**Title: Mechanistic Analysis of Odorant Antagonism in Drosophila Or56a via Computational Docking and Structural Evaluation**

**Objective**

To explore a plausible molecular mechanism underlying the observed antagonism of the Drosophila melanogaster olfactory receptor Or56a by the odorant linalool, using geosmin as a reference agonist. The goal was to use computational tools to model ligand-receptor interactions, assess binding modes, and consider how different poses might influence receptor function.

**Approach and Thought Process**

My goal was not just to identify the best binding pose, but to *understand the range of binding possibilities*, how they relate to receptor activation or inhibition, and how these insights could inform a scalable, real-world discovery pipeline.

I approached the task by replicating a structure-based virtual screening workflow that could plausibly fit into Monarch's early-stage pipeline: combining structure prediction, ligand preparation, docking, and structural analysis.

I was particularly interested in whether linalool and geosmin bind the same pocket, and if not, whether their relative energies and spatial poses could support an allosteric or dual antagonism model. I also wanted to see how pose diversity could inform more robust computational strategies.

Additionally, I wanted to assess the plausibility of the Or56a AlphaFold-predicted pocket by comparing it to experimentally resolved insect olfactory receptors, namely 8V02 and 6C70. By aligning Or56a to these known structures, I hoped to evaluate the structural conservation of potential binding sites and further validate the biological interpretation of my docking results.

**Workflow Summary**

**1. Receptor Preparation**

* Obtained AlphaFold2 structure of DmelOr56a (unrelaxed model)
* Removed unnecessary tags and converted to PDBQT (rigid receptor)

**2. Ligand Preparation**

* Used RDKit to build and optimize 3D structures for geosmin and linalool using UFF
* Converted to PDBQT with Open Babel while retaining geometry and adding Gasteiger charges

**3. Docking with AutoDock Vina**

* Grid center: (0, 0, 0), Grid size: 30 Å cube
* Exhaustiveness: 8
* Both ligands docked flexibly against rigid Or56a

**4. Visualization & Analysis**

* Visualized binding poses using PyMOL
* Identified residues within 4 Å of ligand using selection-based queries
* Compared pose distribution across models 1–9 for each ligand

**5. Structural Alignment**

**A close-up of a colorful structure

AI-generated content may be incorrect.**

* Downloaded experimental OR structures 8V02 (2.9 Å) (Magenta) and 6C70 (3.5 Å) (Yellow)
* Aligned Or56a(Green) to both using PyMOL
* Measured RMSD values:
  + Or56a – 8V02: 5.32 Å
  + Or56a – 6C70: 11.43 Å
  + 8V02 – 6C70: 0.79 Å
* Overlayed all three structures to visualize fold similarity and pocket conservation

**Docking Results Summary (Or56a with Geosmin and Linalool)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Ligand** | **Model** | **Affinity (kcal/mol)** | **Pose Description** |
| Geosmin | 1 | **-6.06** | Deep TM pocket (orthosteric) |
| Geosmin | 4 | -5.08 | Distinct alternative surface pocket |
| Linalool | 1 | **-4.45** | Overlapping geosmin site (weaker fit) |
| Linalool | 5 | -4.23 | Alternate pose in different region |

**Interpretation and Mechanistic Hypothesis**

**1. Competitive Binding (Pose 1):** Both geosmin and linalool bind a common transmembrane pocket. Linalool's weaker binding affinity and altered pose within the same pocket suggest it may block or distort the agonist binding environment, consistent with **competitive antagonism**.

**2. Alternative Binding Site (Pose 4 for geosmin, Pose 5 for linalool):** The presence of energetically viable secondary poses in distinct pockets implies potential **multi-site behavior**. These may represent either:

* Allosteric binding sites that modulate receptor activation
* False positives due to docking uncertainty (though <1 kcal/mol energy difference supports their plausibility)

**3. Dual Antagonism Hypothesis:** Linalool may exert its antagonistic effect via two routes:

* By binding the orthosteric site and preventing geosmin engagement
* By stabilizing an alternate conformation that inhibits channel gating

This mechanism fits well with the poorly understood **non-GPCR** signaling architecture of insect ORs and aligns with the electrophysiological inhibition observed in literature.

**Structural Validation Using 8V02 and 6C70**

Despite only moderate structural similarity between Or56a and 8V02 (RMSD = 5.32 Å), the close alignment between 8V02 and 6C70 (RMSD = 0.788 Å) indicates that insect olfactory receptors—at least their transmembrane domains—follow a conserved structural framework. This is particularly noteworthy considering that both 8V02 and 6C70 come from different insect species (*Apocrypta bakeri*), whereas Or56a is from *Drosophila melanogaster*. Given these species-level differences, some structural divergence is expected. The higher RMSD between Or56a and 6C70 (11.43 Å) likely reflects a combination of phylogenetic distance and receptor subtype differences, and it emphasizes the importance of using multiple references when evaluating predicted structures.

Among the two experimental structures, 8V02 stands out not just for its structural proximity to Or56a but also for its higher resolution (2.9 Å vs 3.5 Å for 6C70), making it a more suitable benchmark for binding site assessment. By overlaying Or56a with 8V02 and visualizing the predicted pocket location, we observed that the docking poses for geosmin and linalool fall within a transmembrane region that is spatially consistent with known odorant-binding cavities. This overlap reinforces the plausibility of the docking predictions and supports the idea that even in the absence of an experimentally solved Or56a structure, AlphaFold-based models can provide structurally reasonable starting points—especially when interpreted in the context of conserved fold families.

These structural comparisons allowed me to place my docking results into a broader biological framework. While some caution is warranted due to species and model limitations, the high conservation between 8V02 and 6C70 provides strong indirect support for using this framework to interpret the behavior of Or56a. It also highlights how reference-based validation can be a practical tool for increasing confidence in predicted pocket locations and guiding mechanistic hypotheses when experimental structures are unavailable.

**Scalability Considerations**

This project underscored how valuable it is to look beyond single docking scores and instead capture the full landscape of ligand poses and binding site environments. In a realistic drug or odorant discovery pipeline, pose diversity and binding topology are often more informative than just choosing the “top-scoring” pose. The fact that geosmin and linalool exhibit multiple plausible binding modes — including ones that likely reflect orthosteric and allosteric interactions — suggests that a scalable pipeline must be designed to preserve and interpret this complexity.

In a real-world setting, I would envision a tiered workflow that starts with high-throughput docking and pose clustering to prioritize ligands not just by affinity, but by the diversity and consistency of their binding modes. These top candidates could then be analyzed for key residue contacts and binding-site overlap, followed by molecular dynamics (MD) simulations to understand conformational flexibility, receptor gating effects, and pose stability.

Finally, to bridge accuracy with scale, machine learning models could be trained not just on docking scores but on a combination of pose clustering outcomes, residue-level interaction fingerprints, and energetics from QM or MD-refined subsets. Such models could learn to recognize antagonistic behaviors based on structural and energetic patterns rather than simplistic thresholds.

From my perspective as a quantum chemist, I believe that even as we aim for scale, it is essential not to lose sight of accuracy and physical interpretability. My own research has focused on combining quantum chemical methods, fragmentation strategies, and machine learning to predict interaction energies in complex systems like water-ion clusters and property prediction in biomolecular systems. These same techniques could be extended to odorant-receptor modeling. While QM is computationally expensive, selective application on key complexes and the integration of delta-learning or hybrid ML/QM workflows could provide more reliable predictions. In the long run, I believe that scalable pipelines will need to reflect a thoughtful balance of physics-based accuracy, chemical intuition, and data-driven generalization — and this project gave me an opportunity to prototype exactly that kind of framework.

**Conclusion and Perspective**

This assignment demonstrates that even with simple docking, we can extract rich mechanistic insight. The ability of linalool to engage Or56a in multiple modes, including one that overlaps the geosmin site, strongly supports its role as a competitive and potentially allosteric antagonist. A flexible docking framework that accounts for these variations will be essential for scalable discovery at Monarch.

Structural alignment with experimental OR structures (8V02 and 6C70) provide further confidence in the biological plausibility of the predicted binding site. The high similarity between 8V02 and 6C70, despite being from different species, emphasizes the robustness of using conserved architecture to guide computational predictions in under-characterized receptors like Or56a.

As a quantum chemist, I approach this protocol with a healthy level of skepticism — particularly around the use of AlphaFold structures and rigid docking as definitive predictors of receptor-ligand behavior. While AlphaFold provides an invaluable starting point, it is not explicitly trained on membrane-bound olfactory receptors or allosteric pocket geometries, and its predictions do not account for conformational flexibility or ligand-induced shifts. Similarly, docking relies on static snapshots of dynamic systems and often oversimplifies the energetic landscape. In a more rigorous setting, I would consider refining the structural model using molecular dynamics to explore relevant conformational ensembles, followed by higher-level quantum mechanical calculations — such as single-point energy evaluations using GFN2-xTB or fragmentation-based DFT — to better capture interaction energetics and polarization effects. While it's computationally impractical to apply QM across large libraries or full receptor systems, a hybrid, top-down approach that combines structure prediction, dynamics sampling, and quantum refinement on key poses could yield far more reliable and interpretable results.

In parallel, I see machine learning as a natural bridge between the scale of docking and the accuracy of quantum chemistry. Once a representative set of ligands–receptor complexes is characterized through MD and QM refinement, that data can be used to train ML models — whether through Δ-learning, interaction fingerprinting, or graph-based architectures — to generalize quantum accuracy across a broader chemical space. This enables a practical middle ground: QM-derived accuracy at ML speed. In fact, my own research focuses on building ML-based force fields and energy predictors for complex many-body systems, which I believe could extend to odorant-receptor interactions if trained properly. Ultimately, I see AlphaFold and docking not as end points, but as entry points to a layered framework where ML models serve as the glue between physical realism and real-world throughput.