



SALÃO DE INICIAÇÃO CIENTÍFICA XXVIII SIC

paz no plural



Evento	Salão UFRGS 2016: SIC - XXVIII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2016
Local	Campus do Vale - UFRGS
Título	Identificação de padrões conformacionais em estruturas experimentais e projeto de meta-heurísticas para a predição da estrutura tridimensional de proteínas
Autor	MARIANA DOS SANTOS OLIVEIRA
Orientador	MARCIO DORN

Identificação de padrões conformacionais em estruturas experimentais e projeto de meta-heurísticas para a predição da estrutura tridimensional de proteínas.

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

Proteínas são longas sequências formadas por aminoácidos, os quais, em condições fisiológicas, se enovelam dando forma a uma estrutura tridimensional (3D) única. A função biológica de cada proteína está fortemente relacionada a sua estrutura 3D, logo o conhecimento da mesma possibilita o estudo mais detalhado dos processos biológicos. Predizer a estrutura terciária de uma proteína a partir da sua sequência de aminoácidos ainda é um problema em aberto para Bioinformática Estrutural. A determinação destas estruturas por métodos experimentais, como RMN e cristalografia, é de alto custo e consomem muito tempo. Desta forma, há uma discrepância entre o número de proteínas com estrutura 3D conhecida e o número de sequências conhecidas, sendo, assim, fundamental o desenvolvimento e utilização de métodos computacionais para a predição da estrutura 3D de proteínas (3D - PSP).

Diferentes estratégias computacionais e algoritmos foram propostos ao longo dos anos a fim de gerar soluções para os problemas 3D - PSP. Essas estratégias podem ser divididas em quatro classes: (i) métodos de primeiro princípios sem informação da base; (ii) métodos de primeiro princípios com informação da base experimental; (iii) métodos de reconhecimento de enovelamentos e alinhamento; (iv) métodos de modelagem comparativa e alinhamento de sequências.

A habilidade de predição de um método é avaliada nos experimentos do CASP (*Critical Assesment of Structure Prediction*), e a análise das últimas edições mostrou que os melhores resultados são obtidos por métodos da classe ii. Frequentemente esses métodos utilizam meta-heurísticas para conduzir o processo de busca de possíveis soluções conformacionais. Meta-heurística é uma classe de algoritmos desenvolvidos para solucionar problemas de otimização devido a habilidade de encontrar soluções satisfatórias/aproximadas com menor esforço para problemas em que a solução exata muitas vezes é inviável.

Computacionalmente, proteínas podem ser descritas por dois ângulos de torção para cada aminoácido: ϕ (Φ) e ψ (Ψ), os quais podem assumir valores de -180° a 180° . Estruturas secundárias flexíveis de proteínas, como alças e voltas, são mais difíceis de serem preditas por não apresentarem padrões comuns, logo pode ser atribuído variadas combinações de valores de ângulos de torção. Isto pode ser observado através do mapa de Ramachandran, em que alças e voltas populam múltiplas áreas. Encontrar valores exatos de ângulos de torção, então, torna-se computacionalmente custoso devido as inúmeras possíveis combinações, entretanto, pode-se diminuir o espaço de busca adotando padrões conformacionais observados em estruturas 3D experimentalmente determinadas.

A partir da informação estrutural experimental armazenada no *Protein Data Bank* (PDB) foram selecionados padrões de estruturas secundárias baseados na presença de estruturas flexíveis (alças e voltas) unindo estruturas rígidas, e, assim, foi criado um banco de dados de padrões de estruturas flexíveis. Utilizando este banco de dados foram gerados indivíduos, os quais formaram a população inicial da meta-heurística desenvolvida que será aplicada ao problema de predição de estruturas tridimensional de proteínas. Os padrões conformacionais contribuem para redução do espaço de busca conformacional da proteína.