

ALTA EXATIDÃO EM ESPECTROMETRIA DE MASSA: DA COMPOSIÇÃO ELEMENTAR À ANOTAÇÃO EM METABOLÓMICA

ANTÓNIO E. N. FERREIRA*

A alta exatidão de massa associada à alta resolução em espectrometria de massa permite o cálculo da composição elementar dos compostos detetados em misturas complexas. O conhecimento da composição elementar restringe o número de resultados nas buscas em bibliotecas de compostos conhecidos, para efeitos de identificação de pequenas moléculas orgânicas. Em relação a compostos desconhecidos, a composição elementar é frequentemente a primeira anotação possível em bases de dados de pequenas moléculas e é o ponto de partida para a elucidação estrutural. Neste artigo é feita uma revisão dos principais passos envolvidos na determinação da composição elementar de pequenas moléculas a partir de valores de massa obtidos em instrumentos de alta exatidão..

INTRODUÇÃO

Durante a última década, o desenvolvimento técnico em espectrometria de massa (MS) resultou num aumento dramático de duas “figures of merit” instrumental: a resolução e a exatidão de massa. A resolução a um dado valor de massa m é geralmente expressa por $m/\Delta m$, em que Δm é a largura a meia altura de um pico e a exatidão de massa é geralmente expressa em ppm , segundo a definição:

$$ppm = \frac{(m_{\text{experimental}} - m_{\text{calculada}})}{m_{\text{experimental}}} 10^6$$

No mercado atual, os espectrómetros do tipo Ressonância Ciclotrónica de Ião com Transformada de Fourier (FT-ICR) conseguem atingir 0,1 a 1 ppm de exatidão de massa e resoluções acima de 1 000 000 [1]. Espectrómetros do tipo Orbitrap conseguem atingir 0,5 a 1 ppm de exatidão de massa e 100 000 de resolução. Embora a alta exatidão de massa seja importante na análise de macromoléculas biológicas e agregados supramoleculares, ela é essencial no campo da identificação sistemática e em larga escala de espécies de baixa massa molecular. Partindo de uma situação em que há um total desconhecimento sobre a estrutura do ião que deu origem a um determinado pico num espectro de MS, a alta exatidão permite restringir

a atribuição de identidade baseada numa busca sistemática realizada sobre uma base de dados de compostos possíveis: regra geral, quanto mais restrito for o intervalo de massa da busca, menor o número de compostos candidatos encontrados.

Uma área de aplicação particularmente importante e de elevado interesse é a metabolómica: a identificação em larga escala de compostos orgânicos de baixa massa molecular com interesse funcional em amostras biológicas complexas. Sob o ponto de vista analítico, em metabolómica, o problema é a necessidade de identificar uma mistura complexa de compostos orgânicos em quantidades quase vestigiais, alguns relativamente instáveis e sob a imposição de um ciclo de trabalho relativamente curto, dada a necessidade de proceder à análise de muitas amostras em tempo útil.

Em geral, a identificação inequívoca de um composto passa pela sua caracterização estrutural, combinando técnicas de espectrometria de massa *tandem* (MS^n) com a informação obtida a partir de outras metodologias (NMR 1D e 2D, espectroscopia de infravermelho, cristalografia de raios X, estudos de reatividade). No entanto, num cenário de identificação em larga escala, é impossível aplicar sistematicamente metodologias complementares para cada um dos compostos presentes num espectro de MS. Por outro lado, a utilização de MS^n pode não ser viável, seja pela lentidão do ciclo de trabalho, seja pelo facto de

não existir uma estratégia universal de elucidação estrutural por MS^n para pequenas moléculas, uma vez que estas não têm a regularidade de repetição de resíduos característica das macromoléculas biológicas (proteínas e ácidos nucleicos). É, por isso, de todo o interesse obter o máximo de informação possível a partir da utilização de apenas um nível de MS.

Neste contexto, a alta exatidão no valor de massa fornece um primeiro nível de restrição no universo da busca para identificação de compostos: ela permite o cálculo da composição elementar de compostos com a consequente possibilidade de pesquisar bibliotecas de compostos a partir de fórmulas elementares. Idealmente, a partir do espectro de massa de uma mistura complexa obteríamos uma única fórmula de composição elementar para cada composto detetado. Na realidade, mesmo em alta exatidão, obtemos um conjunto de composições elementares compatíveis com o valor de massa de cada um dos compostos detetados dentro de um intervalo de exatidão de massa especificado. Este cálculo é feito em vários passos, compreendendo a obtenção da massa monoisotópica de cada espécie, a remoção dos valores de massa de aductos, o cálculo combinatório das fórmulas possíveis e a aplicação de filtros ortogonais de restrição do número de fórmulas calculadas.

Em seguida é apresentado um resumo de cada um desses passos.

* Centro de Química e Bioquímica, Departamento de Química e Bioquímica, Faculdade de Ciências da Universidade de Lisboa
E-mail: aeferreira@fc.ul.pt

REMOÇÃO VIRTUAL DE ADUCTOS

Especialmente em LC/MS é frequente que grande parte dos iões subjacentes a um espectro sejam o resultado de aductos e que poucos picos possam ser associados a iões moleculares. Em modo positivo, por exemplo, são particularmente vulgares os aductos com H^+ , Na^+ e K^+ . Assim, na análise de um espectro de MS, devem ser identificadas ou calculadas as espécies sem aductos, de preferência calculando a massa da espécie molecular neutra. A um aducto com H^+ , por exemplo, deve ser subtraído o valor de 1,007276 Da¹ para obter a espécie neutra. Na interpretação de um espectro de massa é geralmente possível atribuir séries isotópicas à presença destes aductos. Por outro lado, o processo pode ser semiautomatizado recorrendo a “calculadores de aductos”, ferramentas que se podem encontrar no software associado a equipamentos de MS.

CÁLCULO DA MASSA MONOISOTÓPICA

Embora teoricamente o cálculo de fórmulas de composição elementar compatível com uma dada massa possa ser realizado a partir de massas de iões em qualquer estado de carga e qualquer composição isotópica, o procedimento está uniformizado para a utilização da massa monoisotópica da espécie neutra. A determinação de uma massa monoisotópica envolve, em primeiro lugar, a desconvolução do conjunto dos iões polivalentes no caso da ionização química ou por *electrospray*. Daqui resulta a identificação, num espectro de MS, da série isotópica correspondente ao ião monovalente de cada um dos compostos detetados. É essencial poder contar

com um instrumento de alta resolução, com o objetivo de resolver os picos da série isotópica junto à linha de base. Por via computacional é feito um ajuste à distribuição estatística das intensidades dos picos da série isotópica e a massa contendo apenas átomos de isótopos mais baixos para cada elemento, a massa monoisotópica, é assim calculada. O processo está, uma vez mais, frequentemente implementado no *software* de apoio ao equipamento de MS, sendo os algoritmos de cálculo de massa monoisotópica algo sofisticados. Na verdade, estes algoritmos têm de lidar com o facto da distribuição isotópica poder ser muito variada consoante a natureza e o número dos elementos presentes no ião, como está ilustrado na Figura 1.

CÁLCULO DE FÓRMULAS DE COMPOSIÇÃO ELEMENTAR COMPATÍVEIS

A partir da massa monoisotópica do composto neutro, são calculadas fórmulas de composição elementar compatíveis dentro de um intervalo de exatidão de massa. Trata-se de um problema simples de química combinatória. Podem aqui ser usados de uma forma eficiente algoritmos de busca exaustiva (“força bruta”), dada a simplicidade do problema e a grande rapidez de cálculo, característica dos meios computacionais atualmente ao dispor. O processo pode ser realizado num computador pessoal utilizando, por exemplo, o programa MWTWIN [2], ou ferramentas do tipo *Formula generators*, incluídas no software associado aos espectrómetros. Na busca de composições elementares tem de ser especificada uma lista de elementos que podem entrar na constituição das fórmulas,

a massa monoisotópica de referência e um intervalo de exatidão de massa, geralmente especificado em ppm. Na Tabela 1 são apresentadas as contagens resultantes do cálculo de fórmulas matematicamente possíveis em torno de dois valores de massa de referência, para diferentes valores de exatidão de massa. Como se pode observar, o efeito da exatidão de massa é muito significativo. É também interessante observar a distribuição das massas de todas as composições elementares para um dado conjunto de elementos tipicamente encontrados em compostos biológicos. Na Figura 2 estão indicadas as contagens das fórmulas possíveis ao longo do intervalo de 50 a 500 Da. São evidentes duas características: o número de composições calculadas tem um crescimento aproximadamente exponencialmente com o aumento de massa (Figura 2 A) e a maior das fórmulas tende a distribuir-se em torno dos valores unitários de massa (Figura 2 B).

CRITÉRIOS DE EXCLUSÃO

Finalmente, são aplicados critérios de exclusão de fórmulas de modo a reduzir o número de composições elementares calculadas. A necessidade deste refinamento decorre do elevado número de fórmulas que são geralmente obtidas levando em conta exclusivamente o cálculo combinatório de elementos. O crescimento exponencial do número de fórmulas leva a que, por exemplo, se considerarmos os elementos C, H, N, O, S e P, existam 3 572 670 fórmulas de composição elementar no intervalo de massa entre 50 e 500 Da. A alta exatidão pode não ser suficiente para reduzir o número de fórmulas para um valor viável para se realizar uma bus-

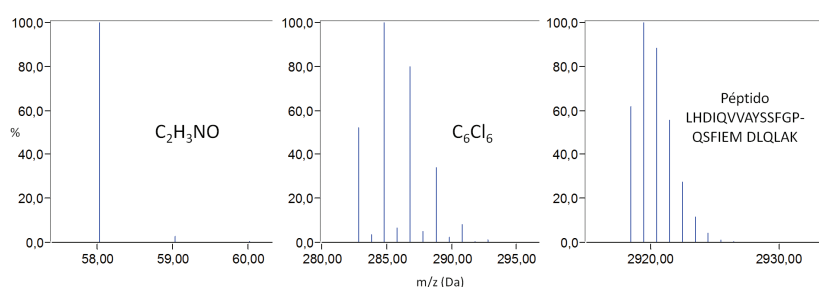


Figura 1 – Distribuição isotópica prevista para diversos compostos puros. Distribuições calculadas com o auxílio do programa MWTWIN [2], ferramenta *Isotopic Distribution Modelling*, para iões do tipo $[M+H]^+$ normalizadas à forma isotópica mais abundante (100 %)

Tabela 1 – Número de fórmulas de composição elementar matematicamente possíveis para diferentes valores de exatidão de massa. Consideraram-se combinações dos elementos C, H, N, O, S e P. Valores calculados com o auxílio do programa MWTWIN [2], ferramenta *Formula Finder*

Massa (Da)	Exatidão de massa (ppm)				
	10	5	1	0,5	0,1
400,002635	347	179	34	18	4
600,001757	2168	1086	218	103	23

ca em bases de dados de compostos de baixa massa molecular [3]. Assim, por exemplo, a uma exatidão de 0,1 ppm, já no limiar do limite que é hoje tecnologicamente possível com os espectrômetros de massa mais poderosos, ainda encontramos 23 fórmulas em torno da massa 600,001757 Da (Tabela 1). Idealmente, obteríamos apenas uma única fórmula de composição elementar para cada valor de massa monoisotópica em alta exatidão.

Os critérios adicionais de exclusão decorrem, por um lado, de considerações teóricas sobre a capacidade de se gerar uma fórmula molecular de um composto orgânico a partir de uma dada composição elementar, e, por outro lado, de regras empíricas que são verificadas pela grande maioria das entradas existentes em bibliotecas de compostos conhecidos.

Kind e Fiehn [4] estudaram a aplicação de alguns destes critérios, avaliando o seu impacto na redução do número de fórmulas candidatas após o cálculo combinatório. Neste estudo, estes autores enunciaram “sete regras de ouro” a aplicar como critérios de seleção. Alguns exemplos destas regras são:

- Número de átomos de cada elemento numa fórmula: o recenseamento da composição elementar dos milhões de entradas em bibliotecas de compostos revela que existem máximos naturais quanto ao número de átomos de cada elemento. Por exemplo, na biblioteca Dictionary of Natural Products, para compostos de massa inferior a 1000 Da, o número de átomos de N não excede 27 [4].

- Regras de LEWIS e SENIOR: estas regras impõem, respetivamente, a obediência à regra do octeto (com a consideração de valências múltiplas) e restrições quanto ao número de átomos e número de valências totais no composto para que possa existir um “grafo” molecular. Estas são regras de natureza teórica, para as quais existem exceções (espécies radicalares no caso da regra de LEWIS), mas que são verificadas pela quase totalidade dos compostos biológicos conhecidos.

- Quocientes H/C e N,O,P,S/C: uma análise à distribuição dos valores destes quocientes em bibliotecas de compostos conhecidos permite definir balizas que contêm quase todos os compostos conhecidos. Este é um filtro de simples implementação, embora não seja dos mais eficazes na exclusão de fórmulas [4].

- Padrões isotópicos: Embora este seja um critério muito poderoso na exclusão de fórmulas, requer informação espectral adicional: a relação entre a intensidade da espécie monoisotópica (no estado de carga monovalente) e a intensidade dos dois picos seguintes na série isotópica. Especificamente, designando por M a massa monoisotópica, é necessário fornecer a proporção entre M+1 e M+2 (consideram-se geralmente as intensidades normalizadas à massa monoisotópica). Os valores experimentais das intensidades destes dois picos de uma série isotópica estão sujeitos a erro experimental, dependendo da sensibilidade e resolução instrumental. Neste critério de exclusão é necessário

especificar, por isso, um desvio padrão para os valores de intensidade do sinal. O critério atribui uma pontuação a cada fórmula, sendo comparado o padrão isotópico previsto com o padrão isotópico experimental. Fórmulas com pontuação muito baixa são eliminadas. Este “filtro ortogonal” é dos mais eficazes na exclusão de fórmulas candidatas, sendo muitas vezes suficiente um valor de 5% no erro das intensidades para se conseguir uma redução de centenas de candidatos a menos de uma dezena [3, 4].

CONCLUSÃO

A possibilidade do cálculo da composição elementar a partir de um valor de massa de alta exatidão já tinha sido antecipada por altura do desenvolvimento de espectrômetros do tipo FT-ICR [5]. No entanto, a sua utilização como elemento informativo capaz de levar à identificação de um composto num espectro de MS complexo só foi possível com o desenvolvimento dos métodos computacionais capazes de automatizar o procedimento e com a disponibilização de bibliotecas de compostos de baixa massa molecular com interesse em determinadas áreas científicas. Embora a explosão combinatória do número de fórmulas possíveis seja o principal obstáculo à identificação em larga escala, sobretudo para massas a partir dos 600 Da, a construção de regras de exclusão e o aproveitamento das características de alto desempenho dos espectrômetros de gama alta atualmente disponíveis aproximam-nos do objetivo de obter uma única composição elementar por cada massa monoisotópica. É de salientar que o conhecimento da composição elementar é útil, não só numa perspetiva de restrição da pesquisa em bibliotecas de compostos, mas também como informação primária a ser atribuída na anotação de compostos desconhecidos. Finalmente, embora não seja possível distinguir isómeros ou compostos isobáricos (compostos com diferentes composições elementares, mas com uma diferença entre massas monoisotópicas não detetável experimentalmente), o conhecimento da composição elementar é um bom

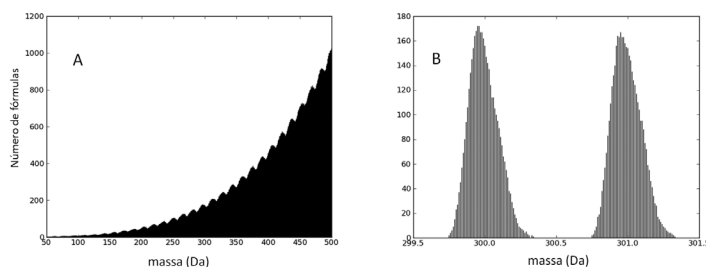


Figura 2 – Número de fórmulas de composição elementar matematicamente possíveis para diferentes valores de massa de referência. Consideraram-se as contagens das combinações dos elementos C, H, N, O, S e P, em intervalos de comprimento 0,01 Da.

A: todo o intervalo de 50 a 500 Da. B: detalhe junto ao valor de massa 300 Da. Fórmulas calculados com o auxílio do programa MWTWIN [2], ferramenta Formula Finder

ponto de partida para as metodologias de elucidação estrutural.

NOTA

¹ Neste texto é usado o símbolo Da (dalton) em vez do símbolo u para designar a unidade de massa atómica unificada, prática corrente em espectrometria de massa aplicada à

metabolómica e proteómica. Ambos os símbolos são atualmente aceites pela IUPAC.

REFERÊNCIAS

- [1] M. Scigelova, M. Hornshaw, A. Giannakopoulos, A. Makarov, *Molecular and Cellular Proteomics* (2011) **10**: M111.009431

- [2] Monroe M: Molecular Weight Calculator. MWTWIN v6.35 <http://www.alchemistmatt.com>. Acedido em 19-01-2012
- [3] T. Kind, O. Fiehn, *BMC Bioinformatics* 7 (2006) 234
- [4] T. Kind, O. Fiehn, *BMC Bioinformatics* 8 (2007) 105
- [5] D.G. Schmid, P. Grosche, H. Bandel, G. Jung, *Biotechnology and Bioengineering* 71 (2000) 149

ATUALIDADE CIENTÍFICA

TESTANDO A PUREZA DE ENANTIÓMEROS

Moléculas que são imagens especulares entre si podem originar efeitos biológicos claramente distintos. Por exemplo, um enantiómero específico pode aliviar sintomas de uma determinada doença, enquanto que o correspondente à sua imagem especular pode ser completamente inativo ou mesmo desencadear efeitos secundários nefastos. Deste modo, os químicos de síntese desenvolvem métodos para produzir seletivamente apenas o enantiómero que lhes interessa. No entanto, estes investigadores podem agora avaliar mais facilmente a eficácia dos seus processos através do recurso a uma nova técnica que possibilita a quantificação rápida da proporção relativa de cada enantiómero presente numa mistura racémica (J. Am. Chem. Soc., DOI: 10.1021/ja205775g).

Para estimar a composição de uma mistura racémica, os especialistas recorrem ao cálculo do excesso enantiomérico. Por exemplo, se uma amostra contém 80% de um enantiómero e, consequentemente, 20% do outro, o excesso enantiomérico é de 60%. Para avaliar com rigor este parâmetro, os químicos utilizam técnicas como cromatografia líquida (HPLC). No entanto, “estes métodos são lentos e não possibilitam uma análise em tempo útil das centenas ou milhares de reações que a indústria farmacêutica ou os investigadores académicos necessitam avaliar durante o desenvolvimento de novos métodos de síntese”, afirma Eric Anslyn, um químico da Universidade do Texas, em Austin.

Anslyn e os seus colaboradores pretendiam desenvolver um meio expedito para a quantificação do excesso enantiomérico de ácidos carboxílicos, que são componentes comuns em medicamentos e polímeros. Em colaboração com James Canary, da Universidade de Nova Iorque, a equipa sintetizou uma molécula aquiral contendo cobre. O átomo de cobre possui um centro de coordenação desocupado e, deste modo, a molécula pode acomodar ácidos carboxílicos quirais. Quando este substrato quiral se liga ao átomo de cobre, os ligandos de quinolona organizam-se em forma de hélice. Esta hélice gira num sentido específico consoante a estereoquímica do substrato.

Quando os investigadores analisaram os complexos através de uma técnica designada por “exciton-coupled circular dichroism”, verificaram que os dois movimentos da hélice absorviam luz polarizada de uma forma diferente. Tal como os enantiómeros, os espectros de cada molécula representavam imagens especulares entre si, produzindo sinais simétricos a um comprimento de onda característico. Assim, os investigadores recolheram espectros de misturas conhecidas de enantiómeros de forma a construírem uma curva de calibração, que posteriormente foi usada para estimar o excesso enantiomérico de um conjunto separado de misturas racémicas.

Através da comparação dos valores de excesso enantiomérico reais de cada mistura com os valores estimados a partir da curva de calibração, os investigadores concluíram que a técnica apresenta erros médios de aproximadamente 3%. Assim, Anslyn afirma “a nossa abordagem não é tão exata como HPLC, mas é 100 vezes mais rápida”. De facto, a equipa pode analisar aproximadamente 100 amostras em duas horas.

(adaptado do artigo de 01/08/2011 de Laura Cassidy: Probing Enantiomer Purity
Chemical & Engineering News –
<http://pubs.acs.org/cen/news/89/i31/8931scene4.html>)

Paulo Brito
(paulo@ipb.pt)
Instituto Politécnico de Bragança

..... Visite-nos em www.spq.pt