

# The GlyGen Upload Interface

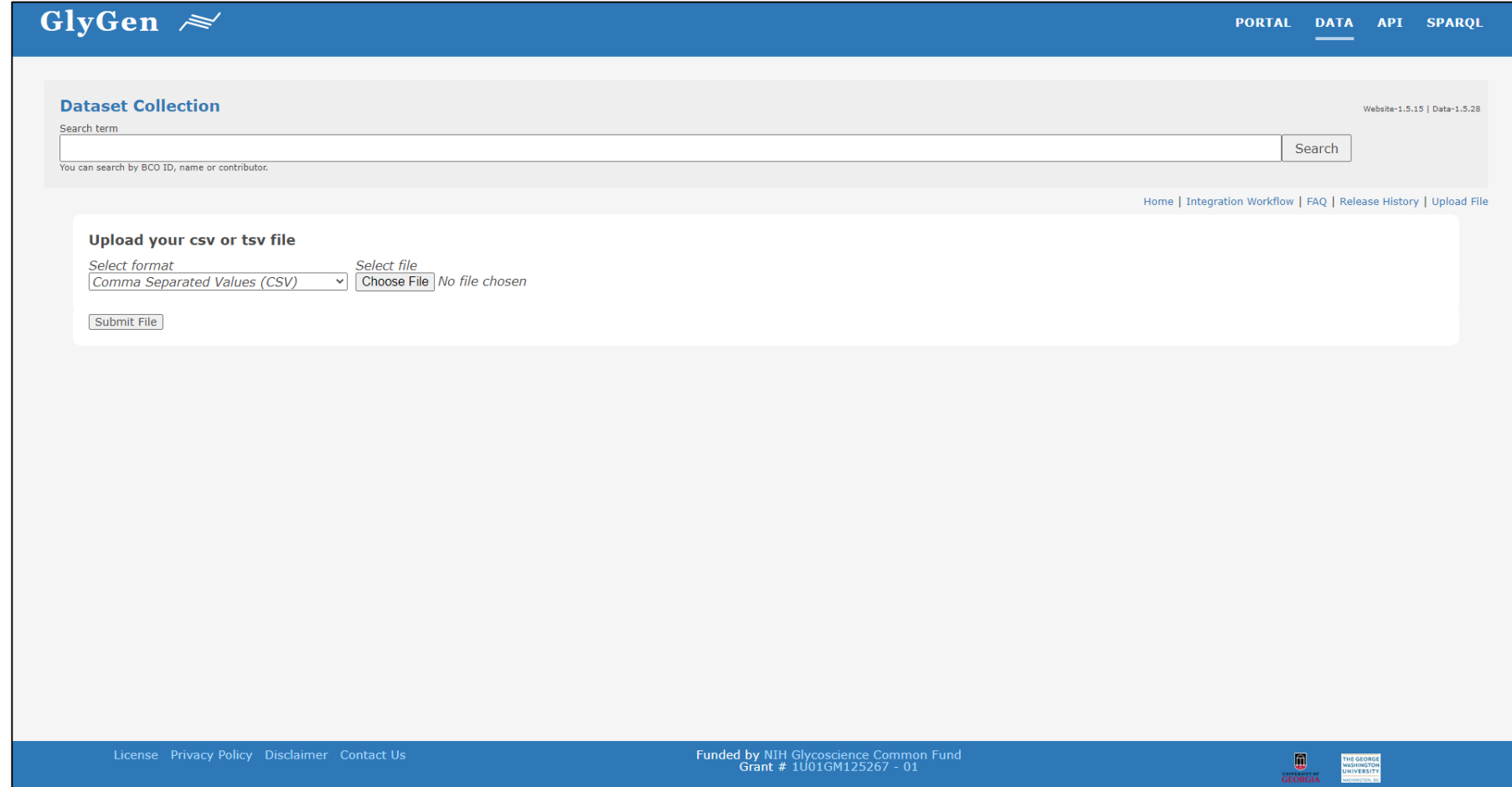
Access the Upload Interface at  
<https://data.glygen.org/upload>

## Overview

Upload CSV or TSV data files to  
GlyGen

Use automated QC checks to  
ensure data integrity

For any issues contact GlyGen  
help:  
<https://glygen.org/contact.html>




The screenshot displays the GlyGen web interface. At the top, a blue header bar contains the GlyGen logo and navigation links for PORTAL, DATA, API, and SPARQL. Below the header, a 'Dataset Collection' section features a search bar with a 'Search' button and a note: 'You can search by BCO ID, name or contributor.' To the right of the search bar, it indicates 'Website-1.5.15 | Data-1.5.28'. A secondary navigation bar includes links for Home, Integration Workflow, FAQ, Release History, and Upload File. The main content area is titled 'Upload your csv or tsv file' and contains a 'Select format' dropdown menu currently set to 'Comma Separated Values (CSV)', a 'Select file' section with a 'Choose File' button and the text 'No file chosen', and a 'Submit File' button. The footer of the page is blue and contains links for License, Privacy Policy, Disclaimer, and Contact Us, along with funding information: 'Funded by NIH Glycoscience Common Fund Grant # 1U01GM125267 - 01'. Logos for the University of California and The George Washington University are also present in the footer.

# Upload a Data File

Choose the file format and select the file to upload

For this example we are uploading a CSV file:  
`glygen_upload_example_file.csv`

**GlyGen** 

PORTAL **DATA** API SPARQL

**Dataset Collection** Website-1.5.15 | Data-1.5.28

Search term

You can search by BCO ID, name or contributor.



[Home](#) | [Integration Workflow](#) | [FAQ](#) | [Release History](#) | [Upload File](#)

**Upload your csv or tsv file**

Select format  Select file  glygen\_upload...mple\_file.csv

[License](#) [Privacy Policy](#) [Disclaimer](#) [Contact Us](#)

Funded by NIH Glycoscience Common Fund  
Grant # 1U01GM125267 - 01

# Review the Basic Sanity Check

The First QC step is called the **Basic Sanity Check**

The report will list the number of rows, columns, and any failed rows that need review

The Basic Sanity Check report

In this example, we have no failed rows that need review, so can proceed to the next step

Upload your csv or tsv file

Select format

Comma Separated Values (CSV)

Select file

Choose File

glygen\_upload...mple\_file.csv

Submit File

Basic Sanity Check

field_name	field_value
rows	7
columns	4
failed_rows	0

Mapping to GlyGen IDs

Protein Site Check

Records Preview

# Mapping to GlyGen IDs

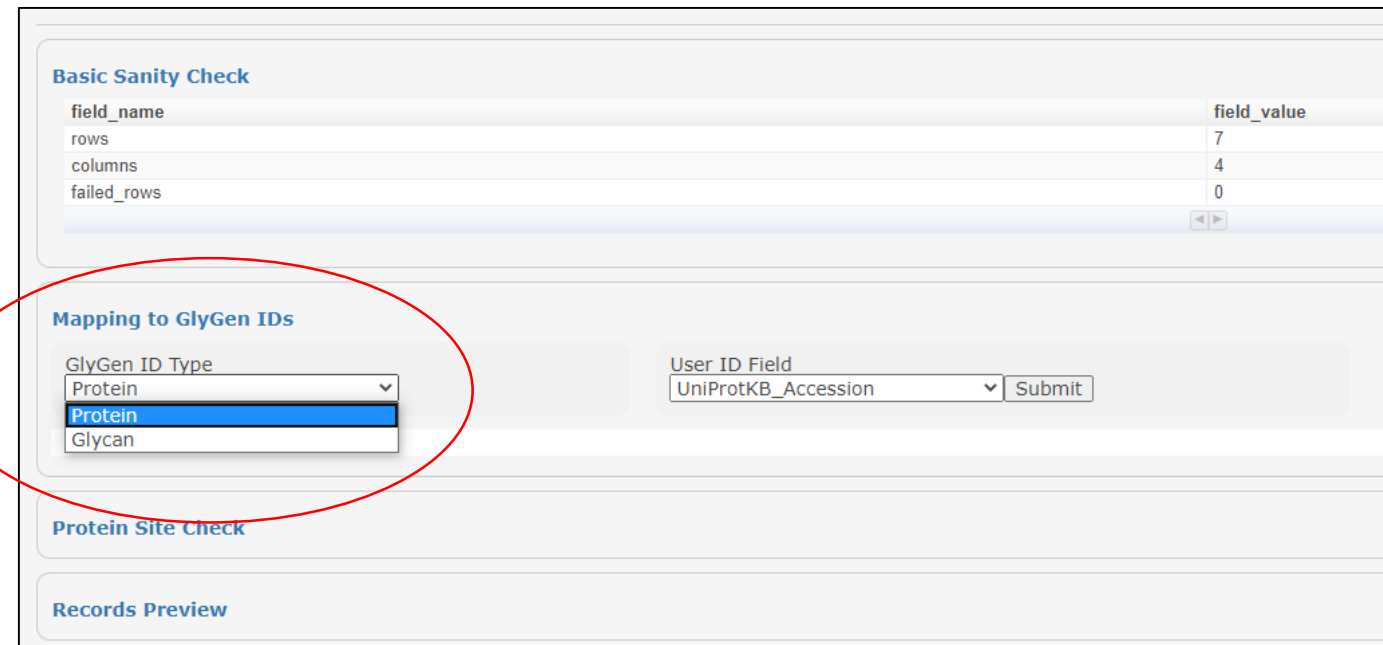
For **Mapping to GlyGen IDs**, a field in the upload file can be mapped to IDs already present in GlyGen

Select the GlyGen ID Type and User ID Field from the drop-down menus, then click Submit

For this example we have chosen:

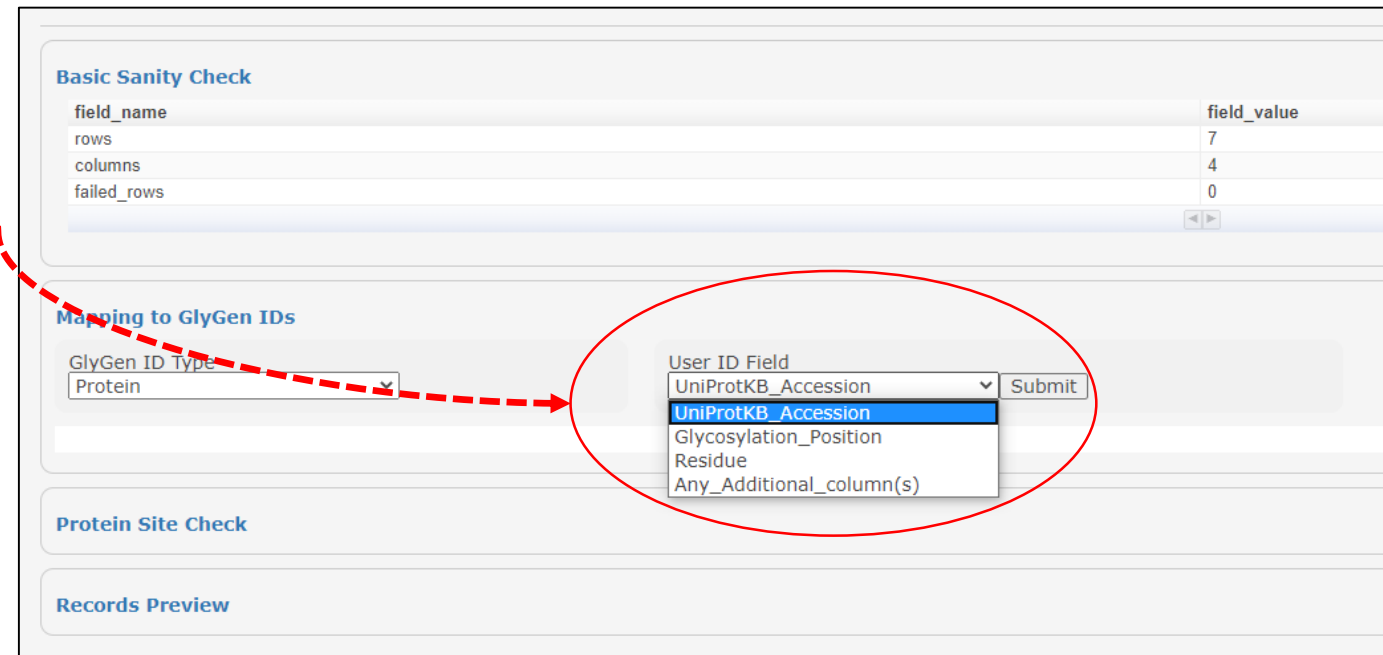
**GlyGen ID Type:** Protein

**User ID Field:** UniProtKB\_Accession



The screenshot shows the 'Mapping to GlyGen IDs' section of a web interface. The 'GlyGen ID Type' dropdown menu is open, displaying three options: 'Protein' (highlighted in blue), 'Protein', and 'Glycan'. The 'User ID Field' dropdown menu is set to 'UniProtKB\_Accession'. A red dashed arrow points from the text box on the left to the 'GlyGen ID Type' dropdown menu.

field_name	field_value
rows	7
columns	4
failed_rows	0



The screenshot shows the 'Mapping to GlyGen IDs' section of a web interface. The 'GlyGen ID Type' dropdown menu is set to 'Protein'. The 'User ID Field' dropdown menu is open, displaying five options: 'UniProtKB\_Accession' (highlighted in blue), 'Glycosylation\_Position', 'Residue', and 'Any\_Additional\_column(s)'. A red dashed arrow points from the text box on the left to the 'User ID Field' dropdown menu.

field_name	field_value
rows	7
columns	4
failed_rows	0

# Protein Site Check

For **Protein Site Check**, evaluate the uploaded residue information for specific positions in the protein

For the Protein Site Check, select from the drop-down menus:  
User Protein ID Field  
User Protein Position Field  
User Protein Residue Field  
then click Submit

For this example we have chosen:

**User Protein ID Field:** UniProtKB\_Accession

**User Protein Position Field:** Glycosylation\_Position

**User Protein Residue Field:** Residue

The screenshot shows the 'Protein Site Check' web interface. It includes a 'Basic Sanity Check' table, a 'Mapping to GlyGen IDs' section, and the 'Protein Site Check' section. The 'Protein Site Check' section has three dropdown menus: 'User Protein ID Field' (set to 'UniProtKB\_Accession'), 'User Protein Position Field' (set to 'Glycosylation\_Position'), and 'User Protein Residue Field' (set to 'Residue'). A red dashed arrow points from the text 'User Protein ID Field' in the instructions to the 'User Protein ID Field' dropdown menu. Below the dropdowns is a 'Submit' button. At the bottom, there is a 'Records Preview' section.

field_name	field_value
rows	7
columns	4
failed_rows	0

**Mapping to GlyGen IDs**

GlyGen ID Type: Protein User ID Field: UniProtKB\_Accession Submit

unmapped_id	comment
-------------	---------

**Protein Site Check**

User Protein ID Field: UniProtKB\_Accession User Protein Position Field: Glycosylation\_Position User Protein Residue Field: Residue Submit

position	residue	comment
----------	---------	---------

**Records Preview**

# Records Preview

With Records Preview, check that the uploaded data is previewed correctly. The report will list all the fields and data from the uploaded file

You can also download the file from the download option in top right corner.

The downloaded file will also provide “qc\_comment” for Qc checks done.

The Records Preview report is displayed in the final section

Basic Sanity Check

field_name	field_value
rows	7
columns	4
failed_rows	0

Mapping to GlyGen IDs

GlyGen ID Type

Protein

User ID Field

UniProtKB\_Accession

Submit

unmapped\_id

comment

Protein Site Check

User Protein ID Field

UniProtKB\_Accession

User Protein Position Field

Glycosylation\_Position

User Protein Residue Field

Residue

Submit

protein\_id

position

residue

comment

Records Preview

UniProtKB_Accession	Glycosylation_Position	Residue	Any_Aditional_column(s)
P07498	143	Thr	-
Q8WWV6	423	Thr	-
Q9NQ40	182	Thr	-
P00747	308	Asn	-
P07911	76	Asn	-
P16112	1599	Ser	-

# Save File

When the file has been uploaded and passed all QC checks, the uploaded file can be saved to GlyGen

Click Save File

Then fill out the information fields:  
First Name  
Last Name  
Email Address  
Institution  
and click Save

Upload your csv or tsv file

Select format

Comma Separated Values (CSV)

Select file

Choose File

glygen\_upload...mple\_file.csv

Submit File

First name

Last name

Email Address

Institute

Save

Basic Sanity Check

field_name	field_value
rows	7
columns	4
failed_rows	0

Mapping to GlyGen IDs

GlyGen ID Type

Protein

User ID Field

UniProtKB\_Accession

Submit

unmapped\_id

comment

Protein Site Check

User Protein ID Field

UniProtKB\_Accession

User Protein Position Field

Glycosylation\_Position

User Protein Residue Field

Residue

Submit

protein\_id

position

residue

comment

Records Preview

UniProtKB_Accession	Glycosylation_Position	Residue	Any_Aadditional_column(s)
P07498	143	Thr	-
Q8WWV6	423	Thr	-
Q9NQ40	182	Thr	-
P00747	308	Asn	-
P07911	76	Asn	-
P18112	1599	Ser	-

Save File