Package: glydb (via r-universe)

June 3, 2026

Title Glycan Structure Database

Version 0.5.0

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 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), glyrepr (>= 0.10.0)

LazyData true

Imports checkmate, cli, glyparse, httr2, purrr, stringr, vctrs

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

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

Date/Publication 2026-06-03 05:56:07 UTC

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

RemoteRef v0.5.0

RemoteSha 47a2c7171c80e839e9786a791b26c15885238596

Contents

`glydb_compositions` *Get Compositions From Glydb Data*

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 `glyrepr::get_mono_type()` 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 "N", "O-GalNAc", "O-GlcNAc", "O-Man", "O-Fuc", "O-Glc". 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.

Note that the `confidence` attribute will be lost after any vector operation like subsetting. Therefore, if used with `glyanno`, the returned value should not be modified manually.

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 *Fully Determined GlyTouCan Glycan Data*

Description

A curated dataset of fully determined glycans from GlyTouCan. "Fully determined" means that all linkages, substituents, anomers, and monosaccharides are fully specified. The dataset is derived from the GlyTouCan v2.11.1 release, with 7,125 glycan structures currently available.

Usage

```
glydb_data
```

Format

A tibble with 7,125 rows and 5 variables:

- `glytoucan_ac`: GlyTouCan accession.
- `glycan_structure`: Glycan structure (glyrepr::glycan_structure()).
- `glycan_composition`: Glycan composition (glyrepr::glycan_composition()).
- `species`: Specie names, separated by semicolons. Unknown species are NAs.
- `glycan_type`: Glycan type, one of "N", "O-GalNAc", "O-GlcNAc", "O-Man", "O-Fuc", "O-Glc".

Source

https://data.glygen.org

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
)
```

Arguments

structure_level	Either "intact", "topological", or "basic". 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 "N", "O-GalNAc", "O-GlcNAc", "O-Man", "O-Fuc", "O-Glc". 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.

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.

Note that the confidence attribute will be lost after any vector operation like subsetting. Therefore, if used with glyanno, the returned value should not be modified manually.

Examples

```
glydb_structures()
glydb_structures(structure_level = "topological")
glydb_structures(structure_level = "basic")
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 *Convert GlyTouCan Accessions to Glycan Structures*

Description

Fetch GlyTouCan accessions from the GlyGen API and parse the returned IUPAC strings as glyrepr::glycan_structure() values.

Usage

```
glytoucan_to_struc(glytoucan_ac)
```

Arguments

glytoucan_ac	A character vector of GlyTouCan accessions.

Value

A glyrepr::glycan_structure() vector. Accessions that cannot be fetched or parsed are returned as NA values in their original positions, and a warning is emitted.

Examples

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

Index