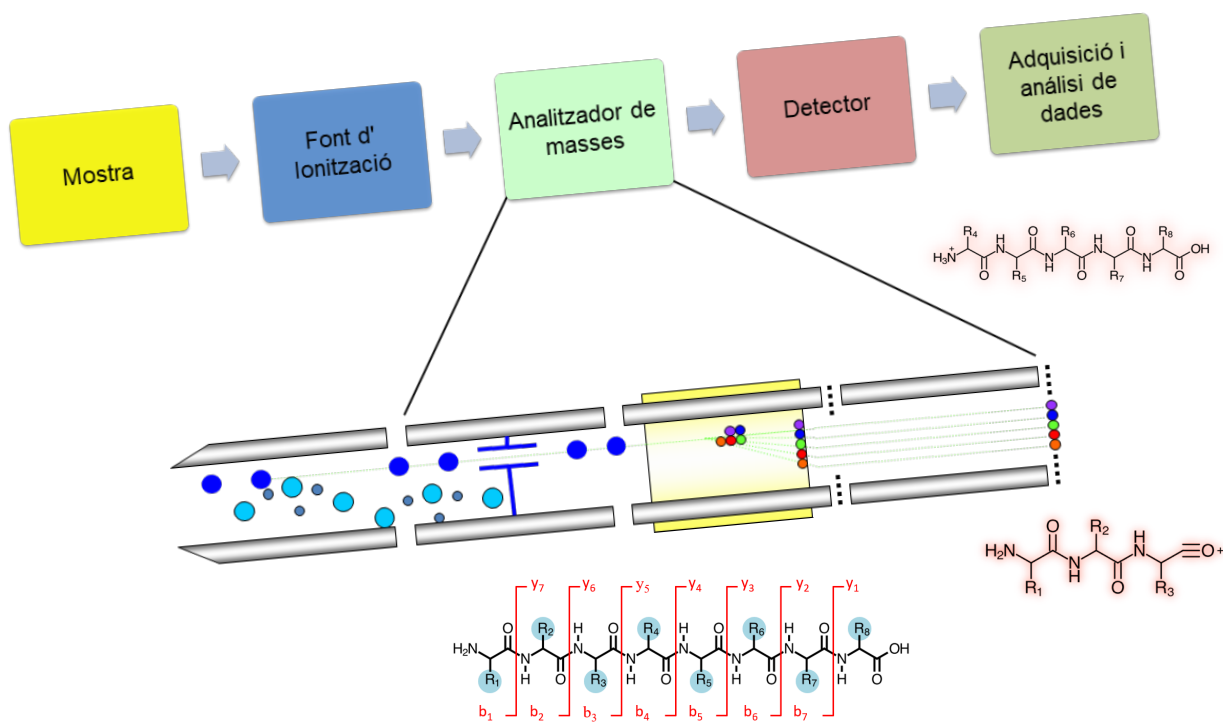
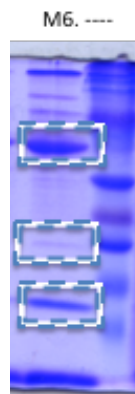
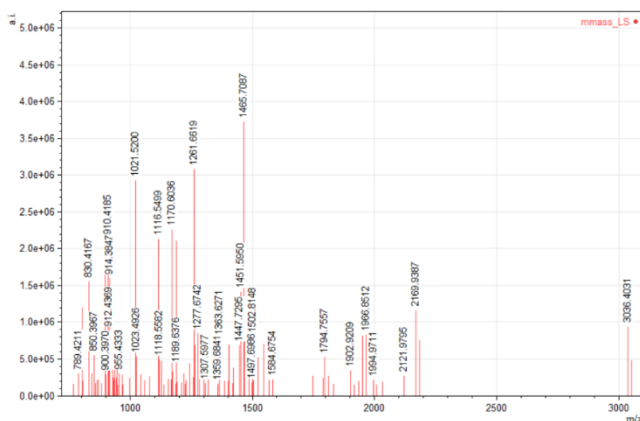


PROBLEMES D'ESPECTROMETRIA DE MASSES



MÈTODES EN BIOQUÍMICA

GRAU BIOQUÍMICA I BIOLOGIA MOLECULAR

María Jesús García Murria

murria@uv.es

En el següent document es presenten els problemes que es treballaran en les sessions de problemes de la part d'espectrometria de masses de l'assignatura de Mètodes en Bioquímica del Grau de Bioquímica i Ciències Biomèdiques.

En un primer lloc, es plantegen els problemes i posteriorment es descriu amb detall com resoldre'ls. Els arxius necessaris es troben a l'aula virtual en la carpeta de problemes d'espectrometria de masses.

Índex:

Sessió 1. Identificació de proteïnes per empremta peptídica utilitzant MASCOT.....pàg. 3

Sessió 2. Seqüenciació de pèptids de *novo*pàg. 4

Explicació detallada del procediment d'Identificació de proteïnes per empremta peptídica..... pàg. 5

1.1 Anàlisi dels espectres.....pàg. 7

1.2 Identificació de la proteïna.....pàg. 9

1.3 Procediment experimental.....pàg. 14

Explicació detallada del procés manual de seqüenciació de pèptids de *novo*. pàg. 17

Sessió 1. Identificació de proteïnes per empremta peptídica.

A partir de les dades del digerit tríptic d'una banda de gel SDS-PAGE, fes la identificació de la proteïna usant l'eina en línia de MASCOT. Segueix les instruccions explicades en classe o detallades en la descripció del problema.

Identifica la proteïna i analitza la informació que s'obté del resultat.

Quina cobertura de la proteïna s'ha aconseguit? Per què no s'ha aconseguit una cobertura del 100%?

Observa la llista de pèptids identificats, com s'obté la $M_r(\text{exp})$ a partir de la Massa Observada? Recorda que les dades s'han obtingut en un espectròmetre MALDI-TOF.

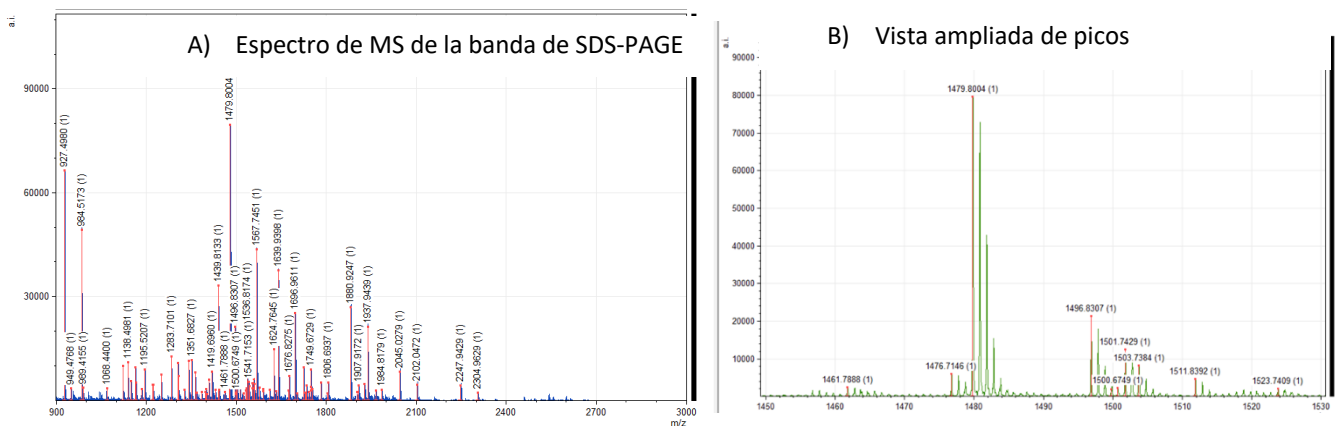
Arxiu: MASCOT_CONTROL.txt

Enllaç a l'eina de cerca:

http://www.matrixscience.com/cgi/search_form.pl?formver=2&*SEARCH=*PMF

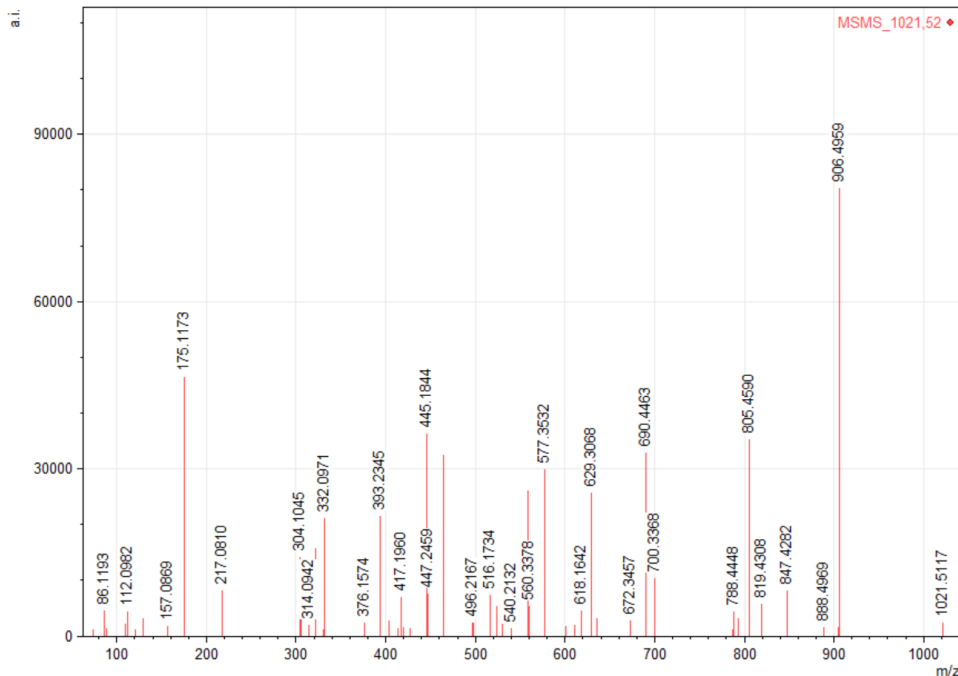
Analitza també l'espectre obtingut. A) Si estem observant l'espectre d'un digerit d'una proteïna purificada, per què no tenen tots els pics la mateixa intensitat?

B) Per què tots els pics tenen el mateix patró on sol el pic etiquetat és el de major intensitat i després d'ell li segueixen una sèrie de pics més de menor intensitat?



Sessió 2. Seqüenciació de pèptids de *novo*.

A partir de l'espectre de fragmentació d'un pèptid dedueix la seua seqüència. Per això has d'identificar els possibles ions y, b i imonio. Calcula la diferència entre pics contigus. Aquell que coincidisca amb un valor de massa d'un residu d'aminoàcid **indicarà el residu** que es troba entre aquests ions.



Aminoàcid	Codi de 3 lletres	Codi d'1 lletra	Massa del residu	Ió imonio
Glicina	Gly	G	57,02	30
Alanina	Ala	A	71,04	44
Serina	Ser	S	87,03	60
Prolina	Pro	P	97,05	70
Valina	Val	V	99,07	72
Treonina	Thr	T	101,05	74
Cisteïna	Cys	C	103,01	76
Isoleucina	Ile	I	113,08	86
Leucina	Leu	L	113,08	86
Asparagina	Asn	N	114,04	87
Àcid aspàrtic	Asp	D	115,03	88
Glutamina	Gln	Q	128,06	101
Lisina	Lys	K	128,09	101
Àcid glutàmic	Glu	E	129,04	102
Metionina	Met	M	131,04	104
Histidina	His	H	137,06	110
Fenilalanina	Phe	F	147,07	120
Arginina	Arg	R	156,01	129
Tirosina	Tyr	Y	163,06	136
Triptòfan	Trp	W	186,08	159
Cisteïna carbamidometilada			160,03	

Explicació Detallada Sessió 1: IDENTIFICACIÓ DE PROTEÏNES MITJANÇANT EMPREMTA PEPTÍDICA UTILITZANT MASCOT

Una vegada efectuat el procés de purificació de Rubisco en les pràctiques de Mètodes en Bioquímica del segon quadrimestre, s'avaluà el procés analitzant cadascuna de les mostres preses al llarg de tot el procés per electroforesi SDS-PAGE.

En la carrera que corresponia a la mostra purificada M6 que provenia del cromatografia d'intercanvi iònic es van detectar més bandes de les esperades. Per a comprovar la presència de Rubisco, es retallà la banda que per migració subunitat gran (LS) i a la subunitat petita (SS), i, a més es va retallar una banda amb una migració intermèdia que apareixia en la majoria dels grups.

El protocol exacte del tractament de les bandes s'adjunta al final de les instruccions d'aquesta sessió.

Breument, cada banda assenyalada del gel d'acrilamida es va tallar amb una fulla, es va destenyir, es va reduir amb DTT i es va bloquejar amb **iodoacetamida**, abans de la digestió mode **Reflector** amb **tripsina**. Directament, el digerit sense concentrar va ser analitzat en de m/z coneguda. **positiu** en un **MALDI-TOF** prèviament calibrat amb estàndards comercials la qual es va Com a control de digestió i detecció es va usar una proteïna estàndard, sotmetre al mateix procés.

En l'aula virtual estan els arxius que contenen la llista de tots els pics exportada directament del MALDI.

Aquests espectres es poden visualitzar en programes lliures com el mmass <http://www.mmass.org/download/> (que no és compatible amb les últimes actualitzacions de Mac). No cal descarregar-se el programa, es pot seguir el procés amb les explicacions del qüestionari.

A continuació, s'explica el procés de visualització i cerca de l'empremta peptídica, usant el programa mmass que es pot descarregar gratuïtament (*instruccions escrites en color blau*) però per a aquells que no puguem o no vulguem descarregar-se el programa es mostraran captures de pantalla del programa.

Les qüestions que haureu de contestar estan emmarcades en un quadrat color daurat.

Una vegada descarregat el programa mmass:

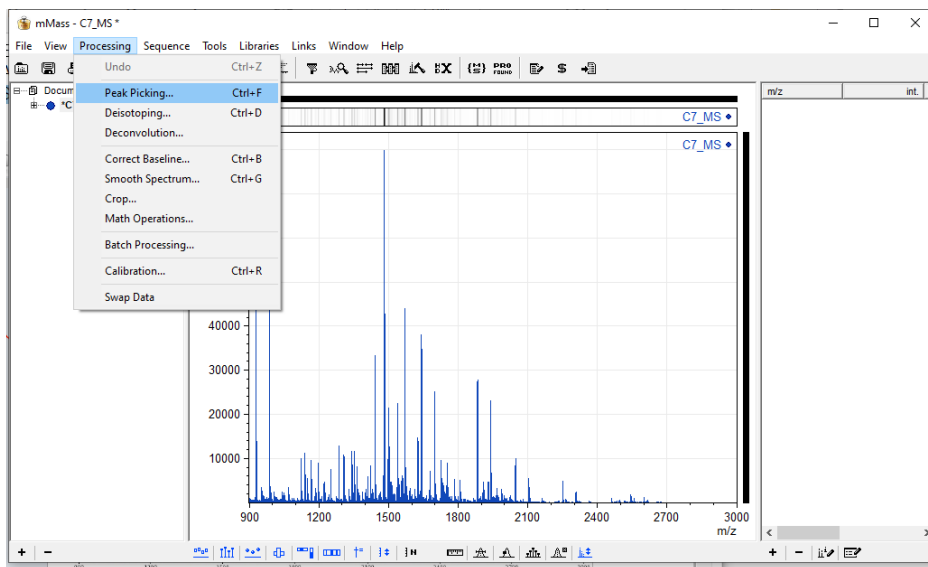
Heu de descarregar les dades de masses dels arxius de l'aula virtual (mmass_Control.txt; mmass_LS.txt).

Començarem analitzant la **mostra de control**:

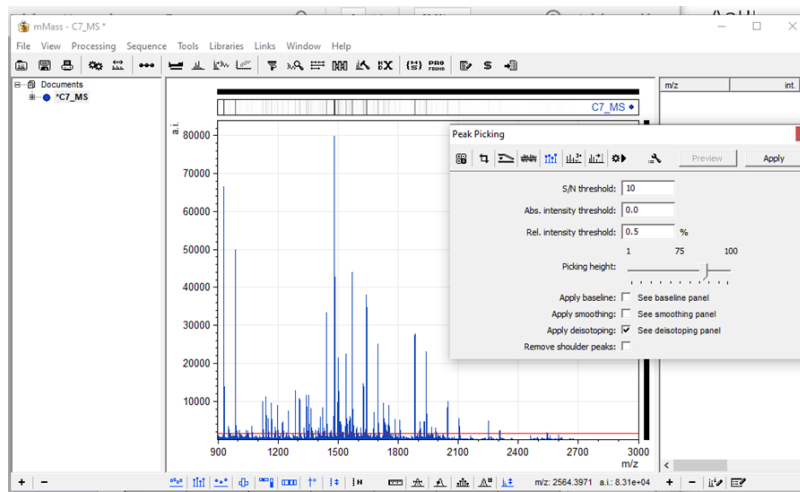
Obriu el document “**mmass_CONTROL.txt**”

S'ha de processar l'espectre per identificar els pèptids.

Processing i després Peak Picking (Ctrl+F)



La selecció dels pics es fa amb els criteris següents, amb un senyal/soroll de 10; un % de límit de detecció de 0.5% d'intensitat relativa i que aplique el “deisotoping” perquè marque només els pics monoisotòpics.

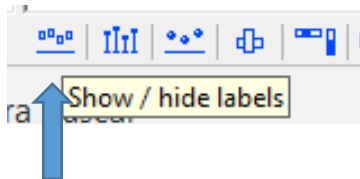
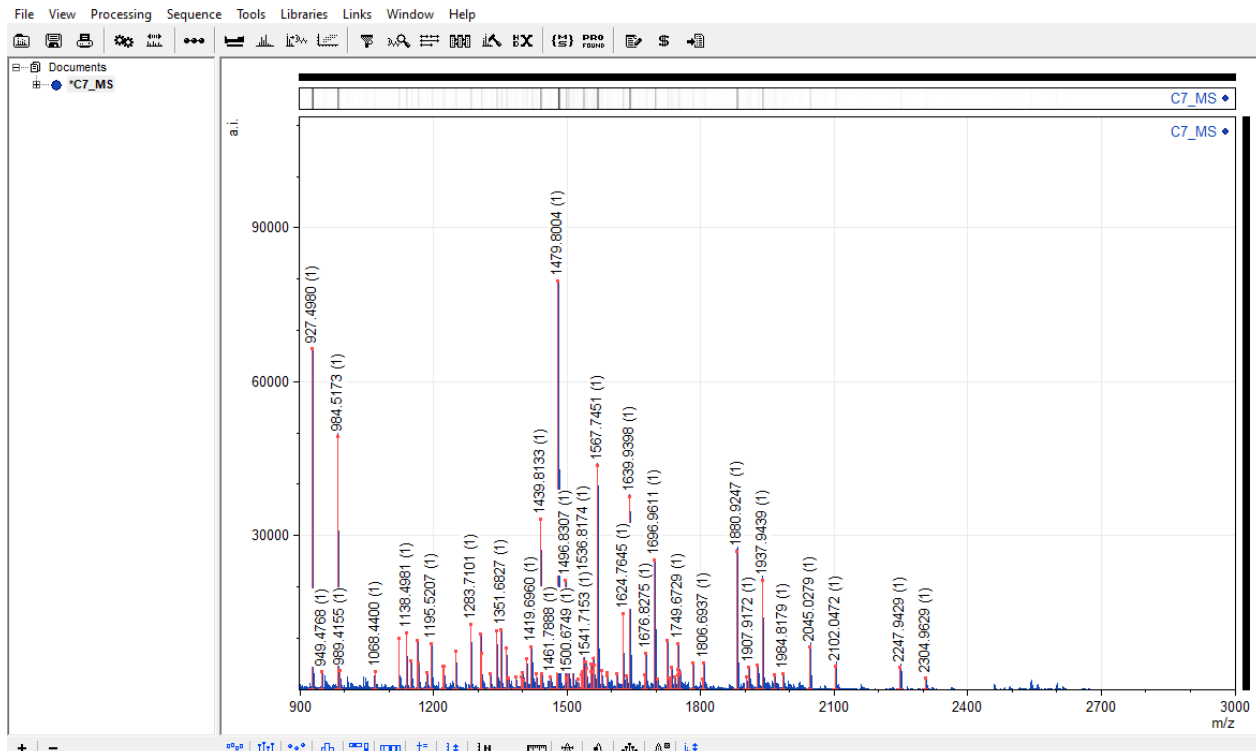


I apareixerà l'espectre amb els pics anotats

1.1 ANÀLISI DELS ESPECTRES

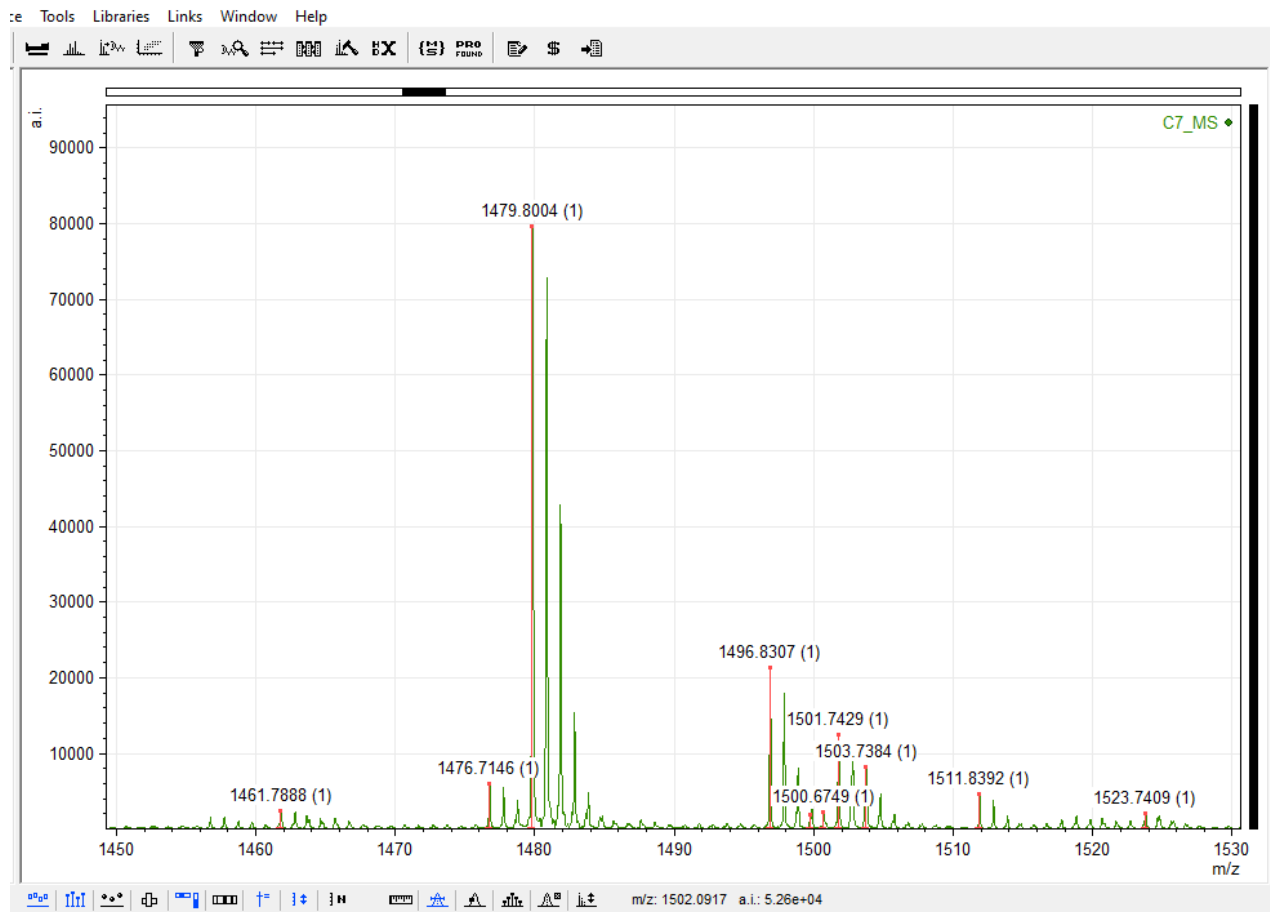
Abans de fer la cerca analitzarem l'espectre.

Si estem observant l'espectre d'un dígerit d'una proteïna purificada, per què no tenen tots els pics la mateixa intensitat?



Per a fer zoom en un pic, useu el botó dret i el cursor per a seleccionar la zona a ampliar. Amb doble clic en l'espectre es torna a la visió global.

En fer zoom en qualsevol pic apareix en tots la mateixa distribució.



Per què tots els pics tenen el mateix patró on sols el pic etiquetat és el de major intensitat i després d'ell li segueixen una sèrie de pics més de menor intensitat?

1.2 IDENTIFICACIÓ DE LA PROTEÏNA

Per a fer la cerca de l'empremta peptídica, es poden usar diferents motors de cerca. Es pot fer des del programa mmass o directament des de la web de MASCOT exportant la llista de pics monoisotòpics de mmass.

Opció 1: Cerca en la web de MASCOT

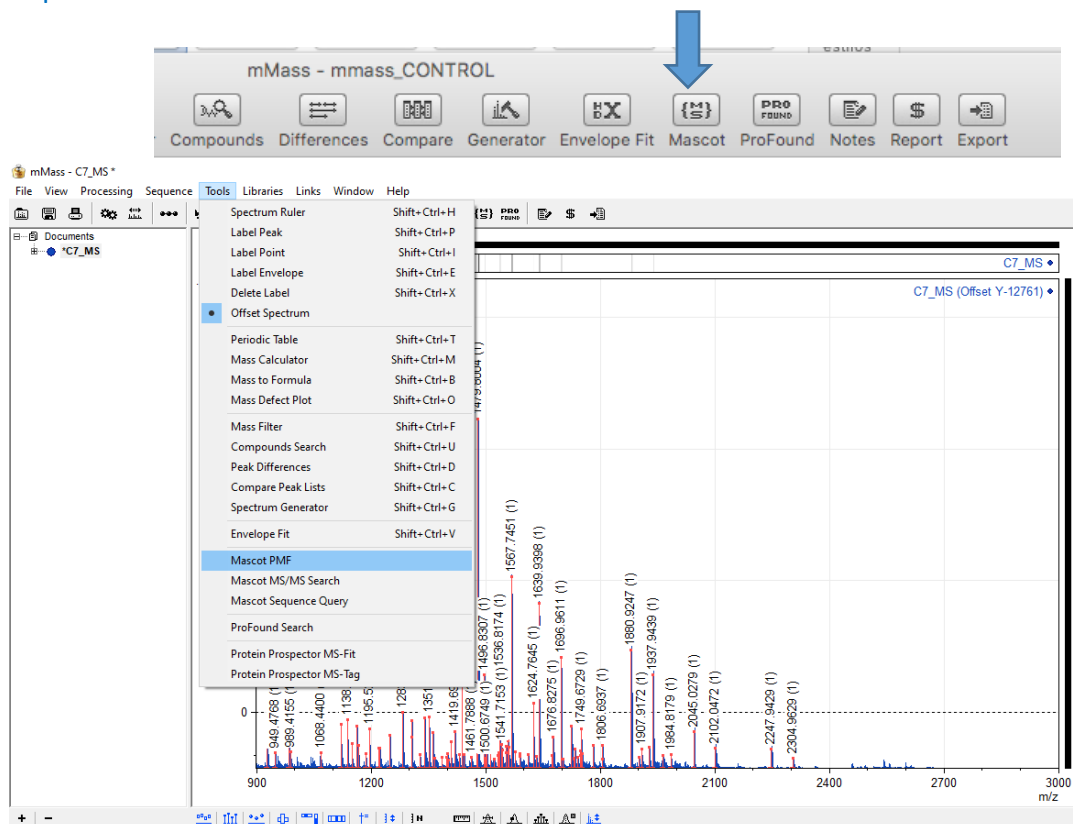
http://www.matrixscience.com/cgi/search_form.pl?formver=2&SEARCH=PMF

Cal descarregar-se de l'aula virtual els arxius que contenen la llista de pics monoisotòpics (**MASCOT_CONTROL.txt**; MASCOT_LS.txt)

Assegureu-vos d'estar en l'opció "**Mascot - Peptide Mass Fingerprinting**". Després d'indicar el nom i el correu (que haureu de validar la primera vegada), seleccioneu els valors i opcions que considereu tenint en compte l'enunciat del principi. *Llegiu altra vegada l'enunciat i penseu quina modificació o quines modificacions hauríeu d'afegir-hi*. Al final d'aquest apartat teniu una explicació i recomanacions per a cada paràmetre que heu d'indicar en la cerca.

Opció 2: Cerca a través de mmass usant MASCOT

Per a fer la cerca en Mascot des del programa, punxeu sobre el botó de Mascot o seleccioneu l'opció "Mascot PMF" a través del menú "Tools".



Paràmetres modificables del motor de Cerca

Base de dades

Si l'organisme en el qual es treballa està ben caracteritzat, es recomana utilitzar la base de dades Swiss-Prot perquè les entrades són d'alta qualitat i estan ben anotades. A més, com que no és redundat, és relativament petita i és més fàcil tenir un resultat estadísticament positiu. Les EST (**Expressed Sequence Tags**) s'usen per a dades de MS/MS si el genoma no està seqüenciat.

Una vegada tingueu la proteïna identificada podríeu provar de fer cerques en altres bases de dades i veure què passa.

Enzim i "missed cleavages"

Seleccioneu la proteasa utilitzada i el nombre de llocs sense tallar que se salta l'enzim. L'habitual és permetre 1-2 llocs, més "misscleavages" podria afectar a obtenir un resultat positiu perquè augmentem el nombre de possibles pèptids teòrics que s'han d'enfrontar al resultat experimental. *Podríeu provar-ho i comprovar-ho.*

Taxonomia:

Fixar la taxonomia accelera i simplifica el resultat perquè no donarà les proteïnes homòlogues d'altres espècies.

A vegades interessa fer una cerca sense fixar la taxonomia per a poder detectar possibles contaminacions o proteïnes que no es troben en la base de dades de l'organisme que treballem.

Modificacions fixes i variables:

La recomanació és, si una modificació és segur que hi és, afegir-la com a fixa.

El nombre de modificacions variables farà que el nombre de pèptids possibles siga major i penalitza a l'hora d'obtenir un resultat significatiu.

En general, com més variables hi haja, **augmenta el nombre de masses de pèptids calculats per a ser enfrontats contra les dades experimentals**, i això disminuirà la puntuació de la identificació i, per tant, la possibilitat de trobar un resultat estadísticament significatiu.

Prot Mass encara que sapiem el rang per la migració en el gel SDS-PAGE, val més no afegir-hi res, i usar la dada de migració electroforètica per a validar el resultat final.

Tolerància:

El valor de tolerància és crític. Amb una tolerància molt ajustada pot ser que no s'obtinga resultat positiu per un problema en el calibratge de les dades. Per contra, una tolerància massa laxa pot generar falsos positius i artefactes.

Habitualment s'estableix una tolerància **50 ppm** per a les cerques d'empremta peptídica. Així i tot, una vegada obtingut un resultat s'ha d'analitzar el gràfic d'error (ppm) vs. Mass (Da) i comprovar que l'error dels pèptids assignats (diferència entre el valor teòric i experimental) està en un rang molt estret.

Mass Value: Seleccioneu el que corresponga. En aquest cas els resultats s'han obtingut en un **MALDI en mode reflector positiu**, per tant, indica en quina forma s'obtenen els fragments.

Data file (només si fas la cerca des de la web) Seleccioneu l'arxiu **MASCOT_CONTROL.TXT** de l'aula virtual o l'obtingut en exportar les dades del mmass *Si es fa directament des del mmass, aquesta opció no apareixerà.*

Start search

(si és la primera vegada que es fa la cerca, s'ha de validar el correu i tornar arrere a fer la cerca)

Cada entrada en la base de dades es digereix *in silico*, es calculen les masses dels pèptids teòrics segons les condicions de la digestió i modificacions indicades i per a cada pèptid es dedueix la seua massa teòrica (M(calc)).

Mitjançant algorismes, es calcula **una puntuació** per a decidir si el nombre de coincidències d'alguna de les proteïnes anotades és suficient perquè no es deguen a l'atzar. En aquest cas, es dona la proteïna com identificada.

Per a fer-ho el programa efectua un càlcul (Expect) de si la coincidència entre les dades experimentals i teòrics es deu a l'atzar i genera una puntuació. Segons els paràmetres de cerca, considera que a partir d'una certa puntuació (Score) el resultat és significatiu i no és degut a l'atzar.

Feu la cerca. En cas de no obtenir un resultat significatiu, canvieu paràmetres seguint les recomanacions anteriors. En cas de tenir informació de la massa o pl podeu comprovar si la identificació quadra amb les dades experimentals

Una vegada tingueu la proteïna identificada amb una puntuació superior al valor que es considera significatiu, es pot punxar sobre el resultat i s'obrirà una pàgina nova. **Dediqueu temps a analitzar el resultat obtingut.**

Al principi apareixerà una llista dels paràmetres seleccionats i de la proteïna identificada.

Al mig, apareixerà la seqüència de la proteïna en negre i en roig els pèptids detectats. A partir d'aquestes dades es calcula **la cobertura** de la seqüència que representaria el percentatge de seqüència coberta a partir dels pèptids detectats.

Protein sequence coverage: 54%

Matched peptides shown in **bold red**.

```
1 MGNVTFISLL LFFSAYSRG VFRRDTHKSE IAHRFYDLGE RHFKGLVLI A
51 FSQILQQCF DEHVLVNBL TEPARTCVAD ESHAGCEKSL HTLFGDLCK
101 VASLRETYGD HADCEKQEP ERNRCPLSHK DSDPDLPLIK PDPNTLCDEF
151 KADEKKFWGK YLVEIARRHP YPIAPLLIY ANKINGVPEE CQADRDGAC
201 LLFRIETMRE KVLASSARQR LRCASIQRFQ ERALKAWSVA RLSQRFFPAE
251 FVEVTKLVID LTRVHRKCCCH GDLLECADDR ADLARKYICDN QDTISSKLRG
```

Quina cobertura de la proteïna s'ha aconseguit?

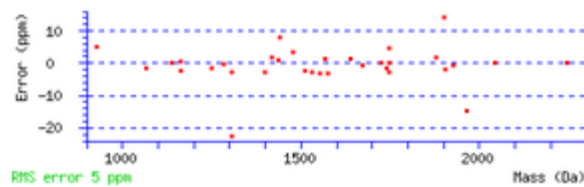
Sota la seqüència apareix una llista dels pèptids detectats ordenats per posició en la seqüència. Cada columna indica:
 Posició inicial i final de l'aminoàcid, la massa observada, la massa experimental (exp), la massa calculada (teòricament a partir de la seqüència d'aminoàcids), la diferència en ppm de Mr(exp)-Mr(calc), els "misscleavages" (M) i la seqüència del pèptid.

Start - End	Observed	Mr(expt)	Mr(calc)	ppm	M	Peptide
35 - 44	1249.6192	1248.6119	1248.6139	-1.57	1	R.FRDLGEENFK.G
66 - 75	1163.6277	1162.6204	1162.6234	-2.53	0	K.LVNELTEFAK.T

Com s'obté la Mr(exp) a partir de la massa observada? Recordeu que les dades s'han obtingut en un espectròmetre MALDI TOF

Finalment, apareix la relació de pics no assignats i la gràfica de l'errada (ppm) respecte a la massa dels ions assignats.

No match to: 949.4768, 984.5173, 989.4155, 1121.4718, 1149.4663, 1185.6135, 1195.5207, 1220.6466, 1223.1500.6749, 1501.7429, 1503.7384, 1523.7409, 1533.7606, 1536.8174, 1541.7153, 1543.7163, 1553.8498, 1558.1804.7205, 1806.6937, 1937.9439, 1984.8179, 2102.0472, 2304.9629



Per què no s'ha aconseguit una cobertura del 100%?

Per a comprovar els pèptids detectables, podeu usar l'eina Peptide Mass d'Expasy ([https://web.expasy.org/peptide mass/](https://web.expasy.org/peptide_mass/)) i així fer una digestió *in silico* de la seqüència que li indiqueu.

Per a això, cal afegir la seqüència de la proteïna (afegint la seqüència d'aminoàcids o indicant l'identificador d'UniprotKb, ex. ALBU_BOVIN) i indicar les modificacions.

Tingueu en compte que en un MALDI-TOF en mode reflector positiu se sol determinar el rang entre 850-2500 Da (amb valors inferiors l'espectre ix contaminat amb pics de la matriu; en valors superiors no se solen detectar ions).

Desmarqueu totes les opcions de "For UniProtKB (Swiss-Prot/TrEMBL) entries only:

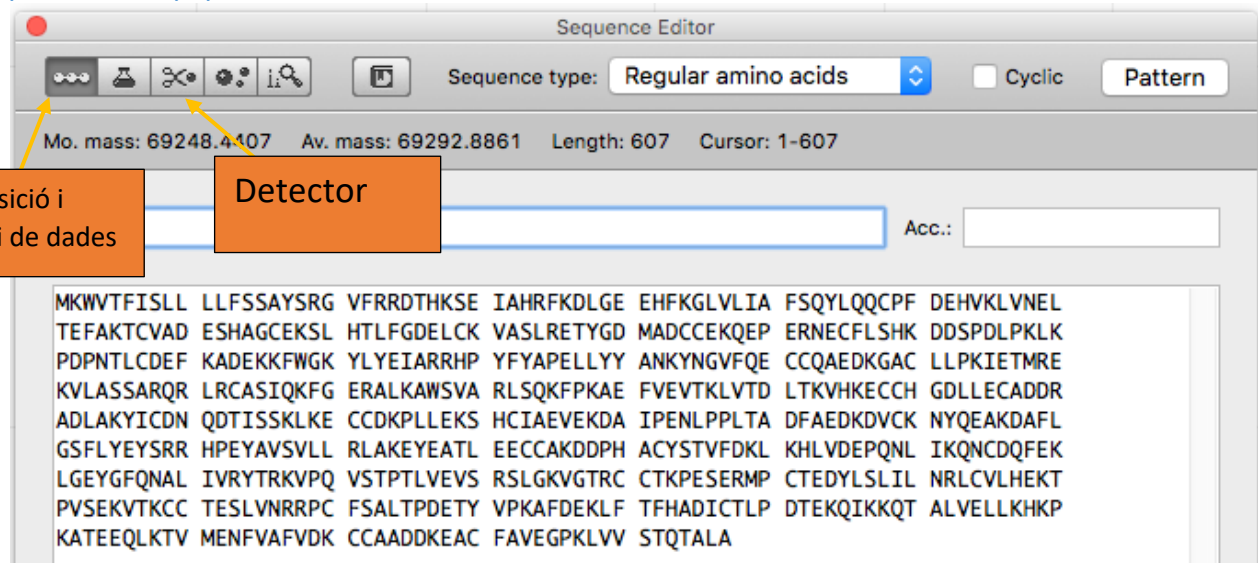
For each peptide display)

Una vegada ho tingueu tot marcat feu clic en «PERFORM the cleavage» (sota el quadrat de la seqüència) per a digerir.

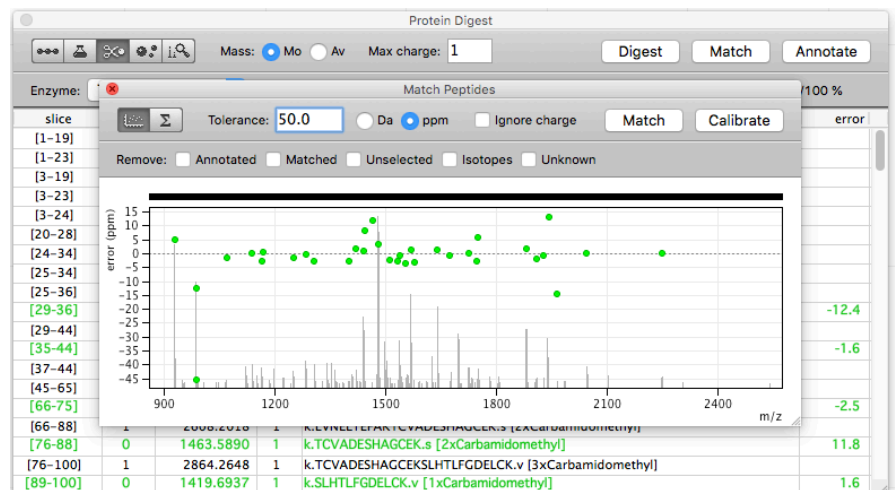
La digestió també es pot realitzar amb el programa mmass i en aquest cas que assigne a cada pic el pèptid que per massa quadraria.

Obriu l'eina i afegiu la seqüència del resultat anterior (la podeu copiar directament del resultat de Mascot)

Moveu-vos pels menús laterals per a afegir modificacions, fer digestió in silico i fer “match” dels pics amb els pèptids.



Una vegada heu fet clic en **Digest**, feu que es marquen els pics experimentals que quadren amb un pèptid (**Match**). Aquests es marcaran en verd. Tanqueu la finestra i de nou sobre la finestra de **Digest** indiqueu **Annotate** perquè mostre la llista de pics anotats.



EXTRES: El programa mmass permet en cas que no s’haja fet, un **calibratge de l’espectre amb estàndards interns de calibratge o les masses dels pèptids autoproteolítics de tripsina porcina (841.50 and 2210.10)**. En aquest cas no és necessari perquè el calibratge ja s’ha fet prèviament.

L’eina de **Peak differences** del programa mmass és molt útil per a ajudar en la seqüenciació de novo manual de pèptids.

Repetiu la cerca per a les bandes LS

Si ho feu directament en Mascot, useu l’arxiu **MASCOT_LS.txt**

Si ho feu a través del mmass, useu l’arxiu **mmass_LS.txt**

Els resultats de SS i contaminant estan en un document a part a l’Aula Virtual.

1.3. PROCEDIMENT EXPERIMENTAL

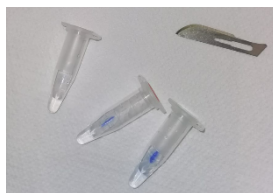
PROTOCOL DIGESTIÓ BANDA DE GEL

ACN: acetonitril

DTT: Ditiotreitòl

IAM: Iodoacetamida

TFA: àcid trifluoroacètic



Sempre que no s'indique el contrari, s'eliminarà tot el volum del reactiu abans d'afegir el següent:

Destinció Coomassie

- 2x10' H₂O/ACN (1:1, v/v)

- ACN quan el gel s'ha deshidratat traure el líquid

- Rehidrateu amb 50 mM NH₄HCO₃, després de 5' afegiu ACN. Espereu 15 min.

- 1x5' ACN



Reducció i bloqueig de les Cisteïnes

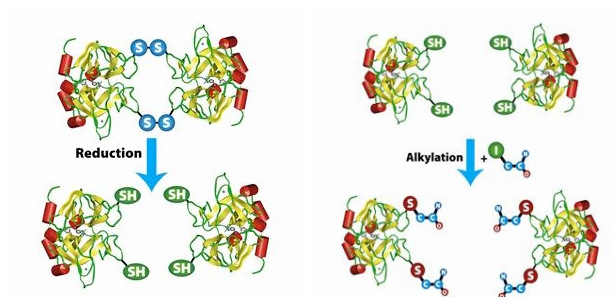
- 1x20' (60 °C) 10mM DTT en 50 mM NH₄HCO₃

- 1x30' (foscor) 55mM IAM en 50 mM NH₄HCO₃

- 1x2' H₂O/ACN

- 1x5' ACN

- Speed-Vac o deixeu assecat completament.



Digestió

- Tripsina (50 ng en 10 µL). Incubeu 30' en gel. Addicioneu 50 µL 50 mM NH₄HCO₃

- tota la nit 37 °C

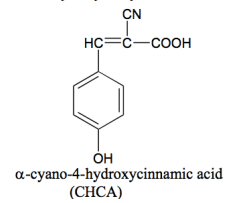
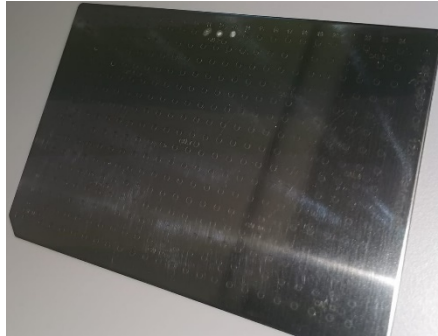
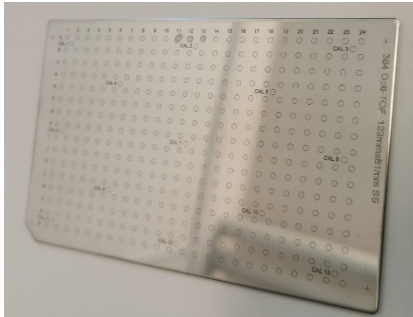
La reacció s'atura:

-3 µL 10%TFA (pH=1)

ADQUISICIÓ DE DADES EN MALDI TOFTOF 5800 (AB Sciex)

1 µL de la solució del digerit es deposita en la placa de MALDI.

Es deixa assecar a l'aire i a continuació s'addicionen 0.5 µL de solució de matriu (5 mg/ml d'ACH (àcid α-ciano-4hidroxicinàmic) en 50% ACN 0.1% TFA).

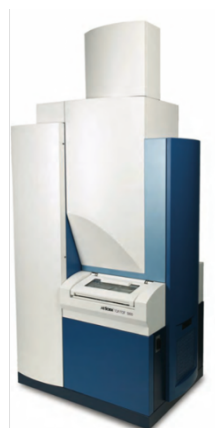


En el MALDI, s'ha de calibrar la geometria de la placa i els mètodes d'adquisició. Per a això, es deposita 1 µL de la solució de calibratge TOFTOF (Referència 4333604, ABSciex), en 13 posicions de la placa. Els mètodes de MSMS es calibren amb l'espectre de fragmentació d'Angiotensina I

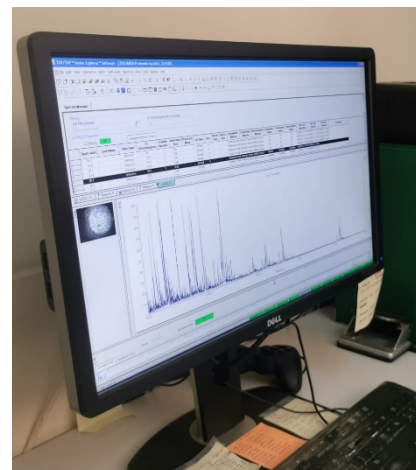
Standard Mixture Component	Final Standard Concentration	CALIBRATION SPECTRUM
des-Arg1-Bradykinin	1.0 pmol/µL	
Angiotensin I	2.0 pmol/µL	
Glu1-Fibrinopeptide B	1.3 pmol/µL	
ACTH (1-17 clip)	2.0 pmol/µL	
ACTH (18-39 clip)	5 pmol/µL	
ACTH (7-38 clip)	3.0 pmol/µL	

S'adquireixen els **espectres de MS en manera reflector positiu**.

El rang de massa analitzat és 850-3500 m/z. La intensitat i el nombre de trets de làser s'ajusten manualment per a optimitzar la resolució i la relació senyal/soroll de l'espectre.



MALDI TOF/TOF 5800



Per a fer MS/MS:

Posteriorment, **se seleccionen automàticament 5 ions**, entre aquells de major intensitat, en aquest cas els que corresponen a contaminacions conegudes es descarten. De cada un d'aquests pèptids s'obté un **espectre de fragmentació**.

Com veieu en el protocol d'adquisició, en realitat s'ha fet un assaig de MS, i després s'han seleccionat els ions MS/MS. En la web del MASCOT en l'apartat de **MS/MS Ions Search** (https://www.matrixscience.com/cgi/search_form.pl?formver=2&SEARCH=ELS MEUS) es podria fer la cerca amb els MS/MS.

Per si voleu provar-ho, en l'aula virtual teniu l'arxiu **MSMSdeLS.txt**

Com a resultat, obtindreu el nom de la proteïna identificada i els pèptids que corresponen als espectres de fragmentació.

1 **RBL_BERBR** 133 Ribulose biphosphate carboxylase large chain (Fragment) OS=Bertiera breviflora OX=43448 GN=rbcL PE=3 SV=1

1.1 [RBL_BERBR](#) **Score** **Mass** **Matches** **Sequences** **emPAI**
 133 52522 4 (4) 4 (4) 0.49 Ribulose biphosphate carboxylase large chain (Fragment) OS=Bertiera breviflora OX=43448
 ▶ **122** same sets of RBL_BERBR

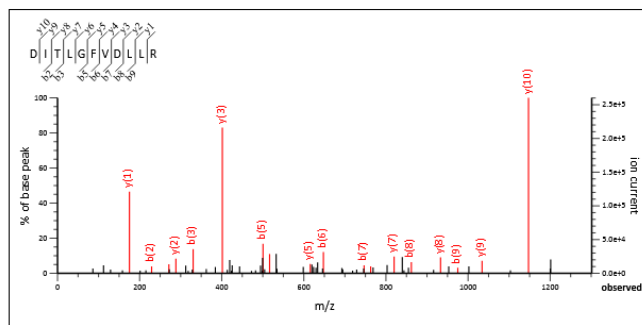
▼ **4 peptide matches (4 non-duplicate, 0 duplicate)**

Auto-fit to window

Query Dupes	Observed	Mr (expt)	Mr (calc)	ppm	M	Score	Expect	Rank	U	Peptide
16	910.4185	909.4112	909.4378	-29.3	0	37	0.0068	▶1	U	R.AVYECLR.G
34	1021.4899	1020.4826	1020.5240	-40.5	0	59	0.00057	▶1	U	K.DTDILAAFR.V
60	1261.6619	1260.6546	1260.7078	-42.2	0	75	2e-05	▶1	U	R.DITLGFVDLLR.D
77	1465.7087	1464.7014	1464.7474	-31.4	0	66	1.9e-05	▶1	U	K.TFQPPHGIQVER.D

Fixeu-vos que genera 122 resultats similars amb la mateixa puntuació, ordenats alfabèticament. Es tracten de pèptids conservats en diferents espècies, entre les quals hi ha *Citrus sinensis*.

En punxar sobre un pèptid apareixerà l'espectre de MS/MS amb l'assignació de cada fragment a l'ió corresponent.



Label all possible matches Label matches used for scoring

Monoisotopic mass of neutral peptide Mr(calc): 1260.7078
 Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
 Ions Score: 75 Expect: 2e-05
 Matches: 18/48 Fragment Ions using 33 most intense peaks (help)

#	a	b	Seq.	y	y ⁺	#
1	88.0393	116.0342	D			11
2	201.1234	229.1183	I	1146.6881	1129.6616	10
3	302.1710	330.1660	T	1033.6041	1016.5775	9
4	415.2551	443.2500	L	932.5564	915.5298	8
5	472.2766	500.2715	G	819.4723	802.4458	7
6	619.3450	647.3399	F	762.4509	745.4243	6
7	718.4134	746.4083	V	615.3824	598.3559	5
8	833.4403	861.4353	D	516.3140	499.2875	4
9	946.5244	974.5193	L	401.2871	384.2605	3
10	1059.6085	1087.6034	L	288.2030	271.1765	2
11			R	175.1190	158.0924	1

Explicació Detallada Sessió 2: SEQÜENCIACIÓ DE PÈPTIDS DE NOVO

La "seqüenciació de pèptids *de novo*" és la seqüenciació de pèptids realitzada sense coneixement previ de la seqüència d'aminoàcids a partir d'un espectre de fragmentació MS/MS.

L'espectre s'obté una vegada se selecciona i aïlla un pèptid precursor en una càmera de col·lisió on en col·lidir amb molècules de gas inert es generen diferents tipus de fragments (Fig.1).

El conjunt de pics que s'observa en un espectre de fragmentació és un reflex de la població d'ions fragments produïts en la cel·la de col·lisió de l'espectròmetre de masses ordenats d'acord amb la seua relació m/z.

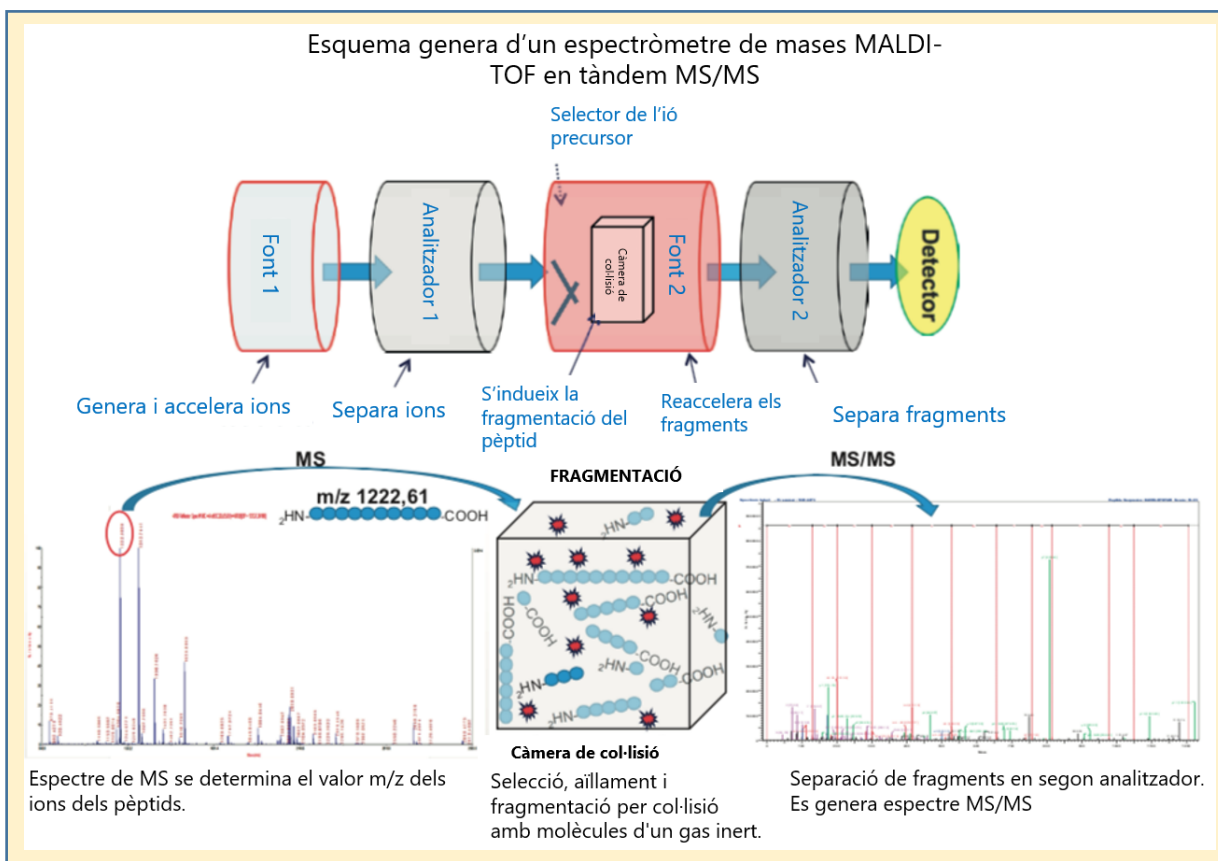


Figura 1. Esquema general d'obtenció d'espectre de fragmentació MS/MS.

Un espectre de MS/MS és un puzzle en el qual no solen estar totes les peces. L'èxit d'una interpretació *de novo* dependrà molt de la "qualitat" de l'espectre que estem analitzant.

Encara que les normes són clares, els espectres no sempre estan complets i no sempre és fàcil. Cal anar amb compte perquè ens podem trobar residus isobàrics i combinacions de residus amb masses similars. A pesar que pràcticament sempre s'usa programari específic per a resoldre'ls, és bo entendre com funciona el procés.

Els **fragments peptídics més comuns** observats en col·lisions de baixa energia són els **ions, b i y** (obtinguts en fragmentar en l'enllaç peptídic) i **ions a** (**fig.2**).

Els pèptids **no es fragmenten seqüencialment**, és a dir, el primer esdeveniment de fragmentació no comença en l'extrem amino i continua seqüencialment eliminant un residu cada vegada. Els **esdeveniments de fragmentació són aleatoris i no seqüencials**.

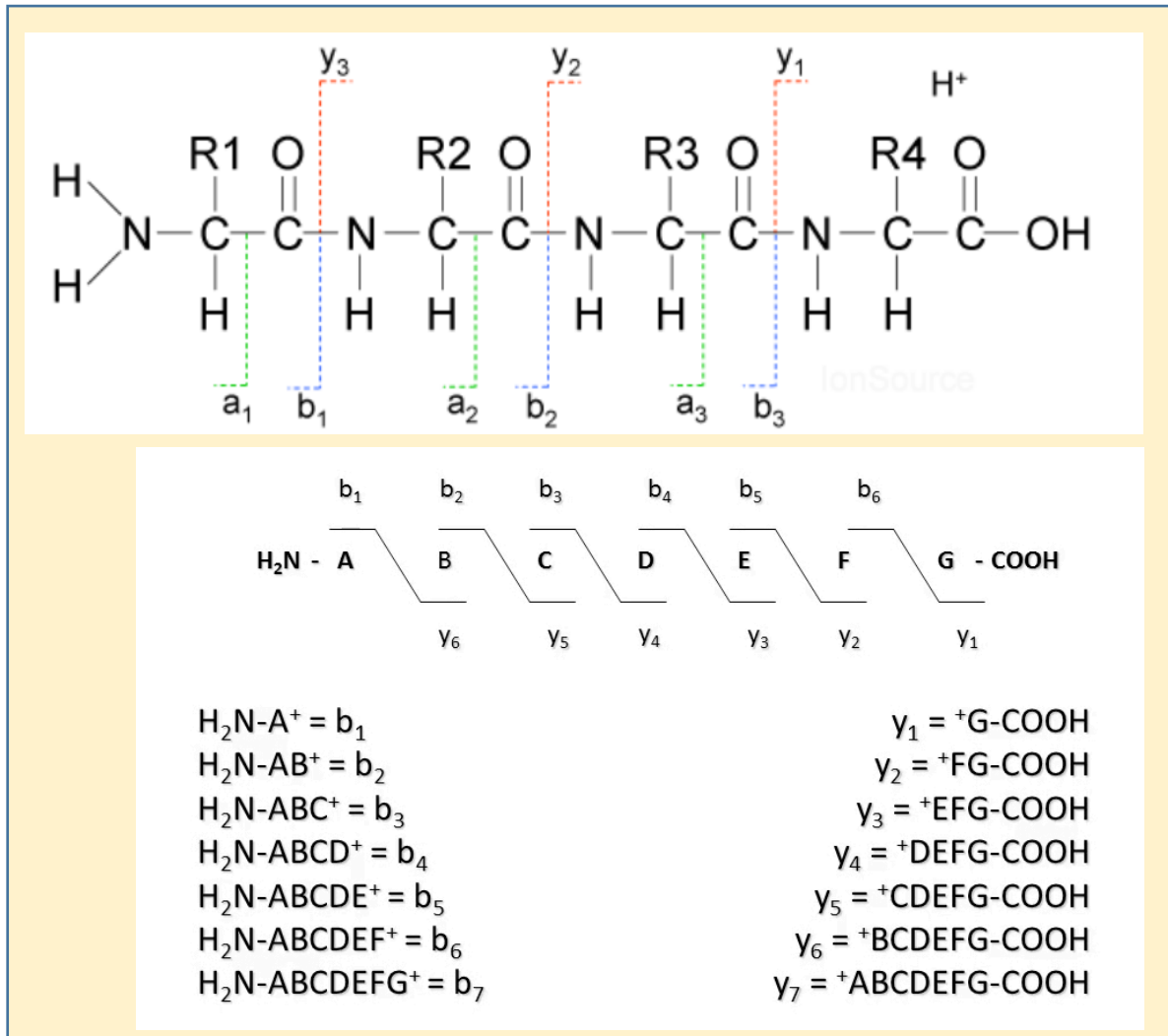
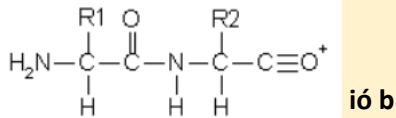


Figura 2. Tipus d'ions principals obtinguts en la dissociació induïda per col·lisions.

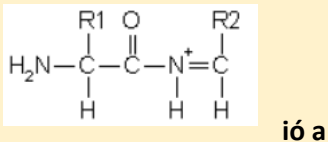
Els **ions b** corresponen als fragments de l'extrem amino. Els pics del fragment b es numeren des de l'extrem amino fins a l'extrem carboxil. *El fragment que conté només l'aminoàcid amino terminal es denomina b1. El fragment que conté els dos primers aminoàcids amino terminals es denomina ió b2, i així successivament. La nomenclatura és fàcil de seguir.*

Massa d'ions **b** = Σ (massa dels residus) + 1 (H⁺) del Nt



Els **ions a** ocorren amb **menor freqüència i abundància** que els ions **b**.

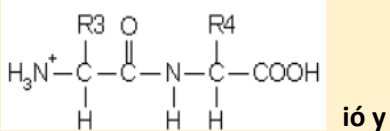
Massa d'ió **és a=b-28u** (la massa del carbonil, C=O)



Els **ions a** s'usen sovint com un **ió diagnòstic per a confirmar els ions b**.

De manera similar, els grups d'ions de **fragments de pèptids** que van des de l'extrem **C-terminal** es denominen "**ions y**". Els **pèptids tríptics** tendeixen a ser més bàsics en l'extrem C-terminal, i per això en general s'identificaran **millor els ions y** que els b.

Massa d'ions **y** = Σ (massa dels residus) + 19 (OH+H+H)



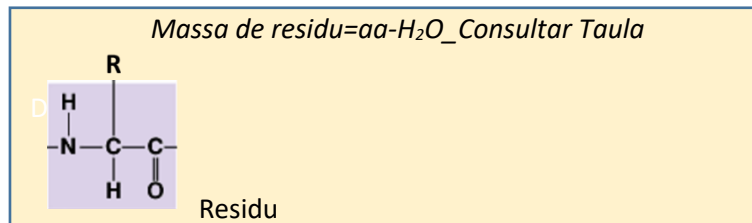
y1 indicarà el residu **Ct**. En cas de pèptids tríptics, la presència de l'ió **147** indica que el pèptid acaba en **Lys** o **175** si es tracta d'**Arg**.

$$y1=R = 156.1 + 19 = 175.1$$

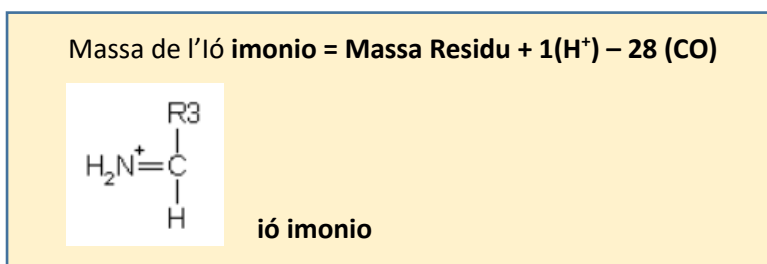
$$y1=K = 128.09 + 19 = 147$$

La seqüència del pèptid es determinarà per la **diferència de massa entre els pics** de l'espectre de MS/MS. Els ions **y** i **b** es trobaran **entremesclats**. La seua identificació permetrà determinar la seqüència, tant cap a en Nt com al Ct.

Cal consultar la **taula I** contínuament per a poder **identificar el residu** que es troba entre els diferents ions **y** o **b**. Una vegada tinguem un ió identificat anirem calculant la diferència entre pics contigus buscant aquell que coincidisca amb el valor de la columna que ens **indicarà el residu** que es troba entre aquests ions.



A més, és possible observar **ions imonio** en l'extrem inferior de l'espectre que **poden donar una pista sobre la composició d'aminoàcids d'un pèptid**. No sempre es detecten tots els ions imonio. Aquests ions són fragments "interns" d'un sol **residu** que perden el grup C = O. Per exemple, per a la lisina que la massa monoisotòpica és d'aproximadament 128 Daltons, el seu ió imonio seria: $MRi + 1 - 28 = 128 + 1 - 28 = 101$



Regles generals per a la seqüenciació

1. En el rang de menor massa es poden detectar els ions imonio que poden donar una idea de la composició d'aminoàcids.
2. Per a saber si el pèptid tríptic **acaba en K o en R** busca els ions diagnòstic **y1 en l'extrem inferior de l'espectre; 147 indica Lys i 175 Arg.**
3. Per a iniciar un projecte de seqüenciació *de novo*, es comença habitualment en l'extrem de massa alta de l'espectre. La seqüència del pèptid es determinarà per la **diferència de massa entre els pics** de l'espectre de MS/MS.
5. Una vegada es coneix la massa d'un ió b o y, l'ió y o b corresponent es pot calcular usant les següents fórmules i buscar l'ió en l'espectre.

(Massa precursor+H)¹⁺=b+y Perquè un ió es detecte, ha d'estar protonat, per tant:

$$y = (M+H)^{1+} - b + 1$$

$$b = (M+H)^{1+} - y + 1$$

Hi ha diferents maneres d'abordar una seqüenciació *de novo*. Es pot començar seguint els **ions y**, ja que aquests fragments **es detecten millor** com s'ha mencionat anteriorment.

1. Busqueu en l'espectre un pic intens, calculeu la diferència entre pics contigus. El que coincidisca amb el valor de la columna (massa del residu) de la taula I ens **indicarà el residu** que es troba entre aquests ions.

penúltim ió $y_{n-1} = (M+H)^{1+}$ - massa residu

Aquest ió y ha de trobar-se entre la massa de residus d'aminoàcid més petita i la més gran, segons taula I, és a dir, entre 57-186 uma.

$(M+H)^{1+}$ - ió observat = massa del residu

2. Una vegada que es troba un **ió y**, calculeu l'ió b corresponent, busqueu-lo, i etiqueteu-lo en l'espectre.
 $b = (M+H)^{1+} - y + 1$
3. Continueu seguint la sèrie d'**ions y** fins a l'extrem inferior de l'espectre de masses. Una vegada s'aconsegueixca l'extrem inferior i s'arriba al y1 (147 o 175) construïu la sèrie d'ions.

Per a cada ió **y** podeu buscar l'ió **b** corresponent o realitzar el procés independentment. Totes aquestes dades han d'encaixar per a ajudar a reafirmar les assignacions.

4. A més, cada vegada que s'identifique **un ió b**,
A -28u es trobarà l'ió **a** (no sempre és present)
A vegades també hi ha **pèrdues d'amoníac i aigua**, -17 i -18u respectivament.
5. Rares vegades s'observa el fragment b1, i això dificulta determinar l'ordre dels dos primers aminoàcids N-terminal en una seqüència peptídica, però que es podrà resoldre amb els **ions y** més llargs

Més consells:

Compte amb els aminoàcids amb massa isobàrica

1. La **leucina i la isoleucina tenen masses isobàriques i no es poden diferenciar** en una col·lisió de baixa energia.
2. La **lisina i la glutamina tenen masses quasi isobàriques**, 128,09496 i 128,05858 respectivament.

Pèrdua d'amoníac i aigua

1. A vegades alguns fragments d'ions **perden amoníac -17 o aigua, -18**.

Encara que en aquest exercici només seguirem els ions y, b, a i els possibles ocasionats per pèrdua d'aigua o grup amoni, també es poden detectar fragments interns. Hi ha taules que ajuden a la seua identificació, però per simplificar no les usarem en aquest exemple. Aquests fragments ens poden ajudar a resoldre conflictes en casos d'espectres de baixa qualitat.

EXEMPLE 1

Començarem amb un exemple de MS/MS obtingut en l'anàlisi de la banda LS (**Fig.3**)

Després de realitzar l'espectre de MS del digerit tríptic, s'han seleccionat els ions més intensos per a ser fragmentats.

Per a aquest exemple analitzarem l'ió 1021,52 que correspondrà a un pèptid de massa 1020,52

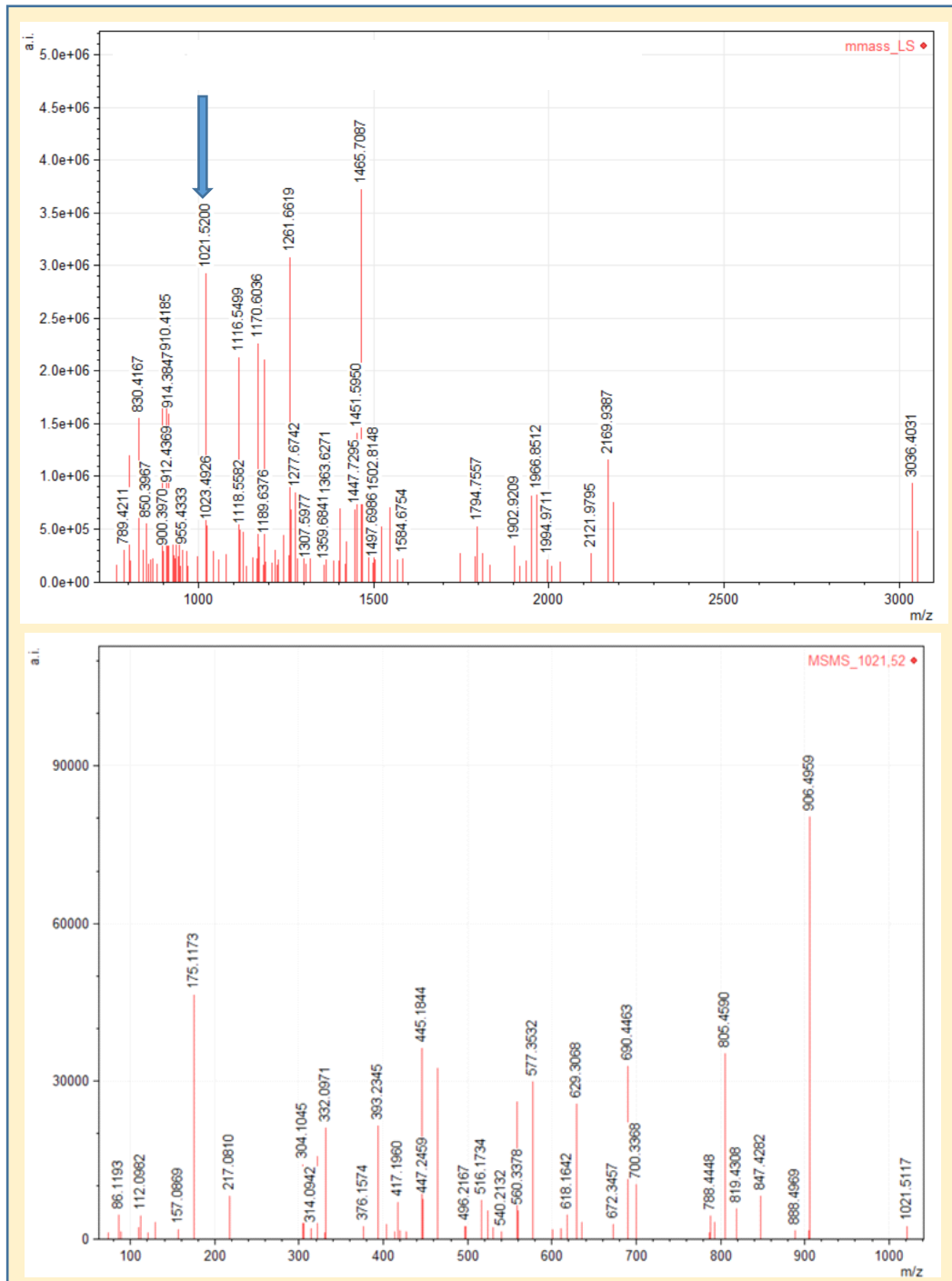


Figura 3. Espectres de masses obtinguts en analitzar la banda LS. A) Espectre de masses B) espectre de MS/MS de l'ió 1021.52

L'espectre de MS/MS el podeu analitzar en el programa mmass (arxiu MSMS_1021.52.txt) o directament a partir de les imatges ampliades i la llista de pics que es troben al final del document.

Podem començar identificant la **sèrie y** que correspon als fragments Ct i que en incorporar els residus bàsics (arginina o lisina) sol ser la més intensa.

Per a començar identificarem si el pèptid acaba en R o K buscant l'ió diagnòstic (175 Arg; 147 Lys).

En aquest cas detectem l'ió **175** que indica que el pèptid acaba en arginina i, per tant, **y1=R**.

Ara anem a la zona de m/z més alta per a identificar els fragments **y** més llargs. Per a això es calculen les diferències entre pics buscant a quin residu correspondria aquesta diferència de massa (Taula I).

Taula I. Valors de massa dels residus i ions imonio.

Aminoàcid	Codi de 3 lletres	Codi d'1 lletra	Massa del residu	Ió imonio
Glicina	Gly	G	57,02	30
Alanina	Ala	A	71,04	44
Serina	Ser	S	87,03	60
Prolina	Pro	P	97,05	70
Valina	Val	V	99,07	72
Treonina	Thr	T	101,05	74
Cisteïna	Cys	C	103,01	76
Isoleucina	Ile	I	113,08	86
Leucina	Leu	L	113,08	86
Asparagina	Asn	N	114,04	87
Àcid aspàrtic	Asp	D	115,03	88
Glutamina	Gln	Q	128,06	101
Lisina	Lys	K	128,09	101
Àcid glutàmic	Glu	E	129,04	102
Metionina	Met	M	131,04	104
Histidina	His	H	137,06	110
Fenilalanina	Phe	F	147,07	120
Arginina	Arg	R	156,01	129
Tirosina	Tyr	Y	163,06	136
Triptòfan	Trp	W	186,08	159
Cisteïna carbamidometilada			160.03	
Metionina oxidada			147.04	

Per ex.

Comencem amb el 1021,5 que correspon al precursor protonat. Com en aquests moments no sabem la longitud del pèptid, el nomenarem com a n.

Calculem la diferència entre pics anteriors

$1021-906=115$ segons la taula **correspondria a un aspàrtic**, per tant, sembla que el residu Nt siga un Asp.

n	Dxxxxxx...R	1021,5
y_{n-1}	xxxxxx...R	906,49

Continuem buscant diferències entre pics (si es tracta d'ions y solen ser més intensos, però, així i tot, podem calcular la diferència entre tots els pics següents).

$906-888=18$ (això indicaria una deshidratació n-H₂O)

$906-847=59$

$906-829=77$

$906-805=101.03$ correspondria a una Thr

n	DTxxxxxx...R	1021,5
y_{n-1}	Txxxxxx...R	906,49
y_{n-2}	xxxxxx...R	805,46

Cerca de l'ió y_{n-3}

$805-788=17$ (això correspondria a una perduda de grup amoni $y_{n-1}-17$)

$805-700=105$

$805-690=115$ que correspondria a D

n	DTDxxxxxx...R	1021,49
y_{n-1}	TDxxxxxx...R	906,49
y_{n-2}	Dxxxxxx...R	805,46
y_{n-3}	xxxxxx...R	690,43

Cerca de l'ió y_{n-4}

$$690.86-672.74=18$$

$$690.86-629.68=61$$

$$690.86-577.69=113$$

690-672=18 (això correspondria a una deshidratació $y_{n-3}-H_2O$)

$$690-635=55$$

$$690-629=61$$

$$690-611=79$$

$$690-601=89$$

690-577=113 (aquesta diferència de massa correspondria a L o I és impossible diferenciar-los)

n	DTDI/Lxxxxx...R	1021,49
y_{n-1}	TDI/Lxxxxx...R	906,49
y_{n-2}	DI/Lxxxxx...R	805,46
y_{n-3}	I/Lxxxxx...R	690,43
y_{n-4}	xxxxx...R	577,34

Cerca de l'ió y_{n-5}

$$577-560=17 (y_{n-4}-NH_3)$$

$$577-558=19$$

$$577-540=37$$

$$577-530=47$$

577-464=113 (de nou aquesta diferència de massa correspondria a L o I és impossible diferenciar-los)

n	D T D I/L I/L xxxxx...R	1021,49
y_{n-1}	T D I/L I/L xxxxx...R	906,49
y_{n-2}	D I/L I/L xxxxx...R	805,46
y_{n-3}	I/L I/L xxxxx...R	690,43
y_{n-4}	I/L xxxxx...R	577,34
y_{n-5}	xxxxx...R	464,26

Cerca de l'ió yn-6

464-447=17 ($y_{n-5}-NH_3$)

464-445=19

464-427=37

464-419=45

464 -393=71 Ala

n	D	T	D	I/L	I/L	A	xxx...R	1021,49
yn-1	T	D	I/L	I/L	A	xxx...R	906,49	
yn-2		D	I/L	I/L	A	xxx...R	805,46	
yn-3			I/L	I/L	A	xxx...R	690,43	
yn-4				I/L	A	xxx...R	577,43	
yn-5					A	xxx...R	464,26	
yn-6						xxx...R	393,22	

Cerca de l'ió yn-7

393-376 =17 ($y_{n-6}-NH_3$)

393-332=61

393-322=71 Correspon a Ala

n	D	T	D	I/L	I/L	A	A	XX...R	1021,49
yn-1	T	D	I/L	I/L	A	A	XX...R	906,49	
yn-2		D	I/L	I/L	A	A	XX...R	805,46	
yn-3			I/L	I/L	A	A	XX...R	690,43	
yn-4				I/L	A	A	XX...R	577,43	
yn-5					A	A	XX...R	464,26	
yn-6						A	XX...R	393,22	
yn-7							XX...R	322,10	

Per a identificar el següent residu, es continua igual, encara que tot apunta que ha de ser el y2 perquè el y1 (175) es troba pròxim. Així i tot, es pot provar de mesurar la diferència entre tots els pics:

322 -314=8

322-305=17 ($y_{n-7}-NH_3$)

322-304=18 ($y_{n-7}-H_2O$)

322 -217=105

322-175=147 Correspon a Phe

Precursor	D T D I/L I/L A A F R	1021,49
y8	T D I/L I/L A A F R	906,49
y7	D I/L I/L A A F R	805,46
y6	I/L I/L A A F R	690,43
y5	I/L A A F R	577,43
y4	A A F R	464,26
y3	A F R	393,22
y2	F R	322,10
y1	R	175,11

Es pot comprovar la presència d'aquests aminoàcids buscant els ions imonio en la taula I. En aquest cas, de D=88 T=74 I/L=86 A=44 F=120 R=129.

Es troben tots, fet que confirma la composició.

El resum d'aquestes assignacions es mostra en la figura 4.

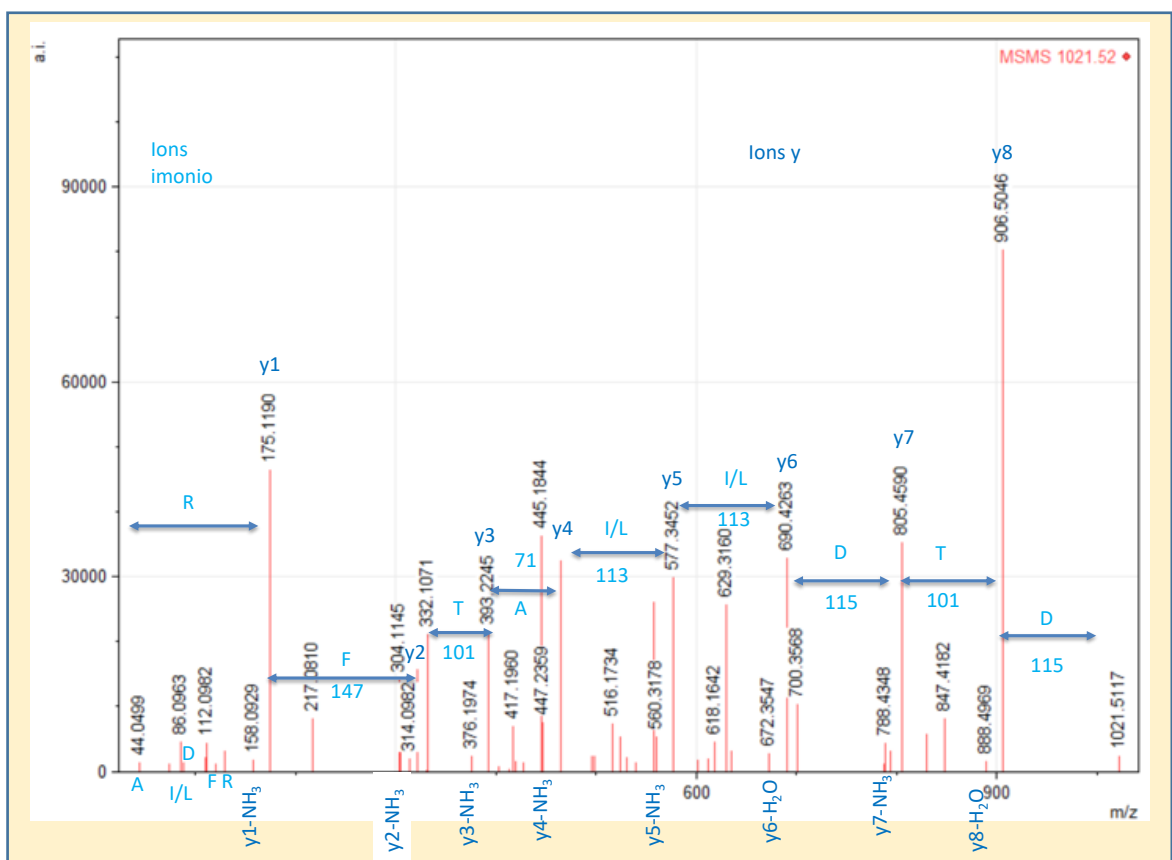


Figura 4. Identificació dels ions imonio i ions y. A partir de la diferència de massa entre els ions y, es pot deduir la seqüència del pèptid.

A més, es pot buscar la **sèrie b** de dues maneres, a partir dels **ions y** o analitzant la diferència entre els pics igual que s'ha fet per a la sèrie y.

Una vegada tenim identificat els ions y, es pot comprovar buscant els ions de la sèrie b

$$b = (M+H)^{1+} - y + 1$$

A partir de y1 es calcula el b8. El precursor és 1021,5 i l'ió y1 és 175

$$b8=1021,5-175.2+1=847$$

Es busca el pic en l'espectre i es marca.

El càlcul per a la determinació de la resta d'ions b es resumeix en taula II i es marquen en la figura 5.

Taula II. Càlcul d'ions b a partir dels ions y identificats prèviament.

seqüència	ió y	massa	càlcul d'ió b	ió b	seqüència
R	y1	175,11	= 1021,5-175,1+1	847	b8
FR	y2	322,18	= 1021,5-322,2+1	700	b7
AFR	y3	393,22	= 1021,5-393,2+1	629	b6
AAFR	y4	464,53	= 1021,5-464,5+1	558	b5
LAAFR	y5	577,35	= 1021,5-577,3+1	445	b4
ILAAFR	y6	690,43	= 1021,5-690,4+1	332	b3
DILAAFR	y7	805,46	= 1021,5-805,5+1	217	b2
TDILAAFR	y8	906,50	= 1021,5-906,5+1	116	b1

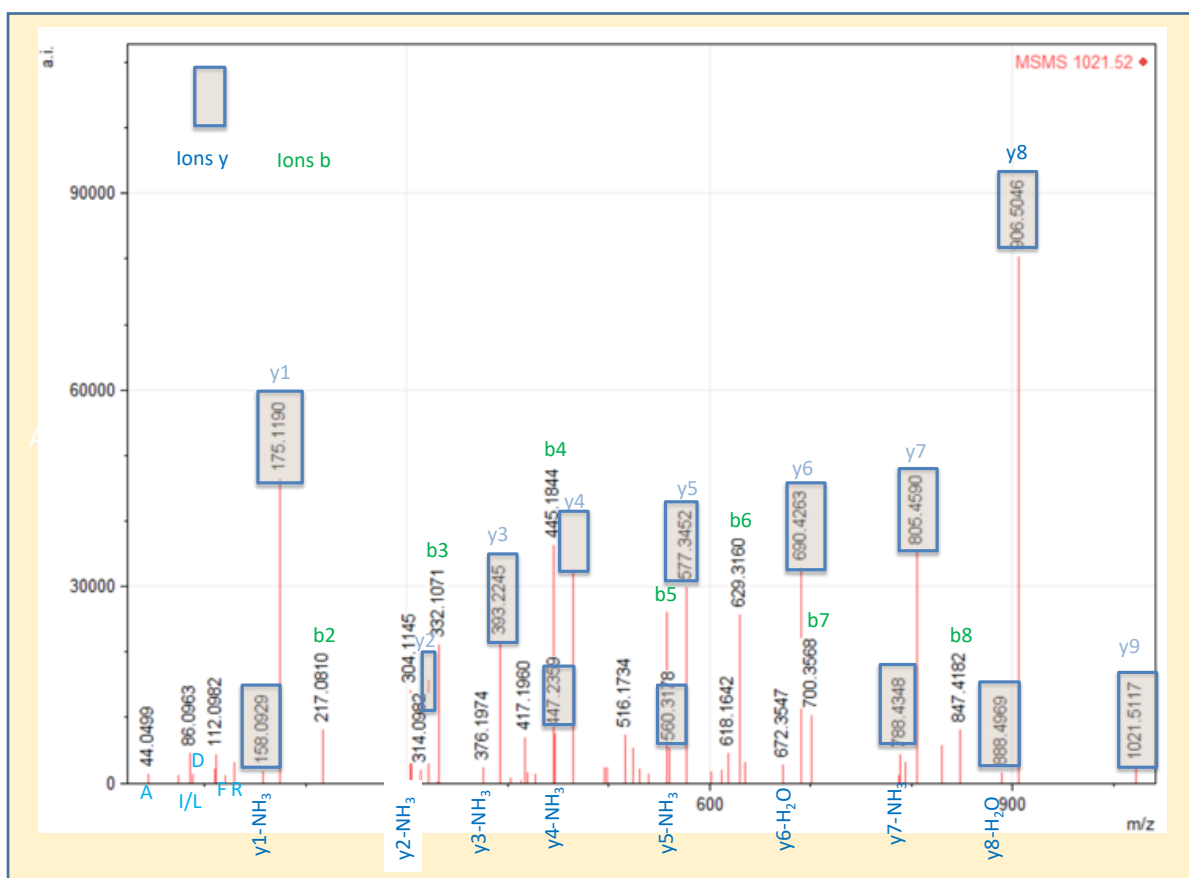


Figura 5. Identificació dels ions b a partir dels ions y identificats prèviament.

A més, es pot buscar els ions b que perden aigua (b-H₂O) (b-18) i els ions a (b-carbonil) (b-28). Els pics identificats en l'espectre es mostren acolorits en la **taula III** i marcats en la **figura 6**.

Taula III. Càlcul d'ions b a partir dels ions y identificats prèviament.

	ions b	b-18 b-H ₂ O	b-28(CO) ions a
b2	217	199	189
b3	332	314	304
b4	445	427	417
b5	558	540	530
b6	629	611	601
b7	700	682	672
b8	847	829	819

Tot això confirmaria que l'assignació realitzada amb els ions y és correcta.

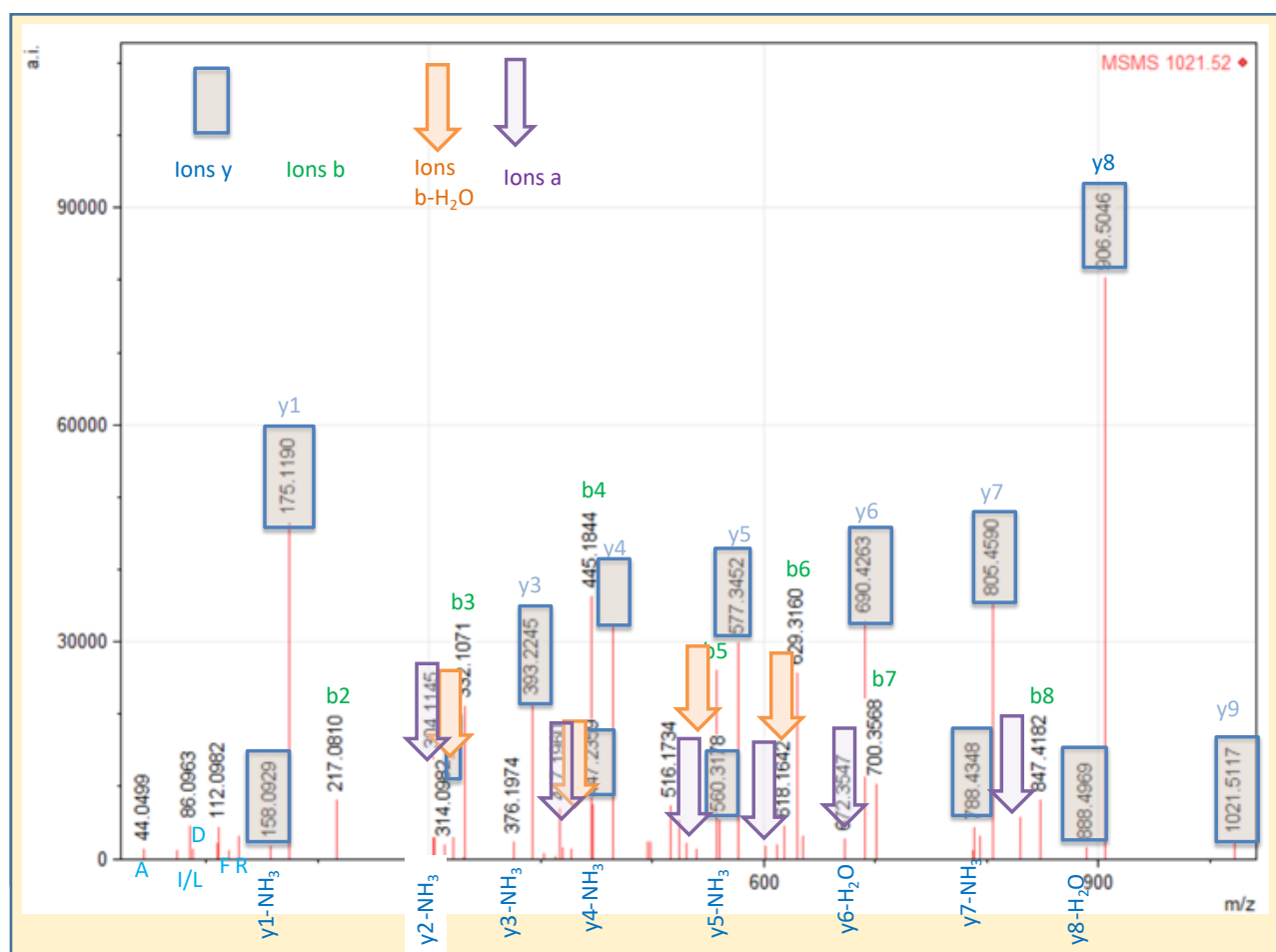


Figura 6. Assignació dels pics de l'espectre a l'ió corresponent.

De manera alternativa, en cas que no es detectaren bé tots els ions y, es podria procedir igual que abans amb els ions b, analitzant la **diferència entre pics que correspon a la massa d'un residu**.

Com l'ió y1 se sol visualitzar fàcilment, serà igualment fàcil de localitzar l'ió **b_{mésllarg}** (b8 en aquest cas)

b8= (Massa del precursor protonada-y1)+H

$$b8=1021,5-175,1+1=847,4$$

En cas que no s'haja detectat el y1 aniríem buscant la diferència entre el precursor i pics de massa superior per a buscar el b més llarg tenint en compte que:

Precursor protonat= b_{mésllarg} +18+aminoàcid en Ct

D'aquesta manera s'aniria buscant l'ió b més llarg mitjançant la diferència de massa entre pics.

$1021,5-847,42-18=156$, cosa que indicaria que el residu Ct és una arginina, com ja havíem comprovat per al y1

Una vegada tenim localitzat l'ió b més llarg (en aquest cas b8=847), es calcula la diferència amb els pics anteriors buscant la coincidència amb la massa d'algun residu segons la **taula II o I**, igual que hem fet per a la sèrie y.

Igual que abans poden aparèixer ions que **perden aigua -18, grup amoni -17** i fins i tot es pot calcular l'ió **a=b-28** (pèrdua del carbonil). Com que aquests càlculs ja estan mostrats en la taula IV, no es realitzaran de nou.

Com veiem que la taula I la massa dels residus va des del 57 al 186, buscarem pics que es troben entre aquestes distàncies. A més, com que ja tenim identificats els **ions y**, no els tindrem en compte ara per a simplificar el càlcul.

Cerca de l'ió b7:

$847-700=147$ correspon amb una Phe

b8	FxxxxxxxR	847
b7	xxxxxxxR	700

Cerca de l'ió b6:

$$700-635=65$$

700-629=71 correspon a una Ala

b8	FAxxxxxR	847
b7	AxxxxxR	700
b6	xxxxxR	629

Cerca de l'ió b5:

629-558=71 Correspon de nou a una Ala

b8	FAAxxxxR	847
b7	AAxxxxR	700
b6	AxxxxR	629
b5	xxxxR	558

Cerca de l'ió b4:

558-496=62.1

558-445=113 correspon a Leu o lle

b8	F A A I/L xxxR	847
b7	A A I/L xxxR	700
b6	A I/L xxxR	629
b5	I/L xxxR	558
b4	xxxR	445

Cerca ió b3:

445-332=113 correspon a Leu o lle

b8	F A A I/L I/L xxR	847
b7	A A I/L I/L xxR	700
b6	A I/L I/L xxR	629
b5	I/L I/L xxR	558
b4	I/L xxR	445
b3	xxR	332

Cerca ió b2

332- 217=115 correspon amb un Asp

b8	F A A I/L I/L DxxR	847
b7	A A I/L I/L DxxR	700
b6	A I/L I/L DxxR	629
b5	I/L I/L DxxR	558
b4	I/L DxxR	445
b3	DxxR	332
b2	xR	217

L'ió b1 no se sol detectar. Amb la sèrie b habitualment no es pot completar els residus finals del Ct, però que sí que es podrien deduir de la resta de dades. D'una banda, se sap que és un pèptid que acaba en Arg, perquè es detecta $\gamma 1$ i l'ió imonio d'Arg. A més, si mirem els ions imonio, es detecta un pic a 74 que correspon a l'ió imonio de la Thr, per tant ja el tindríem identificat.

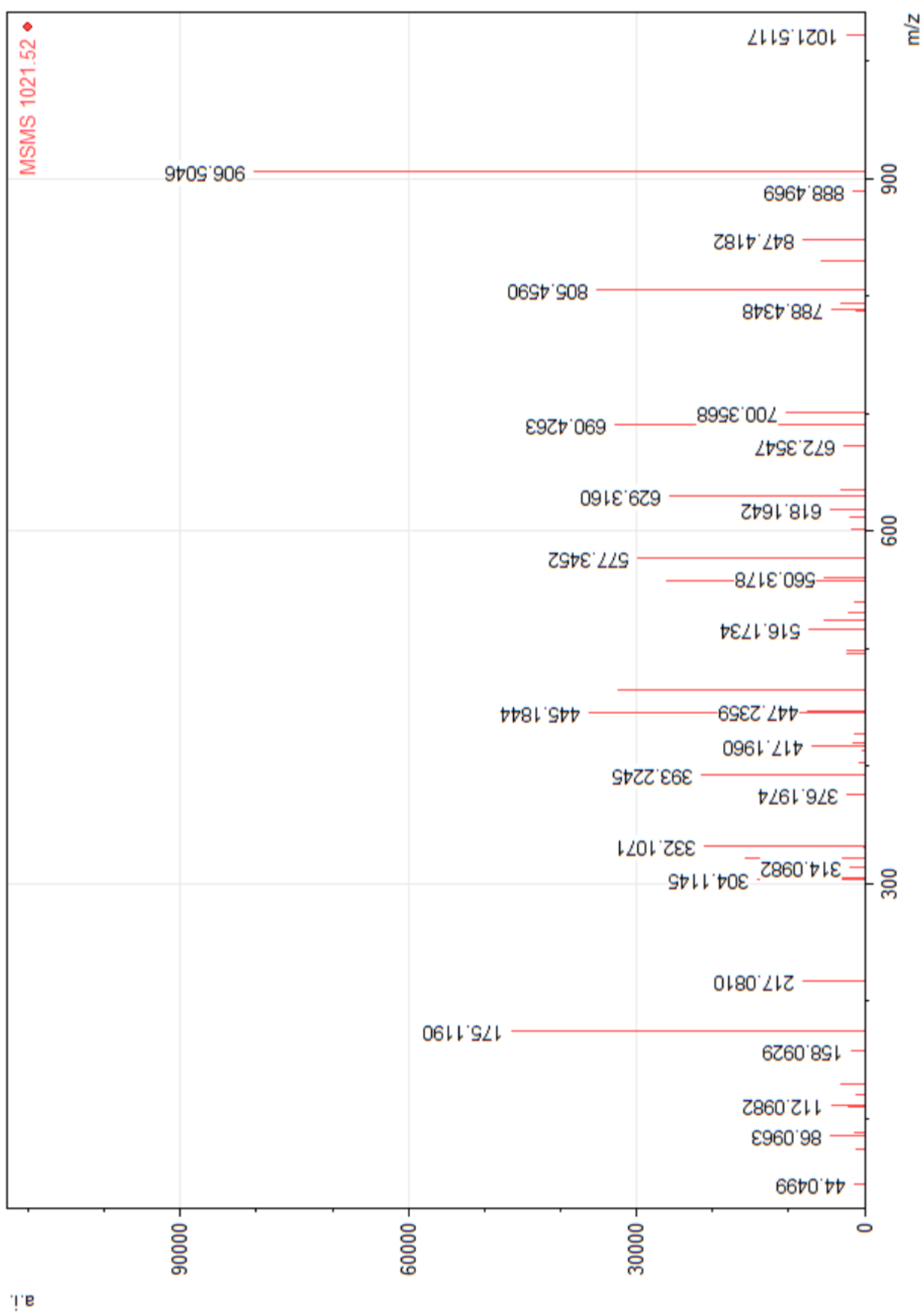
Proveu de fer-ho a partir dels espectres de MSMS del precursor **1021,52** ampliat a continuació.

Si voleu provar amb un altre espectre de seqüència desconeguda, s'adjunten al final la llista de pics i imatges de l'espectre de MSMS de l'ió 1465 del digerit de la banda LS. A l'aula virtual hi ha l'arxiu MSMS_1465.txt que es pot visualitzar en el programa mmass.

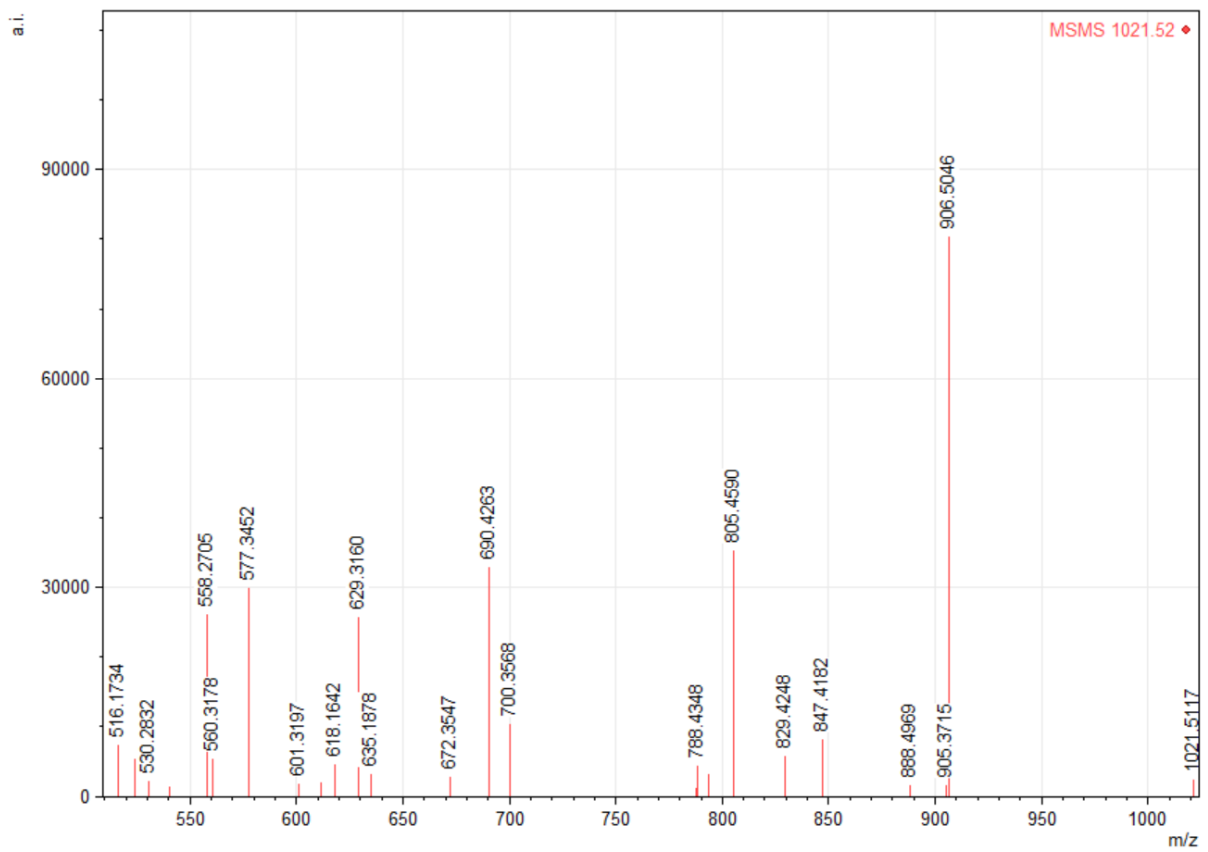
A l'Aula virtual també hi ha els arxius Sequencequery de cada ió, on es troba la llista de pics per a fer la identificació del MSMS usant el motor de cerca de MASCOT en la secció de Sequence Query.

http://www.matrixscience.com/cgi/search_form.pl?formver=2&SEARCH=SQ

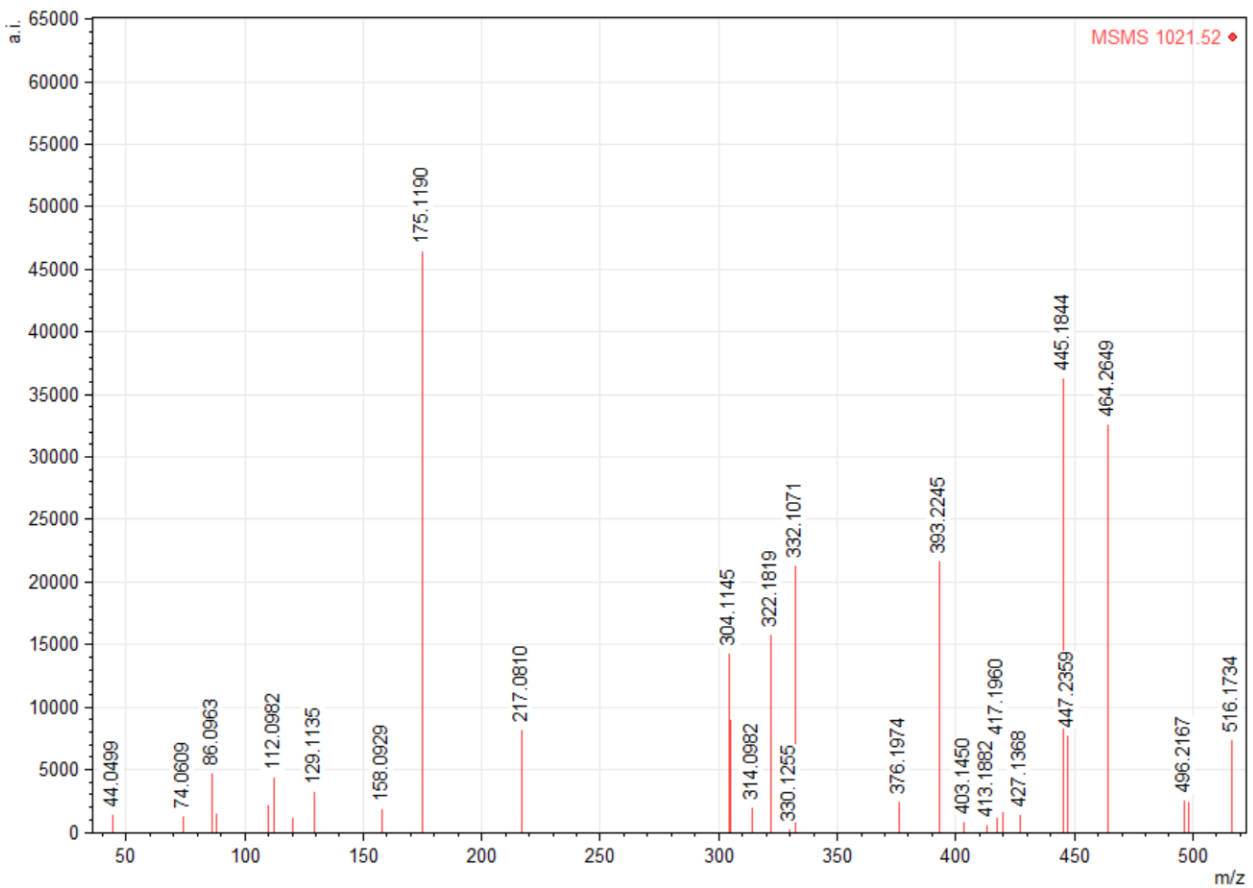
Espectre MSMS de l'ió 1021.52



Ampliació en rang superior de masses



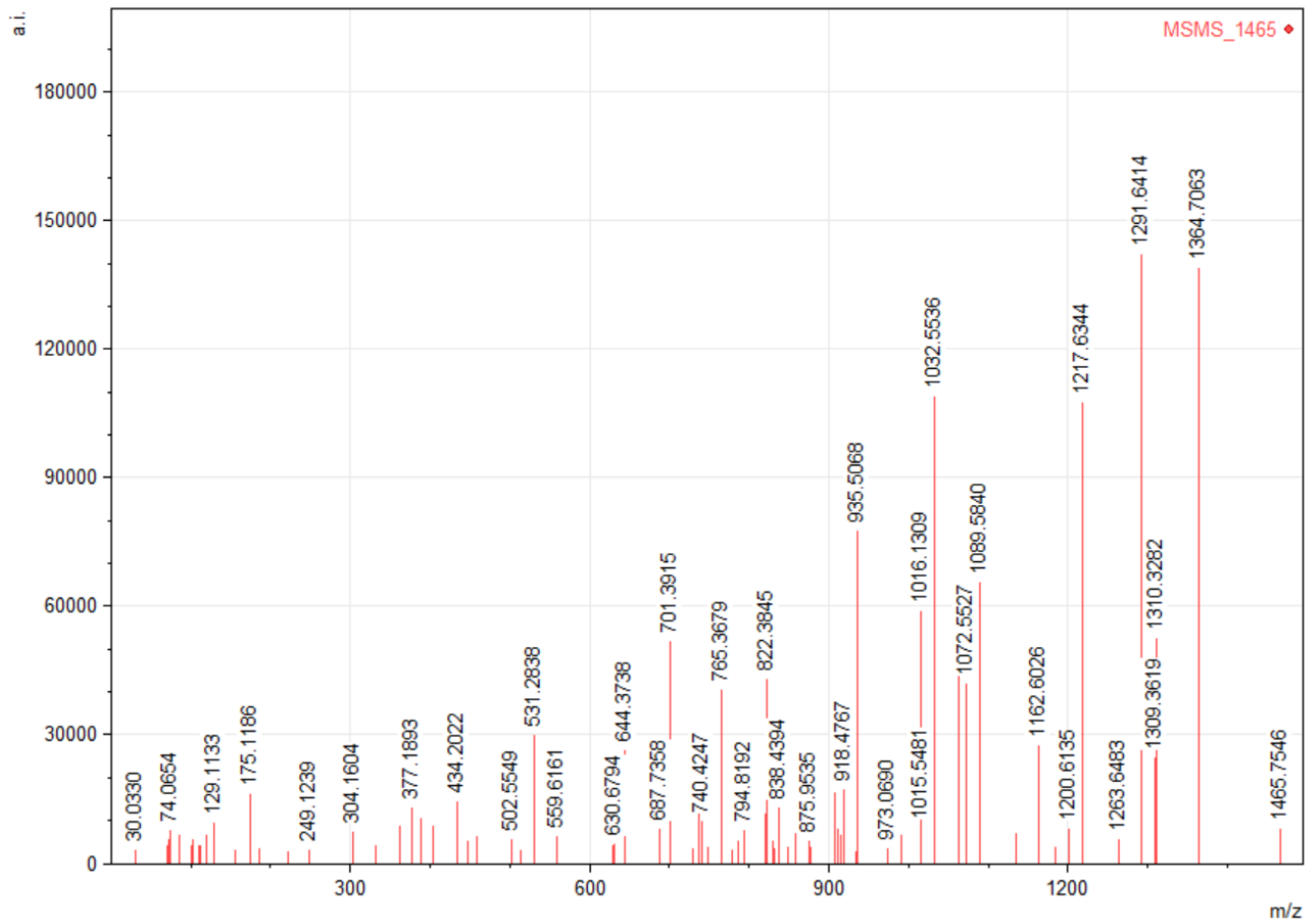
Ampliació en rang inferior de masses



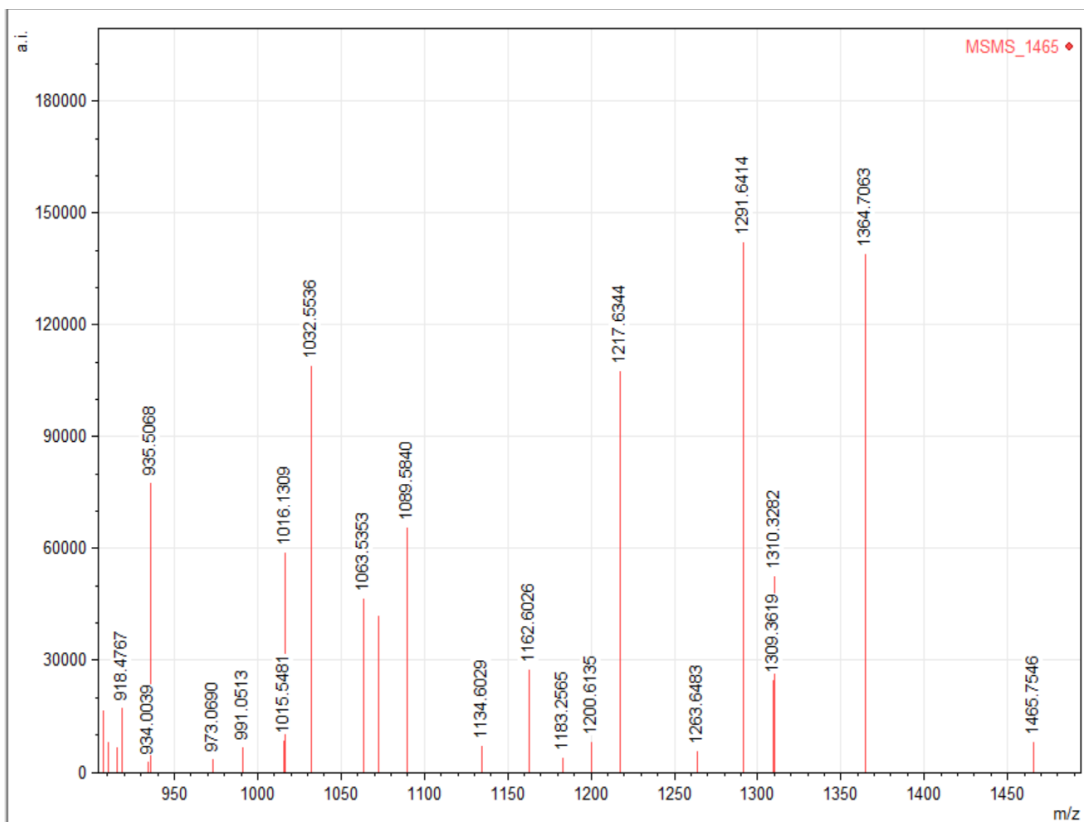
Lista de pics de l'espectre MS/MS del 1021.65

m/z	Intensitat	m/z	Intensitat
44,05	1400,91	496,22	2561,55
74,06	1316,21	498,16	2519,97
86,10	4739,62	516,17	7444,00
88,04	1587,79	524,21	5465,02
110,05	2258,85	530,28	2306,40
112,10	4435,02	540,26	1567,16
120,12	1261,35	558,27	26210,90
129,11	3266,59	560,32	5465,45
158,09	1906,21	577,35	29951,59
175,12	46476,70	601,32	1796,59
217,08	8261,35	611,30	2107,93
304,11	14338,41	618,16	4723,79
305,16	9023,27	629,32	25918,02
314,10	1998,48	635,19	3228,11
322,18	15792,89	672,35	2887,02
330,13	316,29	690,43	32941,83
332,11	21338,41	700,36	10548,65
376,20	2516,83	787,45	1307,81
393,22	21714,72	788,43	4513,89
403,14	825,30	793,35	3311,75
413,19	580,93	805,46	35382,45
417,20	7153,32	829,42	5817,73
419,81	1706,20	847,42	8183,02
427,14	1410,14	888,50	1681,54
445,18	36337,37	905,37	1670,86
447,24	7727,61	906,50	80492,34
464,26	32609,75	1021,51	2500,00

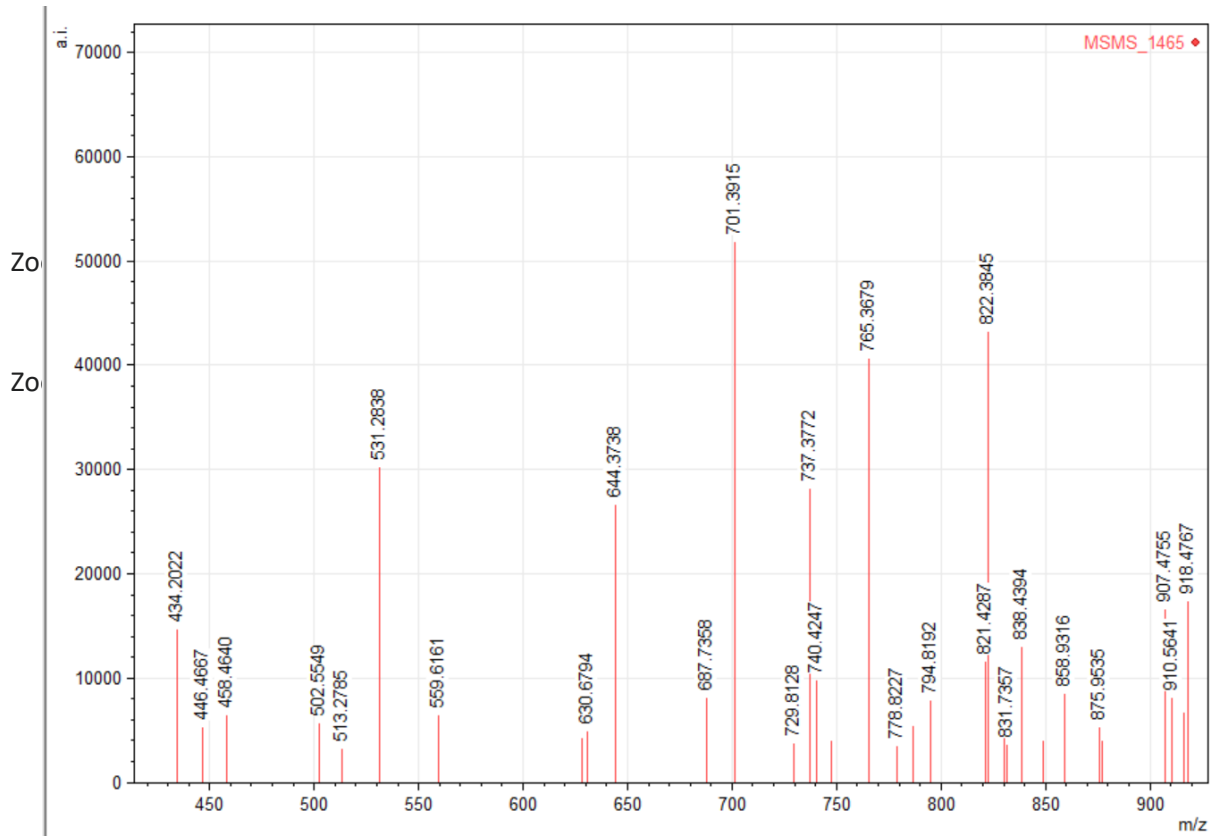
Espectre MS/MS de l'ió 1465



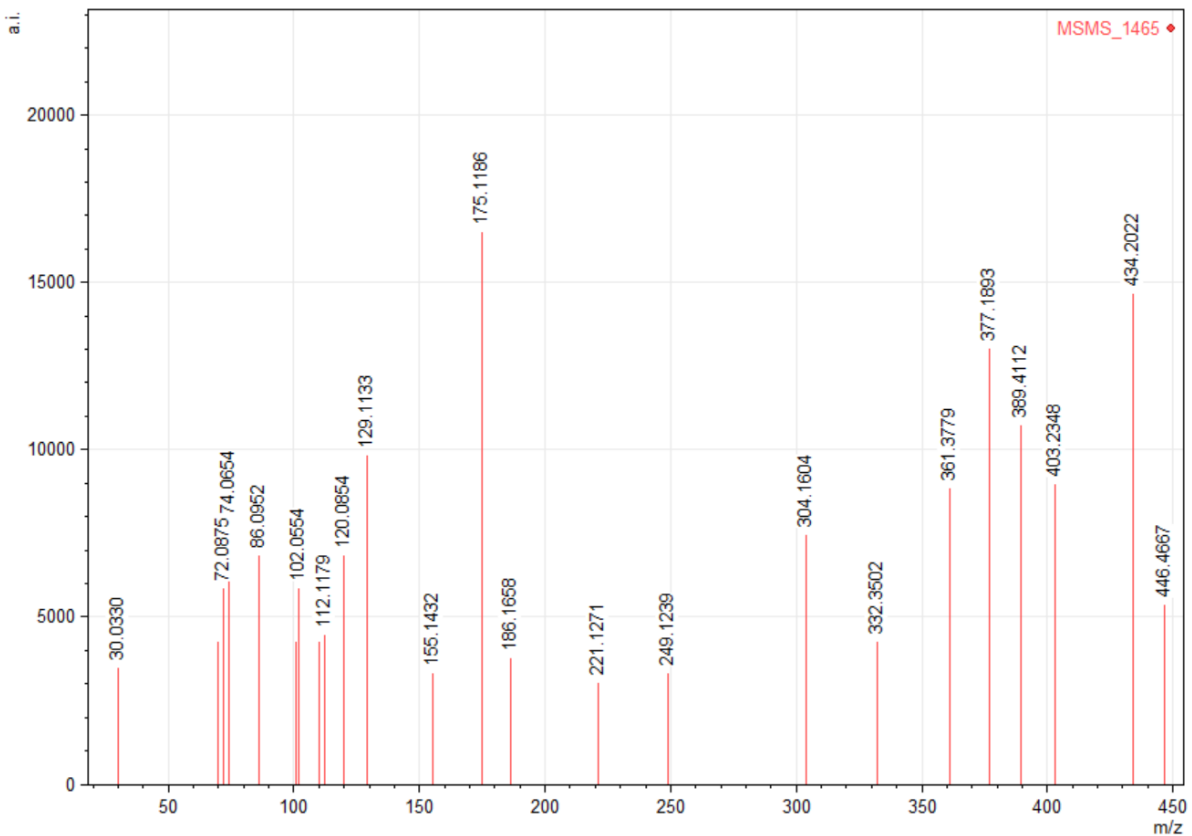
Ampliació en rang superior de masses



Ampliació en rang intermedi de masses



Ampliació en rang inferior de masses



Llista de pics de l'espectre MS/MS del 1465

m/z	Intensitat	m/z	Intensitat
30,03	3500,77	765,37	40693,82
70,07	4287,33	778,82	3514,94
72,09	5865,23	786,85	5407,57
74,07	7865,23	794,82	7957,63
86,10	6865,23	821,43	11695,19
101,08	4262,72	822,38	43325,00
102,06	5865,23	830,62	5555,66
110,07	4287,33	831,74	3664,21
112,12	4494,65	838,44	13076,12
120,09	6865,23	848,90	4098,53
129,11	9865,23	858,93	8590,55
155,14	3336,13	875,95	5366,90
175,12	16528,50	876,96	4103,27
186,17	3808,84	907,48	16700,94
221,13	3055,65	910,56	8168,23
249,12	3336,13	915,98	6780,42
304,16	7488,74	918,48	17408,87
332,35	4279,99	934,00	3017,67
361,38	8852,82	935,51	77846,10
377,19	13052,82	973,07	3784,46
389,41	10747,06	991,05	6718,50
403,23	8987,06	1015,55	8579,05
434,20	14692,92	1016,13	59053,13
446,47	5387,94	1032,55	109144,75
458,46	6553,41	1063,54	46661,22
502,55	5666,03	1072,55	42260,85
513,28	3258,46	1089,58	65960,31
531,28	30326,59	1134,60	7294,84
559,62	6526,93	1162,60	27688,65
628,31	4245,83	1183,26	4173,20
630,68	4911,06	1200,61	8311,59
644,37	26657,08	1217,63	107704,02
687,74	8177,71	1263,65	5778,09
701,39	51899,16	1291,64	142365,94
729,81	3828,94	1309,36	24900,58
737,38	28186,58	1310,33	52845,47
740,42	9888,15	1364,71	139146,88
747,83	3993,73	1465,75	8311,59