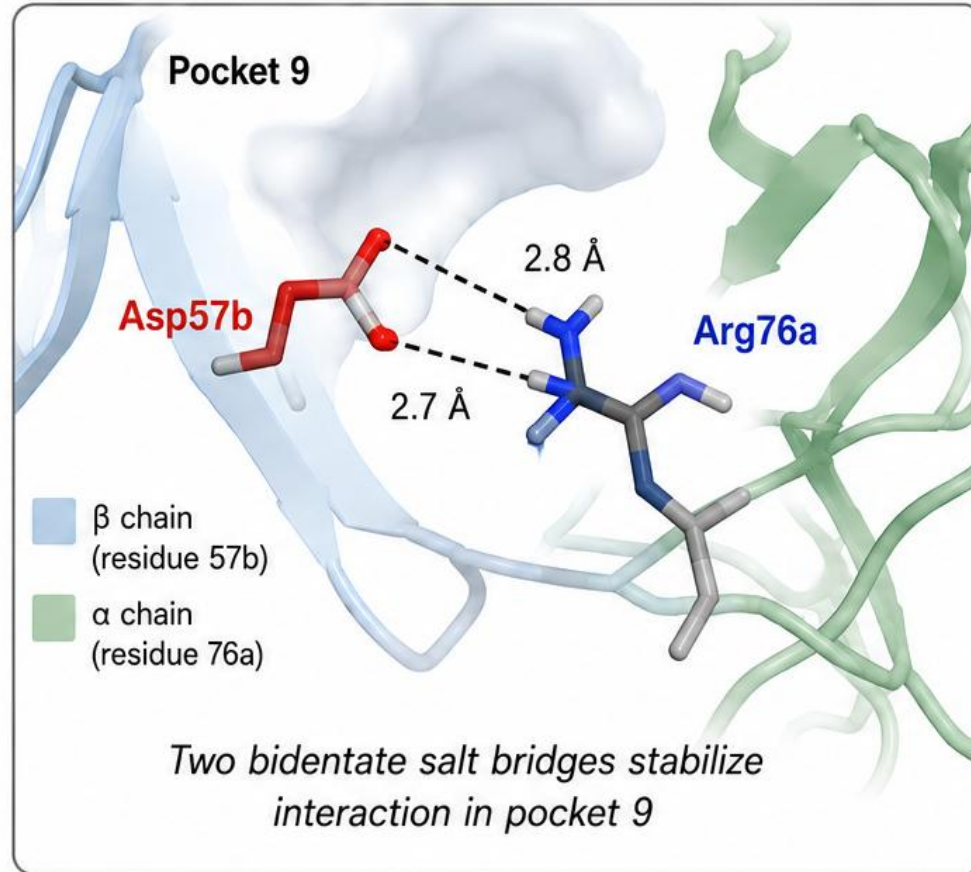
preDQ cannot predict T-cell epitopes.

preDQ cannot predict associations with celiac disease.

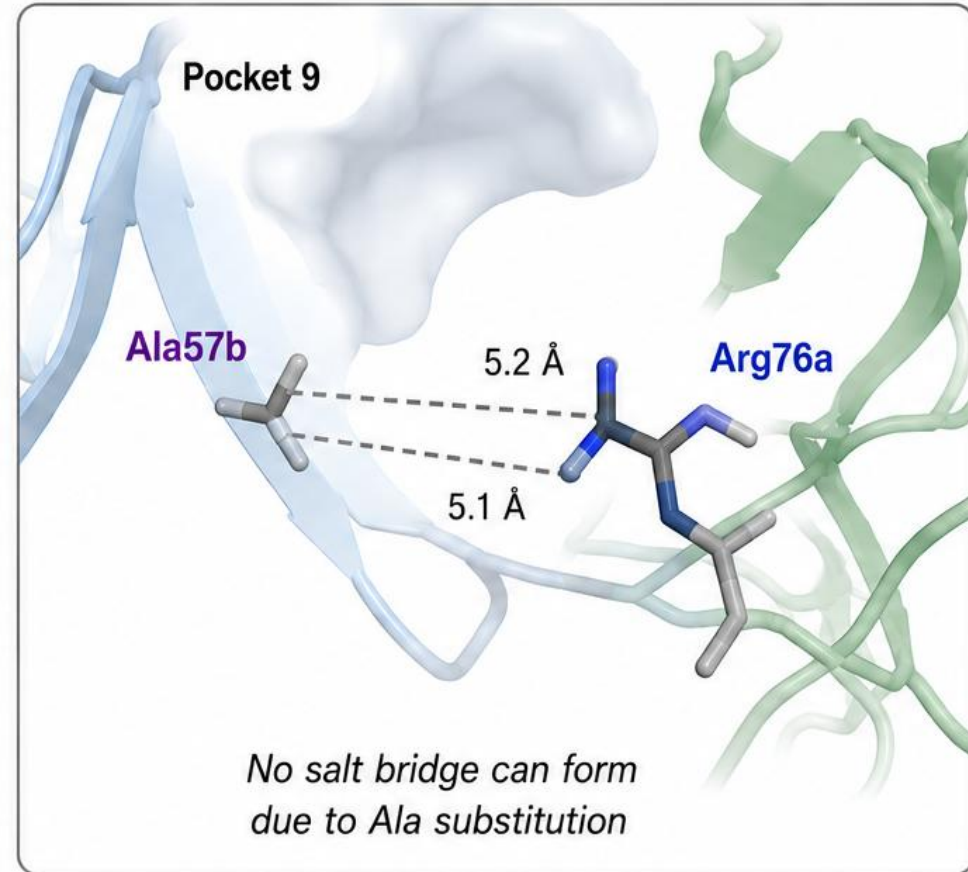
Pocket 9 in HLA-DQ2.5: Asp57b Mutated to Ala

Ala Cannot Form Salt Bridge with Arg76a

Asp57b-containing HLA-DQ proteins



Ala57b-containing HLA-DQ proteins



Summary

- **Asp57b**-containing HLA-DQ proteins form two bidentate salt bridges between **Asp57b** and **Arg76a** in pocket 9.
- Mutation of **Asp57b** to **Ala** eliminates the negative charge and removes the ability to form salt bridges with **Arg76a**.

preDQ v2 Updates

The Example in plain format button was positioned outside the table on the right. Now, It is centred.

[Home](#) | [Method Description](#) | [Contact](#)

preDQ

a tool for peptide binding prediction to HLA-DQ2 and/or HLA-DQ8

[Help](#)

[Example in plain format](#)

1. Insert a single sequence in one CAPITAL letter code. It should contain at least 9 amino acids. The result is displayed in a table.

For prediction of multiple sequences upload a file in fasta format. The result is in a .csv file.

[Choose File](#) | aaa.fasta

3. Choose deamidation position

- ☐ p1 (recommended for HLA-DQ8.1)
- ☐ p4 (recommended for HLA-DQ2.5 and HLA-DQ8.1)
- ☐ p6 (recommended for HLA-DQ2.5 and HLA-DQ8.1)
- ☐ p7 (recommended for HLA-DQ2.5)
- ☐ p9 (recommended for HLA-DQ2.5 and HLA-DQ8.1)

5. Choose the output format

- ☒ Majority voting (at least 3 models out of 5 with positive prediction)

[Predict](#)

2. Choose an allele:

HLA-DQ2.5 (A1*05:01/B1*02:01) ▼

4. Search only peptides containing specific motif (recommended for celiac peptides and HLA-DQ2.5 allele)

- ☒ Q/E-X1-P-X2 motif
- ☒ N/D-X1-P-X2 motif

6. Non-binders

- ☒ include
- ☐ exclude

7. Annotate results with Food Nonamers Database

- ☒ yes
- ☐ no

[Reset form](#)



The front-page design has been rearranged for better screen fit. Now, both X-ray structures of the peptide-HLA protein complexes are displayed with DQ2.5 on the left and DQ8.1 on the right.

preDQ v2 Updates

[Home](#) | [Method Description](#) | [Contact](#)

preDQ

a tool for peptide binding prediction to HLA-DQ2 and/or HLA-DQ8

[Help](#)

[Example in plain format](#)

1. Insert a single sequence in one CAPITAL letter code. It should contain at least 9 amino acids. The result is displayed in a table.

HLA-DQ2.5 (A1*05:01/B1*02:01) ▼

2. Choose an allele:

For prediction of multiple sequences upload a file in fasta format. The result is in a .csv file.

[Choose File](#) | aaa.fasta

3. Choose deamidation position

- ☐ p1 (recommended for HLA-DQ8.1)
- ☐ p4 (recommended for HLA-DQ2.5 and HLA-DQ8.1)
- ☐ p6 (recommended for HLA-DQ2.5 and HLA-DQ8.1)
- ☐ p7 (recommended for HLA-DQ2.5)
- ☐ p9 (recommended for HLA-DQ2.5 and HLA-DQ8.1)

5. Choose the output format

- ☒ Majority voting (at least 3 models out of 5 with positive prediction)

4. Search only peptides containing specific motif (recommended for celiac peptides and HLA-DQ2.5 allele)

- ☒ Q/E-X1-P-X2 motif
- ☒ N/D-X1-P-X2 motif

6. Non-binders

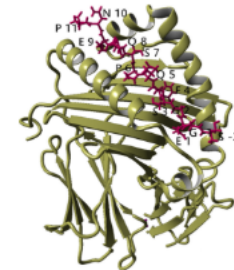
- ☒ include
- ☐ exclude

7. Annotate results with Food Nonamers Database

- ☒ yes
- ☐ no

[Predict](#)

[Reset form](#)



In option 1, the requirement "It should contain at least 9 amino acids" has been added.

Before, if the sequence was shorter than 9 amino acids, an unhandled exception occurred. Now, an error message is displayed.

Before, if the sequence contained a newline character ("\\n"), it was recognized as invalid.

Now, the sequence is trimmed to remove newlines and spaces.

Before, if a non-naturally occurring amino acid occurred in a protein, an unhandled exception occurred.

Now, an error message is displayed, and the invalid proteins are listed at the end of the table.

preDQ v2 Updates

[Home](#) | [Method Description](#) | [Results Interpretation](#) | [Contact](#)

The Help and the Results Interpretation pages have been updated.

preDQ

a tool for peptide binding prediction to HLA-DQ2 and/or HLA-DQ8

[Output Description](#)

[Download CSV](#)

Query:				Allele:					
TKASVGFKA...				HLA-DQ2.5					
Position	Nonamer	Docking-based Prediction	Predict_by_Logo_model	pIC50_by_SVM	pIC50_by_XGboost	pIC50_by_RF	Positive predictions	Known binder ID / Non-binder	Presence in known foods
	cutoff	0	binding score > non-binding score	5.6	5.8	6.0			
1	TKASVGFKA	probable non-binder (not reconized as a binder by majority voting)							Yes
2	KASVGFKAG	probable non-binder (not reconized as a binder by majority voting)							Yes
3	ASVGFKAGV	0.887	1	5.793	5.001	5.696	3		Yes
4	SVGFKAGVK	probable non-binder (not reconized as a binder by majority voting)							Yes
5	VGFKAGVKD	probable non-binder (not reconized as a binder by majority voting)							Yes
6	GFKAGVKDY	probable non-binder (not reconized as a binder by majority voting)							Yes
7	FKAGVKDYK	probable non-binder (not reconized as a binder by majority voting)							Yes

Before, the results were sorted alphabetically. Now, they are displayed in the order of the proteins in the file.

If the protein is longer than 9 amino acids, only the first 9 are displayed in the Query field, followed by an ellipsis (e.g., GLEFLNVAT...).

preDQ v2 Upgrades

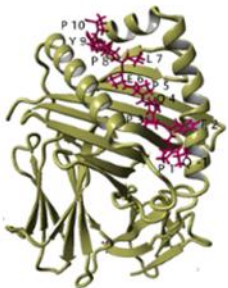
[Home](#) | [Method Description](#) | [Contact](#)

preDQ

a tool for peptide binding prediction to HLA-DQ2 and/or HLA-DQ8

[Help](#)

[Example in plain format](#)



1. Insert a single sequence in one CAPITAL letter code. It should contain at least 9 amino acids. The result is displayed in a table.

2. Choose an allele:

HLA-DQ2.5 (A1*05:01/B1*02:01) ▼

For prediction of multiple sequences upload a file in fasta format. The result is in a .csv file.

[Choose File](#) | aaa.fasta

3. Choose deamidation position

- ☐ p1 (recommended for HLA-DQ8.1)
- ☐ p4 (recommended for HLA-DQ2.5 and HLA-DQ8.1)
- ☐ p6 (recommended for HLA-DQ2.5 and HLA-DQ8.1)
- ☐ p7 (recommended for HLA-DQ2.5)
- ☐ p9 (recommended for HLA-DQ2.5 and HLA-DQ8.1)

4. Search only peptides containing specific motif (recommended for celiac peptides and HLA-DQ2.5 allele)

- ☒ Q/E-X1-P-X2 motif
- ☒ N/D-X1-P-X2 motif

5. Choose the output format

- ☒ Majority voting (at least 3 models out of 5 with positive prediction)

6. Non-binders

- ☒ include
- ☐ exclude

7. Annotate results with Food Nonamers Database

- ☒ yes
- ☐ no

[Predict](#)

[Reset form](#)

In option 4 Search only peptides containing a specific motif, the N/D-X1-P-X2 motif has been added.

In the previous version of preDQ, X1 and X2 were considered as all 20 naturally occurring amino acids. Now, the selection has been refined: X1 includes L, Q, F, S, E, X2 includes Y, F, A, V, Q.

Option 7 Annotate results with Food Nonamers Database is the most import upgrade in v2.

preDQ v2 Upgrades

[Home](#) | [Method Description](#) | [Results Interpretation](#) | [Contact](#)

preDQ

a tool for peptide binding prediction to HLA-DQ2 and/or HLA-DQ8

[Output Description](#)

[Download CSV](#)

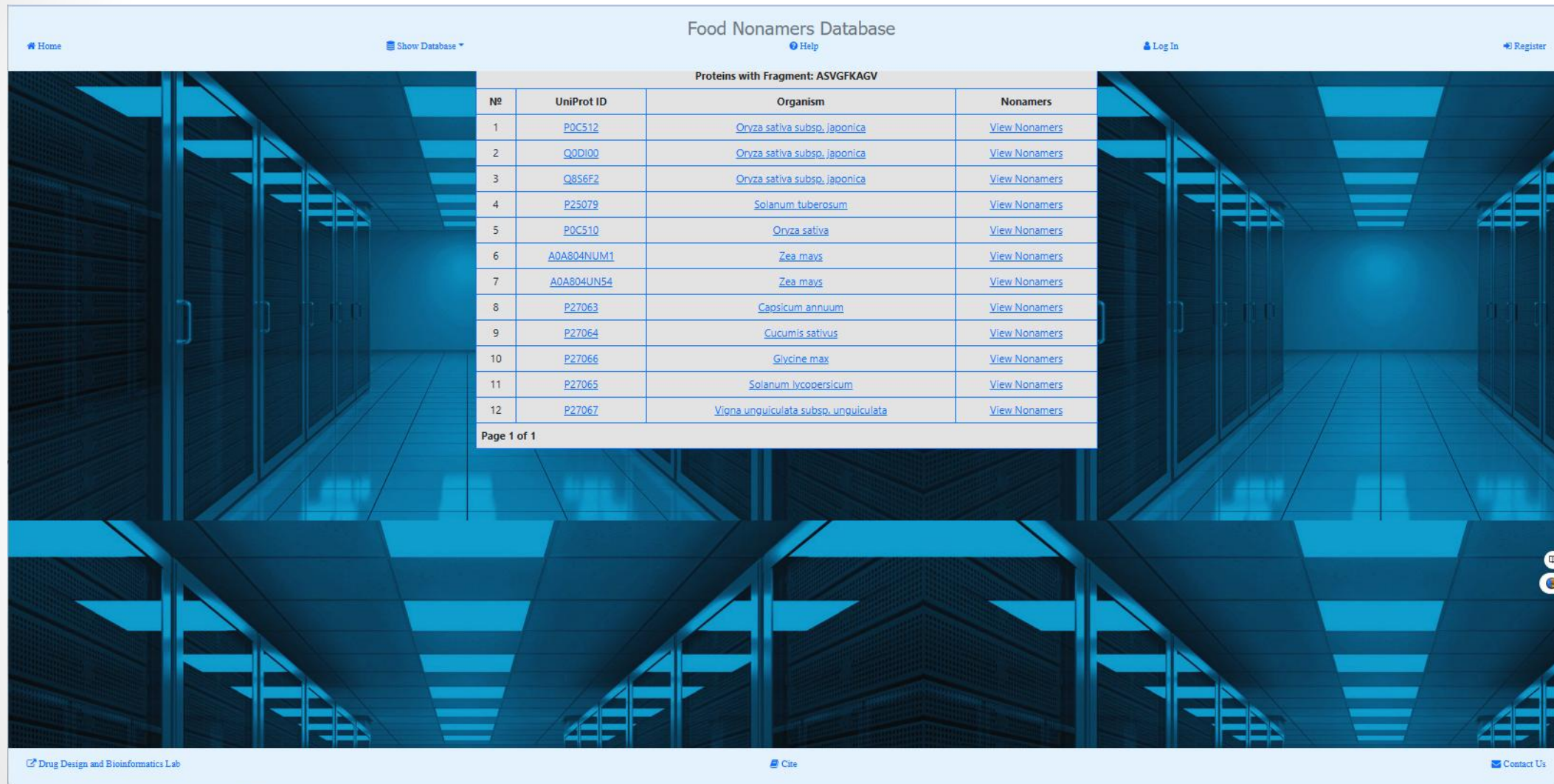
Query:			Allele:						
TKASVGFKA...			HLA-DQ2.5						
Position	Nonamer	Docking-based Prediction	Predict_by_Logo_model	pIC50_by_SVM	pIC50_by_XGboost	pIC50_by_RF	Positive predictions	Known binder ID / Non-binder	Presence in known foods
	cutoff	0	binding score > non-binding score	5.6	5.8	6.0			
1	TKASVGFKA	probable non-binder (not reconized as a binder by majority voting)							Yes
2	KASVGFKAG	probable non-binder (not reconized as a binder by majority voting)							Yes
3	ASVGFKAGV	0.887	1	5.793	5.001	5.696	3		Yes
4	SVGFKAGVK	probable non-binder (not reconized as a binder by majority voting)							Yes
5	VGFKAGVKD	probable non-binder (not reconized as a binder by majority voting)							Yes
6	GFKAGVKDY	probable non-binder (not reconized as a binder by majority voting)							Yes
7	FKAGVKDYK	probable non-binder (not reconized as a binder by majority voting)							Yes

A new column, Presence in Known Foods, is added to the Results table.

The tested peptide is checked against the Food Nonamers Database.

If a match is found, this column provides additional information about the parent proteins and organisms.

Food Nonamers Database



The screenshot displays the Food Nonamers Database interface. At the top, there are navigation links: Home, Show Database, Help, Log In, and Register. The main content area is titled "Proteins with Fragment: ASVGFKAGV". It contains a table with 12 rows of protein data. Each row includes a number, a UniProt ID, the organism name, and a link to view nonamers. The background of the website features a blue-toned image of server racks.

Proteins with Fragment: ASVGFKAGV				
Nº	UniProt ID	Organism	Nonamers	
1	P0C512	Oryza sativa subsp. japonica	View Nonamers	
2	Q0DI00	Oryza sativa subsp. japonica	View Nonamers	
3	Q856F2	Oryza sativa subsp. japonica	View Nonamers	
4	P25079	Solanum tuberosum	View Nonamers	
5	P0C510	Oryza sativa	View Nonamers	
6	A0A804NUM1	Zea mays	View Nonamers	
7	A0A804UN54	Zea mays	View Nonamers	
8	P27063	Capsicum annuum	View Nonamers	
9	P27064	Cucumis sativus	View Nonamers	
10	P27066	Glycine max	View Nonamers	
11	P27065	Solanum lycopersicum	View Nonamers	
12	P27067	Vigna unguiculata subsp. unguiculata	View Nonamers	

Page 1 of 1

Drug Design and Bioinformatics Lab Cite Contact Us

The final version of the database contains 389 organisms, 98 643 proteins, and 20 000 210 nonamers.

Publications

1. Doytchinova I, Dimitrov I, Atanasova M. preDQ – a software tool for peptide binding prediction to HLA-DQ2 and HLA-DQ8. EFSA Supporting Publications 2023, 7, 8108E.
2. Doytchinova I., Atanasova M., Fernandez A., Moreno J., Koning F., Dimitrov I. Modeling peptide-protein interactions by a logo-based method: Application in peptide-HLA binding predictions. Molecules 2024, 29, 284.
3. Atanasova M., Dimitrov I., Fernandez A., Moreno J., Koning F., Doytchinova I. Assessment of novel proteins triggering celiac disease via docking-based approach. Molecules 2024, 29, 138.
4. Doytchinova, I.; Dimitrov. I.; Atanasova. M. preDQ – Update & Upgrade Update and maintenance of the preDQ – software tool for peptide binding prediction to HLA-DQ2 and HLA-DQ8: Re-evaluation of the logo models. EFSA Supporting Publications 2025, EN-9585.

Summary

preDQ is a tool for peptide binding prediction to HLA-DQ2.5 and DQ8.1.

preDQ alone cannot predict T-cell epitopes.

preDQ alone cannot predict associations with celiac disease.