GlycoCD: a repository for carbohydrate-related CD antigens

Sonu Kumar[1], Thomas Lütteke[2] and Reinhard Schwartz-Albiez[1],*

[1]D015, Translational Immunology, German Cancer Research Center, INF 280, D-69120 Heidelberg and [2]Institute of Veterinary Physiology and Biochemistry, Justus-Liebig-University Giessen, D-35392 Giessen, Germany

ABSTRACT

Summary: The open access comprehensive GlycoCD database application is for representation and retrieval of carbohydrate-related clusters of differentiation (CDs). The main objective of this database platform is to provide information about interactions of carbohydrate moieties with proteins that are important for identification of specific cell surface molecule with a focus on the integration of data from carbohydrate microarray databases. GlycoCD database comprises two sections: the carbohydrate recognition CD and glycan CD. It allows easy access through a user-friendly web interface to all carbohydrate-defined CDs and those that interact with carbohydrates along with other relevant information.

Availability: The database is freely available at http://glycosciences.de/glycocd/index.php

Contact: r.s-albiez@dkfz.de

1 INTRODUCTION

The clusters of differentiation (CD) antigens are cell surface-expressed antigens defined by monoclonal antibodies. The great value of the CD nomenclature, validated and approved as part of the International Union of Immunological Societies nomenclature (Zola *et al.*, 2003), is to provide a unified system of classified antigens useful both for biomedical research and clinical practice. The CD antibodies have been categorized during the past 28 years in Human Leucocyte Differentiation Antigens (HLDA) workshops (Heddy Zola, 2007; Mason *et al.*, 2002). The CD antigens are expressed on human hematopoietic cells (T and B lymphocytes, natural killer cells, myeloid cells, erythrocytes and platelets), specific types of endothelial and epithelial cells. The information for CD antigens expands from cellular and tissue expression to biochemical structure, genetic background and to its functional impact in a given cellular environment. The interactions of the immune system are mediated by cell surface glycoproteins and glycolipids (Schwartz-Albiez, 2009). Most of the CD antibodies recognize the protein moiety of surface, membrane integral macromolecules, but some of them detect oligosaccharide structures that are linked to either glycoproteins or glycosphingolipids.

During the HLDA conference 2000, a separate workshop on carbohydrate and lectin CDs was established considering the increasing interest and advance in research in unraveling cellular

protein–glycan interactions (Schwartz-Albiez, 2002). The objective was to initiate a more comprehensive compilation of monoclonal antibody-defined carbohydrate structures and lectins differentially expressed on human leucocytes and to provide an appropriate classification within the CD nomenclature. Some carbohydrate CD antibodies were re-organized into subgroups of CDs because of their reactivity with variations of certain oligosaccharide structures. For example, CD15 (blood group Lewis x, Le^x) received additional subgroups of CD15s (sialyl Lewis x), CD15u (3sulfo Lewis x) and CD15su (6 sulfo-sialyl Lewis x). CD60 was subdivided into CD60a (the oligosaccharide structure of the disialo-ganglioside GD3 and related glycans), CD60b (9-*O*-acetylated sialic acid GD3) and CD60c (7-*O*-acetylated sialic acid GD3). CD75 comprises lactosaminyl sequences into CD75 (non-sialylated lactosamines) and CD75s (sialylated lactosamines) (Mandal, et al. 2012). The oligosaccharides of the CD75s subgroup have a different functional impact as modulators of adhesion to certain lectins such as Siglecs (Cao *et al.*, 2008). Due to the high complexity of oligosaccharide structures, antibodies of a particular carbohydrate CD may bind to variations of the same oligosaccharide type and consequently display different cellular staining patterns. For instance, this may be due to variations in the linkage of the terminal sialic acid (α2-6 versus α2-3), to oligosaccharide chains which extend the glycan epitope (glycotope) recognized by the respective antibody or by varying expression of the glycotope on different glycosphingolipids and/or glycoproteins on a specific cell type. In general, most carbohydrate CDs define histo-blood group antigens, which are characterized either by their different fucosylation or sialylation.

This GlycoCD database now endeavours to provide in-depth information about glycan CD antigens and those CD-defined surface receptors (lectins), which bind to glycan structures.

2 METHODS

GlycoCD is a web-based, platform-independent, manually curated database. All data were collected by comprehensively searching literature, online resources and books. It has been designed as a relational database and currently runs on an Apache server. This database has been implemented using the open source MySQL database (http://www.mysql.com) that serves as the back-end for storing all the carbohydrate-related CD information along with related publications. PHP: Hypertext Preprocessor was used to fetch data from the database, while both PHP and HTML were used to make the web interface. The schematic overview is provided in Fig. 1. The database is presently hosted on a server at glycoscience.de—a glycoinformatics web portal—and it will be manually updated and maintained.
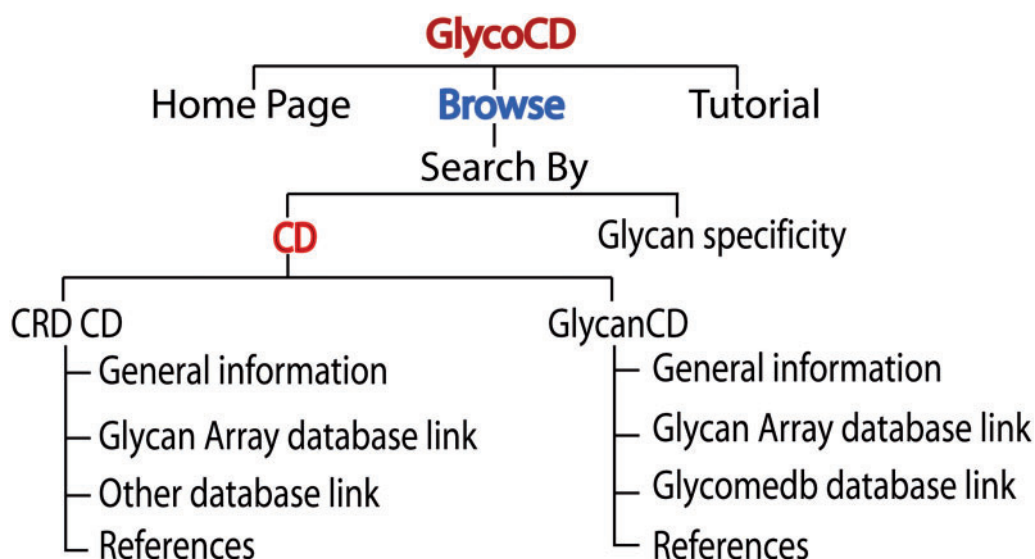
*To whom correspondence should be addressed.

Fig. 1. The schematic overview of GlycoCD database

3 RESULTS AND DISCUSSION

GlycoCD is a collection of lectins and carbohydrates defined by CD antibodies that are differentially expressed primarily on human leukocytes and also on certain endothelial and epithelial cells. The database comprises information on 63 CD antigens (out of 363 of total CD antigens as defined in the HLDA 9 workshop). We have categorized these data within GlycoCD repository into two sections: carbohydrate recognition CD antigens and glycan CDs. For the convenience of the user, glycan-related CDs can be searched either by the respective CD number or the glycan-binding specificity. There is a tutorial section describing the workflow in a simple manner with relevant information about the respective CD. There are online resources available which touch selective fields of CD: http://www.immunologylink.com/cdantigen.html summarizes in short form information about CD antigens/antibodies; http://pathologyoutlines.com/cdmarkers.html provides information about the immunohistology of selected CDs and two databases: SugarBindDB (http://sugarbind.expasy.org/sugarbind/) describes carbohydrate-binding proteins in the field of microbiology and GlycoEpitope (http://www.glyco.is.ritsumei.ac.jp/epitope/) for carbohydrate epitopes in general.

Our new database now provides user-friendly comprehensive information about human CD CRDs and glycan epitopes (glycotopes) and additionally links to other resources.

3.1 Carbohydrate recognition CD

The carbohydrate recognition CD section contains: (1) lectin CDs that are classified within distinct lectin families such as selectins or siglecs. (2) CD antigens other than lectin families that have a carbohydrate recognition domain (CRD) defined by sequence homologies. For some of these, a functional influence of the CRD has not yet been clarified. (3) CD proteins that are related to carbohydrate binding, in the way that interactions with glycan moieties of other antigens influence their functional activities. There are 44-carbohydrate recognition CDs reported in this section.

Users can search these data from the 'Browse' menu. Each set of data in this section contains information about synonyms of this particular CD protein, its structure (also, three dimensional information wherever available), amino acid and nucleic acid sequence, cellular and tissue expression, function (and clinical application), references, carbohydrate and other possible ligands and further links to glycan array data. In glycan array information, it has been linked to the Consortium of Functional Glycomics database (Raman *et al.*, 2006) and the Glyaffinity database (Frank and Schloissnig, 2010). There is also a small feature for finding interactions of non-GlycoCD/molecules with GlycoCD.

3.2 Glycan CD

Here, glycan CDs that define oligosaccharide moieties linked to glycoproteins and/or glycosphingolipids are included. There are 19 CDs reported in this section. Most of the CDs of this section interact with several molecules. We have given information about ligands if available or else linked it to the glycan array database Glyaffinity for relative affinity data.

Users can search these data from the 'Browse' menu. Each set of data in this section contains information about Glycotope (the carbohydrate epitope recognized by the respective CD-antibody), synonyms of the particular CD, its molecular structure, cellular and tissue expression, function (and clinical application), ligands and useful references, further links to GlycomeDB structure database (Ranzinger *et al.*, 2008) with their ids, Glycosciences.de structural database and to glycan array data.

workshops and to David Y.Mason (deceased 2008) who contributed significantly to the establishment of CDs and the organization of the conferences. We remember both not only as excellent leaders in this field but also as warm hearted persons with a great sense of humour.

Conflict of Interest: none declared.

REFERENCES

Cao,Y. *et al.* (2008) Expression of CD175 (Tn), CD175s (sialosyl-Tn) and CD176 (Thomsen–Friedenreich antigen) on malignant human hematopoietic cells. *Int. J. Cancer*, **123**, 89–99.

Frank,M. and Schloissnig,S. (2010) Bioinformatics and molecular modeling in glycobiology. *Cell Mol. Life Sci.*, **67**, 2749–2772.

Zola,H. *et al.* (2007) *Leukocyte and Stromal Cell Molecules: The CD Markers*. John Wiley & Sons, Hoboken, New Jersey.

Mandal,C. *et al.* (2012) Functions and biosynthesis of O-acetylated sialic acids. *Top. Curr. Chem.*, [Epub ahead of print].

Mason,D. *et al.* (2002) CD antigens 2002. *Blood*, **99**, 3877–3880.

Raman,R. *et al.* (2006) Advancing glycomics: implementation strategies at the consortium for functional glycomics. *Glycobiology*, **16**, 82R–90R.

Ranzinger,R. *et al.* (2008) GlycomeDB—integration of open-access carbohydrate structure databases. *BMC Bioinformatics*, **9**, 384.

Schwartz-Albiez,R. (2002) Carbohydrates and lectins. In *Leucocyte Typing VII*. University Press, Oxford, pp. 149–205.

Schwartz-Albiez,R. (2009) Inflammation and glycosciences. In *The Sugar Code*. Wiley-VCH, Weinheim, Germany, pp. 447–463.

Zola,H. *et al.* (2003) Human leucocyte differentiation antigen nomenclature: update on CD nomenclature. Report of IUIS/WHO Subcommittee. *J. Immunol. Methods*, **275**, 1–8.