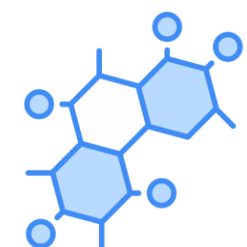# Stage 1. Warm-up Stage: Mol-SFT
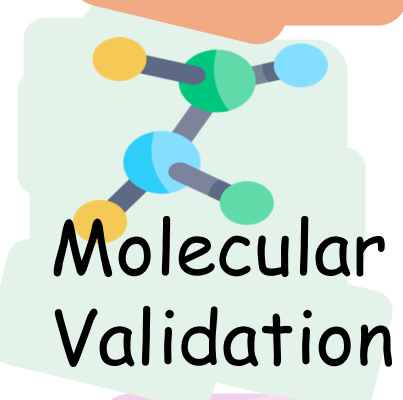
Knowledge-Guided CoT Construction

Caption

Molecule

GPT-4o →

Distillation

Molecular Validation

Response Validation

Molecule Captioning w/ CoT
1. Understand the SELFIES String...
2. Interpretation of Fragments...
3. Analysis of Chemical Characteristics...
4. Generate the Chemical Description...
<answer> The molecule is a ...</answer>

Text-based de novo Molecule Generation w/ CoT
1. Understand the Text Description...
2. Extract Key Structural Features...
3. Combine Fragments and Components...
4. Verify Structural Information...
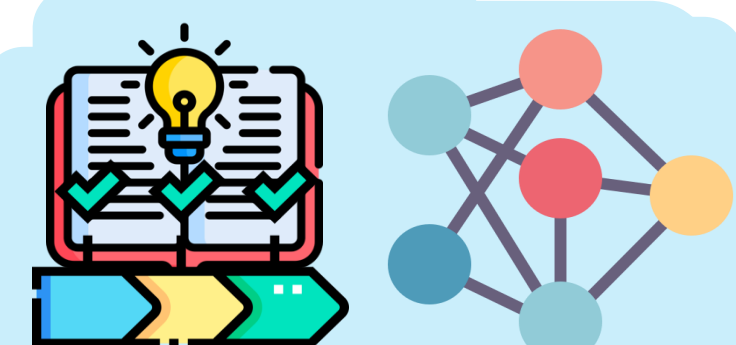
<answer> [C][C@H1][C@@H1][Branch2]...</answer>

Mol-SFT

To generate the SELFIES representation, ...\n\n1. **Identify the Base Molecule**:...\n2. **Oleoyl Chain**: ...The oleoyl chain is represented as [C@@H1][C@@H1][C@@H1][C@@H1][C@@H1][C@@H1][C@@H1][C@@H1]..C@@H1][C@@H1]

Qwen

Cold Start SFT w/ Knowledge-Guided CoT

To reconstruct or deduce the molecular structure…, 1. **Identify the Base Structure**…, which consists of two glycerol moieties linked to oleic acid and a phosphate group…6. **Linking the Oleoyl Chains**:…Each oleoyl chain is a long hydrocarbon chain (C18) with a double bond (C=C) at the appropriate position,…the final representation of the structure is:
<answer>[C][C][C][C]...[=Branch1][C][C]...[C][C][C]</answer>
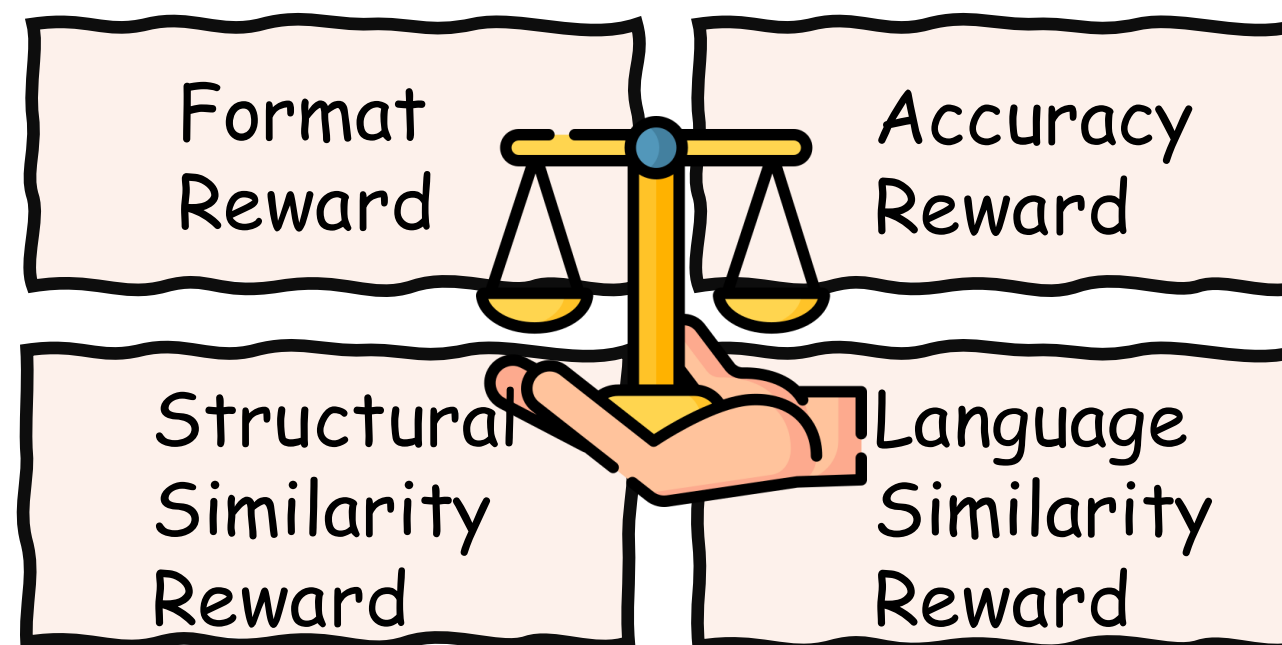
Qwen w/ Cold Start

✅ Learning CoT Reward Format   ✅ Providing Shallow Reasoning Prior

- - -

Stage 2. Reinforcement Learning Stage: Mol-RL

GRPO

Qwen w/ Cold Start →

✅ ❌ ✅ ❌ ❌

Sampled Answer

Multi-Level Reward

Format Reward | Accuracy Reward

Structural Similarity Reward | Language Similarity Reward

To deduce the...2. **Understand the Phosphate Group**: ...the phosphate group has a negative charge (O-)...3. **Incorporate the Oleoyl Groups**: ...(C18) with a double bond...5. **Connect the Components**:...Attach the oleoyl groups to the first and second carbon atoms of each glycerol molecule...Link the phosphate group to the third carbon atom of each glycerol molecule... 7. **Use the Fragments**:...The fragment <|O=P([O-])(O)O|> represents the phosphate...final answer is:
<answer>[C][C][C]...[P][=Branch1][C][=O]...[C]</answer>

✅ Chemical Consistency and Validity   ✅ Deep Global and Local Reasoning