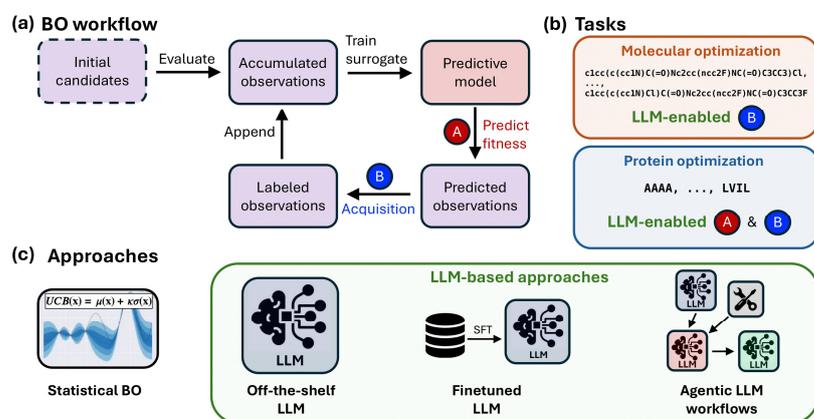


## Conclusion

Improving Bayesian optimization (BO) for biochemical discovery with LLMs requires carefully aligning the approach to the task structure. Reasoning LLMs can generate domain-aware hypotheses, but they often overfixate on irrelevant context. Agentic workflows help by offloading filtering to tools, and finetuned models can mimic Bayesian behavior efficiently.



## Method

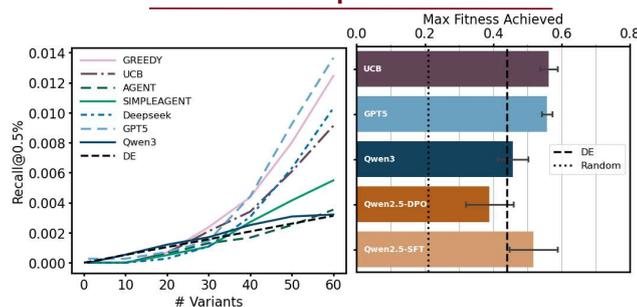
We benchmarked two types of scientific tasks: molecular optimization using string-based representations and optimization of four-residue binding motifs within proteins. Statistical BO was compared against off-the-shelf LLMs, finetuned LLMs, and agentic workflows with a minimal toolset by highest discovered fitness and recall.

## Results

- Off-the-shelf LLMs** (GPT5, Qwen3, Deepseek) performed well on the protein optimization task, but failed on the molecular task, due to SMILES misinterpretation and long-context degradation.
- Agentic workflows** (AGENT) showed inverse pattern. Paradoxically, reducing contextual information improved performance (SIMPLEAGENT), indicating constrained prompts mitigate hyperfixation.
- Smaller non-reasoning models could be **finetuned** (Qwen2.5-SFT/DPO) to emulate BO strategies (Greedy, UCB), exceeding their reasoning counterparts.
- Optimal performance** for each approach tracked closely with the task structure as summarized in the table:

Condition	Off-the-shelf	Agentic Workflow	Fine-tuned (SFT)
<b>Search space</b>	Generative and dense - LLM should reliably generate candidates close to search space	Indifferent - tools adapted for any domain	Structured / predictable - possible to synthesise faithful training data
<b>Representations</b>	Simple	Indifferent	Not evaluated
<b>Budget</b>	Small	Large	Not evaluated
<b>Inference time</b>	Medium-Very High	Medium-High	Low
<b>Failure modes</b>	Long contexts; SMILES misparsing	Over-reliance on tools; reduced goal generation; Pipeline complexity	Generalization to training domain; sensitive to training method

## Protein optimization



## Molecular optimization

