Development of a rapid analytical system for glycans using a multistage tandem mass spectral database

— Toward an era where everyone can analyze glycan structure without specialist knowledge —

Akihiko KAMEYAMA[1*], Norihiro KIKUCHI[2], Shuuichi NAKAYA[3] and Shinji FUNATSU[3]

[Translation from *Synthesiology*, Vol.8, No.4, p.200-213 (2015)]

Conducting glycan analysis requires expertise. This requirement has been a major bottleneck in the progress of glycomics. If glycan analysis can be done easily and rapidly without specialist knowledge, then the development of glycan functional analysis and associated applications is expected to accelerate. Here, we describe the construction of a multistage tandem mass spectral database, and a system for rapid glycan analysis that utilizes this database, as examples of infrastructure development for the advancement of glycoscience.

Keywords : Glycomics, mass spectrometry, database, structural analysis, algorithm, library, isomer

1 Introduction

Many readers may not be familiar with the term glycotechnology (or glycoengineering). In fact, already a quarter of a century has passed since this term was born. When genetic engineering and protein technology that were based on the science of nucleic acids and proteins started to have large impact on society as biotechnology, the lack of knowledge of glycans, which is the third chain of life, became a problem. Against such a background, the academic discipline of glycobiology emerged, and the concept of glycotechnology was born in Japan, with the perception that this knowledge should be actively utilized in the field of biotechnology. Details are described in the beginning of *Tosa Kogaku* that was published (by Sangyo Chosakai) in 1992.[1] After about 10 years, the "Glycogene (GG) Project" was started by the New Energy and Industrial Technology Development Organization (NEDO), followed by the "Structural Glycomics (SG) Project" and the "Medical Glycomics (MG) Project." These glycan projects were led by Project Leader Hisashi Narimatsu for 10 years. The results all the way to product realization of clinical diagnostic drugs are summarized in *Synthesiology* Volume 5 Issue 3 (2012) "Development of basic tools for glycoscience and their application to cancer diagnosis."[2] The first output to society was a clinical diagnostic drug, but the results of this on-going project do not end there. One of the important results was the laying of the foundation for glycan research that was lacking in life sciences, or the infrastructure for "synthesis," "structure," and "function." Glycotechnology that was proposed 25 years ago is taking off now.

The structural analysis of glycans is difficult. The main reason is because, unlike the nucleic acids and proteins for which the primary structure can be known if the sequence is read, there are branch structures, positional isomerisms, stereoisomerisms, and others, and simple sequence decoding is not sufficient to know the glycan structure (Fig. 1). This means that the heart of glycan analysis is how to identify the isomers. The difficulty of glycan structure analysis is described in *Synthesiology* Volume 7 Issue 2 (2014) "Development of lectin microarray, an advanced system for glycan profiling."[3] The glycan structure analysis was a work undertaken by specialists with skilled craftsmanship, and this was a major bottleneck in glycan research. If glycan structure analysis can be conducted easily by anyone, it is expected that the range of glycan research will widen and the clarification of glycan functions that still remain mysterious and their application will rapidly progress. In the SG Project, two approaches were taken for the glycan structure analysis. One is the glycan profiling method[4] where the lectin array, in which various types of proteins (lectin) that can identify the partial glycan structure, is arranged on the slide glass. This method yielded results in the search for disease biomarkers and stem cell markers.[5]-[7] The search for markers demands sensitivity rather than precision, and the highly sensitive lectin microarray that can be prepared easily is utilized effectively. On the other hand, if one wanted to clarify the marker itself at a molecular level or check the content of the multiple glycan types, glycan analysis by mass spectrometry described in this paper is useful. The two methods mutually supplement each other's weaknesses. In this paper, we

1. Biotechnology Research Institute for Drug Discovery, AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba 305-8568, Japan ∗ E-mail: aki-kameyama@aist.go.jp, 2. Business Development Division, Mitsui Knowledge Industry Co., Ltd. Atago Green Hills Mori Tower 2-5-1 Atago, Minato-ku 105-6215, Japan (Present address: Bayer Yakuhin, Ltd. Open Innovation Center Japan 1-6-5 Marunouchi, Chiyoda-ku 100-8265, Japan), 3. Analytical & Measuring Instruments Division, Shimadzu Corporation 1 Nishinokyo Kuwabara-cho, Nakagyo-ku, Kyoto 604-8511, Japan

Original manuscript received December 24, 2014, Revisions received June 22, 2015, Accepted June 25, 2015

discuss the construction of a multistage tandem mass spectral database as one of the infrastructures of glycan structure analysis and the rapid analytical system for glycans utilizing this database.

2 Glycan structure analysis by mass spectrometry (MS)

Since the development of two soft ionization methods, the electro-spray ionization (ESI) and the matrix assisted laser desorption/ionization (MALDI), the use of mass spectrometry rapidly increased in the life science field. Currently, these are used widely in proteome analysis, pharmacokinetic analysis, biomarker search, microorganism identification, and others. When our research was started in the beginning of the 2000s, the proteome analysis was introduced enthusiastically as the method for post-genomic research, and MS played a central role and evolved rapidly. The glycan structure analysis at the time mainly used the database of retention time for various glycans in the high-performance liquid chromatography (HPLC). However, since precise data could be obtained quickly and with high sensitivity, it was considered a matter of time before MS would replace HPLC in glycan structure analysis. Against such a background, before the start of the project, Dr. Koichi Tanaka of Shimadzu Corporation visited the Research Center for Medical Glycoscience, AIST and introduced the MALDI quadrupole ion trap time-of-flight MS (MALDI-QIT-TOF MS) that was a new mass spectrometer developed in the UK (Fig. 2). With this spectrometer, glycans become singly ionized by MALDI, multistage collision induced dissociation (CID) is accomplished by ion trapping, and resolution is high since it is time-of-flight MS. These characteristics are appropriate for glycan structure analysis where it is necessary to determine the isomers with high sensitivity. Therefore, it was decided in the Project to develop a new glycan structure analysis method using this device.

2.1 Development trend at the beginning of the research

In the precise structure analysis of glycans using mass spectrometer, the method employed was careful analysis of fragment ions obtained by high-energy CID[Term 1] in the fast atom bombardment MS (FAB-MS), after methylation of all OH, NH, and COOH groups of the glycan (permethylation). This method could be conducted by a few, limited specialists of glycan MS, and in the omics[Term 2] boom at the beginning of the 2000s, a new simplified method was sought to replace this method. The method that was employed at the practical level was the creation of a database of a calculated fragment list of all known glycan structures and then referencing the peak list of the MS^2 spectrum of the analyzed glycans to this database. Although this was a simple method, it was not possible to identify isomers that was the heart of glycan analysis. The University of New Hampshire was engaging in research of a method that enabled identification including the isomers, but their method necessitated the permethylation of glycans before analysis.[8]-[10] There was no commercially available device or kit forpermethylation, and it was a difficult derivatization method for those without expertise.

2.2 Goal and outcome

Our goal is the development and product realization of a new system where anyone can easily conduct glycan analysis including the identification of isomers. By making such a system available in society, we hope that glycan analysis will become familiar to many life science researchers who have kept a distance from glycans, and by increasing the

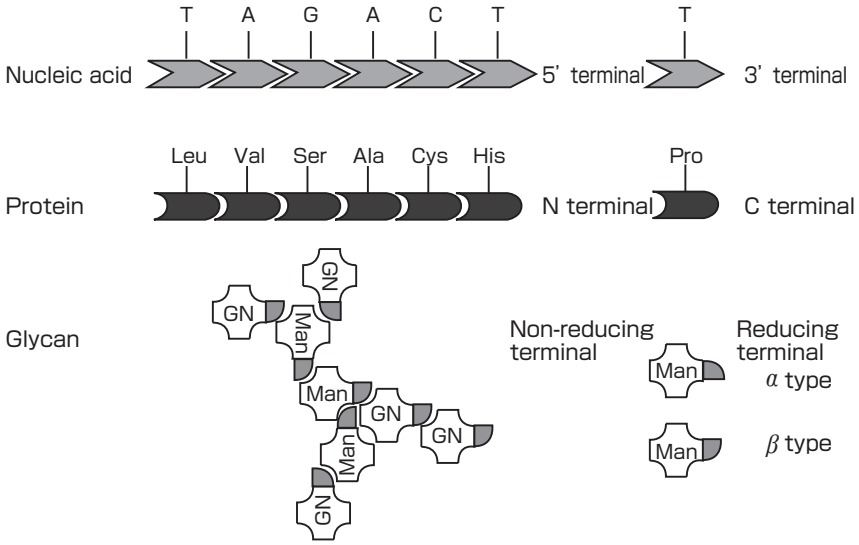

Fig. 1 The reason why it is difficult to conduct sequence analysis of glycan
In glycans, the reducing terminal and non-reducing terminal of sugar bond together. There are four bonding sites (shown as fan-shaped concave parts) in the non-reducing terminal, and there are α and β stereoisomerisms in the bonding sites of the reducing terminal (shown as fan-shaped convex parts). With this configuration, complex isomers with branching structures are formed.

number of people who study glycans, the discoveries of new glycan functions and glycan biomarkers that were formerly unknown are expected to increase. The findings of new glycan functions invite more technological developments, and glycotechnology will spread into the life science research sites. We drew such an outcome.

3 Scenario for the rapid analytical system for glycans

The scenario to achieve the goal did not exist from the beginning. Since the MALDI-QIT-TOF MS was a new type of spectrometer, we did not know what kind of data could be obtained when the glycans were analyzed using this device until we actually did the measurements. First, there was the idea that the laws of fragmentation would be found by analyzing a number of glycans, and then the structure could be estimated from the MS^n [Term 3] spectra based on the laws. We collaborated with the Computational Biology Research Center, AIST, and although we obtained some research results, they did not lead to practical use.[11]-[13]

Under the thinking that identification of isomers is the heart of glycan analysis, we compared the MS^n spectra of multiple glycans with the same molecular mass. It was found that with MS^2 some gave extremely similar spectrum while most isomers gave different spectra when compared up to MS^3.[2] On the other hand, cloning had been done for almost all human glycogenes at that point in the Project, and there was a library of glycosyltranferase[Term 5] coded for each glycogene. Since the glycosyltransferase is extremely specific, it is possible to selectively synthesize a desired isomer by selecting the appropriate enzyme.[14] Therefore, we considered the following scenario (Fig. 3).

First, several types of reference glycan with known structures are purchased, and variations of the reference glycans are increased by extending the sugar chain specifically using glycosyltransferase. Next, the MS^n spectra of each reference glycan are measured and made into a database as fixed values

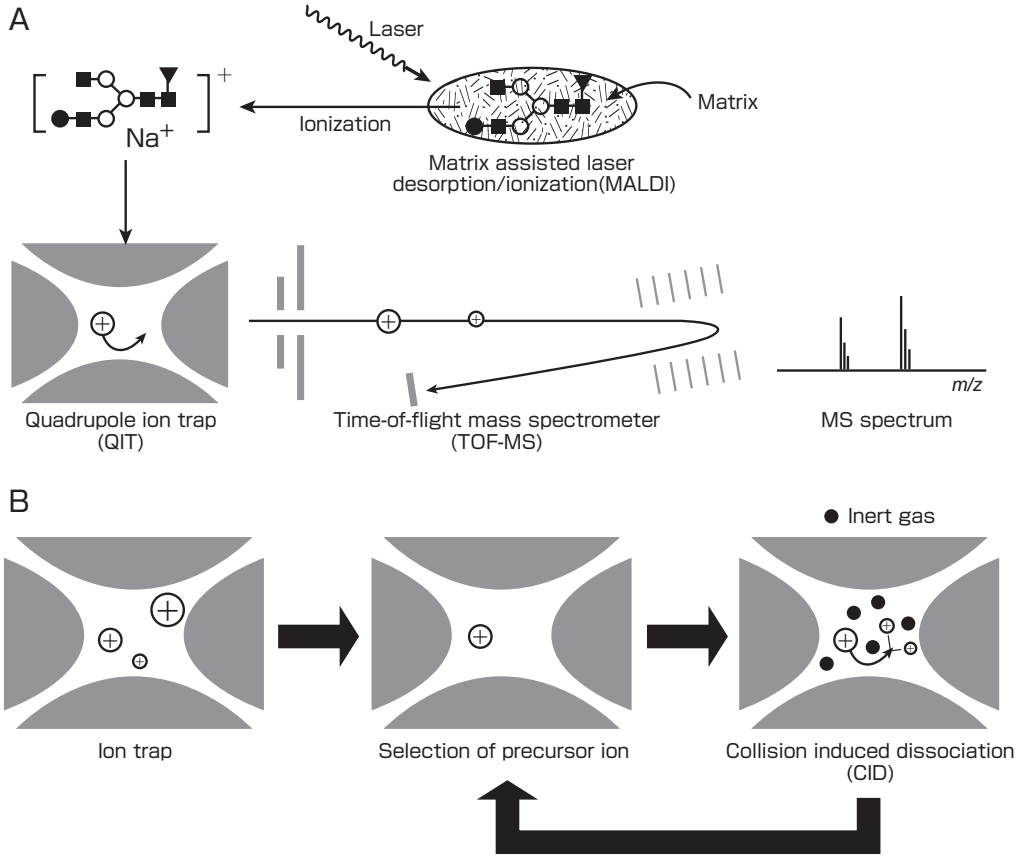

Fig. 2 Schematic diagram of the MALDI-QIT-TOF MS

A: Ions are produced (in the figure, sodium ion adduct is given as an example) when laser is irradiated on the sample that is a mixture of a glycan and a matrix. After the ion is captured in the quadrupole ion trap, it is sent to the time-of-flight mass spectrometer, and the time taken to arrive at the detector is measured. The value where the time is converted to mass-to-charge ratio (*m/z*) is displayed as the MS spectrum.

B: The ion trap captures the ion that is in a certain range of the *m/z* value. When the ion is sent to the TOF-MS, the MS spectrum can be obtained. Also, ions that do not have a certain *m/z* value can be eliminated from the captured ion (selection of precursor ion[Term 4]). When the ion and inert gas collide, the ion dissociates into smaller ions (collision-induced dissociation). The MS^2 spectrum can be obtained by sending the dissociated ion to TOF-MS. If the selection of precursor ion and collision induced dissociation are repeated using the dissociated ion, the MS^3 spectrum is obtained. Theoretically, by repeating the procedure n times, the MS^n spectrum is obtained.

for each glycan. Algorithm to estimate the glycan structure by comparing the MS^n spectra of the analyzed sample and the spectra in the database is developed. Then, the interface software for linking the structure estimation algorithm and the MS operation software is developed. Finally, these are all integrated to create a product, the rapid analytical system for glycans with excellent stability, reproducibility, and ease of use.

4 Development of the elemental technology

To realize the aforementioned scenario, a rapid analytical system for glycans was developed jointly by three parties: AIST, Mitsui Knowledge Industry Co., Ltd., and Shimadzu Corporation. AIST was in charge of the construction of glycan spectra database using the resource from the glycosyltransferase library; Mitsui Knowledge Industry worked on the structure estimation algorithm since it had experience in glycan informatics such as glycogene search; and Shimadzu, the manufacturer of MALDI-QIT-TOF MS, was in charge of the MS and interface software. The details will be explained below.

4.1 Construction of the glycan MS^n spectra database

4.1.1 Selection of the glycan labeling agents

In glycan analysis, generally, fluorescent labeled glycans or permethylated glycans are used, and it is rare to analyze glycans with no derivatization. Therefore, it is necessary to create a database for derivatized reference glycans. In this case, since it was not realistic to prepare several types of derivatives for one glycan, it had to be narrowed down to one. There are several fluorescent labeling agents for glycans, and 2-aminopyridine (PA) is used frequently in Japan, while 2-aminobenzamide (2-AB) is regularly used in Europe and the USA.[15][16] There are personal preferences for fluorescent labels among researchers, and there were people in the NEDO Project who promoted other labeling agents such as pyrene derivatives or 3-aminobenzoicacid (3-AA). As guidance for selecting the labeling agents, there were three points: ionization efficiency in MALDI, information volume of the MS^n spectra obtained by low energy CID,[Term 6] and wide usage in glycan research labs. In other words, focus was placed on sensitive detection, slight structural change that was reflected in the MS^n spectrum, and widespread use by many people. After various deliberations, we decided on PA that was used frequently in Japan because it had good balance in terms of sensitivity and MS^n, and this agent was selected as the glycan labeling agent for the database.

4.1.2 Maintaining the reproducibility of data

To use the database, the reproducibility of data is essential. However, in the tandem MS by low-energy CID, the spectra fluctuated depending on the energy intensity on the precursor ion (CID energy), and since it was impossible to strictly control the energy, it was necessary to devise a way to obtain good reproducible spectra. We focused on the fact that almost the same MS^n spectrum could be obtained every time even if the CID energy fluctuated, if the spectrum was measured at the CID energy when the precursor ion almost disappeared (Fig. 4). Therefore, after various trials, we established a standard of measuring the MS^n spectrum at CID energy at 15 % or less of the maximum peak of precursor ion intensity.[17] Moreover, we ensured that, for each type of precursor ion, the variations in real measurement values could be absorbed by storing two spectra data for MS^2 and three spectra data for high-level spectra of MS^3 or higher where the data tended to fluctuate.

4.1.3 Unification of the measurement mode and the precursor ion species

The mass spectrometer is a device that ionizes the molecules and separates and detects the ion by mass-to-charge ratio. The measurement mode for analyzing positively charged ion is called the positive ion mode, and the mode for analyzing negatively charged ion is called the negative ion mode. When constructing the database, we considered which measurement mode should be selected. While we were attracted to the negative ion mode since there were reports of production of special fragments that might be useful in the glycan structure estimation,[18][19] we selected the positive ion mode since the negative mode was disadvantageous in terms of ionization. In

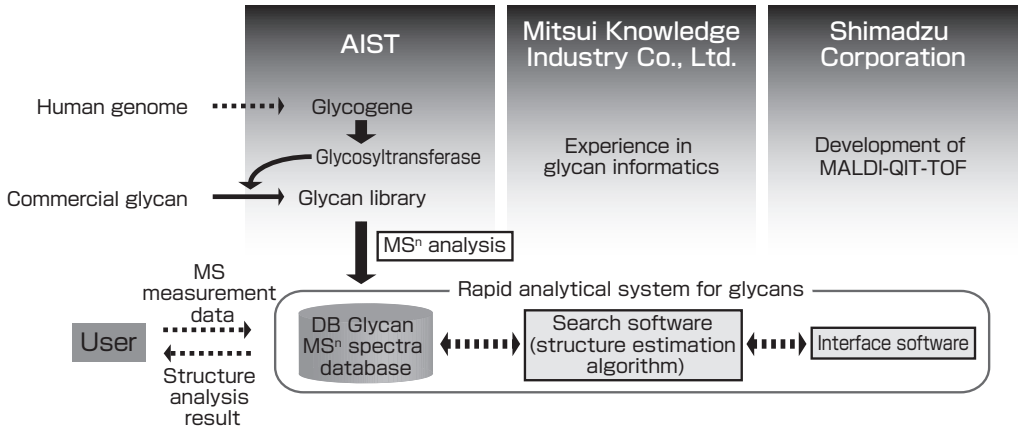

Fig. 3 Elemental technologies, their background, and mutual relationships

ionization, various ions with added protons, sodium ions, and potassium ions could be produced, and there were differences in fragmentation according to the adduct ion. Therefore, it was necessary to determine the ion species. In case of glycans, it was found that the spectra difference readily occurred between the isomers by using the sodium ion adduct as the precursor rather than using the proton adduct as the precursor, and we selected the sodium ion adduct. Also, there were reports that for the proton adduct of fucose containing glycans, fucose rearranged in the ion trap.[20] and the sodium adduct was considered more appropriate.

4.1.4 Selection of the matrix

In MALDI, the selection of the matrix is important. Glycans degrade readily in acid conditions, and it is common that decomposed matters may be generated during ionization if the acid matrix is used. Also, due to the unevenness of sample concentration that may occur when creating the co-crystals of the sample and the matrix, there are problems of so-called "sweet spots" or the area from which the signals can be obtained is limited when irradiating with lasers. Since the data reproducibility is necessary in the database, it is desirable that the measurement be done automatically with no inclusion of operator bias, but sweet spots make that difficult. We considered various matrices, and selected a method of creating homogenous microcrystals with no sweet spots by making the co-crystals and then re-crystalizing by ethanol, using 2,5-dihydroxy benzoic acid (DHB) as the matrix.[21] The DHB was an acid matrix and glycans degraded during ionization in some cases, but this method was best considering the issues of sensitivity and sweet spots.

4.1.5 Glycan preparations for constructing the database

Since there were only a few types of pyridylaminated glycans that were commercially available, we modified the commercial pyridylaminated glycans using the glycosyltransferase at the Research Center for Medical Glycoscience to increase the variation. In the synthesis for the glycan library, Researcher Hiromi Ito of the RCMG (currently at Fukushima Medical University) played a central role. The above measurements were conducted using the samples, and ultimately 2,897 spectra were incorporated into the glycan MS^n spectra database.

4.2 Structure estimation algorithm and search software

4.2.1 Glycan description language

The database stores the structural information of glycans and the MS^n spectra data for those glycans. For the glycan structure, it is necessary to describe the branches and bonds, and these were difficult to handle with the computer. There was no standard method for data description at the beginning of the project. Therefore, we developed the carbohydrate sequence markup language (CabosML) using XML.[Term 7][22] The structure described in the CabosML format was stored in the database, and various applications were developed using the CabosML format for the input data.

4.2.2 Elimination of the noise peaks

With the search software, the structure estimation is done based on the similarity of spectrum form including the peak intensities, by comparing the measured MS spectra of unknown structures and the spectra of known structures stored in the database. If the noise peak is included in the reference spectra data, the search precision decreased due

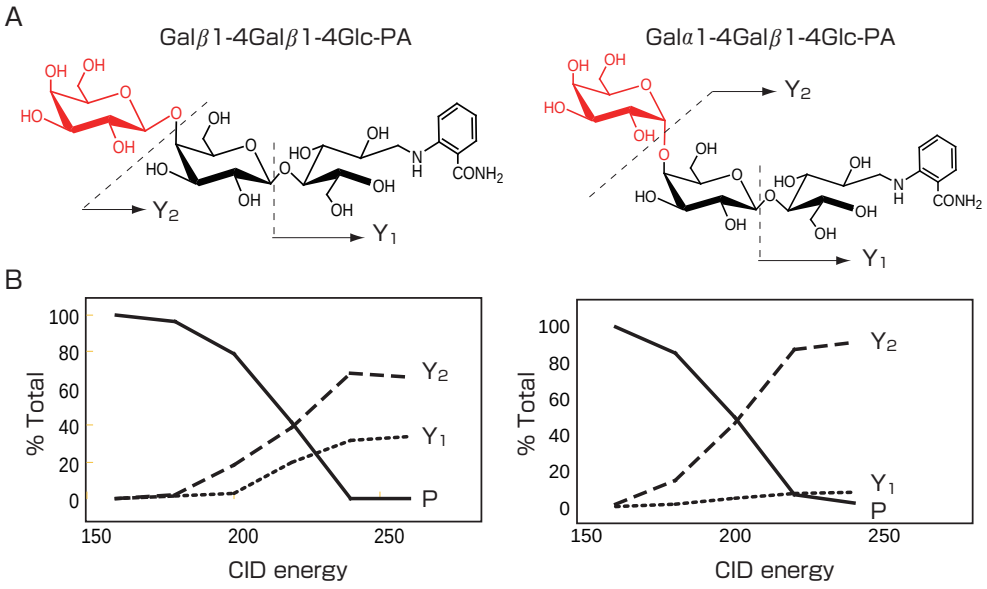

Fig. 4 Fragmentation of the glycans by ion trap

The glycans on right and left are isomers in which only the configurations of terminal galactose (shown in red) are different. The graphs beneath them show the changes in peak intensity of the fragment ion produced from each glycan when the CID energy is increased. In the CID energy where P almost completely disappears (around P < 15 %), the ratio of Y_1 and Y_2 become almost constant. The Y_1/Y_2 ratio that becomes constant is different between the two isomers. P: precursor ion, Y_1, Y_2: fragment ion.

to the noise peak. Therefore, the theoretical fragments were calculated from the glycan structure, and only the peaks that matched the monoisotopic mass[Term 8] of each fragment were extracted from the spectra data. The peak list including the intensity information was set as the spectra data for search.

4.2.3 Detection of the glycan peaks

The MS spectra of the samples include the peaks of foreign materials, decomposed matters, as well as peaks deriving from the matrix, other than the peaks of the glycans that one wishes to study. Therefore, the first step of database retrieval is to detect the glycan peak from the measured spectrum. While there are various methods, in the structure estimation using the database, there is no meaning in analyzing the glycan peak not registered in the database. Therefore, judgment for glycan peak detection was made according to whether the glycan peak was registered or not registered in the database (Fig. 5).

The interface software that will be described later delivers the sample information such as the labeling agent and the adduct[Term 9] along with the peak list. The search software reviews such information, searches for the precursor ion with the same *m/z* value as the registered peak, and notifies the interface software of the found peak as being a glycan-derived peak. If it is not found, the glycan peak is detected in the Extended Search mode described later (see Section 4.2.7).

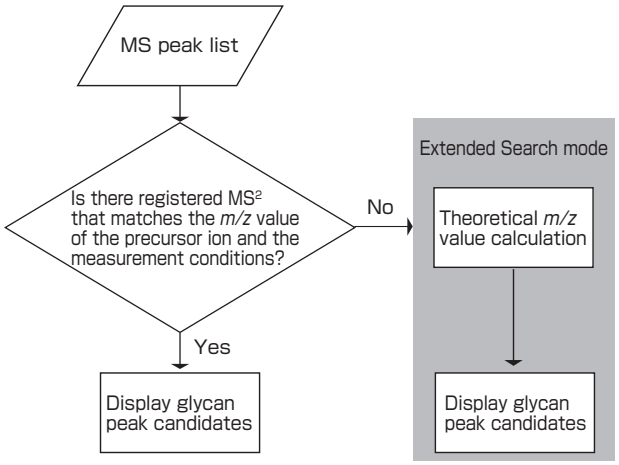

Fig. 5 Logic flow of the glycan peak detection

4.2.4 MS^2 search

In the MS^2 search, the similarities between the MS^2 spectra of the sample and the MS^2 spectra in the database are referenced. The similarity is calculated using the vector composed of the *m/z* value and the peak intensity of each peak, and the glycan with the value at the threshold or higher will be given as the candidate for the estimated structure.

4.2.5 MS^3 search and rapid identification

When there are multiple candidates for the estimated structure, narrowing down is done by measuring MS^3. In MS^3, each peak on the MS^2 spectra will be the precursor ion candidate (Fig. 6). Too much time and labor will be spent if the MS^3 is measured for each peak of the MS^2 spectra and then compared with the spectra in the database. Therefore, "which peak should be measured to narrow down to one candidate" is projected using the MS^3 spectra in the database, and rapid identification is achieved by sending this information to the interface software.

4.2.6 Extended Search mode

As mentioned earlier, we unified the fluorescent labeling agent for glycans to PA for the database construction. Therefore, glycan structure estimation cannot be done if the user is using a different labeling agent. This will be a major weak point when introducing this database to Europe and the USA where other labeling agents are used. Therefore, we solved this problem by devising a search method, without

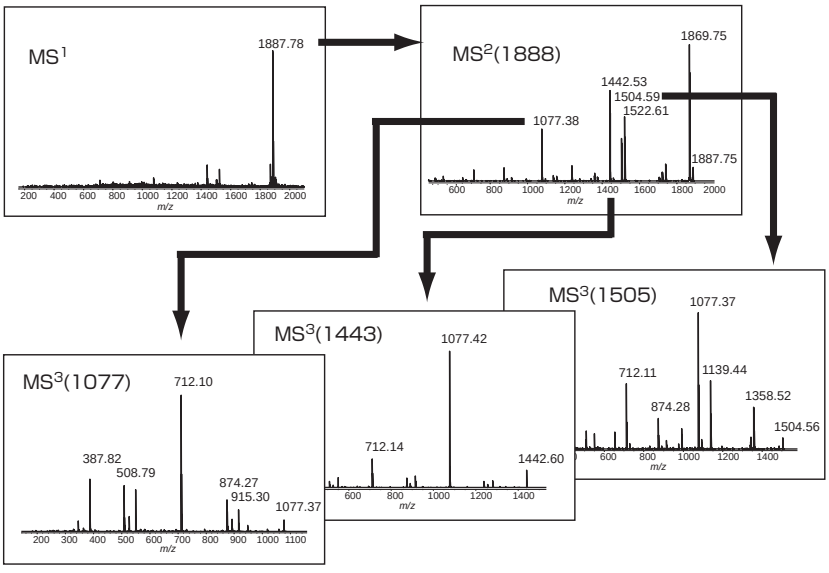

Fig. 6 Tree structure of the multistage MS

adding to the content of the database. This is the method called the Extended Search mode. In the Extended Search mode, the following search is conducted.

4.2.7 Detection of the glycan peak in the Extended Search mode

The user enters the sample information such as the labeling agents and adduct of the glycan to be analyzed. The system calculates the *m/z* values of various glycans that can be conceived theoretically by combining the glycans based on the given information, and creates the aggregation of these *m/z* values. If the *m/z* value of the peak of MS spectrum measured exists within the aforementioned *m/z* value aggregation, that *m/z* value is shown as the glycan peak candidate (Fig. 5 right).

4.2.8 Presentation of the key fragment

Since the labeling agent for MS^2 spectrum of the presented glycan peak candidate is different, the similarity with the spectra in the database cannot be determined. However, there are many fragments from which the labels have fallen off that appear in the MS^2 spectra. By comparing the MS^3 spectra assuming these are precursor ions, it seems to be possible to estimate the structure. What should be noted here is that each peak of MS^2 spectra may not necessarily be composed of a single fragment. Most peaks are composed of a combination of multiple fragments with the same *m/z* value. Comparing this with the MS^3 that sets such a peak as precursor ion is not useful in structure estimation. To conduct structure estimation, the peak composed of a single fragment structure without the labeling agent (key fragment) must be selected as the precursor ion of MS^3. In the N-glycan, the chitobiose section of the reducing terminal is readily cleaved, and the *m/z* value of the fragment cleaved at this section theoretically cannot contain other fragment structures, and therefore it can be used as the key fragment (Fig. 7).

In the search software, the *m/z* value of the key fragment is calculated from the *m/z* value of the precursor ion of the glycan peak candidate, that is, [*m/z* value - (labeled GlcNAc)] or [*m/z* value - (Fuc labeled GlcNAc)], and search is done to see whether this exists in the MS^2 data measured by the user. However, in the structure where the Fuc is bonded to the reducing terminal GlcNAc (Fig. 7A), the peak [*m/z* value - (labeled GlcNAc)] is not formed, and only [*m/z* value - (Fuc labeled GlcNAc)] is given (Fig. 7B, corresponding to the key fragment). Considering these factors, the search is conducted according to the logic flow in Fig. 8.

4.2.9 MS^3 search

When the MS^3 measurement data of a key fragment is sent to the search system, the search is conducted of the MS^3 data in the database to look for items with high similarity. In a case where there are several glycan structure candidates with a degree of similarity over the threshold, the *m/z* value of the precursor ion of MS^4 data that can be used to narrow down the candidates is sent to the interface software to conduct the measurement and search. If the MS^4 candidate does not exist in the database, the search is terminated.

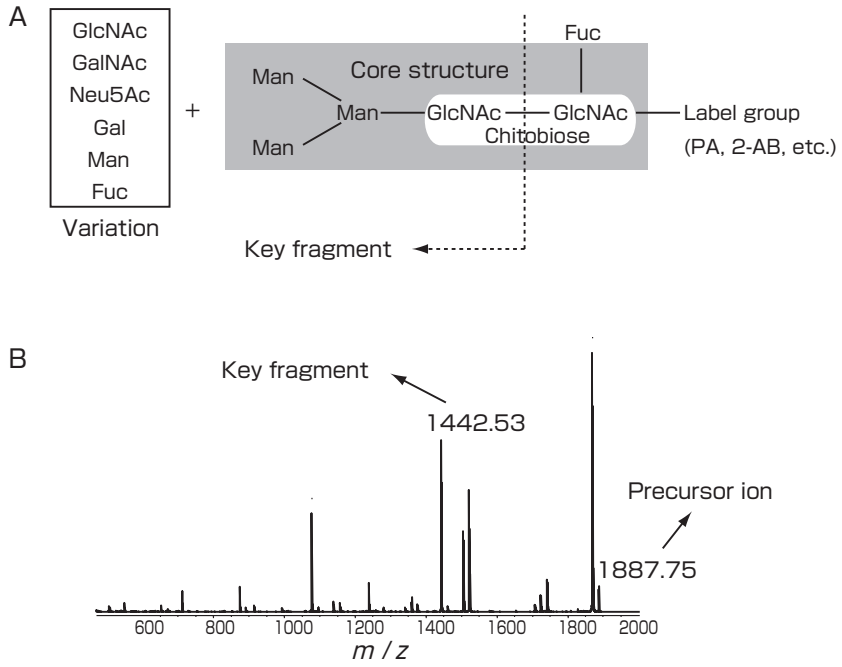

Fig. 7 Structural characteristic of the N-glycan and key fragment

A: For N-glycans, variations occur when various types of sugar in the white area join by various bonding patterns to the Man on the left side of the core structure.

B: In the MS^2 spectrum of N-glycans, the key fragment forms a large peak. This peak is formed by single structure that does not contain the label.

4.3 Interface software

The rapid analytical system for glycans aims at anyone being able to reach the estimated structure of the analyzed samples by measuring the MS spectrum through the navigation by the search software. To realize this, an interface software that can be used "easily by anyone" is necessary, as well as linking the analysis software that controls and operates the mass spectrometer, the search software, and the database. In the interface software we developed, we implemented the function to deliver the data from the search software to the analysis software with a click of a button, as well as the function to deliver the spectra data from the analysis software to the interface software with one click. The outline of the analysis flow using the interface software is shown in Fig. 9.

4.3.1 Measurement aid through the interface software

The measured MS spectrum data is delivered to the interface software after being converted to the peak list format composed of the *m/z* value and peak intensity value of each peak by the analysis software.

In the interface software, the search parameters are entered first. The parameters include the labeling agent, sample information such as adducts, as well as the tolerance setting and others. Next, the parameters set by the user are added to the peak list information obtained from the analyzing software, and are sent to the search software. The interface software receives the list of precursor ions that should be measured next from the search software and displays them. Items are prioritized in the lists, and the precursor ion information selected by the users are sent to the analysis software.

The display method of precursor ions that are measurement candidates was selected based on the visual ease of understanding for the user. Specifically, by organizing the measurement candidate in a tree format by each search ID, the

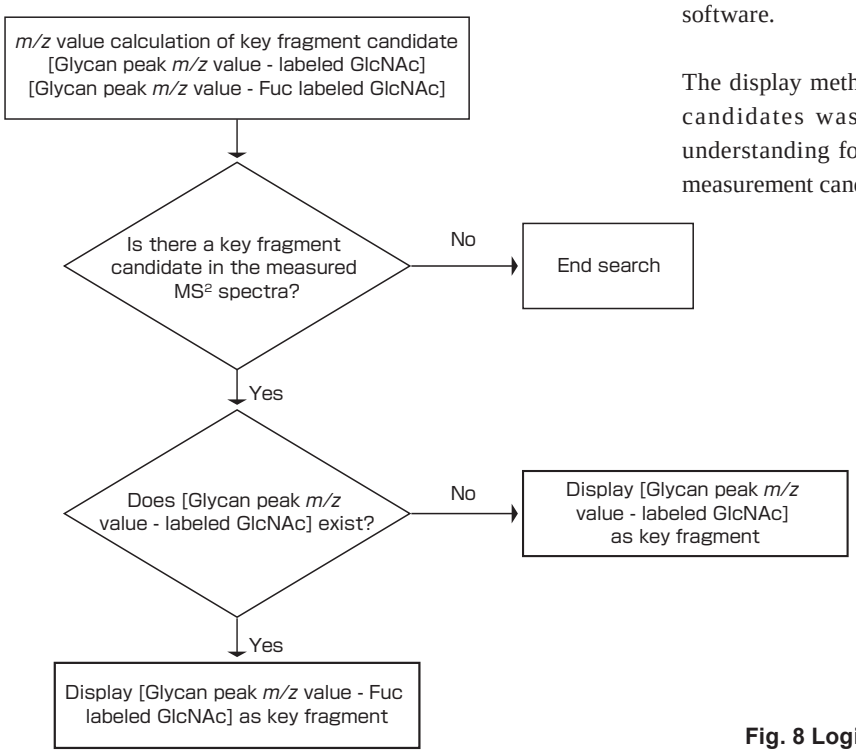

Fig. 8 Logic flow of the key fragment identification

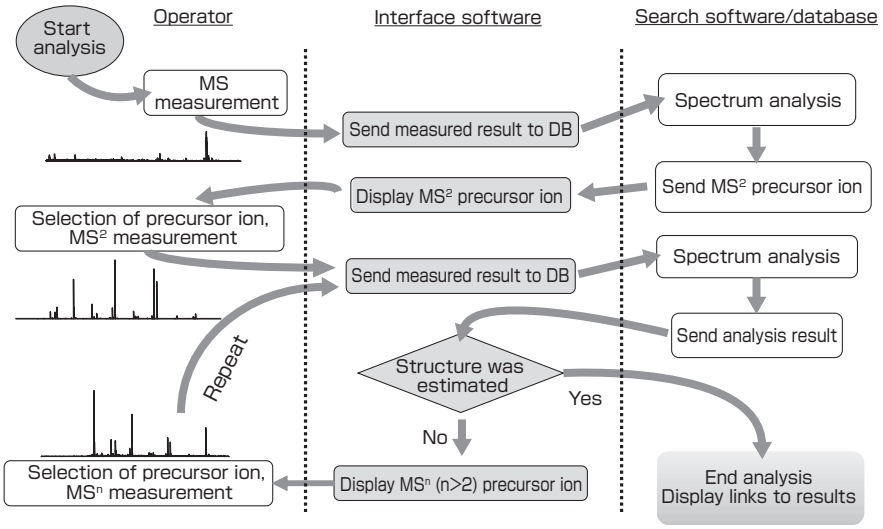

Fig. 9 Analysis flow of the rapid analytical system for glycans

The interface software executes the functions shown in the center, and mediates between the operator and the search software.

multiplier of the MS^n currently measured and the *m/z* values of the precursor ion in the previous step can be understood readily. Also, the icons of each precursor ion have different colors according to the acquisition status of the spectra data (Fig. 10).

In the analysis software, the precursor ion information received is reflected in the setting section of the MS^n analysis condition. The user sets the appropriate CID energy value to execute the MS^n analysis. The MS^n spectrum data obtained is sent again to the interface software to conduct the search.

4.3.2 Displaying the results

The interface software is implemented with the function to display the message from the server. If the structure estimation result is obtained by the search software, a message is displayed and the user discerns the end of the analysis. The structure estimation result is aggregated on the HTML format web page through the search software, and viewing is done on the Internet browser. The interface software receives notification of the web address of the estimation result page from the search software, and with a button click, the browser opens to display the web page for the structure estimation result. The estimation result shows the score in graph format, the estimated glycan structure, and the link to the Japan Consortium for Glycobiology and Glycotechnology Database (JCGGDB) for viewing the information for the related glycan structures (Fig. 11).

5 Intellectual property strategy

For product realization, securing of the intellectual property rights is essential. The patent for this system was filed under the title “Method of identifying sugar chain structure and apparatus for analyzing the same.”[23] This is the patent to claim the method and the apparatus that embodies the method. The method of estimating the structure by matching the spectrum of the analyzed substance to the spectra in the database is not particularly novel, and a prior patent had been already filed, even limited to glycans. After consulting the patent attorney, we decided to emphasize the idea that would be the core of the system. This was based on the idea that “the heart of glycan structure analysis is the identification of isomers,” and it was a method for searching the MS^3 spectrum that is expected to show the most difference in spectra among the isomers in the database when estimating the structure by matching the high-level tandem MS spectrum of MS^3 or higher, and comparing only the MS^3 spectrum with a degree of similarity at a certain value or less. In the actual patent claim, this idea was expanded to MS^n, not just MS^3. Care had to be taken because the feeling of the researcher was that the patent could be obtained just for the method that used the glycan MS^n database that never existed before. This patent was registered in Japan, Germany, and China after a few office actions.[24][25]

For the intellectual property strategy, there is the method of obtaining the rights by claiming the copyright of publication, other than obtaining the patent. The rapid analytical system for glycans is composed of the mass spectrometer, the glycan MS^n spectra database, the structure estimation algorithm, and the interface software. The glycan MS^n database was one of the intellectual foundations, and since it could be used independently, it was created and registered as intellectual property as a unique publication of AIST.[26] The search software with the structure estimation algorithm was registered as a joint publication of Mitsui Knowledge

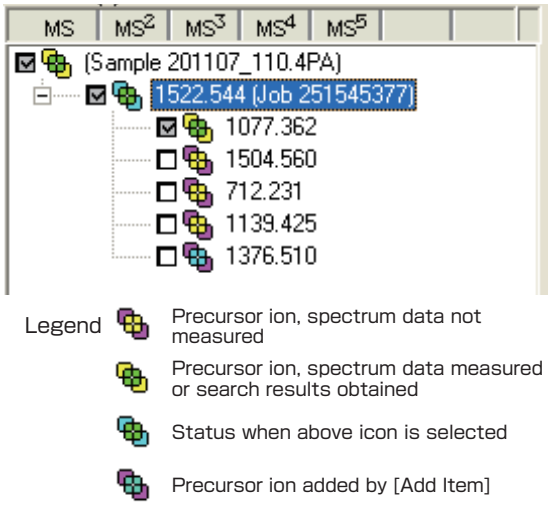

Fig. 10 Display of the precursor ion candidates

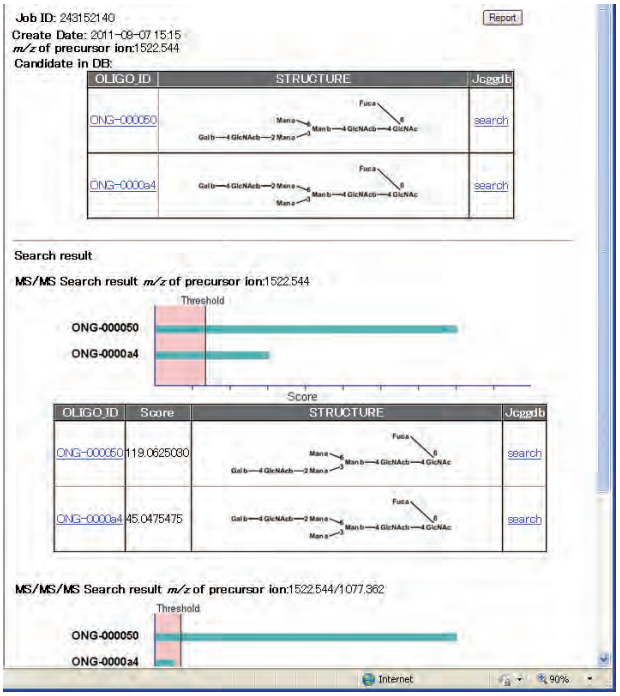

Fig. 11 Screen of the search result

This shows matching glycans where the score bars are within the threshold (red area). When one clicks on [Search] in the Jcggdb section, the page with the information for that glycan opens.

Industry Co., Ltd. and AIST under the title "CASIS."[27] The mass spectrometer and the interface software were already covered by the patent and copyright of Shimadzu Corporation.

Full Research toward product realization after filing the patent was continued in the framework of patent realization joint research that was handled by the Intellectual Property Division of AIST at the time. Product realization might not have happened without this framework.

6 Product realization of the rapid analytical system for glycans

To promote diffusion of this system, Shimadzu Corporation worked on the development of the product as a package of three components (database, search software, and interface software). The MS^n spectra database for glycans in this system was created for the glycans that are synthesized in human cells. Therefore, the focus was set on fields that studied humans, and the antibody pharmaceuticals for which the development was increasing were set as the main target at the time. Joint research for patent realization was started among three parties, AIST, Mitsui Knowledge Industry Co., Ltd., and Shimadzu Corporation.

In the SG Project, since the data collection and investigation were done on one mass spectrometer, in product realization, it was necessary to check whether correct search was conducted after considering the device differences among the spectrometers. Therefore, several devices that were the same as the spectrometer used in the Project were obtained, the data measured by these devices were compared, adjustments in the search algorithm were made to avoid large effect on search due to the data differences among devices, and the method for inspecting the device condition was established. Concurrently, we attempted the expansion of the glycan MS^n spectra database. There was a plan to add more glycan MS^n spectra to the database, mainly for the N-bond glycans containing sialic acid. However, this data collection was stopped because the antibody pharmaceuticals obtained from the Chinese hamster ovaries (CHO) cells, which is the production host necessary for the biopharmaceuticals, contained hardly any sialic acid and because the measurement of glycans containing sialic acid required a separate chemical treatment called methyl esterification. It was decided that the effort should be spent on completing the software. The early commercial version of the Accurate Glycan Analyzer, the rapid analytical system with functions developed during the Project, was released (June 2010).

At the time, this system was capable of estimating the structure of the PA-labeled N-glycans and the sugar alcohol of O-glycans in which reducing terminals were reduced. Although it caught attention of various companies and researchers, there were requests for improvements from people who used labeling methods other than the PA label. Therefore, we planned the development of an improved system that could meet the user demands, such as the function to enable structure estimation for glycans treated with 2-AB or 2-aminobenzoic acid (2-AA) or glycans labeled with various agents other than PA (see aforementioned Extended Search mode), the function where the users themselves can expand the database, and the data verification function that allows search using the data registered by the user.

Considering the required specifications to achieve all of the above functions, it was found that the development budget went beyond the joint research budget, and the narrowing down of the core functions became necessary. Upon discussion among the three parties, we reached the conclusion that priority should be given to incorporating various labels. We decided to develop the improved rapid analytical system for glycans with the Extended Search mode, where the glycan structure estimation could be done using the existing database in the case where user's labeling agent is used, as well as the commonly used glycan labels such as PA, 2-AB, and 2-AA. We were able to release the Accurate Glycan Analyzer 2, the improved rapid analytical system for glycans that can respond to various glycan labeling agents (December 2011).

7 Results and their significance

7.1 Continue on without being swayed by the trend

The first version of the rapid analytical system for glycans was released in June 2010, and the second version that was improved based on user demands was released in December 2011. In retrospect, none of the competitors at the time when we started the development achieved product realization. To get the product released, one must continue on steadily toward product realization without being swayed by the prevailing trend. Pertaining to this Project, in the beginning, the development of a new glycan analysis method using mass spectrometry was the central issue of glycomics, and active research was being done throughout the world. Yet, in about three years, the interest shifted to disease biomarkers. Many people who had been studying the glycan analysis method jumped to the glycan analysis of various clinical samples, and competed in publishing papers and filing patents for biomarkers one after another. The development of the analysis method was incomplete, but they threw that out and shifted to the next trend. Biomarker search became the theme in NEDO's glycan projects, and the authors were also swept up in the wave and had to give up system development under the NEDO Project. Fortunately, research on the rapid analytical system was continued on a separate budget of the Intellectual Property Division from 2007, and finally we achieved the above results. Currently, this system is employed

at 10 universities, companies, and research institutions in six countries (as of April 2015). Also, the actually measured MS^n spectra database was released as an open database. The open version does not involve numerical data but consists of images of MS^n spectra, and this maintains academic usefulness while avoiding the pirating of the rapid analytical system for glycans.

7.2 Power of software backed by experimental science

The product we developed is the world's first glycan analysis system with an actually measured MS^n spectra database. Anyone can easily identify isomers that are the heart of glycan analysis. Currently, there is no other glycan analytical system using an actually measured database other than our product on market. This system was born by combining hardware consisting of a new type of mass spectrometer and three software (spectra database, structure estimation algorithm, and interface software). To maximize the capacity of the hardware, the power of software that captures the on-site demand is necessary. Just as the car navigation system can guide a person to a destination even in an unknown place, this system interactively leads the user who may not know anything about glycan analysis to the estimated structure. However, the development of this system was not done just by information science, and we would like to emphasize that the foundation is the experimental science of glycogenes and glycan analysis.

It should be noted that glycan analysis software using a theoretical database is available on the market, but there is none that is of world standard at this point, including the free products.

7.3 Were we too ahead of our time?

The heart of glycan analysis is how to identify the isomers. We developed this system for that purpose. While it is very clear that the differences in glycans by positional isomerisms or stereoisomerisms lead to different functions in the body, people who study this are the glycobiology researchers who have been studying glycans from the beginning. For ordinary life science researchers, knowing the molecular mass of a glycan is sufficient, and they do not even imagine that there are possibilities in isomers. While this situation may change as the functional clarification of glycans progress, we feel that we were ahead of our time in developing the rapid analytical system for glycans that allows anyone to easily study glycans.

8 Toward the realization of the outcome

The real outcome is the situation where the knowledge of "glycans," which is the third chain of life, is used effectively and routinely in life innovation such as drug discovery and regenerative medicine. In other words, it is a situation where studying glycans is a standard procedure. The lectin microarray and the rapid analytical system are only parts of this goal. In fact, the technological foundation of glycoscience is still fragile. The glycans for which we created the database is limited to human glycans. Mucin, which is the viscous component of the mucus and bodily fluids, is a giant glycoprotein and is said to be related to disease, but the understanding of its function has not progressed because it is difficult to analyze. Podocalyxin that was recently reported as one of the iPS markers is also a mucin-like protein.[28] Currently, research is underway for the analysis of O-glycans that is a mucin glycan. For glycan analysis, it is necessary to pretreat the glycoprotein and derivatize it to a state where its glycan be analyzed, but this pretreatment process is a barrier in the diffusion of glycan research and must be improved (Fig. 12). These issues are only examples, and we face huge obstacles when the issues of "synthesis" and "function" are combined.

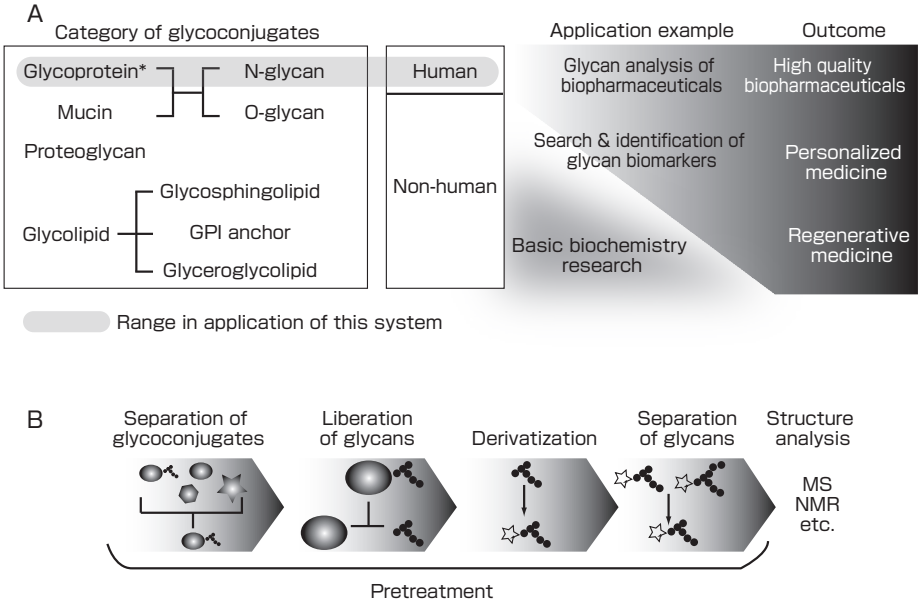

Fig. 12 Positioning of this system in the glycoconjugate research and future issues

A: The glycans exist as glycoconjugates bonded to proteins and lipids. Although there are various glycoconjugates, this system targets only human N-glycans of glycoproteins. ∗Mucin and proteoglycan are also glycoproteins, and therefore, to differentiate from those, here, we refer to glycoproteins with 50 % or less sugar content.

B: Multistage pretreatment is necessary for analyzing the glycoconjugates. This system supports only the final stage or the structural analysis.

Although industrialization is necessary to realize the outcome, and it has been said that glycan research is Japan's stronghold, in the current situation, there are very few cases where this research was released to society as industrial science. In this sense, in the future, AIST that has accumulated various glycan resources (knowledge, technology, and personnel) must not just deepen and develop the technology but must send this industrial technology into society through collaboration with companies. On the other hand, the important requirement for a company to send a product into society is profitability. When one attempts to realize new glycan technology as products, this is the first problem one must face. To expand the market, it is necessary to have the pharmaceutical companies and life science researchers to widely recognize the importance of glycans. To do so, it is necessary to continue to elucidate glycan functions and accumulate case studies with impact. Although this may be a commonplace conclusion, in order to realize the outcome, we believe we must steadily engage in a wide-range of research from basic glycan research to transferring the technology to the companies.

Terminologies

Term 1. High-energy CID: Here, this term refers to the collision-induced dissociation in magnetic sector MS and time-of-flight MS. The molecule is characterized by the fact that although it may be cleaved in various places due to collision, multiple places are not cleaved at the same time.

Term 2. Omics: Research method in which the total biomolecular species in an organism are comprehensively researched. If the subject is genes, it is genomics. If it is protein, it is proteomics. In the beginning of the 2000s, various omics were born including metabolomics and glycomics due to the rapid development of mass spectrometer.

Term 3. MS^n: Multistage tandem MS. See Fig. 2B for MS^n using the ion trap.

Term 4. Precursor ion: Ion selected to be fragmented in the tandem MS.

Term 5. Glycosyltransferase: The enzyme that transfers sugar and extends the sugar chain. It has extremely high specificity for bonding position and stereoisomerism. Therefore, precise synthesis of glycans is possible by selecting the appropriate glycosyltransferase.

Term 6. Low-energy CID: Here, this term refers to the collision-induced dissociation in the quadrupole MS and ion trap MS. The weak bond in the molecule is cleaved. In glycans, fragments where several places are broken are often observed. If a charged glycan labeling agent is used, the ionization efficiency may increase, but only the fragments containing the label are observed and the information volume decreases. Even with a non-charged labeling agent, in the case where there is a large difference between the observation of fragments with and without the labeling agent, the spectra information decreases.

Term 7. XML: Acronym for extensible markup language. A type of computer language to structuralize and describe the text by designating the logical structure and meaning of the text with tags called the markups. Data sharing and processing by program is facilitated by structuralizing and describing the text.

Term 8. Monoisotopic mass: Elements have various natural isotopes. The mass calculated using only the mass of the maximum isotope abundance ratio for each component element of a molecule is the monoisotopic mass.

Term 9. Adduct: Here, this term refers to the additional ions that are generated during the ionization of glycans in MS. There are proton adducts, sodium ion adducts, and others.

References

[1] A. Kobata: *Tosa Kogaku* (Glycotechnology) (Tosa Kogaku Kenkyu Kyogikai, supervised), 3-3, Sangyo chosakai, Tokyo (1992) (in Japanese).

[2] H. Narimatsu: Development of basic tools for glycoscience and their application to cancer diagnosis: A 10-year strategy of the Research Center for Medical Glycoscience of AIST, *Synthesiology*, 5 (3), 190-203 (2012) (in Japanese) [*Synthesiology English edition*, 5 (3), 201-215 (2012)].

[3] J. Hirabayashi: Development of lectin microarray, an advanced system for glycan profiling: From frontal affinity chromatography to evanescent wave excitation fluorescence detection method, *Synthesiology*, 7 (2), 105-117 (2014) (in Japanese) [*Synthesiology English edition*, 7 (2), 105-117 (2014)].

[4] A. Kuno, N. Uchiyama, S. Koseki-Kuno, Y. Ebe, S. Takashima, M. Yamada and J. Hirabayashi: Evanescent-field fluorescence-assisted lectin microarray: a new strategy for glycan profiling, *Nat. Methods*, 2 (11), 851-856 (2005).

[5] A. Kuno, Y. Ikehara, Y. Tanaka, K. Saito, K. Ito, C. Tsuruno, S. Nagai, Y. Takahama, M. Mizokami, J. Hirabayashi and H. Narimatsu: LecT-Hepa: A triplex lectin-antibody sandwich immunoassay for estimating the progression dynamics of liver fibrosis assisted by a bedside clinical chemistry analyzer and an automated pretreatment machine, *Clin. Chim. Acta*, 412 (19-20), 1767-1772 (2011).

[6] A. Matsuda, A. Kuno, H. Matsuzaki, T. Kawamoto, T. Shikanai, Y. Nakanuma, M. Yamamoto, N. Ohkohchi, Y. Ikehara, J. Shoda, J. Hirabayashi and H. Narimatsu: Glycoproteomics-based cancer marker discovery adopting dual enrichment with Wisteria floribunda agglutinin for high specific glyco-diagnosis of cholangiocarcinoma, *J. Proteomics*, 85, 1-11 (2013).

[7] H. Tateno, M. Toyoda, S. Saito, Y. Onuma, Y. Ito, K. Hiemori, M. Fukumura, A. Matsushima, M. Nakanishi, K. Ohnuma, H. Akutsu, A. Umezawa, K. Horimoto, J. Hirabayashi and M. Asashima: Glycome diagnosis of human induced pluripotent stem cells using lectin microarray, *J. Biol. Chem.*, 286 (23), 20345-20353 (2011).

[8] D. Ashline, S. Singh, A. Hanneman and V. N. Reinhold: Congruent strategies for carbohydrate sequencing. 1. Mining structural details by MS^n, *Anal. Chem.*, 77 (19), 6250-6262 (2005).
[9] H. Zhang, S. Singh and V. N. Reinhold: Congruent strategies for carbohydrate sequencing. 2. FragLib: an MS^n spectral library, *Anal. Chem.*, 77 (19), 6263-6270 (2005).
[10] A. J. Lapadula, P. J. Hatcher, A. J. Hanneman, D. J. Ashline, H. Zhang and V. N. Reinhold: Congruent strategies for carbohydrate sequencing. 3. OSCAR: an algorithm for assigning oligosaccharide topology from MS^n data, *Anal. Chem.*, 77 (19), 6271-6279 (2005).
[11] K. Fukui, A. Kameyama, Y. Mukai, K. Takahashi, N. Ikeda, Y. Akiyama and H. Narimatsu: A computational study of structure-reactivity relationships in Na-adduct oligosaccharides in collision-induced dissociation reactions, *Carbohydr. Res.*, 341 (5), 624-633 (2006).
[12] A. Kameyama, S. Nakaya, H. Ito, N. Kikuchi, T. Angata, M. Nakamura, HK. Ishida and H. Narimatsu: Strategy for simulation of CID spectra of N-linked oligosaccharides toward glycomics, *J. Proteome Res.*, 5 (4), 808-814 (2006).
[13] H. Suzuki, A. Kameyama, K. Tachibana, H. Narimatsu and K. Fukui: Computationally and experimentally derived general rules for fragmentation of various glycosyl bonds in sodium adduct oligosaccharides, *Anal. Chem.*, 81 (3), 1108-1120 (2009).
[14] H. Ito, A. Kameyama, T. Sato, K. Kiyohara, Y. Nakahara and H. Narimatsu: Molecular-weight-tagged glycopeptide library: efficient construction and applications, *Angew. Chem. Int. Ed. Engl.*, 44 (29), 4547-4549 (2005).
[15] S. Hase: Pyridylamination as a means of analyzing complex sugar chains, *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.*, 86 (4), 378-390 (2010).
[16] P. J. Domann, A. C. Pardos-Pardos, D. L. Fernandes, D. I. Spencer, C. M. Radcliffe, L. Royle, R. A. Dwek and P. M. Rudd: Separation-based glycoprofiling approaches using fluorescent labels, *Proteomics*, 7 (Suppl 1), 70-76 (2007).
[17] A. Kameyama, N. Kikuchi, S. Nakaya, H. Ito, T. Sato, T. Shikanai, Y. Takahashi, K. Takahashi and H. Narimatsu: A strategy for identification of oligosaccharide structures using observational multistage mass spectral library, *Anal. Chem.*, 77 (15), 4719-4725 (2005).
[18] D. J. Harvey: Collision-induced fragmentation of negative ions from N-linked glycans derivatized with 2-aminobenzoic acid, *J. Mass Spectrom.*, 40 (5), 642-653 (2005).
[19] D. J. Harvey: 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 (5), 622-630 (2005).
[20] M. Wuhrer, C. A. Koeleman, C. H. Hokke and A. M. Deelder: Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments, *Rapid Commun. Mass Spectrom.*, 20 (11), 1747-1754 (2006).
[21] D. J. Harvey: Quantitative aspects of the matrix-assisted laser desorption mass spectrometry of complex oligosaccharides, *Rapid Commun. Mass Spectrom.*, 7 (7), 614-619 (1993).
[22] N. Kikuchi, A. Kameyama, S. Nakaya, H. Ito, T. Sato, T. Shikanai, Y. Takahashi and H. Narimatsu: The carbohydrate sequence markup language (CabosML): an XML description of carbohydrate structures, *Bioinformatics*, 21 (8), 1717-1718 (2005).
[23] A. Kameyama, H. Narimatsu, N. Kikuchi and S. Nakaya: Patent No. 4025850 "Tosa kozo doteihoho oyobi dokaisekisochi" (Method of identifying sugar chain structure and apparatus for analyzing the same), filed Mar 19, 2004, registered Oct 19, 2007 (in Japanese).
[24] A. Kameyama, H. Narimatsu, N. Kikuchi and S. Nakaya: Patent (Germany) No. 112005000598.4 (Method for identification of glycan structure and its analysis device of identifying sugar chain structure and apparatus for analyzing the same), filed Mar 18, 2005, registered Feb 5, 2009 (in German).
[25] A. Kameyama, H. Narimatsu, N. Kikuchi and S. Nakaya: Patent (China) No. ZL200580008027.4 (Method for identification of glycan structure and its analysis device of identifying sugar chain structure and apparatus for analyzing the same), filed Mar 18, 2005, registered Jun 3, 2009 (in Chinese).
[26] A. Kameyama, H. Ito, T. Sato and H. Narimatsu: Database H23PRO-1310 "Glycan Mass Spectral Database ver-2," registered Sep 12, 2011 (in Japanese).
[27] A. Kameyama and N. Kikuchi: Program H23PRO-1311 "CASIS-ver2," registered Sep 14, 2011 (in Japanese).
[28] H. Tateno, A. Matsushima, K. Hiemori, Y. Onuma, Y. Ito, K. Hasehira, K. Nishimura, M. Ohtaka, S. Takayasu, M. Nakanishi, Y. Ikehara, M. Nakanishi, K. Ohnuma, T. Chan, M. Toyoda, H. Akutsu, A. Umezawa, M. Asashima and J. Hirabayashi: Podocalyxin is a glycoprotein ligand of the human pluripotent stem cell-specific probe rBC2LCN, *Stem Cells Transl. Med.*, 2 (4), 265-273 (2013).

Authors

Akihiko KAMEYAMA

Completed the doctor's program at the United Graduate School of Agricultural Science, Gifu University in 1994 (Doctor of Philosophy in Agricultural Science). Worked at the Pharmaceutical Division, Nisshin Oil Mills, Ltd. (currently, Nisshin Oillio Group Ltd.); post-doctoral fellow at the Department of Science, University of Alberta; and Senior Scientist, Amersham Pharmacia Biotech Inc. Became Senior Researcher, AIST in 2005; Team Leader, Research Center for Medical Glycoscience, AIST in 2006; Group Leader, Bioproduction Research Institute, AIST in 2012; and currently, Chief Senior Researcher, Biotechnology Research Institute for Drug Discovery, AIST in April 2015. Visiting Professor, Department of Agriculture, United Graduate School of Agricultural Science, Gifu University since 2013. Has worked continuously on glycan research since university. In this paper, was in charge of writing Subchapter 4.1, Chapters 1~3, 5, 7, and 8, as well as editing the whole paper.

Norihiro KIKUCHI

Completed the master's program at the Department of Applied Biological Science, Graduate School of Science and Technology, Tokyo University of Science in 1999. Obtained Doctor (Medical Science) from Tsukuba University in 2006. General Manager, Personal Genome Department, Mitsui Knowledge Industry Co., Ltd. (until August 2014). Manager, Open Innovation Center, Bayer Yakuhin Ltd. in Japan from September 2014. In this paper, was in charge of writing Subchapter 4.2.

Shuuichi NAKAYA

Completed the master's program at the Graduate School of Frontier Sciences, University of Tokyo in 2002. Assistant Manager, Global Application Development Center, Analytical and Measuring Instrument Division, Shimadzu Corporation. In this paper, was in charge of writing Chapter 6.

Shinji FUNATSU

Graduated from the Faculty of Biotechnology, Fukui Prefectural University in 2001. Assistant Manager, MS Business Unit, Life Science Business Department, Analytical and Measuring Instrument Division, Shimadzu Corporation. In this paper, was in charge of writing Subchapter 4.3.

Discussions with Reviewers

1 Overall

Comment (Tai Kubo, AIST)

Glycomics, or systematic approaches to glycobiology, started far behind proteomics and genomics. It was mainly due to the complexity and diversity of the structure and functions of glycans. However, national projects activated and boosted the glycan research in Japan, and unleashed innovation toward biotechnologies and drug discovery as well. We are now advanced in the field and are leading the world. In the advancement, the researchers at AIST played essential roles and contributed a lot. The authors of this article were in charge of the analysis of glycan structure in the "Structural Glycomics (SG) Project". This paper describes the processes how the authors tackled the encountered problems when they applied mass spectrometry (MS) to analyze the glycan structure, and how they accumulated the knowledge to establish an integrated glycan analysis system with a database. They also created the system with the intellectual property aspect in mind, and it is now open to the public with high utility value. It can be said that this paper is written along the objective of *Synthesiology.*

Comment (Hiroaki Tao, AIST)

This paper is a description of the background, elemental technologies, development scenario, product capability at this point, and future prospect in realizing the outcome, or the product realization of the "glycan analytical system that can analyze the isomers and branch structures." This is necessary for future glycan research and glycotechnology. The specialties of three parties including AIST, an analysis device manufacturer, and an information processing company, as well the strategy to fuse their specialties are aptly described. It contains useful information for researchers and engineers who aim to develop analyzing devices and analytical systems that are necessary in pioneering new academic fields, and I feel the paper is valuable for publication in *Synthesiology.*

2 Research scenario

Comment (Tai Kubo)

The scenario in Chapter 3 is an important point for *Synthesiology.* Please create a schematic diagram that shows the flow from R&D to product realization, including the development of individual elemental technologies described in Chapter 4.

Comment (Hiroaki Tao)

The scenario in Chapter 3 should be described as "3 Scenario for the development of a rapid analytical system for glycans." I think the overall strategy can be understood more readily if you show in a diagram the technologies that the companies already possessed, for example, the human glycan gene cloning and the glycosyltransferase library of AIST and the MALDI-MS of Shimadzu, and the technologies that were newly developed for this system.

Answer (Akihiko Kameyama)

I added "Figure 3. Elemental technologies, their background, and mutual relationships" in Chapter 3.

3 Elemental technology

Comment (Tai Kubo)

In Chapter 4 you describe the processes to solve various problems that you encountered during the development of the glycan structure analysis system, and we can imagine that enormous effort was spent during the process. However, sometimes the explanations are too detailed with specialized terms. Please review the expressions overall with the readers outside of this field in mind. I would suggest that additions of a terminology section and a schematic diagram of the MS method might be helpful for a broad range of readers.

Answer (Akihiko Kameyama)

I added a schematic diagram of the mass spectrometer used in this research to Chapter 2 along with the description. I also added a list of terminology at the end of the paper.

4 Search software and analysis software

Question & Comment (Tai Kubo)

You describe your efforts as you grope toward a solution one by one: for example the reproducible measurement of glycan MS, the structural description language to create a database, the method to extract the glycan peaks in the MS, and so on. All these developments must have been done with universal applications in mind. What is the world standard for search and analysis software? Is there compatibility with databases other than JCGGDB? Please address the comparison and compatibility with other competitors.

Answer (Shuuichi Nakaya)

I don't think there is a world standard for glycan search or analysis software. The commonly used search software is the Glyco Mod Tool that is registered in the ExPASy protein server. It is a tool on the web that can be used free of charge. However, this is not for structural analysis, but it detects the glycan peak in the MS spectra and displays the glycan composition that is estimated from the molecular mass. Recently, there is a payware called SimGlycan (Premier Biosoft) that is becoming popular. This is not a database of actually measured spectra, but is a database of calculated fragments. Other than these, software programs that use mass spectrometry are developed and sold individually by each MS manufacturer. For the search using the interface software, currently, we have no compatibility other than with JCGGDB.

Answer (Akihiko Kameyama)

I think the heart of glycan structure analysis is the identification of isomers. In this case, the database of calculated fragments are useless. I have expressed this point in Subchapter 2.1.

5 Results and their significance

Comment (Hiroaki Tao)

In Subchapter 7.3 "Were we too ahead of our time?," I felt that what you really wanted to say or should do were to describe "the

difficulty encountered by people who pioneer in the forefront, and the will to contribute to glycan research no matter how difficult." Therefore, I think you should describe the impact of introducing this system or an example of effective application.

Answer (Shuuichi Nakaya)

This system has been introduced to overseas drug development companies for the analysis of glycan structural differences in the development of biosimilars.

Answer (Akihiko Kameyama)

I believe the impact of introducing this system and examples of effective application will emerge soon.

6 For the realization of the outcome

Comment (Hiroaki Tao)

In Chapter 8, you describe the research targets to which this system can be applied. I think the readers will be able to readily grasp the overall image of glycan research if you show a diagram that explains this systematically. To do this, I think you should show, for example, the overall image of glycans that represent the glycan types, the glycans to which the system can be applied, and the direction of future R&D for the glycans to which the system cannot be applied.

Answer (Akihiko Kameyama)

I added Fig. 12 "Positioning of this system in the glycoconjugate research and future issues" in Chapter 8.