

RCSB **PDB**
PROTEIN DATA BANK

RCSB.ORG A LIVING DIGITAL
DATA RESOURCE THAT ENABLES
SCIENTIFIC BREAKTHROUGHS

2021

**ANNUAL
REPORT**



DIRECTOR'S MESSAGE

2021 was a remarkable year for the RCSB Protein Data Bank (PDB), albeit one still firmly rooted in the “virtual” world.



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

Director
RCSB Protein Data Bank
University Professor and
Henry Rutgers Chair
Rutgers, The State University
of New Jersey

Adjunct Professor
University of California San Diego

Most significant for our team and user communities were celebrations of the 50th birthday of the PDB on October 20th. On that date in 1971, structural biologists came together in a prescient initiative to establish a single global repository for three-dimensional (3D) macromolecular structure information—the PDB. From its inception, keepers of the PDB archive embraced technical excellence and open access, ensuring its widespread use across the sciences. Fifty years on, many millions of students, educators, and scientists around the world access PDB data at no charge and with no limitations on usage. They explore, understand, and advance our understanding of fundamental biology, biomedicine, bioenergy, and bioengineering/biotechnology, and help discover and develop new life-changing drugs and design new vaccines.

RCSB PDB together with our Worldwide Protein Data Bank (wwPDB) partners planned some geographically-bounded events to

celebrate this important milestone in the history of structural biology and data science. Paradoxically, continued pandemic public health measures allowed us to celebrate with the entire PDB community in ways not previously imagined. The necessary move to virtual events widened our audience dramatically, making for truly international symposia. In all, eight celebratory scientific meetings hosted distinguished speakers who have made tremendous advances in structural biology or bioinformatics. In addition to the international live audience, recordings of many of these presentations are now available online to be enjoyed by many more fans of the PDB.

The year 2021 will also be remembered as the era when protein structure prediction came of age. Shortly after we began our 50th year of PDB operations, Google DeepMind, Inc. announced that their AlphaFold2 software could predict 3D structures (atomic coordinates) for small globular proteins with accuracies comparable to that of lower-resolution

Table of Contents

RCSB PDB SERVICES

	Message from the RCSB PDB Director	PDB and 2021 at a glance	Archive Management	Outreach and Education	Credits and Funding		
PAGE	2	4	6	8	10	11	12
		Celebrating 50 Years of Open Access to the PDB Archive	Deposition and Biocuration	Data Exploration		2021 Publications	

DIRECTOR'S MESSAGE CONT'D

experimental methods. A frequent topic of discussion at PDB 50th meetings was the role open access PDB structures played in fueling this sea change. Successful application of artificial intelligence/deep learning methods to the protein structure prediction problem would not have been possible without the generosity of tens of thousands of PDB depositors who made their structures freely available and provided critical data underpinning protein structure prediction. Today, AlphaFold2 and RoseTTAFold computed structure models for nearly one million proteins are freely available (from AlphaFoldDB and the ModelArchive), and the underlying computer code from both software tools has also been made public. As a global bioscience research community, we are just beginning to explore the myriad opportunities that whole proteome 3D structure prediction will enable.

PDB 50th birthday preparations revealed previously hidden talents among some RCSB PDB team members. We now provide open access to PDB Structural Biology Playing Cards and a PDB-themed board game that gives players the opportunity to explore the process of structure discovery from sample preparation to PDB deposition.

In another 2021 milestone, the Electron Microscopy Data Bank (EMDB), the global open-access repository for electron microscopy map and tomogram information, became an official member of the wwPDB partnership. Revision of the wwPDB Charter at the beginning of the year codified a long-standing relationship between the EMD and wwPDB, focused on electron microscopy data deposition, validation, and biocuration. During 2021 the number of PDB structures coming from electron microscopy methods grew by a ~24% to a total of more than 9,600.

Throughout 2021, wwPDB partners continued to focus considerable efforts on the fight against the COVID-19 pandemic. More than 1,700 SARS-CoV-2 related structures were freely available from the PDB archive at the end of the year. Structural considerations and PDB structures (including those of the closely-related SARS-CoV spike protein) contributed to the design of SARS-CoV-2 spike protein-encoding mRNA vaccines, which became available for clinical use in record time. The very first PDB structure of a SARS-CoV-2 protein, the main protease, was made publicly available in early February 2020. Two years later we now have emergency use authorization of

nirmatrelvir, a medicine in pill form that blocks the action of the main protease and stops the virus cold by preventing essential cleavage of the viral polyprotein early in the infection. Structural biologists deserve much praise for their myriad contributions to turning the tide of the pandemic.

I look forward to the next half century during which structural biosciences and the PDB shine ever more brightly, illuminating the way forward as we address pressing global public health challenges, such as climate change, the next pandemic to cross the species barrier into humans, and the ongoing war on cancer.



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

Remembering John Westbrook

John D. Westbrook Jr. (1957-2021), Research Professor at Rutgers University and Data & Software Architect Lead for the RCSB PDB, passed away suddenly on October 18, 2021.

For more than 25 years John supported Rutgers, RCSB PDB, and millions of data users worldwide with his vision and passion for innovative databases, ontologies, and other technologies for management of complex biological data.

As Data & Software Architect Lead of the RCSB PDB, John was central to the design and development of infrastructure and services to acquire, curate, archive, and deliver 3D macromolecular structure data to the broad community of PDB users.

His work also established the PDBx/mmCIF data dictionary and format as the foundation of the modern Protein Data Bank (PDB) archive.



John D. Westbrook Jr.
(1957–2021)

Acta Cryst (2021) **D77**: 1475-1476
doi: 10.1107/S2059798321011402.

CELEBRATING 50 YEARS OF OPEN ACCESS TO THE PROTEIN DATA BANK ARCHIVE

PROTEIN DATA BANK



In 1971, the structural biology community established the single worldwide archive for macromolecular structure data—the Protein Data Bank (PDB). PDB was the 1st open access digital data resource in all of biology and medicine.

Growing from the initial seven structures to more than 185,000 today, the PDB archive is a leading global resource for experimental data central to scientific discovery. PDB data are repackaged and redistributed by hundreds of data resources and millions of users exploring fundamental biology, energy, and biomedicine. In 2003, the Worldwide Protein Data Bank (wwPDB) was established to jointly manage the single PDB archive of macromolecular structural data that is freely and publicly available to the global community. It consists of regional data centers that act as deposition, data processing, and distribution centers for PDB data.

Throughout 2021, RCSB PDB and the wwPDB celebrated the 50th Anniversary of the PDB. Symposia were held (virtually) around the world, showcasing high-impact applications of protein structural data.

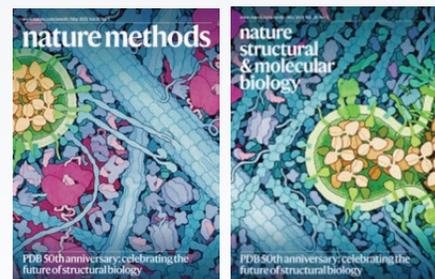
Many videos of presentations are available, including the inaugural PDB50 meeting hosted by ASBMB in May; the two-day Function Follows Form: Celebrating the 50th Anniversary of the Protein Data Bank hosted by the American Crystallographic Association in July; the Biophysical Society-hosted symposium in October; and the Protein Data Bank at 50: Accessing, Understanding, and Assessing PDB Data workshops hosted by the Royal Society of Chemistry Chemical Information and Computer Applications Group in November.

In July 2021, Congressman Frank Pallone, Jr., who represents New Jersey's 6th district that includes Rutgers University, heralded the 50th anniversary of the PDB in the Congressional Record.

JOURNAL COLLECTIONS



Journal of Biological Chemistry. How the Protein Data Bank Changed Biology contains reviews that celebrates scientific areas profoundly influenced by the creation of the PDB; The PDB in *JBC* highlights some seminal PDB structures published in the journal; the corresponding PDB structures in this collection are also featured at [RCSB.org](https://www.rcsb.org)



Nature Collection: PDB 50th Anniversary: celebrating the future of structural biology in *Nature Methods* and *Nature Structural & Molecular Biology*.

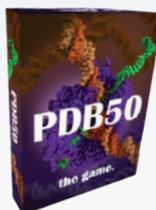


HISTORICAL VIDEO looks at the milestones that shaped the PDB.



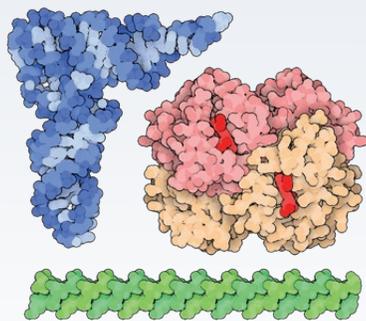
A **TIMELINE** tracks highlights from the PDB and the structural biology community.

The **PDB50 BOARD GAME** explores the process of structure discovery.



STRUCTURAL BIOLOGY PLAYING CARDS celebrate PDB and the structural biologists who have determined and deposited many biomolecular structures.

Visit [RCSB.org/pdb50](https://www.rcsb.org/pdb50) for the archive of all PDB50 celebration materials, including



MOLECULE OF THE MONTH
Fifty Years of Open Access to PDB Structures

PDB AND 2021 AT A GLANCE

Users and Impact

RCSB PDB supports an international community of millions of users, including biologists (in fields such as structural biology, biochemistry, genetics, pharmacology); other research scientists (in fields such as bioinformatics); software developers for data analysis and visualization; students and educators (all levels); media writers, illustrators, textbook authors; and the general public.

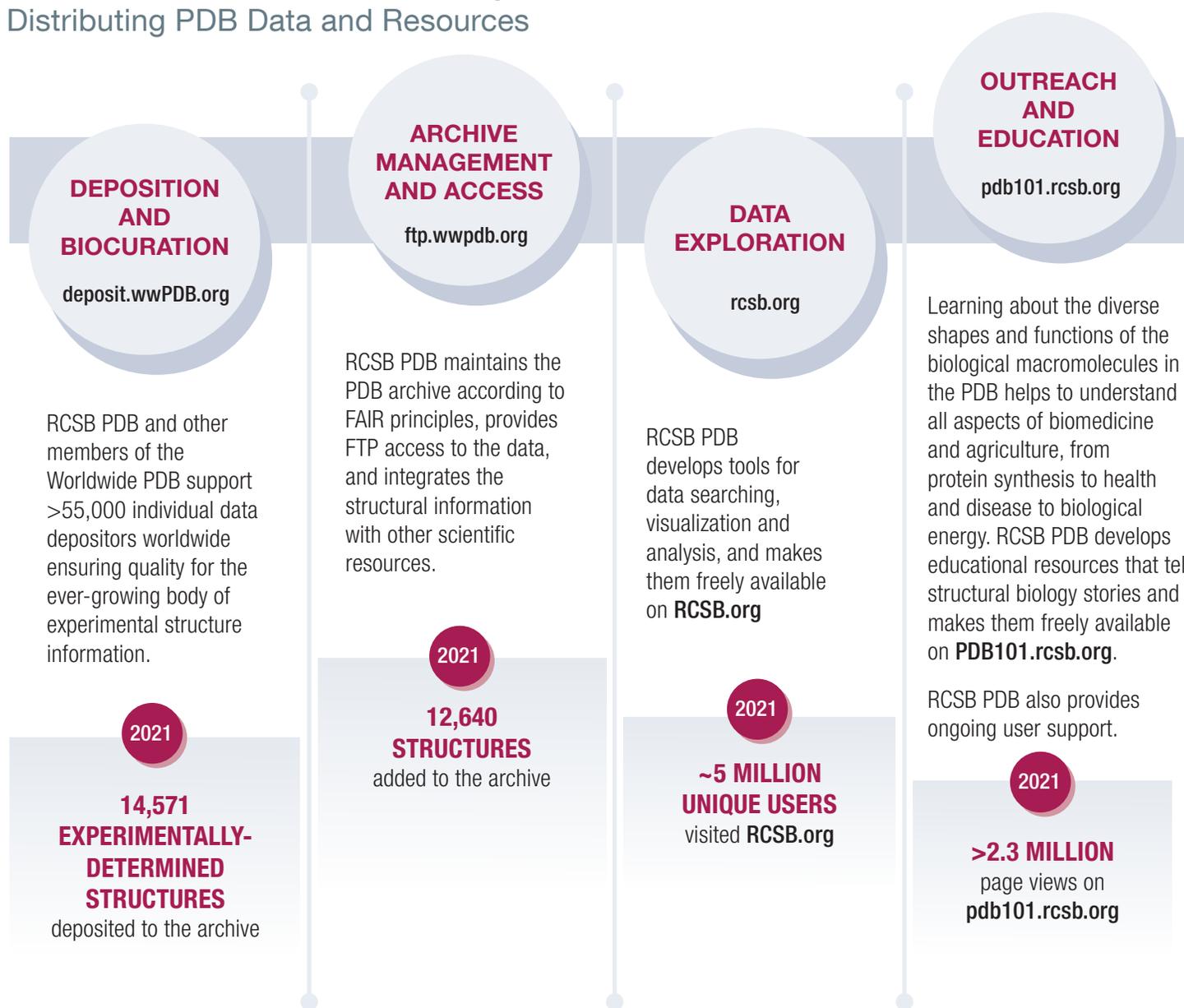
The inaugural RCSB PDB publication (Berman *et al.*, *Nucleic Acids Research* 2000) is one of the top-cited scientific reports of all time.

Beyond using the PDB for safe storage of their data, structural biologists use the PDB for determination of new structures. Homologous structural models are used for molecular replacement. The archive provides models for fitting electron microscopy maps.

Computational biologists utilize the data to: classify and compare structures; understand sequence, structure/function relationships; and predict new structures.

PDB data are important contributors to structure-guided drug design by medicinal and pharmaceutical chemists in both industry and academe. Structural considerations and PDB structures also played a role in the design and success of SARS-CoV-2 mRNA vaccines.

RCSB PDB SERVICES: Sustaining the PDB Core Archive and Distributing PDB Data and Resources



RCSB PDB SERVICES: DEPOSITION AND BIOCURATION

The Worldwide Protein Data Bank (wwPDB) was established to manage a single PDB archive of macromolecular structural data that is freely and publicly available to the global community.



It consists of organizations that act as deposition, data processing and distribution centers for PDB data.

In 2021, 14,571 PDB structures were deposited by researchers from around the world and then prepared for public release by the wwPDB partners RCSB PDB, PDBe, and PDBj. 2,539 new small-molecule ligands and 29 new Biologically Interesting Molecule Reference dictionary items were created.

Biocuration responsibilities are distributed geographically. As the US Data Center, RCSB PDB biocurates structures submitted by scientists working in the Americas and Oceania. During 2021, RCSB PDB processed more than 39% of all incoming structures.

Biocuration of SARS-CoV-2 structures continued to be prioritized; 925 SARS-CoV-2 structures were deposited, biocurated, and released in 2021; a total of 1713 were available by the end of the year.

PDB structures contain 3D atomic coordinates, experimental data, mandatory metadata, authors (e.g., ORCID ID, PI contact information), primary citation, sample preparation, data collection, and structure determination, polymer sequence(s) (proteins, DNA, RNA), and chemical information.

All deposited data undergo expert review and curation. Each structure is examined for self-consistency, standardized using controlled vocabularies, cross-referenced with other biological data resources, and validated for scientific/technical accuracy.

Validation is key to providing high quality data. wwPDB Working Groups and Task Forces include more than 100 academic and industrial volunteers who make recommendations and contribute software tools used to generate wwPDB Validation Reports that assess the quality

and accuracy of every structure stored in the PDB archive. Servers and APIs are provided for independent use by depositors before data submission, and reports are provided during biocuration.

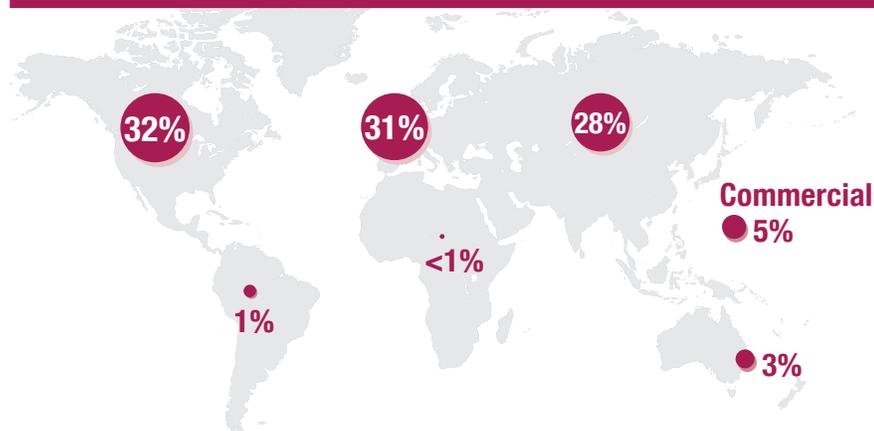
Validation reports can be provided to journal editors and reviewers to help ensure the integrity of peer-reviewed scientific literature. Validation data are also provided publicly to enable meaningful analyses and comparisons across the entire archive.

Reports for new depositions are now provided in PDBx/mmCIF to make validation data more interoperable with the archival format and more “database-friendly”. These reports are easier to interpret, containing a high-level summary and easier access to residue-level information.

Validation reports for all released 3DEM PDB structures and EMDB maps were released in 2021. Recent improvements include images for deposited masks and improved map-model overlay.

2021 Deposition Statistics

14,571 PDB structures were deposited by researchers from around the world



Percentage of Biocurated Structures by wwPDB Processing Site



IMPROVING wwPDB VALIDATION

REPORTS: Validation is a key part of PDB data deposition. Validation Reports are generated to help depositors identify potential discrepancies in their data. During 2021, branched representation and 2D geometric quality images for ligands and carbohydrates, and 3D electron density fits (determined by macromolecular crystallography) for ligands and carbohydrates, restraints validation for NMR, and map volume analysis for 3DEM were added to these reports.

RCSB PDB SERVICES: ARCHIVE MANAGEMENT

The mission of the RCSB PDB is to sustain a unique living data resource of PDB structure information following the FAIR Guiding Principles for scientific data management and stewardship—structure data need to be Findable, Accessible, Interoperable, and Reusable.

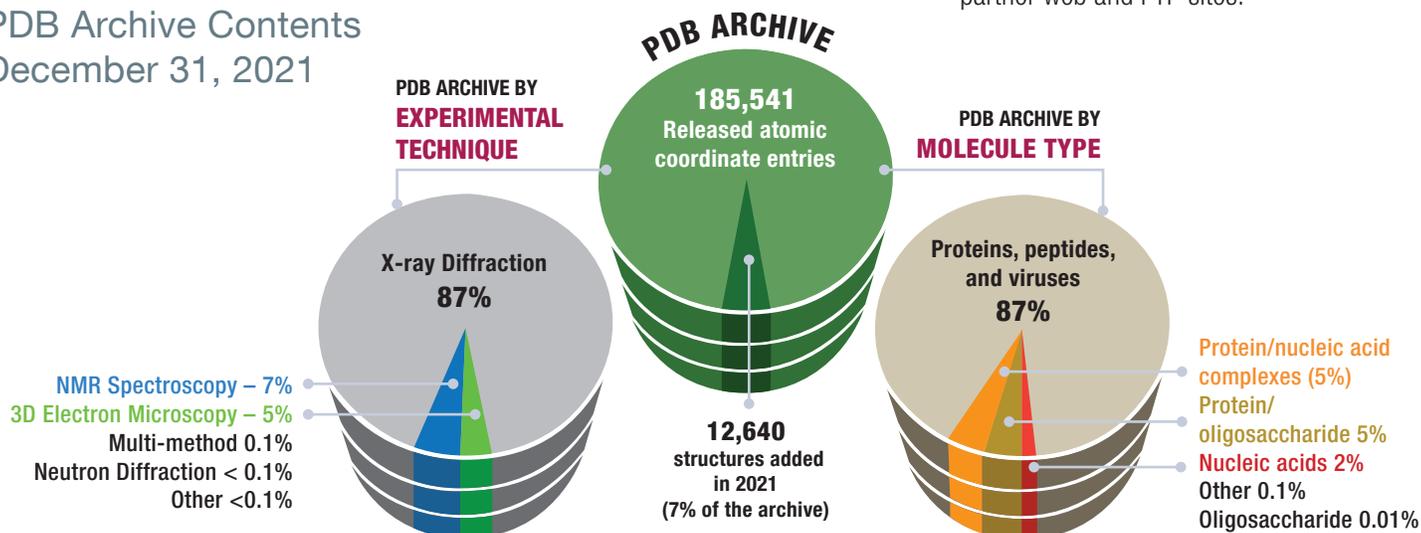
By following these FAIR principles, usage of PDB data and RCSB PDB Services drive patent applications, drug discovery and development, publication of innovative research in scientific disciplines ranging from Agriculture to Zoology, and innovations leading to discovery and development of life-changing biopharmaceutical products.

In 2021, 12,640 new PDB structures were added to the archive, for a total of 185,541 available entries as of January 1, 2022. As wwPDB archive keeper, the RCSB PDB is responsible for safeguarding the PDB archive and maintaining the PDB FTP (<ftp.wwpdb.org>). RCSB PDB coordinates weekly updates of the PDB archive with wwPDB Data Centers in Europe and Japan.

To support **RCSB.org** resources, calculations are run weekly to generate clusters of similar sequences and 3D structures to support search and analysis applications. Data are also integrated with more than 50 external data resources from across the Life Sciences ecosystem.

2.3 billion PDB data files and validation reports were downloaded from all wwPDB partner web and FTP sites.

PDB Archive Contents December 31, 2021



PDB Archive and SARS-CoV-2 Proteins

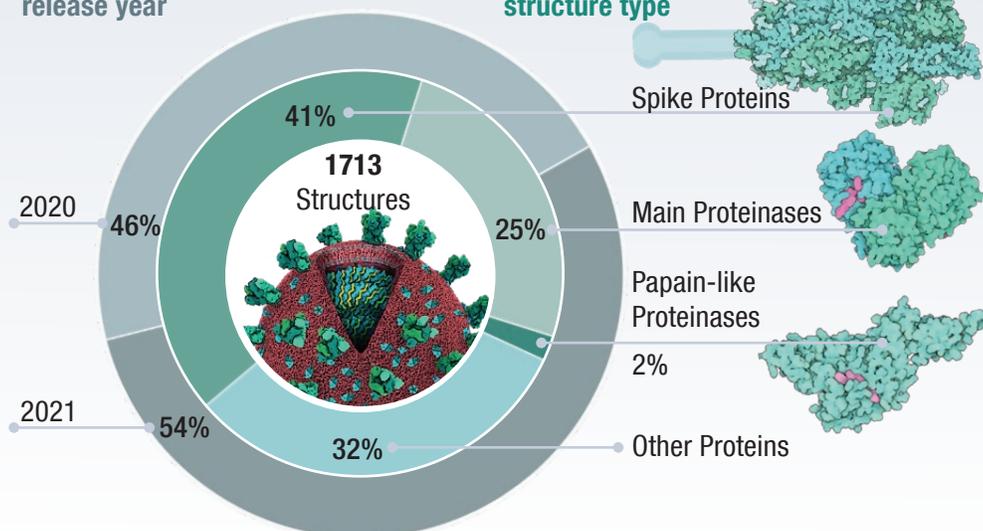
As of the end of 2021, PDB Archive contained over 1,700 SARS-CoV-2 structures. More than 900 were deposited in 2021.

Consistent taxonomy name and ID (**Severe acute respiratory syndrome coronavirus 2; 2697049**) are applied to all COVID-19 structures.

In addition, consistent UniProt referencing is incorporated: **PODTC1, PODTC1, PODTC2, PODTC9.**

Distribution of SARS-CoV-2 structures in the PDB by release year

Distribution of SARS-CoV-2 structures in the PDB by structure type



RCSB PDB SERVICES: DATA EXPLORATION

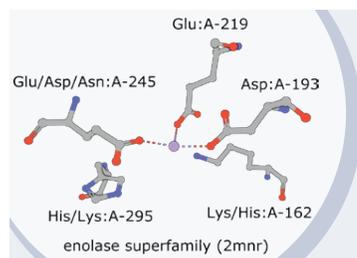
The open-access web portal [RCSB.org](https://www.rcsb.org) supports PDB Data Consumers in the US and around the world with resources for PDB structure access, visualization, and analysis.

The website supports millions of users representing a broad range of skills and interests. In addition to retrieving 3D structure data, PDB users access comparative data, and external annotations, such as information about point mutations and genetic variations.

RCSB PDB services go well beyond the original structure and scientific publication. Each PDB structure is represented by a Structure Summary page that organizes access to important information, including a snapshot of the validation report and other high-level content, annotations,

sequence information, sequence and experimental data. These data are updated weekly, which means that while the corresponding scientific publication remains static, RCSB PDB delivers contemporary views of all structures.

In addition to new features in ligand searches (described on the next page), **RCSB.org** features developed or enhanced in 2021 include:



STRUCTURE MOTIF SEARCH: Similarities between PDB structures can provide valuable functional and evolutionary insights. A new service finds structures with a small number of specific amino acids in specific 3D configurations.



Visual summaries of sequence annotations and structure (**PROTEIN FEATURE VIEW**) and correspondences between PDB entity sequences and reference genomes (**GENOME VIEW**).

Structure ID	Description	Sequence Length	Modeled Residues	Coverage
4HHB.A	Hemoglobin subunit alpha	141	141	99%
10J6.D	Neuroglobin	151	150	93%

ALIGNMENT

RMSD	TM-score	Score	SI%	SS%	Length
2.34	0.78	307.44	23	40	140

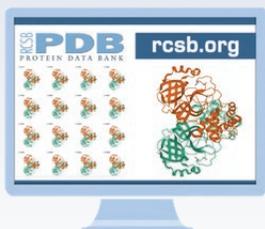
```

1  VLSPADKTNVKAANGKVGAGAGEYGAELERNFLSPTTKTYFFPHFD --- 47
1  MERPE-PELIROSNRAVRSRSPLEHGTVLFAFLFALEPDLPLFQYNGROF 49
48  ---LSHGSAQVKGHGKRVADALTNVAHVDDMPNALGALSDLHAH--KL 91
50  SSPEDSLSSPEFLDHIRKVMVLIDAAVTNVEDLSLEEYLAELGRKRAV 99
92  RYDPPVNFLLSHCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR 141
100 GVKLSSTFVGSLLYMLEKSLGPAFTPATRAANSOLYCAVVQAMSRGWD 149
    
```

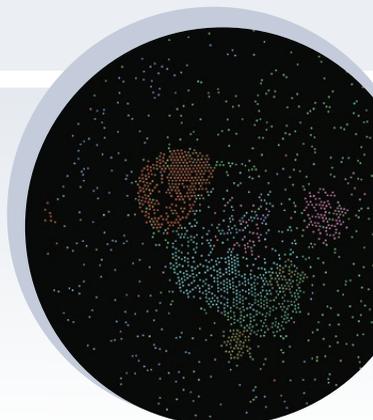
PAIRWISE STRUCTURE ALIGNMENT TOOL: Calculates alignments using different trusted methods and displays sequence alignments and superposed 3D visualization. Comparisons can be made for any protein in the PDB archive and/or uploaded data files including computed structure models.

DATA INTEGRATION: Users can now access information related to PDB structures from glycosylation (GlyYouCan, GlyCosmos, GlyGen) and membrane protein resources (OPM, PDBTM, MemProtMD)

For result sets containing multiple structures representing highly similar proteins, a **NEW GROUPING OPTION** generates a non-redundant search result set based on sequence identity, UniProt ID, or Group depositions.



PDB CITATION MESH NETWORK EXPLORER: Explores co-occurrence networks of MeSH terms associated with PDB entries. This new way of visualizing MeSH terms can provide insights into relationships between PDB structure publications.



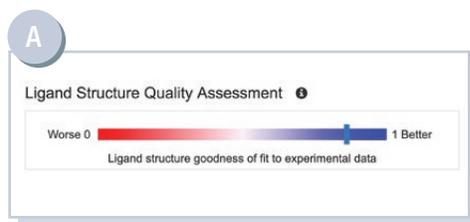
DATA EXPLORATION: EXPLORING SMALL MOLECULES

More than 70% of PDB structures are complexed with small molecule ligands. In 2021, several new RCSB.org features focused on exploration of ligands.

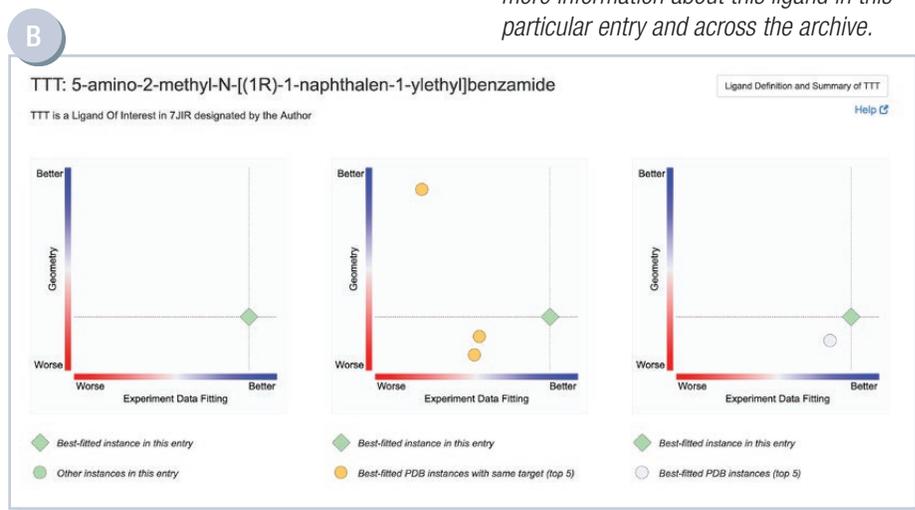
LIGAND STRUCTURE QUALITY ASSESSMENT

New ligand structure quality assessment metrics are available for PDB structures determined by X-ray crystallography. Correlated quality indicators (e.g., RSR, RSCC) have been aggregated into a ranking score that can be used for comparison across the archive. The new indicators allow any user to quickly review ligand structure quality and unambiguously select the ligand/s in a particular PDB structure that will best serve their research or teaching interests.

These complexes represent a range of “quality” in terms of how well the atomic coordinates are supported by experimental data, and how well the ligand 3D structure agrees with known chemistry (e.g., bond lengths, bond angles).



New graphics summarize ligand quality assessments. A) The slider graphic on the Structure Summary page for the PDB structure of a SARS-CoV-2 papain-like proteinase of 7JIR indicates the fitting quality of the ligand of interest. Clicking on the bar will take the user to the Ligands tab (B) for more information about this ligand in this particular entry and across the archive.



LIGAND SEARCHES AND REPORTS

With the new Chemical Sketch Tool, users can draw or edit a molecule (A), then use the resulting SMILES or InChI (B) string to search for matching molecular definitions in the PDB Chemical Component Dictionary.

Chemical descriptors of the molecule currently displayed in the sketch tool

SMILES CC(=O)OC1=CC=CC=C1C(=O)O

InChI InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)(11)12/h2-5H,1H3,(H,11,12)

RCSB.org search results (C) can be displayed as different data types; the Molecular Definitions option will return components as defined in the Chemical Component Dictionary and Biologically Interesting molecule Reference Dictionary. Each definition lists links to the summary page for the component with information and options for download.

AIN

2-(ACETYLOXY)BENZOIC ACID

Formula C9 H8 O4

Molecular Weight 180.157

Type non-polymer

InChIKey BSYNRYMUTXBXSQ-UHFFFAOYSA-N

Identifiers 2-(acetyloxy)benzoic acid, 2-acetyloxybenzoic acid

S7B

2-[(cyclopropylcarbonyl)oxy]-3-methylbenzoic acid

Formula C12 H12 O4

Molecular Weight 220.221

Type non-polymer

InChIKey RHIRFUJRCDAPMZ-UHFFFAOYSA-N

Identifiers 2-[(cyclopropylcarbonyl)oxy]-3-methylbenzoic acid, 2-cyclopropylcarbonyloxy-3-methyl-benzoic acid

M3S

2-(acetyloxy)-3-methylbenzoic acid

Formula C10 H10 O4

Molecular Weight 194.184

Type non-polymer

DOCUMENTATION

is provided for all new features. These RCSB PDB tools and resources provide rich structural views of biological systems to enable breakthroughs in scientific inquiry, medicine, drug discovery, technology and education.

Advanced Search Query Builder

- Full Text
- Structure Attribute
- Chemical Attribute
- Sequence
- Sequence Motif
- Structure Similarity
- Structure Motif
- Chemical

Attribute Se

Sequence E

Structure B:

Chemical S

RCSB PDB SERVICES: OUTREACH AND EDUCATION

PDB-101 (pdb101.rcsb.org) hosts educational materials for a broad audience.

In 2021, PDB-101 ~700,000 users accounted for >2.3 million page views. The most popular feature is the *Molecule of the Month* series that tells stories about molecular structure and function, their diverse roles within living cells, and the growing connections between biology and nanotechnology. This content has inspired readers around the world, and is a regular read for students and researchers alike.

Several *Molecule of the Month* features published in 2021 were co-written by undergraduate and graduate students that

participated in a science communication “boot camp” hosted by RCSB PDB.

Other additions to PDB-101 in 2021 included videos about opioids and experimental methods in structural biology (image 1); a feature to help users understand how carbohydrates are represented in the PDB archive; and 3D print model (image 2) and lesson plan on serum albumin.

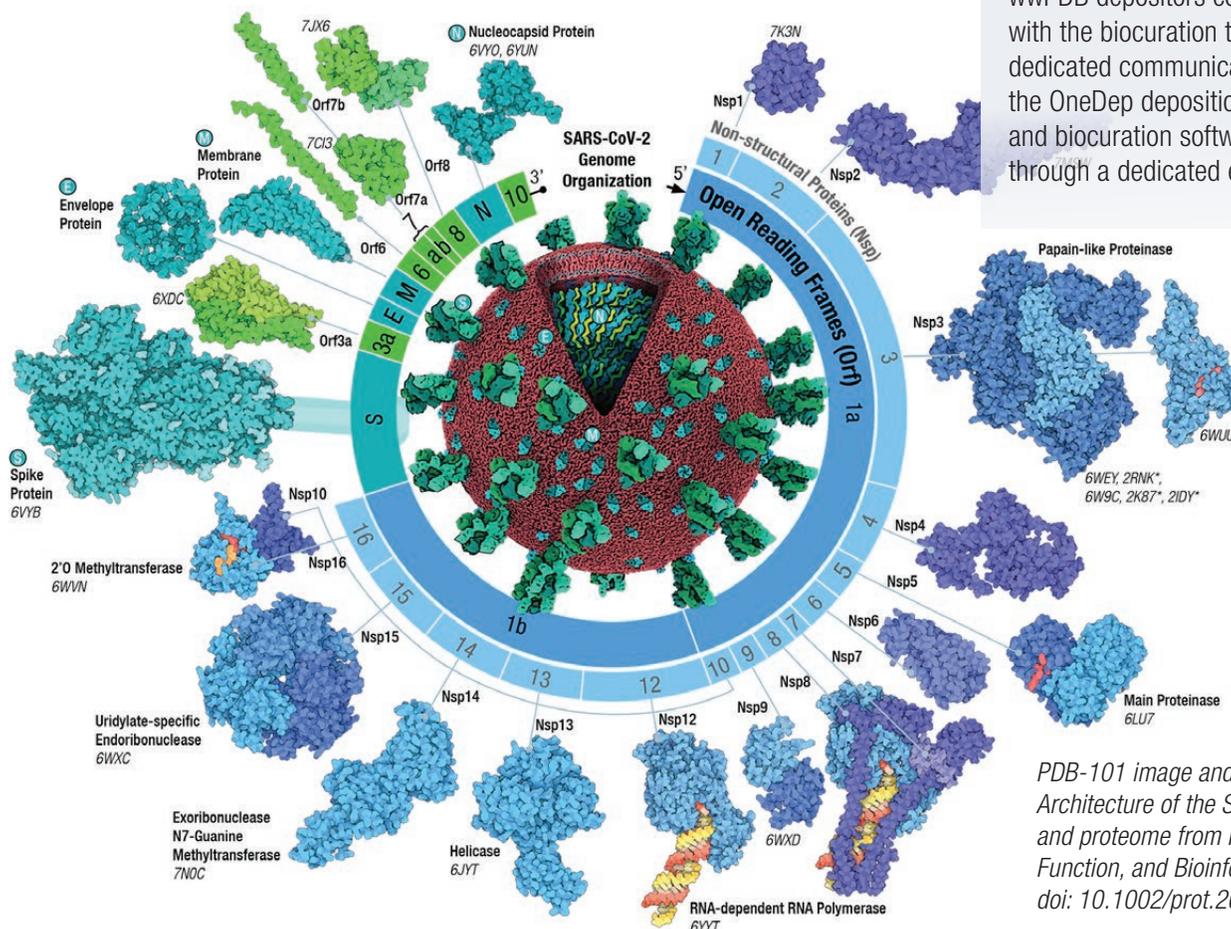
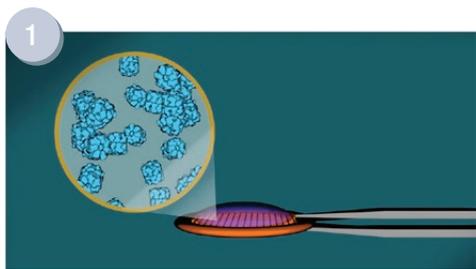
PDB-101 hosts a collection of materials for learning about SARS-CoV-2 at pdb101.rcsb.org/browse/coronavirus.

Customer Service

In the past 50 years, correspondence with users has kept pace with the times. Early on, interactions with depositors and users took place by letter and phone. Beginning in the 1990s, correspondence blossomed as email became available and website technology made PDB data more accessible. Now in the 21st century, customer service has evolved to help users worldwide access the many tools available to explore the rich depth of structural information stored in the PDB.

The **RCSB.org** customer service department engages in email conversations with beginning students, power user experts, and everyone in between. This user input drives development of existing and future features.

wwPDB depositors correspond with the biocuration team using a dedicated communication module in the OneDep deposition, validation, and biocuration software system and through a dedicated email help desk.



PDB-101 image and flyer adapted from *Architecture of the SARS-CoV-2 genome and proteome from Proteins: Structure, Function, and Bioinformatics 2021*; doi: 10.1002/prot.26250

RCSB PDB SERVICES: OUTREACH AND EDUCATION

Health Focus: DRUGS AND THE BRAIN

PDB-101 adopts a bi-annual public health theme to focus development and student engagement. The theme for 2020-2021 was Drugs and the Brain.

The theme was used at Rutgers for an Undergraduate Honors Seminar Course exploring the foundations of structural biology and for an annual RCSB PDB-sponsored video contest.

For this challenge, high school students focused on the molecular mechanisms of drugs for mental disorders and created a variety of animations exploring Alzheimer's Disease, anxiety, depression, and schizophrenia.

Browse PDB-101 materials on Drugs and the Brain at pdb101.rcsb.org/browse/drugs-and-the-brain



The painting "Excitatory and Inhibitory Synapses", 2018 (only fragment shown) from PDB-101 David Goodsell Gallery is one of the resources to learn about nervous system.

The Health Focus for 2022-23 will be Cancer Biology and Therapeutics

2021 PUBLICATIONS

RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences (2021)
Nucleic Acids Research 49: D437–D451
doi: 10.1093/nar/gkaa1038

Molecular storytelling for online structural biology outreach and education (2021)
Structural Dynamics 8: 020401
doi: 10.1063/4.0000077

Painting a Portrait of SARS-CoV-2 (2021)
American Scientist 109: 88-93
doi: 10.1511/2021.109.2.88

Enhanced validation of small-molecule ligands and carbohydrates in the Protein Data Bank (2021)
Structure 29: 393-400.e391
doi: 10.1016/j.str.2021.02.004

CellPAINT: Turnkey Illustration of Molecular Cell Biology (2021)
Frontiers in Bioinformatics 1
doi: 10.3389/fbinf.2021.660936

Impact of structural biologists and the Protein Data Bank on small-molecule drug discovery and development (2021)
Journal of Biological Chemistry 296: 100559
doi: 10.1016/j.jbc.2021.100559

RCSB Protein Data Bank: Architectural Advances Towards Integrated Searching

and Efficient Access to Macromolecular Structure Data from the PDB Archive (2021)
Journal of Molecular Biology 443: 166704
doi: 10.1016/j.jmb.2020.11.003

Synergies between the Protein Data Bank and the community (2021)
Nature Structural & Molecular Biology 28: 400–401
doi: 10.1038/s41594-021-00586-6

Art as a tool for science (2021)
Nature Structural & Molecular Biology 28: 402–403
doi: 10.1038/s41594-021-00587-5

Seeing the PDB (2021)
Journal of Biological Chemistry 296:100742
doi: 10.1016/j.jbc.2021.100742

Open-access data: A cornerstone for artificial intelligence approaches to protein structure prediction (2021)
Structure 29: 515-520
doi: 10.1016/j.str.2021.04.010

Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures (2021)
Nucleic Acids Research 49: W431–W437
doi: 10.1093/nar/gkab314

Transactions from the 70th Annual Meeting of the American Crystallographic Association: Structural Science-New Ways to Teach the Next Generation (2021)
Struct Dyn 8: 040401
doi: 10.1063/4.0000125

Design and proof of concept for targeted phage-based COVID-19 vaccination strategies with a streamlined cold-free supply chain (2021)
Proceedings of the National Academy of Sciences 118: e21057391181
doi: 10.1073/pnas.2105739118

Modernized uniform representation of carbohydrate molecules in the Protein Data Bank (2021)
Glycobiology 31: 1204–1218
doi: 10.1093/glycob/cwab039

wwPDB biocuration: on the front line of structural biology (2021)
Nature Methods 18: 431–432
doi: 10.1038/s41592-021-01137-z

Predicting Proteome-Scale Protein Structure with Artificial Intelligence (2021)
N Engl J Med 385: 2191-2194
doi: 10.1056/NEJMcibr2113027

Picturing science: using art and imagination to explore new worlds (2021)
Biochem (Lond) 43: 32–38
doi: 10.1042/bio_2021_156

Evolution of the SARS-CoV-2 proteome in three dimensions (3D) during the first 6 months of the COVID-19 pandemic (2021)
Proteins Structure, Function, and Bioinformatics Sep 27:10.1002/prot.26250
doi: 10.1002/prot.26250

RCSB PDB is managed by the members of the Research Collaboratory for Structural Bioinformatics: Rutgers, UCSD/SDSC, and UCSF



RCSB PDB is a member of the wwPDB organization
wwPDB.org



PDB is a CoreTrustSeal Board certified Trusted Digital Repository.

CITE RCSB PDB

The Protein Data Bank (2000) *Nucleic Acids Res* **28**: 235-242.
doi: [10.1093/nar/28.1.235](https://doi.org/10.1093/nar/28.1.235).

RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences (2021) *Nucleic Acids Research* **49**: D437–D451
doi: [10.1093/nar/gkaa1038](https://doi.org/10.1093/nar/gkaa1038).

FUNDING

RCSB PDB is funded by National Science Foundation (DBI-1832184), the US Department of Energy (DE-SC0019749), and the National Cancer Institute, National Institute of Allergy and Infectious Diseases, and National Institute of General Medical Sciences of the National Institutes of Health under grant R01GM133198.

FOLLOW US



[/RCSBPDB](https://www.facebook.com/RCSBPDB)



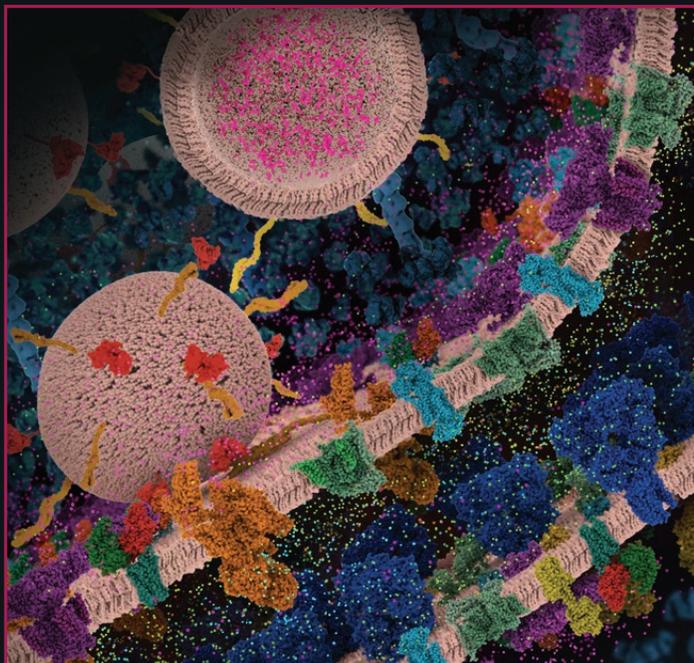
[/RCSBProteinDataBank](https://www.youtube.com/RCSBProteinDataBank)



[/buildmodels](https://twitter.com/buildmodels)



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



ABOUT THE COVER

This illustration shows a fragment of a glutamatergic synapse. It highlights a few spherical pre-synaptic vesicles that carry the neurotransmitter glutamate (shown in pink in the cross-section). The pre-synaptic and post-synaptic membranes are shown with proteins relevant for transmitting and modulating the neuronal signal. The images of the proteins were created based on experimentally-determined structures from the Protein Data Bank. The pre-synaptic membrane features the voltage-gated sodium channel (green; 5x0m, 6j8j), the voltage-gated potassium channel (purple; 6pbx), the sodium-potassium pump (blue; 2zxe), the voltage-gated calcium channel (orange; 6jp5), and opioid receptor with G protein (aqua with dark orange and green; 4dkl). The post-synaptic membrane features in addition the glutamate (AMPA) receptor (navy blue; 5wek), and adenylyl cyclase (yellow; 6r3q). Potassium is shown in magenta, sodium in lime green, and calcium in cyan.

This illustration was adapted from the PDB-101 Video *Opioids and Pain Signaling* released in 2021.