

A fluorescence micrograph showing a network of plant cells. The cell walls are stained with multiple fluorescent dyes, appearing in shades of red, green, and blue. The cells themselves are dark, creating a complex, interconnected pattern of bright, multi-colored lines.

Résolution de mélanges complexes :  
Matrices biologiques complexes

Delphine Pflieger  
LAMBE UMR CNRS 8587  
Université d'Evry Val d'Essonne

# **Illustrations en protéomique**

**Identification de peptides**

**Quantification relative d'échantillons**

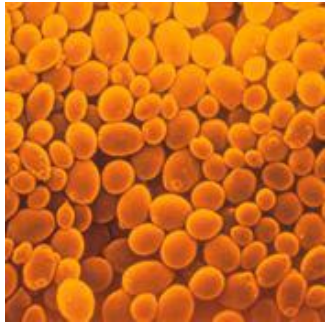
**1) en analyses exploratoires**

**2) en analyses ciblées**

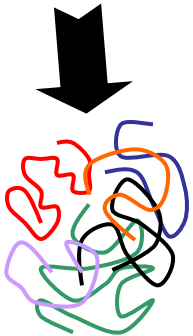
# Identification de peptides

# Différentes stratégies d'analyses d'échantillons protéiques

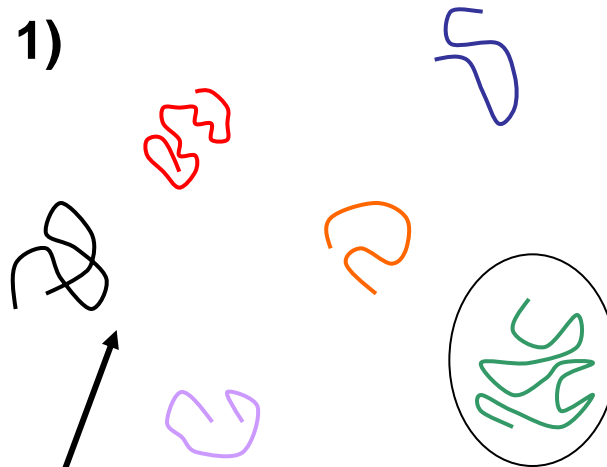
Extraction des protéines



[www.chez.comomgmethode.html](http://www.chez.comomgmethode.html)

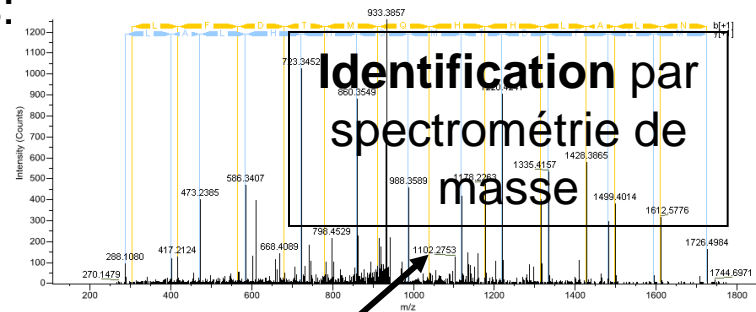


1)



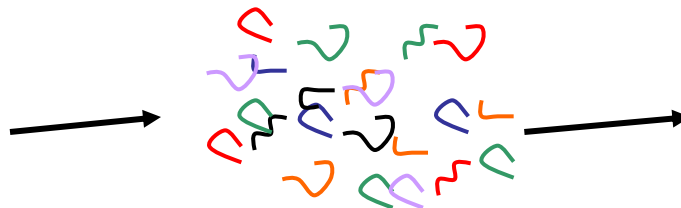
Séparation des protéines:  
2DE, techniques de  
séparation en solution

Digestion  
enzymatique



Identification par  
spectrométrie de  
masse

2)



Digestion du mélange  
protéique

Séparation des peptides  
(LC)

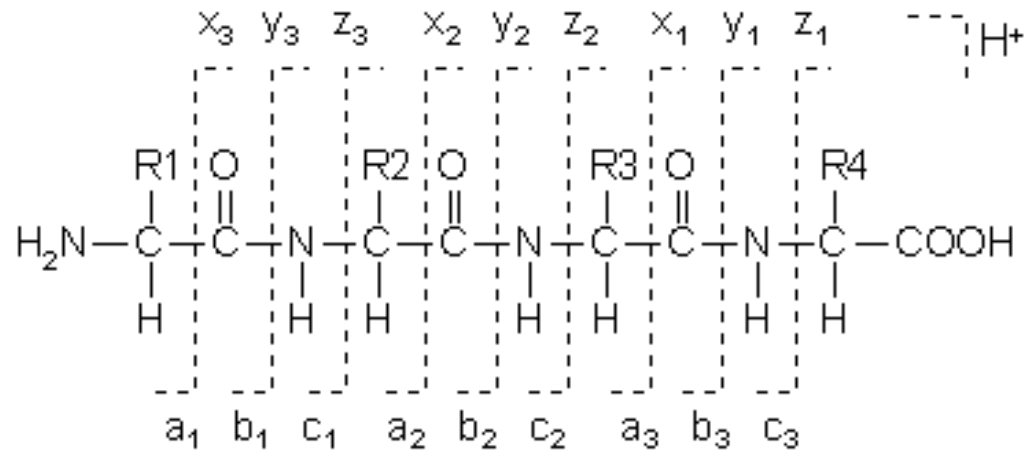
Fractionnement  
de l'échantillon  
(fractionnement  
subcellulaire)

# Règles de fragmentation des peptides

## Fragments contenant :

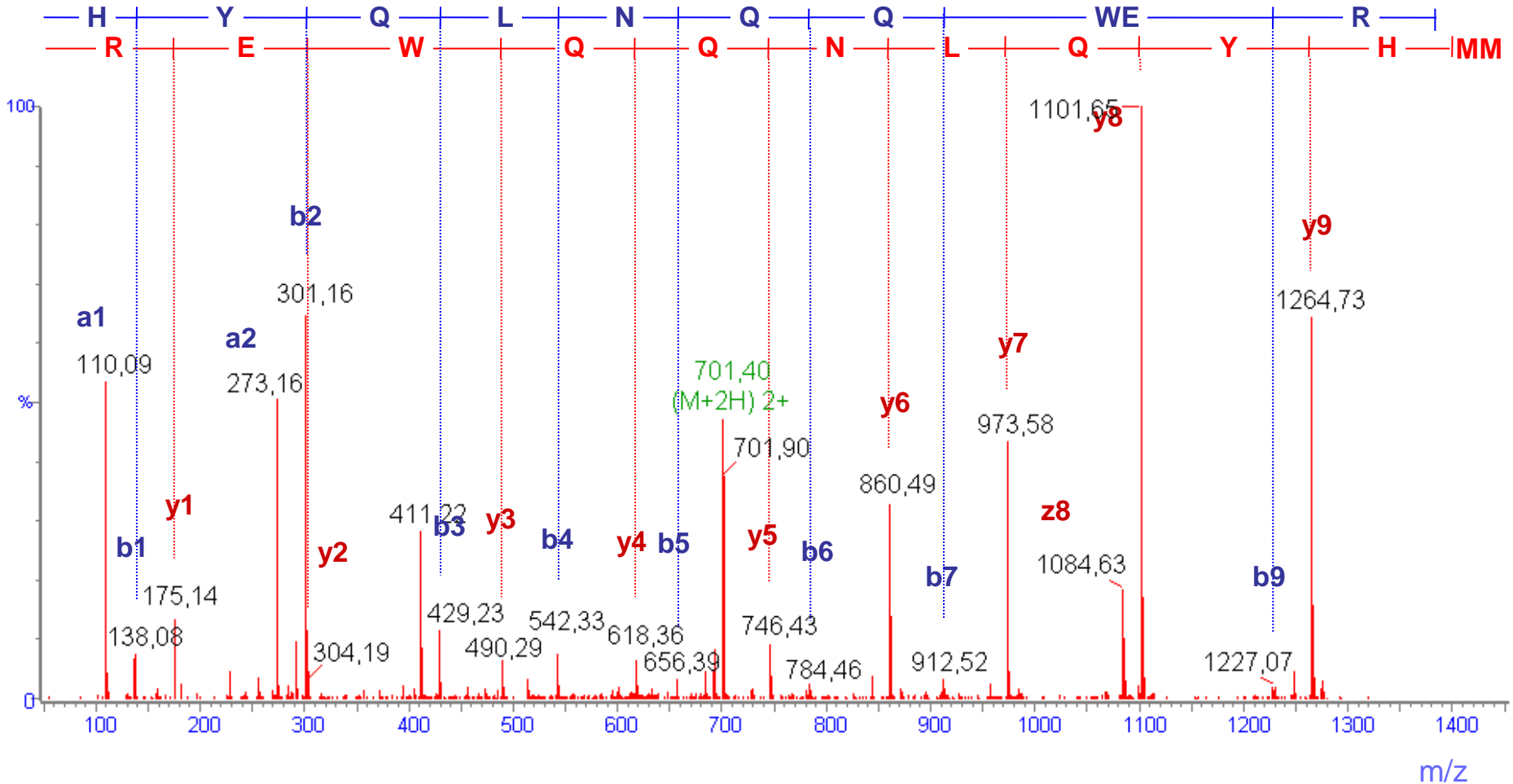
L'extrémité C-terminale du précurseur (x, y, z)

L'extrémité N-terminale du précurseur (a, b, c).



# Identification par informations de séquence

Fragmentation du peptide doublement chargé à  $m/z = 701,40$  Da

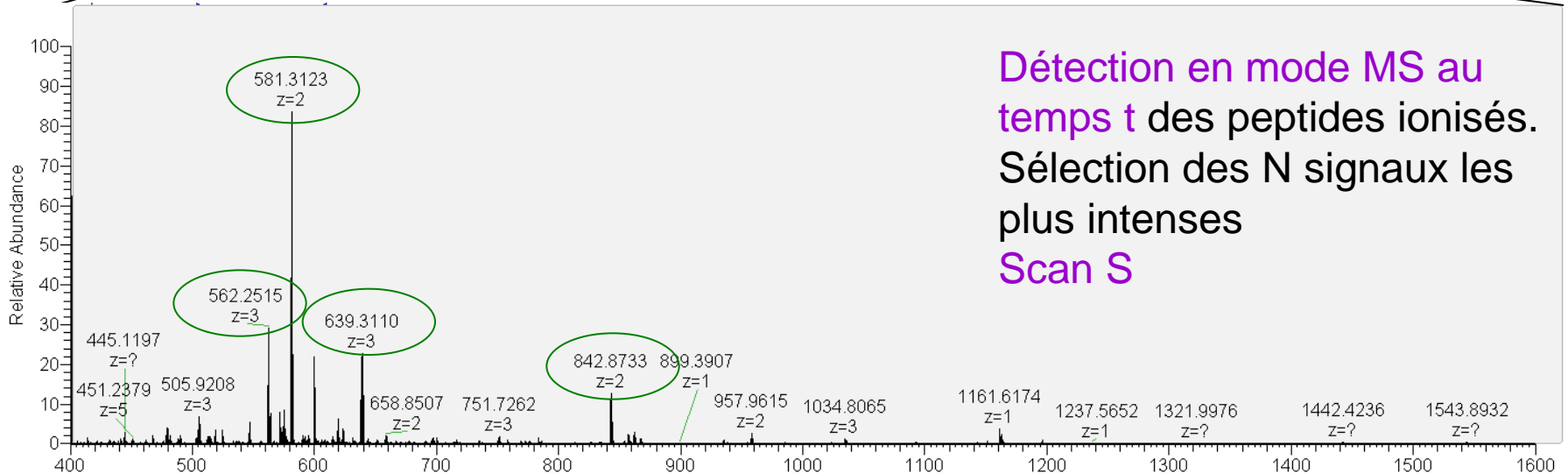
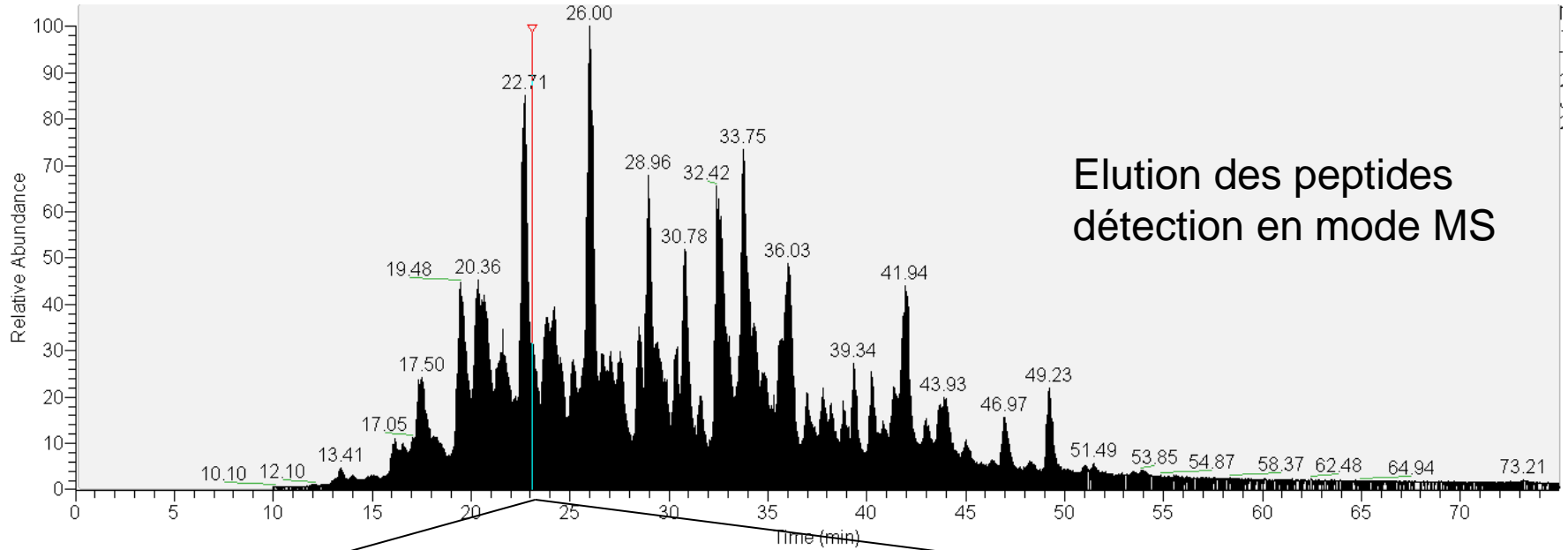


Identification de la séquence HYQLNQQWER

Séquence lue de gauche à droite sur la base de fragments contenant l'extrémité N-term

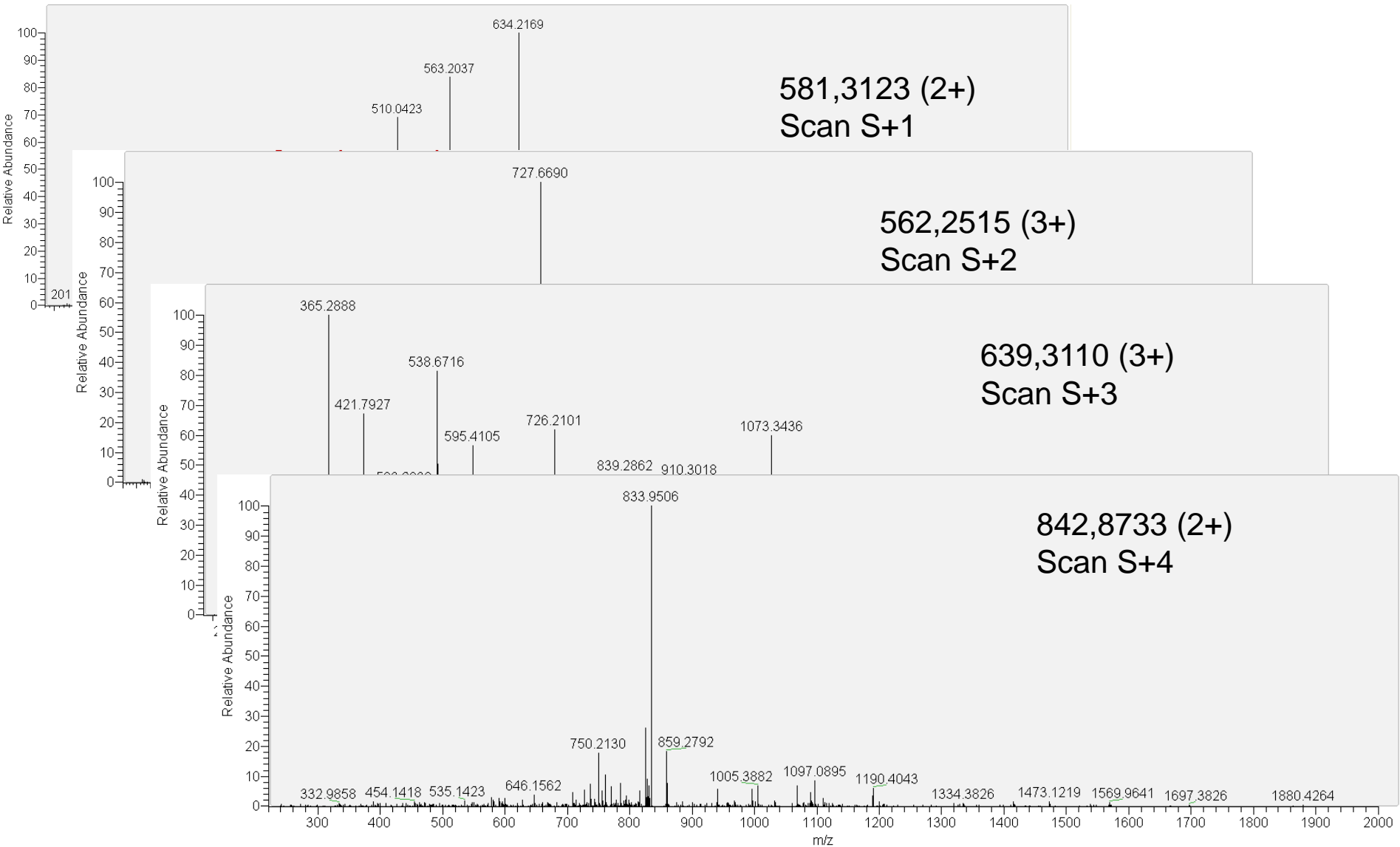
Séquence lue de droite à gauche sur la base de fragments contenant l'extrémité C-term

# Principe des acquisitions LC-MS/MS automatiques (1)



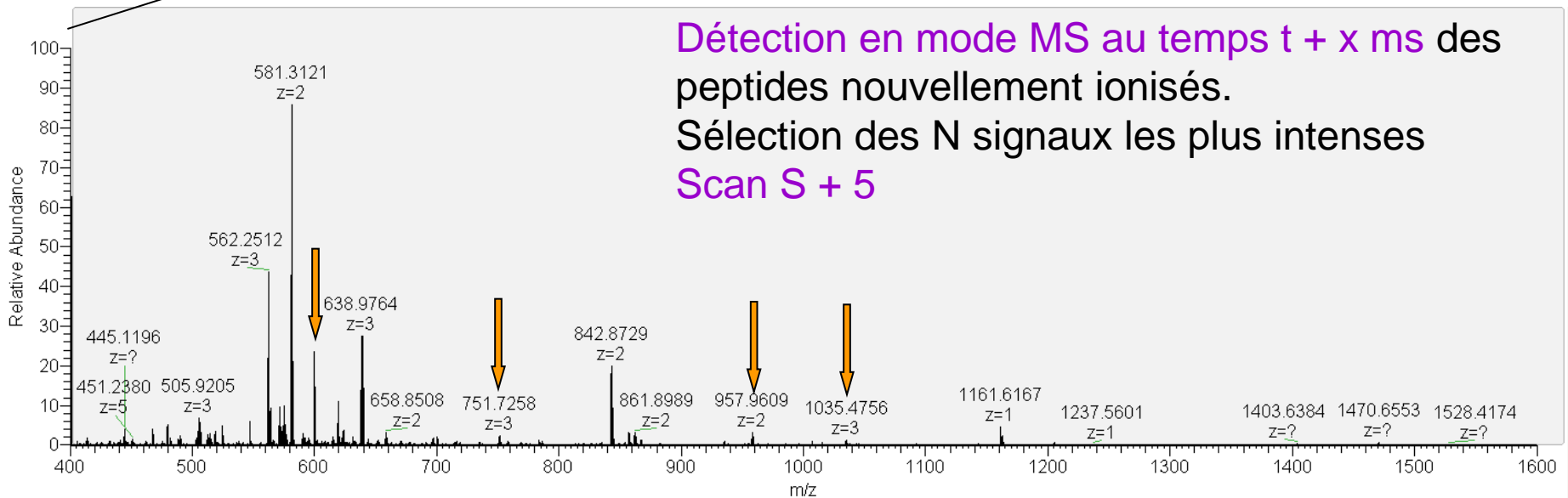
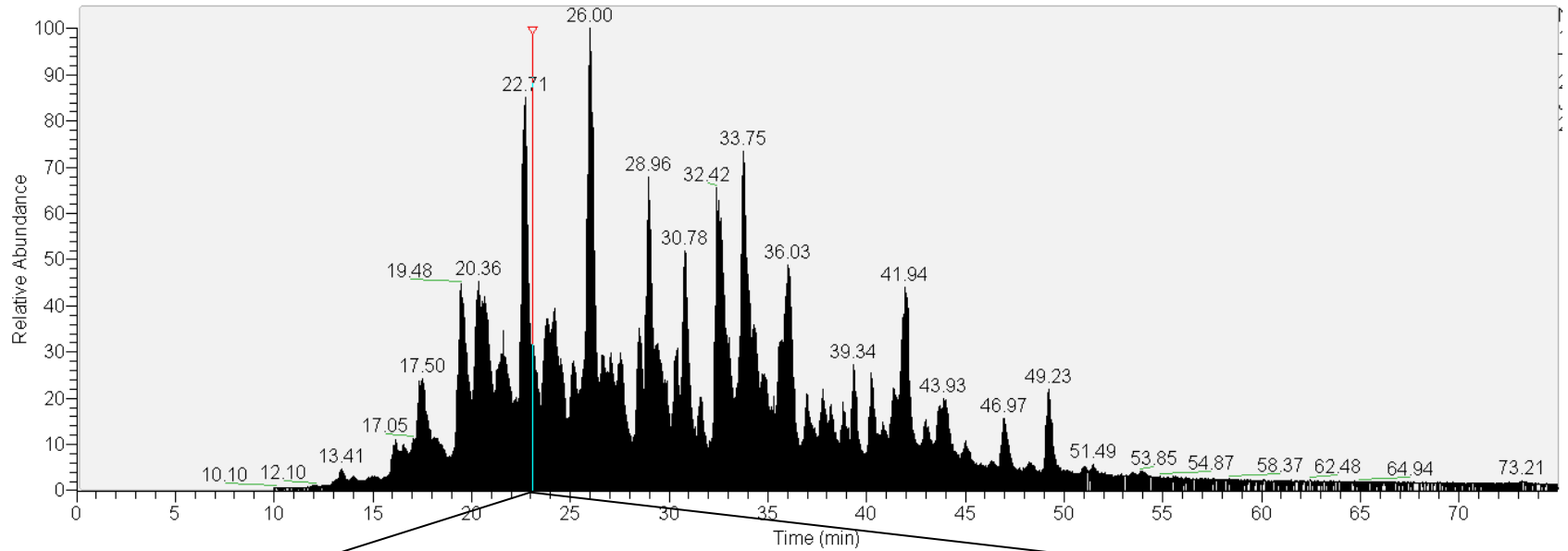
# Principe des acquisitions LC-MS/MS automatiques (2)

## Fragmentation des espèces multichargées sélectionnées

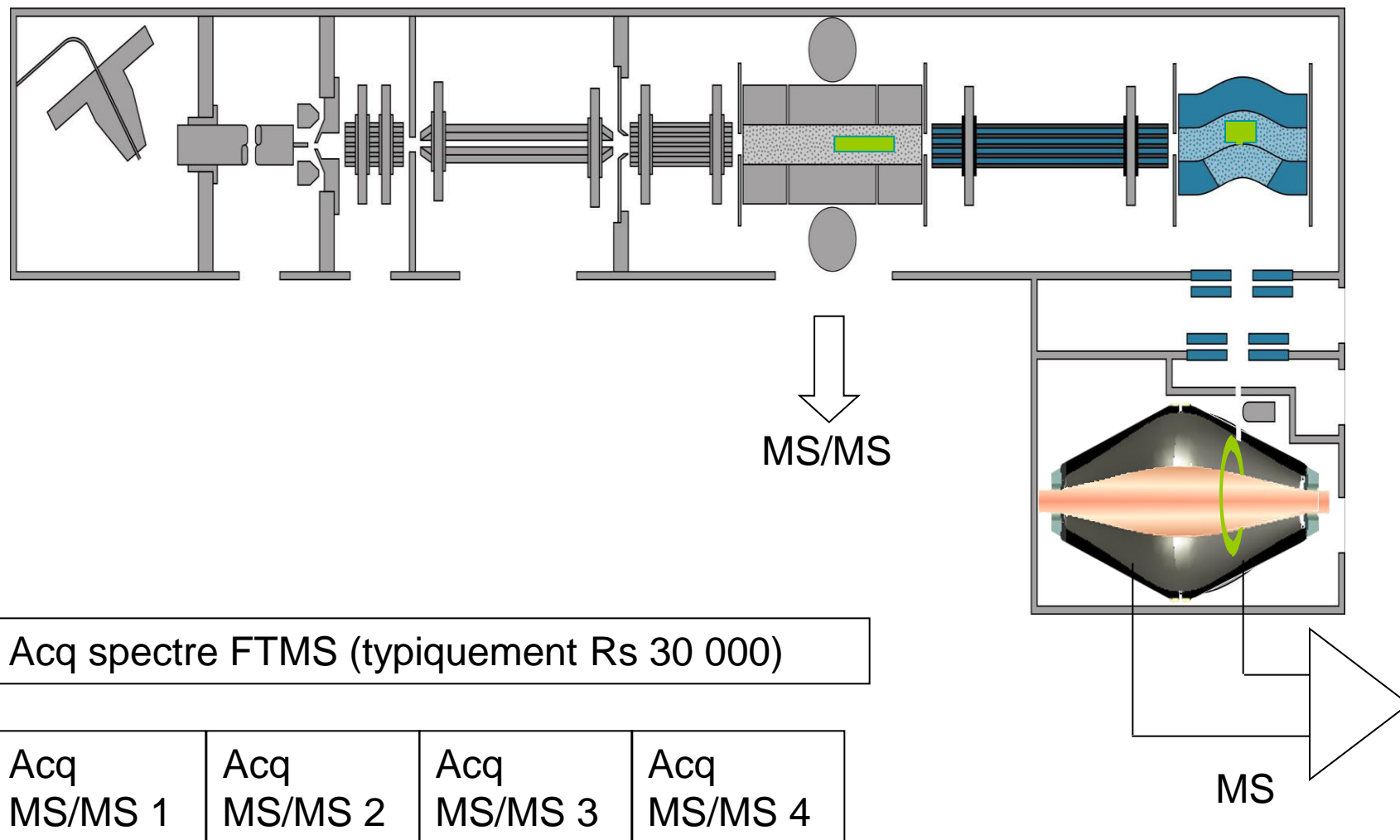




# Principe des acquisitions LC-MS/MS automatiques (3)



# Acquisitions MS et MS/MS en parallèle



Suivant les générations d'appareillage : jusqu'à 20 MS/MS pendant un spectre FTMS

# Recherches dans les bases de données protéiques

## MASCOT MS/MS Ions Search

The screenshot shows the MASCOT MS/MS Ions Search interface with several parameters set. Annotations with arrows point to specific fields, explaining their function and options.

Your name	<input type="text"/>	<b>Base de données à explorer et taxonomie</b> <ul style="list-style-type: none"><li>- organisme d'intérêt</li><li>- tous les organismes (pour détecter des contaminants)</li></ul>
Search title	<input type="text"/>	
Database	TAIR10	
Taxonomy	All entries	
Enzyme	Trypsin	<b>Enzyme utilisée et sites de clivage non couvés</b> : trypsine 2
Fixed modifications	Methyl (C-term) Methyl (DE) Methyl (K) Methylthio (C) NIPCAM (C)	<b>Modifications :</b> <ul style="list-style-type: none"><li>- totales</li><li>- partielles</li></ul>
Quantitation	None	
Peptide tol. ±	5 ppm # <sup>13</sup> C 0	<b>Tolérance sur la mesure des masses :</b> sur les précurseurs et sur les fragments
Peptide charge	2+	
Data file	<input type="text"/> Browse...	
Data format	Mascot generic	
Instrument	ESI-TRAP	<b>Instrument MS :</b> fragments à prendre en compte : <ul style="list-style-type: none"><li>- C-term y, z</li><li>- N-term b, a</li><li>- interne, immoniums, chargés 2+, etc</li></ul>
Decoy	<input checked="" type="checkbox"/>	
<input type="button" value="Start Search ..."/>		

# Recherches dans les bases de données protéiques

## {*MATRIX*} Mascot Search Results

User  
Email  
Search title  
MS data file  
Database  
Taxonomy  
Timestamp  
Significant hits

4. [P24821-03-00-00](#) Mass: 230733 Score: **491** Peptides matched: 8  
(TENA\_HUMAN) Splice isoform 4; Variant Displayed; Conflict Displayed; from P24821 Tenascin precursor

Query	Observ
<input type="checkbox"/> <a href="#">65</a>	<b>701.</b>
<input checked="" type="checkbox"/> <a href="#">109</a>	<b>798.</b>
<input checked="" type="checkbox"/> <a href="#">122</a>	<b>833.</b>
<input checked="" type="checkbox"/> <a href="#">173</a>	<b>675.</b>
<input checked="" type="checkbox"/> <a href="#">184</a>	<b>762.</b>
<input checked="" type="checkbox"/> <a href="#">209</a>	<b>1272.</b>
<input checked="" type="checkbox"/> <a href="#">223</a>	<b>974.</b>
<input checked="" type="checkbox"/> <a href="#">229</a>	<b>1013.</b>

### Peptide View

MS/MS Fragmentation of **LEELENLVSSLR**

Found in **P24821-03-00-00**, (TENA\_HUMAN) Splice isoform 4; Variant Displayed; Conflict Displayed; from P24821 Tenascin precursor (JI)

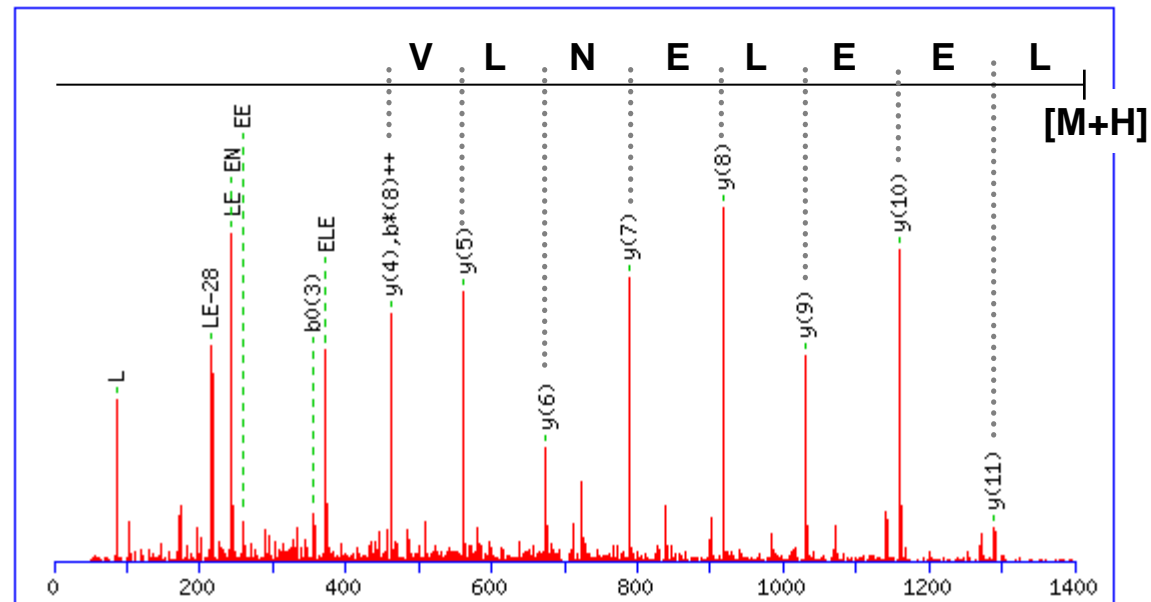
Match to Query 65: 1400.792048 from(701.403300,2+)

From data file E:\Dossier Matrexc 0503 et suite\Analyses LCMS depuis 072003\LCMS\_150304\040315\_0258\_i\_dmp.pkl

Click mouse within plot area to zoom in by factor of two about that point

Or, Plot from  to  Da

[P02545-00-00-00](#)  
[P07996](#)  
[P02543](#)  
[P07355-00-00-00](#)  
[P00761-00-00-00](#)  
[P35555-00-59-00](#)  
[P35441](#)  
[P12110-00-00-00](#)  
[Q9Y6C2](#)  
[Q05793](#)  
[Q28178](#)  
[P16105-00-00-00](#)  
[P12763](#)

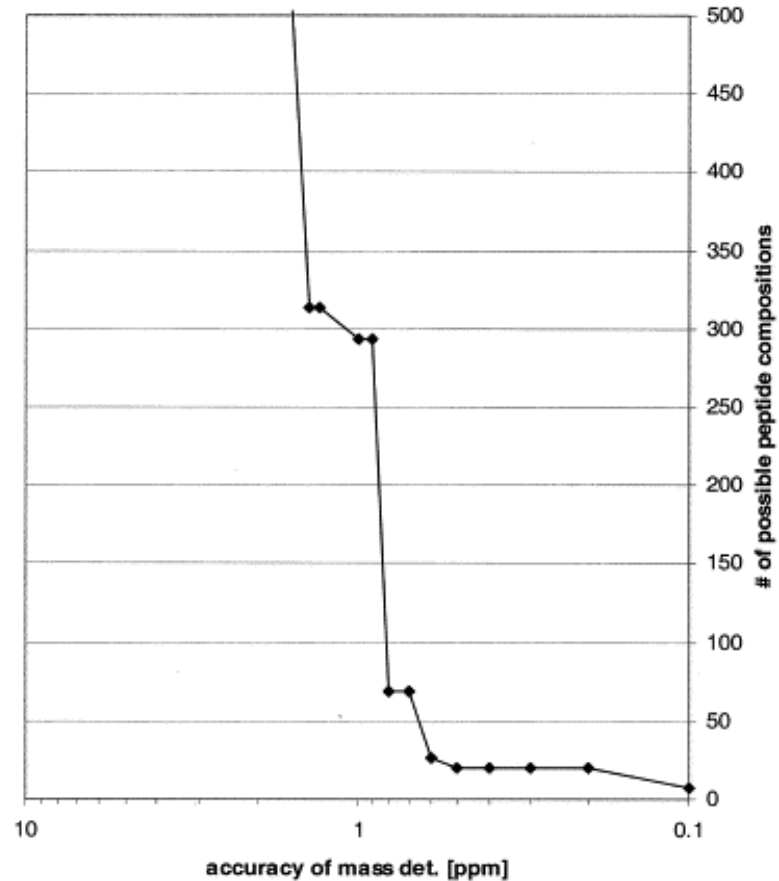


# Intérêt de la haute précision de mesure de masse en MS

Masse déterminée à < 5 ppm sur les précurseurs en routine

< 3 ppm avec l'option « lock mass » sur LTQ-Orbitrap (Olsen JV *et al.* Mol. Cell. Proteomics. 2005;4(12):2010-21)

Encore mieux sur appareils FTICR



Number of possible amino acid compositions of peptides for a given example mass of 1005.4433 u as depending on the accuracy of mass determination.

# Intérêt de la haute précision de mesure de masse en MS/MS ?

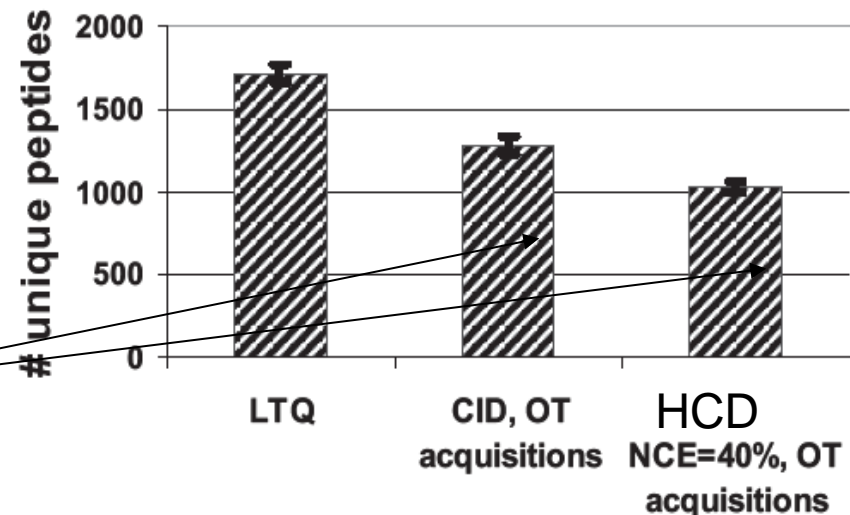
Etude utilisant un LTQ-Orbitrap et un LTQ-FT

Analyzer	Activation method	Peptide amount loaded (mg)	Mean number of tandem mass spectra	SD	Mean ion injection time	SD
OT	CID	0.5	2579	87	378	58
OT	HCD 40%*	0.5	2062	104	530	58
LIT	CID	0.5	4202	12	41	2

Comparaison du nombre de spectres MS<sup>2</sup> acquis et du temps de remplissage des analyseurs.

Comparaison du nombre de séquences identifiées

Dans ces deux cas, pas possible d'acquérir spectres MS et MS/MS en parallèle !



(a)

CID method / Scan function

# Intérêt de la haute précision de mesure de masse en MS/MS ?

Comparaisons similaires aux précédentes faites entre une trappe d'ions et une géométrie Q-TOF

	<i>Ion Trap</i>	<i>Q-TOF</i>
Total number spectra acquired	355,136	363,692
MS/MS acquired	271,507	186,049
Number of extracted files (.pkl)	92,081	129,752
Number of search result files (.spo)	72,028	20,871
Number of validated spectra	10,618	16,072
Number of validated unique peptides	4151	6253
Number of protein IDs	873	994
Number of unvalidated spectra	61,410	4799

Table 1. Comparison between 6330 Ion Trap and 6510 Q-TOF search results and hits.

# Intérêt de la haute précision de mesure de masse en MS/MS ?

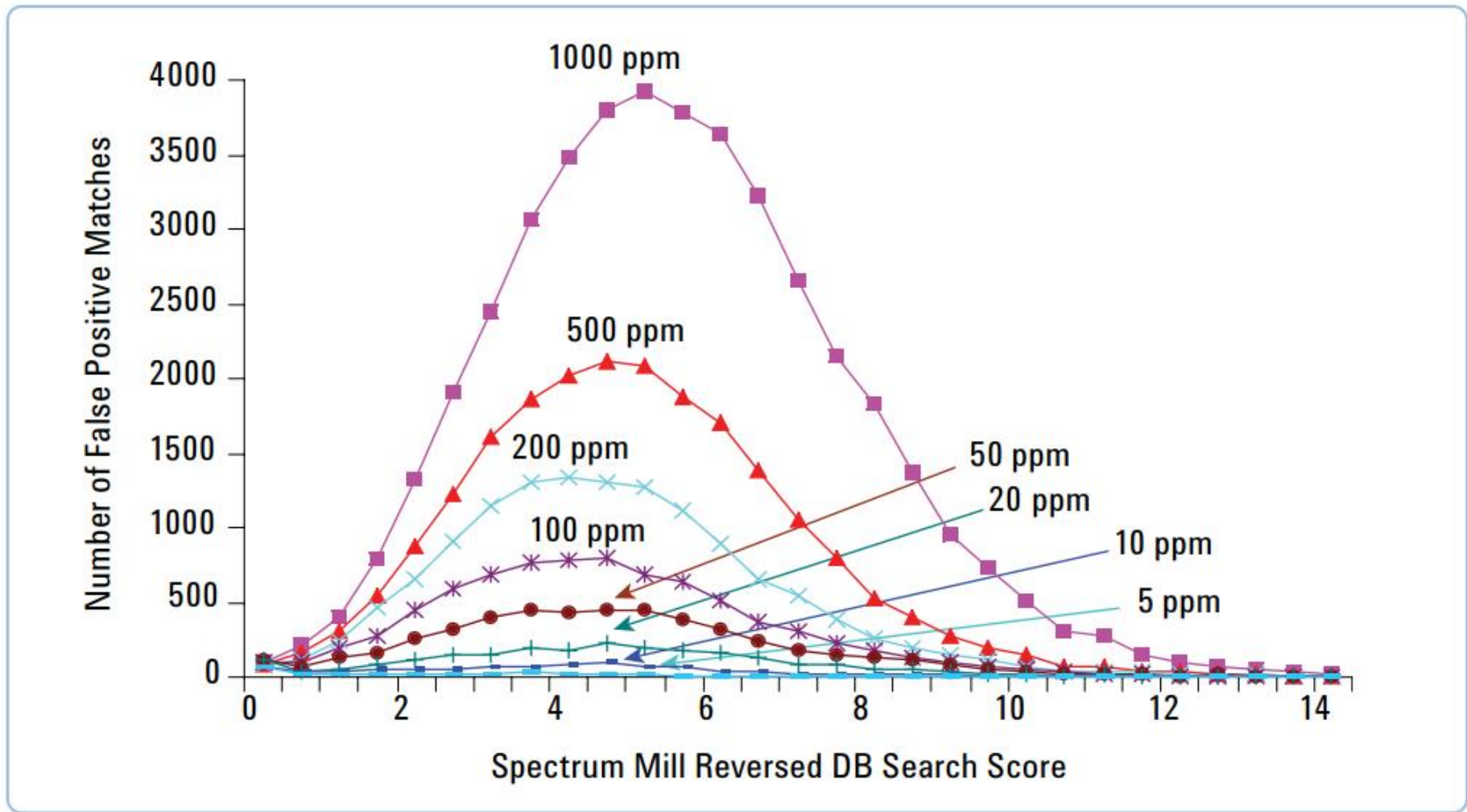


Figure 4. Spectrum Mill score distributions of reversed database hits for variable fragment tolerance at constant 10 ppm precursor tolerance using the 6510 Q-TOF.



# Intérêt de la haute précision de mesure de masse en MS/MS ?

<i>ppm Tolerance</i>		<i>Number of Valid Peptide Matches with % Confidence</i>			<i>Number of Valid Protein Matches with % Confidence (minimum of 2 or more unique peptides/protein)</i>		
<b>Precursor</b>	<b>Fragment</b>	<b>&gt;95%</b>	<b>&gt;75%</b>	<b>&gt;50%</b>	<b>&gt;95%</b>	<b>&gt;75%</b>	<b>&gt;50%</b>
1000	1000	954	2220	3361	102	213	282
5	1000	4181	6590	7645	338	478	524
10	40	6590	11552	15509	478	728	829

**Table 2. The number of confident hits for three different combinations of precursor and fragment mass accuracy settings on the 6510 Q-TOF.**

# Intérêt de la haute précision de mesure de masse en MS/MS ?

Pour les analyses de type  
« Data-Independent Acquisitions »

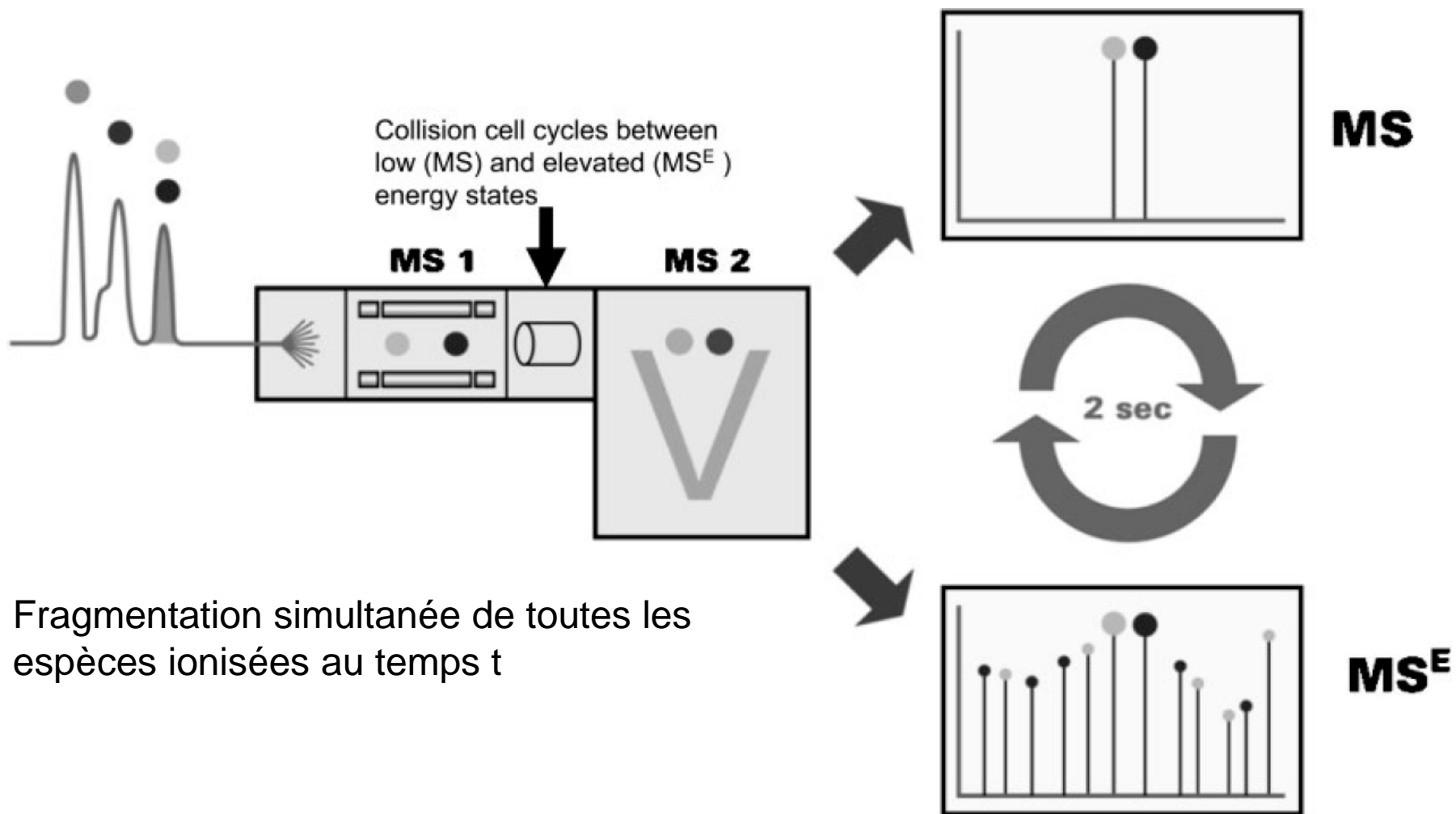
**Voir la revue récente :**

Law KP1, Lim YP.

Expert Rev Proteomics. 2013;10(6):551-66.

# Fragmentation large bande MS-MS<sup>E</sup>

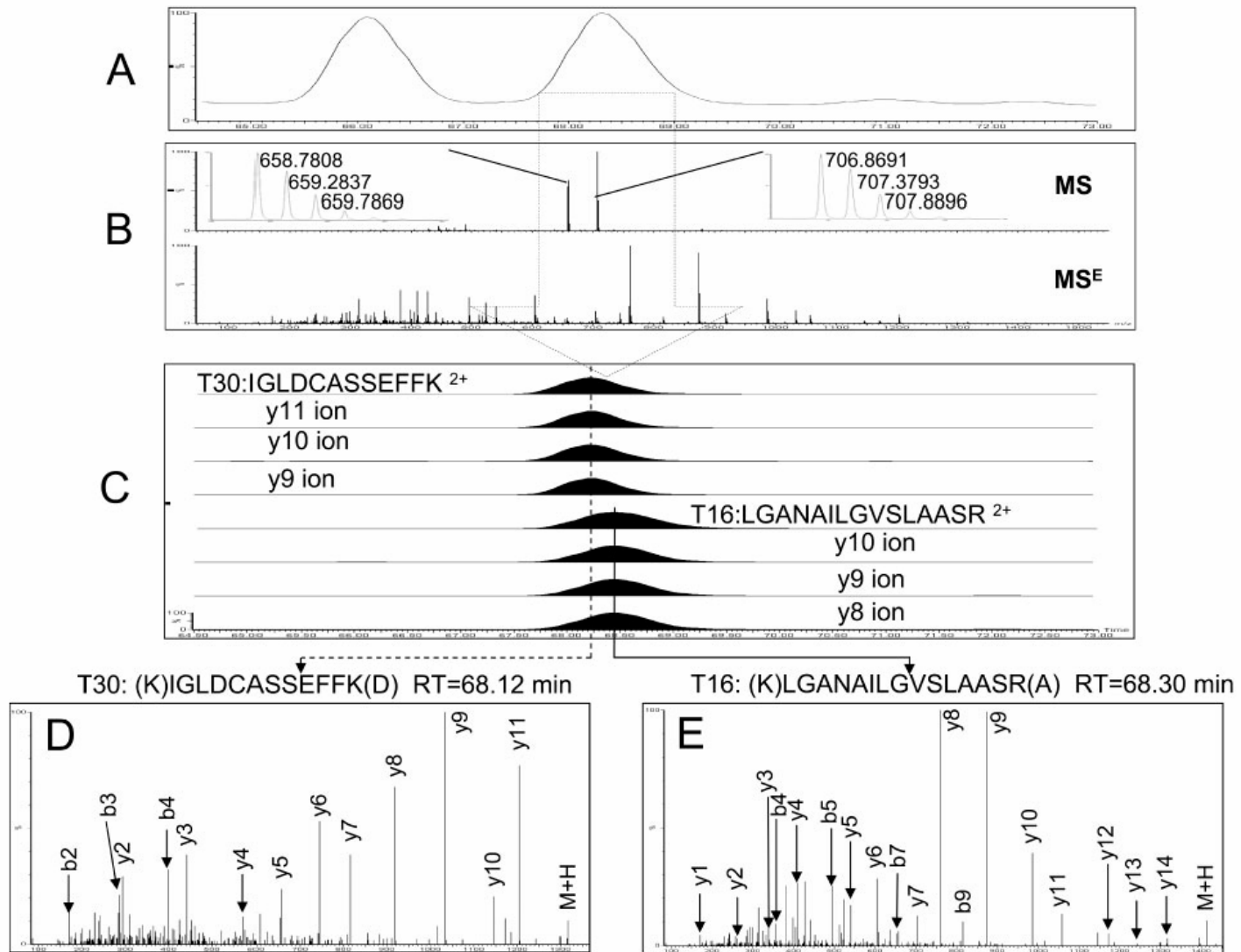
Idée originellement développée sur des appareils de géométrie Q-TOF



MCP 2006, 5, 589-607, Silva *et al.*

RCM 2007, 21, 730-744, Chakraborty *et al.*

# Fragmentation large bande MS-MS<sup>E</sup>



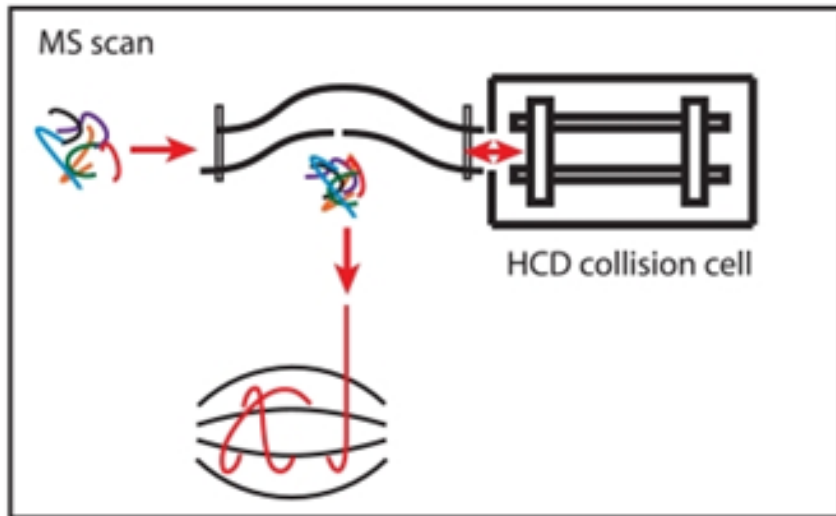
Chaque précurseur est fragmenté de multiples fois, à travers tout son pic d'élution chromatographique

# Méthode AIF: All Ion Fragmentation

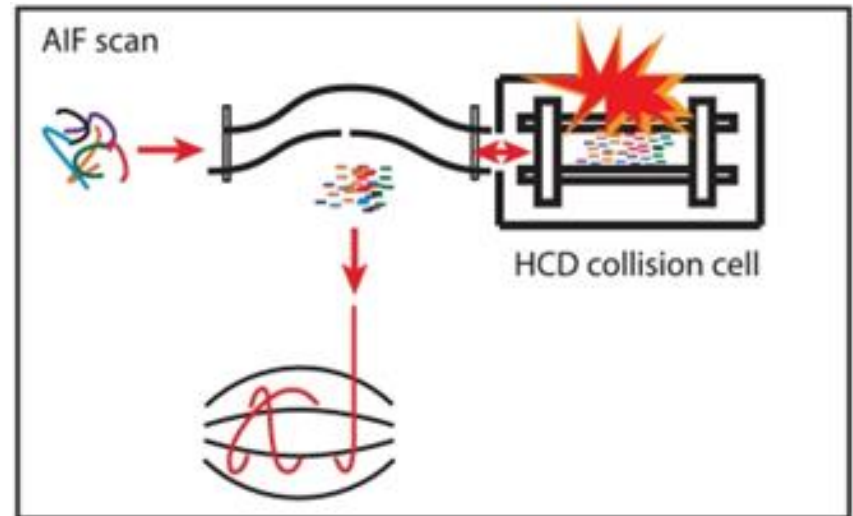
Méthode mise au point pour les appareils de type Exactive (seul analyseur = un Orbitrap)

Rs 100 000 à  $m/z$  200

Recherches dans les bases de données avec 7 ppm de tolérance sur les précurseurs et 15 ppm sur les fragments



Spectre MS

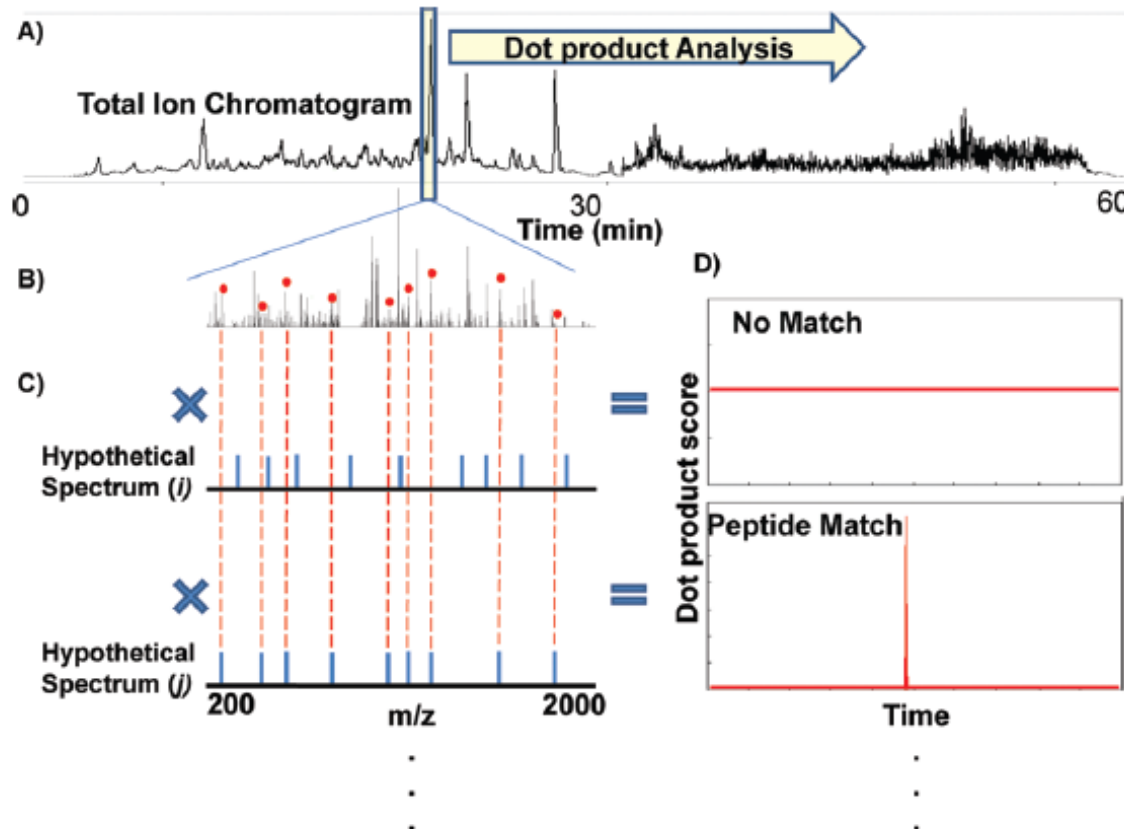


Spectre AIF

Ici aussi, l'association des ions fragments provenant du même précurseur est obtenue en considérant les pics d'élution chromatographique (programme MaxQuant)

# Méthode FT-ARM : Fourier transform-all reaction monitoring

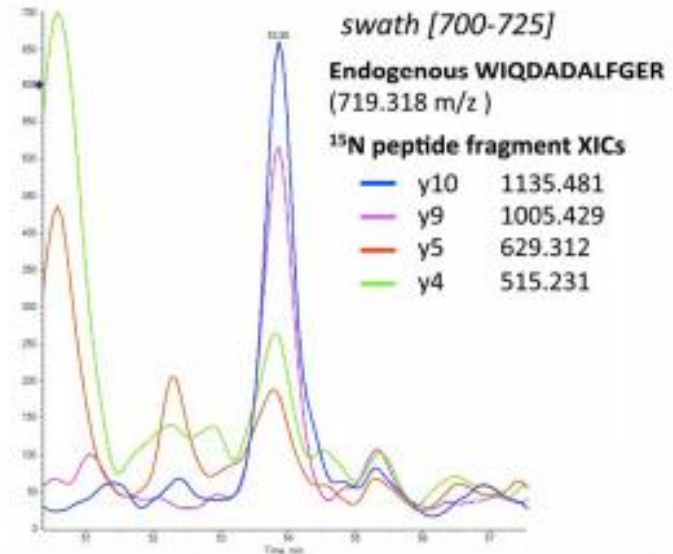
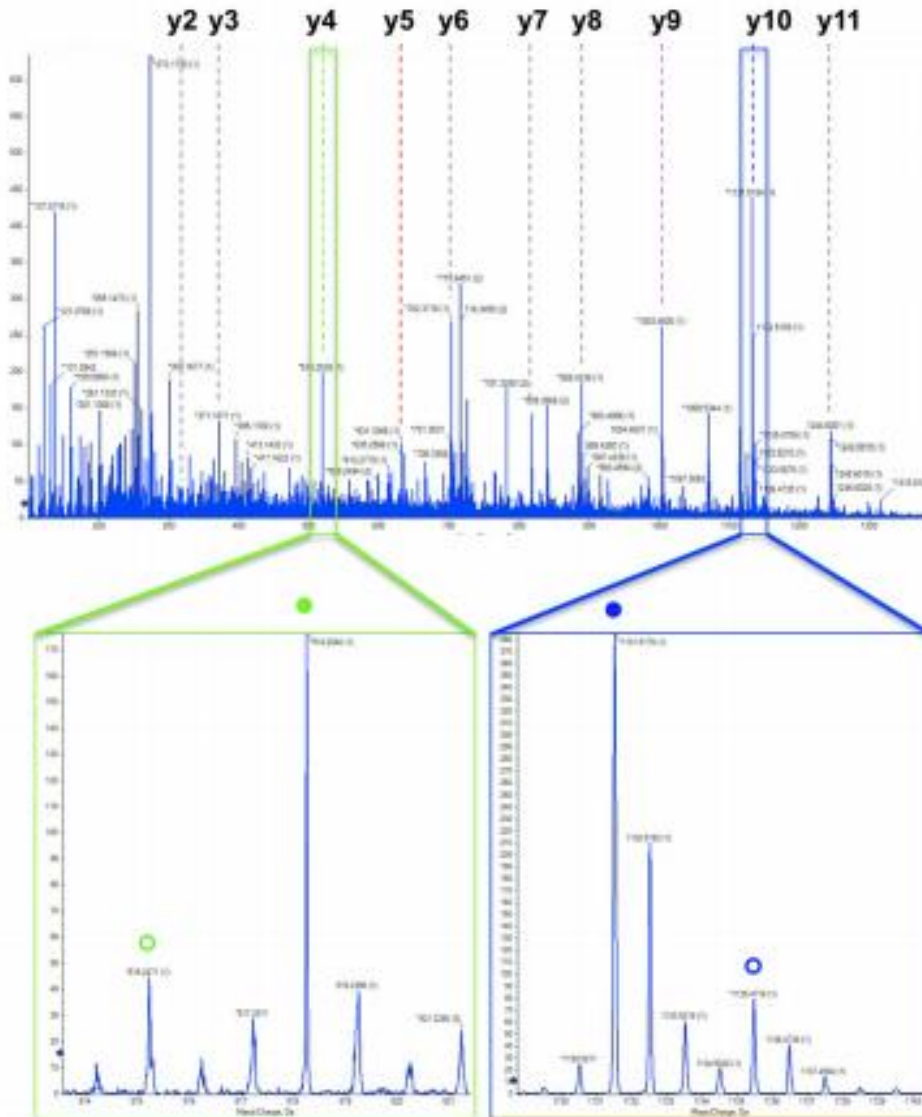
Méthode développée sur des appareils LTQ-FT et LTQ-Orbitrap



- Fenêtre de sélection de précurseurs : 12 ou 100 Da.
- Ici l'information des pics d'éluion chromatographique n'est pas utilisée ; le programme développé utilise la précision de mesure de masse sur les fragments.
- Application : identification et quantification de BSA dans un digestat de levure

# Fragmentation « SWATH »

Méthode développée sur un appareil QTOF (TripleTOF 5600)



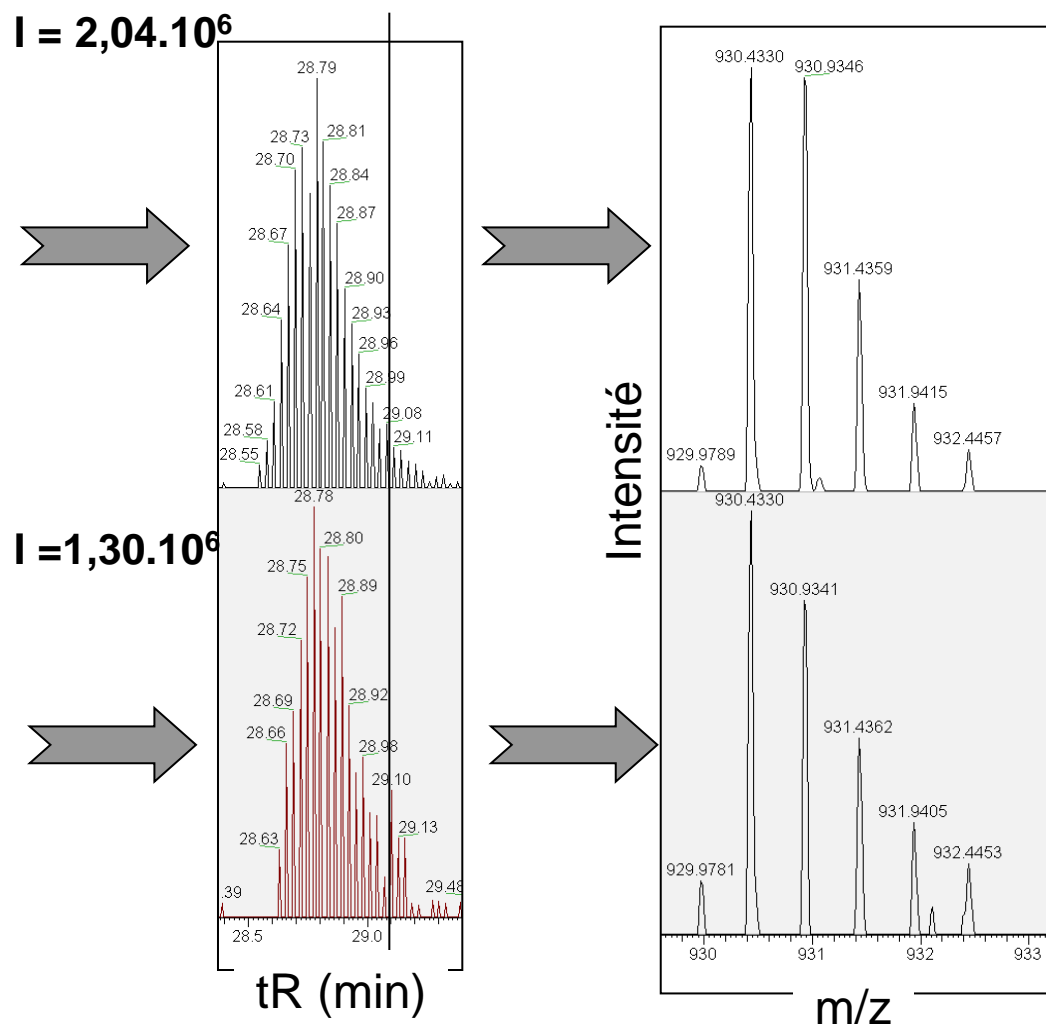
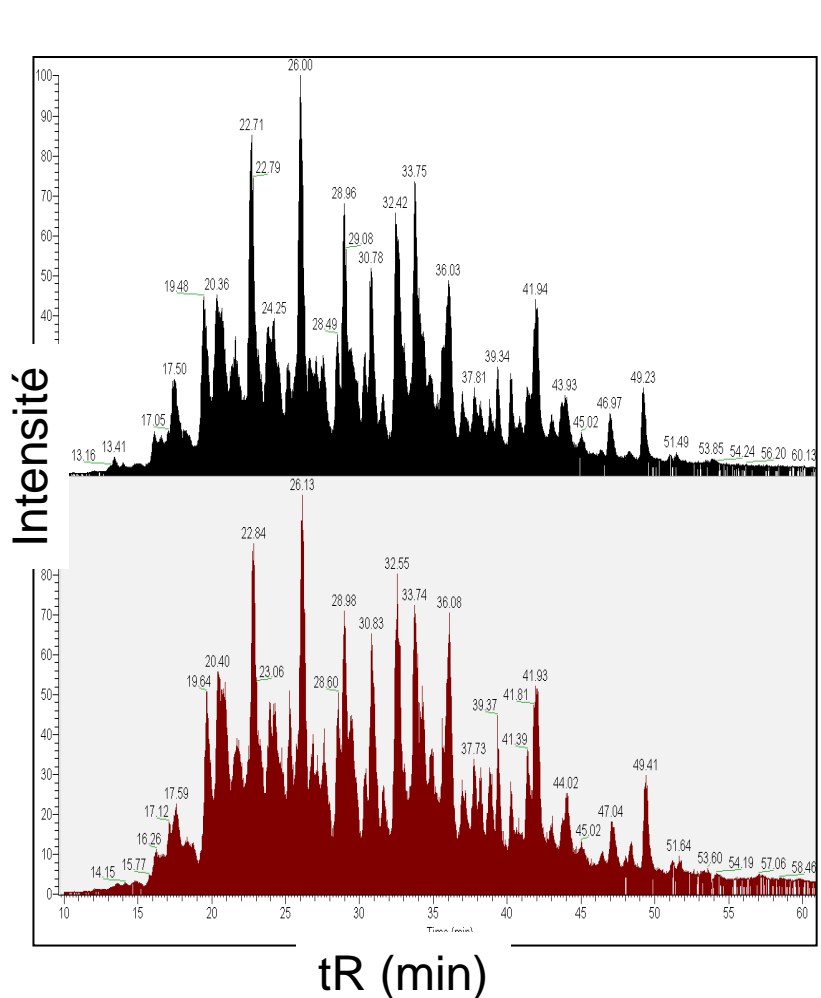
- Fragmentation systématique de plages de 25 Da.
- Recherche dans les spectres MS/MS composites d'ions fragments attendus pour des séquences d'intérêt => possible sur une gamme dynamique de 4 log

# **Quantification relative d'échantillons**

## **1) en analyses exploratoires**



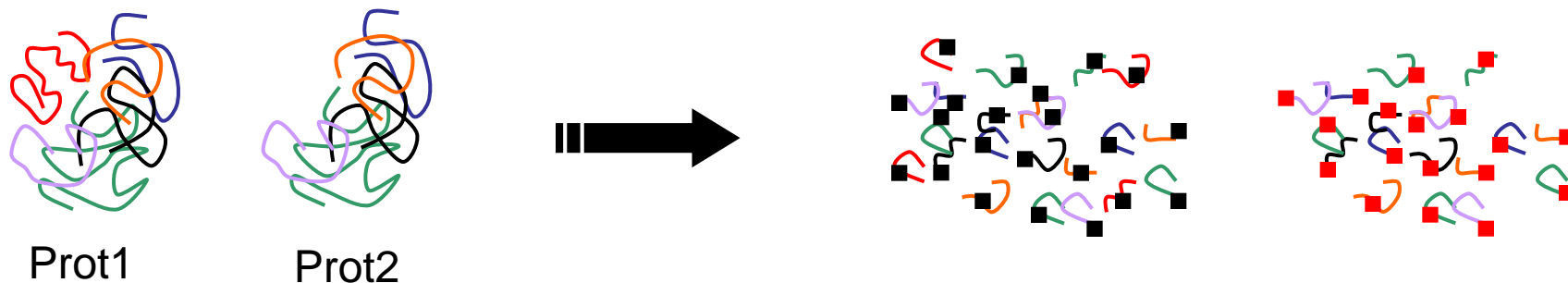
# Analyse différentielle sans marquage des échantillons



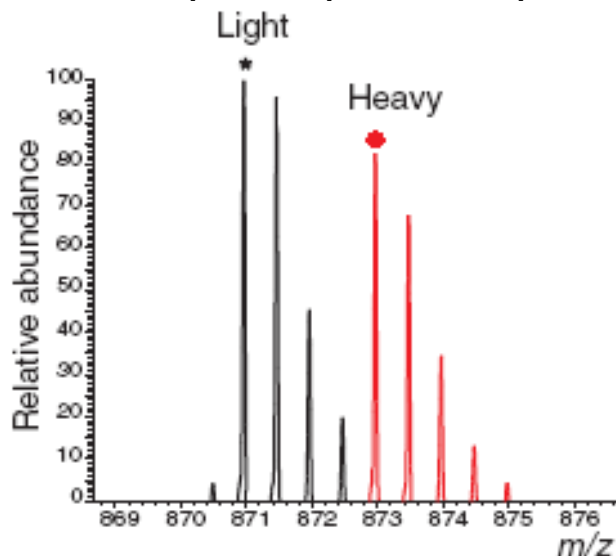
Extraction et alignement des signaux relatifs à chaque ion détecté en MS...

...basée sur la haute précision de mesure de masse des rapports  $m/z$  et sur leurs temps de rétention

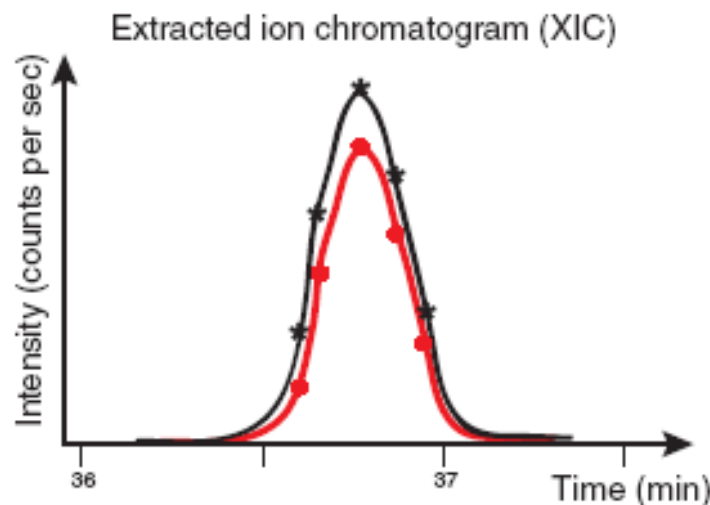
# Analyses semi-quantitatives avec marquage différentiel des échantillons



Les deux formes peptidiques co-éluent et possèdent la même efficacité d'ionisation. Le rapport de leurs intensités correspond au rapport de leurs abondances dans les deux échantillons protéiques comparés.

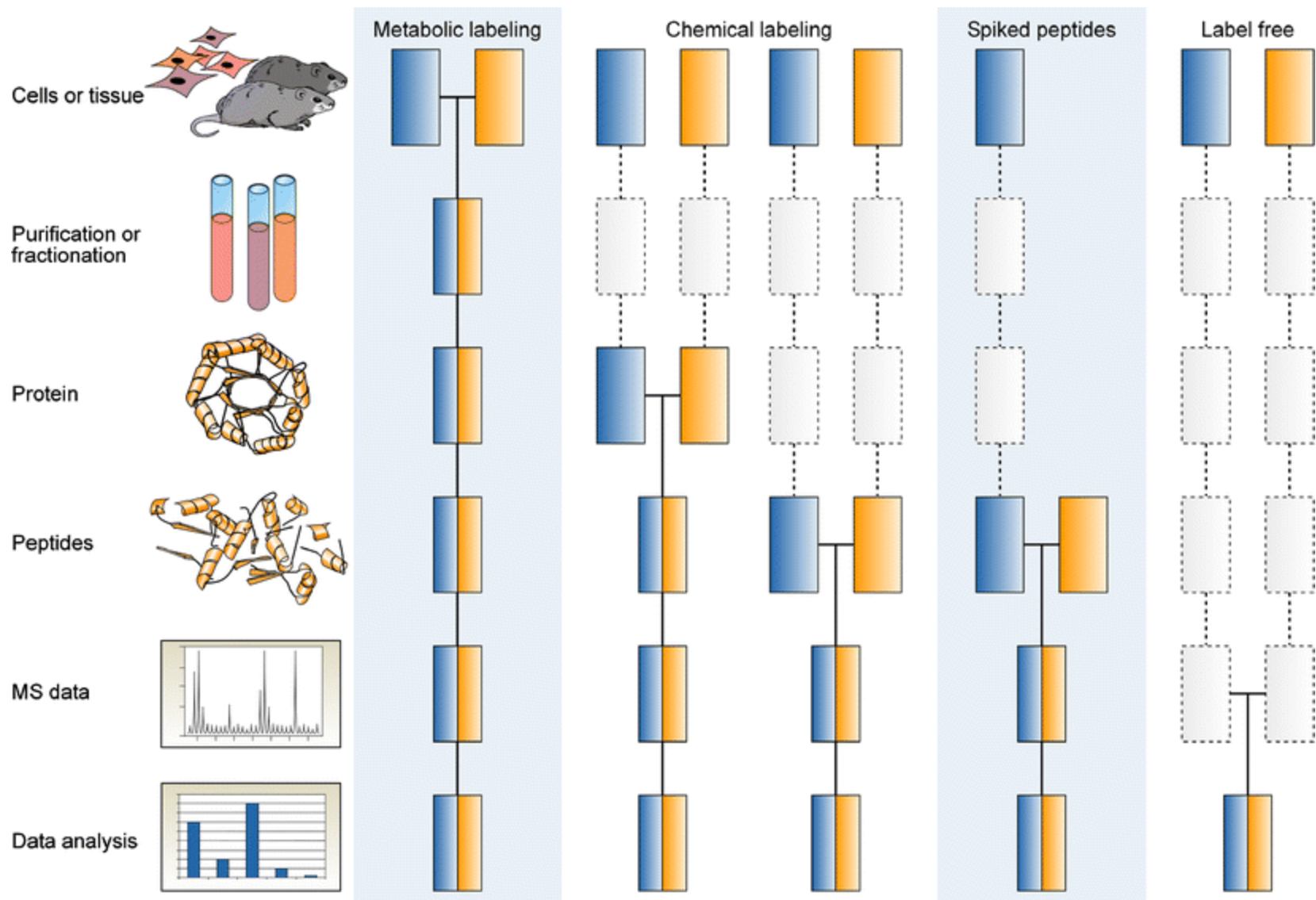


Détection en MS de la paire de signaux au temps t



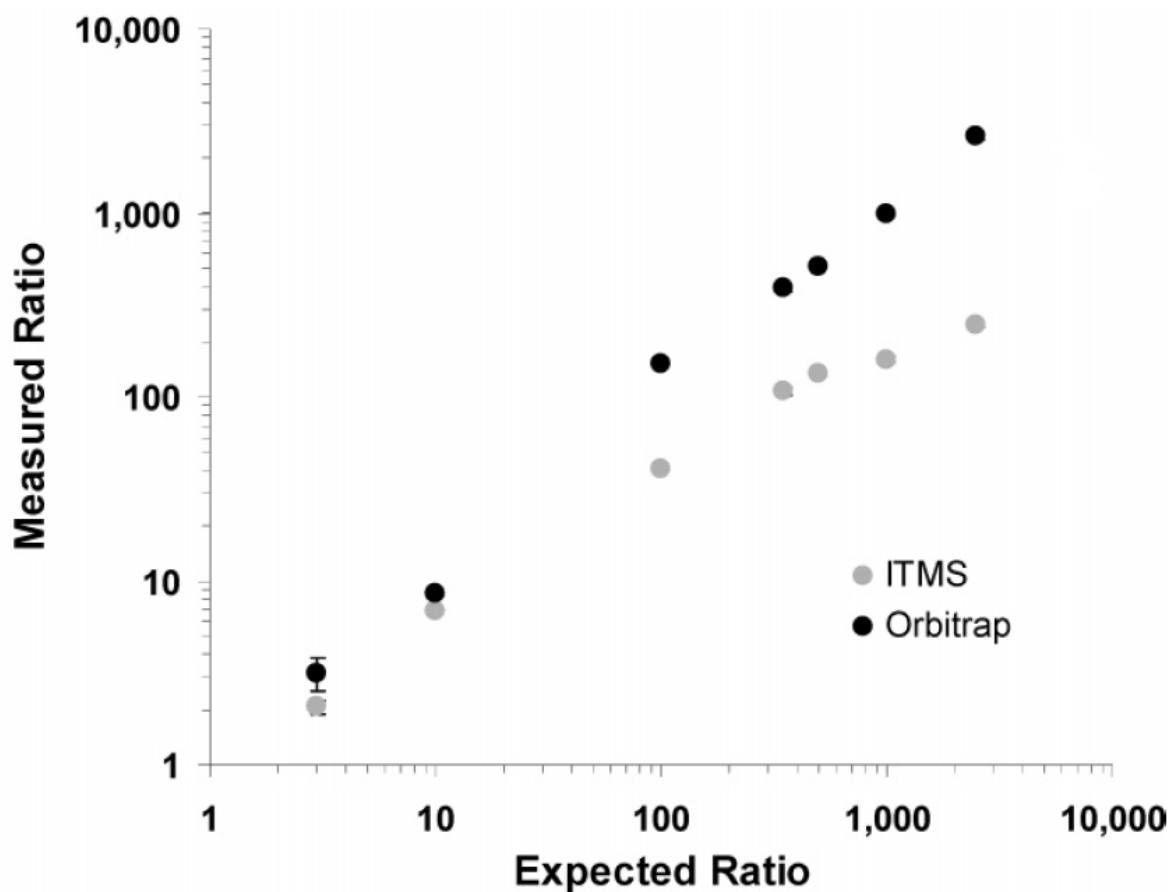
Reconstitution des pics d'éluion pour les deux formes peptidiques

# Stratégies d'analyse semi-quantitative



# Gamme dynamique des mesures quantitatives

## Gamme dynamique au sein d'un même spectre MS (marquage)

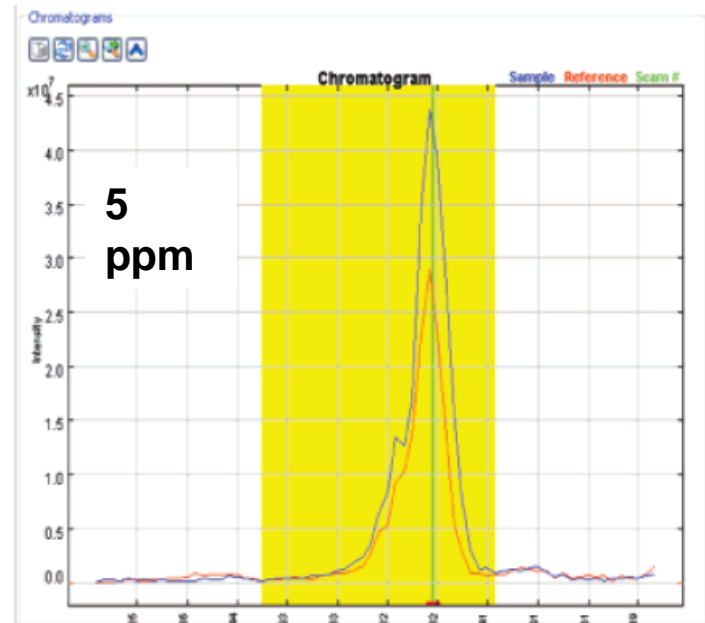
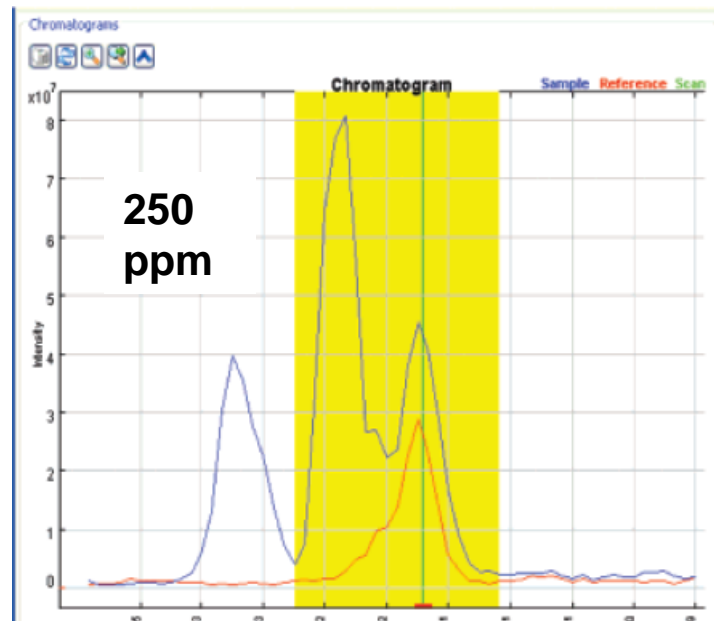
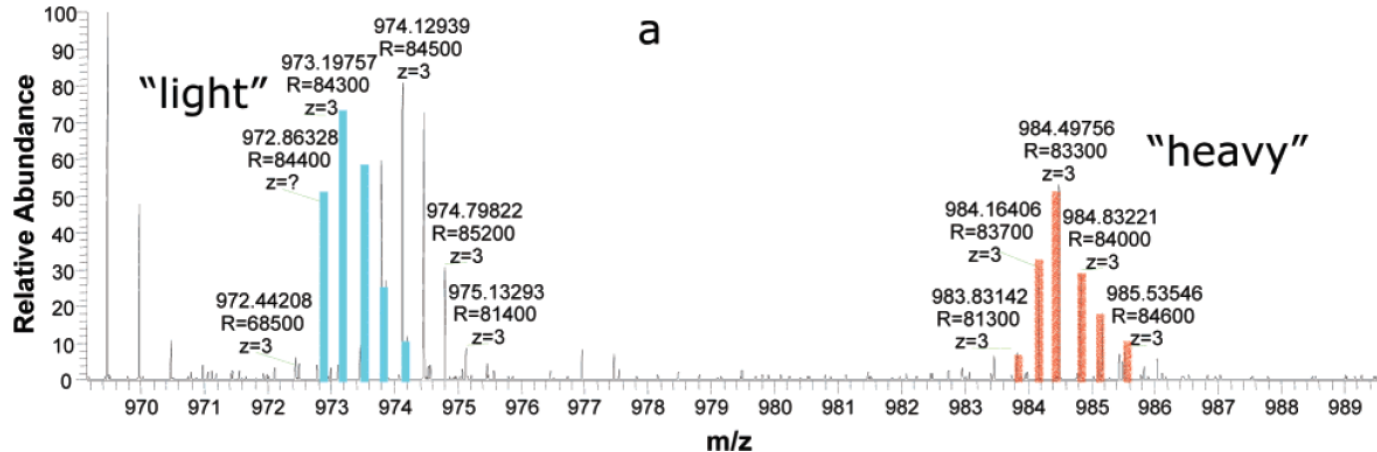


A partir de 10:1, rapports sous-estimés par analyse sur le LTQ.

Gamme dynamique avec le LTQ-Orbitrap : au moins 2500.

# Précision et justesse des mesures quantitatives

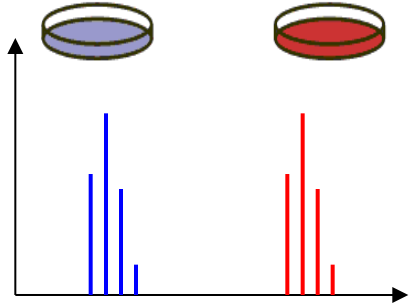
## Effet de la haute précision de mesure de masse



# Précision et justesse des mesures quantitatives

## Corrélation avec le rapport signal sur bruit (S/B)

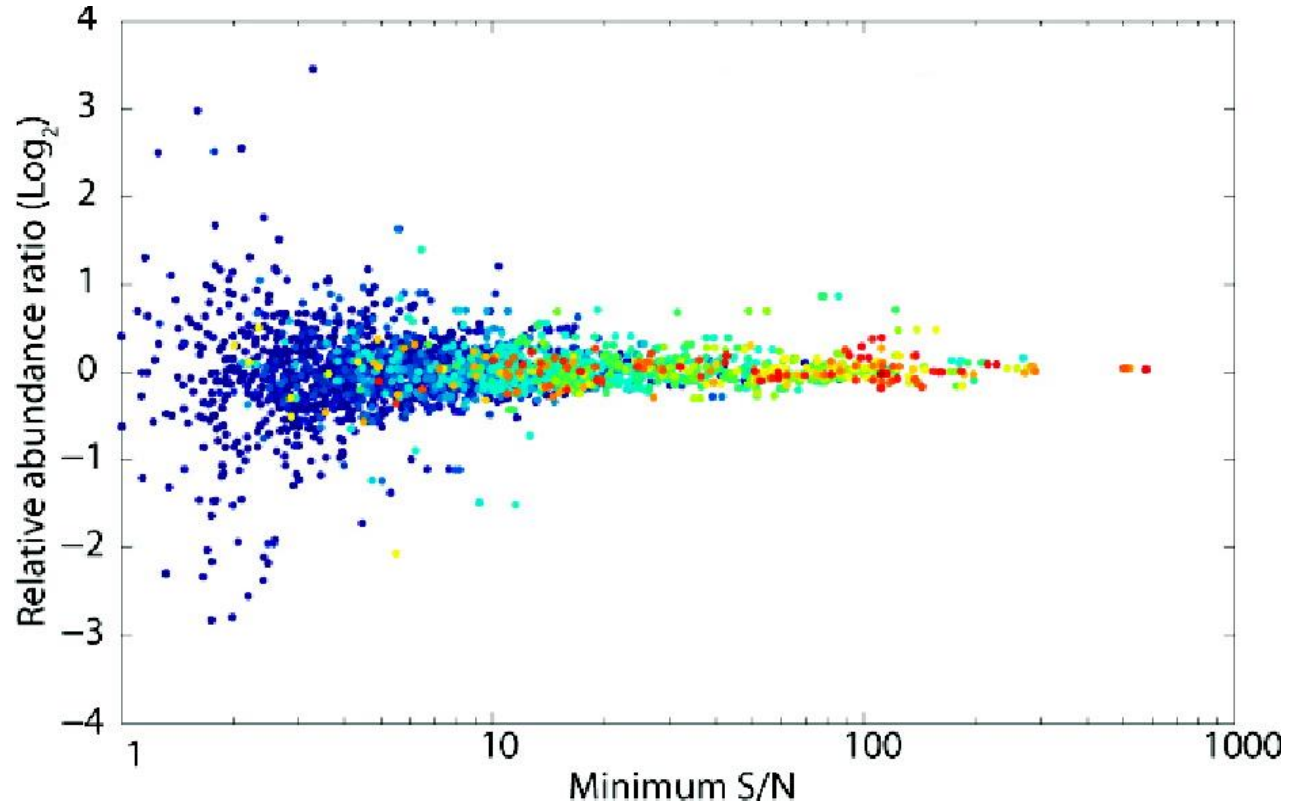
### Codage des Arg et Lys



Mélange 1:1

~1/3 des ions à S/B < 5

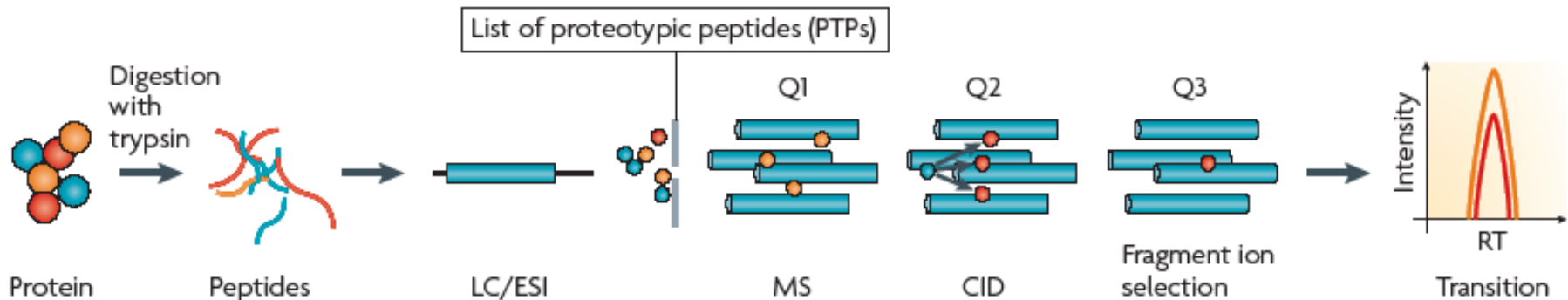
(appareils LTQ-FT ou LTQ-Orbitrap)



# **Quantification relative d'échantillons**

## **2) en analyses ciblées**

# Analyses protéomiques ciblées : analyses par SRM/MRM



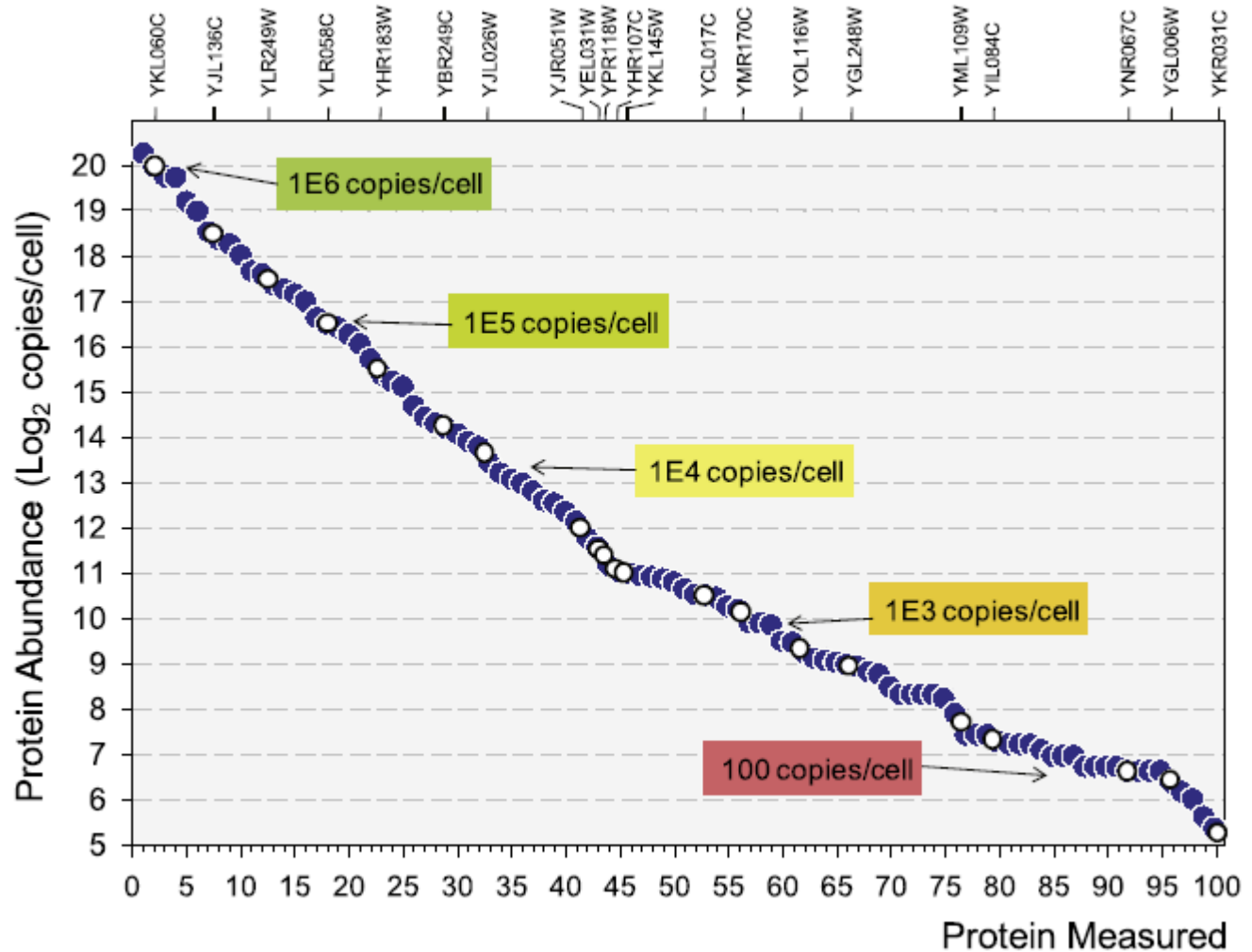
**Pour chaque protéine à quantifier : sélection de peptides protéotypiques, *i.e.* séquences spécifiques de cette protéine et bien identifiables par MS.**

**Détection de paires précurseur-fragments (transitions) spécifiques de chaque peptide protéotypique.**



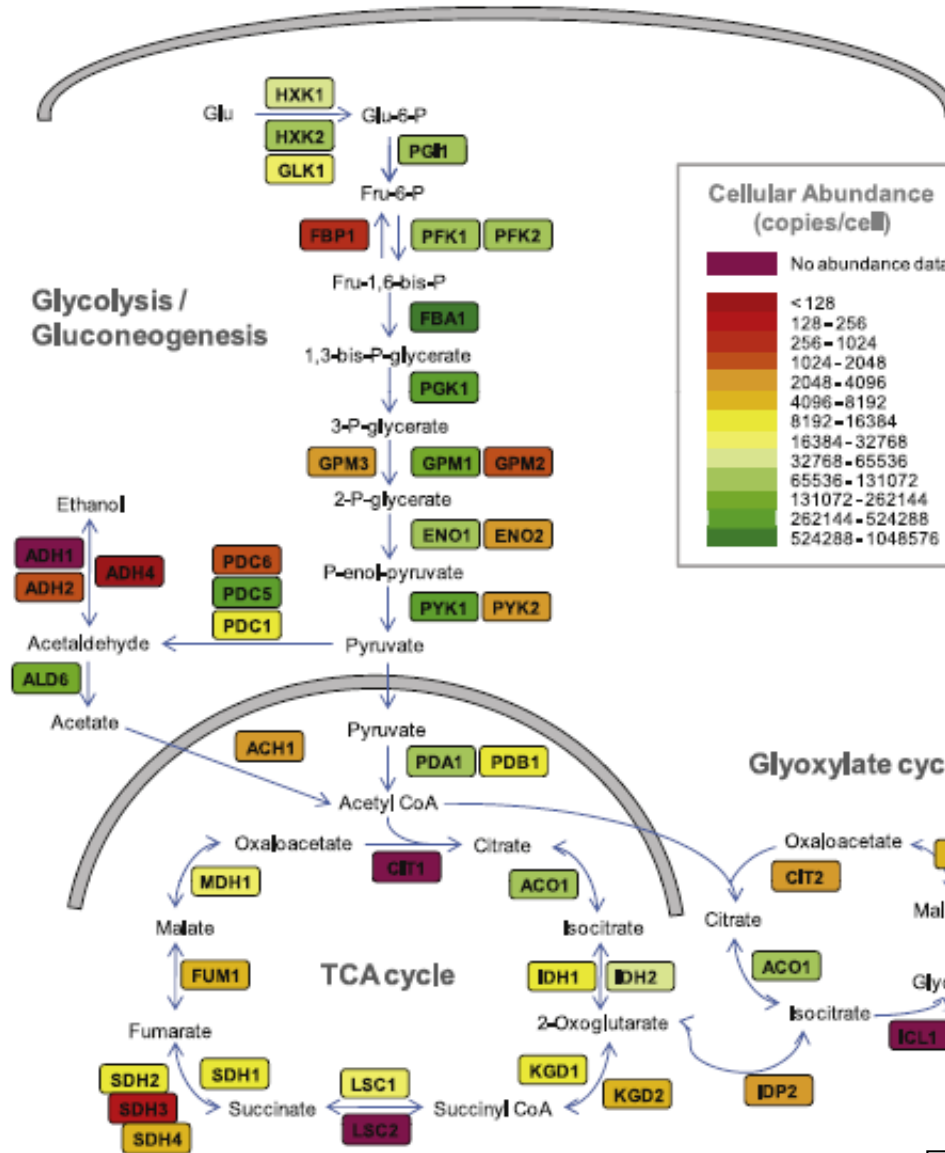
# Analyses protéomiques ciblées : quelle couverture du protéome ?

Abondances  
déterminées  
par WB  
quantitatif



Quantification absolue de certaines protéines (SRM de peptides marqués) :  
confirmation de la possibilité de quantifier des protéines entre  $\sim 40$  et  $\sim 10^6$   
copies par cellule.

# Analyses protéomiques ciblées : protéines impliquées dans le métabolisme du carbone



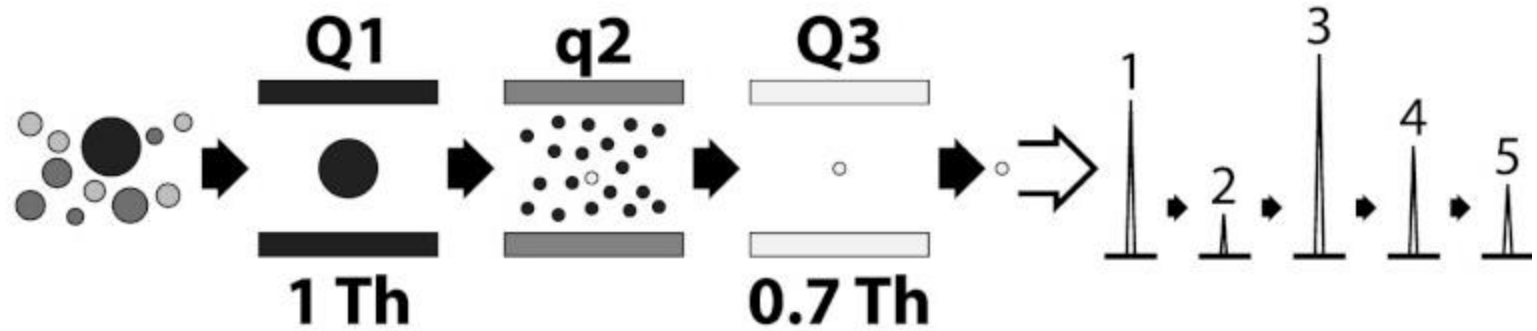
Gradient LC de 30 min

Quantification par :

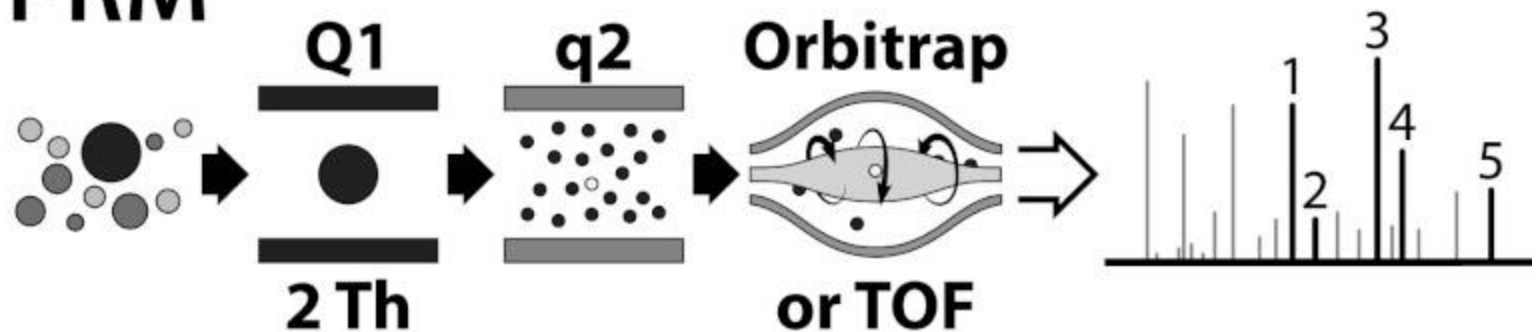
- 2 pept/protéine
- trois transitions SRM
- 8 points de mesure

# Comparaison SRM (appareils QqQ) et PRM (appareils de type QqOrbitrap)

## A SRM

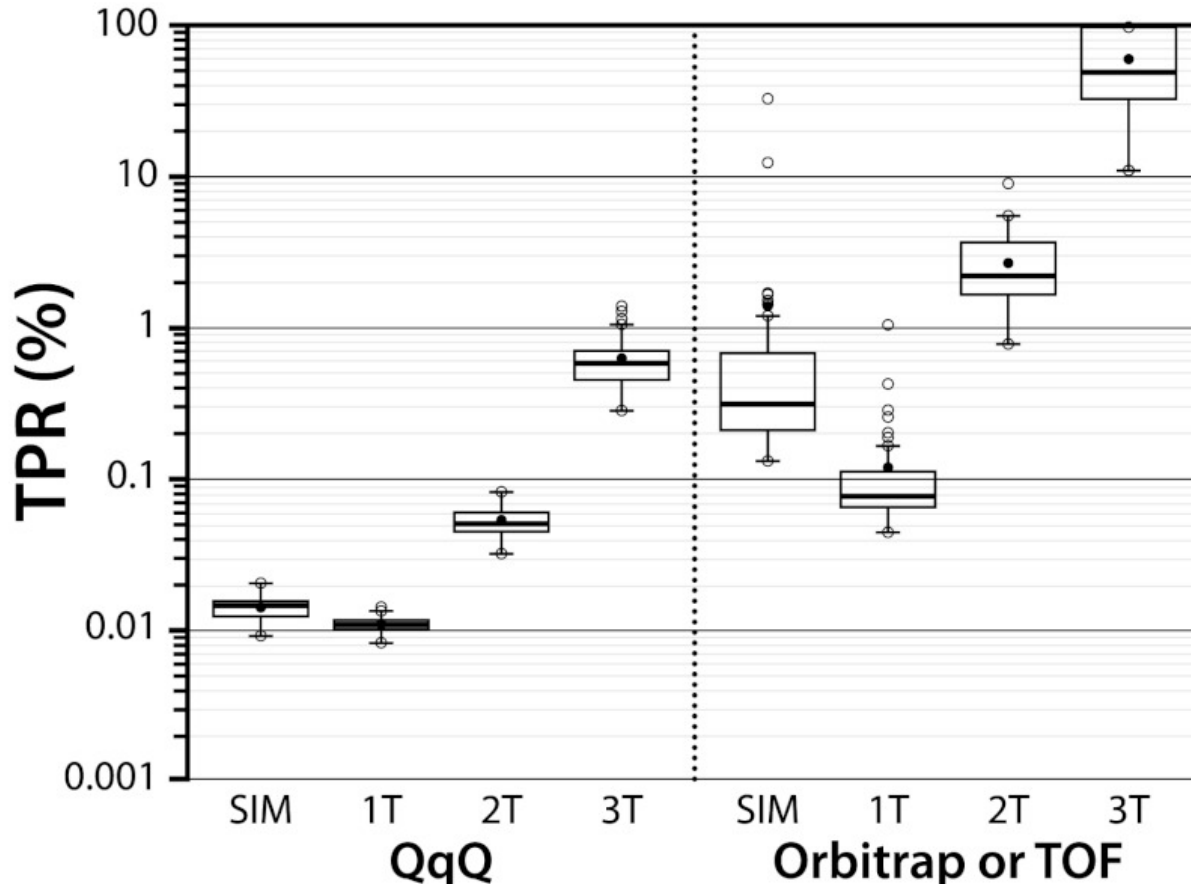


## B PRM



# Comparaison SRM (appareils QqQ) et PRM (appareils de type QqOrbitrap)

**C** Count: 7485 9459 1997 182 334 1249 49 3  
 TPR (%): 0.014 0.011 0.053 0.63 1.4 0.12 2.7 61



Pour 25 peptides d'intérêt :  
 calcul théorique des chances de les identifier correctement lorsqu'ils sont en présence d'un produit de digestion des protéines humaines.

Count: nombre de peptides théoriques vis-à-vis desquels les peptides d'intérêt ne peuvent être discriminés.

Attention : ici, pas d'information de temps de rétention prise en compte

# Comparaison SRM (appareils QqQ) et PRM (appareils de type QqOrbitrap)

**Dans le cas d'analyses de peptides en matrice complexe (extrait protéique de levure digéré) :**

- Meilleure précision de mesure quantitative en SRM, sans doute en raison de la plus haute vitesse de scan du QqQ (environ un facteur 2 par rapport au QqOrbitrap)
- Plus grande gamme dynamique atteinte en PRM qu'en SRM

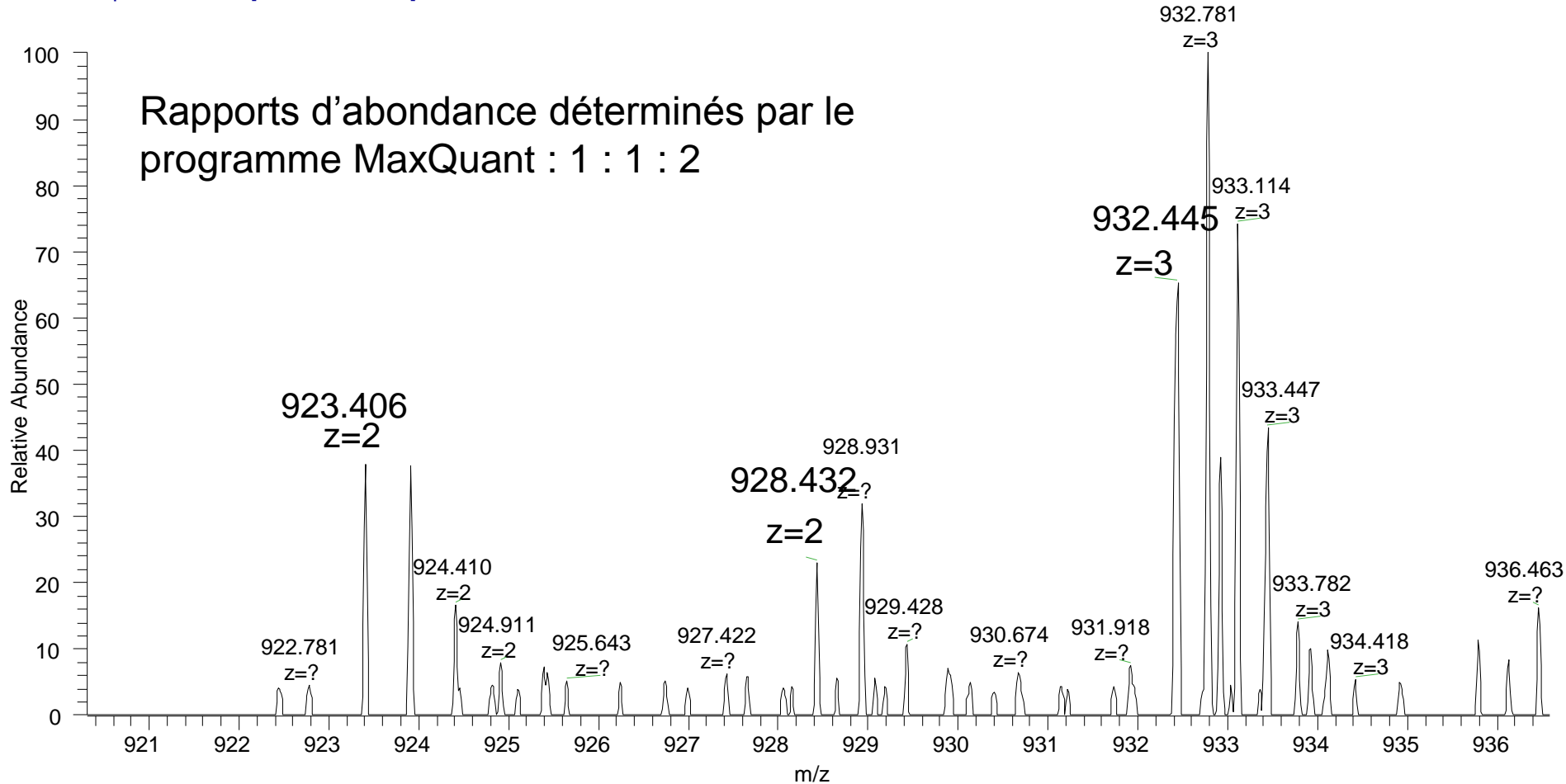
# Ecueils restants (1)

## Exemple d'une limitation bioinformatique :

Quantification relative de trois échantillons marqués différentiellement (SILAC)

silac1\_msa\_el3 #1010 RT: 19.07 AV: 1 NL: 5.13E5  
T: FTMS + p NSI Full ms [400.00-1400.00]

Rapports d'abondance déterminés par le programme MaxQuant : 1 : 1 : 2



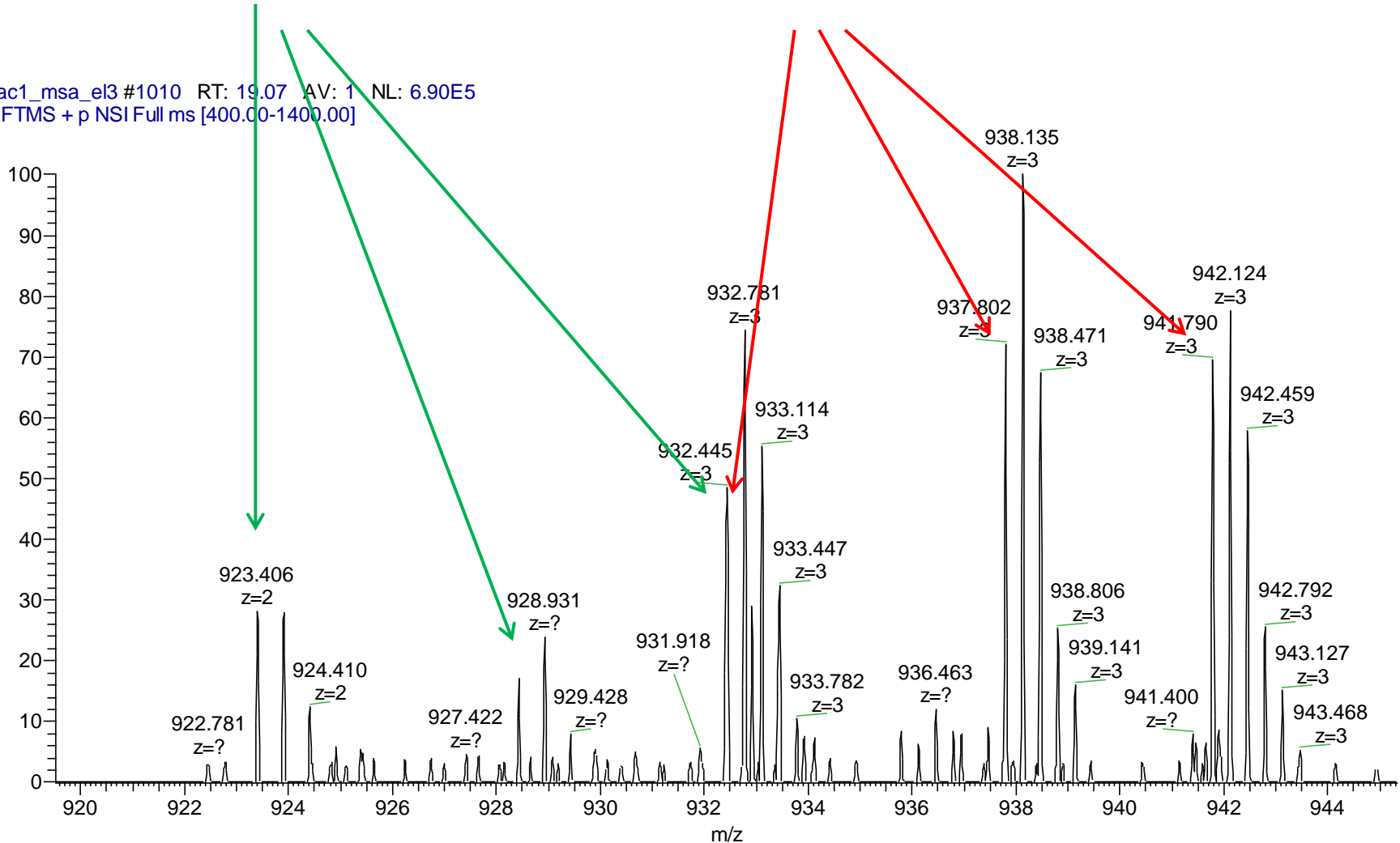
# Ecueils restants (1)

Si on dé-zoome :

Triplet du peptide d'intérêt

Triplet « contaminant »

silac1\_msa\_el3 #1010 RT: 19.07 AV: 1 NL: 6.90E5  
T: FTMS + p NSI Full ms [400.00-1400.00]



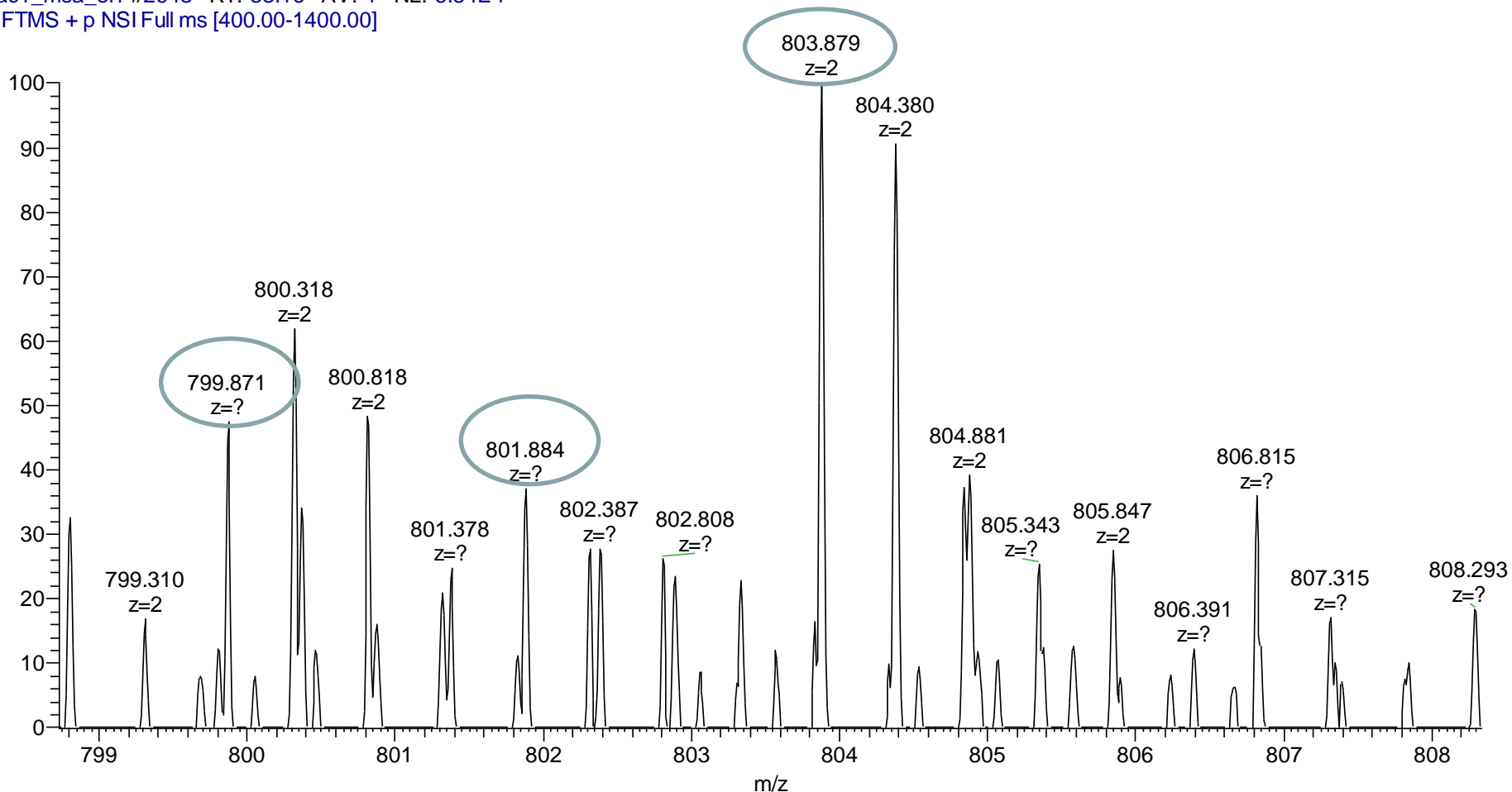
Conclusion : l'état de charge déterminé par la MS (LTQ-Orbitrap) n'est pas pris en compte par MaxQuant !!

# Ecueils restants (2)

Quid est de l'intérêt de la haute résolution ?

En analyse protéomique classique, une haute résolution en MS est assez peu mise à profit. En cas de co-élution et donc de co-fragmentation de deux espèces quasi-isobares, les logiciels d'interprétation des données classiquement utilisés ne permettent généralement pas d'identifier les deux séquences.

silac1\_msa\_el1 #2643 RT: 33.16 AV: 1 NL: 5.84E4  
T: FTMS + p NSI Full ms [400.00-1400.00]





# Petit bilan

Apport indéniable de la haute précision de mesure de masse en mode MS et en mode MS/MS à la fiabilité des identifications de peptides en mélanges complexes.

Performances des appareils FT (FTICR ou Orbitrap) mais aussi des TOF de générations récentes.

Développements de méthodes s'affranchissant de la sélection de précurseurs pour identifier plus d'espèces...y compris si elles ne sont pas visibles en détection MS (cf Gillet *et al.*, Mol. Cell. Proteomics 2012).

En analyses ciblées, intérêt de l'approche PRM pour gagner en facilité de mise en œuvre des analyses, en fiabilité d'identification et en précision de quantification.

Nécessité de démocratisation des programmes bioinformatiques / de davantage de développements pour pleinement mettre à profit les gains en capacités instrumentales.