Project Report: Comparative Molecular Docking of Natural Compounds and Osimertinib against EGFR Family Kinase Domains

1. Executive Summary

This study employed a blind molecular docking approach to evaluate the binding potential of two natural compounds (Curcumin, Apigenin) compared to an FDA-approved tyrosine kinase inhibitor (Osimertinib) against the kinase domains of the EGFR family (ErbB1-4).

Key Finding: While Osimertinib demonstrated the highest selectivity across the family (Range: 1.44 kcal/mol), the natural flavonoid **Apigenin** displayed unexpectedly high binding affinities, outperforming Osimertinib in blind docking simulations against ErbB2 (-9.16 kcal/mol vs -8.37 kcal/mol). Curcumin consistently acted as a moderate, non-selective binder.

2. Objectives

The primary objectives of this computational study were to:

1. **Prepare structural models:** Integrate experimental X-ray crystallography data with AlphaFold predictions to account for missing residues in kinase domains[1111].
2. **Dock bioactive ligands:** Simulate the interaction of Curcumin, Apigenin, and Osimertinib against EGFR, ErbB2, ErbB3, and ErbB4[2].
3. **Analyze selectivity:** Determine if natural compounds show specific binding preferences compared to engineered drugs[3].

3. Methodology

3.1 Protein Structure Preparation

Target structures were retrieved from the RCSB Protein Data Bank (PDB) based on resolution and domain completeness[4].

- **EGFR (ErbB1):** PDB 1M17 (Res: 2.60 Å)
- **ErbB2 (HER2):** PDB 3PP0 (Res: 2.25 Å)
- **ErbB3 (HER3):** PDB 3LMG (Res: 2.80 Å)
- **ErbB4 (HER4):** PDB 3BBT (Res: 2.80 Å)

Structure Refinement:

Experimental structures were cleaned to remove crystallographic artifacts (water, ions). Missing loops and residues were identified and retrieved using the AlphaFold DB API, merging predicted coordinates with experimental backbones to ensure structural completeness 5.

3.2 Ligand Preparation

Structures were retrieved from PubChem and prepared using **RDKit** and **OpenBabel**:

1. **3D Generation:** 2D SDF data was converted to 3D conformers.
2. **Protonation:** Hydrogens were added at physiological pH (7.4)[6].
3. **Minimization:** Geometry was optimized using the MMFF94 force field[7].
4. **Format Conversion:** Structures were converted to PDBQT format, preserving torsion trees for rotatable bonds.

3.3 Molecular Docking Protocol

Engine: AutoDock Vina (v1.2.5+)

Method: Blind Docking[8].

- **Grid Box:** Centered on the geometric center of the protein, with dimensions extended 10Å beyond the protein extents to ensure full surface coverage (unbiased search).
- **Exhaustiveness:** Set to 32 to ensure thorough sampling of the conformational space[9].
- **Receptor State:** Rigid (side chains fixed).

4. Results

4.1 Binding Affinity Matrix (kcal/mol)

Lower values indicate stronger binding.

Ligand	EGFR (1M17)	ErbB2 (3PP0)	ErbB3 (3LMG)	ErbB4 (3BBT)	Average
Apigenin	-8.67	**-9.16**	-8.60	-8.43	**-8.72**

Osimertinib	**-8.69**	-8.37	-7.53	**-8.97**	-8.39
Curcumin	-7.61	-7.73	-7.62	-8.33	-7.82

4.2 Selectivity Analysis

Selectivity is defined as the difference between the strongest (min) and weakest (max) binding energy.

Ligand	Best Target	Worst Target	Selectivity Range (ΔG)	Classification
Osimertinib	ErbB4	ErbB3	**1.44**	Selective
Apigenin	ErbB2	ErbB4	0.74	Non-Selective
Curcumin	ErbB4	EGFR	0.72	Non-Selective

5. Biological Interpretation & Discussion

5.1 Drug vs. Natural Compound Efficacy

Comparison of the FDA-approved drug Osimertinib against natural bioactive compounds revealed distinct profiles:

- **Osimertinib:** As an engineered inhibitor, it displayed the highest single affinity against ErbB4 (-8.97 kcal/mol) and EGFR (-8.69 kcal/mol). This aligns with its biological role as a high-affinity tyrosine kinase inhibitor[10].
- **Apigenin:** Surprisingly, this flavonoid exhibited stronger average binding energies than the drug. Its rigid, planar structure (only 4 rotatable bonds) likely allows it to stack efficiently against surface residues in "blind" docking scenarios, potentially mimicking ATP's adenine ring.
- **Curcumin:** Displayed the weakest interaction energies. Its flexibility (12 rotatable bonds) incurs a higher entropic penalty upon binding, leading to less favorable scores compared to the rigid Apigenin.

5.2 Selectivity Profiles

- **Osimertinib (High Selectivity):** The drug showed a significant preference for ErbB4/EGFR over ErbB3 (Gap: 1.44 kcal/mol). This is consistent with the biological fact that ErbB3 has an impaired kinase domain (pseudokinase)[11111111], lacking the catalytic aspartate residues that typical inhibitors target.
- **Promiscuity of Natural Compounds:** Both Curcumin and Apigenin had narrow selectivity ranges (< 0.75 kcal/mol). This suggests they are "pan-ErbB" binders. In a biological context, this promiscuity implies they may affect multiple signaling pathways simultaneously, which can be beneficial for overcoming resistance but carries a higher risk of off-target toxicity.

5.3 Structural Comparison

- **Ligand Complexity:** Osimertinib (MW ~500) is significantly larger than Apigenin (MW ~270). In blind docking, smaller ligands like Apigenin can access smaller surface cavities that steric hindrance blocks for larger drugs.
- **ErbB2 (HER2) Anomaly:** Apigenin's superior score on ErbB2 (-9.16) vs Osimertinib (-8.37) warrants investigation. ErbB2 has a unique active conformation; Apigenin may be fitting into an allosteric pocket that is accessible in the crystallized conformation of 3PP0.

---

6. Limitations of the Study

1. **Blind Docking Bias:** By searching the entire protein surface, the algorithm may identify high-affinity surface sites that are not biologically relevant (non-functional sites), potentially inflating the scores of "sticky" molecules like Apigenin. Targeted docking into the ATP-binding pocket would likely show a greater advantage for Osimertinib.
2. **Receptor Rigidity:** The simulation treated proteins as rigid bodies. It did not account for the "induced fit" required for large inhibitors like Osimertinib to enter the deep binding pocket.
3. **Solvent Effects:** The docking was performed in a vacuum approximation, ignoring the energetic cost of desolvating the ligand before binding.

7. Conclusion

The docking pipeline successfully modeled the interaction of drug-like and natural molecules against the EGFR family. The results highlight that while **Osimertinib** possesses the selectivity characteristic of a designed therapeutic, the natural compound **Apigenin** exhibits high theoretical affinity for ErbB2, suggesting it may serve as a potent, albeit non-selective, lead compound for further structural optimization.