Task 2: Molecular Docking of EGFR Family Kinase Domains

Objective

In this task, you will perform **blind molecular docking** of small molecules against the **kinase domains of the EGFR family** to compare binding affinities and interaction patterns.

You will integrate:

- Structural biology
- Protein structure modeling
- Molecular docking
- Comparative drug–target analysis

Biological Background (for Students)

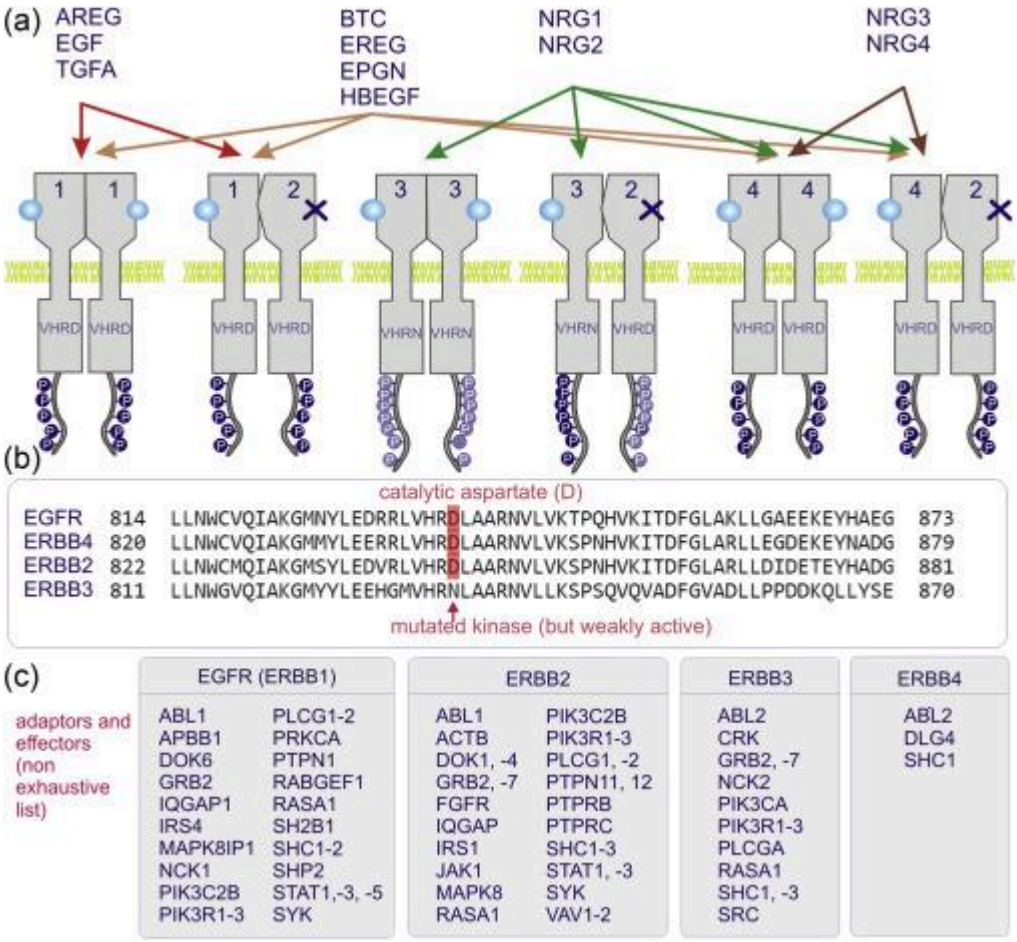
The **EGFR (Epidermal Growth Factor Receptor) family** consists of four closely related receptor tyrosine kinases:

- EGFR / ErbB1
- ErbB2 (HER2)
- ErbB3 (HER3)
- ErbB4 (HER4)

These proteins play a central role in:

- Cell growth and proliferation
- Cancer development
- Targeted cancer therapy

Many anti-cancer drugs, such as **osimertinib**, specifically inhibit the **kinase domain** of EGFR family members.



Overview of the Task

You will:

1. Prepare **complete kinase-domain structures** for EGFR family proteins
2. Prepare **small-molecule ligands**
3. Perform **blind docking**
4. Compare docking results across proteins and ligands

Task 4.1: Download Protein Structures from PDB

Objective

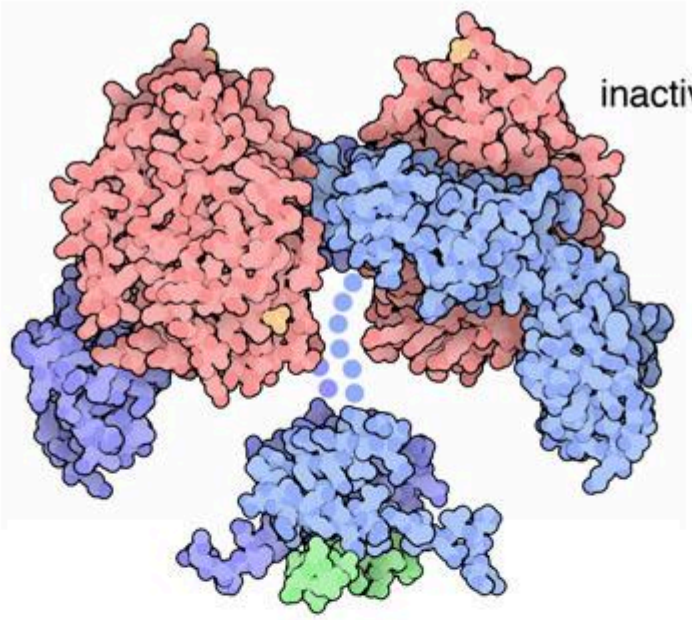
Obtain experimentally resolved structures of EGFR family kinase domains.

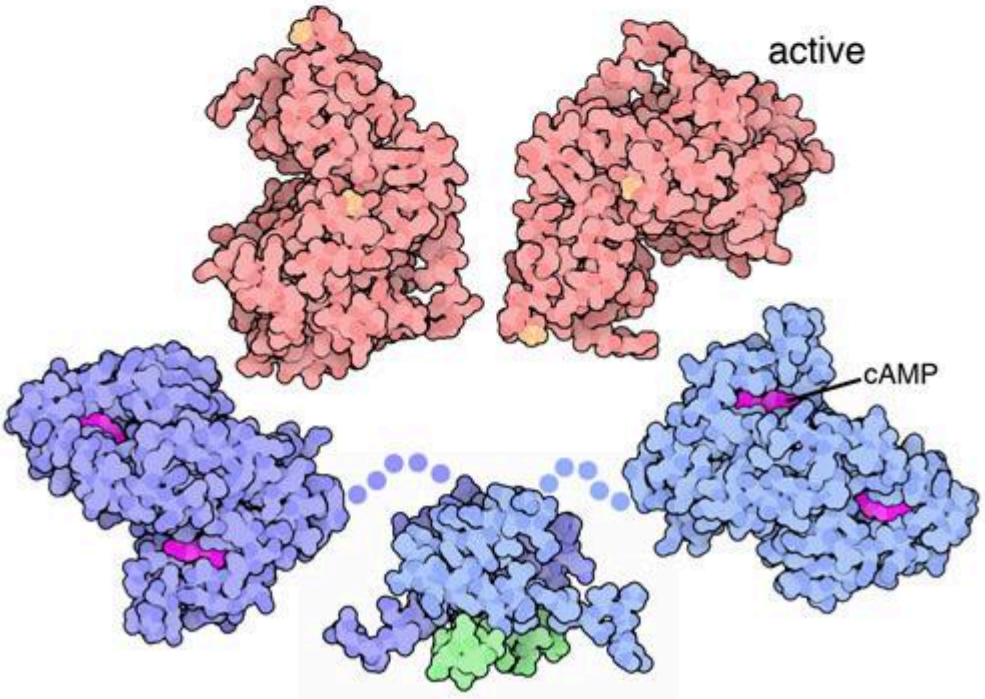
What you must do

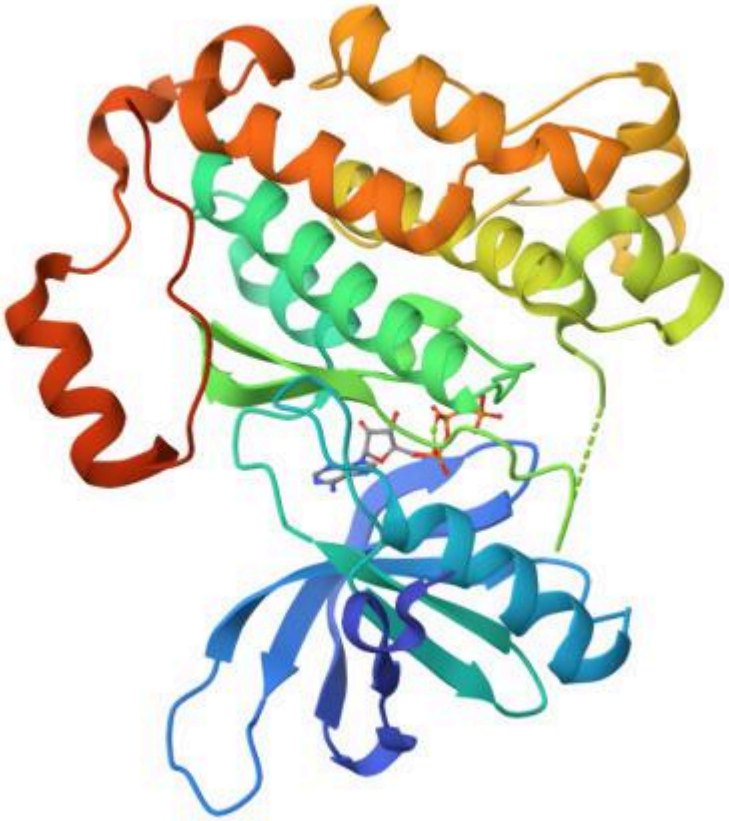
1. Search the **Protein Data Bank (PDB)** for kinase-domain structures of:
    - EGFR (ErbB1)
    - ErbB2
    - ErbB3
    - ErbB4
2. Select structures that:
    - Contain the kinase domain
    - Have reasonable resolution
3. Download the **PDB files**

Required output

- One PDB file per protein
- A table listing:
    - Protein name
    - PDB ID
    - Resolution
    - Missing residues (yes/no)


inactive


active
cAMP



Task 2.2: Complete Protein Structures Using AlphaFold

Objective

Generate **full-length, structurally complete kinase domains** by modeling missing residues.

What you must do

1. Identify **missing residues** in each PDB structure.
2. Use **AlphaFold** (or AlphaFold-derived models) to:
    - Predict the full kinase-domain structure
3. Merge:
    - Experimental PDB structure
    - AlphaFold-predicted regions
4. Clean the final structure:
    - Remove unresolved fragments
    - Ensure correct chain continuity

    - Remove crystallization artifacts (water, ions, ligands)

Required output

- One **final, cleaned PDB file** per protein
- Short description of:
    - Which residues were missing
    - How AlphaFold was used to complete them

Task 2.3: Ligand Preparation

Objective

Prepare bioactive compounds and a known anti-cancer drug for docking.

Ligands to use

Ligand	Type
Curcumin	Natural bioactive compound
Apigenin	Natural bioactive flavonoid
Osimertinib	FDA-approved anti-cancer drug

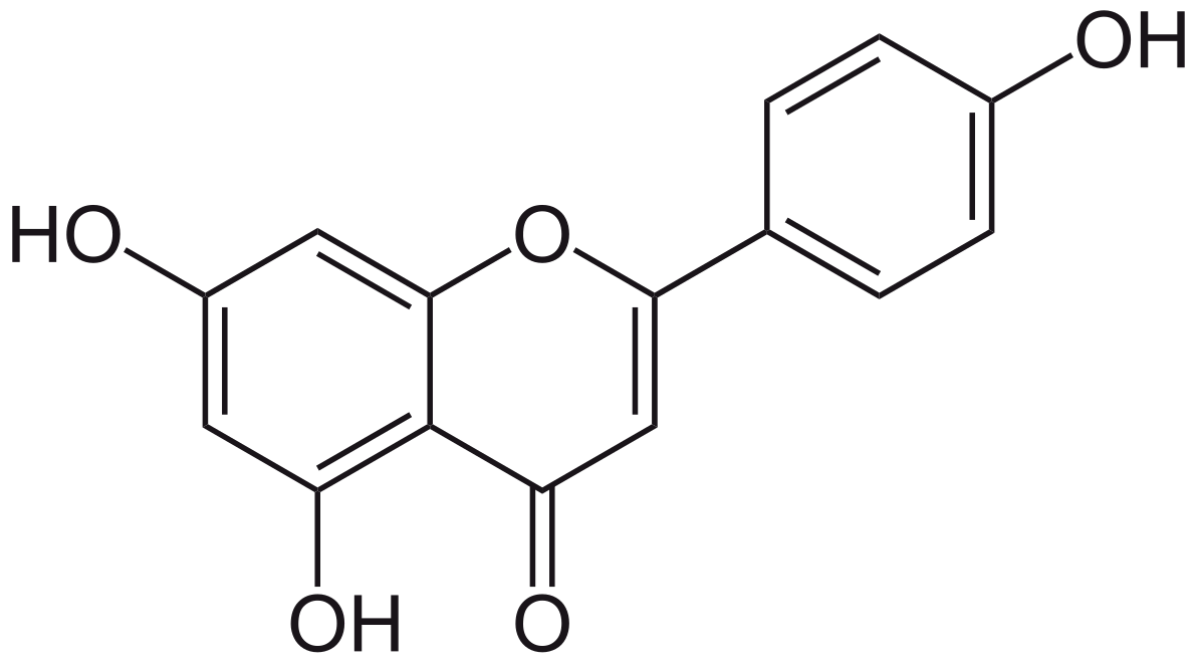
What you must do

1. Download ligand structures from a chemical database (e.g., PubChem)
2. Convert ligands to docking-ready format
3. Perform ligand preparation:
    - Add hydrogens
    - Optimize geometry
    - Assign charges

Required output

- One prepared ligand file per compound
- A table summarizing:
    - Ligand name

- Molecular weight
- Number of rotatable bonds



Task 2.4: Blind Molecular Docking

Objective

Perform **blind docking** of each ligand against each EGFR family kinase domain.

What you must do

1. Define a **blind docking grid** that covers the entire kinase domain
2. Dock:
    - Curcumin
    - Apigenin
    - Osimertinib

    against:
    - EGFR
    - ErbB2

    - ErbB3
    - ErbB4
3. Use the **same docking parameters** for all runs

📌 This results in **12 docking experiments** total.

Required output

For each docking run:

- Docking score / binding energy
- Best-scoring pose
- Docking log file

Task 2.5: Results Comparison and Analysis

Objective

Compare binding behavior across ligands and proteins.

What you must analyze

1. Compare **binding affinities**:
    - Natural compounds vs osimertinib
2. Compare **binding consistency** across EGFR family members
3. Identify:
    - Strongest ligand–protein combinations
    - Selectivity patterns

Required tables and figures

- A **comparison table** of docking scores
- Bar plots or heatmaps of binding energies
- Visualization of top docking poses in the kinase domain

Task 2.6: Biological Interpretation

Required written discussion

You must answer:

- How do curcumin and apigenin compare to osimertinib?
- Are natural compounds predicted to bind all EGFR family members equally?
- What structural features may explain differences in binding?
- What are the **limitations** of docking-based predictions?

Final Deliverables for Task 2

You must submit:

- ✅ Final cleaned protein PDB files
- ✅ Prepared ligand files
- ✅ Docking score tables
- ✅ Docking pose visualizations
- ✅ Comparative analysis and interpretation

Learning Outcomes

After completing this task, you should be able to:

- Prepare protein structures for docking
- Use AlphaFold to model missing regions
- Perform blind molecular docking
- Compare drug-like and natural compounds
- Interpret docking results in a cancer biology context