



VIRTUAL WEBINAR

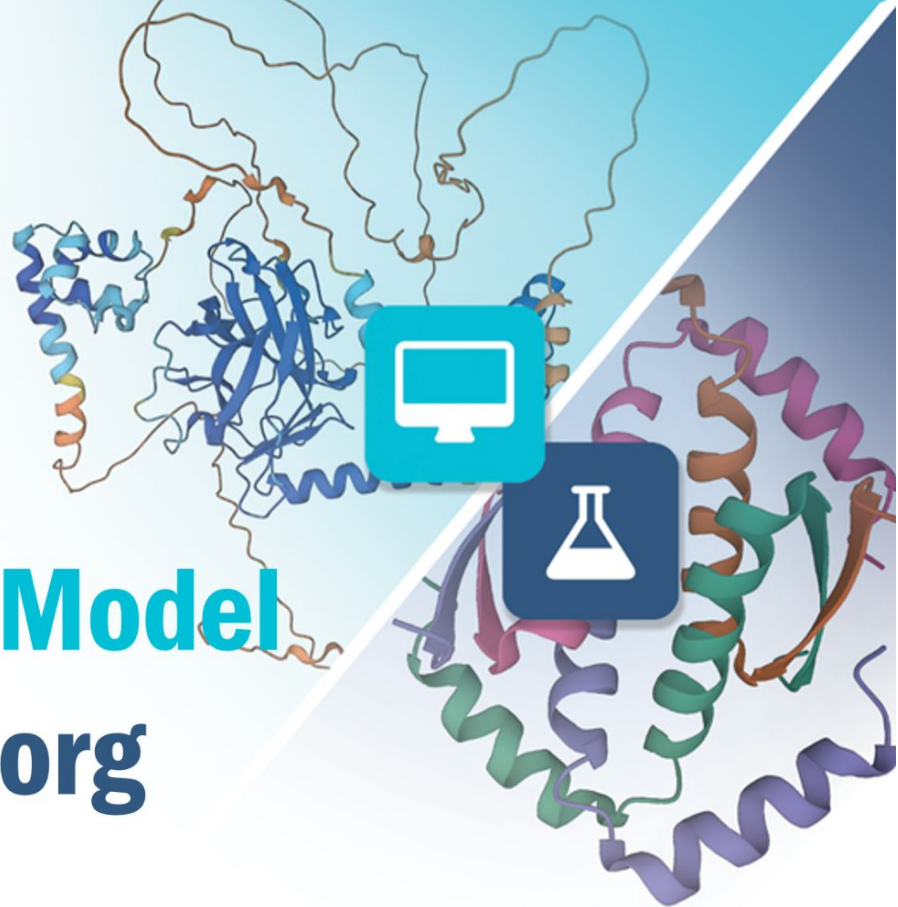
A Deep Dive into

Computed Structure Model

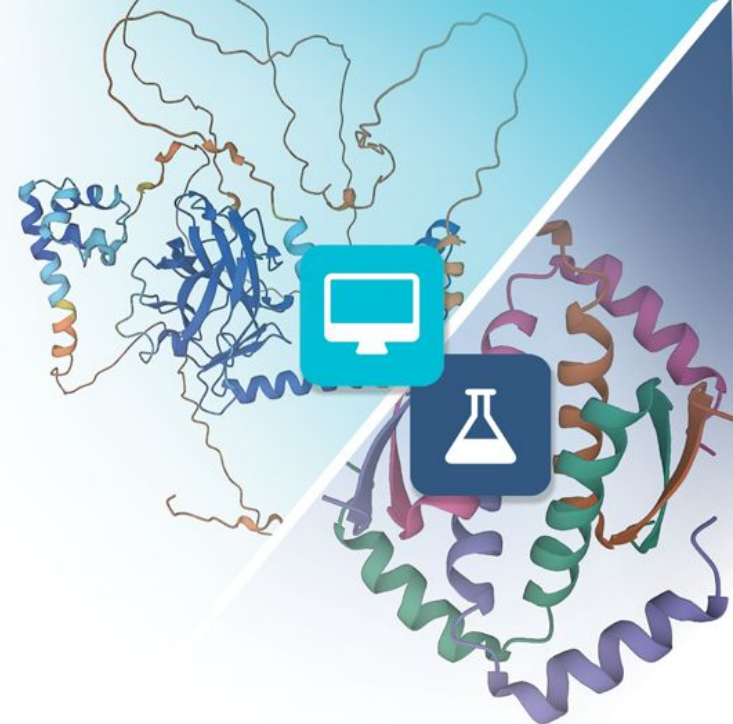
Exploration at [RCSB.org](https://www.rcsb.org)

Tuesday April 30th 2024

9-10am Pacific | 12-1pm Eastern



Welcome



Stephen K. Burley, M.D., D.Phil.

stephen.burley@rcsb.org

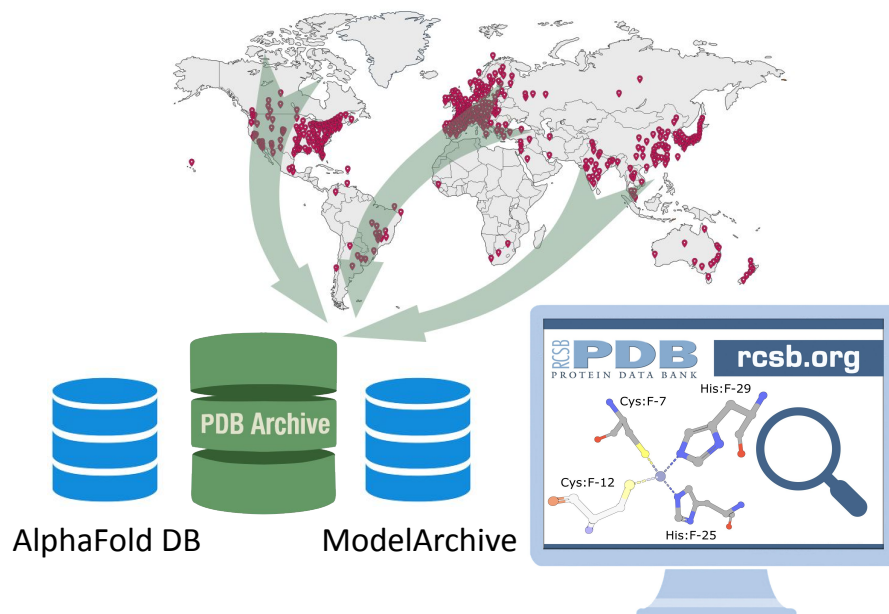
Agenda

9:00 - 9:05 AM PDT 12:00 - 12:05 PM EDT	Welcome and Introduction	Stephen K. Burley, Director
9:05 - 9:25 AM PDT 12:05- 12:25 PM EDT	A case study of Low-density Lipoprotein Receptor Adapter Protein 1 (LDLRAP1)	Yana Rose, Scientific Software Developer & Data Architect
9:25 - 09:45 AM PDT 12:25 - 12:45 PM EDT	Case study of Class II aminoacyl-tRNA synthetases	Sebastian Bittrich, Scientific Software Developer
09:45 - 10:00 AM PDT 12:45 - 1:00 PM EDT	Q&A	Jose Duarte, Brinda Vallat, Yana Rose, Sebastian Bittrich, Joan Segura

RCSB.org: One-Stop-Shop for Public 3D Biostructure Data

- **RCSB.org:** Tools for searching, accessing, visualizing, analyzing, and downloading data
 - Open access to ~220,000 experimental structures of macromolecules
 - >1 million Computed Structure Models (CSMs) predicted using AI/ML methods
- Provenance/reliability of both data types clearly identified

3D structural data from around the world



RCSB.org: Opt-in to include CSMs

RCSB PDB Deposit Search Visualize Analyze Download Learn About Documentation Careers COVID-19 MyPDB Contact us

RCSB PDB PROTEIN DATA BANK

218,500 Structures from the PDB

1,068,577 Computed Structure Models (CSM)

3D Structures Enter search term(s), Entry ID(s), or sequence

Include CSM

Advanced Search | Browse Annotations Help

PDB-101 PDB EMDataResource NAKB wwPDB Foundation PDB-Dev

- Welcome
- Deposit
- Search
- Visualize
- Analyze
- Download
- Learn

RCSB Protein Data Bank (RCSB PDB) enables breakthroughs in science and education by providing access and tools for exploration, visualization, and analysis of:

- Experimentally-determined 3D structures from the **Protein Data Bank (PDB)** archive
- Computed Structure Models (CSM)** from AlphaFold DB and ModelArchive

These data can be explored in context of external annotations providing a structural view of biology.

Explore NEW Features

PDB-101 Training Resources

April Molecule of the Month

PhiX174 YES Complex

Case Study of Low-density Lipoprotein Receptor Adapter Protein 1 (LDLRAP1)



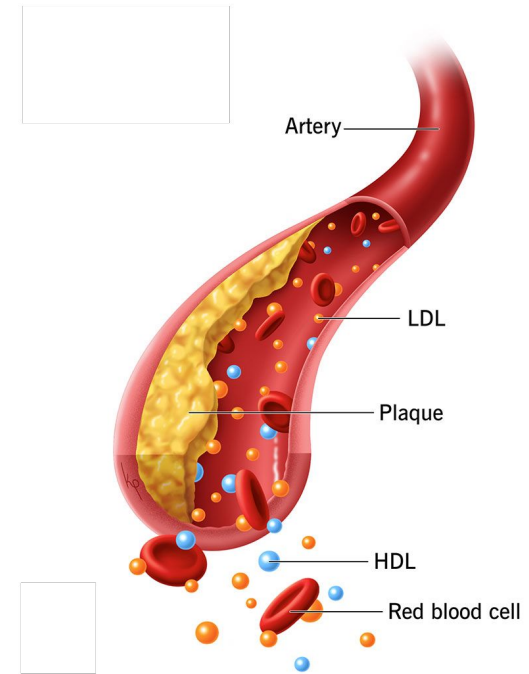
Yana Rose, PhD

yana.rose@rcsb.org

Low-density lipoprotein receptor adapter protein 1 (LDLRAP1)

- LDLRAP1 protein helps remove low-density lipoprotein (LDL) “bad” cholesterol from the bloodstream
- Mutations in the *LDLRAP1* gene has been shown to cause a form of familial hypercholesterolemia due to high levels of LDL
- Familial hypercholesterolemia can cause heart attacks at an early age

Familial Hypercholesterolemia



The structural coverage of the top 5 proteins causing familial hypercholesterolemia

Gene	UniProt ID	Protein	Protein length (residues, n)	Experimental coverage (residues, n)	Experimental coverage (residues, %)	AlphaFold (pLDDT \geq 70) coverage (residues, n)	AlphaFold (pLDDT \geq 70) coverage (residues, %)
LDLRAP1	Q5SW96	Low density lipoprotein receptor adapter protein 1	308	16	5.2	159	51.6
LDLR	P01130	Low-density lipoprotein receptor	860	705	82.0	643	74.8
APOB	P04114	Apolipoprotein B-100	4563	N/A	N/A	-	-
APOE	P02649	Apolipoprotein E	317	298	94.0	218	68.8
PCSK9	Q8NBP7	Proprotein convertase subtilisin/kexin type 9	692	642	92.8	563	81.4

Live demo

[RCSB.org](https://www.rcsb.org)

Browse disease annotations on RCSB.org

Search Query History Browse Annotations MyPDB

ATC Biological Process CARD CATH Cellular Component **Disease Ontology** ECOD Enzyme Classification Genome Location MeSH Molecular Function mpstruc OPM Protein Symmetry
SCOP-e SCOP2 Source Organism

Disease Ontology Browser

Help

The Disease Ontology browser organizes protein sequences by [Mondo Disease Ontology](#) terms describing disease definitions developed by the Monarch Initiative team. The definitions have been mapped to UniProt sequences by the Knowledge Management Center (KMC) for the Illuminating the Druggable Genome (IDG) program and available for PDB entities mapped to UniProt.

Start typing the disease name and select the suggested term

Click to search

Familial Hyperc
familial hypercholesterolemia MONDO:0005439
disease (MONDO:0000001) - [99,602 Polymer Entities]
▼ human disease (MONDO:0700096) - [99,602 Polymer Entities]
▶ acute disease (MONDO:0020683) - [26,346 Polymer Entities]
▶ familial hypercholesterolemia (MONDO:0005439) - [293 Polymer Entities]

[Browse Disease Annotations](#)

Learn [more about Mondo DO](#)

Use Advanced Search to group structures based on the reference sequence

Search Query History Browse Annotations MyPDB

QUERY: Lineage Identifier = "MONDO:0005439" AND Annotation Type = "disease" MyPDB Login Search API

Advanced Search Query Builder [Help](#)

Full Text [?](#)

Structure Attributes [?](#) [Help](#)

Lineage Identifier - Pharos Disease x is MONDO:0005439 + NOT Count x

Add Attribute Add Subquery Remove Subquery

Add Subquery

Chemical Attributes [?](#)

Sequence Similarity [?](#)

Sequence Motif [?](#)

Structure Similarity [?](#)

Structure Motif [?](#)

Chemical Similarity [?](#)

Retrieve results as groups based on UniProt KB Accession and include the CSMs

Return Polymer Entities [?](#) grouped by UniProt Accession [?](#) displaying as Groups [?](#) Include Computed Structure Models (CSM) Count Clear Search

[Use Advanced Search tool to explore individual proteins](#)

Explore search results and navigate to the GSP



Group Cholesteryl ester transfer protein
Group Total Size: 4
Polymer Entities 4 matching query
Best Example 2OBD_1 (Resolution: Best) [Explore in 3D](#)
Group ID P11597
Description Involved in the transfer of neutral lipids, including cholesteryl ester and triglyceride, ...



Group Signal-transducing adaptor protein 1
Group Total Size: 3
Polymer Entities 3 matching query
Best Example 3MAZ_1 (Resolution: Best) [Explore in 3D](#)
Group ID Q9ULZ2
Description In BCR signaling, appears to function as a docking protein acting downstream of TE...



Group Low density lipoprotein receptor adaptor protein 1
Group Total Size: 2
Polymer Entities 2 matching query
Best Example 2G30_2 (Resolution: Best) [Explore in 3D](#)
Group ID Q5SW96
Description Adapter protein (clathrin-associated sorting protein (CLASP)) required for efficient en...

Click to navigate to the Group Summary page



Explore LDLRAP1 protein group

Group Summary Group Sequence

Member 2 of 2 Member 1 of 2

Explore in 3D: [Sequence Alignments](#)
AF_AFQ5SW96F1_1 - Summary | Structure
Name: Low density lipoprotein receptor adapter p...
Organism: Homo sapiens
Experimental Method: NA
Resolution: N/A
Molecular Weight: 33.927 kDa

Explore in 3D: [Sequence Alignments](#)
2G30_2 - Summary | Structure
Name: Low density lipoprotein receptor adapter p...
Organism: Homo sapiens
Experimental Method: X-ray
Resolution: 1.6 Å
Molecular Weight: 1.783 kDa

Low density lipoprotein receptor adapter protein 1

UniProtKB accession: Q5SW96
Grouped By: Matching UniProtKB accession
Group Content: Polymer Entities matching query 2
Go to UniProtKB: [Q5SW96](#)

UniProtKB description: Adapter protein (clathrin-associated sorting protein (CLASP)) required for efficient endocytosis of the LDL receptor (LDLR) in polarized cells such as hepatocytes and lymphocytes, but not in non-polarized cells (fibroblasts). May be required for LDL binding and internalization but not for receptor clustering in coated pits. May facilitate the endocytosis of LDLR and LDLR-LDL complexes from coated pits by stabilizing the interaction between the receptor and the structural components of the pits. May also be involved in the internalization of other LDLR family members. Binds to phosphoinositides, which regulate clathrin bud assembly at the cell surface. Required for trafficking of LRP2 to the endocytic recycling compartment which is necessary for LRP2 proteolysis, releasing a tail fragment which translocates to the nucleus and mediates transcriptional repression (By similarity).

PDB Low density lipoprotein receptor adapter protein
Sequence Alignment in 3D: Q5SW96
Found in patients with hypercholesterolemia; likely pathogenic; dbSNP/rs752949346; VAR_078925 [UNIPROT] | Position: 56

Low density lipoprotein receptor adapter protein 1
AF_AFQ5SW96F1 | Model 1 | Instance AFQ_1 | A | THR 56
UNIP_Q5SW96:56 T
pLDDT Score (1 Residue): 97.58 (very high)

[Explore Group Summary](#)

[Explore Sequence Alignment in 3D](#)

Explore AlphaFold model of LDLRAP1 protein

Structure Summary | Structure | Annotations | Sequence | Genome

Display Files | Download Files | Data API

AF_AFQ5SW96F1

COMPUTED STRUCTURE MODEL OF LOW DENSITY LIPOPROTEIN RECEPTOR ADAPTER PROTEIN 1

AlphaFold DB: AF-Q5SW96-F1

Released in AlphaFold DB: 2021-07-01
Last Modified in AlphaFold DB: 2022-09-30

Organism(s): Homo sapiens
UniProtKB: Q5SW96


There are no experimental data to verify the accuracy of this computed structure model. See Model Confidence metrics below for all regions of the polypeptide chain.

Click to search for proteins with similar 3D structure

Macromolecules

Find similar proteins by: Sequence (by Identity cutoff) | 3D Structure

Entity ID: 1

Molecule	Chains	Sequence Length	Organism	Details	Image
Low density lipoprotein receptor adapter protein 1	A	308	Homo sapiens	Mutation(s): 0 Gene Names: LDLRAP1	

Explore in 3D: [Structure](#) | [Sequence Annotations](#)

Global Symmetry: Asymmetric - C1
Global Stoichiometry: Monomer - A1

[Find Similar Assemblies](#)

[Explore Structure Summary](#)

Combine Structure Similarity and text annotation searches

Refinements

Structure Determination Methodology

- computational (9)
- experimental (2)

CSM Source Database

- AlphaFoldDB (9)

pLDDT Global Quality Score

- 50.0 - 70.0 (2)
- 70.0 - 90.0 (7)

Scientific Name of Source Organism

- Danio rerio (3)
- Homo sapiens (2)
- Mus musculus (2)
- Candida albicans SC5314 (1)
- Dracunculus medinensis (1)
- Fonsecaea pedrosoi CBS 271.37 (1)
- Rattus norvegicus (1)

1 to 2 of 2 Polymer Entities Page 1 of 1 25 Sort by | Score

AF_AFQ5SW96F1: Entity 1 [Download File](#)

Computed structure model of Low density lipoprotein receptor adapter protein 1

AlphaFold DB AF-Q5SW96-F1

pLDDT (global) 70.78

Chain ID A

Organism [Homo sapiens](#)

Macromolecule [Low density lipoprotein receptor adapter protein 1](#)

Structure Match [Align in 3D](#)

[Explore in 3D](#)

6OVF: Entity 1 [Download File](#) [View File](#)

Crystal Structure of the Disabled-2 (Dab2) Dab Homology Domain in Complex with Peptide STA03

Chavez, M., Madden, D.R.

To be published

Released 2020-05-13

Method X-RAY DIFFRACTION 1.95 Å

Chain IDs A, B

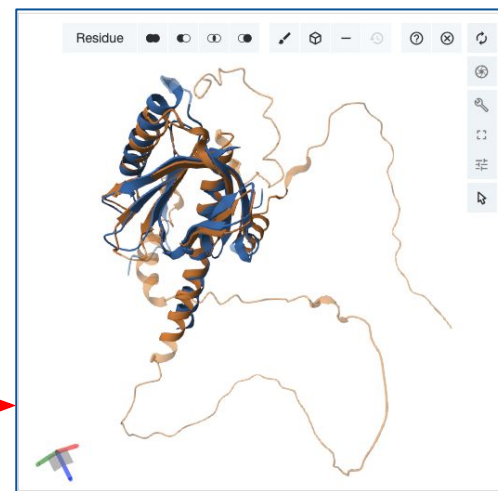
Organism [Homo sapiens](#)

Macromolecule [Disabled homolog 2](#)

Structure Match [Align in 3D](#)

1 to 2 of 2 Polymer Entities Page 1 of 1 25 Sort by | Score

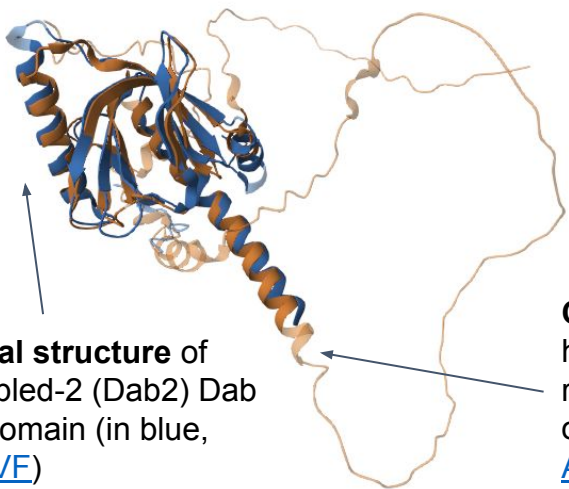
[Navigate to Pairwise Structure Alignment tool](#)



[Search for human proteins with similar shape](#)

Disabled homolog 2 (DAB2) protein

Entry	Chain	RMSD	TM-score	Identity	Equivalent Residues	Sequence Length	Modelled Residues
AF_AFQ5SW96F1	A	-	-	-	-	308	308
6OVF	A	2.43	0.43	18%	148	161	154



Experimental structure of human Disabled-2 (Dab2) Dab Homology Domain (in blue, [PDB ID: 6OVF](#))

Computed structure model of human low density lipoprotein receptor adapter protein 1 (in orange, [AlphaFoldDB ID: AF-Q5SW96-F1](#))

Disabled-2 (Dab2) is a signal-transduction protein that contains phosphotyrosine binding (PTB) domain and, like LDLRAP1, binds to and receives signals from members of the low-density lipoprotein receptor (LDLR) family

Summary

- RCSB.org offers seamless access to a suite of powerful tools designed to explore structure data
- Computed Structure Model (CSM) data provided alongside experimental structures helps to bridge gaps in structural data exploration for proteins with low experimental coverage

Case study of Class II aminoacyl-tRNA synthetases



Florian Kaiser



Sarah Krautwurst



Christoph Leberecht



Dirk Labudde



Sebastian Salentin



V. Joachim Haupt



Michael Schroeder

Sebastian Bittrich, PhD

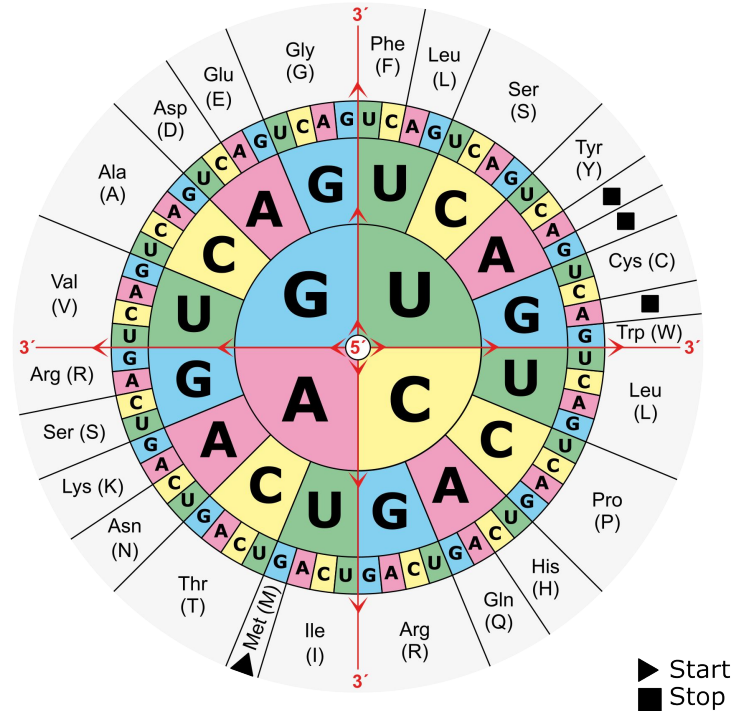
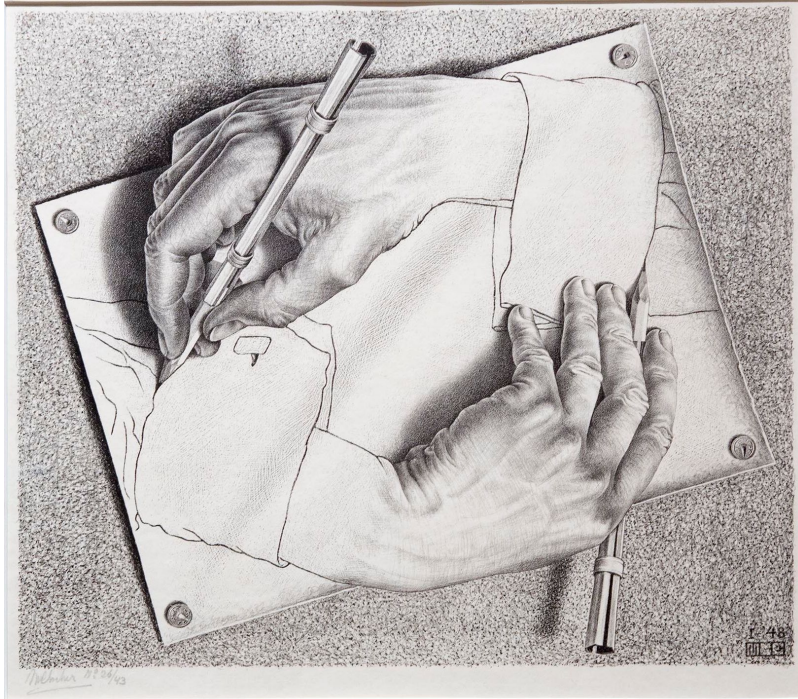
sebastian.bittrich@rcsb.org

Aminoacyl-tRNA Synthetases (aaRS)

- Present in all organisms and all cells
- One implementation per amino acid
- 2-step reaction that uses ATP as energy source



A Self-Referential System

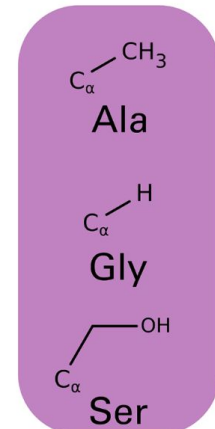
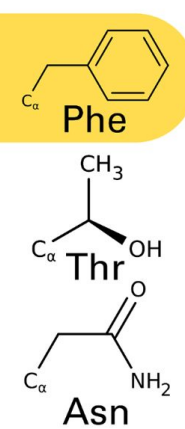
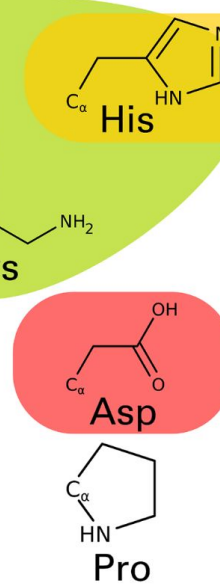
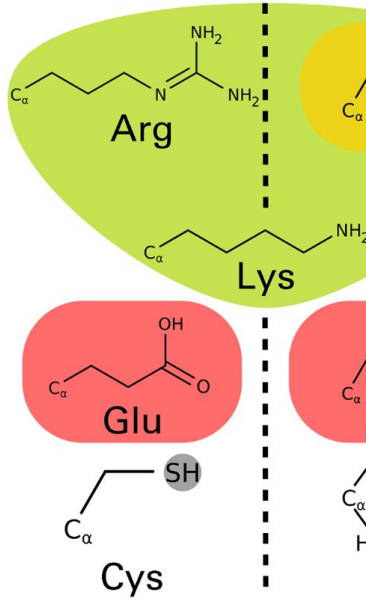
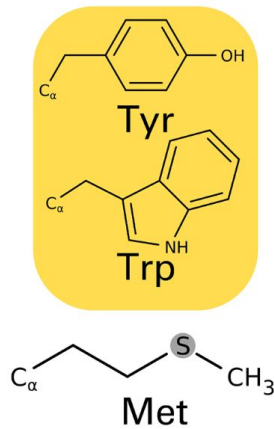
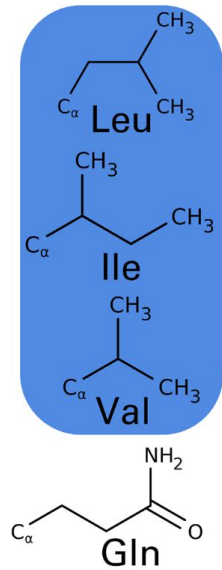


▶ Start
■ Stop

2 Classes of Aminoacyl-tRNA Synthetases

aaRS Class I

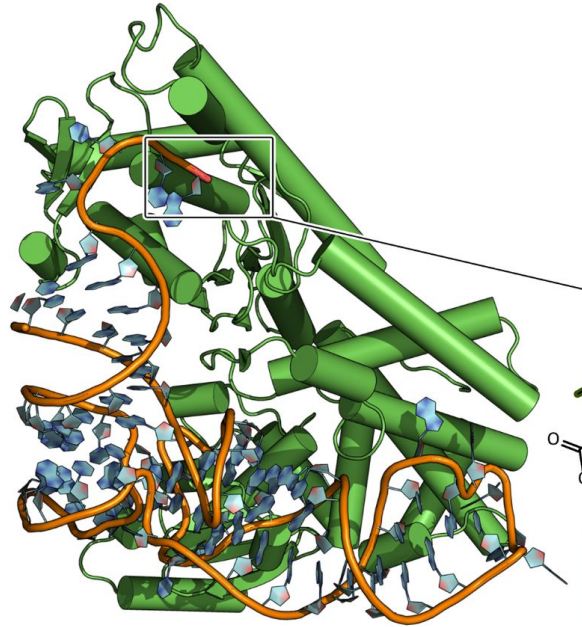
aaRS Class II



- negative
- positive
- tiny
- aromatic
- aliphatic
- sulfur

2 Classes of Aminoacyl-tRNA Synthetases

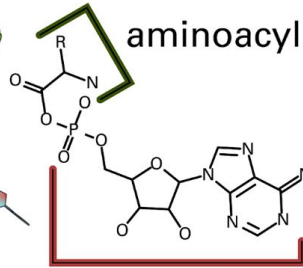
aaRS Class I



ArgRS
PDB:1f7u

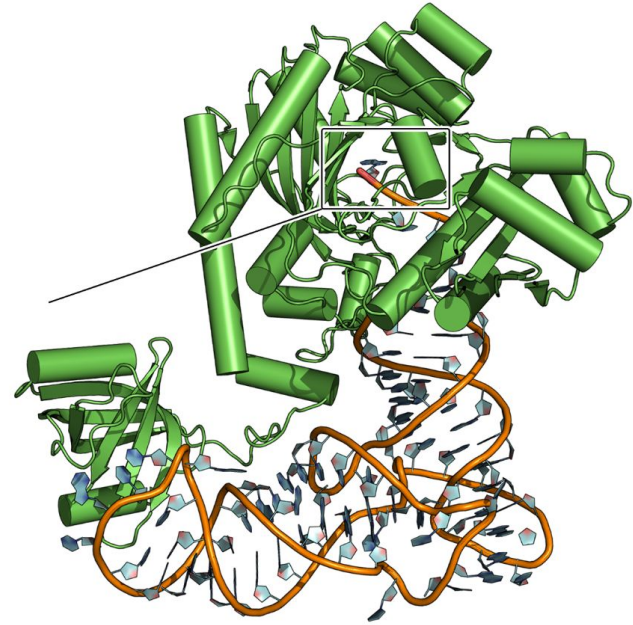
catalytic site

aminoacyl



adenosine monophosphate

aaRS Class II



AspRS
PDB:1c0a

2018: First Archive-Wide Structural Analysis

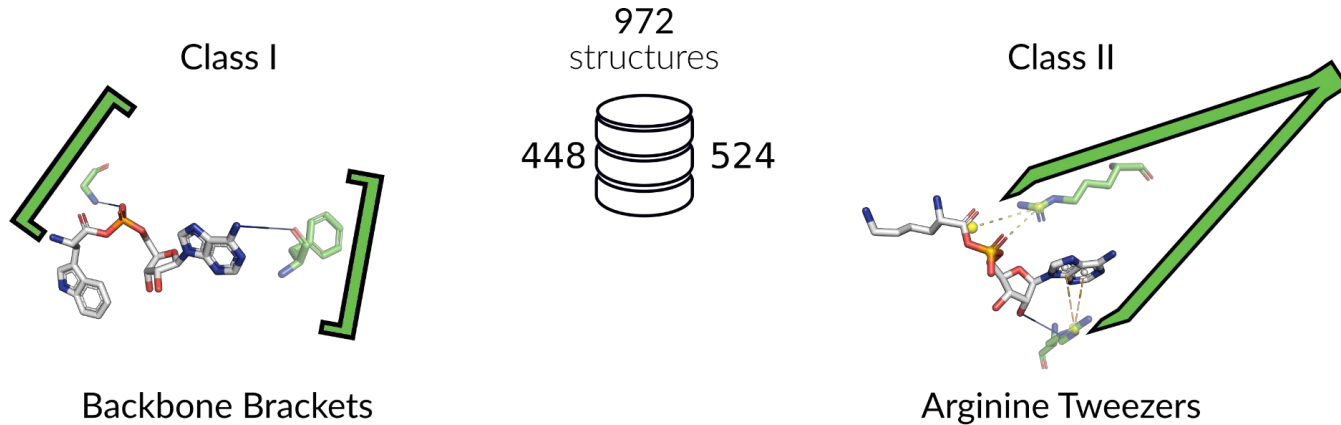
Class I

972
structures

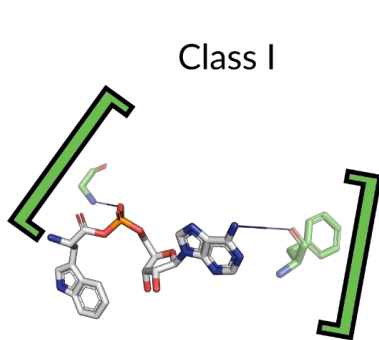
Class II



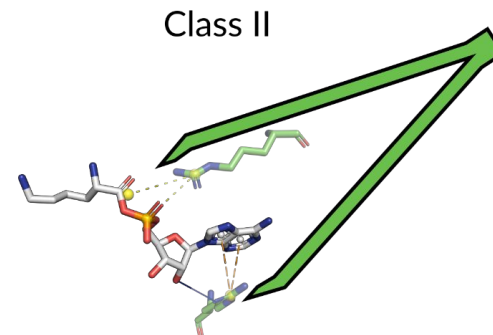
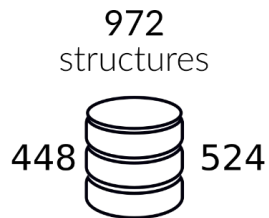
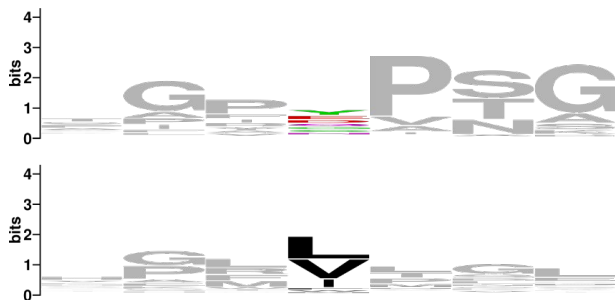
aaRS Class Signatures



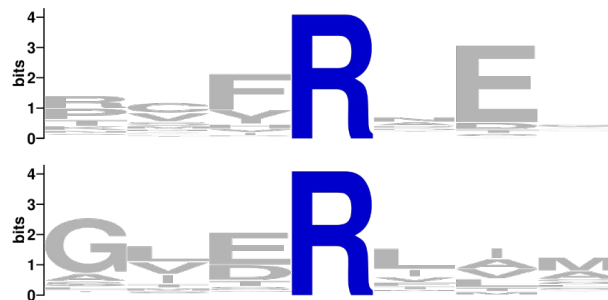
aaRS Class Signatures



Backbone Brackets



Arginine Tweezers



Using 1c0a as Reference for Class II aaRS

Structure Summary | Structure | Annotations | Experiment | Sequence | Genome | Ligands | Versions

Biological Assembly 1 ?

Display Files | Download Files | Data API

1C0A

CRYSTAL STRUCTURE OF THE E. COLI ASPARTYL-TRNA SYNTHETASE : TRNAASP : ASPARTYL-ADENYLATE COMPLEX

PDB DOI: <https://doi.org/10.2210/pdb1C0A/pdb> NAKB: 1C0A

Classification: **LIGASE/RNA**
Organism(s): *Escherichia coli*
Expression System: *Escherichia coli*
Mutation(s): No

Deposited: 1999-07-15 Released: 1999-11-23
Deposition Author(s): Eiler, S., Dock-Bregeon, A.-C., Moulinier, L., Thierry, J.-C., Moras, D.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 2.40 Å
R-Value Free: 0.249
R-Value Work: 0.208
R-Value Observed: 0.208

wwPDB Validation **?** [3D Report](#) [Full Report](#)

Metric	Percentile Ranks	Value
Rfree		0.242
Clashscore		16
Ramachandran outliers		0.5%
Sidechain outliers		6.0%
RSRZ outliers		2.0%
RNA backbone		0.65

Worse | Percentile relative to all X-ray structures | Better
Percentile relative to X-ray structures of similar resolution

Ligand Structure Quality Assessment **?**

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

Global Symmetry: Asymmetric - C1 **?**
Global Stoichiometry: Monomer - A1 **?**

[Find Similar Assemblies](#)

Biological assembly 1 assigned by authors.

Revisiting Class II – 6 Years and 1M CSMs Later

Sequence Search



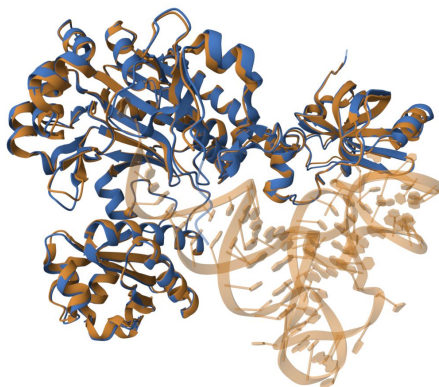
Example Queries

[20 experimental](#)
[7 new since 2018](#)



[796 CSMs](#)

Global Structure Search

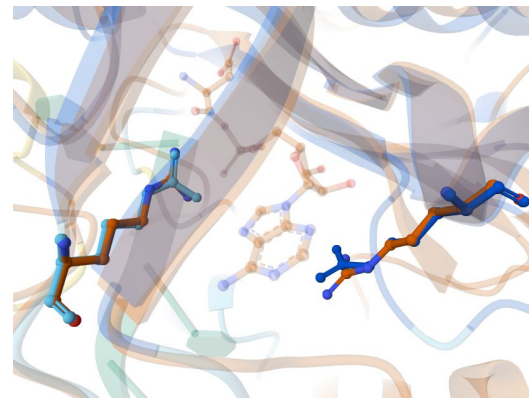


Example Queries

[18 experimental](#)
[6 new since 2018](#)

[715 CSMs](#)

Structure Motif Search



Example Queries

[1,604 experimental](#)
[956 new since 2018](#)

[6,073 CSMs](#)

Live demo


[RCSB.org](https://www.rcsb.org)

Structure Summary Page: Similarity Queries

Macromolecules

Proteins **1** Nucleic Acids / Hybrid **1**

Find similar proteins by: [Sequence](#) (by identity cutoff) [3D Structure](#) **Launch Structure Similarity Search**

Entity ID: 2	100%	Sequence Length	Organism	Details	Image
Molecule	95%				
ASPARTYL TRNA SYNTHETASE	90%	585	Escherichia coli	Mutation(s): 0	
	80%				
	70%				
	60%				
	50%				
UniProt	40%				
Find proteins for P21889	30%	(12))	Explore P21889 ⓘ	Go to UniProtKB: P21889	

Entity Groups ⓘ

Remember to Toggle CSMs if Relevant

Advanced Search Query Builder [?](#) [Help](#)

Full Text [?](#)

Structure Attributes [?](#)

Chemical Attributes [?](#)

Sequence Similarity [?](#) [Help](#)

MRTEYCGQLRLSHVGGQVTLGQWVNRRLDLGSLIFIDMRDREGIVQVFFDPDRADALKASELRNEFCIQVTGTVRARDEKNIINRDMATGEIEVLASSLTIINRADVLPDLSNHVNTTEEARLKRYRDLRRPEMAQRLKT
RAKITSLVRRFMDDHGFLDIETPMLTKATPEGARDYLVP SRVHKGKGYALPQSPQLFKLLMMSGFDRYYQIVKCFRDEDLRADRQPEFTQIDVETSFMTAPQVREVMEALVRHLWLEVKGVDLGDFFVMTFAEAERR

Entry ID Sequence Type [?](#) E-Value Cutoff [?](#) Identity Cutoff % (Integer only) [?](#)

Sequence Motif [?](#)

Structure Similarity [?](#)

Structure Motif [?](#)

Chemical Similarity [?](#)

Return [?](#) grouped by [?](#) [?](#)



CSMs aren't included by default when launching queries from Structure Summary Pages of experimentally-determined entries

Toggle to Include CSMs

Sequence Similarity Search Results

Search Query History Browse Annotations MyPDB

QUERY: Sequence Similarity = MRTEYCGQLRLSHVQQVTLGWNRRRDLGSLIFIDMRDREGIVQVFFDPDRADALKASELNEFCIQVTGVRARDEKINIRDMATGIEIVLASSLTIINRA DVLPLDSNHVNTTEEARLKYRYLDRRPEMAQRLKTRAKITSLVRRFMDHGFLDIETPMLTKATPEGARDYLVPSRVHKGKFYALPQSPQLFKLLMMSGFDRYYQVQKCFRDEDLRADRQPEFT QIDVETFSMTAPQVREVMEALVRHLWLEVKGVLDGDFPVMTFFAEERRYGSDKPDLRNPMELTDVADLLKSEVAFVAGPANDPKGRVAALRVPGGASLTRKQIDEYGNFVKIYGAKGLAYIKVNE RAKGLEGINSPVAKFLNAEIIEDLDRTAQAQDGMIFFGADNKKIVADAMGALRKLKVGKDLGLTDESKWAPLWVIDPMPFEDDGEGGLTAMHHPTSPKDMTAAELKAAPENAVANAYDMVINGYE VGGGSRRIHNGDMQQTVFGILGINEEEEQREKFGLLDALKYGTTPHAGLAFGLDRLTLLTGDNRDVIAPFKTTAAACLMTAEPSFANPTALAELSIQVVK **AND** Sequence Type = Protein **AND** E-Value Cutoff = 0.1 **AND** Identity Cutoff = 30%

MyPDB Login Search API

Advanced Search Query Builder Help

Search Summary This query matches 816 Polymer Entities.

Refinements Structure Determination Methodology

- computational (796)
- experimental (20)

CSM Source Database

- AlphaFoldDB (790)
- ModelArchive (6)

pLDDT Global Quality Score

- 50.0 - 70.0 (4)
- 70.0 - 90.0 (40)
- > 90.0 (752)

Scientific Name of Source Organism

- Thermus thermophilus (8)
- Glycine max (7)
- Schistosoma mansoni (7)
- Zea mays (5)
- Mycobacterium ulcerans str. Harvey (4)
- Spongilla lacustris (4)

Alignment Reference Query

1 to 25 of 816 Polymer Entities Page 1 of 33 25 Sort by Score

1C0A: Entity 2 Download File View File Download Alignment

CRYSTAL STRUCTURE OF THE E. COLI ASPARTYL-TRNA SYNTHETASE : TRNAASP : ASPARTYL-ADENYLATE COMPLEX

Eiler, S., Dock-Bregeon, A.-C., Moulinier, L., Thierry, J.-C., Moras, D.

(1999) EMBO J 18: 6532-6541

Released 1999-11-23

Method X-RAY DIFFRACTION 2.4 Å

Chain ID B [auth A]

Organism Escherichia coli

Macromolecule ASPARTYL TRNA SYNTHETASE

Sequence Match Sequence Identity: 100%, E-Value: 0, Region: 1-585

Explore in 3D

QUERY 1C0A_2

1EQR: Entity 1 Download File View File Download Alignment

Search Summary

Manipulate Search

Result Summary

Sequence Alignment

Structure Similarity Search Results

Search Query History Browse Annotations MyPDB

QUERY: Structure Similarity WHERE (Entry ID = "1C0A" AND Chain ID = "B" AND Shape Match = "Strict") AND Structure Determination Methodology = "computational" MyPDB Login Search API

Advanced Search Query Builder Help

Search Summary This query matches 715 Polymer Entities.

Refinements

- Structure Determination Methodology
 - computational (715)
- CSM Source Database
 - AlphaFoldDB (714)
 - ModelArchive (1)
- pLDDT Global Quality Score
 - 70.0 - 90.0 (18)
 - > 90.0 (697)
- Scientific Name of Source Organism
 - Arabidopsis thaliana (2)
 - Methylobacillus flagellatus KT (2)
 - Streptococcus mutans UA159 (2)
 - Thiobacillus denitrificans ATCC 25259 (2)
 - Acaryochloris marina MBIC11017 (1)
 - Acidiphilium cryptum JF-5 (1)
 - Acidithiobacillus ferrooxidans ATCC 23270 (1)
 - Acidithiobacillus ferrooxidans ATCC 53993 (1)
 - Acidobacterium capsulatum ATCC 51196 (1)

-- Tabular Report -- All Selected Download

1 to 25 of 715 Polymer Entities Page 1 of 29 25 Sort by Score

AF_AFA3N1F3F1: Entity 1 Download File

Computed structure model of Aspartate--tRNA ligase

AlphaFold DB AF-A3N1F3-F1
pLDDT (global) 94.52
Chain ID A
Organism Actinobacillus pleuropneumoniae serovar 5b str. L20
Macromolecule Aspartate--tRNA ligase
Structure Match Align in 3D

Explore in 3D

AF_AFB3GY37F1: Entity 1 Download File

Computed structure model of Aspartate--tRNA ligase

AlphaFold DB AF-B3GY37-F1
pLDDT (global) 94.58
Chain ID A
Organism Actinobacillus pleuropneumoniae serovar 7 str. AP76
Macromolecule Aspartate--tRNA ligase
Structure Match Align in 3D

Explore in 3D

Refine Search

Align Chains

Alignment Visualization

Pairwise Structure Alignment

This tool allows the selection of protein 3D structures for alignment. Use an existing PDB or [Computed Structure Model](#) entry ID, upload a local file with atomic coordinates, or enter a URL of a file on the web

✳ Alignment API

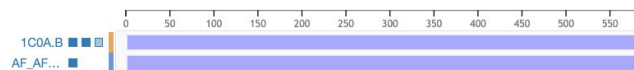
▸ Compare Protein Structures

Entry	Chain	RMSD	TM-score	Identity	Equivalent Residues	Sequence Length	Modelled Residues
1C0A	B	-	-	-	-	585	585
AF_AFA3N1F3F1	A	0.75	0.98	75%	585	591	591

Sequence Alignment in 3D

↓ Export ▾

📄 Copy Link



This view is only linked when searching for similar chains
(and not available when searching for assemblies)

Structure Summary Page: Ligand Focus View

Structure Summary | Structure | Annotations | Experiment | Sequence | Genome | Ligands | Versions

Biological Assembly 1 ?

Display Files | Download Files | Data API

1C0A

CRYSTAL STRUCTURE OF THE E. COLI ASPARTYL-TRNA SYNTHETASE : TRNAASP : ASPARTYL-ADENYLATE COMPLEX

PDB DOI: <https://doi.org/10.2210/pdb1C0A/pdb> NAKB: 1C0A

Classification: **LIGASE/RNA**
Organism(s): *Escherichia coli*
Expression System: *Escherichia coli*
Mutation(s): No

Deposited: 1999-07-15 Released: 1999-11-23
Deposition Author(s): Eiler, S., Dock-Bregeon, A.-C., Moulinier, L., Thierry, J.-C., Moras, D.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 2.40 Å

wwPDB Validation **3D Report** Full Report

Metric	Percentile Ranks	Value
Rfree		0.242
Clashscore		16
Ramachandran outliers		0.5%
Sidechain outliers		6.0%
RSRZ outliers		2.0%
RNA backbone		0.65

Worse | Percentile relative to all X-ray structures | Better
Percentile relative to X-ray structures of similar resolution

Ligand Structure Quality Assessment **3D Focus on Ligand of Interest**

Explore in 3D: Structure | Sequence Annotations | Electron Density | Validation Report | **Ligand Interaction (AMO)**

Global Symmetry: Asymmetric - C1
Global Stoichiometry: Monomer - A1

Find Similar Assemblies

Biological assembly 1 assigned by authors.

Define Structure Motifs using Mol*



Select residues in 3D to include in your search. Make sure to activate the "Selection Mode".

The screenshot shows the Mol* interface with a 3D protein structure. The search panel on the right is open, showing the 'Structure Motif Search' section. A red box highlights the 'Toggle the Selection Mode' button, and another red box highlights the 'Structure Motif Search' controls, including the search input field and the 'Submit Search' button. The search results show two motifs: 1. ARG 217 | B [auth ...] and 2. ARG 537 | B [auth ...].

Controls to Define Structure Motif Queries

Structure Motif Search Results

QUERY: Structure Motif = 1C0A [B_1-217 [ARG] AND B_1-537 [ARG]] AND RMSD Cutoff = 1 AND Atom Pairing = "All Atoms" AND Structure Determination Methodology = "computational"

MyPDB Login Search API

Advanced Search Query Builder Help

Search Summary This query matches 6,073 Assemblies.

Refinements Structure Determination Methodology

- computational (6,073)

CSM Source Database

- AlphaFoldDB (5,974)
- ModelArchive (99)

pLDDT Global Quality Score

- < 50.0 (8)
- 50.0 - 70.0 (247)
- 70.0 - 90.0 (1,210)
- > 90.0 (4,608)

Scientific Name of Source Organism

- Zea mays (95)
- Glycine max (85)
- Homo sapiens (75)
- Danio rerio (72)

-- Tabular Report --

1 to 25 of 6,073 Assemblies Page 1 of 243 25 Sort by Score

AF_AFQ67KH4F1: Assembly Download File

Computed structure model of ATP phosphoribosyltransferase regulatory subunit

AlphaFold DB **AF-Q67KH4-F1**

pLDDT (global) 92.12

Global Symmetry: Asymmetric - C1

Oligomeric Count 1

Structure Motif Match Align in 3D RMSD: 0.26 Å - ARG:A-114, ARG:A-321

Explore in 3D

Align Motif Residues

AF_AFQ8... Assembly Download File

Computed structure model of Histidine--tRNA ligase

AlphaFold DB **AF-Q8EPR9-F1**

pLDDT (global) 93.86

Structure Motif Alignment View

Sequence of 1C0A | CRYSTAL... Chain 1: ASPARTYL... A [auth B]

601 611 620A 630 640 650 660 670
GGAGCGG 4SU AGUUCAG H2U CGG H2U H2U AGAAUACCUGCCU QU0 UCACGCGAGGGG G7M UCGCGGG 5MU PSU CGAGUCCCG PSU CCGUUCGCGCA

11:26:00 Updated Ball & Stick in 9ms.
11:26:00 Updated AF_AFQ67KH4F1 #1: 0.26 Å - Polymers in 0ms.
11:26:00 Updated Cartoon in 2ms.

Import

- Open Files
- Download Structure

Session

Structure

2 structures

Nothing Focused

Measurements

Structure Motif Search

Superposition

Components 2 structures

Preset	+ Add	☰	⌂
1C0A	Ball & Stick	👁	🗑
1C0A - Polymers	Cartoon	👁	🗑
1C0A - Ligands	Ball & Stick	👁	🗑
1C0A - Ions	Ball & Stick	👁	🗑
1C0A - Waters	Ball & Stick	👁	🗑
AF_AFQ67KH4F1 #1: 0.26 Å		👁	🗑
AF_AFQ67KH4F1 #1: 0.26...		👁	🗑

Export Models

Export Animation

Export Geometry

Summary: RCSB.org is a powerful tool for CSM exploration

- Different ways to search for relevant Class II aaRS entries
 - Sequence search **primarily** finds aspartyl-RS
 - Global structure search **primarily** finds aspartyl-RS
 - Structure motif search finds **all Class II aaRS**, even for other amino acids plus ATP-binding paralogs
 - In general: Different types of searching may be relevant and RCSB.org is the **one-stop-shop to run different search types**
- Relevant CSMs identified from [AlphaFold DB](#) and [ModelArchive](#)
 - Numerous aaRS predictions available from model organisms and genomes with global health implications from AlphaFold DB
 - NB: ModelArchive contains i.a. [predictions relevant for climate change](#)

Related Resources

- **AlphaFold:** Highly accurate protein structure prediction with AlphaFold. J. Jumper et al., Nature 596, 583-589 (2021).
 - **AlphaFold DB:** alphafold.ebi.ac.uk
- **RoseTTAfold:** Accurate prediction of protein structures and interactions using a three-track neural network. M. Baek et al., Science 373, 871-876 (2021).
- **ModelArchive:** Computational Structural Biology group at the SIB - Swiss Institute of Bioinformatics and the Biozentrum University of Basel
 - **ModelArchive:** modelarchive.org
- **ModelCIF:** An Extension of PDBx/mmCIF Data Representation for Computed Structure Models. B. Vallat et al., J Mol Biol, 168021 (2023).
- **Open-access data:** A cornerstone for artificial intelligence approaches to protein structure prediction. S. K. Burley, H. M. Berman, Structure 29, 515-520 (2021).



RCSB PDB Team

RCSB PDB RCSB.ORG
PROTEIN DATA BANK info@rcsb.org

Core Operations Funding

US National Science Foundation (DBI-2321666),
National Institute of General Medical Sciences,
National Institute of Allergy and Infectious Disease, and
National Cancer Institute (NIH R01GM133198), and the
US Department of Energy (DE-SC0019749)

Management



RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

UC San Diego

SDSC SAN DIEGO
SUPERCOMPUTER CENTER

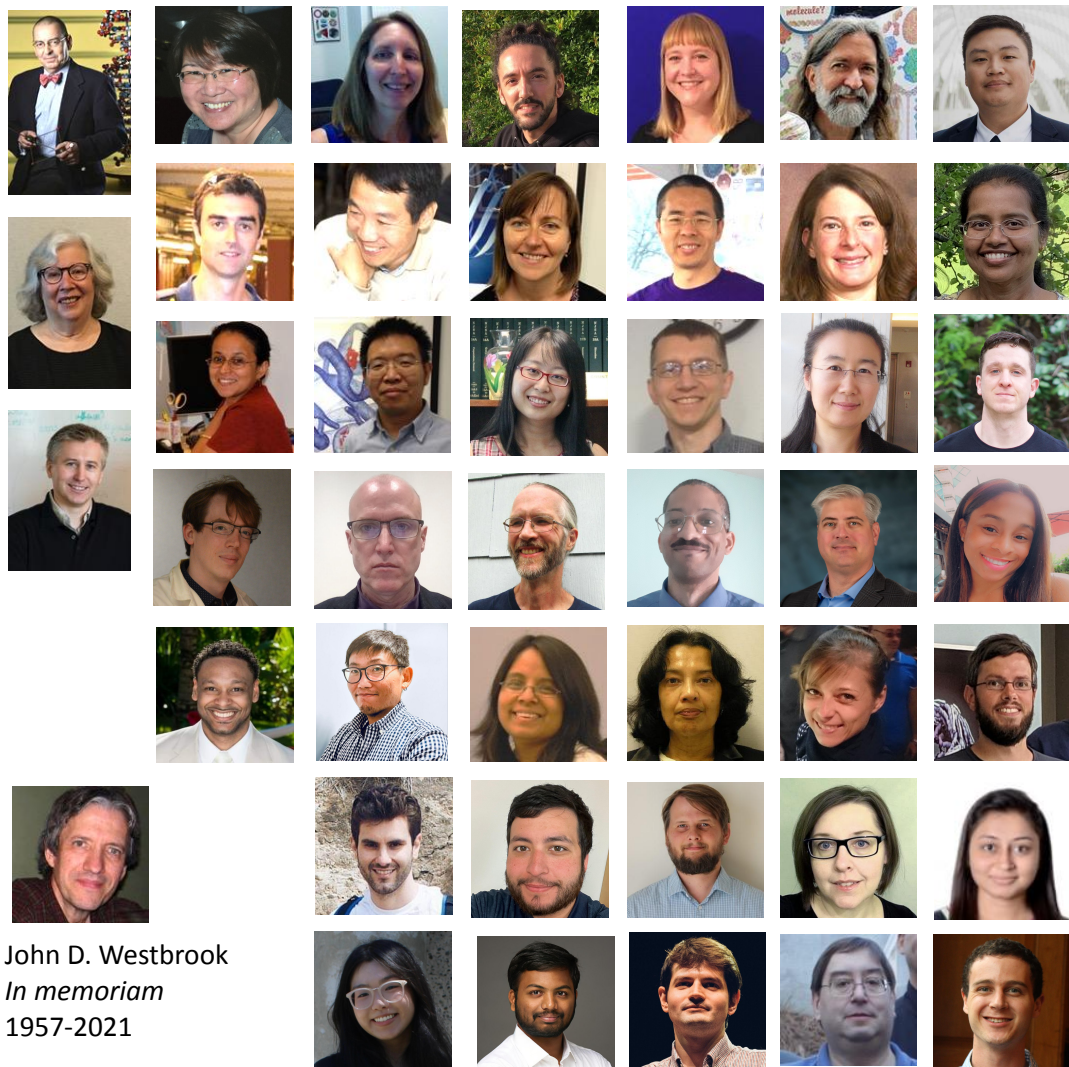
UCSF

University of California
San Francisco



Member of the
Worldwide Protein Data Bank
(wwPDB; wwpdb.org)

Follow us



John D. Westbrook
In memoriam
1957-2021