Shape

Description automatically generated with medium confidence

**Prova 3 - Bioinformática**

Nome:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Modelagem de Estruturas Proteicas com RoseTTAFold2**

A enzima Acetolactato Sintase (ALS) desempenha um papel essencial na síntese de aminoácidos de cadeia ramificada em plantas. A sequência da ALS de *Arabidopsis thaliana* está disponível na base de dados UniProt sob o ID P17597.

>sp|P17597|ILVB\_ARATH Acetolactate synthase MAAATTTTTTSSSISFSTKPSPSSSKSPLPISRFSLPFSLNPNKSSSSSRRRGIKSSSPS SISAVLNTTTNVTTTPSPTKPTKPETFISRFAPDQPRKGADILVEALERQGVETVFAYPG GASMEIHQALTRSSSIRNVLPRHEQGGVFAAEGYARSSGKPGICIATSGPGATNLVSGLA DALLDSVPLVAITGQVPRRMIGTDAFQETPIVEVTRSITKHNYLVMDVEDIPRIIEEAFF LATSGRPGPVLVDVPKDIQQQLAIPNWEQAMRLPGYMSRMPKPPEDSHLEQIVRLISESK KPVLYVGGGCLNSSDELGRFVELTGIPVASTLMGLGSYPCDDELSLHMLGMHGTVYANYA VEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTPHVSVCGDVKLALQ GMNKVLENRAEELKLDFGVWRNELNVQKQKFPLSFKTFGEAIPPQYAIKVLDELTDGKAI ISTGVGQHQMWAAQFYNYKKPRQWLSSGGLGAMGFGLPAAIGASVANPDAIVVDIDGDGS FIMNVQELATIRVENLPVKVLLLNNQHLGMVMQWEDRFYKANRAHTFLGDPAQEDEIFPN MLLFAAACGIPAARVTKKADLREAIQTMLDTPGPYLLDVICPHQEHVLPMIPSGGTFNDV

1. Utilizando o programa **RoseTTAFold2**, disponível no notebook da iniciativa ColabFold (<https://github.com/sokrypton/ColabFold>), faça o download do script e realize a modelagem estrutural da proteína ALS, gerando **cinco modelos** distintos.
   * Certifique-se de usar os parâmetros adequados para obter os melhores resultados de modelagem.
2. Visualize os modelos gerados utilizando ferramentas como **PyMOL** ou a plataforma **TMalign**. Capture imagens (prints) dos cinco modelos para posterior análise e salve os arquivos gerados para validação futura.
3. Valide os cinco modelos utilizando a tabela fornecida em anexo. Identifique qual dos modelos apresenta a melhor qualidade e descreva o critério utilizado para sua seleção. Explique a relevância do modelo escolhido como representativo da proteína.

**Imunoinformática e Predição de Epítopos**

A bioinformática tem sido uma ferramenta fundamental para a descoberta de vacinas baseadas em epítopos, especialmente em contextos de pandemia, onde é necessária uma resposta ágil. Por meio de técnicas avançadas, pesquisadores podem analisar grandes volumes de dados genômicos e proteômicos para identificar rapidamente epítopos candidatos. Durante a pandemia de COVID-19, ferramentas de predição de epítopos contribuíram significativamente para o desenvolvimento de vacinas em tempo recorde.

1. Usando a ferramenta **NetMHCpan - 4.1**, disponível na plataforma DTU, analise a sequência da proteína (ID UniProt: **P26715**). Para isso:

* Selecione **peptídeos de 9 aminoácidos**.
* Crie uma tabela que categorize a quantidade de epítopos classificados como **high binders** e **weak binders** para os seguintes alelos: **HLA-A01:01**, **HLA-A02:02**, **HLA-A03:02**, e **HLA-A24:01**.

1. Usando a ferramenta **NetMHCIIpan - 4.3**, também disponível na plataforma DTU, realize uma análise semelhante:
   * Selecione **peptídeos de 15 aminoácidos**.
   * Crie uma tabela que categorize a quantidade de epítopos classificados como **high binders** e **weak binders** para os mesmos alelos mencionados acima.
2. Ainda na ferramenta **NetMHCIIpan - 4.3**, crie uma tabela com peptídeos e informações de predição dos cinco peptídeos com maior probabilidade de ligação com o MHC-II.