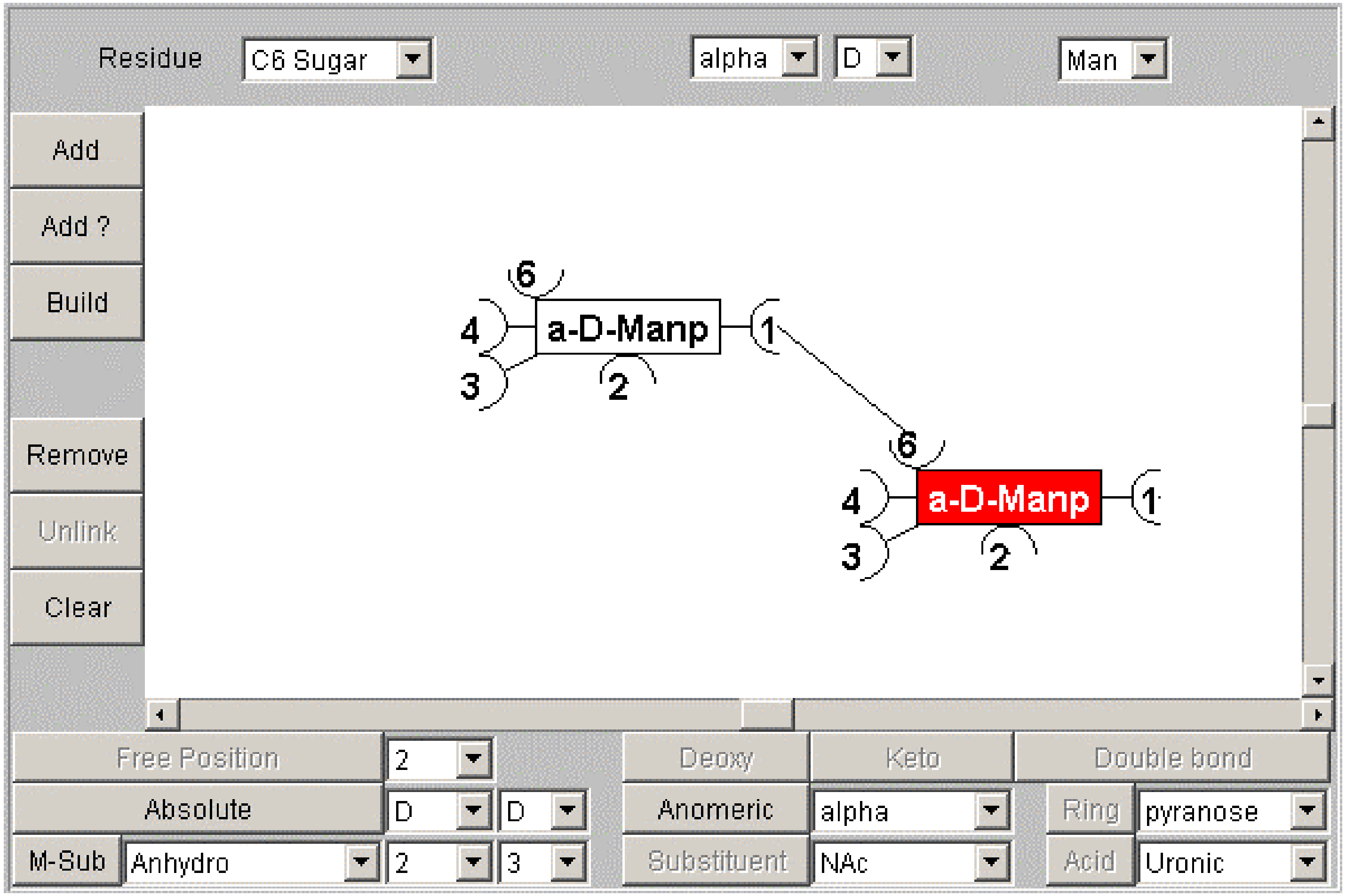
Web resources for the Carbohydrate Chemist

Roland Stenutz
Stockholm University
roland@organ.su.se

SugaBase/CarbBank

http://www.boc.chem.uu.nl/sugabase/databases.html



Residue: C6 Sugar; alpha; D; Man

Free Position: 2

Absolute: D; D

M-Sub: Anhydro; 2; 3

Anomeric: alpha

Substituent: NAc

Ring: pyranose

Acid: Uronic

CarbBank results

Query Results

Your query with the structure:

```
a-D-Manp-(1-6)-a-D-Manp
```

resulted in **459** hit(s)

Only the first 5 out of 459 hits are displayed!

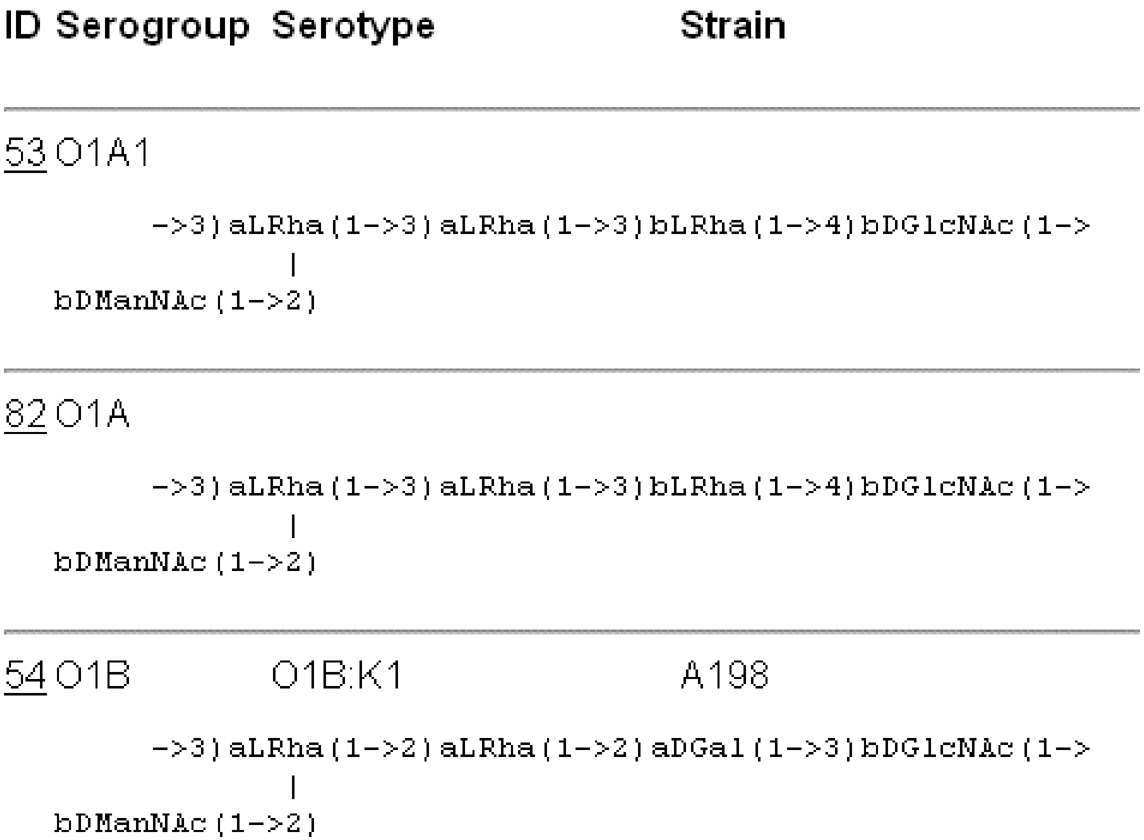
Use '**Back**' button to return to form

```
***** Hit 1 *****
CC: CCSD:4023
AU: Gunnarsson A; Svensson B; Nilsson B; Svensson S
TI: Structural studies on the O-glycosidically linked carbohydrate
chains of glucoamylase G1 from Aspergillus niger
CT: Eur J Biochem (1984) 145: 463-467
FC: 19957e6c
BS: (GS) Aspergillus niger
AG: Threonine
SB: van Kuik A
DA: 06-02-1991
MT: O-linked glycoprotein
PM: glucoamylase G1
SI: CBank:4618
----------------
structure:

 a-D-Manp-(1-6)-a-D-Manp-(1-3)-Ser
================end of record
***** Hit 2 *****
CC: CCSD:8320
AU: Zopf DA; Tsai CM; Ginsburg V
TI: Antibodies against oligosaccharides
CT: ACS Symp Ser (1979) 88: 90-101
```

E. coli O-antigens

http://www.casper.organ.su.se/ECODAB/



ID Serogroup Serotype Strain

---

53 O1A1

```
     ->3)aLRha(1->3)aLRha(1->3)bLRha(1->4)bDGlcNAc(1->
            |
bDManNAc(1->2)
```

---

82 O1A

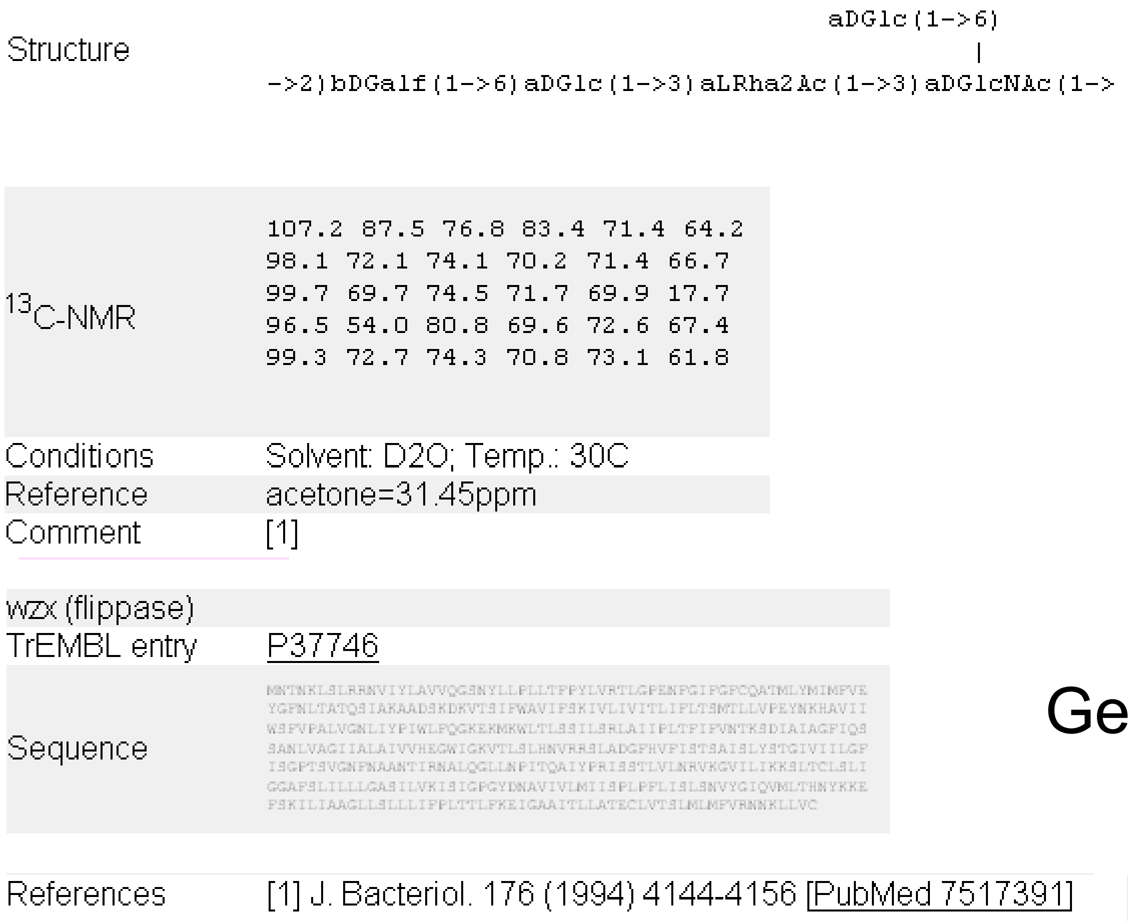
```
     ->3)aLRha(1->3)aLRha(1->3)bLRha(1->4)bDGlcNAc(1->
            |
bDManNAc(1->2)
```

---

54 O1B O1B:K1 A198

```
     ->3)aLRha(1->2)aLRha(1->2)aDGal(1->3)bDGlcNAc(1->
            |
bDManNAc(1->2)
```

Information about O-antigens



Structure

```
                                                aDGlc(1->6)
                                                          |
->2)bDGalf(1->6)aDGlc(1->3)aLRha2Ac(1->3)aDGlcNAc(1->
```

^{13}C-NMR	107.2 87.5 76.8 83.4 71.4 64.2 98.1 72.1 74.1 70.2 71.4 66.7 99.7 69.7 74.5 71.7 69.9 17.7 96.5 54.0 80.8 69.6 72.6 67.4 99.3 72.7 74.3 70.8 73.1 61.8
Conditions	Solvent: D2O; Temp.: 30C
Reference	acetone=31.45ppm
Comment	[1]

wzx (flippase)	
TrEMBL entry	P37746
Sequence	MNTNKLSLRRNVIYLAVVQGSNYLLPLLTFPYLVRTLGPENFGIFGFCQATMLYMIMFVE YGFNLTATQSIAKAADSKDKVTSIFWAVIFSKIVLIVITLIFLTSMTLLVPEYNKHAVII WSFVPALVGNLIYPIWLFQGKEKMKWLTLSSILSRLAIIPLTFIFVNTKSDIAIAGFIQS SANLVAGIIALAIVVHEGWIGKVTLSLHNVRRSLADGFHVPISTSAISLYSTGIVIILGF ISGPTSVGNFNAANTIRNALQGLLNPITQAIYPRISSTLVLNRVKGVILIKKSLTCLSLI GGAFSLILLLGASILVKISIGPGYDNAVIVLMIISPLPFLISLSNVYGIQVMLTHNYKKE FSKILIAAGLLSLLLIFPLTTLFKEIGAAITLLATECLVTSLMLMFVRNNKLLVC

References	[1] J. Bacteriol. 176 (1994) 4144-4156 [PubMed 7517391]

Structure

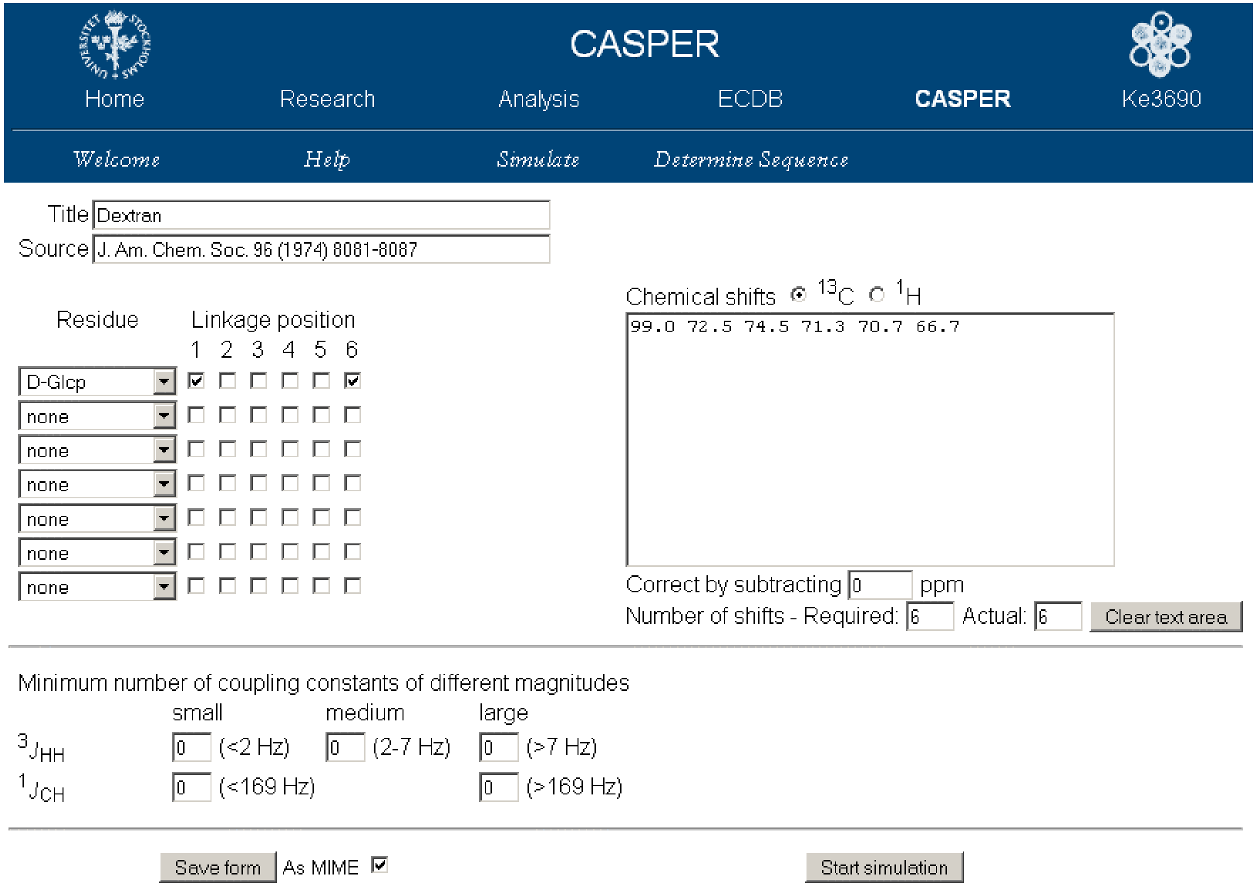
NMR-data

Genetic information

Links to PubMed

CASPER

http://www.casper.organ.su.se/casper



Title: Dextran

Source: J. Am. Chem. Soc. 96 (1974) 8081-8087

Residue	Linkage position 1	2	3	4	5	6
D-Glcp	☑	☐	☐	☐	☐	☑
none	☐	☐	☐	☐	☐	☐
none	☐	☐	☐	☐	☐	☐
none	☐	☐	☐	☐	☐	☐
none	☐	☐	☐	☐	☐	☐
none	☐	☐	☐	☐	☐	☐
none	☐	☐	☐	☐	☐	☐

Chemical shifts ● ^{13}C ○ ^{1}H

```
99.0 72.5 74.5 71.3 70.7 66.7
```

Correct by subtracting 0 ppm

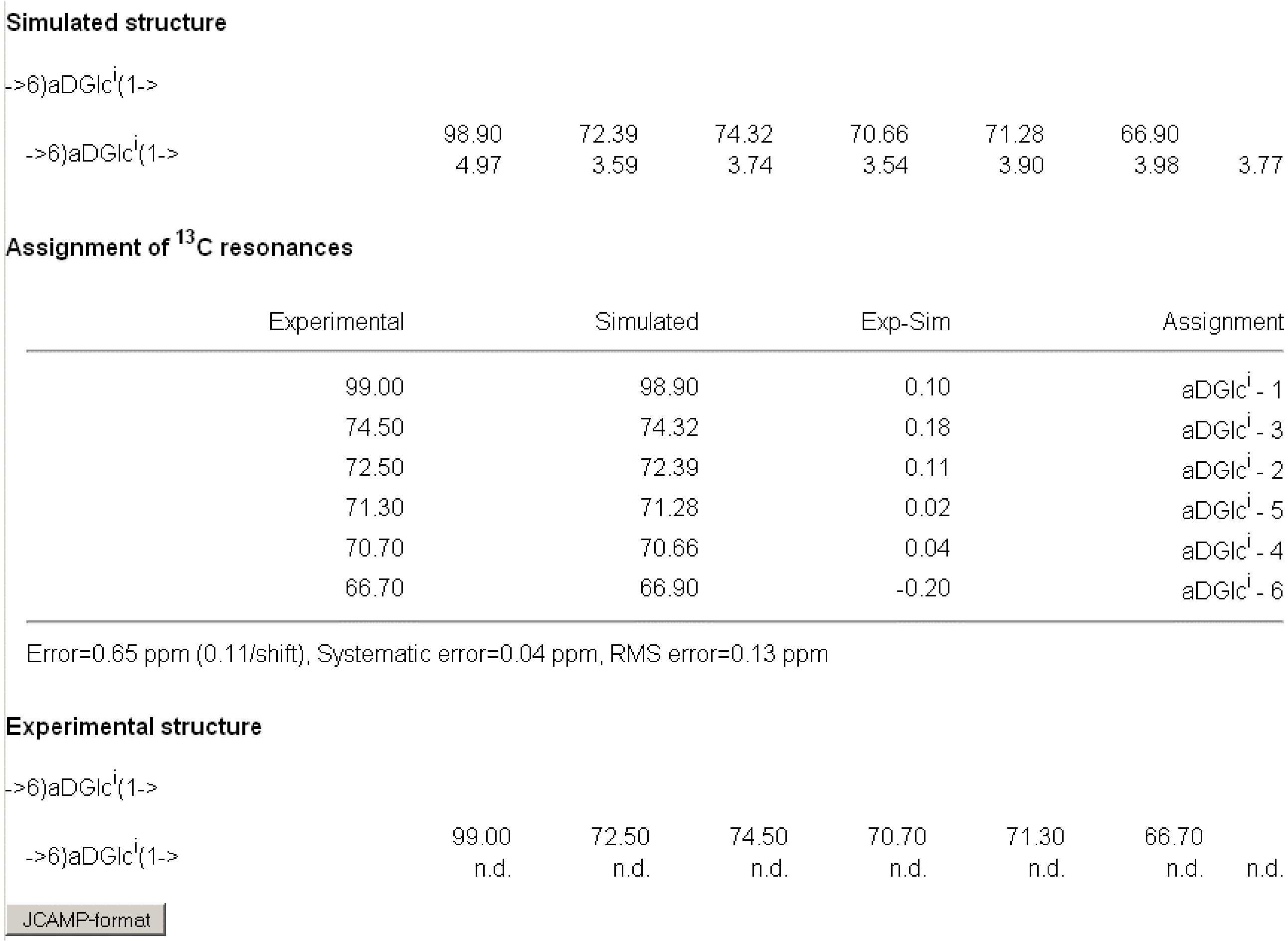
Number of shifts - Required: 6 Actual: 6

Minimum number of coupling constants of different magnitudes

	small	medium	large
$^3J_{HH}$	0 (<2 Hz)	0 (2-7 Hz)	0 (>7 Hz)
$^1J_{CH}$	0 (<169 Hz)		0 (>169 Hz)

As MIME ☑

CASPER



Simulated structure

->6)aDGlc^{i}(1->

->6)aDGlc^{i}(1->	98.90	72.39	74.32	70.66	71.28	66.90	
	4.97	3.59	3.74	3.54	3.90	3.98	3.77

Assignment of ^{13}C resonances

Experimental	Simulated	Exp-Sim	Assignment
99.00	98.90	0.10	aDGlc^{i} - 1
74.50	74.32	0.18	aDGlc^{i} - 3
72.50	72.39	0.11	aDGlc^{i} - 2
71.30	71.28	0.02	aDGlc^{i} - 5
70.70	70.66	0.04	aDGlc^{i} - 4
66.70	66.90	-0.20	aDGlc^{i} - 6

Error=0.65 ppm (0.11/shift), Systematic error=0.04 ppm, RMS error=0.13 ppm

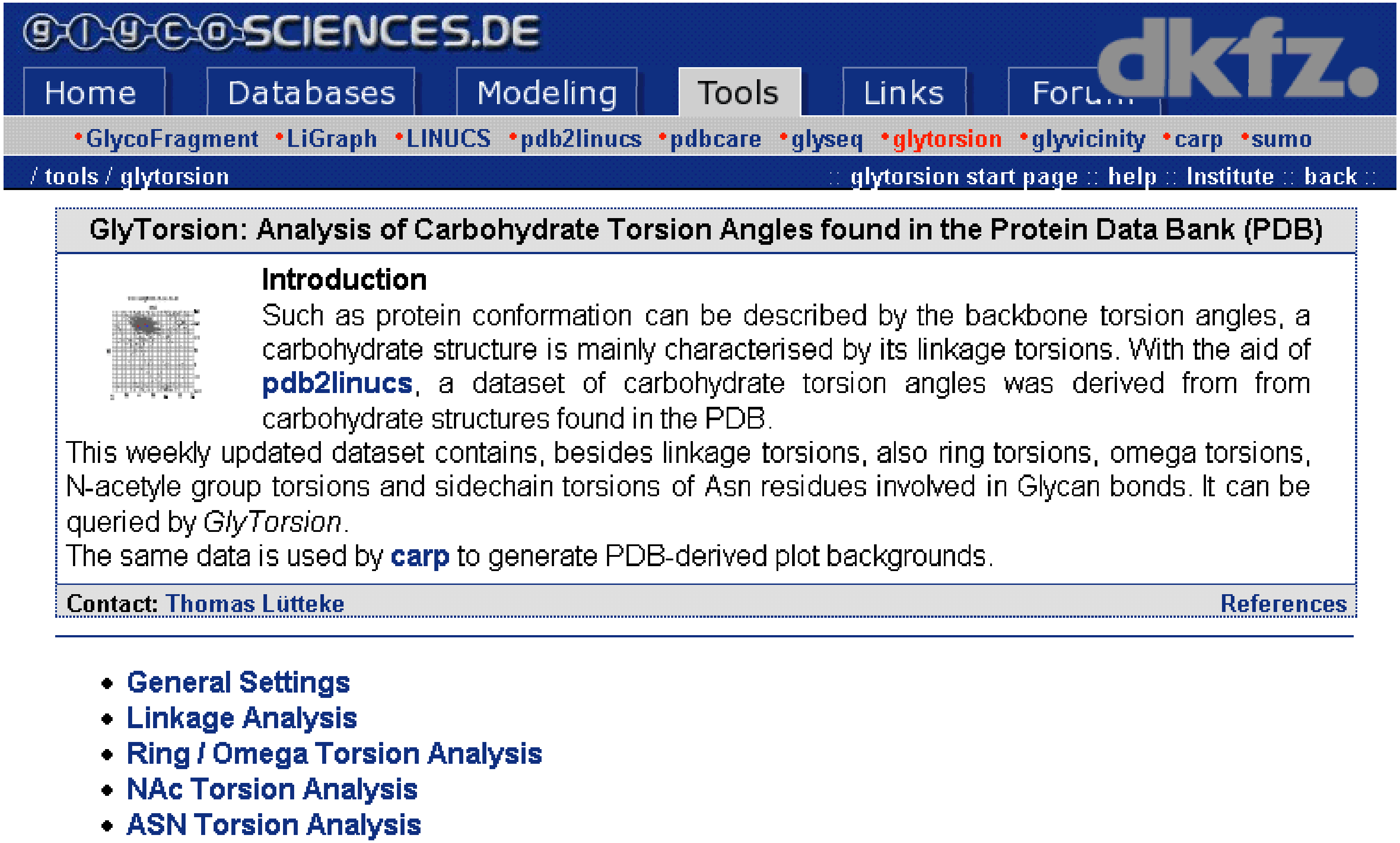
Experimental structure

->6)aDGlc^{i}(1->

->6)aDGlc^{i}(1->	99.00	72.50	74.50	70.70	71.30	66.70	
	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

GlyTorsion

http://www.glycosciences.de/tools/glytorsion/



GlyTorsion: Analysis of Carbohydrate Torsion Angles found in the Protein Data Bank (PDB)

Introduction

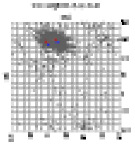
Such as protein conformation can be described by the backbone torsion angles, a carbohydrate structure is mainly characterised by its linkage torsions. With the aid of **pdb2linucs**, a dataset of carbohydrate torsion angles was derived from from carbohydrate structures found in the PDB.

This weekly updated dataset contains, besides linkage torsions, also ring torsions, omega torsions, N-acetyle group torsions and sidechain torsions of Asn residues involved in Glycan bonds. It can be queried by *GlyTorsion*.

The same data is used by **carp** to generate PDB-derived plot backgrounds.

Contact: Thomas Lütteke — References

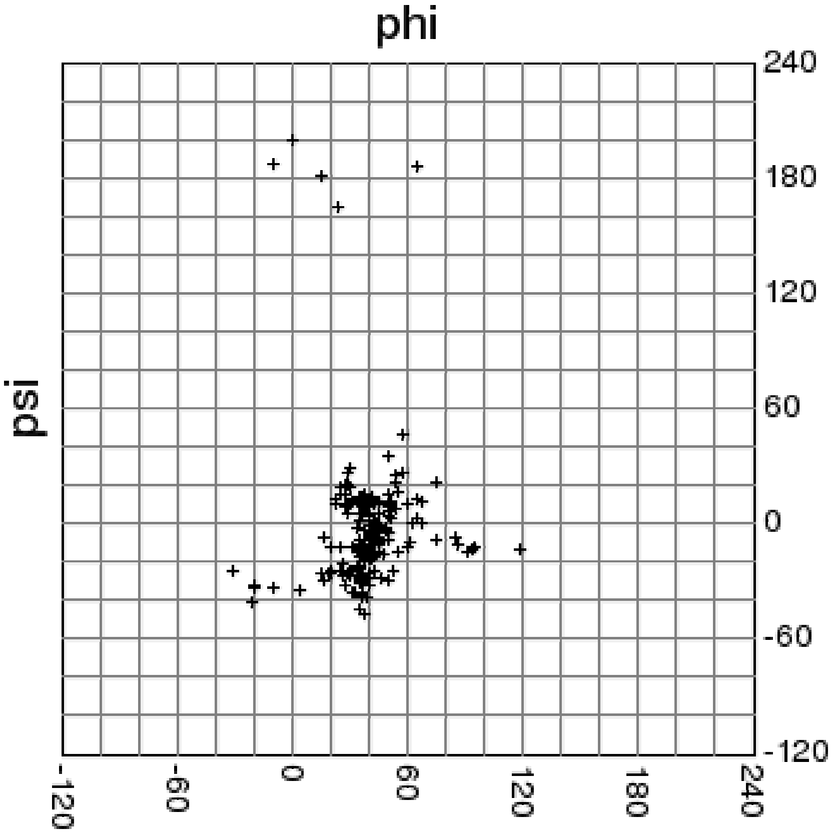
- **General Settings**
- **Linkage Analysis**
- **Ring / Omega Torsion Analysis**
- **NAc Torsion Analysis**
- **ASN Torsion Analysis**



GlyTorsion Analysis Results

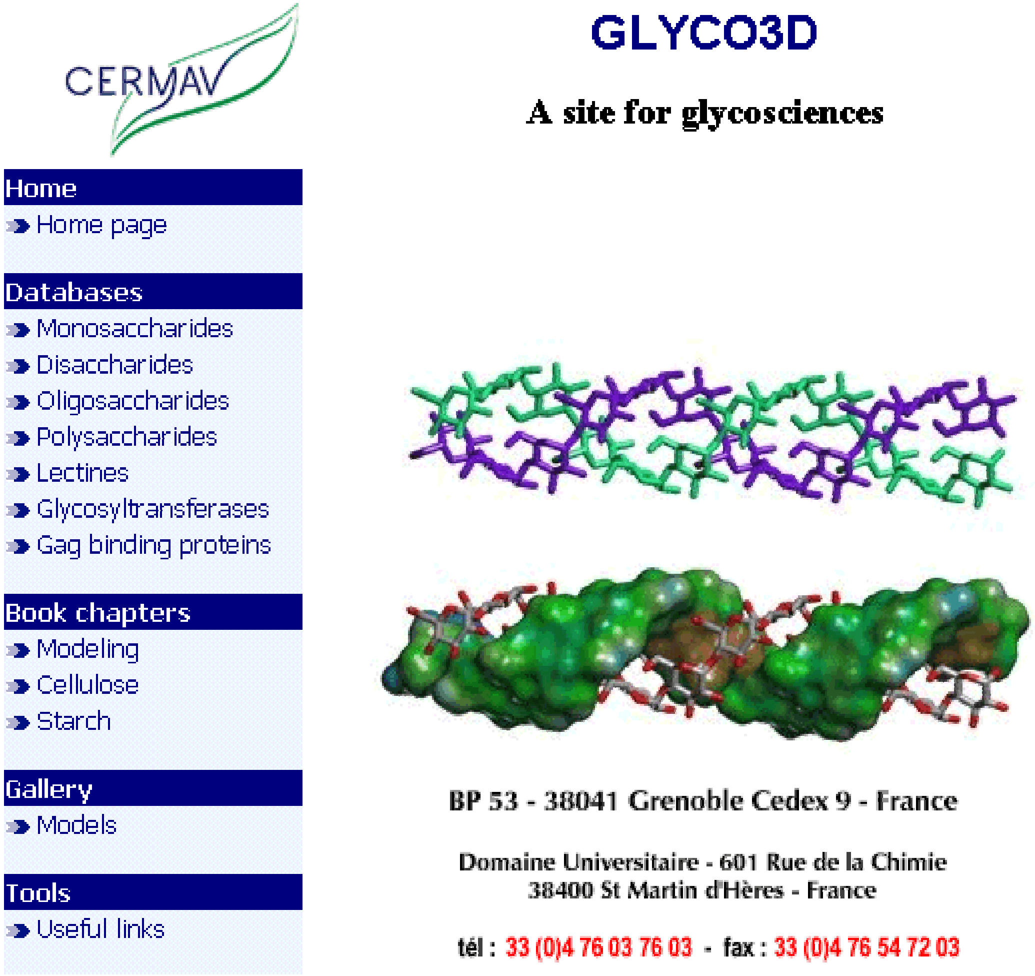
Linkage Torsion Analysis: b-D-Glcp-(1-4)-?-D-Glcp

Torsions analysed: 194



Glyco3D

http://www.cermav.cnrs.fr/glyco3d/index.php



GLYCO3D

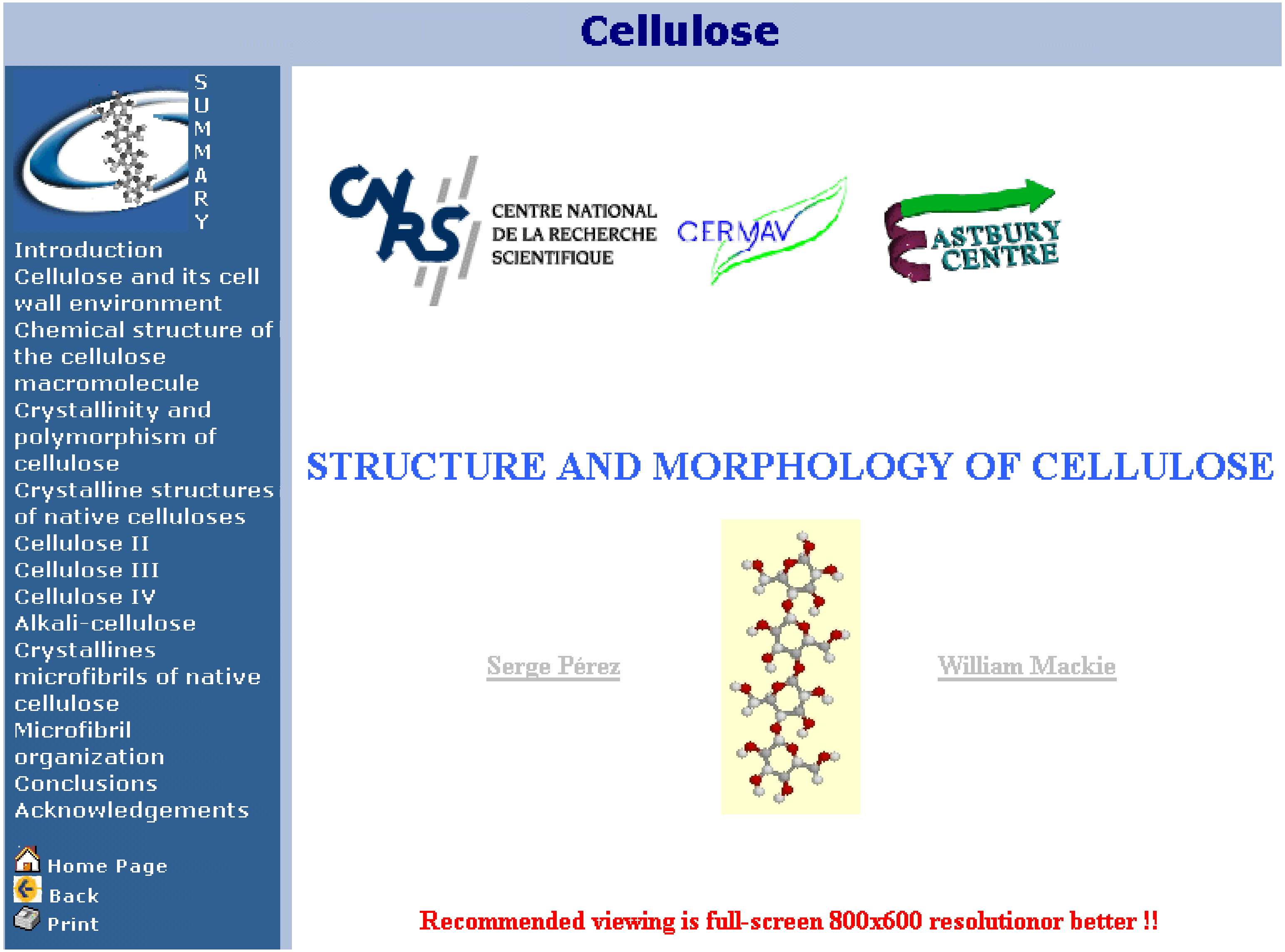
A site for glycosciences

BP 53 - 38041 Grenoble Cedex 9 - France

Domaine Universitaire - 601 Rue de la Chimie
38400 St Martin d'Hères - France

tél : 33 (0)4 76 03 76 03 - fax : 33 (0)4 76 54 72 03

Monographs @ CERMAV



Cellulose

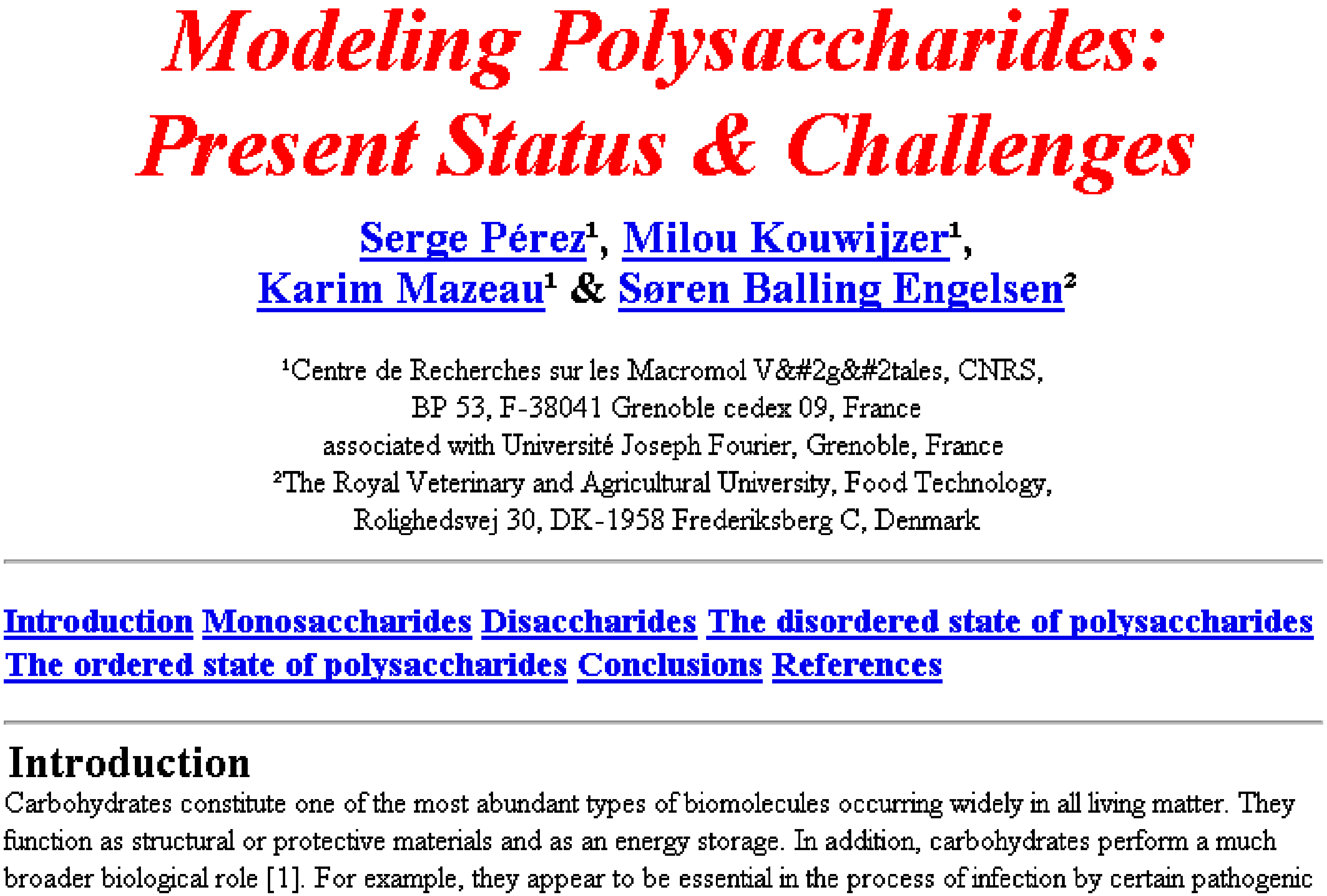
STRUCTURE AND MORPHOLOGY OF CELLULOSE

Serge Pérez

William Mackie

Recommended viewing is full-screen 800x600 resolutionor better !!

Monographs @ CERMAV



Modeling Polysaccharides: Present Status & Challenges

Serge Pérez[1], **Milou Kouwijzer**[1],
Karim Mazeau[1] & **Søren Balling Engelsen**[2]

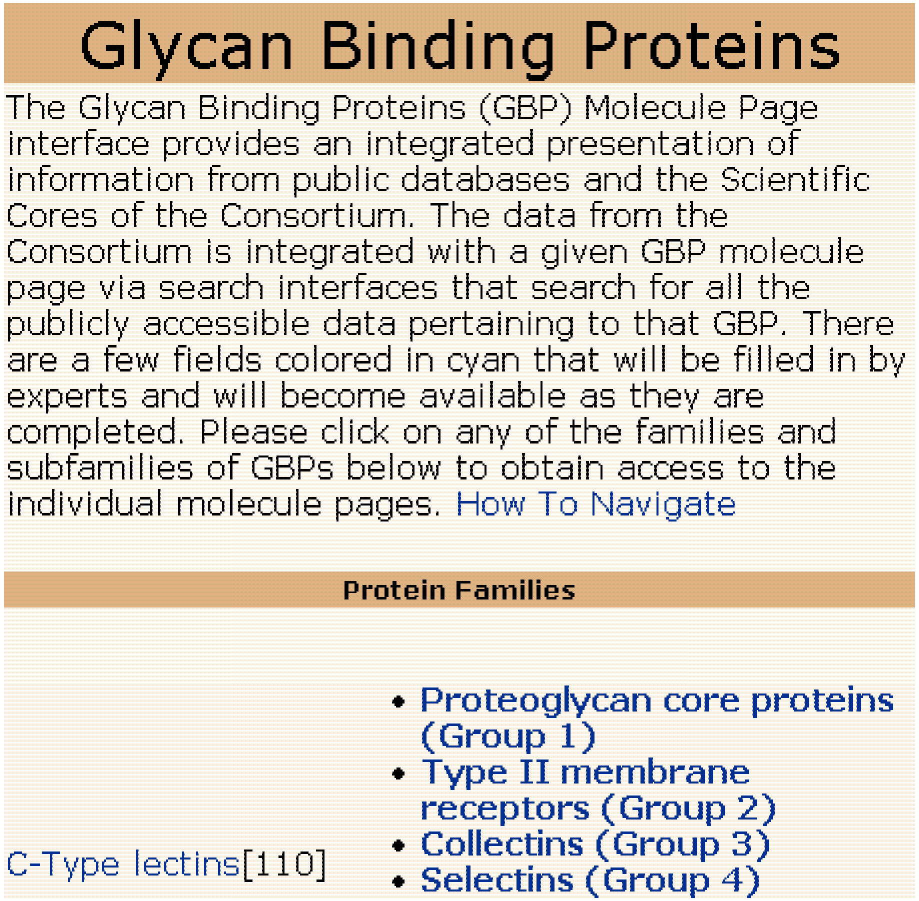
[1]Centre de Recherches sur les Macromol V&#2g&#2tales, CNRS,
BP 53, F-38041 Grenoble cedex 09, France
associated with Université Joseph Fourier, Grenoble, France
[2]The Royal Veterinary and Agricultural University, Food Technology,
Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark

Introduction

Carbohydrates constitute one of the most abundant types of biomolecules occurring widely in all living matter. They function as structural or protective materials and as an energy storage. In addition, carbohydrates perform a much broader biological role [1]. For example, they appear to be essential in the process of infection by certain pathogenic

Glycan Binding Proteins

www.functionalglycomics.org/glycomics/molecule/jsp/gbpMolecule-home.jsp



Glycan Binding Proteins

The Glycan Binding Proteins (GBP) Molecule Page interface provides an integrated presentation of information from public databases and the Scientific Cores of the Consortium. The data from the Consortium is integrated with a given GBP molecule page via search interfaces that search for all the publicly accessible data pertaining to that GBP. There are a few fields colored in cyan that will be filled in by experts and will become available as they are completed. Please click on any of the families and subfamilies of GBPs below to obtain access to the individual molecule pages. How To Navigate

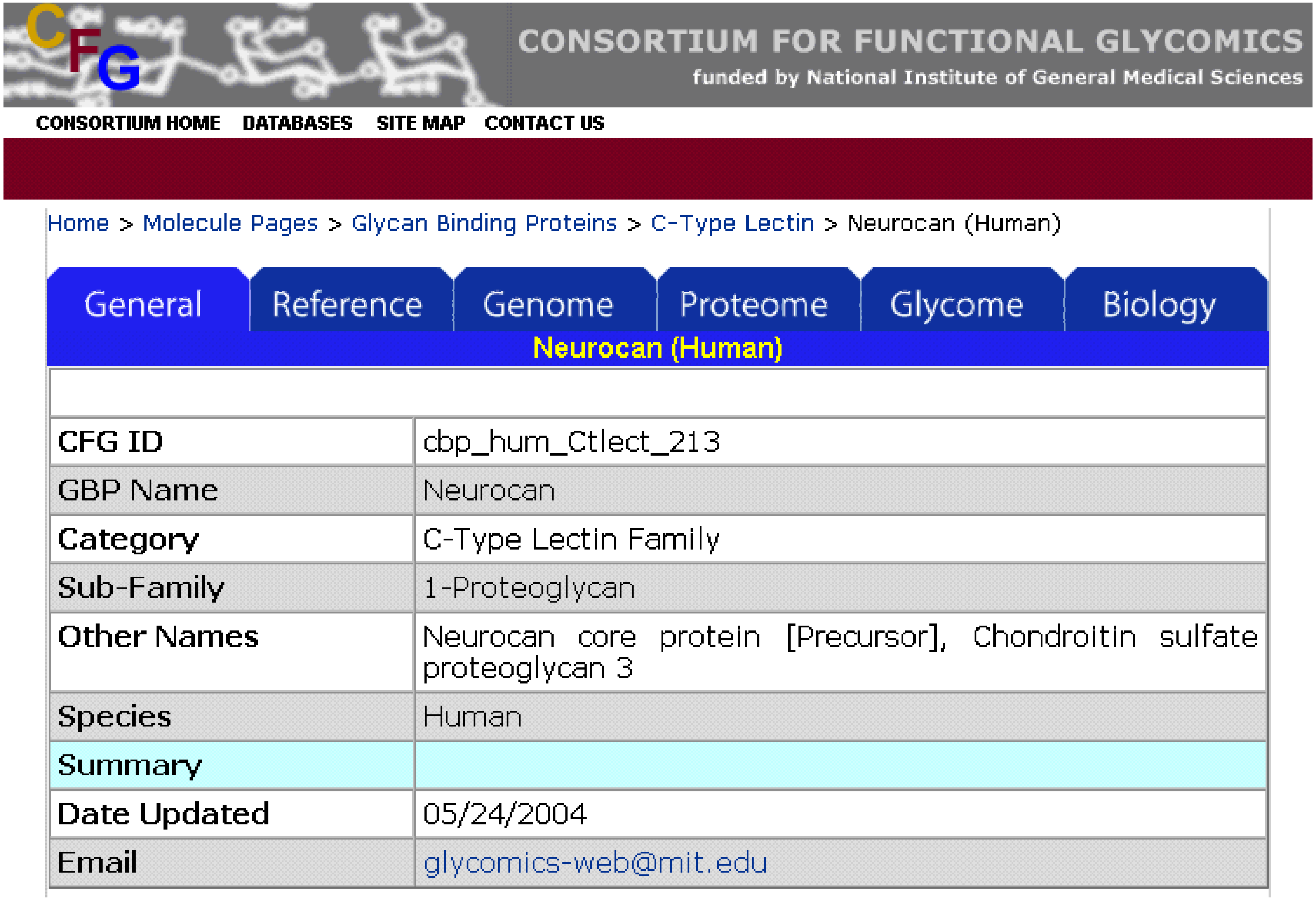
Protein Families

C-Type lectins[110]

- Proteoglycan core proteins (Group 1)
- Type II membrane receptors (Group 2)
- Collectins (Group 3)
- Selectins (Group 4)

Glycan Binding Proteins

www.functionalglycomics.org/glycomics/molecule/jsp/gbpMolecule-home.jsp



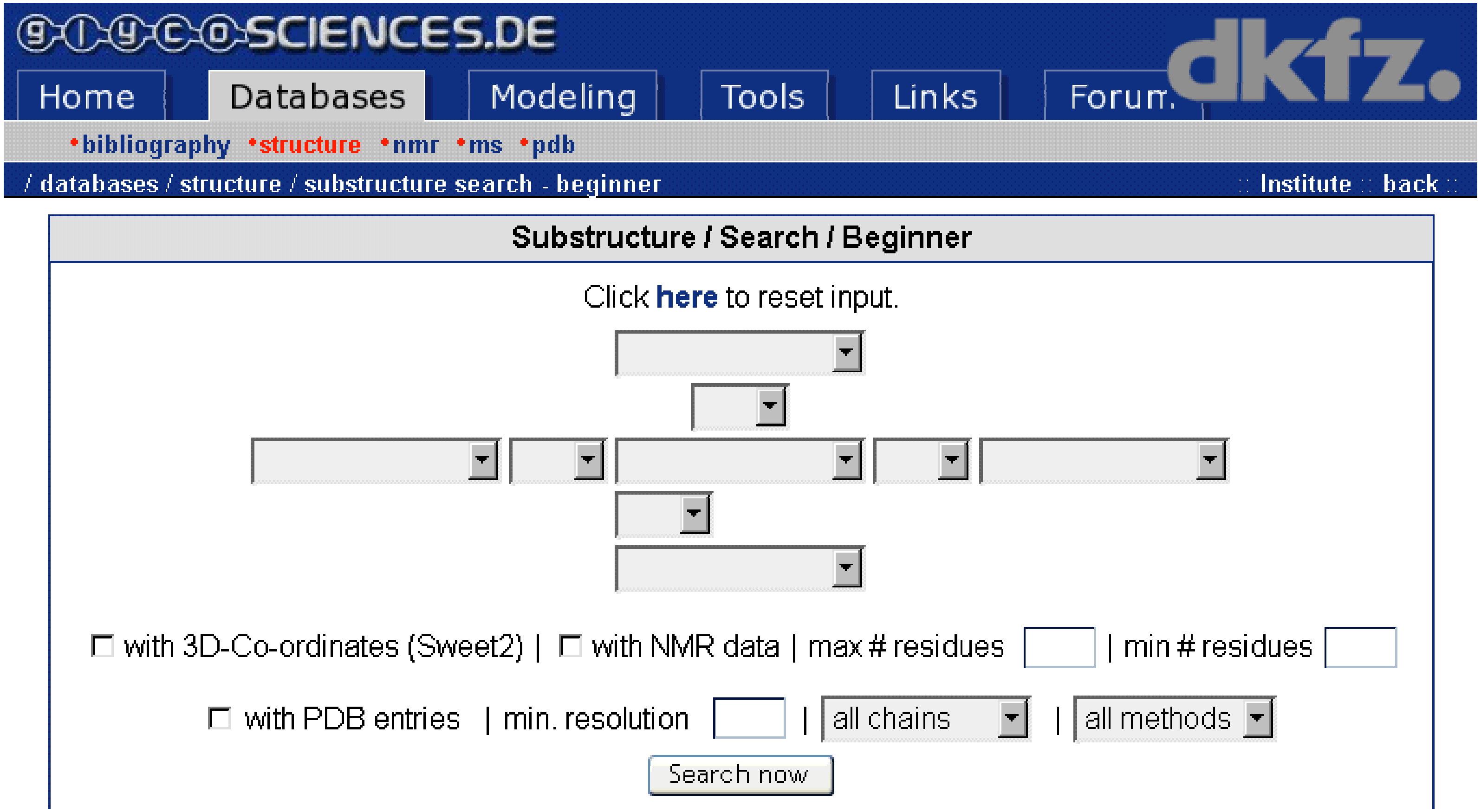
Home > Molecule Pages > Glycan Binding Proteins > C-Type Lectin > Neurocan (Human)

Neurocan (Human)

CFG ID	cbp_hum_Ctlect_213
GBP Name	Neurocan
Category	C-Type Lectin Family
Sub-Family	1-Proteoglycan
Other Names	Neurocan core protein [Precursor], Chondroitin sulfate proteoglycan 3
Species	Human
Summary	
Date Updated	05/24/2004
Email	glycomics-web@mit.edu

Glycosciences.de

http://www.glycosciences.de



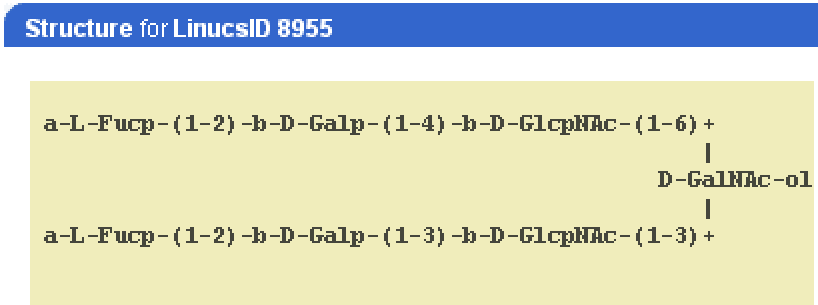
Substructure / Search / Beginner

Click **here** to reset input.

☐ with 3D-Co-ordinates (Sweet2) | ☐ with NMR data | max # residues | min # residues

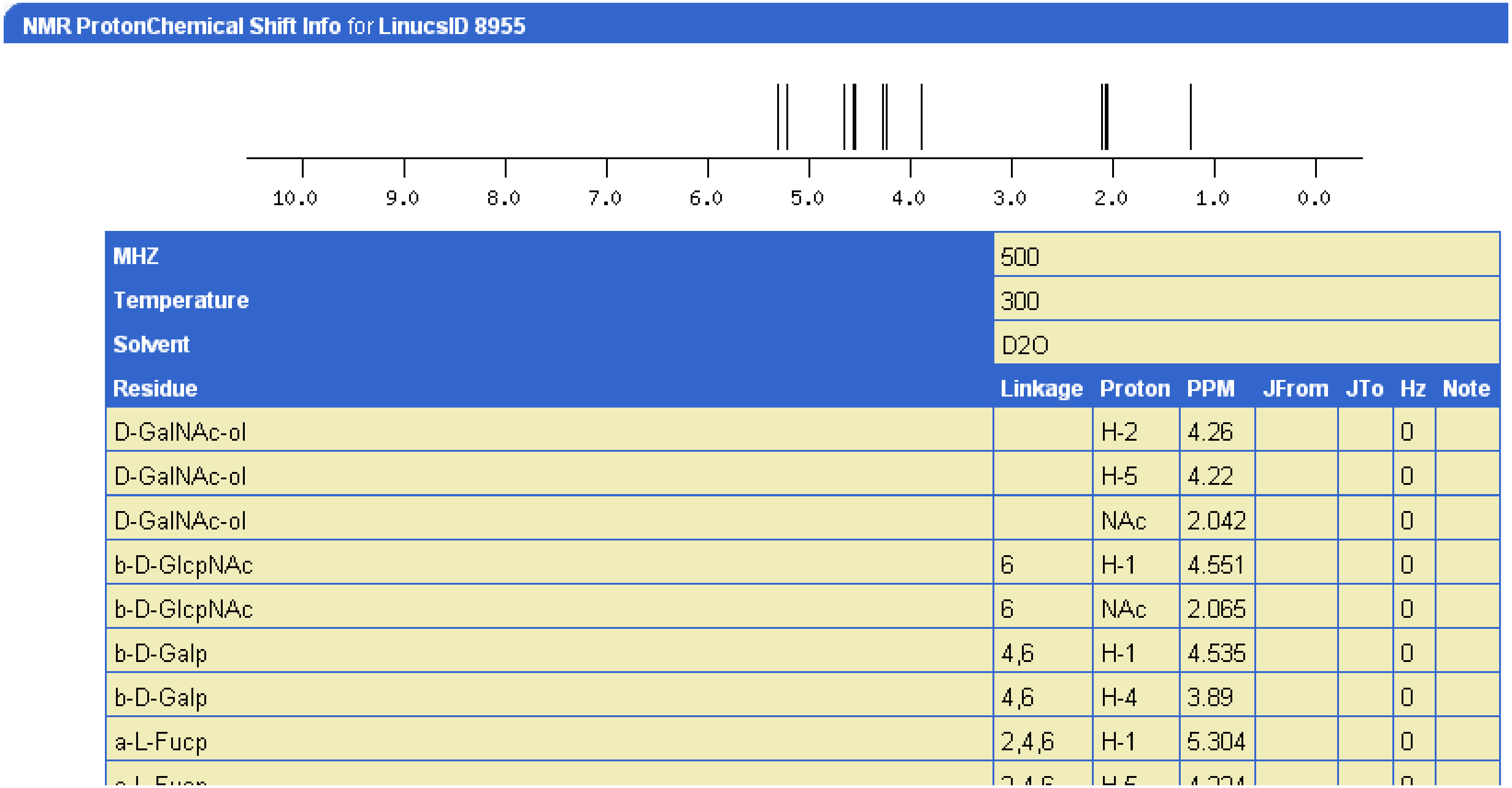
☐ with PDB entries | min. resolution | all chains | all methods

SugaBase - NMR-data



Structure for **LinucsID 8955**

```
a-L-Fucp-(1-2)-b-D-Galp-(1-4)-b-D-GlcpNAc-(1-6)+
                                                |
                                           D-GalNAc-ol
                                                |
a-L-Fucp-(1-2)-b-D-Galp-(1-3)-b-D-GlcpNAc-(1-3)+
```

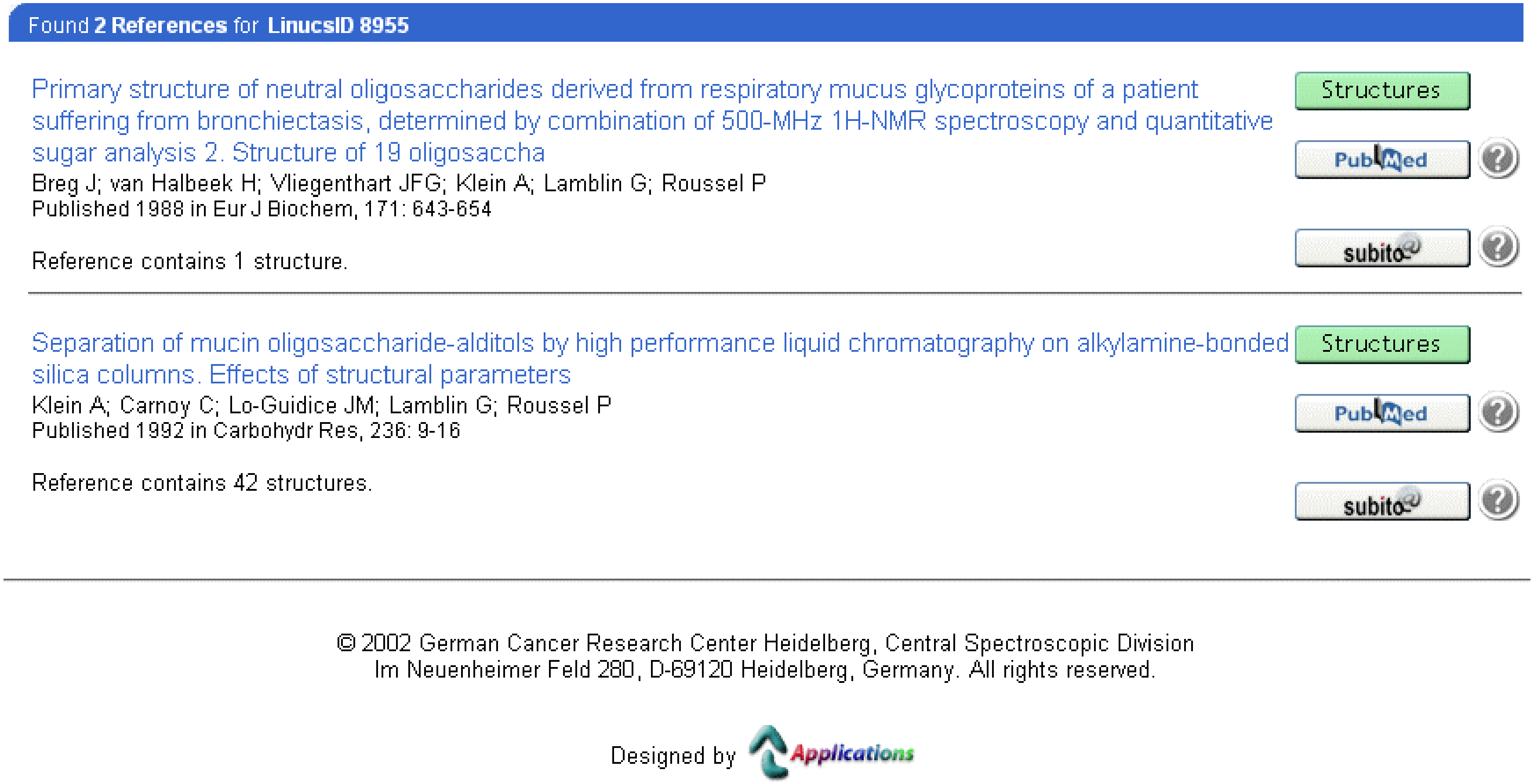


NMR ProtonChemical Shift Info for **LinucsID 8955**

MHZ	500
Temperature	300
Solvent	D2O

Residue	Linkage	Proton	PPM	JFrom	JTo	Hz	Note
D-GalNAc-ol		H-2	4.26			0	
D-GalNAc-ol		H-5	4.22			0	
D-GalNAc-ol		NAc	2.042			0	
b-D-GlcpNAc	6	H-1	4.551			0	
b-D-GlcpNAc	6	NAc	2.065			0	
b-D-Galp	4,6	H-1	4.535			0	
b-D-Galp	4,6	H-4	3.89			0	
a-L-Fucp	2,4,6	H-1	5.304			0	
a-L-Fucp	2,4,6	H-5	[illegible]			0	

SweetDB – Bibliographic information (from CarbBank)



Found **2 References** for **LinucsID 8955**

Primary structure of neutral oligosaccharides derived from respiratory mucus glycoproteins of a patient suffering from bronchiectasis, determined by combination of 500-MHz 1H-NMR spectroscopy and quantitative sugar analysis 2. Structure of 19 oligosaccha
Breg J; van Halbeek H; Vliegenthart JFG; Klein A; Lamblin G; Roussel P
Published 1988 in Eur J Biochem, 171: 643-654

Reference contains 1 structure.

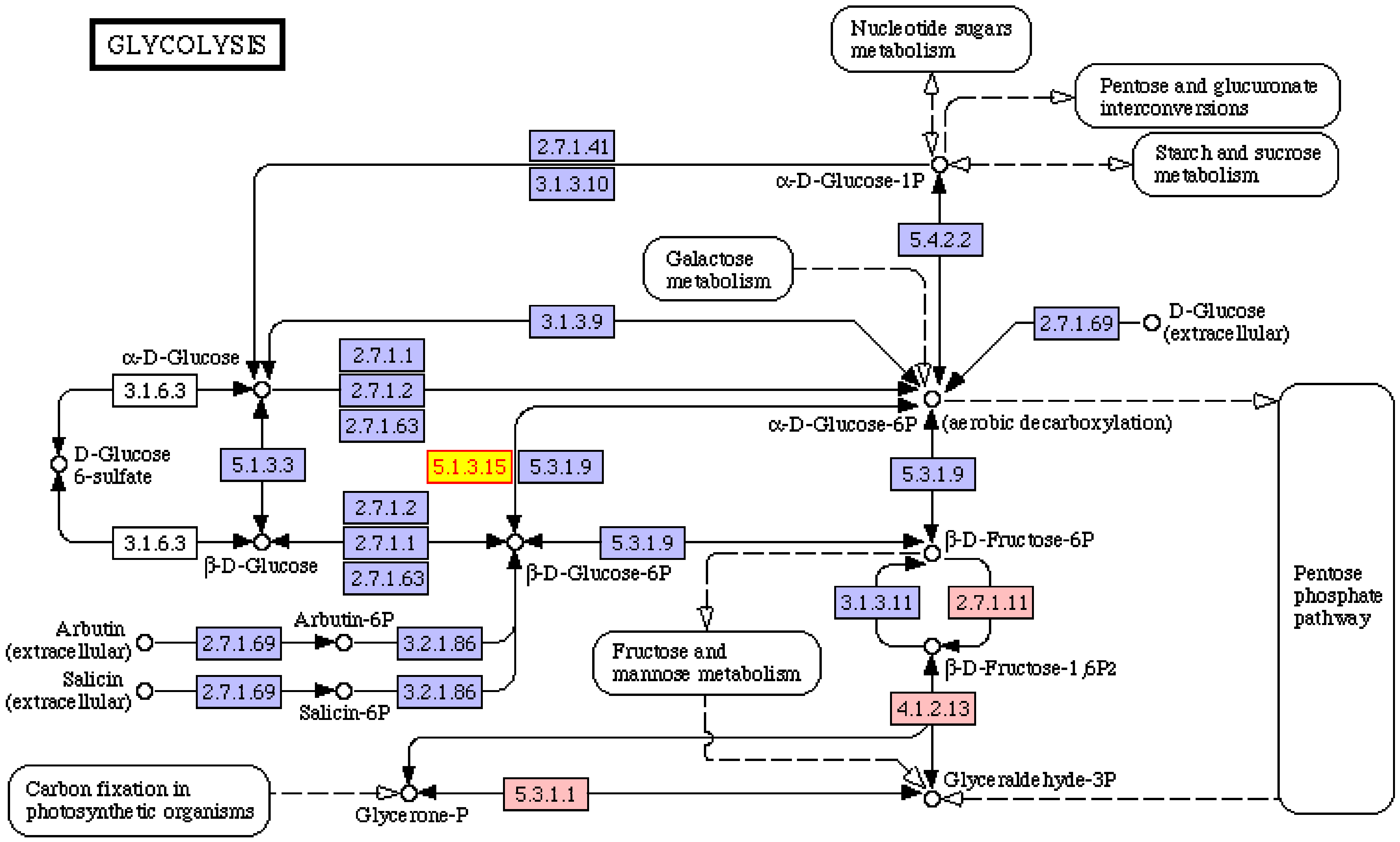
Separation of mucin oligosaccharide-alditols by high performance liquid chromatography on alkylamine-bonded silica columns. Effects of structural parameters
Klein A; Carnoy C; Lo-Guidice JM; Lamblin G; Roussel P
Published 1992 in Carbohydr Res, 236: 9-16

Reference contains 42 structures.

Designed by Applications

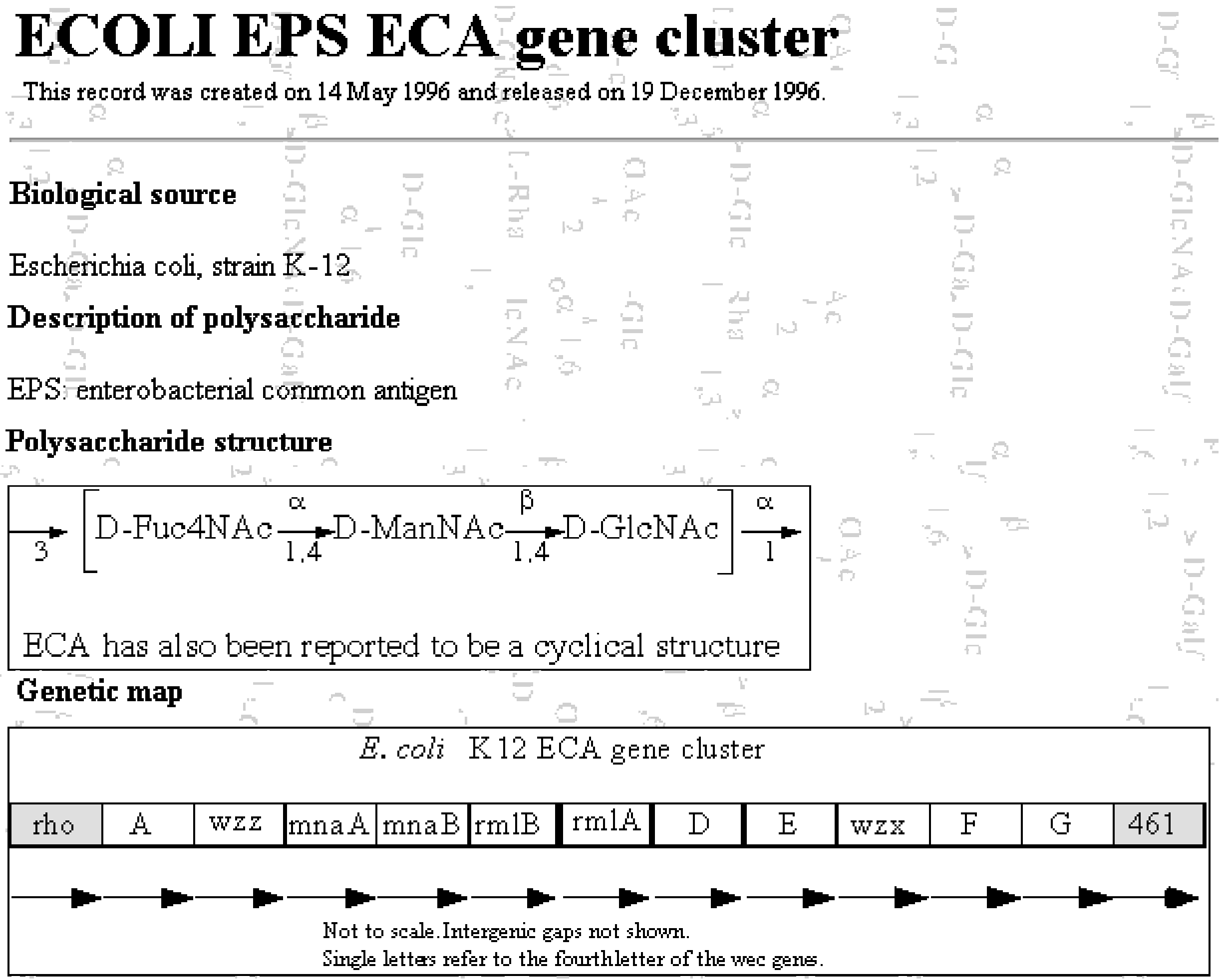
KEGG - Kyoto Encyclopedia of Genes and Genomes

http://www.genome.jp/kegg/



Bacterial Polysaccharide Gene Database

http://www.microbio.usyd.edu.au/BPGD/default.htm



ECOLI EPS ECA gene cluster

This record was created on 14 May 1996 and released on 19 December 1996.

Biological source

Escherichia coli, strain K-12

Description of polysaccharide

EPS: enterobacterial common antigen

Polysaccharide structure

$$\xrightarrow[3]{}\left[\text{D-Fuc4NAc}\xrightarrow[1,4]{\alpha}\text{D-ManNAc}\xrightarrow[1,4]{\beta}\text{D-GlcNAc}\right]\xrightarrow[1]{\alpha}$$

ECA has also been reported to be a cyclical structure

Genetic map

E. coli K12 ECA gene cluster

rho	A	wzz	mnaA	mnaB	rmlB	rmlA	D	E	wzx	F	G	461

Not to scale. Intergenic gaps not shown.
Single letters refer to the fourth letter of the wec genes.

IUPAC nomenclature

http://www.chem.qmul.ac.uk/iupac

INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY

Recommendations on Organic & Biochemical Nomenclature, Symbols & Terminology etc.

If the table below is unusable check this button for an alternative presentation

http://www.chem.qmul.ac.uk/iupac/
World Wide Web material prepared by G. P. Moss
Department of Chemistry, Queen Mary University of London,
Mile End Road, London, E1 4NS, UK
g.p.moss@qmul.ac.uk

To search the database click here.

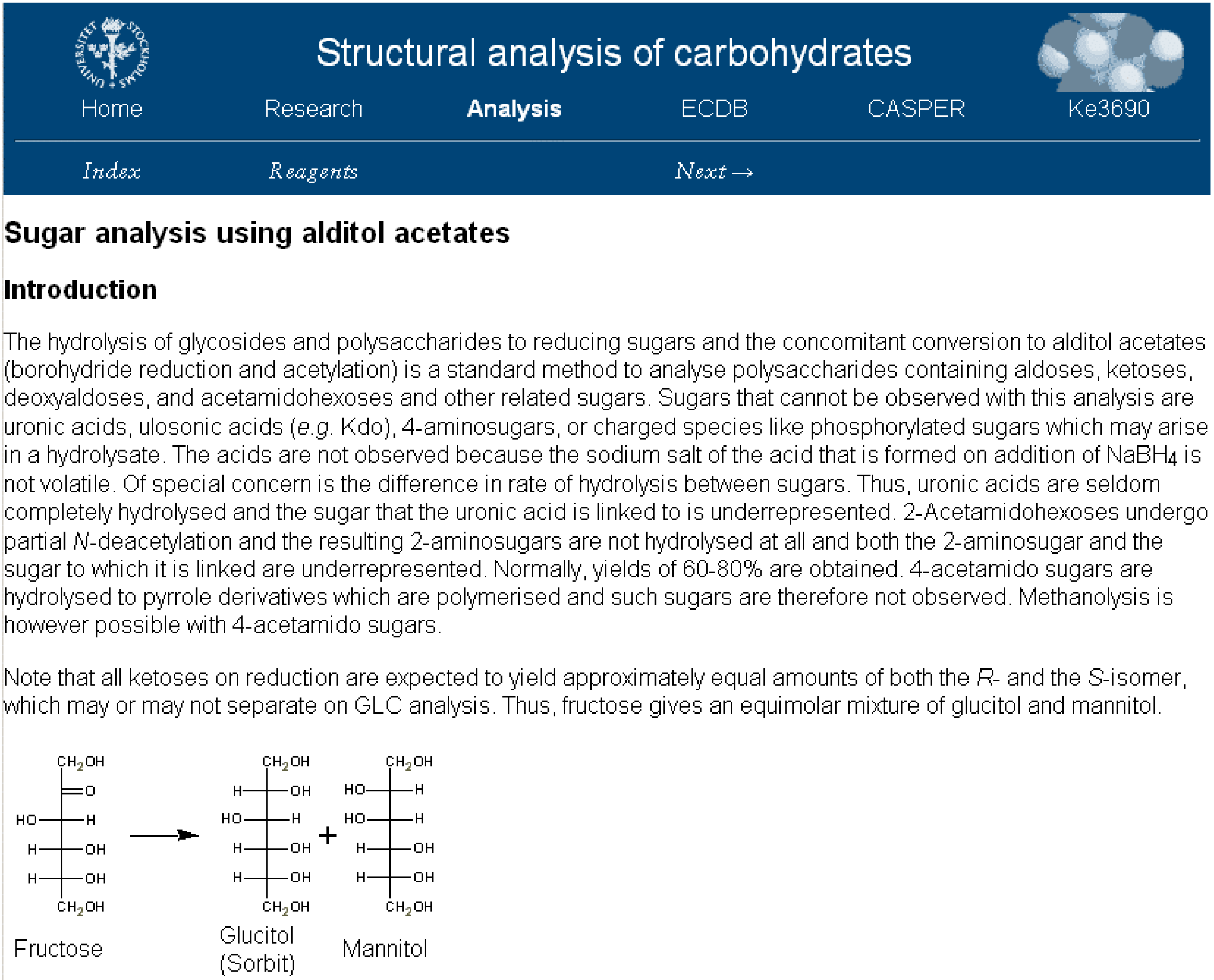
What's Here and What's New	Changes to Published Documents for World Wide Web Presentation	Main IUPAC Home Page and mirror sites in USA, Germany, Japan, Korea, South Africa, UK

Full text of IUPAC Recommendations

Glossary of Organic Class Names	**Glossary of Terms in Physical Organic Chemistry**	**Basic Terminology of Stereochemistry**
Hantzsch-Widman Nomenclature for Heteromonocyclic Rings	**Non-standard Valence States (Lambda Convention)**	**Cyclic compounds with contiguous formal double bonds (delta Convention)**
Fused Ring Nomenclature	**Phane Nomenclature**	**Phane Nomenclature Part II** New
Glossary of Terms in Bioinorganic Chemistry	**Glossary of Medicinal Chemistry Terms**	**Nomenclature of Isotopically Modified Compounds (Section H)**

Structural analysis

http://www.casper.organ.su.se/sop/



Structural analysis of carbohydrates

Sugar analysis using alditol acetates

Introduction

The hydrolysis of glycosides and polysaccharides to reducing sugars and the concomitant conversion to alditol acetates (borohydride reduction and acetylation) is a standard method to analyse polysaccharides containing aldoses, ketoses, deoxyaldoses, and acetamidohexoses and other related sugars. Sugars that cannot be observed with this analysis are uronic acids, ulosonic acids (*e.g.* Kdo), 4-aminosugars, or charged species like phosphorylated sugars which may arise in a hydrolysate. The acids are not observed because the sodium salt of the acid that is formed on addition of $NaBH_4$ is not volatile. Of special concern is the difference in rate of hydrolysis between sugars. Thus, uronic acids are seldom completely hydrolysed and the sugar that the uronic acid is linked to is underrepresented. 2-Acetamidohexoses undergo partial *N*-deacetylation and the resulting 2-aminosugars are not hydrolysed at all and both the 2-aminosugar and the sugar to which it is linked are underrepresented. Normally, yields of 60-80% are obtained. 4-acetamido sugars are hydrolysed to pyrrole derivatives which are polymerised and such sugars are therefore not observed. Methanolysis is however possible with 4-acetamido sugars.

Note that all ketoses on reduction are expected to yield approximately equal amounts of both the *R*- and the *S*-isomer, which may or may not separate on GLC analysis. Thus, fructose gives an equimolar mixture of glucitol and mannitol.

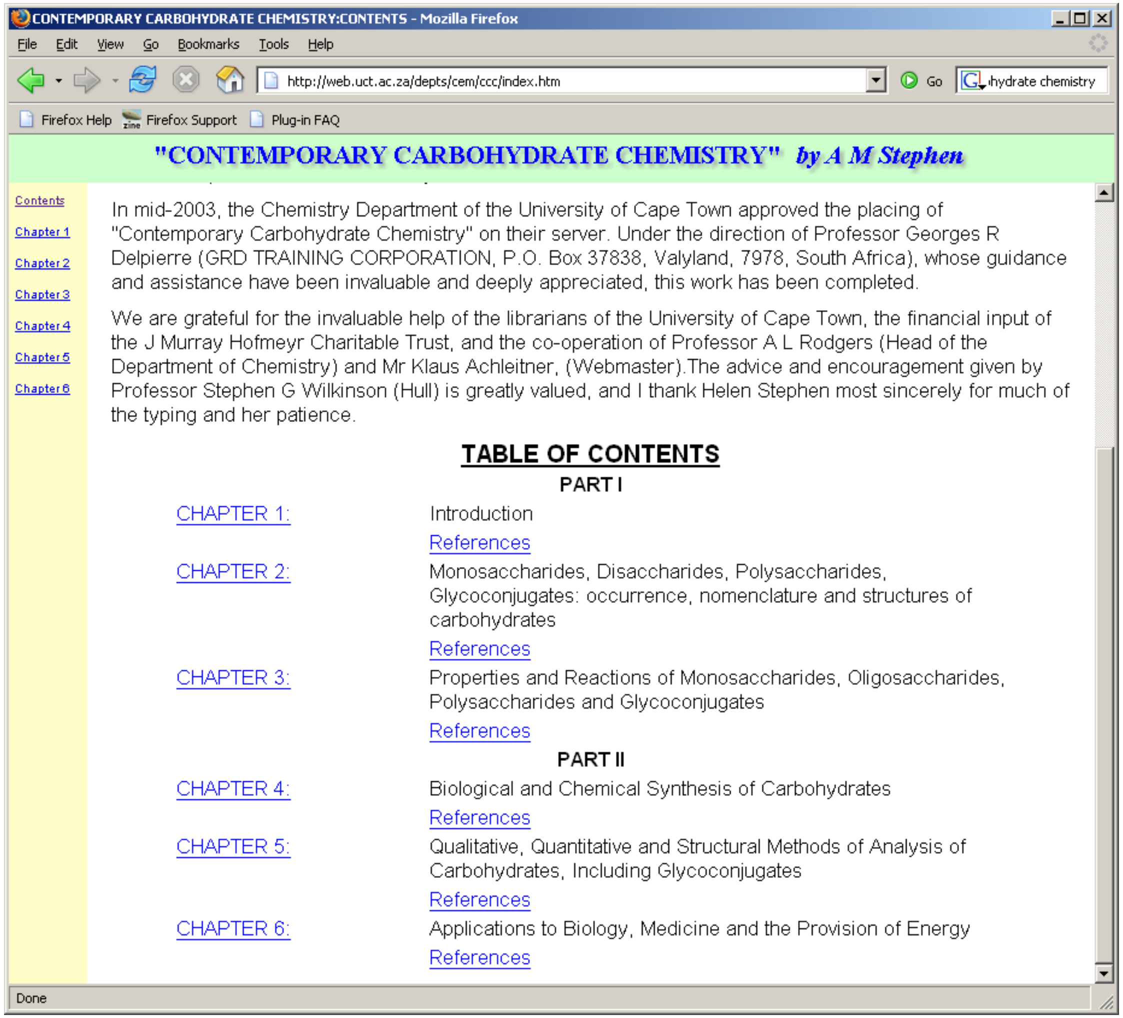
Fructose

Glucitol (Sorbit)

Mannitol

Contemporary Carbohydrate Chemistry

http://web.uct.ac.za/depts/cem/ccc/index.htm



"CONTEMPORARY CARBOHYDRATE CHEMISTRY" *by A M Stephen*

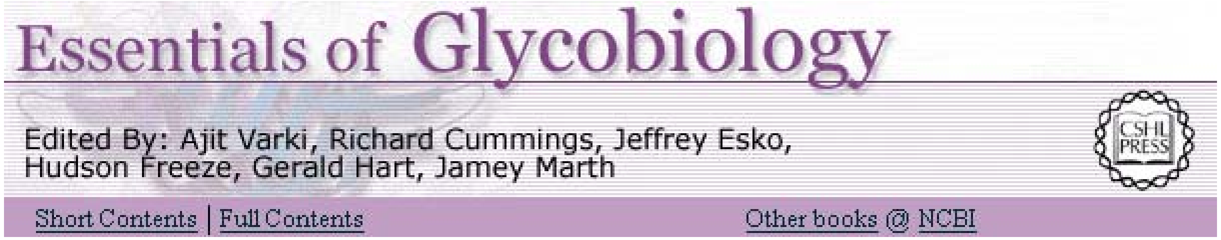
In mid-2003, the Chemistry Department of the University of Cape Town approved the placing of "Contemporary Carbohydrate Chemistry" on their server. Under the direction of Professor Georges R Delpierre (GRD TRAINING CORPORATION, P.O. Box 37838, Valyland, 7978, South Africa), whose guidance and assistance have been invaluable and deeply appreciated, this work has been completed.

We are grateful for the invaluable help of the librarians of the University of Cape Town, the financial input of the J Murray Hofmeyr Charitable Trust, and the co-operation of Professor A L Rodgers (Head of the Department of Chemistry) and Mr Klaus Achleitner, (Webmaster).The advice and encouragement given by Professor Stephen G Wilkinson (Hull) is greatly valued, and I thank Helen Stephen most sincerely for much of the typing and her patience.

TABLE OF CONTENTS

PART I

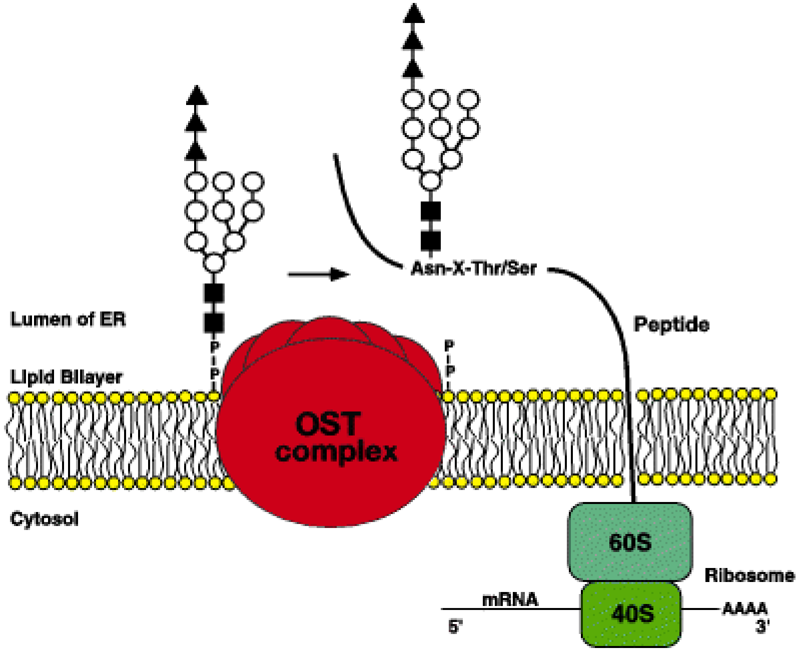
PART II



Essentials of Glycobiology

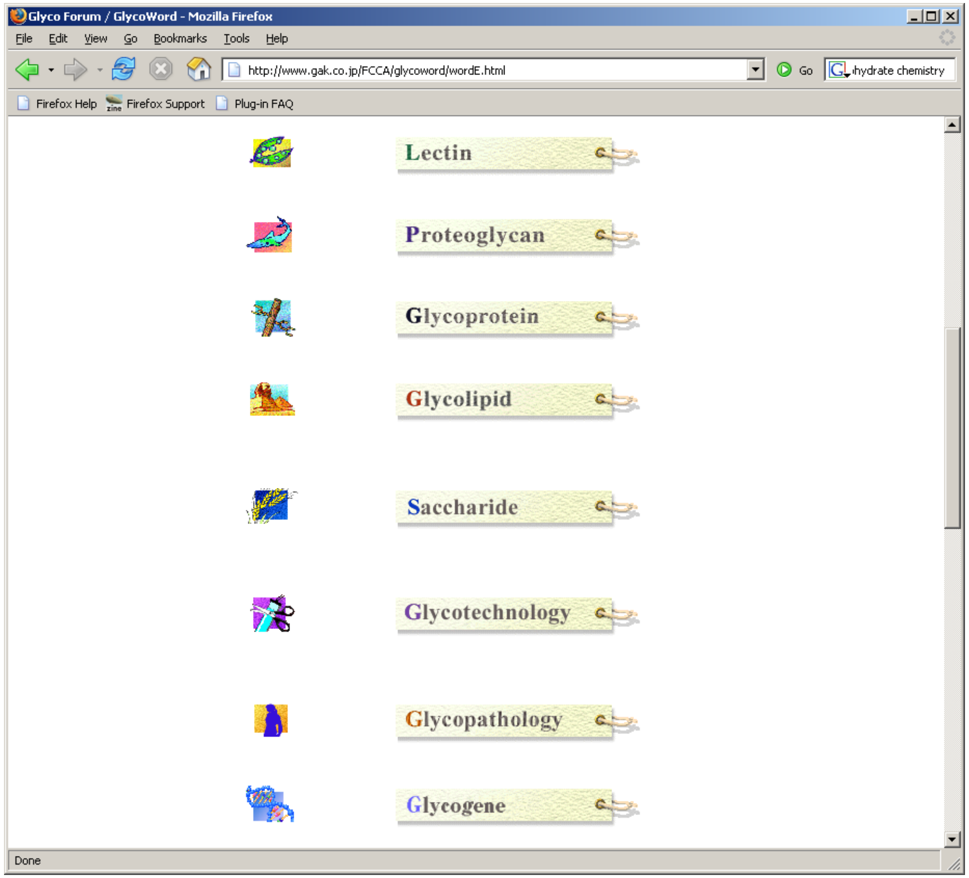
Edited By: Ajit Varki, Richard Cummings, Jeffrey Esko,
Hudson Freeze, Gerald Hart, Jamey Marth

http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowTOC&rid=glyco.TOC&depth=1



GlycoWord

http://www.gak.co.jp/FCCA/glycoword/wordE.html



Lectin

Proteoglycan

Glycoprotein

Glycolipid

Saccharide

Glycotechnology

Glycopathology

Glycogene