Universidade de São Paulo

Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto

Departamento de Física e Matemática

Pós-graduação em Física aplicada à Medicina e Biologia

Inês Regina Silva

Enovelamento protéico: fatores topológicos.

v.1

Ribeirão Preto

2005

Universidade de São Paulo
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto
Departamento de Física e Matemática
Pós-graduação em Física aplicada à Medicina e Biologia

Inês Regina Silva

Enovelamento protéico: fatores topológicos.

Tese apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, da Universidade de São Paulo para obtenção do título de Doutor em Ciências.

Área de Concentração: Física aplicada à Medicina e Biologia
Orientador: Prof. Dr. Antonio Caliri

v.1

Ribeirão Preto
2005

Catalogação-na-Publicação

Serviço de Documentação Bibliográfica

Faculdade de Filosofia, Ciências e Letras da Universidade de São Paulo

SILVA, Inês Regina

Enovelamento protéico: fatores topológicos. Ribeirão Preto, 2005.

75 f.: il.; 30 cm.

Tese de Doutorado, apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto/USP – Área de concentração: Física aplicada à Medicina e Biologia.

Orientador: Antonio Caliri.

1. Enovelamento de proteína. 2. Modelo em rede. 3. Restrições estéricas. 4. Caracterização topológica.

DEDICATÓRIA

Ao meu pai, Hortêncio Silva, à minha mãe, Maria Conceição de Martin Silva, pela minha vida. Com amor, admiração e gratidão por toda paciência que tiveram comigo, por sua compreensão, carinho, presença e incansável apoio ao longo do período de elaboração deste trabalho.

Às minhas irmãs, Rosa Maria Silva Donato e Marli Cristina Silva Vidotti pela compreensão nos momentos estressantes e pelo apoio dado nos momentos difíceis da minha vida.

Aos meus cunhados, Antonio Aparecido Vidotti e Gilberto Donato Junior, pela presença nos momentos críticos.

Ao meu sobrinho Daniel Henrique Silva Donato por me proporcionar momentos de agradáveis brincadeiras e risos e à minha sobrinha Amaitê que está por chegar.

A todos os meus amigos que contribuíram direta ou indiretamente para que eu desenvolvesse este trabalho de doutorado.

AGRADECIMENTOS

Ao Prof. Dr. Antonio Caliri que, nos anos de convivência, muito me ensinou, contribuindo para meu crescimento científico e intelectual. Pela orientação e completo envolvimento com este projeto de pesquisa.

Ao Prof. Dr. José Drugowich de Felício e Prof. Dr. Nelson Augusto Alves, pela atenção inicial durante o período de definição da linha de pesquisa e de minha orientação.

À Universidade de São Paulo pela oportunidade de realização do meu curso de doutorado.

Ao Departamento de Física e Química da Faculdade de Ciências Farmacêuticas de Ribeirão Preto, por disponibilizar os computadores utilizados neste trabalho e uma sala de estudos.

A todas as pessoas que participaram direta ou indiretamente do desenvolvimento deste trabalho de doutorado.

“Se um homem tem um talento e não tem capacidade de usá-lo, ele fracassou.
Se ele tem um talento e usa somente a metade deste, ele fracassou parcialmente.
Se ele tem um talento e de certa forma aprende a usá-lo em sua totalidade, ele triunfou gloriosamente e obteve uma satisfação e um triunfo que poucos homens conhecerão”.

Thomas Wolfe

RESUMO

SILVA, I. R. **Enovelamento protéico: fatores topológicos**. 2005. 75 p. Tese (Doutorado). Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo, Ribeirão Preto, 2005.

O entendimento dos princípios básicos do enovelamento protéico pode conduzir a muitas aplicações importantes. Embora não se conheçam todos os aspectos significativos envolvidos neste problema, experimentos e aproximações teóricas têm produzido avanços relevantes na sua compreensão. Um fato experimental importante tem sido a descoberta de que o logaritmo da taxa de enovelamento log k_f se correlaciona linearmente com parâmetros estruturais globais, como a ordem de contato relativa χ. Com o propósito de contribuir para o entendimento do processo de enovelamento, o objetivo primordial deste trabalho consiste em explicar o porquê de certas proteínas não seguirem o comportamento linear entre log k_f e χ, verificado para outras proteínas da mesma classe (usualmente proteínas pequenas e com termodinâmica descrita pela aproximação de dois estados). Para isso foi necessário identificar os parâmetros topológicos da estrutura nativa que constituíssem importantes determinantes da cinética do enovelamento de proteínas globulares. Também se estudou como as especificidades estéricas dos aminoácidos afetam o processo do enovelamento de proteínas, assim como influenciam na correlação entre a ordem de contato relativo e a taxa de enovelamento.

Empregou-se neste estudo um modelo simplificado em rede cúbica, que foi tratado por meio de simulações Monte Carlo. Um conjunto de 52 estruturas maximamente compactas, correspondendo a cadeias de tamanho $L = 27$ monômeros, foi usado para representar estados nativos; estas estruturas foram escolhidas de forma a representar uma variedade significativa de padrões estruturais, independentemente de χ. Através de uma análise detalhada da influência de parâmetros topológicos das configurações nativas na cinética do enovelamento, conclui-se que a taxa de enovelamento é fortemente dependente daquilo que denominamos aqui como “conteúdo de estruturas tipo-secundárias” da estrutura nativa. Adicionalmente, observou-se que aquela (taxa), independentemente do valor da ordem de contato relativo, é fortemente influenciada pelos padrões confíguracionais e suas combinações presentes na nativa.

Por meio dessa premissa, foi então possível explicar de forma consistente os casos que não obedecem a pretensa relação linear entre log k_f e χ, levando a concluir que o logaritmo da taxa de enovelamento e a ordem de contato relativo são linearmente dependentes somente para aquelas configurações em que há uma certa quantidade equilibrada (que depende de χ) de padrões estruturais, mesclando contatos efetivos de curto alcance (alto conteúdo de estruturas tipo-secundárias), com outros de longo alcance (baixo conteúdo de estruturas tipo-secundárias). Estruturas nativas que quebram este equilíbrio têm sua cinética de enovelamento afetada com respeito à reta de regressão linear ajustada para o conjunto de todas as configurações consideradas. Dessa forma, verificou-se que o mecanismo físico básico que relaciona o conteúdo de estruturas tipo-secundárias e a taxa de enovelamento, envolve o conceito de cooperatividade: se a estrutura nativa é rica em combinações de padrões estruturais ricos em contatos efetivos de curto alcance, o processo de enovelamento é mais rápido porque contatos locais são naturalmente estimulados por flutuações térmicas.

ABSTRACT

SILVA, I. R. **Protein folding: topological determinants**. 2005. 75 p. Thesis (Doctoral). Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo, Ribeirão Preto, 2005.

The understanding of basic principles of the protein folding problem can lead to many important applications. Although not all the involved significant aspects of this problem are known, experiments and theoretical approaches have produced important advances in its understanding. An important experimental fact has been the discovery that the logarithm of the folding rate log k_f correlates linearly with global structural parameters, like the relative contact order χ. In order to contribute for the understanding of folding process, the primordial goal of this work consists in to explain why certain proteins do not follow the linear behavior between log k_f and χ, as verified to other proteins from the same class (usually small two states proteins). For this, it was necessary to identify those topological parameters of the native structure that are important to the folding kinetic of globular protein. It was also studied how steric specificities of the aminoacids affect the protein folding process, as well how they influence the correlation between the relative contact order and the folding rate.

It was employed in this study a simplified cubic lattice model, treated by Monte Carlo simulation. A set of 52 maximum compact structures, corresponding to chains of size L = 27 monomers, was used to represent the native states; these structures were chosen in such a way to represent a significant diversity of structural patterns, independently of χ. Through a detailed analysis of the influence of topological parameters of the native configurations on the folding kinetic, it was concluded that the folding rate is strongly dependent of what we call here as “content of type-secondary” of the native. Additionally, it was observed that log k_f is, independently of χ, strongly influenced by the configurational patterns and its combinations in the native.

Through this premise it was possible to consistently explain the cases that do not obey the pretense linear relation between log k_f and χ, leading to conclude that the logarithm of the folding rate and the relative contact order are linearly related only for those configurations in that there is a certain balanced amount of structural patterns (which depend on χ) mixing short-range effective contacts (high contents of secondary-

type structures) and long-range contacts (low contents of secondary-type structures). Structures that break this balance have its folding kinetic affected with respect to the linear fitting adjusted for the set of all the considered configurations. Of this form, it was verified that basic physical mechanism that relates the content of type-secondary structures and the folding rate involves the cooperativety concept: if the native structure presents combinations of structural standards rich in effective contacts of short-range, the folding process is faster because local contacts are naturally stimulated by thermal fluctuations.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE GRÁFICOS

LISTA DE SIGLAS

LISTA DE SÍMBOLOS

SUMÁRIO