

ATP sintase

ESTRUTURA DE PROTEÍNAS

BIOQUÍMICA 2013/2014

Ana Margarida Pinheiro
pinheiroanamargarida@gmail.com

DOGMA CENTRAL DA BIOLOGIA MOLECULAR

The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that such information cannot be transferred from protein to either protein or nucleic acid.

Crick, F (August 1970). "Central dogma of molecular biology." *Nature* **227** (5258): 561–3



DOGMA CENTRAL DA BIOLOGIA MOLECULAR

- Exceções

Table of the 3 classes of information transfer suggested by the dogma

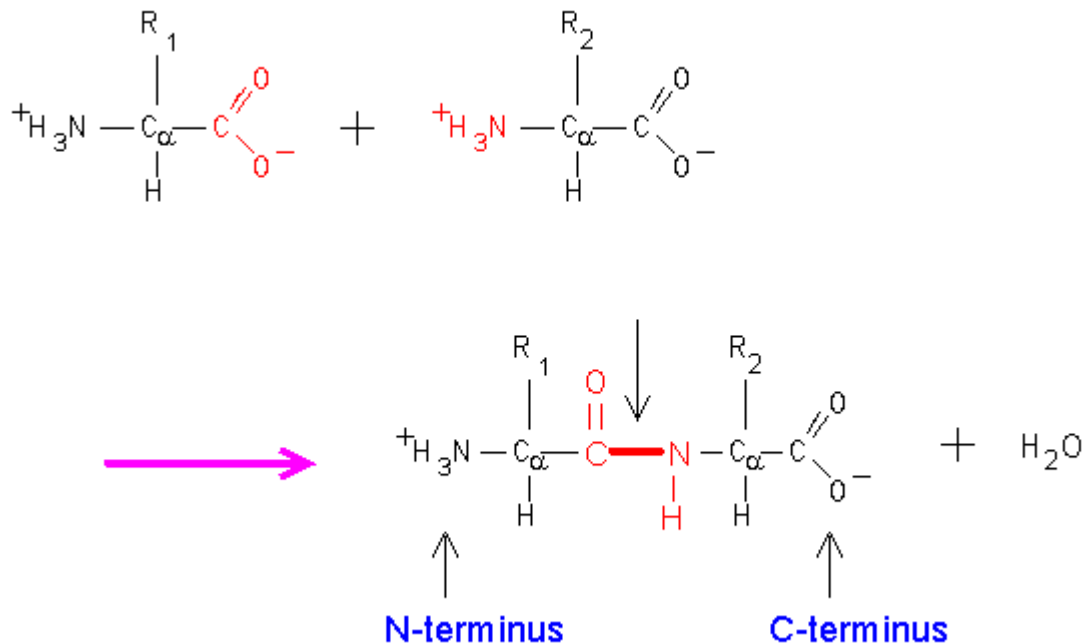
General	Special	Unknown
DNA → DNA	RNA → DNA	protein → DNA
DNA → RNA	RNA → RNA	protein → RNA
RNA → protein	DNA → protein	protein → protein



NÍVEIS DE ORGANIZAÇÃO DAS PROTEÍNAS

○ Estrutura Primária

- Nível de estrutura mais simples
- Dada pela sequência de aminoácidos



NÍVEIS DE ORGANIZAÇÃO DAS PROTEÍNAS

○ NCBI

The screenshot displays the NCBI website homepage. At the top, there is a navigation bar with the NCBI logo, "Resources" and "How To" dropdown menus, and a "Sign in to NCBI" link. Below this is a search bar with a dropdown menu set to "All Databases" and a "Search" button. The main content area is divided into three columns. The left column contains a vertical navigation menu with items like "NCBI Home", "Resource List (A-Z)", "All Resources", "Chemicals & Bioassays", "Data & Software", "DNA & RNA", "Domains & Structures", "Genes & Expression", "Genetics & Medicine", "Genomes & Maps", "Homology", "Literature", "Proteins", "Sequence Analysis", "Taxonomy", "Training & Tutorials", and "Variation". The middle column features a "Welcome to NCBI" section with a paragraph about the center's mission and a list of "Get Started" links: "Tools", "Downloads", "How-To's", and "Submissions". Below this is a "Genotypes and Phenotypes" banner with a diagram of a family tree and text about genome-wide association studies. The right column contains "Popular Resources" (PubMed, Bookshelf, PubMed Central, PubMed Health, BLAST, Nucleotide, Genome, SNP, Gene, Protein, PubChem) and "NCBI Announcements" with several news items dated from February 2014.

www.ncbi.nlm.nih.gov

NCBI Resources How To Sign in to NCBI

NCBI National Center for Biotechnology Information

All Databases Search

NCBI Home

Resource List (A-Z)

All Resources

Chemicals & Bioassays

Data & Software

DNA & RNA

Domains & Structures

Genes & Expression

Genetics & Medicine

Genomes & Maps

Homology

Literature

Proteins

Sequence Analysis

Taxonomy

Training & Tutorials

Variation

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [Research](#) | [NCBI News](#)

Get Started

- [Tools](#): Analyze data using NCBI software
- [Downloads](#): Get NCBI data or software
- [How-To's](#): Learn how to accomplish specific tasks at NCBI
- [Submissions](#): Submit data to GenBank or other NCBI databases

Genotypes and Phenotypes

Data from Genome Wide Association studies that link genes and diseases. See study variables, protocols, and analysis.

Popular Resources

- PubMed
- Bookshelf
- PubMed Central
- PubMed Health
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

NCBI Announcements

Genome Workbench Update 2.7.15 released
Feb 26, 2014

Genome Workbench 2.7.15 has been released. The update includes several new features.
Feb 26, 2014

New CDD Release v.3.11 includes recomputed PSSMs and more
Feb 19, 2014

Conserved Domain Database (CDD) version 3.11 is now available with 506 new domains.
Feb 18, 2014

GenBank has milestone 200th release
Feb 18, 2014

GenBank's 200th release is now available

NÍVEIS DE ORGANIZAÇÃO DAS PROTEÍNAS

○ UniProt

The screenshot shows the UniProt website interface. At the top, there is a navigation bar with the UniProt logo and links for Downloads, Contact, and Documentation/Help. Below this is a search bar with tabs for Search, Blast, Align, Retrieve, and ID Mapping. The search bar is set to 'Protein Knowledgebase (UniProtKB)' and has a 'Query' input field with 'Search', 'Advanced Search', and 'Clear' buttons.

WELCOME

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

What we provide

UniProtKB	Protein knowledgebase, consists of two sections: <ul style="list-style-type: none">★ Swiss-Prot, which is manually annotated and reviewed.★ TrEMBL, which is automatically annotated and is not reviewed. Includes complete and reference proteome sets .
UniRef	Sequence clusters, used to speed up sequence similarity searches.
UniParc	Sequence archive, used to keep track of sequences and their identifiers.
Supporting data	Literature citations , taxonomy , keywords , subcellular locations , cross-referenced databases and more.

Getting started

- [Text search](#)
- [Sequence similarity searches \(BLAST\)](#)
- [Sequence alignments](#)
- [Batch retrieval](#)

NEWS

UniProt release 2014_02 - Feb 19, 2014

Epigenetics in the spotlight | Change of the cross-references to PROSITE and HAMAP | Revision of the UniParc records used in the UniRef databases

- › [Statistics for UniProtKB:](#)
 - [Swiss-Prot](#) · [TrEMBL](#)
- › [Forthcoming changes](#)
- › [News archives](#)

[Follow @uniprot](#) 844 followers

SITE TOUR

Learn how to make best use of the tools and data on this site.

NÍVEIS DE ORGANIZAÇÃO DAS PROTEÍNAS

○ UniProt

www.uniprot.org/blast/uniprot/20140312727GL0UHCY

UniProt > blast/uniprot

Search **Blast** Align Retrieve ID Mapping *

Sequence or UniProt ID:

MDYSYDLEEVEETIDYKDPCKAAAFWGDIALDEEDLANFKIDRIVDLTKHTIHTVSGAA
TNISRPEKGRRTKERRRSREKRASTRSPERVWPDGVIYVVISGNFSGSQRAIFRQAMRH
WEKHTCVTFLERTDEDSYIVFTYRPGCCSYVGRGGGFPQAISIGKNCDKFGIVVHELGH
VIGFWHEHTRPDRDDHVSIIRENIQPGQYFNFLKMEPEEVEVSLGETYDFDSIMHYARNIF
SRGIFLDTILPKYDVNGVRPPIGQRTLRSSGDVAQARKLYKCPACGETLQDSQGNFSSPG
FPNGYSAYMHCVRLSVTPGEKIIILNFTSLDLYRSRLCWYDYIEVRDGFWKKAPLRGRFC

Blast Clear Options

Help
For a sequence similarity search, enter:
• a protein or nucleotide sequence
• a UniProt identifier, e.g.
P00750 or A4_HUMAN or UPI0000000001
More...

Database: UniProtKB Threshold: 10 Matrix: Auto Filtering: None Gapped Hits: yes 250

250 hits for blastp blast on UNIPROTKB sorted by score descending

Browse by taxonomy · keyword · gene ontology · enzyme class or pathway | Map to UniProtKB · UniRef or UniParc

Filter · Overview · Results · Job information · Customize order

Page 1 of 10 | Next

Filter

Dataset: UniProtKB (250) Taxonomy: Bilateria (250) Filter

Graphical overview

Color code for identity 0-100% =

Accession	Entry name	0Query hit707	0Match hit (sqrt scale)2092	Name (Organism)
<input type="checkbox"/>	20140312727GL0UHCY			
<input type="checkbox"/>	P98070 BMP1_XENLA			Bone morphogenetic protein 1 (Xenopus laevis)
<input type="checkbox"/>	Q66K13 Q66K13_XENLA			BMP-1 protein (Xenopus laevis)
<input type="checkbox"/>	O57381 O57381_XENLA			Bone morphogenetic protein 1b (Xenopus laevis)
<input type="checkbox"/>	Q28C16 Q28C16_XENTR			Bone morphogenetic protein 1 (Xenopus tropicalis)

NÍVEIS DE ORGANIZAÇÃO DAS PROTEÍNAS

UniProt

www.uniprot.org/blast/uniprot/20140312727GLOUHCY

Filter · Overview · Results · Job information · Customize order

Page 1 of 10 | Next

Detailed BLAST results

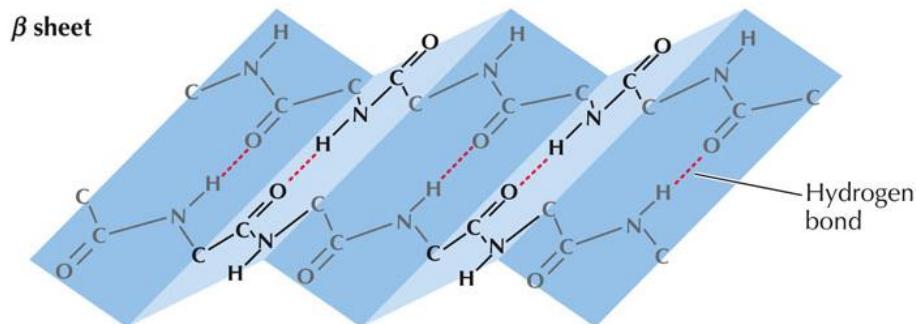
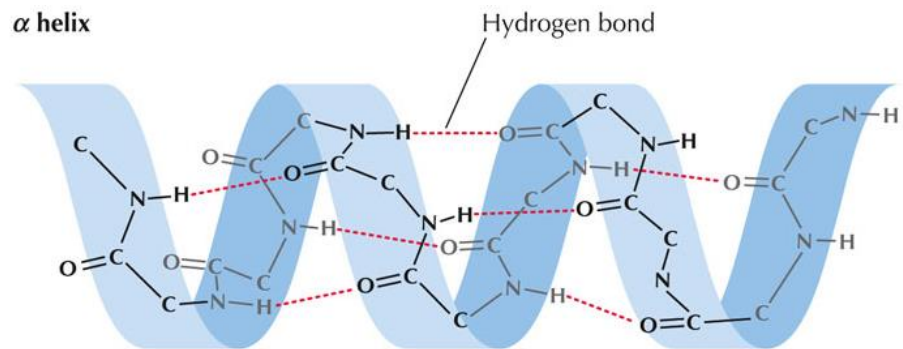
Show hits with 3D data only.
Show hits from complete proteomes only.

Alignments	Entry	Entry name	Status	Protein names	Organism	Length	Identity	Score	E-value	Gene names
	P98070	BMP1_XENLA	★	Bone morphogenetic protein 1	Xenopus laevis (African clawed frog)	707	100.0%	3,851	0.0	bmp1
	Q66K13	Q66K13_XENLA	★	BMP-1 protein	Xenopus laevis (African clawed frog)	735	96.0%	3,732	0.0	bmp1 BMP-1
	O57381	O57381_XENLA	★	Bone morphogenetic protein 1b	Xenopus laevis (African clawed frog)	735	96.0%	3,732	0.0	BMP-1
	Q28C16	Q28C16_XENTR	★	Bone morphogenetic protein 1	Xenopus tropicalis (Western clawed frog) (Silurana tropicalis)	734	96.0%	3,731	0.0	bmp1 TEgg002103.1
	F7CB15	F7CB15_XENTR	★	Uncharacterized protein	Xenopus tropicalis (Western clawed frog) (Silurana tropicalis)	734	96.0%	3,727	0.0	bmp1
	B7ZS79	B7ZS79_XENLA	★	Xtld protein	Xenopus laevis (African clawed frog)	977	99.0%	3,642	0.0	Bmp1
	B7ZS77	B7ZS77_XENLA	★	Xtld protein	Xenopus laevis (African clawed frog)	977	99.0%	3,642	0.0	Bmp1
	Q91925	Q91925_XENLA	★	Xtld protein	Xenopus laevis (African clawed frog)	977	99.0%	3,626	0.0	bmp1
	F7CB21	F7CB21_XENTR	★	Uncharacterized protein	Xenopus tropicalis (Western clawed frog) (Silurana tropicalis)	990	94.0%	3,520	0.0	bmp1
	U3F1G3	U3F1G3_CALJA	★	Bone morphogenetic protein 1 isoform 1	Callithrix jacchus (White-tufted-ear marmoset)	730	80.0%	3,155	0.0	BMP1
	U3D0M6	U3D0M6_CALJA	★	Bone morphogenetic protein 1 isoform 1	Callithrix jacchus (White-tufted-ear marmoset)	730	80.0%	3,147	0.0	BMP1
	U3AUX1	U3AUX1_CALJA	★	Bone morphogenetic protein 1 isoform 1	Callithrix jacchus (White-tufted-ear marmoset)	730	80.0%	3,147	0.0	BMP1
	P13497-2	BMP1_HUMAN	★	Isoform BMP1-1 of Bone morphogenetic protein ...	Homo sapiens (Human)	730	80.0%	3,140	0.0	BMP1 PCOLC
	K7C5B3	K7C5B3_PANTR	★	Bone morphogenetic protein 1	Pan troglodytes (Chimpanzee)	730	80.0%	3,136	0.0	BMP1
	F6RTG5	F6RTG5_MACMU	★	Bone morphogenetic protein 1 isoform 1	Macaca mulatta (Rhesus macaque)	730	80.0%	3,135	0.0	BMP1
	G3X157	G3X157_SARHA	★	Uncharacterized protein	Sarcophilus harrisii (Tasmanian devil) (Sarcophilus lanianus)	1,002	82.0%	3,134	0.0	BMP1
	Q59F71	Q59F71_HUMAN	★	Bone morphogenetic protein 1 isoform 1, varia...	Homo sapiens (Human)	803	80.0%	3,134	0.0	
	H3B7R3	H3B7R3_LATCH	★	Uncharacterized protein	Latimeria chalumnae (West Indian ocean coelacanth)	978	84.0%	3,132	0.0	

NÍVEIS DE ORGANIZAÇÃO DAS PROTEÍNAS

○ Estrutura Secundária

- Arranjo espacial de aminoácidos próximos



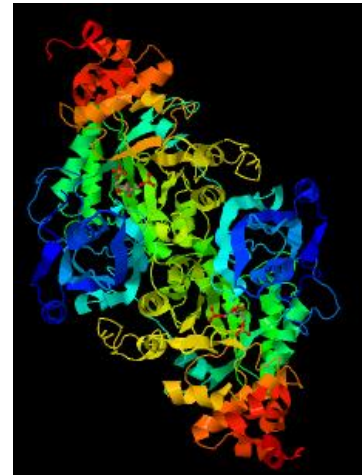
NÍVEIS DE ORGANIZAÇÃO DAS PROTEÍNAS

○ Estrutura Terciária

- Arranjo tridimensional das proteínas no espaço

○ Estrutura Quaternária

- Associação de diferentes cadeias polipeptídicas
- Ligadas entre si por ligações covalentes e não covalentes



RuBisCO presente em algumas bactérias



ESTRUTURA TRIDIMENSIONAL

PDB

www.rcsb.org/pdb/home/home.do

RCSB PDB PROTEIN DATA BANK

A MEMBER OF THE PDB | EMDataBank

An Information Portal to Biological Macromolecular Structures

As of Tuesday Feb 25, 2014 at 4 PM PST there are 98117 Structures | PDB Statistics

Search: Everything | Author | Macromolecule | Sequence | Ligand

e.g., PDB ID, molecule name, author

Search History, Previous Results

Biological Macromolecular Resource

Full Description

Learn: Featured Molecules

Structural View of Biology

Molecule of the Month
Broadly Neutralizing Antibodies

Viruses like HIV and influenza have evolved sneaky methods for evading our immune system. The immune system searches for foreign molecules, but several viruses have found ways to hide their unique parts and masquerade as normal human molecules. They do this in many ways. As viral surface glycoproteins are synthesized in infected cells, they are decorated with the same sugar chains that coat human proteins, providing an effective camouflage. The conserved functional sites of the viral protein are hidden deep in a pocket surrounded by these sugars, and thus are difficult for **antibodies** to reach. In addition, these viruses have error-prone replication machinery, which creates a great diversity in the viral glycoproteins. So unfortunately, once the immune system has found antibodies to recognize the infecting virus, other viruses rapidly mutate to change the site that is recognized.

Full Article

Protein Structure Initiative Featured System
Viroporins

Viruses always seem to find streamlined ways of doing things. Their enzymes are often multifunctional and more compact than their cellular counterparts, and overlapping genes allow them to pack a lot of genetic information into a small space. They have also found a very parsimonious way of creating membrane channels. These simple channels, termed viroporins, are stripped to the bare minimum, with just enough

Available on the App Store

PDB-101

Structural View of Biology
Understanding PDB Data
Molecule of the Month
Educational Resources
Author Profiles

MyPDB

Login to your Account
Register a New Account
MyPDB Help Page

Home

News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us

New Features

Latest release: December 2013

Tabular Reports

View Improved Tabular Reports

Website Release Archive:

RCSB PDB News

Weekly | Quarterly | Yearly

2014-02-25
Video Challenge for High School Students


ESTRUTURA TRIDIMENSIONAL – PORQUÊ ANALISAR?

- Fornece informações importantes acerca de:
 - Função biológica da proteína
 - Mecanismo de acção
 - Relações evolucionárias entre proteínas

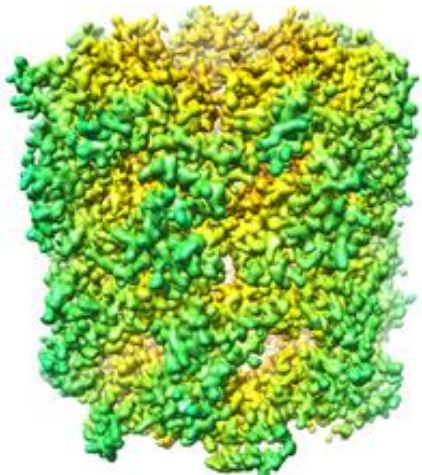


COMO OBTER A ESTRUTURA 3D?

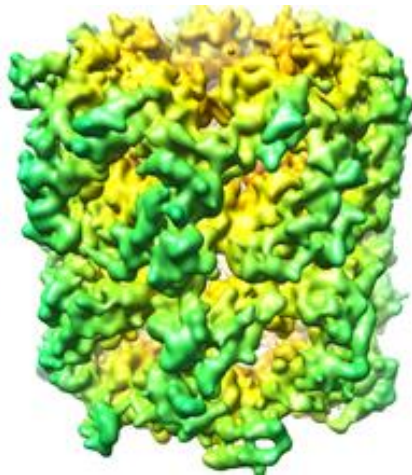
- **Cristalografia de Raio-X** (principal técnica utilizada na resolução de estrutura, resolução de 1-3 Å)
- **NMR** (resolve a estrutura com boa resolução mas está normalmente limitada a pequenas proteínas)
- **Microscopia Crioelectrónica** (aplicada para grandes complexos e “agregados” cuja estrutura não possa ser resolvida por cristalografia; resolução na ordem dos 10 Å)

$$1 \text{ \AA} = 10^{-10} \text{ m}$$


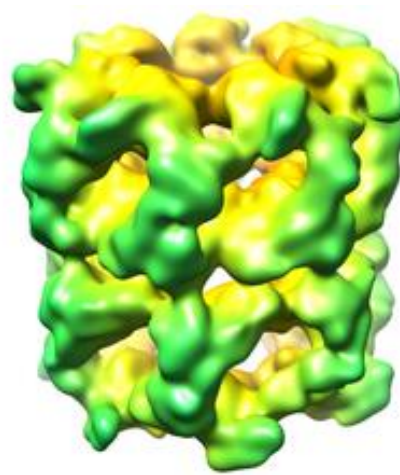
COMO OBTER A ESTRUTURA 3D?



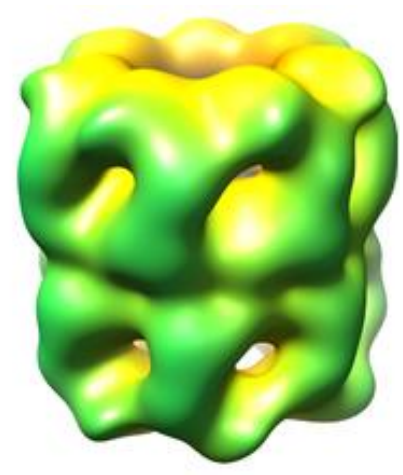
4 Å



8 Å



16 Å

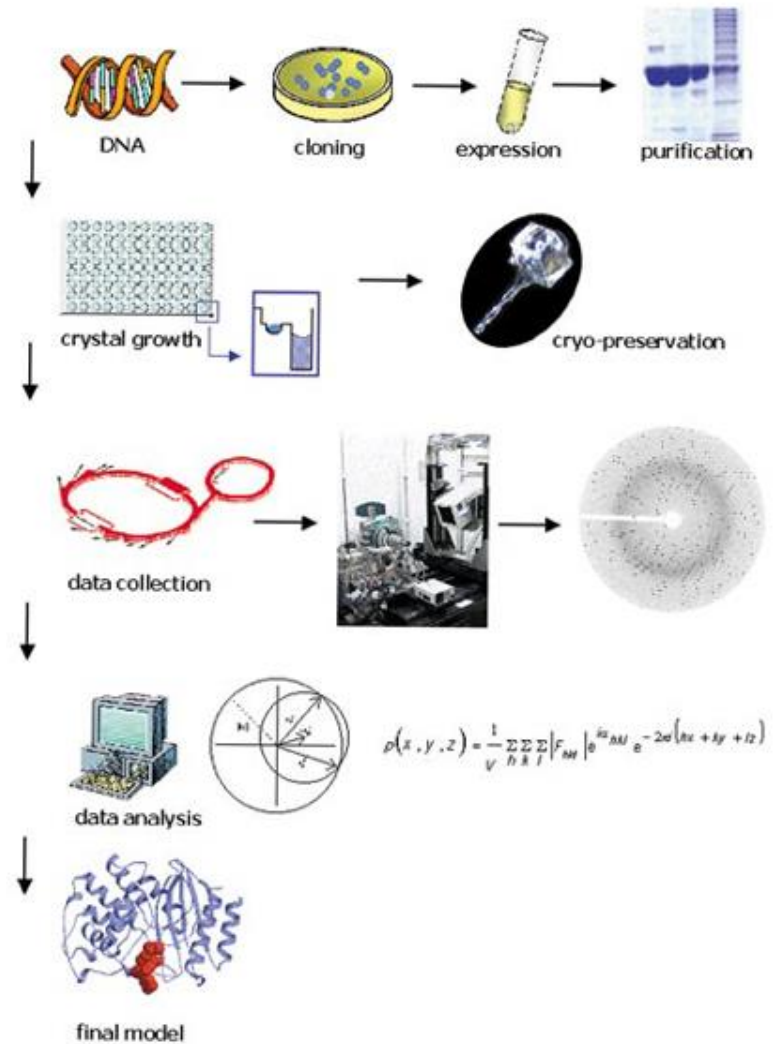


32 Å

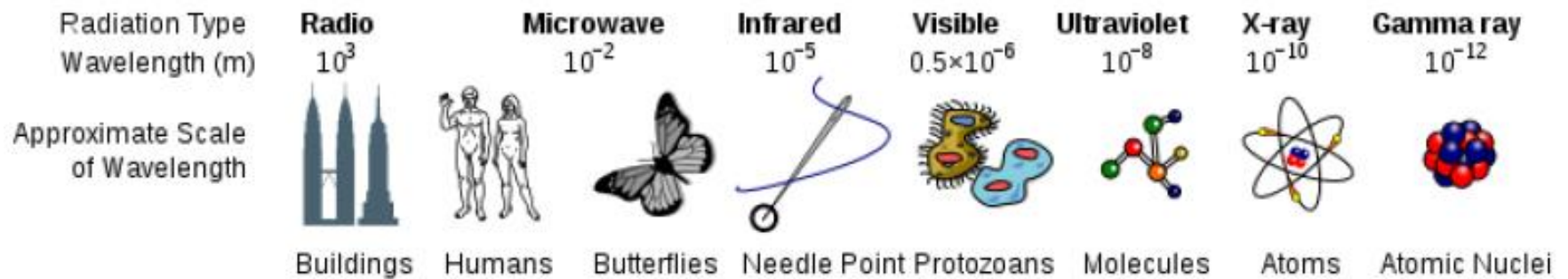


CRISTALOGRAFIA DE RAIO-X

- Produção da proteína e purificação
- Cristalização
- Difraccção de raio-X
- Determinação da estrutura
- Construção e melhoramento do modelo
- Análise da estrutura



CRISTALOGRAFIA DE RAIO-X



CRISTALOGRAFIA DE RAIO-X

**X-RAY
CRYSTALLOGRAPHY**

PREVISÃO DA ESTRUTURA 3D

- Impossibilidade de cristalizar a proteína (dificuldades ao nível da produção recombinante, da purificação, da formação do cristal, ...)

Swiss-Model

Phyre 2



SWISS-MODEL

- Utilizado quando não se consegue obter a estrutura tridimensional da proteína
- Actualmente existem inúmeras proteínas sequenciadas mas sem estrutura conhecida
- Compara a sequência de aminoácidos da proteína *target* com uma de sequência semelhante



SWISS-MODEL - PROCEDIMENTO

- Selecção do *template*
- Alinhamento
- Construção do modelo
- Avaliação do modelo



CHIMERA

- Beta conglutina de *Lupinus albus* (tremoço)

GI: 46451223

- Proteína de reserva das sementes do tremoço
- Apresenta dois domínios de cupina

Partes conservadas de
uma proteína

