



N-GlycoMiner

User Guide



CATALOG

1

Mouse N-glycoproteomics Search

- Tissue-specific
- Spatial-temporal and disease-related Data

2

Literature-Integrated Glycoprotein

- Query
- Comparative disease glycoproteomics

3

Theoretical Plate

- N-Glycopeptides Pool

NGlycoMiner

Home

Mouse N-glycoproteome Search

Literature-Integrated Glycoprotein

Theoretical Plate

Tissue-specific

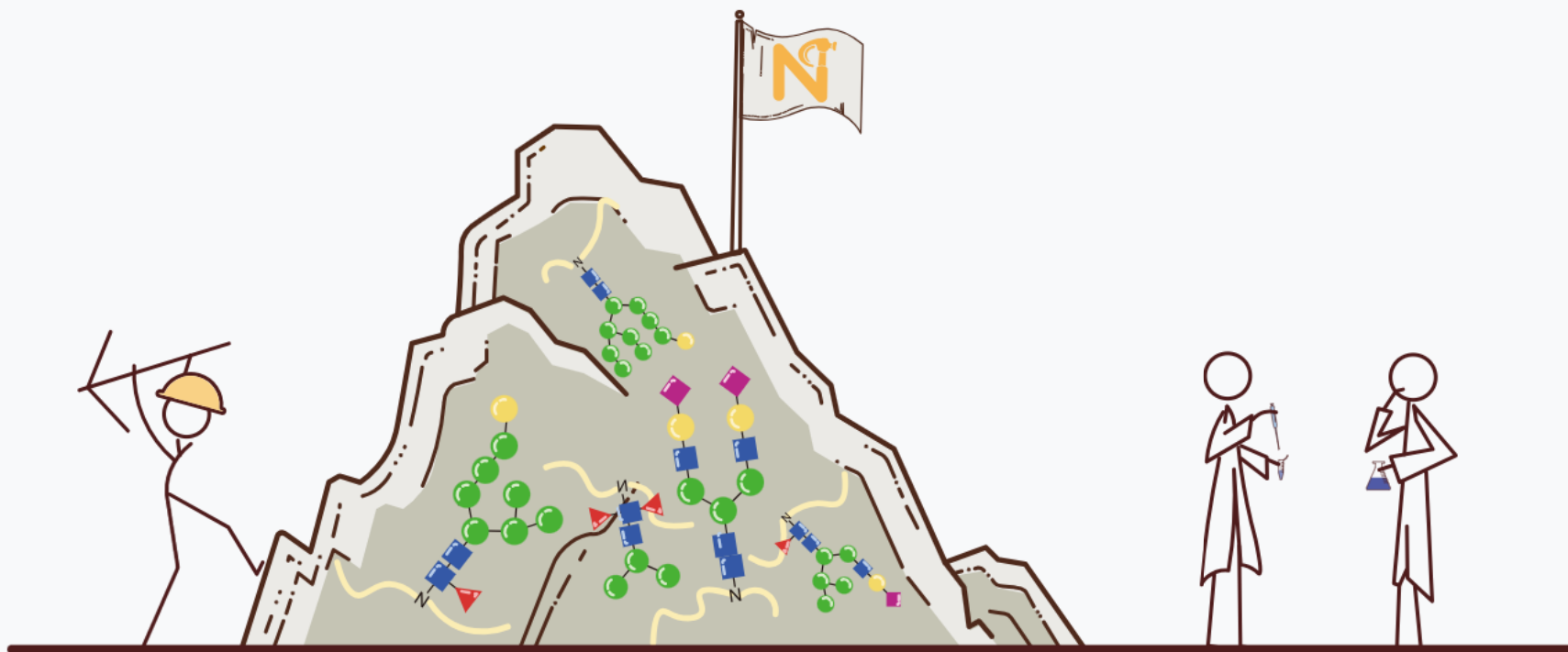
Query

N-Glycopeptides Pool

Spatial-temporal & disease-related
Data

Comparative Disease
Glycoproteomics

WELCOME TO NGLYCOMINER!





CATALOG

1

Mouse N-glycoproteomics Search

- Tissue-specific
- Spatial-temporal and disease-related Data

2

Literature-Integrated Glycoprotein

- Query
- Comparative disease glycoproteomics

3

Theoretical Plate

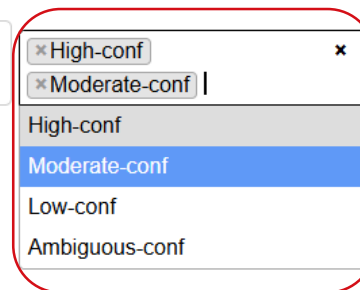
- N-Glycopeptides Pool

Integrated Mouse N-Glycoproteomic Data Query

Type UniprotID/Gene Name/Organ name here

提交

Enter the UniprotID/Gene Name/Organ name to search



A dropdown menu for selecting confidence levels. It has a red border and a close button (x) in the top right corner. The menu is currently open, showing a list of options: 'High-conf', 'Moderate-conf', 'Low-conf', and 'Ambiguous-conf'. The 'Moderate-conf' option is highlighted with a blue background. There is a small vertical line to the right of the 'Moderate-conf' text in the dropdown.

×High-conf	×
×Moderate-conf	
High-conf	
Moderate-conf	
Low-conf	
Ambiguous-conf	

Choose confidence: The level of confidence is determined by the results identified by the four software tools(pGlyco3、StrucGP、MSFragger-Glyco、Glyco-Decipher)

High-Conf : Identified as the same ID by 4 software tools

Moderate-Conf : Identified as the same by 2&3 software tools

Low-Conf : a spectrum only identified by one

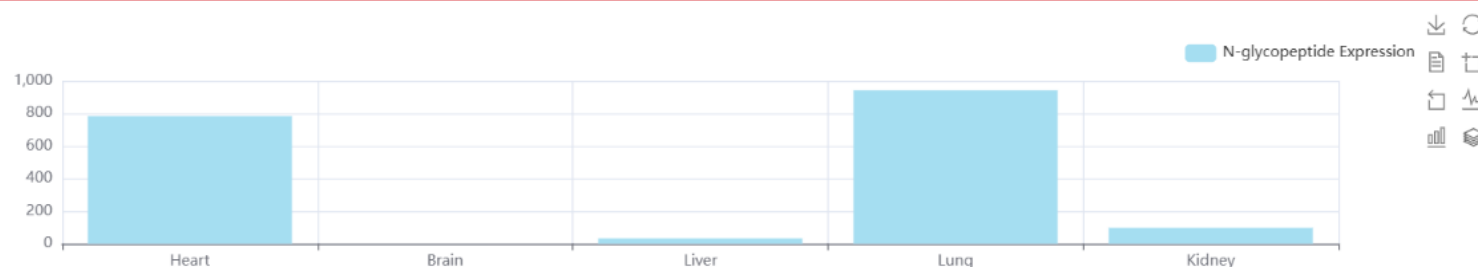
Ambiguous ID : IDs differ across multiple software tools

Integrated Mouse N-Glycoproteomic Data Query

Take CD36, for example

Accession ID: >sp|Q08857|CD36_MOUSE

Tissue specific:



Tissue-specific
distribution of
CD36 protein

Enzyme: **Trypsin+Glu-C** **Trypsin** **Trypsin+Lys-C**

Enrichment Methods: **Sepharose** **ZIC-HILIC**

The digestion method used in the experiment

The enrichment method used in the experiment

[Check it in Uniprot](#)[Complete IDs source file Download](#)

> Here, we will display the top 15 IDs identified for this protein respectively. Detailed Information can be downloaded above.

For this protein, the specific spectrum scan number can be downloaded.

[Spectrum scan number Download](#)

Check in Uniprot: Jump to the Uniprot information page for the corresponding protein
Complete IDs source file Download and Spectrum scan number Download: Download the full ID file and spectrum scan number associated with CD36 for further analysis

The first 15 IDs identified for this protein, Sort by number of spectra

Precursors Proteins Glycoforms Glycosites

UniprotID	Glycosite	Peptide	GlycanComposition	Charge	Modification	Heart	Liver	Brain	Lung	Kidney	Trypsin	Trypsin+Lys-C	Trypsin+Glu-C	ZIC-HILIC	Sepharose	pGlyco3	StrucGP	Glyco_decipher	MSFragger_Glyco
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEK	H(9)N(2)	3	nan	323	24	0	329	34	416	294	0	546	164	710	710	710	710
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEK	H(9)N(2)	4	nan	38	1	0	118	20	81	96	0	100	77	177	177	177	177
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEK	H(8)N(2)	3	nan	36	7	0	79	7	54	75	0	75	54	129	129	129	129
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(9)N(2)	3	nan	43	0	0	78	1	55	22	45	41	81	122	122	122	122
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEKAEMFK	H(9)N(2)	4	nan	42	1	0	61	12	84	32	0	39	77	116	116	116	116
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEKAEMFK	H(9)N(2)	4	23Oxidation;	39	0	0	58	8	76	29	0	42	63	105	105	105	105
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(8)N(2)	3	nan	34	0	0	46	0	36	13	31	31	49	80	80	80	80
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(10)N(2)	3	nan	24	0	0	45	0	29	15	25	14	55	69	69	69	69
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(9)N(2)	2	nan	30	0	0	15	0	25	1	19	28	17	45	45	45	45
>sp Q08857 CD36_MOUSE	79	QFWIFDVQNPDDVAKNSSK	H(7)N(4)F(1)	3	nan	12	0	0	28	0	26	14	0	7	33	40	40	40	40
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(8)N(2)	2	nan	19	0	0	7	0	16	0	10	13	13	26	26	26	26
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEK	H(10)N(2)	3	nan	7	0	0	11	4	12	10	0	19	3	22	22	22	22
>sp Q08857 CD36_MOUSE	79	QFWIFDVQNPDDVAKNSSK	H(5)N(4)A(2)F(1)	3	nan	17	0	0	0	0	17	0	0	9	8	17	17	17	17
>sp Q08857 CD36_MOUSE	417	NLKRPYIVPILWLNETGTIGDEK	H(9)N(2)	4	nan	15	0	0	1	0	16	0	0	10	6	16	16	16	16
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEKAEMFK	H(9)N(2)	3	nan	12	0	0	3	0	15	0	0	13	2	15	15	15	15

The numbers in the table represent the corresponding identified spectrum counts.

[See this protein in curated dataset](#)

Protein Sequence

Hide Sequence

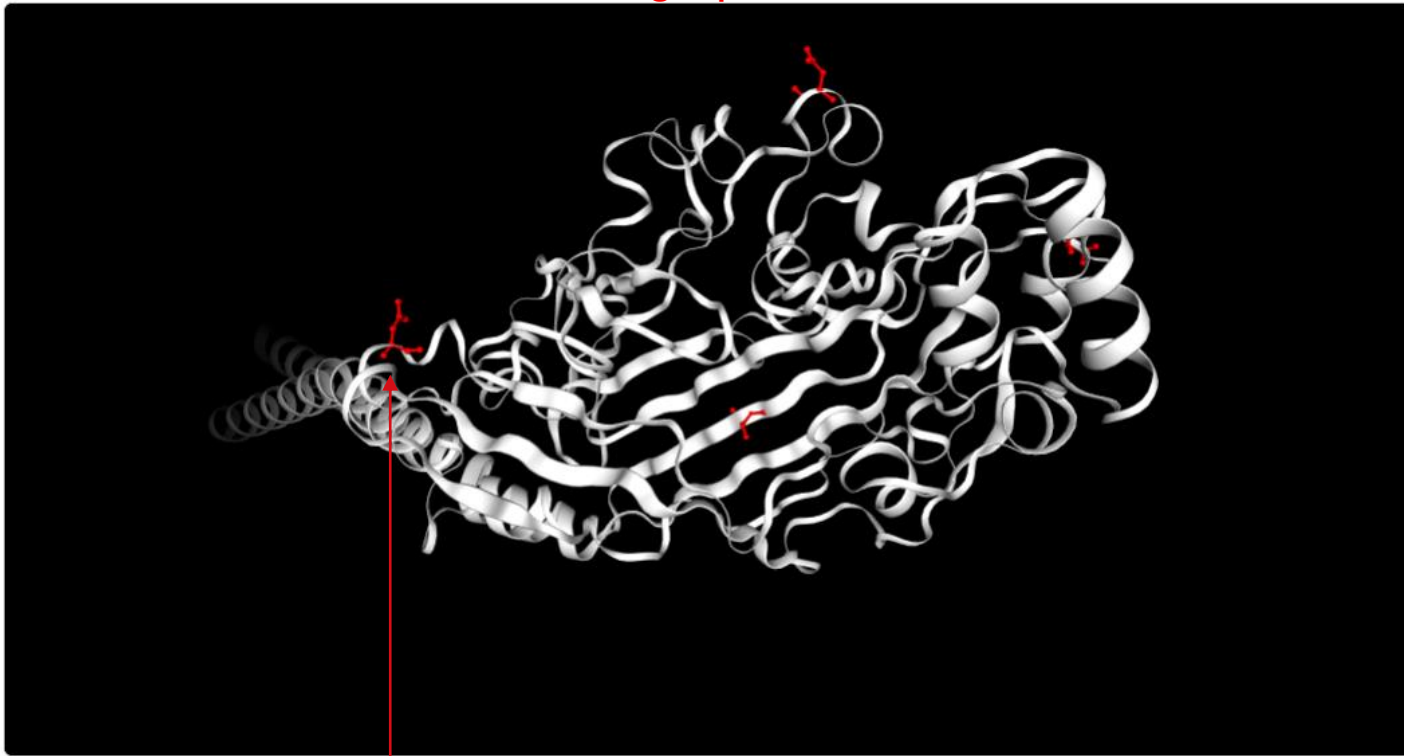
10	20	30	40	50	60	70	80	90	100	110
MGCDRNCGLI	AGAVIGAVLA	VFGGILMPVG	DMLIEKTIKR	EVVLEEGTTA	FKNWVKTGTT	VYRQFWIFDV	QNPDDVAKNS	SKIKVKQRGP	YTYRVRYLAK	ENITQDPEDH
						QFWIFDV	QNPDDVAKNS	SKIK	YLAK	ENITQDPE
120	130	140	150	160	170	180	190	200	210	220
TVSFVQPNGA	IFEPSLSVGT	EDDFTVLNL	AVAAAPHIYQ	NSFVQVVLNS	LIKKSSSMF	QTRSLKELLW	GYKDPFLSLV	PYPISTTVGV	FYPYNDTVDG	VYKVFNGKDN
						SLKELLW	GYKDPFLSLV	PYPISTTVGV	FYPYNDTVDG	VYKVFNGKDN
230	240	250	260	270	280	290	300	310	320	330
ISKVAIIESY	KGKRNL SYWP	SYCDMINGTD	AASFPPFVEK	SRTLRFSSD	ICRSIYAVFG	SEIDLKGIPV	YRFVLPANAF	ASPLQNPDNH	CFCTEKVISN	NCTSYGVLDI
ISKVAIIESY	KGKRNL SYWP	SYCDMINGTD	AASFPPFVEK	SR		GIPV	YRFVLPANAF	ASPLQNPDNH	CFCTEKVISN	NCTSYGVLDI
340	350	360	370	380	390	400	410	420	430	440
GKCKEGKPVY	ISLPHFLHAS	PDVSEPIEGL	HPNEDEHRTY	LDVEPITGFT	LQFAKRLQVN	ILVKPARKIE	ALKNLKRPYI	VPILWLNETG	TIGDEKAEMF	KTQVTGKIKL
GKCK							NLKRPII	VPILWLNETG	TIGDEKAEMF	K
450	460	470	#							
LGMVEMALLG	IGVVMFVAFM	ISYACKSKN	GK							

Identified N-glycosite

Potential N-glycosites not experimentally Verified

Identified N-Glycopeptide sequence

3D infographic of CD36



The atomic structure of Asn were showed in red as the identified N-glycosite

N-GlycoSite

Select a Glycosite ▼

- Highlight 321
- Highlight 417
- Highlight 220
- Highlight 79

GlycanComposition

● Hex ■ HexNAc ▲ Fuc ◆ NeuAc ◆ NeuGc

Check it in AlphaFold PDB for more details

Go to the AlphaFold PDB page

Identified N-glycosite

Potential N-glycosites not experimentally Verified

Identified N-Glycopeptide sequence



N-GlycoSite

Highlight 321

Select a Glycosite

Highlight 321

Highlight 417

Highlight 220

Highlight 79

Load PDB for more details

GlycanComposition

● Hex ■ HexNAc ▲ Fuc ◆ NeuAc ◇ NeuGc

● x7 ■ x4 ▲ x1

● x5 ■ x4 ◆ x2 ▲ x1

● x5 ■ x4 ◆ x1 ◇ x1 ▲ x1

After selecting the glycosylation site, the composition of the glycan chain can be seen on the right

PrecursorsProteinsGlycoformsGlycosites

UniprotID	Glycosite	Peptide	GlycanComposition	Charge	Modification	Heart	Liver	Brain	Lung	Kidney	Trypsin	Trypsin+Lys-C	Trypsin+Glu-C	ZIC-HILIC	Sepharose	pGlyco3	StrucGP	Glyco_decipher	MSFragger_Glyco
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEK	H(9)N(2)	3	nan	323	24	0	329	34	416	294	0	546	164	710	710	710	710
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEK	H(9)N(2)	4	nan	38	1	0	118	20	81	96	0	100	77	177	177	177	177
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEK	H(8)N(2)	3	nan	36	7	0	79	7	54	75	0	75	54	129	129	129	129
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(9)N(2)	3	nan	43	0	0	78	1	55	22	45	41	81	122	122	122	122
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEKAEMFK	H(9)N(2)	4	nan	42	1	0	61	12	84	32	0	39	77	116	116	116	116
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEKAEMFK	H(9)N(2)	4	23Oxidation;	39	0	0	58	8	76	29	0	42	63	105	105	105	105
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(8)N(2)	3	nan	34	0	0	46	0	36	13	31	31	49	80	80	80	80
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(10)N(2)	3	nan	24	0	0	45	0	29	15	25	14	55	69	69	69	69
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(9)N(2)	2	nan	30	0	0	15	0	25	1	19	28	17	45	45	45	45
>sp Q08857 CD36_MOUSE	79	QFWIFDVQNPDDVAKNSSK	H(7)N(4)F(1)	3	nan	12	0	0	28	0	26	14	0	7	33	40	40	40	40
>sp Q08857 CD36_MOUSE	220	VFNGKDNISK	H(8)N(2)	2	nan	19	0	0	7	0	16	0	10	13	13	26	26	26	26
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEK	H(10)N(2)	3	nan	7	0	0	11	4	12	10	0	19	3	22	22	22	22
>sp Q08857 CD36_MOUSE	79	QFWIFDVQNPDDVAKNSSK	H(5)N(4)A(2)F(1)	3	nan	17	0	0	0	0	17	0	0	9	8	17	17	17	17
>sp Q08857 CD36_MOUSE	417	NLKRPYIVPILWLNETGTIGDEK	H(9)N(2)	4	nan	15	0	0	1	0	16	0	0	10	6	16	16	16	16
>sp Q08857 CD36_MOUSE	417	RPYIVPILWLNETGTIGDEKAEMFK	H(9)N(2)	3	nan	12	0	0	3	0	15	0	0	13	2	15	15	15	15

The numbers in the table represent the corresponding identified spectrum counts.

See this protein in curated dataset

Jump to the curated dataset page for this protein (From literature database) . Please see next page.

The curated dataset page of CD36

Accession ID: [>sp|Q08857|CD36_MOUSE](#)

Protein full-name Description: Platelet glycoprotein 4

N-Glycosite(8 in total):

N79; N205; N220; N235; N247; N320; N321; N417

This protein has 301 unique Glycoforms identified.

[All Glycoforms of this protein Download](#)

Download all glycoform information for CD36

#All types of Glycoforms are shown below

Protein	Glycosite	GlycanComposition	peptide	peptide_cal
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(2)Hex(2)Fuc(1)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(2)Hex(5)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(2)Hex(6)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(3)Hex(3)Fuc(2)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(3)Hex(5)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(4)Hex(3)Fuc(3)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(4)Hex(5)NeuGc(2)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(4)Hex(6)NeuGc(1)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(4)Hex(6)NeuGc(1)Fuc(1)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43
>sp Q08857 CD36_MOUSE@Cd36	205	HexNAc(4)Hex(7)NeuGc(1)	DPFLSLVPYPISTTVGVFYPYNDTVDGVYK	43

Page 1 of 31. [Next >>](#) [Last »](#)

[See this protein in our dataset with high-confidence](#)

Go back to the page you just saw



CATALOG

1

Mouse N-glycoproteomics Search

- Tissue-specific
- Spatial-temporal and disease-related Data

2

Literature-Integrated Glycoprotein

- Query
- Comparative disease glycoproteomics

3

Theoretical Plate

- N-Glycopeptides Pool

Spatial-temporal and disease-related N-glycoproteome

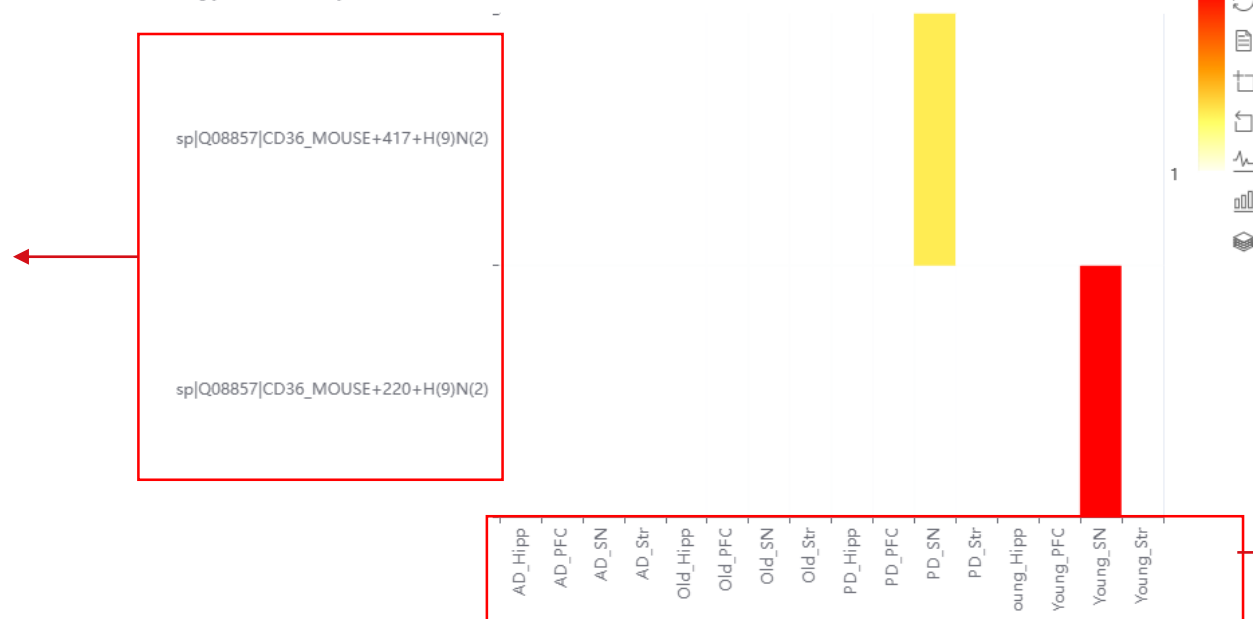
Let's take CD36_MOUSE
as an example

Accession ID: sp|Q08857|CD36_MOUSE

Spatiotemporal expression of N-glycoform

Normalized glycoform intensity

Glycoforms
identified in
CD36 under
different
conditions



The color indicates the
numerical range of
normalized intensities:
from yellow (low
expression) to red (high
expression).

Samples or tissues under
different experimental
conditions involve
different time, space, and
disease states

Protein Sequence

Hide Sequence

Protein sequence and glycosylation site information

10	20	30	40	50	60	70	80	90	100	110
MGCDRNCGLI	AGAVIGAVLA	VFGGILMPVG	DMLIEKTIKR	EVVLEEGTTA	FKNWVKTGTT	VYRQFWIFDV	QNPDDVAKNS	SKIKVKQRGP	YTYRVRYLAK	ENITQDPEDH
120	130	140	150	160	170	180	190	200	210	220
TVSFVQPNGA	IFEPSLSVGT	EDDNFTVLNL	AVAAAPHIYQ	NSFVQVVLNS	LIKKS KSSMF	QTRSLKELLW	GYKDPFLSLV	PYPIS TTVGV	FYPYNDTVDG	VYKVFNGKDN
230	240	250	260	270	280	290	300	310	320	330
ISKVAIIIESY	KGKRNL SYWP	SYCDMINGTD	AASFPPFVEK	SRTLRFSSD	ICRSIYAVFG	SEIDLKGIPV	YRFVLPANAF	ASPLQNPDNH	CFCTEKVISN	NCTSYGVLDI
340	350	360	370	380	390	400	410	420	430	440
GKCKEGKPVY	ISLPHFLHAS	PDVSEPIEGL	HPNEDEHRTY	LDVEPITGFT	LQFAKRLQVN	ILVKPARKIE	ALKNLKRPYI	VPILWLNETG	TIGDEKAEMF	KTQVTGKIKL
450	460	470								
LGMVEMALLG	IGVVMFVAFM	ISYCAKSKN	GK							

 Identified N-glycosite

 Potential N-glycosites not experimentally Verified



CATALOG

1

Mouse N-glycoproteomics Search

- Tissue-specific
- Spatial-temporal and disease-related Data

2

Literature-Integrated Glycoprotein

- Query
- Comparative disease glycoproteomics

3

Theoretical Plate

- N-Glycopeptides Pool

Curated N-glycoproteome data

N-Glycoproteomic information of Homo sapiens\Mus musculus\Fission yeast can be searched

OR QUERY THROUGH SPECIES / SAMPLE / SPECIFIC CONDITION WELL DOCUMENTED 

Curated N-glycoproteome data

Homo sapiens ▼

提交

OR QUERY THROUGH SPECIES / SAMPLE / SPECIFIC CONDITION WELL DOCUMENTED 

The corresponding protein was queried by selecting the species\sample\specific conditions

Homo sapiens ▼	Tissue ▼	ovarian tumor ▼	提交
Select species	Select Type	Select Status	
Homo sapiens	Tissue	ovarian tumor & non-tumor tissue samples	
Mus musculus	Human milk	RA patients & knee OA patients (Fresh knee joint synovial tissues)	
Fission yeast	plasma	Alzheimer's Disease(whole frontal cortexfrom; 10 each with no pathologic AD & 10 each with/without cognitive impairment who had pathologic AD)	
	serum	patients with intrahepatic cholangiocarcinoma (n = 6) and hepatocellular carcinoma	
	Cell line	OSCC (lymph node metastasis (N+) (n=19) & without lymph node metastasis (N0) (n=12))	
	Blood	Pancreatic cancer (pancreatic tissues)	
	Semen	OSCC(from patients with (n = 19) and without (n = 12) lymph node metastasis)	
	Urine	(Alzheimer's Disease)Human brain tissue	
	Standard Glycoprotein	Breast Cancer	
	All types		

[Download Protein List](#)

The list of eligible proteins is available for viewing and downloading

Protein ID	Description
>sp A6NMZ7 CO6A6_HUMAN@COL6A6	Collagen alpha-6(VI) chain
>sp O00469 PLOD2_HUMAN@PLOD2	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2
>sp O00481 BT3A1_HUMAN@BTN3A1	Butyrophilin subfamily 3 member A1
>sp O14498 ISLR_HUMAN@ISLR	Immunoglobulin superfamily containing leucine-rich repeat protein
>sp O14672 ADA10_HUMAN@ADAM10	Disintegrin and metalloproteinase domain-containing protein 10
>sp O14773 TPP1_HUMAN@TPP1	Tripeptidyl-peptidase 1
>sp O15031 PLXB2_HUMAN@PLXNB2	Plexin-B2
>sp O15230 LAMA5_HUMAN@LAMA5	Laminin subunit alpha-5
>sp O15460 P4HA2_HUMAN@P4HA2	Prolyl 4-hydroxylase subunit alpha-2
>sp O60568 PLOD3_HUMAN@PLOD3	Multifunctional procollagen lysine hydroxylase and glycosyltransferase LH3
>sp O75051 PLXA2_HUMAN@PLXNA2	Plexin-A2
>sp O75339 CILP1_HUMAN@CILP	Cartilage intermediate layer protein 1
>sp O75477 ERLN1_HUMAN@ERLIN1	Erlin-1
>sp O75629 CREG1_HUMAN@CREG1	Protein CREG1
>sp O75718 CRTAP_HUMAN@CRTAP	Cartilage-associated protein
>sp O75882 ATRN_HUMAN@ATRN	Attractin
>sp O65202 FKBP5_HUMAN@FKBP5	Peptidyl prolyl cis-trans isomerase FKBP5

Comparative Glycoproteome insights into specific diseases

Condition List:

Dengue patients ▼

Generate Scatter

Dengue patients

colostrum and mature milk comparison

MCF-7 cancer stem cells

Alzheimer's Disease

Pancreatic cancer

hepatocellular carcinoma & intrahepatic cholangiocarcinoma

type 2 diabetes

prediabetic (PDB) subject

high-grade serous ovarian adenocarcinoma

oral squamous cell carcinoma

Rheumatoid arthritis & osteoarthritis

atopic disease(Farm Effect)

Comparative Glycoproteome insights into specific diseases

Condition List:

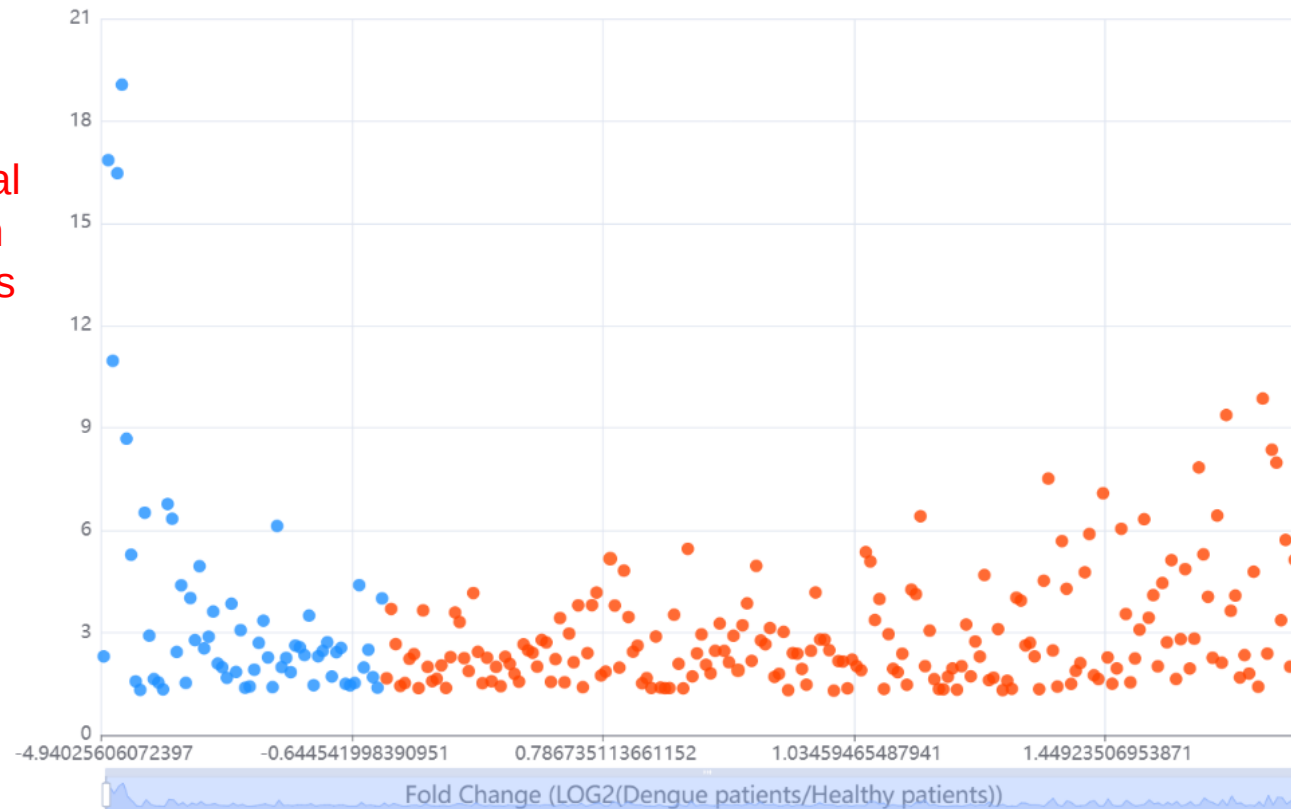
[Generate Scatter](#)

Resource:

Significant changes of glycoforms in dengue patients compared to healthy patients from curated literatures

Scatter Plot for Dengue patients

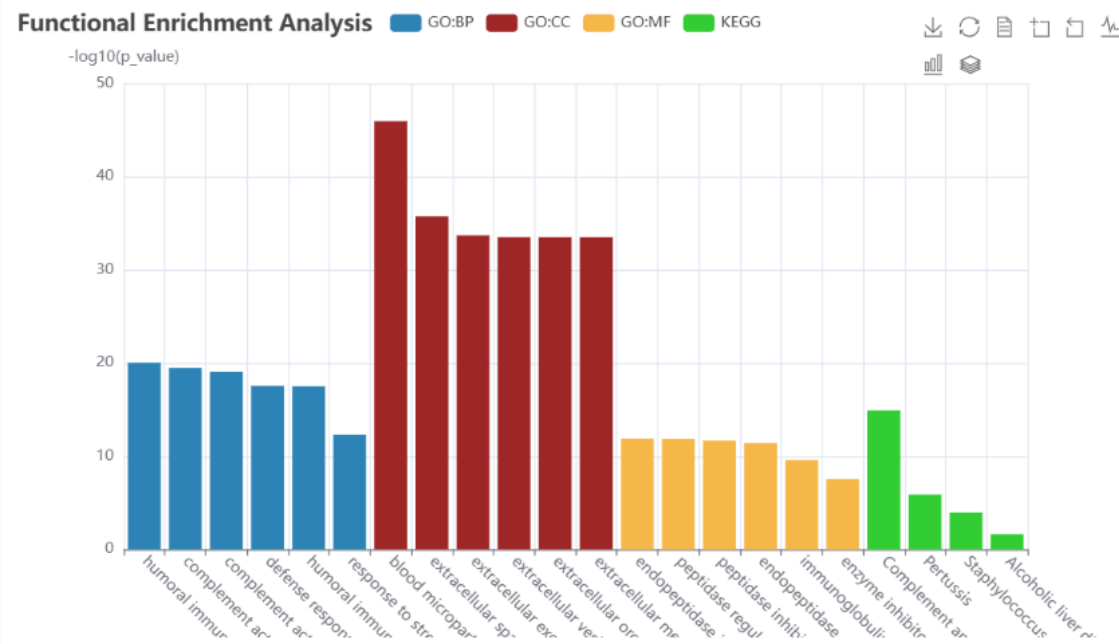
-LOG10(P Value)



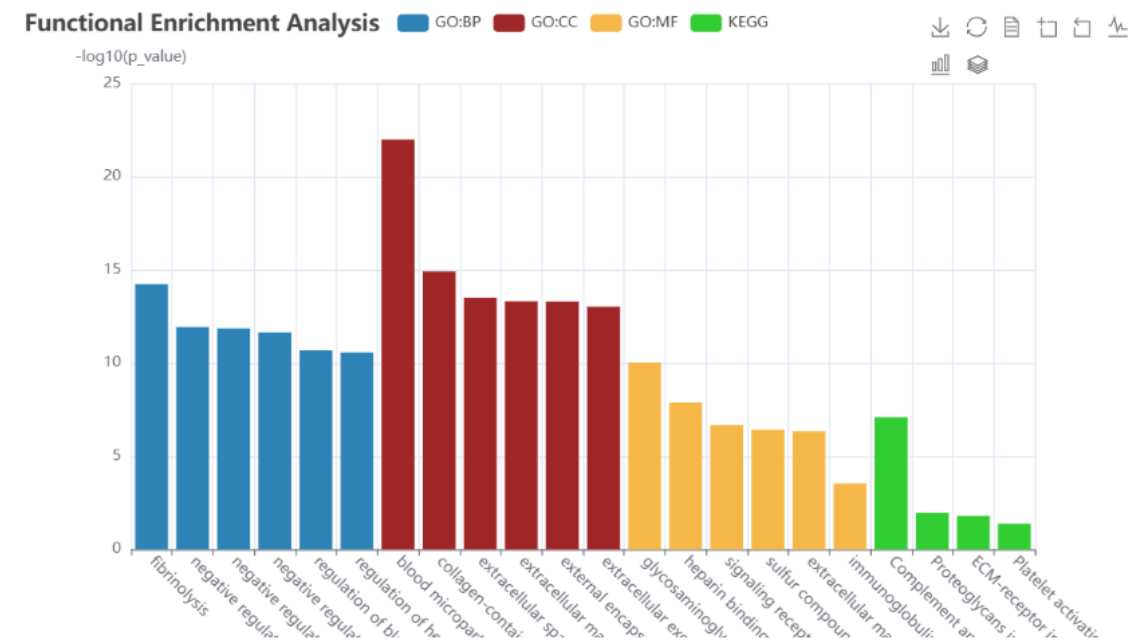
Y-axis: Indicates the statistical significance of the changes in glycoforms, with higher values indicating higher significance.

X-axis: Indicates fold changes in glycoform abundances in dengue patients relative to healthy patients, taking the logarithmic value (Log2).

Up regulated



Down regulated



References:

38949184

Protein	GlycoForm	Note
P02675	P02675@FGB+394+HexNAc(4)Hex(6)Fuc(1)	Down_regulated
P02675	P02675@FGB+394+HexNAc(4)Hex(5)	Down_regulated
P02679	P02679@FGG+78+HexNAc(4)Hex(4)NeuAc(1)	Down_regulated
P02763	P02763@ORM1+93+HexNAc(4)Hex(5)NeuAc(2)	Up_regulated
P02679	P02679@FGG+78+HexNAc(4)Hex(5)	Down_regulated
P02675	P02675@FGB+394+HexNAc(4)Hex(5)NeuAc(1)	Down_regulated

Theoretical Glycopeptides Analysis

Upload FASTA file and it will generate a series of Glycoproteomic analysis charts automatically (**Client Maximum file size: 30MB**)

Upload FASTA file: Q08857_CD36_MOUSE.fasta

☒ **Enzyme:**

☐ **Custom Enzyme:**

Note: input enzyme cleavage direction and sites separated by _ . Use / to separate multiple sites. e.g. C_K/R represents cleavage at K/R on the C-terminal

Max Missed Cleavage Sites:

most common Variable modification

☒ **Methionine oxidation:**

☒ **Acetylation on protein N-term:**

Note, We add Carbamidomethyl as fixed modification at amino acid C

Get Started!



- Trypsin
- Pepsin
- Glu-c
- Chymotrypsin
- Asp-N
- Trypsin+Pepsin
- Trypsin+Glu-c
- Trypsin+Chymotrypsin
- Trypsin+Asp-N

Theoretical analysis tools for N-glycosylated peptides:
By uploading the FASTA file (still using CD36_MOUSE as an example, uniprot id: Q08857), select the enzyme digestion method, Max Missed Cleavage Sites and Variable Modifications for theoretical simulation

N-Glycopeptides Pool

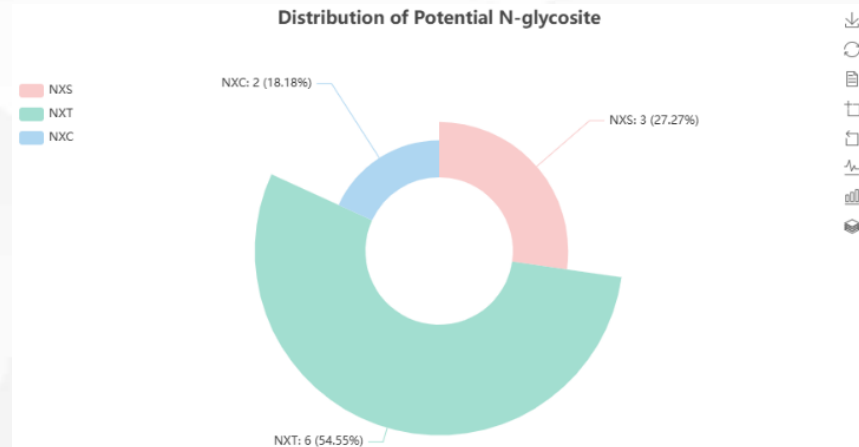
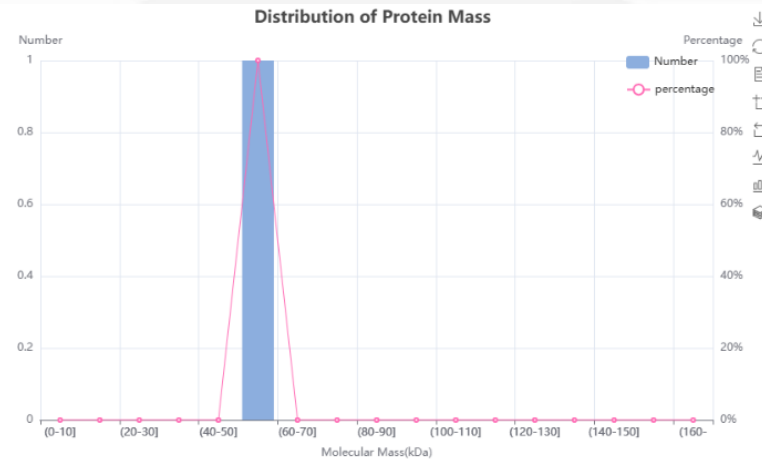
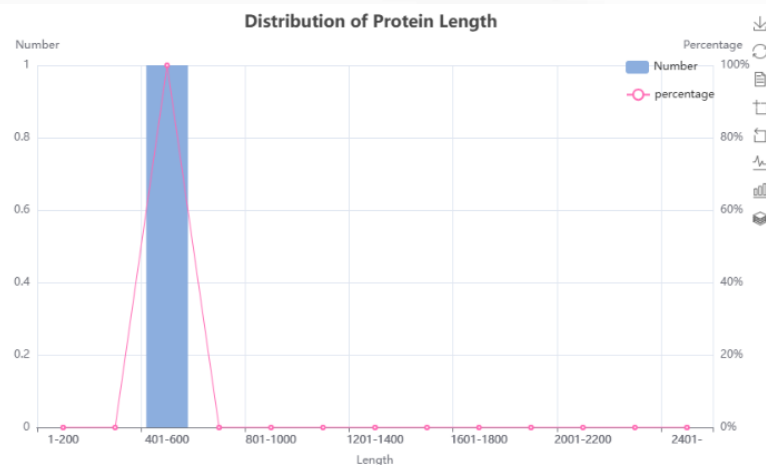
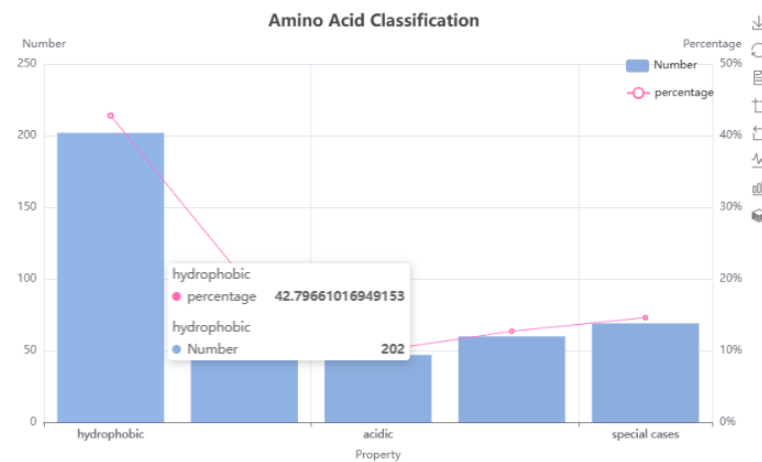
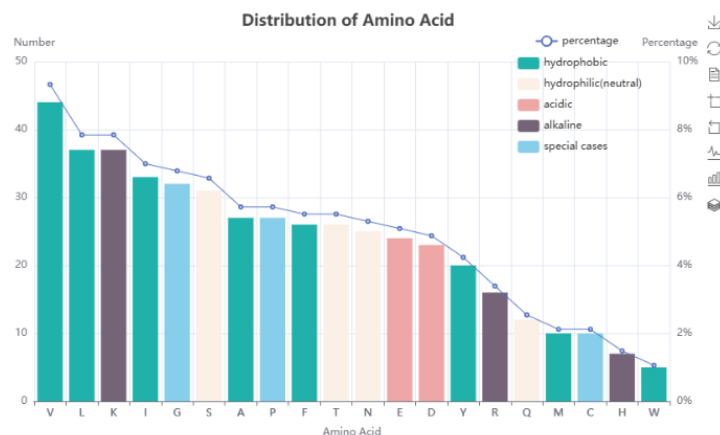
We only consider the primary sequence composition of the proteome

The uploaded FASTA file contains a total of 472 amino acids, and their respective numbers and proportions are shown in the following figures.

Our classification criteria for the properties of amino acids are:

- hydrophobic : A,V,I,L,M,F,Y,W
- hydrophilic(neutral) : S,T,N,Q
- acidic : D,E
- alkalinity : K,R,H
- special_cases : C,U,P,G

The primary structure information of the protein is shown in the table



Theoretical Peptide Library Generation Progress

100%

[Continue to see theoretical peptides](#)

After loading 100%, jump to the theoretical peptide page

We only consider the primary sequence composition of the proteome

The uploaded FASTA file contains a total of 472 amino acids, and their respective numbers and proportions are shown in the following figures.

Our classification criteria for the properties of amino acids are:

- hydrophobic : A,V,I,L,M,F,Y,W
- hydrophilic(neutral) : S,T,N,Q
- acidic : D,E
- alkalinity : K,R,H
- special_cases : C,U,P,G

Distribution of Amino Acid



[Nglycopeptide analysis](#)
[Download Data](#)

See the next page for glycopeptide analysis

If the uploaded FASTA file is too large, it may take a few minutes of waiting time.

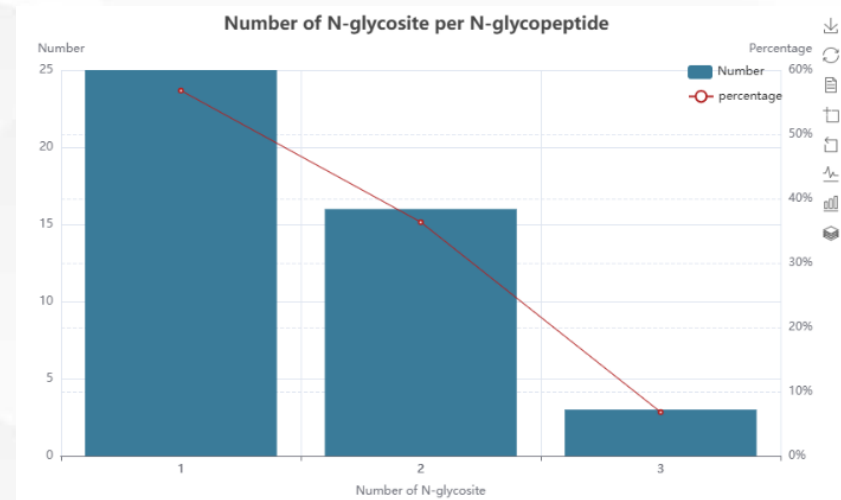
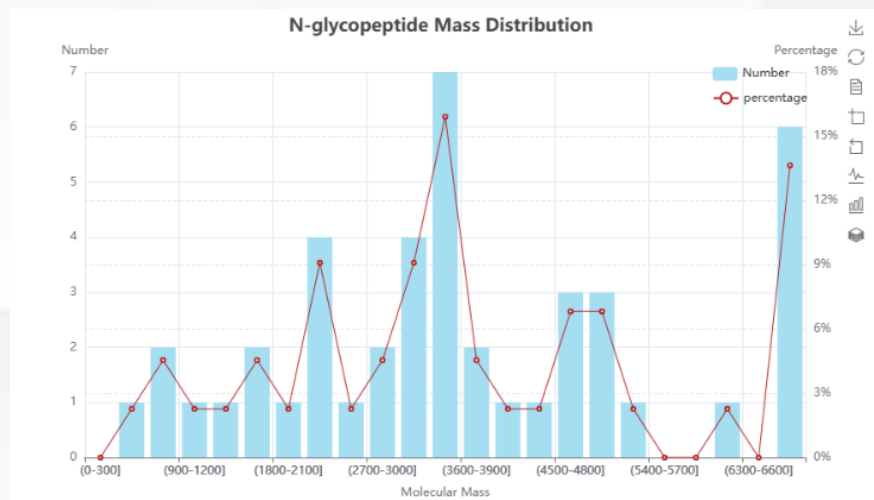
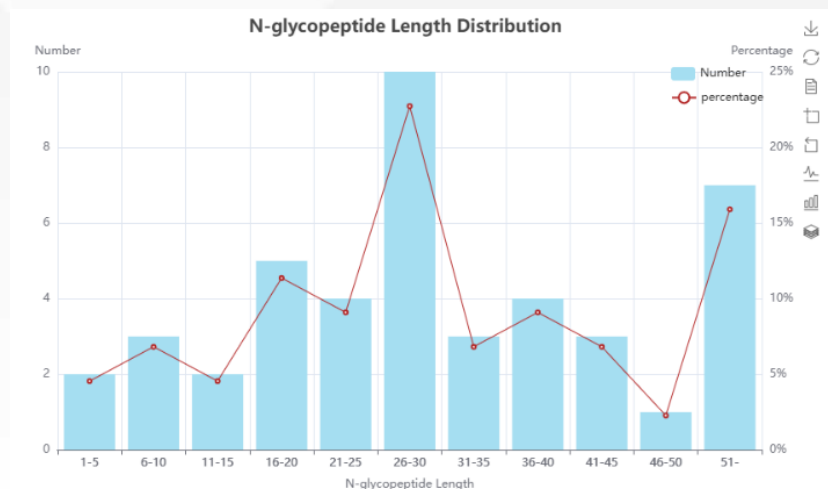
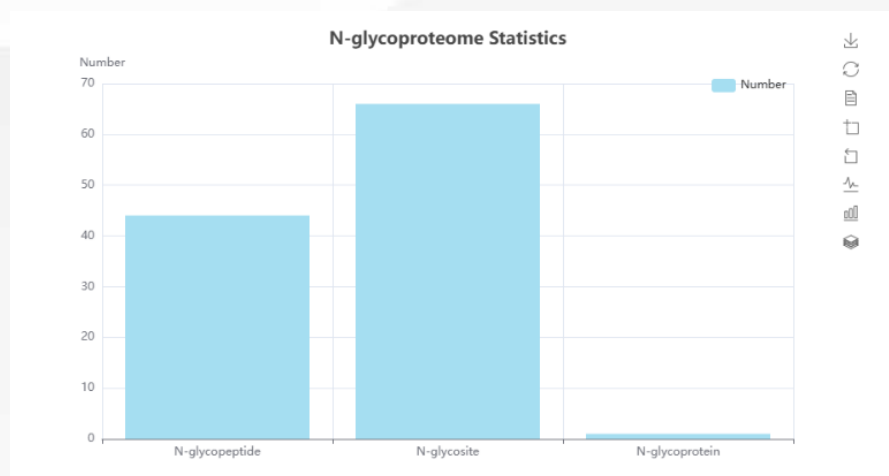
Theoretical N-glycopeptides table display

44 peptides in total

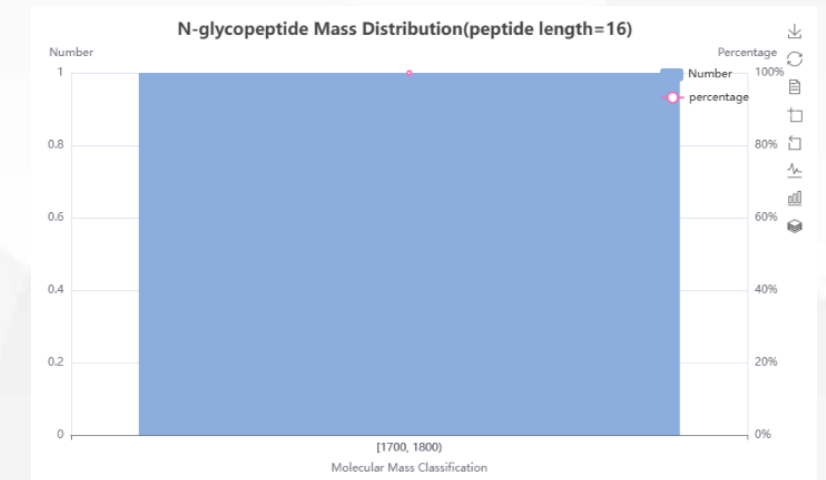
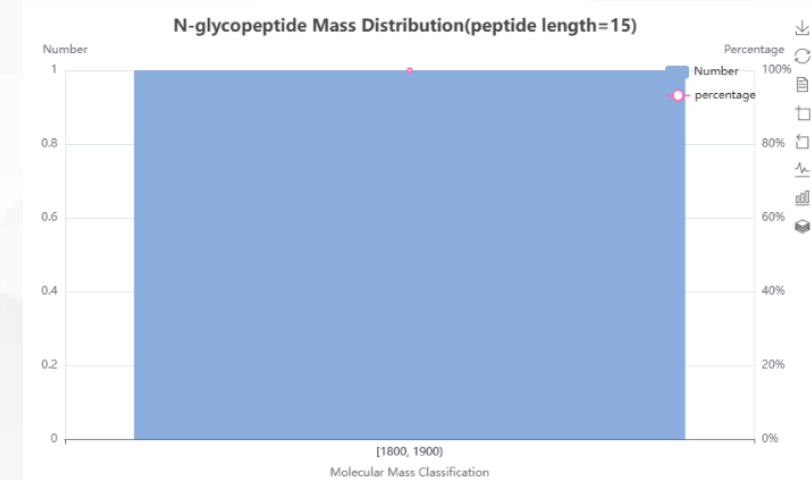
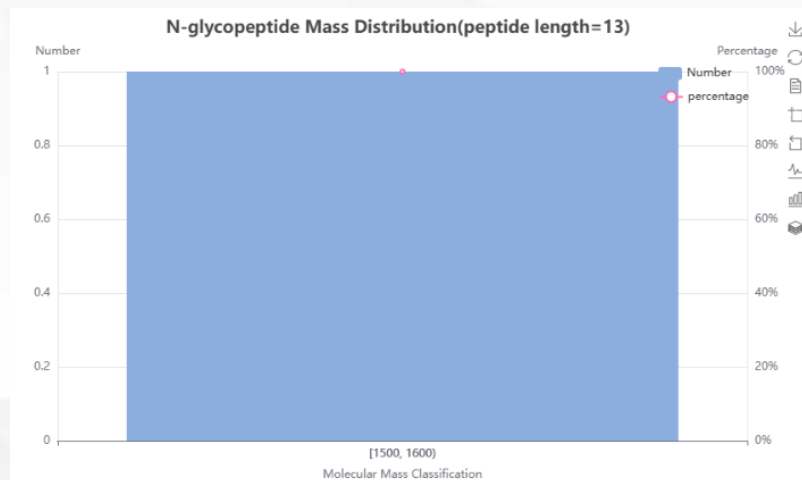
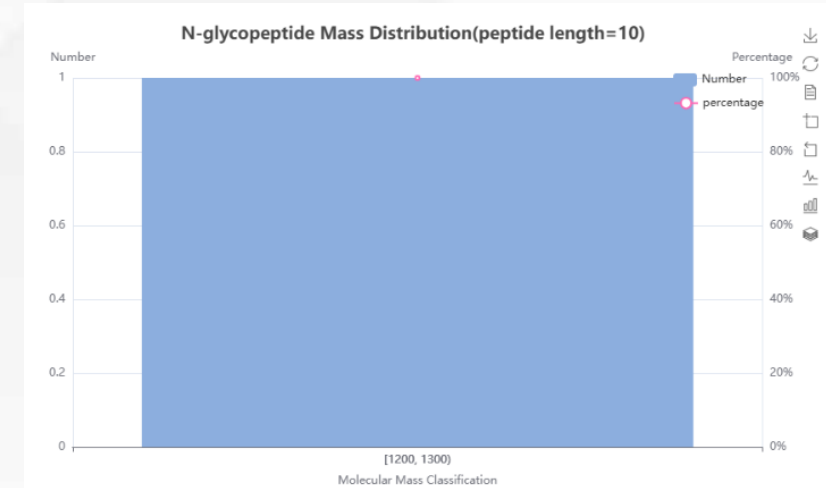
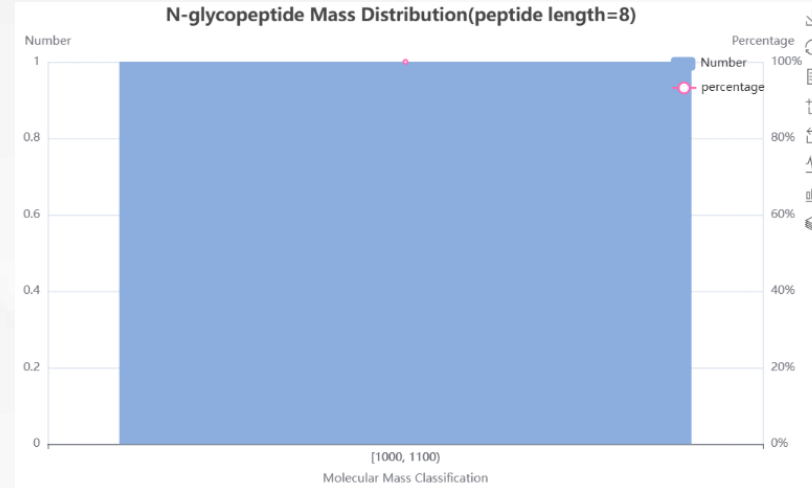
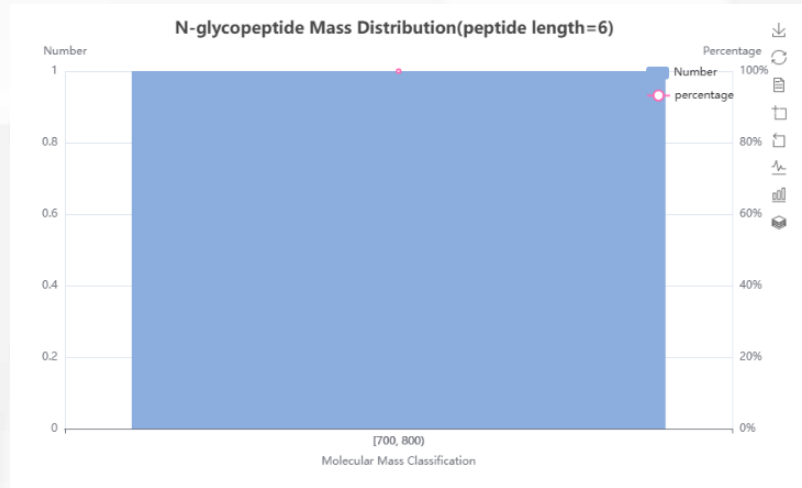
ProteinID & Name	Peptide	Peptide Mass	Peplength	Potential N-glycosite Count	N-glycosite position	IsUnique
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	JSSK	477.2308972	4	1	J1;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	EJITQDPEDHTVSFVQPNGAIFEPSLSVGTEDDJFTVLNLAVAAAPHIYQNSFVQVVLNSLIK	6894.4430302	63	2	J6;J38;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	DPFLSLVPYPISTTVGVFYFYJDTV DGVYK	3408.6931171999995	30	1	J89;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	DJISK	618.3098762000001	5	1	J99;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	JLSYWPSYCDMIJGTDAASFPPFVEK	3067.3499871999998	26	2	J103;J115;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	FVLPANAFASPLQNP DJHCFCTEK	2819.2932161999993	24	1	J145;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	VISJCTSYGVLDIGK	1781.8748992	16	2	J156;J157;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	RPYIVPILWJETGTIGDEK	2356.2552622	20	1	J179;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	QFWIFDVQNPDDVAKJSSK	2341.1226344	19	1	J204;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	JSSKIK	779.4388754	6	1	J208;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	YLAKEJITQDPEDHTVSFVQPNGAIFEPSLSVGTEDDJFTVLNLAVAAAPHIYQNSFVQVVLNSLIK	7430.7514424	67	2	J219;J251;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	EJITQDPEDHTVSFVQPNGAIFEPSLSVGTEDDJFTVLNLAVAAAPHIYQNSFVQVVLNSLIKK	7083.5669444000005	64	2	J282;J314;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	ELLWGYKDPFLSLVPYPISTTVGVFYFYJDTV DGVYK	4359.191849399999	37	1	J373;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	DPFLSLVPYPISTTVGVFYFYJDTV DGVYKVFNGK	4015.0182503999995	35	1	J403;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	VFNGKDJISK	1224.6350094000002	10	1	J423;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	DJISKVAIESYK	1582.8453874	13	1	J428;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	RJLSYWPSYCDMIJGTDAASFPPFVEK	3284.4800493999996	27	2	J441;J453;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	JLSYWPSYCDMIJGTDAASFPPFVEKSR	3371.5120773999997	28	2	J467;J479;	1
>sp Q08857 CD36_MOUSE Platelet glycoprotein 4 OS=Mus musculus OX=10090 GN=Cd36 PE=1 SV=2	GIPVYRFVLPANAFASPLQNP DJHCFCTEK	3565.7133043999993	30	1	J517;	1

N-glycoPeptide Presentation

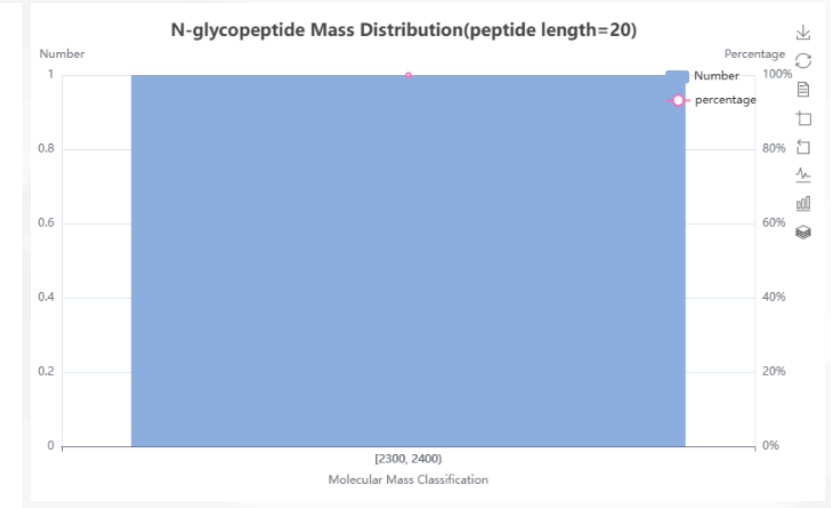
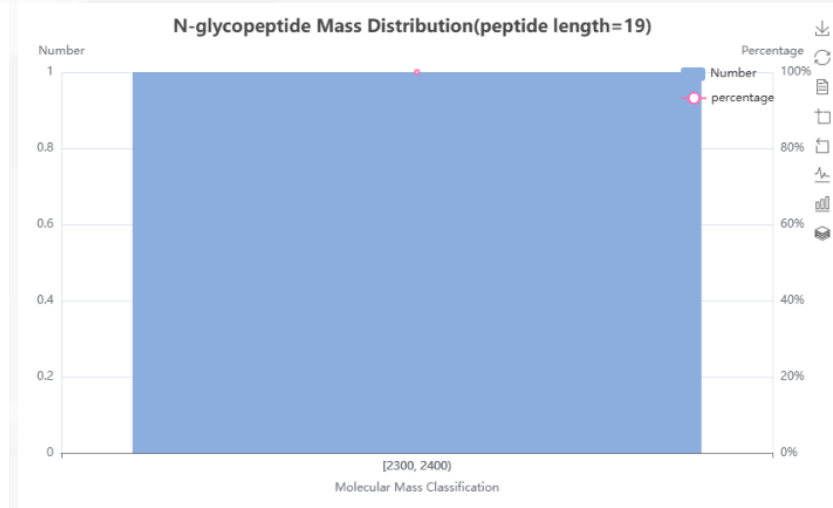
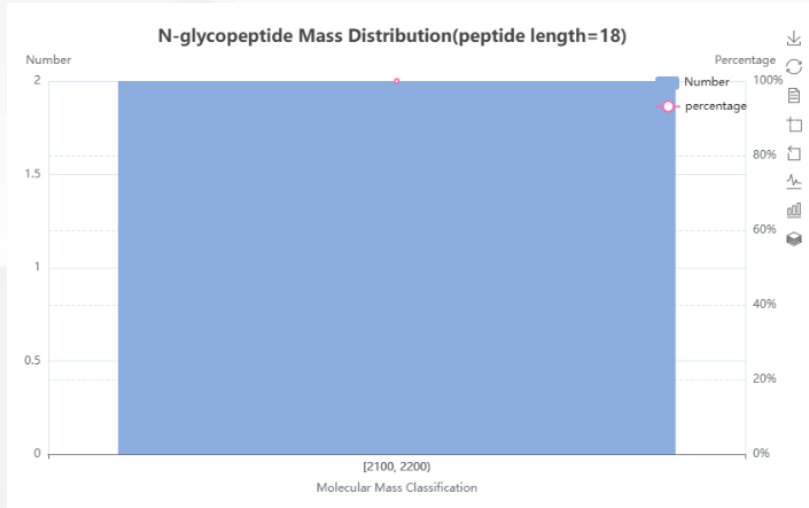
After simulating enzymatic digestion, a total of potential 66 N-glycosites are generated from 44 N-glycopeptides in 1 potential N-glycoproteins



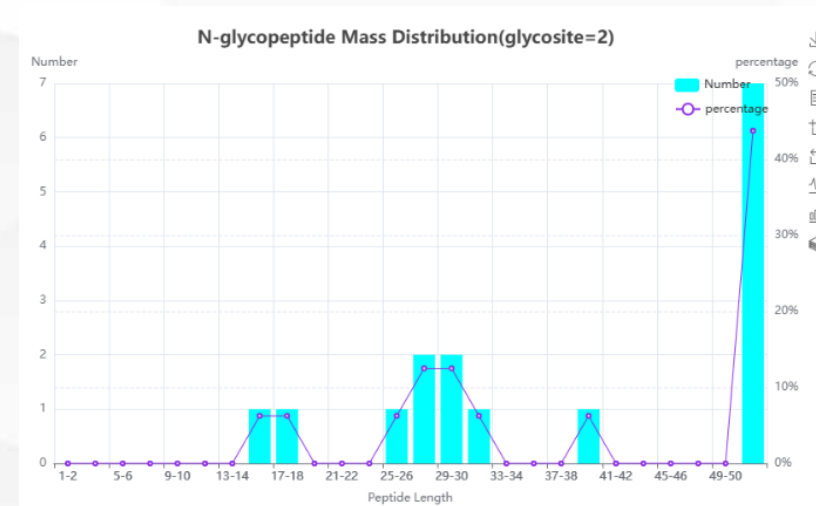
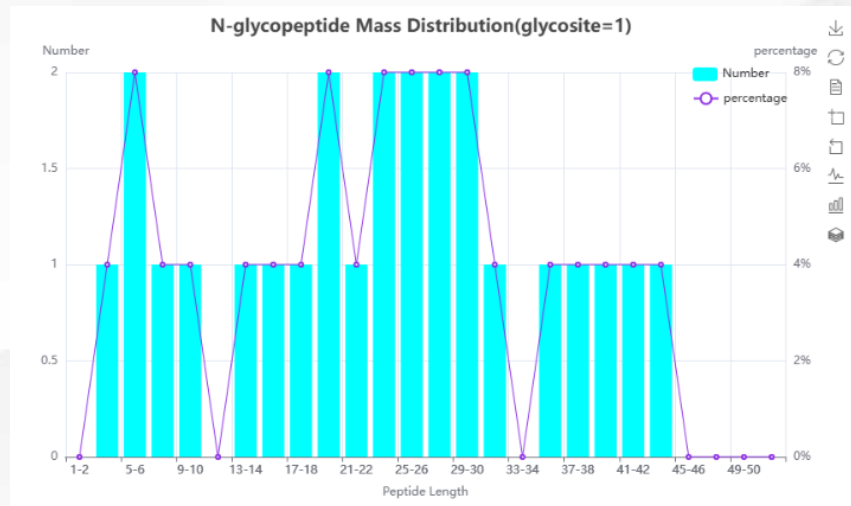
Display of Mass Distribution of N-glycopeptide Length Ranging in Length from 6 to 20



Nglycopeptide analysis



Length Distribution of N-glycopeptides with varying numbers of glycosites





THANKS FOR WATCHING
