**Resumo Tema 1: Identificação, caracterização e quantificação de peptídeos e proteínas por espectrometria de massas;**

As proteínas são as biomoléculas mais abundantemente encontradas nos organismos e são essenciais em quase todos os processos biológicos. Além de papeis estruturais fazem parte de uma infinidade de processos celulares, como metabolismo, biossinalização, regulação genética, síntese de proteínas, transporte de solutos através de membranas, resposta imune e fotossíntese. Adicionalmente, constituem os produtos finais mais importantes das vias de informação e os instrumentos moleculares pelos quais a informação genética é expressa.

A regulação anormal da função proteica é um dos fatores mais relevantes no surgimento de uma doença, daí a importância de compreender o dinamismo destas biomoléculas.

As proteínas formam uma entidade altamente estruturada conhecida como proteoma, cujas proteínas constituintes desempenham suas funções em condições e locais específicos dentro da célula, em associação física ou funcional com outras proteínas ou biomoléculas. O proteôma de uma célula se adapta dinamicamente a mudanças externas e internas ( genéticas) e, assim, define o estado funcional da célula o que determina seus fenótipos. Descrever e compreender o proteoma completo e quantitativo, bem como sua função e dinâmica é o desafio da proteômica.

Bioquímicamente falando proteínas são polímeros de aminoácidos em que cada resíduo de aminoácido se liga ao próximo mediante um tipo de ligação covalente chamada de ligação peptídica.

Uma das técnicas mais amplamente utilizadas na descoberta e caracterização de proteínas é a espectrometria de massas, técnica analítica utilizada para medir a razão massa-carga (m/z) de uma ou mais moléculas presentes em uma amostra em fase gasosa. O espectrometro de massas detecta a presença e abundância de peptídeos (e outras biomoléculas como metabólitos, lipídios, carboidratos, ácidos nucléicos e proteínas) usando

propriedades fundamentais das moléculas, como massa e carga. Quando os peptídeos são ionizados (geralmente através de ganho de prótons), eles são chamados de íons peptídicos.

Todos os espectrômetros de massa têm como mínimo três componentes: uma fonte de íons, um ou mais analisadores de massa e um detector de íons. Na fonte de ionização, as moléculas são convertidas em íons em fase gasosa para que possam ser movidas e manipuladas por campos elétricos e magnéticos externos.Como os espectrômetros de massa só podem analisar em fase gasosa foram desenvolvidos métodos como ionização por electrospray (ESI) a fim de converter peptídeos em fase líquida a íons gasosos. Assim a amostra líquida contendo os peptídeos é bombeada através do orifício de entrada do espectrômetro, mantido em alta tensão. Ao chegar ao emissor da fonte, o fluxo constante do líquido se desintegra em gotículas pequenas, altamente carregadas, das quais a líquido se evapora rapidamente deixando íons peptídicos na fase gasosa. Em proteômica é comum usar sistemas de cromatografia líquida de alto desempenho (HPLC) para separar misturas de peptídeos usando nano fluxo para a injeção da amostra à fonte de íons ESI.

Quanto ao analisador de massa sua principal função é separar íons por suas razões massa-carga (m/z). Fundamentalmente, todos os íons são separados modulando suas trajetórias em campos elétricos. Os analisadores de massa diferem no princípio que usam para

separar íons, e isso define sua aplicação. Analisadores do tipo quadrupolo, analisadores por tempo de voo (TOF) ou analisadores Orbitrap, são os mais comuns em proteômica.

Normalmente o analisador de ións é seguido por uma ‘cela de colisão’ um tipo de analisador onde os íons podem ser fragmentados com colisão com um gás inerte como o N2. O analisador de massa geralmente funciona em conjunto com o sistema de detecção de íons, registrando a intensidade de cada m/z detectada, gerando um espectro de massa. Os espectros resultantes dos íons precursores são chamados MS1 enquanto os espectros dos fragmentos gerados são chamados de MS2 ou MS/MS.

Mas como os espectrômetros de massas sequenciam ou identificam peptídeos? As proteínas podem ser estudadas como moléculas intactas por espectrometria de massa, uma

abordagem proteômica conhecida como  *top-down*. Nesta estratégia

A vantagem da proteômica *top-down* é que torna possível a medição em conjunto de todas as modificações que ocorrem na mesma molécula, permitindo a identificação precisa de uma proteoforma.

No entanto, a abordagem mais amplamente utilizada para mapear proteomas inteiros é a abordagem *bottom-up*. De uma forma geral o fluxo de trabalho da abordagem bottom-up consiste de: 1) extração e lise das proteínas da amostra, 2) digestão das proteínas com uma enzima proteolítica que gera peptídeos com sequências específicas, 3) separação da mistura de peptídeos resultantes por cromatografia de fase reversa, a qual é acoplada à fonte de ionização por electrospray. Os íons precursores com m/z específico são primeiramente isolados e depois fragmentados, o que causa a quebra de ligações peptídicas o que gera espectros de MS/MS com sequências de picos que diferem pela massa de um aminoácido. Esta informação específica permite a identificação da sequência peptídica. Uma sequência de apenas alguns aminoácidos é suficiente para identificar um peptídeo de uma proteína. Geralmente, a identificação das proteínas que compõem um organismo envolve a geração de todos os possíveis espectros de fragmentação teóricos gerados a partir de um banco de dados, que são comparados contra os espectros obtidos experimentalmente.

Três métodos de aquisição de dados, são normalmente usados na proteômica *bottom-up*:

1. Aquisição dependente de dados (DDA), na qual o espectrômetro é configurado alternar entre o registro de todos os íons peptídicos que co eluem em um determinado tempo do gradiente de eluição cromatográfica (Fullscan) e fragmentar apenas íons peptídicos com uma intensidade determinada (MS2). Este modo de aquisição visa alcançar uma cobertura imparcial e completa do proteoma permitindo fazer uma análise não direcionada
2. Aquisição independente de dados (DIA), na qual é feita a fragmentação de todos os peptídeos que eluem da separação cromatográfica.
3. Multiple Reaction Monitoring (MRM) na qual o equipamento é configurado para monitorar uma m/z conhecida e seus produtos de fragmentação, visando a aquisição reproduzível, sensível e simplificada de um subconjunto de peptídeos de interesse. Em espectrometros de alta resolução como os orbitraps esta análise é chamada de Parallel Reaction Monitoring (PRM) na qual é monitorado o espectro inteiro de MS2 de cada íon precursor selecionado.Este método de aquisição permite uma análise proteômica direcionada (target).

As análises proteômicas atuais tem como objetivo não apenas a identificação, mas a quantificação de proteínas para comparar diferentes condições do organismo (saudável ou patológica), abundância de uma ou mais proteínas em diferentes tipos celulares, organelas e tecidos, e mudanças do proteôma provocadas por estímulos internos (ex, silenciamiento de genes) ou externos (uso de um medicamentos, mudanças de temperatura, etc).

Assim, a quantificação de peptídeos pode seguir duas estratégias:

-Quantificação label-free (LFQ), na qual os sinais dos peptídeos (geralmente MS1) são extraídos dos dados brutos, normalizados e comparados entre as condições em estudo. É experimentalmente o mais simples e geralmente a abordagem mais econômica, no entanto, esta estratégia tem maior variação de quantificação devido a diferenças na pureza do peptídeo o desempenho do instrumento o que pode impactar as comparações entre amostras individuais

-Quantificação label, na qual são usados isótopos estáveis para marcar diferentes condições experimentais. Os isótopos podem ser introduzidos metabolicamente, como é o caso de aminoácidos contendo isotopos pesados em culturas celulares (SILAC), que marcam o proteoma completo do sistema biológico em estudo. A quantificação relativa entre as condições estudadas, resulta da comparação de espectros MS1 de proteínas “leves” naturais combinadas com as proteínas pesadas marcadas no início da preparação da amostra.

Porém, atualmente a marcação química é a mais comumente usada para fazer quantificação relativa de proteínas. Esta consiste no uso de tags isobáricos que permanecem sem modificação, mas cuja distribuição de isótopica é revelada após a fragmentação. No caso peptídeos provenientes da digestão são marcados combinados antes da LC-MS/MS. Durante a fragmentação (CID ou HCD) é liberado um íon repórter diferente para cada amostra biológica,cuja intensidades revela a abundâncias relativas das proteínas entre as condições experimentais analisadas. A vantagem do uso de marcação para quantificação é a comparação de peptídeos que têm exatamente o mesmo comportamento físico-químico,mas têm diferenças previsíveis na massa.

A etapa final da análise proteômica bottom-up é a análise estatística e interpretação biológica dos dados. Uma vez calculadas as quantidades relativas de proteínas, várias ferramentas podem ser usadas para analisar estatisticamente e interpretar os resultados. Alguns programas como Perseus, Proteome Discoverer, TransProteomic Pipeline e MS Stats são projetados para processamento e análise estatística de dados proteômicos. Os dados também podem ser analisados ​​manualmente usando linguagem de programação como python ou R. De um modo geral, num experimento proteômico comparativo, são removidos contaminantes da lista de proteínas identificadas, e os dados das abundâncias são transformados a um logaritmo para dar às quantidades uma distribuição normal, depóis é feita a correção da variabilidade técnica entre execuções, e análise de mudanças usando testes estatísticos como teste t/ANOVA/ regressão linear. Após a conclusão da análise estatística, os resultados podem ser interpretados biologicamente. O banco de dados Gene Ontology é um recurso popular para estudar vias metabólicas que são alteradas no estudo. Outra ferramenta de análise busca por é o STRING e o CytoScape que fornecem informação sobre relações ou interações conhecidas entre proteínas de interesse.

Por outro lado, os avanços na espectrometria de massas atingiram agora um estado em que é possível uma infinidade de aplicações como a detecção de modificações pós traducionais (PTMs) e análises de células individuais (Single-cell proteomics), o que tem feito uma contribuição importante na medicina translacional, particularmente na identificação e uso rotineiro de biomarcadores.

Uma das ferramentas mais inovadoras é o estudo do proteoma de células individuais (Single-cell proteomics) enquanto retém toda a informação espacial do ambiente celular revelando diretamente a dinâmica intercelular, como interações receptor-ligante entre as células e seu microambiente.

Outro tipo de análise passível de ser feita por espectrometria de massas é o estudo de modificações pós-traducionais (PTMs). PTMs ocorrem durante ou após a biossíntese de proteínas e aumentam a diversidade funcional do proteoma. Algumas PTMs como fosforilação, acetilação, metilação, glicosilação, ubiquitinação, sumoilação (entre muitas outras modificações) influenciam todos os aspectos da biologia celular. A proteômica baseada em espectrometria de massas é a abordagem mais utilizada para análise de PTMs. No entanto, devido à baixa abundância e à natureza lábil de muitas PTMs, se faz necessário o enriquecimento de peptídeos modificados o para análise por espectrometria de massas