

Python Programming & Computational Structural Biology

Computational Biology – Spring 2021

What's it about ?

Computational analysis and study of the structure and function of biological molecules (macromolecules and small molecules)

- Computational analysis and study of biological molecules (macromolecules and small molecules)
- Storage and management of structural information
- Computational algorithms for structure search and comparison
- Computational algorithms for the prediction of 2D and 3D structures of macromolecules
- Computational algorithms for the ***simulation*** of the physical and chemical behavior of biological (Biomolecular Simulation)
- Computational algorithms for protein-ligand analysis and discovery of new drugs (Computational Drug Design)

What will happen?

Learning

- Introduction of structural databases and formats of molecular representation
- ~~• Introduction to methods of macromolecule structure determination~~
- Prediction of secondary and tertiary structure of proteins
- ~~• Biomolecular Simulation Methods~~
- ~~• Introduction to protein-ligand interactions, computational docking and virtual screening~~

Doing

- Introduction to Python Programming and Jupyter notebooks and how to use them for analysis of sequences and structures of biological macromolecules
- Learning the basics of working in the Jupyter Hub and Linux environments
- How to use the molecular visualizer PyMOL to analyze, display and build structures of molecules (and make pretty pictures!)
- Constructing a protein model by comparative (homology) modelling
- ~~• Running a docking simulation with the Autodock Vina software~~

Who am I ?

- My name is [Paulo Martel](#)
- I teach Bioinformatics, Computational Biology, Drug Design, Protein Structure and Enzymology (far too much!)
- My research focus around computational analysis of protein structure and function using biomolecular simulation techniques, namely Molecular Dynamics, Comparative Modelling and Protein Electrostatics
- My office number is [3.12](#) in [Building 8](#) (but you probably won't go there...)
- My email is pmartel@ualg.pt
- Feel free to contact me or ~~drop by~~ if you have any questions regarding the course and its content
- This course will be taught in remote learning mode, but the final exam will be given in-person, in a classroom (in-person learning is back in UALG from April, 19th)

Before we start

- Most of computational activities will take place on a **remote server** with address <https://compbio-2021.ddns.net>
- You can login to that server from any computer you wish
- There are accounts for you on this server using, with login name same as your student number (without the leadin “a”).
- Login in with your student number as user and mas password **#123<your_student_number>** ... we can change those passwords later if you prefer.
- The environment we are using is called [Jupyter Hub](#) and provides access both to a [Python Notebook](#) interface and one or more [login terminals](#) on the server
- You will be required to install the following software on your personal computer:
 - PyMOL (you can download it from the site <https://www.pymol.org>)
 - ~~MGLTools (the instructor will provide a link for download)~~

Biomolecular Structures

Structure Leads to Function

Sequence

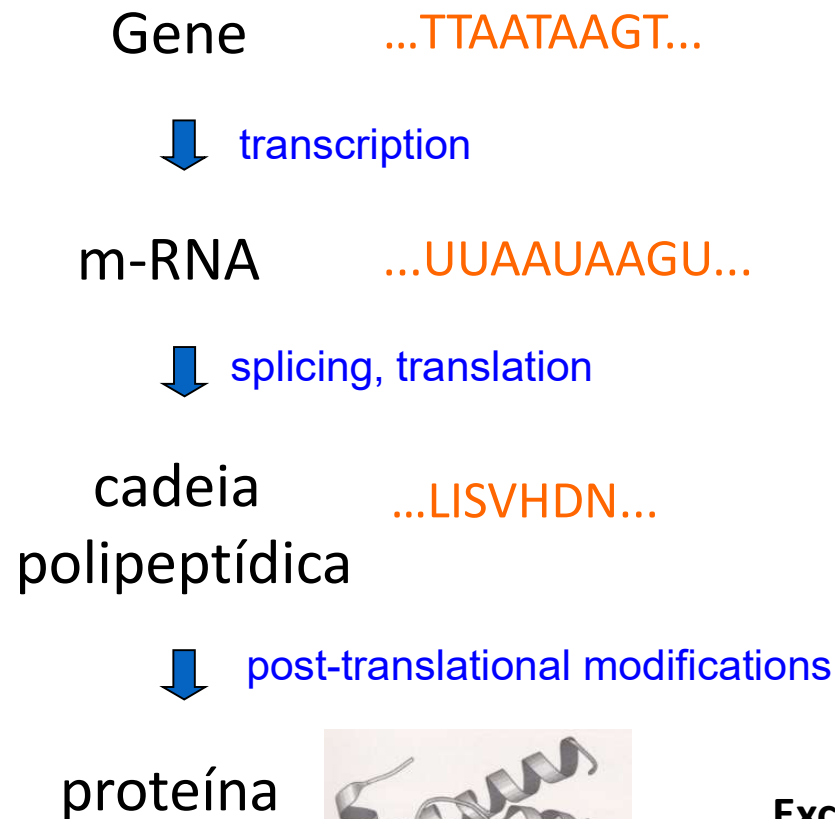


Structure

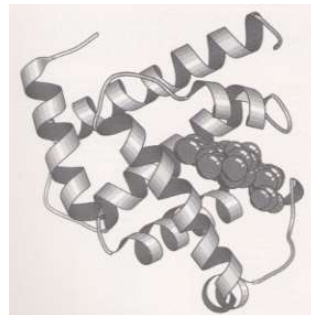


Function

Flow of Biological Information

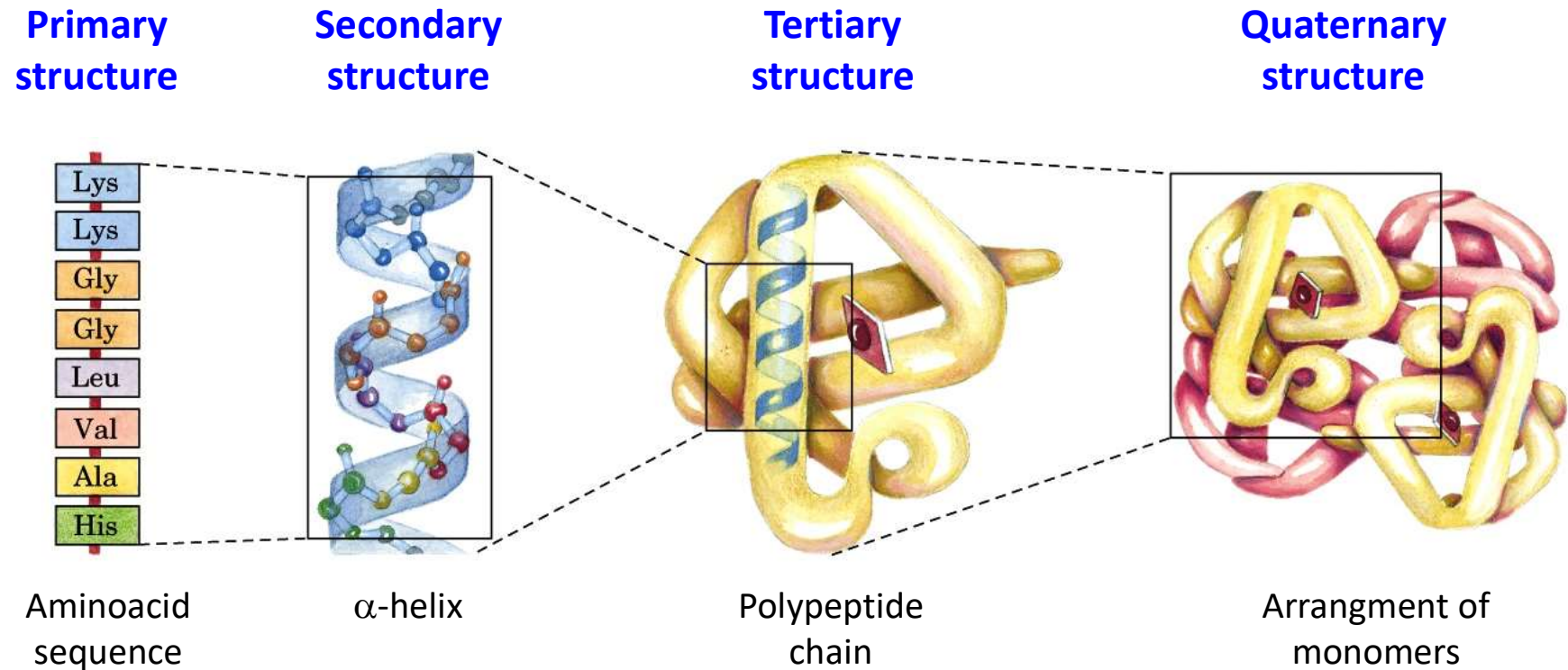


Central Dogma of
Molecular Biology



Exceptions: RNA viruses, prions, ribozymes (?)

Levels of structural organization

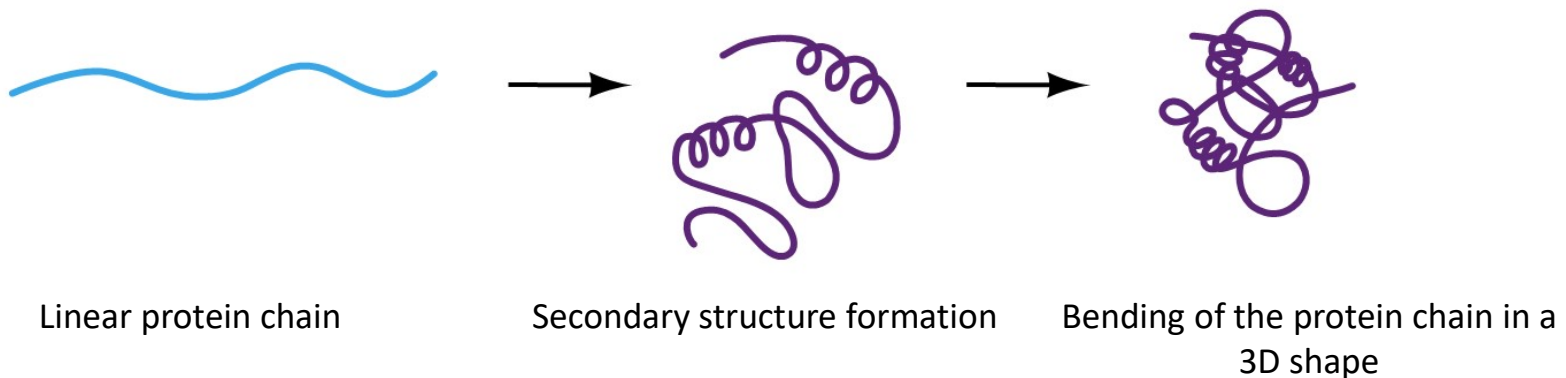


Sequence determines structure

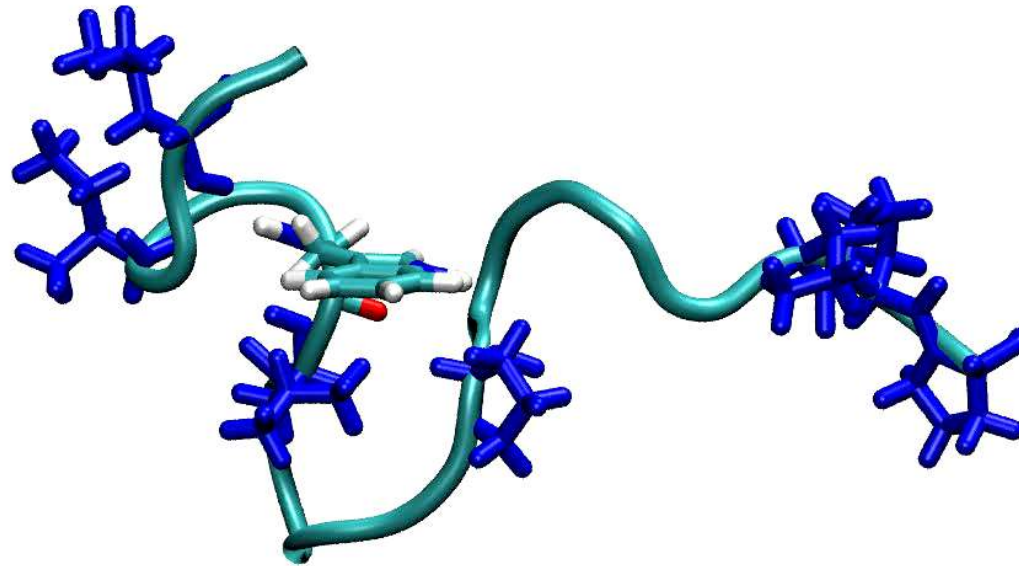
The tridimensional structure of proteins arises due to the physico-chemical forces acting between atoms in the polypeptide chain and the solvent. Many proteins will spontaneously acquire their *native* structure following ribosomal synthesis. This process is called *protein folding*.

Predicting the native structure of a protein given its sequence is one of the most fundamental problems of modern day molecular biology.
(*Folding problem*)

Protein folding:



Computing folding



Molecular Dynamics simulation of the folding mechanism of the mini-protein ***trp-cage***

Representing Structures

The representation of molecular structure is much more complex and storage intensive than sequence

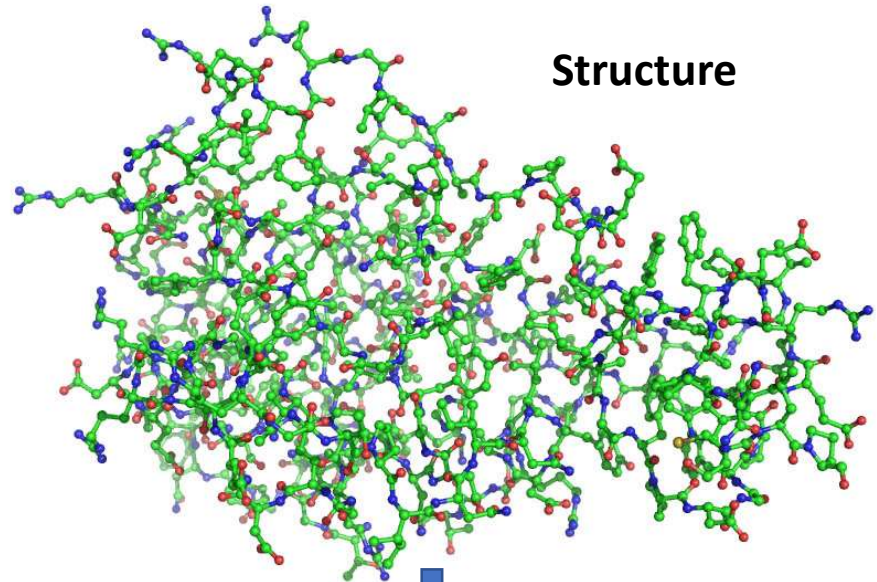
Sequence

...AVAGGATILVHNQDAGEPAIVLAFG...



Simple sequence of
one-letter symbols

Structure



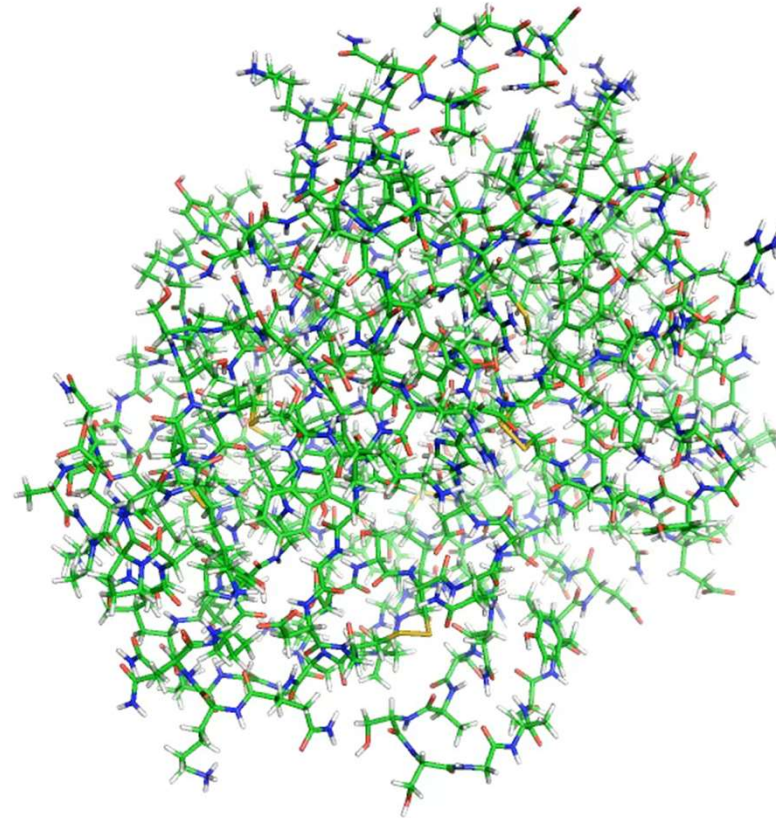
XYZ **coordinates** of each atom,
their **types** and **connectivity**

Example

Human Trypsin

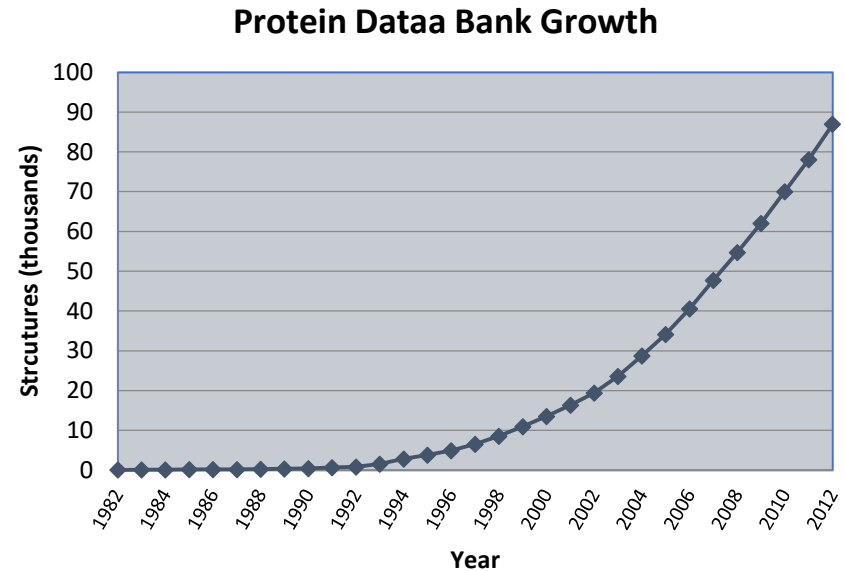
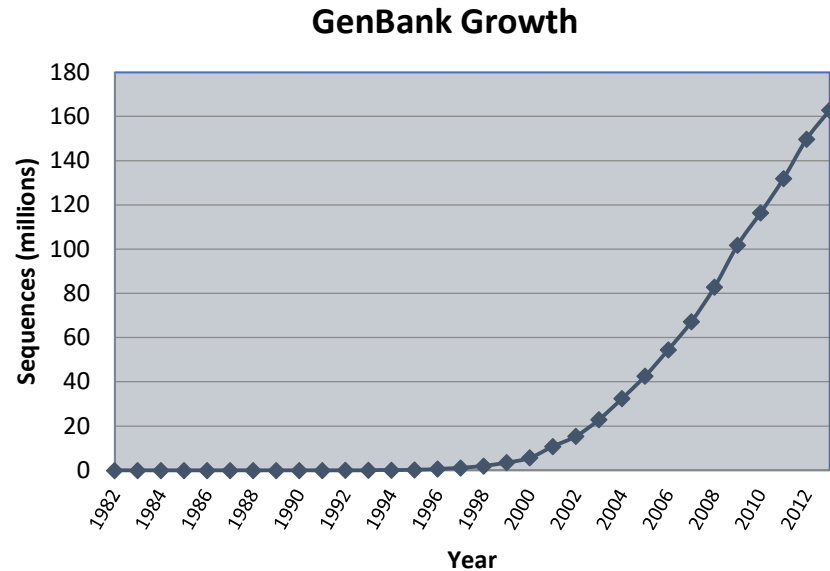
```
>sp|P07477|TRY1_HUMAN Trypsin-1 OS=Homo sapiens OX=9606 GN=PRSS1  
MNPLLILTFVAAALAAPFDDDDKIVGGYNCEENSVPYQVSLNSGYHFCGGSLINEQWVVS  
AGHCYKSRIQVRLGEHNIEVLEGNEQFINAAKIIRHPQYDRKTLNNDIMLIKLSRAVIN  
ARVSTISLPTAPPATGTKCLISGWGNTASSGADYPDELQCLDAPVLSQAKCEASYPGKIT  
SNMFCVGFLEGGKDSCQGDGSGGPVVCNGQLQGVVSWGDGCAQKNKPGVYTKVYNYVKWIK  
NTIAANS
```

247 aminoacids



3415 atoms

Sequence *versus* structure



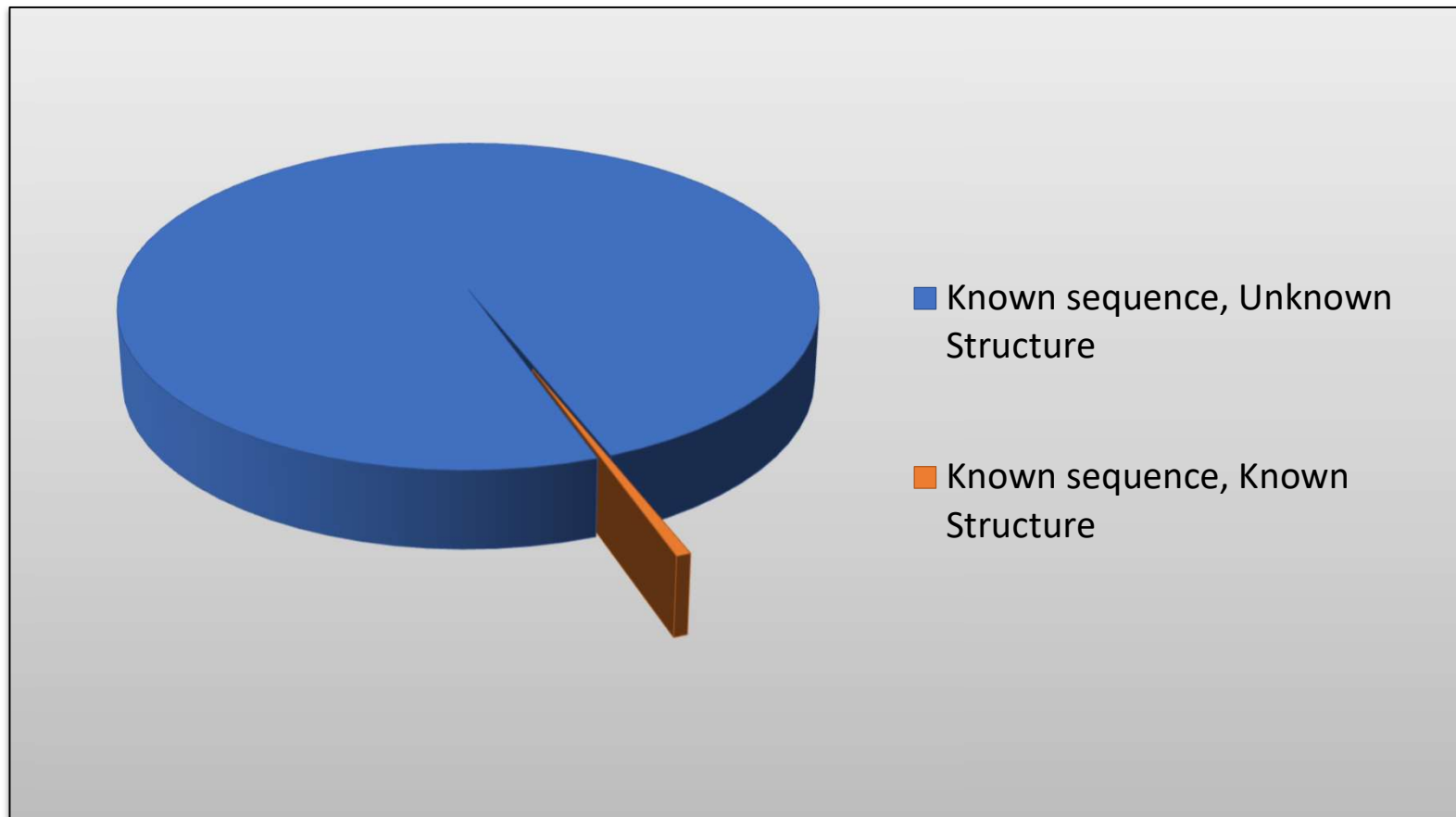
milions of sequences *versus* **thousands** of structures!

In 1982: 172 know structures and 602 sequences

Today (March 2019): 149,886 structures and 212,260,377 sequences!!

In conclusion: Sequencing is way faster than structure determination (the number of proteins of **known** sequence and **unknown** structure is growing very rapidly)!

Most known proteins have unknown sequence !



The importance of structure prediction

The vast (and steadily growing) number of proteins of unknown structure puts a heavy demand on ever faster methods for 3D structure determination. Due to their intricacies, such methods simply cannot cope with the fast pace of sequencing, and the gap widens more and more. This situation is not likely to change, ever.

So what can we do ?

We need to be able *predict* the 3D structure of proteins from their aminoacid sequence. In general terms this is a very hard computational problem, but there are special situations where it is very feasible. That is currently our best hope of coping up.

Prediction of the 3D structure of proteins is thus one of the most fundamental problems of bioinformatics / computational biology.

Biomolecular Databases and Formats

Macro vs. Small Molecules

- **Macromolecules:** contain the structure of biological macromolecules. The primary source is the **Protein Data Bank** (<https://www.rcsb.org>).
- **Small molecules:** databanks containing formulas, structures and other information relative to small molecules* . Some of the most important are:
 - PubChem (NCBI) :<https://pubchem.ncbi.nlm.nih.gov/>
 - ZINC (purchasable compounds): <https://zinc.docking.org/>
 - Drugbank (pharma oriented): <https://www.drugbank.ca/>
 - CCDC (crystallographic structures): <https://www.ccdc.cam.ac.uk/>

* The definition of small molecule may vary (~1000-atom limit)

Macromolecules

Macromolecular Databanks

- **Primary databanks:** contain the **raw** information, usually with a set of tools that can be accessed through a portal. Example: **Protein Data Bank** (<https://www.rcsb.org>).
- **Secondary Databanks:** specialized views, collections or filters on the primary databanks. Example: The PDBind database (<http://www.pdbbind.org.cn>)

- The development of molecular structure determination techniques lead to the accumulation of a large body of structures of proteins and nucleic acids (~160000)
- For the most part, those structures were solved using two structure determination methods, X-ray crystallography and nuclear magnetic resonance (NMR)
- The body of structures is stored in the public accessible Protein Databank (PDB) <http://www.rcsb.org>

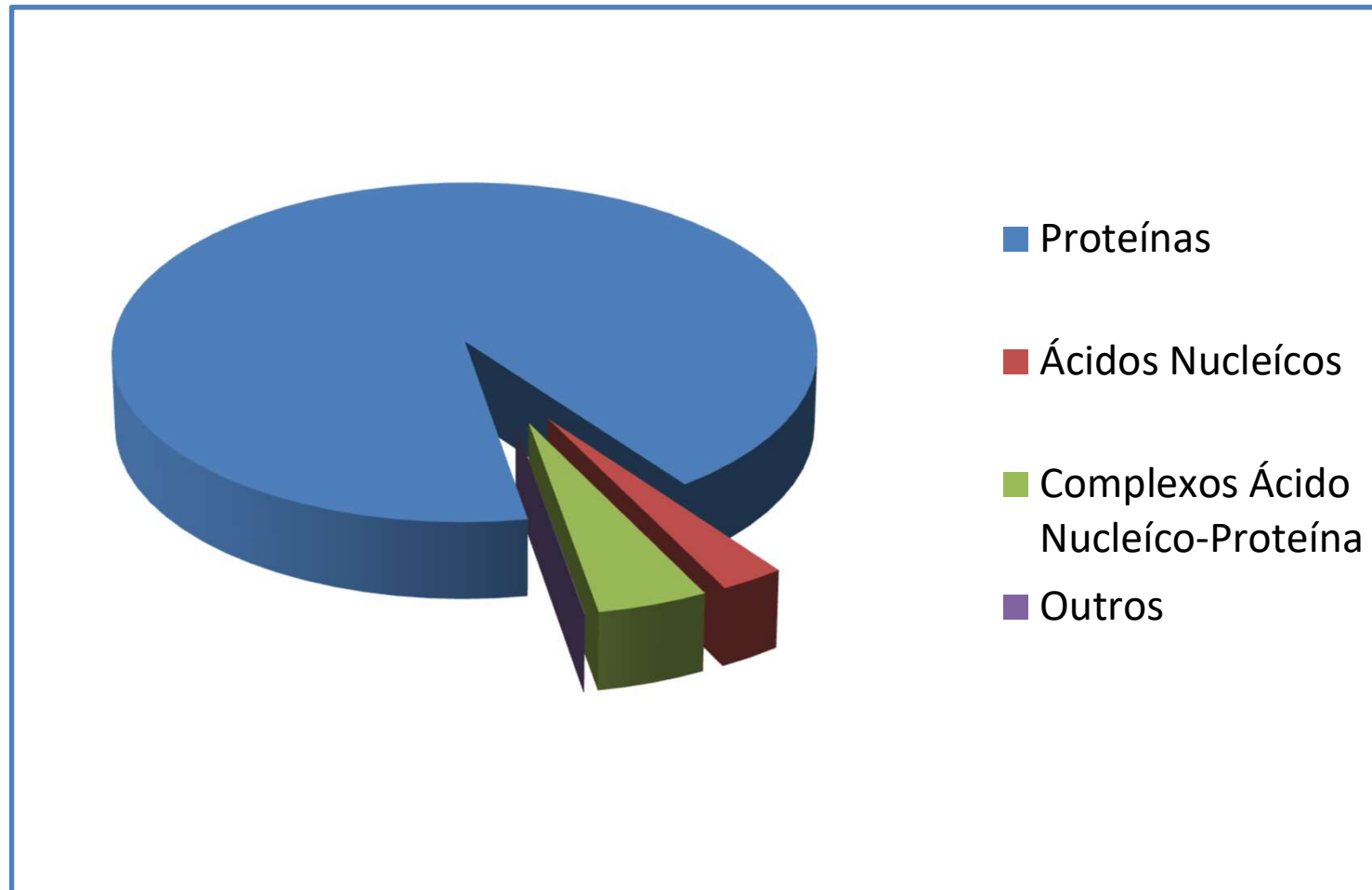
The Protein Data Bank

- O Protein Data Bank (PDB) foi criado em 1971 por E.Meyer e W.Hamilton, do Brookhaven National Laboratory (USA), contendo no início 7 estruturas!
- A gestão do PDB foi transferida em 1998 para os membros do RCSB (Research Collaboratory in Structural Bioinformatics) dos quais a Universidade de Rutgers é o site principal. O PDB (<http://www.rcsb.org>) é um banco de dados de acesso livre.
- Contendo inicialmente estruturas de proteínas, o PDB contém hoje em dia outros tipos de moléculas, tais como ácidos nucleicos, lípidos e polissacáridos.
- Número total de estruturas em 12/12/2019: **158,180**

Técnica experimental	Proteínas	Ácidos nucleicos	Complexos Ac.Nuc./Proteína	Outros	Total
Cristalografia de raios X	132004	2073	6787	8	140872
NMR	11248	1306	262	8	12824
Microscopia electrónica	2974	33	1021	0	4028
Outras	281	4	6	13	304
Combinação	144	5	2	1	152
Total	146651	3421	8078	30	158180

Dados de 12/12/2019 em <http://www.rcsb.org>

The Protein Databank contains different types of biological macromolecules

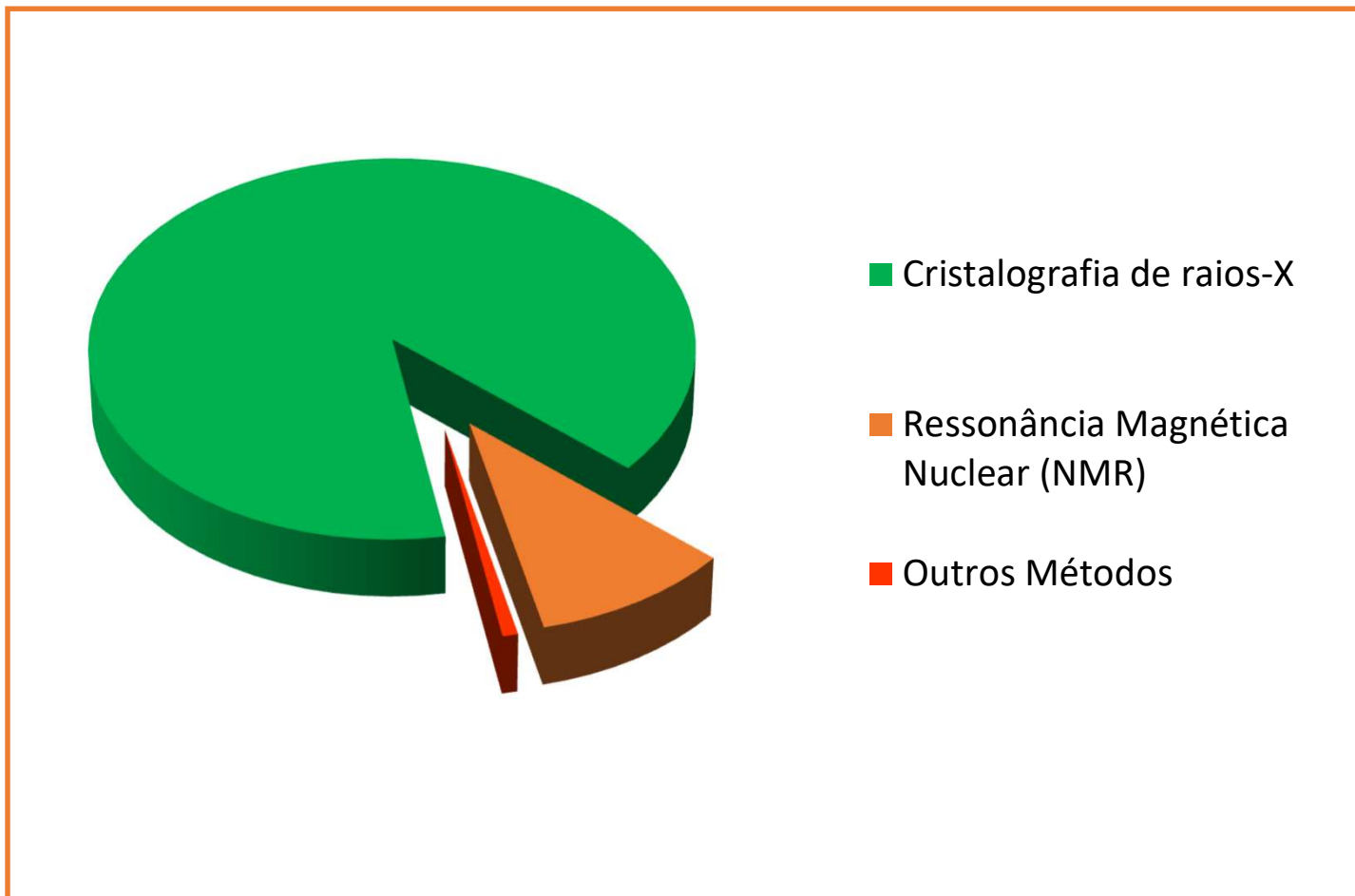


Where does all the structural information come from ?

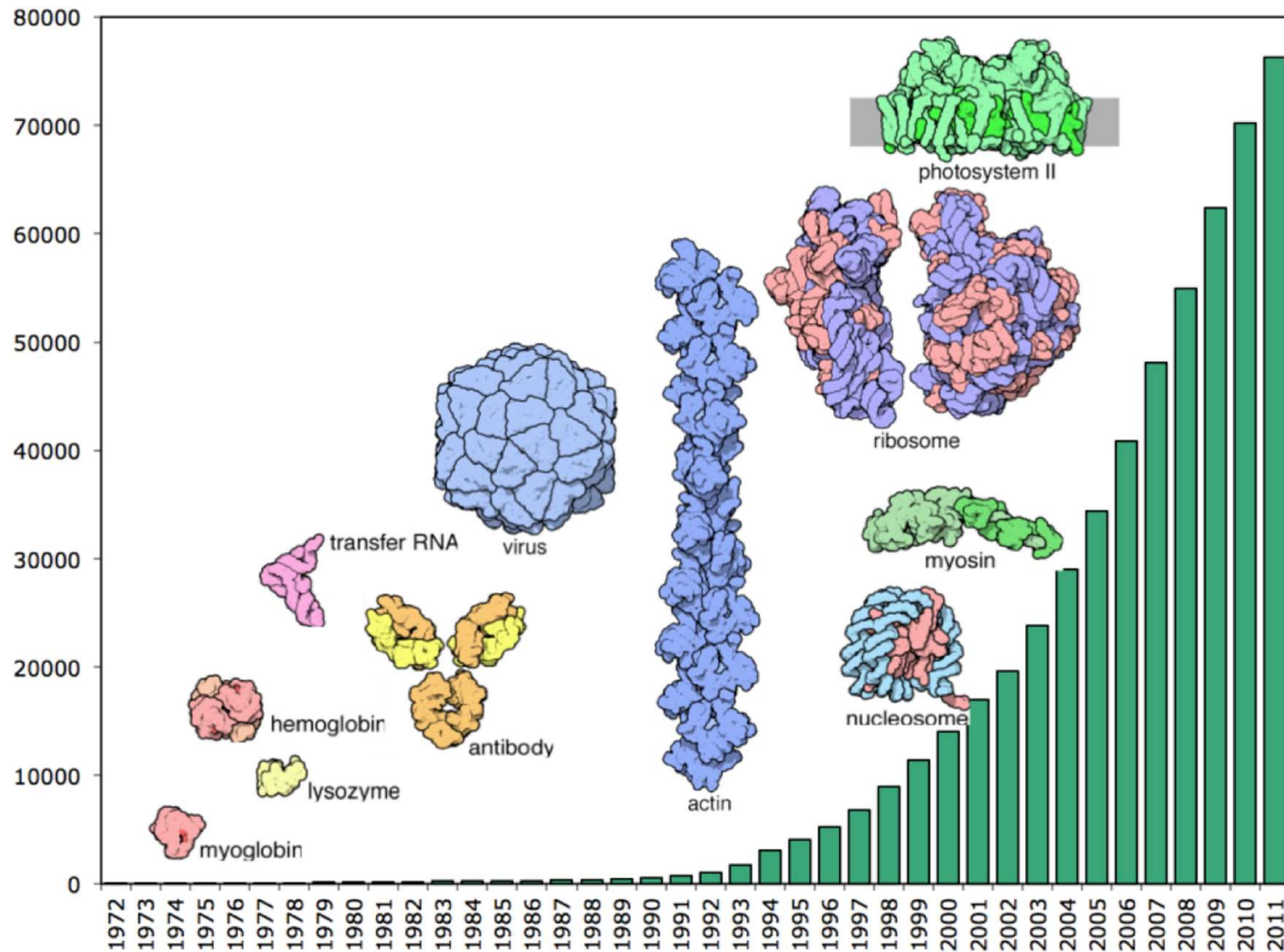
It's a combination of various types of data:

- Molecular geometry and bond theory
- Small molecule geometry
- Experimental Methods of Macromolecular structure determination
 - ❖ X-ray crystallography
 - ❖ Nuclear Magnetic Resonance (NMR)
 - ❖ Other methods (Cryo-EM, neutron diffraction, etc)

Most structures have been solved by X-ray
cristallography

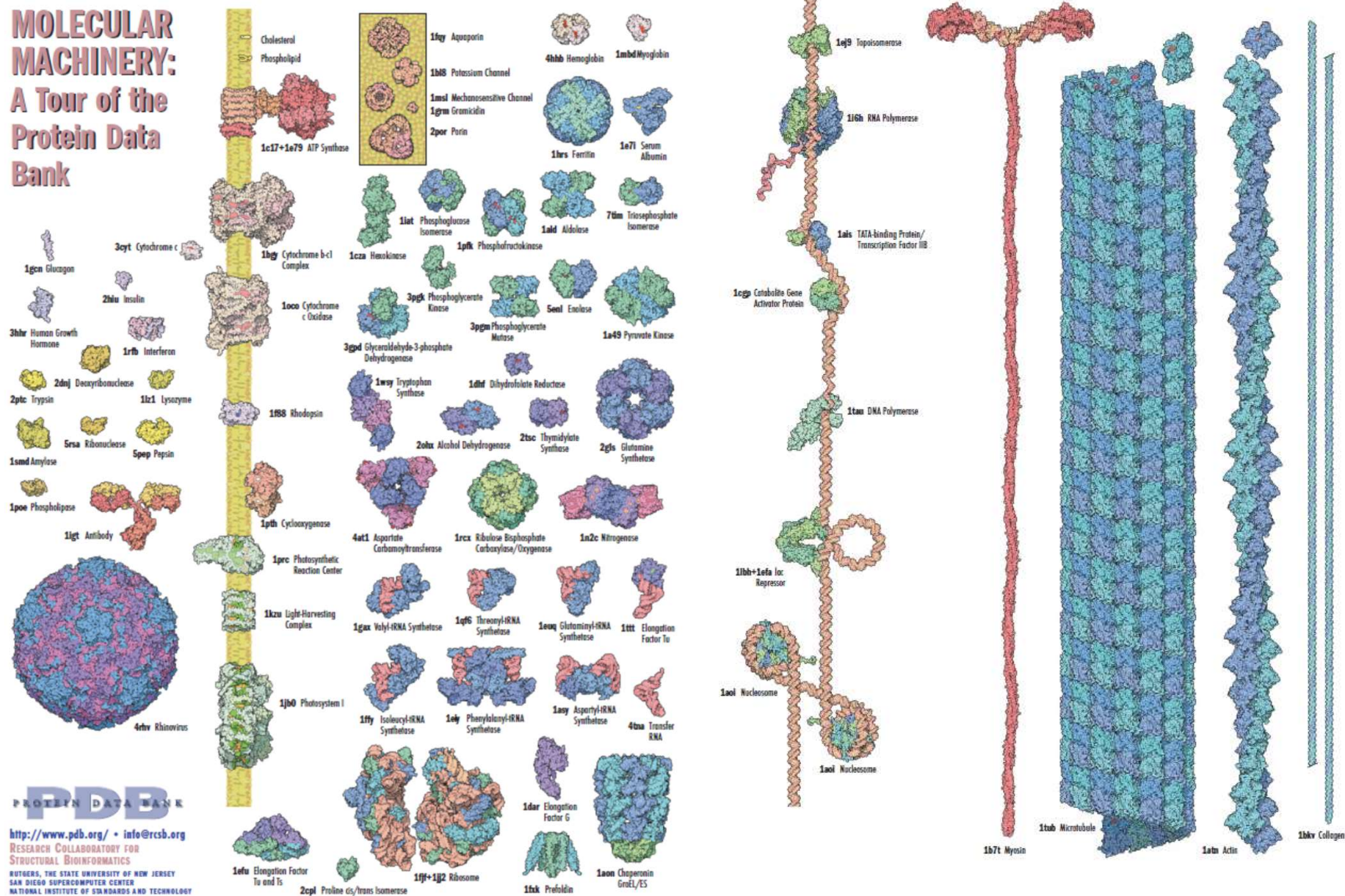


Progress in the determination of macromolecular structures



The PDB contains a very diverse set of structures !

MOLECULAR MACHINERY: A Tour of the Protein Data Bank



PDB Web Portal

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB

RCSB PDB

PROTEIN DATA BANK

158180 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

Search by PDB ID, author, macromolecule, sequence, or ligands

Go

Advanced Search | Browse by Annotations

PDB-101

Worldwide Protein Data Bank

EMDataResource

EMBL-EBI

Worldwide Protein Data Bank Foundation

f

t

w

y

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn


A Structural View of Biology

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

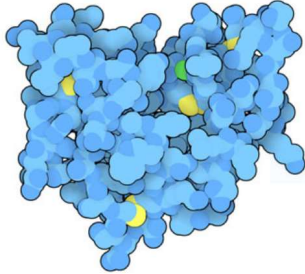
As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

Video: How Enzymes Work



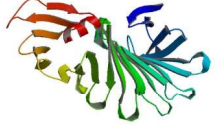
November Molecule of the Month



Phospholipase A2

Latest Entries

As of Tuesday, Nov 26 2019






6J6B

Borrelia burgdorferi OspA via surface entropy reduction (form2)

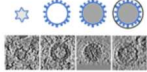
PDB Entry

Features & Highlights

-  New EM map validation in OneDep
Additional validation for electron microscopy maps helps users identify potential discrepancies.
-  Improved resolution of DOIs for PDB entries
Access new wwPDB summary pages for released PDB entries with PDB DOIs
-  Mandatory PDBx/mmCIF format files submission for MX depositions
Submission of PDBx/mmCIF format files for crystallographic depositions to the PDB will be mandatory from July 1st 2019 onward. PDB format files will no longer be accepted for deposition of structures solved by MX techniques.

News

Publications

-  Education Corner: Gaming Structural Biology for General Audiences (Part 2)
Learn about Deep Learning, Citizen Science & Puppies from Diamond Light Source's Michele Darrow » 11/26/2019
- Introducing Mol* » 11/19/2019
- New Papers on Molecular Visualization » 11/12/2019
- New EM map validation in OneDep » 11/05/2019
- Education Corner: Gaming Structural Biology for General Audiences (Part 1) » 10/29/2019
- PDB Turns 48 » 10/20/2019
- Happy Birthday, Irving Geis » 10/18/2019
- From the Bench to Molecule of the Month » 10/15/2019

PDB at a Glance 48974 Distinct Protein Sequences | 44467 Structures of Human Sequences | 11504 Nucleic Acid Containing Structures | More Statistics

<https://www.rcsb.org>

PDB Web Portal

Paulo

RCSB PDB: Homepage x

Secure | <https://www.rcsb.org>

Apps Acad AWS Free Cell Bits PDF Calendario 2017-18 VTC NiceHash LuckPool Supernova WTM WhatToMine D pmartel Other bookmarks

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB

RCSB PDB 137692 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

PDB-101 **WORLDWIDE PDB** **EMDataBank** **NUCLEIC ACID DATABASE** **Worldwide Protein Data Bank Foundation**

Janela de pesquisa

Search by PDB ID, author, macromolecule, sequence, or ligand **Go**

Advanced Search | Browse by Annotations

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

A Structural View of Biology

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

New Video: What is a Protein?

February Molecule of the Month

EPSP Synthase and Weedkillers

Latest Entries As of Tuesday Feb 13 2018

Features & Highlights

News Publications

Meet RCSB PDB at AAAS

Learn how RCSB PDB is Sustaining A Living Digital

PDB Web Portal

RCSB PDB - Query Results

www.pdb.org/pdb/results/results.do?tabtoShow=Current&qrid=B13E001B

Gnuplot surprisi... Gnuplot tricks N Old Nabble - G... Music Production Bioinformatics 10.10.20.1/form... + Pocket Other bookmarks

Search | All Categories: lysozyme

Query Parameters: Text Search for: lysozyme

Other search suggestions:

Query Refinements: Select an item or pie chart

Organism Taxonomy Exp. Method X-ray Resolution Release Date Polymer Type

Enzyme Classification SCOP Classification

Refine Query with Advanced Search Remove Similar: Select Percent Similarity

5 Related Molecule of the Month articles

PDB Pioneers Nanobodies Adrenergic Receptors Antibodies Lysozyme

Showing 1 - 25 of 1346 Results Results: 25 Page: 1 of 54

Filter: Check All View: Detailed Reports: Select one.. Sort: Relevance

1JKB HUMAN LYSOZYME MUTANT WITH GLU 35 REPLACED BY ALA

Authors: Muraki, M., Harata, K., Goda, S., Nagahora, H.

Release: 1997-05-15 Classification: Lysozyme

Experiment: X-RAY DIFFRACTION with resolution of 1.66 Å Residue Count: 130

Compound: 1 Polymer [Display Full Polymer Details | Display for All Results]
1 Ligand [Display Full Ligand Details | Display for All Results]

Citation: Importance of van der Waals contact between Glu 35 and Trp 109 to the catalytic action of human lysozyme. (1997) Protein Sci. 6: 473-476 [Display Full Abstract | Display for All Results]

Molecule of the Month: Nanobodies, PDB Pioneers, Antibodies, Lysozyme

Search Hit: Classification: LYSOZYME

1JKC HUMAN LYSOZYME MUTANT WITH TRP 109 REPLACED BY PHE

Authors: Muraki, M., Harata, K., Goda, S., Nagahora, H.

Release: 1997-05-15 Classification: Lysozyme

PDB Web Portal

RCSB Protein Data Bank

www.pdb.org/pdb/explore/explore.do?structureId=1JKB

Gnuplot surprisi... Gnuplot tricks Old Nabble - G... Music Production Bioinformatics 10.10.20.1/form... + Pocket Other bookmarks

PDB-101 Hide

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

MyPDB Hide

Login to your Account
Register a New Account
Query Results (1346)
Query History (1)

Home Hide

News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers
External Links
Sitemap
New Website Features

Deposition Hide

All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamlines/Facilities
Related Tools

Tools Hide

Download Files
Compare Structures
File Formats
RESTful Web Services
Widgets

Help Hide

Launch Help System
Display Settings
Video Tutorials
Glossary of Terms
RCSB PDB Mobile

Summary Sequence Annotations Seq. Similarity 3D Similarity Literature Biol. & Chem. Methods Geometry Links

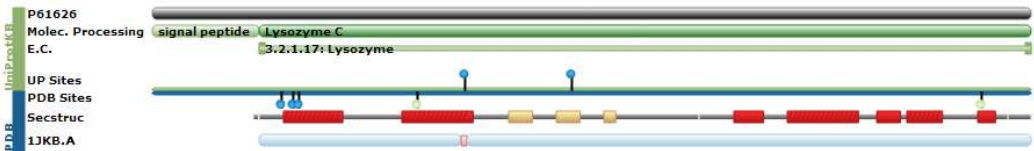
HUMAN LYSOZYME MUTANT WITH GLU 35 REPLACED BY ALA

DOI:10.2210/pdb1jkb/pdb

Primary Citation

Importance of van der Waals contact between Glu 35 and Trp 109 to the catalytic action of human lysozyme.
Muraki, M., Goda, S., Nagahora, H., Harata, K.
Journal: (1997) Protein Sci. 6: 473-476
PubMed: 9041653
PubMedCentral: PMC2143631
DOI: 10.1002/pro.5560060227
Search Related Articles in PubMed
PubMed Abstract:
The importance of van der Waals contact between Glu 35 and Trp 109 to the active-site structure and the catalytic properties of human lysozyme (HL) has been investigated by site-directed mutagenesis. The X-ray analysis of mutant HLs revealed that both...
[Read More & Search PubMed Abstracts]

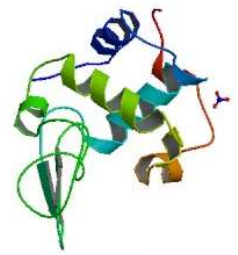
Molecular Description Hide

Classification: Lysozyme
Structure Weight: 14786.71
Molecule: LYSOZYME
Polymer: 1 **Type:** protein **Length:** 130
Chains: A
EC#: 3.2.1.17
Mutation: E35A
Organism: Homo sapiens
UniProtKB: Protein Feature View | Search PDB | P61626


Source Hide

1JKB

Display Files
Download Files
Share this Page

Biological Assembly

View in 3D More Images...
Biological assembly 1 assigned by authors
Downloadable viewers:
Simple Viewer Protein Workshop
Kiosk Viewer

MyPDB Personal Annotations Hide

To save personal annotations, please login to your MyPDB account.

Deposition Summary Hide

Authors: Muraki, M., Harata, K., Goda, S., Nagahora, H.
Deposition: 1996-11-13
Release: 1997-05-15
Last Modified (REVDAT): 2009-02-24

Revision History Hide

Mouse over text for details

Download th structure in PDB format

The screenshot shows the RCSB Protein Data Bank (PDB) website interface. The main content area displays the entry for **HUMAN LYSOZYME MUTANT WITH GLU 35 REPLACED BY ALA** (DOI: 10.2210/pdb1jkb/pdb). The left sidebar contains navigation links for PDB-101, MyPDB, Home, Deposition, Tools, and Help. The right sidebar contains links for MyPDB Personal Annotations, Deposition Summary, and Revision History. A dropdown menu is open on the right, showing various download options. The 'PDB File (Text)' option is highlighted with a red box.

FASTA Sequence
PDB File (Text)
PDB File (gz)
mmCIF File
mmCIF File (gz)
PDBML/XML File
PDBML/XML File (gz)
Structure Factor (Text)
Structure Factor (gz)
Biological Assembly (gz) (A)

Primary Citation
Importance of van der Waals contact between Glu 35 and Trp 109 to the catalytic action of human lysozyme.
Muraki, M., Goda, S., Nagahora, H., Harata, K.
Journal: (1997) Protein Sci. 6: 473-476
PubMed: 9041653
PubMedCentral: PMC2143631
DOI: 10.1002/pro.5560060227
Search Related Articles in PubMed

PubMed Abstract:
The importance of van der Waals contact between Glu 35 and Trp 109 to the active-site structure and the catalytic properties of human lysozyme (HL) has been investigated by site-directed mutagenesis. The X-ray analysis of mutant HLs revealed that both...
[Read More & Search PubMed Abstracts]

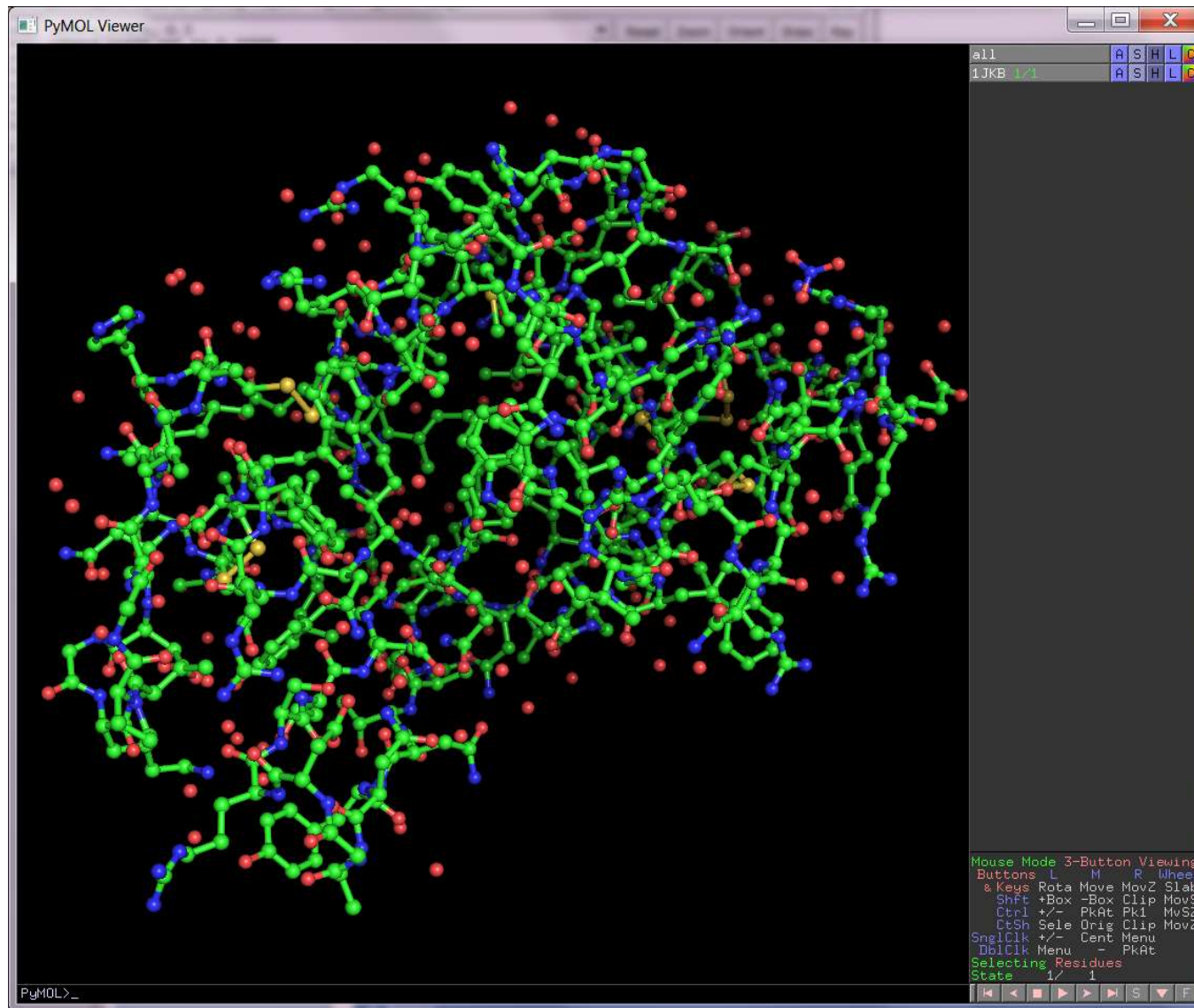
Molecular Description
Classification: Lysozyme
Structure Weight: 14786.71
Molecule: LYSOZYME
Polymer: 1 **Type:** protein **Length:** 130
Chains: A
EC#: 3.2.1.17
Mutation: E35A
Organism: Homo sapiens
UniProtKB: Protein Feature View | Search PDB | P61626

Downloadable viewers:
Simple Viewer Protein Workshop
Kiosk Viewer

Deposition Summary
Authors: Muraki, M., Harata, K., Goda, S., Nagahora, H.
Deposition: 1996-11-13
Release: 1997-05-15
Last Modified (REVDAT): 2009-02-24

Revision History
Mouse over text for details

Viewing the structure with a molecular visualizer (PyMOL)



Formatos de representação da estrutura

- A representação da estrutura molecular em bancos de dados passa pela descrição das **coordenadas atômicas**, do **tipo de átomo**, e das **ligações químicas** presentes.
- A descrição do tipo de átomos e ligações que os unem designa-se como **topologia** da molécula.
- No caso das proteínas, a topologia dos 20 aminoácidos standard pode ser assumida *a priori*, pois a estrutura dos aminoácidos é conhecida
- A topologia de outras moléculas, tais como grupos prostéticos , deverá ser especificada
- O formato “tradicional” de representação de estrutura no Protein Data Bank é o formato **PDB**.

Formato da informação no Protein Data Bank

- A informação contida no Protein Databank inclui coordenadas atómicas, topologias de ligação (descrição das ligações químicas), nomes dos átomos e grupos químicos, dados associados ao processo de determinação experimental da estruturas e outras informações sobre a função, ligandos, propriedades, etc...
- Presentemente a informação no PDB está disponível nos seguintes formatos:
 - **pdb file:** O formato “flat file”, um tipo de ficheiro chamado “ficheiro PDB”. Estes ficheiros são os mais utilizados pelos softwares de manipulação e visualização de estruturas e têm geralmente a extensão “.pdb”
 - **mmCIF:** - um formato mais poderoso e estruturado que o ficheiro PDB, ainda não tendo sido largamente adoptado
 - **XML:** - extended mark-up language, um formato muito geral de representação de informação, compatível com um vasto número de aplicações de software.
 - **mmtf:** formato binário altamente compacto, ilegível para humanos, mas ocupando muito menos espaço e oferecendo rápida transmissão, leitura e escrita.

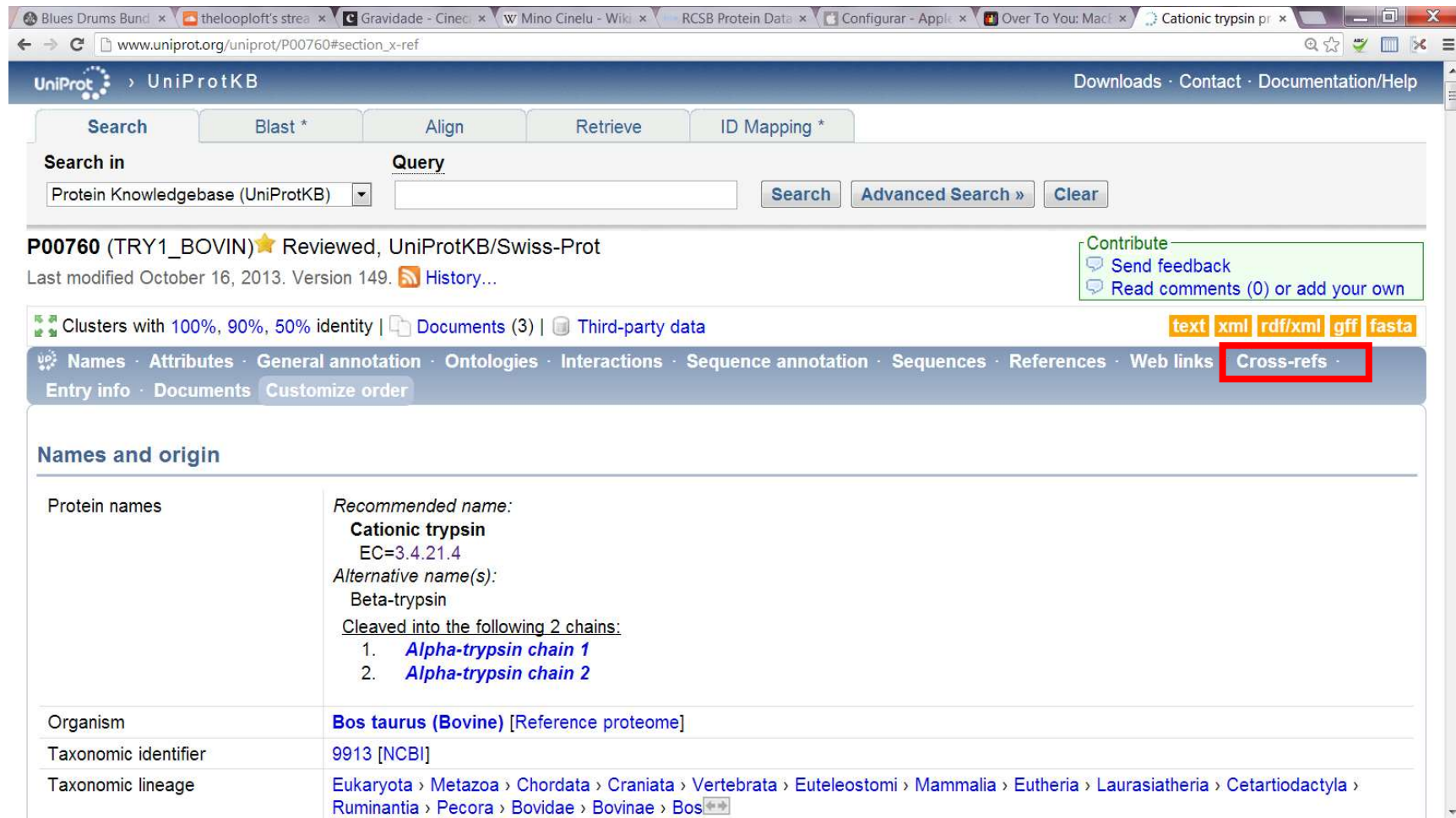
Formato do ficheiro PDB

```
HEADER      METAL BINDING PROTEIN                      21-AUG-03   1Q8H
TITLE       CRYSTAL STRUCTURE OF PORCINE OSTEOCALCIN
COMPND      MOL_ID: 1;
COMPND      2 MOLECULE: OSTEOCALCIN;
COMPND      3 CHAIN: A
SOURCE      MOL_ID: 1;
SOURCE      2 ORGANISM_SCIENTIFIC: SUS SCROFA;
SOURCE      3 ORGANISM_COMMON: PIG
KEYWDS      HELIX-TURN-HELIX-TURN-HELIX, PAPER-CLIP, HYDROXYAPATITE
KEYWDS      2 CRYSTAL SURFACE BINDING PROTEIN, CALCIUM BINDING PROTEIN,
KEYWDS      3 BONE GLA PROTEIN
EXPDTA      X-RAY DIFFRACTION
AUTHOR      Q.Q.HOANG,F.SICHERI,A.J.HOWARD,D.S.YANG
REVDAT      1 11-NOV-03 1Q8H 0
JRNL        AUTH  Q.Q.HOANG,F.SICHERI,A.J.HOWARD,D.S.YANG
JRNL        TITL   BONE RECOGNITION MECHANISM OF PORCINE OSTEOCALCIN
JRNL        TITL 2 FROM CRYSTAL STRUCTURE.
JRNL        REF    NATURE                      V. 425   977 2003
JRNL        REFN   ASTM NATUAS   UK ISSN 0028-0836
REMARK      1
REMARK      2
REMARK      2 RESOLUTION. 2.00 ANGSTROMS.
REMARK      3
REMARK      3 REFINEMENT.
REMARK      3   PROGRAM      : CNS 1.1
REMARK      3   AUTHORS      : BRUNGER, ADAMS, CLORE, DELANO, GROS, GROSSE-
.....
ATOM        1  N   PRO A  13      10.210  29.966  44.935  1.00 38.06
ATOM        2  CA  PRO A  13       9.718  29.013  43.919  1.00 37.33
ATOM        3  C   PRO A  13       9.566  29.662  42.541  1.00 37.52
ATOM        4  O   PRO A  13       9.275  30.855  42.444  1.00 38.00
ATOM        5  CB  PRO A  13       8.383  28.488  44.434  1.00 37.68
ATOM        6  CG  PRO A  13       7.919  29.624  45.336  1.00 36.60
ATOM        7  CD  PRO A  13       9.196  30.126  45.995  1.00 36.47
ATOM        8  N   ASP A  14       9.777  28.879  41.483  1.00 36.83
ATOM        9  CA  ASP A  14       9.671  29.384  40.116  1.00 36.13
.....
MASTER      299    0    6    3    0    0    0    6  378    1  38    4
END
```

Cabeçalho

Coordenadas

Interligação entre Uniprot e PDB



The screenshot shows the UniProtKB entry for P00760 (TRY1_BOVIN). The 'Cross-refs' tab is highlighted in red. The entry is a reviewed protein from the Protein Knowledgebase (UniProtKB/Swiss-Prot), last modified on October 16, 2013, version 149. The protein is Cationic trypsin, EC=3.4.21.4, and is cleaved into two chains: Alpha-trypsin chain 1 and Alpha-trypsin chain 2. The organism is Bos taurus (Bovine) [Reference proteome]. The taxonomic lineage is Eukaryota > Metazoa > Chordata > Craniata > Vertebrata > Euteleostomi > Mammalia > Eutheria > Laurasiatheria > Cetartiodactyla > Ruminantia > Pecora > Bovidae > Bovinae > Bos.

Blues Drums Bund x thelooploft's stre x Gravidade - Cinec x Mino Cinelu - Wiki x RCSB Protein Data x Configurar - Appl x Over To You: Mac x Cationic trypsin pr x

www.uniprot.org/uniprot/P00760#section_x-ref

UniProt > UniProtKB Downloads · Contact · Documentation/Help

Search Blast * Align Retrieve ID Mapping *

Search in Protein Knowledgebase (UniProtKB) Query Search Advanced Search » Clear

P00760 (TRY1_BOVIN) ★ Reviewed, UniProtKB/Swiss-Prot
Last modified October 16, 2013. Version 149. History...

Contribute
Send feedback
Read comments (0) or add your own

Clusters with 100%, 90%, 50% identity | Documents (3) | Third-party data text xml rdf/xml gff fasta

Names · Attributes · General annotation · Ontologies · Interactions · Sequence annotation · Sequences · References · Web links Cross-refs ·

Entry info · Documents Customize order

Names and origin

Protein names	<i>Recommended name:</i> Cationic trypsin EC=3.4.21.4 <i>Alternative name(s):</i> Beta-trypsin <i>Cleaved into the following 2 chains:</i> 1. Alpha-trypsin chain 1 2. Alpha-trypsin chain 2
Organism	Bos taurus (Bovine) [Reference proteome]
Taxonomic identifier	9913 [NCBI]
Taxonomic lineage	Eukaryota > Metazoa > Chordata > Craniata > Vertebrata > Euteleostomi > Mammalia > Eutheria > Laurasiatheria > Cetartiodactyla > Ruminantia > Pecora > Bovidae > Bovinae > Bos

Interligação entre Uniprot e PDB

Blues Drums Bundle x thelooploft's stream x W Mino Cinelu - Wikip x RCSB Protein Data B x Configurar - Apple S x Over To You: MacBoi x Cationic trypsin prec x

www.uniprot.org/uniprot/P00760#section_x-ref

Names · Attributes · General annotation · Ontologies · Interactions · Sequence annotation · Sequences · References · Web links · Cross-refs · Entry info · Documents Customize order

Cross-references

Sequence databases

<input checked="" type="radio"/> EMBL	BC134797 mRNA. Translation: AAI34798.1 .
<input type="radio"/> GenBank	BC146041 mRNA. Translation: AAI46042.1 .
<input type="radio"/> DDBJ	D38507 mRNA. Translation: BAA07516.1 .
IPI	IPI00706427 .
PIR	TRBOTR . A90164.
RefSeq	NP_001107199.1 . NM_001113727.1 .
UniGene	Bt.91423 .

3D structure databases

<input checked="" type="radio"/> PDBe	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
<input type="radio"/> RCSB PDB	1AQ7	X-ray	2.20	A	24-246	[>]
<input type="radio"/> PDBj	1AUJ	X-ray	2.10	A	24-246	[>]
	1AZ8	X-ray	1.80	A	24-246	[>]
	1BJU	X-ray	1.80	A	24-246	[>]
	1BJV	X-ray	1.80	A	24-246	[>]
	1BTP	X-ray	2.20	A	18-246	[>]
	1BTW	X-ray	1.70	A	18-246	[>]
	1BTX	X-ray	1.70	A	18-246	[>]
	1BTY	X-ray	1.50	A	18-246	[>]
	1BTZ	X-ray	2.00	A	18-246	[>]

Blues Drums Bundle x thelooploft's stream x Mino Cinelu - Wikipe x RCSB Protein Data B x Configurar - Apple S x Over To You: MacBoi x RCSB Protein Data B x

www.pdb.org/pdb/explore/explore.do?structureId=1AQ7

RCSB PDB PROTEIN DATA BANK

A MEMBER OF THE **PDB** | **EMDataBank**

An Information Portal to Biological Macromolecular Structures

As of Tuesday Oct 29, 2013 at 5 PM PDT there are 95113 Structures | [PDB Statistics](#) | [Email](#) | [RSS](#) | [Help](#)

Search **Advanced** **Browse**

Everything Author Macromolecule Sequence Ligand ?

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

[Search History](#), [Previous Results](#)

PDB-101 **Hide**

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

MyPDB **Hide**

Login to your Account
Register a New Account
MyPDB Help Page

Home **Hide**

News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers
External Links
Sitemap
New Website Features

Summary 3D View Sequence Annotations Seq. Similarity 3D Similarity Literature Biol. & Chem. Methods Geometry Links

TRYPSIN WITH INHIBITOR AERUGINOSIN 98-B

1AQ7

[Display Files](#) [Download Files](#) [Share this Page](#)

DOI:10.2210/pdb1aq7/pdb

Primary Citation

Atomic Structure of the Trypsin-Aeruginosin 98-B Complex

Sandler, B., Murakami, M., Clardy, J.

Journal: (1998) J.Am.Chem.Soc. **120**: 595-596

PubMed ID is not available

Molecular Description **Hide**

Classification: Hydrolase/hydrolase Inhibitor

Structure Weight: 23979.18

Molecule:	TRYPSIN	Type:	protein	Length:	223
Polymer:	1				
Chains:	A				
EC#:	3.4.21.4				

Biological Assembly ?

