

○山田 一作<sup>1</sup>、松原 正陽<sup>1</sup>、土屋 伸一郎<sup>1</sup>、木下 聖子<sup>2</sup>

1.公益財団法人 野口研究所 糖鎖情報科学研究室、2.創価大学 糖鎖生命システム融合研究所

**GlycoNAVI**  
Support System for Carbohydrate Research



# 背景

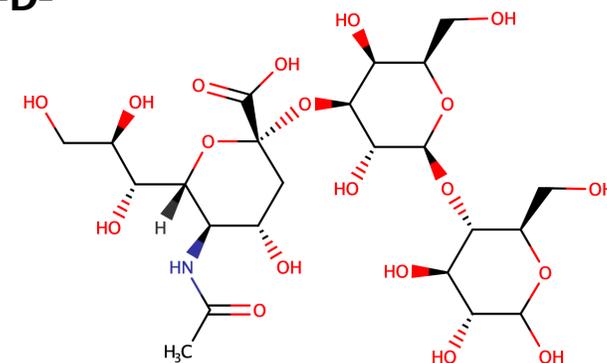
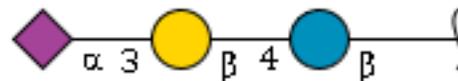
糖鎖研究において糖鎖は名称、略号、記号など多様な方法により記述されている。これらは糖鎖の研究者であれば各々の分野においては理解できる。しかし、略号で糖鎖を理解している研究者が、化学構造式で描かれた糖鎖構造を**同じ糖鎖であるかを判断することは容易ではない。**

$\alpha$ -N-acetylneuraminy-(2→3)- $\beta$ -D-galactosyl-(1→4)- $\beta$ -D-glucopyranose

5-acetamido-3,5-dideoxy-D-glycero- $\alpha$ -D-galacto-non-2-ulopyranonosyl-(2→3)- $\beta$ -D-galactopyranosyl-(1→4)- $\beta$ -D-glucopyranose

Neu5Ac( $\alpha$ 2-3)Gal( $\beta$ 1-4)Glc $\beta$ 1

GM3



名称

略号

記号

化学構造式

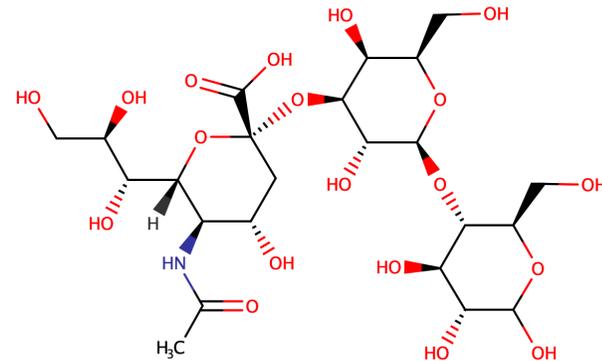
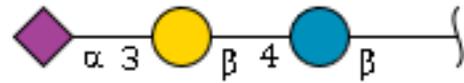
# 目的

研究者が理解可能な方法（今は 文字列）により、糖鎖を記述しデータを検索可能とすることを目的とする。

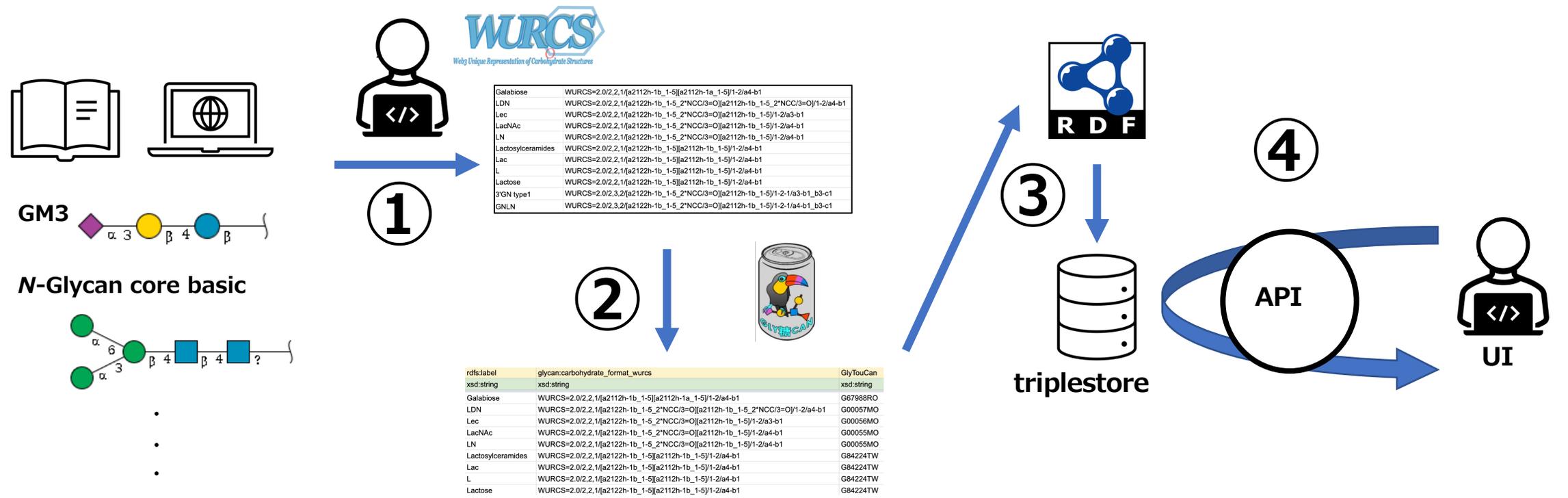
$\alpha$ -N-acetylneuraminyl-(2→3)- $\beta$ -D-galactosyl-(1→4)- $\beta$ -D-glucopyranose

Neu5Ac( $\alpha$ 2-3)Gal( $\beta$ 1-4)Glc $\beta$ 1

GM3



# 手順



略号、慣用名、モチーフ名など

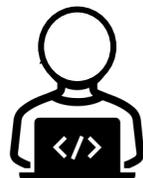


# ①: 糖鎖の画像をWURCSへ変換

## WURCSとは?

- 糖鎖構造の線形文字列表記法、糖鎖構造を一意の文字列で表現可能

糖鎖の略号に対応する糖鎖画像を**GlycanBuilder2**を用いてWURCSへと変換

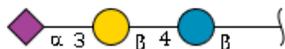


WURCS=2.0/2,2,1/[a2112h-1b\_1-5][a2112h-1a\_1-5]/1-2/a4-b1

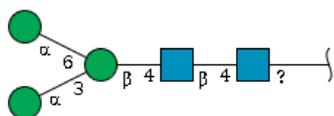
人力

Google Sheets

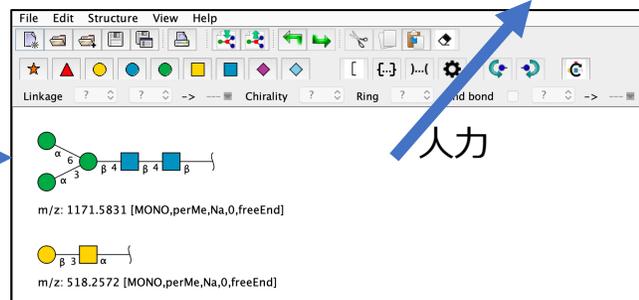
GM3



N-Glycan core basic



人力



人力

**GlycanBuilder2**

開発している糖鎖構造編集ソフト

Galabiose	WURCS=2.0/2,2,1/[a2112h-1b_1-5][a2112h-1a_1-5]/1-2/a4-b1
LDN	WURCS=2.0/2,2,1/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5_2*NCC/3=O]/1-2/a4-b1
Lec	WURCS=2.0/2,2,1/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5]/1-2/a3-b1
LacNAc	WURCS=2.0/2,2,1/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5]/1-2/a4-b1
LN	WURCS=2.0/2,2,1/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5]/1-2/a4-b1
Lactosylceramides	WURCS=2.0/2,2,1/[a2122h-1b_1-5][a2112h-1b_1-5]/1-2/a4-b1
Lac	WURCS=2.0/2,2,1/[a2122h-1b_1-5][a2112h-1b_1-5]/1-2/a4-b1
L	WURCS=2.0/2,2,1/[a2122h-1b_1-5][a2112h-1b_1-5]/1-2/a4-b1
Lactose	WURCS=2.0/2,2,1/[a2122h-1b_1-5][a2112h-1b_1-5]/1-2/a4-b1
3'GN type1	WURCS=2.0/2,3,2/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5]/1-2-1/a3-b1_b3-c1
GNLN	WURCS=2.0/2,3,2/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5]/1-2-1/a4-b1_b3-c1

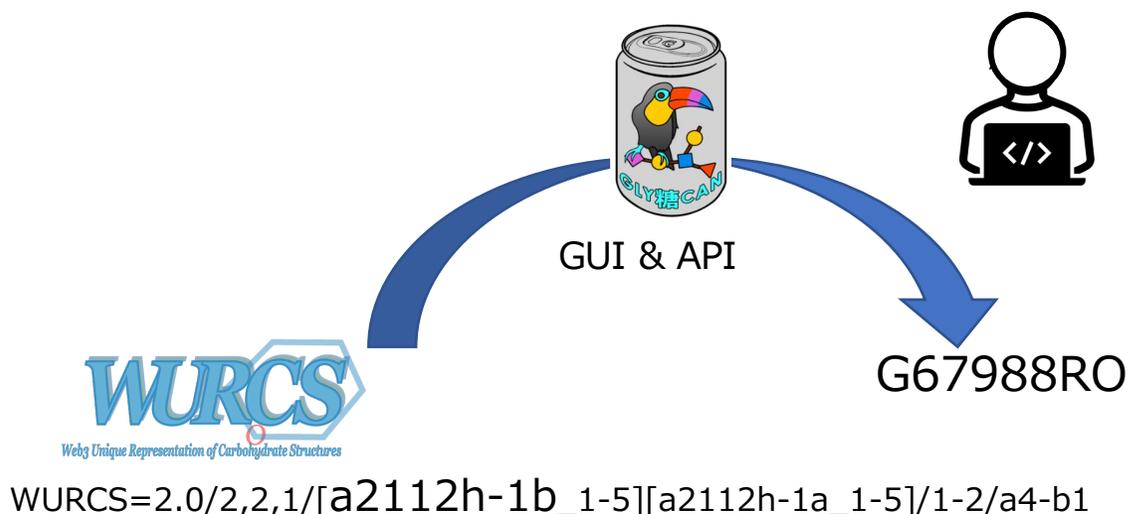
# ② : GlyTouCanのアクセス番号付与

GlyTouCanとは？



糖鎖構造に固有のアクセス番号を付与するための国際糖鎖構造リポジトリ

WURCSを国際糖鎖構造リポジトリGlyTouCanに登録しアクセス番号を付与



Google Sheets

rdfs:label	glycan:carbohydrate_format_wurcs	GlyTouCan
Galabiose	WURCS=2.0/2,2,1/[a2112h-1b_1-5][a2112h-1a_1-5]/1-2/a4-b1	G67988RO
LDN	WURCS=2.0/2,2,1/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5_2*NCC/3=O]/1-2/a4-b1	G00057MO
Lec	WURCS=2.0/2,2,1/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5]/1-2/a3-b1	G00056MO
LacNAc	WURCS=2.0/2,2,1/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5]/1-2/a4-b1	G00055MO
LN	WURCS=2.0/2,2,1/[a2122h-1b_1-5_2*NCC/3=O][a2112h-1b_1-5]/1-2/a4-b1	G00055MO
Lactosylceramides	WURCS=2.0/2,2,1/[a2122h-1b_1-5][a2112h-1b_1-5]/1-2/a4-b1	G84224TW
Lac	WURCS=2.0/2,2,1/[a2122h-1b_1-5][a2112h-1b_1-5]/1-2/a4-b1	G84224TW
L	WURCS=2.0/2,2,1/[a2122h-1b_1-5][a2112h-1b_1-5]/1-2/a4-b1	G84224TW
Lactose	WURCS=2.0/2,2,1/[a2122h-1b_1-5][a2112h-1b_1-5]/1-2/a4-b1	G84224TW

# ③ : Resource Description Framework(RDF)化

## データのRDF化を実施

### Google Sheets

rdfs:label	glycan:carbohydrate_format_wurcs	GlycoCan
xsd:string	xsd:string	xsd:string
Galactose	WURCS=2.0/2.1/[a2122h-1b_1-5][a2122h-1b_1-5]/1-2/a4-b1	G07098RO
LDN	WURCS=2.0/2.1/[a2122h-1b_1-5_2*NCC/3=O][a2122h-1b_1-5_2*NCC/3=O]/1-2/a4-b1	G00057MO
Lec	WURCS=2.0/2.1/[a2122h-1b_1-5_2*NCC/3=O][a2122h-1b_1-5]/1-2/a4-b1	G00056MO
LacNAc	WURCS=2.0/2.1/[a2122h-1b_1-5_2*NCC/3=O][a2122h-1b_1-5]/1-2/a4-b1	G00055MO
LN	WURCS=2.0/2.1/[a2122h-1b_1-5_2*NCC/3=O][a2122h-1b_1-5]/1-2/a4-b1	G00059MO
Lactosylceramides	WURCS=2.0/2.1/[a2122h-1b_1-5][a2122h-1b_1-5]/1-2/a4-b1	G84224TW
Lac	WURCS=2.0/2.1/[a2122h-1b_1-5][a2122h-1b_1-5]/1-2/a4-b1	G84224TW
L	WURCS=2.0/2.1/[a2122h-1b_1-5][a2122h-1b_1-5]/1-2/a4-b1	G84224TW
Lactose	WURCS=2.0/2.1/[a2122h-1b_1-5][a2122h-1b_1-5]/1-2/a4-b1	G84224TW



CSV, TSV



RDF化

### RDF/Turtle

```

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix sio: <http://semanticscience.org/resource/> .
@prefix glycan: <http://purl.jp/bio/12/glyco/glycan#> .
@prefix can: <http://rdf.glyconavi.org/CAN/> .
    
```

```

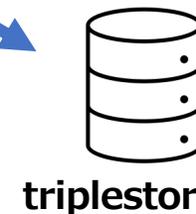
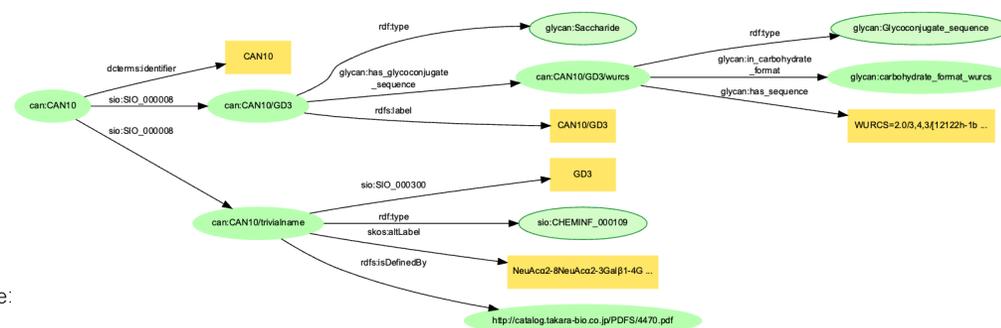
can:CAN10
  dcterms:identifier "CAN10" ;
  sio:SIO_000008 <http://rdf.glyconavi.org/CAN/CAN10/GD3> ;
  sio:SIO_000008 <http://rdf.glyconavi.org/CAN/CAN10/trivialname>
    <http://rdf.glyconavi.org/CAN/CAN10/trivialname>
      sio:SIO_000300 "GD3" ;
      a sio:CHEMINF_000109 ;
      skos:altLabel "NeuAc α 2-8NeuAc α 2-3Gal β 1-4Glc β 1-1Cer" ;
      rdfs:isDefinedBy <http://catalog.takara-bio.co.jp/PDFS/4470.pdf> .

<http://rdf.glyconavi.org/CAN/CAN10/GD3>
  a glycan:Saccharide ;
  glycan:has_glycoconjugate_sequence

<http://rdf.glyconavi.org/CAN/CAN10/GD3/wurcs> ;
  rdfs:label "CAN10/GD3" .

<http://rdf.glyconavi.org/CAN/CAN10/GD3/wurcs>
  a glycan:Glycoconjugate_sequence ;
  glycan:in_carbohydrate_format glycan:carbohydrate_format_wurcs ;
  glycan:has_sequence "WURCS=2.0/3,4,3/[12122h-1b_1-5][12122h-1b_1-5][a6d21122h-2a_2-6_5*NCC/3=O]/1-2-3-3/a4-b1_b3-c2_c8-d2" .
    
```

### RDF Schema



triplestore

# ④ : データ検索 - SPARQList -

データ検索



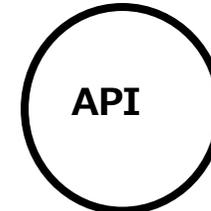
SPARQL 記述支援ツール



SPARQL クエリの結果をクライアントに返す REST API を提供するためのミドルウェア

SPARQList

検索用API作成



<https://wiki.lifesciencedb.jp/mw/SPARQList>

# 結果：糖鎖を略号などの文字列で検索可能

GM3



WURCS  
Web3 Unique Representation of Carbohydrate Structures



GlycoNAVI  
Support System for Carbohydrate Research

WORLDWIDE  
PDB  
PROTEIN DATA BANK



GlyGen

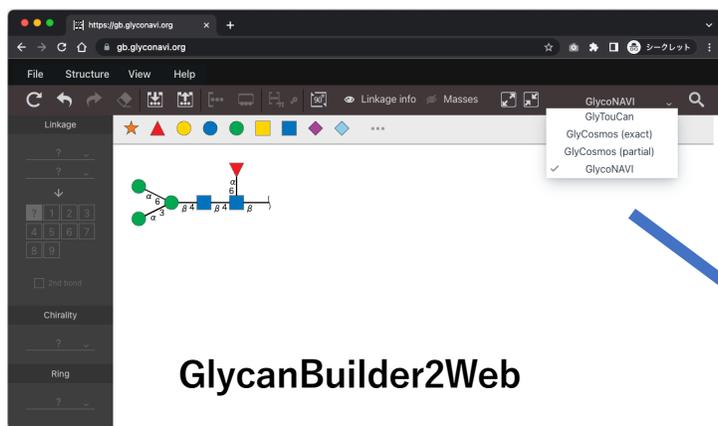
PubChem



# やさしい糖鎖データ検索を目指して

## WURCSによる糖鎖構造の統合化

### 糖鎖の構造



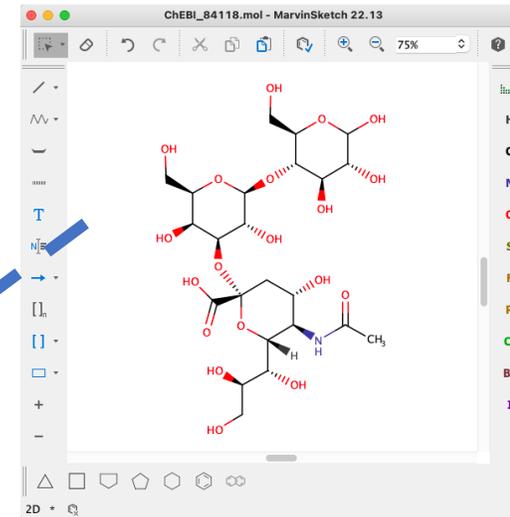
### 糖鎖構造の名前

GM3

**WURCS**  
Web3 Unique Representation of Carbohydrate Structures



### 糖鎖の化学構造



# まとめ

多様な分野の研究者が糖鎖のデータを検索する際に、適した方法を提供することは利用者にとってやさしいサイトにする上で重要である。

我々は、糖鎖構造を描画するツール (SugarDrawer, GlycanBuilder2Web) を開発することにより、検索したい糖鎖構造を描画する方法でデータベース検索する手法を提供してきた。

今回、糖鎖構造をあらわす「文字列」を利用して、データベース検索する手法を開発した。これにより、WURCSを利用しているデータベースのデータを検索することが可能な新しい方法を提供することができた。