

# POCKETMÓN

## SBI-PYTHON PROJECT

### EXAMPLES

## AUTHORS

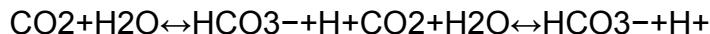
Marta García · marta.garcia38@estudiant.upf.edu  
Karim Hamed · karim.hamed01@estudiant.upf.edu  
Ivon Sánchez · ivon.sanchez01@estudiant.upf.edu



Universitat  
Pompeu Fabra  
Barcelona

# 1. 2WEG - HUMAN CARBONIC ANHYDRASE II

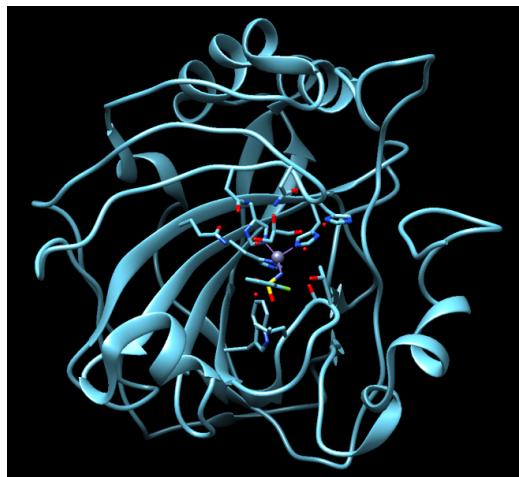
The protein structure 2WEG represents human carbonic anhydrase II (CA II), a zinc metalloenzyme belonging to the  $\alpha$ -class of carbonic anhydrases. CA II plays a central role in maintaining acid-base balance in physiological systems by catalyzing the reversible hydration of carbon dioxide:



The enzyme is encoded by the *CA2* gene and is highly expressed in erythrocytes, kidneys, lungs, and the gastrointestinal tract. The structure of 2WEG was determined using X-ray crystallography at an ultra-high resolution of 1.1 Å, allowing detailed examination of the active site and ligand interactions [1].

## Binding Site Architecture

The active site of CA II is a deep conical cavity located in the center of the protein and is coordinated by a catalytically essential zinc ion ( $\text{Zn}^{2+}$ ). This zinc ion is tetrahedrally coordinated by the imidazole groups of three histidine residues: His94, His96, and His119. The fourth coordination site is typically occupied by a water molecule or hydroxide ion, which acts as a nucleophile in the hydration of  $\text{CO}_2$  [2].



**Figure 1a:** 2WEG structural conformation with its ligand (Zn)

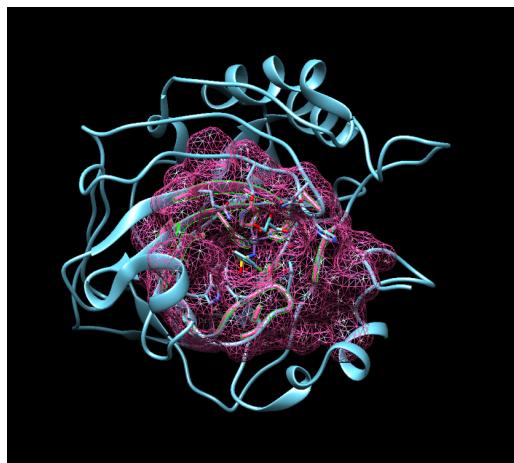
In the 2WEG structure, the enzyme is complexed with FBV (4-fluorobenzenesulfonamide), a sulfonamide-based inhibitor. FBV binds in the active site, displacing the zinc-bound water molecule and coordinating directly to the zinc ion via its deprotonated sulfonamide group. This mode of binding is characteristic of

classical carbonic anhydrase inhibitors and is crucial for understanding structure-activity relationships in drug design [3].

The hydrophobic and hydrophilic regions of the binding pocket are also well-defined. Residues such as Val121, Phe131, and Leu198 contribute to the hydrophobic lining, while Thr199 and Glu106 form part of the polar environment that stabilizes ligand binding through hydrogen bonding and electrostatic interactions [4].

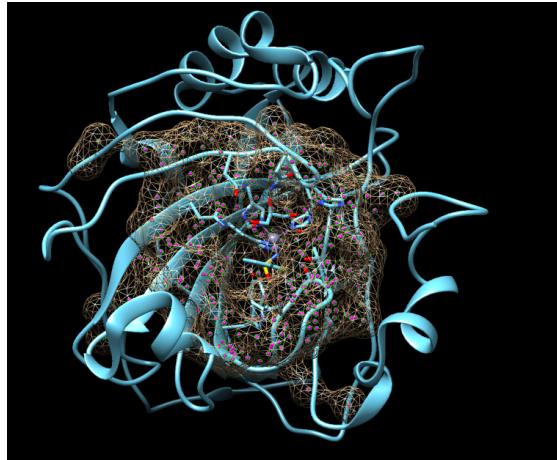
### Prediction

The binding pocket in the PDB-derived structure (Figure 1b) is well-defined and displays the canonical features of carbonic anhydrase II's active site. A zinc ion is tetrahedrally coordinated by His94, His96, and His119, with a sulfonamide-based inhibitor (FBV) occupying the fourth coordination position. The pocket is deeply embedded within a conical cavity, surrounded by polar residues like Thr199 and Glu106, and hydrophobic residues such as Val121 and Phe131, facilitating strong inhibitor binding through both hydrophobic interactions and hydrogen bonding. [2, 3]



**Figure 1b:** 2WEG PDB with PDBBind's dataset pocket

In contrast, the predicted model (Figure 1c) shows a binding site that closely overlaps the general spatial position of the native pocket, yet with subtle differences in surface contour and residue orientation. While the zinc-binding site is correctly recovered, minor shifts are observed in the positions of coordinating residues, most notably His94 and Thr199, which may slightly alter binding geometry. Additionally, the surface mesh in the model appears more diffuse, suggesting reduced precision in pocket boundaries or possible overprediction of solvent-accessible regions.



**Figure 1c:** 2WEG PDB with our model's prediction

## 2. 1AFK - BOVINE PANCREATIC RIBONUCLEASE A

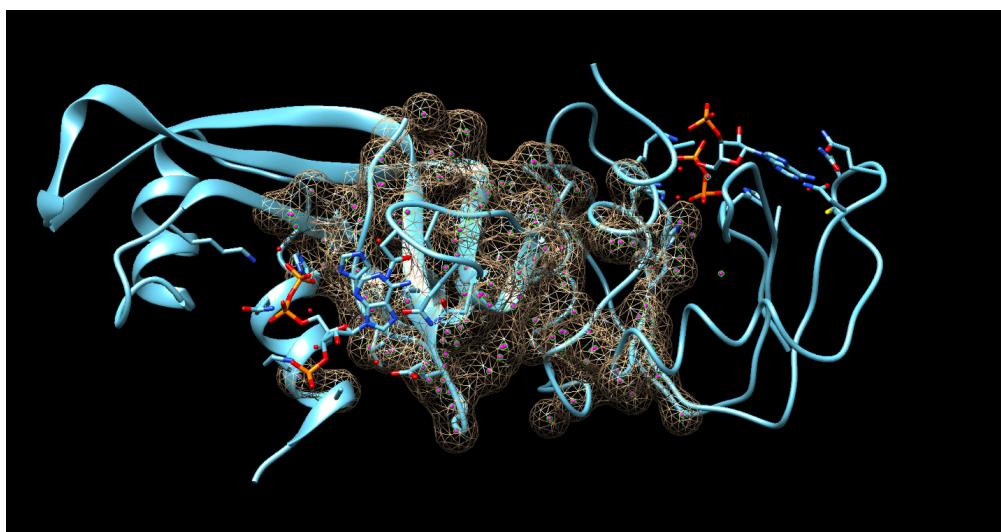
High-resolution (1.7 Å) crystal structures have been determined for bovine pancreatic ribonuclease A (RNase A) complexed with 5'-diphosphoadenosine 3'-phosphate (ppA-3'-p) and 5'- diphosphoadenosine 2'-phosphate (ppA-2'-p), as well as for a native structure refined to 2.0 Å. [5]

Figure 2a displays the 1AFK PDB structure with annotated residues that are part of the experimentally validated binding site. These residues directly interact with the ligand or are in close spatial proximity to the binding cleft. The ligands are accommodated in a cleft formed by several loops and adjacent surface-exposed residues, aligning with known characteristics of nucleotide-binding proteins. The high-resolution data allows us to precisely identify side-chain orientations and hydrogen bond interactions with the ligand, reinforcing the structural basis for specificity.



**Figure 2a:** 1AFK pdb with known binding sites

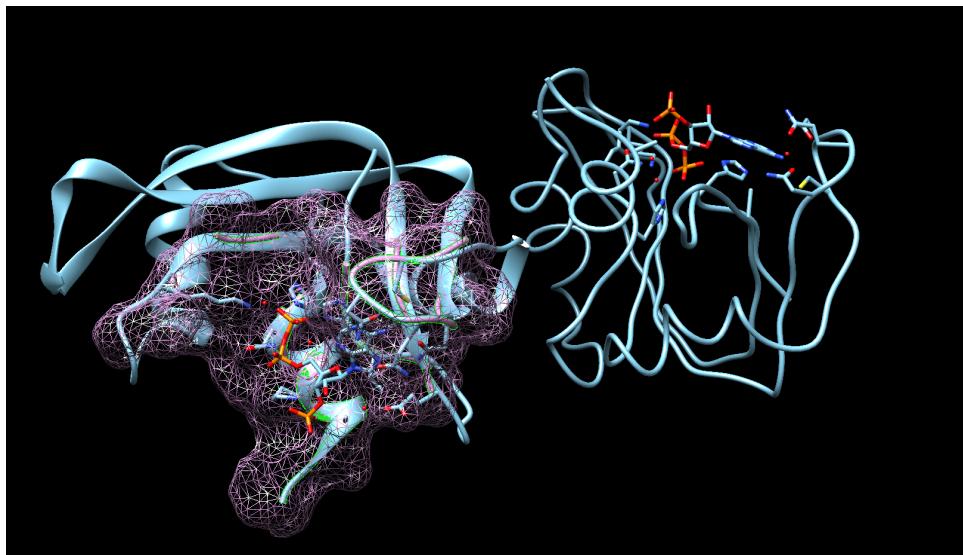
Using our prediction model, we can see in Figure 2b a mesh that shows the position of this prediction. As we can see, even though the majority of our predicted residues in the binding site are not the same as the ones in the known binding site, most are situated spatially close to residues that belong to the known binding site. Moreover, we can see that a lot of the residues from the predicted and known binding sites are located in loop regions of the protein, which makes sense given the fact that the residues situated in loop regions have a bigger degree of conformational flexibility and surface accessibility compared to residues in secondary structure elements such as alpha helices or beta strands. This flexible nature of the loop regions makes them more prone to participate in the binding of other molecules. This also applies to the rest of the examples presented in this document.



**Figure 2b:** 1AFK PDB with our model's prediction

Figure 2c shows the ligand-binding pocket as defined by the PDDBBind database, superimposed on the 1AFK structure. PDDBBind provides curated binding site annotations based on a wide array of protein-ligand complexes, facilitating standardization across computational studies. When visualized in Chimera, the PDDBBind pocket aligns well with the known binding site and also shares considerable overlap with the predictions made by our model.

The agreement between our predicted binding site and the PDDBBind-defined pocket supports the predictive power of our method. Even in cases where specific residue matches are absent, the spatial coherence and topological context of the predictions align well with biological expectations.



**Figure 2c:** 1AFK PDB with PDBBind's dataset pocket

### 3. 4R3W - ASADH

The aspartate pathway is essential for the production of the amino acids required for protein synthesis and of the metabolites needed in bacterial development. This pathway also leads to the production of several classes of quorum-sensing molecules that can trigger virulence in certain microorganisms. The second enzyme in this pathway, aspartate-semialdehyde dehydrogenase (ASADH), is absolutely required for bacterial survival and has been targeted for the design of selective inhibitors. Fragment-library screening has identified a new set of inhibitors that, while they do not resemble the substrates for this reaction, have been shown to bind at the active site of ASADH. Structure-guided development of these lead compounds has produced moderate inhibitors of the target enzyme, with some selectivity observed between the Gram-negative and Gram-positive orthologs of ASADH. However, many of these inhibitor analogs and derivatives have not yet achieved the expected enhanced affinity. Structural characterization of these enzyme-inhibitor complexes has provided detailed explanations for the barriers that interfere with optimal binding. Despite binding in the same active-site region, significant changes are observed in the orientation of these bound inhibitors that are caused by relatively modest structural alterations. Taken together, these studies present a cautionary tale for issues that can arise in the systematic approach to the modification of lead compounds that are being used to develop potent inhibitors. [6]

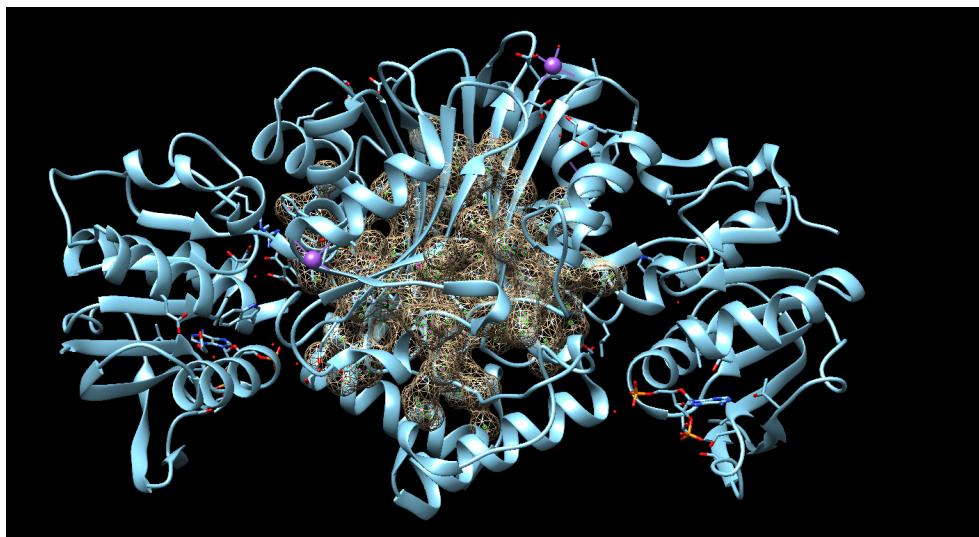
Figure 3a displays the crystal structure of *aspartate-semialdehyde dehydrogenase* (ASADH) from 4R3W reveals a well-defined active site that plays a critical role in the

aspartate pathway. Highlighted residues indicate the catalytic core and substrate-binding pocket. This site is crucial for bacterial survival and is the target of fragment-based inhibitor design.



**Figure 3a:** Active Site of ASADH from 4R3W Structure

This panel shows the superimposition of multiple fragment-based inhibitors identified through screening. Although these compounds do not structurally mimic the native substrate, they occupy overlapping regions of the active site. Their interactions with conserved residues illustrate the potential for non-substrate-like molecules to act as inhibitors, despite their limited affinity.



**Figure 3b:** Binding of Fragment-Derived Inhibitors in ASDAH Active Site

Figure 2c shows that despite binding to the same catalytic region, structural comparisons reveal significant variability in the orientation of inhibitor analogs. Even minor chemical modifications lead to altered binding poses, as visualized here. This highlights the challenge of structure-guided inhibitor optimization, where small molecular changes can cause unpredictable shifts in binding geometry and reduce efficacy. When visualized in Chimera, the PDBBind pocket aligns well with the known binding site and also shares considerable overlap with the predictions made by our model.

The strong spatial overlap between our predicted binding site and the PDBBind-defined pocket highlights the accuracy of our prediction method. Even when exact residue matches are lacking, the predicted regions maintain structural and contextual consistency with known biological binding sites.



**Figure 3c:** PDBBind's pocket with ASDAH (4R3W PDB)

## REFERENCES:

- [1] PDB entry 2WEG – Protein Data Bank in Europe (PDBe).  
<https://www.ebi.ac.uk/pdbe/entry/pdb/2weg>
- [2] Lindskog, S. (1997). Structure and mechanism of carbonic anhydrase. *Pharmacology & Therapeutics*, 74(1), 1–20.  
[https://doi.org/10.1016/S0163-7258\(96\)00198-2](https://doi.org/10.1016/S0163-7258(96)00198-2)
- [3] Supuran, C. T. (2008). Carbonic anhydrases: Novel therapeutic applications for inhibitors and activators. *Nature Reviews Drug Discovery*, 7, 168–181.  
<https://doi.org/10.1038/nrd2467>
- [4] Alterio, V. et al. (2012). Crystal structure of the human carbonic anhydrase II-FBV complex at 1.1 Å resolution: Implications for drug design. *Bioorganic & Medicinal Chemistry Letters*, 22(2), 859–863. <https://doi.org/10.1016/j.bmcl.2011.12.015>
- [5] Crystal Structures of Ribonuclease A Complexes with 5'-Diphosphoadenosine 3'-Phosphate and 5'-Diphosphoadenosine 2'-Phosphate at 1.7 Å Resolution.  
<https://pubmed.ncbi.nlm.nih.gov/9154942/>
- [6] A cautionary tale of structure-guided inhibitor development against an essential enzyme in the aspartate-biosynthetic pathway.  
<https://journals.iucr.org/paper?S1399004714023979>