



IEDB Customized Data Exports

Presented by: Kelly Wheeler, Senior Software Engineer

What was the previous export?

Current Filters: ✖ Epitope Structure: Linear Sequence ✖ Linear Sequence: SIINFEKL ✖ Include Positive Assays

Epitopes (3) Antigens (1) Assays (535) Receptors (331) References (211)

Go To Records Starting At: 1200

3 Records Found Page 1 of 1 Export Results

Details	Epitope	Antigen	Organism		
58560	SIINFEKL	Gal d 2	Gallus gallus (chicken)		
114230	SIINFEKL + SCM(K7)	Gal d 2	Gallus gallus (chicken)		
1309579	SIINFEKL + BIOT(L8)	Gal d 2	Gallus gallus (chicken)	1	1

3 Records Found Page 1 of 1 25 Per Page Export Results

Go To Records Starting At: 1200

Export to CSV file. [?](#)

Export to CSV file with IRIs. [?](#)



Current Filters: Epitope Structure: Linear Sequence Linear Sequence: SIINFEKL Include Positive Assays

Your export is complete, it should download shortly or can be accessed [Here](#) for 24 hours after completion.
Full downloads of the entire database can be found [Here](#)

.csv only with dual headers

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Related O	Related O	Related	
2	Epitope ID	Object Type	Description	Epitope M	Epitope Modification(s)	Starting P	Ending Po	Non-pept	Epitope S	Antigen Name	Antigen Accession	Parent Protein	Parent Protein Accession	Organism Name	Parent Organism	Parent Organism ID	Epitope C	Epitope R	Object Ty	Descrip	
3	58560	Linear peptide	SIINFEKL			258	265			ovalbumin	AAA48998.1	Gal d 2	P01012	Gallus gallus	Gallus gallus	9031					
4	114230	Linear peptide	SIINFEKL + SCM(K7)	K7	Side chain modification	258	265			ovalbumin	AAA48998.1	Gal d 2	P01012	Gallus gallus	Gallus gallus	9031				The epitope is modified with trinitrophenyl	
5	1309579	Linear peptide	SIINFEKL + BIOT(L8)	L8	Biotin BIOT	257	264			ovalbumin	0705172A	Gal d 2	P01012	Gallus gallus	Gallus gallus	9031					
6																					
7																					
8																					
1	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Related Object	Related Object	Related Object	Related Object
2	Epitope ID	Object Type	Description	Epitope Modified Residue(s)	Epitope Modification(s)	Starting Position	Ending Position	Non-peptidic epitope	Epitope Accession	Epitope Synonyms	Antigen Name	Antigen Accession	Parent Protein	Parent Protein Accession	Organism Name	Parent Organism	Parent Organism ID	Epitope C	Epitope R	Object Type	Description
3	"58560"	"Linear peptide"	"SIINFEKL"	"", "", "258", "265"	"", "", "ovalbumin"	"AAA48998.1"	"Gal d 2"	"P01012"	"Gallus gallus"	"Gallus gallus"	"9031"	"", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", ""	
4	"114230"	"Linear peptide"	"SIINFEKL + SCM(K7)"	"K7"	"Side chain modification"	"258"	"265"	"", "", "ovalbumin"	"AAA48998.1"	"Gal d 2"	"P01012"	"Gallus gallus"	"Gallus gallus"	"9031"	"Gallus gallus"	"Gallus gallus"	"9031"	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"The epitope is modified with trinitrophenyl on Lys at
5	"1309579"	"Linear peptide"	"SIINFEKL + BIOT(L8)"	"L8"	"Biotin BIOT"	"257"	"264"	"", "", "ovalbumin"	"0705172A"	"Gal d 2"	"P01012"	"Gallus gallus"	"Gallus gallus"	"9031"	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", "", ""	"", "", "", "", "", "", "", "", "", "", "", "", ""
6																					

Why Change?

- Limited to .csv file format
- Fixed set of export fields
- Dual headers only
- Limited explanation of what each field means
- Limited information on the query performed
- Lack of information on when the query was performed

Introducing the New Customized Data Exports

Goals:

- Expand the usability of the exports
- Provide more information on each field
- Allow more customization to the user
- Provide more information linking the exported data to the query

New Customized Data Exports

Current Filters: **X** Epitope Structure: Linear Sequence **X** Linear Sequence: SIINFEKL **X** Include Positive Assays

Epitopes (3) Antigens (1) Assays (586) Receptors (331) References (229)

Go To Records Starting At 1200

3 Records Found Page 1 of 1 Export Results 25 Per Page

IEDB ID	Epitope	Antigen	Organism	References	# Assays
58560	SIINFEKL	Gal d 2	Gallus gallus (chicken)	226	578
114230	SIINFEKL + SCM(K7)	Gal d 2	Gallus gallus (chicken)	4	7
422946	SIINFEKL + BIOT(L8)	Gal d 2	Gallus gallus (chicken)	1	1

3 Records Found Page 1 of 1 25 Per Page Go To Records Starting At 1200 Export Results



Current Filters: **X** Epitope Structure: Linear Sequence **X** Linear Sequence: SIINFEKL **X** Include Positive Assays

Epitopes (3) Antigens (1) Receptors (331) References (229)

3 Records Found

IEDB ID	Epitope	# References	# Assays
58560	SIINFEKL	226	578
114230	SIINFEKL + SCM(K7)	4	7
422946	SIINFEKL + BIOT(L8)	1	1

3 Records Found 25 Per Page Export Results

EXPORT TO FILE

File Format: **.XLSX**

Header Row Format: **Double Headers**

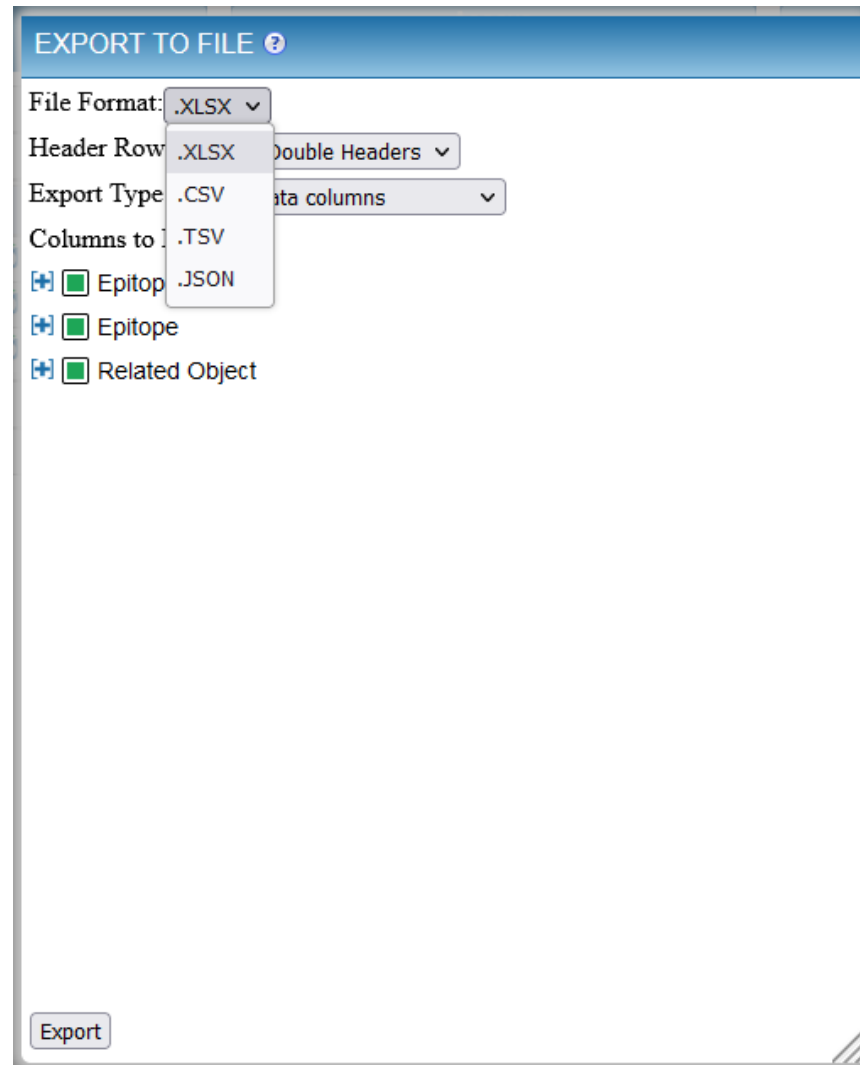
Export Type: **Full, all data columns**

Columns to Include:

- Epitope ID
- Epitope
- Related Object

Export

Multiple File Formats



.XLS – Excel format – Most user friendly

	A	B	C	D	E	F	G	H	I	J	K	L
1	Epitope ID	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope	Epitope
2	IEDB IRI	Object Type	Name	Modified Residue(s)	Modifications	Starting Position	Ending Position	IRI	Synonyms	Source Molecule	Source Molecule IRI	Molecule Parent
3	https://www.iedb.org/ep/Linear peptide		SIINFEKL			258	265			ovalbumin	http://www.ncbi.nlm.nih.gov/protein/AAA48998.1	Gal d 2
4	https://www.iedb.org/ep/Linear peptide		SIINFEKL + SCM(K7)	K7	Side chain modification	258	265			ovalbumin	http://www.ncbi.nlm.nih.gov/protein/AAA48998.1	Gal d 2
5	https://www.iedb.org/ep/Linear peptide		SIINFEKL + BIOT(L8)	L8	Biotin BIOT	257	264			ovalbumin	http://www.ncbi.nlm.nih.gov/protein/0705172A	Gal d 2

.CSV – Comma Separated Values – Current format – Programmer friendly

```
["Epitope ID", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Epitope", "Related Object", "Related Object", "Related Object", "Related Object", "Related Object", "Related Ob
"IEDB IRI", "Object Type", "Name", "Modified Residue(s)", "Modifications", "Starting Position", "Ending Position", "IRI", "Synonyms", "Source Molecule", "Source Molecule IRI", "Source Molecule IRI", "Molecule Parent", "Molecule Parent IRI", "Source Organism", "Source Organism IRI",
https://www.iedb.org/epitope/58560, "Linear peptide", "SIINFEKL", "", "258,265", "ovalbumin", "http://www.ncbi.nlm.nih.gov/protein/AAA48998.1", "Gal d 2", "http://www.uniprot.org/uniprot/P01012", "Gallus gallus", "http://purl.obolibrary.org/obo/NCBITaxon_90
https://www.iedb.org/epitope/114230, "Linear peptide", "SIINFEKL + SCM(K7)", "K7", "Side chain modification", "258,265", "ovalbumin", "http://www.ncbi.nlm.nih.gov/protein/AAA48998.1", "Gal d 2", "http://www.uniprot.org/uniprot/P01012", "Gallus gallus", "htt
https://www.iedb.org/epitope/422946, "Linear peptide", "SIINFEKL + BIOT(L8)", "L8", "Biotin|BIOT", "257,264", "ovalbumin", "http://www.ncbi.nlm.nih.gov/protein/0705172A", "Gal d 2", "http://www.uniprot.org/uniprot/P01012", "Gallus gallus", "http://purl.obolibrary.org/obo/NCBITaxon_90"]
```

.TSV – Tab Separated Values – Similar to CSV but with tabs – Programmer friendly

```
"Epitope ID"   "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Epitope"  "Related Object"    "Related Object"    "Related Object"    "Related Object"
"IEDB IRI"    "Object Type"  Name     "Modified Residue(s)"  Modifications  "Starting Position"    "Ending Position"    IRI     Synonyms    "Source Molecule"    "Source Molecule IRI"    "Molecule Parent"    "Molecule Parent IRI"    "Source Organism"    "Source Organism IRI"
https://www.iedb.org/epitope/58560  "Linear peptide"    "SIINFEKL"    "258,265"    "ovalbumin"    "http://www.ncbi.nlm.nih.gov/protein/AAA48998.1"    "Gal d 2"    "http://www.uniprot.org/uniprot/P01012"    "Gallus gallus"
https://www.iedb.org/epitope/114230  "Linear peptide"    "SIINFEKL + SCM(K7)"    "K7"    "Side chain modification"    "258,265"    "ovalbumin"    "http://www.ncbi.nlm.nih.gov/protein/AAA48998.1"    "Gal d 2"
https://www.iedb.org/epitope/422946  "Linear peptide"    "SIINFEKL + BIOT(L8)"    "L8"    "Biotin|BIOT"    "257,264"    "ovalbumin"    "http://www.ncbi.nlm.nih.gov/protein/0705172A"    "Gal d 2"    "http://www.uniprot.org/uniprot/P01012"    "Gallus gallus"
```

.JSON – JavaScript Object Notation – A programming friendly format used to transport data

```
{
  "Search Parameters": [
    "Epitope Structure: Linear Sequence",
    "Linear Sequence: SIINFEKL",
    "Include Positive Assays",
    "Export Type: Full, all data columns"
  ],
  "Export Date": "October 26, 2023",
  "Data": [
    {
      "Epitope ID - IEDB IRI": "https://www.iedb.org/epitope/58560",
      "Epitope - Object Type": "Linear peptide",
      "Epitope - Name": "SIINFEKL",
      "Epitope - Modified Residue(s)": "",
      "Epitope - Modifications": "",
      "Epitope - Starting Position": "258",
      "Epitope - Ending Position": "265",
      "Epitope - IRI": "",
      "Epitope - Synonyms": "",
      "Epitope - Source Molecule": "ovalbumin",
      "Epitope - Source Molecule IRI": "http://www.ncbi.nlm.nih.gov/protein/AAA48998.1",
      "Epitope - Molecule Parent": "Gal d 2",
      "Epitope - Molecule Parent IRI": "http://www.uniprot.org/uniprot/P01012",

```


Multiple Header Formats

EXPORT TO FILE ?

File Format: .XLSX

Header Row Format: Double Headers

Export Type: Full, all

Columns to Include: Single Header

Epitope ID

Epitope

Related Object

Export

Double Headers: Two header rows with a section and field row

Epitope ID	Epitope IEDB IRI	Epitope Object Type	Epitope Name	Epitope Modified Residue(s)	Epitope Modifications	Epitope Starting Position	Epitope Ending Position	Epitope IRI
https://www.iedb.org/ef/Linear%20peptide		SIINFEKL				258	265	
https://www.iedb.org/ef/Linear%20peptide		SIINFEKL + SCM(K7)	K7		Side chain modification	258	265	
https://www.iedb.org/ef/Linear%20peptide		SIINFEKL + BIOT(L8)	L8		Biotin BIOT	257	264	

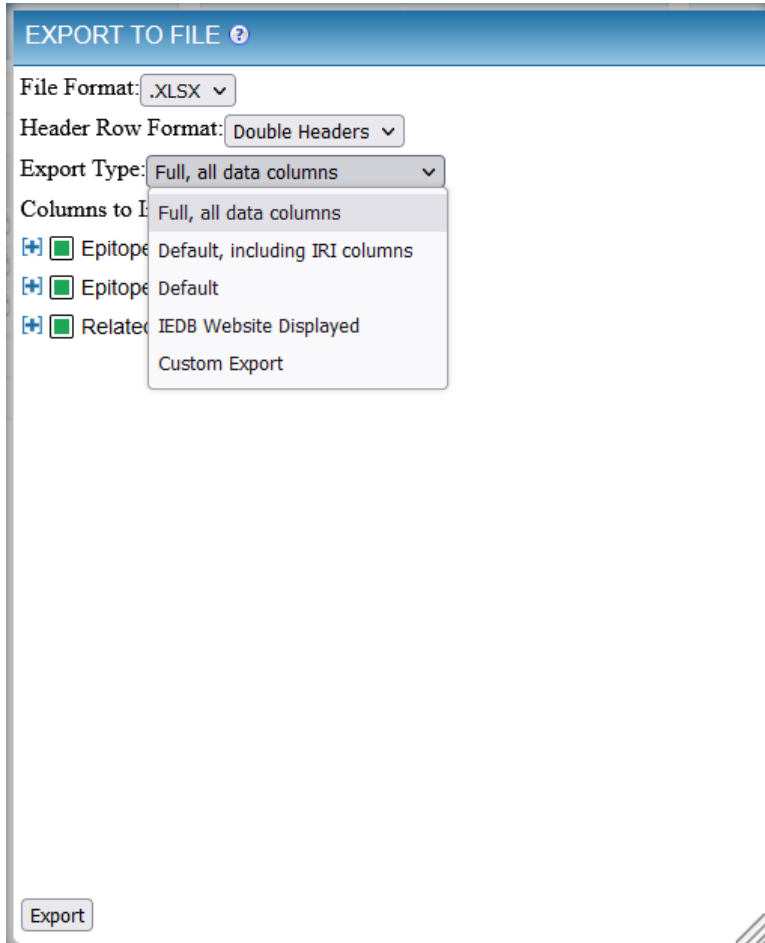
Single Header: Single header row with section and field merged

Epitope ID - IEDB IRI	Epitope - Object Type	Epitope - Name	Epitope - Modified Resic	Epitope - Modifications	Epitope - Starting Posicik	Epitope - Ending Positio	Epitope - IRI
https://www.iedb.org/ef/Linear%20peptide		SIINFEKL			258	265	
https://www.iedb.org/ef/Linear%20peptide		SIINFEKL + SCM(K7)	K7	Side chain modification	258	265	
https://www.iedb.org/ef/Linear%20peptide		SIINFEKL + BIOT(L8)	L8	Biotin BIOT	257	264	

No Headers: Simple data centric format with no headers

https://www.iedb.org/ef/Linear%20peptide		SIINFEKL			258	265														
https://www.iedb.org/ef/Linear%20peptide		SIINFEKL + K7	Side chain		258	265		ovalbumin	http://www.Gal d 2	http://www.Gallus gall	http://purl	Gallus gall	http://purl	obolibrary.org/obo/NCBITaxon_						
https://www.iedb.org/ef/Linear%20peptide		SIINFEKL + L8	Biotin BIOT		257	264		ovalbumin	http://www.Gal d 2	http://www.Gallus gall	http://purl	Gallus gall	http://purl	obolibrary.org/obo/NCBITaxon_						

Export Types



Export types are preselected subsets of fields to be exported

Available Export Types:

- Full, all data columns: All the available columns available for that export
- Default: A preselected 'default' set of columns selected by the IEDB team
- Default, including IRI columns: An expanded set of the Default column set with IRIs included where available
- IEDB Website Displayed: A set of columns that match what is shown in the tab being exported
- Custom Export: User selected set of columns

IEDB Website Displayed

EXPORT TO FILE ⓘ

File Format: .XLSX ▾

Header Row Format: Double Headers ▾

Export Type: IEDB Website Displayed ▾

Columns to Include:

- Epitopes
 - IEDB ID
 - Epitope
 - Antigen
 - Organism
 - # References
 - # Assays

Export

- IEDB Website Displayed is a special Export Type. It directly reflects the fields shown on the selected results page tab
- Due to its unique nature its fields can't be modified or added to other exports

Custom Export

The screenshot shows the 'EXPORT TO FILE' dialog box with the following settings:

- File Format: .XLSX
- Header Row Format: Double Headers
- Export Type: Custom Export

Under 'Columns to Include:', there is a tree view with the following items:

- Epitope ID
 - IEDB IRI
- Epitope
 - Object Type
 - Name
 - Modified Residue(s)
 - Modifications
 - Starting Position
 - Ending Position
 - IRI
 - Synonyms
 - Source Molecule
 - Source Molecule IRI
 - Molecule Parent
 - Molecule Parent IRI
 - Source Organism
 - Source Organism IRI

An 'Export' button is located at the bottom left of the dialog.

- Allows for fine grain customization of your export set of columns
- Section headers are shown with a [+] icon that can be expanded to show the individual columns beneath. Similar functionality to the trees in the finders
- Full sections can be selected or removed by clicking on the box next to the section header
- Full green boxes indicate a selected column/section. Half green boxes indicate that at least 1 column under that section is selected

Field Selection Examples

Epitope

EXPORT TO FILE ?

File Format: .XLSX ▾

Header Row Format: Double Headers ▾

Export Type: Full, all data columns ▾

Columns to Include:

- Epitope ID
 - IEDB IRI
- Epitope
 - Object Type
 - Name
 - Modified Residue(s)
 - Modifications
 - Starting Position
 - Ending Position
 - IRI
 - Synonyms
 - Source Molecule
 - Source Molecule IRI
 - Molecule Parent
 - Molecule Parent IRI
 - Source Organism
 - Source Organism IRI

Export

Antigen

EXPORT TO FILE ?

File Format: .XLSX ▾

Header Row Format: Double Headers ▾

Export Type: Full, all data columns ▾

Columns to Include:

- Antigen
 - Antigen Name
 - Antigen IRI
 - Organism Name
 - Organism IRI
 - # Epitopes
 - # Assays
 - # References

Export

Field Selection Examples

Assay

EXPORT TO FILE ?

File Format: .XLSX ▾

Header Row Format: Double Headers ▾

Export Type: Full, all data columns ▾

Columns to Include:

- Assay ID
- Reference
- Epitope
- Related Object
- Host
- 1st in vivo Process
- 1st immunogen
- 2nd in vivo Process
- 2nd immunogen
- In vitro Process
- in vitro Responder Cell
- in vitro Stimulator Cell
- in vitro immunogen
- Adoptive Transfer
- Immunization
- Assay
- Effector Cell

Export

Receptor

EXPORT TO FILE ?

File Format: .XLSX ▾

Header Row Format: Double Headers ▾

Export Type: Full, all data columns ▾

Columns to Include:

- Receptor
 - Group IRI
 - IEDB Receptor ID
 - Reference Name
 - Type
- Reference
- Epitope
- Assay
- Chain 1
 - Type
 - Organism IRI
 - Nucleotide Sequence
 - Curated V Gene
 - Calculated V Gene
 - Curated D Gene
 - Calculated D Gene
 - Curated J Gene

Export

Field Selection Examples

Reference

EXPORT TO FILE ⓘ

File Format: .XLSX ▾

Header Row Format: Double Headers ▾

Export Type: Full, all data columns ▾

Columns to Include:

- Reference ID
 - IEDB IRI
- Reference
 - Type
 - PMID
 - Submission ID
 - Alternate IRIs
 - Authors
 - Journal
 - Date
 - Title

Export

Additional Improvements

- Inclusion of query parameters in the Excel export via a new sheet, and the JSON export
- Inclusion of export date in the Excel and JSON exports
- Inclusion of field descriptions in the Excel export via a new sheet

Export Parameters Example

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Epitope Structure: Linear Sequence														
2	Linear Sequence: SIINFEKL														
3	Include Positive Assays														
4															
5															
6	Export Date: October 26, 2022														
7															
8															
9															

- Includes the full list of the parameters used on the IEDB website to generate this export
- Also will include the date the export was ran for future use
- Available in the Excel and JSON formats

Field Help Tab Example

	A	B	C	D	E	F	G	
1	Section							Epito
2	Header	IEDB IRI	Object Type	Name	Modified Residue(s)	Modifications	Starting Position	Ending Pos
3	Example	http://www.iedb.org/epitope/1056	Linear peptide	AELLVALENQHTIDL			441	455
4	Example	http://www.iedb.org/epitope/2136	Linear peptide	AIYHTENAYSVVSSHYNR			207	225
5	Other possible values	http://www.iedb.org/epitope/570	Linear peptide	ACKRPGSGGFFSRLN	R4	Citrullination CITR	138	152
6	Description of Field	The url for this epitope on the IEDB site.	The object type of the epitope, which may include Linear and	A short description of the epitope, including its sequence or	The amino acid and its position within	The type of post-transla	The starting position of	The ending
7	Help Topic Link		http://curationwiki.iedb.org/wiki/index.php/Data_Field_Des	http://curationwiki.iedb.org/wiki/index.php/Data_Field_Des	http://curationwiki.iedb.org/wiki/index	http://curationwiki.iedb	http://curationwiki.iedb	http://curationwiki.iedb
8								
9								
10								

- A new tab in the Excel export that offers information on each field selected and exported
- Each field has the section and header name matching the header format selected
- Below this are 2 pre-selected examples chosen by our team to demonstrate what is in the IEDB
- After this is another possible value to show what that field could include in case our above examples are blank or we want to show other possible values
- Next is a brief description of what that field represents
- Finally, there is a link to an IEDB help article for that field if available