

# Molecular Explorations of Cancer Biology and Therapeutics at PDB-101

David S. Goodsell,<sup>1,2,3,4</sup>, Shuchismita Dutta,<sup>1,2,3</sup>, Maria Voigt,<sup>1,2</sup> Christine Zardecki,<sup>1,2\*</sup>  
and Stephen K. Burley<sup>1,2,3,5</sup>

<sup>1</sup>Research Collaboratory for Structural Bioinformatics Protein Data Bank, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

<sup>2</sup>Institute for Quantitative Biomedicine, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

<sup>3</sup>Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901, USA

<sup>4</sup>Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

<sup>5</sup>Research Collaboratory for Structural Bioinformatics Protein Data Bank, San Diego Supercomputer Center, University of California, La Jolla, CA 92093, USA

\*Corresponding Author: christine.zardecki@rcsb.org

The Protein Data Bank (PDB) was established in 1971 as the first open access digital data resource in biology [1]. It currently houses nearly 190,000 three dimensional (3D) structures of biological macromolecules (proteins, DNA, and RNA, and their complexes with one another and small molecule drugs, etc.), providing valuable insights into biochemical and biological function. Now in its 51<sup>st</sup> year of continuous operations, the PDB has become a leading global resource for experimental data central to discovery across fundamental biology, biomedicine, energy sciences, and bioengineering/biotechnology [2]. By providing open access to 3D structure data for the molecules of life, the PDB archive enables understanding of how normal cell growth is controlled, and how cancer cells bypass these essential controls to proliferate and metastasize. Open access to PDB data also facilitates discovery and development of novel anti-cancer agents, many of which are the product of structure-guided drug design using the tools of macromolecular crystallography [3].

PDB-101 (PDB101.rcsb.org) is an online web portal designed for educators, students, and the general public to promote exploration of the shapes, interactions, and of proteins and nucleic acids [4] ("101", as in an entry-level course). It was established in 2011 by the RCSB Protein Data Bank [2, 5, 6], a global resource for the advancement of research and education in biology and medicine and the US data center for the Worldwide PDB partnership [7, 8].

Every two years, the RCSB PDB selects a biennial public health theme to focus PDB-101 content development and educator/student engagement. The theme for 2022 and 2023 explores the molecular mechanisms of *Cancer Biology and Therapeutics* and is reflected across the web portal.

The *Molecule of the Month* article series, launched by RCSB PDB in 2000, has introduced millions to the shape and function of 3D structures archived in the PDB. Each of the 270 installments freely available on PDB101.rcsb.org includes an introduction to the structure and function of a PDB molecule, discussion of the relevance of the molecule to human health and welfare, and

suggestions as to how PDB-101 visitors can view these structures and access further details. While topical features such as Coronavirus and Opioid Receptors draw large audiences, articles related to the topics commonly addressed in classrooms (e.g., hemoglobin, catalase) continue to be highly accessed year after year.

In January 2022, an intensive science communication “boot camp” was co-hosted virtually by the Rutgers Institute for Quantitative Biomedicine (IQB) and the RCSB PDB. Undergraduate and graduate students researched pre-selected topics in Cancer Biology and Therapeutics, including HER2/neu and Trastuzumab [9]; Nicotine, Cancer, and Addiction [10]; Non-Homologous End Joining Supercomplexes [11]; Pyruvate Kinase M2 and the Warburg Effect [12]; Secretory Antibodies [13]; and Vascular Endothelial Growth Factor (VegF) and Angiogenesis [14]. Boot camp participants hailed from the following colleges and universities around the United States: Fort Lewis College, CO; Hampton University, VA; Hunter College, NY; Nova Southeastern University, FL; Montclair State University, NJ; Pennsylvania State University, PA; Rutgers University, NJ; and University of Puerto Rico-Mayaguez, PR.

For each of the Cancer-related *Molecule of the Month* articles, a team of four students identified exemplar structures from the PDB archive to help tell the molecular story and then created stunning molecular visualizations (Figure 1). Articles were co-authored under the guidance of Rutgers IQB faculty (coauthors D.S.G., S.D., and S.K.B.), and then externally reviewed for accuracy by biomedical researchers expert in the subject matter.

PDB-101 also hosts an annual Video Challenge for High School students (<https://pdb101.rcsb.org/events/video-challenge/the-challenge>). For the 2022 challenge, participating students were asked to explore two molecular signaling systems important in cancer: the p53/p21 pathway and the EGFR/Ras pathway. In all, 38 teams created short videos, each telling a coherent story explaining scientific concepts and touching on public health aspects of cancer, such as screening, prevention, and awareness. External expert reviewers selected the winning videos based on scientific content and communication effectiveness, while the public was invited to vote for a “Viewer’s Choice” video. Award winning entries can be found on the PDB-101 web portal at <https://pdb101.rcsb.org/events/video-challenge/2022-awards>.

All cancer-related educational materials available from PDB-101 can be found at <https://pdb101.rcsb.org/browse/cancer>. To support and provide context for these materials, PDB-101 also offers diverse educational curricula and activities to introduce young scientists to structural biology, comprehensive search and browse capabilities to help visitors find information, and a “Guide to Understanding PDB Data” to help visitors access and explore primary structural data stored in the PDB archive.

## ACKNOWLEDGEMENTS

We thank the students who participated in the 9<sup>th</sup> Annual IQB Winter Boot Camp and the 2022 PDB101 Video Challenge. RCSB PDB core operations are jointly funded by the National Science Foundation (DBI-1832184, PI: S.K. Burley), the US Department of Energy (DE-SC0019749, PI: S.K. Burley), and the National Cancer Institute, the National Institute of Allergy and Infectious Diseases, and the National Institute of General Medical Sciences of the National Institutes of Health (R01GM133198, PI: S.K. Burley).

## AUTHORSHIP

All authors contribute to the development of PDB-101 resources and to the writing of this editorial. DSG is the primary creator of the Molecule of the Month series, with SD and SKB contributing to the IQB Boot Camp articles. MV is the lead of the Video Challenge project.

## COMPETING INTERESTS

The authors declare no competing interests.

## REFERENCES

- 1 Protein Data Bank. Crystallography: Protein Data Bank. *Nature (London), New Biol* 1971; 233: 223-223.
- 2 Burley SK, Bhikadiya C, Bi C, Bittrich S, Chen L, Crichtlow G *et al.* 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. *Nucleic Acids Research* 2021; 49: D437-D451.
- 3 Westbrook JD, Soskind R, Hudson BP, Burley SK. Impact of Protein Data Bank on Anti-neoplastic Approvals. *Drug Discov Today* 2020; 25: 837-850
- 4 Zardecki C, Dutta S, Goodsell DS, Lowe R, Voigt M, Burley SK. PDB-101: Educational resources supporting molecular explorations through biology and medicine. *Protein Sci* 2022; 31: 129-140.
- 5 Burley SK, Bhikadiya C, Bi C, Bittrich S, Chen L, Crichtlow GV *et al.* RCSB Protein Data Bank: Celebrating 50 years of the PDB with new tools for understanding and visualizing biological macromolecules in 3D. *Protein Sci* 2022; 31: 187-208.
- 6 Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H *et al.* The Protein Data Bank. *Nucleic Acids Res* 2000; 28: 235-242.
- 7 Berman HM, Henrick K, Nakamura H. Announcing the worldwide Protein Data Bank. *Nature Structure Biology* 2003; 10: 980.
- 8 wwPDB consortium. Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res* 2019; 47: D520-D528.
- 9 De Leon Cruz S, Herrod A, Park KH, Wu Wu A, Goodsell DS, Burley SK. HER2/neu and Trastuzumab *RCSB PDB Molecule of the Month* 2022: doi: 10.2210/rcsb\_pdb/mom\_2022\_4.
- 10 Choudhry K, Muse D, De Maio DP, Soto Acevedo AA, Goodsell DS, Dutta S. Nicotine, Cancer, and Addiction. *RCSB PDB Molecule of the Month* 2022: doi: 10.2210/rcsb\_pdb/mom\_2022\_5.

11 Diaz-Figueroa G, Egozi M, Jannath S, Maddy J, Goodsell DS. Non-Homologous End Joining Supercomplexes. *RCSB PDB Molecule of the Month* 2022: doi: 10.2210/rcsb\_pdb/mom\_2022\_7.

12 Ahmed; F, Ash; J, Patel; T, Sanders; A, Goodsell; DS, Dutta S. Pyruvate Kinase M2. *RCSB PDB Molecule of the Month* 2022: doi: 10.2210/rcsb\_pdb/mom\_2022\_6.

13 Brooks L, Colón-Colón C, Patel A, Polat A, Goodsell DS. Secretory Antibodies. *RCSB PDB Molecule of the Month* 2022: doi: 10.2210/rcsb\_pdb/mom\_2022\_8.

14 Cartagena; E, Gelashvili; M, Keyes; J, Rosenzweig; E, Goodsell; DS, Burley SK. Vascular Endothelial Growth Factor (VegF) and Angiogenesis. *RCSB PDB Molecule of the Month* 2022: doi: 10.2210/rcsb\_pdb/mom\_2022\_3.

15 Cho HS, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW, Jr. *et al.* Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* 2003; 421: 756-760.

16 Aertgeerts K, Skene R, Yano J, Sang BC, Zou H, Snell G *et al.* Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein. *J Biol Chem* 2011; 286: 18756-18765.

17 Mineev KS, Bocharov EV, Pustovalova YE, Bocharova OV, Chupin VV, Arseniev AS. Spatial structure of the transmembrane domain heterodimer of ErbB1 and ErbB2 receptor tyrosine kinases. *J Mol Biol* 2010; 400: 231-243.

## FIGURE LEGENDS

### **Figure 1: Images from cancer-focused Molecule of the Month articles.**

(left) Artistic conception of VegF signaling. VegF (magenta, top left) arrives at the potential site of a new blood vessel by traveling through the blood plasma (tan). VegF brings together two copies of VegFR (top center, lavender/yellow) to form an active dimer. Active VegFR then initiates a signal cascade that leads to intracellular phosphorylation of many proteins, including cadherin (green). The phosphorylated cadherins separate, making room for new blood vessels. Full image available at PDB-101 and in the article on Vascular Endothelial Growth Factor (VegF) and Angiogenesis.[14] (right) Trastuzumab (red/pink) antibody bound to HER2 (blue). The cell membrane is shown schematically in gray. The illustration is built from three PDB structures: the extracellular domain with the antigen-binding fragment (Fab) of trastuzumab (1n8z [15]), the kinase domain inside the cell (3pp0 [16]), and the transmembrane domain (2ks1 [17]). Image available from the article on HER2/neu and Trastuzumab [9].

Filename: Oncogene\_Editorial\_PDB101\_07202022.docx  
Directory: /Users/zardecki/Downloads  
Template: /Users/zardecki/Library/Group  
Containers/UBF8T346G9.Office/User  
Content.localized/Templates.localized/Normal.dotm  
Title:  
Subject:  
Author: Christine Zardecki  
Keywords:  
Comments:  
Creation Date: 7/20/22 9:43:00 AM  
Change Number: 4  
Last Saved On: 7/20/22 10:45:00 AM  
Last Saved By: Christine Zardecki  
Total Editing Time: 5 Minutes  
Last Printed On: 2/1/23 9:30:00 AM  
As of Last Complete Printing  
Number of Pages: 5  
Number of Words: 1,621  
Number of Characters: 22,236 (approx.)