SHIMADZU

New plug-ins for freely available Mass++ software to identify biomolecules

Kentaro Morimoto[1], Takashi Nishikaze[1], Satoshi Tanaka[1] , Masaki Murase[1] , Shin-ichi Utsunomiya[1] , Shigeki Kajihara[1] , Tsuyoshi Tabata[2] , Ken Aoshima[2] , Yoshiya Oda[2] , Koichi Tanaka[1]
1 Koichi Tanaka Laboratory of Advanced Science and Technology, Shimadzu Corporation, Kyoto, Japan. 2 Eisai Product Creation systems Eisai Co., Ltd, Tsukuba, Japan

Eisai Co., Ltd.

ASMS2013
MP18-358

ms3d FIRST Program

1. Introduction

Mass++ is plug-in type software to analyze mass spectrometric (MS) data and freely available from http://www.first-ms3d.jp/english/. Although Mass++ can read various types of MS data formats, Mass++ didn't have identification tools except MassBank (http://www.massbank.jp/index.html?lang=en) [1] search for metabolite identification. We have developed peptide identification plug-ins "Mascot plug-in" and "X!Tandem plug-in" to utilize already existing search engines such as Mascot and X!Tandem [2]. In addition, we developed original plug-ins

"Glycan Analysis plug-in" to identify *N*-glycans,
"SIMSE plug-in" to identify various compounds by *de novo* sequencing,
"MS^n search plug-in" to identify peptides using MS^n (n > 2) spectra.

We have evaluated those plug-ins performances.

2. Method and result

2-1. Mass++

Mass++ is a plug-in style software. Mass++ has many functions for MS data analysis such as peak detection, smoothing, baseline subtraction, background subtraction, quantitation and so on. We introduce new functions to identify various compounds.

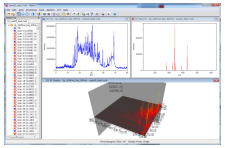


Fig 1. Mass++ main dialog

Target compound	Analysis tool in Mass++
Metabolite	MassBank plug-in
Protein, peptide	Mascot plug-in, X!tandem plug-in, MS^n search plug-in, SIMSE plug-in
Nucleic acid	SIMSE plug-in
Sugar chain	Glycan Analysis plug-in, SIMSE plug-in

Table 1 Plug-in list in Mass++

2-2. Glycan Analysis plug-in

Glycan Analysis plug-in incorporates an iterative algorithm to calculate *m/z* of fragment ions from glycan structural database and to match experimental MS/MS spectra with them. Glycan identification is performed according to the following procedures.

1) calculate theoretical fragment ions for glycan structure in database
2) select "diagnostic ions"[3-5] from theoretical fragment ions
3) match *m/z* of diagnostic ions with experimental MS/MS spectra

When an ion of experimental MS/MS spectrum is the same *m/z* (within a certain variation) as a diagnostic ion, the corresponding glycan structure is assigned to the MS/MS spectrum as a candidate.

Diagnostic ions: fragment ions that are diagnostic of specific structural features and that are not easily obtainable by other means.

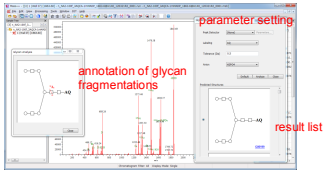
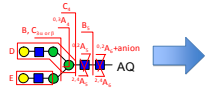


Fig 2. Glycan Analysis plug-in on Mass++



Commercially-available *N*-glycan standards and *N*-glycans released from glycoproteins as human IgG, RNaseB, Fetuin and HER2 were successfully identified by Glycan Analysis plug-in. Negative-ion MS/MS spectra of *N*-glycans were acquired by AXIMA Resonance, MALDI-QIT instrument (Shimadzu/Kratos, U.K.) after on-target 3-aminoquinoline (3AQ) labeling using a liquid matrix [6, 7].

calculated *m/z*	annotation	diagnostic ions for spectral matching
179.06	C_1	
281.09	$^{0,2}A_2$	
382.14	C_2	
424.15	Gal+GlcNAc+59($^{1,3}A_2$)	
466.16	E	x
526.18	B_3	
544.19	C_3	
616.21	$^{0,3}A4$	
670.22	D-H_2O	
688.23	D	x
1233.43	C_4	
1275.44	$^{2,4}A_5$	x
1317.45	$^{0,2}A_5$-H_2O	
1335.46	$^{0,2}A_5$	
1418.50	B_5	
1478.52	$^{2,4}A_6$	x
1520.53	$^{0,2}A_6$-H_2O	
1618.50	$^{0,2}A_6$+anion-H_2O	
1636.51	$^{0,2}A_6$+anion	

Table 2. *m/z* of theoretical fragmentation

2-2. SIMSE plug-in

SIMSE is a *de novo* sequencing tool for not only peptide but also nucleic acids and glycans using dynamic programming algorithm [8]. We applied SIMSE plug-in to identify peptide sequence, nucleic acid sequence and composition of *N*-glycan.

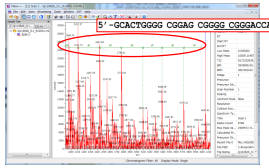


Fig 3-1. SIMSE plug-in on Mass++

Synthetic nucleic acid sequence was determined from ISD spectrum. Glycan composition and partial peptide sequence of *N*-glycopeptide were determined from MS/MS spectrum of tryptic digests of human transferrin. ISD spectrum was acquired by AXIMA Performance, MALDI TOF-TOF instrument and MS/MS spectrum was acquired by AXIMA Resonance.

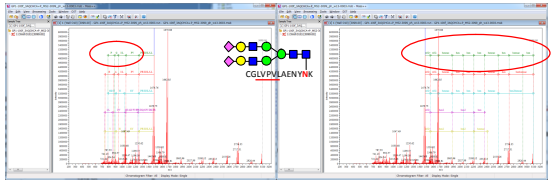


Fig 3-2. SIMSE plug-in on Mass++

SIMSE successfully identified nucleic acids, glycan composition, and a part of peptide sequence.

2-3. MS^n search plug-in

For peptide identification, we use Mascot and X!Tandem plug-ins. When it is difficult to identify peptides from MS/MS spectra because of their poor fragmentation or overlapping precursor ions, we acquire MS^3 spectra and then apply MS^n search plug-in, which can identify peptides with higher confidence.

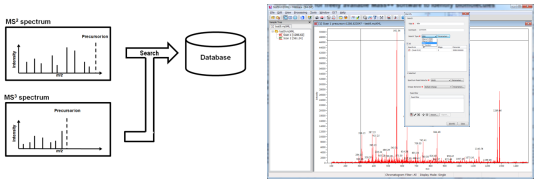


Fig 4. MS^n search plug-in on Mass++

MS^n search is a peptide database search algorithm using MS^n (n > 1) spectra[9]. It has an advantage of identifying a mixture of MS/MS spectra which have close two *m/z* of precursor ions (Fig 5). I applied MSn search for some of mixture peptides which were difficult to identify conventional database search such as Mascot or X!Tandem. We applied database search for trypsin digested proteins, Phosphorylase B, Enolase and Ovalbumin. We chose MS1 peaks which were not identified by conventional database search such as Mascot and X!Tandem using only MS/MS spectra. These peaks were mixture of peptide ions as below.

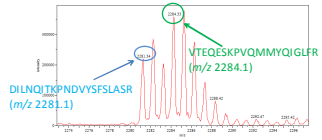


Fig 5. Mixture peak of peptides

Mixture of peptide ions were successfully identified by MS^n search plug-in. MS^n spectra were acquired by AXIMA Resonance.

Protein name	Peptide sequence	*m/z*	E-value (MS^n search)
Phosphorylase B	GYNAQEYYDRIPELR	1886.9	0.13
	LITAIGDVVNHDPVVGDR	1890.0	1.98
Phosphorylase B	LLSYVDDEAFIR	1440.7	0.0012
	VLYPNDNFFEGK	1442.6	2.35
Enolase	VNQIGTLSESIK	1286.7	0.19
	NVNDVIAPAFVK	1288.7	1.71
Ovalbumin	DILNQITKPNDVYSFSLASR	2281.1	0.0063
	VTEQESKPVQMMYQIGLFR	2284.1	0.21

Table 3. Identified peptides using MS^n search plug-in

Table3 shows results of MS^n search. Because these spectra were complicated by mixing peptide ions, conventional database search doesn't work well.

3. Conclusions

We developed and evaluated glycan structure database search, *de novo* sequencing for nucleic acid and peptide database search using MS^n data. Users can analyze various compounds using only Mass ++ and its rich plug-ins.
Mass++ has following two advantage.
1)various format reading
2)various compound identification
Therefore, Mass++ will support every user related to mass spectrometry.

Reference

[1] Horai et al, *J. Mass Spectrom*., 45, 703-714 (2010)
[2] Craig et al, *Bioinformatics*, 20, 1466-1467 (2004)
[3] Harvey, D. J. *J. Am. Soc*. Mass Spectrom 2005, 16, 622-630
[4] Harvey, D. J. *J. Am. Soc*. Mass Spectrom 2005, 16, 631-646
[5] Harvey, D. J. *J. Am. Soc*. Mass Spectrom 2005, 16, 647-659
[6] Kaneshiro et al, *Anal. Chem*. 83, 3663-3667 (2011)
[7] Nishikaze et al, *Anal. Chem*. 84, 9453-9461 (2012)
[8] Murase et al, ASMS 58th , TP148
[9] Morimoto et al, ASMS 59th , MP445

Acknowledgement

This research is granted by the Japan Society for Promotion of Science (JSPS) through its "Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)" initiated by the Council for Science and Technology Policy (CSTP).