

O IMPACTO DO PARÂMETRO DE DIFUSÃO NO TEMPO DE PASSAGEM DO PROCESSO DE ENOVELAMENTO DE PROTEÍNAS

RICOTTA, R. M.

*Fatec São Paulo - Departamento de Ensino Geral
regina.ricotta@fatec.sp.gov.br*

The Impact of the Diffusion Parameter on the Passage Time of the Protein Folding Process

Eixo Tecnológico: Controle e Processos Industriais.

Resumo

As proteínas são entidades funcionais em todos os sistemas vivos, formadas por uma sequência linear de aminoácidos. Para exercer sua função biológica a proteína é sintetizada e assume uma estrutura tridimensional. À sequência linear associamos o estado desenovelado, atingindo seu estado funcional numa estrutura tridimensional, referida como estado enovelado. Para chegar a esta estrutura final, esta sequência linear pode passar por conformações intermediárias para se tornar funcional. O comportamento cinético de enovelamento é importante para identificar como o tempo de passagem do estado desenovelado para o estado enovelado é regulado pelo estado intermediário. O estado enovelado corresponde ao estado de energia mínima que é o mais acessível do ponto de vista cinético. Recentemente, um método matemático para resolver a equação de Fokker Plank (EFP) possibilitou a análise da cinética de enovelamento de proteínas, através da construção da evolução temporal da densidade de probabilidade. Uma função potencial tri-estável simétrica foi usada para descrever os estados desenovelado e enovelado da proteína, bem como um estado intermediário da proteína. Neste trabalho são revisados os principais pontos da metodologia, baseada no formalismo algébrico da Mecânica Quântica Supersimétrica (MQS), e são apresentados os últimos resultados sobre a cinética de evolução do sistema caracterizado em termos do parâmetro de difusão.

Palavras-chave: *Mecânica Quântica Supersimétrica, Método Variacional, Enovelamento de proteínas, Difusão.*

Abstract

Proteins are functional entities in all living systems, formed by a linear sequence of amino acids. To exercise its biological function, the protein is synthesized and assumes a three-dimensional structure. We associate the unfolded state with the linear sequence, reaching its functional state in a three-dimensional structure, referred to as the folded state. To arrive at this final structure, this linear sequence can undergo intermediate conformations to become functional. The kinetic behavior of folding is important to identify how the transition time from the unfolded state to the folded state is regulated by the intermediate state. The folded state corresponds to the minimum energy state that is the most accessible from the kinetic point of view. Recently, a mathematical method to solve the Fokker Plank equation (FPE) allowed the analysis of protein folding kinetics, through the construction of the temporal evolution of the probability density. A symmetric tri-stable potential function was used to describe the unfolded and folded states of the protein, as well as an intermediate state of the protein. This article reviews the main points of the methodology, based on the algebraic formalism of Supersymmetric Quantum Mechanics (SQM), and presents new results on the kinetics of evolution of the system characterized in terms of the diffusion parameter.

Keywords: *Supersymmetric Quantum Mechanics, Variational Method, Protein folding, Diffusion.*

1. Introdução

Proteínas são estruturas funcionais codificadas por uma sequência de aminoácidos. No estado desenovelado a proteína apresenta uma configuração linear de aminoácidos e é sintetizada em uma estrutura tridimensional enovelada para desempenhar funções específicas

Anais da VII Mostra de Docentes em RJJ

no organismo. Para chegar a essa estrutura final, essa sequência linear pode passar por conformações intermediárias, que são estados de transição nos quais a proteína pode não atingir seu estado enovelado e se aglomerar, criando estados amorfos, fibras amorfas que podem causar uma série de problemas relacionados ao mau funcionamento de proteínas. A importância de estudar enovelamento reside em compreender como uma estrutura linear desenovelada atinge uma estrutura enovelada tridimensional e funcional. O assunto tem sido amplamente estudado nas últimas décadas sob o conceito de funil de enovelamento, onde o cenário energético tem o formato de um funil, [1]-[3].

Em um trabalho anterior, [4], foi introduzido um modelo matemático consistente para descrever fisicamente o processo biológico de enovelamento de proteínas. A abordagem considera o processo como um modelo de difusão inspirado no conceito de funil de enovelamento, visando analisar seu comportamento dinâmico. O processo de enovelamento de proteínas é descrito pela equação de Fokker-Planck, EFP, associada a uma energia livre descrita por uma função potencial simétrica tri-estável $V(x)$. Por sua vez, a EFP pode ser mapeada para uma equação do tipo Schrödinger, ES, [5]-[6], ou seja, ambas as equações compartilham o mesmo espectro. Neste ponto a metodologia da Mecânica Quântica Supersimétrica, MQS, associada ao método variacional, [7]-[8], é utilizada para obter o espectro aproximado de energia e autofunções da ES e para avaliar a densidade de probabilidade dependente do tempo, $P(x, x_0, t)$, solução FPE, onde x é a coordenada da reação, [9], e a coordenada x_0 está associada à proteína no estado desenovelado.

A energia livre, dada pela função potencial tri-estável $V(x)$, é uma função simétrica que possui mínimos laterais com a mesma profundidade (poços simétricos) que podem ser interpretados, respectivamente, como os estados da proteína enovelada e desenovelada; o mínimo central está relacionado a uma conformação proteica intermediária. A cinética do processo de difusão foi caracterizada pelo cálculo da população de partículas do poço direito (estado enovelado). O tempo necessário para a evolução da população do sistema desde o seu estado inicial até o poço à direita é utilizado como o tempo característico de passagem do sistema até o estado de enovelamento da proteína. Os resultados em [4] são consistentes com aqueles esperados em problemas de difusão semelhantes, [10].

Neste trabalho é apresentada uma breve revisão da metodologia (Seção 2.2) mostrando a ligação entre a EFP e a ES, dada em termos da energia livre $V(x)$. O modelo é ilustrado por uma função específica de energia livre, um estudo de caso diferente daquele apresentado em [4], (Seção 2.3). A Seção 3 contém um mapeamento da correlação entre o parâmetro de difusão e o tempo de passagem para o enovelamento e sua influência no perfil simétrico de energia livre realizado. O objetivo deste estudo foi analisar a forma como o aumento da difusão impacta o tempo de passagem pelo estado intermediário para o estado enovelado da proteína. A Seção 4 contém os resultados e as conclusões, seguida pelas considerações finais.

2. Materiais e métodos

2.1. Materiais

O processo de enovelamento proteico é considerado um processo de difusão, descrito pela EFP. A distribuição de probabilidade, solução da EFP, é encontrada por um mapeamento em uma ES, cujas soluções são obtidas pelo método variacional associado à MQS. A energia livre a ser utilizada é descrita por um potencial tri-estável. Por ser dependente do tempo, a distribuição de probabilidade descreve um tempo específico para a dinâmica de

Anais da VII Mostra de Docentes em RJJ

desenvolvimento de proteínas através de uma estimativa do tempo de passagem em função da coordenada de reação. Verifica-se o comportamento da população em direção ao terceiro poço, o que caracteriza o estado enovelado, em função do tempo, como um decaimento exponencial característico de processos difusivos com força direcional, [4].

2.2. Metodologia: Formalismo da EFP e a ES

A EFP, que descreve a evolução temporal da distribuição de probabilidade $P(x,t)$ em sistemas de difusão é dada por:

$$\frac{d^2}{dx^2} \Psi(x) - \frac{1}{2Q} \left(\frac{f(x)^2}{2Q} + \frac{df(x)}{dx} \right) \Psi(x) = \frac{\lambda}{Q} \Psi(x) \quad (1)$$

x é a variável característica do sistema, a coordenada de reação (número de contatos nativos); t é a variável de tempo; Q é o coeficiente de difusão e $f(x)$ representa uma força externa (força motriz) atuando no meio, está associada à energia livre do meio, o potencial tri-estável $V(x)$,

$$f(x) = -\frac{d}{dx} V(x). \quad (2)$$

Escrevendo a probabilidade $P(x,t)$ como um produto de uma função de posição e uma função dependente de t ,

$$P(x, t) = \Psi(x) e^{-\lambda t} \quad (3)$$

pode-se mostrar que as soluções da EFP são soluções de uma equação do tipo Schrödinger independente do tempo, ES, dada por

$$\frac{d^2}{dx^2} \Psi(x) - \frac{1}{2Q} \left(\frac{f(x)^2}{2Q} + \frac{df(x)}{dx} \right) \Psi(x) = \frac{\lambda}{Q} \Psi(x) \quad (4)$$

onde λ é proporcional à energia E . Expandindo numa base ortonormal, obtemos a distribuição de probabilidade dada por

$$P(x, t|x_0, t_0) = \frac{\Psi_0(x)}{\Psi_0(x_0)} \sum_{n=0}^{\infty} \Psi_n(x) \cdot \Psi_n(x_0) \cdot e^{-\lambda_n(t-t_0)} \quad (5)$$

onde λ é a função de onda do estado fundamental e x_0 é a posição inicial. A ES usada na MQS, escrita em unidades reduzidas, é dada por

$$-\frac{d^2}{dx^2} \Psi(x) + \underbrace{\left(W_1(x)^2 - \frac{dW_1(x)}{dx} + E_0^{(1)} \right)}_{V_{SE}(x)} \Psi(x) = E \Psi(x) \quad (6)$$

onde $V_{SE}(x)$ é a função potencial de Schrödinger definido em termos da função superpotencial, [7]. Comparando as equações (4) e (6) e considerando que a relação de $f(x)$ com $V(x)$ dada pela Eq. (2), obtemos

$$W_1(x) = \frac{1}{2Q} \frac{dV(x)}{dx}, \quad (7)$$

ou seja, o potencial da EFP está relacionado com o superpotencial $W_1(x)$ da MQS. Além disso, a energia E está relacionada ao parâmetro como

$$E = \frac{\lambda}{Q} \quad (8)$$

Assim, a metodologia da MQS associada ao método variacional pode ser utilizada, [4]. Neste ponto é importante observar a relação entre a função potencial da ES, Eq. (6) com o parâmetro de difusão Q , através do superpotencial na Eq. (7),

$$V_{SE}(x) = W_1(x)^2 - \frac{dW_1(x)}{dx} + E_0^{(1)}. \quad (9)$$

Em outras palavras, ao utilizar a metodologia MQS o espectro é explicitamente dependente do valor da constante de difusão Q .

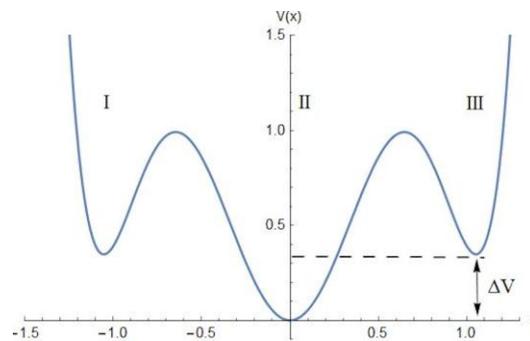
2.3. Estudo de Caso para um parâmetro de difusão Q fixo

O potencial tri-estável utilizado é dado por

$$V(x) = ax^6 - 8,93851x^4 + 5,42373x^2 \quad (10)$$

ilustrado na Fig. 1. Os mínimos laterais (V_{min}) têm a mesma profundidade (poços simétricos) e são interpretados, respectivamente, como os estados desenovelado (poço esquerdo) e enovelado (poço direito) da proteína, e o mínimo central está relacionado a um conjunto de conformações intermediárias da proteína, com $\Delta V = V(0) - V_{min}$. A escolha dos parâmetros foi feita de forma a um potencial simétrico com $V(0) = 0$, como em [10].

Fig. 1 - Representação da função potencial tri-estável, Eq. (10) com o respectivo valor de a .



Fonte: [4].

Uma vez definida a energia livre dada pelo potencial $V(x)$, voltamos à construção do espectro da ES equivalente, solução da Eq. (6), com constante de difusão fixa, $Q = 0,5$. Assim, utilizando a metodologia da MQS associada ao método variacional, [7]-[8], o espectro aproximado de energias e autofunções foi calculado e é mostrado pela Tab. 1 e Tab. 2. Deve-

Anais da VII Mostra de Docentes em RJI

se ressaltar que como estamos lidando com um método aproximado, o número de termos na expansão de probabilidade, Eq. (5), foi fixado em seis termos da série, ($n = 0, \dots, 5$), já que a contribuição do próximo termo exponencial é várias ordens de grandeza menor que o termo anterior e, portanto, pode ser desprezado.

Tab. 1 - Valores do espectro de energia da ES quando a função potencial (energia livre) é $V(x) = 3,90456 x^6 - 8,93851 x^4 + 5,42373 x^2$.

n	0	1	2	3	4	5
λ_n	0	1,0361	1,9797	9,2517	17,3589	27,4762

Fonte: Elaboração própria.

Tab. 2 - Funções de onda do espectro de energia da ES quando a função potencial (energia livre) é $V(x) = 3,90456 x^6 - 8,93851 x^4 + 5,42373 x^2$.

$$\Psi_0^{(1)}(x) = 1,021e^{(-5,42373x^2+8,93851x^4-3,90456x^6)}$$

$$\Psi_1^{(1)}(x) = e^{(2,73824x^2-3,05175x^4-0,777117x^6)}x(1,99656 - 8,75317x^2 + 10,4419x^4)$$

$$\Psi_2^{(1)}(x) = e^{(-3,68104x^2+1,08564x^4-2,13171x^6)}(-0,904686 + 5,15415x^2 - 9,43439x^4 + 33,953x^6 - 199,101x^8 + 303,278x^{10})$$

$$\Psi_3^{(1)}(x) = e^{(-3,50991x^2-1,40729x^4-0,64738x^6)}(-5,23999x + 1,07696x^3 + 20,5988x^5 - 74,5237x^7 + 66,0048x^9 - 36,7167x^{11} + 73,6961x^{13} + 69,4599x^{15})$$

$$\Psi_4^{(1)}(x) = e^{(-4,601x^2-1,49668x^4-0,559156x^6)}(0,928068 - 17,7878x^2 + 1,05918x^4 + 5,10868x^6 - 84,9985x^8 - 0,86348x^{10} - 24,6105x^{12} + 118,086x^{14} + 180,684x^{16} + 126,316x^{18} + 44,1023x^{20})$$

$$\Psi_5^{(1)}(x) = e^{(-5,52086x^2-1,54808x^4-0,53713x^6)}(4,7712x + 0,962939x^3 - 63,0876x^5 - 155,71x^7 - 181,732x^9 + 13,8384x^{11} + 423,621x^{13} + 738,097x^{15} + 712,503x^{17} + 438,794x^{19} + 175,5x^{21} + 42,7624x^{23} + 5,15394x^{25})$$

Fonte: Elaboração própria.

3. Difusão

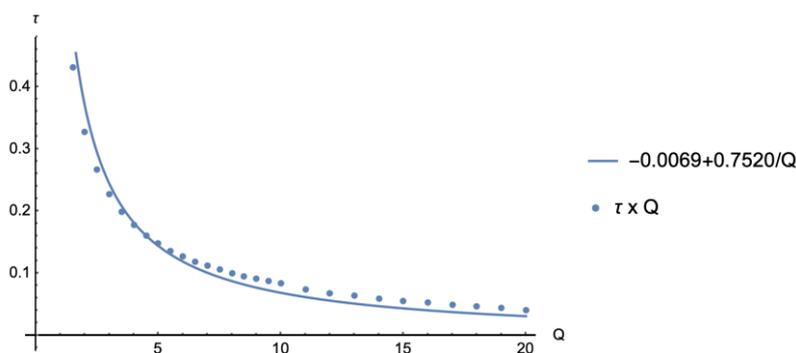
Utilizando a metodologia desenvolvida em [4] para o processo de enovelamento de proteínas, o tempo de passagem característico τ foi avaliado para diferentes valores do parâmetro de difusão Q para a energia livre definida pela Eq. (10), ilustrada pela Fig. 1. Para cada valor fixo de Q no intervalo $0,4 < Q < 20$, foi calculado o tempo de passagem τ da evolução da população do poço da esquerda (estado desenovelado) até o poço direito (estado enovelado), partindo da posição inicial fixa, conforme mostrado na Fig. 2 (curva pontilhada).

Anais da VII Mostra de Docentes em RJJ

A Fig. 2 também mostra o melhor ajuste analítico para os resultados (linha contínua), tendo sido obtido um comportamento proporcional a uma função de $1/Q$.

A Fig. 2 mostra que o tempo de passagem τ diminui à medida que a difusão aumenta, como esperado. O comportamento geral da curva de τ versus Q é uma função proporcional a $1/Q$ compatível com aquela obtida por outro método, que utiliza a aproximação do estado estacionário, [10].

Fig. 2 - Tempo de passagem τ versus constante de difusão Q para (Figura 1) para resultados analíticos (linha pontilhada) e melhor ajuste (linha sólida).



Fonte: Elaboração própria.

4. Resultados e Discussão

O ponto principal deste trabalho foi determinar explicitamente a dependência do tempo de passagem (τ) com o parâmetro de difusão Q . A curva geral τ versus Q obtida (Fig. 2) é uma função proporcional a $1/Q$ que é compatível com aquela obtida por outro método, [10].

Os resultados obtidos reforçam a aplicação do método matemático MQS proposto para problemas de envelhecimento de proteínas, principalmente na determinação da densidade de probabilidade $P(x, t)$, resolvendo a EFP através de sua relação com a ES. O tempo de passagem do processo de envelhecimento é um ingrediente importante para a cinética da reação e os resultados são consistentes com os obtidos em [10]. Nesta referência apenas o estado fundamental da ES é utilizado o que faz com que a densidade de probabilidade dependa apenas de x e não explicitamente de t , ou seja, este método utiliza apenas o estado estacionário. Nossa abordagem utiliza mais termos na expansão da Eq. (5), que faz com que a dependência temporal na distribuição de probabilidade $P(x, t)$ apareça explicitamente, permitindo o cálculo do tempo de passagem.

5. Considerações finais

Este trabalho faz parte da colaboração com Prof. Dr. Elso Drigo Filho e Prof. Dr. Jorge Chahine do IBILCE, campus de São José do Rio Preto da UNESP e Prof. Dr. Marcelo Tozo Araujo da UNILAGO, São José do Rio Preto, SP. O trabalho em andamento envolve a utilização de dados de energia livre provenientes de simulação de dinâmica molecular de proteínas. Através de um ajuste analítico é possível se obter função que descreve a energia livre e, assim calcular o espectro de funções de onda e energias através da metodologia aqui utilizada.

Anais da VII Mostra de Docentes em RJI

Tendo obtido as soluções da ES e conseqüentemente da EFP, o tempo de passagem para o envelhecimento pode ser realizado.

Referências

- [1] WOLYNES, P. G. et al. Navigating the Folding Routes, **Science**, v.267, p.1619-1620. 1995. <https://doi.10.1126/science.7886447>
- [2] BEST, R. B., HUMMER, G. Diffusion models of protein folding, **Phys. Chem. Chem. Phys.** v.13, p. 16902-16911. 2011. DOI: 10.1039/c1cp21541h
- [3] KRIVOV, S. V., The Free Energy Landscape Analysis of Protein, (FIP35) Folding Dynamics, **J. Phys. Chem.**, v. B 115, p. 12315–12324. 2011. [dx.doi.org/10.1021/jp208585r](https://doi.org/10.1021/jp208585r).
- [4] DRIGO FILHO, E. et al. Probability distribution to obtain the characteristic passage time for different tri-stable potentials. **Physica A: Statistical Mechanics and its Applications**, v.606, p.128121. 2022. <https://doi.org/10.1016/j.physa.2022.128121>.
- [5] RISKEN, H., **The Fokker-Planck Equation Methods of Solution and Applications**, Springer, (1996).
- [6] TOMÉ, T., OLIVEIRA, M. J., **Stochastic Dynamics and Irreversibility**, Springer, (2015).
- [7] DRIGO FILHO, E.; RICOTTA, R. M. Morse potential energy spectra through the variational method and supersymmetry. **Phys. Lett. A**, v.269, n.5-6, p.269-276. 2000. [https://doi:10.1016/S0375-9601\(00\)00267-X](https://doi:10.1016/S0375-9601(00)00267-X).
- [8] BORGES, G. R. P. et al. Variational supersymmetric approach to evaluate Fokker–Planck probability, **Physica A**, v.389, p.3892-3899. 2010. <https://doi.org/10.1016/j.physa.2010.05.027>.
- [9] POLOTTO, F. et al., Supersymmetric quantum mechanics method for the Fokker-Planck equation with applications to protein folding dynamics, **Physica A**, v.493, p.286-300. 2018. <https://doi.org/10.1016/j.physa.2017.10.021>.
- [10] WAGNER, C. et al., Intermediates can accelerate protein folding, **Proc. Natl. Acad. Sci. USA**, v.96 p. 6716-6721. 1999. <https://doi.org/10.1073/pnas.96.12.6716>.