

Identificação de proteínas através de espectrometria de massa

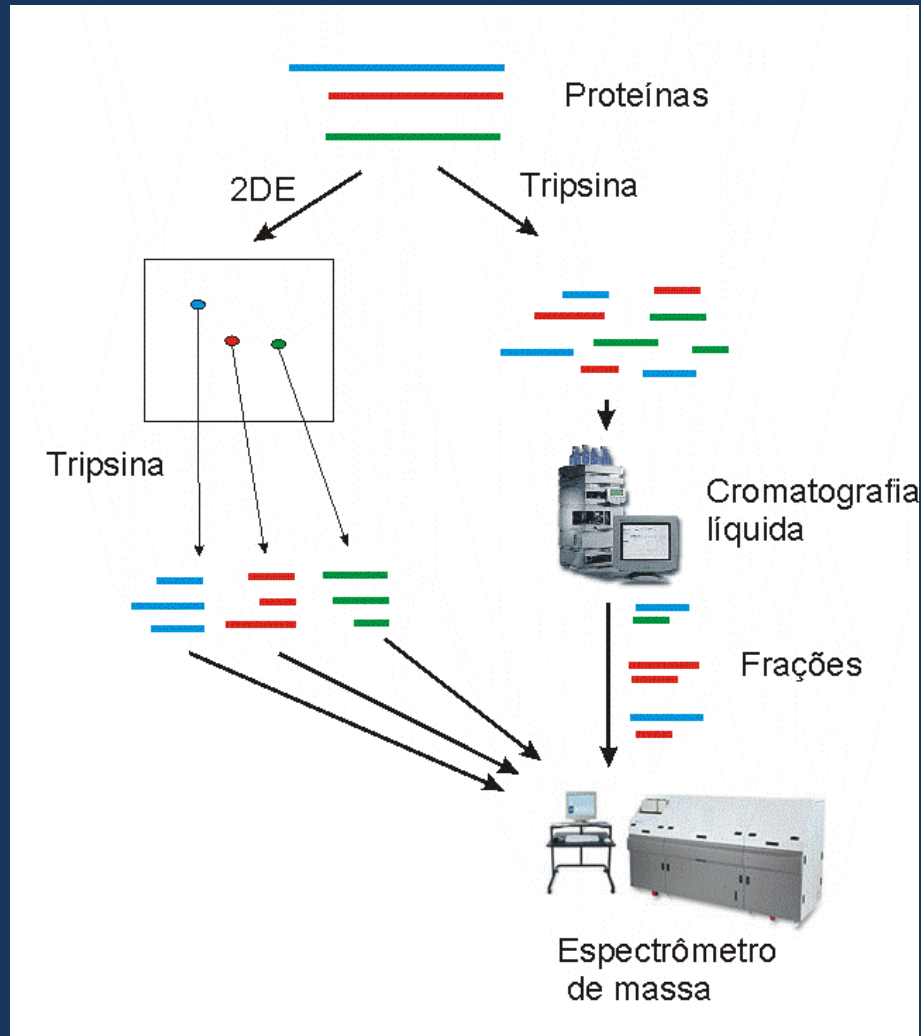
Proteomica

- Durante o curso vimos que é possível realizar medidas de presença e abundancia de moléculas de mRNA e deste modo realizar inferências sobre a presença de proteínas que são os verdadeiros efetores da maioria dos processos celulares.
- Entretanto existem mecanismos pós-transcrionais que podem afetar o nível de proteínas e que não são levados em conta quando medimos o nível de mRNAs
- Os dados de mRNA dão indicação de presença de uma proteína, mas não dão nenhuma informação sobre a localização da proteína dentro da célula.

Proteomica

- Existem metodologias utilizando espectrometria de massa que permitem a identificação de seqüências protéicas a partir da medida de sua massa
- A conjunção desta metodologia com técnicas de separação de proteínas permitem o estudo do conteúdo protéico de um organismo, constituindo assim a chamada proteomica
- Técnicas de fracionamento celular permitem que diferentes porções da célula sejam separadas e que seu conteúdo seja determinado

Proteômica

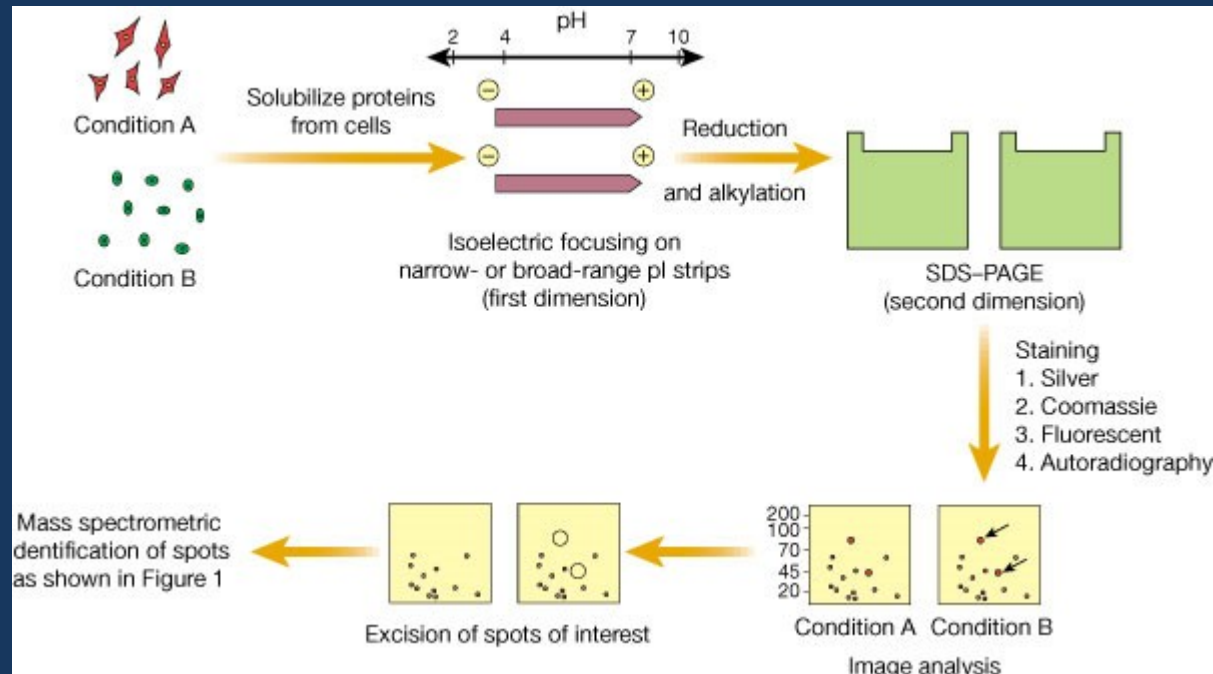


A separação de proteínas pode ser realizada através de métodos eletroforeticos ou cromatográficos

Tratamento com proteases permitem com que peptídeos de pequeno tamanho sejam formados

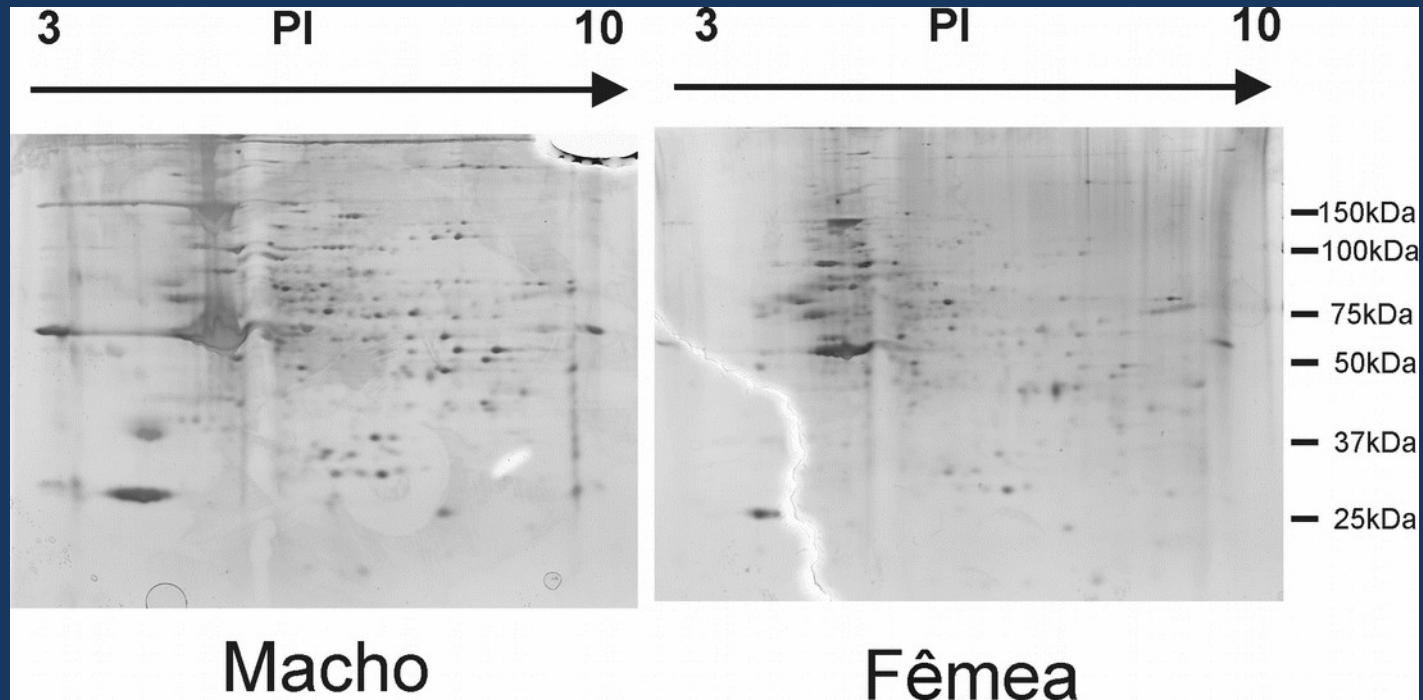
Analises com espectrometria de massa permitem com que estes peptídeos (e não a proteína inteira) sejam identificados

Eleetroforese em duas dimensões



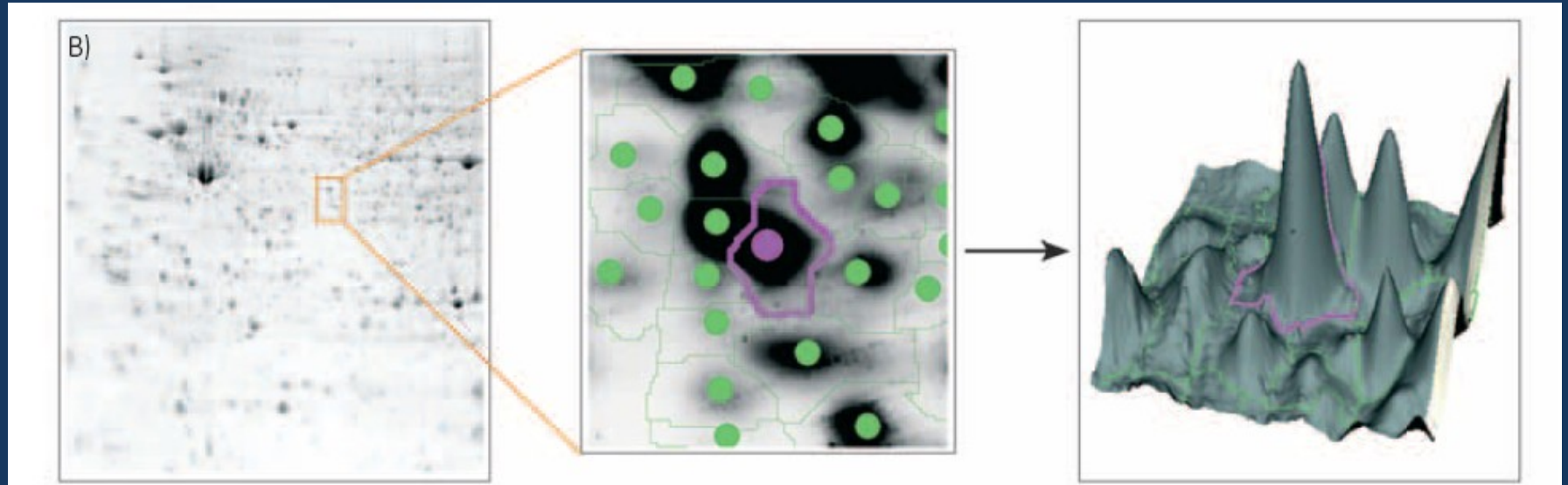
Proteínas podem ser separadas pelo seu ponto isoelétrico através de um procedimento inicial de isoeletrofocalização e depois separados em uma segunda dimensão pelo seu peso molecular através de SDS-PAGE

Eletroforese em duas dimensões



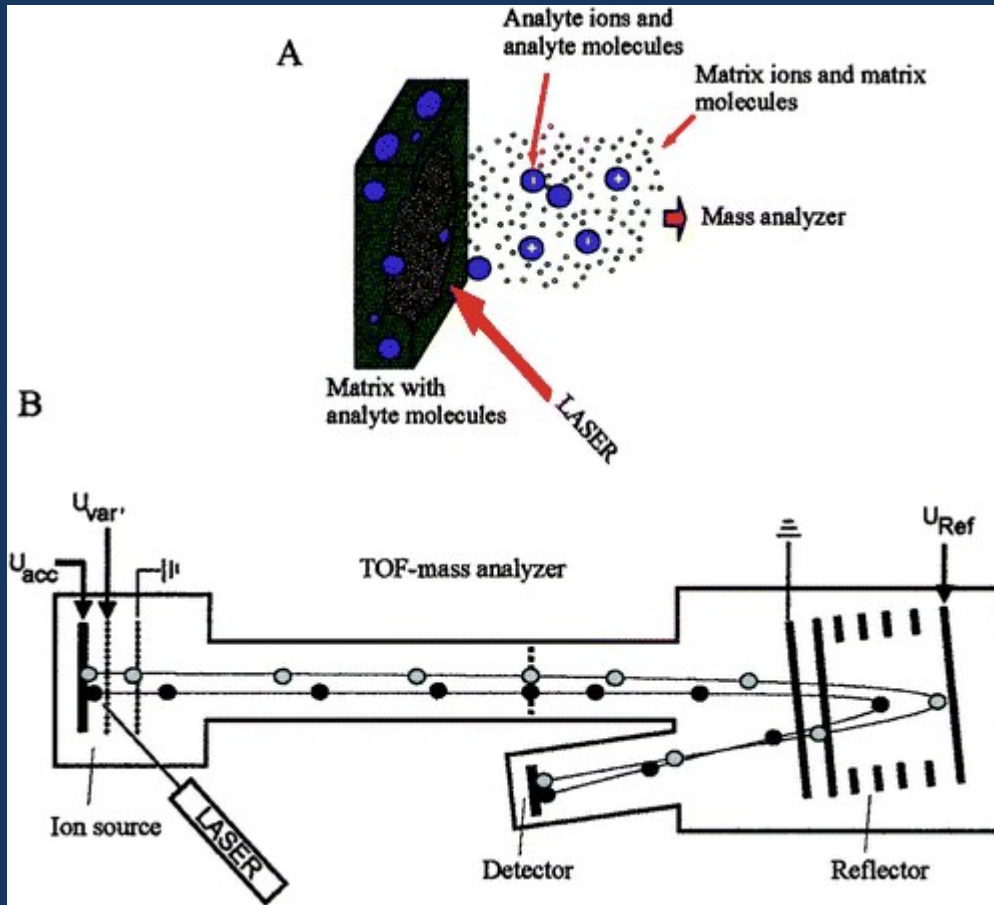
- Exemplo de separação em eletroforese em duas dimensões. Cada ponto representa uma diferente proteína que possui PI e peso molecular determinados

Analise de géis 2D



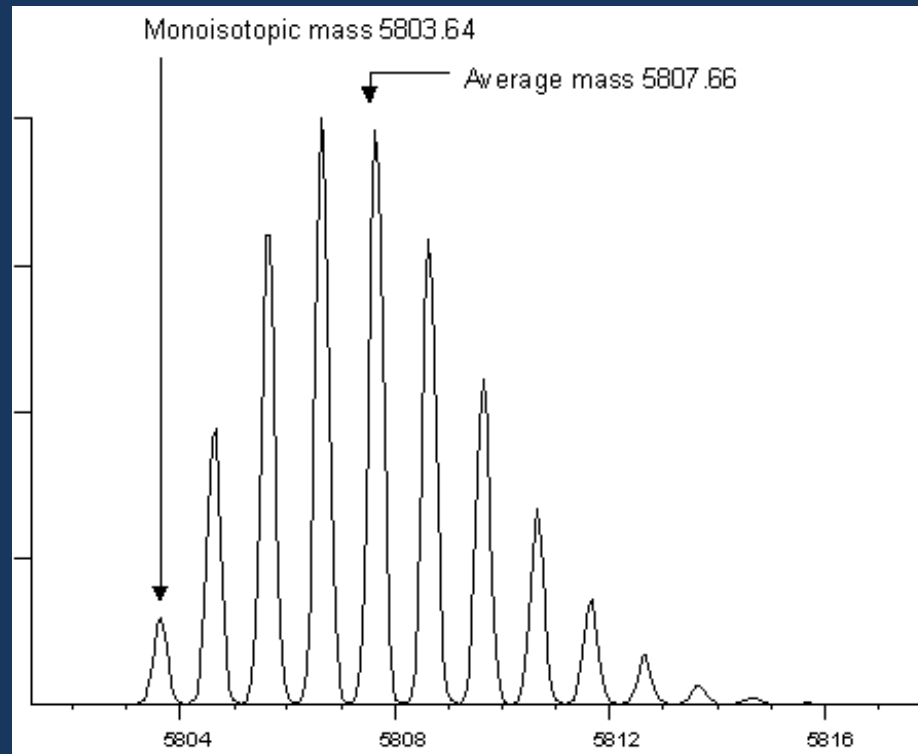
- Programas de reconhecimento de imagem detectam os pontos presente no gel e realizam análises densitométricas, permitindo a comparação entre diferentes amostras

Espectrometria de massa



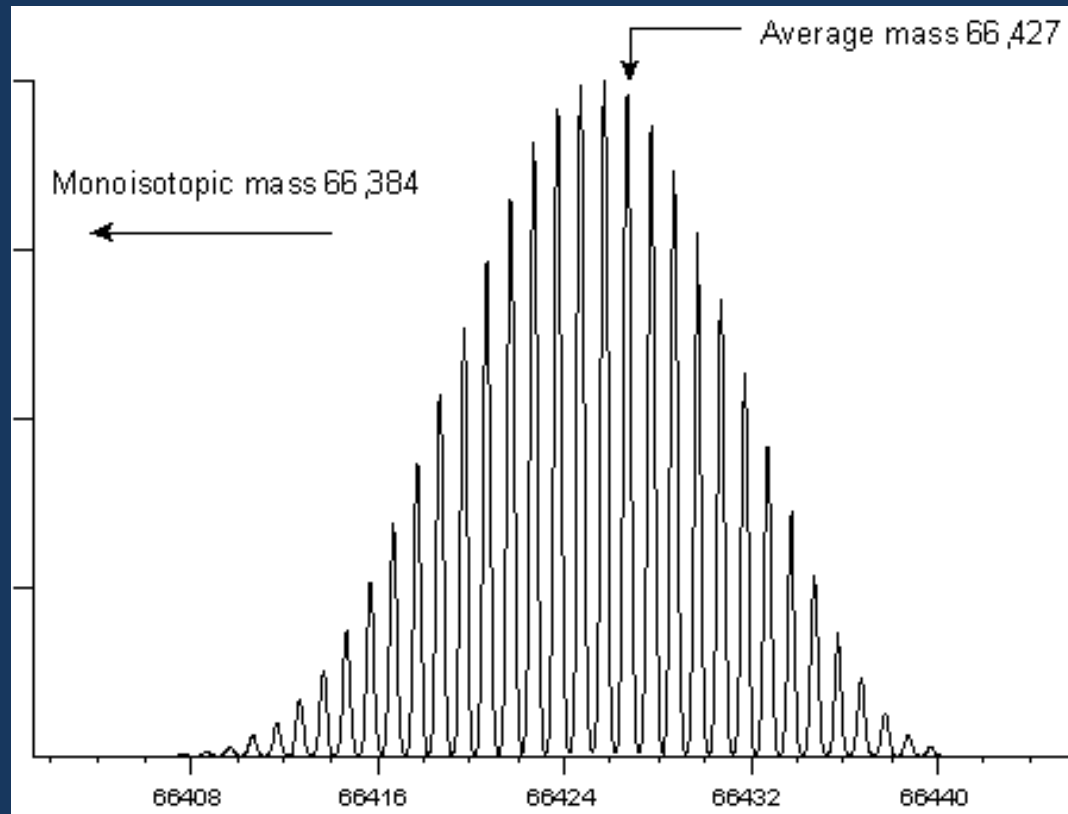
- Moléculas dos peptídeos são ionizadas utilizando diferentes métodos
- Moléculas ionizadas são aceleradas e diferentes métodos espectrométricos podem ser utilizados para medida da sua massa

Espectrometria de massa



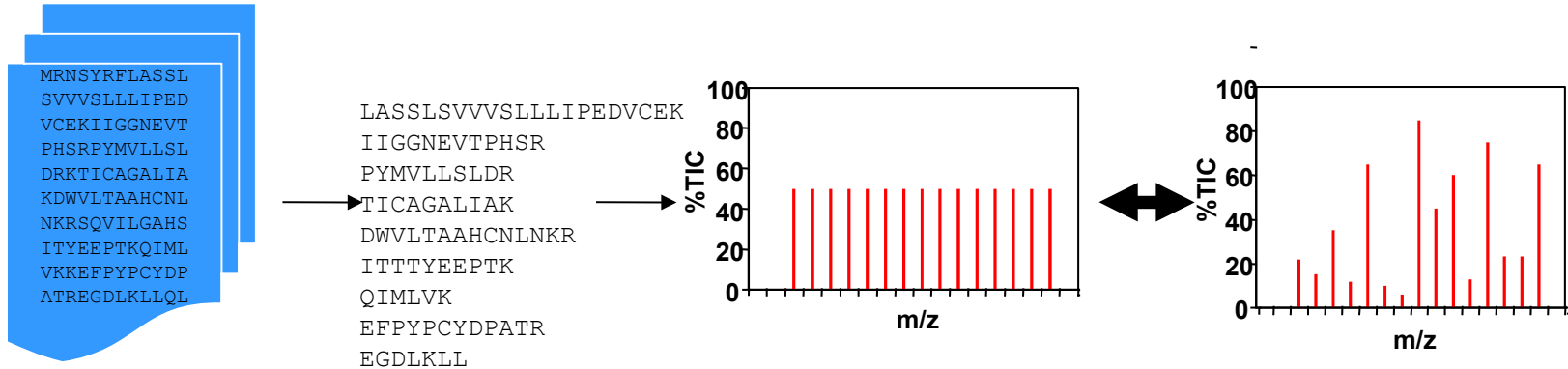
Cada peptídeo gerará diversos picos devido a presença de isótopos na molécula. De modo geral, considera-se a massa monoisotópica como medida da massa da molécula para análises posteriores

Espectrometria de massa



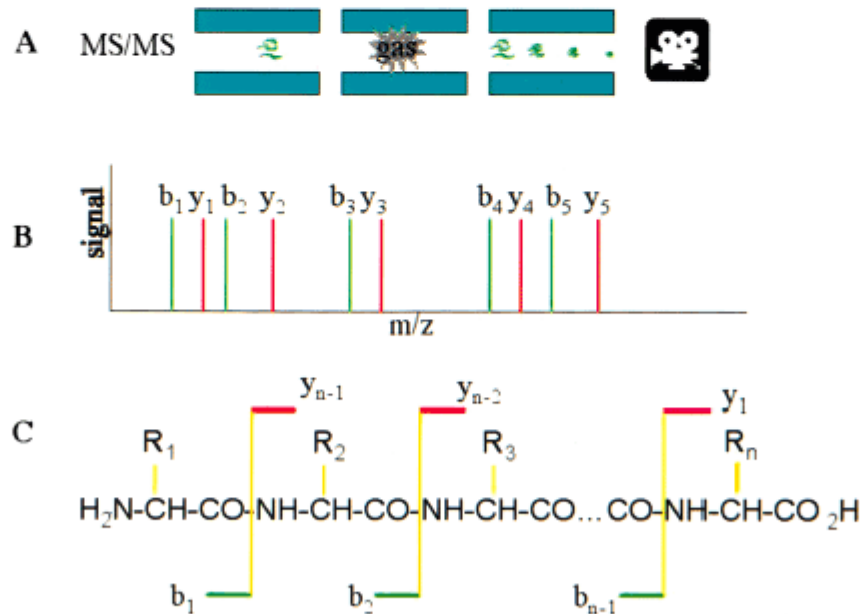
No entanto para moléculas muito grandes é muito difícil detectar moléculas com a massa isotópica e por isso a massa média passa a ser considerada

Analise de dados de MS



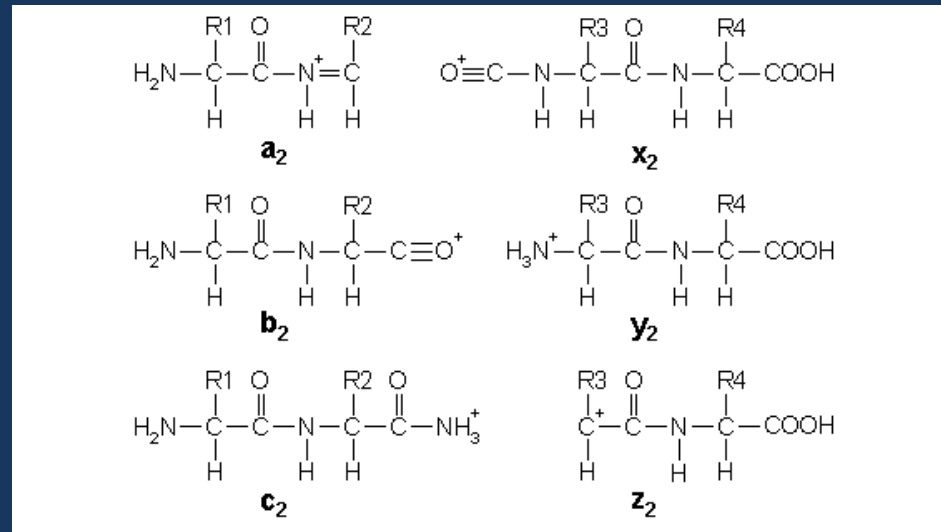
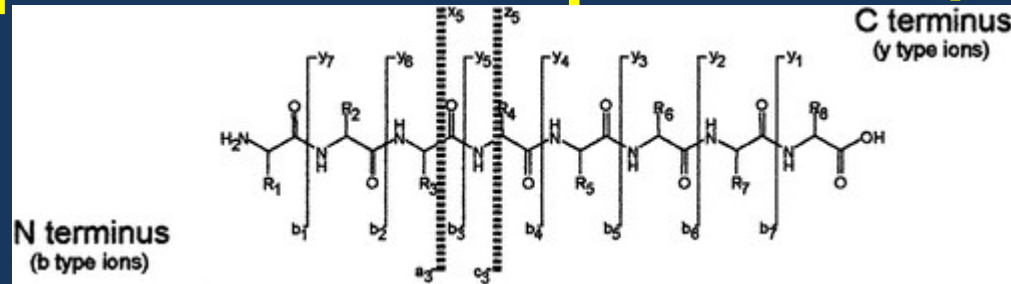
- A partir dos dados experimentais das massas detectadas para os peptídeos é possível realizar análises comparando o espectro obtido com os espectros teóricos de proteínas depositadas em bancos de dados
- Apesar deste tipo de análise se útil na identificação de proteínas ainda podem existir casos muito ambiguos

Espectroscopia MS/MS



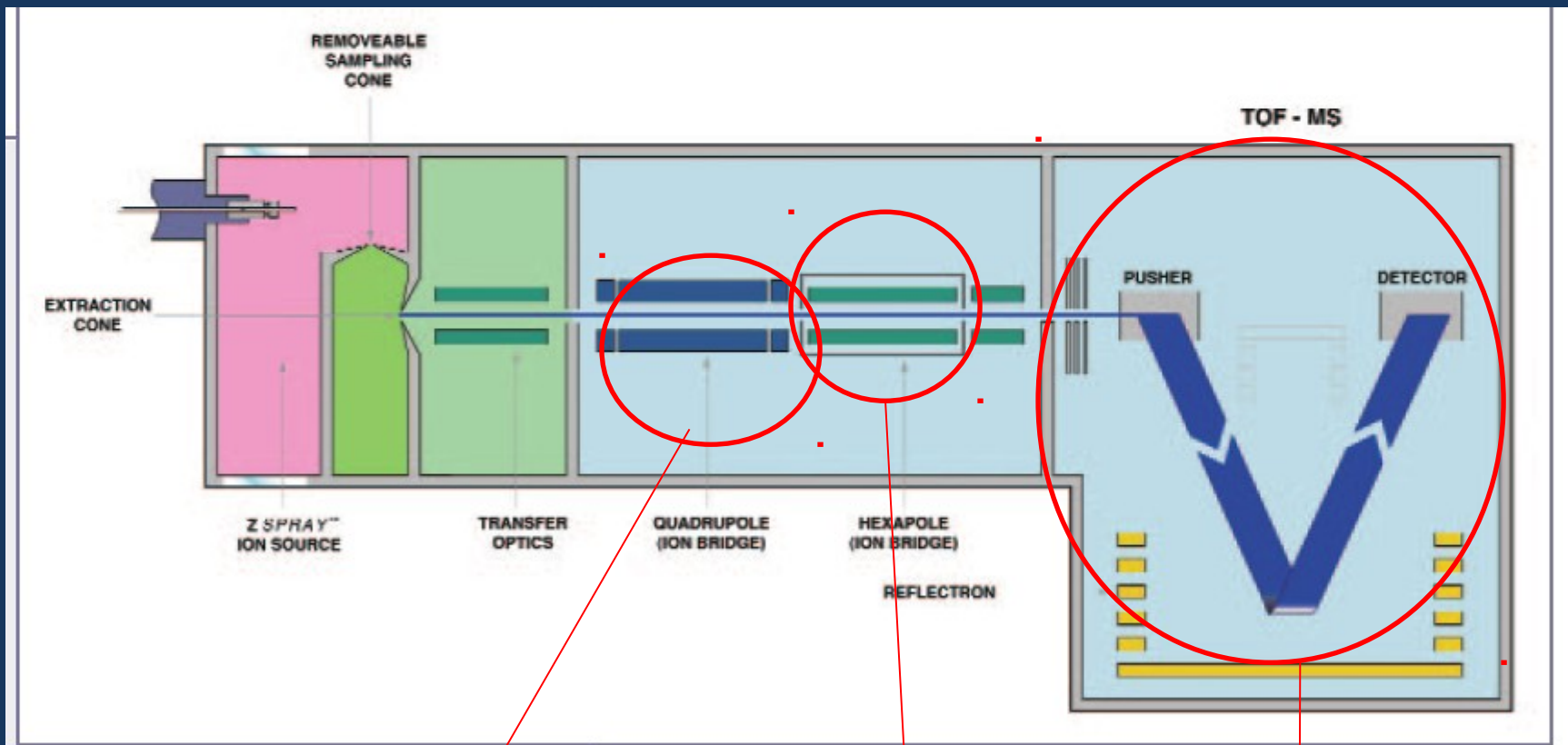
A espectrometria MS/MS permite a aquisição de maiores informações sobre o peptídeo pois inclui uma etapa adicional de colisão do fragmento que permite a detecção de novos fragmentos formados a partir deste primeiro

Espectroscopia MS/MS



- a fragmentação dos peptídeos tende a ocorrer na ligação peptídica fazendo que se formem íons das series b e y.

Espectroscopia MS/MS

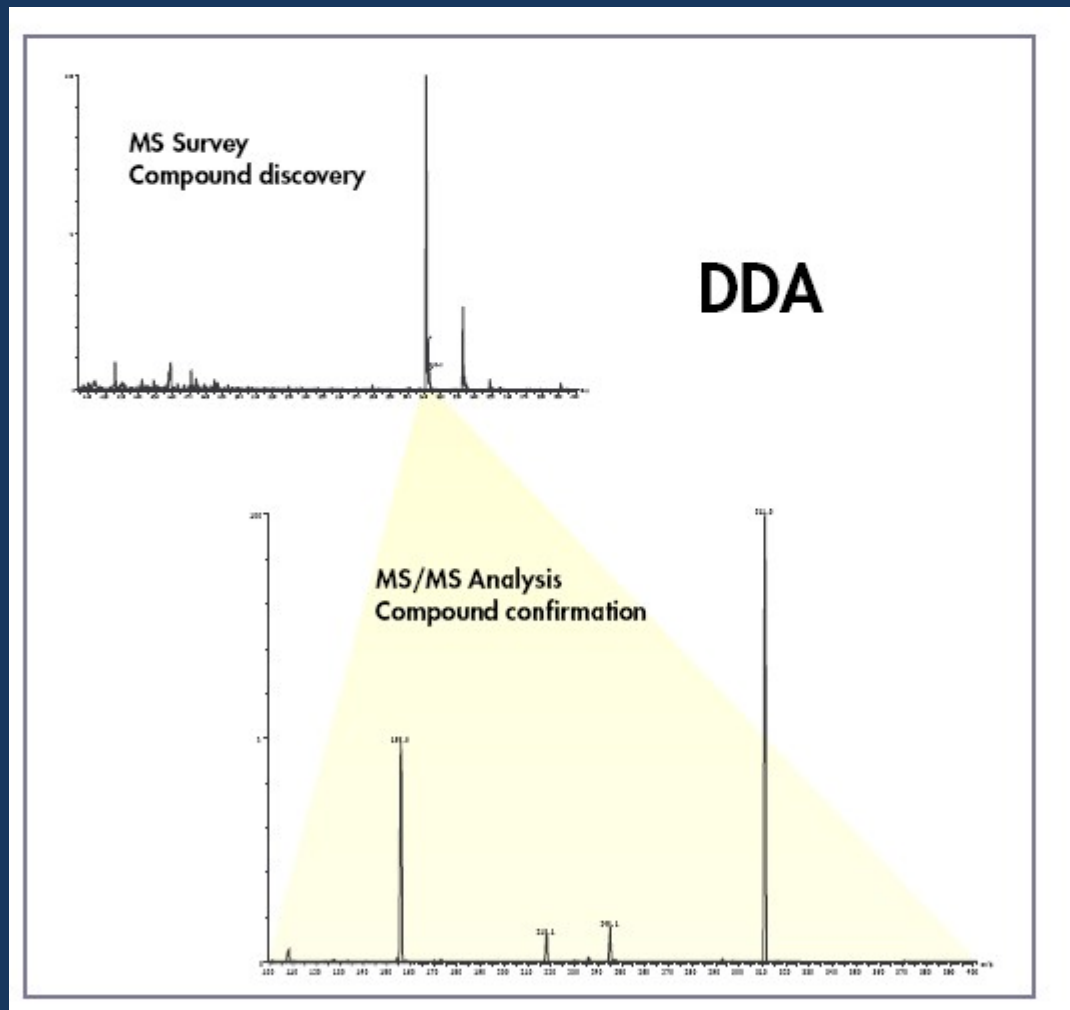


Realiza seleção de
fragmento de tamanho
desejado

Câmara de colisão

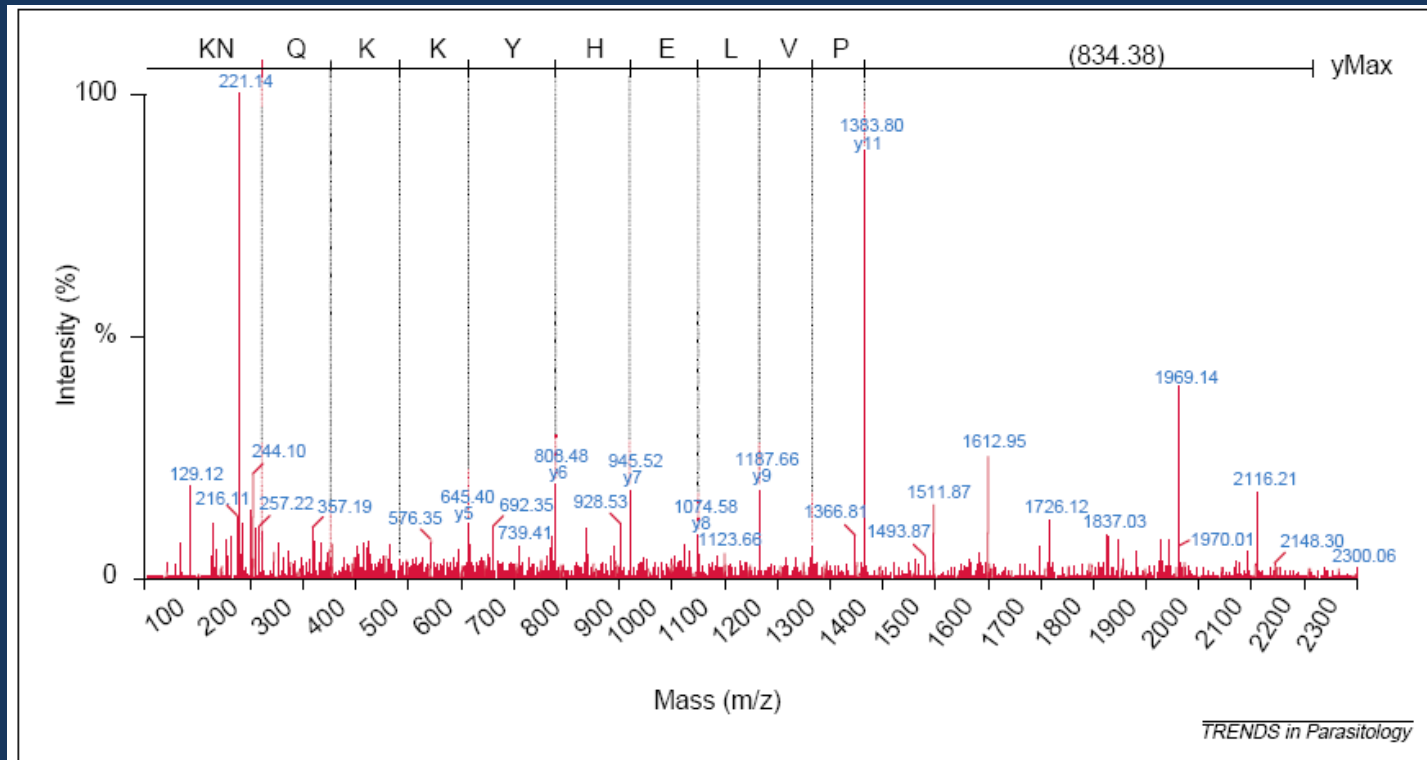
Deteção de fragmento
formados

Espectroscopia MS/MS



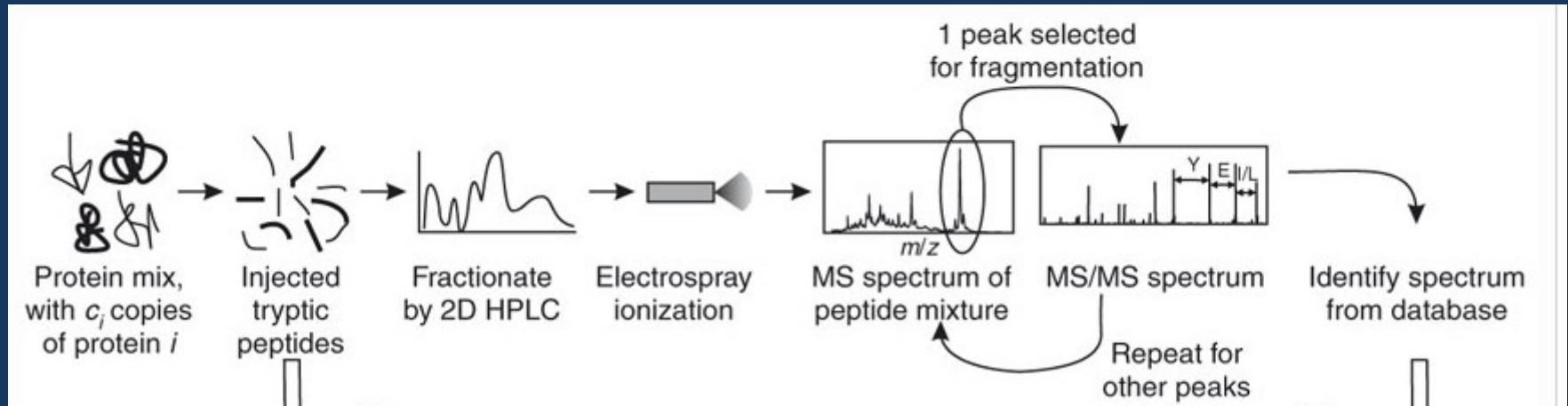
- Espectro inicial de massa passa a ter espectros associados ao resultado da fragmentação do íon detectado

Espectroscopia MS/MS



Espectro de fragmentação permite seqüenciamento do peptídeo
Seqüência obtida é muito curta para atribuição de função-
necessidade de comparação com bases de proteínas

Shotgun proteomics



- Abordagem que realiza uma tripsinização de todas as proteínas de um extrato protéico e realiza posterior separação dos peptídeos gerados através de técnicas cromatográficas
- Técnica muito mais simples que separação por eletroforese 2D, mas fornece menos informações sobre a proteína

Busca em banco de dados

- Uma análise em larga escala do proteoma de um organismo pode gerar entre centenas e milhares de espectros de massa.
- Deste modo são necessárias ferramentas que permitam a interpretação e integração destes dados
- Programas do espectrômetro de massa já analisam o espectro e fornecem uma lista de valores que correspondem a massa apurada para cada íon
- Programas adicionais irão utilizar estas listas para procurar padrões semelhantes em bancos de espectros teóricos formados a partir de bancos de dados de proteínas

MASCOT

- Programa de busca em bases de dados a partir de um conjunto de espectros de massa (versão web somente em bases de dados predefinidas, versão local (Paga) tem possibilidade de utilizar bases próprias)
- Fornece escores de modo que é possível realizar uma avaliação da significância do resultado
- Permite a integração de dados em casos de buscas em experimento de shotgun proteomics

MASCOT

MASCOT MS/MS Ions Search

Database: **MSDB**

Enzyme: Trypsin

Fixed modifications: Carbamidomethyl (C)

Variable modifications: Acetyl (K), Acetyl (N-term), Acetyl (Protein N-term), Amidated (C-term), Amidated (Protein C-term)

Peptide tol. \pm 1.2 Da, # ^{13}C 0, MS/MS tol. \pm 0.6 Da

Peptide charge 1+, Monoisotopic, Average

Data file: D:\Ricardo\Ricardo-SaoCarlos\ Arquivo...

Data format: Mascot generic, Precursor: m/z

Instrument: Default, Error tolerant: ☐

Decoy: ☐ Report top: AUTO hits

Start Search ... Reset Form

Banco de dados a ser procurado

Modificações existentes nos aminoácidos

Taxa de tolerância de erro

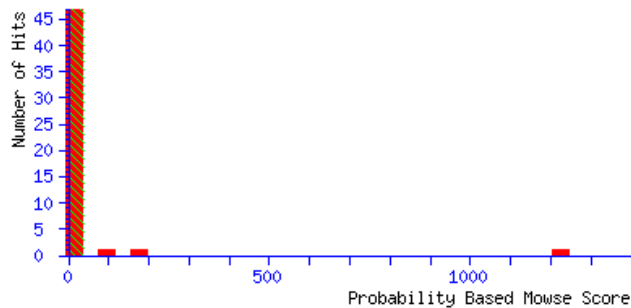
MASCOT

Probability Based Mowse Score

Ions score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event.

Individual ions scores > 38 indicate identity or extensive homology ($p < 0.05$).

Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



- Gráfico indica a probabilidade de que as proteínas identificadas não são um evento randomico

MASCOT

1. [CH60 HUMAN](#) Mass: 61016 Score: 1225 Queries matched: 31
 60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa chaperonin) (CPN60) (Heat shock)
☐ Check to include this hit in error tolerant search

| Query | Observed | Mr(expt) | Mr(calc) | Delta | Miss | Score | Expect | Rank | Peptide |
|--|-----------|-----------|-----------|---------|------|-------|---------|------|---|
| <input checked="" type="checkbox"/> 11 | 417.1822 | 832.3498 | 832.3828 | -0.0329 | 0 | 45 | 0.016 | 1 | K.APGFGDNR.K |
| <input checked="" type="checkbox"/> 12 | 422.7433 | 843.4720 | 843.5066 | -0.0346 | 0 | 46 | 0.017 | 1 | K.VGEVIVTK.D |
| <input checked="" type="checkbox"/> 13 | 430.7328 | 859.4510 | 859.4837 | -0.0327 | 0 | 36 | 0.15 | 1 | K.IPAMTIAR.N + Oxidation (M) |
| <input checked="" type="checkbox"/> 15 | 451.2499 | 900.4853 | 900.5280 | -0.0428 | 0 | 52 | 0.0039 | 1 | K.LSDGVAVLK.V |
| <input checked="" type="checkbox"/> 16 | 456.7806 | 911.5467 | 911.5804 | -0.0337 | 0 | 59 | 0.00056 | 1 | K.VGLQVVAVK.A |
| <input checked="" type="checkbox"/> 21 | 480.7447 | 959.4748 | 959.5036 | -0.0288 | 0 | 45 | 0.017 | 1 | R.VTDALNATR.A |
| <input checked="" type="checkbox"/> 24 | 595.7855 | 1189.5565 | 1189.6012 | -0.0447 | 0 | (57) | 0.0011 | 1 | K.EIGNIISDAMK.K |
| <input checked="" type="checkbox"/> 25 | 603.7720 | 1205.5294 | 1205.5962 | -0.0668 | 0 | 60 | 0.00048 | 1 | K.EIGNIISDAMK.K + Oxidation (M) |
| <input checked="" type="checkbox"/> 26 | 608.3099 | 1214.6052 | 1214.6507 | -0.0455 | 0 | 73 | 2.2e-05 | 1 | K.NAGVEGSLIVEK.I |
| <input checked="" type="checkbox"/> 27 | 617.2857 | 1232.5569 | 1232.5885 | -0.0316 | 0 | 81 | 4e-06 | 1 | K.VGGTSDVEVNEK.K |
| <input checked="" type="checkbox"/> 31 | 672.8375 | 1343.6605 | 1343.7085 | -0.0480 | 0 | 64 | 0.00016 | 1 | R.TVIEQSWGSFK.V |
| <input checked="" type="checkbox"/> 34 | 714.8884 | 1427.7623 | 1427.8058 | -0.0435 | 0 | (65) | 0.00014 | 1 | R.GVMLAVDAVIAELK.K |
| <input checked="" type="checkbox"/> 35 | 714.8938 | 1427.7730 | 1427.8058 | -0.0327 | 0 | (73) | 2.1e-05 | 1 | R.GVMLAVDAVIAELK.K |
| <input checked="" type="checkbox"/> 36 | 722.8849 | 1443.7552 | 1443.8007 | -0.0455 | 0 | 75 | 1.2e-05 | 1 | R.GVMLAVDAVIAELK.K + Oxidation (M) |
| <input checked="" type="checkbox"/> 37 | 722.8934 | 1443.7722 | 1443.8007 | -0.0285 | 0 | (73) | 2.2e-05 | 1 | R.GVMLAVDAVIAELK.K + Oxidation (M) |
| <input checked="" type="checkbox"/> 39 | 752.8643 | 1503.7141 | 1503.7490 | -0.0349 | 0 | 90 | 4.3e-07 | 1 | K.TLNDELEIEGGMK.F |
| <input checked="" type="checkbox"/> 40 | 760.8461 | 1519.6777 | 1519.7439 | -0.0662 | 0 | (89) | 4.7e-07 | 1 | K.TLNDELEIEGGMK.F + Oxidation (M) |
| <input checked="" type="checkbox"/> 45 | 640.3281 | 1917.9625 | 1918.0636 | -0.1010 | 0 | 102 | 2.1e-08 | 1 | K.ISSIQSIVPALEIANHR.K |
| <input checked="" type="checkbox"/> 46 | 960.0327 | 1918.0509 | 1918.0636 | -0.0127 | 0 | (87) | 5.1e-07 | 1 | K.ISSIQSIVPALEIANHR.K |
| <input checked="" type="checkbox"/> 48 | 1019.5106 | 2037.0067 | 2037.0153 | -0.0087 | 0 | 52 | 0.0015 | 1 | R.IQEIEQLDVTTSEYEK.E |
| <input checked="" type="checkbox"/> 51 | 1057.0537 | 2112.0929 | 2112.1323 | -0.0394 | 0 | 116 | 6.8e-10 | 1 | R.ALMLQGVDLLADAVAVTMGPK.G |
| <input checked="" type="checkbox"/> 52 | 1065.0399 | 2128.0653 | 2128.1272 | -0.0619 | 0 | (72) | 1.7e-05 | 1 | R.ALMLQGVDLLADAVAVTMGPK.G + Oxidation (M) |
| <input checked="" type="checkbox"/> 53 | 1065.0623 | 2128.1100 | 2128.1272 | -0.0172 | 0 | (26) | 0.63 | 1 | R.ALMLQGVDLLADAVAVTMGPK.G + Oxidation (M) |
| <input checked="" type="checkbox"/> 54 | 1073.0477 | 2144.0809 | 2144.1221 | -0.0412 | 0 | (93) | 1.3e-07 | 1 | R.ALMLQGVDLLADAVAVTMGPK.G + 2 Oxidation (M) |
| <input checked="" type="checkbox"/> 58 | 789.1062 | 2364.2968 | 2364.3264 | -0.0296 | 0 | (56) | 0.00052 | 1 | R.KPLVIIAEDVDGEALSTLVNLR.L |
| <input checked="" type="checkbox"/> 59 | 1183.1570 | 2364.2994 | 2364.3264 | -0.0270 | 0 | (65) | 5.3e-05 | 1 | R.KPLVIIAEDVDGEALSTLVNLR.L |
| <input checked="" type="checkbox"/> 60 | 789.1094 | 2364.3063 | 2364.3264 | -0.0201 | 0 | 95 | 6.3e-08 | 1 | R.KPLVIIAEDVDGEALSTLVNLR.L |
| <input checked="" type="checkbox"/> 61 | 828.1238 | 2481.3495 | 2481.3942 | -0.0447 | 0 | (26) | 0.42 | 1 | R.TALLDAAGVASLLTTAEVVVTEIPK.E |
| <input checked="" type="checkbox"/> 62 | 828.1322 | 2481.3748 | 2481.3942 | -0.0194 | 0 | 48 | 0.003 | 1 | R.TALLDAAGVASLLTTAEVVVTEIPK.E |
| <input checked="" type="checkbox"/> 64 | 854.0588 | 2559.1545 | 2559.2413 | -0.0868 | 0 | 75 | 5.1e-06 | 1 | K.LVQDVANNTEEAGDGTITATVLAR.S |
| <input checked="" type="checkbox"/> 65 | 1038.5031 | 3112.4873 | 3112.5023 | -0.0150 | 0 | 13 | 6.7 | 1 | K.DMAIATGGAVFGEGLTLNLEDVQPHDLGK.V + Oxidation (M) |

- Resultado apresenta todos os espectros que podem ser associados a um peptídeo da sequência descrita

MASCOT

2. [CH60 DROME](#) Mass: 60771 Score: 174 Queries matched: 4
60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa chaperonin) (CPN60) (Heat shock
☐ Check to include this hit in error tolerant search

| Query | Observed | Mr(expt) | Mr(calco) | Delta | Miss | Score | Expect | Rank | Peptide |
|--------------------|-----------|-----------|-----------|---------|------|-------|---------|------|-------------------------------|
| 11 | 417.1822 | 832.3498 | 832.3828 | -0.0329 | 0 | 45 | 0.016 | 1 | K.APGFGDNR.K |
| 27 | 617.2857 | 1232.5569 | 1232.5885 | -0.0316 | 0 | 42 | 0.032 | 2 | R.VGGSSEVEVNEK.K |
| 59 | 1183.1570 | 2364.2994 | 2364.3264 | -0.0270 | 0 | 12 | 11 | 2 | R.KPLVIAEDIDGEALSTLVVNR.L |
| 64 | 854.0588 | 2559.1545 | 2559.2413 | -0.0868 | 0 | 75 | 5.1e-06 | 1 | K.LVQDVANNTNEEAGDGTITATVLAR.A |

- Sequências em preto são aquelas que tem um rank maior que 1. Isto é existe um peptídeo diferente no banco de dados que possui um escore melhor em relação a esse espectro do que o mostrado aqui. Deste modo, é improvável que este predição esteja correta