

How to use the Protein Data Bank (PDB)

Loren Williams
Georgia Tech
(revised 1/23/2023)

Biologists and biochemists use sequence databases, structure databases, literature databases, etc. The database we will learn here is called the Protein Database (PDB). The PDB has all known 3D structures of proteins, DNAs and RNAs. To find the PDB on the web, type 'PDB' into google, and go to the first link returned, which is:

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

You need to either fetch or download a structure (i.e., the PDB coordinates) of the protein that you intend to studying.

Each structure is contained in a pdb entry specified by a unique entry code that does not carry any information (for example 1H97.pdb). A PDB file is a text file with the xyz coordinates of all the atoms in the protein (one protein has lots and lots of atoms).

Example of two lines of a pdb file

ATOM	1	N	ALA	A	1	5.089	4.202	28.188	1.00	42.31	N
ATOM	2	CA	ALA	A	1	4.695	2.911	28.829	1.00	41.76	C

atom type

How to use the PDB.

aa

chain

occupancy

Ask Questions:

How many structures from the virus SARS-CoV-2 have been deposited in the pdb over the last year (as of 1/23/2023) with resolution better than (i.e., less than) 2.2 Å?

Start by Googling PDB

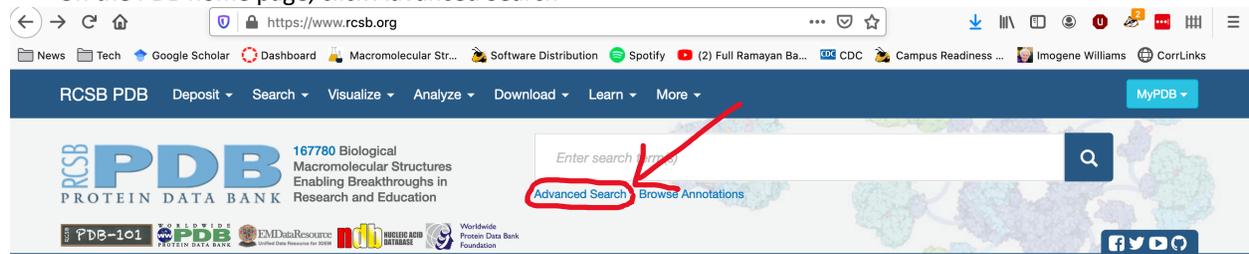
The first hit

www.rcsb.org

[RCSB PDB: Homepage](#)

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data according to agreed upon standards. The RCSB PDB also provides a variety of ...

On the PDB home page, click Advanced Search



- Welcome
- Deposit
- Search
- Visualize

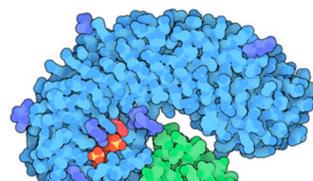
A Structural View of Biology

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.

August Molecule of the Month



Contact Us

Search Query History Browse Annotations MyPDB

QUERY: Structure Keywords HAS ANY OF WORDS "SARS-Cov-2"

JSON MyPDB Login

Advanced Search Query Builder Help

- Full Text
- Structure Attributes
- Chemical Attributes
- Sequence Similarity

Include the quotes

QUERY: Structure Keywords HAS ANY OF WORDS "SARS-Cov-2"

JSON MyPDB Login

Advanced Search Query Builder Help

Full Text

"SARS-Cov-2" Count

Add Term Add Subquery Remove Subquery

Add Subquery

QUERY: Structure Keywords HAS ANY OF WORDS "SARS-Cov-2"

JSON MyPDB Login

Advanced Search Query Builder Help

Full Text

"SARS-Cov-2" Count

Add Term Add Subquery Remove Subquery

Add Subquery

Structure Attributes

AND

-- Type to filter and/or select an attribute --

ID(s) and Keywords

Structure Details

Computed Structure Model Details

Chemical Attributes

Entry Features

Advanced Search Query Builder Help

Full Text

"SARS-Cov-2" Count

Add Term Add Subquery Remove Subquery

Add Subquery

Structure Attributes

AND

-- Type to filter and/or select an attribute --

ID(s) and Keywords

Structure Details

Structure Title

Chemical Attributes

Structure Author

Structure Author ORCID

Sequence Similarity

Deposit Date

Sequence Motif

Release Date

Structure Similarity

Revision Date

Structure Motif

Structural Genomics Project Name

Structural Genomics Project Center Initials

Advanced Search Query Builder

Help

Full Text: "SARS-Cov-2" Count
Add Term Add Subquery Remove Subquery
Add Subquery

Structure Attributes: Deposit Date x > Enter a date between 8/10/1972 and 1/8/2023 + NOT Count x
Add Attribute Add Subquery Remove Subquery
Add Subquery

Chemical Attributes
Sequence Similarity
Sequence Motif



Advanced Search Query Builder

Help

Full Text: "SARS-Cov-2" Count
Add Term Add Subquery Remove Subquery
Add Subquery

Structure Attributes: Deposit Date x > 01/23/2022 + NOT Count x
Add Attribute Add Subquery Remove Subquery
Add Subquery



Advanced Search Query Builder

Help

Full Text: "SARS-Cov-2" Count
Add Term Add Subquery Remove Subquery
Add Subquery

Structure Attributes: Deposit Date x > 01/23/2022 + NOT Count x
AND -- Type to filter and/or select an attribute --
AND / OR ID(s) and Keywords
Structure Details
Computed Structure Model Details
Chemical Attributes Entry Features
Polymer Molecular Features
Sequence Similarity Polymer Instance (Chain) Features
Sequence Motif Nonpolymer Molecular Features
Structure Similarity Oligosaccharide/Branched Molecular Features
Assembly Features
Structure Motif Methods
Chemical Similarity X-ray Method Details
X-ray Data Collection Details
Cell Dimensions and Space Group

Return Structures Include Computed Structure Models (CSM) Count Clear Search



Advanced Search Query Builder

Help

Full Text: "SARS-Cov-2" Count
Add Term Add Subquery Remove Subquery
Add Subquery

Structure Attributes: Deposit Date x > 01/23/2022 + NOT Count x
AND Data Collection Resolution x < 2.2 + NOT Count x
AND / OR Add Attribute Add Subquery Remove Subquery
Add Subquery



Advanced Search Query Builder Help

Full Text ?

"SARS-Cov-2" Count

Add Term Add Subquery Remove Subquery

Add Subquery

Structure Attributes Help

AND

Deposit Date	x	>	01/23/2022	+ NOT	Count	x
AND Data Collection Resolution	x	<	2.2	+ NOT	Count	x

AND / OR Add Attribute Add Subquery Remove Subquery

Add Subquery

Chemical Attributes ?

Sequence Similarity ?

Sequence Motif ?

Structure Similarity ?

Structure Motif ?

Chemical Similarity ?

Return Structures ? grouped by No Grouping ?

Include Computed Structure Models (CSM) ?

361 Clear Search

361 structures with resolution better than (i.e., less than) 2.2 Å related to SARS-Cov-2 were deposited between Jan 23, 2022 and Jan 23, 2023.

They are listed for you.

Refinements

Structure Determination Methodology

- experimental (361)

Scientific Name of Source Organism

- Severe acute respiratory syndrome coronavirus 2 (348)
- Homo sapiens (23)
- Lama glama (5)
- Vicugna pacos (3)
- synthetic construct (3)
- Enterovirus A71 (2)
- Severe acute respiratory syndrome coronavirus (2)
- Escherichia coli (1)
- Macaca mulatta (1)
- Middle East respiratory syndrome-related coronavirus (1)
- [More...](#)

Taxonomy

- Riboviria (354)
- Eukaryota (31)
- other sequences (3)
- Bacteria (1)

Experimental Method

- X-RAY DIFFRACTION (361)
- NEUTRON DIFFRACTION (3)

Polymer Entity Type

- Protein (361)

Refinement Resolution (Å)

- 0.5 - 1.0 (16)
- 1.0 - 1.5 (193)
- 1.5 - 2.0 (104)
- 2.0 - 2.5 (50)

Release Date

- 2020 - 2024 (361)

Enzyme Classification Name

- Hydrolases (330)
- Transferases (312)
- Lyases (312)

Membrane Protein Annotation

- mpstruc (1)

1 to 25 of 361 Structures

Page 1 of 15

Sort by | Score

7U09 ← entry code

Crystal Structure of C13B8 Fab in complex with SARS-CoV-2 S fusion peptide

Tortorici, M.A., Veesler, D.

(2022) Science **377**: 735-742

Released 2022-08-03

Method X-RAY DIFFRACTION 2.1 Å ← resolution

Organisms Homo sapiens
Severe acute respiratory syndrome coronavirus 2

Macromolecule Heavy chain Fab C13B8 (protein)
Light chain Fab C13B8 (protein)
SARS-CoV-2 S fusion peptide (protein)

7U0E

Crystal Structure of C13C9 Fab in complex with SARS-CoV-2 S fusion peptide

Tortorici, M.A., Veesler, D.

(2022) Science **377**: 735-742

Released 2022-08-03

Method X-RAY DIFFRACTION 2.1 Å

Organisms Homo sapiens
Severe acute respiratory syndrome coronavirus 2

Macromolecule Heavy chain Fab C13C9 (protein)
Light chain Fab C13C9 (protein)
SARS-CoV-2 S fusion peptide (protein)

7U0A

Crystal Structure of C77G12 Fab in complex with SARS-CoV-2 S fusion peptide

Tortorici, M.A., Veesler, D.

(2022) Science **377**: 735-742

Released 2022-08-03

Method X-RAY DIFFRACTION 1.7 Å

Organisms Homo sapiens
Severe acute respiratory syndrome coronavirus 2

Macromolecule Heavy chain Fab C77G12 (protein)
Light chain Fab C77G12 (protein)
SARS-CoV-2 S fusion peptide (protein)

Unique Ligands CL

7Y9N

an engineered 5-helix bundle derived from SARS-CoV-2 S2 in complex with HR2P