



Evento	Salão UFRGS 2015: SIC - XXVII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2015
Local	Porto Alegre - RS
Título	Desenvolvimento de Métodos e Estratégias Computacionais para a Identificação de Padrões Conformacionais em Proteínas
Autor	MARIANA DOS SANTOS OLIVEIRA
Orientador	MARCIO DORN

Desenvolvimento de Métodos e Estratégias Computacionais para a Identificação de Padrões Conformacionais em Proteínas

Mariana Oliveira Orientador: Márcio Dorn
Universidade Federal do Rio Grande do Sul

Proteínas são longas sequências de resíduos de aminoácidos que em condições fisiológicas adotam uma estrutura tridimensional (3D) única, a qual esta associada a sua função na célula. A predição dessa estrutura 3D pode ser obtida de maneira experimental, através das técnicas de cristalografia por difração de raio-x e de RMN. Entretanto, essas técnicas são de alto custo e podem levar muito tempo, assim, a utilização de métodos computacionais e algoritmos capazes de prever a estrutura nativa dessas proteínas é fundamental.

A predição 3D de uma proteína unicamente a partir de sua sequência de resíduos de aminoácidos ainda é um problema para a Ciência da Computação, Matemática, Química, Física e Biologia. O principal desafio é entender como a informação codificada na sequência linear de resíduos de aminoácidos de uma proteína é traduzida em estrutura tridimensional. Na literatura, há vários métodos já propostos para resolver esse problema, os quais podem ser divididos em métodos baseados de primeiros princípios (*ab initio* e predição *de novo*) e métodos baseados em conhecimentos (*threading* e modelagem comparativa).

Nas últimas edições do CASP , verificou-se que os melhores resultados foram encontrados utilizando os métodos baseados em conhecimento. Esses métodos são dependentes de informações estruturais determinadas experimentalmente, ou seja, padrões estruturais e regras relacionadas ao processo de enovelamento de proteínas. Independente das informações preliminares utilizadas, ainda há muito a ser obtido de bases experimentais, visto que o espaço de busca conformacional de uma proteína é vasto. Desta maneira, existe a necessidade da construção de estratégias computacionais para identificação e extração de padrões conformacionais de modelos experimentais, a fim de melhorar a predição de métodos baseados em conhecimento.

Nesse trabalho, estão sendo desenvolvidos métodos e estratégias conformacionais para identificação de padrões conformacionais em proteínas determinadas experimentalmente e armazenadas no *Protein Data Bank* (PDB). Na primeira etapa utilizou-se o STRIDE, para determinar os ângulos diédros Phi (ϕ) e Psi (ψ) e conteúdo de estrutura secundária de proteínas modelo extraídas do PDB. Os ângulos Phi e Psi representam as torções do esqueleto peptídico responsáveis pelo enovelamento da proteína. Através de resultados preliminares foi possível identificar padrões estruturais e preferenciais em estruturas modelo obtidas do PDB. Padrões estruturais representam resíduos de aminoácido com mesma estrutura secundária e padrões preferenciais ligações entre diferentes estruturas secundárias. Foram aplicados estratégias de clusterização nos dados experimentais a fim de encontrar preferencias conformacionais. Os dados obtidos do PDB foram clusterizados por proximidade dos ângulos (Phi e Psi) e, após, por proximidade estrutural, feita a partir dos valores de RMSD. Como trabalho futuro pretende-se estudar a aplicação de outras técnicas de clusterização para identificar padrões conformacionais e analisar a aplicabilidade destes padrões em métodos de predição baseados em conhecimento.