

Análise Fractal de Proteínas de Repetição

Fernando S. Silva (Doutorando - MCTI), fermath@gmail.com;

Marcelo A. Moret (Orientador - MCTI), mamoret@gmail.com;

Faculdade SENAI CIMATEC

Palavras Chave: *Fractalidade, Enovelamento, Flory, Proteína.*

Introdução

As proteínas são longas sequências de aminoácidos que desempenham diversas funções biológicas específicas dentro de um organismo. Uma parte significativa das proteínas têm segmentos repetidos de aminoácidos podendo ou não ser múltiplos inteiros.

De acordo com Marcotte et al (1999) cerca de 14% de todas as proteínas possuem segmentos repetidos. Tais repetições de aminoácidos quando ocorre *in tandem* (lado a lado) desencadeiam estruturas terciárias regulares, exatamente pela sua multiplicidade, de diversos tamanhos e funções acarretando em diversas classificações de famílias de proteínas de repetição (ANDRADE; PEREZ-IRATXETA; PONTING, 2001).

Uma grande vantagem das proteínas de repetição apontada por Mosavi et al 2004 são os percursos dobráveis que são passíveis de projetar, em geral, não encontrado nas proteínas globulares. Para compreender a melhor rota de dobramento das proteínas se faz necessário elaborar ferramentas para deduzir ou prever a estrutura tridimensional da sequência de aminoácidos.

Diversos estudos, como por exemplo Moret et al (2005), apontam o dobramento de proteínas como um processo estocástico. Sendo que, as forças hidrofóbicas são as mais relevantes para entender enovelamento, ou seja, acionam a dobragem das proteínas. As proteínas reais enovelam-se em um tempo muito pequeno, comparativamente para uma busca aleatória pelo estado de menor energia global, isto é, dobramento de proteínas está sob controle termodinâmico, onde o único fator determinante na forma dobrada da sequência será a menor energia livre.

Para sistemas com lei de potência $M \propto R^D$, o modelo de campo médio de Flory tem a energia total, soma de um energia elástica mais uma energia repulsiva, como função do raio R .

Neste trabalho coletamos 1213 proteínas da base de dados *Protein Data Bank* (PDB) e reunidas em seis famílias conforme sugestão de Andrade (2001). No PDB fica disponibilizado livremente arquivos com

a localização espacial, em angstrom ($\text{\AA} = 10^{-10}$), de cada átomo de uma proteína no estado nativo, cujas estruturas foram determinadas por métodos experimentais.

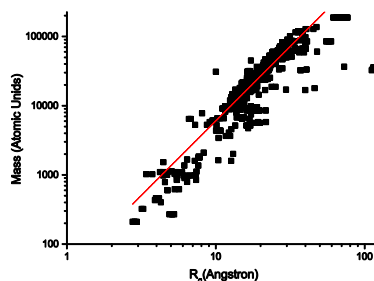
Estas famílias possuem repetições, lado a lado, de parte de sua sequência de aminoácidos com comprimento que variam de 20 a 50 aminoácidos, são elas: Ankyrin, β -Propeller, β -Tetrafoil, Tetratricopeptides, Armadillo HEAT e Leucine-rich.

Vamos analisar a Fractalidade das proteínas como ferramenta para entender o enovelamento, investigando a auto-similaridade do conjunto de proteínas utilizando a estratégia apontada por Moret et al (2005).

Resultados e Discussão

A relação *Massa-Raio* das 1213 proteínas foi analisada obtendo valor $D=2,15 \pm 0,025$, conforme pode ser visualizado na Figura 1. O coeficiente de correlação $R=0,923$ mostra que a estrutura dobrada compacta apresenta um comportamento médio elegante, o que está bem distante de $\delta = 2,47 \pm 0,03$ das proteínas globulares (MORET et al, 2009).

Figura 1: Relação massa-tamanho das 1213 proteínas, $D=2,15 \pm 0,025$, com coeficiente de correlação de Pearson $R=0,923$.



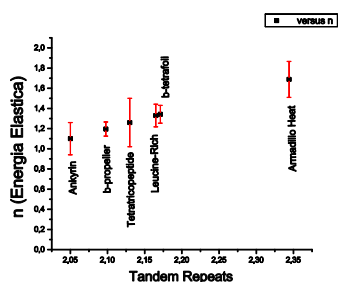
Fonte: Própria.

Da dimensão fractal observada, a não linearidade da força elástica encontrada $F_{el} = -1,075 R_g^{0,3}$ para $n = 1,3$, justifica a sua forma alongada no estado enovelado devido à pouca retração. Neste modelo, tanto a força hidrofóbica quanto a força de exclusão são fracas em longas distâncias quando comparadas

com as mesmas forças nas proteínas globulares conforme apontado em Moret et al, (2009).

A partir do modelo de campo médio de Flory, os n 's das famílias pesquisadas das proteínas com repetição mostram que as forças de restituição não obedecem à lei de Hooke, conforme apresentado na Figura 2.

Figura 2: A força elástica $F_{el} = \frac{1}{2} n R_g^{n-1}$, $n < 2$, gera pouca energia de interação de longo alcance.



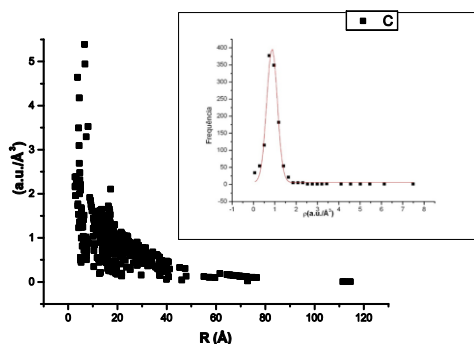
Fonte: Própria.

O método de cálculo de Massa-Raio consiste em sobrepor esferas sobre as proteínas em que as massas ficam aproximadamente contidas em esferas determinadas pelo raio de giro R_g , este mede a compactação de uma proteína com energia mínima e tem origem no centro de massa.

O expoente da relação Massa-Raio encontrado é muito próximo da estrutura de um arame amassado numa pequena cavidade bidimensional, estudado por Gomes (2008), mas podendo formar laços que é impossível no caso bidimensional.

A dimensão fractal encontrada sugere a compactação das famílias estudadas. Para verificar se a distribuição coletada têm densidades que seguem uma distribuição normal medimos a densidade $\rho_j = M_j/V_j$ da j -ésima proteína de repetição, onde V_j é o volume aproximado pela esfera pelo seu raio de giro (R_g)_j tende à média de compactação no valor de $\rho = 0,86 \text{ u.m.a./}\mu^3$ comportando-se como uma distribuição normal conforme a Figura 3, próxima da obtida a partir de estruturas cristalinas.

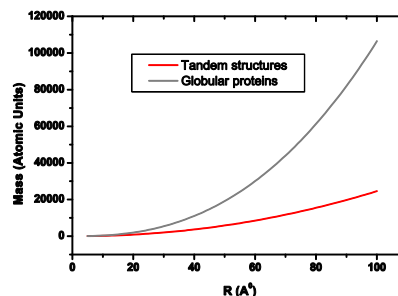
Figura 3: Comportamento da densidade média de compactação de 1.213 proteínas, $\rho = 0,86 \text{ u.m.a./}\mu^3$ com 0.7 de desvio padrão.



Fonte: Própria.

Em média, quanto maior a massa de uma proteína de repetição menor sua massa, isto é, as proteínas de repetição seria próxima a uma folha lisa de papel conforme Figura 4.

Figura 4: A densidade das proteínas de repetição diminuem com o raio de giro, ao contrário das proteínas globulares.



Fonte: Própria.

Conclusões

Ao estudar 1213 proteínas de repetição em tandem foi possível estabelecer que volumes e massas se comportam como objetos fractais, Figuras 1 e 2, e a dobra nativa como um fractal. A partir do modelo de campo médio de Flory justificamos a forma alongada das proteínas de repetição no estado enovelado pois, o expoente $n=2D-3 < 2$, $E_{el} \propto R_g^{2D-3}$, com $D=2.15$, mostra que a força de restituição não obedecem à lei de Hooke.

Por último, a forma das proteínas de repetição, conforme a Figura 4, assemelham-se a folhas de papel ligeiramente amassada, com alta densidade, Figura 3. A dimensão encontrada revela que nenhuma escala de hidrofobicidade pode ser relevante a estas proteínas, visto que, a geometria fractal se comporta entre a de um arame amassado e das esferas empacotadas aleatoriamente no limiar da percolação.

Referências

- ANDRADE, Miguel A.; PEREZ-IRATXETA, Carolina; PONTING, Chris P. Protein repeats: structures, functions, and evolution. *Journal of structural biology*, v. 134, n. 2, p. 117-131, 2001.
- GOMES, M. A. F.; BRITO, V. P.; COELHO, A. S. O.; DONATO, C. C. Plastic deformation of 2D crumpled wires. *Journal of Physics D: Applied Physics*, v. 41, n. 23, p. 235408, 2008.
- MARCOTTE, E. M; PELLEGRINI, M., Yeates, T. O.; EISENBERG, D. A census of protein repeats. *Journal of molecular biology*, v. 293, n. 1, p. 151-160, 1999.
- MOSAVI, L. K.; CAMMETT, T. J.; DESROSIERS, D. C.; PENG, Z. Y. The ankyrin repeat as molecular architecture for protein recognition. *Protein Science*, v. 13, n. 6, p. 1435-1448, 2004.
- MORET, M. A. ; MIRANDA, J. G. V.; NOGUEIRA JÚNIOR, E.; SANTANA, M. C.; ZEBENDE, G. F. Self-similarity and protein chains. *Physical Review E*, v. 71, n. 1, p. 012901, 2005.
- MORET, M. A; SANTANA, M. C., ZEBENDE, G. F.; PASCUTTI, P. G. Self-similarity and protein compactness. *Physical Review E*, v. 80, n. 4, p. 041908, 2009.