A nossa tarefa consiste em, dado uma sequência a pesquisar (identificador ou em formato fasta), procurar na base de dados PrositeScan ou CDD (Conserved Domains database) os motifs associados a essa sequência.

O output será uma estrutura de dados com informação de motifs Prosite ou CDD que estejam presentes na sequência.

Ou seja, informação como:

- lista de motifs e a sua localização;

- o seu score;

- mais informação será adicionada em breve…

Recursos:

REST API do scanProsite e BioPython. Possível adição de API para CDD

Para melhor perceber este trabalho, é importante perceber algum conceitos, como por exemplo, motifs, que são padrões existentes numa sequência biológica e que são partilhados por várias sequências dada a função biológica associada (ex: locais de ligação de proteínas regulatórios ao DNA; em sequências de proteínas a representação de domínios conservados de proteínas com funções biológicas determinadas).