

Package: glydb (via r-universe)

October 16, 2026

Title Glycan Structure Database

Version 0.7.2

Description Provides a comprehensive database of glycan structures from GlyTouCan, including fully determined glycan structures with complete linkage, substituent, anomer, and monosaccharide information. This database serves as a foundational resource for the glycoverse ecosystem, enabling glycan structure analysis, comparison, and research applications.

License MIT + file LICENSE

Suggests glyparse (>= 0.8.1), testthat (>= 3.0.0), tibble

Config/testthat/edition 3

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

URL <https://glycoverse.github.io/glydb/>

Depends R (>= 4.1)

LazyData true

LazyDataCompression bzip2

Imports glyrepr (>= 1.1.0), checkmate, cli, purrr, stringr, vctrs

Config/pak/sysreqs libglpk-dev libicu-dev libxml2-dev

Repository <https://glycoverse.r-universe.dev>

Date/Publication 2026-10-16 04:57:02 UTC

RemoteUrl <https://github.com/glycoverse/glydb>

RemoteRef v0.7.2

RemoteSha af614eaf3b6748fe0a8edc9f65b7bed8c61d1ada

Contents

glydb_compositions	2
glydb_data	3
glydb_species	4
glydb_structures	4
glytoucan_to_struc	5
struc_to_glytoucan	6
Index	7

glydb_compositions	<i>Get Compositions From Glydb Data</i>
--------------------	---

Description

Get unique glycan compositions from `glydb_data` as a `glyrepr::glycan_composition()` vector.

Usage

```
glydb_compositions(  
  mono_type = "concrete",  
  species = NULL,  
  glycan_type = NULL,  
  mono_range = NULL  
)
```

Arguments

mono_type	Either "generic" or "concrete". Default is "concrete". See <code>glyrepr::get_mono_type()</code> for details.
species	A string of specie names. See <code>glydb_species()</code> for available specie names. Default is NULL, which means glycans from all species are included.
glycan_type	A string of glycan types. Can be "HMO", "N", "GSL", "GAG", "O", "O-GalNAc", "O-GlcNAc", "O-Man", "O-Fuc", "O-Glc", or "GPI". When "O", all O-linked glycans are included. Specific O-glycan types only include that subtype. Default is NULL, which means glycans of all types are included.
mono_range	A named list for filtering compositions by monosaccharide counts. Each element should be an integer vector of length 2 specifying the minimum and maximum count for that monosaccharide. Monosaccharides not specified will be excluded (count = 0). Use NULL for no filtering. See examples for usage.

Value

A `glyrepr::glycan_composition()` vector, with a confidence attribute as a numeric vector with the same length.

Confidence

The returned value has a confidence attribute: a numeric vector of the same length as the result containing log-transformed citation counts for each glycan in glydb_data. When multiple glycans are aggregated into lower-resolution structures or compositions, the maximum confidence score is retained.

The confidence attribute is preserved by vector operations that subset, reorder, repeat, combine, or replace glycans. When vectors containing the same glycan with different confidence values are combined, the maximum confidence is used for every occurrence of that glycan.

Examples

```
glydb_compositions()
glydb_compositions(mono_type = "generic")
glydb_compositions(species = "Homo sapiens")
glydb_compositions(glycan_type = "N")
glydb_compositions(glycan_type = "N", mono_range = list(Hex = c(5L, 10L)))
glydb_compositions(mono_range = list(Hex = c(3L, 9L), HexNAc = c(2L, 6L)))
```

glydb_data

GlyGen Glycan Data

Description

A curated dataset glycan structures from GlyGen Data v2.12.1, with 49,728 glycan structures currently available.

Usage

```
glydb_data
```

Format

A tibble with 49,728 rows and 6 variables:

- glytoucan_ac: GlyTouCan accession.
- glycan_structure: Glycan structure (glyrepr::glycan_structure()).
- glycan_composition: Glycan composition (glyrepr::glycan_composition()).
- species: Species names, separated by semicolons. Unknown species are NAs.
- glycan_type: Glycan type, such as "HMO", "N", "GSL", "GAG", "O", "O-GalNAc", "O-GlcNAc", "O-Man", "O-Fuc", "O-Glc", or "GPI".
- confidence: Confidence score used to rank glycans.

Source

<https://data.glygen.org/ln2data/releases/data/v-2.12.1/reviewed/>

glydb_species

Get Supported Species From Glydb Data

Description

Get a character vector of supported species from [glydb_data](#).

Usage

```
glydb_species()
```

Value

A character vector of supported species.

Examples

```
glydb_species()
```

glydb_structures

Get Structures From Glydb Data

Description

Get unique glycan structures from [glydb_data](#) as a [glyrepr::glycan_structure\(\)](#) vector.

Usage

```
glydb_structures(
  structure_level = "intact",
  species = NULL,
  glycan_type = NULL,
  mono_range = NULL,
  mono_type = "concrete"
)
```

Arguments

structure_level

Either "intact" or "topological". Default is "intact". See [glyrepr::get_structure_level\(\)](#) for details.

species

A string of specie names. See [glydb_species\(\)](#) for available specie names. Default is NULL, which means glycans from all species are included.

glycan_type	A string of glycan types. Can be "HMO", "N", "GSL", "GAG", "O", "O-GalNAc", "O-GlcNAc", "O-Man", "O-Fuc", "O-Glc", or "GPI". When "O", all O-linked glycans are included. Specific O-glycan types only include that subtype. Default is NULL, which means glycans of all types are included.
mono_range	A named list for filtering structures by monosaccharide counts. Each element should be an integer vector of length 2 specifying the minimum and maximum count for that monosaccharide. Monosaccharides not specified will be excluded (count = 0). Use NULL for no filtering. See examples for usage.
mono_type	Either "generic" or "concrete". Default is "concrete". See glyrepr::get_mono_type() for details.

Value

A [glyrepr::glycan_structure\(\)](#) vector, with a confidence attribute as a numeric vector with the same length.

Confidence

The returned value has a confidence attribute: a numeric vector of the same length as the result containing log-transformed citation counts for each glycan in glydb_data. When multiple glycans are aggregated into lower-resolution structures or compositions, the maximum confidence score is retained.

The confidence attribute is preserved by vector operations that subset, reorder, repeat, combine, or replace glycans. When vectors containing the same glycan with different confidence values are combined, the maximum confidence is used for every occurrence of that glycan.

Examples

```
glydb_structures()
glydb_structures(structure_level = "topological")
glydb_structures(mono_type = "generic")
glydb_structures(structure_level = "topological", mono_type = "generic")
glydb_structures(species = "Homo sapiens")
glydb_structures(glycan_type = "N")
glydb_structures(glycan_type = "N", mono_range = list(Hex = c(5L, 10L)))
glydb_structures(mono_range = list(Hex = c(3L, 9L), HexNAc = c(2L, 6L)))
```

glytoucan_to_struc	<i>Convert GlyTouCan Accessions to Glycan Structures</i>
--------------------	--

Description

Look up GlyTouCan accessions in the bundled [glydb_data](#).

Usage

```
glytoucan_to_struc(glytoucan_ac)
```

Arguments

glytoucan_ac A character vector of GlyTouCan accessions.

Value

A `glyrepr::glycan_structure()` vector. Accessions without a bundled match are returned as NA values in their original positions, and a warning is emitted.

Examples

```
glytoucan_to_struc("G17689DH")
```

struc_to_glytoucan	<i>Convert Glycan Structures to GlyTouCan Accessions</i>
--------------------	--

Description

Match glycan structures to accessions in the bundled `glydb_data`.

Usage

```
struc_to_glytoucan(strucs)
```

Arguments

strucs A `glyrepr::glycan_structure()` vector or a character vector of IUPAC-condensed glycan structures.

Details

Missing anomeric positions are filled with `glyrepr::fill_anomer_pos()` in both the input and bundled structures before matching. When multiple bundled rows normalize to the same structure, the accession from the first row is returned.

Value

A character vector of GlyTouCan accessions. Structures without a bundled match are returned as NA values in their original positions, and a warning is emitted.

Examples

```
struc_to_glytoucan(glydb_data$glycan_structure[1])
```

Index

* datasets

- glydb_data, [3](#)
- glydb_compositions, [2](#)
- glydb_data, [2](#), [3](#), [4–6](#)
- glydb_species, [4](#)
- glydb_species(), [2](#), [4](#)
- glydb_structures, [4](#)
- glyrepr::fill_anomer_pos(), [6](#)
- glyrepr::get_mono_type(), [2](#), [5](#)
- glyrepr::get_structure_level(), [4](#)
- glyrepr::glycan_composition(), [2](#)
- glyrepr::glycan_structure(), [4–6](#)
- glytoucan_to_struc, [5](#)
- struc_to_glytoucan, [6](#)