Bioinformatics, 31(13), 2015, 2217–2219
doi: 10.1093/bioinformatics/btv110
Advance Access Publication Date: 23 February 2015
Applications Note

OXFORD

Structural bioinformatics

GlycanAnalysis Plug-in: a database search tool for *N*-glycan structures using mass spectrometry

Kentaro Morimoto[1,*], Takashi Nishikaze[1,*], Akiyasu C. Yoshizawa[1,*], Shigeki Kajihara[1], Ken Aoshima[2], Yoshiya Oda[2] and Koichi Tanaka[1]

[1]Koichi Tanaka Laboratory of Advanced Science and Technology, Shimadzu Corporation, Kyoto, Japan and [2]Eisai Product Creation Systems, Eisai Co., Ltd., Tsukuba, Ibaraki, Japan

*To whom correspondence should be addressed.
Associate Editor: Anna Tramontano

Received on October 3, 2014; revised on January 23, 2015; accepted on February 15, 2015

Abstract

Summary: Tandem mass spectrometry (MS/MS or MS^n) is a potent technique for characterizing *N*-glycan structures. GlycanAnalysis searches a glycan database to support the identification of glycan structures from MS/MS spectra. It also calculates diagnostic ions of glycan structures registered in a glycan database (GlycomeDB or KEGG GLYCAN) and searches for MS/MS spectra of *N*-glycans that match diagnostic ions to determine the structures. This program functions as a plug-in for Mass++, a freeware mass spectrum visualization and analysis program.
Availability and implementation: The executable files of Mass++ are available for free at http://www.first-ms3d.jp/english/. The GlycanAnalysis plug-in is included in the standard package of Mass++ for Windows.
Contact: k-morimt@shimadzu.co.jp or nishikaz@shimadzu.co.jp or acyshzw@shimadzu.co.jp
Supplementary information: Supplementary material are available at *Bioinformatics* online.

1 Introduction

N-Glycosylation is an important post-translational modification of proteins and has a wide variety of biological roles. *N*-Glycans form branching-tree structures consisting of two main antennae, 6-antenna and 3-antenna. Since these structures are assumed to be related to their biological functions, it is important to characterize the glycan structures.

Mass spectrometry (MS) is a powerful technique for analyzing *N*-glycans. The structures can be deduced using tandem MS (MS/MS). Typical *N*-glycan analysis using MS is performed in positive-ion mode. In contrast, Harvey's method can be adopted in negative-ion mode, which has proven useful for less ambiguous structural clarification (Harvey, 2005a,b,c). Harvey described the methodology to identify the glycan structure uniquely with the observed structure-specific 'diagnostic ions,' which are included in only negative-ion MS/MS spectra. Although various excellent computer programs (e.g. GlycoWorkBench—Ceroni *et al.*, 2008) have been developed for interpreting MS/MS spectra of *N*-glycans, they do not consider highly specific features of negative-ion mode analysis (e.g. diagnostic ions) (see Supplementary Material for details).

Recently, we developed more sensitive methods for analyzing *N*-glycan structures in negative-ion mode using matrix-assisted laser desorption/ionization-MS with unique chemical labeling (Kaneshiro *et al.*, 2011; Nishikaze *et al.*, 2012a, b). We thus developed GlycanAnalysis software to provide support to routine processes in manual interpretation of mass spectrometric data for *N*-glycan structures, using diagnostic ions searching against the public glycan structural database. For this software, we developed an algorithm to calculate *m/z* values of fragment ions, including diagnostic ions, and to search for structures matching MS/MS spectra in glycan structural databases. GlycanAnalysis is available as a free plug-in of Mass++, pluggable freeware for analyzing and visualizing MS data (Tanaka *et al.*, 2014).

2 Methods

Acidic *N*-glycans readily form negative-ion species $[M\text{-}H]^-$, while neutral *N*-glycans do not. Thus, ionizing neutral *N*-glycans as anion-adducted forms is a common approach in sensitive *N*-glycan analysis in negative-ion mode (Harvey 2005a; Kaneshiro *et al.*, 2011; Nishikaze *et al.*, 2012a, b; Rohmer *et al.*, 2010). Diagnostic

fragment ions can be obtained from both deprotonated ($[M-H]^-$) and anion-adducted ($[M+anion]^-$) forms.

To determine glycan structures, the *m/z* values of generated ions are calculated from the theoretical structure of *N*-glycans, and the diagnostic ions from MS/MS spectra are sought as either deprotonated or anion-adducted forms (NO_3^- or $H_2PO_4^-$ adducts). The search process consists of repeating the following four steps (see Supplementary Material for details).

1. Read the structure data of an entry from the glycan database.
2. Calculate the *m/z* values of theoretical fragment ions for glycan structures in all entries. This process does not generate all possible fragment ions (e.g. performing a brute-force search), but possible fragment ions specific to only negative-ion *N*-glycan fragmentation.
3. Read measured mass spectrometric raw data to automatically detect peaks and their intensities. Users themselves need not count generating peak lists.
4. Search for matched precursor ions and diagnostic ions in the MS/MS spectrum by comparing *m/z* values. A diagnostic ion is identical to a D-ion that reflects the 6-antenna structure, except when glycan structures have bisecting GlcNAc.
5. When D-ions (or D–H_2O ions) are detected in the input peak list, the corresponding glycan structure is added to the result list, the number of matching ions (A, B, C-series and E-ion) is counted. The matching score is then calculated by the ratio of its intensity to the total intensity.

As the glycan structure database, the subsets of GlycomeDB (Ranzinger *et al.*, 2011) and KEGG GLYCAN (Hashimoto *et al.*, 2006) (a KEGG user license is required) are available.

3 Implementation and usage

GlycanAnalysis runs as a plug-in of Mass++, a free software platform for visualizing and analyzing mass spectrometric data, which runs in the Windows environment; hence, all functions of GlycanAnalysis are performed on Mass++ as its functions.

For the standard analysis, the user first opens the target MS/MS data from the Mass++ file menu, performs peak detection to detect ion peaks and then opens the GlycanAnalysis plug-in. At this stage, parameters (mass-error tolerance, labeling method and anion type) must be set; the input value of the labeling method is used for annotating and calculating the *m/z* values of precursor ions. After the database search, candidates for the target glycan structure are listed in the dialog window (Fig. 1). Descriptions of annotations for the selected glycan structure in the result list are displayed in the Mass++ spectrum window. Clicking the text in the Mass++ spectrum window presents the detailed glycan structure as the annotation of glycan fragmentation in another window.

GlycanAnalysis is implemented in C++, and the graphical user interface is designed with wxWidgets, as with the Mass++ main program. GlycanAnalysis is included in the Mass++ standard package and can be downloaded for free from http://www.first-ms3d.jp/english/ under the Mass++ license.

4 Discussion

GlycanAnalysis is an easy-to-use tool for supporting the structure identification of *N*-glycans; its distinguishing characteristic is the ability to perform a database search using the specific features of *N*-glycan anion spectra. Users have the option to select a suitable peak-detection algorithm implemented in Mass++ and to optimize parameters because the performance of retrieving candidates depends on these parameters. In addition, the annotation indicated in the MS/MS spectrum window of Mass++ can be utilized to determine which peaks are used for the retrieval.

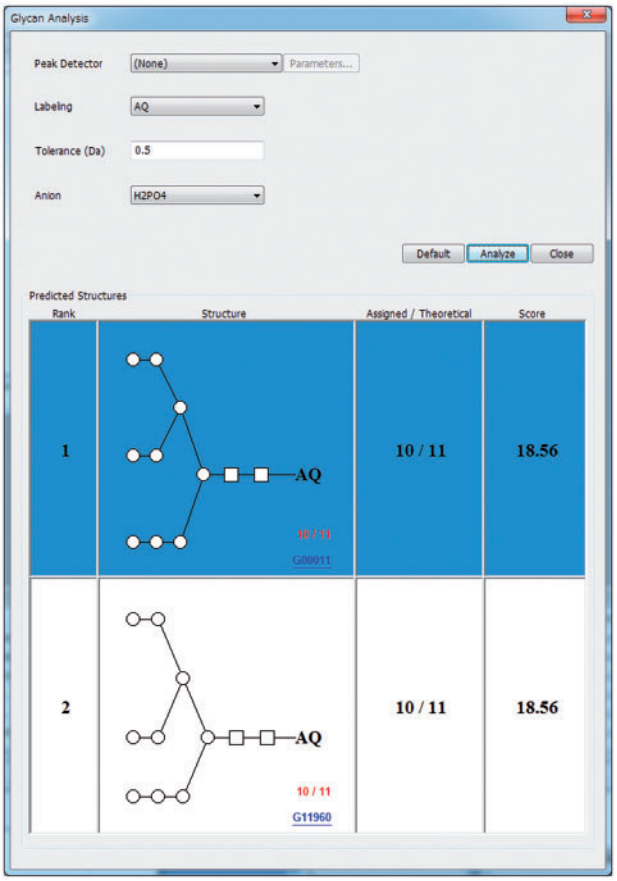

Fig. 1. Example of database search results of a high-mannose *N*-glycan. All predicted candidate structures are displayed in the lower half of the dialog window with IDs and scores. The Assigned/Theoretical column indicates the number of matched theoretical fragment ions and the number of all theoretical fragment ions. The Score column indicates the ratio of intensity of diagnostic ions to total intensity of all theoretical fragment ions. (Also see Supplementary Fig. S8 for the annotation window.)

The database search function is the direct implementation of our developed method but depends only on the existence of a diagnostic ion without any consideration of peak intensity. Similarly, the corresponding score of the retrieved glycan entry is currently defined as the number of matching ions. This scoring definition thus has room for improvement; the intensity of peaks and its related data can be employed. This indicates one future direction for improving GlycanAnalysis.

Acknowledgements

We thank René Ranzinger for permission to distribute the reduced *N*-glycan database generated from GlycomeDB.

Funding

This study was funded by the Japan Society for the Promotion of Science (JSPS) through its Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program).

Conflict of Interest: none declared.

References

Ceroni,A. *et al.* (2008) GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.*, **7**, 1650–1659.

Harvey,D.J. (2005a) Fragmentation of negative ions from carbohydrates: part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from *N*-linked carbohydrates. *J. Am. Soc. Mass Spectrom.*, **16**, 622–630.

Harvey,D.J. (2005b) Fragmentation of negative ions from carbohydrates: part 2. Fragmentation of high-mannose *N*-linked glycans. *J. Am. Soc. Mass Spectrom.*, **16**, 631–646.

Harvey,D.J. (2005c) Fragmentation of negative ions from carbohydrates: part 3. Fragmentation of hybrid and complex *N*-linked glycans. *J. Am. Soc. Mass Spectrom.*, **16**, 647–659.

Hashimoto,K. *et al.* (2006) KEGG as a glycome informatics resource. *Glycobiology*, **16**, 63R–70R.

Kaneshiro,K. *et al.* (2011) Highly sensitive MALDI analyses of glycans by a new aminoquinoline-labeling method using 3-aminoquinoline/α-cyano-4-hydroxycinnamic acid liquid matrix. *Anal. Chem.*, **83**, 3663–3667.

Nishikaze,T. *et al.* (2012a) Sensitive analyses of neutral *N*-glycans using anion-doped liquid matrix G_3CA by negative-ion matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.*, **84**, 6097–6103.

Nishikaze,T. *et al.* (2012b) Structural analysis of *N*-glycans by the glycan-labeling method using 3-aminoquinoline-based liquid matrix in negative-ion MALDI-MS. *Anal. Chem.*, **84**, 9453–9461.

Ranzinger,R. *et al.* (2011) GlycomeDB—a unified database for carbohydrate structures. *Nucleic Acids Res.*, **39**, D373–D376.

Rohmer,M. *et al.* (2010) 3-Aminoquinoline acting as matrix and derivatizing agent for MALDI MS analysis of oligosaccharides. *Anal. Chem.*, **82**, 3719–3726.

Tanaka,S. *et al.* (2014) Mass++: a visualization and analysis tool for mass spectrometry. *J. Proteome Res.*, **13**, 3846–3853.