Simulation of Entry Flow: In-Silico Modeling Use Case

1. User Initiates Interaction:

User: Types into the interface: "I want to design a new molecule to inhibit the activity of 'TargetX' in 'OrganismY'."

2. Target Identification Agent Engages:

- **Input:** User query: "design a new molecule to inhibit the activity of 'TargetX' in 'OrganismY'."
- **Action:** Parses the query and identifies "TargetX" and "OrganismY".
- **Dummy UniProt Search:** Sends a dummy query to a simulated UniProt database for "TargetX" in "OrganismY".
- **Dummy UniProt Response:** Receives the following dummy information:
  - Protein Name: TargetX
  - UniProt ID: P12345
  - Dummy Sequence: "ABCDEFG..." (a placeholder sequence)
  - Potential PDB IDs: ["1ABC", "1DEF", "2GHI"] (representing different structural snapshots or related structures)

3. Information Enrichment Agent Activates:

- **Input:** TargetX, OrganismY, UniProt ID: P12345, Potential PDB IDs: ["1ABC", "1DEF", "2GHI"]
- **Action:**
  - **Dummy PubChem Search:** Sends a dummy query to a simulated PubChem for "TargetX inhibitor" and UniProt ID "P12345".
  - **Dummy PubChem Response:** Returns a dummy list of 3 existing compounds with:
    - Dummy Compound 1: SMILES "C1=CC=CC=C1O", PubChem CID 101
    - Dummy Compound 2: SMILES "N#CC=CC#N", PubChem CID 102
    - Dummy Compound 3: SMILES "CC(=O)C1CC1", PubChem CID 103
  - **Dummy ChEMBL Search:** Sends a dummy query to a simulated ChEMBL for bioactivity data related to "TargetX" and UniProt ID "P12345".
  - **Dummy ChEMBL Response:** Returns dummy bioactivity data for 2 compounds:
    - ChEMBL Compound A: ChEMBL ID CHEMBL100, IC$_{50}$ = 100 nM
    - ChEMBL Compound B: ChEMBL ID CHEMBL101, K$_i$ = 50 nM
  - **Dummy PDB Retrieval:** "Attempts" to retrieve 3D structures for PDB IDs "1ABC", "1DEF", and "2GHI" from a simulated PDB. Let's say it successfully retrieves all three.

4. Property Prioritization Agent Considers:

- **Input:** Target information, existing ligands (dummy), bioactivity data (dummy).
- **Action:** Based on the goal (inhibition) and the characteristics of the dummy existing ligands (some are aromatic, some have nitriles, etc.), the agent might prioritize:
  - Potential for specific interactions with a dummy active site (e.g., hydrogen bonding, hydrophobic interactions).
  - Dummy drug-like property ranges (e.g., Molecular Weight between 200-500 Da).

5. Molecule Generation/Retrieval Agent Proceeds:

- **Input:** Prioritized properties, dummy existing ligands.
- **Action:**
  - **Decision (Dummy):** Let's say the agent decides to perform *de novo* design based on a generic scaffold and incorporating some features from the dummy existing ligands.
  - **Dummy RDKit Usage:** Uses a simulated RDKit to generate 5 dummy novel molecules with varying structures. For example:
    - Dummy Molecule 1: SMILES "C1CNCCC1CC(=O)O"
    - Dummy Molecule 2: SMILES "c1ccccc1NCCC"
    - Dummy Molecule 3: SMILES "O=C(N)c1ccncc1"
    - Dummy Molecule 4: SMILES "ClC1=CC=CC=C1"
    - Dummy Molecule 5: SMILES "CC(C)C(=O)NCC"

6. 3D Structure Generation Agent Works:

- **Input:** SMILES strings of the 5 dummy novel molecules.
- **Action: Dummy RDKit Usage:** Uses a simulated RDKit to generate 3D coordinates for each of the 5 dummy molecules.

7. Molecular Docking Agent Executes:

- **Input:** 3D structures of the 5 dummy molecules and the 3 dummy PDB structures ("1ABC", "1DEF", "2GHI").
- **Action: Dummy Docking:** Runs a simulated docking program (e.g., AutoDock Vina) for each of the 5 molecules against each of the 3 protein structures.
- **Dummy Output:** Generates dummy binding scores (e.g., in kcal/mol) for each molecule-protein complex.

8. Property Prediction Agent Analyzes:

- **Input:** SMILES strings of the 5 dummy novel molecules.
- **Action: Dummy RDKit Usage:** Uses a simulated RDKit to calculate dummy molecular properties for each molecule (e.g., Molecular Weight, LogP, Number of H-bond donors/acceptors).

9. Filtering and Ranking Agent Evaluates:

- **Input:** Dummy docking scores and dummy molecular properties.
- **Action:** Applies dummy filtering criteria (e.g., binding score < -7.0 kcal/mol, Molecular Weight < 450 Da) and ranks the remaining molecules based on a dummy scoring function (e.g., primarily based on average binding score across the three PDB structures, with a penalty for unfavorable properties).
- **Dummy Output:** Ranks the 5 dummy molecules. Let's say the ranking is: Molecule 3 > Molecule 1 > Molecule 5 > Molecule 2 > Molecule 4.

10. Result Presentation Agent Shows Output:

- **Input:** Ranked list of dummy molecules with their dummy data.
- **Action:** Presents a dummy output to the user:

Rank	Dummy Molecule (SMILES)	Dummy Avg. Binding Score (kcal/mol)	Dummy MW	Dummy LogP	...
1	O=C(N)c1ccncc1	-8.2	121.14	0.5	...
2	C1CNCCC1CC(=O)O	-7.8	129.17	0.1	...
3	CC(C)C(=O)NCC	-7.5	115.18	1.2	...
4	c1ccccc1NCCC	-6.5	133.20	2.1	...
5	ClC1=CC=CC=C1	-5.9	112.56	2.6	...

It might also show 2D structures (dummy renderings) and mention the dummy PDB structures used.

11. Feedback Loop Agent Waits:

- **User:** Might provide feedback like "Show me more molecules similar to the top-ranked one" or "Filter for molecules with higher LogP".
- **Action:** The system would then use this dummy feedback to adjust parameters and reiterate the process (e.g., instruct the Molecule Generation Agent to generate more analogs of "O=C(N)c1ccncc1" or the Filtering Agent to apply a different LogP range).