

Estrutura Tridimensional de Proteínas

- 4.1 Visão geral sobre a estrutura das proteínas 115
- 4.2 Estrutura secundária das proteínas 119
- 4.3 Estruturas terciária e quaternária das proteínas 125
- 4.4 Desnaturação e enovelamento das proteínas 143

As proteínas são moléculas grandes. O esqueleto covalente de uma proteína clássica é formado por centenas de ligações simples. Como é possível a livre rotação entre várias dessas ligações, a proteína consegue, em princípio, assumir um número de conformações praticamente incontáveis. Entretanto, cada proteína tem uma função química e uma estrutura específica, sugerindo que cada uma delas tenha uma estrutura tridimensional única (Figura 4-1). Quão estável é essa estrutura, quais fatores guiam sua formação e o que a mantém unida? No final de 1920, várias proteínas foram cristalizadas, incluindo a hemoglobina (M_r 64.500) e a enzima urease (M_r 483.000). Como, geralmente, o arranjo ordenado das moléculas em um cristal pode ocorrer somente se as unidades moleculares forem idênticas, a descoberta de que várias proteínas poderiam ser cristalizadas era a evidência de que até mesmo proteínas muito grandes são entidades químicas separadas com estruturas únicas. Essa conclusão revolucionou o pensamento sobre as proteínas e suas funções, mas o conhecimento que ela gerou foi incompleto. A estrutura proteica é sempre ma-

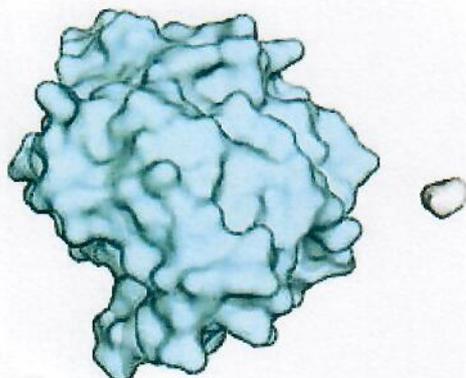


FIGURA 4-1 Estrutura da enzima quimotripsina, uma proteína globular. A molécula de glicina (em cinza) é representada para comparação de tamanho. As estruturas tridimensionais conhecidas das proteínas estão arquivadas no *Protein Data Bank* – PDB (ver Quadro 4-4). A imagem mostrada aqui foi elaborada utilizando os dados do PDB ID 6GCH.

leável, algumas vezes de forma surpreendente. Mudanças na estrutura podem ser tão importantes para a função da proteína quanto a estrutura por si só.

Neste capítulo, será examinada a estrutura das proteínas. São enfatizados seis temas. Primeiro, a estrutura tridimensional, ou seja, as estruturas de uma proteína são determinadas por sua sequência de aminoácidos. Segundo, a função de uma proteína típica depende de sua estrutura. Terceiro, a maior parte das proteínas isoladas existem em um ou em um pequeno número de formas estruturalmente estáveis. Quarto, as forças mais importantes de estabilização das estruturas específicas de uma dada proteína são as interações não covalentes. Quinto, dentre desse enorme número de estruturas únicas das proteínas, podem ser reconhecidos alguns padrões estruturais comuns, que ajudam a organizar o entendimento sobre a arquitetura das proteínas. Sexto, as estruturas proteicas não são estáticas. Todas as proteínas passam por mudanças na conformação variando desde sutis até bastante significativas. Partes de muitas proteínas possuem estruturas não discerníveis. Para algumas proteínas, a ausência de estrutura definida é fundamental para sua função.

4.1 Visão geral sobre a estrutura das proteínas

O arranjo espacial dos átomos em uma proteína ou qualquer parte da proteína é chamado de **conformação**. As conformações possíveis de uma proteína ou de qualquer segmento proteico incluem qualquer estado estrutural que ela possa assumir sem a quebra de suas ligações covalentes. Uma mudança conformacional pode ocorrer, por exemplo, pela rotação sobre as ligações simples. Das várias conformações teoricamente possíveis para uma proteína com centenas de ligações simples, uma ou (mais comumente) poucas predominam em condições biológicas. A necessidade de múltiplas conformações estáveis reflete as mudanças que devem ocorrer na proteína quando ela se liga a outras moléculas ou catalisa reações. As conformações que existem em determinadas condições são, normalmente, aquelas termodinamicamente mais estáveis – isto é, aquelas com energia livre de Gibbs (G) menores. Proteínas dobradas, em qualquer uma de suas conformações funcionais, são chamadas de proteínas **nativas**.

Para a grande maioria das proteínas, uma estrutura em particular ou um pequeno grupo de estruturas é crucial para a função. No entanto, em muitos casos, partes das proteínas carecem de estruturas perceptíveis. Esses seg-

mentos proteicos são intrinsecamente desordenados. Em alguns casos, proteínas inteiras são intrinsecamente desordenadas, e ainda assim funcionais.

Quais os princípios que determinam as conformações mais estáveis de uma proteína típica? Uma compreensão da conformação de proteínas pode ser construída, passo a passo, a partir da discussão sobre estrutura primária no Capítulo 3, passando pela consideração das estruturas secundárias, terciárias e quaternárias. A essa abordagem clássica, é preciso acrescentar a ênfase mais recente dada aos padrões comuns e classificáveis de enovelamento, variavelmente chamados de estruturas supersecundárias, enovelamentos ou motivos, que estabelecem um importante contexto organizacional para esse esforço complexo. A título de introdução, serão apresentados alguns princípios básicos.

A conformação de uma proteína é estabilizada por interações fracas

No contexto da estrutura de proteínas, o termo **estabilidade** pode ser definido como a tendência em manter a conformação nativa. Proteínas nativas são apenas marginalmente estáveis: o ΔG que separa os estados dobrados e não dobrados em proteínas comuns, sob condições fisiológicas, está na faixa de apenas 20 a 65 kJ/mol. Uma dada cadeia polipeptídica pode, teoricamente, assumir inúmeras conformações e, como resultado, o estado não dobrado de uma proteína é caracterizado por um alto grau de entropia conformacional. Essa entropia, junto com as interações de ligações de hidrogênio dos diversos grupos da cadeia polipeptídica com o solvente (água), tendem a manter o estado não dobrado. As interações químicas que contrabalançam esses efeitos e estabilizam a conformação nativa incluem ligações dissulfeto (covalentes) e interações fracas (não covalentes), descritas no Capítulo 2: ligações de hidrogênio e interações hidrofóbicas e iônicas.

Várias proteínas não têm ligações dissulfeto. O ambiente dentro da maioria das células é altamente redutor devido à alta concentração de agentes redutores, como a glutatona, e a maior parte das sulfidrilas permanece então no estado reduzido. Fora da célula, o ambiente é frequentemente mais oxidante e a formação de dissulfeto é mais provável de ocorrer. Em eucariotos, as ligações dissulfeto são encontradas, principalmente, em proteínas secretadas, extracelulares (p. ex., o hormônio insulina). As ligações dissulfeto também são incomuns em proteínas de bactérias. Entretanto, bactérias termofílicas, assim como arqueobactérias, geralmente apresentam várias proteínas com ligações dissulfeto, que as estabilizam. Presumivelmente, isso é uma adaptação para a vida a altas temperaturas.

Para todas as proteínas de todos os organismos, as interações fracas são especialmente importantes para o enovelamento das cadeias polipeptídicas em suas estruturas secundárias e terciárias. A associação de múltiplos polipeptídeos para formar estruturas quaternárias também tem como base estas interações fracas.

Aproximadamente 200 a 460 kJ/mol são necessários para quebrar uma ligação covalente simples, enquanto interações fracas podem ser rompidas com apenas 0,4 a 30 kJ/mol. Individualmente, uma ligação covalente, como as

ligações dissulfeto conectando regiões distintas de uma única cadeia polipeptídica, é claramente muito mais forte que uma interação fraca. Entretanto, por serem muito numerosas, são as interações fracas que predominam como forças estabilizadoras da estrutura proteica. Em geral, a conformação proteica de energia livre mais baixa (i.e., de conformação mais estável) é aquela com o número máximo de interações fracas.

A estabilidade de uma proteína não é simplesmente o somatório das energias livres de formação das diversas interações fracas internas. Para cada ligação de hidrogênio formada em uma proteína durante seu enovelamento, uma ligação de hidrogênio (de força equivalente) entre o mesmo grupo e a água é quebrada. A estabilidade resultante da contribuição de uma dada ligação de hidrogênio, ou a *diferença* de energia livre entre os estados dobrado e não dobrado, deve ser próxima de zero. Interações iônicas podem ser tanto estabilizadoras, quanto desestabilizadoras. Portanto, é preciso olhar em outros lugares para entender por que uma determinada conformação nativa é favorável.

A partir de um exame cuidadoso da contribuição das interações fracas na estabilidade das proteínas, fica evidente que as **interações hidrofóbicas** geralmente predominam. A água pura contém moléculas de H_2O formando uma rede de ligações de hidrogênio. Nenhuma outra molécula tem o potencial de ligação de hidrogênio da água, e a presença de outras moléculas na solução aquosa rompe estas ligações de hidrogênio. Quando a água envolve uma molécula hidrofóbica, o arranjo ótimo de ligação de hidrogênio resulta em uma camada altamente estruturada, ou **camada de solvatação**, de água em torno da molécula (ver Figura 2-7). O aumento da ordem das moléculas de água na camada de solvatação está correlacionado com uma redução desfavorável na entropia da água. Entretanto, quando grupos apolares se agrupam, o tamanho da camada de solvatação diminui, porque cada grupo não mais expõe toda sua superfície à solução. O resultado é um aumento favorável de entropia. Como descrito no Capítulo 2, esse aumento de entropia é a principal força termodinâmica que rege a associação de grupos hidrofóbicos em solução aquosa. Cadeias laterais de aminoácidos hidrofóbicos tendem a se agrupar no interior das proteínas, longe da água (pense em uma gota de óleo na água). A sequência de aminoácidos da maioria das proteínas, assim, apresenta um conteúdo significativo de cadeias laterais de aminoácidos hidrofóbicos (especialmente Leu, Ile, Val, Phe e Trp). Posicionam-se de forma a se aglomerar quando a proteína é dobrada, formando um núcleo hidrofóbico da proteína.

Sob condições fisiológicas, a formação de ligações de hidrogênio em uma proteína é em grande parte dirigida pelo mesmo efeito entrópico. Grupos polares normalmente podem formar ligações de hidrogênio com a água, e, por isso, são solúveis em água. Entretanto, o número de ligações de hidrogênio por unidade de massa normalmente é maior para a água pura do que para qualquer outro líquido ou solução, e há limites de solubilidade até para as moléculas mais polares, pois sua presença causa uma diminuição no número total de ligações de hidrogênio por unidade de massa. Portanto, uma camada de solvatação, até certo ponto, também se forma em torno de moléculas polares. Apesar

de a energia de formação de uma ligação de hidrogênio intramolecular entre dois grupos polares em uma macromolécula ser em grande parte anulada pela eliminação de tais interações entre esses grupos polares e a água, a liberação da água estruturada, na forma de interações intramoleculares, garante a força entrópica que leva ao enovelamento. A maior parte da variação de energia livre na forma de interações fracas dentro da proteína é assim originária do aumento de entropia na solução aquosa circundante, resultante do confinamento das superfícies hidrofóbicas. Isso mais do que contrabalança a grande perda de entropia conformacional, pois o polipeptídeo é limitado à sua conformação dobrada.

As interações hidrofóbicas são importantes na estabilização da conformação: o interior de uma proteína geralmente é um núcleo altamente empacotado de cadeias laterais de aminoácidos hidrofóbicos. Também é importante que cada grupo polar ou carregado no interior da proteína tenha um par adequado para fazer ligação de hidrogênio ou interação iônica. Uma ligação de hidrogênio parece contribuir pouco para a estabilidade de uma estrutura nativa, mas a presença de grupos que fazem ligações de hidrogênio sem par no núcleo hidrofóbico de uma proteína pode ser tão *desestabilizadora* que conformações contendo esse grupo são termodinamicamente insustentáveis. A variação favorável de energia livre resultante da combinação de vários desses grupos com parceiros na solução que os circunda pode ser maior do que a diferença de energia livre entre os estados dobrados e não dobrados. Além disso, ligações de hidrogênio entre grupos em uma proteína se formam cooperativamente (a formação de uma torna mais provável a formação da próxima) em estruturas secundárias repetidas que otimizam as ligações de hidrogênio, como descrito a seguir. Dessa forma, as ligações de hidrogênio normalmente têm um importante papel na condução do processo de enovelamento de proteínas.

A interação entre grupos carregados com cargas opostas, que formam um par iônico ou uma ponte salina, pode exercer tanto um efeito estabilizante quanto desestabilizante na estrutura da proteína. Como no caso das ligações de hidrogênio, cadeias laterais de aminoácidos carregados interagem com a água e com sais quando a proteína não está dobrada, e a perda dessas interações deve ser considerada quando se avalia o efeito da ponte salina na estabilidade geral de uma proteína dobrada. Entretanto, a força de uma ponte salina aumenta à medida que se desloca para um ambiente com constante dielétrica mais baixa, ϵ (p. 50): do solvente aquoso polar (ϵ próximo a 80) para o interior apolar da proteína (ϵ próximo a 4). Pontes salinas, especialmente aquelas parcial ou totalmente internas na proteína, podem, assim, proporcionar uma estabilização significativa da estrutura de uma proteína. Essa tendência explica o aumento da ocorrência de pontes salinas internas nas proteínas de organismos termofílicos. Interações iônicas também limitam a flexibilidade estrutural e conferem uma singularidade a uma determinada estrutura proteica que as interações hidrofóbicas não específicas não conseguem proporcionar.

No ambiente atômico altamente empacotado de uma proteína, mais um tipo de interação fraca pode ter um efeito significativo – as interações de van der Waals (p. 54).

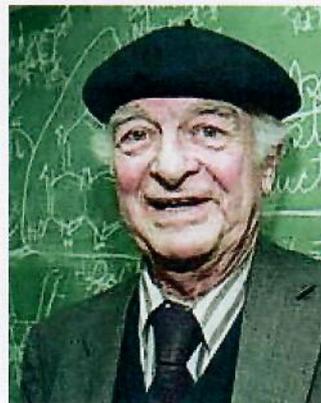
As interações de van der Waals são interações dipolo-dipolo envolvendo os dipolos elétricos permanentes de grupos, tal como as carbonilas, dipolos transitórios derivados das flutuações das nuvens de elétrons em torno de qualquer átomo, e dipolos induzidos pela interação de um átomo com outro que contém um dipolo permanente ou transitório. À medida que os átomos interagem um com o outro, essas interações dipolo-dipolo fornecem uma força intermolecular atrativa que opera apenas sobre uma distância intermolecular limite (0,3 a 0,6 nm). As interações de van der Waals são fracas e individualmente contribuem pouco para a estabilidade da proteína em geral. No entanto, em uma proteína bem empacotada ou na interação de uma proteína com outra proteína ou com outra molécula em uma superfície complementar, o número de tais interações pode ser substancial.

A maioria dos padrões estruturais resumidos neste capítulo reflete duas regras simples: (1) resíduos hidrofóbicos estão basicamente escondidos no interior da proteína, longe da água, e (2) o número de ligações de hidrogênio dentro da proteína é maximizado, reduzindo assim o número de grupos capazes de fazer ligações de hidrogênio e grupos iônicos que não estão adequadamente pareados. As proteínas de membrana (examinadas no Capítulo 11) e proteínas intrinsecamente desordenadas ou que têm segmentos intrinsecamente desordenados seguem regras diferentes. Isso reflete suas funções ou ambientes específicos, mas as interações fracas ainda são elementos estruturais importantes. Por exemplo, proteínas solúveis, mas com segmentos intrinsecamente desordenados, são ricas em cadeias laterais de aminoácidos carregados (especialmente Arg, Lys e Glu) ou pequenos (Gly e Ala), gerando pouca ou nenhuma oportunidade para formação do núcleo hidrofóbico estável.

A ligação peptídica é rígida e planar

Arquitetura das proteínas – Estrutura primária As ligações covalentes também impõem importantes restrições na conformação de um polipeptídeo. No final de 1930, Linus Pauling e Robert Corey iniciaram uma série de estudos que lançaram os fundamentos do entendimento atual sobre estrutura de proteínas. Eles começaram com uma cuidadosa análise da ligação peptídica.

Os carbonos α de resíduos adjacentes de aminoácidos são separados por três ligações covalentes, arranjados na



Linus Pauling, 1901-1994



Robert Corey, 1897-1971

forma $C_{\alpha}-C-N-C_{\alpha}$. Estudos de difração de raios X de cristais de aminoácidos e de dipeptídeos e tripeptídeos simples mostraram que a ligação peptídica $C-N$ é de alguma forma mais curta que a ligação $C-N$ de uma amina simples, e que os átomos associados à ligação peptídica são planares. Isso indicava a ressonância ou o compartilhamento parcial de dois pares de elétrons entre o oxigênio carbonílico e o nitrogênio da amida (Figura 4-2a). O oxigênio tem uma carga parcial negativa e o hidrogênio ligado ao nitrogênio tem uma carga líquida parcial positiva, formando um pequeno dipolo elétrico. Os seis átomos do **grupo peptídico** estão em um único plano, com o átomo de oxigênio do grupo carbonílico *trans* ao átomo de hidrogênio do nitrogênio da amida. A partir destas observações, Pauling e Corey concluíram que as ligações peptídicas $C-N$ não podem girar livremente, devido ao seu caráter parcial de ligação dupla. A rotação é permitida ao redor das ligações $N-C_{\alpha}$ e $C_{\alpha}-C$. O esqueleto de uma cadeia polipeptídica pode, então, ser descrito como uma série de planos rígidos, com planos consecutivos compartilhando um ponto comum de rotação no C_{α} (Figura 4-2b). As ligações peptídicas rígidas limitam a variação de conformações possíveis para uma cadeia polipeptídica.

A conformação da ligação peptídica é definida por três ângulos diedros (também conhecidos como ângulos de torção), chamados de ϕ (phi), ψ (psi) e ω (ômega), que refletem a rotação sobre cada uma das três ligações que se repetem no esqueleto peptídico. Um ângulo diedro é o ângulo da interseção de dois planos. No caso dos peptídeos, os planos são definidos pelos vetores das ligações do esqueleto peptídico. Dois vetores de ligações sucessivas descrevem um plano. Três vetores de ligações sucessivas descrevem dois planos (o vetor da ligação central é comum a ambos; Figura 4-2c), e o ângulo entre esses dois planos é medido para descrever a conformação da proteína.

CONVENÇÃO-CHAVE: Os ângulos diedros importantes para um peptídeo são definidos por três vetores das ligações que conectam quatro átomos consecutivos da cadeia principal (esqueleto peptídico) (Figura 4-2c): ϕ envolve as ligações $C-N-C_{\alpha}-C$ (com a rotação ocorrendo entre a ligação $N-C_{\alpha}$), e ψ envolve as ligações $N-C_{\alpha}-C-N$. Ambos, ϕ e ψ , são definidos como $\pm 180^\circ$ quando o polipeptídeo está completamente estendido e todos os grupos peptídicos estão no mesmo plano (Figura 4-2d). Quando se observa ao longo do vetor da ligação central, na direção da flecha do

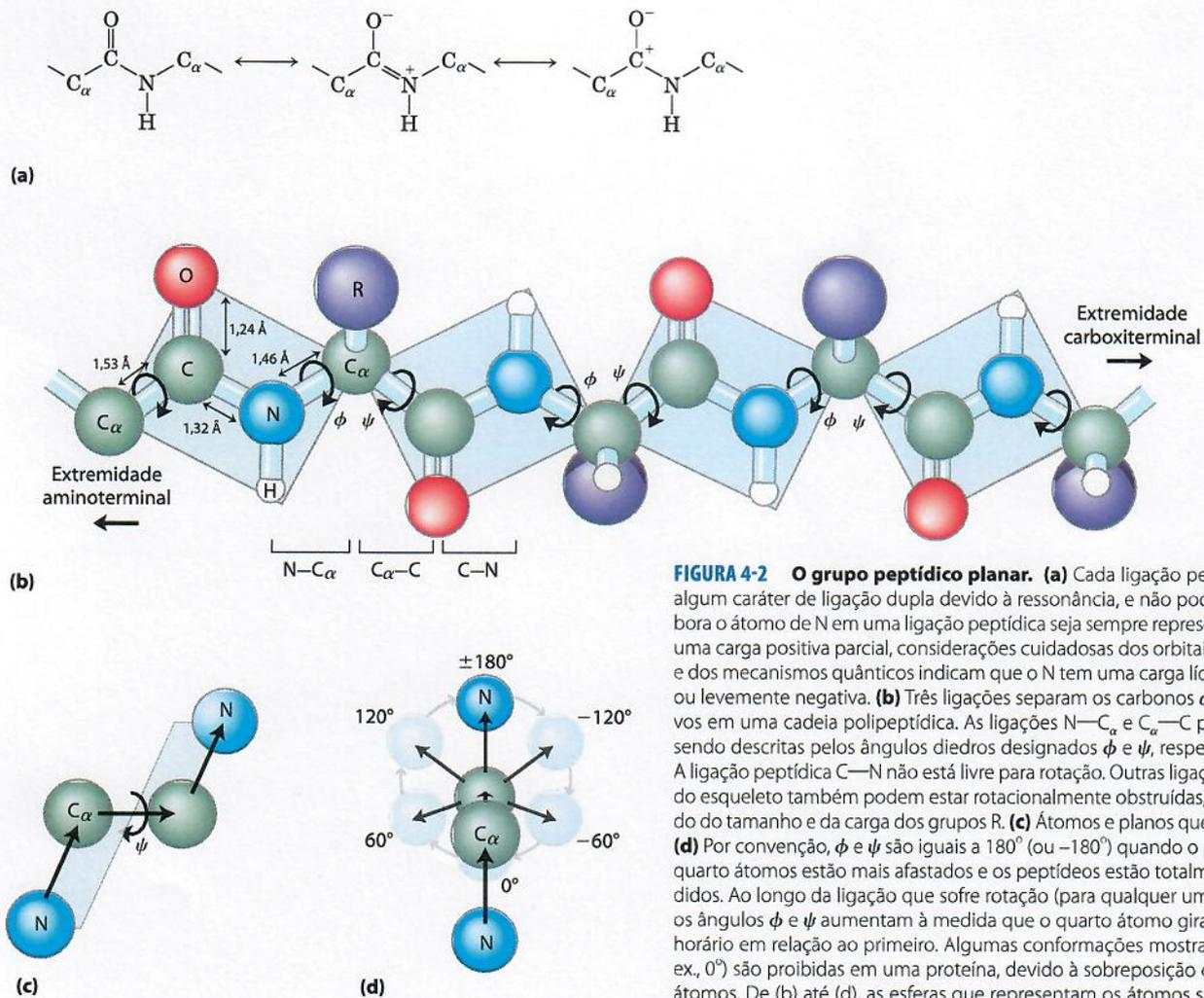


FIGURA 4-2 O grupo peptídico planar. (a) Cada ligação peptídica tem algum caráter de ligação dupla devido à ressonância, e não pode girar. Embora o átomo de N em uma ligação peptídica seja sempre representado com uma carga positiva parcial, considerações cuidadosas dos orbitais de ligação e dos mecanismos quânticos indicam que o N tem uma carga líquida neutra ou levemente negativa. (b) Três ligações separam os carbonos α consecutivos em uma cadeia polipeptídica. As ligações $N-C_{\alpha}$ e $C_{\alpha}-C$ podem girar, sendo descritas pelos ângulos diedros designados ϕ e ψ , respectivamente. A ligação peptídica $C-N$ não está livre para rotação. Outras ligações simples do esqueleto também podem estar rotacionalmente obstruídas, dependendo do tamanho e da carga dos grupos R. (c) Átomos e planos que definem ψ . (d) Por convenção, ϕ e ψ são iguais a 180° (ou -180°) quando o primeiro e o quarto átomos estão mais afastados e os peptídeos estão totalmente estendidos. Ao longo da ligação que sofre rotação (para qualquer um dos lados), os ângulos ϕ e ψ aumentam à medida que o quarto átomo gira no sentido horário em relação ao primeiro. Algumas conformações mostradas aqui (p. ex., 0°) são proibidas em uma proteína, devido à sobreposição espacial dos átomos. De (b) até (d), as esferas que representam os átomos são menores do que os raios de van der Waals para esta escala.

vetor (como mostrado na Figura 4-2c para ψ), os ângulos diedros aumentam à medida que o átomo distal (quarto átomo) gira no sentido horário (Figura 4-2d). A partir das posições $\pm 180^\circ$, o ângulo diedro aumenta de -180° para 0° , o ponto no qual o primeiro e o quarto átomos estão eclipsados. A rotação pode continuar de 0° a $+180^\circ$ (mesma posição que -180°) para retornar a estrutura ao ponto de partida. O terceiro ângulo diedro, ω , nem sempre é considerado. Ele envolve as ligações $C_\alpha-C-N-C_\alpha$. A ligação central nesse caso é a ligação peptídica, cuja rotação é restrita. A ligação peptídica está, normalmente (99,6% do tempo), na conformação *trans*, restringindo o ω a um valor de $\pm 180^\circ$. Em um caso raro de ligação peptídica *cis*, $\omega = 0^\circ$. ■

A princípio, ϕ e ψ podem ter qualquer valor entre $+180^\circ$ e -180° , mas diversos valores são proibidos por impedimento estérico entre os átomos do esqueleto polipeptídico e as cadeias laterais dos aminoácidos. A conformação na qual ambos, ϕ e ψ , são 0° (Figura 4-2d) é proibida por esta razão: essa conformação é apenas um ponto de referência para a descrição dos ângulos diedros. Valores permitidos de ϕ e ψ tornam-se evidentes quando ψ é colocado em um gráfico *versus* ϕ , no **diagrama de Ramachandran (Figura 4-3)**, introduzido por G. N. Ramachandran. Os diagramas de Ramachandran são ferramentas muito úteis e de uso frequente para testar a qualidade de estruturas tridimensionais de proteínas depositadas em bancos de dados internacionais.

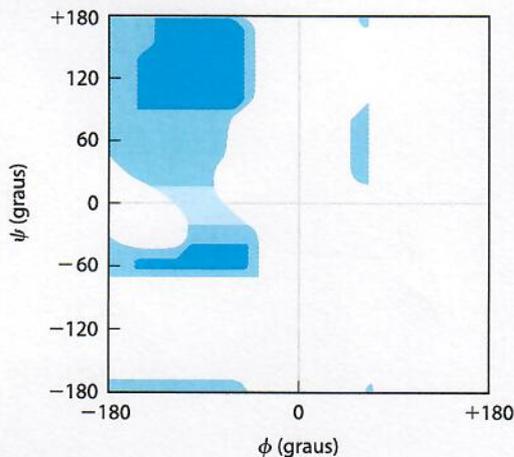


FIGURA 4-3 Diagrama de Ramachandran para resíduos L-Ala. A conformação dos peptídeos é definida pelos valores de ϕ e ψ . Conformações consideradas possíveis são aquelas que envolvem pouco ou nenhum impedimento estérico, com base nos cálculos dos raios de van der Waals conhecidos e dos ângulos diedros. As áreas coloridas em azul-escuro representam as conformações que não envolvem sobreposição estérica se os raios de van der Waals de cada átomo estão modelados como esferas rígidas, e, portanto, são totalmente permitidas; o azul médio indica as conformações permitidas, se for possível a aproximação dos átomos mais 0,1 nm, um leve choque; o azul-claro indica as conformações que são admitidas se for possível uma pequena flexibilidade (poucos graus) no ângulo diedro ω que descreve a própria ligação peptídica (geralmente presa a 180°). As regiões em branco são conformações não permitidas. A assimetria do diagrama é resultante da estereoquímica L dos resíduos de aminoácidos. Os diagramas para outros resíduos L com cadeias laterais não ramificadas são quase idênticos. Os limites permitidos para resíduos ramificados como Val, Ile e Thr são um pouco menores do que para Ala. O resíduo Gly, que é estericamente menos impedido, apresenta um limite bem mais amplo de conformações permitidas. O limite para os resíduos Pro é muito mais restrito porque seu ϕ é limitado entre -35° e -85° pela cadeia lateral cíclica.

RESUMO 4.1 Visão geral sobre a estrutura das proteínas

- ▶ Uma proteína típica geralmente tem uma ou mais estruturas tridimensionais, ou conformações que refletem sua função. Algumas proteínas têm segmentos intrinsecamente desordenados.
- ▶ A estrutura da proteína é estabilizada em grande parte por múltiplas interações fracas. As interações hidrofóbicas, derivadas do aumento da entropia da água circundante quando moléculas ou grupos apolares estão agrupados, são os principais contribuintes para a estabilização da forma globular da maioria das proteínas solúveis; as interações de van der Waals também contribuem. As ligações de hidrogênio e interações iônicas são otimizadas nas estruturas termodinamicamente mais estáveis.
- ▶ Ligações covalentes não peptídicas, particularmente ligações dissulfeto, são importantes na estabilização da estrutura de algumas proteínas.
- ▶ A natureza das ligações covalentes no esqueleto polipeptídico estabelece restrições à estrutura. A ligação peptídica tem um caráter parcial de ligação dupla, que mantém todo o grupo peptídico de seis átomos em uma configuração planar rígida. As ligações $N-C_\alpha$ e $C_\alpha-C$ podem girar para definir os ângulos diedros ϕ e ψ , respectivamente.
- ▶ O diagrama de Ramachandran é uma descrição visual das combinações dos ângulos diedros ϕ e ψ permitidos em um esqueleto peptídico ou não permitidos devido a impedimentos estéricos.

4.2 Estrutura secundária das proteínas

O termo **estrutura secundária** se refere a qualquer segmento de uma cadeia polipeptídica e descreve o arranjo espacial de seus átomos na cadeia principal, sem considerar a posição de suas cadeias laterais ou sua relação com outros segmentos. Uma estrutura secundária *comum* ocorre quando cada ângulo diedro, ϕ e ψ , permanece igual, ou quase igual, ao longo do segmento. Existem alguns tipos de estruturas secundárias que são particularmente estáveis e ocorrem extensamente em proteínas. As mais conhecidas são as hélices α e as conformações β ; outro tipo comum é a volta β . Quando um padrão regular não é observado, a estrutura secundária algumas vezes é chamada de indefinida ou espiral aleatória. Esta última, entretanto, não descreve adequadamente a estrutura desses segmentos. O curso da maioria dos esqueletos polipeptídicos em uma proteína típica não é aleatório; ele é altamente específico e inalterado na estrutura e na função de uma proteína em particular. Aqui, a discussão se concentra nas estruturas mais comuns.

A hélice α é uma estrutura secundária comum em proteínas

Arquitetura proteica – hélice α Pauling e Corey sabiam da importância das ligações de hidrogênio na orientação de grupos polares como os grupos $C=O$ e $N-H$ das ligações peptídicas. Também conheciam os resultados experimentais de William Astbury, que, em 1930, conduziu estudos pio-

neiros de proteínas com raios X. Astbury demonstrou que proteínas que formam os cabelos e os espinhos do porco-espinho (a proteína fibrosa α -queratina) têm uma estrutura regular que se repete a cada 5,15 a 5,2 Å. (O ångström, Å, em homenagem ao físico Anders J. Ångström, é igual a 0,1 nm. Apesar de não ser uma unidade do SI, ela é universalmente utilizada pelos biólogos estruturais para descrever as distâncias atômicas – é aproximadamente o tamanho de uma ligação C—H comum.) Com essa informação e seus dados sobre ligação peptídica, e com a ajuda de modelos construídos de forma precisa, Pauling e Corey iniciaram a determinação das conformações prováveis das moléculas de proteínas.

O primeiro avanço ocorreu em 1948. Pauling foi um professor visitante na Universidade de Oxford, ficou doente e se recolheu a seu apartamento por alguns dias para descansar. Entediado com a leitura disponível, Pauling pegou alguns papéis e lápis para trabalhar em uma estrutura estável plausível que poderia ser adotada por uma cadeia polipeptídica. O modelo que ele desenvolveu, confirmado mais tarde no trabalho com Corey e o colaborador Herman Branson, foi o arranjo mais simples que a cadeia polipeptídica pode assumir que maximiza o uso de ligações de hidrogênio internas. É uma estrutura helicoidal que Pauling e Corey chamaram de **hélice α** (Figura 4-4). Nessa estrutura, o esqueleto polipeptídico é firmemente enrolado em torno de um eixo imaginário desenhado longitudinalmente no centro

da hélice, e os grupos R dos resíduos de aminoácidos se projetam para fora do esqueleto helicoidal. A unidade que se repete forma uma volta de hélice, que se estende por cerca de 5,4 Å ao longo do eixo, levemente maior do que a periodicidade observada por Astbury na análise por raios X da queratina do cabelo. Os átomos do esqueleto dos resíduos de aminoácidos em uma hélice α típica têm um grupo característico de ângulos diedros que definem a conformação da hélice α (Tabela 4-1), e cada volta de hélice é formada por 3,6 resíduos de aminoácidos. Os segmentos de hélice α em proteínas normalmente se desviam um pouco desses ângulos diedros, podendo até variar dentro de um mesmo segmento, gerando curvaturas ou torções do eixo da hélice. Pauling e Corey consideraram as variantes da hélice α voltadas tanto para direita quanto para esquerda. A elucidação da estrutura tridimensional subsequente da hélice α da mioglobina e de outras proteínas mostrou que a hélice α voltada para direita é a forma comum (Quadro 4-1). As hélices α estendidas voltadas para esquerda são teoricamente menos estáveis e não foram observadas em proteínas. A hélice α é a estrutura predominante nas α -queratinas. De forma geral, cerca de um quarto de todos os resíduos de aminoácidos das proteínas é encontrado em hélices α . A fração exata varia muito de uma proteína para outra.

Por que as hélices α se formam mais facilmente do que qualquer outra conformação possível? A resposta encontra-se, em parte, no uso otimizado das ligações de hidrogênio

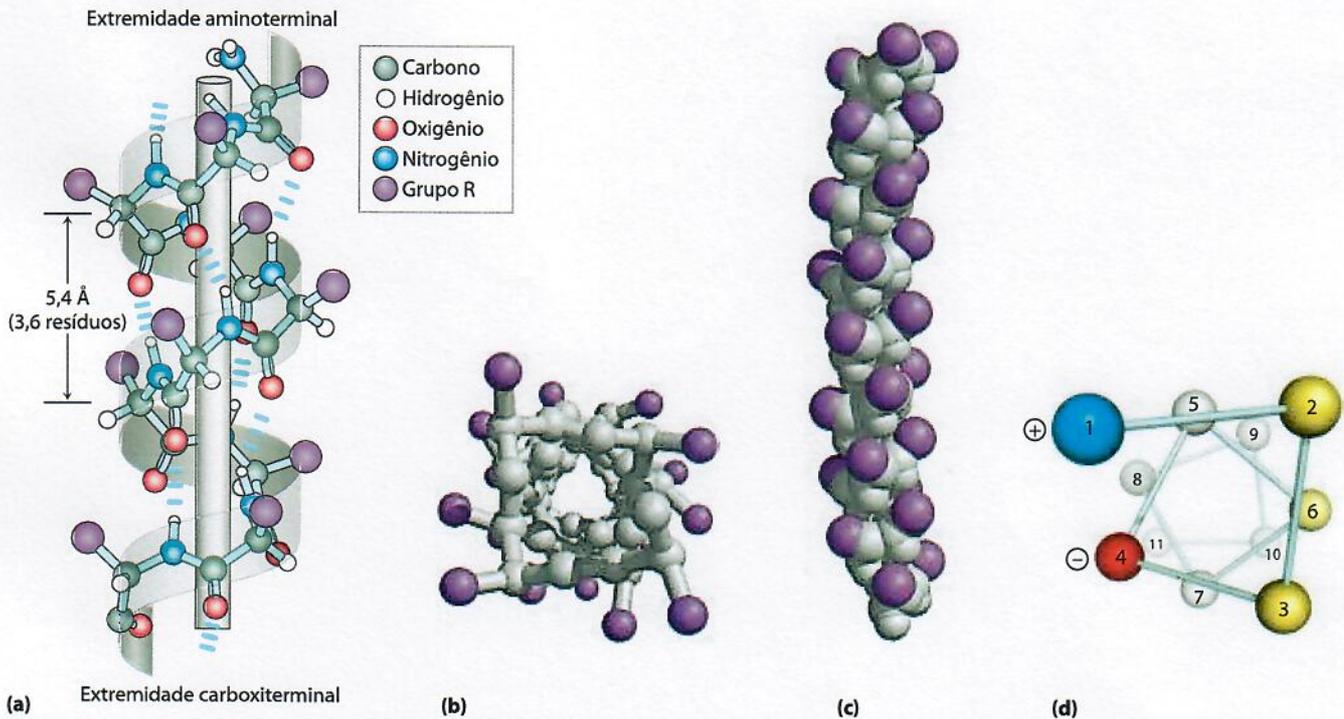


FIGURA 4-4 Modelos de hélice α , mostrando os diferentes aspectos de sua estrutura. (a) Modelo de esfera e bastão mostrando as ligações de hidrogênio internas da cadeia. A unidade que se repete forma uma volta da hélice: 3,6 resíduos. (b) Hélice α vista de uma de suas extremidades, ao longo do eixo central (obtida a partir do PDB ID 4TNC). Observe as posições dos grupos R, representados pelas esferas roxas. Observe modelo de esfera e bastão, que ressalta o arranjo helicoidal, dá uma falsa impressão de que a hélice é oca, pois as esferas não mostram os raios de van der Waals de cada um dos áto-

mos. (c) Como este modelo de volume atômico mostra, os átomos no centro da hélice α estão em contato estreito. (d) Projeção da rotação helicoidal de uma hélice α . Esta representação é colorida para a identificação de superfícies com determinadas propriedades. Os resíduos em amarelo, por exemplo, podem ser hidrofóbicos e fazer parte de uma interface entre a hélice mostrada aqui e outra parte do mesmo ou de outro polipeptídeo. Os resíduos vermelhos (negativo) e azuis (positivo) ilustram o potencial de interação de cadeias laterais de cargas opostas, separadas por dois resíduos na hélice.

TABELA 4-1 Ângulos ϕ e ψ ideais para estruturas secundárias comuns em proteínas

Estrutura	ϕ	ψ
Hélice α	-57°	-47°
Conformação β		
Antiparalela	-139°	+135°
Paralela	-119°	+113°
Tripla hélice de colágeno	-51°	+153°
Volta β tipo I		
<i>i</i> + 1*	-60°	-30°
<i>i</i> + 2*	-90°	0°
Volta β tipo II		
<i>i</i> + 1	-60°	+120°
<i>i</i> + 2	+80°	0°

Nota: Nas proteínas reais, os ângulos diedros frequentemente são um pouco diferentes desses valores ideais.
*Os ângulos *i* + 1 e *i* + 2 são aqueles para o segundo e terceiro resíduos de aminoácidos na volta β , respectivamente.

internas. A estrutura é estabilizada por uma ligação de hidrogênio entre o átomo de hidrogênio ligado ao átomo de nitrogênio eletronegativo de uma ligação peptídica e o átomo de oxigênio eletronegativo da carbonila do quarto aminoácido no lado aminoterminal da ligação peptídica (Figura 4-4a). Na hélice α , cada ligação peptídica (exceto aquelas próximas às extremidades da hélice α) participa de tais ligações de hidrogênio. Cada volta sucessiva da hélice α é mantida por voltas adjacentes por três ou quatro ligações de hidrogênio, conferindo uma significativa estabilidade à estrutura global. Nas extremidades de um segmento α -helicoidal, sempre há três ou quatro grupos carbonila ou amino que não podem participar desse padrão helicoidal de ligações de hidrogênio. Esses podem estar expostos ao solvente circundante, onde suas ligações de hidrogênio com

a água ou com outras partes da proteína podem proteger a hélice e proporcionar os parceiros necessários para a ligação de hidrogênio.

Outros experimentos mostraram que uma hélice α pode ser formada em polipeptídeos constituídos de L ou D-aminoácidos. Entretanto, todos os resíduos devem ser de um mesmo estereoisômero; um D-aminoácido irá romper a estrutura regular formada por L-aminoácidos, e vice-versa. A forma mais estável de uma hélice α de D-aminoácidos é a voltada para a esquerda.

PROBLEMA RESOLVIDO 4-1 Estrutura secundária e dimensões de uma proteína

Qual é o comprimento de uma cadeia polipeptídica de 80 resíduos de aminoácidos em uma única hélice α ?

Solução: Uma hélice α ideal tem 3,6 resíduos por volta, e o avanço ao longo do eixo helicoidal é de 5,4 Å. Portanto, o avanço sobre o eixo, para cada resíduo de aminoácido, é de 1,5 Å. O comprimento do peptídeo é, portanto, 80 resíduos \times 1,5 Å/resíduo = 120 Å.

A sequência de aminoácidos afeta a estabilidade da hélice α

Nem todos os polipeptídeos podem formar uma hélice α estável. Cada resíduo de aminoácido em um polipeptídeo tem uma propensão intrínseca de formar uma hélice α (Tabela 4-2), consequência das propriedades de seu grupo R e como elas interferem na capacidade de seus átomos de conexão da cadeia principal em aceitar os ângulos ϕ e ψ característicos. A alanina apresenta a melhor tendência a formar hélices α na maioria dos sistemas-modelo experimentais.

A posição de um resíduo de aminoácido em relação a seus vizinhos também é importante. Interações entre cadeias laterais dos aminoácidos podem estabilizar ou desestabilizar a estrutura α -helicoidal. Por exemplo, se uma cadeia polipeptídica possui uma longa sequência de resíduos

QUADRO 4-1 MÉTODOS Distinção entre o giro para a direita e o giro para a esquerda

Existe um método simples para determinar se uma estrutura helicoidal gira para a direita ou para a esquerda. Faça uma associação com suas duas mãos, com os polegares esticados, apontando para longe de você. Olhando para a sua mão direita, pense em uma hélice se enrolando ao longo de seu polegar direito, na direção dos seus outros dedos, dobrados conforme mostrado na figura (sentido horário). A hélice resultante gira para a direita. Sua mão esquerda representará uma hélice que gira para a esquerda no sentido anti-horário à medida que se enrola em seu polegar.

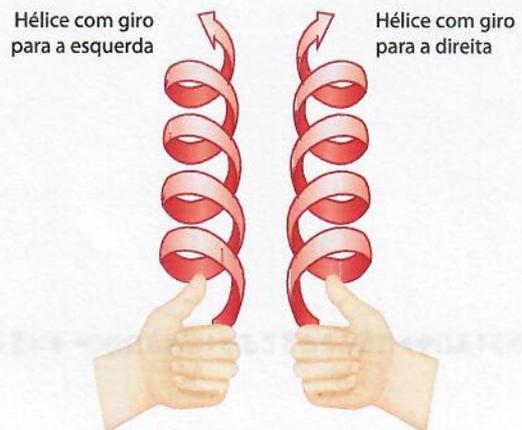


TABELA 4-2 Tendência dos resíduos de aminoácidos em assumir a conformação de hélice α

Aminoácido	$\Delta\Delta G^\circ$ (kJ/mol)*	Aminoácido	$\Delta\Delta G^\circ$ (kJ/mol)*
Ala	0	Leu	0,79
Arg	0,3	Lys	0,63
Asn	3	Met	0,88
Asp	2,5	Phe	2,0
Cys	3	Pro	> 4
Gln	1,3	Ser	2,2
Glu	1,4	Thr	2,4
Gly	4,6	Tyr	2,0
His	2,6	Trp	2,0
Ile	1,4	Val	2,1

Fontes: Dados (exceto para a prolina) de Bryson, J.W., Betz, S.F., Lu, H.S., Suich, D.J., Zhou, H.X., O'Neil, K.T. e DeGrado, W.F. (1995) Protein design: a hierarchic approach. *Science* **270**, 935. Dados sobre a prolina de Myers, J.K., Pace, C.N. e Scholtz, J.M. (1997) Helix propensities are identical in proteins and peptides. *Biochemistry* **36**, 10, 926.

* $\Delta\Delta G^\circ$ é a diferença de variação de energia livre, relativa àquela para a alanina, necessária para que os resíduos de aminoácidos assumam a conformação em hélice α . Valores maiores refletem uma grande dificuldade para assumir a conformação em hélice. O conjunto de dados é derivado de múltiplos experimentos e sistemas experimentais.

Glu, esse segmento da cadeia não irá formar uma hélice α em pH 7,0. Os grupos carboxílicos, carregados negativamente, dos resíduos Glu adjacentes repelem-se mutuamente de forma tão forte que impedem a formação da hélice α . Pela mesma razão, se existem muitos resíduos Lys e/ou Arg, com grupos R carregados positivamente em pH 7,0, eles também se repelem, impedindo a formação da hélice α . O volume e a forma dos resíduos Asn, Ser, Thr e Cys também podem desestabilizar uma hélice α se estiverem muito próximos na cadeia.

A torção de uma hélice garante que ocorram interações críticas entre a cadeia lateral de um aminoácido e a cadeia lateral do terceiro (às vezes quarto) resíduo adiante, para ambos os lados da hélice. Isso fica claro quando a hélice α é representada como roda helicoidal (Figura 4-4d). Os aminoácidos carregados positivamente costumam ser encontrados a três resíduos de distância dos aminoácidos carregados negativamente, possibilitando a formação de pares iônicos. Dois aminoácidos aromáticos com frequência são espaçados de forma semelhante, resultando em uma interação hidrofóbica.

Uma restrição à formação da hélice α é a presença de resíduos de Pro e Gly, que apresentam a menor propensão em formar hélices α . Na prolina, o átomo de nitrogênio faz parte de um anel rígido (ver Figura 4-8), e a rotação sobre a ligação N—C $_{\alpha}$ não é possível. Dessa forma, um resíduo Pro gera uma torção, que desestabiliza a hélice. Além disso, o átomo de nitrogênio do resíduo Pro em uma ligação peptídica não tem hidrogênio para participar em ligações com outros resíduos. Por essas razões, a prolina raramente é encontrada em uma hélice α . A glicina com frequência não ocorre em hélices por outro

motivo: ela apresenta maior flexibilidade conformacional do que os outros aminoácidos. Polímeros de glicina tendem a formar estruturas espiraladas bem diferentes de uma hélice α .

Um último fator que afeta a estabilidade de uma hélice α é a identidade dos resíduos de aminoácido próximos às extremidades do segmento α -helicoidal do polipeptídeo. Existe um pequeno dipolo elétrico em cada ligação peptídica (Figura 4-2a). Esses dipolos estão alinhados através das ligações de hidrogênio da hélice, resultando em um dipolo livre ao longo do eixo helicoidal que aumenta com o comprimento da hélice (Figura 4-5). As cargas parcialmente positivas e negativas do dipolo da hélice ocorrem nos grupos amino e carbonil próximos às extremidades amino e carboxiterminal, respectivamente. Por isso, aminoácidos carregados negativamente costumam ser encontrados próximos à extremidade aminoterminal do segmento helicoidal, onde apresentam interações estabilizantes com a carga positiva do dipolo da hélice; um aminoácido positivamente carregado na extremidade aminoterminal desestabilizaria o sistema. O oposto é verdade para a extremidade carboxiterminal do segmento helicoidal.

Resumindo, cinco tipos de restrições influenciam a estabilidade de uma hélice α : (1) a tendência intrínseca de um resíduo de aminoácido de formar uma hélice α ; (2) as interações entre os grupos R, especialmente aqueles espaçados por três (ou quatro) aminoácidos; (3) os volumes de grupos R adjacentes; (4) a ocorrência de resíduos Pro e Gly; e (5) interações entre os resíduos de aminoácidos das extremidades do segmento helicoidal e o dipolo elétrico inerente da hélice α . A tendência de um determinado segmento de uma cadeia polipeptídica de formar uma hélice α depende, portanto, da identidade e da sequência de resíduos de aminoácidos do segmento.

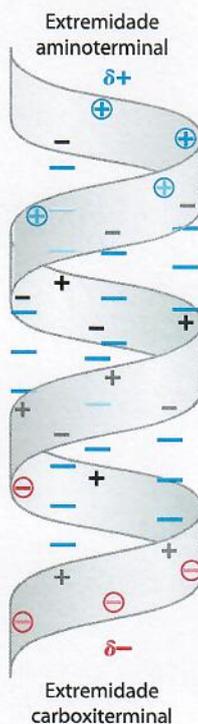


FIGURA 4-5 Dipolo da hélice. O dipolo elétrico da ligação peptídica (ver Figura 4-2a) é transmitido ao longo do segmento α -helicoidal pelas ligações de hidrogênio intracadeia, resultando em um dipolo da hélice. Nesta ilustração, os componentes amino e carbonil de cada ligação peptídica estão indicados com os símbolos + e -, respectivamente. Os componentes amino e carbonil das ligações peptídicas próximas às extremidades da região α -helicoidal, que não participam de ligações de hidrogênio, estão circulares e mostrados em cores.

As conformações β organizam as cadeias polipeptídicas em forma de folha

Arquitetura proteica – folha β Em 1951, Pauling e Corey reconheceram um segundo tipo de estrutura recorrente, a **conformação β** . Essa é uma conformação mais estendida das cadeias polipeptídicas, e sua estrutura é de novo definida pelos esqueletos dos átomos arranjados de acordo com grupo característico de ângulos diedro (Tabela 4-1). Na conformação β , o esqueleto da cadeia polipeptídica está estendido em forma de zigue-zague, em vez de em estrutura helicoidal (Figura 4-6). O arranjo de vários segmentos lado a lado, os quais estão na conformação β , é chamado de **folha β** . A estrutura em zigue-zague dos segmentos polipeptídicos individuais dá origem a uma aparência pregueada da folha em geral. As ligações de hidrogênio são formadas entre segmentos adjacentes da cadeia polipeptídica, dentro da folha. Os segmentos que formam a folha β normalmente estão próximos na cadeia polipeptídica, mas também podem estar bem distantes uns dos outros na sequência linear

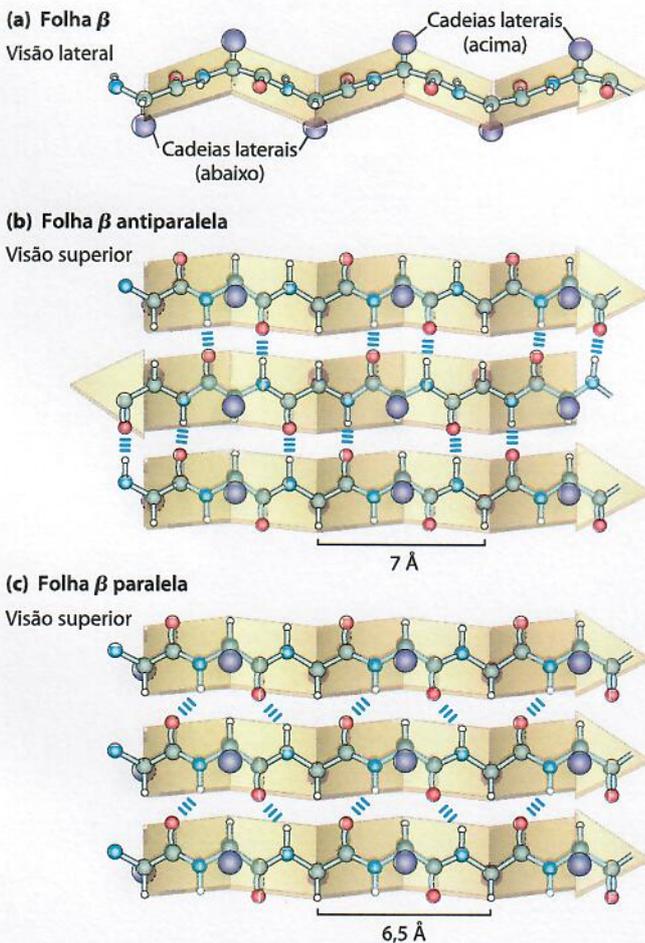


FIGURA 4-6 A conformação β das cadeias polipeptídicas. Estas visões (a) lateral e (b, c) superior mostram os grupos R saindo do plano da folha β e enfatizam a forma pregueada formada pelos planos das ligações peptídicas. (Um nome alternativo para esta estrutura é folha β pregueada.) As ligações de hidrogênio entre as cadeias adjacentes também são mostradas. A orientação das cadeias adjacentes (setas), do aminoterminal para carboxiterminal, pode ser a mesma ou oposta, formando (b) folhas β antiparalelas ou (c) folhas β paralelas.

do polipeptídeo; eles podem até estar em cadeias polipeptídicas diferentes. Os grupos R dos aminoácidos adjacentes se projetam da estrutura em zigue-zague em direções opostas, criando um padrão alternado que pode ser observado na visão lateral da Figura 4-6.

As cadeias polipeptídicas adjacentes em uma folha β podem ser tanto paralelas quanto antiparalelas (apresentando uma orientação aminocarboxiterminal igual ou oposta, respectivamente). As estruturas são, de alguma forma, semelhantes, apesar de o período de repetição ser menor na conformação paralela (6,5 Å, versus 7 Å para a antiparalela) e o padrão das ligações de hidrogênio ser diferente. As ligações de hidrogênio intersegmentos são alinhadas (ver Figura 2-5) na folha β antiparalela, enquanto elas são distorcidas ou não alinhadas na variante paralela. As estruturas ideais exibem os ângulos de ligação dados na Tabela 4-1; esses valores variam um pouco nas proteínas verdadeiras, resultando em uma variação estrutural, conforme visto anteriormente para as hélices α .

Voltas β são comuns em proteínas

Arquitetura proteica – voltas β Em proteínas globulares, que apresentam estrutura dobrada compacta, alguns resíduos de aminoácidos estão em voltas ou alças onde a cadeia polipeptídica inverte sua direção (Figura 4-7). Esses são elementos conectores que ligam estruturas sucessivas de hélices α e conformações β . As **voltas β** são particularmente comuns e conectam as extremidades de dois segmentos adjacentes de uma folha β antiparalela. A estrutura é uma volta de 180° que envolve quatro resíduos de aminoácidos, com o oxigênio carbonílico do primeiro resíduo formando uma ligação de hidrogênio com o hidrogênio do grupo amino do quarto resíduo. Os grupos peptídicos dos dois resíduos centrais não participam de nenhuma ligação de hidrogênio inter-resíduos. Vários tipos de voltas β têm sido descritas, cada uma definida pelos ângulos ϕ e ψ das ligações que ligam os quatro resíduos de aminoácidos que formam a volta em particular (Tabela 4-1). Os resíduos Gly e Pro frequentemente ocorrem em voltas β ; o primeiro porque é pequeno e flexível, e o último porque as ligações peptídicas envolvendo o nitrogênio imino da prolina facilmente assumem configuração *cis* (Figura 4-8), forma particularmente acessível em uma volta fechada. Os dois tipos de voltas β , mostrados na Figura 4-7, são os mais comuns. As voltas normalmente são encontradas próximas à superfície das proteínas, onde os grupos peptídicos dos dois resíduos de aminoácidos centrais da alça podem fazer ligação de hidrogênio com a água. Muito menos comum é a volta γ , uma volta com três resíduos e ligação de hidrogênio entre o primeiro e o terceiro resíduo.

Estruturas secundárias comuns têm ângulos diedros característicos

As hélices α e as conformações β são as principais estruturas secundárias que se repetem em um grande número de proteínas, apesar de existirem outras estruturas que se repetem em algumas proteínas especializadas (um exemplo é o colágeno; ver Figura 4-13). Cada tipo de estrutura secundária pode ser completamente descrito pelos ângulos diedros ϕ e ψ , associados a cada resíduo. Como mos-

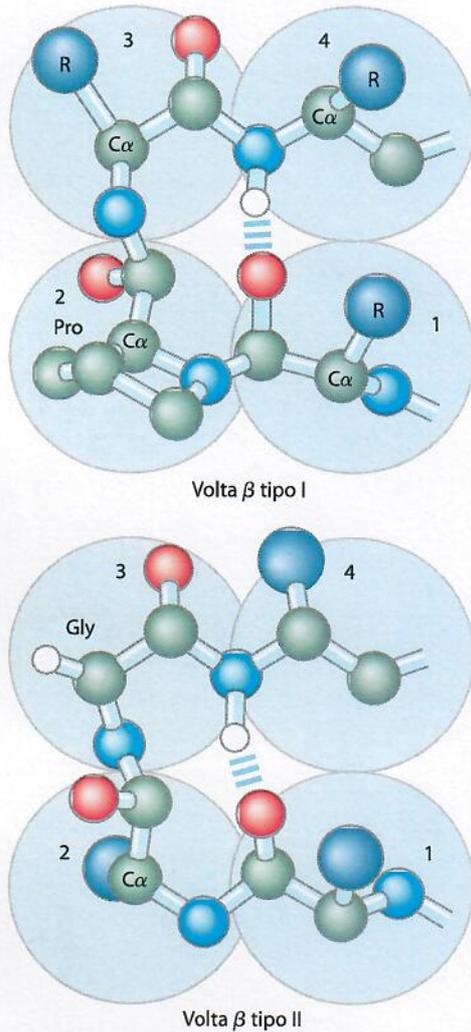


FIGURA 4-7 Estruturas de voltas β . Voltas β dos tipos I e II são as mais comuns; distinguem-se pelos ângulos ϕ e ψ adotados pelo esqueleto peptídico na volta (ver Tabela 4-1). As voltas do tipo I ocorrem duas vezes mais do que as voltas do tipo II. As voltas β do tipo II normalmente possuem uma Gly como terceiro resíduo. Observe a ligação de hidrogênio entre os grupos peptídicos do primeiro e do quarto resíduos da volta. (Cada um dos resíduos de aminoácidos está identificado por grandes círculos azuis. Nem todos os átomos de H estão mostrados.)

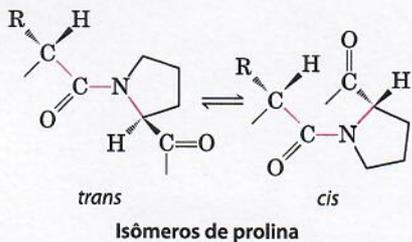


FIGURA 4-8 Isômeros *trans* e *cis* de uma ligação peptídica envolvendo o nitrogênio imino da prolina. Das ligações peptídicas entre resíduos de aminoácidos que não Pro, mais de 99,95% estão na configuração *trans*. Para as ligações peptídicas envolvendo o nitrogênio imino da prolina, no entanto, cerca de 6% estão na configuração *cis*; muitas delas ocorrem em voltas β .

trado no diagrama de Ramachandran, os ângulos diedros que definem a hélice α e a conformação β se encontram em uma região restrita de estruturas estericamente permitidas (**Figura 4-9a**). A maioria dos valores de ϕ e ψ obtidos de estruturas de proteínas conhecidas cai nas regiões esperadas, com alta concentração próximo aos valores preditos para as hélices α e as conformações β (**Figura 4-9b**). O único resíduo de aminoácido normalmente encontrado fora dessas regiões é a glicina. Como sua cadeia lateral é pequena, o resíduo Gly pode assumir diversas conformações estericamente proibidas para os demais aminoácidos.

As estruturas secundárias comuns podem ser identificadas por dicroísmo circular

Qualquer forma de assimetria estrutural em uma molécula leva a diferenças de absorção da luz circularmente polari-

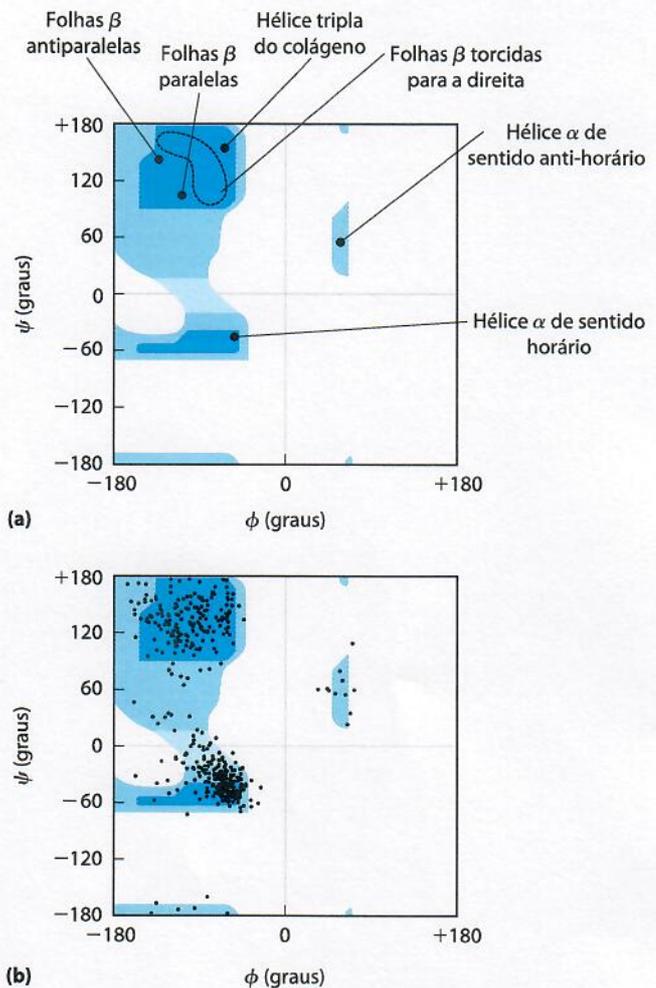


FIGURA 4-9 Diagrama de Ramachandran mostrando uma variedade de estruturas. (a) Os valores de ϕ e ψ de várias estruturas secundárias permitidas estão sobrepostos no diagrama da Figura 4-3. Apesar de as hélices α de sentido anti-horário com vários resíduos de aminoácidos serem teoricamente permitidas, elas não têm sido observadas em proteínas. (b) Os valores de ϕ e ψ para todos os aminoácidos, exceto Gly, da enzima piruvato-cinase (isolada de coelho) estão sobrepostos no diagrama das conformações teoricamente permitidas (**Figura 4-3**). Os pequenos e flexíveis resíduos de Gly foram excluídos, pois em geral caem fora da região esperada (em azul).

zada para a esquerda em relação à direita. A medida desta diferença é chamada de **espectroscopia de dicroísmo circular (CD, de circular dichroism)**. Uma estrutura ordenada, como de uma proteína dobrada, resulta em um espectro de absorção que pode ter picos ou regiões com valores tanto positivos, quanto negativos. Para proteínas, os espectros são obtidos na região de UV distante (190 a 250 nm). A entidade que absorve luz, ou cromóforo, nessa região é a ligação peptídica; um sinal é obtido quando esta ligação peptídica está em um ambiente dobrado. A diferença em coeficiente de extinção molar (ver Quadro 3-1) para a luz circularmente polarizada para a esquerda e para a direita ($\Delta\epsilon$) é colocada no gráfico em função do comprimento de onda. As hélices α e as conformações β têm espectros de CD característicos (**Figura 4-10**). Com o espectro de CD, os bioquímicos podem determinar se as proteínas estão dobradas corretamente, estimar a fração da proteína que assume qualquer uma das duas estruturas secundárias comuns e monitorar as transições entre os estados dobrados e não dobrados.

RESUMO 4.2 Estrutura secundária das proteínas

- ▶ Estrutura secundária é o arranjo espacial dos átomos da cadeia principal em um determinado segmento da cadeia polipeptídica.
- ▶ As estruturas secundárias regulares mais comuns são as hélices α , as conformações β e as voltas β .
- ▶ A estrutura secundária de um segmento polipeptídico pode ser completamente definida se seus ângulos ϕ e ψ são conhecidos para todos os aminoácidos do segmento.
- ▶ A espectroscopia de dicroísmo circular é um método para a identificação das estruturas secundárias comuns e o monitoramento do enovelamento das proteínas.

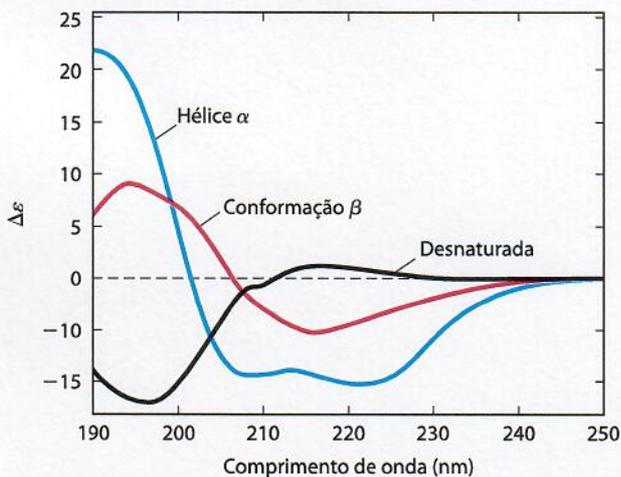


FIGURA 4-10 Espectroscopia de dicroísmo circular (CD). Este espectro mostra a polilisina inteiramente como hélice α , como conformação β ou desnaturada, espiral aleatória. A unidade do eixo y é uma simplificação das unidades comumente utilizadas nos experimentos de CD. Como as curvas são diferentes para as hélices α , as conformações β e a desnaturada, o espectro de CD fornece a determinada proteína uma estimativa da fração da proteína, formada pelas duas estruturas secundárias mais comuns. O espectro de CD da proteína nativa pode servir como referência para o estado enovelado, e é útil no monitoramento da desnaturação, ou de mudanças conformacionais resultantes de alterações nas condições de solução.

4.3 Estruturas terciária e quaternária das proteínas

Arquitetura proteica – Introdução da estrutura terciária O arranjo tridimensional total de todos os átomos de uma proteína é chamado de **estrutura terciária**. Enquanto o termo “estrutura secundária” se refere ao arranjo espacial dos resíduos de aminoácidos adjacentes em um segmento polipeptídico, a estrutura terciária inclui aspectos de *alcance mais longo* da sequência de aminoácidos. Aminoácidos que estão bem distantes na sequência polipeptídica e em diferentes tipos de estruturas secundárias podem interagir na estrutura da proteína completamente dobrada. A localização das curvaturas (incluindo as voltas β) nas cadeias polipeptídicas e sua direção e seu ângulo são determinados pelo número e pela localização de resíduos específicos que tendem a formá-las, como Pro, Thr, Ser e Gly. Segmentos da cadeia polipeptídica que interagem entre si são mantidos em suas posições terciárias características por diferentes tipos de interações fracas (e algumas vezes por ligações covalentes, como ligações dissulfeto) entre os segmentos.

Algumas proteínas contêm duas ou mais cadeias polipeptídicas distintas, ou subunidades, que podem ser idênticas ou diferentes. O arranjo destas subunidades proteicas em complexos tridimensionais constitui a **estrutura quaternária**.

Considerando esses níveis mais altos de estrutura, é conveniente designar dois grandes grupos nos quais muitas proteínas podem ser classificadas: **proteínas fibrosas**, com cadeias polipeptídicas arranjadas em longos filamentos ou folhas, e **proteínas globulares**, com cadeias polipeptídicas dobradas em forma esférica ou globular. Os dois grupos são estruturalmente distintos. As proteínas fibrosas em geral são formadas por um único tipo de estrutura secundária, e sua estrutura terciária é relativamente simples. As proteínas globulares normalmente contêm diversos tipos de estruturas secundárias. Os dois grupos também se diferenciam funcionalmente: as estruturas que garantem suporte, forma e proteção externa aos vertebrados são feitas de proteínas fibrosas, enquanto as enzimas e as proteínas reguladoras em sua maioria são proteínas globulares.

As proteínas fibrosas são adaptadas às funções estruturais

Arquitetura proteica – Estrutura terciária da proteína fibrosa A α -queratina, o colágeno e a fibroína da seda ilustram bem a relação entre a estrutura da proteína e sua função biológica (Tabela 4-3). As proteínas fibrosas compartilham propriedades que dão força e/ou flexibilidade às estruturas nas quais ocorrem. Em cada caso, a unidade estrutural fundamental é um elemento simples de estrutura secundária que se repete. Todas as proteínas fibrosas são insolúveis em água, propriedade conferida pela alta concentração de resíduos de aminoácidos hidrofóbicos tanto no interior quanto na superfície da proteína. Essas superfícies hidrofóbicas estão em grande parte escondidas, visto que muitas cadeias polipeptídicas similares são reunidas formando um complexo supramolecular. A simplicidade estrutural das proteínas fibrosas as torna particularmente interessantes para ilus-

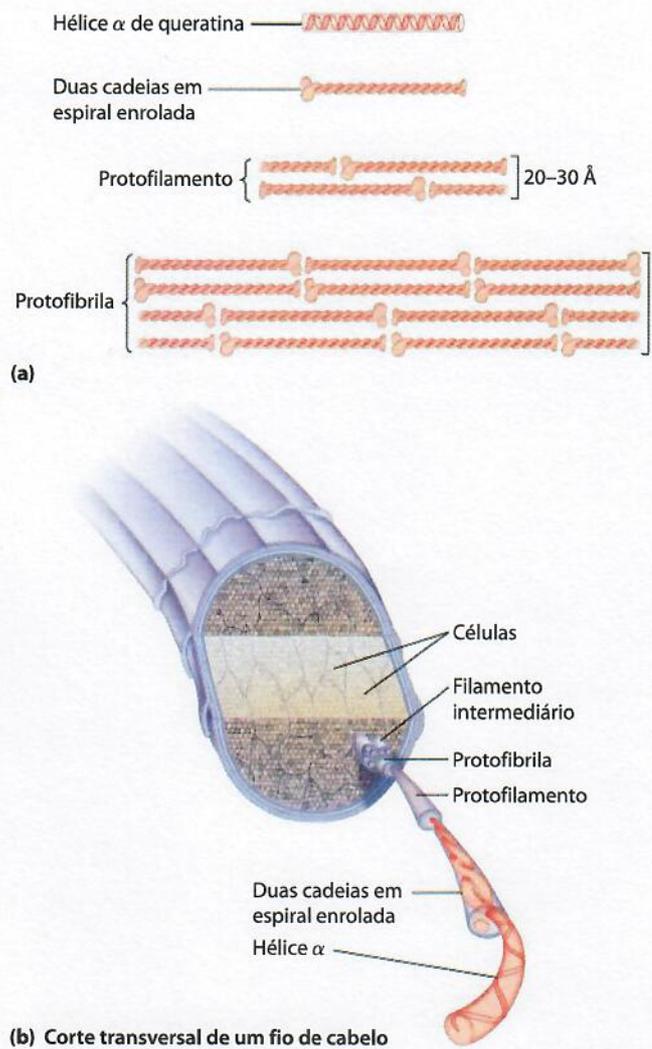
trar alguns dos princípios fundamentais da estrutura proteica discutidos anteriormente.

α -Queratina As α -queratinas foram desenvolvidas para força. Encontradas somente em mamíferos, essas proteínas constituem praticamente todo o peso seco de cabelos, pelos, unhas, garras, penas, chifres, cascos e grande parte da camada mais externa da pele. As α -queratinas fazem parte de uma família mais ampla de proteínas chamadas de proteínas de filamento intermediário (FI). Outras proteínas FI são encontradas no citoesqueleto de células animais. Todas as proteínas FI têm função estrutural e compartilham das características estruturais das α -queratinas.

A hélice da α -queratina é uma hélice α voltada para a direita, a mesma hélice encontrada em várias outras proteínas. Francis Crick e Linus Pauling, no início de 1950, sugeriram, independentemente, que as hélices α da queratina estavam arranjadas na forma de espiral enrolada. Duas fibras de α -queratina, orientadas em paralelo (com seus aminoterminais na mesma extremidade), são enroladas uma sobre a outra, formando uma espiral enrolada supertorcida. A supertorção amplifica a força da estrutura como um todo, assim como as fibras são trançadas para formar uma corda forte (Figura 4-11). A torção do eixo de uma hélice α para formar uma espiral enrolada explica a discrepância entre os 5,4 Å por volta preditos por Pauling e Corey, e as estruturas repetidas a cada 5,15 a 5,2 Å observadas na difração de raios X do cabelo (p. 120). O sentido da hélice das estruturas supertorcidas é anti-horário, oposto ao da hélice α . As superfícies onde as duas hélices α se tocam são formadas por resíduos de aminoácidos hidrofóbicos, e seus grupos R se entrelaçam em um padrão regular interconectado. Isso permite um arranjo próximo das cadeias polipeptídicas dentro da estrutura supertorcida no sentido anti-horário. Não é surpresa, portanto, que a α -queratina seja rica em resíduos hidrofóbicos Ala, Val, Leu, Ile, Met e Phe.

Um polipeptídeo na α -queratina com as cadeias em espiral enrolada tem uma estrutura terciária relativamente simples, dominada por uma estrutura secundária α -helicoidal, com seu eixo enrolado em uma super-hélice anti-horária. O entrelaçamento de dois peptídeos α -helicoidais é um exemplo de estrutura quaternária. Espirais enroladas desse tipo são elementos estruturais comuns em proteínas filamentosas e na proteína muscular miosina (ver Figura 5-27). A estrutura quaternária da α -queratina pode ser bem complexa. Várias espirais enroladas podem ser associadas em grandes complexos supramoleculares, como o arranjo da α -queratina para formar o filamento intermediário do cabelo (Figura 4-11b).

A resistência das proteínas fibrosas é aumentada pelas ligações covalentes entre as cadeias polipeptídicas nas "cordas" multi-helicoidais e entre cadeias adjacentes em



(b) Corte transversal de um fio de cabelo

FIGURA 4-11 Estrutura do cabelo. (a) A α -queratina do cabelo é uma longa hélice α com elementos mais densos próximos às extremidades amino e carboxiterminais. Os pares destas hélices são enrolados em um sentido anti-horário para formar as duas cadeias em espiral enrolada. Estas, por sua vez, se combinam em estruturas mais complexas, chamadas de protofilamentos e protofibrilas. Cerca de quatro protofibrilas – 32 moléculas de α -queratina ao todo – se combinam para formar um filamento intermediário. As duas cadeias em espiral enrolada das várias subestruturas também parecem estar entrelaçadas, mas a orientação destas torções e outros detalhes estruturais não são conhecidos. (b) O cabelo é um conjunto de filamentos de α -queratina, formado pelas subestruturas mostradas em (a).

um arranjo supramolecular. Nas α -queratinas, as ligações entre as cadeias que estabilizam a estrutura quaternária são ligações dissulfeto (Quadro 4-2). Nas α -queratinas mais

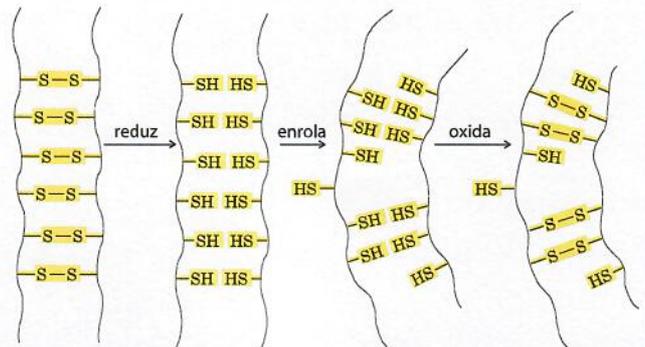
TABELA 4-3 Estrutura secundária e propriedades de algumas proteínas fibrosas

Estrutura	Características	Exemplos de ocorrência
Hélices α , ligadas por ligações dissulfeto	Estruturas de proteção insolúveis e resistentes, com dureza e flexibilidade variáveis	α -Queratina de cabelos, penas, unhas
Conformação β	Filamentos macios e flexíveis	Fibroínas de seda
Hélice tripla de colágeno	Grande resistência à tração, sem elasticidade	Colágeno de tendões, matriz óssea

QUADRO 4-2 Ondulação permanente é engenharia bioquímica

Quando os cabelos são expostos ao calor úmido, eles podem ser esticados. No nível molecular, as hélices α da α -queratina do cabelo são esticadas até que cheguem a uma conformação β completamente estendida. Com o resfriamento espontâneo, elas retornam para a conformação α -helicoidal. A “esticabilidade” característica das α -queratinas e suas inúmeras ligações dissulfeto constituem a base dos processos da ondulação permanente. Os cabelos que serão cacheados ou encrespados são inicialmente enrolados ao redor de uma forma com um formato adequado. Uma solução de um agente redutor, normalmente um composto contendo um grupo tiol ou sulfidril (—SH), é então aplicada com calor. O agente redutor rompe as ligações transversais pela redução de cada ligação dissulfeto em dois resíduos Cys. O calor úmido quebra as ligações de hidrogênio e provoca o desenrolamento da estrutura da cadeia polipeptídica. Depois de um tempo, a solução redutora é removida, e um agente oxidante é adicionado para fazer *novas* ligações dissulfeto entre pares de resíduos Cys de cadeias adjacentes, mas não dos mesmos pares de antes do tratamento. Depois do cabelo lavado e resfriado, as cadeias polipeptídicas voltam à sua conformação

helicoidal. As fibras do cabelo agora enrolam da forma desejada porque as novas ligações dissulfeto transversais exercem uma torção ou rotação sobre os feixes de hélices das fibras. O mesmo processo pode ser usado para o alisamento dos cabelos naturalmente crespos. Um alisamento permanente de cabelo não é permanente de fato, pois o cabelo cresce; e no cabelo novo que substituirá o antigo, a α -queratina tem o padrão natural de ligações dissulfeto.



duras e resistentes, como as dos chifres dos rinocerontes, até 18% dos resíduos são cisteínas envolvidas em ligações dissulfeto.

Colágeno Assim como as α -queratinas, o colágeno evoluiu para garantir resistência. Ele é encontrado nos tecidos conectivos como os tendões, as cartilagens, a matriz orgânica dos ossos e na córnea dos olhos. A hélice de colágeno é uma estrutura secundária única, bem diferente da hélice α . Ela gira para esquerda e tem três resíduos de aminoácidos por volta (Figura 4-12 e Tabela 4-1). O colágeno também é uma espiral enrolada, mas com estruturas terciária e quaternária distintas: três polipeptídeos separados, chamados de cadeias α (não confundir com hélices α), são supertorcidos uns sobre os outros (Figura 4-12c). No colágeno, a torção super-helicoidal tem sentido horário, oposto ao da hélice anti-horária das cadeias α .

Existem vários tipos de colágenos nos vertebrados. Normalmente, eles contêm em torno de 35% de Gly, 11% de Ala e 21% de Pro e 4-Hyp (4-hidroxiprolina, um aminoácido incomum; ver Figura 3-8a). A gelatina comestível é derivada do colágeno. Tem baixo poder nutricional como proteína, pois o colágeno tem uma quantidade muito baixa dos aminoácidos essenciais à dieta humana. O conteúdo de aminoácidos incomuns do colágeno está relacionado com as restrições estruturais únicas de sua hélice. A sequência de aminoácidos no colágeno geralmente é uma repetição de uma unidade tripeptídica, Gly-X-Y, onde X normalmente é uma Pro e Y em geral é uma 4-Hyp. Somente os resíduos Gly podem ser acomodados nas junções muito apertadas entre as cadeias α individuais (Figura 4-12d). Os resíduos Pro e 4-Hyp permitem a torção acentuada da hélice do colágeno. A sequência de aminoácidos e a estrutura quaternária supertorcida do colágeno permitem uma compactação

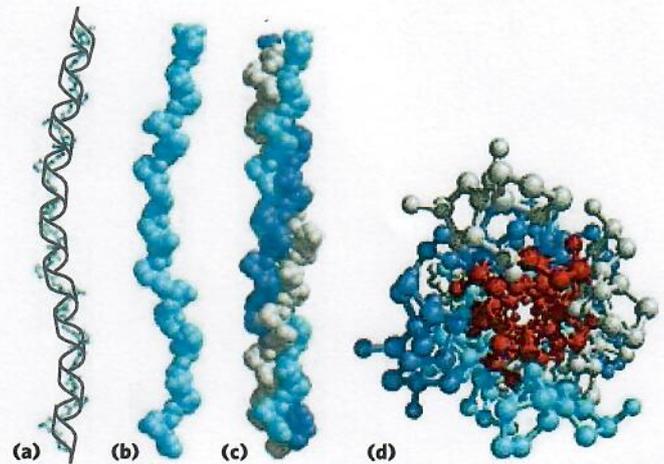


FIGURA 4-12 Estrutura do colágeno. (Obtida do PDB ID 1CGD.) (a) A cadeia α do colágeno tem uma estrutura secundária repetitiva que é única desta proteína. A sequência tripeptídica que se repete Gly-X-Pro ou Gly-X-4-Hyp adota uma estrutura helicoidal anti-horária com três resíduos por volta. A sequência repetida para gerar este modelo é a Gly-Pro-4-Hyp. (b) Modelo de volume atômico da mesma cadeia α . (c) Três destas hélices (mostradas aqui em cinza, azul e roxo) se enrolam uma sobre as outras, no sentido horário. (d) A super-hélice de colágeno formada por três cadeias, mostrada a partir de uma das extremidades, está representada em um modelo de esfera e bastão. Os resíduos de Gly são mostrados em vermelho. A glicina, por ser pequena, é necessária para uma junção firme na região onde as três cadeias estão em contato. As esferas nesta ilustração não representam os raios de van der Waals dos átomos individuais. O centro da super-hélice de três cadeias não é oco, como aparece aqui, mas sim firmemente compacto.

muito justa de seus três polipeptídeos. A 4-hidroxiprolina tem um papel importante na estrutura do colágeno – e na história do ser humano (Quadro 4-3).

QUADRO 4-3



MEDICINA

Por que marinheiros, exploradores e universitários devem comer frutas e vegetais frescos

“... por um azar, junto com a insalubridade do país, onde nunca cai uma gota de chuva, fomos acometidos pela “doença”, que era tal que toda a carne de nossos braços murchou, e a pele de nossas pernas ficou com manchas escuras, com pedaços bolorentos, como uma bota velha; e uma carne esponjosa surgiu nas gengivas daqueles que pegaram a doença, e ninguém escapou dela, indo direto para as garras da morte. O sinal era o seguinte: quando o nariz começava a sangrar, então a morte estava próxima...”

— Memórias do Lorde de Joinville, 1300 d.C.

Essa passagem descreve a situação do exército de Luís IX no fim da Sétima Cruzada (1248-1254), quando o exército, enfraquecido pelo escorbuto, foi destruído pelos egípcios. Qual a natureza desse mal que acometeu esses soldados do décimo terceiro século?

O escorbuto é causado pela falta de vitamina C, ou ácido ascórbico (ascorbato). A vitamina C é necessária para, entre outras coisas, a hidroxilação da prolina e da lisina no colágeno. O escorbuto é uma doença caracterizada pela degeneração do tecido conectivo, sendo que sua manifestação em estágio avançado inclui inúmeras pequenas hemorragias causadas por vasos sanguíneos frágeis, perda dos dentes, difícil cicatrização de feridas e reabertura de feridas antigas, dor e degeneração dos ossos, e, no final, falência cardíaca. Casos mais brandos de deficiência de vitamina C são acompanhados de fadiga, irritabilidade e aumento da gravidade das infecções do trato respiratório. A maioria dos animais sintetiza grande quantidade de vitamina C pela conversão da glicose em ascorbato em quatro etapas enzimáticas. Contudo, no curso da evolução, os humanos e alguns outros animais – gorilas, cobaias e morcegos frugívoros – perderam a última enzima dessa rota, devendo obter o ascorbato da dieta. A vitamina C é encontrada em uma grande variedade de frutas e vegetais. Até 1800, entretanto, ela estava ausente nos alimentos desidratados e outros suprimentos alimentares estocados para o inverno ou para longas viagens.

O escorbuto foi registrado pelos egípcios em 1500 a.C., e descrito em escritos de Hipócrates, que datam do quinto século a.C. Entretanto, ele não tornou-se uma notícia pública até as viagens de descobrimento europeias de 1500 a 1800. A primeira navegação ao redor do mundo (1519-1522), liderada por Ferdinand Magellan, acabou com a perda de mais de 80% de sua tripulação para o escorbuto. Durante a segunda viagem de Jacques Cartier para explorar o rio St. Lawrence (1535-1536), a tripulação foi ameaçada por um completo desastre, até que um americano nativo ensinou os homens a fazer chá de cedro (contendo vitamina C) que curava e prevenia o escorbuto. As epidemias de escorbuto nos invernos da Europa foram gradativamente sendo eliminadas no século XIX, à medida que a cultura de batata, originária da América do Sul, foi disseminada.

Em 1747, James Lind, cirurgião escocês da Marinha Real, realizou o primeiro estudo clínico controlado registrado na história. Durante uma longa viagem no navio de guerra *HMS Salisbury*, Lind selecionou 12 marinheiros com escorbuto e os dividiu em grupos de dois. Todos os 12 receberam a mesma dieta, exceto que cada um dos grupos recebeu um remédio diferente, dentre os recomendados na época, para o escorbuto. Os marinheiros que receberam limões e laranjas se recuperaram e voltaram ao trabalho. Os marinheiros que receberam suco de maçã fervida tiveram uma pequena melhora. Os outros continuaram a piorar. O *Tratado sobre escorbuto* de Lind foi publicado em 1753, mas a inércia permaneceu na Marinha real por mais 40 anos. Em 1795, o Ministério da Marinha Britânica finalmente ordenou que fosse dada uma ração com suco concentrado de lima ou limão para todos os marinheiros britânicos. O escorbuto continuou a ser um problema em algumas partes do mundo até 1932, quando o cientista húngaro Albert Szent-



James Lind, 1716–1794; cirurgião da marinha, 1739–1748

-Györgyi, e W. A. Waugh e C. G. King da Universidade de Pittsburgh isolaram e sintetizaram o ácido ascórbico.

O ácido L-ascórbico (vitamina C) é um pó branco, inodoro e cristalino, totalmente solúvel em água e relativamente insolúvel em solventes orgânicos. O pó seco, longe da luz, é estável por um tempo considerável. O consumo diário apropriado dessa vitamina ainda é alvo de debates. O valor recomendado nos Estados Unidos é 90 mg (na Austrália e no Reino Unido, o recomendado é 60 mg; na Rússia, são recomendados 125 mg). Além das frutas cítricas e praticamente todas as outras frutas frescas, boas fontes de vitamina C incluem pimentas, tomates, batatas e brócolis. A vitamina C de frutas e vegetais é destruída por supercozimento ou longa estocagem.

Por que o ascorbato é tão necessário para uma boa saúde? De particular interesse é seu papel na formação de colágeno. Como dito no texto, o colágeno é formado por unidades repetidas de tripeptídeos Gly-X-Y, dos quais X e Y em geral são Pro e 4-Hyp – o derivado de prolina (4R)-L-hidroxiprolina, que tem um papel importante no entrelaçamento das fibras do colágeno e na manutenção de sua estrutura. O anel de prolina normalmente é encontrado em duas conformações, chamadas de C_γ-endo e C_γ-exo (Figura Q-1). A estrutura da hélice de colágeno necessita de resíduos de Pro na posição Y para estar na conformação C_γ-exo, sendo essa conformação favorecida pela substituição do hidroxil no C-4 da 4-Hyp. A estrutura de colágeno também requer que o resíduo Pro na posição X tenha uma conformação C_γ-endo, e a introdução

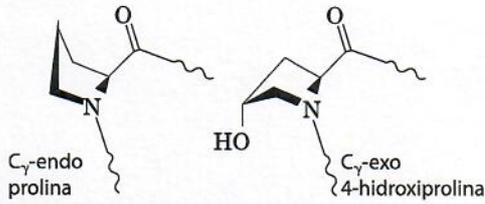


FIGURA Q-1 Conformação C_γ-endo da prolina e conformação C_γ-exo da 4-hidroxiprolina.

da 4-Hyp nessa posição pode desestabilizar a hélice. Na ausência de vitamina C, as células não conseguem hidroxilar a Pro da posição Y. Isso leva a uma instabilidade do colágeno e aos problemas no tecido conectivo observados no escorbuto.

A hidroxilação de resíduos específicos de Pro no pró-colágeno, o precursor do colágeno, requer a ação da enzima prolil 4-hidroxilase. Essa enzima (*M_r* = 240.000) apresenta-se como um tetrâmero α₂β₂ em todos os vertebrados. A atividade de hidroxilação da prolina está nas subunidades α. Cada subunidade α contém um átomo de ferro não hemínico (Fe²⁺), e a enzima faz parte de uma classe de hidroxilases que necessitam de α-cetoglutarato em suas reações.

Em uma reação normal da prolil-4-hidroxilase (Figura Q-2a), uma molécula de α-cetoglutarato e uma de O₂

se ligam à enzima. O α-cetoglutarato é oxidativamente decarboxilado para formar CO₂ e succinato. O átomo de oxigênio remanescente é então usado para hidroxilar o resíduo de Pro apropriado no pró-colágeno. Nenhum ascorbato é necessário nessa reação. Entretanto, a prolil-4-hidroxilase também catalisa uma decarboxilação oxidativa de α-cetoglutarato não acoplada à hidroxilação da prolina (Figura Q-2b). Durante essa reação, o Fe²⁺ do grupamento heme se oxida, inativando a enzima e impedindo a hidroxilação da prolina. O ascorbato consumido nessa reação é necessário para restaurar a atividade enzimática – pela redução do ferro do grupamento heme.

O escorbuto permanece sendo um problema ainda hoje, não somente em regiões remotas onde os alimentos nutritivos são escassos, mas, surpreendentemente, nos campi de universidades americanas. Os únicos vegetais consumidos pelos estudantes são aqueles em saladas refogadas, e os dias passam sem que esses jovens consumam frutas frescas. Um estudo de 1998, com 230 estudantes da Arizona State University, revelou que 10% deles tinham sérias deficiências de vitamina C, e dois estudantes tinham níveis de vitamina C tão baixos que provavelmente tinham escorbuto. Somente metade dos estudantes nessa pesquisa consumia a quantidade diária recomendada de vitamina C.

Então, consuma vegetais e frutas frescos.

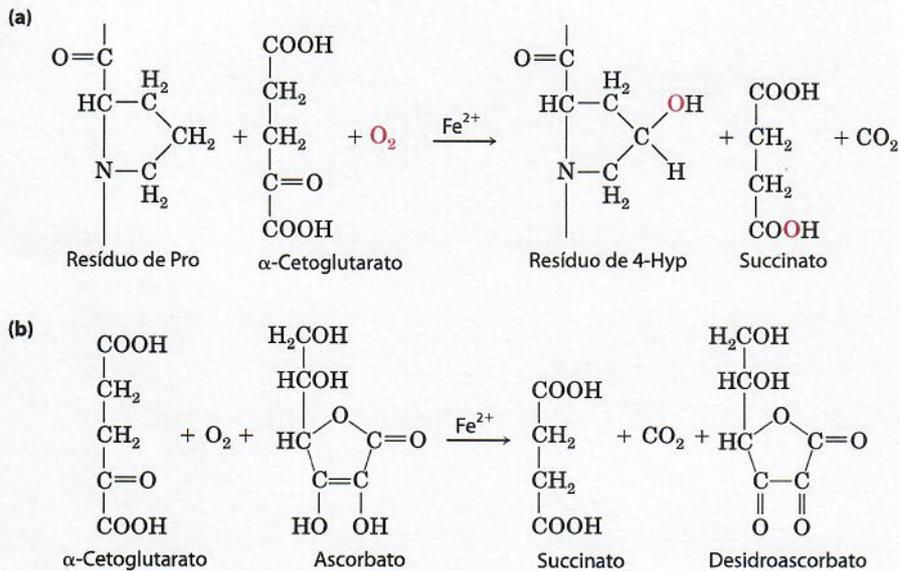


FIGURA Q-2 Reações catalisadas pela prolil-4-hidroxilase. (a) Reação usual, acoplada à hidroxilação da prolina, que não necessita de ascorbato. O destino dos dois átomos de oxigênio do O₂ é mostrado em vermelho. (b) Reação não acoplada, na qual o α-cetoglutarato sofre decarboxilação oxidativa sem a hidroxilação da prolina. O ascorbato é consumido estequiometricamente neste processo, à medida que é convertido a desidroascorbato, prevenindo a oxidação do Fe²⁺.

O empacotamento compacto das cadeias α na hélice tripla de colágeno garante uma resistência elástica maior do que aquela de uma barra de aço de mesma seção transversal. As fibrilas de colágeno (Figura 4-13) são conjuntos de estruturas supramoleculares formados por triplas hélices de moléculas de colágeno (algumas vezes chamadas de moléculas de tropocolágeno) associadas em uma grande variedade de formas para garantir diferentes graus de resistência elástica. As cadeias α e as fibrilas das moléculas de colágeno são interligadas por tipos incomuns de ligações covalentes envolvendo Lys, HyLys (5-hidroxisilina; ver Figura 3-8a), ou resíduos His presentes em algumas posições X e Y. Essas ligações criam resíduos de aminoácidos incomuns, como o desidro-hidroxisilina. O caráter cada vez mais rígido e quebradiço do tecido conectivo envelhecido resulta do acúmulo de ligações covalentes transversais nas fibrilas de colágeno.



Um mamífero tem mais de 30 variantes estruturais de colágeno, específicas para cada tecido e diferentes em sequência e função. Alguns defeitos genéticos na estrutura do colágeno em humanos ilustram a estreita relação entre a sequência de aminoácidos e a estrutura tridimensional nesta proteína. A osteogênese imperfeita é caracterizada pela formação anormal dos ossos em bebês; pelo menos oito variações dessa condição, em diferentes graus de severidade, ocorrem na população humana. A síndrome de Ehlers-Danlos é caracterizada por articulações soltas, e pelo menos seis variantes ocorrem em humanos. O compositor Niccolò Paganini (1782-1840) era famoso por sua destreza aparentemente impossível de tocar violino. Ele sofria de uma variante da síndrome Ehlers-Danlos que lhe rendeu efetivamente juntas duplas. Algumas variantes podem ser letais, outras causam problemas por toda vida.

Todas as variantes de ambas as condições resultam da substituição de um resíduo de aminoácido com um grupo R volumoso (como Cys ou Ser) por um resíduo de Gly de uma cadeia α em uma ou outra proteína colágena (um resíduo Gly diferente é substituído em cada uma das doenças). Essas substituições de um único resíduo têm um efeito catastrófico na função do colágeno, pois interrompem a repetição de Gly-X-Y, que garante ao colágeno sua estrutura helicoidal única. Dada sua importância na hélice tripla do colágeno (Figura 4-12d), a Gly não pode ser substituída por outro resíduo de aminoácido sem um efeito substancialmente deletério na estrutura do colágeno. ■

Fibroína da seda A fibroína, a proteína da seda, é produzida por insetos e aranhas. Suas cadeias polipeptídicas estão predominantemente na conformação β . A fibroína é rica em resíduos Ala e Gly, permitindo um grande empacotamento das folhas β e um arranjo entrelaçado dos grupos R (Figura

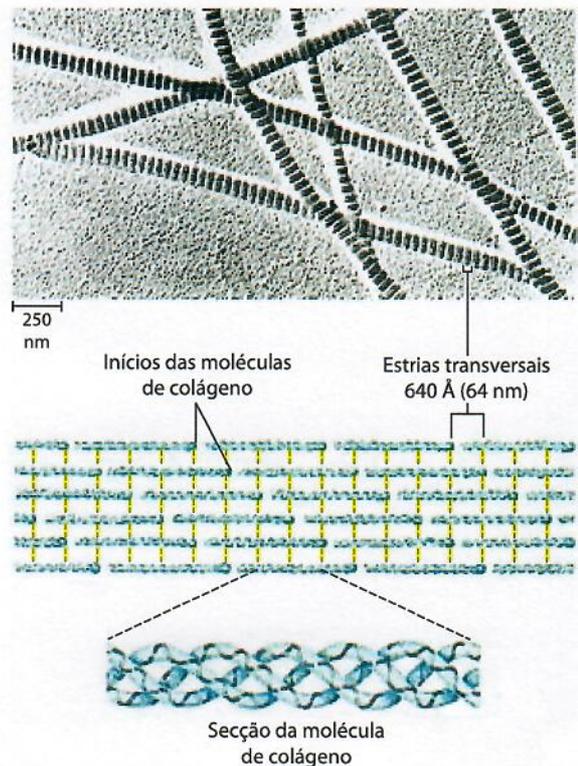


FIGURA 4-13 Estrutura das fibrilas de colágeno. O colágeno (M_r 300.000) é uma molécula em forma de bastão, com cerca de 3.000 Å de comprimento e apenas 15 Å de largura. Suas três cadeias α helicoidais entrelaçadas podem ter sequências diferentes; cada cadeia tem aproximadamente 1.000 resíduos de aminoácidos. As fibrilas são feitas de moléculas de colágeno alinhadas de forma escalonada e com ligações cruzadas que garantem resistência. O alinhamento específico e o grau de ligações transversais variam com o tecido e formam estrias transversais características, vistas por microscopia eletrônica. No exemplo mostrado aqui, o alinhamento dos grupos iniciais de cada quatro moléculas produz estrias distantes 640 Å (64 nm) umas das outras.

4-14). A estrutura global é estabilizada por extensivas ligações de hidrogênio entre todas as ligações peptídicas dos polipeptídeos de cada cadeia β , assim como pela otimização das interações de van der Waals entre as cadeias. A seda não é elástica porque sua conformação β já está bem estendida (Figura 4-6). Entretanto, a estrutura é flexível, pois as cadeias estão unidas por inúmeras interações fracas, em vez de ligações covalentes como ligações dissulfeto nas α -queratinas.

A diversidade estrutural reflete a diversidade funcional nas proteínas globulares

Em uma proteína globular, segmentos diferentes das cadeias polipeptídicas (ou de múltiplas cadeias polipeptídicas) se doam uns sobre os outros, gerando uma forma mais compacta do que a observada para as proteínas fibrosas (Figura 4-15). O enovelamento também garante a diversidade estrutural necessária às proteínas para realizar um grande leque de funções biológicas. Proteínas globulares incluem enzimas, proteínas transportadoras, proteínas motoras, proteínas reguladoras, imunoglobulinas e proteínas com muitas outras funções.

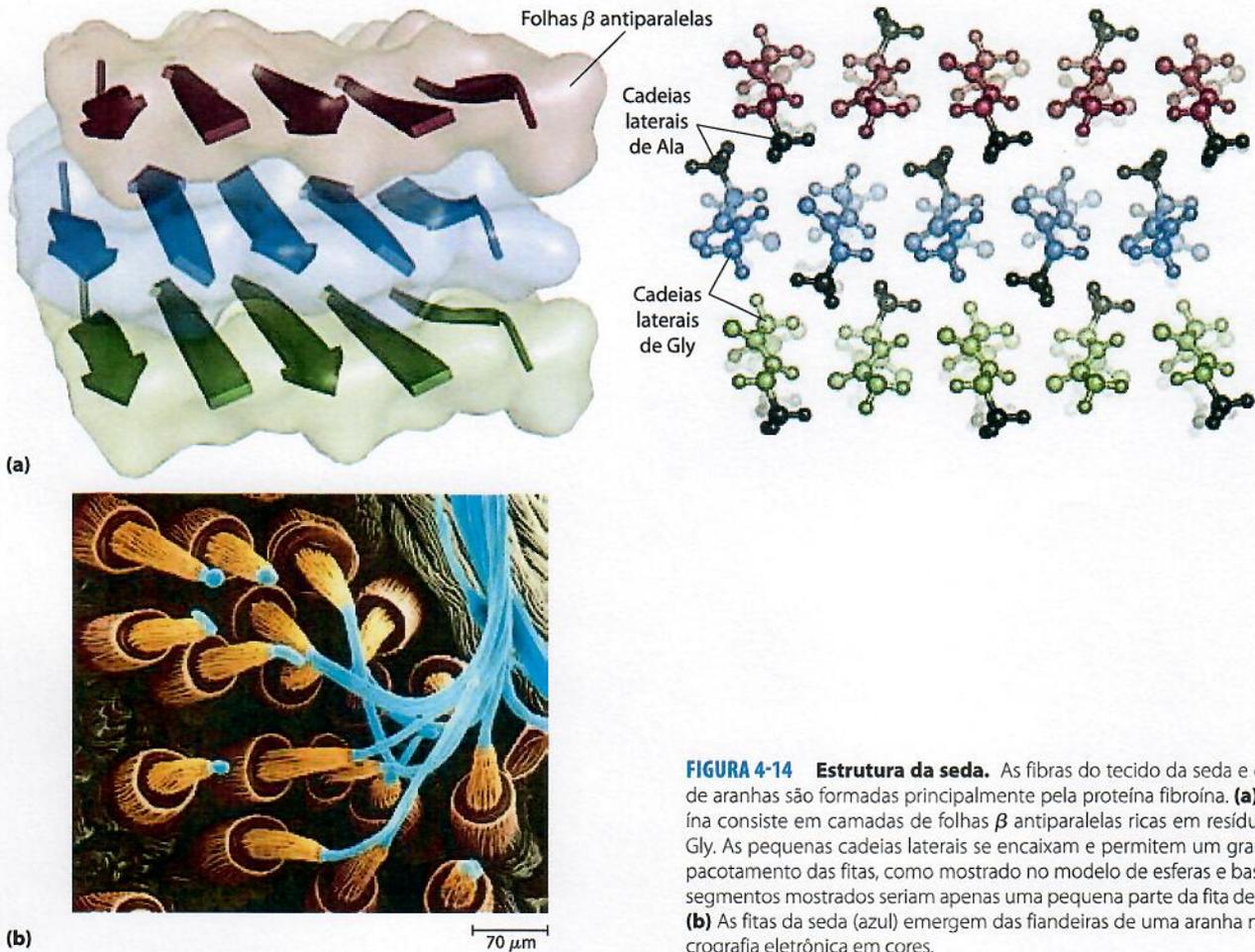


FIGURA 4-14 Estrutura da seda. As fibras do tecido da seda e das teias de aranhas são formadas principalmente pela proteína fibroína. **(a)** A fibroína consiste em camadas de folhas β antiparalelas ricas em resíduos Ala e Gly. As pequenas cadeias laterais se encaixam e permitem um grande empacotamento das fitas, como mostrado no modelo de esferas e bastões. Os segmentos mostrados seriam apenas uma pequena parte da fita de fibroína. **(b)** As fitas da seda (azul) emergem das fiandeiras de uma aranha nesta micrografia eletrônica em cores.

A presente discussão sobre proteínas globulares começa com os princípios vislumbrados a partir das primeiras estruturas de proteínas elucidadas. Segue-se uma detalhada descrição e classificação comparativa de subestruturas proteicas. Tais discussões são possíveis somente devido à grande quantidade de informações disponíveis na Internet, a partir de bancos de dados públicos, particularmente do Banco de Dados de Proteínas (Quadro 4-4).

Conformação β
 $2.000 \times 5 \text{ \AA}$

Hélice α
 $900 \times 11 \text{ \AA}$

Forma globular nativa
 $100 \times 60 \text{ \AA}$

FIGURA 4-15 As estruturas proteicas globulares são compactas e variadas. A albumina sérica humana (M_r , 64.500) tem 585 resíduos em uma única cadeia. São dadas aqui as dimensões aproximadas de uma de suas cadeias polipeptídicas se ela ocorresse em uma conformação β , ou em uma hélice α . Também é mostrado o tamanho da proteína em sua forma globular nativa, determinada por cristalografia por raios X; a cadeia polipeptídica deve estar com um enovelamento muito compacto para caber nestas dimensões.

A mioglobina forneceu os indícios iniciais sobre a complexidade da estrutura globular proteica

Arquitetura proteica – Estrutura terciária de proteínas globulares pequenas, II. Mioglobina O primeiro avanço no entendimento da estrutura tridimensional de uma proteína globular veio com os estudos de difração por raios X da mioglobina, feitos por John Kendrew e seus colaboradores, em 1950. A mioglobina é uma proteína ligadora de oxigênio relativamente pequena (M_r , 16.700) das células musculares. Ela funciona tanto para a estocagem de oxigênio quanto para facilitar a difusão do oxigênio nos tecidos musculares em contração. A mioglobina contém uma única cadeia polipeptídica de 153 resíduos de aminoácidos de sequência conhecida e um único grupo ferro-protoporfirina, ou heme. O mesmo grupo heme encontrado na mioglobina é encontrado na hemoglobina, a proteína ligadora de oxigênio dos eritrócitos, sendo responsável pela coloração vermelha amarronzada tanto da mioglobina quanto da hemoglobina. A mioglobina é particularmente abundante nos músculos de mamíferos mergulhadores, como as baleias, as focas e os botos – tão abundante que os músculos destes animais são marrons. A estocagem e a distribuição do oxigênio pela mioglobina do músculo permitem que animais que mergulham permaneçam submersos por um longo período. As

QUADRO 4-4 O banco de dados de proteínas

Atualmente, o número de estruturas proteicas tridimensionais conhecidas está na casa das dezenas de milhares e mais do que dobra a cada dois anos. Essa abundância de informações está revolucionando o entendimento sobre a estrutura de proteínas, a relação estrutura-atividade e as rotas evolutivas pelas quais as proteínas chegaram ao seu estado atual, que podem ser vistas nas semelhanças que aparecem entre as famílias à medida que os bancos de dados de proteínas são examinados e classificados. Um dos recursos mais importantes disponíveis aos bioquímicos é o **Protein Data Bank (PDB)**, ou banco de dados de proteínas; www.pdb.org.

O PDB é um arquivo de estruturas tridimensionais de macromoléculas biológicas determinadas experimentalmente, contendo quase todas as estruturas macromoleculares (proteínas, RNA, DNA, etc.) elucidadas até o momento. A cada estrutura é atribuído um código de identificação (um código de quatro caracteres, chama-

do de PDB ID). Tais identificadores são fornecidos nas legendas das figuras para cada uma das estruturas derivadas do PDB ilustradas nesse texto, de forma que os estudantes e os professores podem explorar essas mesmas estruturas. Os arquivos de dados no PDB descrevem as coordenadas espaciais de cada átomo cuja posição foi determinada (muitas das estruturas catalogadas não estão completas). Arquivos de dados adicionais fornecem informações de como as estruturas foram determinadas e sua precisão. As coordenadas atômicas podem ser convertidas em uma imagem da macromolécula com a ajuda de programas de visualização de estruturas. Os estudantes são incentivados a acessar o PDB e explorar as estruturas usando programas de visualização referenciados no banco de dados. Os arquivos de estruturas macromoleculares também podem ser baixados e explorados em seu computador, usando programas livres, tal como o Jmol.

atividades da mioglobina e de outras moléculas de globina são investigadas em detalhe no Capítulo 5.

A **Figura 4-16** mostra diversas representações estruturais da mioglobina, ilustrando como a cadeia polipeptídica é dobrada nas três dimensões – sua estrutura terciária. O grupo vermelho envolvido pela proteína é o heme. O esqueleto da molécula de mioglobina consiste em oito segmentos de hélices α relativamente retas, interrompidas por dobras, algumas delas sendo voltas β . A hélice α mais longa tem 23 resíduos de aminoácidos, a mais curta apenas 7. Todas as hélices são voltadas para a direita. Mais de 70% dos resíduos de mioglobina estão nessas regiões α -helicoidais. As análises por raios X revelaram a posição precisa de cada um dos grupos R, que preenchem quase todo o espaço dentro da cadeia dobrada, não ocupada por átomos do esqueleto.

Várias conclusões importantes podem ser tiradas a partir da estrutura da mioglobina. As posições das cadeias late-

rais dos aminoácidos refletem uma estrutura cuja estabilidade é resultante de suas interações hidrofóbicas. A maioria dos grupos R hidrofóbicos está no interior da molécula, escondidos da exposição à água. Exceto dois resíduos, todos os outros grupos R polares estão na superfície externa da molécula e estão hidratados. A molécula de mioglobina é tão compacta que em seu interior só há lugar para quatro moléculas de água. Esse denso núcleo hidrofóbico é típico das proteínas globulares. A fração de espaço ocupado pelos átomos em um líquido orgânico é de 0,4 a 0,6. Em uma proteína globular, essa fração é de cerca de 0,75, comparável com aquela do cristal (em um cristal típico, a fração é de 0,70 a 0,78, próximo do máximo teórico). Nesse ambiente compacto, as interações fracas são fortalecidas e reforçam umas às outras. Por exemplo, as cadeias apolares no núcleo estão tão próximas que as interações de curto alcance do tipo van der Waals têm uma contribuição significativa na estabilização das interações hidrofóbicas.

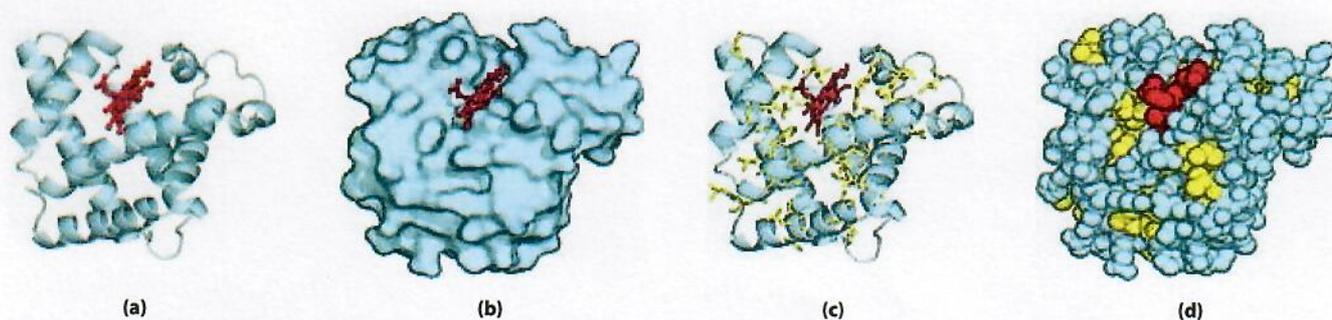


FIGURA 4-16 Estrutura terciária da mioglobina de cachalote. (PDB ID 1MBO) A orientação da proteína é igual de (a) a (d); o grupo heme é mostrado em vermelho. Além de ilustrar a estrutura da mioglobina, esta figura mostra exemplos de diversas formas de apresentar a estrutura de uma proteína. **(a)** Esqueleto polipeptídico na representação de fita, introduzido por Jane Richardson, destacando as estruturas secundárias. As regiões de hélice α são evidentes. **(b)** Imagem da superfície da proteína; ela é útil na visualização de

fechas, onde outras moléculas podem se ligar à proteína. **(c)** Representação da estrutura em fita, incluindo as cadeias laterais (em amarelo) dos resíduos hidrofóbicos Leu, Ile, Val e Phe. **(d)** Modelo de volume atômico com todas as cadeias laterais dos aminoácidos. Cada átomo está representado por uma esfera que cobre o seu raio de van der Waals. Os resíduos hidrofóbicos são novamente mostrados em amarelo; a maioria está no interior da proteína, e por isso não está visível.

A dedução da estrutura da mioglobina confirmou algumas expectativas e introduziu alguns elementos novos à estrutura secundária. Como predito por Pauling e Corey, todas as ligações peptídicas estão em configuração planar *trans*. As hélices α da mioglobina forneceram a primeira evidência experimental direta da existência desse tipo de estrutura secundária. Três dos quatro resíduos Pro são encontrados em dobras. O quarto resíduo Pro ocorre em uma hélice α , onde ele cria uma torção necessária para o firme empacotamento da hélice.

O grupo heme planar repousa sobre uma fenda na molécula de mioglobina. O átomo de ferro no centro do grupo heme apresenta duas posições de ligação (coordenação) perpendiculares ao plano do heme (Figura 4-17). Uma delas se liga ao grupo R de um resíduo de His na posição 93; a outra corresponde ao local onde a molécula de O_2 se liga. Dentro dessa fenda, a acessibilidade do grupo heme ao solvente é bem restrita. Isso é importante para a função, pois grupos heme livres em uma solução oxigenada são rapidamente oxidados do estado ferroso (Fe^{2+}), que é ativo na ligação reversível com o O_2 , para o estado férrico (Fe^{3+}), que não se liga ao O_2 .

À medida que as diferentes estruturas da mioglobina foram resolvidas, os pesquisadores foram capazes de observar as mudanças estruturais que acompanham a ligação do oxigênio ou de outra molécula, e, assim, pela primeira vez, de entender a correlação entre a estrutura de uma proteína e sua função. Centenas de proteínas são agora submetidas a análises semelhantes. Hoje, a espectroscopia de ressonância magnética nuclear (RMN) e outras técnicas suplementam os dados de difração de raios X, fornecendo mais informações sobre as estruturas das proteínas (Quadro 4-5). Adicionalmente, o sequenciamento de DNA genômico de muitos organismos (Capítulo 9) identificou milhares de genes que codificam proteínas de sequências conhecidas, mas ainda de função desconhecida; esse trabalho continua a passos largos.

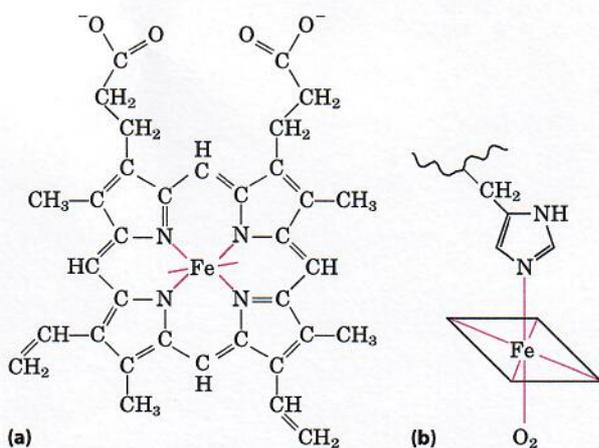


FIGURA 4-17 O grupo heme. Este grupo está presente na mioglobina, hemoglobina, citocromos e muitas outras proteínas (as proteínas heme). **(a)** O heme consiste em uma estrutura de anel orgânico complexo, a protoporfirina, que liga um átomo de ferro no seu estado ferroso (Fe^{2+}). O átomo de ferro tem seis ligações coordenadas, quatro delas no plano da molécula da porfirina (e ligadas a ela) e duas perpendiculares a ela. **(b)** Na mioglobina e na hemoglobina, uma das ligações perpendiculares está ligada a um átomo de nitrogênio de um resíduo His. A outra está "vaga", servindo de sítio de ligação para a molécula de O_2 .

As proteínas globulares têm uma diversidade de estruturas terciárias

A partir do que se sabe sobre estruturas terciárias de centenas de proteínas globulares, é claro que a mioglobina ilustra apenas uma das diversas formas em que a cadeia polipeptídica pode se dobrar. A Tabela 4-4 mostra as proporções de hélices α e conformações β (expressas em porcentagem de resíduos em cada um dos tipos) em diversas proteínas globulares pequenas e de cadeia única. Cada uma dessas proteínas tem uma estrutura diferente, adaptada à sua função biológica particular, mas como um todo elas compartilham diversas propriedades importantes com a mioglobina. Cada uma delas está dobrada de forma compacta e, em cada caso, as cadeias laterais hidrofóbicas dos aminoácidos estão orientadas na direção do centro da proteína (longe da água), e suas cadeias laterais hidrofílicas se encontram na superfície. As estruturas também são estabilizadas por inúmeras ligações de hidrogênio e algumas interações iônicas.

Para os estudantes iniciantes, as estruturas terciárias muito complexas das proteínas globulares – algumas muito maiores do que a mioglobina – são mais facilmente acessadas concentrando-se nos padrões estruturais comuns, recorrentes em diferentes proteínas, geralmente não relacionadas. A estrutura tridimensional de uma proteína globular típica pode ser considerada um conjunto de segmentos polipeptídicos em hélices α e conformações β , ligados por segmentos conectores. A estrutura pode ser, então, definida pela forma como os segmentos estão empilhados uns sobre os outros, e como estão arranjados os segmentos que os conectam.

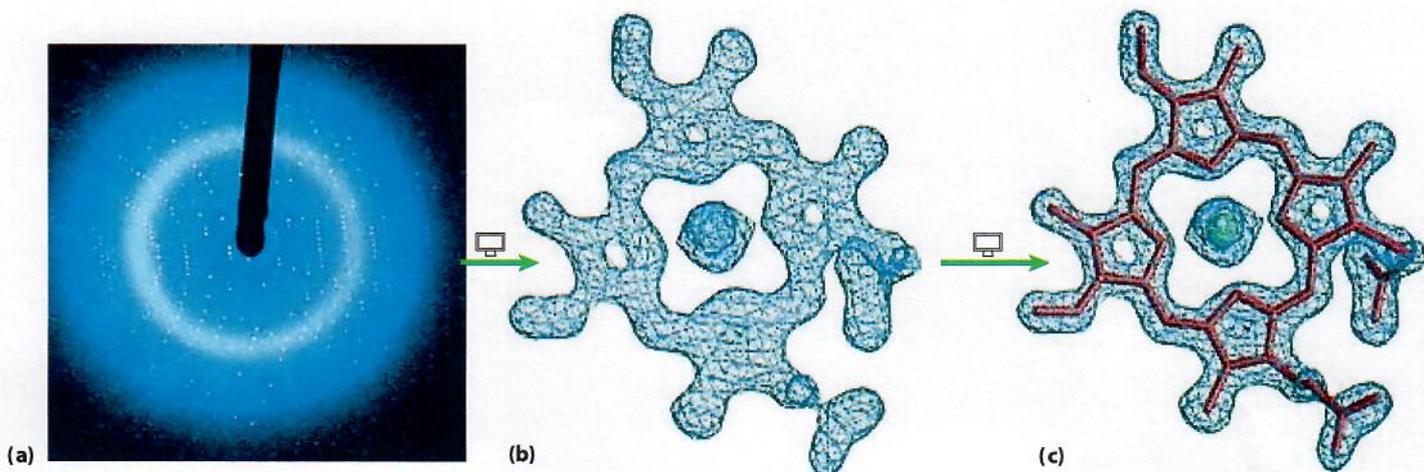
Para entender uma estrutura tridimensional completa, é preciso analisar seu padrão de enovelamento. Primeiro, definem-se dois importantes termos que descrevem os padrões de estruturas de proteínas ou elementos de uma cadeia polipeptídica, seguindo então para as regras de enovelamento.

TABELA 4-4 Proporção aproximada de hélices α e conformações β em algumas proteínas de cadeia única

Proteína (número total de resíduos)	Resíduos (%)*	
	Hélice α	Conformação β
Quimotripsina (247)	14	45
Ribonuclease (124)	26	35
Carboxipeptidase (307)	38	17
Citocromo <i>c</i> (104)	39	0
Lisozima (129)	40	12
Mioglobina (153)	78	0

Fonte: Dados de Cantor, C.R. e Schimmel, P.R. (1980) *Biophysical Chemistry*, Part I: *The Conformation of Biological Macromolecules*, p. 100, W.H. Freeman and Company, New York.

*As proporções das cadeias polipeptídicas que não fazem parte de hélices α e conformações β consistem em curvaturas e espirais irregulares ou regiões estendidas. Segmentos de hélice α e conformação β algumas vezes diferem um pouco de suas dimensões e geometrias usuais.

QUADRO 4-5 MÉTODOS Métodos de determinação da estrutura tridimensional de uma proteína**Difração de raios X**

O espaçamento dos átomos em um retículo cristalino pode ser determinado pela medida da localização e da intensidade dos pontos produzidos por um feixe de raios X, de um dado comprimento de onda, em um filme fotográfico, depois desse feixe ser difratado pelos elétrons dos átomos. Por exemplo, a análise por raios X de cristais de cloreto de sódio mostra que os íons Na^+ e Cl^- estão dispostos em um arranjo cúbico simples. O espaçamento de diferentes tipos de átomos em moléculas orgânicas complexas, até mesmo as grandes como as proteínas, também pode ser analisado por métodos de difração de raios X. Entretanto, essa técnica de análise de cristais de moléculas complexas é muito mais trabalhosa do que a de cristais de sais simples. Quando o padrão de repetição do cristal é uma molécula tão grande quanto uma proteína, o número de átomos da molécula resulta em milhares de pontos de difração que precisam ser analisados por computador.

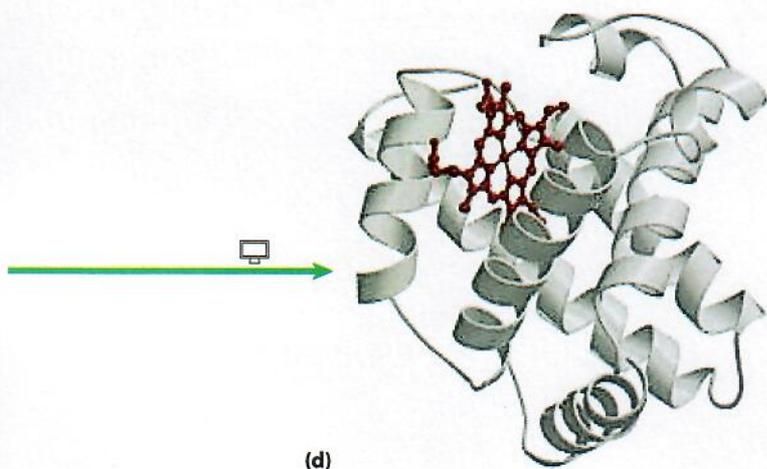
Considere como as imagens são geradas em um microscópio óptico. A luz de uma fonte pontual é focalizada em um objeto. O objeto espalha as ondas de luz, e estas ondas espalhadas são reagrupadas por uma série de lentes para gerar uma imagem aumentada do objeto. O tamanho mínimo do objeto cuja estrutura pode ser determinada desta forma – isto é, o poder de resolução do microscópio – é determinado pelo comprimento de onda de luz, neste caso, luz visível, com um comprimento de onda entre 400 e 700 nm. Objetos menores do que a metade do comprimento de onda da luz incidente não têm resolução. Para resolver objetos tão pequenos quanto as proteínas, é necessário usar raios X, com comprimentos de onda na faixa de 0,7 a 1,5 Å (0,07 a 0,15 nm). Entretanto, não há lentes que possam reagrupar os raios X para formar a imagem. Em vez disso, o padrão de difração dos raios X é coletado diretamente, e a imagem é reconstruída por técnicas matemáticas.

A quantidade de informação obtida em uma cristalografia por raios X depende do grau de organização estru-

tural da amostra. Alguns parâmetros importantes foram obtidos nos primeiros estudos de padrão de difração de proteínas fibrosas arranjadas de forma regular no cabelo e na lã. Entretanto, os feixes ordenados formados pelas proteínas fibrosas não são cristais – as moléculas estão alinhadas lado a lado, mas não estão todas orientadas na mesma direção. Informações estruturais tridimensionais mais detalhadas sobre proteínas necessitam de um cristal de proteína altamente ordenado. As estruturas de muitas proteínas ainda permanecem desconhecidas, simplesmente porque se mostraram difíceis de cristalizar. Os profissionais comparam a preparação de cristais com manter unida uma pilha de bolas de boliche com uma fita de celofane.

Operacionalmente, existem vários passos em uma análise estrutural por raios X (Figura Q-1). Um cristal é colocado em um feixe de raios X, entre a fonte de raios X e o detector, e um conjunto de pontos, chamados de reflexões, é gerado. Os pontos são gerados pela difração do feixe de raios X, e cada átomo da molécula faz uma contribuição para cada ponto. Um mapa de densidade eletrônica da proteína é reconstruído a partir do padrão total de difração, por uma técnica matemática chamada de transformada de Fourier. O computador age, na verdade, como “lente computacional”. Um modelo da estrutura é então construído de acordo com o mapa de densidade eletrônica.

John Kendrew verificou que o padrão de difração de raios X da mioglobina cristalina (isolada de músculos do cachalote) é muito complexo, com cerca de 25.000 reflexões. A análise computacional dessas reflexões ocorreu em etapas. A resolução aumentou a cada etapa, até que em 1959 as posições de quase todos os átomos diferentes de hidrogênio foram determinadas. A sequência de aminoácidos da proteína, obtida por análise química, foi consistente com o modelo molecular. A estrutura de milhares de proteínas, a maioria delas mais complexa que a mioglobina, foi desde então determinada com um nível de resolução semelhante.



(d)

É claro que o ambiente físico em um cristal não é idêntico ao ambiente em solução ou em uma célula viva. O cristal impõe uma média temporal e espacial na estrutura deduzida de sua análise, e os estudos de difração de raios X fornecem poucas informações sobre os movimentos moleculares no interior da proteína. A conformação das proteínas em um cristal pode, a princípio, ser afetada por fatores não fisiológicos como contatos proteína-proteína dentro do cristal. Entretanto, quando estruturas derivadas da análise de cristais são comparadas com informações estruturais obtidas de outras formas (como RMN, descrita a seguir), as estruturas derivadas do cristal quase sempre representam a conformação funcional da proteína. A cristalografia por raios X pode ser aplicada com sucesso a proteínas grandes demais para serem estruturalmente analisadas por RMN.

Ressonância magnética nuclear

Uma vantagem dos estudos de ressonância magnética nuclear (RMN) é que eles são realizados com as macromoléculas em solução, enquanto a cristalografia é limitada a moléculas que podem ser cristalizadas. A RMN também pode esclarecer o lado dinâmico da estrutura da proteína, incluindo mudanças conformacionais, enovelamento proteico e interação com outras moléculas.

A RMN é a manifestação do momento angular do *spin* nuclear, uma propriedade quântico-mecânica dos núcleos atômicos. Somente alguns átomos, incluindo ^1H , ^{13}C , ^{15}N , ^{19}F e ^{31}P , apresentam o tipo de *spin* nuclear que gera sinal na RMN. O *spin* nuclear gera um dipolo magnético. Quando um campo magnético estático e forte é aplicado em uma solução contendo um único tipo de macromolécula, os dipolos magnéticos são alinhados no campo em uma das duas orientações, paralela (baixa energia) ou antiparalela (alta energia). Um pequeno ($\sim 10 \mu\text{s}$) pulso de energia eletromagnética de frequência adequada (a frequência de ressonância, que está dentro da faixa de radiofrequência) é aplicado, em certos ângulos, aos núcleos alinhados no campo magnético.

FIGURA Q-1 Etapas da determinação da estrutura da mioglobina de cachalote via cristalografia por raios X. **(a)** O padrão da difração por raios X é gerado a partir de um cristal da proteína. **(b)** Os dados obtidos pelo padrão de difração são utilizados para calcular um mapa de densidade eletrônica tridimensional. Somente a densidade eletrônica do heme, uma pequena parte da estrutura, é mostrada aqui. **(c)** As regiões de maior densidade eletrônica revelam a localização do núcleo atômico, esta informação é utilizada para construir a estrutura final. Aqui, a estrutura do heme é modelada em seu mapa de densidade eletrônica. **(d)** A estrutura completa da mioglobina de cachalote, incluindo o heme (PDB ID 2MBW).

Alguma energia é absorvida à medida que o núcleo muda para o estado de maior energia, e o espectro de absorção resultante contém informações sobre a identidade do núcleo e seu ambiente químico nas imediações. Os dados de diversos desses experimentos com a amostra são reunidos, aumentando a relação sinal-ruído, e é gerado um espectro de RMN, como o da Figura Q-2.

O ^1H é particularmente importante nos experimentos de RMN, pois é altamente sensível e naturalmente abundante. Para macromoléculas, o espectro de RMN de ^1H pode ser bem complicado. Até mesmo uma pequena proteína pode conter centenas de átomos de ^1H , normalmente gerando um espectro de RMN unidimensional muito complexo para ser analisado. A análise estrutural de proteínas se tornou possível com o advento das técnicas de RMN bidimensionais (Figura Q-3). Esses métodos permitem a medida do acoplamento, dependente da distância, dos *spins* nucleares de átomos espacialmente próximos (o efeito Overhauser nuclear [NOE], em um método chamado de NOESY) ou o acoplamento dos *spins* nucleares de átomos conectados por ligações covalentes (espectroscopia de correlação total, ou TOCSY).

A tradução de um espectro de RMN bidimensional em uma estrutura tridimensional completa é um processo muito trabalhoso. Os sinais de NOE fornecem algumas informações sobre as distâncias entre cada um dos áto-

(Continua na próxima página)

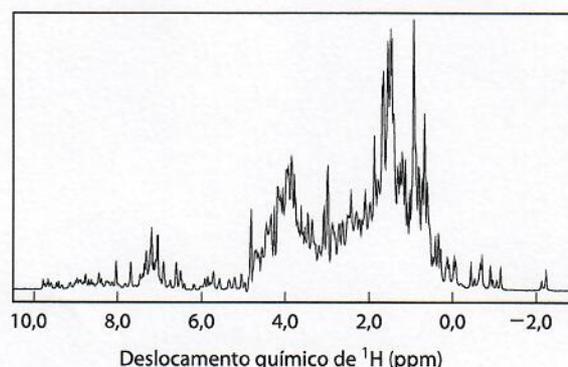


FIGURA Q-2 Espectro de RMN unidimensional da globina do verme-de-sangue marinho (*Glycera* sp.). Esta proteína e a mioglobina de cachalote são análogos estruturais muito próximos, pertencendo à mesma família estrutural e compartilhando a função de transporte de oxigênio.

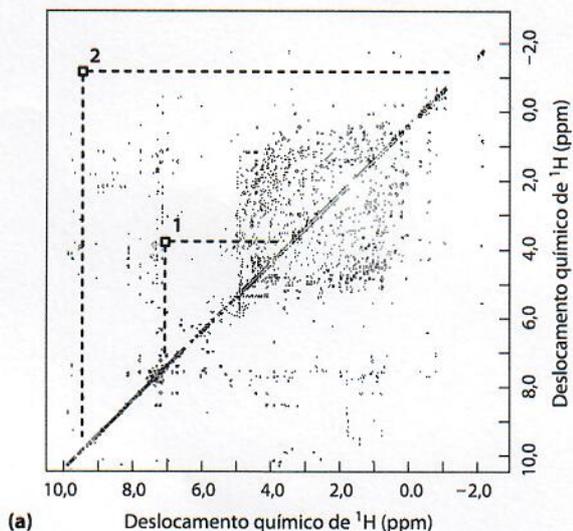
QUADRO 4-5 MÉTODOS Métodos de determinação da estrutura tridimensional de uma proteína (*Continuação*)

mos, mas para esses valores de distâncias serem úteis, os átomos que originam cada sinal devem ser identificados. Os experimentos complementares de TOCSY podem ajudar na identificação de quais sinais de NOE refletem átomos que estão ligados por ligações covalentes. Certos padrões de sinais de NOE têm sido associados com estruturas secundárias como hélices α . A engenharia genética (Capítulo 9) pode ser usada para preparar proteínas que contenham isótopos raros como ^{13}C ou ^{15}N . Os novos sinais de RMN produzidos por esses átomos e o acoplamento com os sinais de ^1H resultantes dessa substituição auxiliam na designação de cada um dos sinais de ^1H NOE. O processo também é facilitado pelo conhecimento da sequência do polipeptídeo.

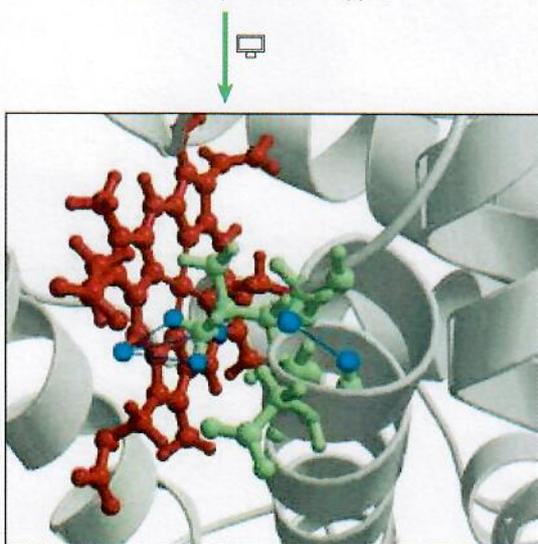
Para gerar a representação de uma estrutura tridimensional, os pesquisadores colocam em um computador as restrições de distância, junto com as restrições geométricas conhecidas como quiralidade, raios de van

der Waals e distância e ângulo de ligação. O computador gera, então, um conjunto de estruturas altamente relacionadas que representa um conjunto de conformações consistentes com os valores observados no NOE (Figura Q-3c). A incerteza das estruturas geradas por RMN é em parte um reflexo das vibrações moleculares dentro de uma proteína, discutidas em mais detalhes no Capítulo 5. O erro experimental normal também influi.

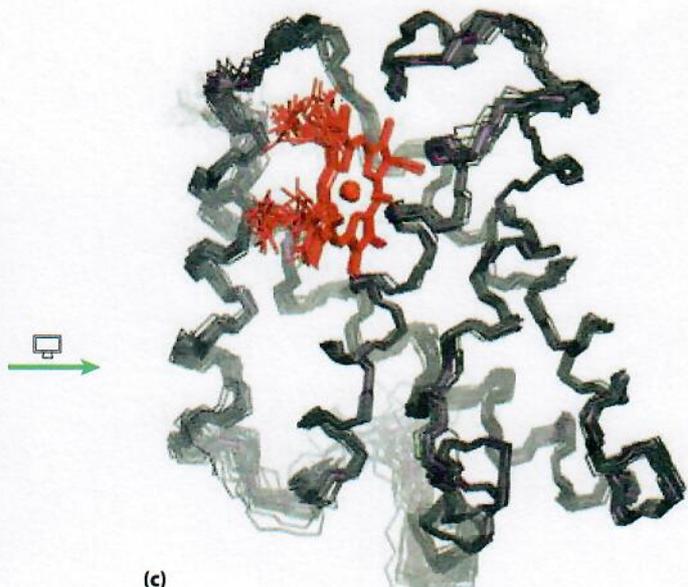
As estruturas proteicas determinadas tanto por cristalografia por raios X quanto por RMN geralmente são bem coincidentes. Em alguns casos, a localização exata de algumas cadeias laterais de aminoácidos específicos da superfície da proteína é diferente, normalmente devido aos efeitos relacionados com o empacotamento entre proteínas adjacentes no cristal. Juntas, as duas técnicas são as responsáveis pelo rápido aumento da disponibilidade de informações estruturais sobre as macromoléculas de células vivas.



(a)



(b)



(c)

FIGURA Q-3 Uso de RMN bidimensional para determinar a estrutura tridimensional da globina, a mesma proteína usada para gerar os dados da Figura Q-2. A diagonal em um espectro de RMN bidimensional é equivalente ao espectro unidimensional. Os picos fora da diagonal são sinais de NOE gerados pelas interações de curto alcance dos átomos de ^1H , que podem gerar sinais bem distantes em um espectro unidimensional. Duas destas interações são identificadas em (a) e suas identidades estão mostradas em linhas azuis em (b) (PDB ID 1VRF). São desenhadas três linhas para a interação 2 entre um grupo metil da proteína e um hidrogênio do heme. O grupo metil gira rapidamente de forma que cada um dos seus três hidrogênios contribui igualmente para a interação e para o sinal de RMN. Tal informação é usada para determinar a estrutura tridimensional completa (PDB ID 1VRE), como em (c). As múltiplas linhas que mostram o esqueleto da proteína em (c) representam o conjunto de estruturas consistentes com os dados de distância da RMN. A semelhança estrutural com a mioglobina (Figura Q-1) é evidente. As proteínas estão orientadas da mesma forma nas duas figuras.

O primeiro termo é **motivo**, também chamado de **estrutura supersecundária** ou **enovelamento**. Um motivo ou enovelamento é um padrão de enovelamento identificável, envolvendo dois ou mais elementos da estrutura secundária e a conexão (ou conexões) entre eles. Um motivo pode ser muito simples, tal como dois elementos de estrutura secundária dobrados um sobre o outro e representa apenas uma pequena parte de uma proteína. Um exemplo é uma **alça β - α - β** (Figura 4-18a). Um motivo também pode ter uma estrutura bem elaborada, envolvendo um grande número de segmentos proteicos dobrados juntos, como o barril β (Figura 4-18b). Em alguns casos, um único e grande motivo pode abranger toda a proteína. Os termos “motivo” e “enovelamento” são frequentemente usados como sinônimos, embora “enovelamento” seja mais comumente aplicado aos padrões de enovelamento ligeiramente mais complexos. Os termos incluem qualquer padrão favorável de enovelamento, sendo útil na descrição desses padrões. O segmento definido como um motivo pode, ou não, ser estável independentemente. Já foi mencionado um motivo bem estudado, a espiral enrolada da estrutura da α -queratina, também encontrado em outras proteínas. O arranjo característico de oito hélices α na mioglobina é reproduzido em todas as globinas e é chamado de enovelamento de globina. Observe que um motivo não é um elemento estrutural hierárquico que fica entre a estrutura secundária e a terciária. Ele é simplesmente um padrão de enovelamento. O termo sinônimo “estrutura supersecundária” pode, portanto, causar confusão, pois sugere uma hierarquia.

O segundo termo que descreve padrões estruturais é o **domínio**. Um domínio, como definido por Jane Richardson em 1981, é uma parte da cadeia polipeptídica que é independentemente estável ou pode se movimentar como uma entidade isolada em relação ao resto da proteína. Polipeptídeos com mais de algumas centenas de resíduos de aminoácidos normalmente se dobram em dois ou mais domínios, algumas vezes com diferentes funções. Muitas vezes, um domínio de uma proteína grande irá conservar sua estrutura tridimensional nativa mesmo quando separado (p. ex., por uma clivagem proteolítica) do resto da cadeia polipeptídica. Em uma proteína com múltiplos domínios, cada um

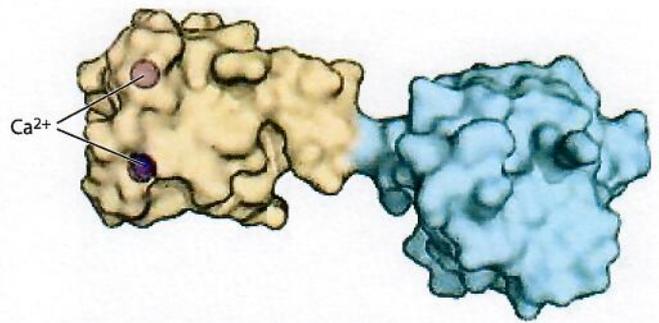


FIGURA 4-19 Domínios estruturais do polipeptídeo troponina C. (PDB ID 4TNC) Esta proteína ligadora de cálcio encontrada no músculo tem dois domínios ligadores de cálcio separados, mostrados aqui em marrom e em azul.

deles pode aparecer como uma porção globular diferente (Figura 4-19); mais frequentemente, extensos contatos entre os domínios tornam difícil o discernimento entre eles. Domínios diferentes em geral têm funções diferentes, como a ligação de pequenas moléculas ou interação com outras proteínas. Proteínas pequenas normalmente têm somente um domínio (o domínio é a proteína).

O enovelamento dos polipeptídeos é sujeito a uma série de limitações físicas e químicas, e várias regras foram propostas a partir de estudos de padrões comuns de enovelamento proteico.

1. As interações hidrofóbicas dão uma grande contribuição para a estabilidade da estrutura de proteínas. O ocultamento dos grupos R dos aminoácidos hidrofóbicos, de modo a excluir a água, necessita de pelo menos duas camadas de estrutura secundária. Motivos simples como a alça β - α - β (Figura 4-18a) criam essas camadas.
2. Quando ocorrem juntas em uma proteína, as hélices α e as folhas β geralmente são encontradas em camadas estruturais diferentes. Isso porque o segmento do esqueleto polipeptídico em conformação β (Figura 4-6) não faz, com facilidade, ligações de hidrogênio com uma hélice α adjacente a ele.
3. Segmentos adjacentes na sequência de aminoácidos normalmente se posicionam de forma adjacente na estrutura dobrada. Segmentos distantes do polipeptídeo podem se aproximar na estrutura terciária, mas não é a regra.
4. A conformação β é mais estável quando os segmentos individuais são levemente torcidos para a direita. Isso influencia tanto o arranjo das folhas β derivado dos segmentos supertorcidos quanto a forma da conexão do polipeptídeo entre elas. Duas folhas β paralelas, por exemplo, devem ser conectadas por uma fita cruzada (Figura 4-20a). Em princípio, essa fita pode ter uma conformação no sentido horário ou anti-horário, mas em proteínas o mais comum é no sentido horário. As conexões voltadas para a direita tendem a ser mais curtas do que as conexões voltadas para esquerda, e tendem a dobrar-se com um ângulo menor, o que as torna mais fáceis de serem formadas. A torção de folhas

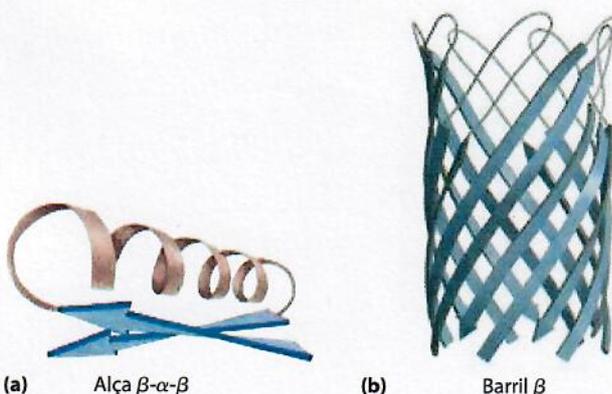


FIGURA 4-18 Motivos estruturais. (a) Um motivo simples, a alça β - α - β . (b) Um motivo mais elaborado, o barril β . Este barril β é o único domínio da α -hemolisina (toxina que mata a célula, criando um orifício em sua membrana) da bactéria *Staphylococcus aureus* (derivada do PDB ID 7AHL).

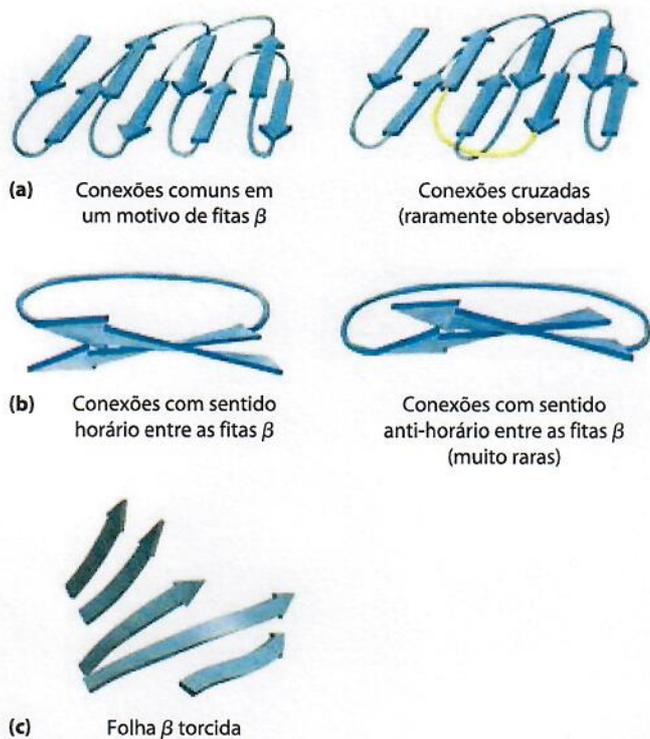


FIGURA 4-20 Padrões de enovelamento estável em proteínas. (a) Conexões entre fitas β em arranjos de folhas β . As fitas aqui são visualizadas a partir de uma extremidade, sem torção. As linhas mais grossas representam as conexões nas extremidades mais próximas do observador; as linhas mais finas representam as conexões nas extremidades mais distantes das fitas β . As conexões em uma das extremidades (p. ex., aquelas próximas ao observador) raramente se cruzam umas com as outras. Um exemplo de tal cruzamento raro está ilustrado na fita amarela da estrutura a direita. (b) Devido à torção para a direita nas fitas β , as conexões entre as fitas em geral são no sentido horário. As conexões no sentido anti-horário devem assumir ângulos mais agudos e são mais difíceis de serem formadas. (c) Esta folha β torcida pertence a um domínio da fotoliase (proteína que repara certos tipos de danos ao DNA) de *E. coli* (derivado do PDB ID 1DNP). As alças de conexão foram removidas para que se foque no enovelamento da folha β .

β também resulta em uma torção característica das estruturas formadas por vários desses segmentos, como pode ser observado no barril β (Figura 4-18b) e na folha β torcida (Figura 4-20c), que forma o núcleo de várias estruturas maiores.

Seguindo essas regras, motivos complexos podem ser construídos a partir daqueles mais simples. Por exemplo, uma série de alças β - α - β arranjadas de tal forma que as fitas β de um barril formem um motivo particularmente comum e estável, o **barril α/β** (Figura 4-21). Nessa estrutura, cada segmento β paralelo é conectado ao seu vizinho por um segmento α -helicoidal. Todas as conexões são voltadas para direita. O barril α/β é encontrado em diversas enzimas, geralmente com um sítio de ligação (de um cofator ou substrato) na forma de uma fenda próxima a uma das extremidades do barril. Observe que domínios com padrão semelhante de enovelamento são ditos ter o mesmo motivo, mesmo que suas hélices α e fitas β constituintes possam diferir em tamanho.

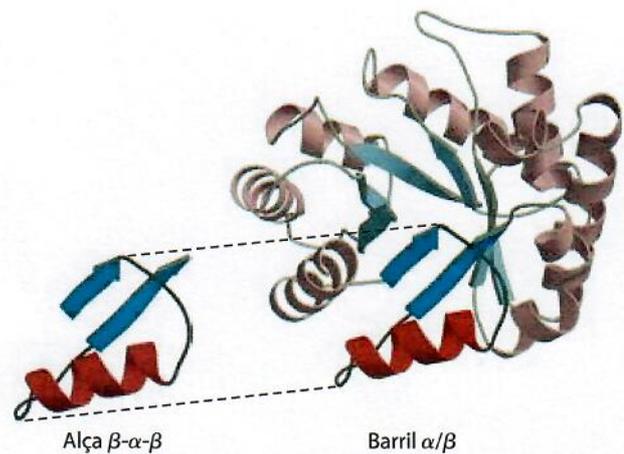


FIGURA 4-21 A construção de domínios maiores a partir dos menores. O barril α/β é um domínio que ocorre comumente, sendo formado por repetições do motivo alça β - α - β . Este barril α/β é um domínio da piruvato-cinase (enzima glicolítica) de coelho (derivada do PDB ID 1PKN).

Motivos de proteínas são as bases da classificação estrutural

Arquitetura proteica – Estrutura terciária de proteínas globulares grandes, IV. Classificação estrutural das proteínas. Como se vê, a compreensão da complexidade da estrutura terciária é facilitada pela consideração das subestruturas. Seguindo essa ideia, os pesquisadores organizaram o conteúdo completo de bancos de dados proteicos, de acordo com uma hierarquia dos níveis estruturais. Todos esses bancos de dados se baseiam nos dados e nas informações depositados no Protein Data Bank (PDB, Banco de Dados de Proteínas). O banco de dados Structural Classification of Proteins (SCOP, ou Classificação Estrutural de Proteínas) é um bom exemplo desta importante tendência na bioquímica. No mais alto nível de classificação, o banco de dados SCOP (<http://scop.mrc-lmb.cam.ac.uk/scop>) utiliza um esquema já em uso, com quatro classes de estruturas de proteínas: **toda α** , **toda β** , **α/β** (com segmentos α e β intercalados ou alternados) e **$\alpha + \beta$** (com regiões α e β um pouco separadas). Cada classe inclui dezenas a centenas de diferentes combinações de enovelamentos (motivos), construídas a partir de subestruturas gradativamente identificáveis. Alguns desses arranjos subestruturais são bem comuns, outros foram encontrados em apenas uma proteína. A **Figura 4-22** mostra a variedade desses motivos distribuídos entre as quatro classes de estrutura de proteínas. Essa é apenas uma pequena amostra das centenas de motivos conhecidos. Entretanto, o número de padrões de enovelamento não é infinito. À medida que a velocidade na qual novas estruturas de proteínas são elucidadas, a fração dessas estruturas que contém novos motivos tem diminuído. Ao todo, devem existir menos de 1.000 motivos diferentes. A Figura 4-22 também mostra como as proteínas podem ser organizadas em função da presença dos vários motivos. Os níveis de organização mais superiores, **classe** e **enovelamento**, são puramente estruturais. Abaixo do nível de enovelamento (ver o código de cores na Figura 4-22), a classificação tem como base as relações evolutivas.

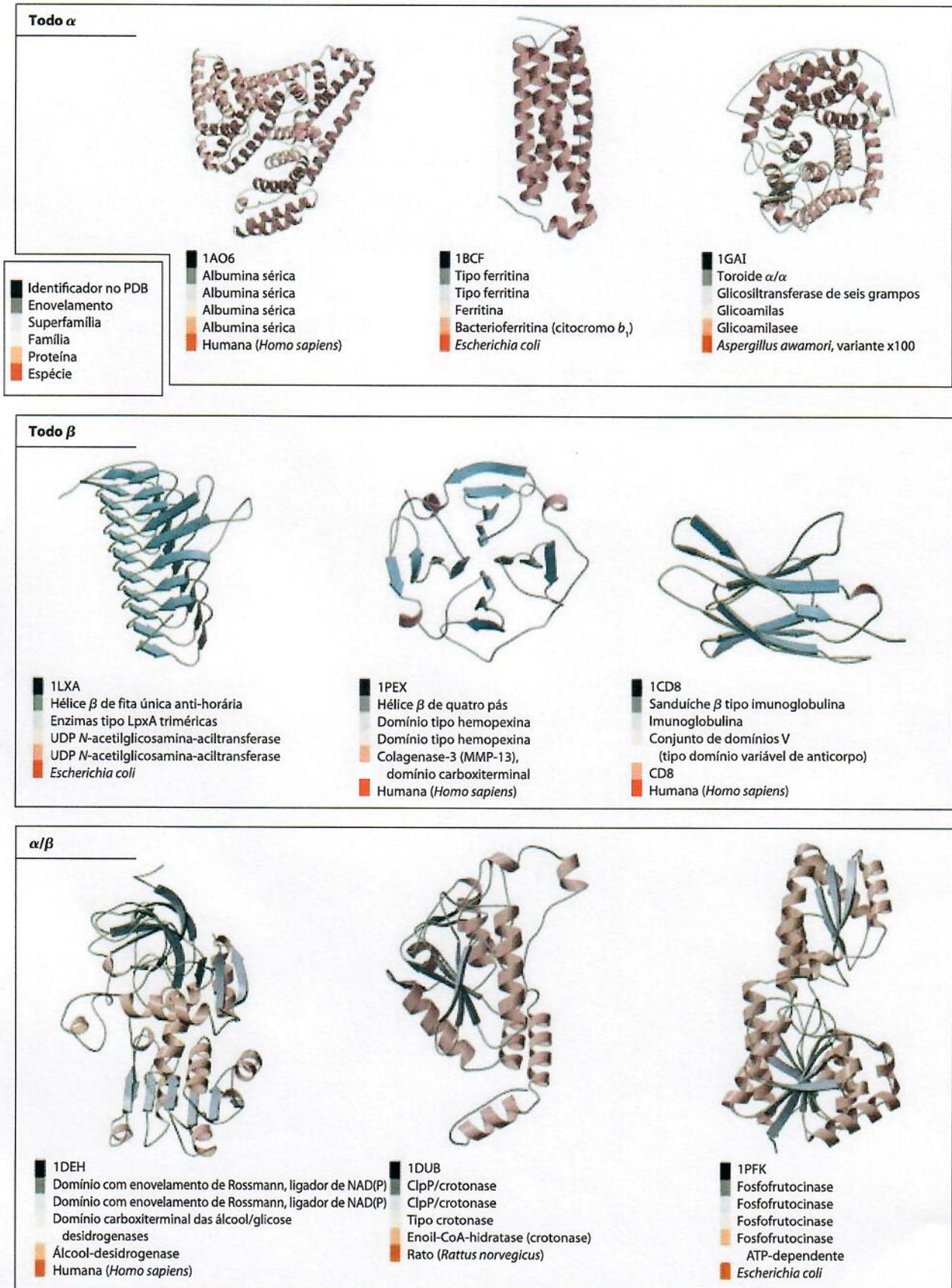


FIGURA 4-22 Organização das proteínas com base nos motivos. Mostrados aqui estão apenas alguns das centenas de motivos estáveis conhecidos, divididos em quatro classes: todo α , todo β , α/β e $\alpha + \beta$. Os dados de classificação estrutural do banco de dados SCOP (Structural Classification of Proteins, <http://scop.mrc-lmb.cam.ac.uk/scop>) também são fornecidos

(ver código de cores). O identificador no PDB (listado inicialmente para cada estrutura) corresponde ao código único de acesso de cada estrutura depositada no Protein Data Bank (www.pdb.org). O barril α/β (ver Figura 4-21) é outro motivo α/β particularmente comum. **(Continua na próxima página)**

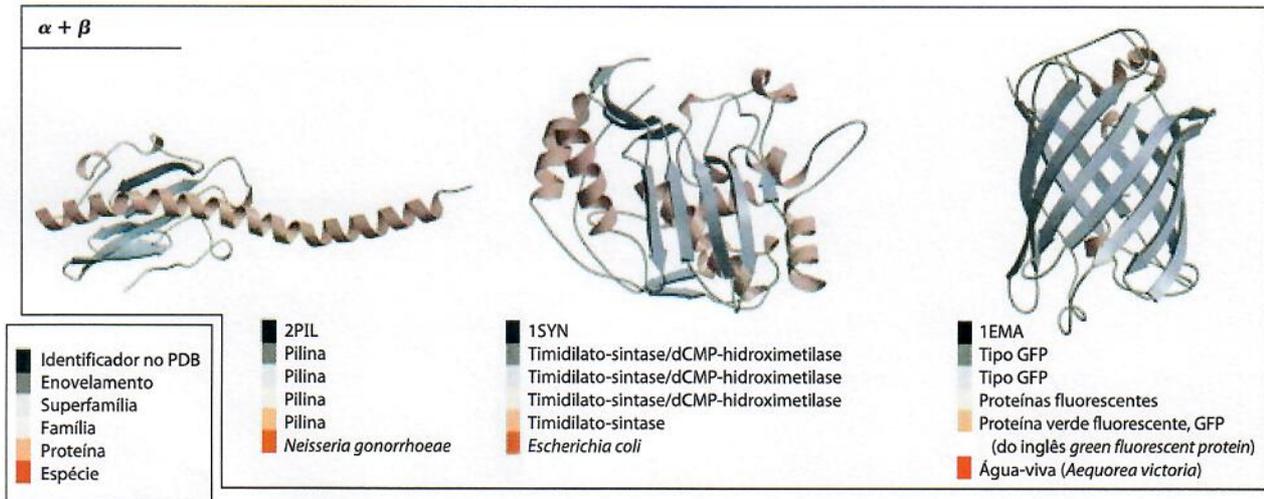


FIGURA 4-22 (Continuação)

Vários exemplos de domínios recorrentes ou motivos estruturais estão disponíveis, revelando que a estrutura terciária das proteínas é mais conservada do que suas sequências de aminoácidos. A comparação entre as estruturas de proteínas pode, então, fornecer muita informação sobre sua evolução. Proteínas com significativa semelhança na estrutura primária e/ou com estrutura terciária e funções semelhantes fazem parte de uma mesma **família proteica**. Uma relação evolutiva forte normalmente é evidente em uma família de proteínas. Por exemplo, a família globina apresenta muitas proteínas diferentes com similaridade estrutural e de sequência com a mioglobina (como visto nas proteínas usadas nos exemplos do Quadro 4-5 e no Capítulo 5). Duas ou mais famílias com pouca similaridade na sequência de aminoácidos algumas vezes utilizam o mesmo motivo estrutural geral e apresentam semelhanças funcionais; essas famílias são agrupadas como **superfamílias**. Uma relação evolutiva entre as famílias em uma superfamília é considerada provável, ainda que as diferenças temporais e funcionais – resultantes de diferentes pressões adaptativas – possam ter suprimido muitas das relações reveladoras entre sequências. Uma família de proteínas pode estar distribuída em todos os três domínios de vida celular: Bacteria, Archaea e Eukarya, sugerindo uma origem remota. Muitas proteínas envolvidas no metabolismo intermediário e no metabolismo dos ácidos nucleicos e proteínas se encaixam nessa categoria. Outras famílias podem estar presentes somente em um pequeno grupo de organismos, indicando que aquela estrutura surgiu mais recentemente. Rastrear a história natural dos motivos estruturais, usando as classificações estruturais de bancos de dados como o SCOP, fornece um grande complemento às análises de sequência na investigação da relação evolutiva. O banco de dados SCOP é manualmente gerenciado, com o objetivo de colocar as proteínas na sua rede evolutiva correta, com base nas características estruturais conservadas.

Motivos estruturais se tornaram especialmente importantes na definição de famílias e superfamílias de proteínas. Sistemas aperfeiçoados de classificação e elucidação de proteínas inevitavelmente levaram à elucidação de novas relações funcionais. Dada a importância das proteínas nos

sistemas vivos, essas comparações estruturais podem ajudar no esclarecimento de todos os aspectos da bioquímica, da evolução de proteínas individuais até a história evolutiva de rotas metabólicas completas.

A estrutura quaternária varia de dímeros simples a grandes complexos

Arquitetura proteica – Estrutura quaternária Muitas proteínas têm múltiplas subunidades polipeptídicas (de duas a centenas). A associação das cadeias polipeptídicas pode servir a uma variedade de funções. Muitas proteínas com múltiplas subunidades podem ter funções reguladoras; a ligação de pequenas moléculas pode afetar a interação entre as subunidades, causando grandes mudanças na atividade da proteína em resposta a pequenas mudanças na concentração de substrato ou moléculas reguladoras (Capítulo 6). Em outros casos, subunidades distintas assumem funções diferentes, mas relacionadas, como catálise e regulação. Algumas associações, como nas proteínas fibrosas consideradas anteriormente neste capítulo e nas proteínas dos capsídeos dos vírus, têm, principalmente, funções estruturais. Algumas associações proteicas muito grandes formam sítios de reações complexas envolvendo múltiplas etapas. Por exemplo, cada ribossomo, o local onde ocorre a síntese de proteínas, incorpora dúzias de subunidades de proteínas com certo número de moléculas de RNA.

Uma proteína de múltiplas subunidades também é chamada de **multímero**. Um multímero com apenas algumas subunidades é chamado de **oligômero**. Se um multímero apresenta subunidades não idênticas, a estrutura total da proteína pode ser assimétrica e bem complicada. Entretanto, a maioria dos multímeros apresenta subunidades idênticas, ou grupos de subunidades não idênticas que se repetem, geralmente em arranjos simétricos. Como apresentado no Capítulo 3, as unidades estruturais de repetição em uma proteína multimérica, tanto de um único tipo de subunidade quanto de um grupo de subunidades, são chamadas de **protômero**. Muitas vezes são utilizadas letras gregas para distinguir as diferentes subunidades que fazem parte do protômero.

A primeira proteína oligomérica que teve sua estrutura tridimensional determinada foi a hemoglobina (M_r 64.500), que contém quatro cadeias polipeptídicas e quatro grupos prostéticos heme, nos quais os átomos de ferro se apresentam no estado ferroso (Fe^{2+}) (Figura 4-17). A porção proteica, a globina, consiste em duas cadeias (141 resíduos em cada cadeia α) e duas cadeias β (146 resíduos em cada cadeia). Observe que neste caso, α e β não se referem às estruturas secundárias. Em uma prática, que pode ser confusa para o aluno iniciante, as letras gregas α e β (γ e δ , e outras) são frequentemente usadas para distinguir dois tipos diferentes de subunidades em uma proteína com múltiplas subunidades, independentemente de quais tipos de estruturas secundárias podem predominar nas subunidades. Como a hemoglobina é quatro vezes maior do que a mioglobina, foi necessário muito mais tempo e esforço para resolver sua estrutura tridimensional por análise de raios X, o que foi finalmente obtido por Max Perutz, John Kendrew e seus colegas, em 1959. As subunidades da hemoglobina estão organizadas em pares simétricos (Figura 4-23). Cada par tem uma subunidade α e uma subunidade β . A hemoglobina pode, portanto, ser descrita tanto como tetrâmero quanto dímero de protômeros $\alpha\beta$. O papel dessas subunidades distintas na função da hemoglobina é amplamente discutido no Capítulo 5.

Algumas proteínas ou segmentos proteicos são intrinsecamente desordenados

Apesar de décadas de progresso no entendimento da estrutura proteica, muitas proteínas não podem ser cristalizadas, tornando difícil a determinação de suas estruturas tridimensionais por métodos considerados clássicos (ver Quadro 4-5). Mesmo quando cristalizadas, partes da proteína estão suficientemente desordenadas no cristal de forma que a estrutura determinada não inclui essas partes. Algumas vezes, isso se deve a características sutis da estrutura que tornam a cristalização difícil. No entanto, a razão pode ser mais simples: algumas proteínas ou segmentos proteicos carecem de estruturas ordenadas em solução.

O conceito de que algumas proteínas funcionam na ausência de uma estrutura definida é um produto da reavaliação dos dados envolvendo muitas proteínas diferentes. Um terço de todas as proteínas humanas pode ser desestruturado ou possuem segmentos desestruturados significativos. Todos os organismos possuem algumas proteínas que se encaixam nessa categoria. As **proteínas intrinsecamente desordenadas** têm propriedades distintas das proteínas estruturadas clássicas. Elas carecem de um núcleo hidrofóbico e, ao contrário, são caracterizadas por alta densidade de aminoácidos carregados, como Lys, Arg e Glu. Resíduos de Pro também são proeminentes, já que eles tendem a romper estruturas ordenadas.

A desordem estrutural e a alta densidade de cargas podem facilitar a atividade de algumas proteínas, como espaçadores, isoladores ou elementos de ligação em estruturas maiores. Outras proteínas desordenadas são sequestradoras, ligando íons e moléculas pequenas em solução e servindo de reservatórios ou depósitos de lixo. No entanto, muitas proteínas intrinsecamente desordenadas são o núcleo de importantes redes de interações proteicas. A falta de uma estrutura ordenada pode facilitar um tipo de promiscuidade funcional, permitindo a interação de uma proteína com múltiplos parceiros. Algumas proteínas intrinsecamente desordenadas atuam por inibir a atividade de outras proteínas por um mecanismo não usual: por envolver-se em torno de suas proteínas-alvo. Uma proteína desordenada pode ter algumas ou mesmo dúzias de proteínas parceiras. A desordem estrutural permite ao inibidor proteico se envolver em torno de múltiplos alvos de diferentes formas. A proteína



Max Perutz, 1914–2002 (esquerda) e John Kendrew, 1917–1997

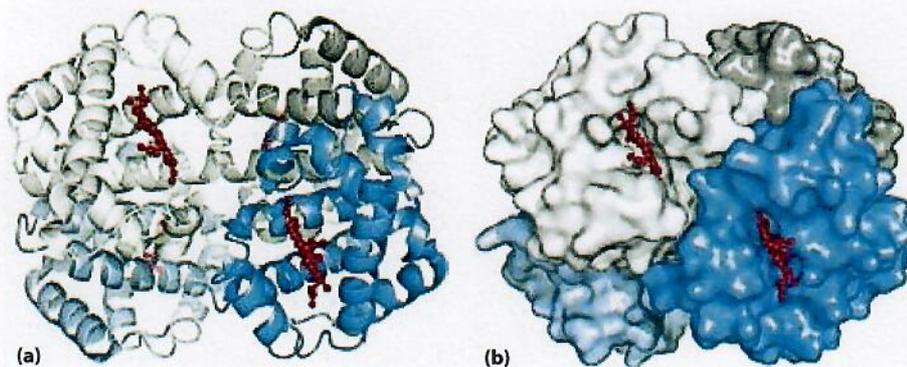


FIGURA 4-23 Estrutura quaternária da desóxi-hemoglobina. (PDB ID 2HHB) A análise por difração de raios X da desóxi-hemoglobina (hemoglobina sem as moléculas de oxigênio ligadas aos grupos heme) mostra como as quatro subunidades polipeptídicas estão agrupadas. **(a)** Representação na forma de fitas revela os elementos estruturais secundários da estrutura e a posição de todos os cofatores hemes. **(b)** Modelo de contorno de superfície mostra a cavidade em que os cofatores hemes estão ligados e ajuda a visualizar o empacotamento das subunidades. As subunidades α são mostradas em tons de cinza, e as subunidades β , em tons de azul. Observe que os grupos heme (em vermelho) estão relativamente distantes entre si.

intrinsecamente desordenada p27 exerce um papel-chave no controle da divisão celular em mamíferos. Essa proteína carece de estrutura definida quando em solução. Ela se enrola em torno e assim inibe a atividade de várias enzimas chamadas proteínas-quinases (ver Capítulo 6) que facilitam a divisão. A estrutura flexível da p27 a permite se acomodar às suas proteínas-alvo distintas. Células tumorais humanas, que são simplesmente células que perderam a capacidade de controlar a divisão celular normalmente, geralmente possuem níveis reduzidos de p27; quanto menor o nível de p27, pior é o prognóstico para o paciente de câncer. Da mesma forma, as proteínas intrinsecamente desordenadas estão frequentemente presentes como concentradores ou suportes no centro de redes proteicas que constituem as vias de sinalização. Essas proteínas ou partes delas podem interagir com muitos parceiros diferentes. Elas frequentemente assumem uma estrutura quando interagem com outras proteínas, mas as estruturas que elas assumem podem variar de acordo com o parceiro de ligação. A proteína de mamíferos p53 também é fundamental no controle da divisão celular. Ela tem segmentos estruturados e não estruturados, e os segmentos distintos interagem com dúzias de outras proteínas. Uma região não estruturada da p53 na extremidade carboxil interage com pelo menos quatro parceiros diferentes e assume uma estrutura distinta em cada complexo (**Figura 4-24**).

RESUMO 4.3 Estruturas terciária e quaternária das proteínas

- ▶ A estrutura terciária é a estrutura tridimensional da cadeia polipeptídica. Muitas proteínas se encaixam em uma ou duas classes de proteínas em geral, com base na estrutura terciária: fibrosa e globular.
- ▶ As proteínas fibrosas, utilizadas principalmente para funções estruturais, são formadas por elementos de estrutura secundária que se repetem.
- ▶ As proteínas globulares têm estruturas terciárias mais complicadas, geralmente contendo diversos tipos de estruturas secundárias na mesma cadeia polipeptídica. A primeira estrutura de proteína globular a ser determinada, por métodos de difração de raios X, foi a estrutura da mioglobina.
- ▶ As estruturas complexas das proteínas globulares podem ser analisadas pela observação dos padrões de enovelamento chamados de motivos (também chamados de enovelamentos ou estruturas supersecundárias). Os milhares de estruturas proteicas conhecidos geralmente são formados por um repertório de apenas poucas centenas de motivos. Os domínios são regiões de uma cadeia polipeptídica que podem se dobrar de forma estável e independente.
- ▶ A estrutura quaternária resulta de interações entre as subunidades de proteínas com múltiplas subunidades

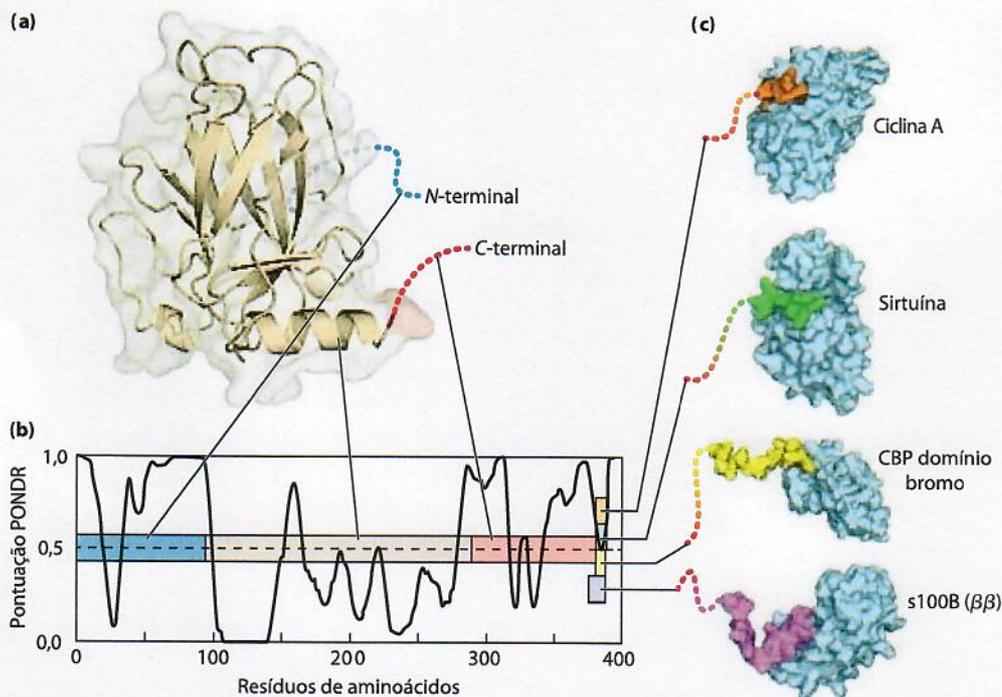


FIGURA 4-24 Ligação da região carboxiterminal intrinsecamente desordenada da proteína p53 aos seus ligantes. (a) A proteína p53 é composta por vários segmentos diferentes (PDB ID 1XQH). Apenas o domínio central é bem organizado. (b) A sequência linear da proteína p53 está representada na barra colorida. O gráfico sobreposto apresenta um diagrama da pontuação PONDR (Predictor of Natural Disordered Regions; Previsor de Regiões Naturalmente Desordenadas) versus a sequência da proteína. PONDR é um dos melhores algoritmos disponíveis para prever a probabilidade que um determinado resíduo de aminoácido está em uma região de desordem intrínseca, com base na sequência de aminoácidos ao redor e na

composição de aminoácidos. Uma pontuação de 1,0 indica uma probabilidade de 100% que a proteína estará desordenada. Nesta estrutura proteica, o domínio central em bege está organizado. As regiões aminoterminal (azul) e carboxiterminal (vermelho) estão desordenadas. A extremidade da região carboxiterminal tem múltiplos ligadores e ela se dobra quando se liga a cada um deles; no entanto, a estrutura tridimensional assumida quando ocorre a ligação é diferente para cada uma das interações mostradas, e por isso o segmento carboxiterminal (resíduos de 11 a 20) está mostrado em cor diferente em cada complexo (ciclina A, PDB ID 1H26; sirtuína, PDB ID 1MA3; CBP domínio bromo, PDB ID 1JSP; s100B ($\beta\beta$), PDB ID 1DT7).

(multiméricas) ou grandes associações de proteínas. Algumas proteínas multiméricas possuem unidades repetidas, formadas por uma única subunidade ou grupos de subunidades, cada unidade chamada protômero.

- ▶ Algumas proteínas ou segmentos proteicos são intrinsecamente desordenados, carecendo de estruturas definidas. Essas proteínas têm composições distintas de aminoácidos que permitem uma estrutura mais flexível. Algumas dessas proteínas desordenadas funcionam como componentes estruturais ou sequestradoras; outras podem interagir com muitas proteínas parceiras, servindo como inibidores versáteis ou como componentes centrais de redes de interação de proteínas.

4.4 Desnaturação e enovelamento das proteínas

As proteínas têm uma existência surpreendentemente precária. Como foi visto, a conformação de uma proteína nativa é apenas marginalmente estável. Além disso, a maioria das proteínas deve manter certa flexibilidade conformacional para funcionar. A manutenção contínua do grupo ativo de proteínas celulares, necessárias em um dado conjunto de condições, é chamada **proteostase**. A proteostase celular requer a atividade coordenada de vias para síntese e enovelamento de proteínas, o redobramento de proteínas parcialmente desdobradas e o sequestro e degradação de proteínas irreversivelmente desdobradas. Em todas as células, essas redes envolvem centenas de enzimas e proteínas especializadas.

Como visto na **Figura 4-25**, a vida de uma proteína engloba muito mais do que sua síntese e degradação. A estabilidade marginal da maioria das proteínas pode produzir um balanço tênue entre os estados dobrados e desdobrados. À medida que as proteínas são sintetizadas nos ribossomos (Capítulo 27), elas devem dobrar-se em sua conformação nativa. Algumas vezes isso ocorre espontaneamente, porém mais frequentemente isso ocorre com a assistência de enzimas e complexos especializados chamados chaperonas. Muitos desses mesmos auxiliares do enovelamento atuam para redobrar proteínas que se tornaram transitoriamente desdobradas. As proteínas inapropriadamente dobradas frequentemente expõem superfícies hidrofóbicas que as tornam “pegajosas”, conduzindo à formação de agregados inativos. Esses agregados podem perder suas funções normais, mas não são inertes; seu acúmulo nas células situa-se no centro de doenças que vão de diabetes a doenças de Parkinson e Alzheimer. Não surpreendentemente, todas as células elaboraram vias de reciclagem e/ou degradação de proteínas irreversivelmente deformadas.

Agora serão abordadas as transições entre os estados dobrados e desdobrados, e a rede de vias que controlam essas transições.

A perda de estrutura da proteína resulta na perda de função

As estruturas proteicas evoluíram para atuar em determinados ambientes celulares. Condições diferentes daquelas da célula podem resultar em mudanças estruturais grandes ou pequenas na proteína. A perda de estrutura tridimensional suficiente para causar a perda de função é chamada de

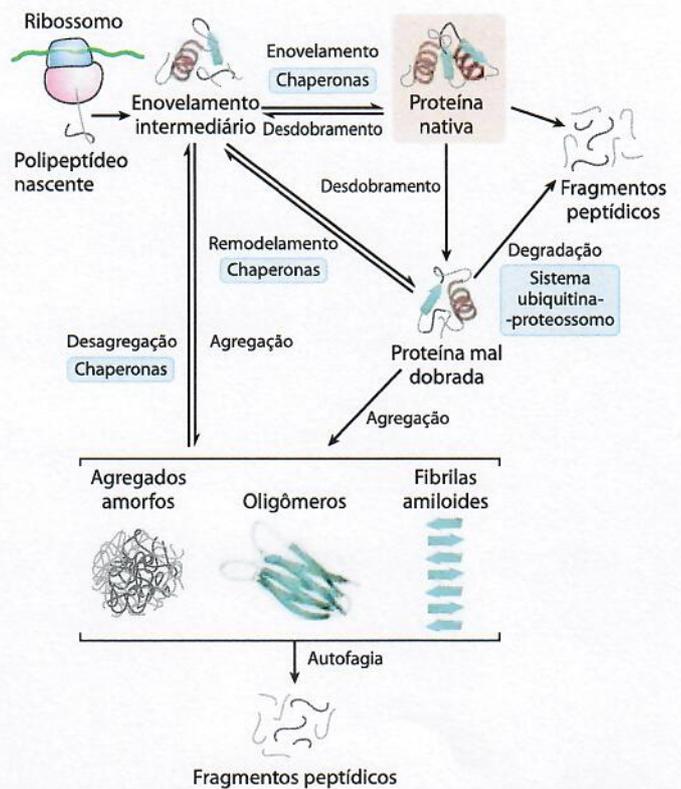


FIGURA 4-25 Vias que contribuem para proteostase. Três tipos de processos contribuem para proteostase. Primeiro, as proteínas são sintetizadas no ribossomo. Segundo, múltiplas vias contribuem para o enovelamento proteico, muitas dessas envolvem a atividade de complexos chamados chaperonas. As chaperonas (incluindo as chaperoninas) também contribuem para o redobramento de proteínas parcialmente ou transitoriamente desdobradas. Finalmente, proteínas irreversivelmente desdobradas estão sujeitas ao sequestro e à degradação por várias vias adicionais. As proteínas desdobradas e os intermediários proteicos enovelados que escapam do controle de qualidade das chaperonas e das vias de degradação podem agregar, formando tanto agregados desordenados quanto agregados organizados do tipo amiloide que contribuem para doenças e processos de envelhecimento.

desnaturação. O estado desnaturado não necessariamente corresponde ao desdobração completo da proteína e à randomização da conformação. Na maioria das condições, as proteínas desnaturadas existem como um conjunto de estados parcialmente dobrados.

A maioria das proteínas pode ser desnaturada pelo calor, que tem efeitos complexos nas muitas interações fracas da proteína (principalmente sobre as ligações de hidrogênio). Se a temperatura é aumentada lentamente, a conformação da proteína em geral permanece intacta até que, em uma estreita faixa de temperatura, ocorre uma perda abrupta da estrutura (e da função) (**Figura 4-26**). A mudança repentina sugere que o desdobração é um processo cooperativo: a perda de estrutura em uma parte da proteína desestabiliza as outras partes. Os efeitos do aquecimento sobre as proteínas não são facilmente preditos. As proteínas altamente termorresistentes de bactérias termofílicas e de arqueobactérias evoluíram para funcionar na temperatura das fontes termais (~100°C). Mesmo assim, as estruturas dessas proteínas diferem pouco de suas proteínas homólogas de bactérias como a *Escherichia coli*. Não se sabe ao

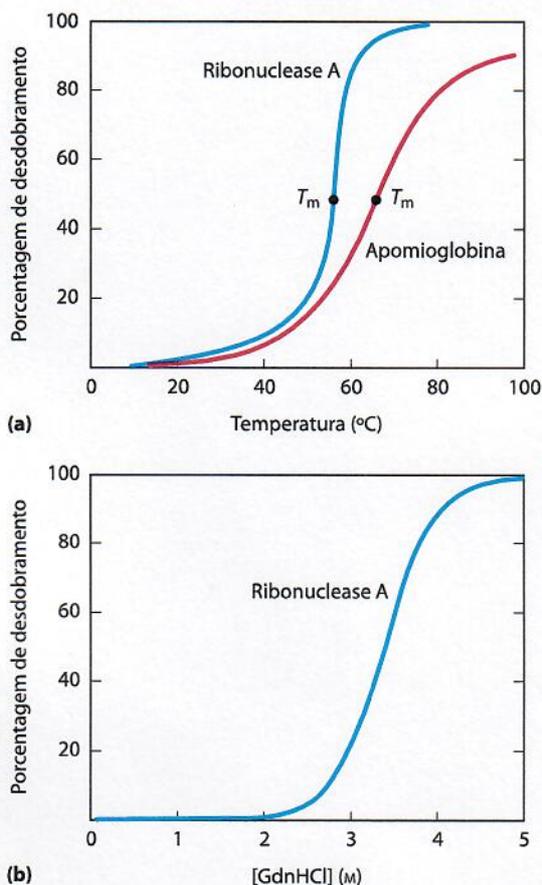


FIGURA 4-26 Desnaturação de proteínas. Os resultados correspondem a proteínas desnaturadas por duas modificações diferentes em seus ambientes. Em cada caso, a transição entre o estado dobrado e o estado não dobrado é abrupta, sugerindo cooperatividade neste processo. **(a)** Desnaturação térmica da apomioglobina (mioglobina sem o grupo prostético heme) de cavalo e da ribonuclease A (com suas ligações dissulfeto intactas; ver Figura 4-27). O ponto médio da faixa de temperatura em que ocorre a desnaturação é chamado de temperatura de fusão, ou T_m . A desnaturação da apomioglobina foi monitorada por dicroísmo circular (ver Figura 4-10), que mede a quantidade de estrutura helicoidal na proteína. A desnaturação da ribonuclease A foi acompanhada pelo monitoramento das mudanças na fluorescência intrínseca da proteína, que é afetada pelas mudanças no ambiente dos resíduos de Trp. **(b)** Desnaturação da ribonuclease A, com suas ligações dissulfeto intactas, pelo hidrocloreto de guanidina (GdnHCl), monitorado por dicroísmo circular.

certo como essas pequenas diferenças garantem a estabilidade estrutural em altas temperaturas.

As proteínas também podem ser desnaturadas por pHs extremos, por certos solventes orgânicos miscíveis, como álcool ou acetona, por certos solutos como ureia e hidrocloreto de guanidina, ou por detergentes. Cada um desses agentes desnaturantes representa um tratamento relativamente brando, já que nenhuma ligação covalente da cadeia polipeptídica é rompida. Solventes orgânicos, ureia e detergentes atuam, principalmente, rompendo as interações hidrofóbicas que mantêm o núcleo estável das proteínas globulares; a ureia também rompe as ligações de hidrogênio; extremos de pH alteram a carga líquida da proteína, causando repulsão eletrostática e o rompimento de algumas ligações de hidrogênio. As estruturas desnaturadas resultantes desses vários tratamentos não são, necessariamente, as mesmas.

A desnaturação frequentemente leva à precipitação de proteínas, uma consequência da formação de agregados proteicos pela exposição de superfícies hidrofóbicas associadas. Os agregados são, com frequência, altamente desordenados. O precipitado proteico visto após ferver um ovo é um exemplo. Agregados mais ordenados também são observados em algumas proteínas, como será visto.

A sequência de aminoácidos determina a estrutura terciária

A estrutura terciária de uma proteína globular é determinada por sua sequência de aminoácidos. A prova mais importante disso vem dos experimentos que mostram que a desnaturação de algumas proteínas é reversível. Certas proteínas globulares desnaturadas por temperatura, extremos de pH ou agentes desnaturantes reassumem suas estruturas nativas e suas atividades biológicas se retornarem às condições nas quais a conformação nativa é estável. Esse processo é chamado de **renaturação**.

Um exemplo clássico é a desnaturação e a renaturação da ribonuclease A, demonstradas por Christian Anfinsen nos anos 1950. A ribonuclease A purificada desnatura completamente em uma solução concentrada de ureia, na presença de um agente redutor. O agente redutor rompe as quatro ligações dissulfeto, resultando em oito resíduos Cys, e a ureia rompe as interações hidrofóbicas de estabilização, desta forma liberando todo o polipeptídeo de sua conformação dobrada. A desnaturação da ribonuclease é acompanhada por uma completa perda de atividade catalítica. Quando a ureia e o agente redutor são removidos, a ribonuclease desnaturada, aleatoriamente enrolada, se dobra de modo espontâneo em sua estrutura terciária correta, com restauração total de sua atividade catalítica (Figura 4-27). O reenovelamento da ribonuclease é tão preciso que as quatro ligações dissulfeto intramoleculares são restabelecidas na molécula renaturada nas mesmas posições da ribonuclease nativa. Mais tarde, resultados semelhantes foram obtidos utilizando a ribonuclease A cataliticamente ativa quimicamente sintetizada. Isso elimina a possibilidade de que algum contaminante minoritário da preparação da ribonuclease purificada por Anfinsen tenha contribuído para a renaturação da enzima, afastando assim qualquer dúvida que poderia restar de que essa enzima se dobra espontaneamente.

O experimento de Anfinsen forneceu a primeira evidência de que a sequência de aminoácidos de uma cadeia polipeptídica contém todas as informações necessárias para o enovelamento da cadeia em sua estrutura tridimensional nativa. Trabalhos posteriores demonstraram que somente uma minoria de proteínas, muitas delas pequenas e estáveis por natureza, se dobrará espontaneamente em sua forma nativa. Apesar de todas as proteínas terem o potencial de se dobrar em sua estrutura nativa, muitas delas necessitam de alguma assistência.

Os polipeptídeos dobram-se rapidamente por um processo gradual

Nas células vivas, as proteínas são construídas, a partir dos aminoácidos, em uma velocidade muito alta. Por exemplo, as células de *E. coli* podem fazer uma molécula de proteína

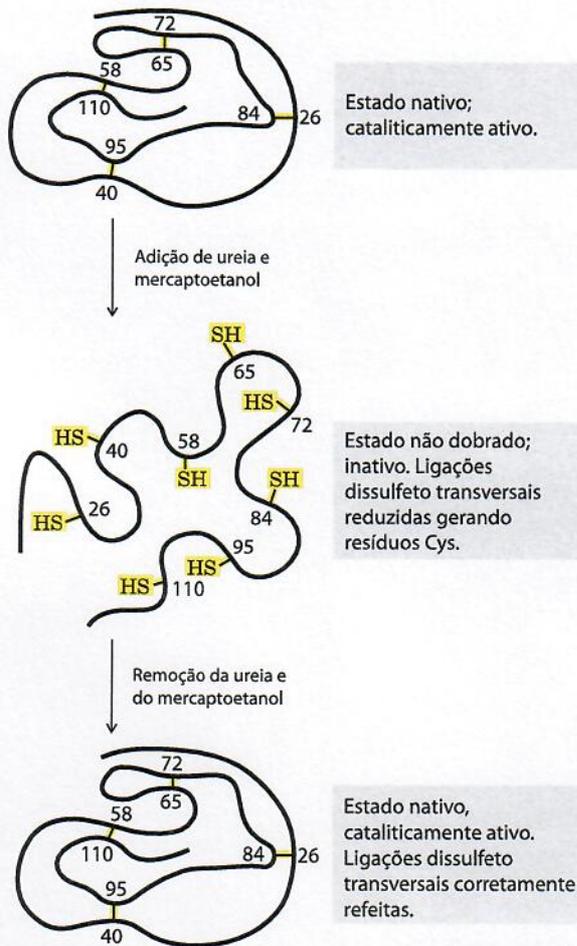


FIGURA 4-27 Renaturação da ribonuclease desnaturada e desdobrada. A ureia desnatura a ribonuclease, e o mercaptoetanol (HOCH₂CH₂SH) a reduz, rompendo as ligações dissulfeto e liberando os oito resíduos de Cys. A renaturação envolve o restabelecimento correto das ligações dissulfeto transversais.

completa, biologicamente ativa, com 100 resíduos de aminoácidos, em 5 segundos a 37°C. No entanto, a síntese das ligações peptídicas no ribossomo não é suficiente; a proteína deve dobrar-se.

Como a cadeia polipeptídica chega à sua conformação nativa? Suponha-se, de forma conservadora, que cada um dos resíduos de aminoácidos pode assumir, em média, 10 conformações diferentes, resultando em 10¹⁰⁰ conformações diferentes do polipeptídeo. Suponha-se, também, que a proteína se dobra espontaneamente por um processo aleatório no qual ela testa todas as conformações possíveis em torno de cada uma das ligações de seu esqueleto até encontrar sua forma nativa, biologicamente ativa. Se cada conformação fosse testada no menor tempo possível (~10⁻¹³ segundos, ou o tempo necessário para uma única vibração molecular), levaria aproximadamente 10⁷⁷ anos para testar todas as possíveis conformações. Claramente, o enovelamento de proteínas não é um processo completamente aleatório, de tentativa e erro. Deve haver atalhos. Esse problema foi apontado pela primeira vez por Cyrus Levinthal, em 1968, sendo algumas vezes chamado de paradoxo de Levinthal.

A rota de enovelamento de uma cadeia polipeptídica grande é muito complicada. No entanto, rápido progresso tem sido feito nesse campo, o suficiente para produzir algoritmos robustos que podem frequentemente predizer a estrutura de proteínas menores, com base em suas sequências de aminoácidos. As principais vias de enovelamento são hierárquicas. As estruturas secundárias locais se formam primeiro. Certas sequências de aminoácidos se dobras prontamente em hélices α ou folhas β, guiadas por restrições como aquelas revisadas na discussão sobre estruturas secundárias. Interações iônicas, envolvendo grupos carregados que normalmente estão próximos na sequência linear da cadeia polipeptídica, podem ter um papel importante no direcionamento dos primeiros passos de enovelamento. O arranjo de estruturas locais é seguido por interações de longo alcance entre, por exemplo, elementos da estrutura secundária que se aproximam para formar estruturas enoveladas estáveis. As interações hidrofóbicas exercem um papel significativo ao longo do processo, já que a agregação das cadeias laterais de aminoácidos apolares fornece uma estabilização entrópica a intermediários e, por fim, à estrutura enovelada final. O processo continua até a formação de domínios completos e até que todo o polipeptídeo esteja dobrado (Figura 4-28). Proteínas nas quais predominam as interações de curto alcance (entre pares de resíduos geralmente localizados próximos um do outro na sequência polipeptídica) tendem a se dobrar mais rapidamente do que

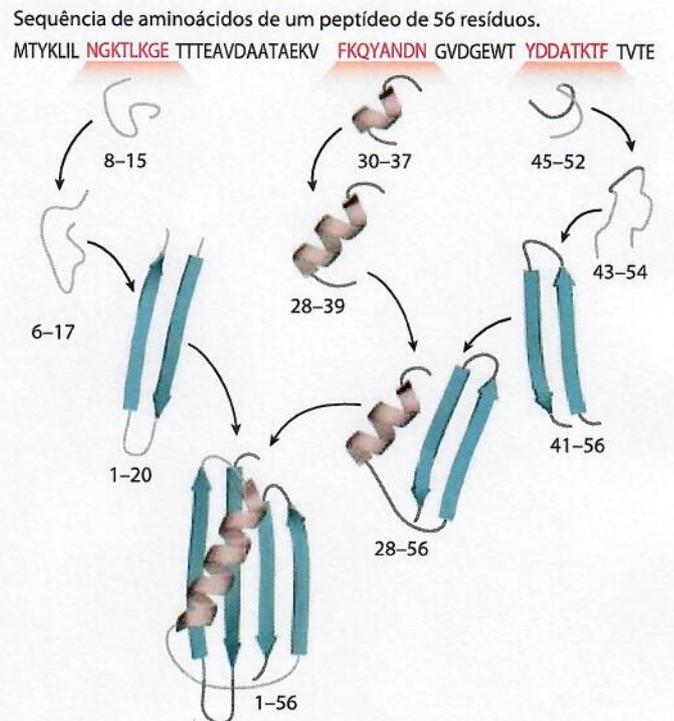


FIGURA 4-28 Uma via de enovelamento proteico definida para uma proteína pequena. Uma via hierárquica é mostrada com base em modelagem computacional. Primeiro, pequenas regiões de estrutura secundária são agrupadas e, então, gradualmente incorporadas em estruturas maiores. O programa usado para este modelo tem tido muito êxito na predição da estrutura tridimensional de pequenas proteínas a partir de sua sequência de aminoácidos. Os números indicam os resíduos de aminoácidos desse peptídeo de 56 resíduos que adquiriu sua estrutura final em cada uma das etapas mostradas.

proteínas com padrões de enovelamento mais complexos e com mais interações de longo alcance entre os diferentes segmentos. À medida que as proteínas com múltiplos domínios são sintetizadas, domínios próximos à extremidade aminoterminal (sintetizadas primeiro) podem se dobrar antes do polipeptídeo inteiro ter sido montado.

Termodinamicamente, o processo de enovelamento pode ser visto como um tipo de funil de energia livre (Figura 4-29). Os estados não dobrados são caracterizados por um alto grau de entropia conformacional e energia livre relativamente alta. À medida que o enovelamento ocorre, o estreitamento do funil reflete a diminuição do espaço conformacional que deve ser procurado, à medida que a proteína se aproxima de seu estado nativo. As pequenas depressões ao longo das paredes do funil de energia livre representam intermediários semiestáveis, que podem tornar o processo de enovelamento um pouco mais lento. No fundo do funil, o conjunto de intermediários dobrados é reduzido a uma única conformação nativa (ou a um pequeno conjunto de conformações nativas). Os funis podem ter uma variedade de formas dependendo da complexidade da via de enovelamento, da existência de intermediários semiestáveis, e do potencial de intermediários em particular se reunirem em agregados de proteínas erroneamente dobradas (Figura 4-29).

A estabilidade termodinâmica não é igualmente distribuída na estrutura da proteína – a molécula tem regiões de relativamente alta estabilidade e outras de estabilidade baixa ou desprezível. Por exemplo, uma proteína pode ter dois domínios estáveis ligados por um segmento inteiramente desordenado. Regiões de menor estabilidade podem permitir à proteína alterar sua conformação entre dois ou mais estados. Como será visto nos próximos dois capítulos, variações na estabilidade em regiões de uma proteína são, com frequência, essenciais para sua função. Proteínas ou

segmentos proteicos intrinsecamente desordenados não se dobram de forma alguma.

À medida que o entendimento sobre estrutura e enovelamento de proteínas evolui, programas computacionais sofisticados para a predição da estrutura das proteínas a partir da sequência de aminoácidos têm sido desenvolvidos. A predição da estrutura tridimensional das proteínas é uma das especialidades da bioinformática, e os progressos nessa área são monitorados com um teste bianual chamado de competição CASP (Critical Assessment of Structural Prediction, ou Avaliação Crítica da Predição Estrutural). Grupos de diversas partes do mundo tentam prever a estrutura de certa proteína (cuja estrutura já foi determinada, mas ainda não foi publicada). Os grupos que têm mais sucesso são convidados a apresentar seus resultados na conferência da CASP. O sucesso desses esforços está melhorando rapidamente.

Algumas proteínas se dobram de forma assistida

Nem todas as proteínas se dobram espontaneamente à medida que são sintetizadas dentro da célula. O enovelamento de muitas proteínas necessita de **chaperonas**, proteínas que interagem com polipeptídeos parcialmente dobrados ou dobrados de forma incorreta, facilitando os mecanismos de enovelamento correto ou garantindo um microambiente adequado para ocorrer o enovelamento. Vários tipos de chaperonas moleculares foram encontrados em organismos que vão desde bactérias até humanos. As duas principais famílias de chaperonas, ambas bem estudadas, são a família da **Hsp70** e as **chaperoninas**.

A família de proteínas Hsp70 geralmente tem massa molecular próxima de 70.000 e são abundantes em células submetidas a altas temperaturas (por isso, a denomina-

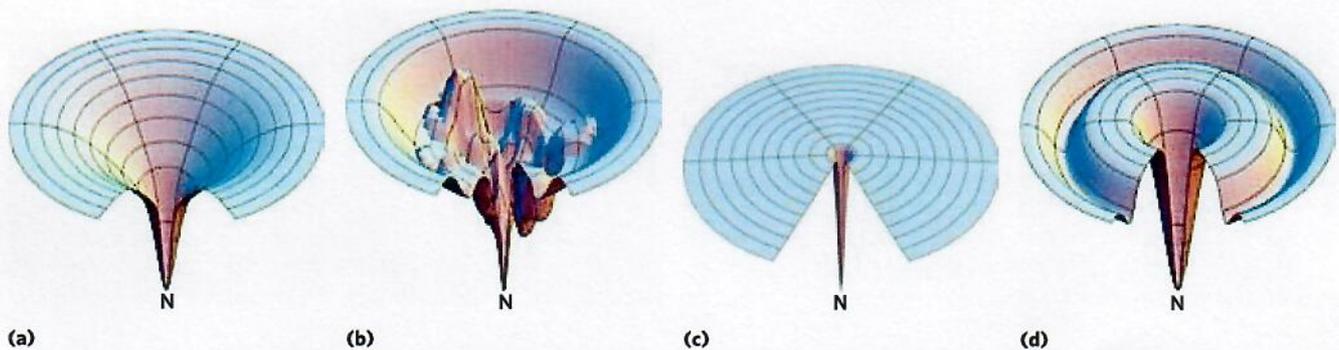


FIGURA 4-29 Termodinâmica do enovelamento proteico, mostrado como funil de energia livre. À medida que as proteínas se dobram, o espaço conformacional que pode ser explorado pela estrutura fica restrito. Este é modelado como funil termodinâmico tridimensional, com ΔG representado como profundidade e com a estrutura nativa (N) no fundo (ponto de menor energia-livre). O funil para determinada proteína pode ter uma variedade de formas, dependendo do número e dos tipos de intermediários de enovelamento nas vias de enovelamento. Qualquer intermediário de enovelamento com estabilidade significativa e um tempo de vida finita estaria representado como um mínimo de energia-livre local – uma depressão na superfície do funil. **(a)** Um funil simples, mas relativamente amplo e suave representa uma proteína com múltiplas vias de enovelamento (ou seja, a ordem em que as partes distintas da proteína se dobram seria de alguma forma aleatória), mas que assume suas estruturas tridimensionais sem

intermediários de enovelamento com estabilidade significativa. **(b)** Este funil representa uma proteína mais comum com múltiplos intermediários de enovelamento possíveis com estabilidade significativa nas várias vias que levam à estrutura nativa. **(c)** Uma proteína com estrutura nativa estável, essencialmente sem intermediários de enovelamento com estabilidade significativa, e apenas uma ou muito poucas vias de enovelamento produtivo está mostrada como um funil com uma depressão estreita levando à forma nativa. **(d)** Uma proteína com intermediários de enovelamento com estabilidade substancial em praticamente qualquer via que leva ao estado nativo (ou seja, uma proteína na qual um motivo ou domínio particular sempre se dobra rapidamente, mas as outras partes da proteína se dobram mais lentamente e em ordem aleatória) está representada por um funil com uma depressão principal ao redor da depressão que leva à forma nativa.

ção “proteínas de choque térmico”, do inglês *heat shock proteins*, de M_r 70.000, ou Hsp70). As proteínas Hsp70 se ligam a regiões não dobradas do peptídeo rico em resíduos hidrofóbicos. Dessa forma, essas chaperonas “protegem” da desnaturação pela temperatura tanto as moléculas de proteína quanto as novas moléculas de peptídeo que estão sendo sintetizadas (e ainda não estão dobradas). As proteínas Hsp70 também bloqueiam o enovelamento de certas proteínas que devem permanecer não dobradas até que sejam deslocadas através da membrana (como descrito no Capítulo 27). Algumas chaperonas também facilitam o arranjo quaternário de proteínas oligoméricas. As proteínas Hsp70 se ligam aos polipeptídeos e os liberam em um ciclo que utiliza a energia da hidrólise do ATP e envolve diversas outras proteínas (incluindo uma classe chamada de Hsp40). A Figura 4-30 ilustra um enovelamento assistido por chaperona, como o elucidado para as chaperonas Hsp70 e Hsp40 eucarióticas. A ligação de um

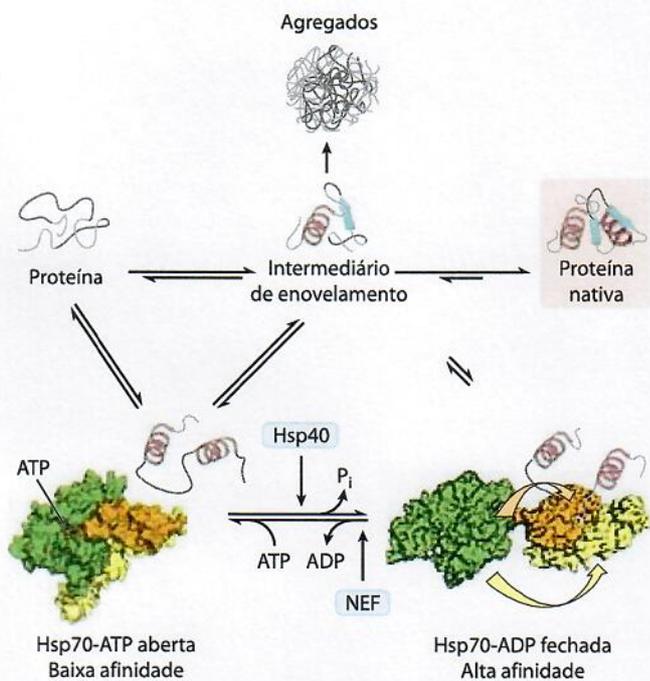


FIGURA 4-30 As chaperonas no enovelamento proteico. A rota pela qual as chaperonas da classe Hsp70 se ligam aos polipeptídeos e os liberam é ilustrada pelas chaperonas eucarióticas Hsp70 e Hsp40. As chaperonas não promovem diretamente o enovelamento da proteína, mas evitam a formação de agregados de peptídeos não dobrados. As proteínas não dobradas ou parcialmente dobradas primeiro se ligam à forma aberta da Hsp70 ligada ao ATP (PDB ID 2QXL). A Hsp40 então interage com esse complexo e induz a hidrólise do ATP produzindo a forma fechada do complexo (derivado de PDB ID 2KH0 e 1DKZ), na qual os domínios coloridos em cor de laranja e amarelo estão juntos como as duas partes de uma mandíbula, mantendo presas partes da proteína desdobrada. A dissociação do ADP e a reciclagem de Hsp70 requer a interação com outra proteína, um fator de troca de nucleotídeo (NEF). Em uma população de moléculas de polipeptídeos, alguma fração das moléculas liberadas após a ligação transitória das proteínas parcialmente dobradas pela Hsp70 assumirá a conformação nativa. As remanescentes são novamente ligadas pela Hsp70 ou desviadas para o sistema das chaperoninas (Hsp60; ver Figura 4-31). Em bactérias, as chaperonas Hsp70 e Hsp40 são chamadas DnaK e DnaJ, respectivamente. DnaK e DnaJ foram primeiro identificadas como proteínas necessárias para a replicação *in vitro* de certas moléculas de DNA viral (por isso a designação “Dna”).

peptídeo não dobrado à chaperona Hsp70 pode quebrar um agregado proteico ou prevenir a formação de um novo agregado. Quando o peptídeo ligado é liberado, ele tem a chance de retomar o enovelamento à sua estrutura nativa. Se o enovelamento não ocorrer rápido o suficiente, o polipeptídeo pode ser ligado de novo e o processo se repete. Alternativamente, o polipeptídeo ligado a Hsp70 pode ser entregue a uma chaperonina.

Chaperoninas são complexos elaborados de proteínas necessários para o enovelamento de algumas proteínas celulares que não se dobras espontaneamente. Em *E. coli*, estima-se que de 10 a 15% das proteínas celulares necessitam do sistema de chaperoninas residentes, chamado de GroEL/GroES, para o enovelamento nas condições normais (mais de 30% precisam de assistência quando as células são submetidas ao estresse por calor). Em eucariotos, o sistema de chaperoninas análogo chama-se Hsp60. As chaperoninas se tornaram conhecidas quando foram consideradas necessárias para o crescimento de certos vírus bacterianos (por isso, a designação “Gro”, de *growth*). Essa família de proteínas é estruturada como uma série de anéis de múltiplas subunidades, formando duas câmaras orientadas uma de costas para a outra. Uma proteína desdobrada primeiro se liga a uma superfície hidrofóbica exposta próxima a extremidade apical de uma câmara GroEL. A proteína é então presa dentro da câmara, fechada transitóriamente, pela “tampa” GroES (Figura 4-31). A GroEL sofre mudanças conformacionais substanciais, associadas à hidrólise lenta de ATP que também regula a ligação e a liberação de GroES. Dentro da câmara, uma proteína tem cerca de 10 segundos para se dobrar – o tempo necessário para a hidrólise do ATP ligado. A restrição de uma proteína dentro da câmara previne a agregação proteica inapropriada e também restringe o espaço conformacional que a cadeia polipeptídica pode explorar à medida que ela se enovela. A proteína é liberada quando GroES se dissocia, mas pode religar rapidamente para outro ciclo se o enovelamento não tiver sido completo. As duas câmaras em um complexo GroEL se alternam na ligação e liberação dos substratos polipeptídicos. Em eucariotos, o sistema Hsp60 utiliza um sistema semelhante para o enovelamento de proteínas. No entanto, em vez da tampa GroES, protruções dos domínios apicais das subunidades se flexionam e fecham a câmara. O ciclo hidrolítico do ATP também é mais lento nos complexos Hsp60, dando mais tempo à proteína retida na câmara para o enovelamento.

Finalmente, as rotas de enovelamento de algumas proteínas requerem duas enzimas que catalisam reações de isomerização. A **proteína dissulfeto-isomerase (PDI)** é uma enzima amplamente distribuída que catalisa a interconversão, ou modificação, de ligações dissulfeto até que as ligações da conformação nativa sejam formadas. Entre suas funções, a PDI catalisa a eliminação de intermediários dobrados com ligações dissulfeto transversais inadequadas. A **peptídeo-prolil-*cis-trans*-isomerase (PPI)** catalisa a interconversão de isômeros *cis* e *trans* de ligações peptídicas de resíduos de Pro (Figura 4-8), que pode representar um passo lento no enovelamento de proteínas que contêm algumas ligações peptídicas com Pro na conformação *cis*.

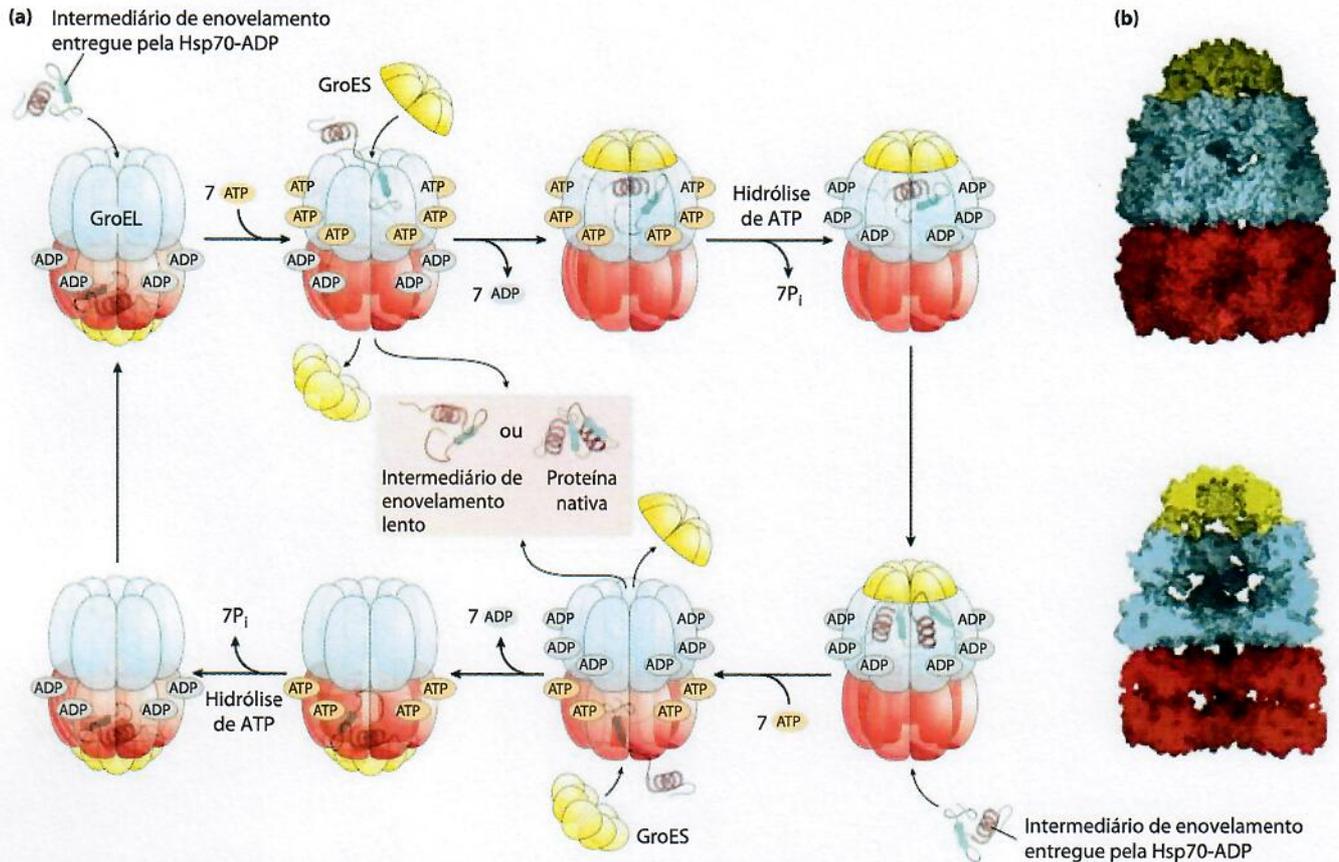


FIGURA 4-31 As chaperoninas no enovelamento de proteínas. (a) Mecanismo de ação proposto para as chaperoninas GroEL (membro da família de proteínas Hsp60) e GroES de *E. coli*. Cada complexo GroEL consiste em dois anéis heptaméricos que formam duas câmaras grandes (cada subunidade com *M_r* 57.000). A GroES, também um heptâmero (subunidades de *M_r* 10.000), bloqueia a entrada de uma das câmaras na GroEL depois da ligação, dentro da câmara, de uma proteína desdobrada. A câmara com a proteína desdobrada é designada como *cis*; o oposto é designado *trans*. O enove-

lamento ocorre no interior da câmara *cis*, durante o tempo que leva para hidrolisar os 7 ATP ligados às subunidades no anel heptamérico. As moléculas de GroES e de ADP então se dissociam, e a proteína é liberada. As duas câmaras do sistema GroEL/Hsp60 alternam a ligação e facilitam o enovelamento das proteínas clientes. (b) Imagens da superfície e corte longitudinal do complexo GroEL/GroES (PDB ID 1AON). O corte (abaixo) ilustra o grande espaço interno no qual as outras proteínas se ligam.

Defeitos no enovelamento proteico fornecem a base molecular para uma ampla gama de doenças genéticas em humanos

Apesar da participação de muitos processos que assistem o enovelamento proteico, erros podem ocorrer. De fato, o enovelamento proteico errado é um problema substancial para a célula, e um quarto ou mais de todos os polipeptídeos sintetizados é destruído por não se dobrar corretamente. Em alguns casos, os erros de enovelamento causam ou contribuem para o desenvolvimento de doenças graves.

Muitas condições, incluindo diabetes do tipo 2, doença de Alzheimer, doença de Huntington e doença de Parkinson, estão associadas com um mecanismo de enovelamento errôneo: uma proteína solúvel normalmente secretada pela célula é secretada em um estado de enovelamento errado, sendo convertida em uma fibra **amiloide** extracelular insolúvel. As doenças são coletivamente chamadas de **amiloidoses**. As fibras são altamente ordenadas e não ramificadas, com um diâmetro de 7 a 10 nm e alto conteúdo de estruturas do tipo folha β. Os segmentos β são orientados de forma perpendicular ao eixo da fibra. Em algumas fibras amiloides, a estrutura

geral apresenta duas camadas de folhas β longas, como aquela mostrada para o peptídeo β-amiloide, na **Figura 4-32**.

Muitas proteínas podem assumir a estrutura da fibrila amiloide como alternativa para sua conformação dobrada normal, e a maioria dessas proteínas tem grande concentração de resíduos de aminoácidos aromáticos em uma região central das folhas β ou hélice α. As proteínas são secretadas em uma conformação não totalmente dobrada. O núcleo (ou parte dele) se dobra em uma folha β antes que o resto da proteína se dobre corretamente, e as folhas β de duas ou mais moléculas de proteínas com enovelamento incompleto se associam para começar a formação de uma fibrila amiloide. A fibrila cresce no espaço extracelular. Outras partes da proteína, então, se dobras de forma diferente, ficando externas ao núcleo de folhas β da fibrila em formação. O efeito dos aminoácidos aromáticos na estabilização da estrutura é mostrado na Figura 4-32c. A maioria das moléculas proteicas se dobra corretamente; por isso, o surgimento dos sintomas das amiloidoses geralmente é muito lento. Se uma pessoa herda uma mutação, como a substituição de um resíduo aromático em uma posição que favorece a formação de fibrilas amiloides, os sintomas da doença podem começar mais cedo.

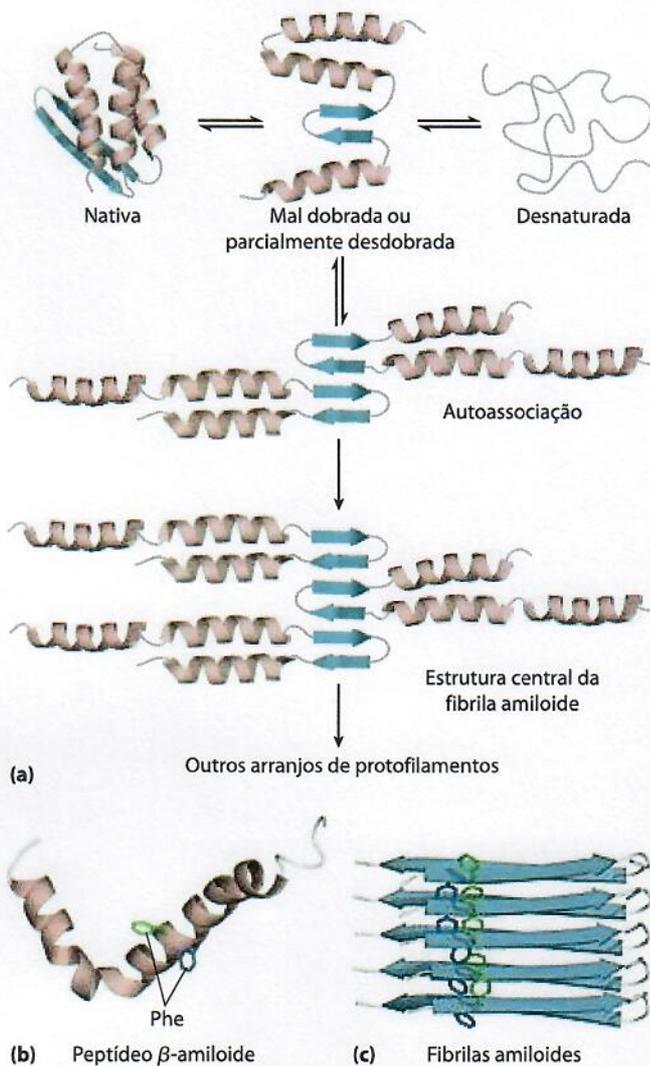


FIGURA 4-32 Formação das fibrilas amiloides causadoras de doenças. (a) Moléculas de proteínas cuja estrutura normal inclui regiões de folhas β passam por enovelamento parcial. Em um pequeno número de moléculas, antes do enovelamento completo, as regiões de folhas β de um polipeptídeo se associam com a mesma região de outro polipeptídeo, formando o núcleo de um amiloide. Outras moléculas de proteína lentamente se associam ao amiloide e o estendem para formar uma fibrila. (b) O peptídeo β-amiloide inicia como dois segmentos α-helicoidais de uma proteína maior. A clivagem proteolítica desta proteína maior libera o peptídeo β-amiloide relativamente instável, que perde sua estrutura α-helicoidal. Ele pode, então, lentamente ser organizado em fibrilas amiloides (c), que contribuem para as placas características no exterior do tecido nervoso das pessoas com doença de Alzheimer. As cadeias laterais aromáticas mostradas aqui exercem um papel significativo na estabilização da estrutura amiloide. A estrutura amiloide é rica em folhas β, com as fitas β arranjadas perpendicularmente ao eixo da fibrila amiloide. O peptídeo β-amiloide assume a forma de duas camadas de folhas β paralelas estendidas. Alguns peptídeos formadores de amiloide podem se dobrar na forma de hélices β voltadas para esquerda (ver Figura 4-22).

Em eucariotos, as proteínas destinadas à secreção iniciam seu enovelamento no retículo endoplasmático (RE; conferir via no Capítulo 27). Quando ocorrem condições de estresse ou quando a síntese proteica ameaça sobrecarregar a capacidade de enovelamento proteico do RE, pode ocorrer o acúmulo de proteínas desdobradas. Essas condições disparam a resposta a proteínas não dobradas (UPR,

de *unfolded protein response*). Um conjunto de reguladores transcricionais que constituem a UPR alinha os vários sistemas por aumentar a concentração de chaperonas no RE ou por diminuir a taxa global de síntese proteica, ou ambos. Os agregados amiloides que se formam antes da UPR podem ser removidos. Alguns são degradados por **autofagia**. Nesse processo, eles são primeiro encapsulados em uma membrana, e, então, o conteúdo da vesícula resultante é degradado após a fusão com um lisossomo citosólico. Alternativamente, as proteínas erroneamente dobradas podem ser degradadas por um sistema de proteases chamado sistema ubiquitina-proteossomo (descrito no Capítulo 27). Defeitos em qualquer um desses sistemas diminuem a capacidade de lidar com proteínas erroneamente dobradas e aumentam a propensão para o desenvolvimento de doenças amiloides.

Algumas amiloidoses são sistêmicas, envolvendo vários tecidos. A mais simples é causada pela deposição de fibrilas formadas por cadeias leves de imunoglobulinas com enovelamento incorreto (ver Capítulo 5), ou fragmentos de cadeias leves derivados da degradação proteolítica. A idade média de aparecimento é de 65 anos. Os pacientes têm sintomas como fadiga, rouquidão, inchaço e perda de peso, e muitos morrem dentro de um ano após o diagnóstico. Os rins e o coração geralmente são os órgãos mais afetados. Algumas amiloidoses estão associadas com outros tipos de doenças. Indivíduos com certas infecções crônicas ou doenças inflamatórias como artrite reumatoide, tuberculose, fibrose cística e alguns cânceres podem ter um grande aumento de secreção de um polipeptídeo pró-amiloide chamado de proteína amiloide sérica A (SAA, do inglês *serum amyloid A*). Essa proteína, ou fragmentos dessa proteína, se deposita no tecido conectivo do baço, dos rins, do fígado e ao redor do coração. Indivíduos com essa condição, conhecida como amiloidose sistêmica secundária, apresentam inúmeros sintomas, dependendo dos órgãos inicialmente afetados. A doença geralmente é fatal dentro de poucos anos. Mais de 80 amiloidoses estão associadas com mutações na transtirretina (proteína que se liga aos hormônios da tireoide e os transporta, distribuindo-os pelo corpo e cérebro). Uma grande variedade de mutações nessa proteína leva à deposição amiloide, concentrada ao redor de diferentes tecidos, produzindo assim diferentes sintomas. As amiloidoses também estão associadas com mutações hereditárias nas proteínas lisozima, cadeia α do fibrinogênio A e apolipoproteínas A-I e A-II, todas descritas nos capítulos posteriores.

Algumas doenças amiloides estão associadas com órgãos específicos. A proteína pró-amiloide geralmente é secretada somente pelos tecidos afetados, e sua alta concentração local leva à deposição de amiloide ao redor do tecido (embora algumas proteínas possam estar sistematicamente distribuídas). Um lugar comum de deposição de amiloide é próximo às células das ilhotas β pancreáticas, responsáveis pela secreção de insulina e regulação do metabolismo de glicose (ver Figura 23-26). A secreção pelas células β de um pequeno peptídeo (37 aminoácidos), chamado de polipeptídeo amiloide das ilhotas (PPAI, ou IAPP, do inglês *islet amyloid polypeptide*), pode levar à deposição de amiloide ao redor das ilhotas, destruindo gradualmente as células. Um adulto saudável tem de 1 a 1,5 milhão de células β pancreáticas. Com a perda progressiva dessas células, a home-

QUADRO 4-6



MEDICINA

Morte por envelamento errado: as doenças priônicas

Uma proteína cerebral dobrada de forma errada parece ser o agente causador de doenças cerebrais neurodegenerativas raras em mamíferos. Talvez a mais conhecida seja a encefalopatia espongiforme bovina (EEB, ou BSE, do inglês *bovine spongiform encephalopathy*; também conhecida como doença da vaca louca). Doenças relacionadas incluem a kuru e a doença de Creutzfeldt-Jakob em humanos, *scrapie* em ovinos, e doença debilitante crônica em cervos e alces. Essas doenças também são conhecidas como encefalopatias espongiformes porque o cérebro doente frequentemente se torna cheio de buracos (Figura Q-1). A deterioração progressiva do cérebro leva a um espectro de sintomas neurológicos, incluindo perda de peso, comportamento errático, problemas de postura, equilíbrio e coordenação, e perda da capacidade cognitiva. Essas doenças são fatais.

Nos anos de 1960, pesquisadores descobriram que amostras de agentes causadores de doença pareciam não conter ácidos nucleicos. Naquela época, Tikvah Alper sugeriu que o agente fosse uma proteína. Inicialmente, a ideia pareceu uma heresia. Todos os agentes causadores de doenças conhecidos até aquele momento – vírus, bactérias, fungos, e assim por diante – continham ácidos nucleicos, e sua virulência estava relacionada à reprodução genética e propagação. Entretanto, quatro décadas de investigações, realizadas mais notavelmente por Stanley Prusiner, forneceram evidências de que as encefalopatias espongiformes são diferentes.

Os agentes infecciosos foram identificados como uma única proteína (M_r 28.000), que Prusiner apelidou de proteína **prion** (PrP). O nome foi derivado de *proteinaceous infectious* (proteína infecciosa), mas Prusiner

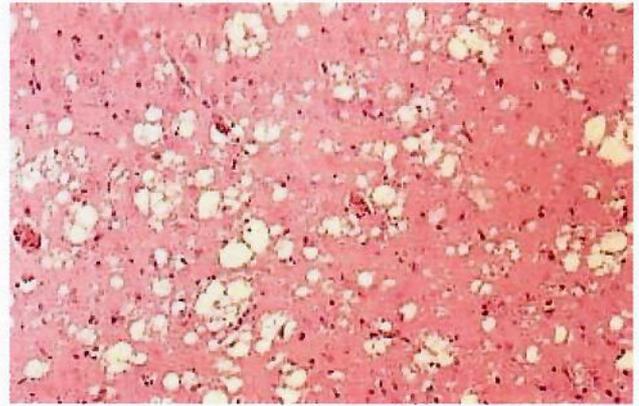


FIGURA Q-1 Secção corada de córtex cerebral, da necropsia de um paciente com a doença de Creutzfeldt-Jakob, mostra a degeneração espongiforme (vacuolar), a característica neuro-histológica mais comum. Os vacúolos amarelados são intracelulares e ocorrem, majoritariamente, em processos pré e pós-sinápticos de neurônios. Os vacúolos nesta secção variam de 20 a 100 μm em diâmetro.

achou que “prion” soava melhor do que “proin”. A proteína prion é um constituinte normal do tecido cerebral em todos os mamíferos. Seu papel não é conhecido em detalhes, mas deve ter uma função de sinalização molecular. Linhagens de camundongos sem o gene para PrP (e, por isso, sem a proteína) sofrem efeitos maléficos não óbvios. A doença ocorre somente quando a PrP celular normal, ou PrP^C, apresenta uma conformação alterada chamada de PrP^{Sc} (Sc significa *scrapie*). A estrutura de PrP^C caracteriza-se por duas hélices α . A estrutura de

ostase de glicose é afetada e, no final, quando 50% ou mais células são perdidas, a condição evolui para diabetes melito do tipo 2 (independente de insulina).

As doenças de deposição amiloide que induzem neurodegeneração, especialmente em adultos idosos, são uma classe especial de amiloidoses localizadas. A doença de Alzheimer está associada com deposição amiloide extracelular pelos neurônios, envolvendo o peptídeo amiloide β (Figura 4-32b), derivado de uma grande proteína transmembrana (proteína precursora de amiloide β) encontrada na maioria dos tecidos humanos. Quando ele faz parte de uma proteína maior, o peptídeo é composto por dois segmentos α -helicoidais que atravessam a membrana. Quando os domínios externos e internos são clivados por proteases específicas, o peptídeo β -amiloide relativamente instável deixa a membrana e perde sua estrutura α -helicoidal. Ele pode então tomar a forma de duas camadas de folhas β paralelas estendidas, que lentamente podem se reunir em fibrilas amiloides (Figura 4-32c). O depósito dessas fibras amiloides parece ser a causa primária da doença de Alzheimer, mas um segundo tipo de agregado do tipo amiloide, envolvendo a proteína chamada tau,

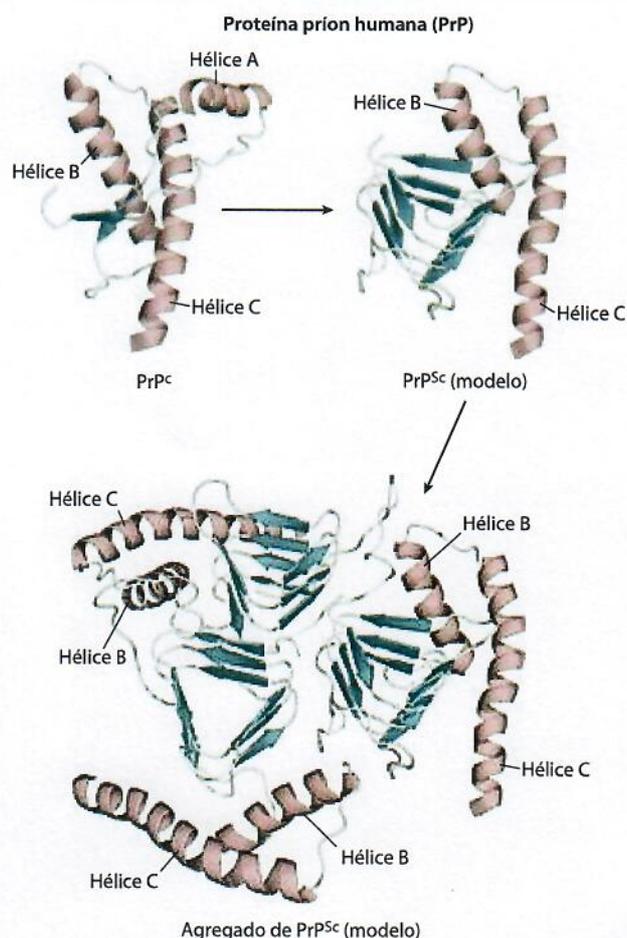
também ocorre intracelularmente (nos neurônios) em pessoas com doença de Alzheimer. Mutações hereditárias na proteína tau não resultam em Alzheimer, mas causam demência frontotemporal e parkinsonismo (condição com sintomas que lembram a doença de Parkinson), que pode ser igualmente devastadora.

Várias outras condições neurodegenerativas envolvem agregação intracelular de proteínas com envelamento errado. Na doença de Parkinson, a forma mal dobrada da proteína α -sinucleína se agrega em massas esféricas filamentosas, chamadas de corpos de Lewy. A doença de Huntington envolve a proteína huntingtina, que tem uma longa repetição de poliglutaminas. Em alguns indivíduos, essa repetição é maior do que o normal, ocorrendo um tipo de agregação intracelular mais sutil. Notavelmente, quando proteínas mutantes humanas envolvidas nas doenças de Parkinson e Huntington são expressas em *Drosophila melanogaster*, as moscas demonstram degeneração expressa como deterioração dos olhos, tremores e morte precoce. Todos esses sintomas são altamente suprimidos se a expressão da chaperona Hsp70 também estiver aumentada.

PrP^{Sc} é muito diferente, com a maior parte da estrutura convertida em folhas β semelhante à amiloide (Figura Q-2). A interação da PrP^{Sc} com a PrP^C converte a última em PrP^{Sc}, iniciando um efeito dominó no qual cada vez mais proteínas do cérebro se convertem na forma causadora da doença. O mecanismo pelo qual a presença de PrP^{Sc} resulta em encefalopatia espongiforme não é conhecido.

A forma hereditária das doenças priônicas, uma mutação no gene que codifica a PrP, produz uma mudança em um resíduo de aminoácido que acredita-se tornar a conversão de PrP^C em PrP^{Sc} mais favorável. Um completo entendimento das doenças priônicas aguarda novas informações de como uma proteína prião afeta as funções cerebrais. Informações estruturais sobre PrP estão começando a fornecer indicativos quanto ao processo molecular que permite às proteínas priônicas interagirem de forma a alterar sua conformação (Figura Q-2).

FIGURA Q-2 Estrutura do domínio globular de uma PrP humana (PDB ID 1QLX) e dos modelos de PrP^{Sc} mal dobrada, na conformação causadora de doença e de um agregado PrP^{Sc}. As hélices α estão marcadas para ajudar a ilustrar a mudança conformacional. A hélice A é incorporada na estrutura da folha β da conformação mal dobrada.



O mal enovelamento proteico não necessariamente significa que a formação de amiloide cause doenças sérias. Por exemplo, a fibrose cística é causada por defeitos em uma proteína de membrana chamada de reguladora transmembrana da fibrose cística (CFTR, do inglês *cystic fibrosis transmembrane conductance*), que atua como canal de íons cloreto. A mutação mais comum que causa a fibrose cística é uma deleção de um resíduo Phe na posição 508 da CFTR, que causa o enovelamento inapropriado da proteína. A maior parte dessa proteína é então degradada, e sua função normal é perdida (ver Quadro 11-2). Muitas das doenças relacionadas com mutações no colágeno (p. 130) também causam um enovelamento defeituoso. Um tipo particular e famoso de enovelamento errado de proteína pode ser visto nas doenças priônicas (Quadro 4-6). ■

RESUMO 4.4 Desnaturação e enovelamento das proteínas

► A manutenção do estado estável da coleção de proteínas celulares ativas necessárias em um conjunto de condições específicas – chamada proteostase – envolve um conjunto elaborado de vias e processos que dobram, redobram e degradam cadeias polipeptídicas.

- A estrutura tridimensional e a função da maioria das proteínas podem ser destruídas pela desnaturação, demonstrando uma relação entre estrutura e função. Algumas proteínas desnaturadas podem renaturar espontaneamente para formar proteínas biologicamente ativas, mostrando que as estruturas terciárias são determinadas pela sequência de aminoácidos.
- O enovelamento de proteínas nas células é geralmente hierárquico. Inicialmente, regiões de estrutura secundária podem se formar, o que é seguido pelo enovelamento em motivos e domínios. Grandes conjuntos de intermediários dobrados são rapidamente conduzidos a uma única conformação nativa.
- Para muitas proteínas, o enovelamento é facilitado por chaperonas Hsp70 e por chaperoninas. A formação de ligações dissulfeto e a isomerização *cis-trans* de ligações peptídicas contendo Pro são catalisadas por enzimas específicas.
- O enovelamento errado é a base molecular de uma grande variedade de doenças humanas, incluindo as amiloidoses.