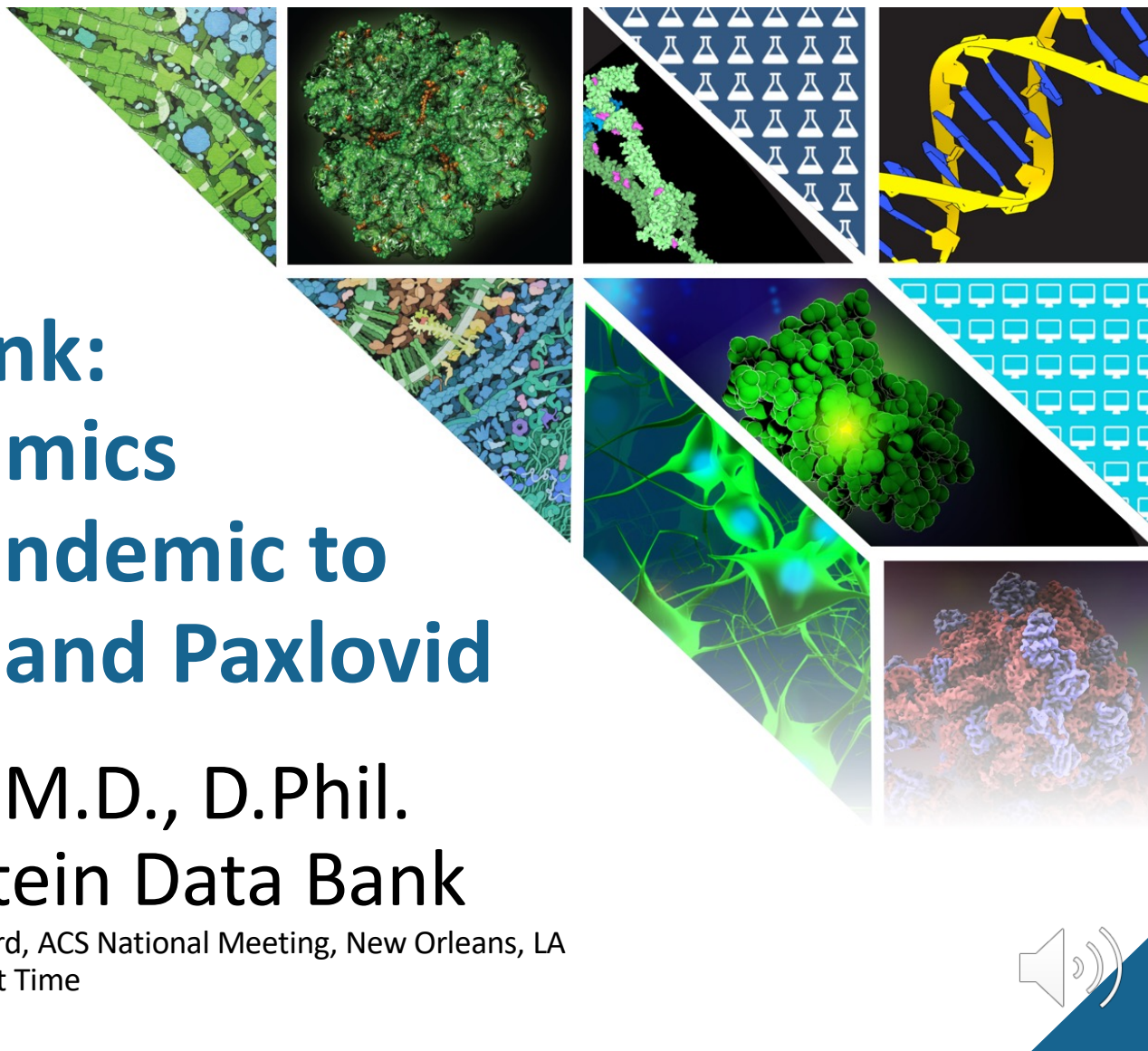


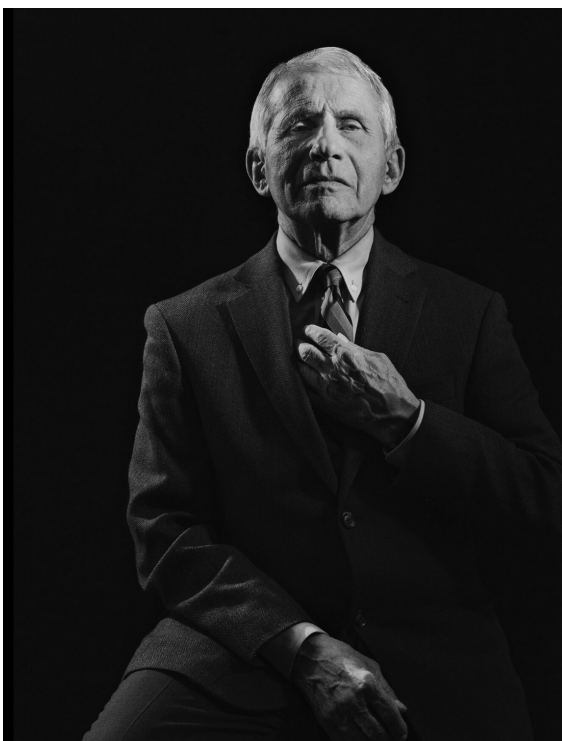
Protein Data Bank: From Two Epidemics to the Global Pandemic to mRNA Vaccines and Paxlovid

Stephen K. Burley, M.D., D.Phil.
Director, RCSB Protein Data Bank

Celebrating the Life and Legacy of Olga Kennard, ACS National Meeting, New Orleans, LA
March 17th 2024 8:15-8:35am Eastern Daylight Time



Dr. Anthony Fauci Loves the PDB!



Wallace-Wells: And what about pandemic preparedness more generally? Let's say we're working from scratch and designing the system at a white board. What reforms are needed?

Fauci: Do you have two weeks to talk?

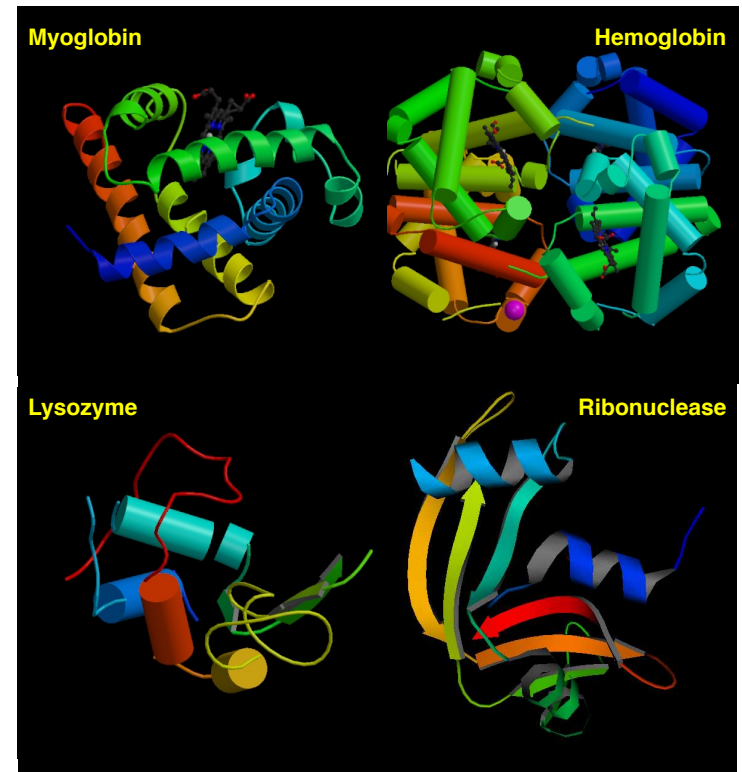
If you look at what worked for us, it was on the science side: the extraordinary investments that were made for decades before the emergence of SARS-CoV-2. First, the work in platform technology that led to essentially a revolution in how we make vaccines. No.2 is structure-based immunogen design. That helped with antiviral design, too — that has been the most underrated part of our response. I mean, show me a person who's vaccinated, got infected, took Paxlovid and died. I can't find anybody.

<https://www.nytimes.com/interactive/2023/04/24/magazine/dr-fauci-pandemic.html>



Protein Data Bank (Established 1971)

- PDB 1st online Open Access digital data resource in all of biology
- Founded 1971 with 7 protein structures
- Single global **archive** for protein and DNA/RNA experimental structures
- **Open Access >215,000 structures!**
- wwPDB Partnership founded in 2003
- Members: RCSB PDB (US), PDBe (EMBL-EBI), PDBj (Japan), and PDBc (China); plus EMDB (3DEM) and BMRB (NMR)

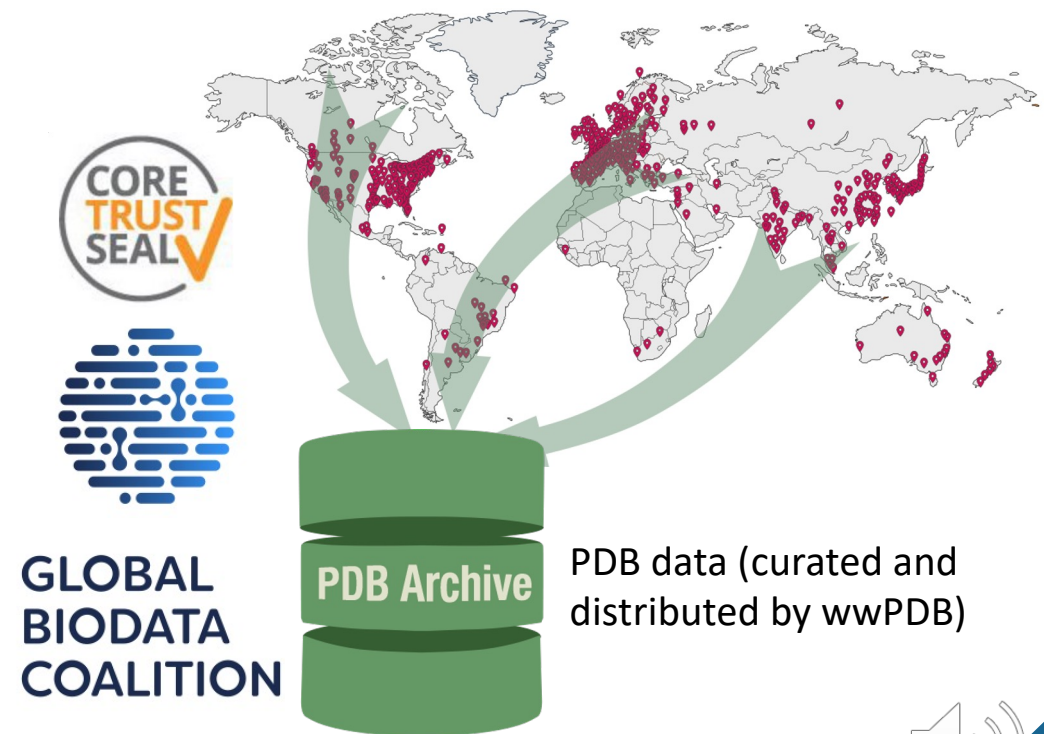


Structures that Inspired Launch of the PDB

Protein Data Bank (1971) *Nature New Biology* 233, 223.
Worldwide Protein Data Bank (2019) *Nucleic Acids Research* 47, D520–D528.

Protein Data Bank is “Crucial for Sustaining the Broader Biodata Infrastructure”

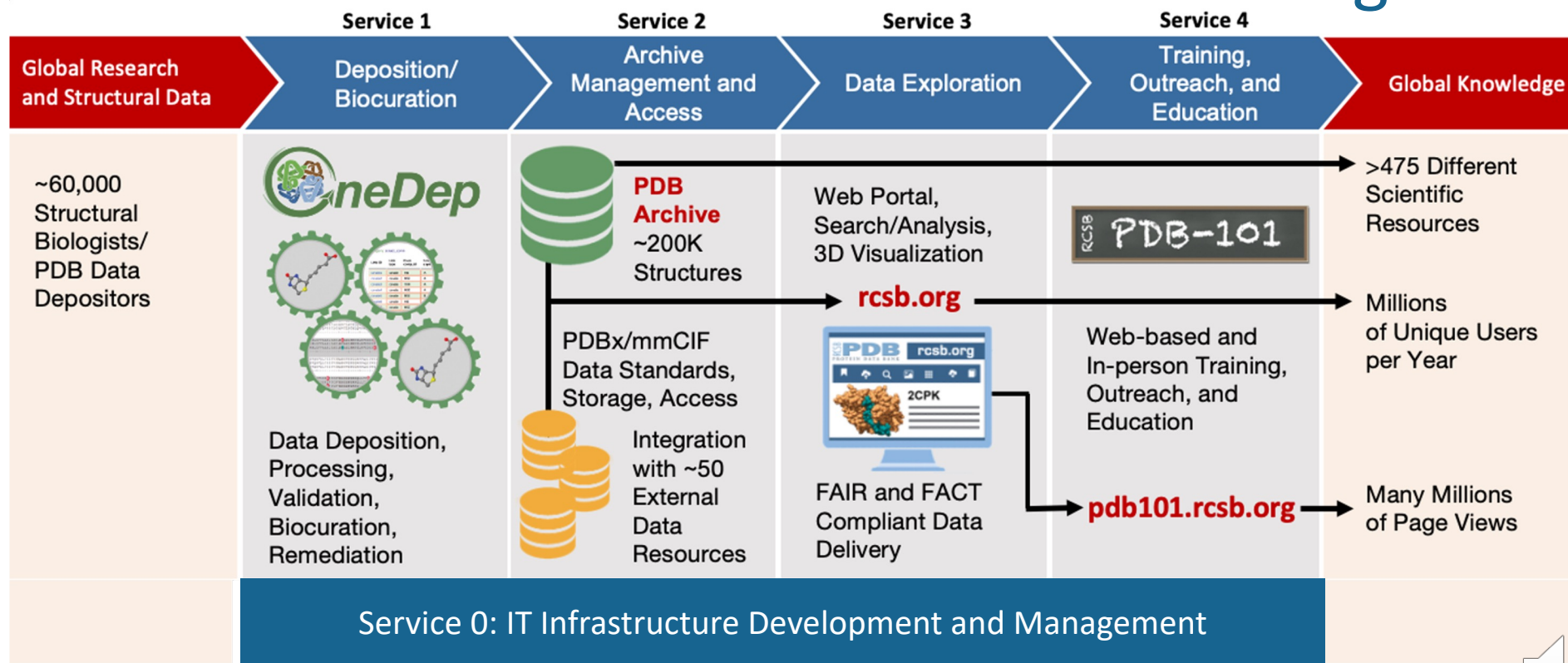
- Single FAIR/FACT compliant global archive providing Open Access to public domain experimental 3D biostructures
- PDB data distributed under the Creative Commons CC0 License
- RCSB Protein Data Bank is the US wwPDB Data Center jointly supported by NSF, NIH, and DOE



Protein Data Bank (1971) *Nature New Biology* 233, 223.
Worldwide Protein Data Bank (2019) *Nucleic Acids Research* 47, D520–D528.

RCSB Protein Data Bank (RCSB PDB) Services

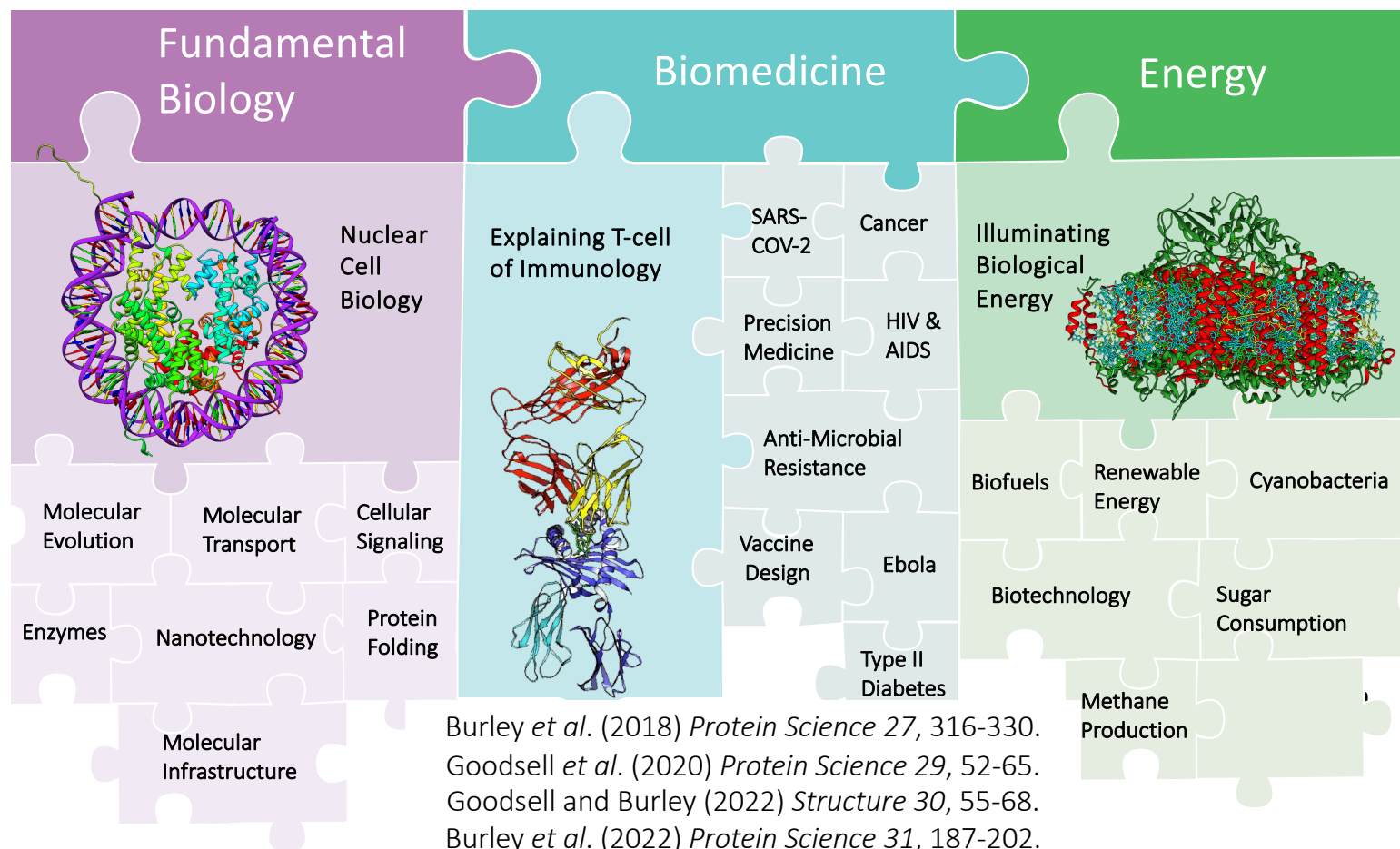
Convert Global Data into Global Knowledge



Burley *et al.* (2022) *Protein Science* 31, E4482.
 Burley *et al.* (2023) *Nucleic Acids Research* 51, D488–D508.



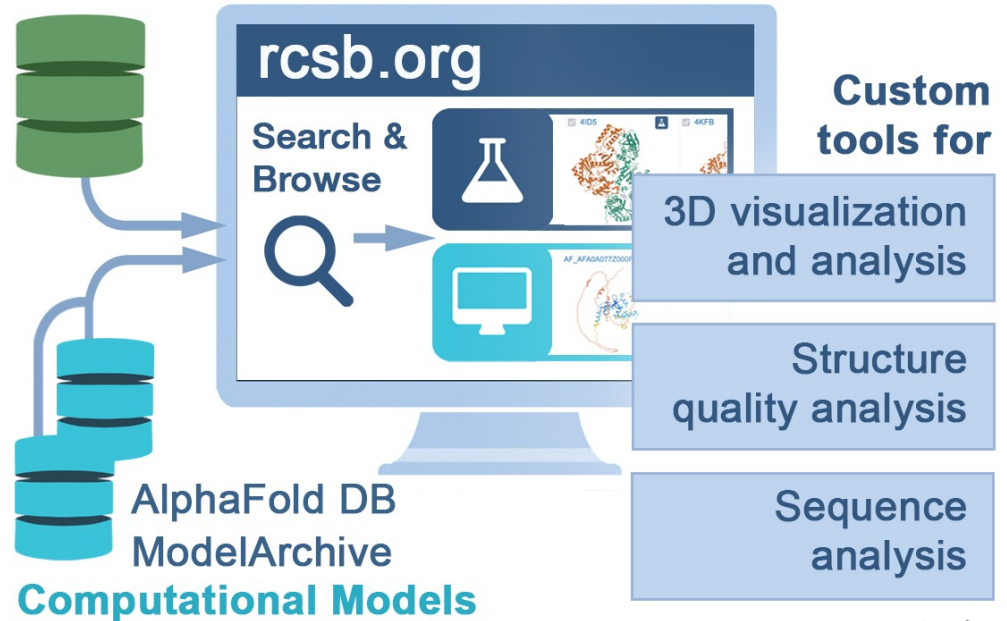
RCSB PDB Impact Across the Biosciences



RCSB.org Research-focused Web Portal: One-Stop-Shop for Public 3D Biostructure Data

- RCSB.org delivers
 - >215,000 PDB structures
 - >1 million Computed Structure Models (CSMs) from AlphaFold DB and the ModelArchive
- RCSB.org data exploration and visualization tools used by many millions of researchers, educators, and students worldwide
- Provenance/reliability of both data types are clearly identified

Experimental Models
Protein Data Bank

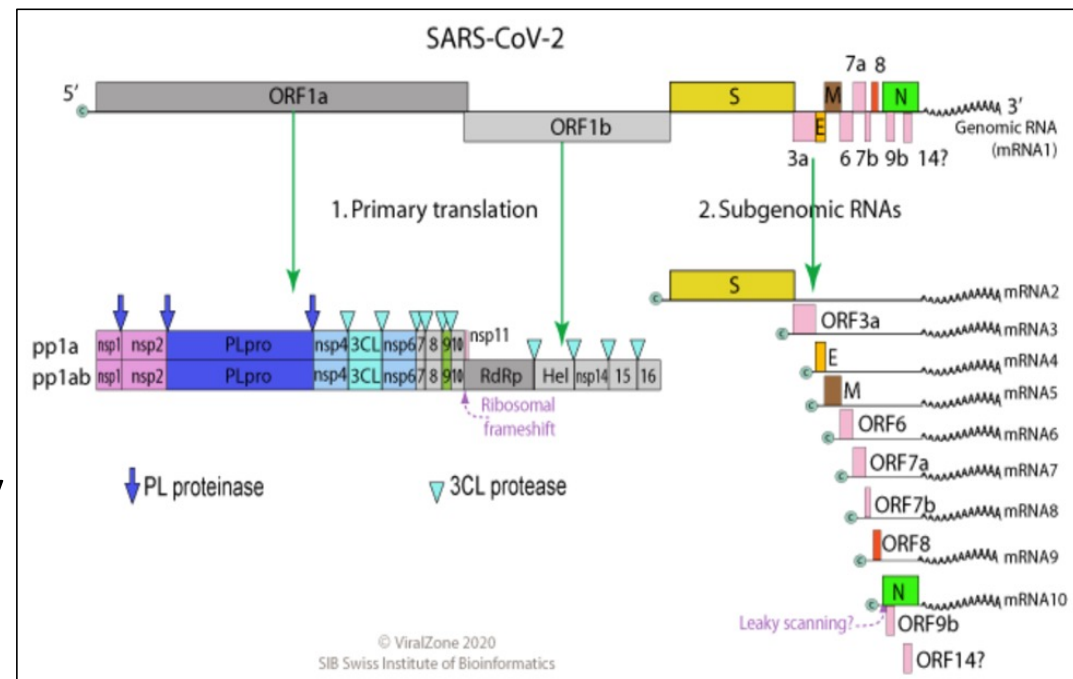


Burley et al. (2023) *Nucleic Acids Research* 51, D488-D508.



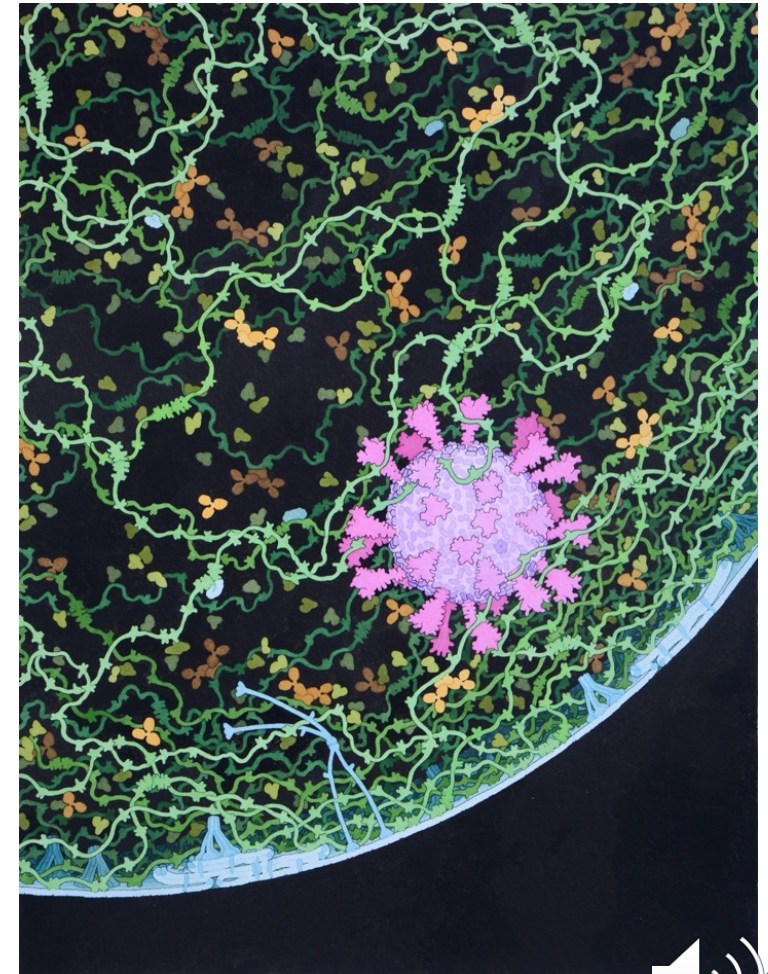
Coronavirus (SARS-CoV-2) Genome Organization

- Viral genome is a single-stranded, positive-sense, 5'-capped, 3'-polyadenylated messenger RNA
- Non-structural proteins expressed as polyproteins requiring enzymatic cleavage by the Main Protease (MP) and the Papain-Like (PL) Proteinase



PDB Essential for Responding to Emerging Viruses

- SARS-CoV Epidemic 2002
>170 SARS-CoV structures→PDB
- MERS-CoV Epidemic 2012
>100 MERS-CoV structures→PDB
- COVID-19 Pandemic 2019
~4,000 SARS-CoV-2 structures→PDB
- Effective mRNA vaccines designed and antiviral agents discovered using PDB structures of SARS-CoV, MERS-CoV, and SARS-CoV-2 proteins

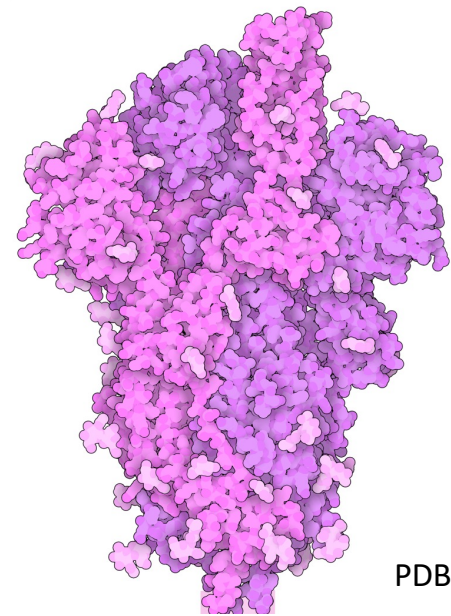


Respiratory Droplet, 2020; David S. Goodsell



Structure-Based Vaccine Design: Spike Protein

- Spike Protein
 - ~1,600 3DEM/Crystal structures
 - All Down and 1 Up/2 Down Trimers
 - Post-fusion Trimers
 - Complexes with ACE2, Fabs, etc.
- mRNA vaccine design relied on PDB structures of SARS-CoV and MERS-CoV spike proteins
- ~5.5 billion vaccinated!
- Tens of millions of lives were saved!



PDB ID 6vsb

Vaccine Discovery Target
Antibody Discovery Target

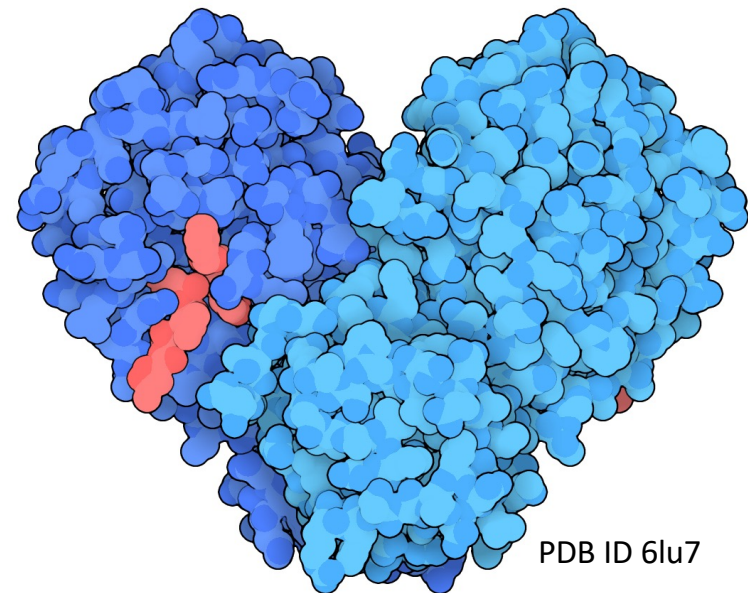
Wrapp *et al.* (2020) *Science* 367, 1260-1263.

SARS-CoV-2: PDB Structures of Drug Targets

PDB Structures

- Nsp5/Main Protease (Mpro)
 - >750 Apo/Co-crystal structures
 - Target of Pfizer's nirmatrelvir (+ritonavir=Paxlovid)
- Paxlovid is the more effective than the repurposed RNA polymerase inhibitors from Merck and Gilead
- N.B.: Paxlovid cannot be prescribed safely to all infected individuals because of drug-drug interactions with ritonavir

Nsp5 - Main Protease (Mpro)



Drug Discovery Target
Symmetric Homodimer; Two Active Sites

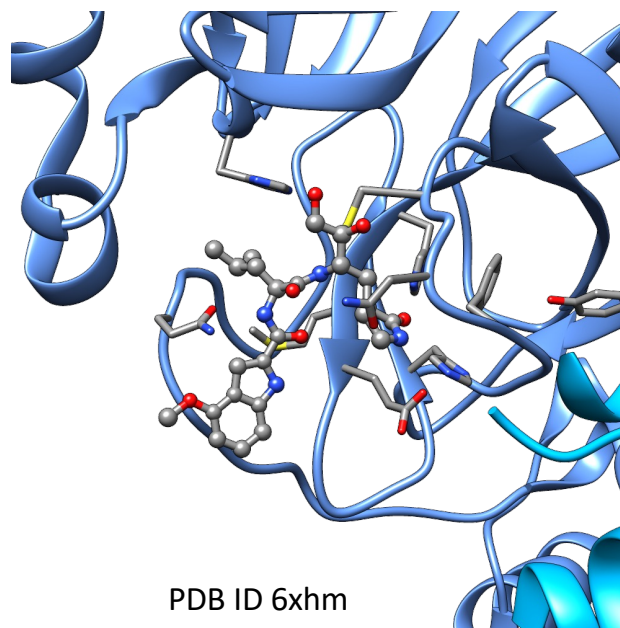
Jin et al. (2020) *Nature* 582, 299-293.

Structure-Based Drug Discovery: PF-00835231

PDB Structures

- Nsp5/Main Protease (Mpro)
 - >750 Apo/Co-crystal structures
 - Target of Pfizer's nirmatrelvir (+ritonavir=Paxlovid)
- In the 2000s, Pfizer used SBDD to discover PF-00835231 as a potential SARS-CoV antiviral
- Project halted in the mid-2000s (no commercial market!)

PF-00835231 Mechanism of Action



PF-00835231 bound to the Mpro Active Site

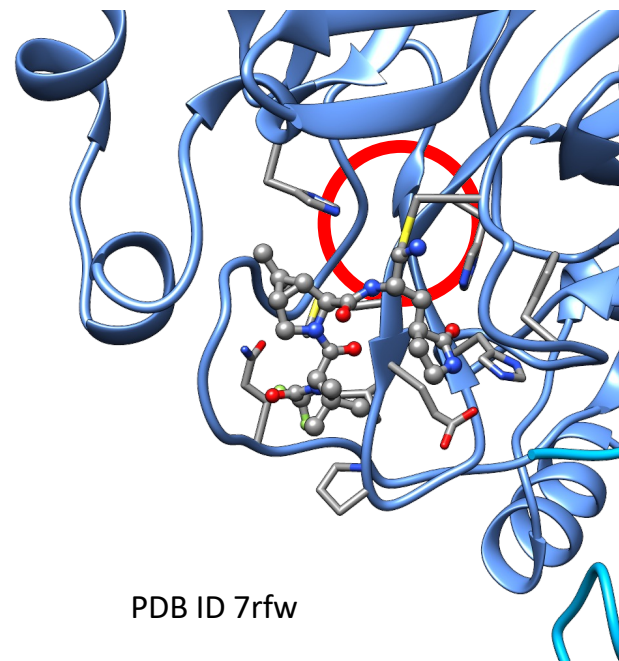
Hoffman *et al.* (2020) *Journal of Medicinal Chemistry* 63, 12725-12747.

Structure-Based Drug Discovery: Nirmatrelvir

PDB Structures

- Nsp5/Main Protease (Mpro)
 - >750 Apo/Co-crystal structures
 - Target of Pfizer's nirmatrelvir (+ritonavir=Paxlovid)
- In 2020, Pfizer used SBDD to turn PF-00835231 into nirmatrelvir as a SARS-CoV-2 antiviral
- N.B.: Also active against SARS-CoV and MERS-CoV Main Proteases

Nirmatrelvir Mechanism of Action



PDB ID 7rfw

Nirmatrelvir bound to the Mpro Active Site

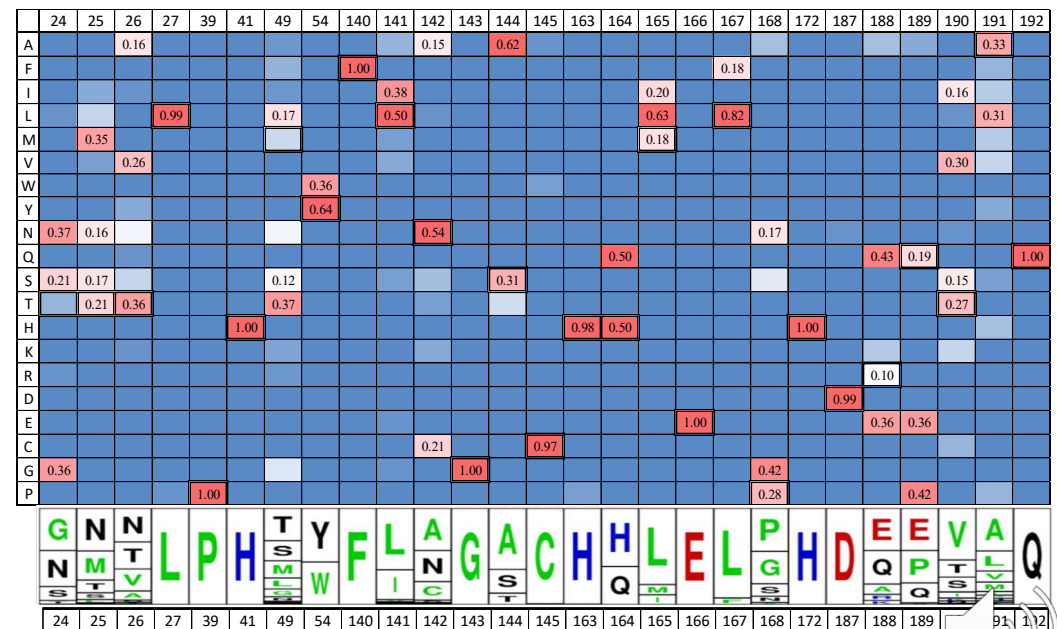
Owen *et al.* (2021) *Science* 374, 1586-1593.

Understanding Coronavirus Mpro Evolution

Nsp5 (Mpro) Conservation

- Nsp5 (Mpro) highly conserved across all known coronaviruses
- Nsp5 (Mpro) active sites are even more highly conserved
- Nsp5 (Mpro) peptide cut sites are also highly conserved
- Double selection pressure!

Nsp5 (Mpro) Active Site Residues



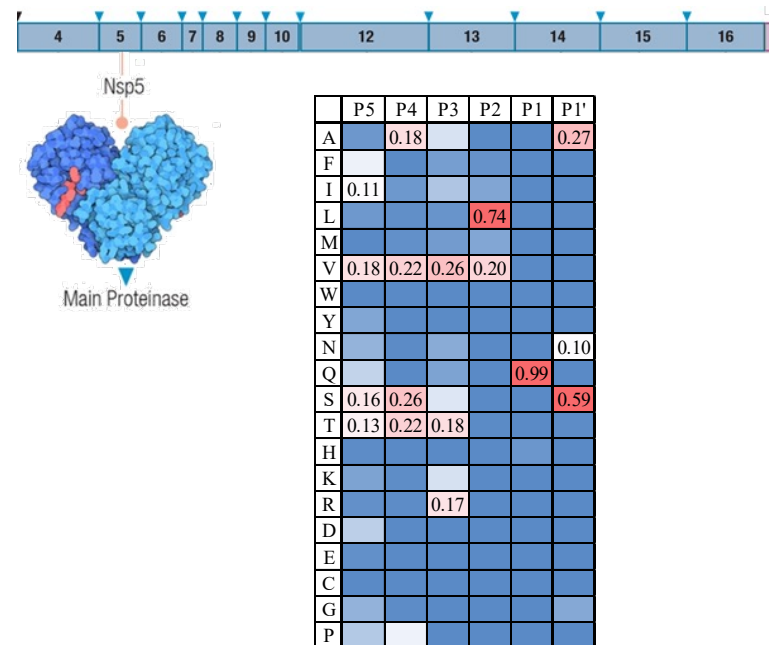
Lubin *et al.* (2023) In Preparation.

Understanding Coronavirus Mpro Evolution

Nsp5 (Mpro) Conservation

- Nsp5 (Mpro) highly conserved across all known coronaviruses
- Nsp5 (Mpro) active sites are even more highly conserved
- Nsp5 (Mpro) polyprotein cut sites also highly conserved
- Double selection pressure!

Nsp5 (Mpro) Polyprotein Cut Sites



Lubin *et al.* (2023) In Preparation.

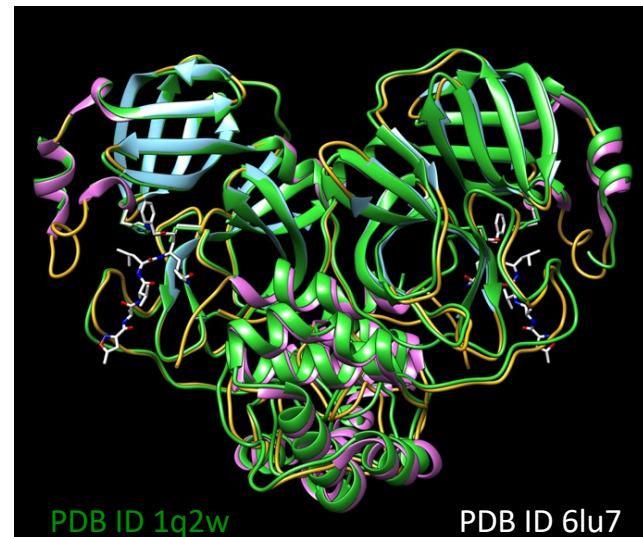
Failure of the Free Market: Sobering Postscript

Nsp5 (MPro) Conservation

- Nsp5 (Mpro) highly conserved
- SARS-CoV-2 and SARS-CoV Mpro structurally similar (Active Site R.M.S.D.<0.5Å)
- **If Pfizer had been incentivized to continue in the mid-2000s we would already have had a drug to combat SARS-CoV-2!**

Burley (2020) *Nature* 580, 167.

SARS-CoV-2 and SARS-CoV Nsp5 (Mpro)



Investment of US\$250M in a SARS-CoV MPro inhibitor could have saved tens of millions of lives and trillions of US\$ in economic losses

RCSB PDB Team

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

Core Operations Funding

National Science Foundation (DBI-1832184),
National Institute of General Medical Sciences,
National Institute of Allergy and Infectious
Disease, and National Cancer Institute
(NIH R01GM133198), and the
Department of Energy (DE-SC0019749)

Management

RUTGERS

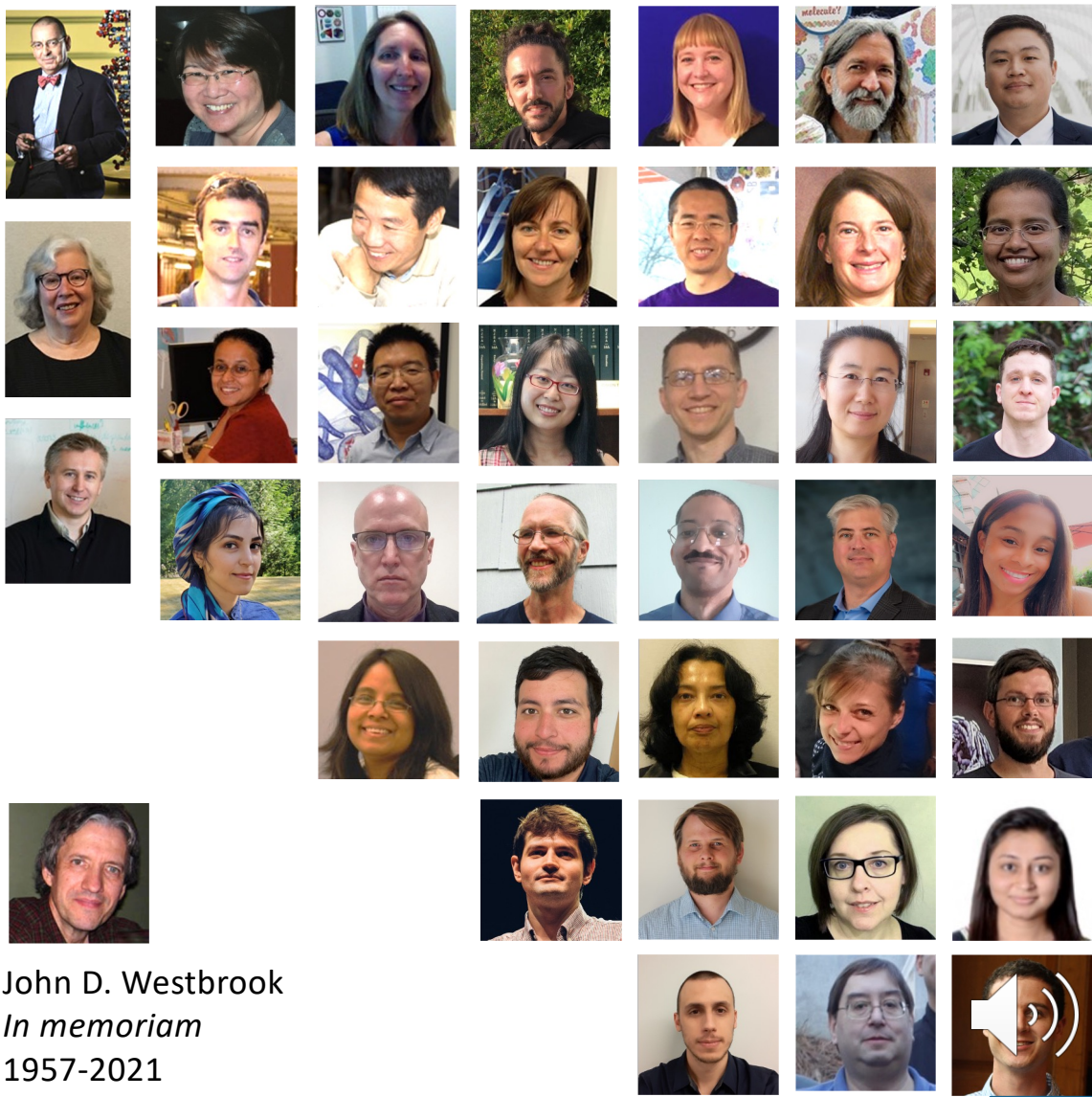
UC San Diego
SDSC SAN DIEGO
SUPERCOMPUTER CENTER

UCSF
University of California
San Francisco

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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

Follow us



John D. Westbrook
In memoriam
1957-2021

