



UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE BIOQUÍMICA E IMUNOLOGIA



PREDIÇÃO DE ESTRUTURAS PROTEICAS E ANÁLISES FUNCIONAIS

Ma. Iasmin Cartaxo Taveira

Ribeirão Preto

09/05/2024

Sumário

- Revisão sobre proteínas
 - Aminoácidos, Estrutura primária a quaternária
- Que informações podemos extrair da estrutura primária de proteínas?
- Metodologias de predição de estruturas proteicas
 - Predição por homologia e *ab initio*
 - Refinamento de estruturas
- Docking e redirecionamento de fármacos



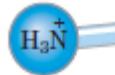
Revisão de proteínas

- Aminoácidos

TABELA 3-1 Propriedades e convenções associadas a aminoácidos comuns encontrados em proteínas

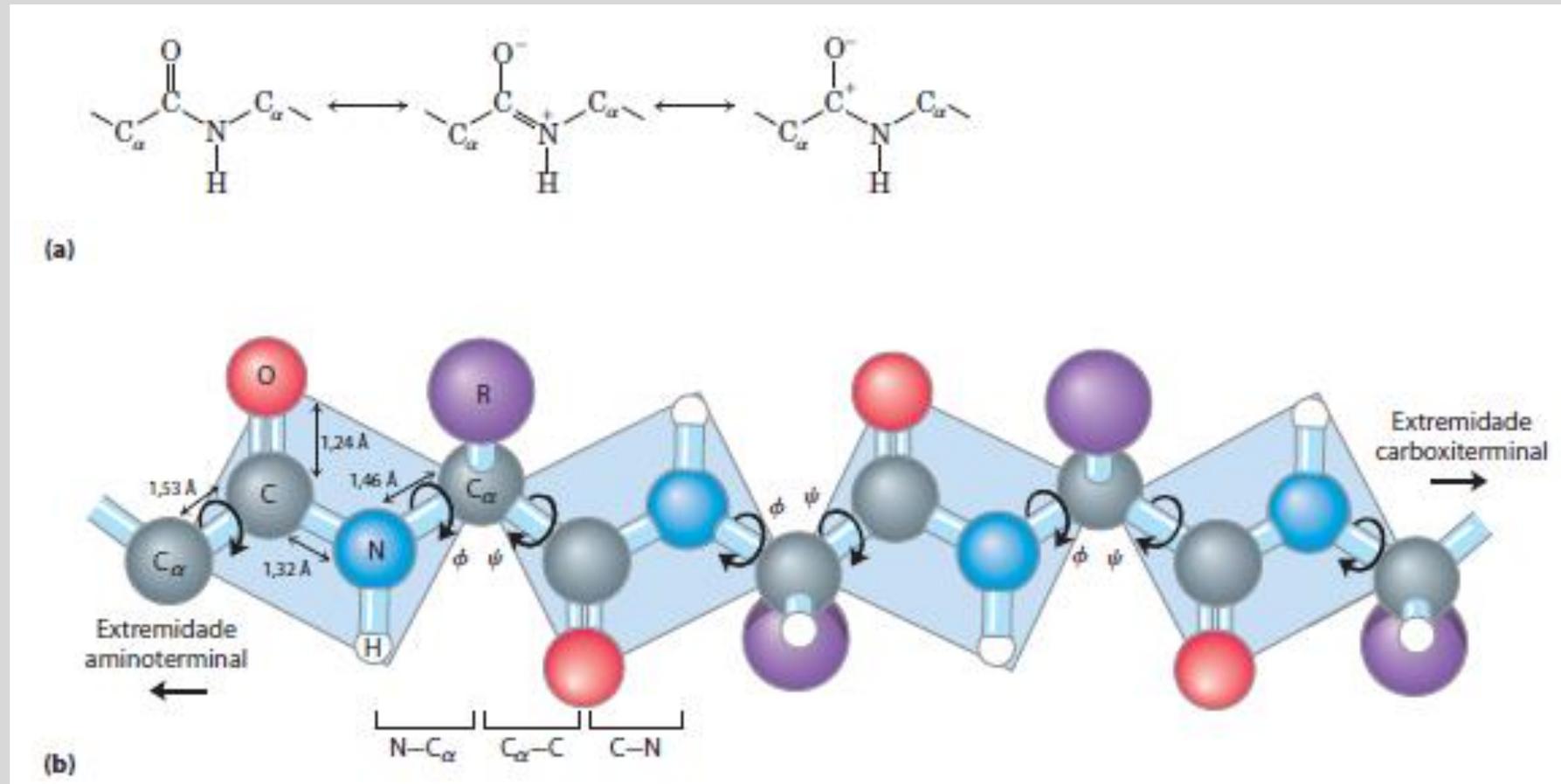
Aminoácido	Abreviação/ símbolo	M_r^*	Valores de pK_a			pI	Índice de hidropatia [†]	Ocorrência em proteínas (%) [†]
			pK_1 (—COOH)	pK_2 (—NH ₃ ⁺)	pK_R (grupo R)			
Grupos R alifáticos, apolares								
Glicina	Gly G	75	2,34	9,60		5,97	-0,4	7,2
Alanina	Ala A	89	2,34	9,69		6,01	1,8	7,8
Prolina	Pro P	115	1,99	10,96		6,48	-1,6	5,2
Valina	Val V	117	2,32	9,62		5,97	4,2	6,6
Leucina	Leu L	131	2,36	9,60		5,98	3,8	9,1
Isoleucina	Ile I	131	2,36	9,68		6,02	4,5	5,3
Metionina	Met M	149	2,28	9,21		5,74	1,9	2,3
Grupos R aromáticos								
Fenilalanina	Phe F	165	1,83	9,13		5,48	2,8	3,9
Tirosina	Tyr Y	181	2,20	9,11	10,07	5,66	-1,3	3,2
Triptofano	Trp W	204	2,38	9,39		5,89	-0,9	1,4
Grupos R polares, não carregados								
Serina	Ser S	105	2,21	9,15		5,68	-0,8	6,8
Treonina	Thr T	119	2,11	9,62		5,87	-0,7	5,9
Cisteína [‡]	Cys C	121	1,96	10,28	8,18	5,07	2,5	1,9
Asparagina	Asn N	132	2,02	8,80		5,41	-3,5	4,3
Glutamina	Gln Q	146	2,17	9,13		5,65	-3,5	4,2
Grupos R carregados positivamente								
Lisina	Lys K	146	2,18	8,95	10,53	9,74	-3,9	5,9
Histidina	His H	155	1,82	9,17	6,00	7,59	-3,2	2,3
Arginina	Arg R	174	2,17	9,04	12,48	10,76	-4,5	5,1
Grupos R carregados negativamente								
Aspartato	Asp D	133	1,88	9,60	3,65	2,77	-3,5	5,3
Glutamato	Glu E	147	2,19	9,67	4,25	3,22	-3,5	6,3

*Os valores de M_r refletem as estruturas como mostradas na Figura 3-5. Os elementos da água (M_r , 18) são removidos quando o aminoácido é incorporado a um polipeptídeo.



Revisão de proteínas

- A ligação peptídica



Revisão de proteínas

- Estrutura primária

Q86WU2 · LDHD_HUMAN

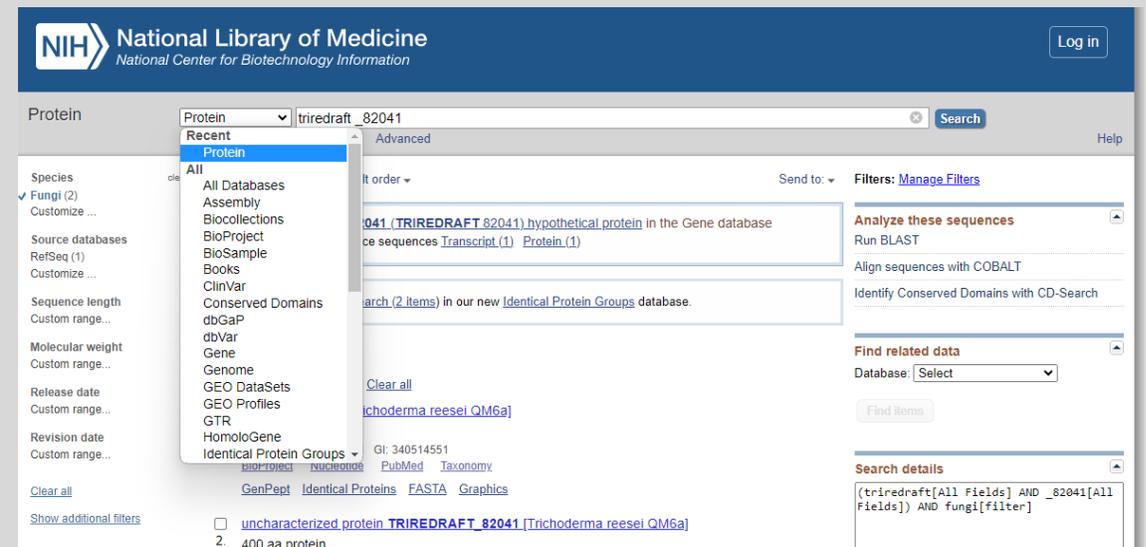
```
>sp|Q86WU2-1|LDHD_HUMAN Isoform 1 of Probable D-lactate dehydrogenase,  
mitochondrial OS=Homo sapiens  
MARLLRSATWELFPWRGYCSQKAKGELCRDFVEALKAVVGGSHVSTAAVVREQHGRDESV  
HRCEPPDAVVWPQNVEQVSRLAALCYRQGVPIIPFGTGTGLEGGVCAVQGGVVCVNLTHMD  
RILELNQEDFSVVVEPGVTRKALNAHLRDSGLWFPVDPGADASLCGMAATGASGTNAVRY  
GTMRDNVLNLEVVLDPGRLHTAGRGRHFRFGFWPEIPHHTAWYSPCVSLGRRKSAAGYN  
LTGLFVGSEGTLGLITATTLRLHPAPEATVAATCAFPSVQAAVDSTVHILQAAVVARIE  
FLDEVMMDACNRYSKLNCLVAPTLFLEFHGSQQALEEQLQRTEEIVQQNGASDFSWAKEA  
EERSRLWTARHNAWYAALATRPGCKGYSTDVCVPISRLPEIVVQTKEDLNASGLTGSIVG  
HVGDGNFHCILLVNPDDAEELGRVKAFAEQLGRRALALHGTCTGEHGIGMGKRQLLQEEV  
GAVGVETMRQLKAVLDPQGLMNPVKVL
```

Sequência em formato FASTA da lactato desidrogenase

Revisão de proteínas

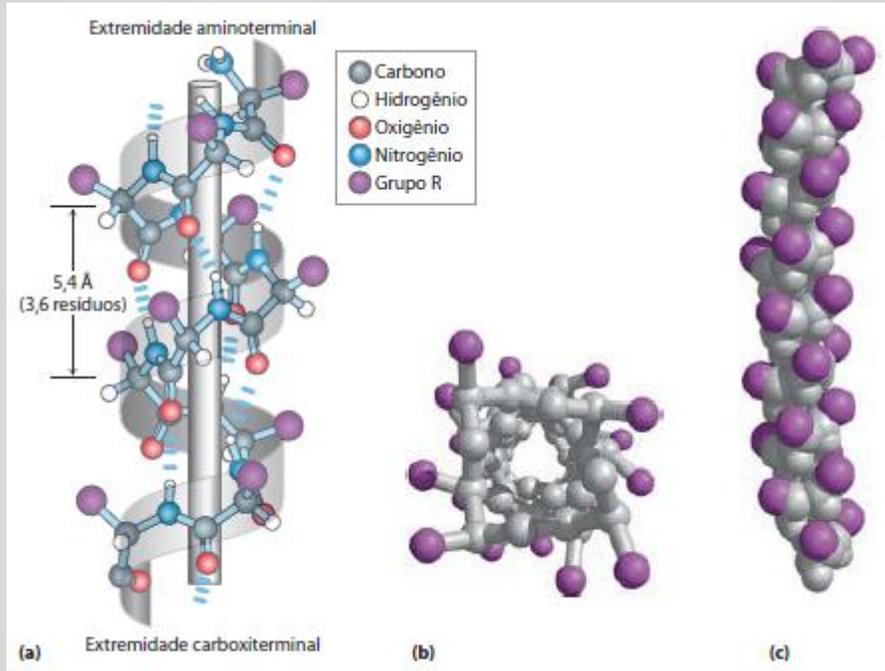
- Onde podemos obter as sequências de proteínas?

Bancos de dados!

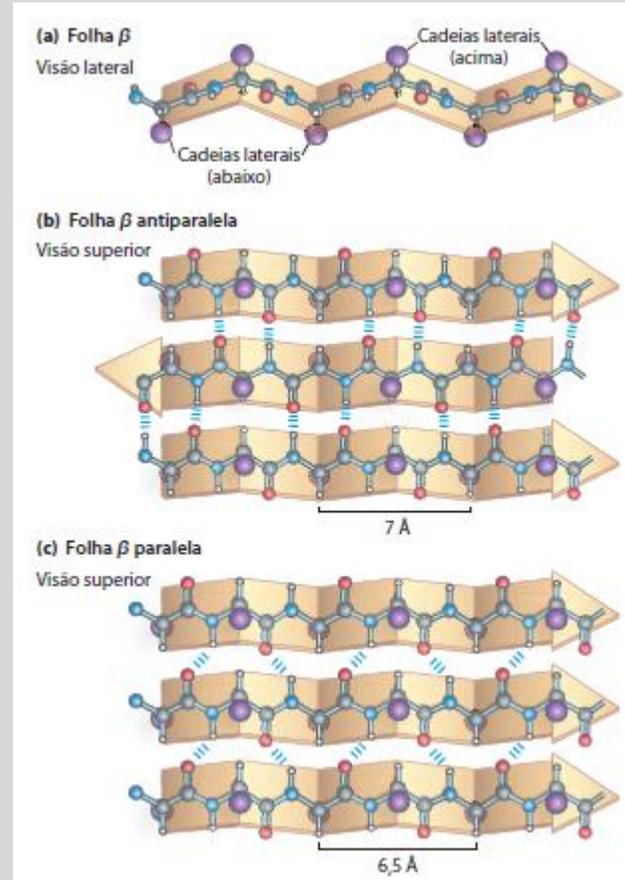


Revisão de proteínas

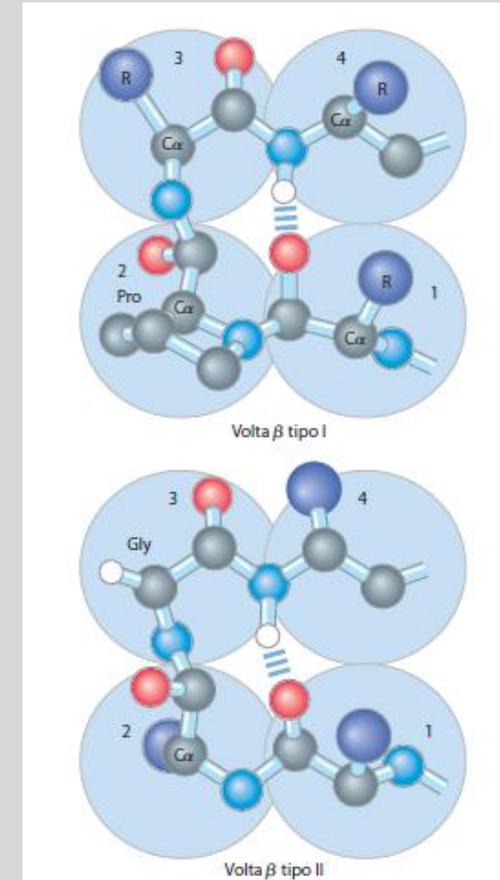
- Estrutura secundária



α -Hélice



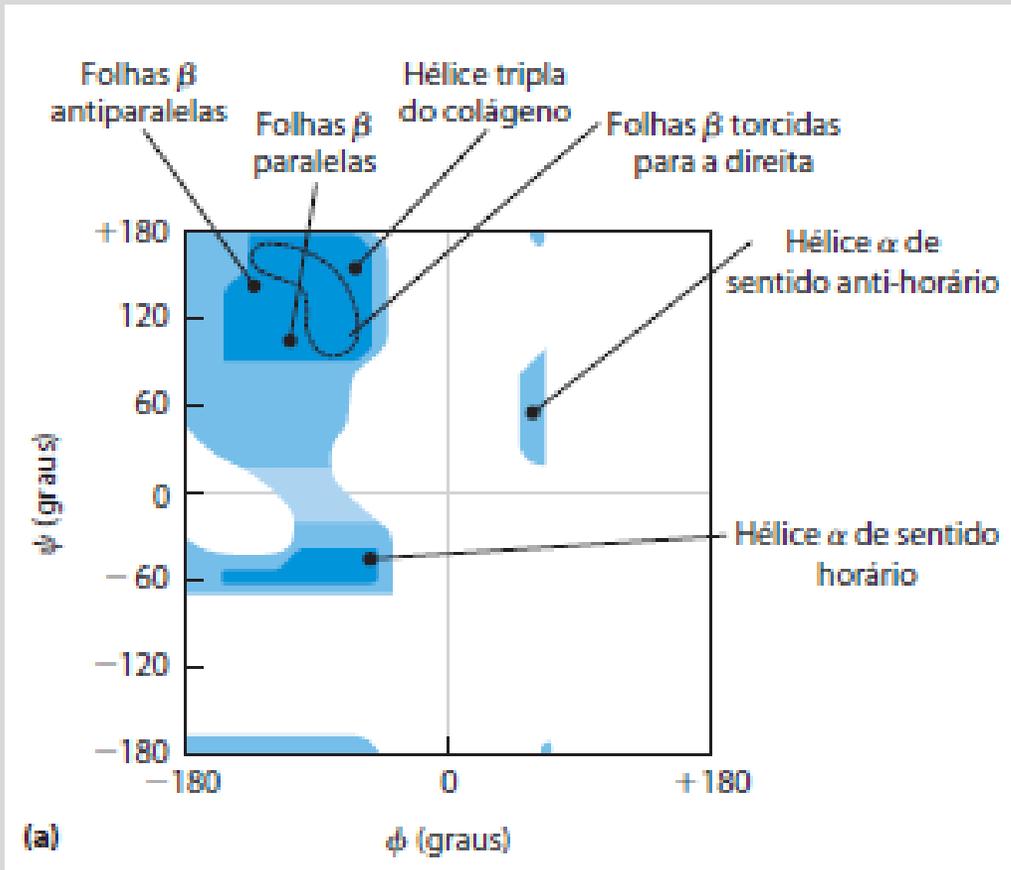
β -Folha



Volta β

Revisão de proteínas

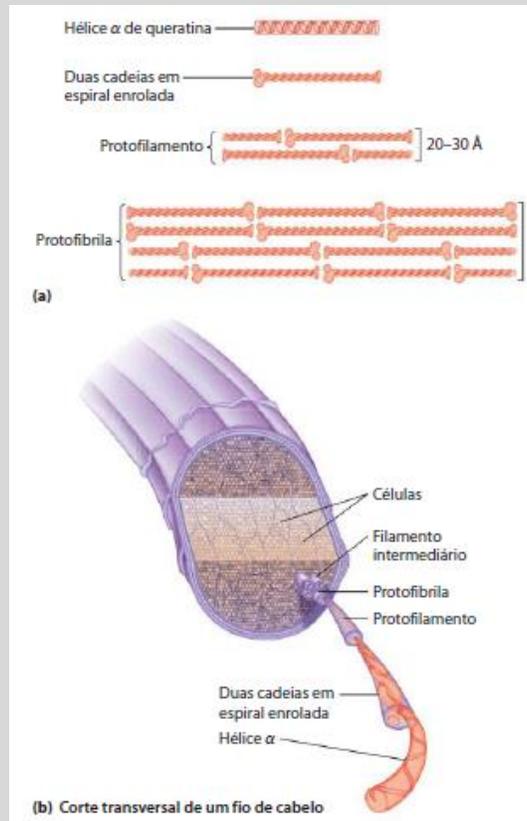
- Estrutura secundária
 - O Gráfico de Ramachandran



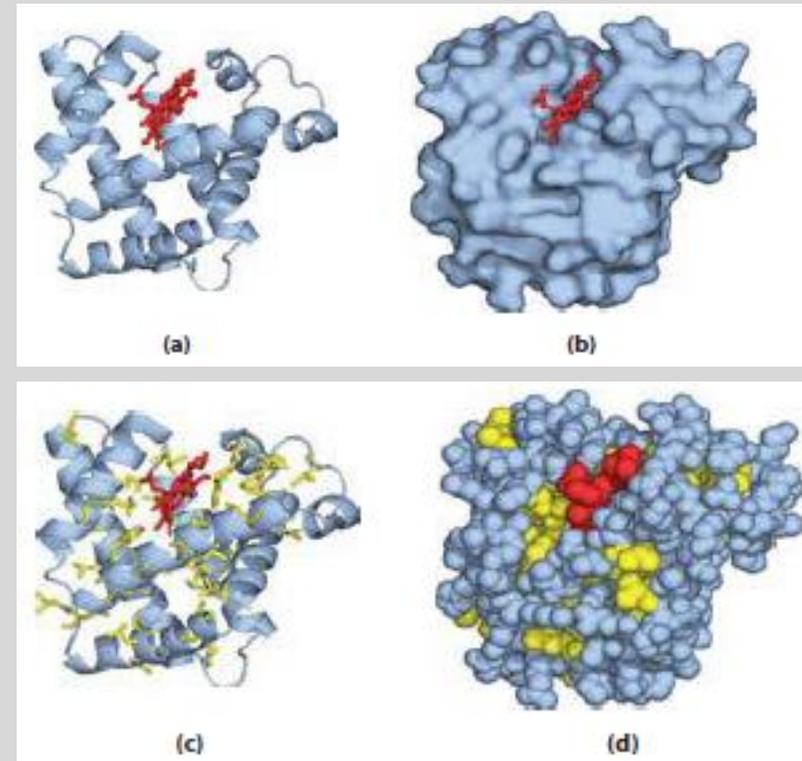
A conformação dos peptídeos é definida pelos valores de phi e psi. Conformações consideradas possíveis são aquelas que envolvem pouco ou nenhum impedimento estérico, com base nos cálculos dos raios de van der Waals conhecidos e dos ângulos diedros. As áreas coloridas em azul-escuro representam as conformações que não envolvem sobreposição estérica se os raios de van der Waals de cada átomo estão modelados como esferas rígidas, e, portanto, são totalmente permitidas; o azul médio indica as conformações permitidas, se for possível a aproximação dos átomos do aminoácido; o azul-claro indica as conformações que são admitidas se for possível uma pequena flexibilidade (poucos graus) no ângulo diedro que descreve a ligação peptídica. As regiões em branco são conformações não permitidas. A assimetria do diagrama é resultante da estereoquímica L dos resíduos de aminoácidos.

Revisão de proteínas

- Estrutura terciária



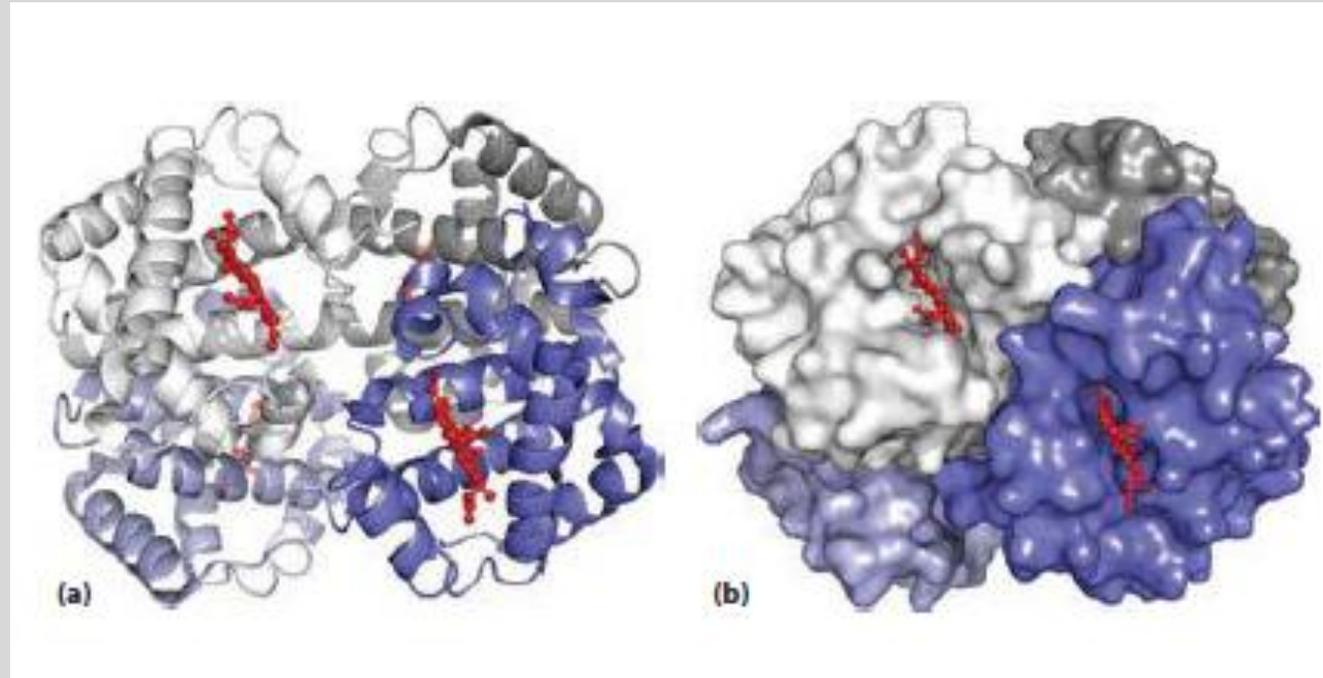
Proteínas fibrosas



Proteínas globulares

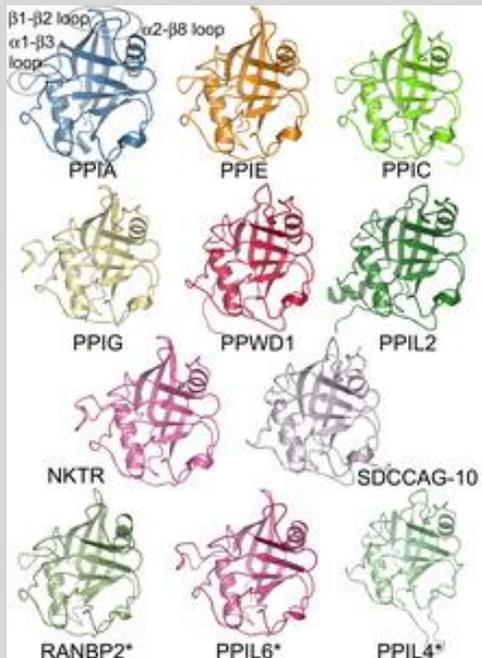
Revisão de proteínas

- Estrutura quaternária
 - Formada pela interação de estruturas terciárias em um complexo proteico

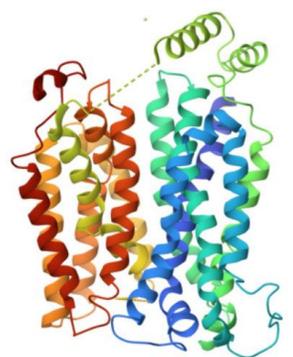


Que informações podemos extrair da estrutura primária de proteínas?

1. Identificação de homólogos
2. Identificação de famílias de proteínas
3. Sequências sinais
4. Sequências consenso



Transp. açúcar



Transp. nitrato

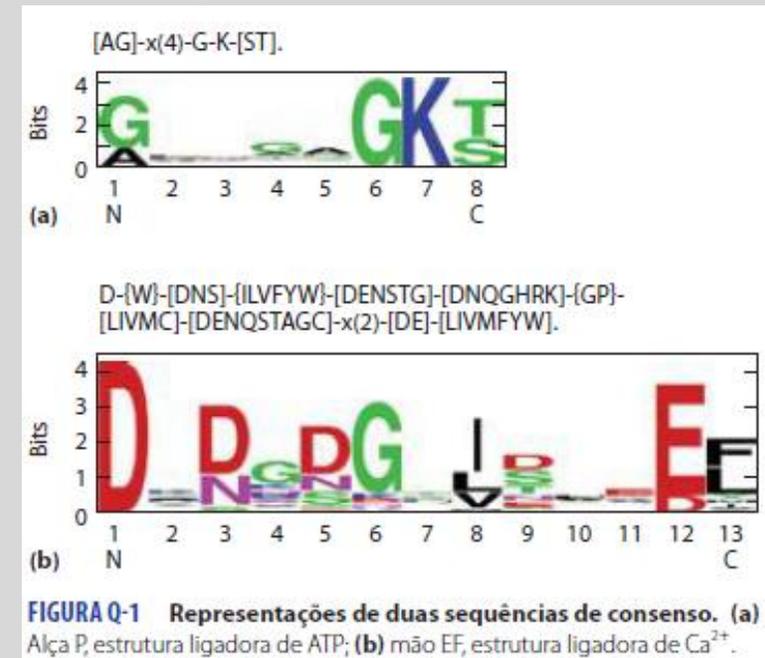
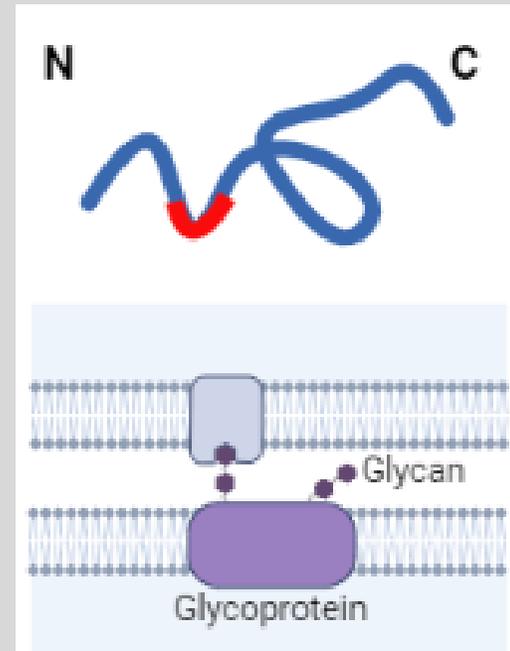
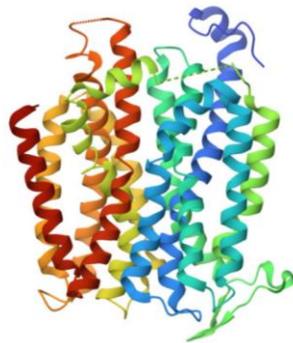
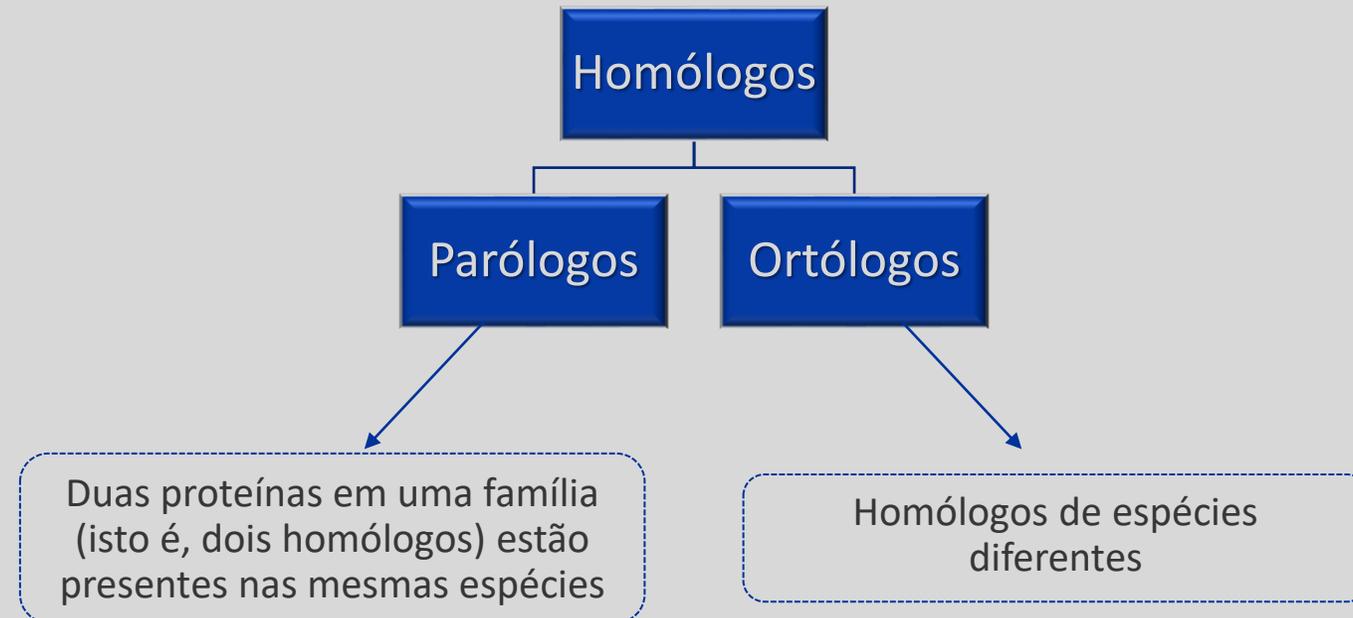


FIGURA Q-1 Representações de duas sequências de consenso. (a) Alça P, estrutura ligadora de ATP; (b) mão EF, estrutura ligadora de Ca^{2+} .

Que informações podemos extrair da estrutura primária de proteínas?

1. Identificação de homólogos

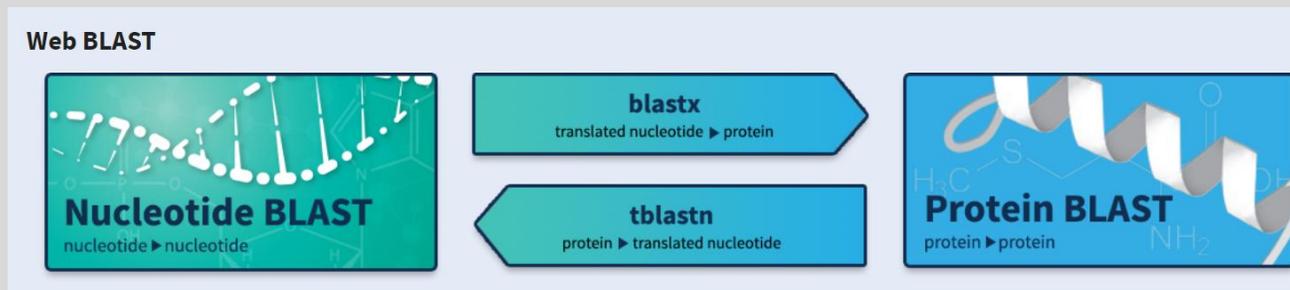
- Proteínas membros das mesmas famílias são denominados proteínas homólogas ou homólogos.



Que informações podemos extrair da estrutura primária de proteínas?

1. Identificação de homólogos

- Proteínas membros das mesmas famílias são denominados proteínas homólogas ou homólogos.



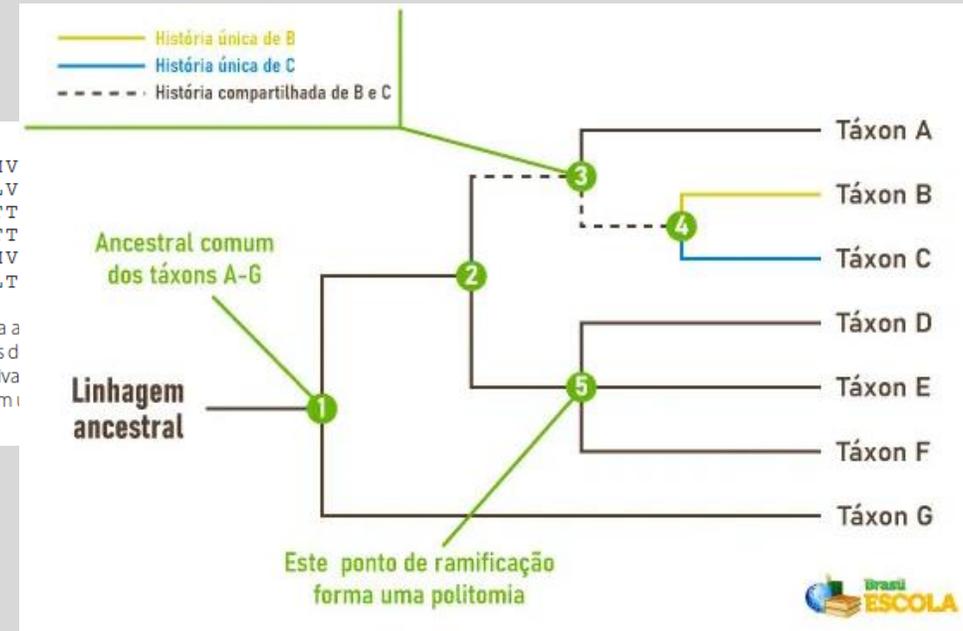
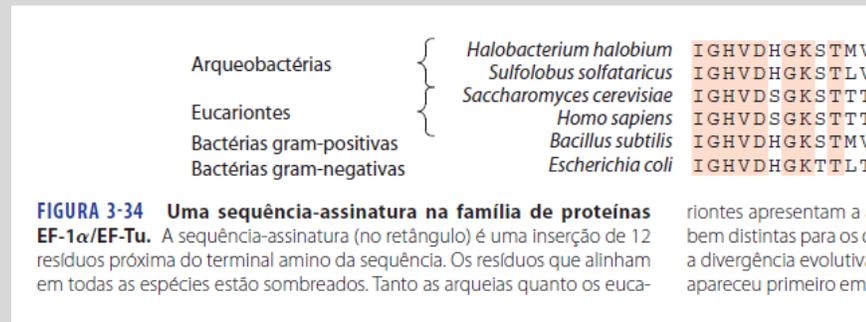
```
Escherichia coli TGNRTIAVYDLGGGTFDISIIEIDEVDGEKTFEVLATNGDTHLGGEDFDSRLIHYL  
Bacillus subtilis DEDQTILLYDLGGGTFDVSILELGDG TFEVRSTAGDNRLGGDDFDQVIIDHL
```

Intervalo

Que informações podemos extrair da estrutura primária de proteínas?

1. Identificação de homólogos

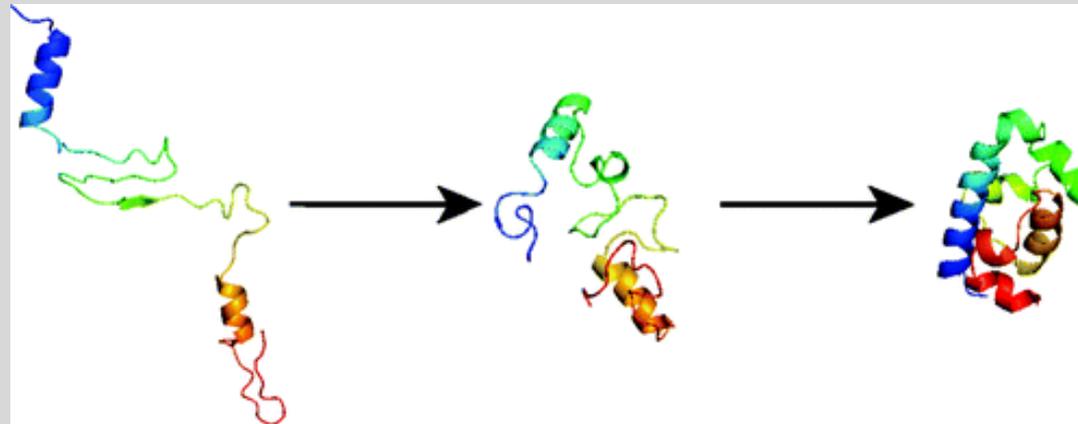
- Proteínas membros das mesmas famílias são denominados proteínas homólogas ou homólogos.



Que informações podemos extrair da estrutura primária de proteínas?

“A **função** de cada proteína depende de sua **estrutura tridimensional**, que, por sua vez, **é determinada** em grande parte por sua **estrutura primária**. Portanto, a informação transmitida por uma sequência de proteínas é limitada apenas por nossa própria compreensão dos princípios estruturais e funcionais.”

“As semelhanças de estruturas tridimensionais revelam algumas vezes relações evolutivas nas quais a homologia e sequências foi apagada pelo tempo.”



Fonte: G. Bowman, V. Voelz, e V. S. Pande/Reprodução ICTP

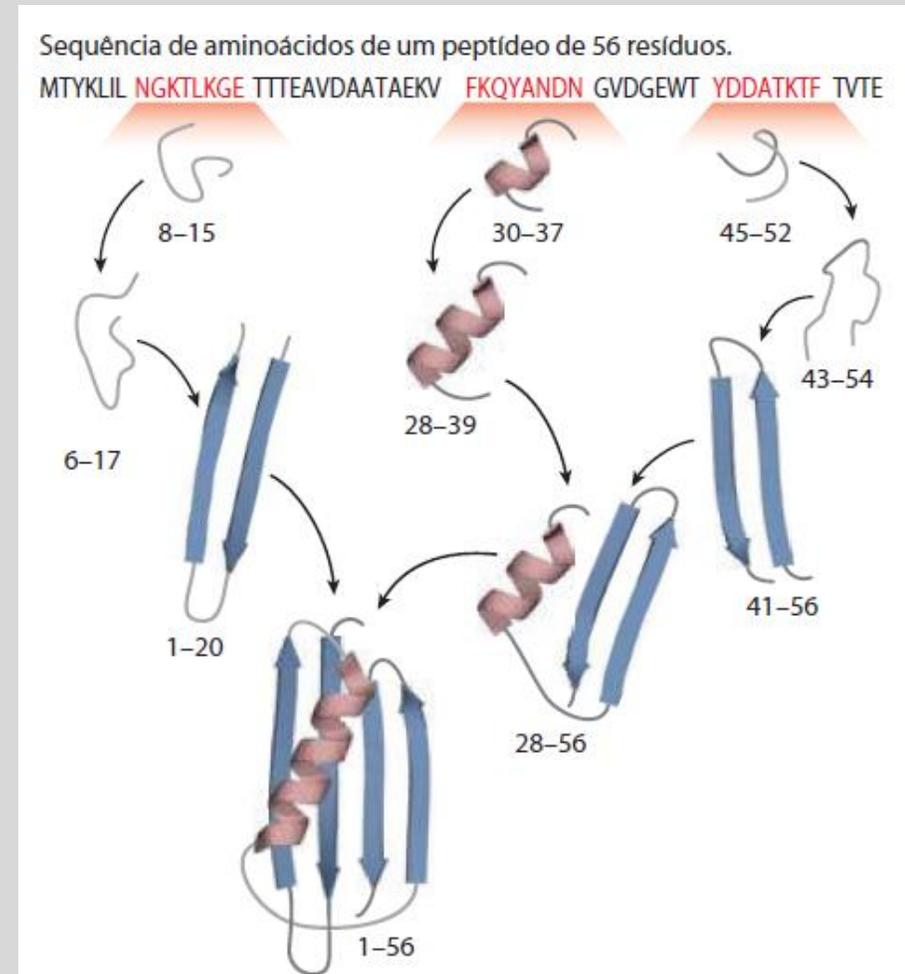
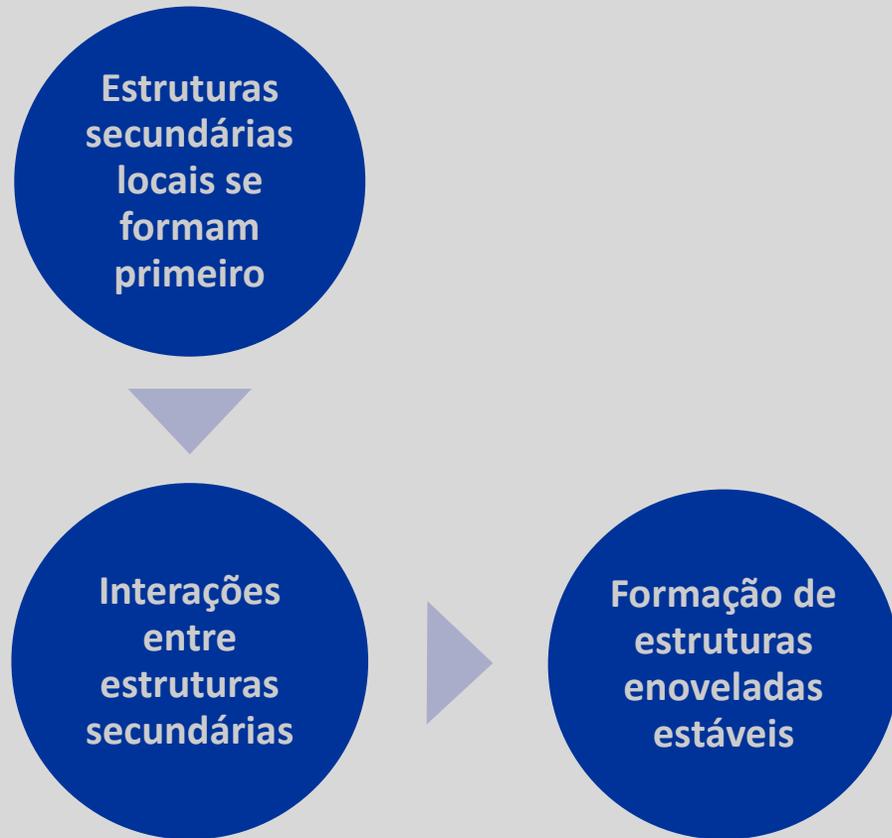
Metodologias de predição de estruturas proteicas

1. Que proteína queremos modelar?



Metodologias de predição de estruturas proteicas

- O enovelamento proteico
 - As principais vias de enovelamento são hierárquicas.



Metodologias de predição de estruturas proteicas

- O enovelamento proteico
 - O processo de enovelamento pode ser visto como um tipo de funil de energia livre.

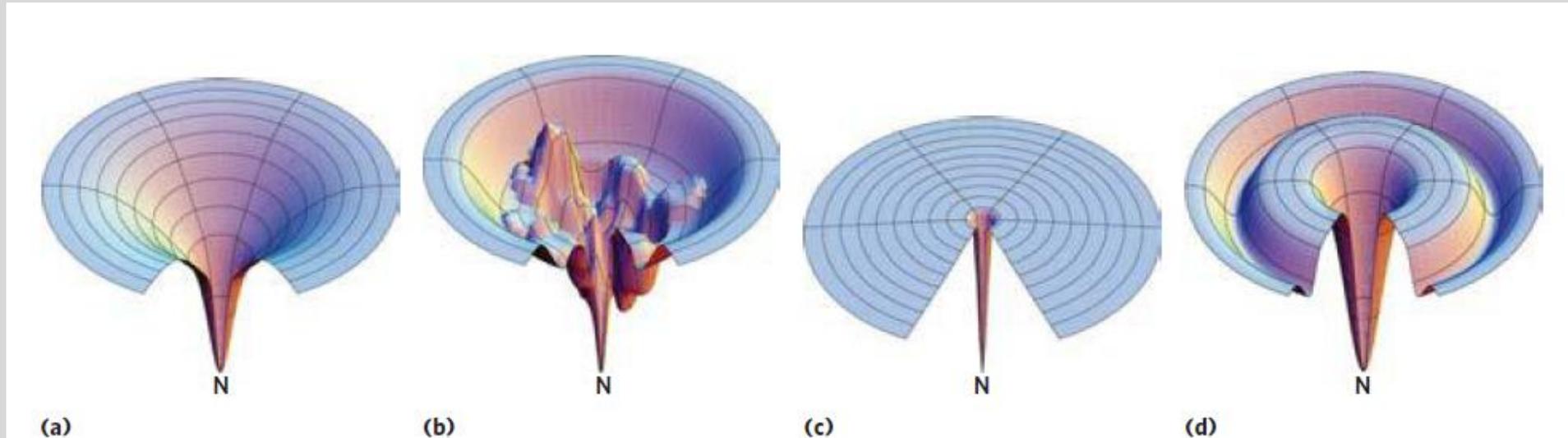


FIGURA 4-29 Termodinâmica do enovelamento proteico, mostrado como funil de energia livre. À medida que as proteínas se dobras, o espaço conformacional que pode ser explorado pela estrutura fica restrito. Este é modelado como funil termodinâmico tridimensional, com ΔG representado como profundidade e com a estrutura nativa (N) no fundo (ponto de menor energia-livre). O funil para determinada proteína pode ter uma variedade de formas, dependendo do número e dos tipos de intermediários de enovelamento nas vias de enovelamento. Qualquer intermediário de enovelamento com estabilidade significativa e um tempo de vida finita estaria representado como um mínimo de energia-livre local – uma depressão na superfície do funil. **(a)** Um funil simples, mas relativamente amplo e suave representa uma proteína com múltiplas vias de enovelamento (ou seja, a ordem em que as partes distintas da proteína se dobras seria de alguma forma aleatória), mas que assume suas estruturas tridimensionais sem

intermediários de enovelamento com estabilidade significativa. **(b)** Este funil representa uma proteína mais comum com múltiplos intermediários de enovelamento possíveis com estabilidade significativa nas várias vias que levam à estrutura nativa. **(c)** Uma proteína com estrutura nativa estável, essencialmente sem intermediários de enovelamento com estabilidade significativa, e apenas uma ou muito poucas vias de enovelamento produtivo está mostrada como um funil com uma depressão estreita levando à forma nativa. **(d)** Uma proteína com intermediários de enovelamento com estabilidade substancial em praticamente qualquer via que leva ao estado nativo (ou seja, uma proteína na qual um motivo ou domínio particular sempre se dobra rapidamente, mas as outras partes da proteína se dobras mais lentamente e em ordem aleatória) está representada por um funil com uma depressão principal ao redor da depressão que leva à forma nativa.

Metodologias de predição de estruturas proteicas

- **Ab initio**

Inteligência artificial



Padrões de enovelamento conhecidos com base nos motivos de enovelamento



Modelo proteico

- **Homologia**

Inteligência artificial



Consulta estruturas similares em bancos de dados e constrói a estrutura com base no enovelamento de um homólogo



Modelo proteico

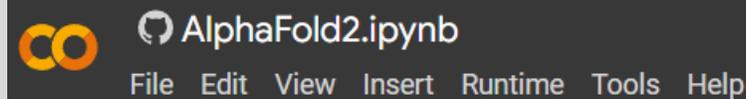
Metodologias de predição de estruturas proteicas

- Predição de estruturas proteicas

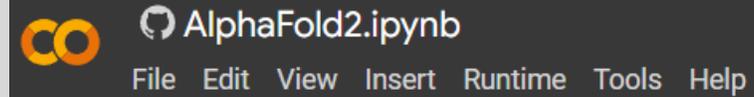


Robetta

 robetta.bakerlab.org



- Refinamento



GalaxyWEB

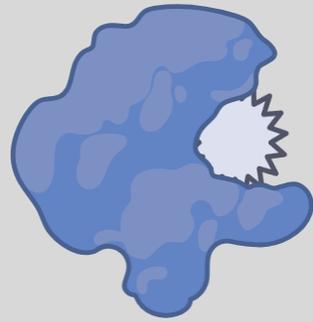
A web server for protein structure prediction, refinement, and related methods
Galux Inc. & Computational Biology Lab, Department of Chemistry, Seoul National University

Metodologias de predição de estruturas proteicas

- **Validação**

- **RMSD:**
 - Fornece o desvio médio entre os átomos correspondentes de duas proteínas;
- **MOLPROBITY:**
 - Gráfico de Ramachandran;
- **VERIFY 3D:**
 - Determina a compatibilidade de um modelo atômico (3D) com sua própria sequência de aminoácidos (1D) atribuindo uma classe estrutural com base em sua localização e ambiente (alfa, beta, loop, polar, apolar etc) e comparando os resultados com boas estruturas;
- **Q-MEAN:**
 - O modelo qualitativo de análise de energia (Q-MEAN) explora distribuições de distância de estruturas de proteínas determinadas experimentalmente que são homólogas ao modelo que está sendo avaliado.

Docking molecular e o redirecionamento de fármacos

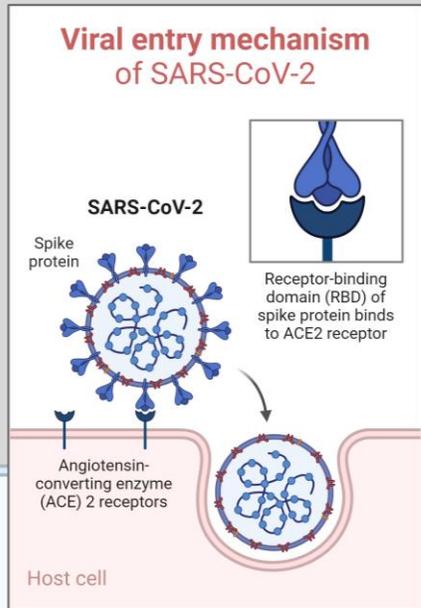


Enzima

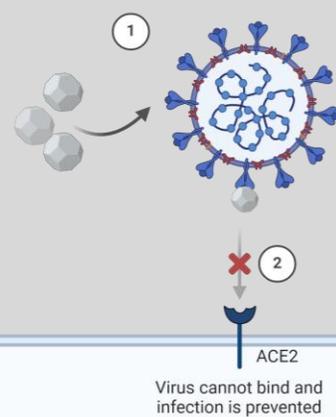


Ligante

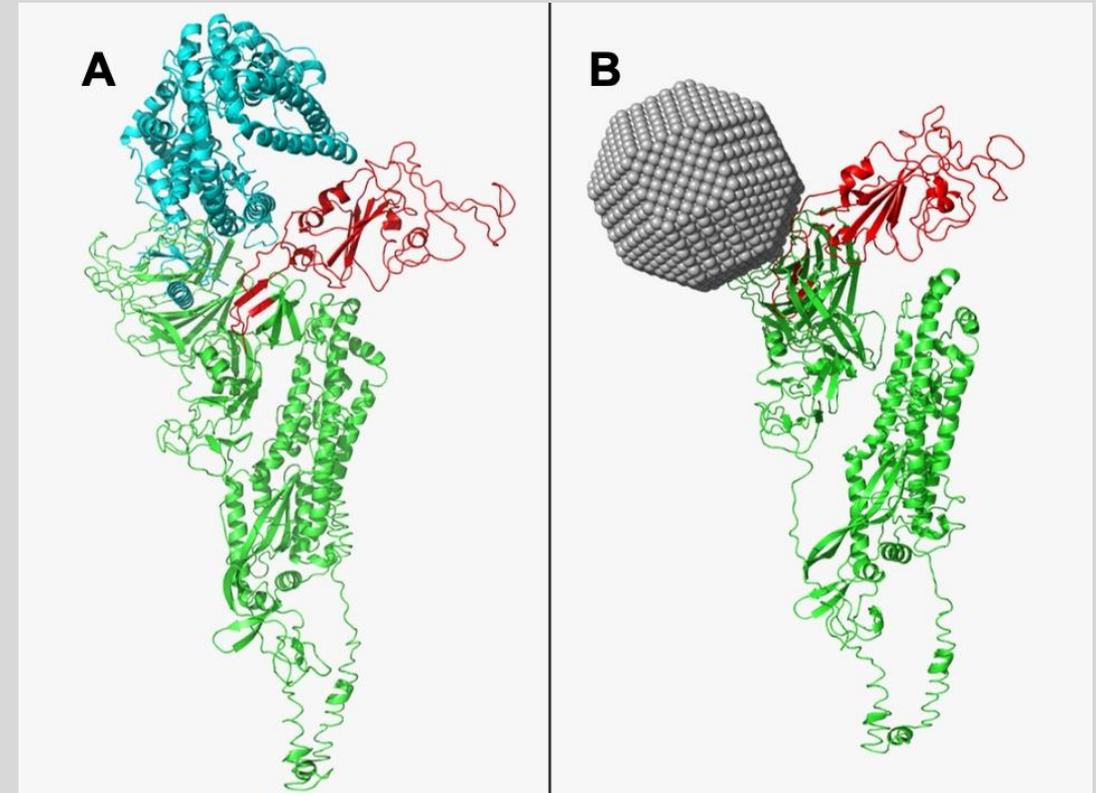
Proposed therapeutic treatments for COVID-19 targeting SARS-CoV-2 viral entry mechanism



Target RBD of spike protein:



Host cell



Dados não publicados



OBRIGADA!!!



Ma. Iasmin Cartaxo Taveira

Ribeirão Preto

09/05/2024