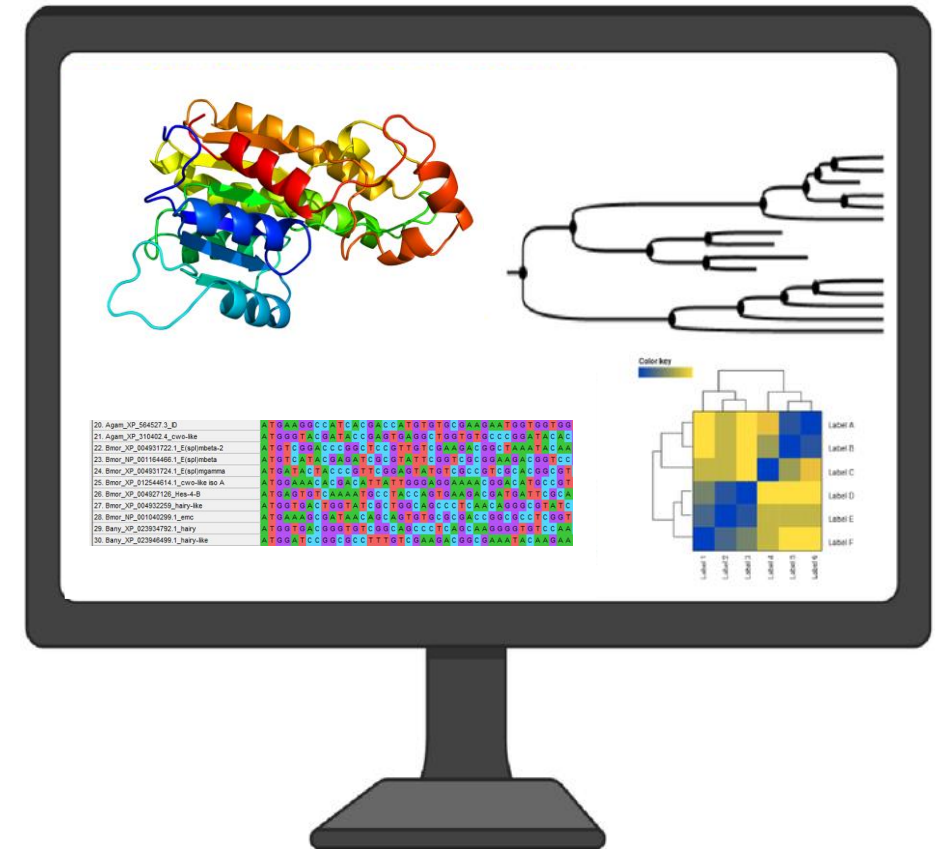


Treinamento Ferramentas Bioinformática



Msc. Ricardo Roldan
Dra. Roberta Lane

Alinhamento

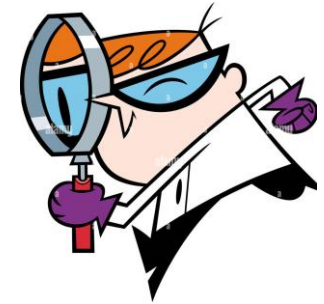
Por que fazer alinhamento?

MFGE8 Anolis_Ref	A	T	G	G	G	T	G	A	T	T	C	C	T	G	C	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 BWA	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 NGM	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 SOAP	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 Trinity	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 Velvet	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G

Alinhamento

Por que fazer alinhamento?

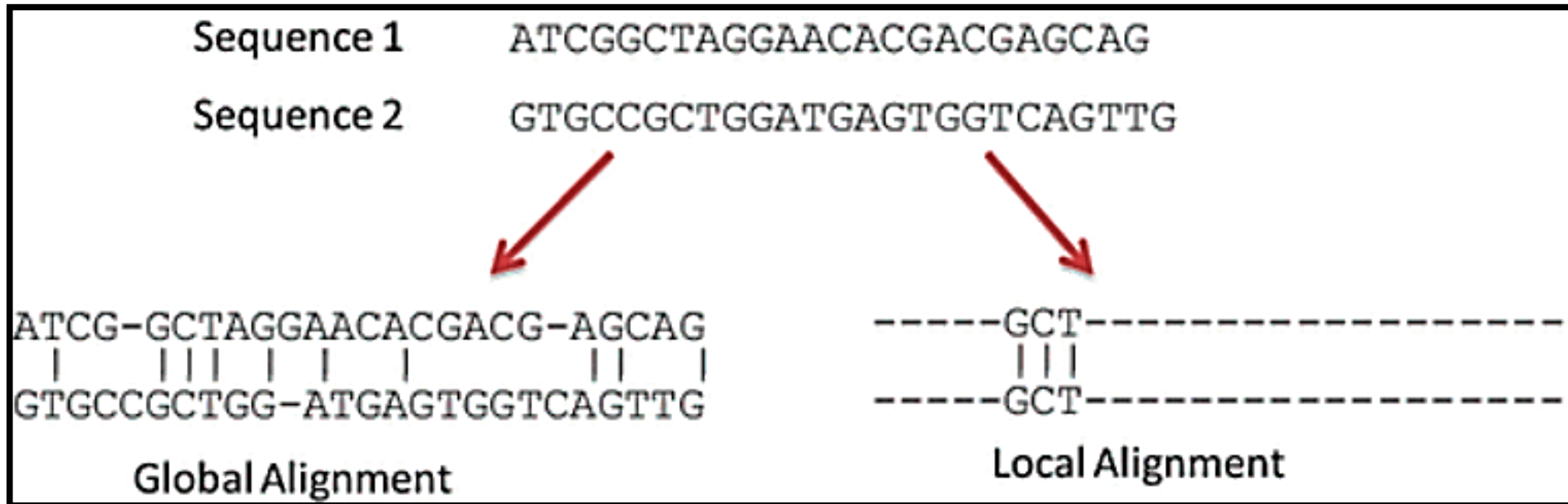
- > Comparar amostra vs DB;
- > Identificar o que é sua amostra;
- > Identificar motivos e domínios;
- > Permite criar filogenias e fenética



MFGE8 Anolis_Ref	A	T	G	G	G	T	G	A	T	T	C	C	T	G	C	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 BWA	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 NGM	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 SOAP	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 Trinity	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G
MFGE8 Velvet	N	N	N	G	G	T	G	A	T	T	C	C	T	G	T	G	A	A	G	T	G	A	A	T	C	A	A	T	G	T	C	T	G

Tipos de Alinhamento

Extensão



- Tentativa de alinhar toda a extensão das sequências (**Início ao fim**);
- Utilizada para comparação de genes ou proteínas com funções similares.

- Alinhamento com apenas a região de maior similaridade;
- Utilizada para identificação de padrões e domínios conservados.

Tipos de Alinhamento

Quantidade

Local Alignment

Pairwise Sequence Alignment

Target Sequence

5' ACTACTAGATTACTTACGGATCAGGTACTTTAGAGGCTTGCAACCA 3'

|||| ||||| ||||| |||||

Query Sequence

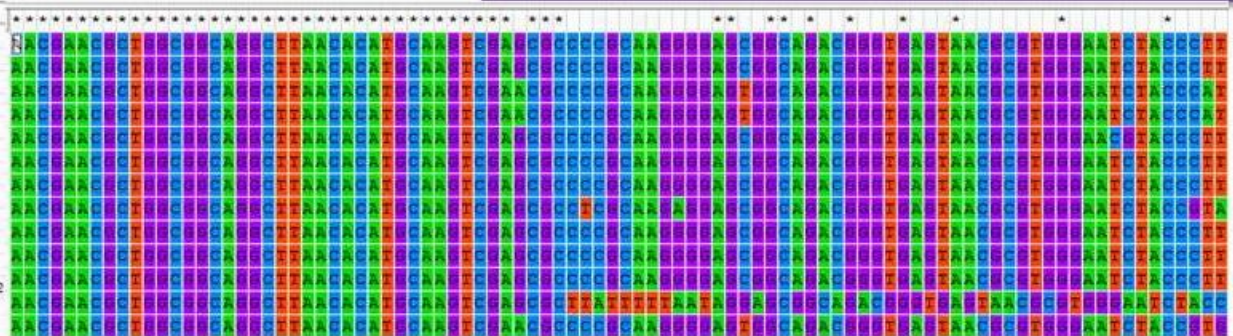
5' TACTCACGGATGAGGTACTTTAGAGGC 3'

Query Sequence

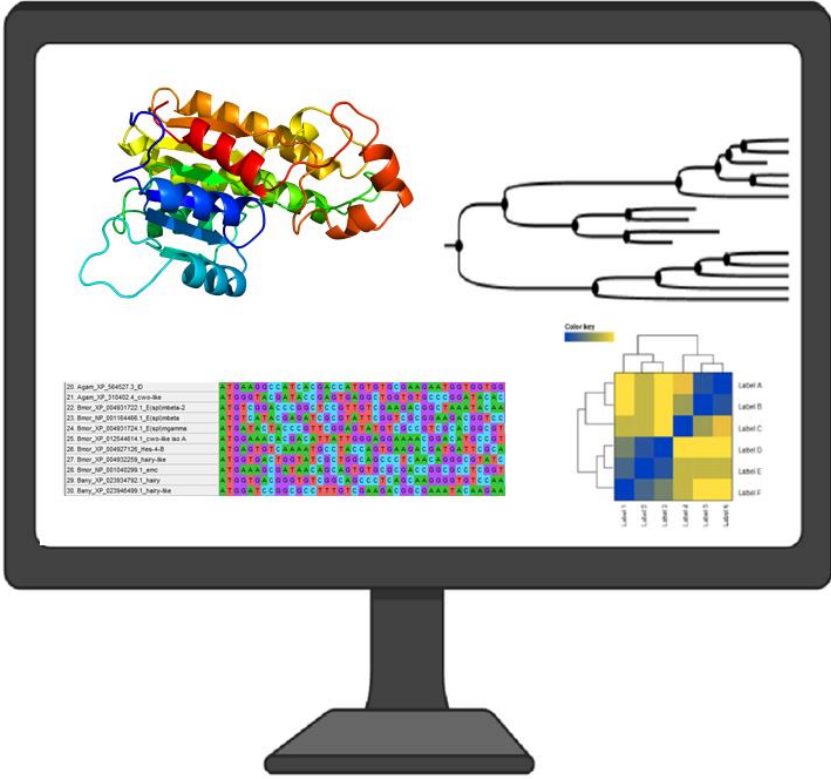
Multiple Sequence Alignment (MSA)

Species/Abbrv

1. Rhizobium leguminosarum_bv._viciae_3841_g115254414
2. Sinorhizobium medicae_WSM419_g150026743
3. Agrobacterium fabrum_str._C58_g159139455
4. Agrobacterium fabrum_str._C58_g159140696
5. Rhizobium etli_CIAI_652_g190694918
6. Rhizobium leguminosarum_bv._trifolii_WSM2304_g209533368
7. Agrobacterium radiobacter_K84_g221721649
8. Agrobacterium vitis_S4_g221737306
9. Sinorhizobium fredii_NGR234_g227339586
10. Rhizobium leguminosarum_bv._trifolii_WSM1325_g240856645
11. Sinorhizobium meliloti_1021_g30407155
12. Candidatus Liberibacter solanacearum_CLso-ZC1_g313495152
13. Agrobacterium_sp._H13-3_g325062059



Atividade

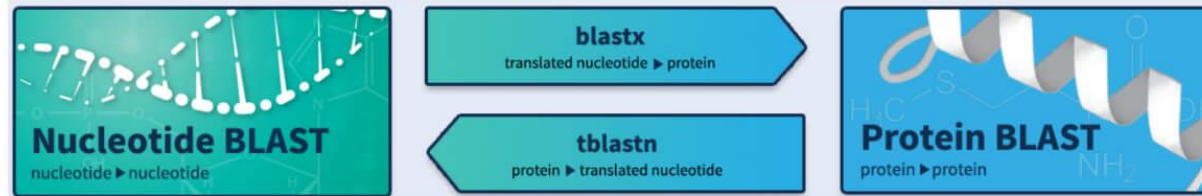


Atividade



BLAST

Basic Local Alignment Search Tool



<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

>Desconhecido

KDCKRESNTFPGICITKPPCRKACIREKFTDGHCSKILRRCLCTKPC

Atividade



<https://www.megasoftware.net/>

Motivos e Domínios

Motivos e Domínios

Motivos:

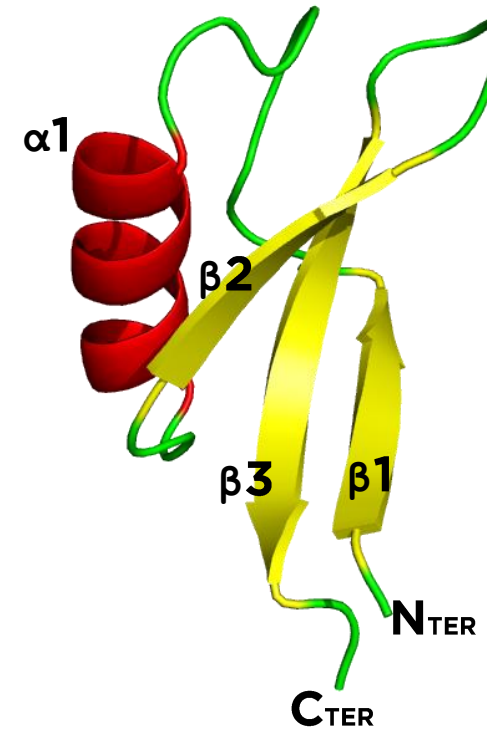
- “Motivo” ~ “Padrão”

*Ex: Defensinas vegetais : $\beta\alpha\beta\beta$

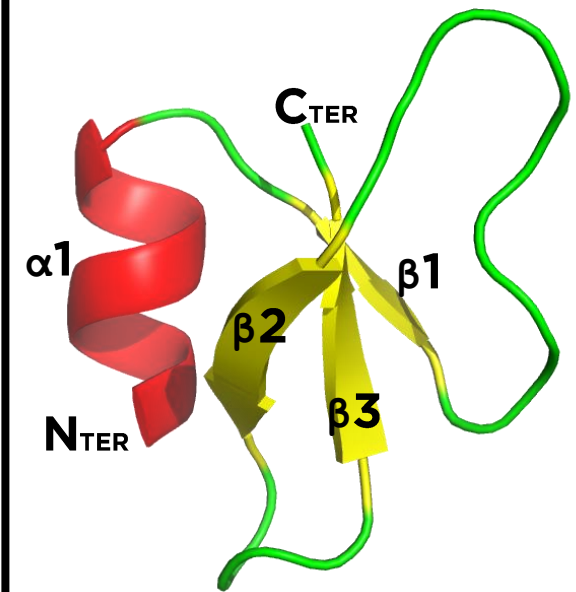
Defensinas humanas : $\alpha\beta\beta\beta$

Domínios:

- Regiões funcionais e/ou estruturais das proteínas;
- Proteínas podem apresentar 1 ou mais domínios.



Defensina vegetal



Defensina humana

Motivos e Domínios

Motivos:

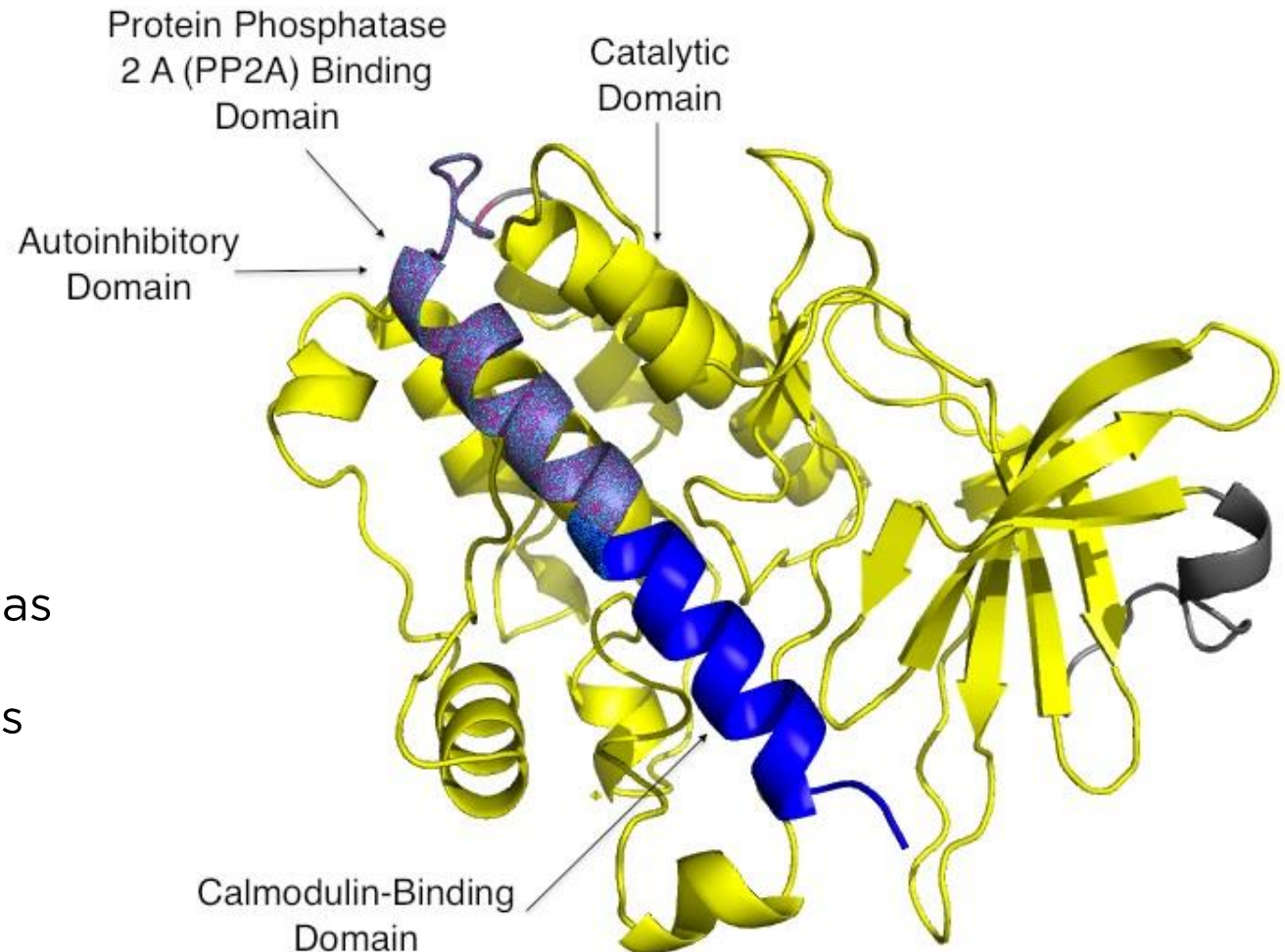
- “Motivo” ~ “Padrão”

*Ex: Defensinas vegetais : $\beta\alpha\beta\beta$

Defensinas humanas : $\alpha\beta\beta\beta$

Domínios:

- Regiões funcionais e/ou estruturais das proteínas;
- Proteínas podem apresentar 1 ou mais domínios.



Atividade



Interpro

<https://www.ebi.ac.uk/interpro/>



CD-Search (NCBI)

<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>

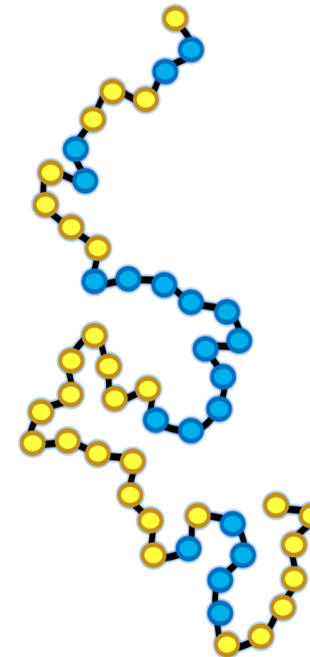
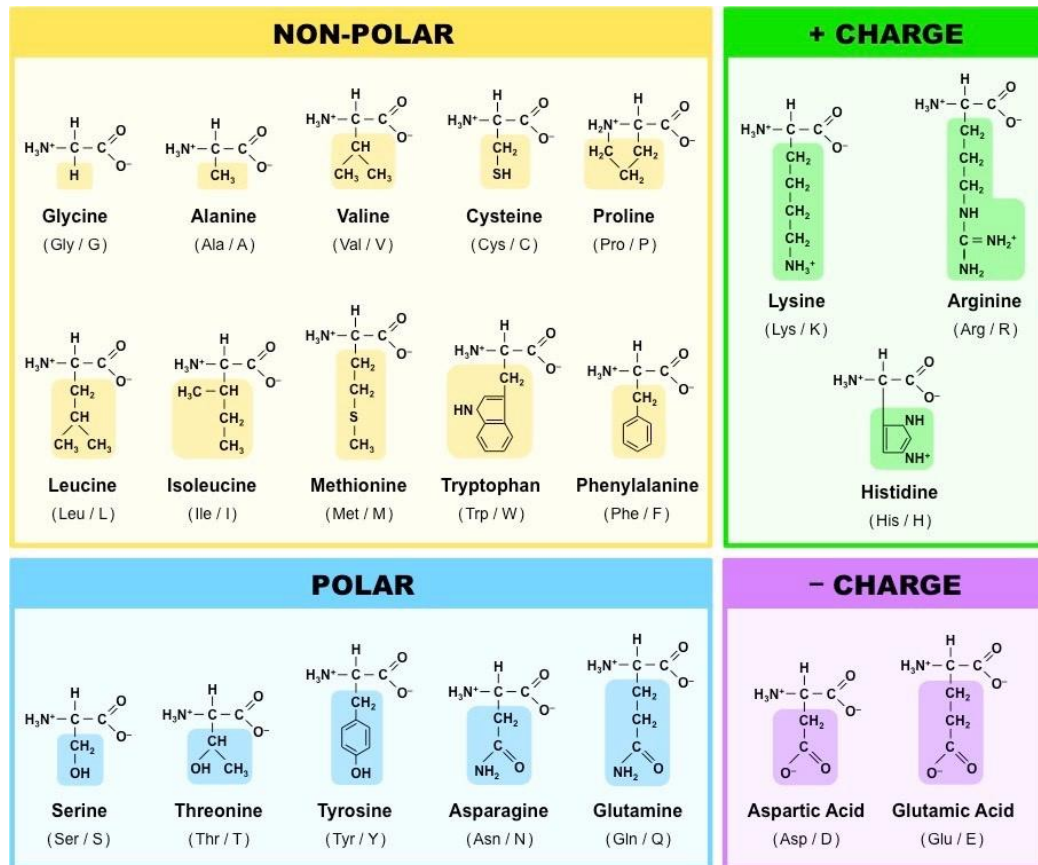
Modelagem proteínas

Modelagem proteínas

Por que estudar estrutura das proteínas?



Proteínas e seus aminoácidos

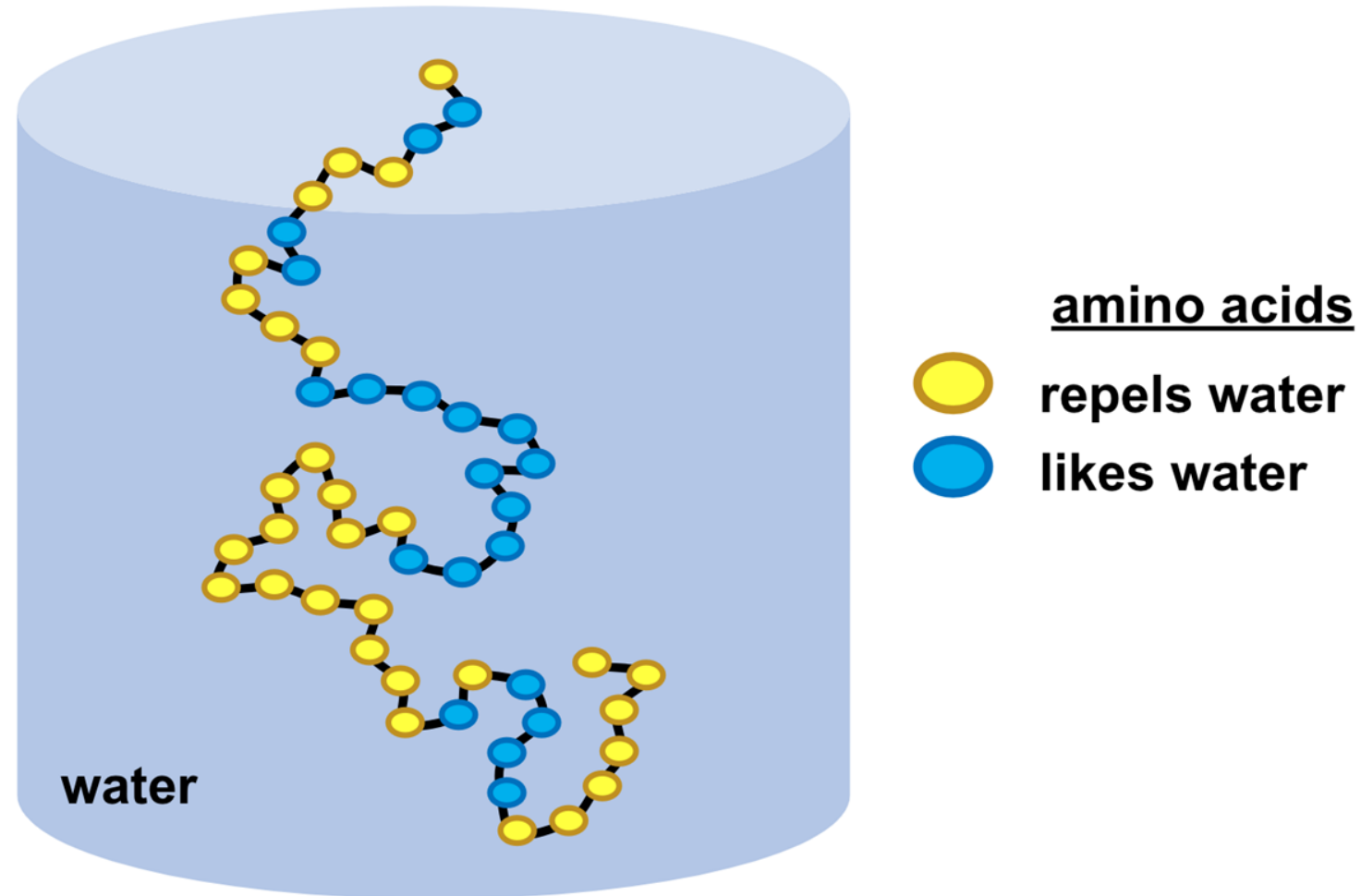


amino acids

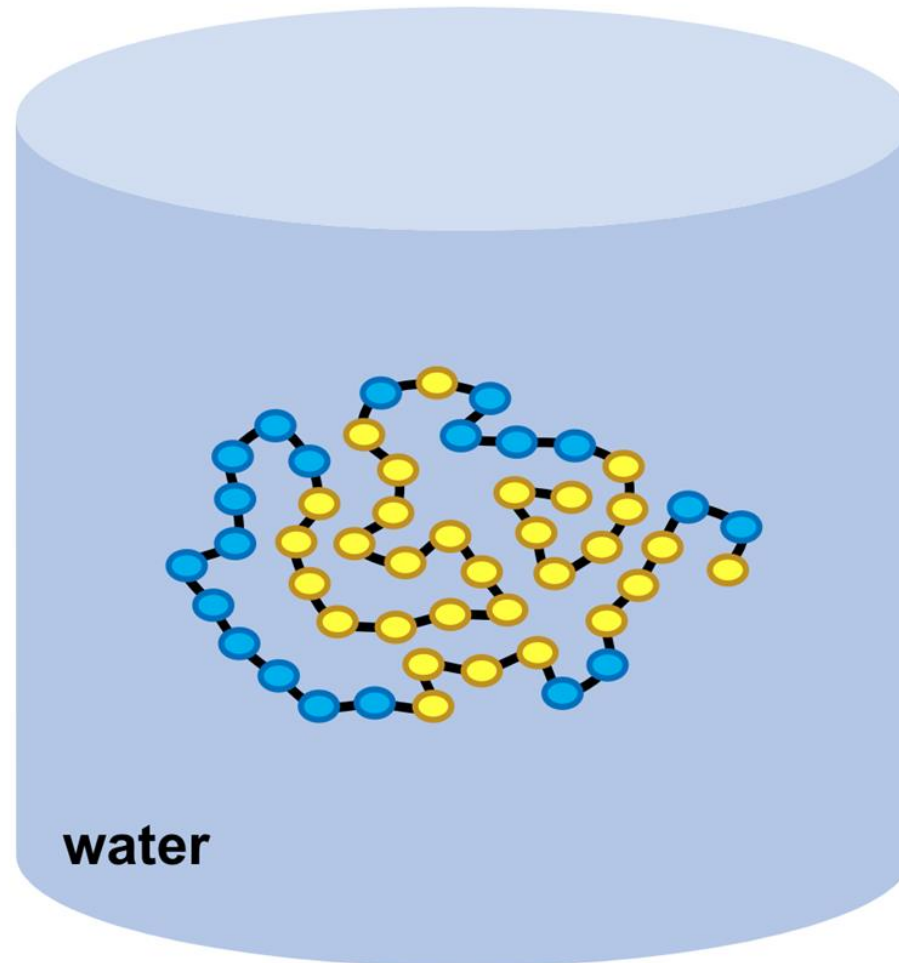
● repels water

● likes water

Proteínas e seus aminoácidos



Proteínas e seus aminoácidos



amino acids

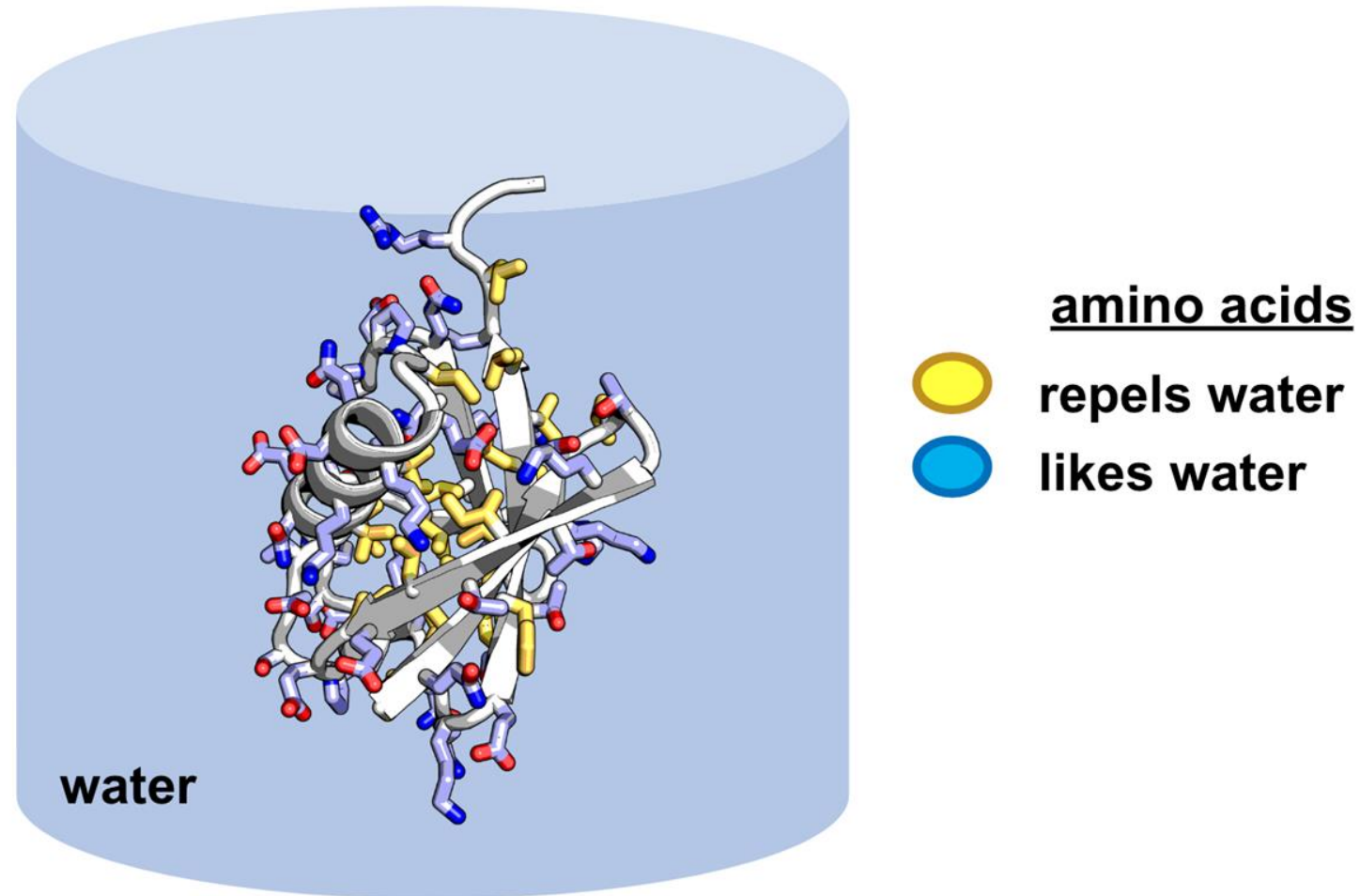


repels water



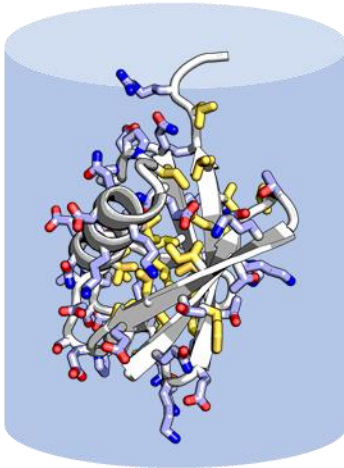
likes water

Proteínas e seus aminoácidos

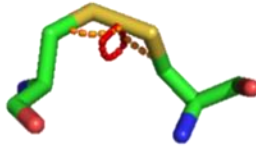


Fatores responsáveis pelo dobramento

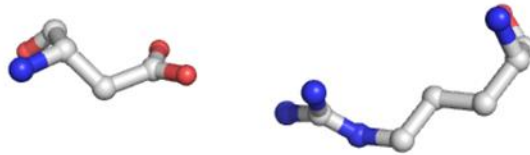
Solvation



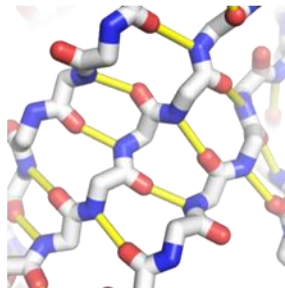
Disulfide bond



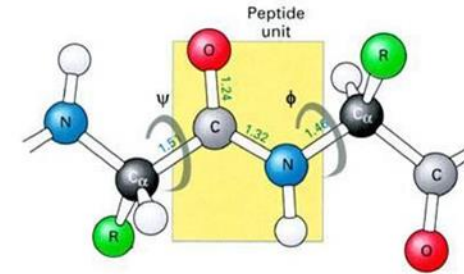
Ionic bond



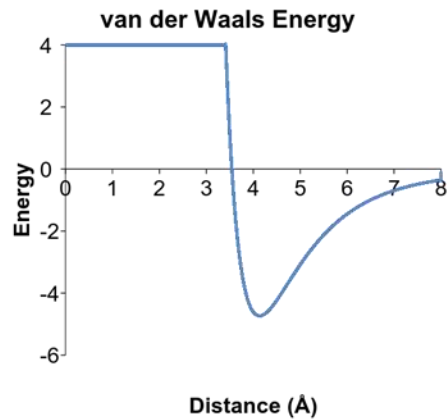
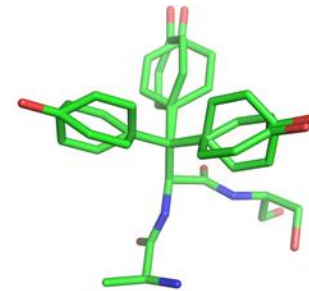
Hydrogen bond



Backbone conformations



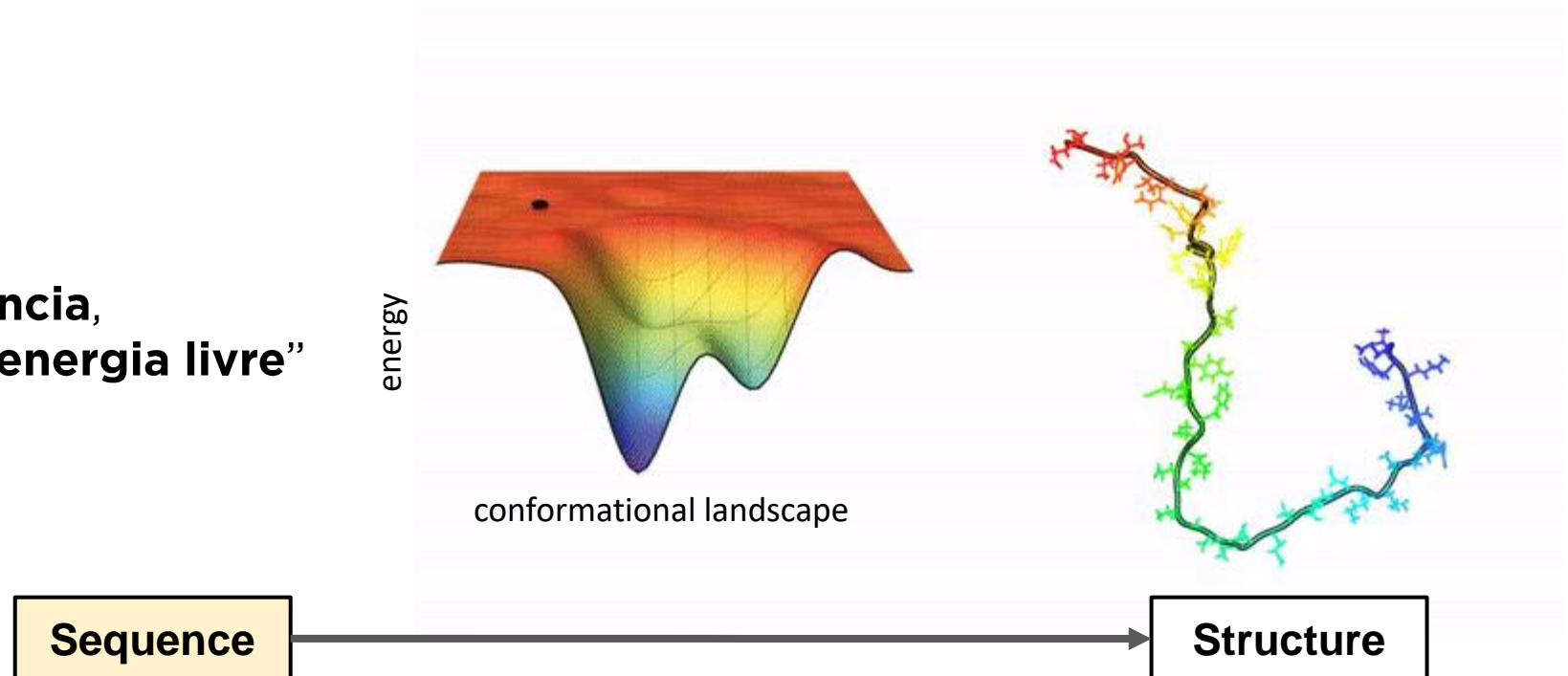
Sidechain conformations



Dr. Sergey Ovchinnikov

Dobramento e a problemática da modelagem

“Para uma dada **sequência**,
encontre a **estrutura** de **menor energia livre**”



[Video credit: C. Fennell]

Dill, K.A. and MacCallum, J.L., 2012. The protein-folding problem, 50 years on. *science*, 338(6110), pp.1042-1046.

Slide credit : Sergey Ovchinnikov

Modelagem comparativa ("Modelagem por homologia")

Modeller

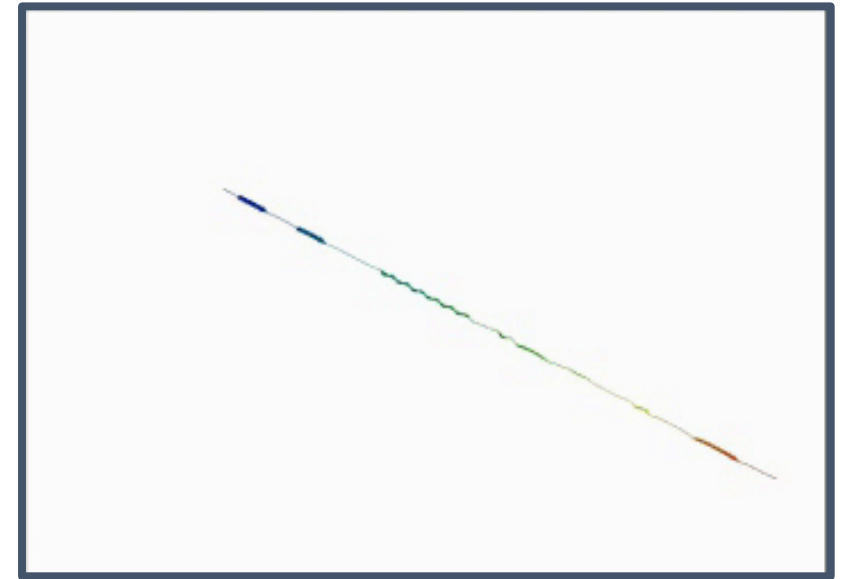
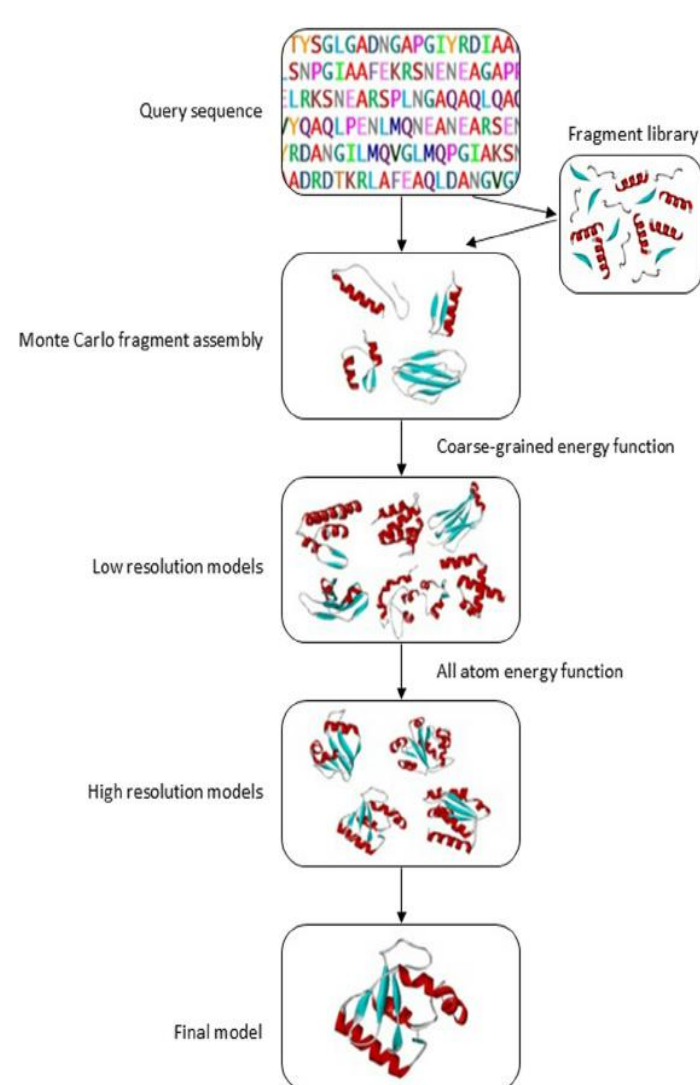
Program for Comparative Protein
Structure Modelling by Satisfaction
of Spatial Restraints

A	I	L	V	G	S	M	P	R	R	D	G	M	E	R	K	D	L	L	K	A	N	V	K	I	F	K	C	Q	G	A
V	E	V	C	P	V	D	C	F	Y	E	G	P	N	F	L	V	I	H	P	D	E	C	I	D	C	A	L	C	E	P
G	A	C	K	P	E	C	P	V	N	I	I	Q	G	S	-	-	Y	A	I	D	A	D	S	C	I	D	C	G	S	
C	-	-	I	A	C	G	A	C	K	P	E	C	P	V	N	I	I	Q	G	S	-	-	I	Y	A	I	D	A	D	S

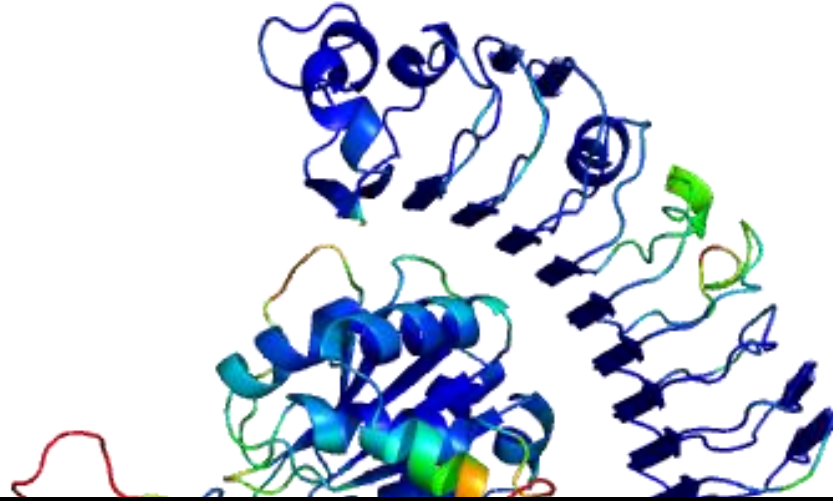


Utiliza uma **estrutura resolvida** experimentalmente como **template** para gerar **modelos**

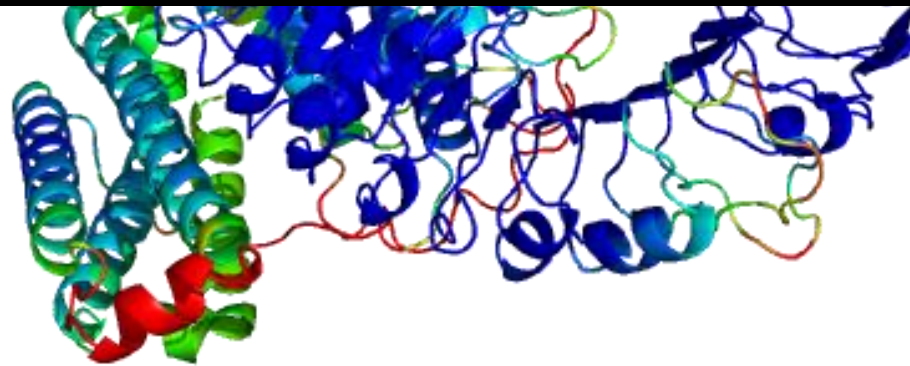
Modelagem *ab initio* ("do zero")



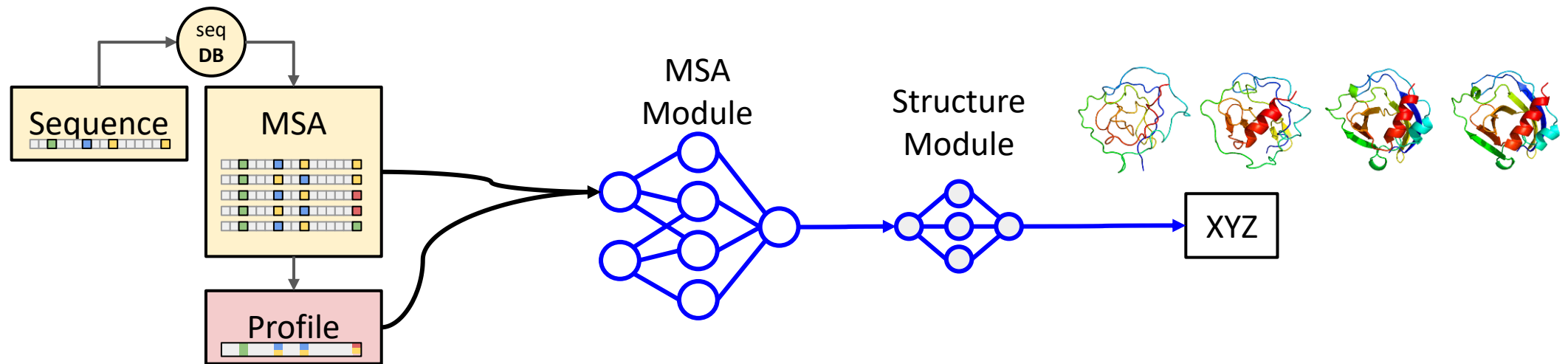
Inteligência artificial



ALPHAFOLD



Inteligência artificial e modelagem



Jumper J. et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*

Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, ..., Baker D. 2021. Acc. pred. of protein struct. and inter. using a 3-track NN. *Science*

Slide credit : Sergey Ovchinnikov

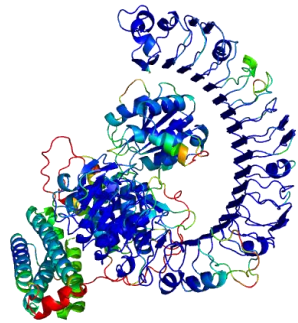
Atividade



PyMOL → Ferramenta de visualização (“**.pdb**”)

<https://pymol.org/2/>

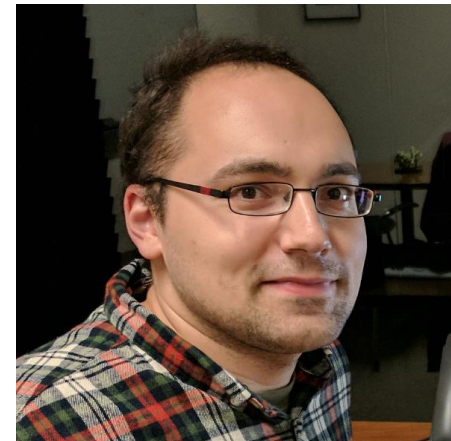
Atividade



ALPHAFOLD

colab

<https://github.com/sokrypton/ColabFold>



Sergey Ovchinnikov
Harvard University

**Obrigado
e
Boa Jornada!**

Contato:



r.s.rolدانfilho@gmail.com



<https://www.linkedin.com/in/ricardoroldan17/>