

Cal Poly Humboldt

Digital Commons @ Cal Poly Humboldt

IdeaFest 2023

2023

Enhanced Learning in Biochemistry Using the Protein Data Bank and 3D Molecular Modeling in ChimeraX

Frank Cappuccio

Cal Poly Humboldt, fc412@humboldt.edu

Jenny Cappuccio

Cal Poly Humboldt, jc3534@humboldt.edu

Christopher Arias

Cal Poly Humboldt, ca190@humboldt.edu

Ezekiel Buchert

Cal Poly Humboldt, eib12@humboldt.edu

Johana Cruz-Lopez

Cal Poly Humboldt, jac295@humboldt.edu

See next page for additional authors

Follow this and additional works at: <https://digitalcommons.humboldt.edu/ideafest2023>

Recommended Citation

Cappuccio, Frank; Cappuccio, Jenny; Arias, Christopher; Buchert, Ezekiel; Cruz-Lopez, Johana; Fernandez, Cesar; Kowalski, Elizabeth; Tesla Nicholas, Elizabeth; Onnenga, Gabrielle; and Pierson, Kellie, "Enhanced Learning in Biochemistry Using the Protein Data Bank and 3D Molecular Modeling in ChimeraX" (2023). *IdeaFest 2023*. 96.

<https://digitalcommons.humboldt.edu/ideafest2023/96>

This Article is brought to you for free and open access by Digital Commons @ Cal Poly Humboldt. It has been accepted for inclusion in IdeaFest 2023 by an authorized administrator of Digital Commons @ Cal Poly Humboldt. For more information, please contact kyle.morgan@humboldt.edu.

Authors

Frank Cappuccio, Jenny Cappuccio, Christopher Arias, Ezekiel Buchert, Johana Cruz-Lopez, Cesar Fernandez, Elizabeth Kowalski, Elizabeth Tesla Nicholas, Gabrielle Onnenga, and Kellie Pierson

Christopher Arias, Ezekiel Buchert, Johana Cruz Lopez, Cesar Fernandez, Elizabeth Kowalski, Elisabet Tesla Nichols, Gabrielle Onnenga, Kellie Pierson, Jenny A. Cappuccio, and Frank E Cappuccio
Department of Chemistry, Cal Poly Humboldt, Arcata, CA, USA

Introduction

- RCSB PDB (RCSB.org) is the US data center for the global Protein Data Bank (PDB) archive of 3D structure data for large biological molecules (proteins, DNA, and RNA) essential for research and education in fundamental biology, health, energy, and biotechnology. (3)
- The Protein Data Bank (PDB) was the first open access digital data resource in biology and medicine.(3)

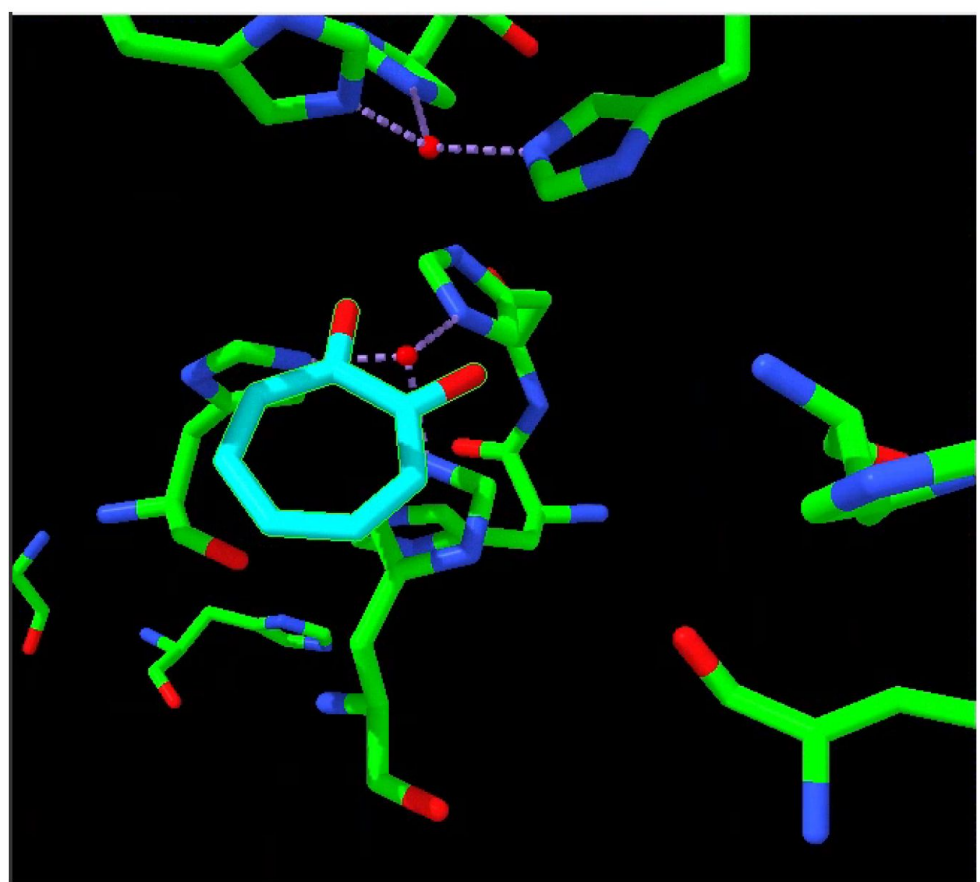


- Macromolecular structures are used in biotechnology, medicine and environmental studies to develop new cures for disease, and better biofuel production for example.
- Molecular visualization software ChimeraX was developed by UCSF faculty as a free program funded by the National Institutes of Health. (1, 2)

Methods

- Faculty developed hands-on student tutorial curriculum to demonstrate how to utilize PDB data.
- Students learned how to make 3D molecular representations of this data using UCSF ChimeraX software.
- A threaded curriculum biotechnology related protein (tyrosinase) was chosen to connect with previous knowledge in other courses.

Figure 1. Excerpt of student submissions for the hands-on PDB tutorial. This student utilizes their knowledge of charge to interpret the binding of an inhibitor to the tyrosinase enzyme active site, an enzyme causing browning in mushrooms. PDB 2Y9X. (5)



Inhibitor is the 7-membered cyan ring ^A. It seems to be attracted to the Cu+ ions – which makes sense given the resonance and more negative Oxygen atoms in the ring being attracted to the Cu+.

- Students applied their knowledge to a protein of their choice creating both a Quad Chart presentation and a 3D printed surface model of the protein (4)
- Quad Chart presentations are used to pitch an idea or quickly and inform the audience for funding or further studies.

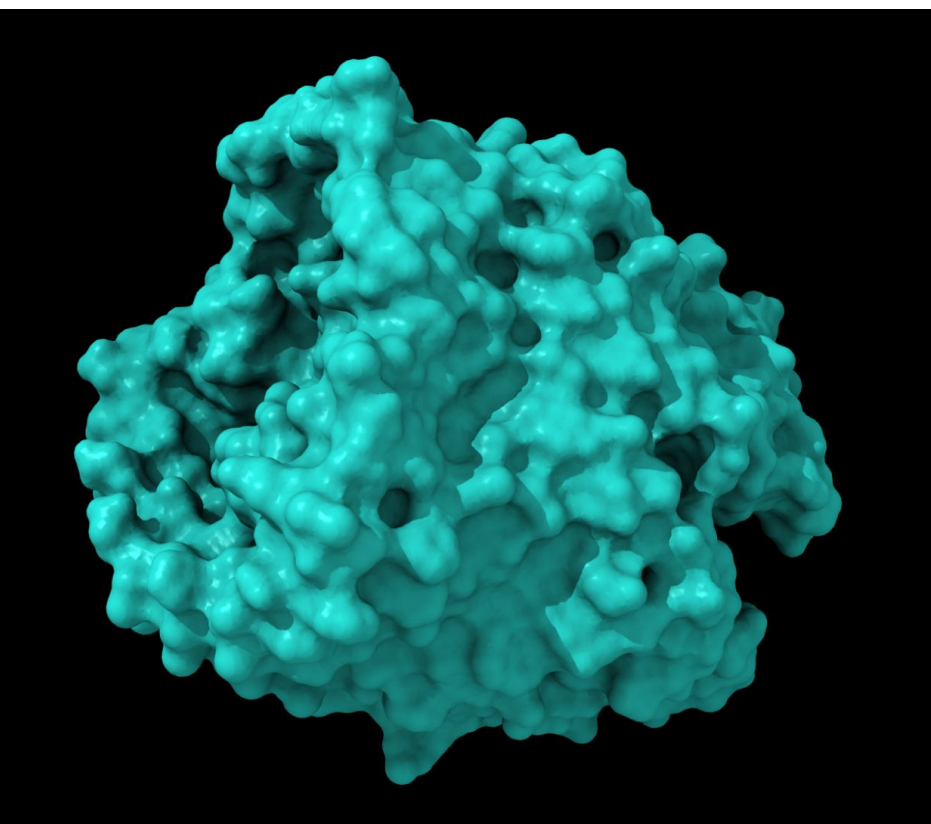


Figure 2. Student surface model submissions .stl file for the 3D printing of their protein choice, cobalamin (vit B12) transporter BtuB from the bacterium E. coli PDB 1NQH. (2003) Nat Struct Biol 10: 394-401

Student Products

Acetylcholinesterase PDB ID: 1ACJ

General & Structural Information

- Enzyme (serine hydrolase) that cleans up used acetylcholine¹
- acetylcholine carries signals from nerve → muscle cells; opens receptors triggers contraction
- breakdown: acetylcholine → acetic acid + choline
- Found in synapses between nerve/muscle cells²
- Structural motifs³
- 2: Thirty α-helices, two β sheets, and numerous D loops connecting the two
- 3: primarily connections by H bonds
- 4: homodimer

Functional Information

- Active site: Serine 200⁴
- Inhibitors: animal poisons/toxins, sarin gas, insecticides⁵
- No enzyme = too much acetylcholine = muscle paralysis
- Oddly, this can be good for AD patients
- Increased receptor stimulus remedies memory loss⁶
- Isomers: AChE, AChE₂, AChE₃
- Generated by 3' end splicing of AChE gene⁷
- Same catalytic domain, but unique C-terminal end
- Possible applications: development of anti-acetylcholinesterase drug⁸

Fun Facts

- 80 μs reaction time - one of the fastest of all enzymes⁹
- First isolated from electric fish (ex. Torpedo ray), which have extensive nervous connections & electric signaling organs

References

1. "Enzyme (serine hydrolase) that cleans up used acetylcholine¹"
2. "Found in synapses between nerve/muscle cells²"
3. "Structural motifs³"
4. "Active site: Serine 200⁴"
5. "Inhibitors: animal poisons/toxins, sarin gas, insecticides⁵"
6. "Increased receptor stimulus remedies memory loss⁶"
7. "Generated by 3' end splicing of AChE gene⁷"
8. "Possible applications: development of anti-acetylcholinesterase drug⁸"
9. "80 μs reaction time - one of the fastest of all enzymes⁹"

Dendrotoxin-I : PDB# 1DEM

Cesar Fernandez 3/9/23 PDB Assignment

MW: 7170 g/mol

Z: 1 beta sheet and 2 alpha helix

A presynaptic neurotoxins which blocks potassium channels

(targeting certain subtypes of voltage-gated potassium channels found in neurons, is designed to boost the release of acetylcholine at neuromuscular junctions.)

Binding sites in red And antiprotease areas

Gene source: Dendroaspis polylepis polylepis (Black mamba)

Meant to destroy nerve tissue Yet a useful protein for pharmacological research for ion channel proteins.

Produces Acetylcholine / and has anti protease functionality.

1. Bank, R. C. S. B. P. D. 1DEM: Protease inhibitor homologues as potassium channel blockers. RCSB PDB.

2. Gasparini, S., Datta, J. M., Lecocq, A., Prikazhd, S., Zinn-Justin, S., Young, L. C., de Meester, G. C. L., Rowan, J. G., Hargreaves, A. L., and Merson, A. (1998). Definition of the functional site of α-dendrotoxin. *Journal of Biological Chemistry* 273, 25339–25403.

3. Adoni, C., Nishida, H., Ito, T., Nishikawa, Y., Kimura, T., Saitoh, K., S. and Yamashita, T. (2000). Structural basis for the biological activity of dendrotoxin I, a potent potassium channel blocker. *Bioplayers* 54, 44–57.

Human Hemoglobin 1CH4

Ezekiel Buchert

Total Structure Weight: 64.55 kDa

Atom Count: 4,997

Modelled Residue Count: 572

Deposited Residue Count: 574

The hemoglobin protein consists of two beta chains and two alpha chains. Each chain contains an iron atom. Oxygen bonds these atoms of iron, and the molecule embedded into the red blood cells transports oxygen throughout the bloodstream. Carbon monoxide can irreversibly bond to the heme groups that contain the iron. This will poison the action of the molecule and can cause death.

This protein is capable of being synthesized by E. coli. This protein is capable of being identified via Iron-60 in the heme group

Bank RCSB PDB 1CH4: Modelled-substituted Chimeric Hemoglobin Beta alpha (P12395). In: RCSB PDB. <https://www.rcsb.org/structure/1CH4>. Accessed 04/06/2023

Shen, T., Fujikawa, M., Yamanaka, T., Iwata, K., Ishihara, K., Mochizuki, Y. Crystal structure of a protein with an artificial covalent cross-link, modelled-substituted chimeric hemoglobin beta alpha, at 2.5 Å resolution. *J. Mol. Biol.* 369, 566–570 (2003).

Muller, T., Lippert, C. L., Dreyer, C., Thiemann, M., Gier, R. C., Mochizuki, Y., & Bryant, D. M. (2017). VADT: modeling structural similarities between macromolecular complexes. *Nucleic Acid Res.* 45, 1010–1018. doi: 10.1093/nar/nkx028

Evans, D. J. (2013). Carbon monoxide poisoning. *J. Soc. Med.* 203, 100–102. doi: 10.1177/0962280212460000

Bank, R. C. S. B. P. D. 1CH4: Modelled-substituted Chimeric Hemoglobin Beta alpha (P12395). In: RCSB PDB. <https://www.rcsb.org/structure/1CH4>. Accessed 04/06/2023

Tau Tubulin Kinase I: PDB ID #4NFM

Kellie Pierson, 03/08/2023, PDB Assignment

General Features:

- Serine/Threonine Protein Kinase
- Part of the Casein Kinase 1 (CK1) family
- 1257 amino acids, 45 alpha-helices, 52 beta strands
- Conserved kinase domain
- MW: 126 kDa; N-terminal kinase MW: 41 kDa; C-terminal ubiquitin-associated domain MW: 85 kDa
- Isoforms: TTBK1A and TTBK1B (differ in N-terminal sequences)
- TTBK1A is longer and the predominant form expressed in the brain
- Monomer; potential to dimerize
- Kinase domain: conserved catalytic core w/ ATP-binding site & catalytic residue for phosphorylation
- Localized to cytoplasm

TTBK1 Function & Substrate Specificity:

Phosphorylation

Implicated in Wnt signaling pathway; regulates cell growth, differentiation, and cell fate determination.

Known to phosphorylate a variety of substrates including tau proteins and tubulins; involved in microtubule assembly and stability; axonal transport, and neuronal development

Disease Implications:

Spinocerebellar ataxia type 11; neurodegenerative disorder

Alzheimer's disease; neurodegenerative disorder

TTBK1 is involved in hyperphosphorylation of tau protein in both of these disorder

Hyperphosphorylation of tau proteins contribute to formation of neurofibrillary tangles

References:

1. "Human tau tubulin kinase 1 (TTBK1)" RCSB Protein Data Bank (February 6, 2014). DOI: 10.2201/4NFM/4NFM
2. "TTBK1 in complex with inhibitor" RCSB Protein Data Bank (September 25, 2013). DOI: 10.2201/4NFM/4NFM
3. Hinks, et al (2021). *Journal of medicine chemistry*, 64(9), 6358-6380 DOI: 10.1021/acs.jmedchem.1c00582

MHETase (PDB ID: 6QGA)

By: Gabby Onnenga

Protein Structure:

- Found in bacteria: from *Idionella sakaiensis*¹
- 3,349 amino acids residues²
- Hydrolase enzyme family → α/β hydrolases¹
- Contains 6 chains:²
- 2 domains – 3 chains each. A dimer.
- Lid domain is unique
- Many ligands bound to chains to activate enzyme
- Benzoic acid, MPD, Sulfate, Calcium ion, Chloride

Protein Function:

- 2nd enzyme needed to break down PET polymers 1,3
- 6QGA breaks stable bonds of MHET creating components EG and TPA¹
- Induced fit mechanism¹
- Once attached to MHET, 180° rotation around F415
- K_m = 7.3 μM
- Various mutations show increase in K_m¹

Fun Facts!

- Enzyme discovered at a recycling facility – bacteria on plastics
- EG and TPA are digestible and metabolized by bacteria AND are building blocks for a new round of PET synthesis¹
- Can provide new pathways for more renewable energy with further study

References:

1. "Enzyme discovered at a recycling facility – bacteria on plastics"
2. "EG and TPA are digestible and metabolized by bacteria AND are building blocks for a new round of PET synthesis¹"
3. "Can provide new pathways for more renewable energy with further study"

Tetrahydrocannabinolic acid (THCA) synthase (PDB: 3VTE)

Elisabet Tesla Nicholas

2° structure: divided in 2 domains with a flavin adenine dinucleotide (FAD) in between¹

Domain I: 8 alpha helices & 8 beta sheets covalently bound to FAD.

Domain II: 8 beta sheets surrounding 5 alpha helices.

FAD covalently bound to His114 and Cys176.

FAD also bound by hydrogen bonds.

Biological Functions:

- THCA synthase uses FAD as a catalyst for the oxidative cyclization of the monoterpene of cannabigerolic acid (CBGA).¹
- May contribute to self-defence of Cannabis plants since the products of the reaction with CBGA (THCA & H2O2) are both cytotoxic.¹

Fun Facts!

- Decarboxylation of THCA produces THC²
- THCA may possibly be used to synthesize pharmaceutical grade THC²

References:

1. Yoshinari Shoyama et al. (2012). *PLoS ONE* 7(12): e43636. doi:10.1371/journal.pone.0043636
2. Memon-Sana G (2016). *Cannabis & Cannabinoid Research*, Vol. 1, pp. 124-130
3. Saper, S. K. (2018). *Journal of Cannabis Research*, Vol. 1, pp. 178-182

Zika Virus :PDB ID# 5IRE

Christopher Arias, 03/08/23, PDB Assignment

General Features:

- Contains 504 amino acid residues
- Secondary structure: contains 30 antiparallel beta sheets and 10 alpha helices
- Molecular weight: 190.1 kDa
- Carbohydrate moiety functions as an attachment site of the virus to host cells.
- Is part of flaviviridae (flavivirus), which is a positive, single stranded RNA virus. Can be found in ticks and mosquitoes

Function:

- The virus is the causal agent of neurological disorders like the Guillain-Barre syndrome and microcephaly in newborns.
- The mechanism of Zika is to cross the placental barrier and the blood brain barrier.

Visual Representation:

"Fun" Facts:

- Zika virus was discovered in 1947, in an infected monkey in the Zika forest of Uganda
- Linked to birth defects
- Causes neurological symptoms
- Primary spread through infected mosquitoes

References:

1. "Zika virus was discovered in 1947, in an infected monkey in the Zika forest of Uganda"
2. "Linked to birth defects"
3. "Causes neurological symptoms"
4. "Primary spread through infected mosquitoes"

Conclusions

- 3D modeling is a valuable interactive tool to understand the complexity of biological macromolecules informing function.
- Student displayed in-depth understanding of protein structure and binding in hands-on activities and presentations of protein structures
- The knowledge of biological molecular structure in drug design and biotechnology is an essential skill set for modern scientists to meet the grand challenges of the future.

Acknowledgments

We would like to thank Tim W. Miller the Cal Poly Humboldt Digital and Media Librarian for assistance in printing the 3-D surface models.

We would like to acknowledge funding from a CSU Program for Education and Research in Biotechnology (CSUPERB) curriculum development grant.

Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

References

1. [UCSF ChimeraX: Structure visualization for researchers, educators, and developers](#). Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. *Protein Sci.* 2021 Jan;30(1):70-82.
2. [UCSF ChimeraX: Meeting modern challenges in visualization and analysis](#). Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. *Protein Sci.* 2018 Jan;27(1):14-25.
3. The Protein Data Bank H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne (2000) *Nucleic Acids Research*, 28: 235-242. <https://doi.org/10.1093/nar/28.1.235>
4. PDB-101: Educational resources supporting molecular explorations through biology and medicine. Christine Zardecki, Shuchismita Dutta, David S. Goodsell, Robert Lowe, Maria Voigt, Stephen K. Burley. (2022) *Protein Science* 31: 129-140 <https://doi.org/10.1002/pro.4200>
5. Crystal Structure of Agaricus Bisporus Mushroom Tyrosinase: Identity of the Tetramer Subunits and Interaction with Tropone. [Ismaya, W.T., Rozeboom, H.J., Weijn, A., Mes, J.J., Fusetti, F., Wichers, H.J., Dijkstra, B.W.](#) (2011) *Biochemistry* 50: 5477