Teaching Structure-Based Drug Discovery Through a Guided Molecular Docking Workflow

Guided Molecular Docking Module
Validation and Comparative Analysis Using Acetylcholinesterase (PDB ID: 4M0E)

1. Rationale and Educational Context

Concepts such as molecular recognition, noncovalent interactions, and protein structure are foundational in chemistry and biochemistry education. However, these topics are frequently taught in isolation, limiting students' ability to integrate structural reasoning with predictive modeling.

This module introduces a structured molecular docking workflow that moves from experimental structural evidence to computational validation and comparative prediction. The emphasis is not on software mechanics, but on analytical interpretation and critical evaluation.

The activity is designed for upper-level undergraduate or early graduate courses in chemistry, biochemistry, medicinal chemistry, or structural biology.

2. Learning Objectives

By the end of this module, students should be able to:

- Interpret enzyme-ligand interactions within a crystallographic framework
- Define a rational docking search space
- Validate a docking protocol using redocking and RMSD
- Compare predicted binding modes of distinct ligands
- Critically assess docking scores and methodological limitations

3. Scientific Background

Acetylcholinesterase (AChE) contains a deep aromatic binding cavity that stabilizes ligands primarily through π–π stacking and hydrophobic interactions. The crystallographic structure 4M0E contains the enzyme in complex with dihydrotanshinone I (1YL), providing a validated structural reference for docking protocol calibration.

This module proceeds in two phases:

1. Protocol validation via redocking

2. Predictive docking of a clinically approved inhibitor (donepezil)

4. Experimental Structure Analysis

Students first examine the crystallographic complex to identify structural features of the active site and characterize experimentally observed interactions. As shown in Figure 1, the active-site forms a deep aromatic cavity responsible for ligand stabilization.

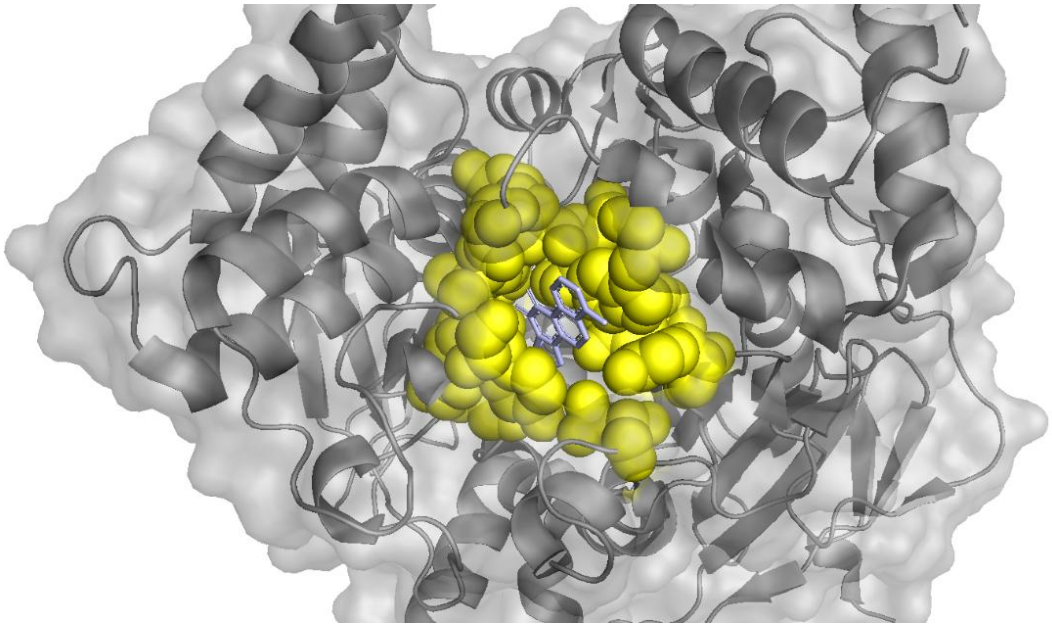

Figure 1: Three-dimensional structure of the target protein. The active site cavity is highlighted, indicating the region selected for docking analyses.

Key questions:

- What geometric features define the aromatic binding cavity?
- Which residues line the cavity and contribute to aromatic stabilization?
- What types of noncovalent interactions stabilize the ligand within the binding site?

5. Docking Workflow

5.1 Protein Preparation

• Chains B, C, and D were removed, retaining only chain A

• Crystallographic water molecules were deleted

• Non-relevant heteroatoms and auxiliary residues not directly involved in the binding site were removed

• Missing loop regions were reconstructed using the *Model/Refine Loops* tool in UCSF Chimera (MODELLER integration)

• Polar hydrogens were added

• Gasteiger charges were assigned

• The receptor was saved in PDBQT format

5.2 Ligand Preparation

Dihydrotanshinone I (1YL) was removed from the crystal structure (4M0E). Donepezil obtained from ChEMBL (ID: ChEMBL502), converted to 3D, minimized, and prepared.

To prepare donepezil structure, the following command was used:

obabel input.smi -O output.pdb --gen3d

obabel input.pdb -O output_min.pdb –minimize --ff MMFF94 --step500

To convert ligand structures into PDBQT format, the following command was used:

obabel input.pdb -O output.pdbqt --addhydrogens --partialcharge gasteiger

5.3 Grid Definition

The docking grid was centered on the crystallographic binding site, as illustrated in Figure 2, and dimensioned to fully encompass the aromatic binding cavity.

Chimera > Tools > Surface/Binding analysis > Autodock Vina.

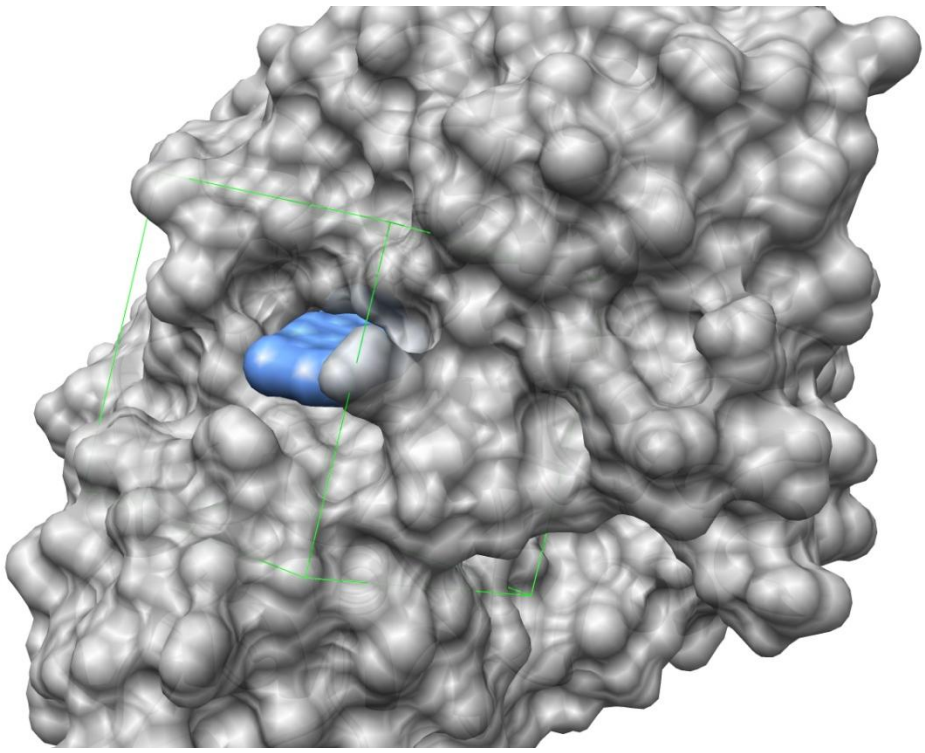

Figure 2: Surface representation of the protein showing the defined grid box (green) centered on the active site cavity used for molecular docking calculations.

5.4 Docking Execution (AutoDock Vina)

Docking calculations were performed using AutoDock Vina with a predefined configuration file containing receptor, ligand, grid center coordinates, and box dimensions. Command:

vina --config config_validation.txt / config_donepezil.txt

Configuration files (config_validation.txt and config_donepezil.txt), provided in the teaching package.

Each configuration file specifies:

- The receptor and ligand files
- The grid center coordinates
- The grid box dimensions
- The exhaustiveness parameter

Identical grid parameters were used for both validation and predictive docking to ensure methodological consistency.

6. Validation Through Redocking

The crystallographic ligand was removed and re-docked using the defined grid parameters. The predicted pose was aligned to the experimental structure, and RMSD was calculated excluding hydrogen atoms. Structural alignment and RMSD calculation were performed in PyMOL using:

rms_cur 1YL and not elem H, 1YL_docking and not elem H

An RMSD value of 0.219 Å was obtained. As shown in Figure 3, the superposition of the crystallographic and redocked poses demonstrates accurate reproduction of the experimental binding mode. An RMSD below 2.0 Å is generally considered indicative of successful protocol validation.

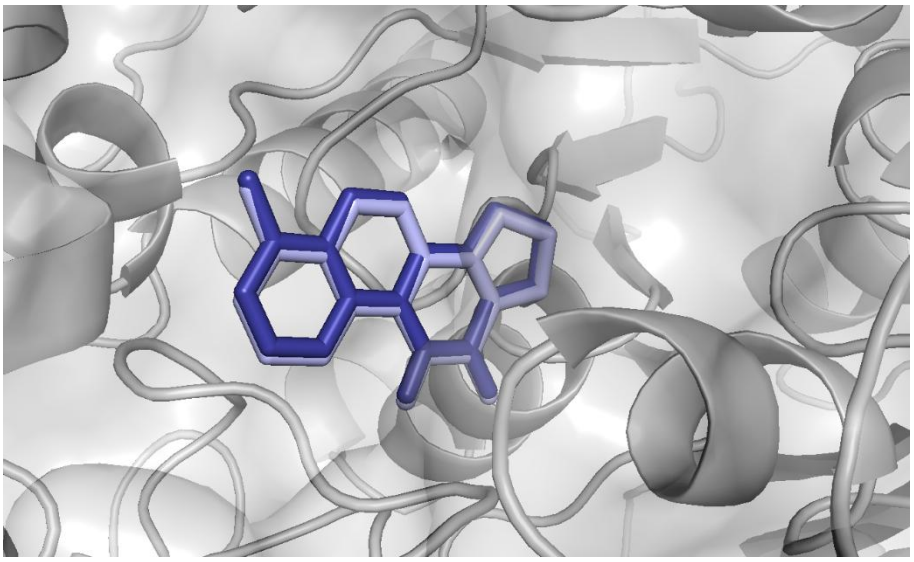

Figure 3: Superposition of the crystallographic ligand (lightblue) and the redocked (deepblue) pose within the active site, demonstrating the accuracy of the docking protocol.

Importantly, redocking validates protocol consistency but does not guarantee predictive reliability for new ligands.

7. Predictive Docking of Donepezil

Docking was performed using identical grid parameters to ensure methodological consistency and isolate ligand-dependent effects. Table 1 shows docking scores, and the structural comparison is shown in Figure 4.

Table 1: Docking scores of the co-crystallized ligand (1YL) and donepezil within the active site, enabling comparison of predicted binding affinities.

Ligand	Vina Score (kcal/mol)
1YL	-11.99
Donepezil	-9.50

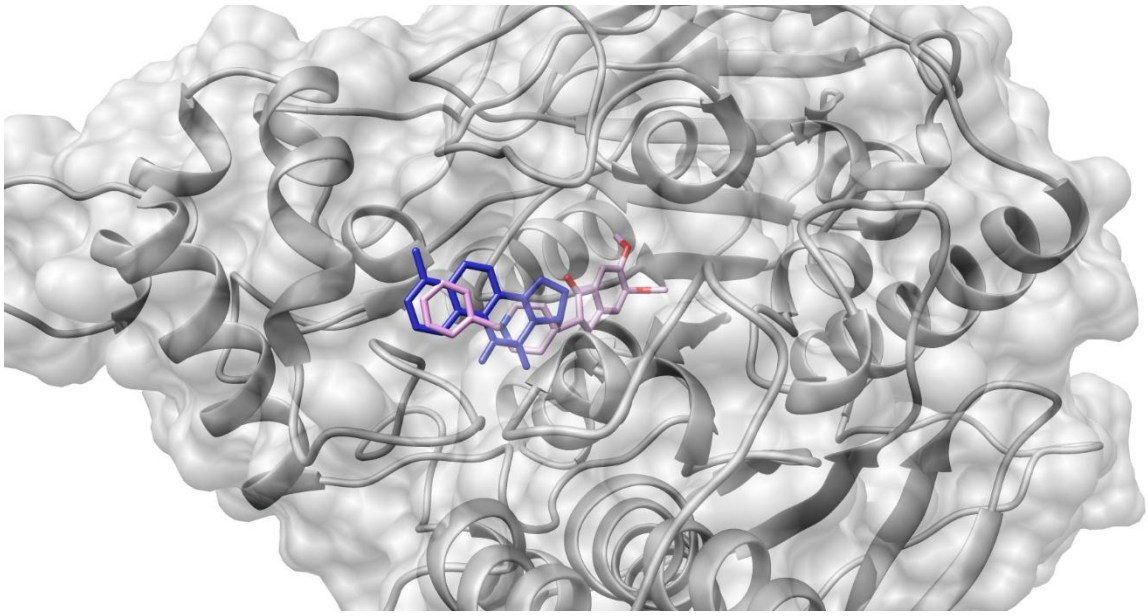

Figure 4: Overlay of the co-crystallized ligand (1YL) and donepezil within the active site, highlighting similarities in binding orientation and interaction pattern.

Despite the more favorable score for dihydrotanshinone I, docking energy does not directly equate to biological efficacy. Structural plausibility and interaction patterns must be evaluated.

8. Interaction Analysis

For visualization and analysis of protein–ligand interactions, 2D interaction diagrams were generated using the ProteinsPlus web server, while 3D interaction representations were produced with PLIP (Protein–Ligand Interaction Profiler). The

interaction profiles of the co-crystallized ligand and donepezil are shown in Figures 5-8.

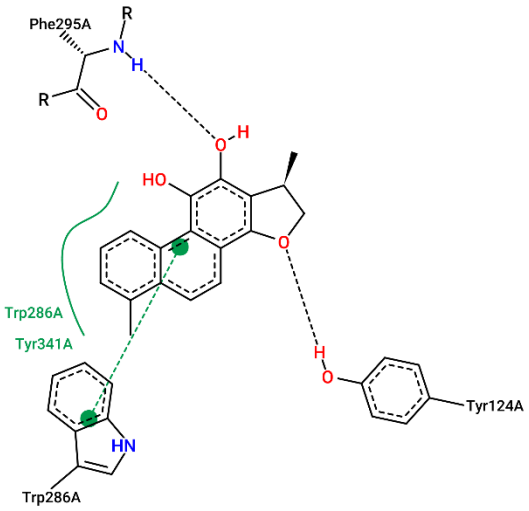

Figure 5: Two-dimensional interaction diagram of the co-crystallized ligand (1YL) within the active site, highlighting hydrogen bonds, π–π stacking interactions, and hydrophobic contacts with key residues.

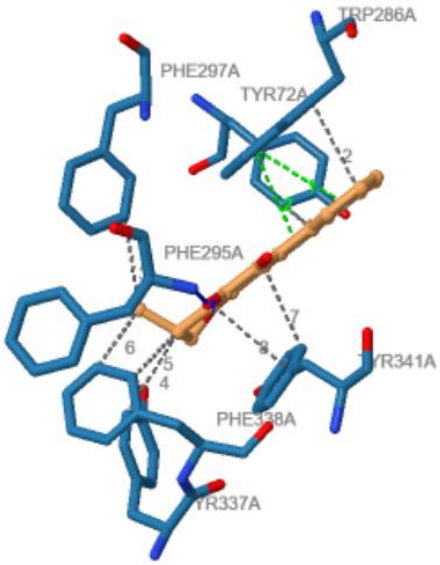

Figure 6: Three-dimensional binding mode of the co-crystallized ligand (1YL) in the active site.

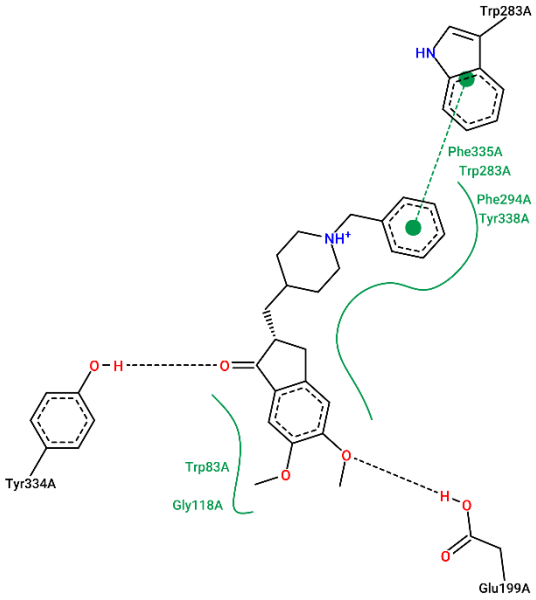

Figure 7: Two-dimensional interaction profile of donepezil within the active site, illustrating hydrogen bonding, π–π stacking, and hydrophobic interactions with surrounding residues.

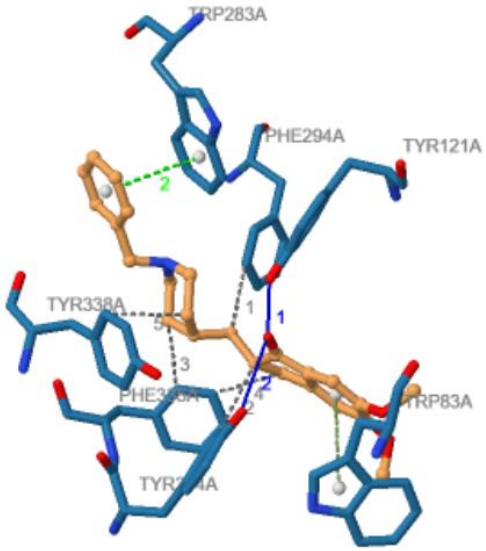

Figure 8: Three-dimensional representation of donepezil bound to the active site.

Comparative evaluation highlights differences in spatial occupancy and interaction networks, reinforcing that docking scores alone are insufficient descriptors of molecular recognition.

9. Critical Reflection and Discussion

Students are encouraged to evaluate:

- Why RMSD is an appropriate validation metric
- The limitations of rigid-receptor docking
- The assumptions embedded in scoring functions
- The difference between binding prediction and pharmacological activity
- The possibility of false positives

Docking is framed as a probabilistic modeling tool rather than definitive evidence of binding efficacy.

10. Implementation Guide

Estimated duration: 2–3 hours

Required software:

- UCSF Chimera
- AutoDock Tools
- AutoDock Vina
- PyMOL
- Open Babel

Optional interaction analysis tools:

- PLIP (Protein–Ligand Interaction Profiler) <https://plip- tool.biotec.tu-dresden.de>
- ProteinsPlus web server <https://proteinsplus.com>

Provided materials:

- Prepared receptor file
- Ligand files
- Configuration files
- High-resolution structural figures

The module is adaptable to upper-level undergraduate or early graduate courses and requires only foundational knowledge in chemistry or biochemistry.