

# Algoritmo evolutivo para o problema de predição de proteínas, considerando a interação soluto-solvente no critério de avaliação.

Christiane Regina Soares Brasil<sup>1</sup> (PG)\*, Alexandre Cláudio Botazzo Delbem<sup>1</sup> (PQ), Telma Woerle da Silva<sup>2</sup> (PG), Daniel Rodrigo Ferraz Bonetti<sup>1</sup> (PG). Email: christiane@icmc.usp.br

<sup>1</sup> Universidade de São Paulo, Instituto de Ciências Matemáticas e Computacionais, São Carlos - SP. Avenida Trabalhador São-Carlense, 400.

Palavras Chave: proteína, estruturas terciárias, algoritmo evolutivo, solvente, área de acessibilidade ao solvente.

## Introdução

O problema de predição de estruturas terciárias (PSP, *Protein Structure Prediction*) é um grande desafio da Bioquímica, sendo fundamental para a compreensão da causa e solução de muitas doenças. Os métodos de cristalografia e RMN propõem-se a auxiliar na obtenção de estruturas terciárias de proteínas, no entanto apresentam suas limitações. A cristalografia fotografa as proteínas, mas em condições em que estas não são encontradas na natureza (ou seja, sem o contato com o solvente). Por outro lado, a RMN trabalha com proteínas imersas em solvente, mas a desvantagem é que apresentam bons resultados somente em proteínas pequenas (com dezenas ou centenas de aminoácidos). Diversas pesquisas estão sendo realizadas para solucionar essas desvantagens. No contexto computacional, os Algoritmos Evolutivos<sup>1</sup> (AEs) são procedimentos heurísticos bastante adequados para o PSP, e têm sido largamente utilizados para a pesquisa nessa área.

A proposta deste trabalho é desenvolver um algoritmo evolutivo ab-initio usando modelo *full-atom*<sup>2</sup>, voltado para o problema de predição de estruturas terciárias da proteína, considerando como critérios de avaliação a energia interna da proteína, assim como a interação proteína-solvente. O procedimento ab initio não precisa de nenhum conhecimento a priori, pretendendo obter-se uma estrutura terciária de uma proteína a partir somente de sua seqüência de aminoácidos.

## Resultados e Discussão

Existem diversos métodos para a avaliação da interação proteína-solvente, classificando-se em modelos discretos (explícitos) e contínuos (implícitos)<sup>3</sup>. Neste trabalho, a princípio, optou-se em utilizar os modelos contínuos para o cálculo da área de acessibilidade ao solvente<sup>4</sup>.

Neste projeto foram usados funções do pacote TINKER<sup>5</sup> (para energias intra-moleculares) e STRIDE<sup>6</sup> (para área de superfície de acessibilidade

ao solvente - SASA). No algoritmo ProtPred<sup>7</sup> foi inserida uma nova função de avaliação, cuja finalidade é calcular a área de acessibilidade da proteína ao solvente, e a partir disso, pode-se verificar a influência desse critério nos resultados obtidos (tamanho da população = 200 e número de iterações = 200).

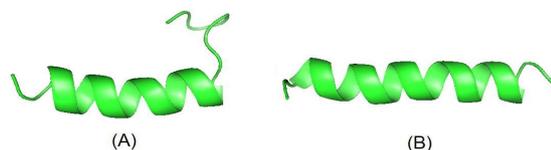


Figura 1. 1A11 sem SASA (A) e 1A11 com SASA (B).

## Conclusões

O novo critério (Figura 1(B)) avaliado juntamente com a energia interna da proteína (força de van der Waals e eletrostática) apresentou 1.79 Å em RMSD comparada a estrutura nativa, sendo que a Figura 1(A) tem 5.16Å em RMSD. Ainda que essa proteína seja pequena, resultados encorajam os trabalhos futuros abordarem a energia livre do sistema (energia de solvatação) como sendo um novo critério de avaliação em algoritmos evolutivos voltados ao PSP, avaliando não somente a área de acessibilidade, mas a energia.

## Agradecimentos

Ao suporte financeiro proveniente da Fundação de Apoio à Pesquisa de São Paulo (FAPESP).

<sup>1</sup> Deb, K. Multi-objective optimization using evolutionary algorithms. **2001**.

<sup>2</sup> McGarrah, D.; Judson, R. *Journal of Computational Chemistry*. **1993**, *14*, 1385-1395.

<sup>3</sup> Jaramillo, A.; Wodak, S.J. *Biophysical Journal*, **2005**, *88*, 156-171.

<sup>4</sup> Gaudio, A.; Takahata, Y. *Computers Chem.*, **1992**, *16*, 277-284.

<sup>5</sup> Ponder, J. Tinker: Software tools for molecular design. *Washington University, Saint Louis*, **2001**.

<sup>6</sup> Frishman, D.; Argos, P. *Proteins: structure, function and genetics*, **1995**, *23*, 566-579.

<sup>7</sup> Lima, T. W.; Gabriel, P. H. R.; Faccioli, R. A.; Delbem, A. C. B.; Silva, I. N. **2007**, *CEC2007*.