

Ressonância magnética nuclear de proteínas

Luciana Coutinho de Oliveira,¹ Layara Akemi Abiko,² Roberto Kopke Salinas¹

¹Departamento de Bioquímica; ²Departamento de Química Fundamental

Instituto de Química da Universidade de São Paulo

Introdução

Ressonância Magnética Nuclear (RMN) é a única técnica espectroscópica capaz de resolver a estrutura de macromoléculas biológicas em solução e com resolução atômica. RMN também fornece informações detalhadas acerca dos movimentos internos de macromoléculas (dinâmica) em diferentes escalas de tempo, e é extremamente poderosa para o estudo de aspectos termodinâmicos, cinéticos e conformacionais de interações entre proteínas ou entre proteínas e outras moléculas como ácidos nucleicos, açúcares, ou metabólitos.¹ Estas informações são complementares às informações estruturais obtidas por cristalografia de proteínas.

Este capítulo contém uma breve introdução à RMN aplicada ao estudo estrutural de proteínas em solução. O texto inicia-se com uma descrição dos aspectos físicos do fenômeno de RMN, o deslocamento químico e o efeito de relaxação. Estes conceitos facilitam a compreensão de porque RMN fornece informações estruturais tão detalhadas e permitem compreender quais são os limites da técnica. A segunda parte do texto refere-se a aspectos práticos. Quais são os pré-requisitos para obter uma boa amostra de proteína para RMN? Quais são os principais experimentos utilizados para atribuir as ressonâncias, e como o cálculo estrutural é realizado? O capítulo termina com uma pequena introdução ao uso de RMN para o estudo de dinâmica e interações envolvendo proteínas.

O spin nuclear e o fenômeno de RMN

O spin, assim como a massa e a carga elétrica, são propriedades intrínsecas de partículas como o núcleo atômico e o elétron. A analogia clássica para o spin é o momento angular, uma grandeza apresentada por objetos em rotação, a qual pode ser visualizada como um vetor orientado ao longo do eixo ao redor do qual o objeto gira. Diferente do momento angular macroscópico, o momento angular de spin não é produzido pela rotação da partícula, mas trata-se de uma propriedade intrínseca, de natureza quântica. A existência de carga elétrica e de um momento angular de spin diferente de zero faz com que o núcleo atômico possua momento magnético, μ , e comporte-se de forma análoga a um pequeno dipolo magnético.² Dipolos magnéticos, quando colocados na presença de um campo magnético externo, acoplam com o campo. A energia de interação entre um dipolo magnético e o campo externo é dada pelo produto escalar entre μ e \mathbf{B} :

$$E = -\mu \cdot \mathbf{B}. \quad (1)$$

O momento angular de spin nuclear assume valores discretos de acordo com o número quântico I , o qual determina quantos estados energéticos serão gerados pela interação com um campo magnético externo (\mathbf{B}). Dessa forma, núcleos que apresentam número quântico de spin $I = 1/2$ (por exemplo ^{13}C , ^{15}N , ^1H), assumem dois estados de energia na presença de \mathbf{B} . Estes dois

estados são conhecidos como α ($m=+1/2$) e β ($m=-1/2$), cujas energias dependem da orientação do momento magnético em relação a **B** (**Figura 1A**):

$$E_{\alpha} = -\frac{1}{2} \hbar \gamma B,$$

e

$$E_{\beta} = +\frac{1}{2} \hbar \gamma B,$$

onde γ é a razão giromagnética, uma grandeza fundamental em RMN, e \hbar é a constante de Planck. Núcleos que não apresentam spin ($I = 0$) como o ^{12}C , não são ativos na RMN, enquanto que núcleos com número quântico de spin maior que $1/2$, como por exemplo o isótopo mais abundante do nitrogênio (^{14}N , $I = 1$), apresentam mais de dois estados de energia e, conseqüentemente, espectros de RMN mais complexos.

A presença do campo magnético externo, **B**, induz a formação de um pequeno excesso de spins no estado de menor energia, o estado α . O grau de polarização é proporcional à diferença de energia entre os estados α e β . Esta diferença é, por sua vez, diretamente proporcional à razão giromagnética (γ) e à magnitude do campo externo (**B**) (**Figura 1B**), e corresponde à frequência de ressonância do núcleo, ω_0 :

$$\begin{aligned} \Delta E &= E_{\beta} - E_{\alpha} = \hbar \gamma B \\ \frac{\Delta E}{\hbar} &= \gamma B = -\omega_0 \end{aligned} \quad (3)$$

Estando a amostra polarizada, o método de RMN baseia-se na aplicação de uma radiação eletromagnética com frequência equivalente a ω_0 , de forma a promover transições entre os diferentes estados de energia do spin nuclear na presença de **B**. Mas é importante notar que a energia de acoplamento entre o momento magnético nuclear e o campo externo é muito pequena quando comparada à energia térmica. Por exemplo, a diferença de energia magnética entre as duas orientações do momento magnético do núcleo de hidrogênio em um campo de 11,74 T (equivalente a um instrumento de 500 MHz) é $3,3 \times 10^{-25}$ J, enquanto que a energia térmica na temperatura ambiente é $k_B T = 4,1 \times 10^{-21}$ J, em que k_B é a constante de Boltzmann. Esta comparação demonstra que a polarização dos spins nucleares é um efeito de magnitude muito pequena e por isso difícil de detectar. Por este motivo, a técnica de RMN é pouco sensível quando comparada a outras técnicas espectroscópicas, e exige grande quantidade de amostra (concentrações da ordem de 0,2 a 1 mM de proteínas). A baixa sensibilidade da RMN, aliado ao interesse em realizar estudos estruturais de amostras biológicas que quase sempre são obtidas em baixas concentrações, é um dos principais estímulos ao desenvolvimento de magnetos supercondutores cada vez mais potentes. O maior campo magnético obtido em instrumentos comerciais atualmente é 1000 MHz (23,5 T).

Características de um espectro de RMN

Será que todos os spins nucleares em uma molécula irão apresentar exatamente a mesma frequência de ressonância? Os spins sentem campos magnéticos locais ligeiramente diferentes de **B**. Este efeito é devido à conformação e à estrutura química da molécula que modificam a distribuição eletrônica local. As nuvens eletrônicas locais são influenciadas pelo campo externo levando ao surgimento de campos magnéticos induzidos que podem ser orientados na direção oposta ou na mesma direção do campo externo.

Dessa forma os diferentes spins nucleares em uma molécula apresentarão frequências de ressonância ligeiramente diferentes de acordo com a conformação e a estrutura química. Estas pequenas diferenças são extraordinariamente informativas sobre a estrutura molecular, conformação e estereoquímica.

A fim de comparar medidas realizadas em campos magnéticos diferentes, as frequências de ressonância são geralmente expressas na forma de deslocamentos químicos em relação a uma frequência de referência. O deslocamento químico é expresso em unidades de *partes por milhão*, ou ppm:

$$\delta = \frac{\omega_0 - \omega_{rf}}{\omega_{rf}} \times 10^6, \quad (4)$$

onde ω_0 é a frequência de ressonância do spin nuclear em questão e ω_{rf} é a frequência de referência.

O aspecto de um espectro de RMN é influenciado não apenas pela separação dos sinais devido aos deslocamentos químicos, mas também pelos desdobramentos provocados pelo acoplamento escalar, e pelas propriedades de relaxação dos spins em questão. Estas questões são discutidas a seguir.

O acoplamento escalar (J) é um efeito de origem quântica que surge devido a uma interação entre dois spins, I e S, que estão a poucas ligações químicas de distância um do outro.³ Dessa forma, a transição de energia do spin I será desdobrada em duas, que dependem do fato de o spin S encontrar-se no estado α ou β . O mesmo desdobramento é observado para as transições do spin S (**Figura 1C**). A diferença de frequência entre as duas linhas é chamada de *constante de acoplamento escalar*, J. Esta constante permite inferir a conectividade entre núcleos via ligações químicas. A magnitude de J é independente do campo magnético externo, mas depende da natureza dos núcleos envolvidos, do número de ligações que os separam, além do ângulo entre estas ligações. A presença deste acoplamento dá origem a multipletos no espectro, ou seja, a desdobramentos de linhas dependentes de J.

O aspecto do espectro de RMN também é influenciado pela relaxação dos spins em questão. A detecção do sinal de RMN envolve excitação das transições do spin nuclear por pulsos de radiofrequência. O sinal detectado corresponde à energia emitida pelo sistema ao retornar ao estado de equilíbrio. O processo de retorno ao equilíbrio é denominado de *relaxação*, e é caracterizado por duas constantes de tempo, T_1 (*tempo de relaxação da magnetização longitudinal*) e T_2 (*tempo de relaxação da magnetização transversal*). Estes dois parâmetros são extremamente importantes em RMN, e dependem do tamanho da molécula na forma do *tempo de correlação rotacional* (τ_c) (**Figura 1D**). Mais importante, a largura dos sinais de RMN é diretamente proporcional a $1/T_2$. Moléculas grandes possuem T_2 muito curto e, portanto, linhas alargadas até o limite em que o sinal não pode mais ser detectado.³ Este fenômeno explica porque é difícil estudar proteínas grandes por RMN em solução.

Os tempos de relaxação (T_1 e T_2) dependem não apenas do movimento global da molécula mas também da dinâmica interna, ou seja, dos movimentos internos que a molécula experimenta em diferentes escalas de tempo. Dessa forma medidas de parâmetros de relaxação de RMN podem ser utilizadas para obter informações sobre a dinâmica interna de proteínas.^{2,3} Este assunto será discutido em detalhes no tópico “dinâmica de proteínas”.

Preparação e adequação da amostra

Proteínas de até 20-30 kDa são geralmente adequadas para serem estudadas por RMN em solução. O estudo de proteínas maiores é mais difícil pois o tempo de relaxação T_2 torna-se muito curto, o que limita a sensibilidade dos experimentos. Aliado a isto, o maior número de sinais aumenta a complexidade e sobreposição espectral.

Os métodos modernos de RMN heteronuclear valem-se da manipulação simultânea dos spins de ^1H , ^{15}N e ^{13}C .^{2,3} Devido à baixa abundância natural é necessário enriquecer a amostra com ^{15}N e ^{13}C . Isto é normalmente realizado através da produção da proteína de interesse de forma recombinante na presença de fontes de nitrogênio e carbono enriquecidas com ^{15}N e ^{13}C . Para a produção da proteína recombinante enriquecida isotopicamente em bactérias utiliza-se, em geral, $^{15}\text{NH}_4\text{Cl}$ e ^{13}C -glicose em meio mínimo.

A amostra de RMN consiste usualmente de 550 μL da proteína purificada, concentrada a 0,2 - 1 mM e enriquecida com ^{15}N ou com ^{15}N e ^{13}C de acordo com o experimento que se deseja realizar. A amostra é acondicionada em um tubo de vidro com largura de 5 mm. Uma grande variedade de sistemas tampão pode ser utilizada. O tampão fosfato é interessante porque não possui hidrogênios cujos sinais poderiam se sobrepor aos sinais da proteína, entretanto, outros tampões também podem ser utilizados no caso de experimentos heteronucleares. É desejável que o pH da amostra não seja muito maior do que 7, pois, neste caso, a velocidade de troca dos hidrogênios HNs com o solvente aumenta, levando ao alargamento de linha. Sais absorvem radiação eletromagnética, o que degrada a sensibilidade dos experimentos. Desta forma,

concentrações de sais muito elevadas (maiores que 0,2 M) devem ser evitadas. Uma forma eficiente de contornar este problema é o uso de tubos de RMN com diâmetro menor (3mm) ou com formato achatado, porque nestes casos a quantidade total de sais na frente da bobina de detecção é diminuída.

O estudo de moléculas maiores que 30 kDa é facilitado pelo emprego de enriquecimento isotópico com ^2H . Neste caso a contribuição da interação dipolar (quando o campo magnético de um spin afeta o campo magnético local de outro spin) para a relaxação dos spins de ^{13}C é minimizada, uma vez que o ^2H possui menor razão giromagnética do que ^1H ($\gamma_{\text{D}}/\gamma_{\text{H}} \sim 1/6,5$).² Portanto, o enriquecimento com ^2H contribui para diminuir a largura de linha dos sinais e aumentar a sensibilidade dos experimentos que envolvem spins de ^{13}C e ^{15}N . Proteínas podem ser enriquecidas em até 50-70 % de ^2H ao serem produzidas em bactérias crescidas em meio mínimo em $^2\text{H}_2\text{O}$. A perdeuteração da proteína (substituição de mais de 95% dos prótons não lábeis por deutério) pode ser atingida utilizando-se além de $^2\text{H}_2\text{O}$, glicose uniformemente enriquecida com deutério como fonte de carbono.

O experimento de ^{15}N -HSQC

O espectro unidimensional de ^1H de uma proteína apresenta sinais provenientes de todos os núcleos de ^1H da molécula (**Figura 1E**). Hidrogênios da ligação peptídica, de anéis aromáticos, ou alifáticos apresentam frequências de ressonância em regiões bastante típicas, no entanto, a grande sobreposição de sinais dificulta a aplicação de espectros unidimensionais para o estudo conformacional de proteínas.

Uma forma de resolver o problema da sobreposição espectral é a aquisição de experimentos bidimensionais. O experimento ^{15}N -HSQC (*heteronuclear single quantum correlation*) resulta em um espectro bidimensional que contém os sinais dos hidrogênios HNs separados em uma segunda dimensão correspondente à frequência de ^{15}N (**Figura 1F**). Assim, cada pico no espectro de ^{15}N -HSQC possui duas frequências, a do hidrogênio ^1HN e a frequência do núcleo de ^{15}N ao qual ele está diretamente ligado. Portanto, o espectro de ^{15}N -HSQC apresentará um pico para cada grupo HN presente na proteína, ou seja um pico para cada aminoácido com exceção da prolina, e um pico para cada grupo NH presente nas cadeias laterais de Asp, Asn, Arg, Lys, His e Trp.

O espectro de ^{15}N -HSQC é a impressão digital de uma proteína. A qualidade do espectro permite avaliar a adequação da amostra para estudos estruturais por RMN em solução. Um dos pré-requisitos para uma boa amostra é possuir picos bastante dispersos (**Figura 1F**). A boa dispersão dos picos ao longo da dimensão do hidrogênio (^1H) é a primeira indicação de que a proteína está bem enovelada. Proteínas desenoveladas apresentam picos de correlação ^1H - ^{15}N agrupados em torno de 8 ppm na dimensão de ^1H (**Figura 1G**). O segundo pré-requisito para uma boa amostra é que todos os picos esperados devem ser observados no espectro de ^{15}N -HSQC. A

ausência de picos para alguns aminoácidos, em geral, indica a presença de equilíbrio conformacional em regime de troca intermediária na escala de tempo de RMN. Neste caso, parâmetros como temperatura, pH e condições do tampão podem ser variados com o objetivo de estabilizar a estrutura e observar o número de picos esperado no espectro.

Sinais mais estreitos também podem ser obtidos explorando-se o efeito TROSY (*Transverse Relaxation-Optimized Spectroscopy*). O efeito TROSY origina-se a partir da observação de que as diferentes componentes do multipletto devido ao acoplamento escalar 1J entre ^1H e ^{15}N possuem larguras de linha diferentes devido a interações construtivas ou destrutivas entre diferentes mecanismos de relaxação. Estes mecanismos de relaxação, embora geralmente correlacionados, contribuem de maneira desigual para as diferentes componentes do multipletto ^1H - ^{15}N , fazendo com que estas relaxem com velocidades bastante diferentes. O experimento TROSY consiste na seleção da componente do multipletto com menor velocidade de relaxação, correspondendo, portanto, à linha mais estreita. Este procedimento aumenta significativamente a resolução e a sensibilidade dos espectros de RMN em altos campos magnéticos (> 14 Tesla), possibilitando o estudo de proteínas maiores que 30 KDa.⁴

Assinalamento das ressonâncias da cadeia principal

A informação mais importante para a caracterização estrutural de uma proteína por RMN é a identificação dos picos no espectro de ^{15}N -HSQC. A atribuição das ressonâncias dos átomos da cadeia principal costuma ser realizada através da análise de espectros multidimensionais heteronucleares. Os experimentos principais são HNCO e HN(CA)CO, HNCA e HN(CO)CA, HNCACB e CBCA(CO)NH tridimensionais, cujas dimensões correspondem às frequências de ^1H , ^{13}C e ^{15}N . A aquisição destes espectros requer uma amostra duplamente enriquecida com ^{15}N e ^{13}C . Ao contrário do experimento de ^{15}N HSQC, cuja aquisição leva tipicamente 5 - 40 minutos, a aquisição de experimentos tridimensionais pode levar de 1 a 2 dias. A aquisição de um conjunto de dados completo pode levar até 10 dias. Durante este intervalo de tempo a amostra deve permanecer estável no espectrômetro. Esta condição é difícil de ser satisfeita para muitas proteínas, especialmente levando-se em conta que os espectros são adquiridos a temperaturas quase sempre maiores que 25°C . O aparecimento de picos espúrios no espectro ^{15}N -HSQC pode indicar degradação, enquanto que o alargamento das linhas com concomitante perda de intensidade é sinal de agregação. A aquisição de um espectro de ^{15}N -HSQC antes e após a aquisição do conjunto de dados tridimensionais é fundamental para aferir a estabilidade da amostra. Atualmente, os métodos que permitem a reconstrução de espectros multidimensionais adquiridos na forma de amostragem não linear das dimensões indiretas têm se tornando cada vez mais robustos. A aplicação destes novos métodos de reconstrução permite diminuir significativamente o tempo de aquisição de espectros multidimensionais.

A nomenclatura adotada para os experimentos de tripla ressonância indica os spins correspondentes a cada uma das dimensões de frequência do espectro, em que CO indica carbonilas, CA, carbonos α e, CB, carbonos β . O HNCO, por exemplo, começa com a excitação de ^1H , em seguida, a polarização é transferida para o ^{15}N via acoplamento escalar ($^1J_{\text{HN}}$), e posteriormente para o spin ^{13}CO do aminoácido anterior via acoplamento escalar ($^1J_{\text{NCO}}$) (**Figura 1H**). O resultado é um espectro tridimensional, cujas dimensões correspondem aos deslocamentos químicos de ^1HN e ^{15}N de um aminoácido e ^{13}CO do aminoácido vizinho. Os spins listados entre parênteses não correspondem a uma dimensão de frequência, mas são utilizados para transferir polarização entre diferentes spins. Por exemplo, o espectro de HN(CA)CO contém correlações entre os spins ^1HN , ^{15}N de um aminoácido e ^{13}CO dele mesmo e do aminoácido vizinho. Análise do HN(CA)CO em conjunto com o HNCO permite identificar as ressonâncias do ^{13}CO de cada aminoácido e do seu vizinho anterior (**Figura 1H**). Dessa forma é possível encontrar correlações sequenciais ao longo de toda a cadeia polipeptídica até a presença de uma prolina.

Ainda que sequências de picos ^{15}N -HSQC correspondentes a aminoácidos adjacentes tenham sido identificadas pelo procedimento descrito acima, é difícil conectar esta informação à sequência de aminoácidos da proteína. Esta conexão pode ser estabelecida conhecendo-se o tipo de aminoácido para cada pico no espectro ^{15}N -HSQC. Deslocamentos químicos de $^{13}\text{C}\alpha$ e $^{13}\text{C}\beta$ são muito úteis para determinar o tipo do aminoácido. Esta informação pode ser obtida por meio da análise conjunta dos experimentos CBCA(CO)NH e HNCACB, os quais permitem conectar os assinalamentos existentes para a cadeia principal com as ressonâncias do $^{13}\text{C}\beta$ da cadeia lateral de um aminoácido e seu vizinho.

Deslocamentos químicos de ^1HN , ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ e ^{13}CO são extremamente informativos sobre a estrutura secundária da proteína. Existe uma forte correlação entre o deslocamento químico secundário, expresso como o desvio do deslocamento químico observado em relação àquele típico de *random coil*, e os ângulos diedrais Φ e ψ que caracterizam a conformação do esqueleto polipeptídico. Deslocamentos químicos secundários positivos para $^{13}\text{C}\alpha$ e negativos para $^{13}\text{C}\beta$ correlacionam com uma alfa-hélice, enquanto que o comportamento inverso é observado para fitas beta.⁵ Excelentes previsões para os ângulos diedrais do esqueleto polipeptídico são obtidas a partir de dados de deslocamentos químicos secundários em combinação com informações de homologia de sequência de aminoácidos e buscas em bancos de dados de estruturas cristalográficas pelo programa TALOS.⁶

Assinalamento das ressonâncias das cadeias laterais

Assinalamentos das ressonâncias dos átomos das cadeias laterais são necessários para calcular a estrutura de proteínas a partir de dados de RMN. Uma das formas de se obter estes assinalamentos é através da análise de experimentos do tipo TOCSY (*total correlation*

spectroscopy). O TOCSY é um espectro bidimensional que permite identificar todos os hidrogênios correlacionados entre si por acoplamento escalar, e dessa forma identificar os sistemas de spins completos de cada aminoácido. Em proteínas maiores que aproximadamente 10 kDa, a obtenção de correlações via ^1H - ^1H TOCSY é pouco eficiente. Nestes casos, a melhor forma de obter correlações é realizando o TOCSY em ^{13}C conforme implementado nos experimentos tridimensionais CCH-TOCSY e HCCH-TOCSY. Nestes experimentos, a polarização dos hidrogênios das cadeias laterais é transferida para o núcleo de ^{13}C ligado covalentemente a cada hidrogênio. Esta magnetização é então transferida para outros núcleos de ^{13}C pertencentes ao mesmo sistema de spins via TOCSY. O espectro resultante será um espectro tridimensional em que se observam correlações entre os spins de ^{13}C (CCH-TOCSY) ou de ^1H (HCCH-TOCSY) de uma mesma cadeia lateral.

Informações estruturais

A principal fonte de informação estrutural em RMN é o NOE (efeito nuclear Overhauser). O NOE é um efeito de relaxação cruzada, envolvendo a transferência de polarização de um núcleo de hidrogênio para outro via interação dipolar. O efeito NOE é de curto alcance e depende da distância entre os dois núcleos elevada à sexta potência ($1/r^6$). Dessa forma, a intensidade do NOE pode ser utilizada para calcular a distância entre os dois núcleos de hidrogênio desde que eles estejam a até 5Å de distância um do outro.

Distâncias obtidas a partir da análise de NOEs são informações geométricas sobre a estrutura da proteína. Hidrogênios que estão próximos no espaço podem ser identificados em experimentos 2DNOE, ou ainda experimentos tridimensionais em que a terceira dimensão corresponde à frequência de ^{15}N ou ^{13}C . Uma proteína apresenta geralmente 10 a 20 NOEs por aminoácido, sendo que a maior parte destes NOEs ocorre entre núcleos de hidrogênio que pertencem ao mesmo aminoácido ou aminoácidos adjacentes. Hidrogênios localizados em aminoácidos que estão longe um do outro na sequência primária mas próximos na estrutura tridimensional também são identificados através do efeito NOE. NOEs de curto e médio alcance são informativos sobre a estrutura secundária da proteína enquanto que os de longo alcance fornecem restrições que permitem elucidar a estrutura tridimensional da molécula.

Outra fonte de informação estrutural obtida a partir da análise de dados de RMN são ângulos diedrais, em particular o ângulo Φ do esqueleto polipeptídico. Este ângulo pode ser obtido através de medidas da constante de acoplamento escalar entre os hidrogênios amídicos (^1HN) e $\text{H}\alpha$ ($^3J_{\text{HNH}\alpha}$). Predições para Φ e ψ do esqueleto polipeptídico podem ser obtidas a partir da análise de deslocamentos químicos, em particular $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ e ^{13}CO , com auxílio do programa TALOS.⁶ Estas predições são frequentemente utilizadas no cálculo estrutural.

Cálculo da estrutura

O cálculo de uma estrutura de proteínas a partir de dados de RMN corresponde à busca de conformações consistentes com todos os dados experimentais fornecidos, em particular NOEs e ângulos diedrais.

O primeiro desafio é realizar a atribuição dos NOEs. Em princípio este problema é trivial, desde que todos os deslocamentos químicos sejam conhecidos. No entanto, a maior parte dos NOEs são ambíguos, isto é, apresentam mais de uma possibilidade de assinalamento. O assinalamento correto só é identificado conhecendo-se a estrutura tridimensional da molécula. Por isso, a estrutura de RMN e o assinalamento de NOEs são realizados de forma iterativa. Inicialmente, um subconjunto de NOEs é assinalado e utilizado para calcular uma estrutura de baixa resolução. Em seguida, este modelo inicial é utilizado para identificar um número maior de NOEs que serão utilizados para calcular um segundo modelo, e assim por diante. Este processo pode ser automatizado conforme implementado em diferentes programas entre os quais destacam-se XPLOR-NIH,⁷ CYANA⁸ e ARIA.⁹ Estes programas utilizam dinâmica molecular e *simulated annealing* como métodos de busca no espaço conformacional.

Os dados experimentais derivados de NOEs e ângulos diedrais são impostos como restrições através de potenciais de penalidade. Após o cálculo é importante refinar as estruturas em uma caixa de moléculas de água explícitas. Este refinamento costuma resolver problemas de estereoquímica e sobreposição de átomos que podem aparecer devido ao uso de campos de força simplificados. O refinamento em água pode ser realizado com o protocolo RECOORD¹⁰ para CNS.¹¹

As estruturas de RMN são geralmente apresentadas como um conjunto dos 10-20 modelos de menor energia e menor número de violações das restrições experimentais (**Figura 1I**). A precisão da estrutura é estimada a partir do cálculo do desvio quadrático médio das coordenadas em relação à melhor estrutura do conjunto (RMSD). As regiões da estrutura que apresentam pior precisão revelam a ausência de um número suficiente de dados experimentais. Esta ausência pode ser provocada pela flexibilidade intrínseca da molécula uma vez que se não há estrutura, tampouco haverá NOEs.

Dinâmica de proteínas

Proteínas são flexíveis, ou seja, visitam diferentes conformações em diferentes escalas de tempo. Estes movimentos podem ser uma característica intrínseca da molécula e também podem ser essenciais para a sua função biológica. A compreensão mais completa da estrutura da proteína requer conhecimento da dinâmica interna.

A dinâmica interna de proteínas afeta diferentes observáveis de RMN. Medidas de parâmetros de relaxação de spins de ¹⁵N (T₁ e T₂, e o chamado NOE heteronuclear entre ¹H e ¹⁵N) são extremamente informativas sobre os movimentos internos do esqueleto polipeptídico em escala de tempo rápida (sub-nanosegundos) ou lenta (microsegundos a milissegundos). Um

exemplo qualitativo é encontrado na **Figura 1I**, em que a região N-terminal da proteína é altamente flexível e, conseqüentemente, desordenada. Medidas do NOE heteronuclear de ^{15}N são qualitativamente consistentes com alta mobilidade na região N-terminal da proteína (representada em alaranjado na **Figura 1I**).

Interações entre proteínas e ligantes

RMN pode ser utilizada para obter informações detalhadas sobre a interface de complexos formados por proteínas e outros ligantes. A abordagem mais simples é observar alterações no espectro de ^{15}N -HSQC de uma proteína induzidas a partir da interação com um ligante específico. Uma vez que os assinalamentos dos picos no espectro de ^{15}N -HSQC da proteína isolada são conhecidos, é possível mapear aqueles aminoácidos envolvidos na interação com o ligante. Para isso é possível realizar um simples experimento de titulação. Neste experimento, alíquotas de uma solução do ligante não marcado são adicionadas ao tubo de RMN contendo a proteína marcada com ^{15}N , e após a adição de cada alíquota adquire-se um espectro de ^{15}N -HSQC.

Um complexo cuja afinidade não é muito alta, ou seja, com constante de dissociação (K_d) da ordem de dezenas de micromolar geralmente está em regime de troca rápida na escala de tempo de RMN. Isto significa que os picos dos aminoácidos localizados na região de interface do complexo caminham no espectro da posição da proteína livre para a posição da proteína ligada (**Figura 1J**). Este tipo de experimento permite identificar facilmente os aminoácidos que estão envolvidos na interação.

Uma outra situação será encontrada caso a afinidade entre proteína e ligante seja muito alta. Neste caso, a interação pode estar em regime de troca lenta na escala de tempo de RMN. Os picos da proteína livre desaparecem, enquanto que os picos da proteína ligada aparecem à medida que o ligante vai sendo adicionado. Neste caso não é possível identificar os aminoácidos envolvidos na interação através do experimento de titulação. É preciso adquirir experimentos de tripla ressonância e realizar o assinalamento completo das ressonâncias da cadeia principal da proteína em complexo.

Uma situação um pouco mais complicada é encontrada quando o complexo formado encontra-se em troca intermediária na escala de tempo de RMN. Neste caso as linhas espectrais apresentam-se alargadas, e podem, eventualmente, desaparecer. Neste caso, uma opção é realizar os experimentos em temperaturas diferentes, a fim de aumentar (temperaturas maiores) ou diminuir (temperaturas menores) a velocidade de troca para recuperação das linhas espectrais.

Bibliografia

1. Wüthrich, K. NMR studies of structure and function of biological macromolecules (Nobel Lecture). *J. Biomol. NMR* **27**, 13–39 (2003).
2. Cavanagh, J. *Protein NMR spectroscopy*. (Academic Press, 2007).
3. Rule, G. S. & Hitchens, T. K. *Fundamentals of Protein NMR Spectroscopy*. (Springer, 2006).
4. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12366–12371 (1997).
5. Spera, S. & Bax, A. Empirical correlation between protein backbone conformation and C.alpha. and C.beta. ¹³C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**, 5490–5492 (1991).
6. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223 (2009).
7. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson. San Diego Calif 1997* **160**, 65–73 (2003).
8. López-Méndez, B. & Güntert, P. Automated Protein Structure Determination from NMR Spectra. *J. Am. Chem. Soc.* **128**, 13112–13122 (2006).
9. Linge, J. P., Habeck, M., Rieping, W. & Nilges, M. ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **19**, 315–316 (2003).
10. Nederveen, A. J. *et al.* RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins* **59**, 662–672 (2005).
11. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).

Legenda para Figura 1

O momento magnético do núcleo pode orientar-se a favor (menor energia) ou contra (maior energia) o campo magnético externo **B (A)**; Sistema de energia de um spin isolado na presença de um campo magnético externo B **(B)**; Sistema de energia correspondente a dois spins acoplados por acoplamento escalar J (esquerda); a interação de acoplamento escalar resulta no desdobramento das linhas do espectro (direita) **(C)**; Tempos de relaxação do spin nuclear do próton amídico (^1HN), T_1 e T_2 , em função do tempo de correlação da molécula e assumindo um campo magnético externo de 800 MHz (frequência de ^1H) **(D)**; Espectro de RMN 1D de hidrogênio de ubiquitina a 800 MHz, 298K **(E)**; espectro de ^{15}N -HSQC da ubiquitina adquirido a 800 MHz, 298K **(F)**; Espectro de ^{15}N -HSQC de uma proteína intrinsecamente desordenada adquirido a 800 MHz **(G)**; Caminhos de transferência de polarização nos experimentos HNCO e HN(CA)CO; plano de um espectro HNCO e HN(CA)CO mostrando correlações entre os spins ^1HN , ^{15}N e ^{13}CO intra e inter-residuais **(H)**; Ensemble de 20 modelos de RMN representando a região N-terminal sem estrutura em alaranjado, o que é consistente com valores baixos de $\{^1\text{H}\}$ - ^{15}N NOE **(I)**; Experimento de titulação acompanhado por ^{15}N -HSQC mostrando um pico que se desloca devido à formação de um complexo em troca rápida na escala de tempo de RMN **(J)**.