

# Prática 10 - Identificação de Proteínas Através da Análise de Espectros de Massa

Ricardo de Marco      Alessandro S. Nascimento

23 de Novembro de 2016

## 1 Introdução

Esta atividade tem como objetivo familiarizar o aluno com *softwares* de identificação de proteínas através de espectrometria de massa. Neste exercício, iremos utilizar o **MASCOT** para trabalhar com duas metodologias diferentes: Na de “*peptide mass fingerprint*”, a proteína é identificada a partir de um conjunto de massas representando diferentes peptídeos resultantes da tripsinização da proteína; Na análise de MS/MS, cada peptídeo é fragmentado e os valores de massa dos fragmentos resultantes ajudam na identificação deste peptídeo.

Para simular resultados reais, iremos utilizar programas que a partir de sequências protéicas são capazes de calcular as massas esperadas para os peptídeos formados, bem como programas mais especializados que são capazes de calcular os valores esperados para um espectro de fragmentação. Lembre-se que os valores que iremos obter são exatos e que uma amostra real apresentará muito mais ruído. Estes valores representariam nossos dados que serão utilizados no **MASCOT** para a identificação das proteínas.

## 2 Identificação de proteínas por *peptide mass fingerprint*

1. Busque a sequência proteica 3EXD\_A no banco do NCBI;
2. Vá ao site do programa [PeptideMass](#). Este programa simula o corte da sua proteína por uma protease de sua escolha e calcula as massas dos peptídeos gerados.
3. Cole sua sequência em formato FASTA no espaço reservado.
4. No item “*the peptide masses are with cysteines treated with:*” selecione iodoacetamide. Isso irá simular o tratamento das cisteínas reduzidas com um agente modificador que prevenirá a formação de pontes dissulfeto entre os peptídeos.
5. Mantenha os parâmetros  $[M+H]^+$  e monoisotopic. Isso significa que será calculada a massa do peptídeo mais um próton e que a massa considerada é do pico monoisotópico.

6. Mantenha a enzima a ser utilizada com tripsina e aperte o botão *perform the cleavage*;
7. O resultado será uma tabela com todos os peptídeos e suas respectivas massas. Peptídeos com massa abaixo de 500 daltons não são listados, pois normalmente esta é uma faixa mais problemática para análise, pois está sujeita a um maior ruído devido à presença de contaminantes não protéicos.
8. Abra o site do [MASCOT](#) e selecione a opção *peptide mass fingerprint* – > *perform search*;
9. Coloque seu nome e um e-mail válido nos campos indicados e selecione a “database”: SwissProt;
10. No campo “*fixed modification*” escolha carbamidomethyl (C) e aperte o botão «”. Com isso você está indicando ao programa que você espera que todas as cisteínas tenham sido modificadas com a reação com a iodoacetamida. O campo ao lado, “*variable modifications*”, indica modificações que podem ou não ter ocorrido em sua amostra (como por exemplo oxidação de metioninas), no caso deste exercício, não iremos marcar nenhuma modificação deste tipo.
11. Mantenha os parâmetros [M+H]<sup>+</sup> e monoisotopic, para serem concordantes com aqueles utilizados para na geração da lista de peptídeos.
12. Escolha de modo randômico o valor de massa molecular de 8 peptídeos gerados para sua proteína e copie-os no campo query.
13. Aperte o botão “*Start search*”
14. Na tela de resultados vá ao campo do lado do botão “*format as*” e selecione “*protein summary*”. Após isso aperte o botão “*Format as*”.
15. Examine os resultados apresentados. Os resultados de identificação obtidos são significativos?
16. Reduza o número de peptídeos cujos valores de massa foram adicionados ao MASCOT e descubra qual o número mínimo de peptídeos necessários para uma identificação significativa de sua proteína
17. A este número mínimo de peptídeos, adicione o valores de massas de dois peptídeos de uma outra proteína não relacionada com a primeira (simulando assim um contaminação de sua amostra com outra proteína) e refaça a predição no mascot. **Foi possível realizar uma identificação efetiva de alguma das proteínas?** (Q1)

### 3 Identificação de proteínas por MS/MS

1. Selecione um peptídeo de cada proteína utilizada na sequência anterior
2. Vá a página do programa [prospector](#).

3. Introduza a sua sequência de aminoácidos no espaço entre os termos “N term” e “C term”. Caso sua sequência possuir uma cisteína, substitua a letra C por C(Carbamidomethyl).
4. Selecione apenas as series de íons b e y.
5. Aperte o botão “induce fragmentation”
6. O resultado gerado informa as massas de todos os íons formados. Abra o excel ou o programa de planilhas do computador. Crie um arquivo seguindo os passos (modelo visual abaixo):
  - Na primeira linha do arquivo, escreva BEGIN IONS.
  - Na segunda linha, escreva TITLE=Nome do query que você quiser
  - Na terceira linha, escreva PEPMASS=XXXXX, onde X é o valor de massa do peptídeo não fragmentado.
  - Nas linhas abaixo, cole os valores fornecidos na tabela gerada pelo programa Prospector, em uma única coluna.
  - Após os dados, escreva END IONS na última linha do primeiro query. Para colocar um segundo query, repita os mesmo procedimentos (a partir da linha BEGIN IONS...) na linha logo abaixo à que você escreveu END IONS, e termine também com a linha END IONS.
7. Salve este arquivo no formato .txt. Para isso, você precisará, em "*salvar como*", selecionar o tipo de arquivo como texto (MS-DOS), que é o formato txt.
8. Abra o site do [mascot](#) e selecione a opção *MS/MS Ion*; Search:
9. No campo “*fixed modification*” escolha carbamidomethyl (C). Com isso você está indicando ao programa que você espera que todas as cisteínas tenham sido modificadas com a reação com a iodoacetamida
10. Selecione peptide charge 1+;
11. No espaço “*data file*” selecione o arquivo TXT que você gerou;
12. Selecione a “*database*”: SwissProt
13. Aperte o botão *start search*;
14. **Analise os resultados obtidos. Você foi capaz de identificar a proteína? (Q2)**
15. **Volte a seu arquivo txt e escolha randomicamente metade dos valores de massas resultante da fragmentação e retire-os do arquivo. Retire as linhas em branco. Salve o arquivo e refaça as análises. Você foi capaz de identificar as proteínas ainda assim? (Q3)**

b			y
---	1	C(Carbamidomethyl)	28
258.0907	2	P	27 3071.4311
387.1333	3	E	26 2974.3784
458.1704	4	A	25 2845.3358
571.2545	5	L	24 2774.2987
718.3229	6	F	23 2661.2146
846.3815	7	Q	22 2514.1462
943.4342	8	P	21 2386.0876
1030.4662	9	S	20 2289.0348
1177.5347	10	F	19 2202.0028
1290.6187	11	L	18 2054.9344
1347.6402	12	G	17 1941.8503
1478.6807	13	M	16 1884.8289



BEGIN IONS	
TITLE=Spectrum 1	
PEPMASS=602.3374	
258.0907	3000
387.1333	3000
458.1704	3000
571.2545	3000
718.3229	3000
846.3815	3000
943.4342	3000
1030.466	3000
1177.535	3000
1290.619	3000
1347.64	3000
1478.681	3000
1602.337	3000
1774.299	3000
1884.829	3000
END IONS	

Massa do peptídeo não fragmentado

Esta coluna indica a intensidade do sinal que neste caso foi arbitrariamente colocada como 3000