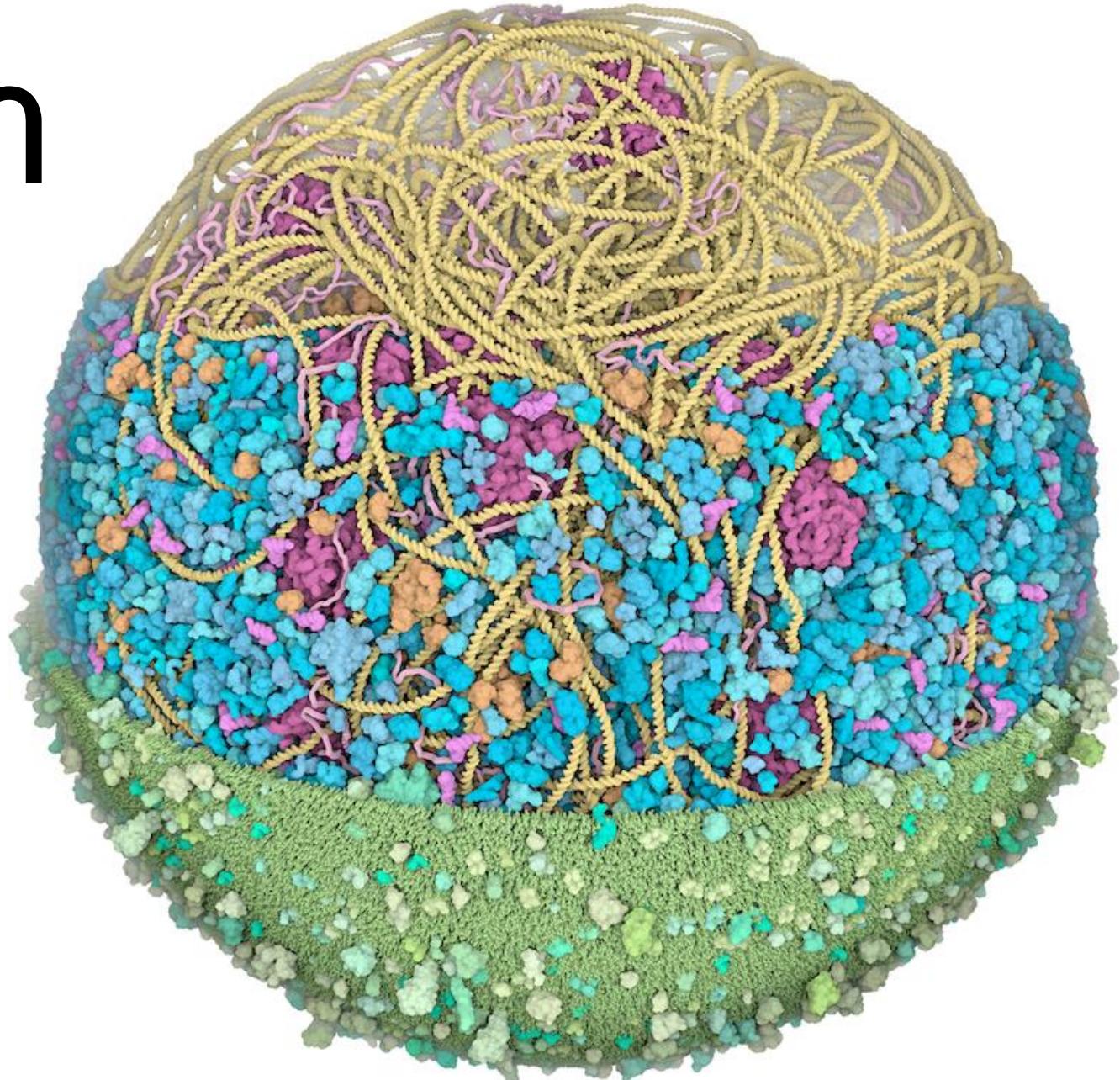
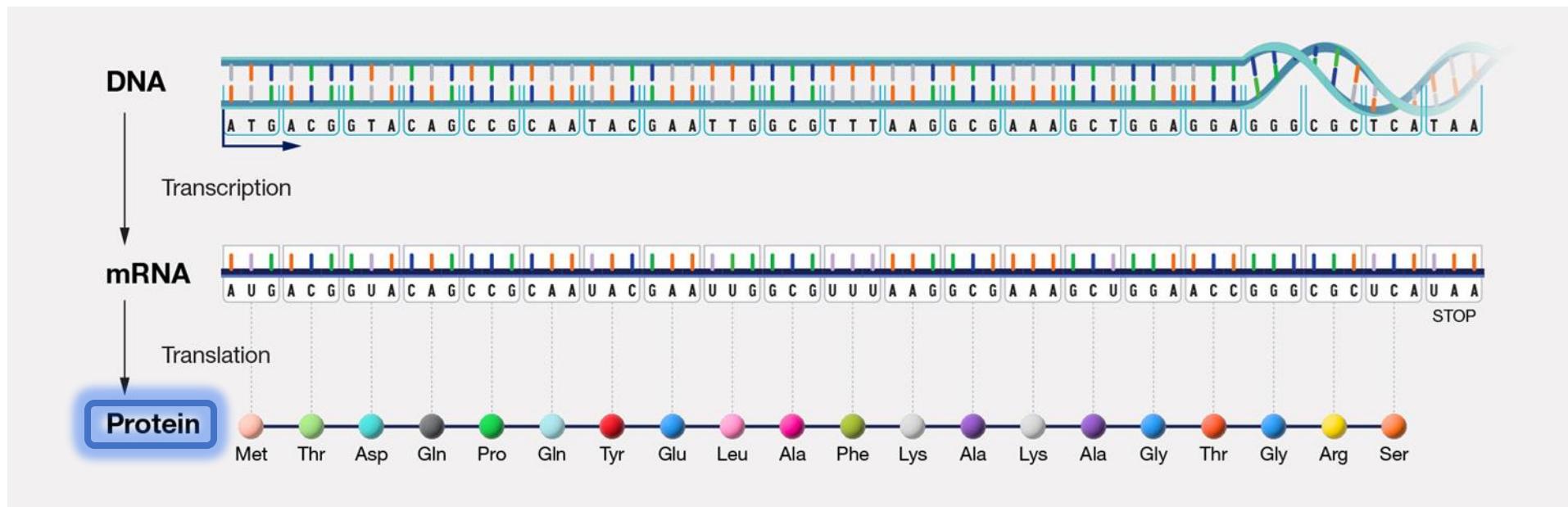


Modelagem de Proteínas com IA

Madson Aragão, MSc
Bioinformática, UFMG
Belo Horizonte, 2025

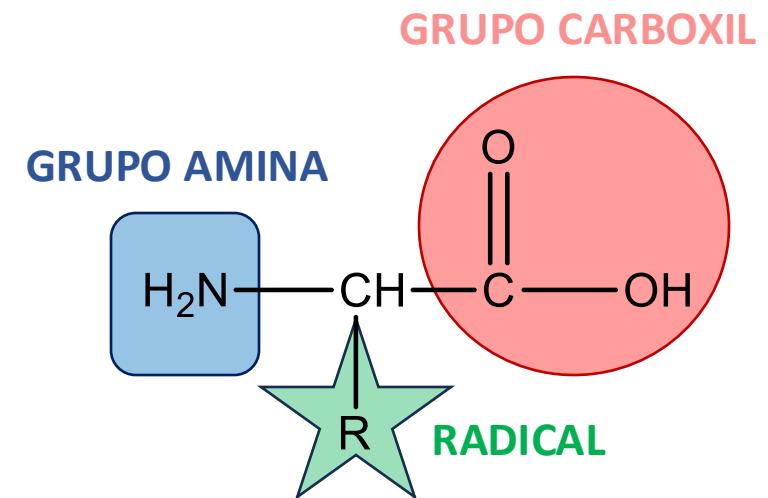
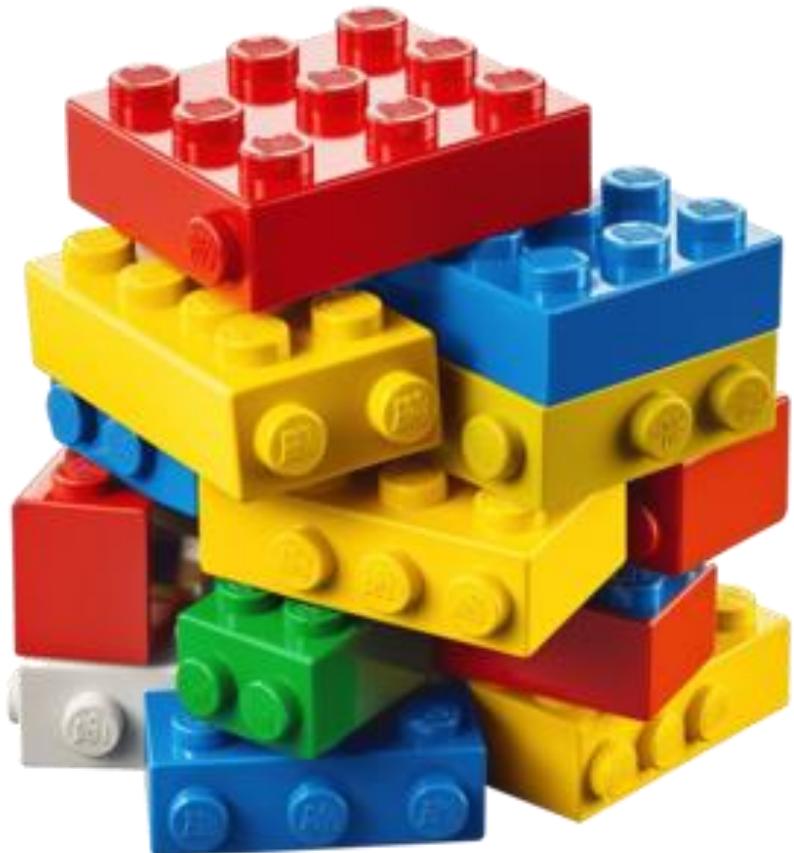


DOGMA CENTRAL DA BIOLOGIA MOLECULAR



aprox. 20.000 genes codificantes de proteínas
aprox. 80.000-400.000 proteínas

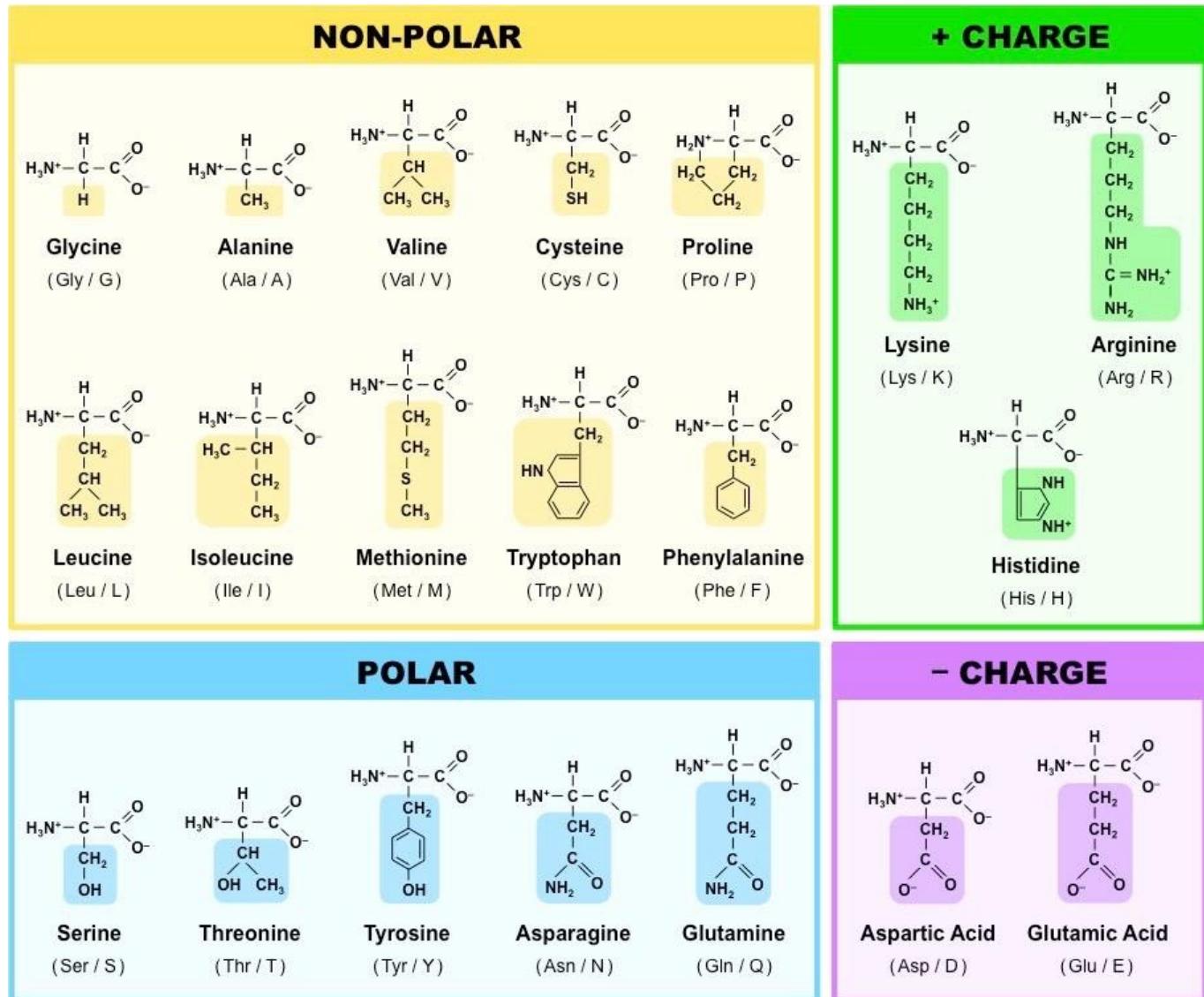
AMINOÁCIDOS (MONÔMERO NATURAIS)

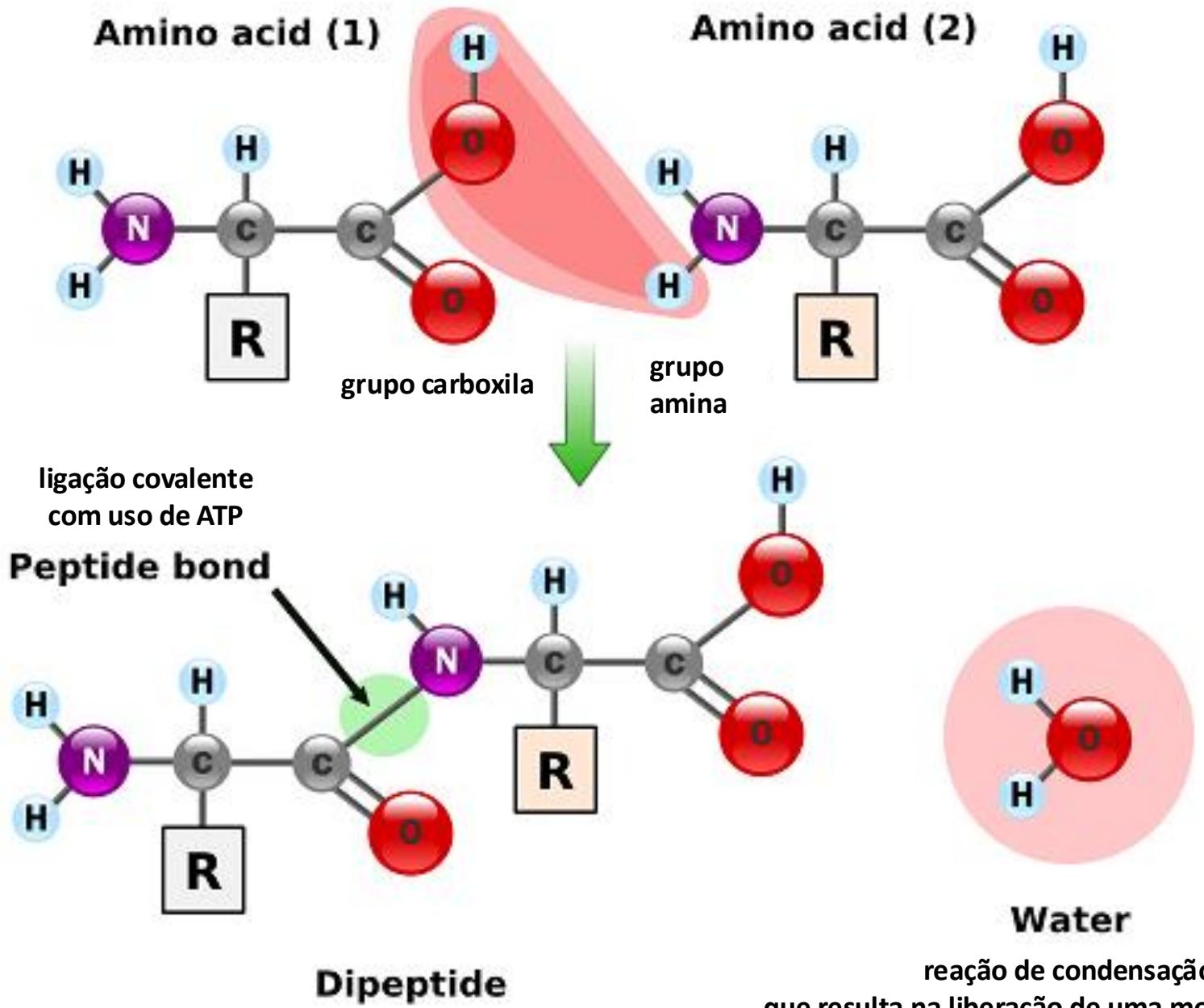


estrutura básica de um aminoácido

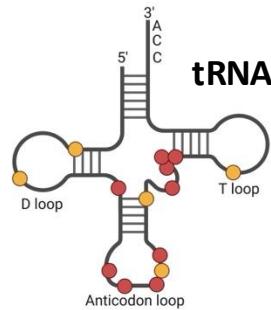
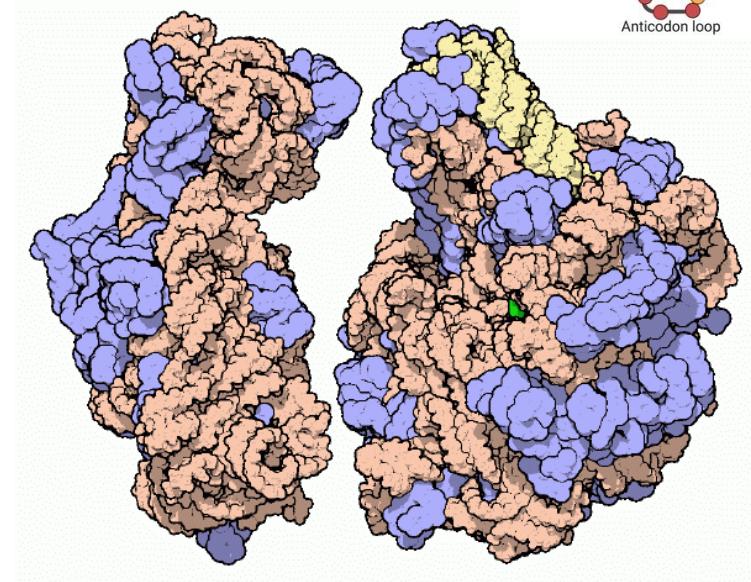
AMINOÁCIDOS (-R)

- 20 aminoácidos naturais;
- Algumas variações surgem por modificações químicas;
- Nós humanos sintetizamos apenas 11 aminoácidos: Alanina, Asparagina, Ácido aspártico, Ácido glutâmico, Glutamina, Serina;
- Os outros 9 aminoácidos, **chamados de aminoácidos essenciais**, são obtidos através da alimentação;
- Os aminoácidos podem ser classificados de acordo com suas propriedades físico-químicas.

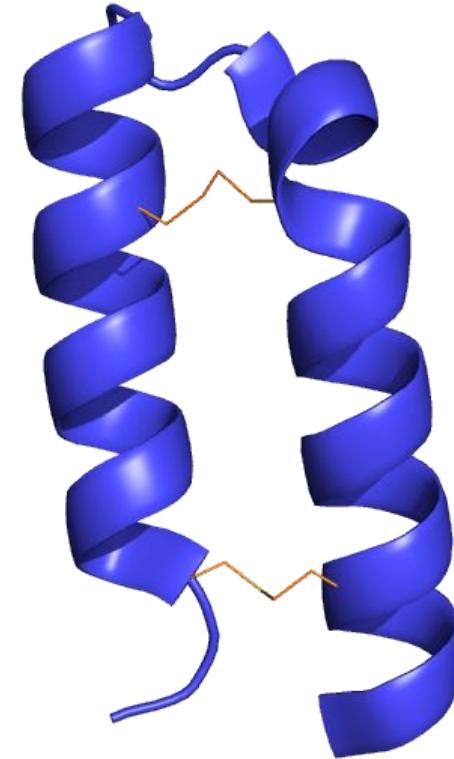
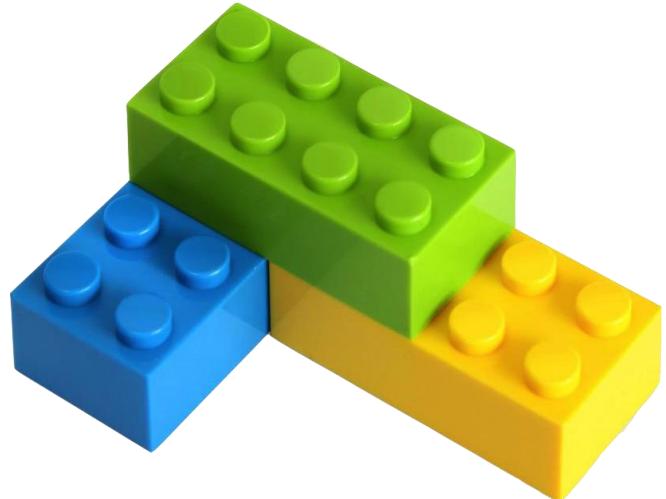




atividade catalítica do ribossomo



PEPTÍDEO (<50)



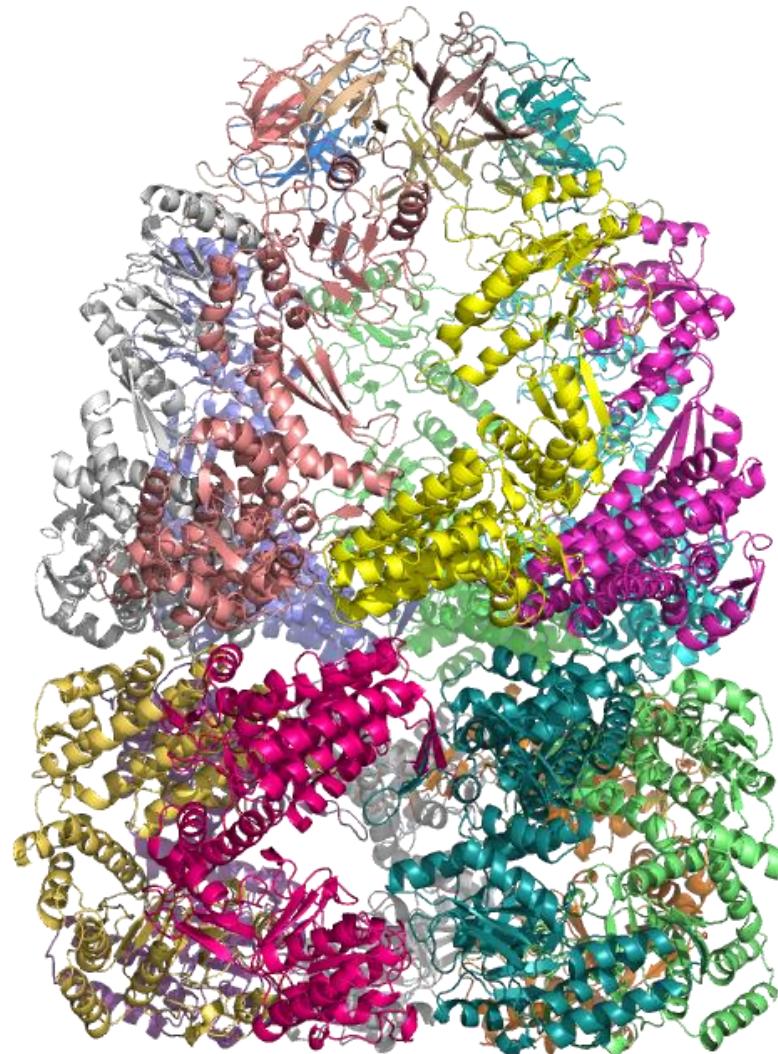
Alpha-helical hairpin (PDB- 1EI0; 38 aa)
Homo sapiens

PROTEÍNA (>50)



Chitinase B (PDB- 1E6Z; 499 aa)
Serratia marcescens

PROTEÍNAS (COMPLEXOS)



Chaperonin (PDB- 1AON; 6.500 aa)
Escherichia coli

APLICAÇÃO DA BIOINFORMÁTICA FOCADA EM PROTEÍNAS E ML

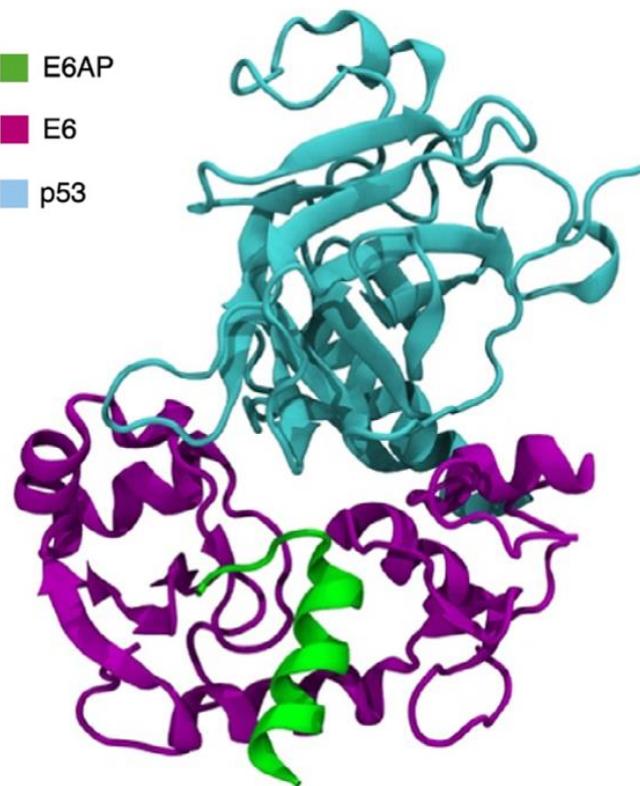
> *Biopolymers*. 2022 Oct;113(10):e23524. doi: 10.1002/bip.23524. Epub 2022 Aug 18.

Association strength of E6 to E6AP/p53 complex correlates with HPV-mediated oncogenesis risk

Matheus Vitor Ferreira Ferraz ^{1 2}, Isabelle Freire Tabosa Viana ¹, Danilo Fernandes Coêlho ^{1 2}, Carlos Henrique Bezerra da Cruz ³, Maíra de Arruda Lima ¹, Madson Allan de Luna Aragão ¹, Roberto Dias Lins ^{1 2}

Affiliations + expand

PMID: 36281776 DOI: 10.1002/bip.23524



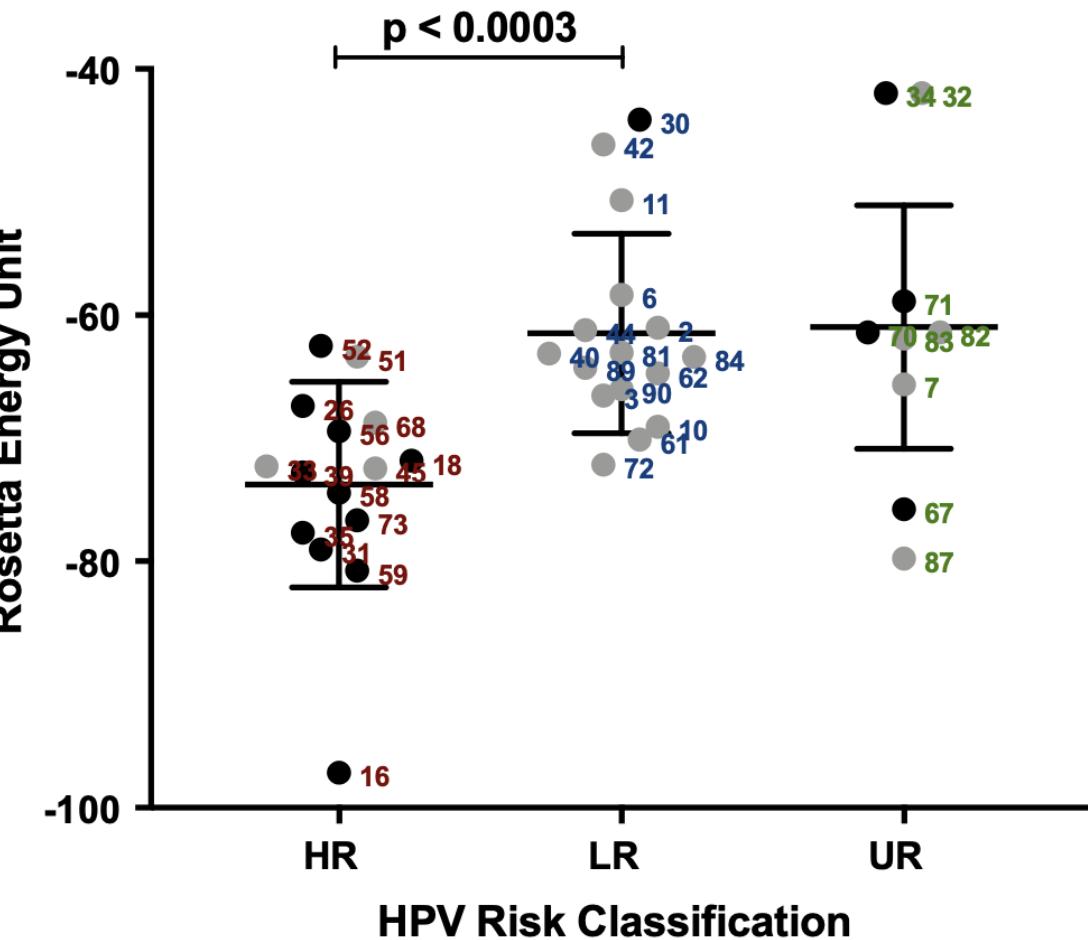
FULL TEXT LINKS

WILEY Full Text Article

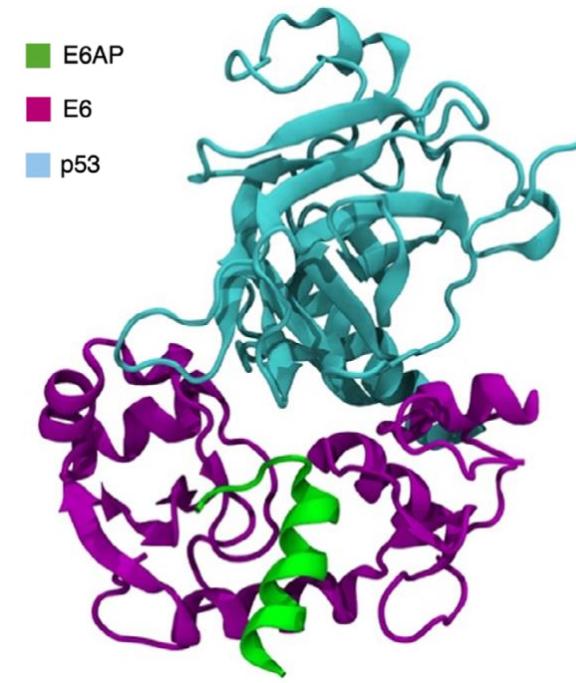
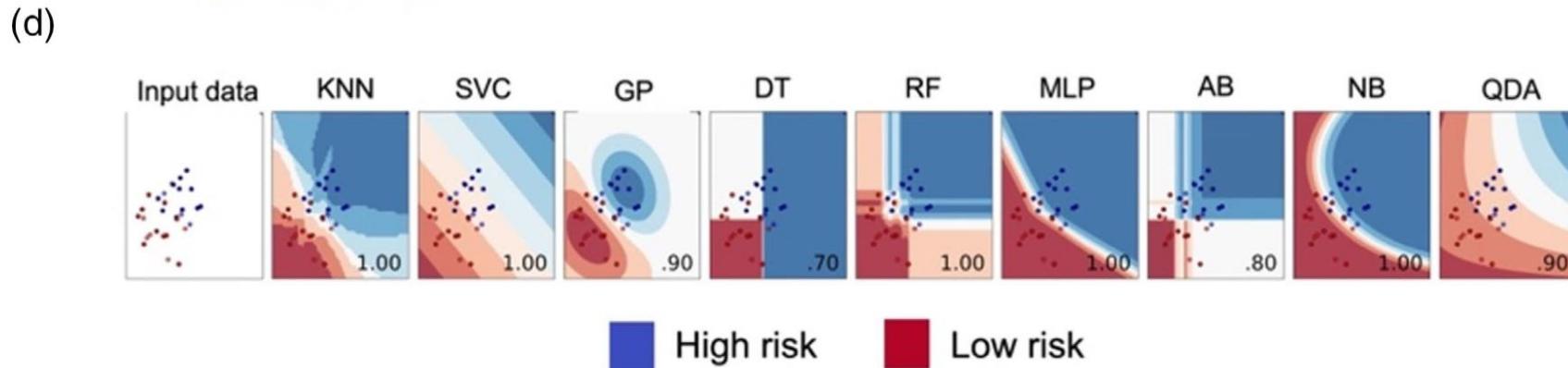
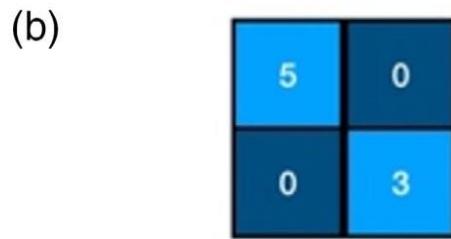
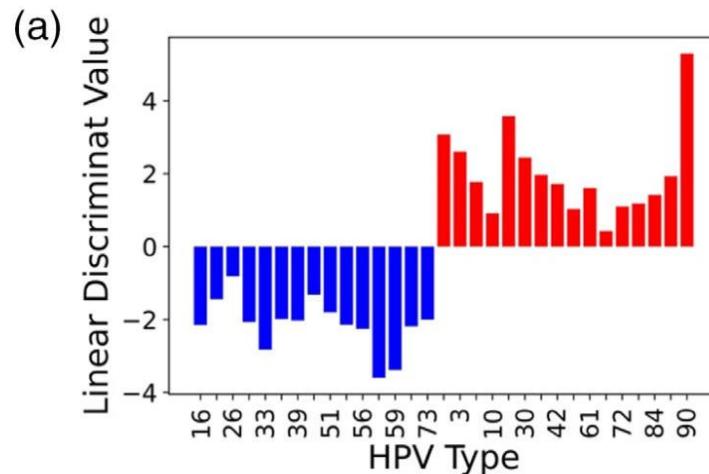
ACTIONS

Cite

Collections



APLICAÇÃO DA BIOINFORMÁTICA FOCADA EM PROTEÍNAS E ML



APLICAÇÃO DA BIOINFORMÁTICA FOCADA EM PROTEÍNAS E ML

> *Biopolymers*. 2022 Oct;113(10):e23524. doi: 10.1002/bip.23524. Epub 2022 Aug 18.

Association strength of E6 to E6AP/p53 complex correlates with HPV-mediated oncogenesis risk

Matheus Vitor Ferreira Ferraz ^{1 2}, Isabelle Freire Tabosa Viana ¹, Danilo Fernandes Coêlho ^{1 2},
Carlos Henrique Bezerra da Cruz ³, Maíra de Arruda Lima ¹, Madson Allan de Luna Aragão ¹,
Roberto Dias Lins ^{1 2}

Affiliations + expand

PMID: 36281776 DOI: [10.1002/bip.23524](https://doi.org/10.1002/bip.23524)

FULL TEXT LINKS



ACTIONS

“ Cite

Collections

Principais Achados:

- **Interações mais fortes** da proteína E6 com E6AP/p53 estão associadas a um maior risco de oncogênese;
- **Machine learning (ML) e deep learning (DL) preveem o risco de tipos de HPV ainda não classificados**, expandindo a vigilância de câncer relacionado ao HPV;
- A dinâmica molecular revelou que **fatores termodinâmicos são essenciais na oncogênese associada ao HPV**.

Impacto do Estudo:

- Melhora nas **estratégias de prevenção (vacinas e rastreamento)**;
- Novas ferramentas preditivas para **classificação de tipos de HPV**;
- Potencial para **desenvolver terapias que bloqueiem a interação E6/E6AP/p53**.

ESTRUTURA DE PROTEÍNAS

NÍVEIS ORGANIZACIONAIS

Hemoglobina
(PDB- 1HGA)

Pro
Ala
Asp
Lys
Arg

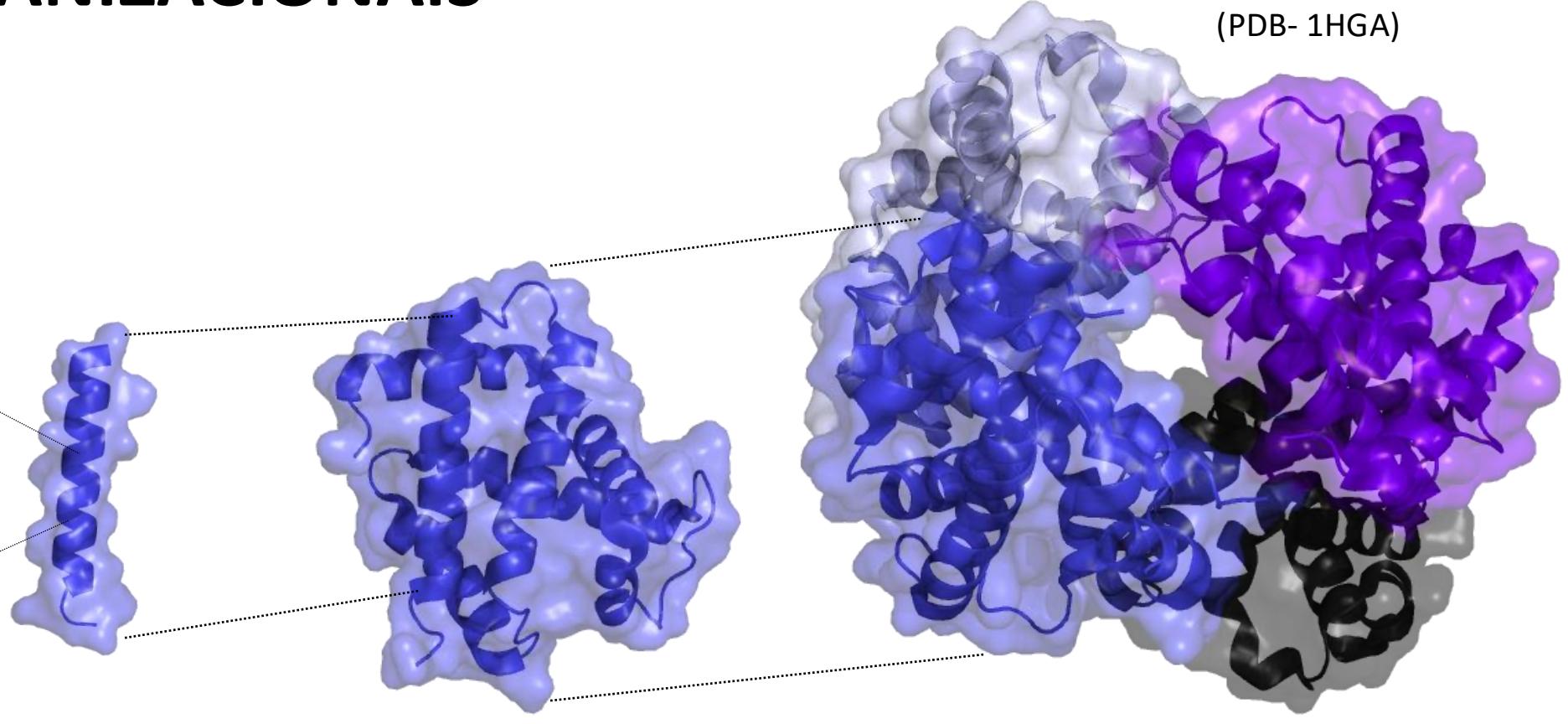
:

Primária

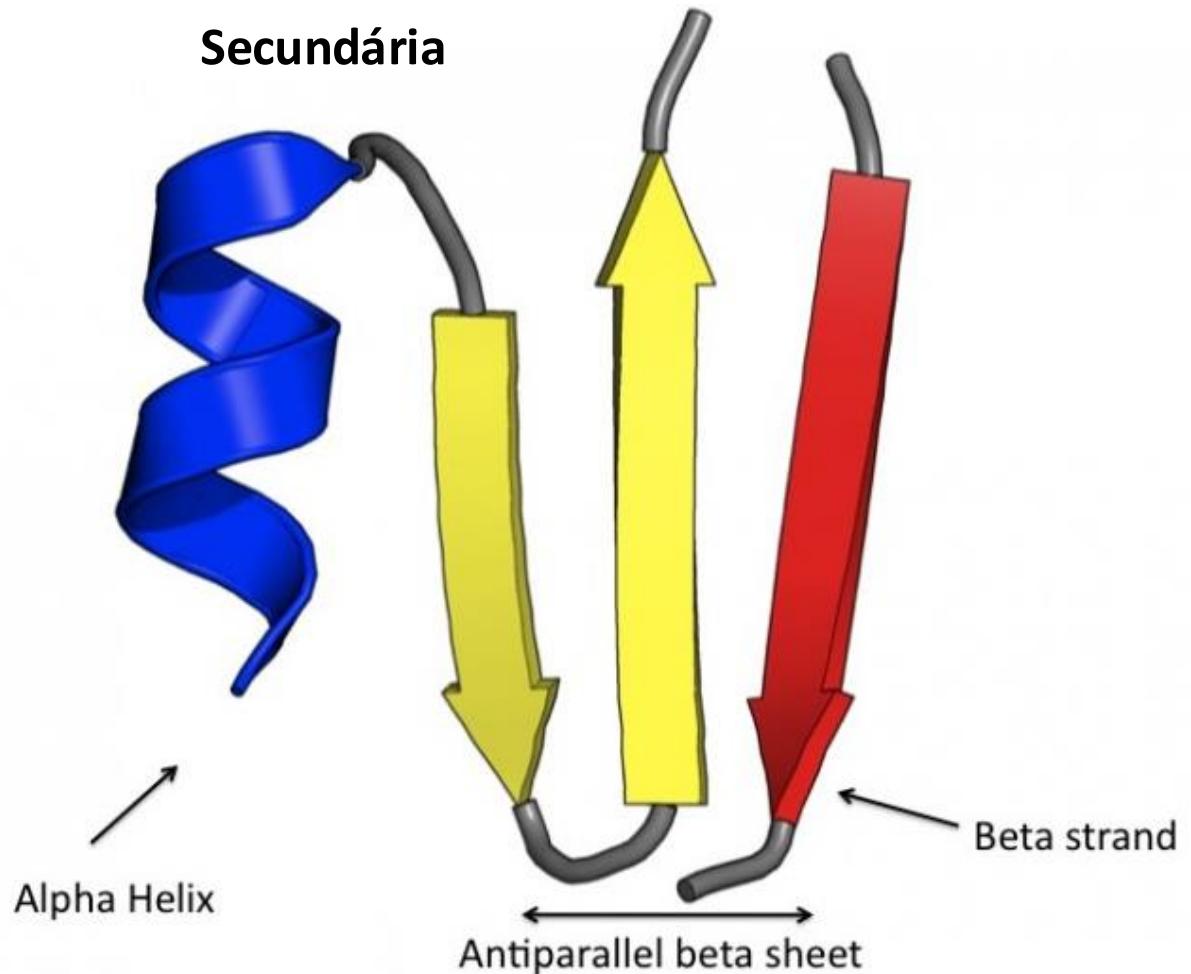
Secundária

Terciária

Quaternária

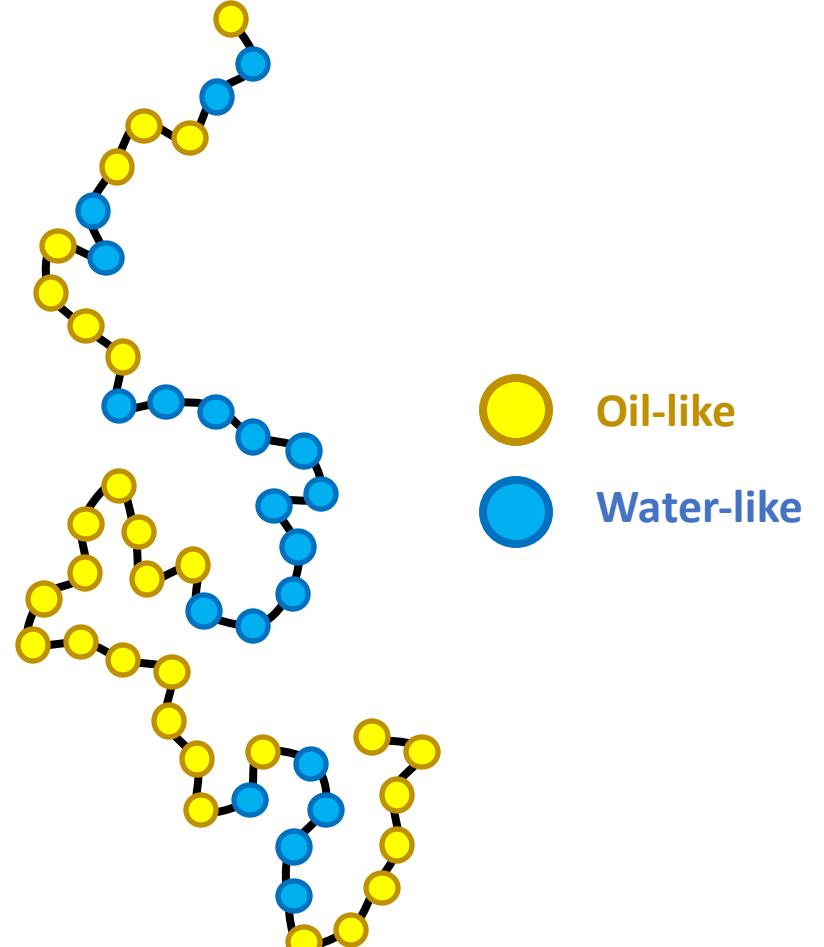
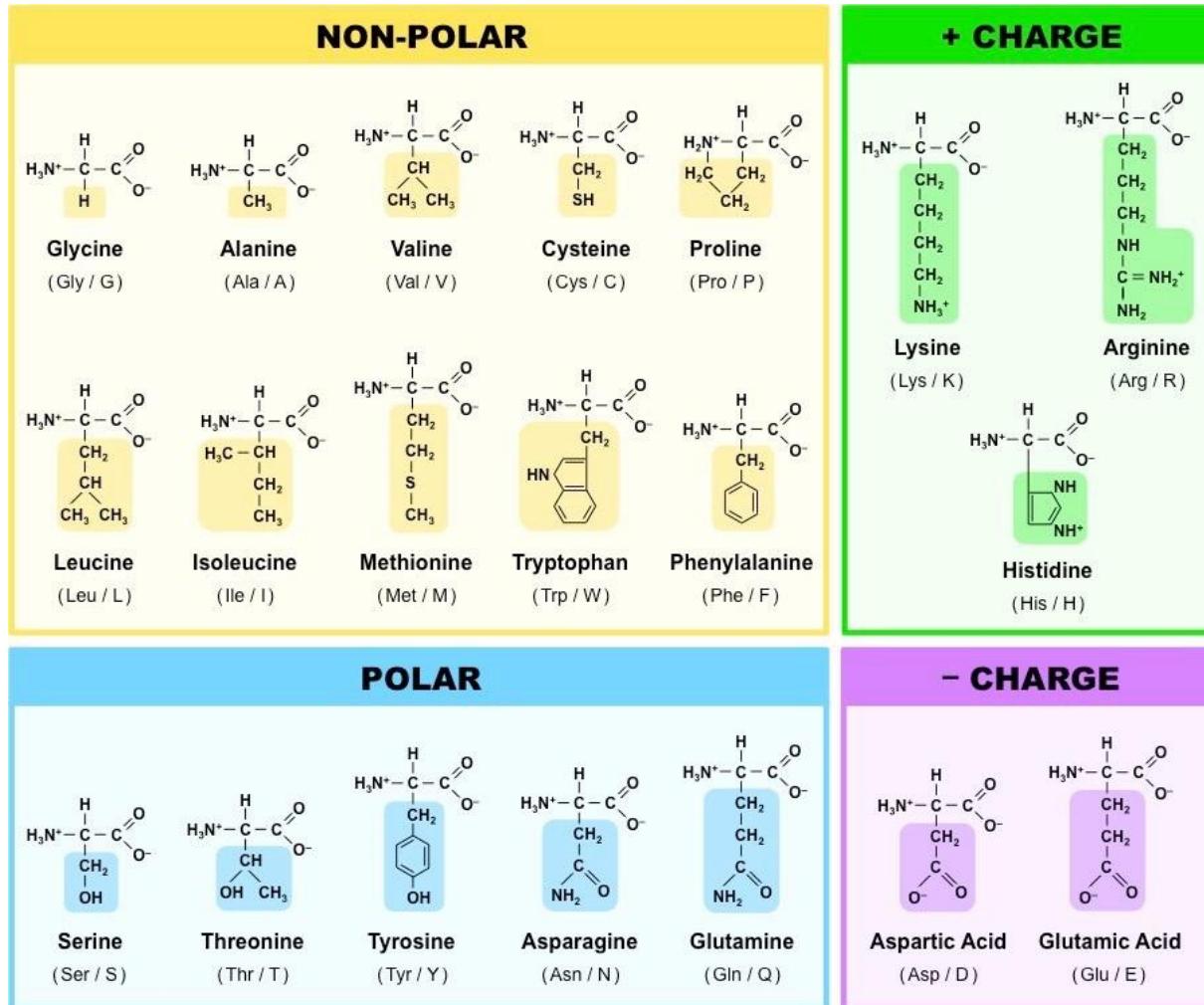


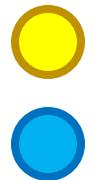
NÍVEIS ORGANIZACIONAIS



Quaternária

PROTEÍNAS & AMINOÁCIDOS

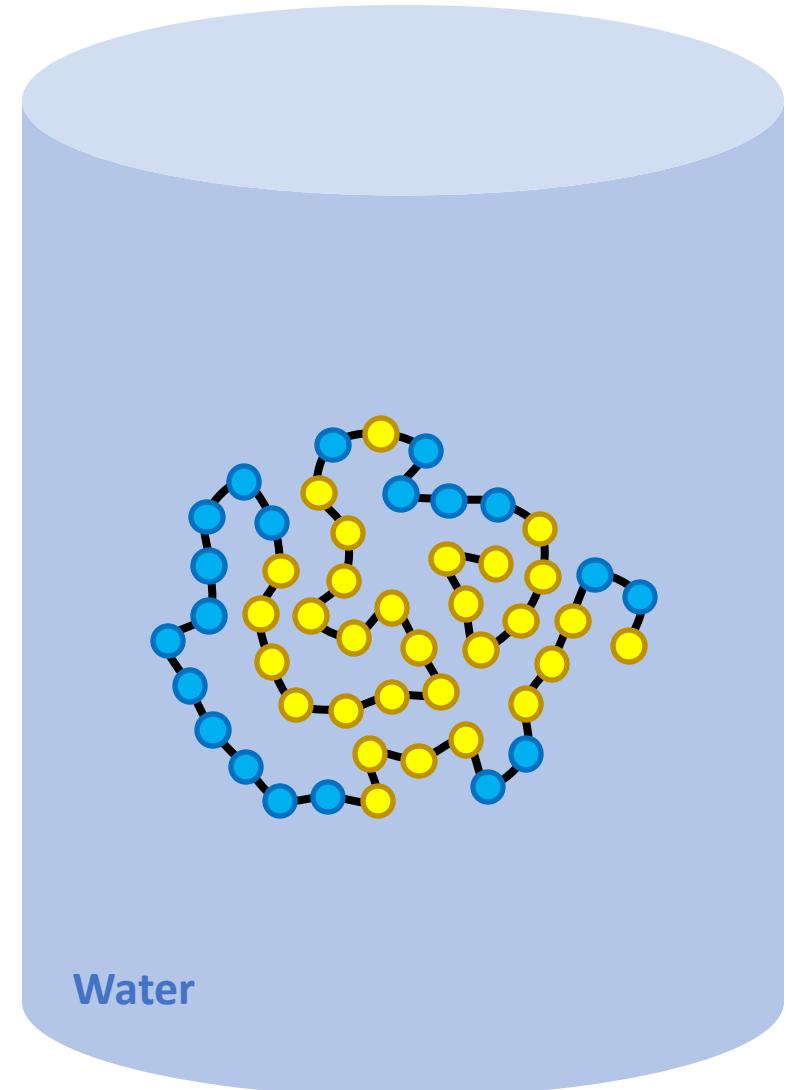
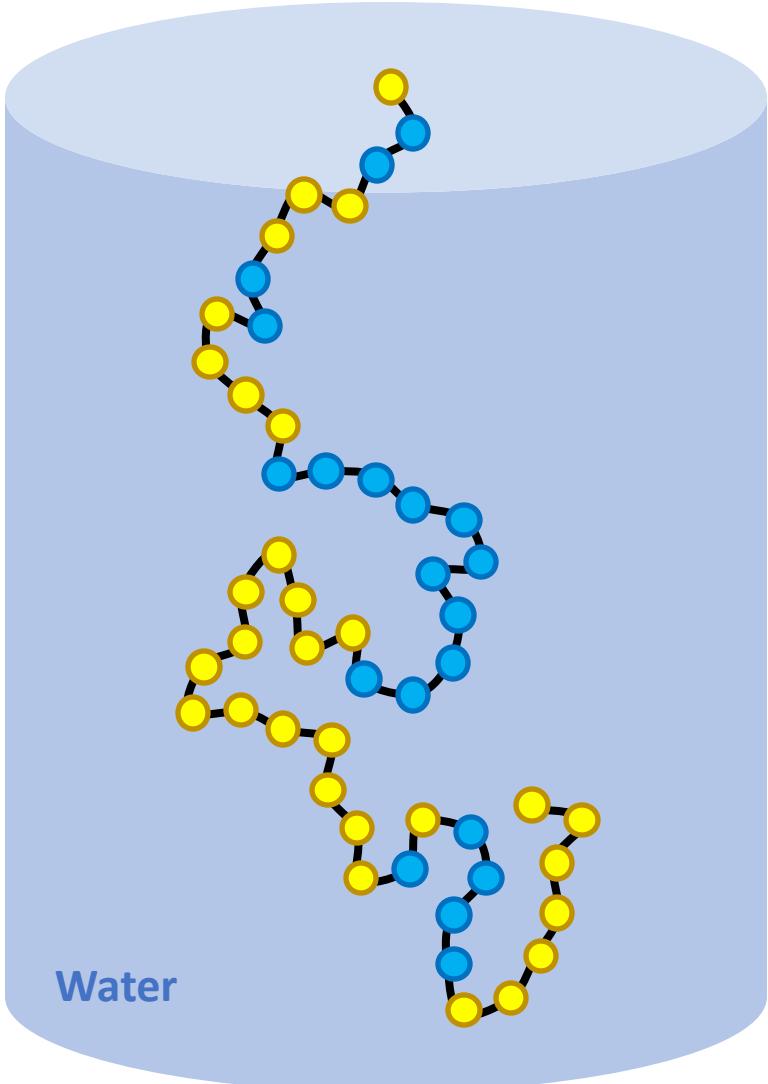


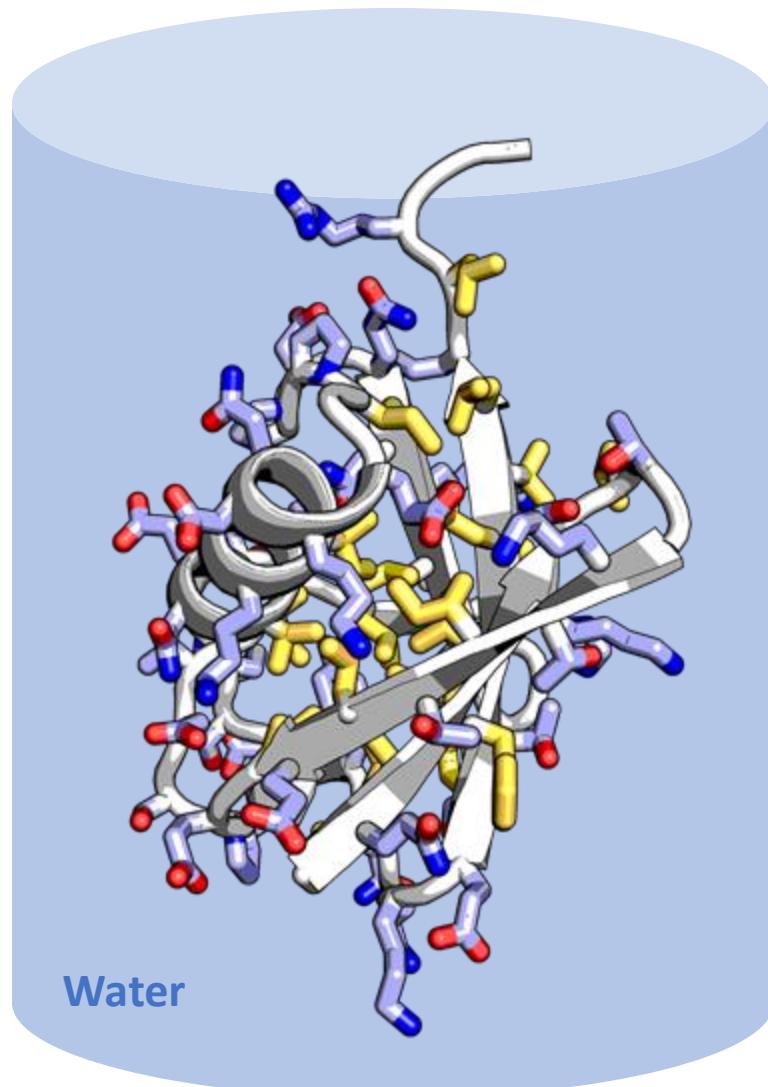


Oil-like



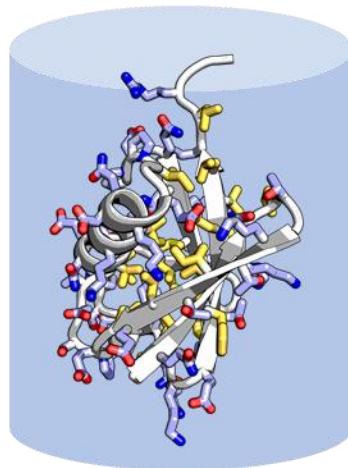
Water-like



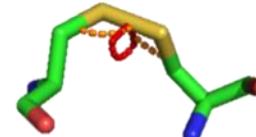


FATORES RESPONSÁVEIS PELO DOBRAMENTO

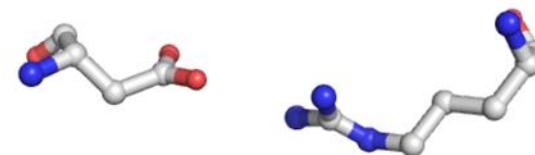
Solvation



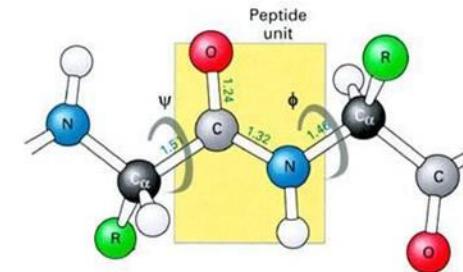
Disulfide bond



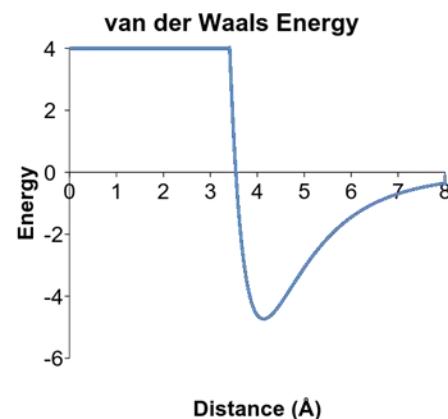
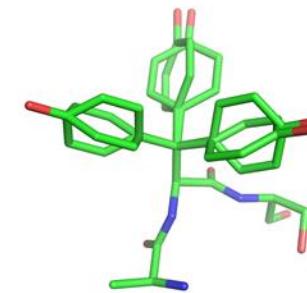
Ionic bond



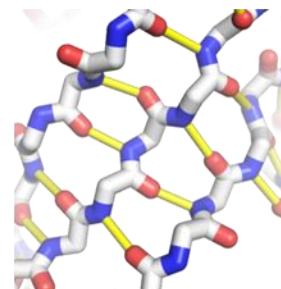
Backbone conformations



Sidechain conformations

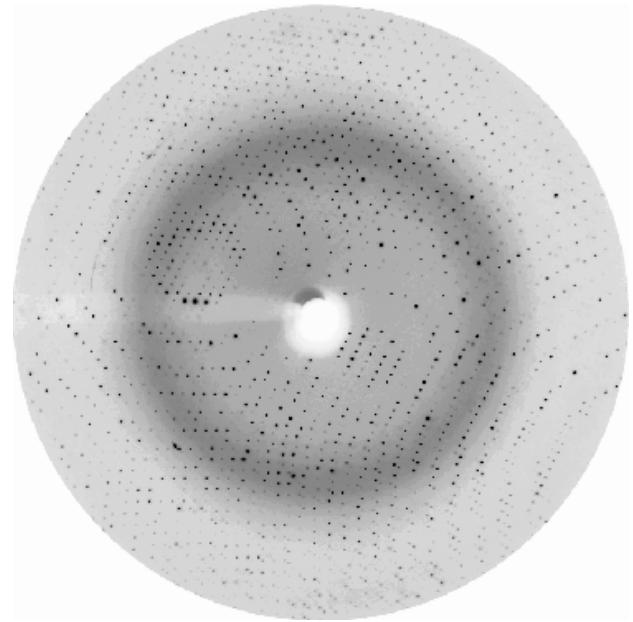


Hydrogen bond



Validação:

- Cristalografia por raio-x
- Ressonância magnética nuclear
- Microscopia crio-eletrônica



+210 mil proteínas depositadas



Structure Summary

Structure

Annotations

Experiment

Sequence

Genome

Ligands

Versions

Display Files ▾

Download Files ▾

Data API

Biological Assembly 1



4CQK

Crystal structure of ligand-bound NaD1

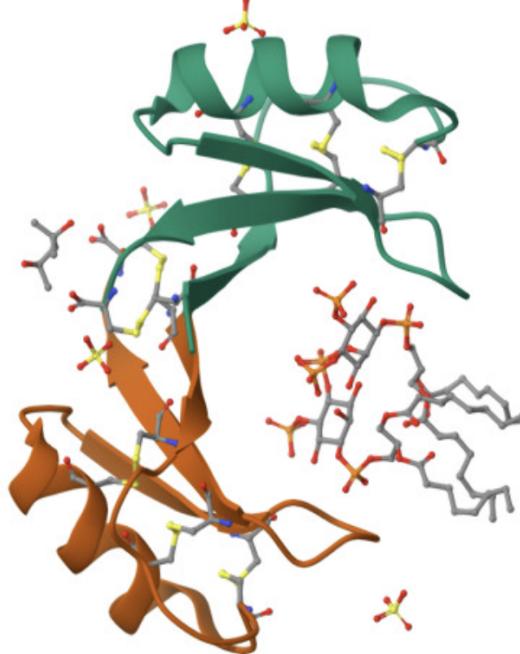
PDB DOI: <https://doi.org/10.2211/pdb4CQK/pdb>

Classification: PLANT PROTEIN

Organism(s): Nicotiana alata

Mutation(s): No

Membrane Protein: Yes OPM



Deposited: 2014-02-17 Released: 2014-04-16

Deposition Author(s): Lay, F.T., Mills, G.M., Poon, I.K.H., Baxter, A.A., Hulett, M.D., Kvansakul, M.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 1.60 Å

R-Value Free: 0.183

R-Value Work: 0.155

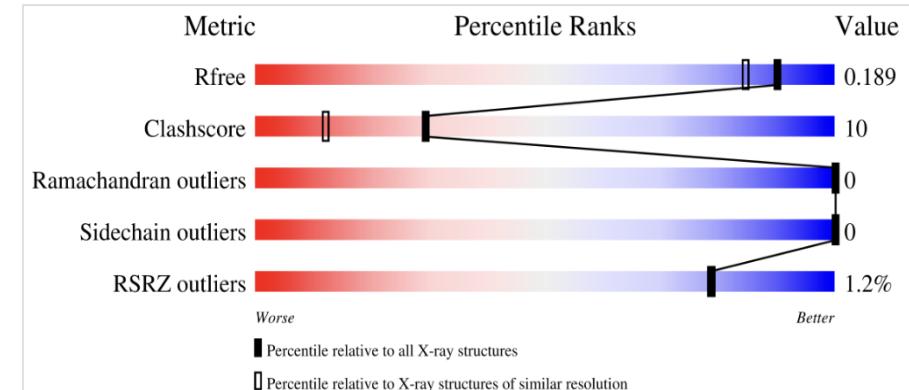
R-Value Observed: 0.156

Starting Model: experimental

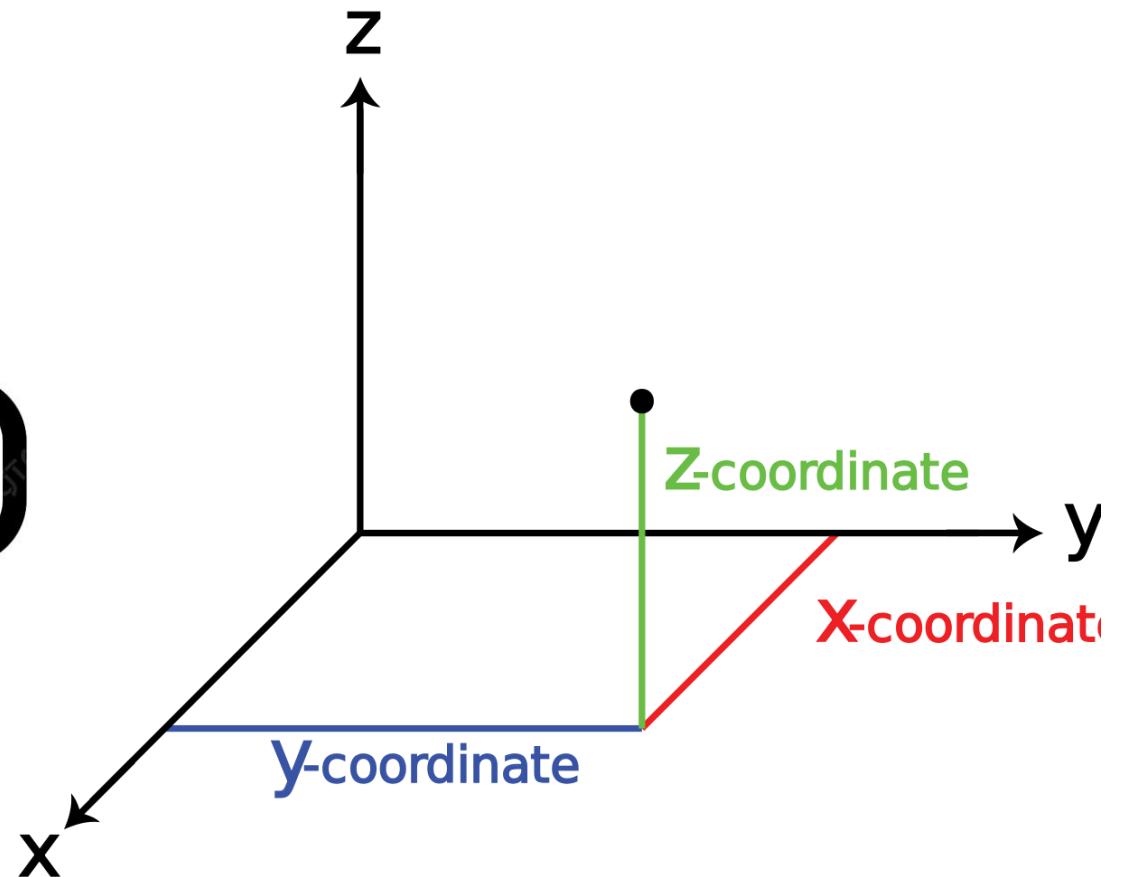
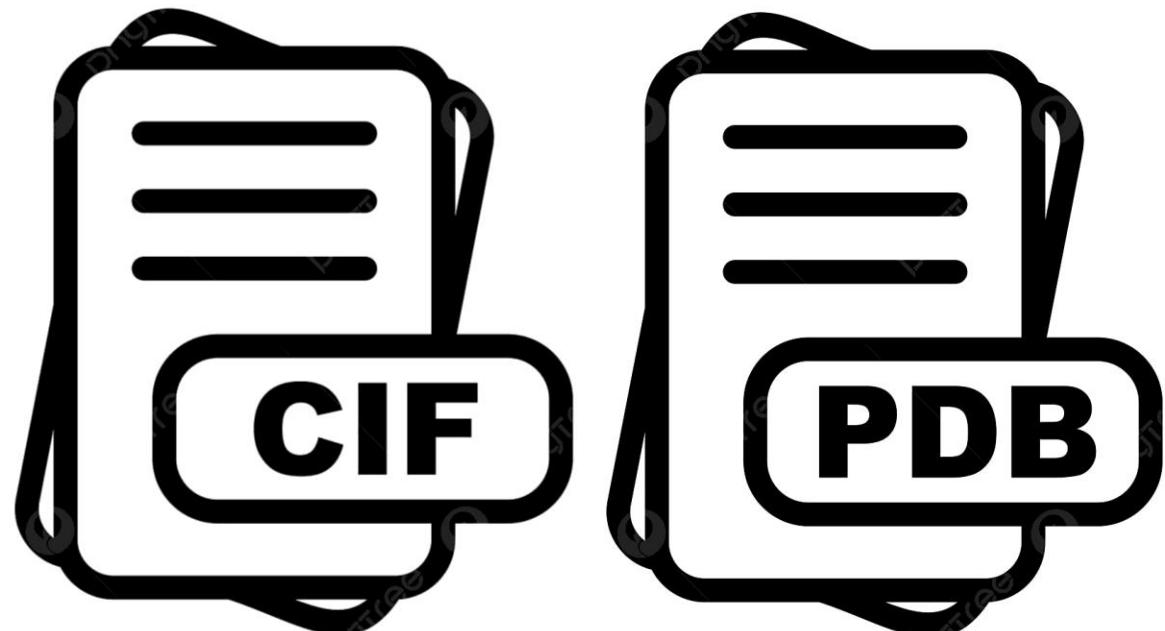
[View more details](#)

wwPDB Validation

3D Report Full Report

 Explore in 3D: [Structure](#) | [Sequence Annotations](#) |[Electron Density](#) | [Validation Report](#) | [Ligand Interaction \(PIO\)](#) |[Predict Membrane](#)

MAS ONDE/COMO ESSAS COORDENADAS SÃO ARMAZENADAS?



MAS ONDE/COMO ESSAS COORDENADAS SÃO ARMAZENADAS?

6sl6.pdb — Edited

Downloads — vi 6sl6.pdb — 80x24

HEADER	CELL CYCLE	18-AUG-19	6SL6
TITLE	P53 CHARGED CORE		
COMPND	MOL_ID: 1;		
COMPND	2 MOLECULE: CELLULAR TUMOR ANTIGEN P53;		
COMPND	3 CHAIN: A;		
COMPND	4 SYNONYM: ANTIGEN NY-CO-13, PHOSPHOPROTEIN P53, TUMOR SUPPRESSOR P53;		
COMPND	5 ENGINEERED: YES;		
COMPND	6 MUTATION: YES		
SOURCE	MOL_ID: 1;		
SOURCE	2 ORGANISM_SCIENTIFIC: HOMO SAPIENS;		
SOURCE	3 ORGANISM_COMMON: HUMAN;		
SOURCE	4 ORGANISM_TAXID: 9606;		
SOURCE	5 GENE: TP53, P53;		
SOURCE	6 EXPRESSION_SYSTEM: ESCHERICHIA COLI;		
SOURCE	7 EXPRESSION_SYSTEM_TAXID: 562		
KEYWDS	TUMOR SUPPRESSOR, AGGREGATION PRONE, MUTANT P53, CELL CYCLE		
EXPDTA	X-RAY DIFFRACTION		
AUTHOR	R.GALLARDO, T.LANGENBERG, J.SCHYMKOWITZ, F.ROUSSEAU, C.ULENS		
REVDAT	3 24-JAN-24 6SL6 1 REMARK		
REVDAT	2 23-SEP-20 6SL6 1 JRNL		
REVDAT	1 11-MAR-20 6SL6 0		
JRNL	AUTH T.LANGENBERG, R.VAN DER KANT, N.LOUROS, E.MICHELS,		
JRNL	AUTH 2 R.DURAN-ROMANA, B.HOUBEN, R.CASSIO, H.WILKINSON, T.GARCIA,		
JRNL	AUTH 3 C.ULENS, J.VAN DURME, F.ROUSSEAU, J.SCHYMKOWITZ		
JRNL	TITL THERMODYNAMIC AND EVOLUTIONARY COUPLING BETWEEN THE NATIVE		
JRNL	TITL 2 AND AMYLOID STATE OF GLOBULAR PROTEINS.		
JRNL	REF CELL REP V. 31 07512 2020		
JRNL	REFN ESSN 2211-1247		
JRNL	PMID 32294448		
JRNL	DOI 10.1016/J.CELREP.2020.03.076		
REMARK	2 RESOLUTION. 1.67 ANGSTROMS.		
REMARK	3		
REMARK	3 REFINEMENT.		
REMARK	3 PROGRAM : PHENIX 1.15.2-3472		
REMARK	3 AUTHORS : PAUL ADAMS, PAVEL AFONINE, VINCENT CHEN, IAN		
REMARK	3 DAVIS, KRESHNA GOPAL, RALF GROSSE-KUNSTLEVE,		
REMARK	3 LI-WEI HUNG, ROBERT IMMORINO, TOM IOERGER,		
REMARK	3 AILIE MCCOY, ERIK MCKEE, NIGEL MORIARTY,		
REMARK	3 REETAL PAI, RANDY READ, JANE RICHARDSON,		
REMARK	3 DAVID RICHARDSON, TOD ROMO, JIM SACCHETTINI,		
REMARK	3 NICHOLAS SAUTER, JACOB SMITH, LAURENT		
REMARK	3 STORONI, TOM TERWILLIGER, PETER ZWART		
REMARK	3 REFINEMENT TARGET : NULL		
REMARK	3		
REMARK	3 DATA USED IN REFINEMENT.		
REMARK	3 RESOLUTION RANGE HIGH (ANGSTROMS) : 1.67		
REMARK	3 RESOLUTION RANGE LOW (ANGSTROMS) : 42.48		
REMARK	3 MIN(FOBS/SIGMA_FOBS) : 1.930		
REMARK	3 COMPLETENESS FOR RANGE (%) : 97.9		
REMARK	3 NUMBER OF REFLECTIONS : 28523		
REMARK	3		
REMARK	3 FIT TO DATA USED IN REFINEMENT.		
REMARK	3 R VALUE (WORKING + TEST SET) : 0.164		
REMARK	3 R VALUE (WORKING SET) : 0.163		
REMARK	3 FREE R VALUE : 0.180		
REMARK	3 FREE R VALUE TEST SET SIZE (%) : 4.970		
REMARK	3 FREE R VALUE TEST SET COUNT : 1419		
REMARK	3		
REMARK	3 FIT TO DATA USED IN REFINEMENT (IN BINS).		
REMARK	3 BIN_RESOLUTION RANGE COMPL. NWORK NFREE RWORK RFREE		
REMARK	3 1 42.4810 - 3.5939 0.99 2883 159 0.1599 0.1641		
REMARK	3 2 3.5939 - 2.8528 0.99 2788 139 0.1461 0.1656		
REMARK	3 3 2.8528 - 2.4922 0.99 2753 142 0.1597 0.1826		
REMARK	3 4 2.4922 - 2.2644 0.99 2724 135 0.1598 0.1710		
REMARK	3 5 2.2644 - 2.1821 0.99 2714 146 0.1486 0.1993		
REMARK	3 6 2.1821 - 1.9782 0.98 2691 151 0.1660 0.1693		
REMARK	3 7 1.9782 - 1.8791 0.98 2675 146 0.1688 0.1913		
REMARK	3 8 1.8791 - 1.7973 0.98 2681 136 0.1697 0.2028		

CIF OU PDB? O HÁ DE NOVO? QUEM ESTÁ USANDO



230,444 Structures from the PDB



1,068,577 Computed Structure Models (CSM)

Enter search term(s), Entry ID(s), Ligand ID or sequence



[Advanced Search](#) | [Browse Annotations](#)



News

2025

2024

2023

2022

2021

2020

2019

2018

2017

2016

2015

2014

2013

2012

2011

2010

2009

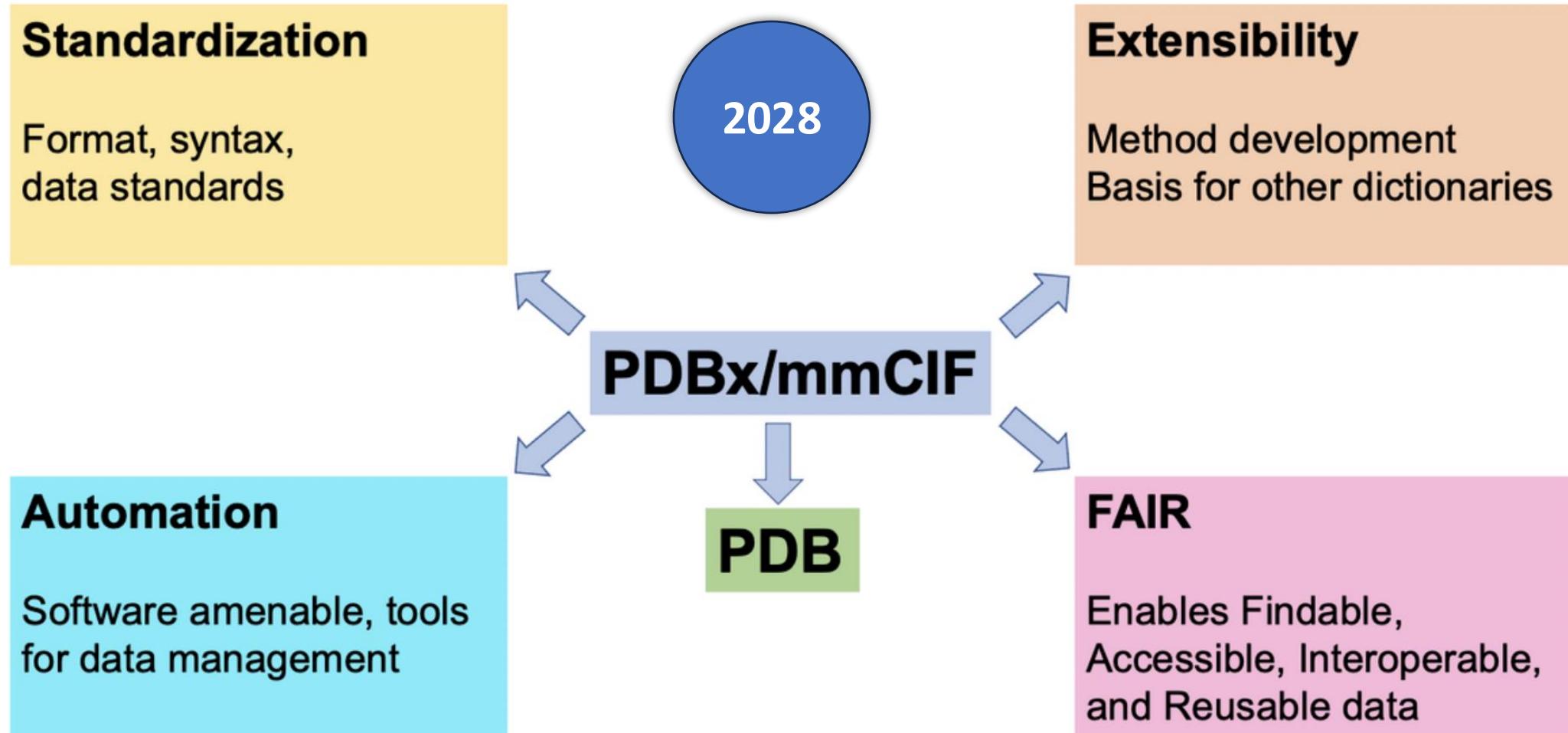
2008

[◀ Previous Article](#)

[Next Article ▶](#)

Announcing the New PDBx/mmCIF User Guide

CIF OU PDB? O HÁ DE NOVO? QUEM ESTÁ USANDO



CIF OU PDB? O HÁ DE NOVO? QUEM ESTÁ USANDO

The PDB format

```
JRNL      TITL  COMPARISON OF THE THREE-DIMENSIONAL STRUCTURES OF
JRNL      TITL 2 RECOMBINANT HUMAN H AND HORSE L FERRITINS AT HIGH
JRNL      TITL 3 RESOLUTION
JRNL      REF   J.MOL.BIOL.          V. 268    424 1997
JRNL      REFN  ASTM JMOBAK  UK ISSN 0022-2836          0070
```

The mmcif format

mmCIF: macromolecular
Crystallographic
Information
File

This is an extension of the Crystallographic Information File (CIF) data representation (used for describing small molecule structures) to describe macromolecules.

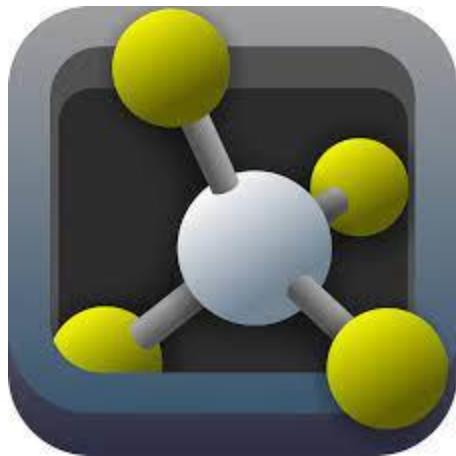
```
_citation.id                         primary
_citation.title

;Comparison of the three-dimensional structures
of recombinant human H and horse L ferritins at
high resolution.

;

_citation.journal_abbrev           J.Mol.Biol.
_citation.journal_volume          268
_citation.page_first               424
_citation.page_last                448
_citation.year                     1997
_citation.journal_id_ASTM         JMOBAK
_citation.country                  UK
_citation.journal_id_ISSN          0022-2836
_citation.journal_id_CSD           0070
_citation.book_publisher           ?
_citation.pdbx_database_id_PubMed  9159481
```

VAMOS CONHECER ALGUMAS PROTEÍNAS

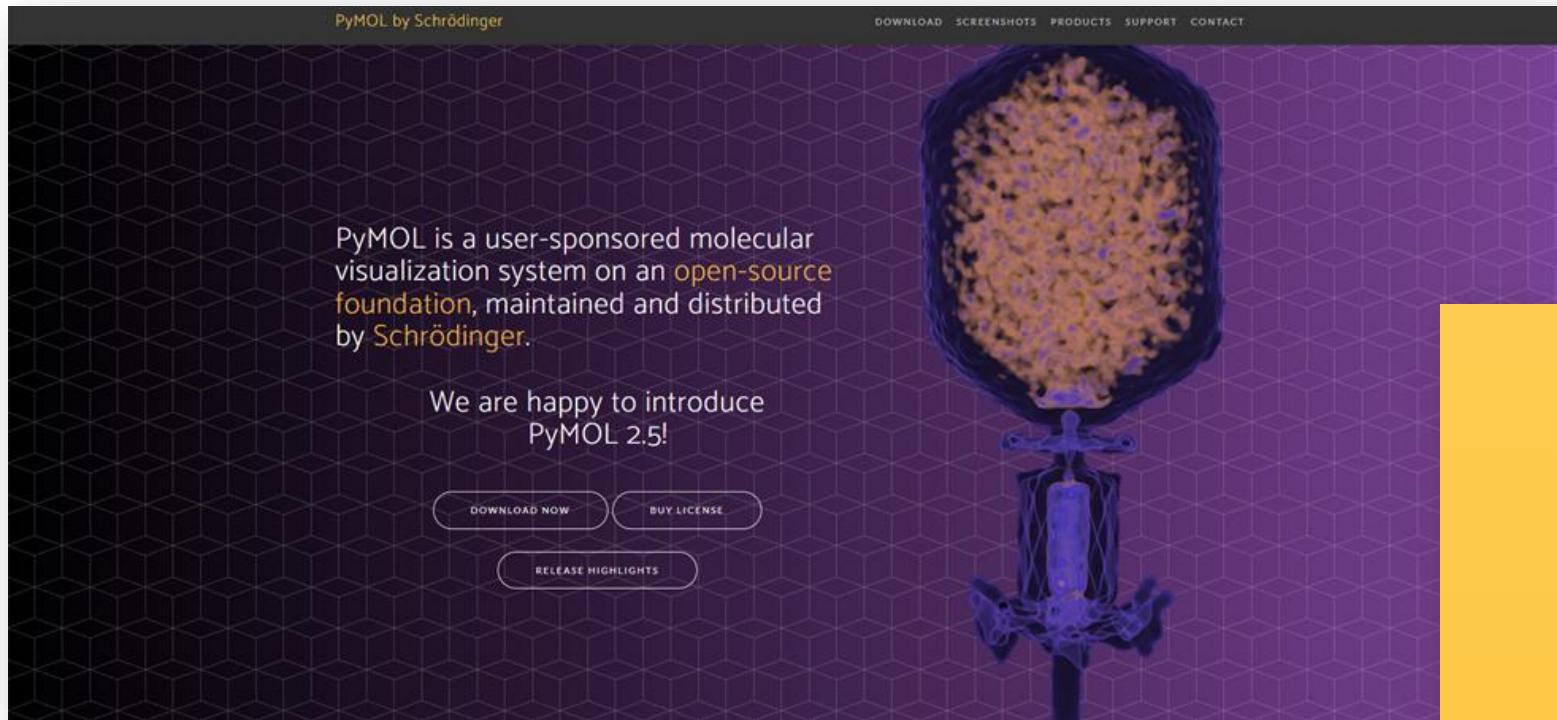


PyMOL

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

FERRAMENTA DE VISUALIZAÇÃO PARA ARQUIVOS **PDB/mmCIF/CIF**

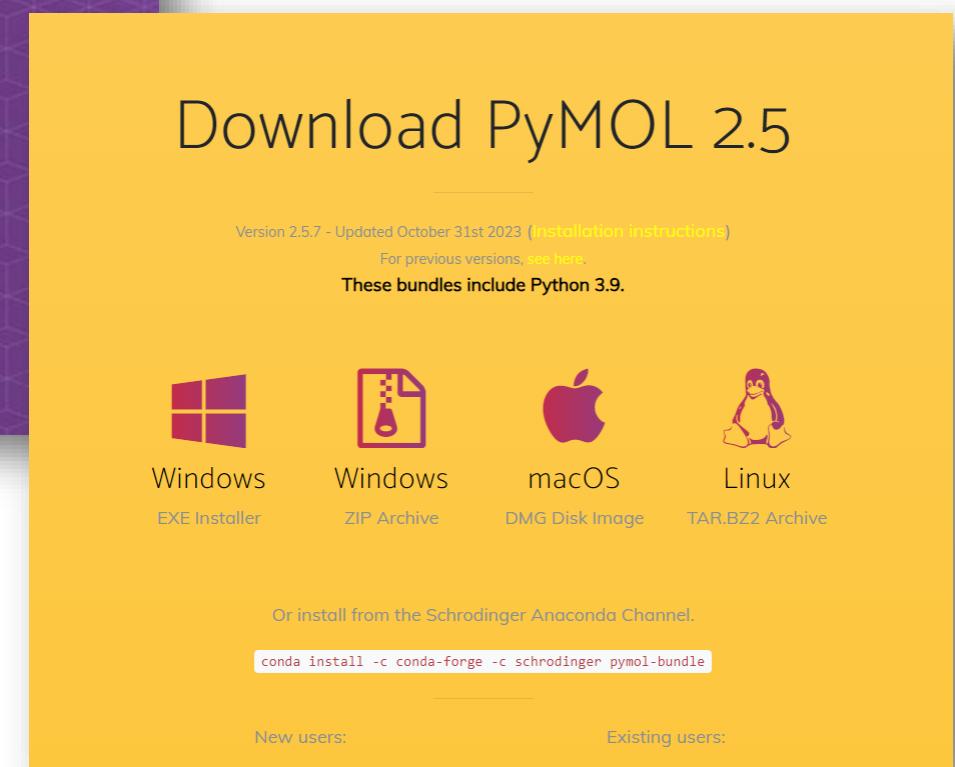
ATIVIDADE



The PyMOL website homepage features a dark purple background with a hexagonal grid pattern. At the top, the header "PyMOL by Schrödinger" is followed by a navigation bar with links: DOWNLOAD, SCREENSHOTS, PRODUCTS, SUPPORT, and CONTACT. Below the header, a large image of a molecular structure is shown, composed of yellow spheres and grey lines. To the left of the image, text reads: "PyMOL is a user-sponsored molecular visualization system on an open-source foundation, maintained and distributed by Schrödinger." Below this, a message says: "We are happy to introduce PyMOL 2.5!" with three buttons: "DOWNLOAD NOW", "BUY LICENSE", and "RELEASE HIGHLIGHTS".



PyMOL



The "Download PyMOL 2.5" page has a yellow background. It starts with the text "Version 2.5.7 - Updated October 31st 2023 ([Installation instruction](#))". Below this, it says "For previous versions, see [here](#). These bundles include Python 3.9." It then shows download links for Windows, macOS, and Linux, each with a corresponding icon: a Windows logo, an Apple logo, and a Linux logo (Tux). Under each icon, there are two links: "Windows EXE Installer" and "Windows ZIP Archive" for Windows; "macOS DMG Disk Image" and "macOS TAR.BZ2 Archive" for macOS; and "Linux TAR.BZ2 Archive" for Linux. At the bottom, it says "Or install from the Schrodinger Anaconda Channel." with a code snippet: `conda install -c conda-forge -c schrodinger pymol-bundle`. Finally, it has sections for "New users:" and "Existing users:".

ATIVIDADE

Windows
EXE Installer

Windows
ZIP Archive

macOS
DMG Disk Image

Linux
tar.gz

Or install from the Schrodinger Anaconda

```
conda install -c conda-forge -c schrodinger
```

New users:

[BUY LICENSE](#)



PyMOL

[DOWNLOAD](#) [SCREENSHOTS](#) [PRODUCTS](#) [SUPPORT](#) [CONTACT](#)

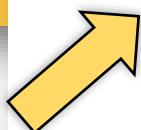
PyMOL by Schrödinger
Buy PyMOL

Order PyMOL and AxPyMOL Online

Please select your intended use from the list below to receive access within one business day:

- **Commercial**
for corporate organizations
- **Government**
for government institutions
- **Academic/Non-profit**
for academic and non-profit researchers/educators
- **Student/Teacher**

Educational subscriptions to PyMOL are available at no cost to full-time students and educators teaching full-time students. However, there is not an educational subscription for AxPyMOL at this time. Under this agreement, Schrödinger has no obligation to furnish updates, fix bugs, provide documentation, and meet other support needs. Please recognize that long-term reliance upon PyMOL for educational purposes should merit a subscription purchase in order to secure essential services and baseline accountability.



ATIVIDADE

(a licença chega por e-mail, use preferencialmente o institucional)

Registration For Educational-Use-Only PyMOL Builds

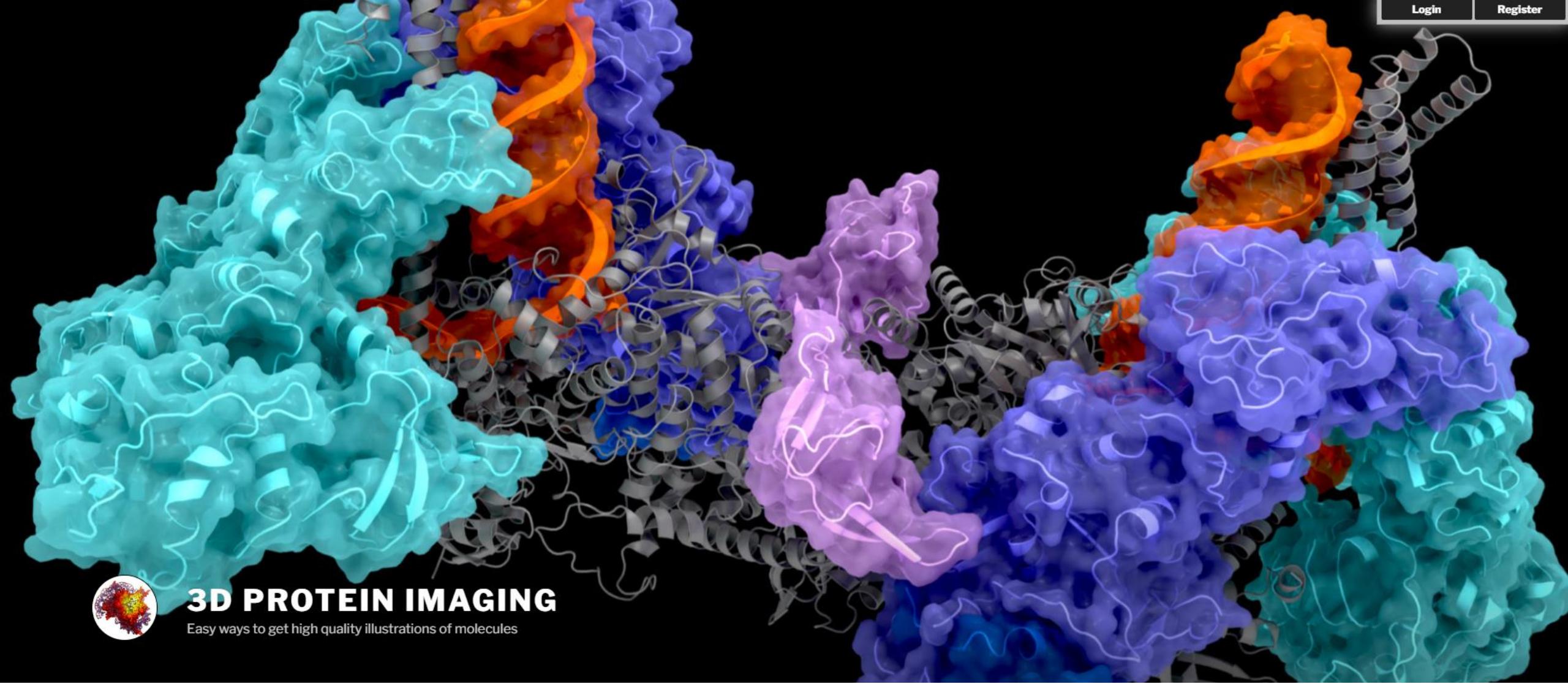
Schrödinger offers **Educational-use-only** PyMOL builds available at no cost to **teachers and high school and college students** (including online courses, homeschooling, etc.) for classroom instruction, homework assignments, and to provide a means for creating high quality figures. Please note that it is not provided for the purposes of academic research or publication.

-> [FAQ \(Frequently Asked Questions\)](#)

The Educational-use-only PyMOL builds are provided "AS IS" with no obligation to grant download access, fix bugs, furnish updates, provide documentation, or meet any other need related to the educational-use PyMOL builds.

If you intend to use PyMOL products for academic research or publication, please purchase an Academic PyMOL subscription, which includes access to technical support, screencasts, and additional resources. See <http://pymol.org/academic>.

I am a:	<input type="button" value="▼"/>
Your First Name:	<input type="text"/>
Your Last Name:	<input type="text"/>
Your Email Address:	<input type="text"/>
Your Telephone Number:	<input type="text"/>
Institution:	<input type="text"/>
Comments (optional):	<input type="text"/>
<input type="button" value="Continue"/>	

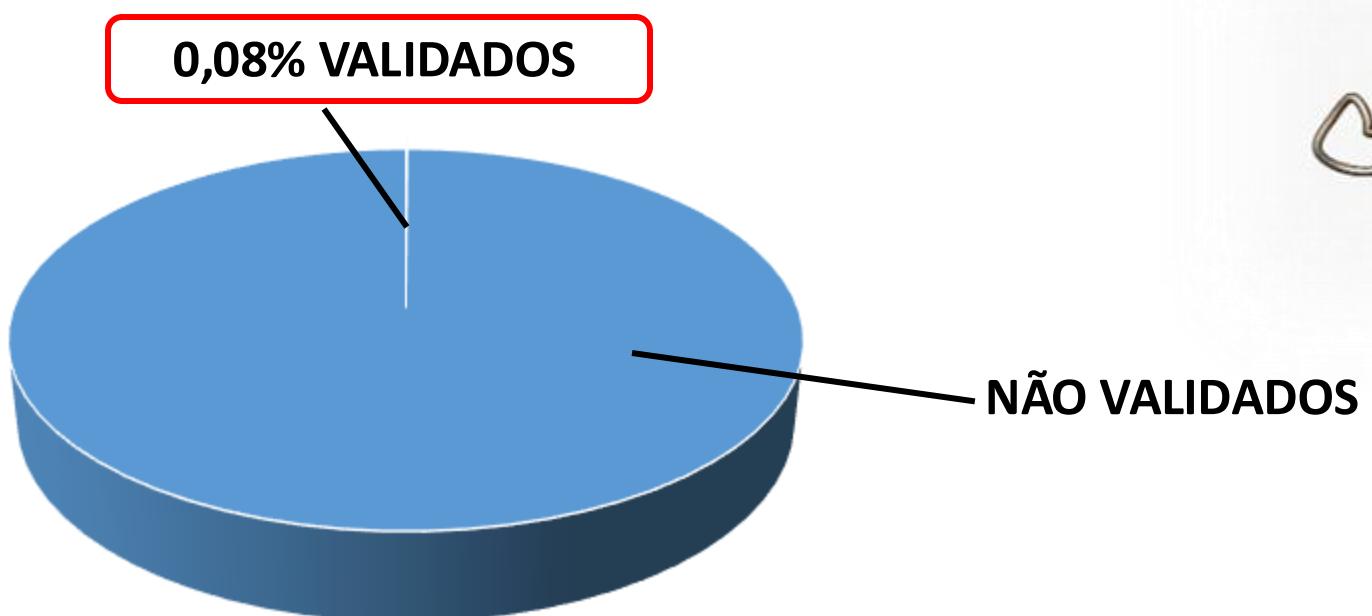


“Plano B” de visualização molecular – Servidor online

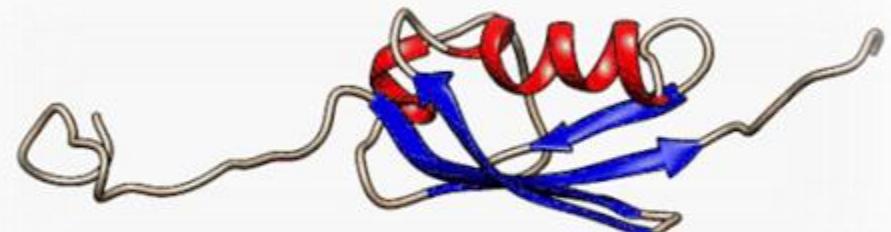
<https://3dproteinimaging.com/protein-imager/>



+285 milhões proteínas
depositadas



+217 mil proteínas com
estruturas validadas





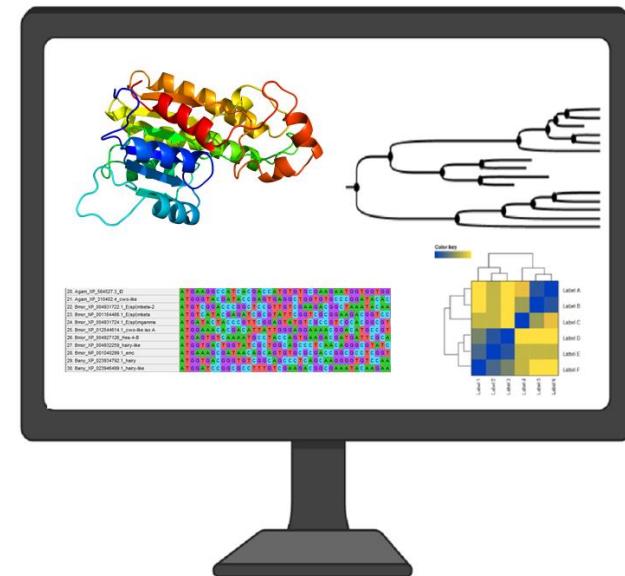
+285 milhões proteínas
depositadas



+217 mil proteínas com
estruturas validadas

Validar experimentalmente: Caro + Demorado

Modelagem computacional tornou-se acessível



MODELAGEM DE PROTEÍNAS

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
S	A	C	K	P	E	C	P	V	N	I	Q	G	S	-	-	Y	A	I	D	A	S	C	I	D	D	G	S			
C	-	-	I	A	C	G	A	C	K	P	E	C	P	V	N	I	Q	G	S	-	-	I	Y	A	I	D	A	D	S	



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

MODELAGEM COMPARATIVA ("MODELAGEM POR HOMOLOGIA")

 BIOZENTRUM
University of Basel
The Center for Molecular Life Sciences

SWISS-MODEL

Modelling Repository Tools Documentation Log in Create Account

SWISS-MODEL will be unavailable on 8th January, between 11:00 and 12:00 CET. Long running jobs started before 11:00 CET will be lost.

SWISS-MODEL

is a fully automated protein structure homology-modelling server, accessible via the [ExPasy web server](#), or from the program DeepView (Swiss Pdb-Viewer).

The purpose of this server is to make protein modelling accessible to all life science researchers worldwide.

[Start Modelling](#)

Repository

Every week we model all the sequences for thirteen core species based on the latest UniProtKB proteome. Is your protein already modelled and up to date in [SWISS-MODEL Repository](#)?

Search SWISS-MODEL Repository



<https://swissmodel.expasy.org/>

ATIVIDADE

?

>desconhecido

KDCKRESNTFPGICITKPPCRKACIREKFTDHCSKILRRCLCTKPC

MODELAGEM COMPARATIVA ("MODELAGEM POR HOMOLOGIA")

 BIOZENTRUM
University of Basel
The Center for Molecular Life Sciences

SWISS-MODEL

Modelling Repository Tools Documentation Log in Create Account

SWISS-MODEL will be unavailable on 8th January, between 11:00 and 12:00 CET. Long running jobs started before 11:00 CET will be lost.

SWISS-MODEL

is a fully automated protein structure homology-modelling server, accessible via the [ExPasy web server](#), or from the program DeepView (Swiss Pdb-Viewer).

The purpose of this server is to make protein modelling accessible to all life science researchers worldwide.

[Start Modelling](#)

Repository

Every week we model all the sequences for thirteen core species based on the latest UniProtKB proteome. Is your protein already modelled and up to date in [SWISS-MODEL Repository](#)?

Search SWISS-MODEL Repository



<https://swissmodel.expasy.org/>

MODELAGEM COMPARATIVA ("MODELAGEM POR HOMOLOGIA")

SWISS-MODEL

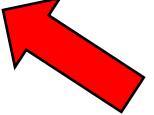
Modelling Repository Tools Documentation Log in Create Account

SWISS-MODEL will be unavailable on 8th January, between 11:00 and 12:00 CET. Long running jobs started before 11:00 CET will be lost.

Start a New Modelling Project 

Target Sequence(s):
(Format must be FASTA,
Clustal,
plain string, or a valid
UniProtKB AC)

Paste your target sequence(s) or UniProtKB AC here

 **1º Colar sequência de interesse**

+ Upload Target Sequence File... Validate

Project Title: Untitled Project

Email: Optional

Supported Inputs 

Sequence(s)
Target-Template Alignment
User Template
DeepView Project

Search For Templates Build Model

By using the SWISS-MODEL server, you agree to comply with the following [terms of use](#) and to cite the corresponding [articles](#).

 **2º Buscar estruturas de referência ("templates")**

MODELAGEM COMPARATIVA ("MODELAGEM POR HOMOLOGIA")

Summary Templates 22 Models Project Data ▾

Template Results ⓘ

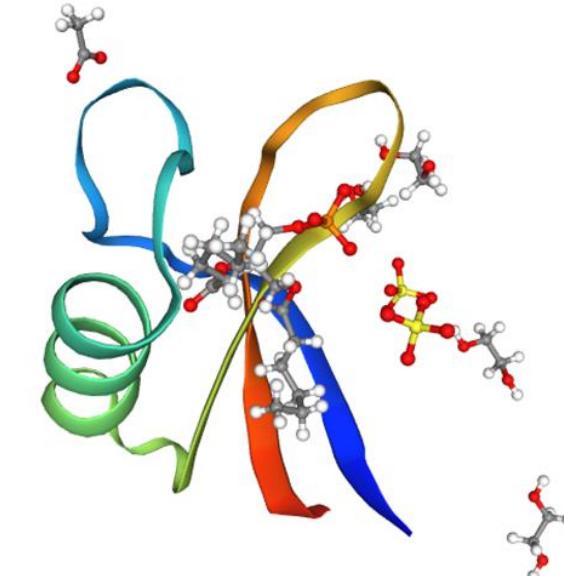
GMQE = Estimativa de qualidade (0 - 1)

	Coverage	GMQE	QSQE	Identity	Method	Oligo State	Ligands
<input type="checkbox"/> C0HK49.1.A Defensin NsD7 AlphaFold DB model of DEF_NICSU (gene: DEF_NICSU, organism: Nicotiana suaveolens (Australian tobacco))	0.98	-	100.00	AlphaFold v2	monomer ✓	None	
<input checked="" type="checkbox"/> 5kk4.1.C NsD7 Crystal Structure of the Plant Defensin NsD7 bound to Phosphatidic Acid	0.90	0.43	100.00	X-ray, 1.7Å	homo-hexamer ✓	6 x 44E ⚡	
<input type="checkbox"/> 5kk4.1.F NsD7 Crystal Structure of the Plant Defensin NsD7 bound to Phosphatidic Acid	0.89	0.43	100.00	X-ray, 1.7Å	homo-hexamer ✓	6 x 44E ⚡	
<input type="checkbox"/> 5kk4.1.E NsD7 Crystal Structure of the Plant Defensin NsD7 bound to Phosphatidic Acid	0.89	0.43	100.00	X-ray, 1.7Å	homo-hexamer ✓	6 x 44E ⚡	
<input type="checkbox"/> 4cqk.6.B FLOWER-SPECIFIC DEFENSIN Crystal structure of ligand-bound NaD1	0.88	0.45	91.49	X-ray, 1.6Å	homo-dimer ✓	2 x P1O ⚡	
<input type="checkbox"/> 4aaZ.1.A FLOWER-SPECIFIC DEFENSIN X-ray structure of Nicotiana alata Defensin 1 NaD1	0.88	-	91.49	X-ray, 1.4Å	monomer ✓	None	
<input type="checkbox"/> 5kk4.1.A NsD7	0.88	-	91.49	X-ray, 1.4Å	monomer ✓	None	

Build Models 1

Clear Selection

Método = De que forma foi feito o template



MODELAGEM COMPARATIVA ("MODELAGEM POR HOMOLOGIA")

All Projects

Untitled Project Created: today at 02:51

Summary Templates 22 Models 2 Project Data ▾

Model Results Order by: GMQE

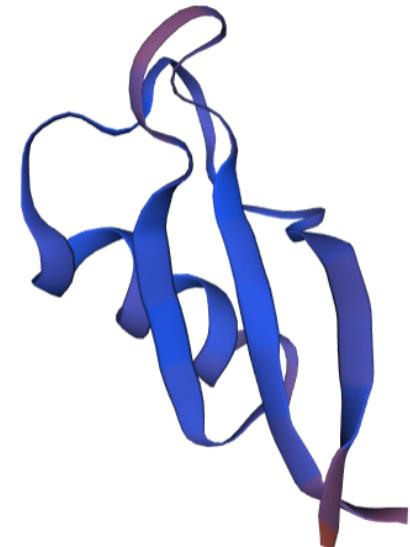
1

Baixar modelo em formato “PDB”

Model 02 ▾ Oligo-State Monomer (requested by user)
Structure Assessment GMQE 0.91 QMEANDisCo Global: 0.84 ± 0.12 Compare

QMEANDisCo Local
QMEAN Z-Scores
Template 5kk4.1.C_NsD7 Seq Identity 100.00% Coverage
Crystal Structure of the Plant Defensin NsD7 bound to Phosphatidic Acid

Model-Template Alignment



AGORA É A SUA VEZ!

 **BIOZENTRUM**
University of Basel
The Center for Molecular Life Sciences

SWISS-MODEL

SWISS-MODEL will be unavailable on 8th January, between 11:00 and 12:00 CET. Long running jobs started before 11:00 CET will be lost.

SWISS-MODEL
is a fully automated protein structure homology-modelling server, accessible via the [ExPasy web server](#), or from the program DeepView (Swiss Pdb-Viewer).
The purpose of this server is to make protein modelling accessible to all life science researchers worldwide.

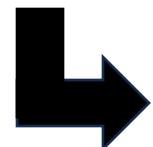
[Start Modelling](#)

Repository
Every week we model all the sequences for thirteen core species based on the latest UniProtKB proteome. Is your protein already modelled and up to date in [SWISS-MODEL Repository](#)?

Search SWISS-MODEL Repository



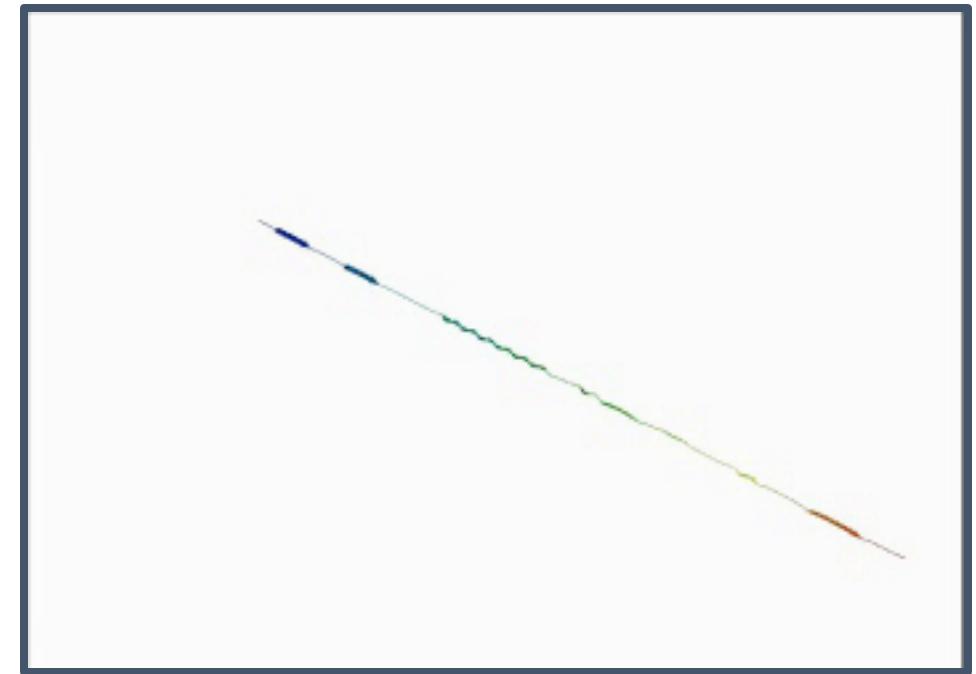
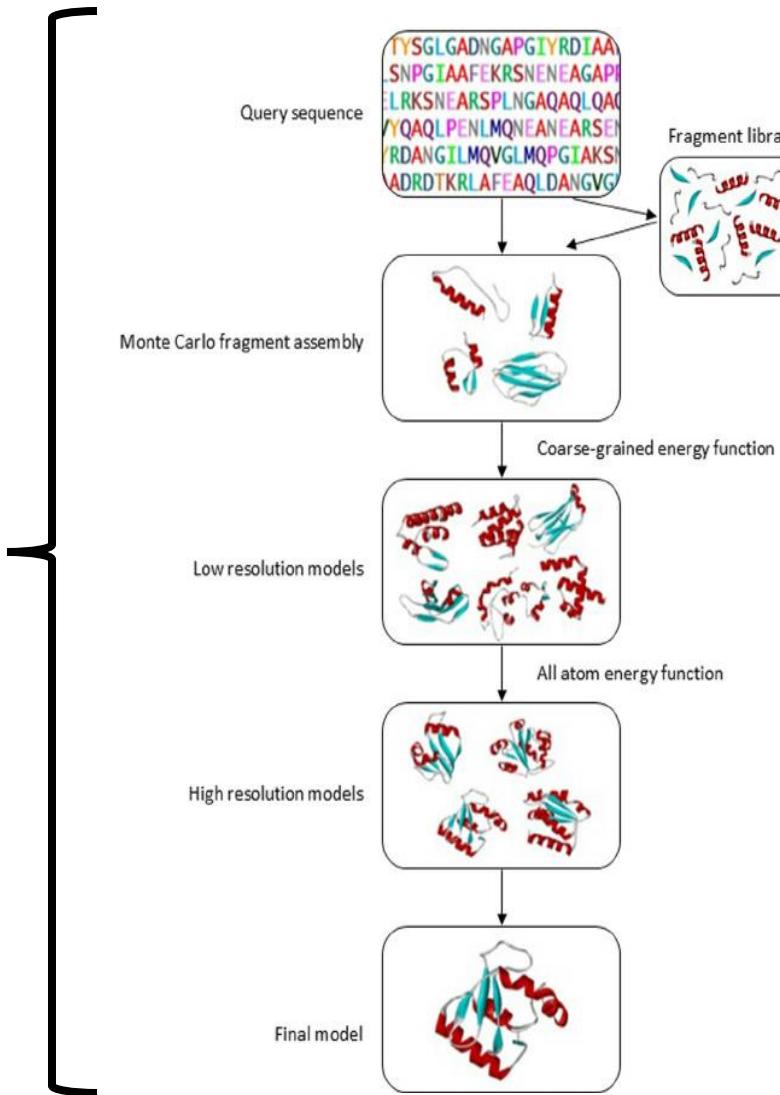
AEARTCESQTHFKFGTCLSETNCANVCKTEGFTGGDCRGLRRRCFCTRHC



?

Ache um modelo 3D
pelo SWISS-MODEL

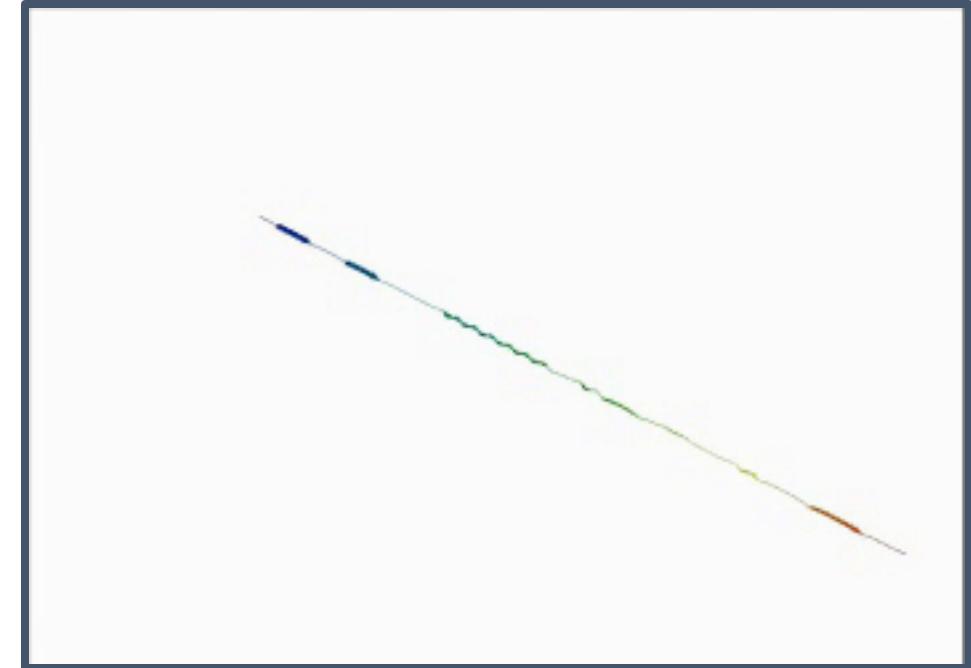
MODELAGEM AB INITIO ("DO ZERO")



MODELAGEM AB INITIO ("DO ZERO")



- **Alternativa** para proteínas completamente desconhecidas
- Alternativas para **caracterização de proteínas de novos patógenos** (ex: SARS-CoV-2)
- **Limitações:** Computacionalmente caro e baixo desempenho para proteínas > 200 AA



MODELAGEM AB INITIO ("DO ZERO")



DAVID BAKER
BAKER LAB

CRIADORES DO ROSETTA
E ROSETTAFOLD

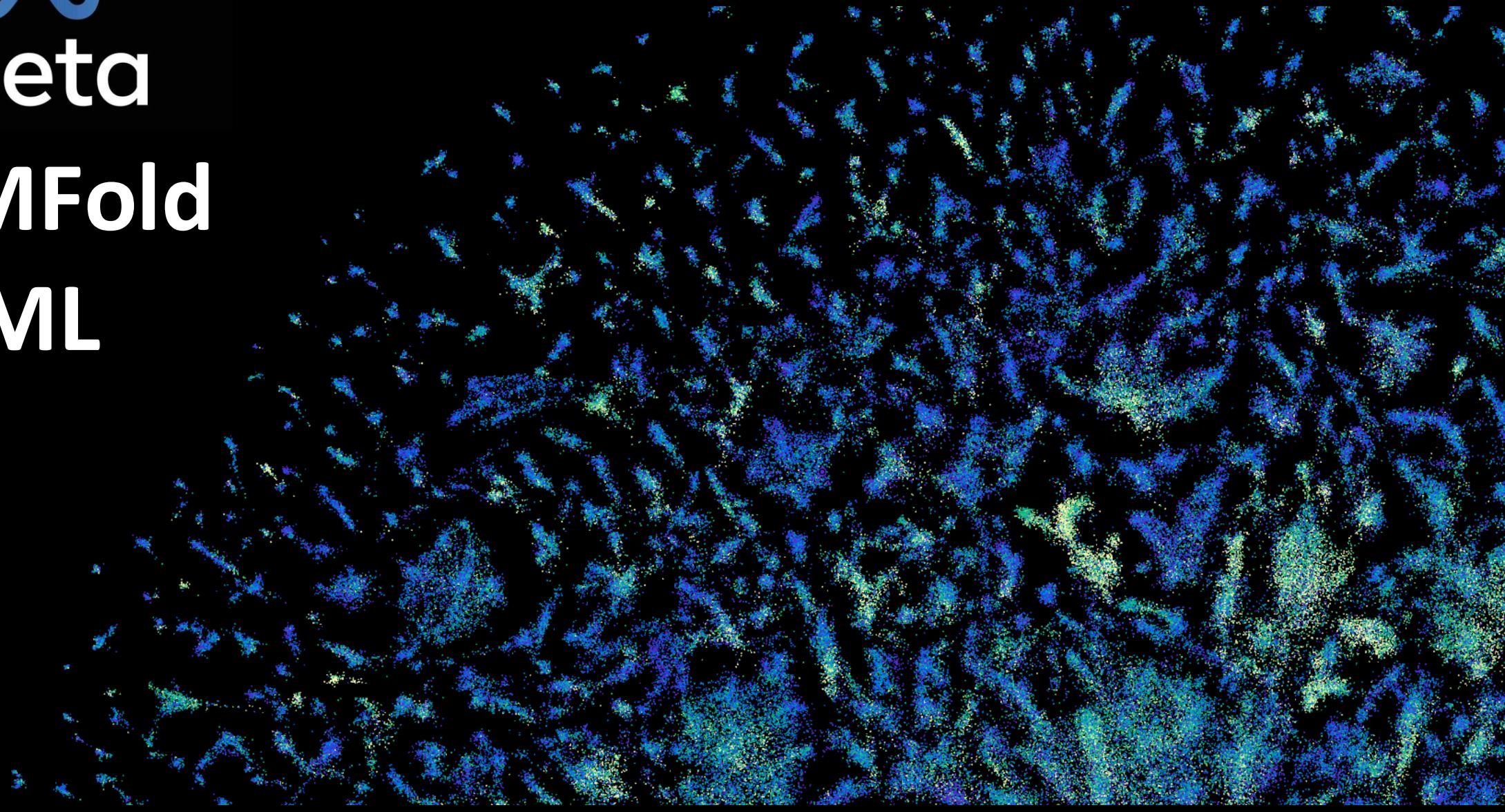


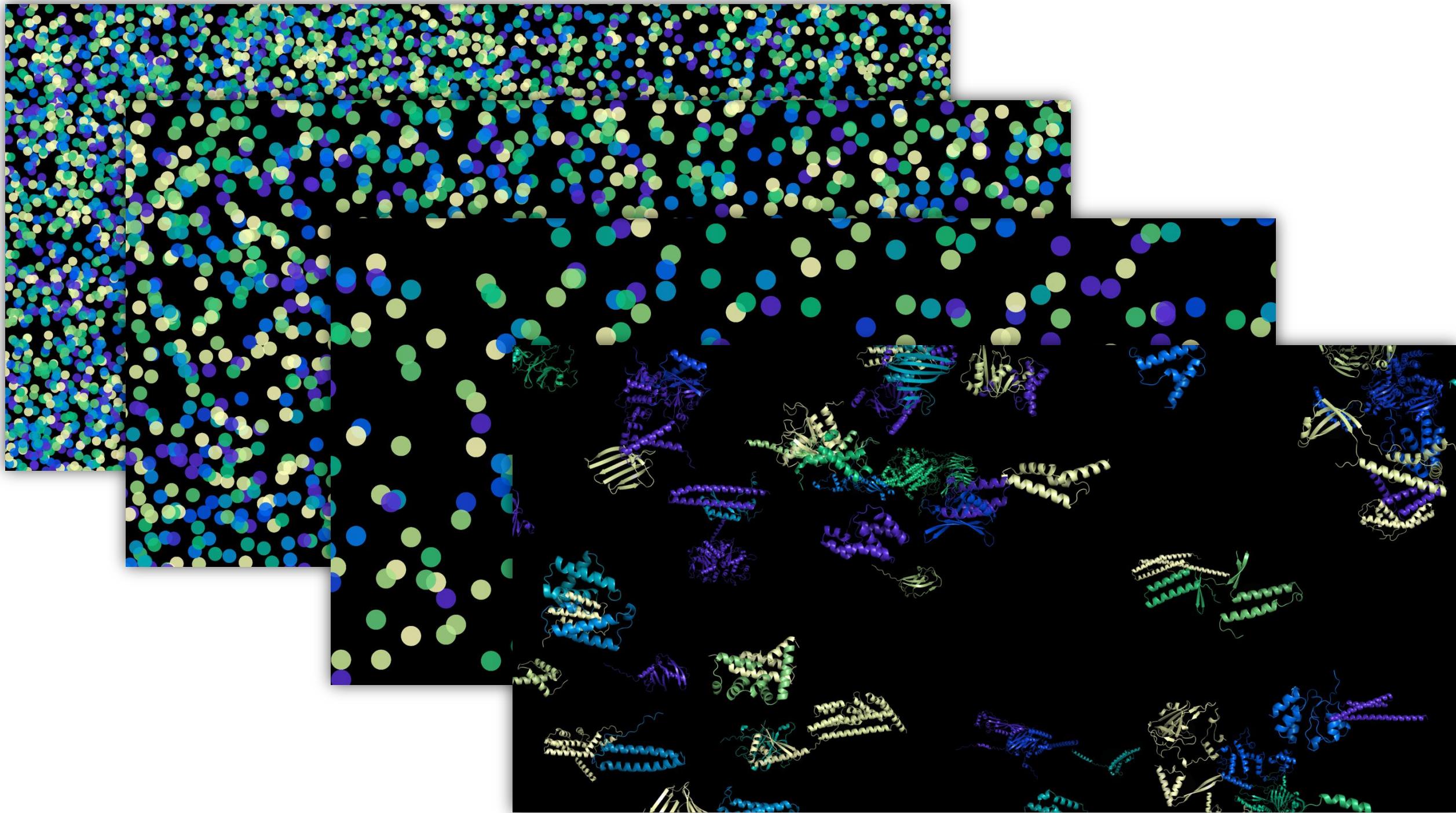
ESMFold

(PLATAFORMA META)



Meta
ESMFold
ML





ESMFold

<https://esmatlas.com/>

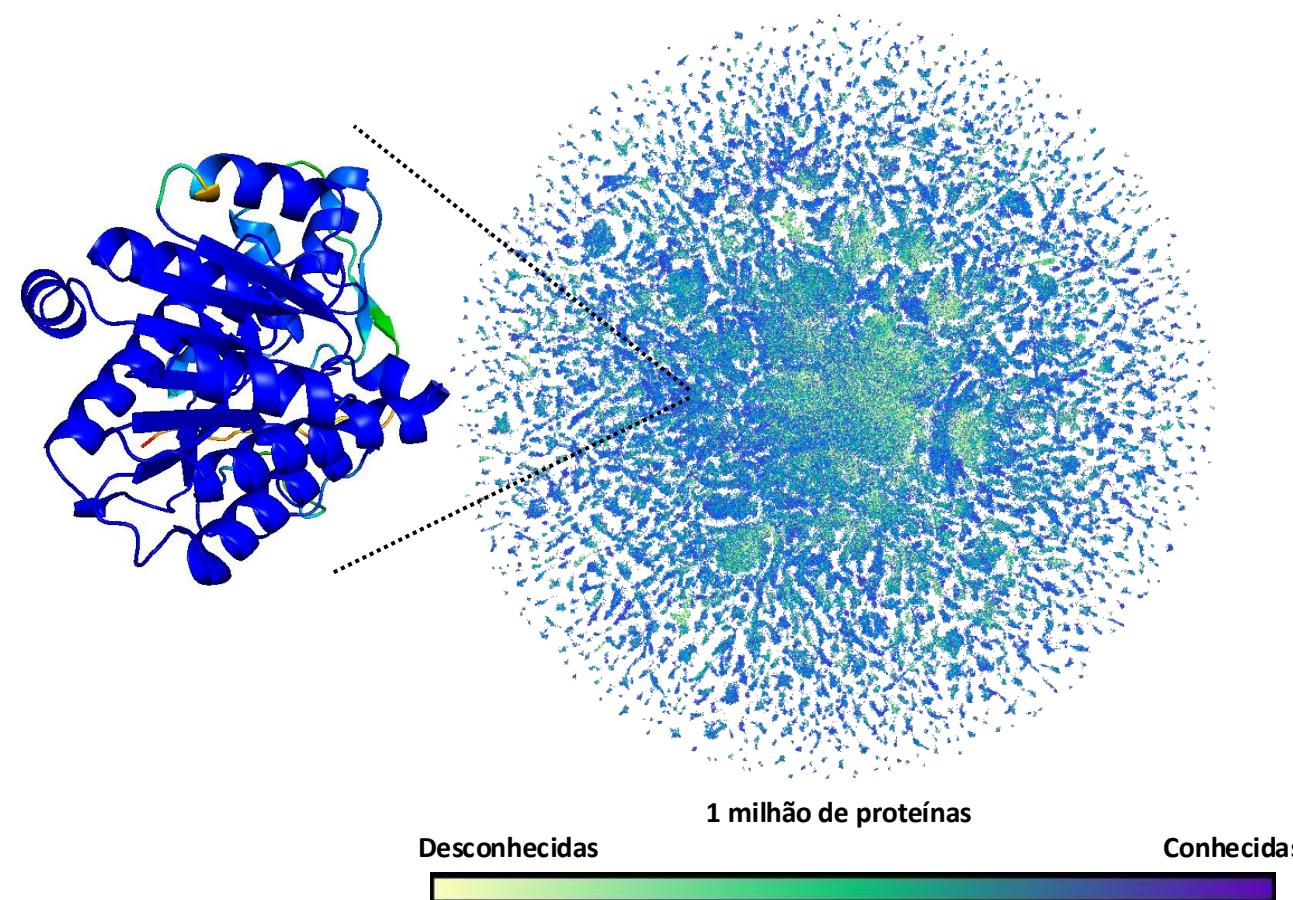
- “*Evolutionary Scale Modeling*”
- Modelo de inteligência que não utiliza alinhamentos múltiplos (\neq AlphaFold)
- Se fundamenta na arquitetura de linguagem de proteínas, baseado em dados de evolução
- Ele é treinado em grandes conjuntos de dados de sequências de proteínas, e DL para prever estruturas

Mais rápido que o AlphaFold, porém...

- Recomendado apenas para pequenas sequências
- Baixo score de confiabilidade para proteínas complexas



ESM Metagenomic Atlas



ESM Metagenomic Atlas

An open atlas of 772 million predicted
metagenomic protein structures

Explore →

Fold sequence ↗

Read blog post ↗

Read research paper ↗

Fold Sequence

Colar sequênciа de aminoácidos

Fold Sequence 

>



/

Try an example:

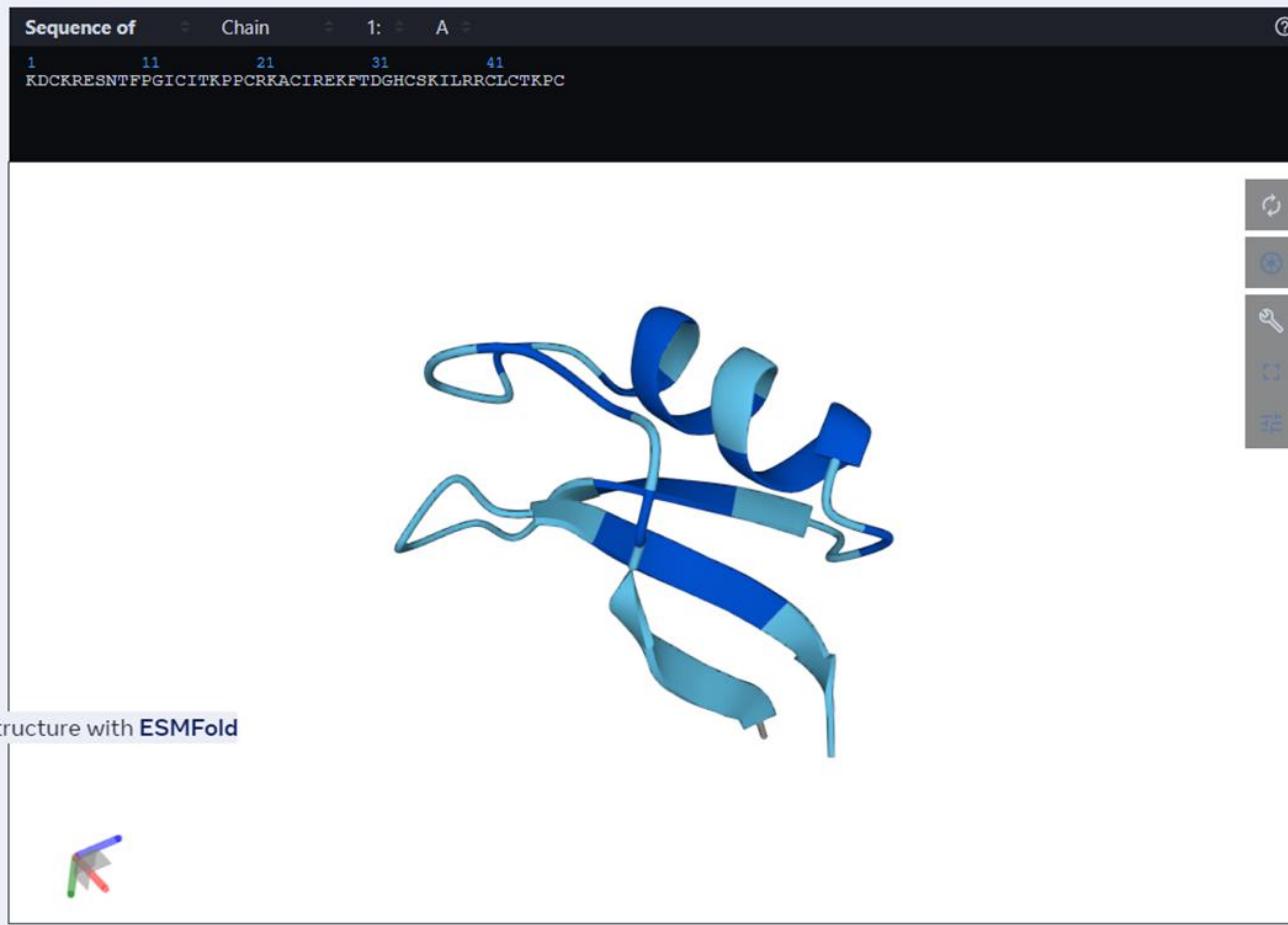
Plastic degradation protein - PETase

Antifreeze protein - 1EZG

AI-generated protein - 8CYK

7-bladed propeller fold - Neuraminidase

>unnamed



Baixar predição de estrutura

Download

[PDB file](#)

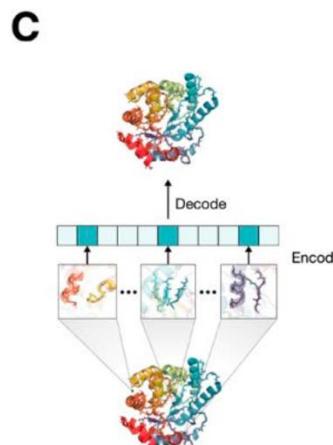
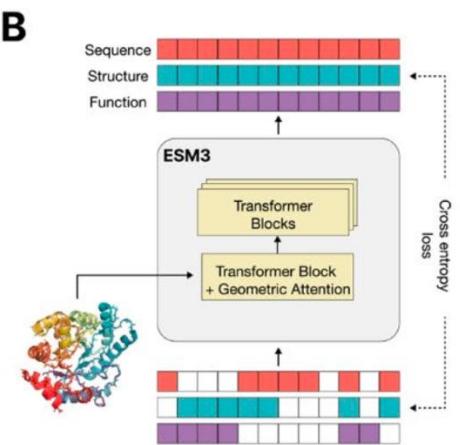
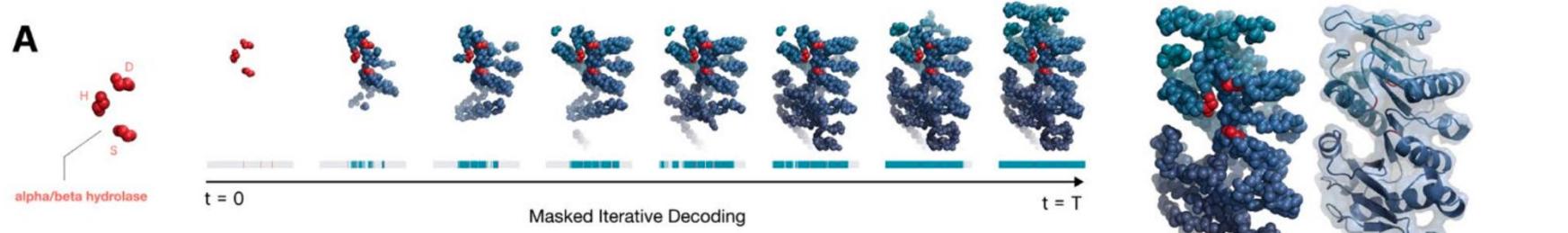
[Sequence](#)

The predicted structure is colored by local prediction confidence (pLDDT) per amino acid location. Blue indicates confident predictions ($p\text{LDDT} > 0.9$), while red indicates low confidence ($p\text{LDDT} < 0.5$).

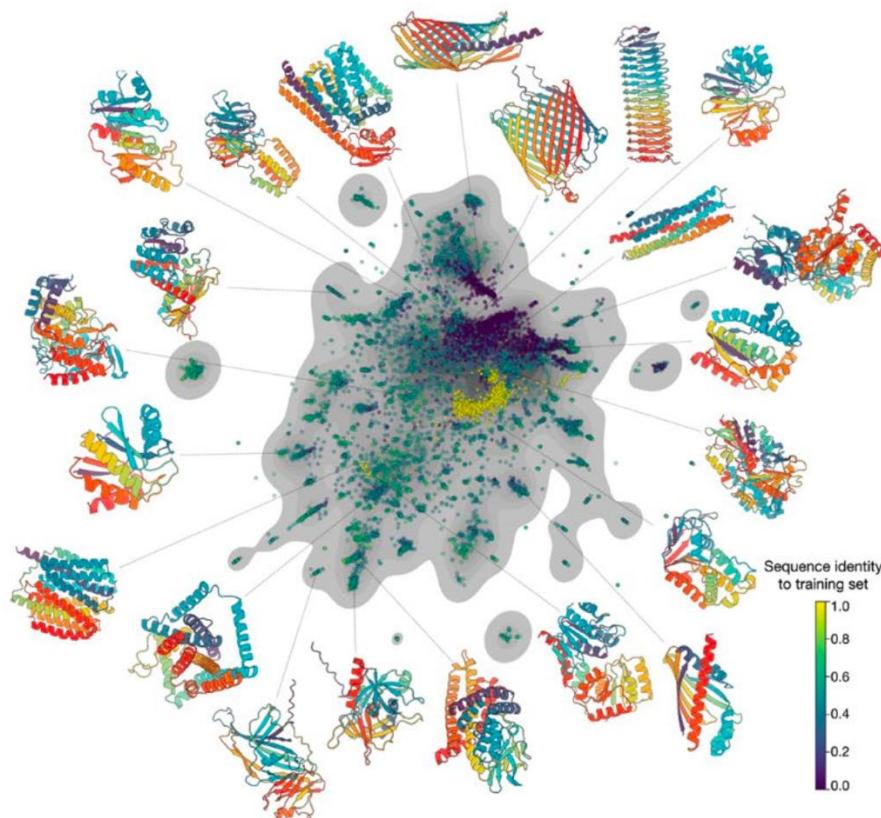
Meta AI

ESM Cambrian

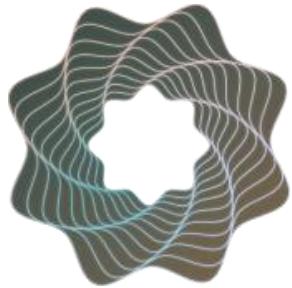
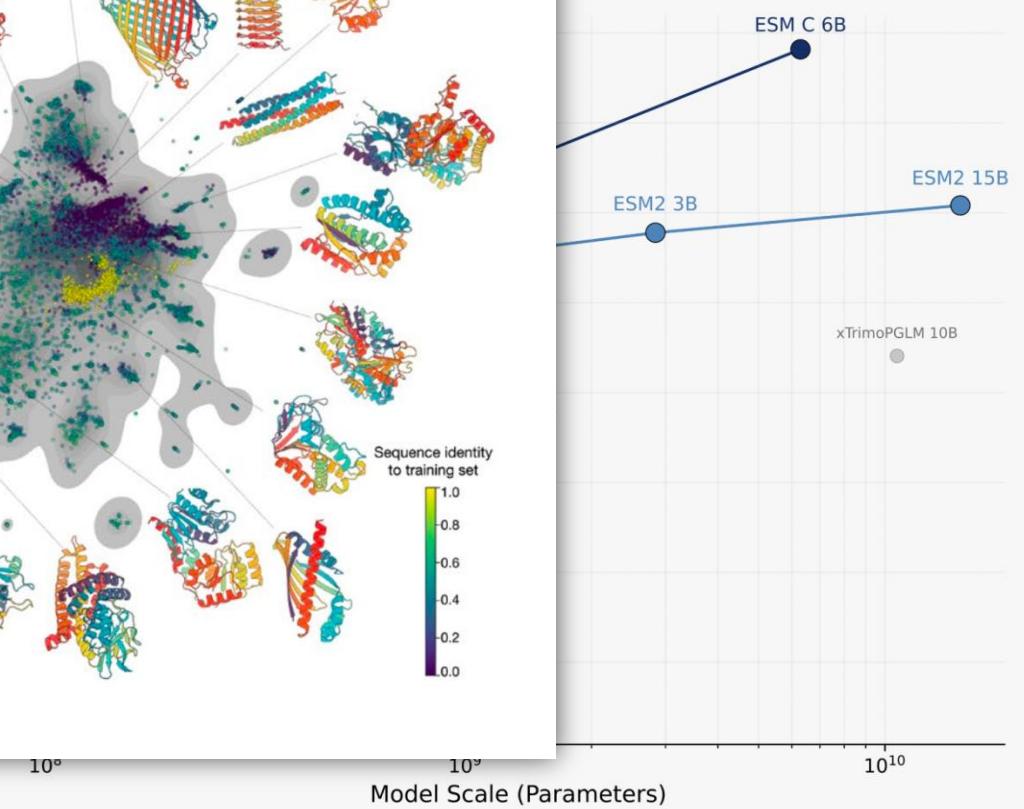
(Evolutionary Scale AI)



E



performance





APPS

🔍 Predict from Sequence

➤ Fold with ESM3

☒ Generate with ESM3

Learn

📊 ESM-Quickstart

☰ Tutorials

API

📊 Console

☰ Reference

HELP

ⓘ FAQ

✉ Contact Us

☰ Terms of Service

Welcome to EvolutionaryScale Forge.

We are excited to see what you'll build with the ESM family of models on Forge:

- **ESM3**, our frontier generative model for proteins
- **ESM C**, our state of the art representation learning model for property prediction applications

See the full list of models available [here](#). More to come soon!

Apps

Generate

Prompt ESM3 with sequence and structure constraints to generate novel proteins

Fold

Prediction of the protein structure directly from ESM3 output tokens.

Predict

Predict protein properties from a sequence

Developer Resources

Console

Get an API key and see credits and available models

Get started with ESM

[Quickstart guide](#)

Tutorials

[Learn how to use ESM](#)



HOLOCENO (2025): 14 (C)

CAMBRIANO (+/- 500 milhões de anos atrás): 20-25 (C) -> Não havia gelo na Terra

APPS

Q Predict from Sequence

> Fold with

Generate

Learn

ESM-Qui

Tutorials

API

Console

References

HELP

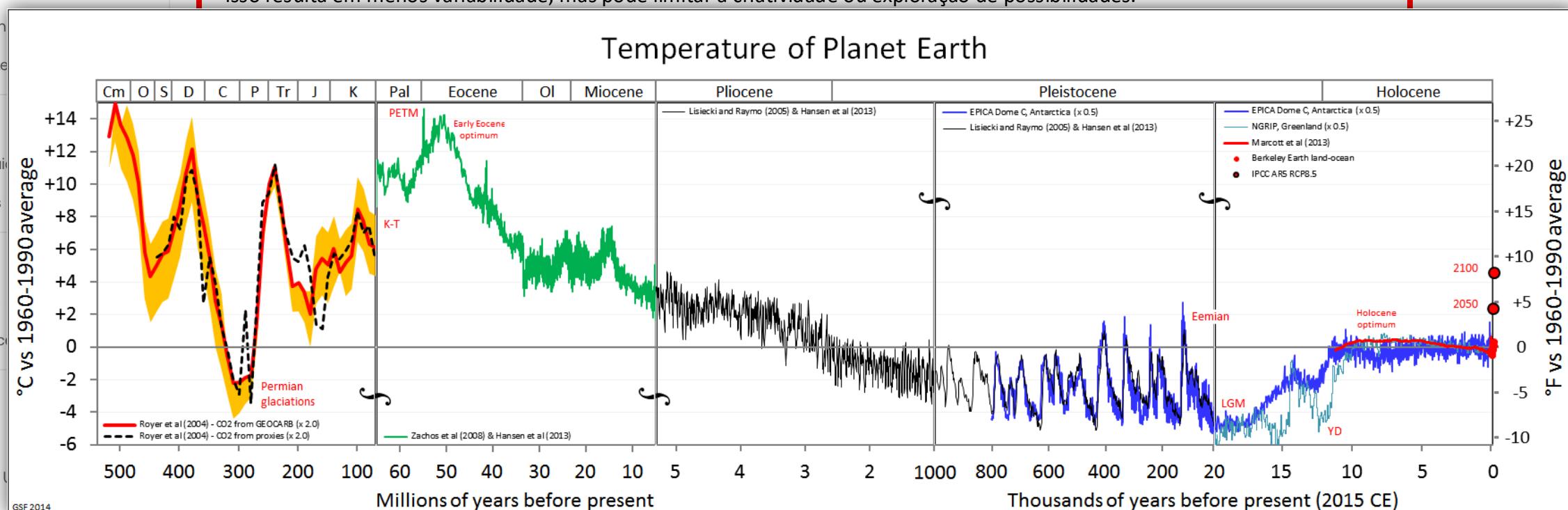
FAQ

Contact U

Terms of Service

Baixa temperatura (próxima de 0):

- Faz com que o modelo gere sequências altamente confiáveis e previsíveis, escolhendo valores com maior probabilidade.
- Isso resulta em menos variabilidade, mas pode limitar a criatividade ou exploração de possibilidades.



```
protein = model.generate(protein, GenerationConfig(track="structure", num_steps=8))
protein.to_pdb("./generation.pdb")
# Then we can do a round trip design by inverse folding the sequence and recomputing the structure
protein.sequence = None
protein = model.generate(protein, GenerationConfig(track="sequence", num_steps=8))
protein.coordinates = None
protein = model.generate(protein, GenerationConfig(track="structure", num_steps=8))
protein.to_pdb("./round_tripped.pdb")
```



KDCKRESNTFPGICITKPPCRKACIREKFTDGHCSKILRRCLCTKPC

Re-run Predict

APPS

[Q Predict from Sequence](#)[▷ Fold with ESM3](#)[Generate with ESM3](#)

Learn

[ESM-Quickstart](#)[Tutorials](#)

API

[Console](#)[Reference](#)

HELP

[FAQ](#)[Contact Us](#)[Terms of Service](#)

APPS

[Q Predict from Sequence](#)[▷ Fold with ESM3](#)[Generate with ESM3](#)

Learn

[ESM-Quickstart](#)[Tutorials](#)

API

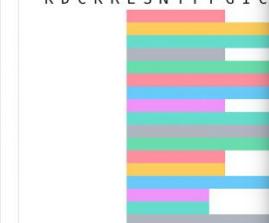
[Console](#)[Reference](#)

HELP

[FAQ](#)[Contact Us](#)[Terms of Service](#)● Very High Confidence ● High Confidence ● Low Confidence ● Very Low Confidence

Download PDB

Function

Pos: 0 Sequence 1 ▾
| 0 K D C K R E S N T F P G I C | 10

Secondary Structure

SASA

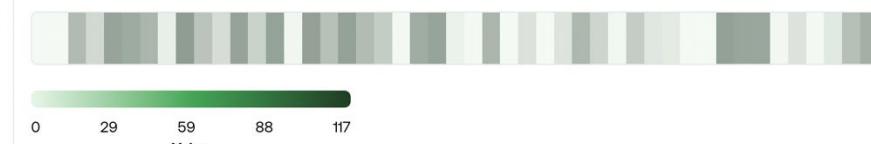
Output

pTM

Secondary Structure

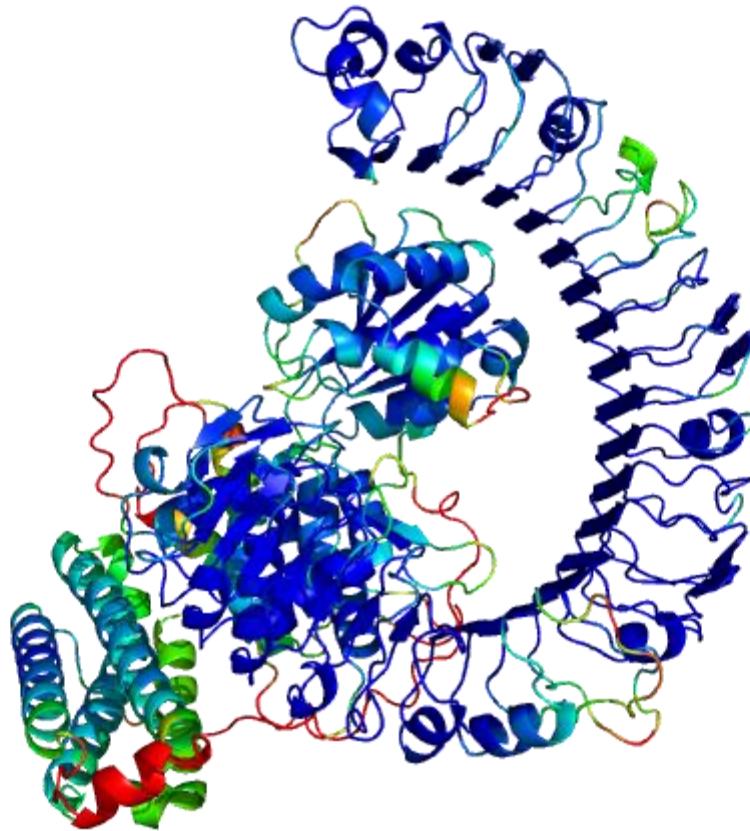


SASA



Screenshot

ALPHAFOLD

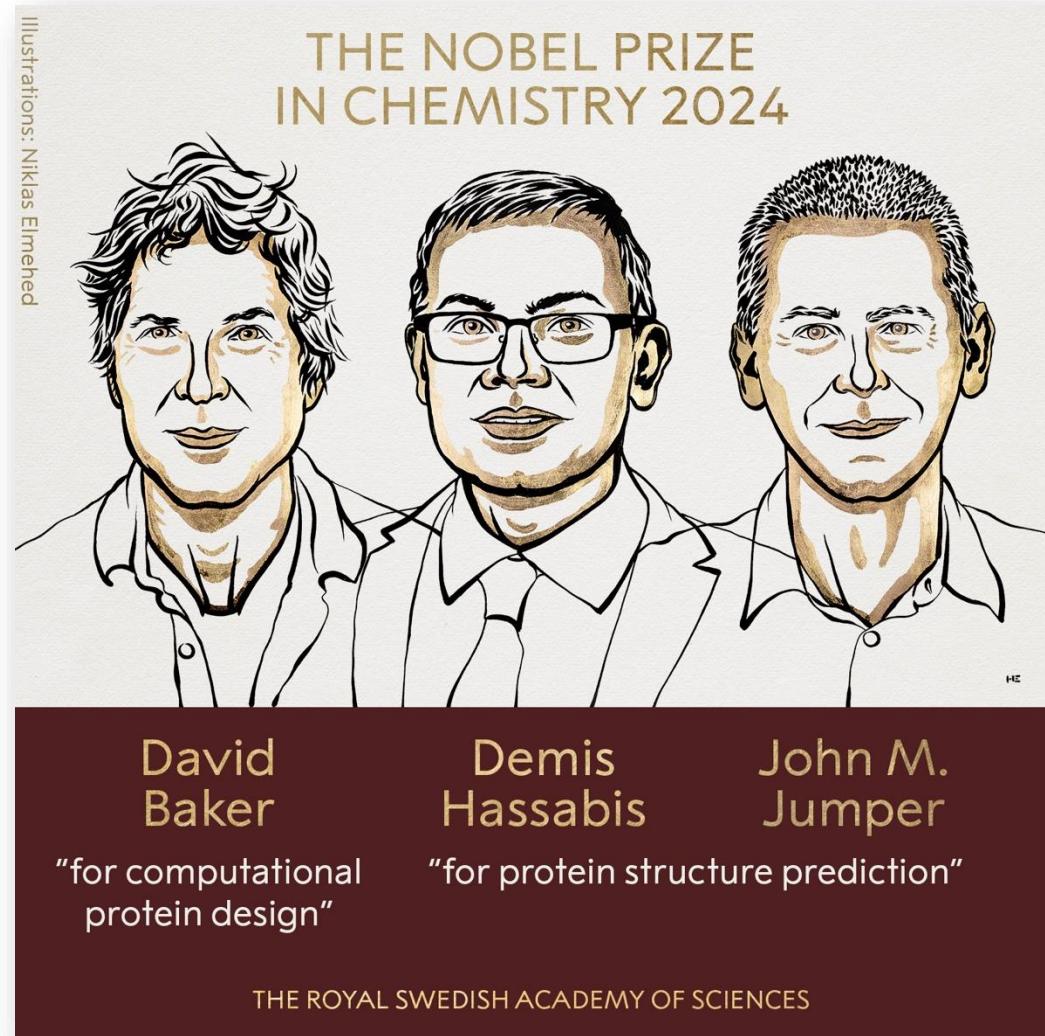


Google DeepMind

ALPHAFOLD



CRIADORES DO ALPHAFOLD

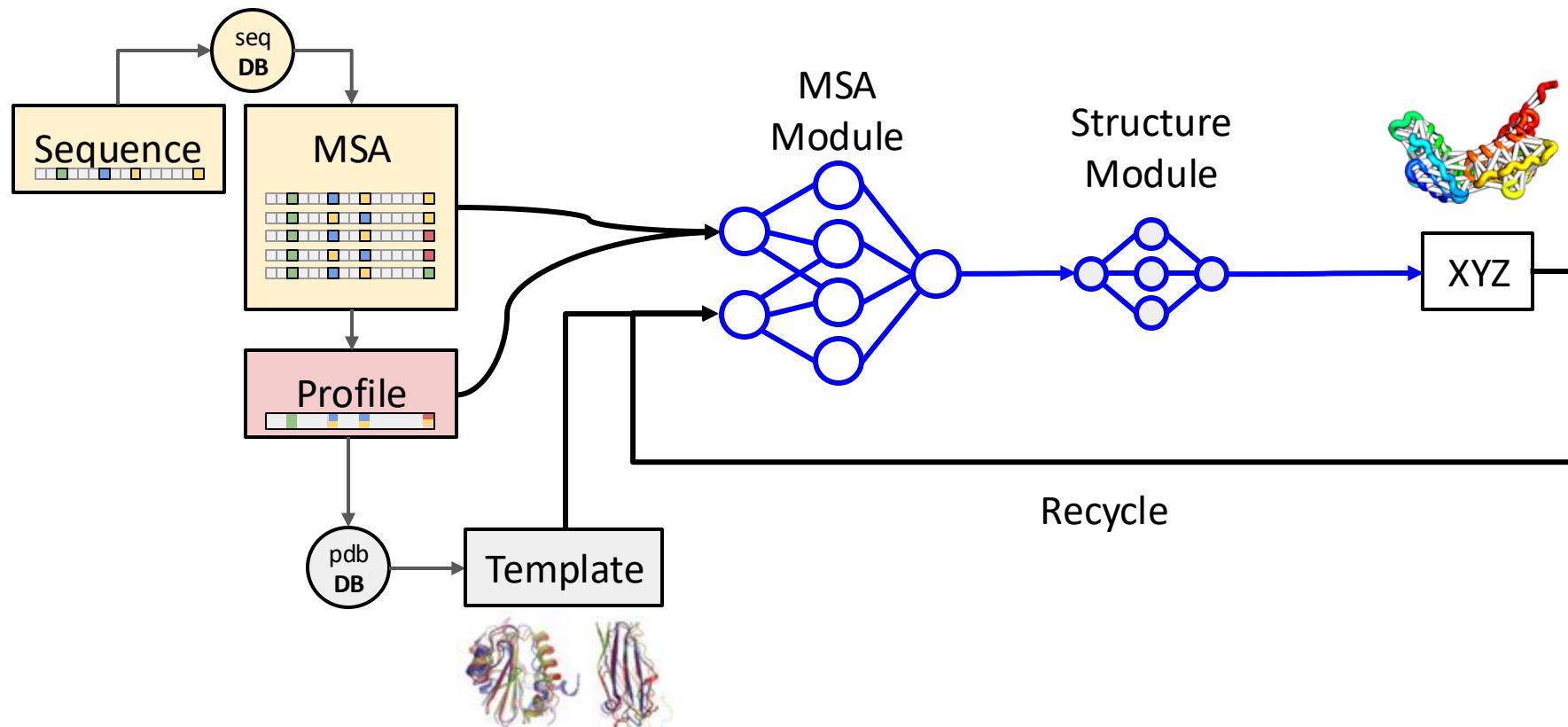


THE ROYAL SWEDISH ACADEMY OF SCIENCES

INTELIGÊNCIA ARTIFICIAL E MODELAGEM

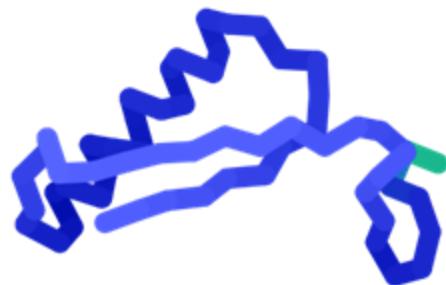
Jumper J. et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*

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*

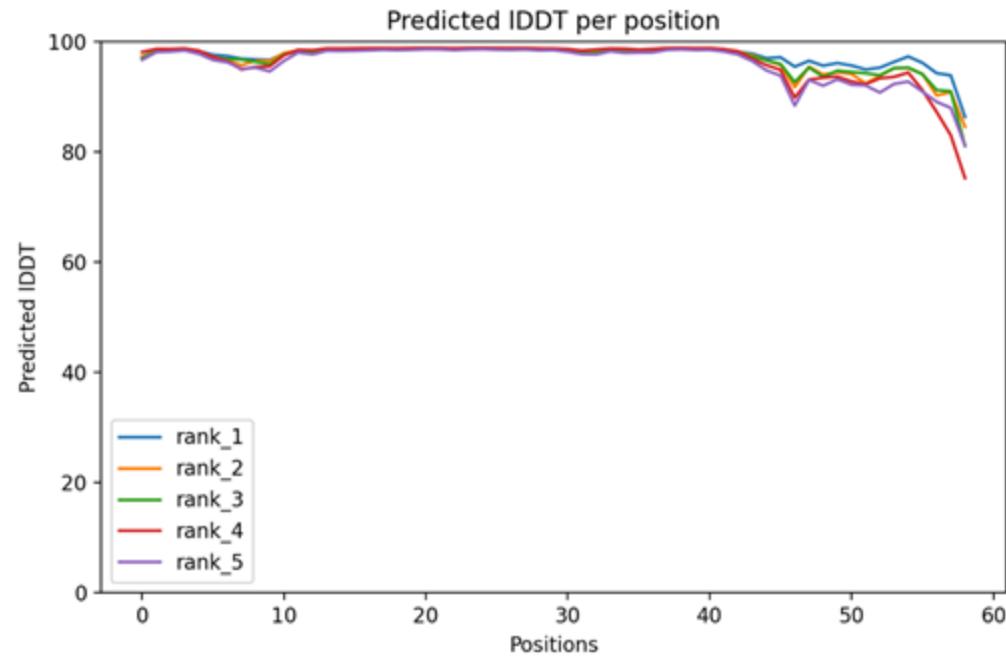


MÉTRICA DE CONFIABILIDADE (AF)

- pLDDT – “per-residue Local Distance Difference Test”
 - 0 a 100 (quanto maior, melhor)
 - **Muito baixo** (<50), **Baixo** (60), **OK** (70), **Confiável** (80), **Muito alto**(>90)



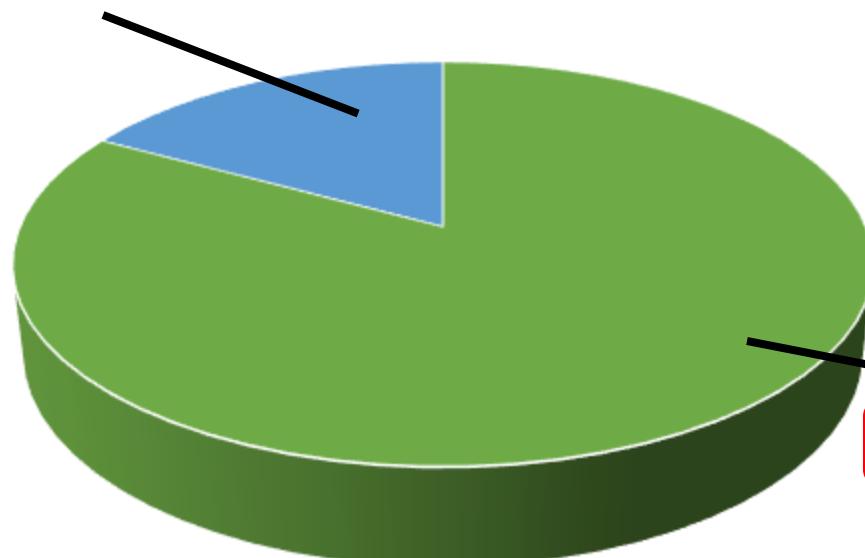
pLDDT 96.1





+285 milhões proteínas
depositadas

Sem modelo



Com modelos



+214 milhões proteínas
preditas

ATIVIDADE



AlphaFold Server

Powered by AlphaFold 3

AlphaFold 3 model is a Google DeepMind and Isomorphic Labs collaboration

<https://golgi.sandbox.google.com/>

ALPHAFOLD 3 SERVER

AlphaFold Server BETA Server About FAQs

Remaining jobs: 18

AlphaFold Server allows you to model a structure consisting of many biological molecules [Learn more ^](#)

- Remaining jobs refresh each day
- Jobs can be up to 5,000 tokens - see more details on token calculation, accepted formats, seed selection and other features in our [FAQ](#)
- ⚙️ Use the entity bar to chemically modify proteins and nucleic acids
- 💬 Get in touch with the AlphaFold team if you have any questions

Explore these examples of structures to see it in action – try them out without using your quota until you begin editing!

Protein-RNA-Ion: PDB 8AW3 Protein-Glycan-Ion: PDB 7BBV Protein-DNA-Ion: PDB 7RCE

Ok, got it

1º Colar sequência de interesse

2º (opcional) definir número de cópias

Molecule type: Protein Copies: 1

Ex:
1 – monômero
2 – dímero
.

AEARTCESQTHKFKGTCLSETNCANVCKTEGFTGGDCRGLRRRCFCTRHC

Upload JSON Clear

Save job

Continue and preview job

3º Clique em continuar

The screenshot shows the AlphaFold 3 Server interface. At the top, there's a navigation bar with 'AlphaFold Server' (BETA), 'Server' (selected), 'About', and 'FAQs'. On the right are icons for help, light/dark mode, and user profile. Below the navigation is a message about remaining jobs (18). A central panel describes the service and lists usage tips. Below that is a section for exploring examples with links to PDB entries. The main work area has a 'Copies' input set to 1, with a red arrow pointing to it from the '2º (opcional)' callout. To the right of the sequence input is a red arrow pointing to the 'Continue and preview job' button from the '3º' callout. A large red arrow points from the 'Ex:' text block down to the sequence input. The sequence itself is highlighted with a red dashed box. A 'Save job' button is also visible. At the bottom, a note says 'Teste: AEARTCESQTHKFKGTCLSETNCANVCKTEGFTGGDCRGLRRRCFCTRHC'.

Teste: AEARTCESQTHKFKGTCLSETNCANVCKTEGFTGGDCRGLRRRCFCTRHC

ALPHAFOLD 3 SERVER

Job name
2024-07-15_15:45

We've generated a job name for you, please edit

Seed: Auto

Type Copies Sequence

Protein 1 AEARTCESQTHKFKGTCLSETNCANVCKTE... (length 50)

Remaining jobs: 18

[Go back and edit this job](#) [Confirm and submit job](#)

(Opcional) edite o nome do trabalho

Clique para confirmar

Baixar os melhores modelos gerados

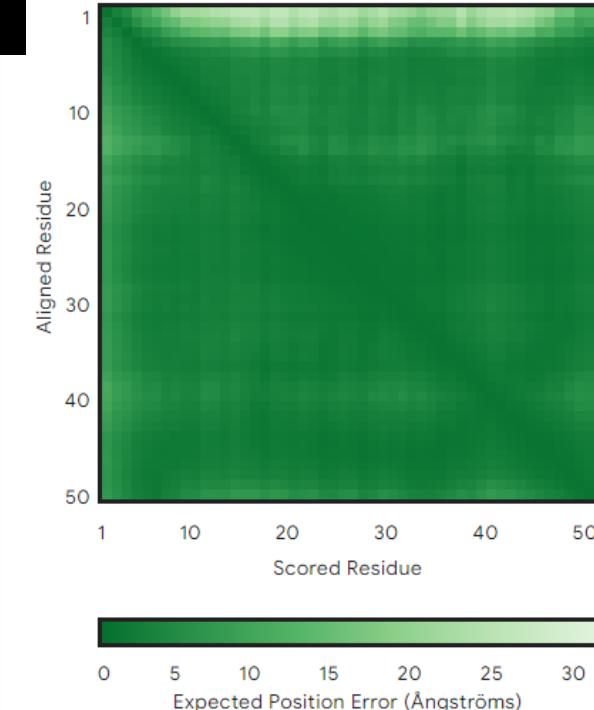
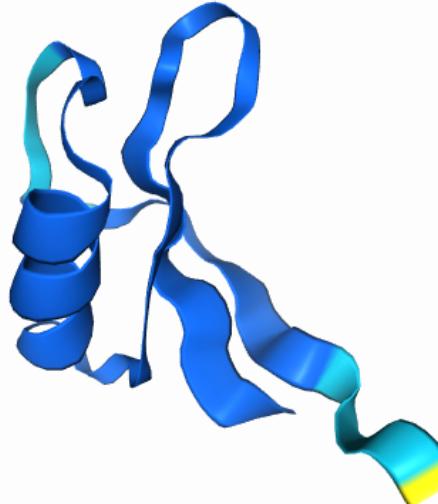
2024-07-15_15:14

[← Back](#)[Download](#)[Clone and reuse](#)[Feedback on structure](#)Very high ($p\text{DDT} > 90$)Confident ($90 > p\text{DDT} > 70$)Low ($70 > p\text{DDT} > 50$)Very low ($p\text{DDT} < 50$)

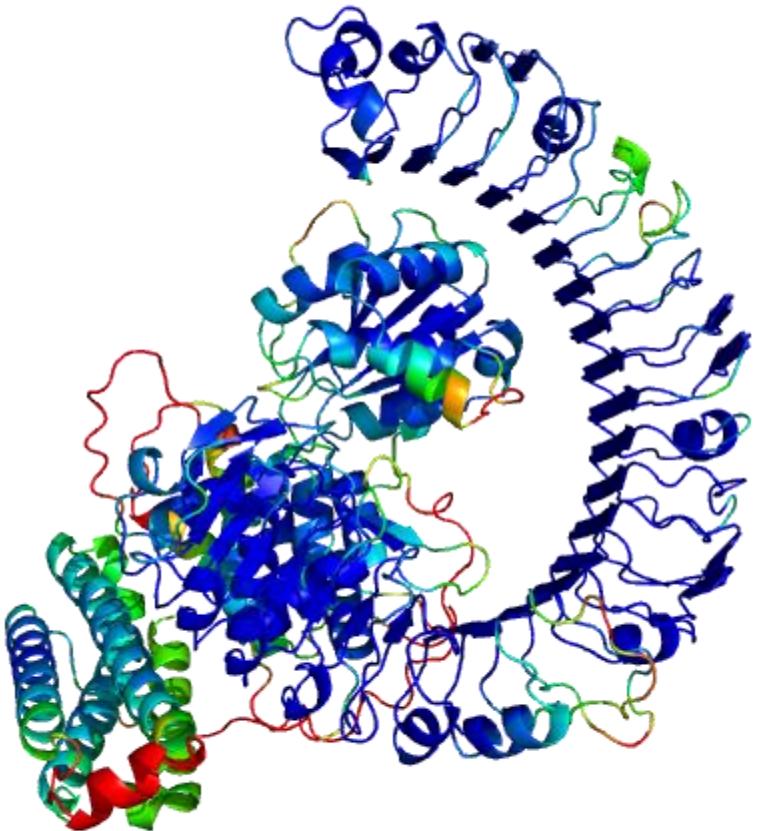
$p\text{DDT}$ scores de confiabilidade

 $\text{ipTM} = -$ $p\text{TM} = 0.71$ [learn more](#)

Pontuação de modelagem prevista comparada a estruturas verdadeiras.
($p\text{TM} > 0.5 \rightarrow$ dobramento previsto é similar à estruturas verdadeiras)

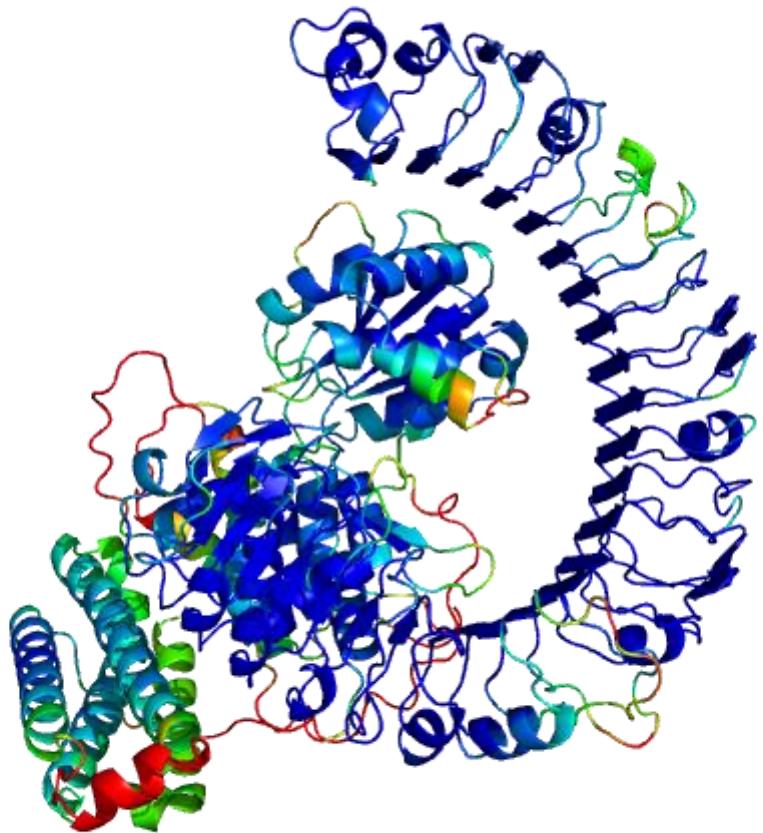


PAE: Erro de alinhamento previsto
("quanto menor, melhor")



**MAS O QUÃO BOM O AF3 É
QUANDO COMPARADO COM
MÉTODOS EXPERIMENTAIS?**

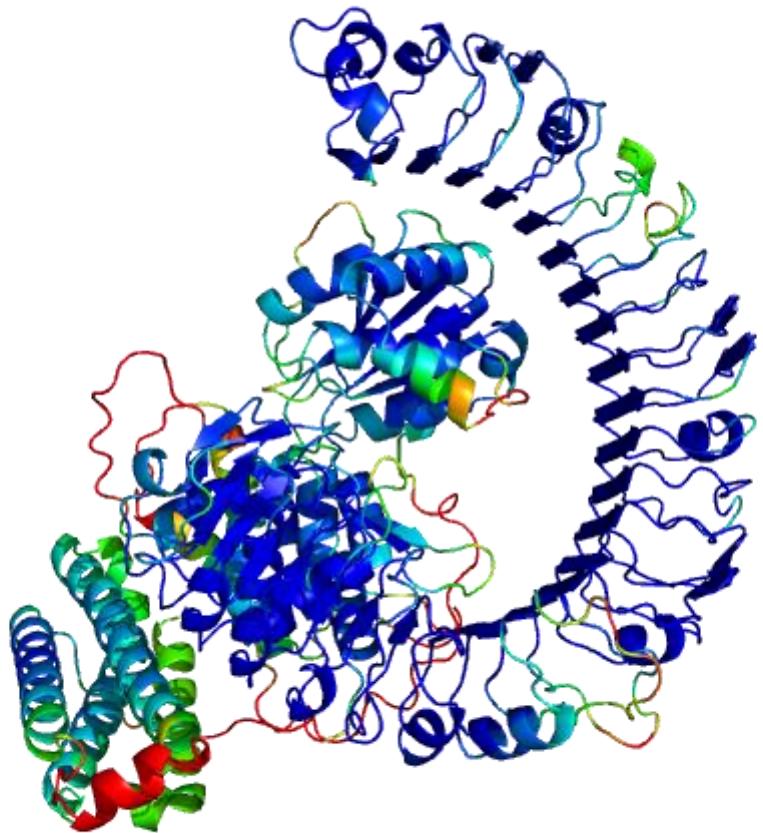
VAMOS VER?



p53 EXPERIMENTAL (X-RAY)

X

p53 MODELADA PELO AF3



THERE WE GO!

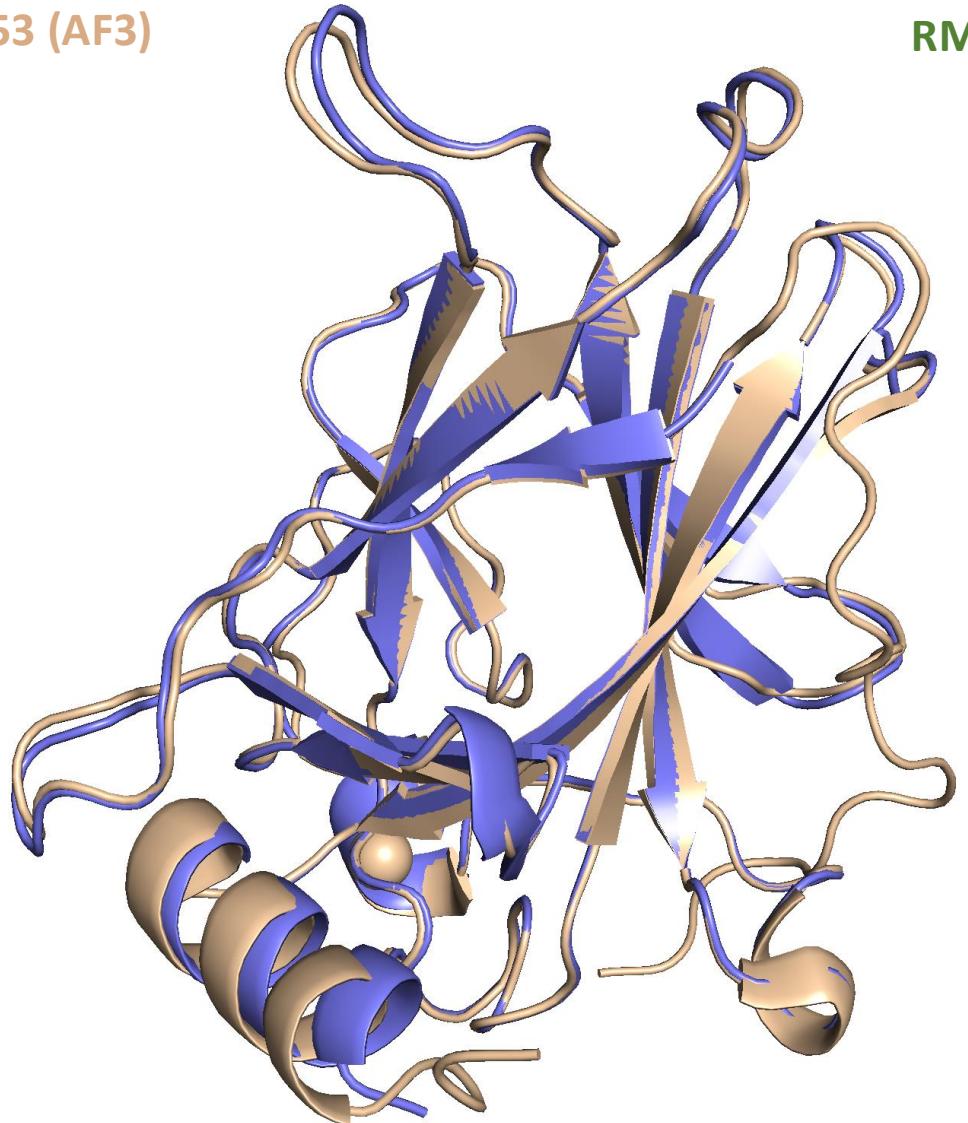


p53 (X-RAY)

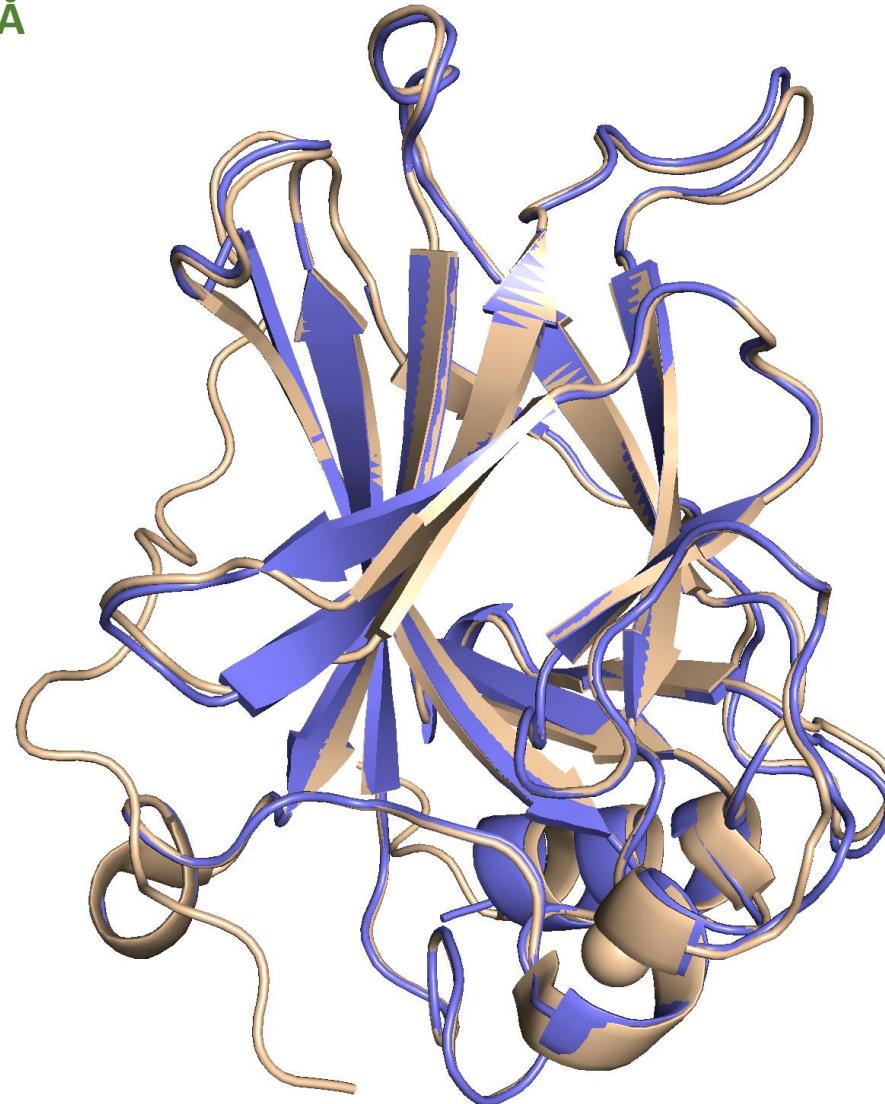


p53 (AF3)

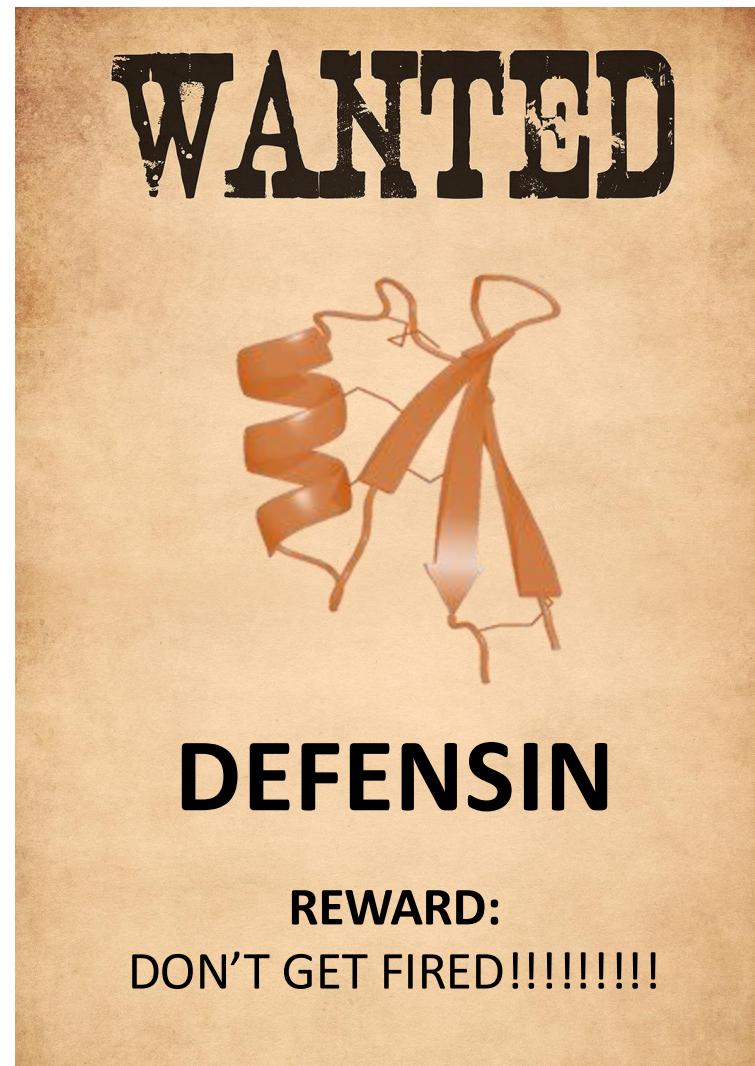
P53 CORE DOMAIN
RMSD = 0.209 Å



180°



AGORA É A SUA VEZ



AGORA É A SUA VEZ

CDS
(Cajanus cajan)



ATGGACAAGGCACGATTGGGCTTCTCGTATTGTTGATTCTCCTGCTTCAGATGGTGGTAGAAC
GGAGGGAAAGGCCTGCGAACATCAAAGAGCCATCGTTAAAGGGATGTGCTAAGCGATCACAACTGCGG
TTCCGTTGCCATGTTGAAGGCTTCACAGGTGGCAAATGTCATGGATTCCGTAGACGCTTTCTGCACTA
AGCACTGTTAG

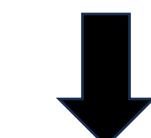
ORFFinder
[\(https://www.ncbi.nlm.nih.gov/orffinder/\)](https://www.ncbi.nlm.nih.gov/orffinder/)

ou

Expasy – translate tool
[\(https://web.expasy.org/translate/\)](https://web.expasy.org/translate/)



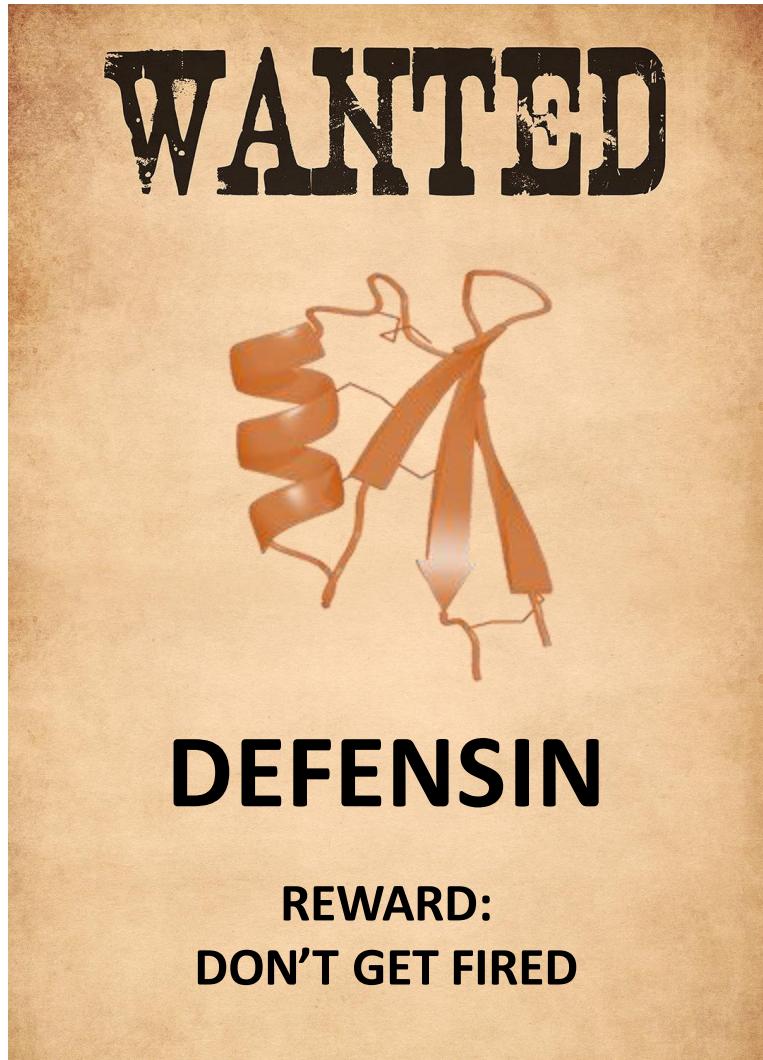
CDSearch
<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>



Proteína
(AF3)



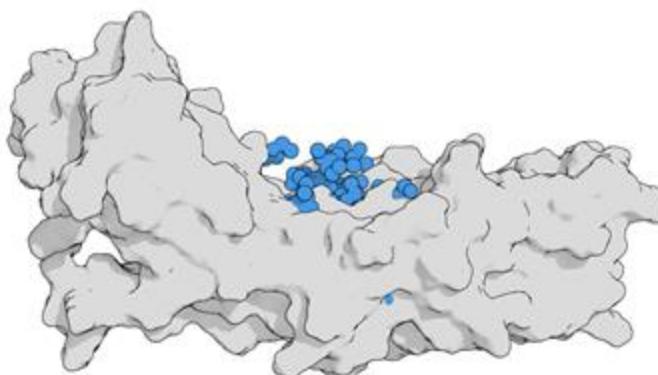
AGORA É A SUA VEZ



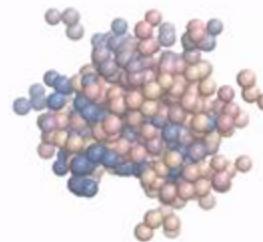
- Qual a sequência completa?
- Qual a família do domínio conservado?
- Qual a sequência do domínio conservado?
- A estrutura tridimensional do domínio foi predita como esperado (compare com a imagem do anúncio)?
- Vamos conversar...

CONSIDERAÇÕES FINAIS

- Elucidar estruturas de proteínas foi um **desafio** por muito tempo, em razão da **longa duração e alto custo** para validação experimental;
- **Ferramentas computacionais** tornaram possível **predizer a estrutura** a partir da sequência de aminoácidos. No entanto, inicialmente era **dependente de *templates* depositados** e da qualidade desses dados;
- Na era da **inteligência artificial** e avanço tecnológico com redes neurais, é possível **gerar modelos tridimensionais** com acentuada **confiabilidade** e em um **curto período** de tempo;
- Apesar desse avanço, a **validação experimental continua sendo o padrão ouro**.



THAT'S ALL, FOLKS!



E QUALQUER DÚVIDA,
É SÓ MANDAR UM E-MAIL OU MENSAGEM



madsondeluna@gmail.com
madsondeluna@ufmg.com



[https://www.linkedin.com/in/
/madsonaragao/](https://www.linkedin.com/in/madsonaragao/)



[https://github.com/
madsondeluna](https://github.com/madsondeluna)