# GlyComb: a novel glycoconjugate data repository that bridges glycomics and proteomics

## Description

Protein glycosylation is a general post-translational modification involved in cell membrane formation, is crucial to dictate proper conformation of many membrane proteins, retaining stability on some secreted glycoproteins, and playing a role in cellâ??cell adhesion. Therefore, the analysis of post-translational modifications of proteins, including glycosylation, is essential to understand the exact functions and interactions of each protein molecule. To conduct large-scale analyses more efficiently, it is essential to promote the accumulation, sharing, and reuse of experimental and analytical data in accordance with the FAIR (Findability, Accessibility, Interoperability, and Re-usability) data principles. However, there was no FAIR data repository for storing and sharing information on glycoconjugates, including glycopeptides and glycoproteins, in a standardised format.

The authors have developed GlyComb (https://glycomb.glycosmos.org) as a new standardized data repository for glycoconjugate data. GlyComb can assign a unique identifier to a set of glycosylation information associated with a specific peptide sequence or UniProt ID. By standardising glycosylation data using GlyComb identifiers and coordinating with existing web resources such as GlyTouCan and GlycoPOST, a comprehensive system for data submission and sharing between researchers can be established. This article describes how GlyComb can integrate the variety of glycoconjugate data already registered in existing data repositories to gain a better understanding of the available glycopeptides and glycoproteins and their glycosylation patterns. We also explain how this system can be used as a basis for a better understanding of glycan function.
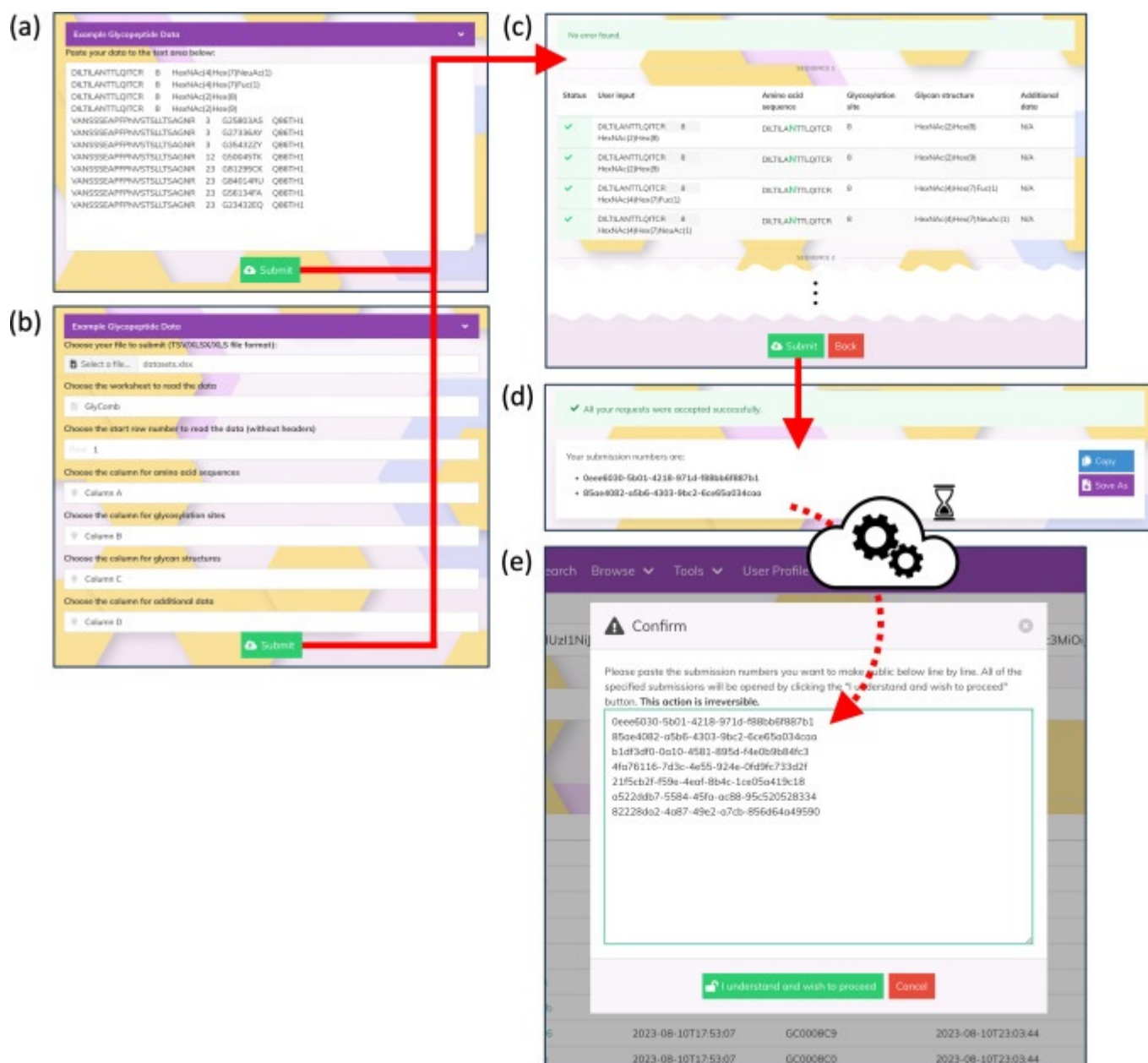
Workflow for registering and making glycopeptide entries publicly available in GlyComb. Researchers access the entry registration screen ( https://glycomb.glycosmos.org/registration) after logging into GlyComb with their Google account. They can either (a) copy and paste the entries they want to register from the clipboard into the text area on the screen, or (b) select an MS Excel file or TSV file to upload. When uploading a file, they can choose the worksheet to be read or select the columns to be read. (c) A confirmation screen to verify that the submission content is correct will be displayed. (d) GlyComb displays a submission number instead of the GlyComb ID, which is the accession number when the submission of an input entry is completed. Researchers can later make their registered entries open to the public by using these submission numbers. (e) A confirmation screen to make each registered entry available to the public (https://glycomb.glycosmos.org/user_profile). Through batch processing on the GlyComb server, they are assigned unique GlyComb IDs within a few hours after the input entries are submitted and each entry is ready to be published. By entering the submission

numbers generated when the glycopeptide or glycoprotein entries were submitted, one per line, and clicking the button at the bottom of the pop-up window, researchers can make multiple GlyComb entries they have submitted open to the public at once.

**Category**

1. News