Celebrating

Lysozyme is a small enzyme that protects us from infections by attacking the protective cell walls of bacteria. It was the first enzyme structure to be determined via X-ray crystallography and one of the historic structures that set the foundation for the PDB archive.

This image was painted by Irving Geis, an artist who helped illuminate the field of structural biology with his iconic images. This particular illustration appeared on the cover of *Scientific American* in 1966.

Used with permission from the Howard Hughes Medical Institute (www.hhmi.org). All rights reserved.
The HIV-1 trans-activation response (TAR) element is an RNA element required for virus replication. TAR adopts a highly dynamic structure. The extent and timescale of the molecular motion is revealed by NMR. Knowledge of the dynamics is aiding the discovery of small molecules that bind TAR as potential therapeutics against HIV.

This image is a long exposure capture of an experimentally derived atomic-resolution ensemble of conformations. The ensemble was generated by using molecular dynamics simulations to generate a library of conformations and then using information about the anisotropy of motions from residual dipolar couplings measured by NMR to guide selection of conformers from the library.
A virus infects living cells and forces them to create many new copies of the pathogen. Viruses are often very simple, composed of a small genome that encodes only a few proteins, including a coat that finds a cell and infects it with the genome. Atomic structures have revealed both the inner workings of viruses and new avenues for the creation of drugs and vaccines to fight them.

Many viruses are icosahedral in form, which can be modeled in 3D using Origami, as shown here.
Photosensitive epilepsy is not common, however it still infiltrates people’s lives and affects their experiences every day, with seizures triggered by flashing or flickering lights. Glutamate, a type of amino acid, is a major excitatory neurotransmitter that plays a key role in the initiation and spread of seizure activity. This can be controlled by blocking the uptake of glutamate through glutamate receptors, which are, therefore, promising targets for epilepsy therapy.

Here, the light bulb filament is represented by a 3D image of the glutamate receptor protein, highlighting its significance in fighting seizures triggered by light.
Zika virus infects people around the globe. For most, the virus causes a mild illness that is quickly fought off by the immune system. But a connection between Zika infection in pregnant women and birth defects has underscored the need to find ways to fight the disease.

The first atomic-level structure of the Zika virus, shown here in pink, came from the lab of the pioneering structural biologist Michael Rossmann. Structures in the PDB archive help researchers better understand how the virus enters human cells after mosquito-borne or sexual transmission.
Cytochrome c is an essential protein in the electron transport chain, a collection of protein-ligand complexes involved in production of cellular energy. It is used to shuttle electrons between complexes using a bound heme molecule, shown in red on this sculpture, to pick up and pass on these electrons.

Cytochrome c was the first structure deposited to the PDB from Asia. The model shown here was created from Japanese wood when the structure was first determined at 4 Å resolution in 1971. The atomic coordinates at a higher resolution of 2.5 Å were deposited to PDB in 1975.

<table>
<thead>
<tr>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>31</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>
Breast cancer is the most common cancer among women. Understanding the complex underlying molecular processes involved in this disease is vital to finding therapeutics to combat this cancer. One important target protein in the quest to understand this disease is BRCA1 (Breast cancer type 1 susceptibility protein), a tumor suppressor protein, found predominantly in breast tissue and responsible for repairing DNA.

Depicted here is the BRCA1 protein interacting with DNA.
Protein structure determines the nature of the interactions with other proteins that are important in biological processes. Not all strong protein-protein interactions require well-defined structure, however. The complex between linker histone H1 and Prothymosin exploits highly charged intrinsically-disordered regions (IDRs) and electrostatic interactions averaged over many conformations to form an ultra-high affinity complex with extensive structural fluctuations on extremely fast timescales.

NMR spectroscopy, single-molecule FRET, and molecular dynamics simulations were integrated to characterize the dynamic properties of the disordered complex.
Gap junctions are tubular arrangements of proteins that connect two cells together, allowing exchange of ions and small molecules. They are essential in neural activity and cell synchronization, especially in cardiac muscle, where they regulate the cardiac cycle.

The gap junction structure shown here consists of twelve connexin-26 proteins and was determined at 3.5 Å resolution by X-ray crystallography. The hourglass-like shape resembles a Japanese hand drum known as a Tsuzumi, also shown in the image.
Oxygen is one of the quintessential molecules of life. Upon entering the lungs, oxygen binds hemoglobin, an iron-rich protein in blood, which carries it to the tissues in the body. Hemoglobin is made up of four polypeptide chains, two alpha and two beta chains, each containing a heme group bound to an iron atom. Oxygen binds reversibly to these iron atoms and is transported through blood. It is the oxygenated state of hemoglobin, called oxyhemoglobin, that gives blood its bright red color, whereas the reduced deoxygenated state is purplish blue in color.

Hemoglobin was one of the very first protein structures to be determined by X-ray crystallography, earning Max Perutz the 1962 Nobel Prize in Chemistry.
Using triple-helices, collagen forms long ropes and tough sheets that lend structure and elasticity to our bodies. This illustration depicts Type IV collagen-containing basement membrane, which supports skin and many organs. The globular heads at one end of the molecule bind strongly together and at the other end, four collagen molecules associate together through their tails, forming an X-shaped complex. Using these two interactions, type IV collagen forms an extended network (light blue). Two other molecules—cross-shaped laminin (blue-green) and long, snaky proteoglycans (green)—fill in the spaces, forming a dense sheet.

The first atomic resolution structure of a collagen triple-helix was determined in the laboratory of Helen M. Berman, one of the original founders of the PDB archive.
A capsid is the protein shell of a virus that protects and conveys the viral genetic material (DNA or RNA). These complex structures can be investigated by integrating multiple structural biology approaches that probe different length scales (commonly referred to as integrative methods). The atomic-resolution structure of HIV-1 capsid protein (CA) tubes was determined by magic-angle-spinning NMR spectroscopy integrated with low resolution cryoEM data and molecular dynamics simulations. The structure of a single polypeptide chain was calculated using interatomic distance and angle restraints determined by NMR. An atomic-resolution structure of the CA hexamer was derived by integrating the NMR restraints and the low-resolution (8.0 Å) cryo-EM map of the CA hexamer in a tubular assembly. The final structure of a CA tube was determined by data-guided molecular dynamics simulations.

The landmark PDB structure of the HIV capsid shown here and released in 2013 contains a record 2,440,800 explicitly-modeled atoms (PDB ID 3J3Q).
This research was described in L. Salmon et al. (2013) J. Am. Chem. Soc. 135: 5457-5466.

**Primary data**: PDB Structure 1anr


doi: 10.1093/nar/24.20.3974

---

**Primary data**: PDB Structure 5i1f


doi: 10.1016/j.neuron.2016.08.012

---

**Primary data**: PDB Structure 1uf2


doi: 10.1016/j.str.2003.08.012

---

**Primary data**: PDB Structure l1f


doi: 10.1016/j.neuron.2016.08.012

---

**Primary data**: PDB Structure 5ire


doi: 10.1126/science.aaf5316

---

Photo credit: Takahiro Kudou, PDBj

The model is owned by the Institute for Protein Research, Osaka University.
The image shows snapshots from the ensemble of the complex obtained from a coarse-grained molecular dynamics simulation and rendered using PyMol, where the local charge is shown in blue (for positive) and red (for negative), highlighting the polyelectrolyte properties of the complex. Representative structures of the complex are illustrated to stress the diversity of the ensemble. The image background was made by combining hand-painting and deep learning for artistic style transfer to remind us that even this dynamic protein complex is functional in the environment of a cell, where a multitude of other structures impact its behavior. These cellular structures have been drawn to resemble peaks in an artistic NMR spectrum to emphasize the importance of NMR to biology.

**Primary data:**
BMRB entries 27216, 27215


doi: 10.1038/nature25762.

**Image credit and additional information:**
Tengyu Zhao, a 14-year-old student from Perse School, Cambridge, UK, created this artwork. She has a keen interest in biology and aspires to become a marine biologist. This dry-point etch print was inspired by the crystal structure of tumor suppressor gene BRCA1. Using DNA as the central piece, she attempts to portray the connections of nucleotide bases of DNA.

**Primary data:**
PDB Structure 3k16


*Structure* 18:167-76.

**Image credit:**
PDBj

**Primary data:**
PDB Structures 3j3q, 3j3y


doi: 10.1038/nature12162

* 3D structural data were the primary inspiration for each of the illustrations in this calendar. The data were enriched by integrating additional scientific information to create contextual visualizations, or by creating visual metaphors based on intrinsic qualities of the 3D structure. In some cases three-dimensional objects reflecting the overall shape of the protein were created and photographed. To explore any of the PDB structures used as inspiration for these illustrations in 3D, choose one of the wwPDB member pages–rcsb.org, pdbe.org, or pdbj.org–and enter the PDB structure ID into the Search interface. To explore the BMRB entries, enter the BMRB entry ID on bmrb.io.

**Primary data:**
PDB Structures 1cag


*Science* 266:75-81.
doi: 10.1126/science.7956699

These artworks were created as part of the PDB Art project, a collaboration between Protein Data Bank in Europe (PDBe), The Art Society CANTAB, and The Art Society GVANTA, who together work with local school students and art departments to support the creation of artworks inspired by the molecules of life. Find out more at PDBe.org/art.
In the late 1950s, scientists began to decipher the 3D shapes of proteins at the level of individual atoms. The implications of these new structures for science inspired the new field of Structural Biology. Imagining the potential research enabled by archiving and sharing data from these experiments moved the scientific community to action.

In June 1971, a symposium on *Structure and Function of Proteins at the Three Dimensional Level* was held at Cold Spring Harbor Laboratory. With its lively conversations, debates, and planning for the future, that meeting defined the beginning of the PDB as an archive for the experimentally-determined 3D structures of biological macromolecules. Later that year, the Protein Data Bank archive was established as the first open-access, digital resource for biology.¹

As PDB enters 2021, the archive contains and supports online access to ~170,000 of biomacromolecular structures determined via macromolecular crystallography, Nuclear Magnetic Resonance spectroscopy, and 3D Electron Microscopy by researchers from around the world. These structures help researchers understand fundamental biology, biomedicine, biological energy, and biotechnology, and other research pertaining to global prosperity and environmental sustainability.

2021 marks the 50th anniversary of the PDB archive

The PDB archive is managed by the Worldwide Protein Data Bank (wwPDB), a consortium of organizations that host deposition, annotation, and distribution centers for PDB data and collaborate on a variety of projects and outreach efforts.²,³

The Worldwide Protein Data Bank Foundation supports the outreach activities of the wwPDB, including symposia and events that will be held in 2021. Individual and institutional donations are critical for future wwPDB outreach efforts.