



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	Informação estrutural das proteínas para a predição de eventos proteolíticos
Autor	MATHEUS DE BASTOS BALBÉ E GUTIERRES
Orientador	JOSE ARTUR BOGO CHIES

Informação estrutural das proteínas para a predição de eventos proteolíticos.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Orientador: José Artur Bogo Chies

Orientado: Matheus de Bastos Balbé Gutierrez

O proteossomo é um intrincado complexo proteico com atividade proteolítica capaz de reconhecer proteínas poliubiquitinadas, desdobrar sua estrutura tridimensional, e direcioná-las para degradação no seu interior catalítico. Esta via de ubiquitinação com posterior clivagem pelo proteossomo constitui o principal mecanismo de degradação de proteínas em células eucarióticas. Este compreende importantes funções tanto na degradação de polipeptídeos anormais, quanto no processamento e consequente apresentação de antígenos na resposta imunológica celular. Ao ingressar na câmara catalítica, o substrato não mais apresenta sua estrutura nativa, embora não seja possível afirmar que o mesmo se encontre totalmente linearizado. Existe uma série de exemplos de inibidores de proteossomos (na sua forma complexada) apresentando estruturação em nível secundário, o que evidencia a existência de espaço e condições para que a mesma ocorra. Se no interior da região catalítica o substrato apresenta alguma estrutura secundária específica atuando como uma propriedade biológica importante na apresentação de antígenos, isto seria de extrema valia no desenho de abordagens imunoterapêuticas. Para verificarmos isso, com o auxílio dos preditores de clivagem Netchop e ProteaSMM, realizamos uma série de testes estatísticos partindo da estrutura cristalográfica de proteínas humanas e virais com o intuito de avaliar se há alguma correlação entre os pontos preditos de clivagem com os aminoácidos contidos em estruturas secundárias específicas. É importante comentar que tanto o Netchop 3.1 quanto o ProteaSMM foram treinados sobre um conjunto de proteínas digeridas em ensaios *in vitro* e não utilizam qualquer informação estrutural para a codificação dos aminoácidos. Nossos resultados preliminares, contudo, apontam que aminoácidos em estruturas de fita beta estavam sobrerrepresentados nas posições com alta probabilidade de clivagem em detrimento de aminoácidos em regiões de voltas. Para investigar a possível influência da estrutura do substrato no evento de clivagem, foi realizada a construção de métodos de aprendizado de máquina utilizando, além de descritores físico-químicos dos aminoácidos, informações como área acessível ao solvente e probabilidade de formar cada um dos três grandes grupos de estrutura secundária (voltas, hélices e fitas beta). Para isso, optamos por construir e treinar redes neurais por meio de dois algoritmos: *backpropagation* e *NeuroEvolution of Augmenting Topologies* (NEAT), que diferem entre si pelo fato de que no primeiro as redes possuem uma topologia fixa, enquanto que NEAT parte de uma topologia inicial que cresce em complexidade determinada por uma função de *fitness*. Para as redes treinadas sob o paradigma *backpropagation*, utilizamos a biblioteca PyBrain, que possui módulos úteis para problemas envolvendo aprendizado de máquina, enquanto para o método NEAT uma versão na linguagem Python vem sendo desenvolvida. Este método consiste no treinamento das redes por evolução, onde redes semelhantes são divididas em espécies e competem entre si, reproduzindo-se os melhores de cada espécie para formar a próxima geração. A avaliação das redes se dá por meio de curvas ROC, que permitem avaliar graficamente o desempenho de um classificador binário. As primeiras redes já treinadas com o algoritmo *backpropagation* possuem uma camada de entrada variando de 15 a 36 neurônios, pois cada aminoácido é representado pelas probabilidades de formar fita beta, hélice ou alça e as janelas variavam de 5 a 12 aminoácidos, uma camada interna com 80 neurônios e um único neurônio de saída, com todas as conexões possíveis. Para o conjunto de validação utilizamos dados de ensaios de degradação de peptídeos provenientes da literatura, e obtivemos redes com desempenho equivalente ao de uma ferramenta de predição de clivagem amplamente avaliada e utilizada (NetChop 3.1). O fato surpreendente é que utilizamos um único descritor, indicando que a estrutura secundária parece ser realmente importante neste fenômeno biológico.