



000 001 002 003 004 005 006 007 008 009 010 011 012 013 014 015 016 017 018 019 020 021 022 023 024 025 026 027 028 029 030 031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047 048 049 050 051 052 053 CHEM-R: LEARNING TO REASON AS A CHEMIST

Anonymous authors

Paper under double-blind review

ABSTRACT

Although large language models (LLMs) have significant potential to advance chemical discovery, current LLMs lack core chemical knowledge, produce unreliable reasoning trajectories, and exhibit suboptimal performance across diverse chemical tasks. To address these challenges, we propose ***Chem-R***, a generalizable ***Chemical Reasoning*** model designed to emulate the deliberative processes of chemists. Chem-R is trained through a three-phase framework that progressively builds advanced reasoning capabilities, including: 1) Chemical Foundation Training, which establishes core chemical knowledge. 2) Chemical Reasoning Protocol Distillation, incorporating structured, expert-like reasoning traces to guide systematic and reliable problem solving. 3) Multi-task Group Relative Policy Optimization that optimizes the model for balanced performance across diverse molecular- and reaction-level tasks. This structured pipeline enables Chem-R to achieve state-of-the-art performance on comprehensive benchmarks, surpassing leading large language models, including Gemini-2.5-Pro and DeepSeek-R1, by up to 32% on molecular tasks and 48% on reaction tasks. Meanwhile, Chem-R also consistently outperforms the existing chemical foundation models across both molecular and reaction level tasks. These results highlight Chem-R’s robust generalization, interpretability, and potential as a foundation for next-generation AI-driven chemical discovery.

1 INTRODUCTION

Large Language Models (LLMs) have recently emerged as a transformative force in scientific discovery (Bai et al., 2025; Ma et al., 2024; Shojaee et al., 2024; Hatakeyama-Sato et al., 2023; Xia et al., 2025). Within the field of chemistry, LLMs demonstrate exceptional potential by learning expressive representations and knowledge of molecular structures and chemical reactions directly from large-scale datasets. This capability enables them to support a wide array of tasks, including molecular property prediction, reaction outcome estimation, retrosynthetic route planning, and reagent selection (Zhang et al., 2024b; Tan et al., 2025; Zhao et al., 2024; Jiang et al., 2025). However, the lack of chemical structured and reliable reasoning processes in current LLMs leads to suboptimal performance and limited interpretability on complex chemical problems.

Specifically, current LLMs encounter three fundamental challenges in performing chemical reasoning. **Challenge 1:** Current LLMs often lack the essential “chemical fundamentals”, leading to frequent mistakes in molecular representations and reaction rules, which undermines reliability at the initial reasoning stage (Zhong et al., 2024; Liu et al., 2025). As illustrated in Figure 1 (Challenge 1), several Chain-of-Thought (CoT) on different tasks generated by DeepSeek-R1 (Guo et al., 2025) demonstrates that the model may fail to recognize basic SMILES (Weininger, 1988) and IUPAC (Kuhn et al., 2004), which undermines the reliability of any subsequent reasoning. **Challenge 2:** The model’s reasoning process is fundamentally flawed because it is unsystematic, failing to adhere to the coherent, step-by-step workflow of an expert (Ouyang et al., 2023; Bran et al., 2025). This lack of a structured approach makes the reasoning unreliable and prone to factual errors. This lack of structure results in a confusing and untrustworthy reasoning process, as exemplified in Figure 1 (Challenge 2) where the model misidentifies fundamental functional groups and generates a flawed, unstructured line of reasoning. **Challenge 3:** As shown in Figure 1 (Challenge 3), even when models are guided by explicit reasoning patterns, their performance across diverse molecular and reaction level tasks remains highly imbalanced, with strong tasks dominating and weaker tasks

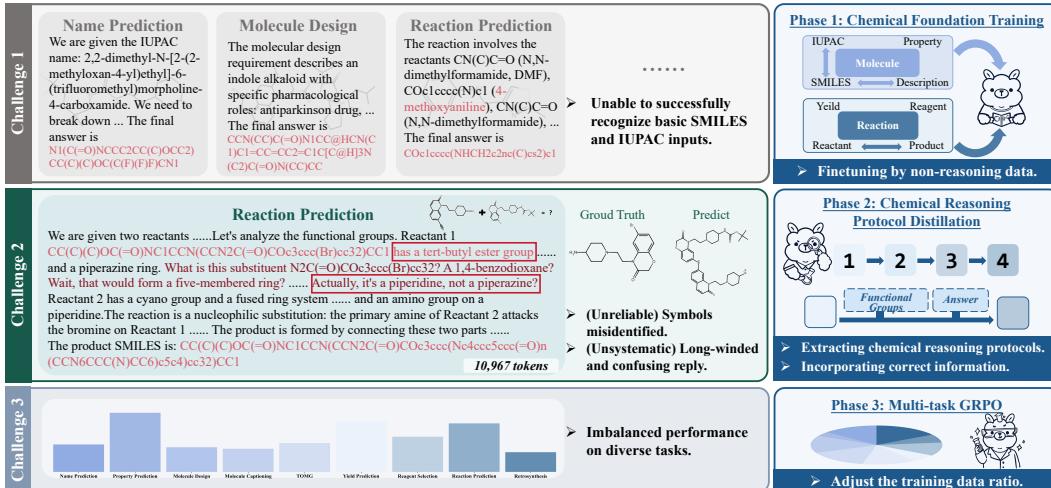


Figure 1: **Challenges and the proposed Chem-R solution.** The left panel highlights three key deficiencies observed in current reasoning models. To overcome these limitations, we introduce a three-phase training framework, illustrated on the right. This strategy is designed to first build a solid chemical foundation (Phase 1), then instill correct, step-by-step reasoning pathways (Phase 2), and finally, optimize for balanced, multi-task proficiency (Phase 3).

underrepresented. Together, these issues highlight that effective chemical reasoning requires domain knowledge, reliable and structured thought, and balanced generalization across heterogeneous tasks.

To address these challenges, we propose Chem-R, a unified framework comprising three phases that enables structured reasoning in molecular- and reaction-level tasks. As shown in Fig. 1, Chem-R follows a three-phase training paradigm, where each phase systematically mitigates one of the aforementioned bottlenecks. **Phase 1: Chemical Foundation Training** equips the model with robust chemical fundamentals by fine-tuning on large-scale non-reasoning corpora, covering both molecular representations (*e.g.*, SMILES, IUPAC) and reaction-level patterns, thereby reducing elementary errors. **Phase 2: Chemical Reasoning Protocol (CRP) Distillation** leverages structured protocols to guide a general-purpose teacher model toward expert-level chemical reasoning, subsequently distilling these strategies into a student model. In this process, expert-like protocols are converted into reusable, modular workflows that facilitate coherent and interpretable problem-solving. As illustrated in bottom row of Fig. 1, the reasoning process can be enhanced by extracting structured Chemical Reasoning Protocols, incorporating correction information, and providing targeted hints to mitigate common errors. **Phase 3: Multi-task Group Relative Policy Optimization (Multi-task GRPO)** further enhances the learned reasoning paradigm across heterogeneous tasks, employing a curriculum-like weighting scheme to prevent strong-task dominance and improve performance balance. Together, these three phases form a principled pipeline that not only reduces low-level mistakes, but also enables the model to generate chemically sound, structured, and explainable reasoning across both molecular and reaction domains.

Our main contributions are summarized as follows:

(1) We propose Chem-R, a unified three-phase framework that enables structured and generalizable chemical reasoning across both molecular and reaction level tasks. **Phase 1 (Chemical Foundation Training)** equips the model with robust chemical fundamentals by pre-training on large-scale non-reasoning corpora. **Phase 2 (CRP Distillation)** introduces Chemical Reasoning Protocols (CRP) distilled from a teacher model, providing modular and interpretable workflows that guide problem solving. **Phase 3 (Multi-task GRPO)** applies GRPO with a curriculum-like weighting scheme to enhance and balance performance across heterogeneous tasks.

(2) We comprehensively evaluate the model on four widely used benchmarks, including Chem-LLMBench (Guo et al., 2023), ChEBI-20 (Edwards et al., 2022), TOMG-Bench (Li et al., 2024a), and USPTO (Schneider et al., 2016). Our evaluation spans two major families of tasks, namely molecular- and reaction-level tasks, covering nine macro-tasks and 25 sub-tasks in total. Across this diverse suite, Chem-R consistently achieves state-of-the-art performance. For instance, compared with ChemDFM-v1.0-13B (Zhao et al., 2025c), Gemini-2.5-Pro (Comanici et al., 2025), and DeepSeek-R1 (Guo et al., 2025), Chem-R improves by 33%, 32%, and 44% on Name Prediction (Exact Match), and by 53%, 50%, and 52% on Yield Prediction (Accuracy), respectively. These sub-

108 stantial gains highlight Chem-R’s ability to deliver both reliable accuracy and robust generalization
109 across heterogeneous molecular and reaction tasks.
110

112 2 RELATED WORK 113

114 2.1 REASONING FOR LLMs 115

116 Generating a CoT (Wei et al., 2022; Kojima et al., 2022) significantly improves the ability of LLMs
117 to perform complex reasoning. To elicit high-quality reasoning chains, various strategies have
118 been proposed, including rejection sampling (Liu et al., 2023a; Tong et al., 2024), reward mod-
119 eling (Ouyang et al., 2022; Zhang et al., 2025b), and preference learning (Rafailov et al., 2023; Pang
120 et al., 2024). More recently, DeepSeek-R1 (Guo et al., 2025) has shown that complex reasoning be-
121 haviors (Gandhi et al., 2025) can emerge from simple rule-based reinforcement learning, particularly
122 when initialized with a cold start phase using CoT data.

123 However, a common limitation of these approaches is their reliance on outcome-based supervision,
124 which can produce unstructured, inconsistent and flawed reasoning chains (Arcuschin et al., 2025;
125 Chen et al., 2025), a critical risk in scientific applications. To address this, process-level supervision
126 provides fine-grained feedback on each intermediate step (Lightman et al., 2024; Wang et al., 2024;
127 Zhang et al., 2024a; 2025a). Another strategy involves multi-model systems where verifier models
128 scrutinize the reasoning process of a primary generator model (Du et al., 2023; Kirchner et al., 2024;
129 Baker et al., 2025).

131 2.2 LLMs FOR CHEMISTRY 132

133 The application of LLMs is driving a paradigm shift in chemistry. Early work demonstrated that
134 generalist models possess latent chemical knowledge (Hatakeyama-Sato et al., 2023; Sallam et al.,
135 2024), paving the way for specialized models fine-tuned for chemistry-specific tasks. These include
136 models like ChemLLM (Zhang et al., 2024b), ChemMLLM (Tan et al., 2025), Chem3DLM (Jiang
137 et al., 2025), ChemDFM (Zhao et al., 2025c) and others (Liu et al., 2023c; Zhang et al., 2025c; Li
138 et al., 2025c), which handle tasks ranging from molecular captioning to reaction analysis.

139 More recent advancements have focused on complex reasoning and cross-domain integration. Rea-
140 soning models such as ether0 (Narayanan et al., 2025) and ChemDFM-R (Zhao et al., 2025b), trained
141 via reinforcement learning, exhibit strong performance across diverse chemical tasks and provide
142 transparent, interpretable outputs. In parallel, scientific foundation models like NatureLM (Xia
143 et al., 2025) and Intern-S1 (Bai et al., 2025) have been trained on large-scale data from various
144 scientific fields. These models can handle a diverse range of inputs spanning biology, chemistry,
145 and materials science. Despite these advances, the progress of foundation models in chemistry lags
146 significantly behind that in high-resource domains like mathematics and code, largely due to the
147 relative scarcity of specialized scientific data for training (Bai et al., 2025). To overcome this data
148 bottleneck, we introduce a specialized pipeline designed to synthesize high-quality CoT data on par
149 with expert-level annotations for model training.

151 3 METHOD 152

153 **Overview.** We propose Chem-R, a three-phase framework explicitly designed to endow LLMs
154 with structured and generalizable chemical reasoning capabilities. **Phase 1** (Section 3.1) estab-
155 lishes a chemistry-aware foundation by supervised fine-tuning on large-scale non-reasoning cor-
156 pora, grounding the model in valid molecular and reaction representations. **Phase 2** (Section 3.2)
157 introduces Chemical Reasoning Protocol (CRP) Distillation, which transfers structured and reusable
158 reasoning workflows from a teacher model into a compact student model. **Phase 3** (Section 3.3) em-
159 ploys Multi-task Group Relative Policy Optimization (Multi-task GRPO) to further enhance and
160 balance performance across diverse molecular- and reaction-level tasks. Together, these phases
161 form a principled pipeline that transforms ad-hoc CoT traces into chemically sound, interpretable,
and broadly generalizable reasoning.

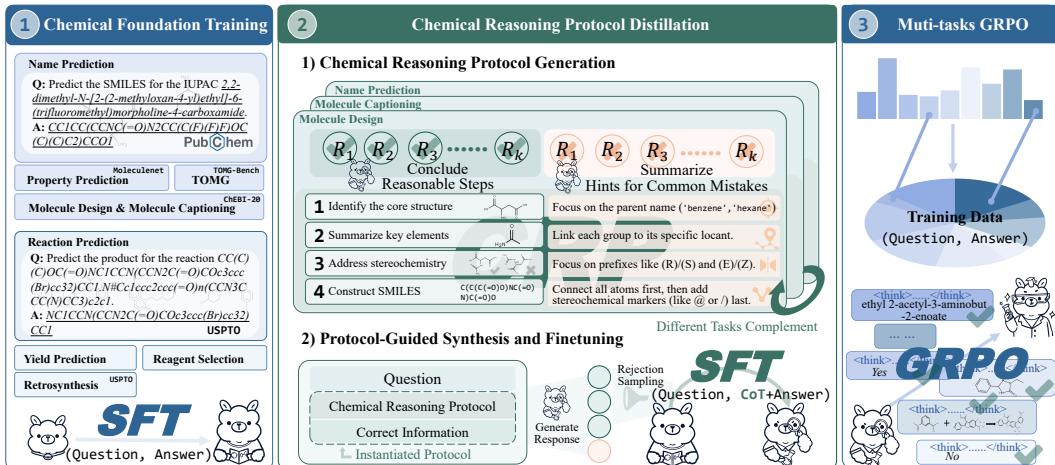


Figure 2: **The overall pipeline of Chem-R.** The model is trained in three phases. **1) Chemical Foundation Training:** Instills basic chemical knowledge using question-answer pairs. **2) Chemical Reasoning Protocol Distillation:** Teaches structured reasoning by fine-tuning on protocol-guided CoT. **3) Multi-task GRPO:** Refines reasoning skills across all tasks using reinforcement learning.

3.1 PHASE 1: CHEMICAL FOUNDATION TRAINING

Establishing a reliable chemical LLM necessitates the integration of domain-specific knowledge in molecular representations and reaction notation. General-purpose corpora (e.g., Wikipedia, textbooks) are inadequate in this regard, as they rarely capture the syntactic rules of SMILES strings or the systematic regularities of IUPAC nomenclature, let alone the canonical mapping between different molecular descriptors (Luo et al., 2022; Taylor et al., 2022; Irwin et al., 2022). To this end, Phase 1 establishes a chemistry-aware foundation by supervised fine-tuning (SFT) on large-scale non-reasoning corpora $\mathcal{D}_{\text{chem}}$, thereby grounding the model in chemically valid input–output behaviors.

Formally, $\mathcal{D}_{\text{chem}}$ is a paired chemistry corpus, $\mathcal{D}_{\text{chem}} = \{(x_i, y_i)\}_{i=1}^N$, where x_i represents a structured chemical input (e.g., a SMILES string, an IUPAC name, or a reaction query), and y_i is the corresponding labels (e.g., a canonical IUPAC name, a valid SMILES string, or the main product of a chemical reaction). $\mathcal{D}_{\text{chem}}$ encompasses both molecular- and reaction-level supervision.

At the **molecule level**, the corpus aligns alternative descriptors of the same compound. This enables the model to master not only the bidirectional translation between SMILES and IUPAC forms, but also the mapping from a molecular structure to its textual description. Such examples teach the model chemically consistent string generation and reduce elementary notational errors.

At the **reaction level**, the corpus encodes prototypical transformations, mapping reactants to their products or reagents and specifying the functional roles of reagents and conditions. Although such instances require only static mapping rather than explicit reasoning, they provide essential priors that prevent chemically implausible outcomes.

In this phase, the model is trained via supervised fine-tuning (SFT) to internalize the syntax and semantics of $\mathcal{D}_{\text{chem}}$, as illustrated in Figure 2 (left). This chemistry-aware initialization substantially reduces representational errors and serves as the basis for structured reasoning in subsequent phases. Examples of data used in the first phase can be found in Appendix B.

3.2 PHASE 2: CHEMICAL REASONING PROTOCOL DISTILLATION

While Phase 1 equips the model with foundational chemical knowledge by training on correct question-answer pairs, it does not yet instill the ability to perform reliable and structured reasoning. A conventional approach to bridge this gap is to distill CoT data from a more powerful teacher model and then train the student model. However, this direct distillation often perpetuates the exact problems we seek to eliminate; as shown in Figure 1, even advanced models can produce chaotic and unreliable reasoning trajectories. To address this challenge and ensure the distilled reasoning is of high quality, we introduce Phase 2 (as shown in middle pannel of Figure 2): CRP Distillation. This phase consists of two stages: (I) generating an expert-like reasoning protocol, and (II) using

216 this protocol to guide the synthesis of high-quality CoT data for finetuning. The overall objective is
 217 to teach Chem-R a systematic and reliable reasoning methodology.
 218

219 **I. Chemical Reasoning Protocol Generation.** It creates the expert’s blueprint for chemical reasoning.
 220 For a given task, we use a simple prompt to have the teacher model generate multiple,
 221 varied responses. From this collection of responses, we then sample k positive examples (reasoning
 222 paths leading to the correct answer, as R_i in Figure 2) and k negative examples (those leading to
 223 incorrect answers). The teacher model systematically analyzes these positive examples to conclude
 224 a generalizable, step-by-step reasoning template. Concurrently, the teacher model also examines
 225 failed reasoning attempts to identify common mistakes, summarizing them as cautionary guidance
 226 attached to each step of the protocol. This process results in a strong thinking guide for each task.
 227 Furthermore, we create a more holistic and robust final reasoning guide by merging the cautionary
 228 guidance from analogous steps across different tasks, enriching the protocol for any given step with
 229 insights from as many relevant contexts as possible.
 230

231 **II. Protocol-Guided Synthesis and Finetuning.** With the expert protocol established, the second
 232 stage focuses on data synthesis and student model training. For each question, we guide the teacher
 233 model by providing it with an *Instantiated Protocol*, a combination of the task’s governing CRP and
 234 the reliable correct information (*i.e.*, functional groups and final answer). This prompts the model to
 235 produce a detailed CoT that strictly adheres to the protocol’s structured steps. To ensure the absolute
 236 quality and logical fidelity of this synthetic data, we implement a Rejected Sampling mechanism.
 237 Specifically, the answer portion of a generated CoT (*e.g.*, tokens included in `<answer>` tag) is
 238 removed, and the model must regenerate the answer based solely on the preceding reasoning. Only
 239 those CoT paths where the regenerated answer matches the original correct answer are retained,
 240 guaranteeing that the reasoning logically and consistently leads to the correct solution. Finally, this
 241 curated dataset of pristine (Question, CoT + Answer) pairs is used to fine-tune Chem-R via
 242 SFT, effectively teaching it to internalize and replicate a reliable and interpretable reasoning process.
 243

3.3 PHASE 3: MULTI-TASK GRPO

244 While Phase 2 equips the model with structured reasoning protocols, ensuring their robust execution
 245 across heterogeneous tasks remains non-trivial. In particular, naive multi-task training tends to
 246 favor easier or high-resource tasks, causing weaker tasks to be underrepresented and resulting in
 247 imbalanced performance. To overcome this issue, we introduce a *Multi-task GRPO* scheme, which
 248 enhances protocol-guided reasoning while explicitly enforcing balance across tasks.
 249

250 Let \mathcal{T} denote the task mixture. For each task $t \in \mathcal{T}$, we estimate its validation accuracy s_t after
 251 Phase 2, and assign a sampling probability p_t that up-weights weaker tasks:
 252

$$p_t = \frac{(1 - s_t)^\alpha}{\sum_{t' \in \mathcal{T}} (1 - s_{t'})^\alpha}, \quad (1)$$

255 where $\alpha \geq 0$ controls the strength of reweighting. This adaptive curriculum ensures that difficult
 256 or underperforming tasks contribute proportionally more updates, thereby mitigating strong-task
 257 dominance and fostering balanced improvement. For each sampled question q , we roll out G re-
 258 sponses $\{o_i\}_{i=1}^G$ using the current policy $\pi_{\theta_{\text{old}}}$. Each token $o_{i,t}$ within a trajectory is optimized under
 259 a clipped-ratio surrogate with KL regularization:

$$261 \quad J_{\text{GRPO}}(\theta) = \mathbb{E}_{q, \{o_i\}_{i=1}^G} \left[\frac{1}{G} \sum_{i=1}^G \sum_{t=1}^{|o_i|} \min \left(\frac{\pi_\theta(o_{i,t}|q)}{\pi_{\theta_{\text{old}}}(o_{i,t}|q)} A_i, \right. \right. \\ 262 \quad \left. \left. \text{clip} \left(\frac{\pi_\theta(o_{i,t}|q)}{\pi_{\theta_{\text{old}}}(o_{i,t}|q)}, 1 - \epsilon, 1 + \epsilon \right) A_i \right) - \beta D_{\text{KL}}(\pi_\theta \| \pi_{\text{ref}}) \right]. \quad (2)$$

266 Here, $o_{i,t}$ is the t -th token of the i -th response o_i , which has length $|o_i|$, ϵ is a hyperparameter that
 267 defines the clipping range, A_i is the normalized group advantage, and $D_{\text{KL}}(\pi_\theta \| \pi_{\text{ref}})$ is a KL diver-
 268 gence regularizer, weighted by β , that penalizes deviation from a reference policy π_{ref} (initialized
 269 from Phase 2). As for our reward design, we do not use any format-based rewards. The accuracy-
 based rewards are task-specific, with detailed calculations provided in Section 4.1.

270 4 EXPERIMENTS
271272 4.1 EXPERIMENTAL SETUP
273

274 **Benchmarks.** We collect four widely used chemical benchmarks, *ChemLLMBench* (Guo et al.,
275 *ChEBI-20* (Edwards et al., 2022), *TOMG-Bench* (Li et al., 2024a), and *USPTO* (Schneider
276 et al., 2016), with their official splits to ensure a fair comparison. Based on these, we organize the
277 evaluation into two families: *molecular tasks* and *reaction tasks*, covering 9 macro-tasks with 25
278 sub-tasks in total. **Molecular tasks** include (1) *name prediction* (IUPAC \leftrightarrow SMILES); (2) *property
279 prediction* on BBBP, HIV, Tox21, ClinTox, and BACE (Wu et al., 2018); (3) *molecule design* from
280 text to SMILES; (4) *molecule captioning* from SMILES to text; (5) *text-based open molecule generation*
281 includes molecule editing (with functional group addition, replacement, or removal), molecule
282 optimization (guided toward target LogP, MR, and QED) and customized molecule generation (by
283 atom count, bond count, and functional-group count). **Reaction tasks** include (6) *yield prediction*
284 for Buchwald–Hartwig and Suzuki reactions; (7) *reagent selection* for reactant, solvent, and ligand
285 in multiple-choice form (8) *reaction prediction* and (9) *retrosynthesis*. More detailed descriptions of
286 the tasks can be found in Appendix B.

287 **Evaluation Metrics.** We adopt task-specific evaluation metrics aligned with prior work. For *name
288 prediction* (1), we report exact match between predicted and reference strings. For *property prediction*
289 (2) and *yield prediction* (6), which are binary classification tasks, we use average Accuracy
290 across datasets. For *molecule design* (3), we measure exact match on the generated SMILES. For
291 *molecule captioning* (4), we compute BLEU-4 to evaluate text generation quality. For *text-based
292 open molecule generation* (5), which covers editing, optimization, and customized generation, we
293 report weighted accuracy over constraints such as functional groups, atom counts, and property tar-
294 gets. For *reagent selection* (7), we evaluate by multiple-choice accuracy. For *reaction prediction*
295 (8) and *retrosynthesis* (9), we use exact match on canonical SMILES, with unordered set matching
296 for multi-product or multi-reactant cases separated by “.”. All SMILES and IUPAC comparisons
297 are performed after canonicalization to ensure consistency. For a more comprehensive analysis,
298 supplementary metrics for these tasks are also reported in the Appendix D.

299 **Baselines.** We group baselines into five families, with the first four reported in the main ta-
300 bles and the fifth provided in Appendix D. The first group consists of *general foundation mod-
301 els*. These include Llama-3.1-8B-Instruct, Llama-3.3-70B (Dubey et al., 2024), and GPT-4o (Hurst
302 et al., 2024). This group establishes the capability of non-chemistry-adapted systems. The sec-
303 ond group is *general reasoning models*. Examples are Gemini-2.5-Pro (Comanici et al., 2025),
304 DeepSeek-R1 (Guo et al., 2025), and QWQ-32B (Yang et al., 2025). These models test whether
305 generic reasoning gains transfer to chemistry. The third group contains *chemical foundation models*.
306 These are ChemLLM-DPO-20B (Zhang et al., 2024b), ChemDFM-v1.0-13B, and ChemDFM-v1.5-
307 8B (Zhao et al., 2025c). This set of models emphasizes chemistry knowledge coverage without
308 explicit multi-step reasoning optimization. The fourth group includes *chemical reasoning models*,
309 such as ether0-24B (Narayanan et al., 2025) and our Chem-R, which target process-level reasoning
310 and interpretability. To account for task-specific nuances, we additionally compare against a fifth
311 group, *task-specialized models*, in the appendix. We evaluate these under each benchmark’s standard
312 protocol with unified normalization and scoring scripts for fairness and reproducibility.

313 **Implementation Details.** We select Llama-3.1-8B-Instruct as our base model, and Llama-3.3-
314 70B-Instruct as our teacher model. The detailed data configurations, hyperparameter settings and
315 specific prompts are provided in Appendix C.

316 4.2 MAIN RESULTS
317

318 As shown in Table 1, our 8B model sets a new state-of-the-art across a diverse range of chemical
319 benchmarks. It surpasses not only general-purpose models like Gemini-2.5-Pro and other chemical
320 foundation models like ether0-24B, but also outperforms task-specific specialist models in Property
321 Prediction, Molecule Design, and TOMG tasks (see Appendix D). While the non-reasoning chemi-
322 cal model ChemDFM-v1.5-8B achieves a higher score in the Molecule Design task with direct out-
323 puts, Chem-R provides interpretable, step-by-step reasoning chains, offering critical explainability

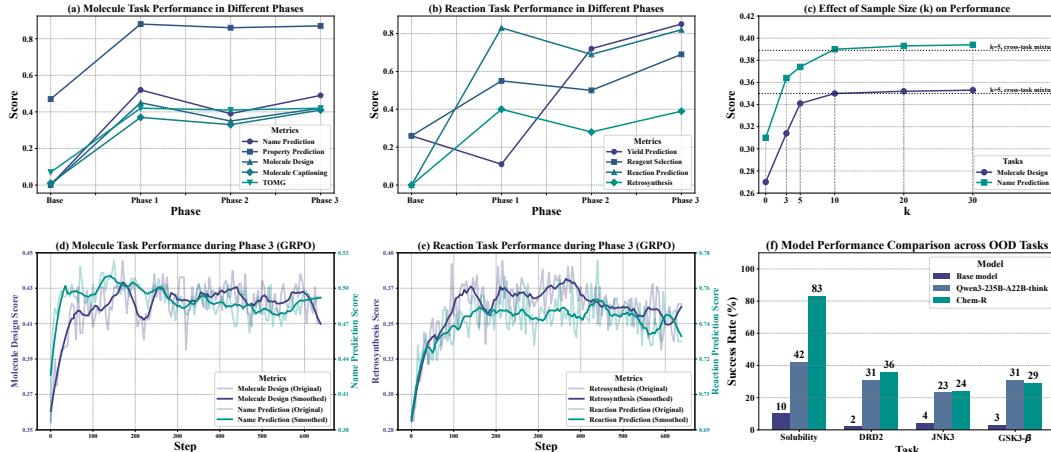


Figure 3: **Comprehensive evaluation of Chem-R.** (a) Molecule task performance in different phases. (b) Reaction task performance in different phases. (c) Effect of sample size (k) on performance. (d) Molecule task performance during phase 3. (e) Reaction task performance during phase 3. (f) Model performance comparison across OOD tasks in ChemCoTBench (Li et al., 2025a).

for scientific discovery. The model’s most significant advances are in reaction-related tasks, where its performance represents a paradigm shift. Chem-R achieves a score of 0.85 in Yield Prediction (*more than doubling* the next-best score) and 0.39 in Retrosynthesis, *a nearly threefold* improvement over the strongest baseline of 0.15. These results validate that our methodology enables superior chemical reasoning within an efficient 8B parameter model. Additionally, the model’s performance at various training phases is shown in Figures 3(a) and (b). For the performance curve during the multi-task GRPO phase, please refer to Figures 3(d) and (e).

Table 1: Performance of different models on chemistry-related tasks. The score for each major task is the average of its subtasks. Column headers use short names: **Name** = Name Prediction (evaluated by Exact Match), **Prop.** = Property Prediction (Accuracy), **Design** = Molecule Design (Accuracy), **Capt.** = Molecule Captioning (BLEU-4), **TOMG** = Tasks in TOMG-Bench (Weighted Accuracy), **Yield** = Yield Prediction (Accuracy), **Reag.** = Reagents Selection (Accuracy), **React.** = Reaction Prediction (Accuracy), **Retro** = Retrosynthesis (Exact Match). For each column: the **best** and **second-best** models are highlighted.

Model	Molecule Tasks					Reaction Tasks				
	Name	Prop.	Design	Capt.	TOMG	Yield	Reag.	React.	Retro	
<i>General Foundation Models</i>										
GPT-4o	0.01	0.68	0.07	0.01	0.32	0.20	0.26	0.04	0.00	
Llama-3.1-8B-Instruct	0.00	0.47	0.00	0.01	0.07	0.26	0.26	0.00	0.00	
Llama-3.3-70B	0.01	0.64	0.03	0.02	0.30	0.22	0.38	0.03	0.00	
<i>General Reasoning Models</i>										
Gemini-2.5-Pro	0.17	0.56	0.29	0.04	–	0.35	0.27	0.04	0.15	
DeepSeek-R1	0.05	0.63	0.22	0.04	–	0.33	0.13	0.34	0.13	
QWQ-32B	0.01	0.71	0.03	0.04	0.30	0.29	0.39	0.01	0.00	
<i>Chemical Foundation Models</i>										
ChemLLM-DPO-20B	0.00	0.49	0.00	0.03	0.10	0.21	0.08	0.02	0.00	
ChemDFM-v1.5-8B	0.14	0.74	0.53	0.10	0.12	0.37	0.35	0.50	0.07	
ChemDFM-v1.0-13B	0.16	0.78	0.42	0.27	0.27	0.32	0.37	0.31	0.04	
<i>Chemical Reasoning Models</i>										
ether0-24B	0.15	0.64	0.30	0.03	0.03	0.03	0.21	0.65	0.04	
Chem-R-8B (Ours)	0.49	0.87	0.42	0.41	0.42	0.85	0.69	0.82	0.39	

4.3 ABLATION STUDY

Phase-wise Contributions. To understand the unique contribution of each training phase, we systematically removed individual phases while holding all other variables constant. The results, shown in Table 2 (A), confirm that all three phases are essential and work synergistically. First, *Phase 1*

(*Foundation Training*) establishes the model’s fundamental understanding of chemistry. Removing this phase severely degrades performance, as seen by the Name Prediction score dropping from 0.49 to 0.14. While critical, this phase does not by itself enable the chain-of-thought reasoning necessary for explainability. The ability to reason emerges in *Phase 2 (CRP Distillation)*, which introduces the core reasoning framework. Without it, the model fails at performing complex reasoning; for example, performance on Reaction Prediction collapses to zero when both Phase 1 and 2 are removed. Lastly, *Phase 3 (Multi-task GRPO)* acts as a crucial refinement stage. Building on the skills from the previous phases, it delivers consistent improvements, boosting the Reagent Selection score from 0.50 to 0.69.

Components of Instantiated Protocol in Phase 2. We analyze the two core components of our CRP Distillation in Phase 2: the task’s governing CRP and *Correct Information*. As detailed in Table 2 (B) and cases in Appendix E, the *CRP* is crucial for improving accuracy by enforcing a logical structure on the reasoning. Its absence results in a consistent performance decline across tasks; for example, the Retrosynthesis score decreases from 0.28 to 0.20 (compared with Chem-R w/o Phase 3). Furthermore, incorporating the ground-truth *Correct Information* is essential for generating a high-quality, large-scale dataset of reasoning chains. Removing this component leads to a severe degradation in performance, with the Reaction Prediction score dropping from 0.69 to only 0.13. Therefore, the CRP provides the indispensable reasoning architecture, while the Correct Information ensures that architecture is used to teach truth, not sophisticated error.

Single-task vs. Multi-task Training in Phase 2. As shown in Table 2 (C), we compare specialized *Single-task* models against a unified *Multi-task* model in Phase 2. While *Single-task* models (*i.e.*, 9 models in total) achieve high scores on their respective tasks, such as 0.75 in Reaction Prediction, confirming the quality of our distilled CoT data, the *Multi-task* model demonstrates clear positive transfer. It outperforms *Single-task* models on related tasks like Reagent Selection (0.50 vs. 0.46) and Retrosynthesis (0.28 vs. 0.26).

Effect of Sample Size k in Phase 2. We investigate the effect of the sample size, k , used to generate the CRP in Phase 2. As shown in Figure 3 (c), performance on both Name Prediction and Molecule Design improves rapidly as k increases, but the gains begin to diminish significantly after k reaches approximately 10. This indicates that simply collecting more samples for a single task yields limited returns. Crucially, we find that a small number of samples (*e.g.*, $k=5$) is sufficient when we enhance the protocol with our cross-task mixture strategy. This efficient approach allows us to achieve a high level of performance, as indicated by the dotted lines, without the need for extensive sampling. This is a critical advantage for complex tasks where successful reasoning paths are often too scarce to collect in large numbers. [More detailed analysis about CRP is shown in Appendix C.5.](#)

Uniform vs. Weighted Sampling in Phase 3. A comparison between *Uniform* and our *Weighted* sampling in Phase 3 demonstrates the effectiveness of the latter. The results, presented in Table 2 (D), indicate that allocating more training focus to tasks the model finds more difficult yields significant performance gains within the same number of training steps. Notably, the score for the challenging Retrosynthesis task improved from 0.33 to 0.39, and the Reagent Selection score rose from 0.63 to 0.69.

Generalization to Out-of-Distribution Tasks. To assess our model’s generalization capabilities, we evaluate it on four out-of-distribution (OOD) molecule optimization tasks (Solubility, DRD2, JNK3, and GSK3) from ChemCoTBench (Li et al., 2025a). While standard benchmarks often focus on familiar properties, we tested performance on three more challenging protein activity-related properties. We intentionally select these tasks because they are not part of our training data and differ significantly from our training objectives, providing a robust test of generalization. As shown in Figure 3 (f), the baseline model (Llama-3.1-8B-Instruct) performs poorly, confirming its inability to generalize. In contrast, Chem-R achieves a massive leap in performance across all four tasks, for example, improving the success rate on Solubility from 10% to 83%.

Table 2: **Ablation study** across 9 chemistry tasks (25 sub-tasks). Columns follow the same short-hand as the main table. **Gray** cells denote Chem-R’s performance.

Ablation Variant	Molecule Tasks					Reaction Tasks			
	Name	Prop.	Design	Capt.	TOMG	Yield	Reag.	React.	Retro
(A) Phase-wise Contributions									
Chem-R w/o Phase 1	0.14	0.87	0.20	0.34	0.38	0.87	0.58	0.54	0.34
Chem-R w/o Phase 2	0.53	0.88	0.43	0.41	0.41	0.80	0.59	0.82	0.39
Chem-R w/o Phase 3	0.39	0.86	0.35	0.33	0.41	0.85	0.50	0.69	0.28
Chem-R w/o Phase 1&2	0.00	0.67	0.03	0.17	0.30	0.80	0.51	0.00	0.00
Chem-R w/o Phase 1&3	0.10	0.85	0.27	0.27	0.37	0.84	0.55	0.48	0.19
Chem-R w/o Phase 2&3	0.52	0.88	0.43	0.40	0.42	0.11	0.55	0.83	0.40
Chem-R w/o Phase 1&2&3	0.00	0.47	0.00	0.01	0.07	0.26	0.26	0.00	0.00
(B) Components of Instantiated Protocol in Phase 2									
Phase 2 w/o CRP	0.31	0.84	0.27	0.33	0.42	0.69	0.43	0.67	0.20
Phase 2 w/o Correct Information	0.04	0.79	0.07	0.02	0.27	0.67	0.38	0.13	0.05
(C) Single-task vs. Multi-task Training in Phase 2									
Single-task	0.41	0.87	0.40	0.32	0.41	0.84	0.46	0.75	0.26
Multi-task	0.39	0.86	0.35	0.33	0.41	0.85	0.50	0.69	0.28
(D) Uniform vs. Weighted Sampling in Phase 3									
Uniform	0.43	0.87	0.38	0.39	0.41	0.84	0.63	0.81	0.33
Chem-R-8B (Ours)	0.49	0.87	0.42	0.41	0.42	0.85	0.69	0.82	0.39

Human Expert Evaluation. To qualitatively assess CoTs, we conduct a human expert evaluation. We had chemistry PhDs rate 50 randomly sampled responses from Chem-R and three strong baselines (Gemini-2.5-Pro, DeepSeek-R1, ether0) on a 1-to-5 scale across six dimensions, with a detailed rubric available in the Appendix D. The averaged scores are visualized in a radar chart for comparison, as shown in Figure 4. Chem-R receives the highest scores across all six metrics, from Chemical Soundness and Logical Coherence to Expert-level Insight, validating the effectiveness of our structured reasoning protocol. This comprehensive feedback confirms that Chem-R not only achieves high accuracy but also generates reasoning chains that are significantly more reliable, interpretable, and aligned with expert thinking. This all-around superiority underscores its potential as a more trustworthy and dependable tool for chemical inquiry.

5 CONCLUSION

Chem-R establishes a new state-of-the-art in chemical reasoning by uniquely emulating the deliberative thought processes of expert chemists. Our novel three-phase training framework systematically builds foundational knowledge, distills structured reasoning protocols, and optimizes for balanced performance across a wide array of tasks. This approach enables Chem-R to significantly outperform leading models, including ChemDFM-v1.5-8B and Gemini-2.5-Pro, achieving unprecedented gains, particularly in complex reaction prediction and retrosynthesis tasks. Beyond superior accuracy, Chem-R’s primary advantage lies in its ability to generate interpretable, logically coherent, and chemically sound reasoning chains, a quality validated by human expert evaluation. By producing reliable and explainable outputs, Chem-R provides a powerful and trustworthy foundation for accelerating the next generation of AI-driven chemical discovery.

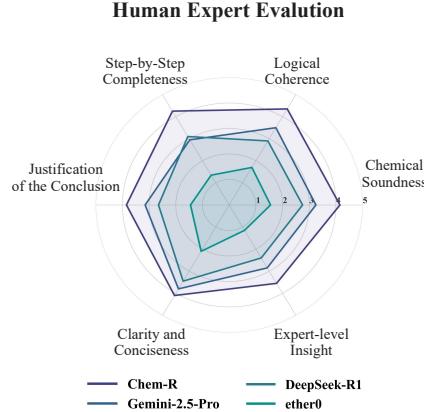


Figure 4: Radar chart of the human expert evaluation.

486 REFERENCES
487

488 Iván Arcuschin, Jett Janiak, Robert Krzyzanowski, Senthooran Rajamanoharan, Neel Nanda, and
489 Arthur Conmy. Chain-of-thought reasoning in the wild is not always faithful. *arXiv preprint*
490 *arXiv:2503.08679*, 2025.

491 Lei Bai, Zhongrui Cai, Maosong Cao, Weihan Cao, Chiyu Chen, Haojiong Chen, Kai Chen,
492 Pengcheng Chen, Ying Chen, Yongkang Chen, et al. Intern-s1: A scientific multimodal foun-
493 dation model. *arXiv preprint arXiv:2508.15763*, 2025.

494 Bowen Baker, Joost Huizinga, Leo Gao, Zehao Dou, Melody Y Guan, Aleksander Madry, Wojciech
495 Zaremba, Jakub Pachocki, and David Farhi. Monitoring reasoning models for misbehavior and
496 the risks of promoting obfuscation. *arXiv preprint arXiv:2503.11926*, 2025.

497 Andres M Bran, Theo A Neukomm, Daniel P Armstrong, Zlatko Jončev, and Philippe Schwaller.
498 Chemical reasoning in llms unlocks steerable synthesis planning and reaction mechanism eluci-
499 dation. *arXiv preprint arXiv:2503.08537*, 2025.

500 He Cao, Zijing Liu, Xingyu Lu, Yuan Yao, and Yu Li. Instructmol: Multi-modal integration
501 for building a versatile and reliable molecular assistant in drug discovery. *arXiv preprint*
502 *arXiv:2311.16208*, 2023.

503 Yanda Chen, Joe Benton, Ansh Radhakrishnan, Jonathan Uesato, Carson Denison, John Schulman,
504 Arushi Soman, Peter Hase, Misha Wagner, Fabien Roger, et al. Reasoning models don't always
505 say what they think. *arXiv preprint arXiv:2505.05410*, 2025.

506 Connor W Coley, Regina Barzilay, Tommi S Jaakkola, William H Green, and Klavs F Jensen. Pre-
507 diction of organic reaction outcomes using machine learning. *ACS central science*, 3(5):434–443,
508 2017.

509 Gheorghe Comanici, Eric Bieber, Mike Schaeckermann, Ice Pasupat, Noveen Sachdeva, Inderjit
510 Dhillon, Marcel Blistein, Ori Ram, Dan Zhang, Evan Rosen, et al. Gemini 2.5: Pushing the
511 frontier with advanced reasoning, multimodality, long context, and next generation agentic capa-
512 bilities. *arXiv preprint arXiv:2507.06261*, 2025.

513 Yilun Du, Shuang Li, Antonio Torralba, Joshua B Tenenbaum, and Igor Mordatch. Improving
514 factuality and reasoning in language models through multiagent debate. *Forty-first International*
515 *Conference on Machine Learning*, 2023.

516 Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha
517 Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. The llama 3 herd of models.
518 *arXiv e-prints*, pp. arXiv–2407, 2024.

519 Carl Edwards, Tuan Lai, Kevin Ros, Garrett Honke, Kyunghyun Cho, and Heng Ji. Translation
520 between molecules and natural language. *arXiv preprint arXiv:2204.11817*, 2022.

521 Yin Fang, Xiaozhuan Liang, Ningyu Zhang, Kangwei Liu, Rui Huang, Zhuo Chen, Xiaohui Fan, and
522 Huajun Chen. Mol-instructions: A large-scale biomolecular instruction dataset for large language
523 models. *arXiv preprint arXiv:2306.08018*, 2023.

524 Kanishk Gandhi, Ayush Chakravarthy, Anikait Singh, Nathan Lile, and Noah D Goodman. Cogni-
525 tive behaviors that enable self-improving reasoners, or, four habits of highly effective stars. *arXiv*
526 *preprint arXiv:2503.01307*, 2025.

527 Samuel Genheden, Amol Thakkar, Veronika Chadimová, Jean-Louis Reymond, Ola Engkvist, and
528 Esben Bjerrum. Aizynthfinder: a fast, robust and flexible open-source software for retrosynthetic
529 planning. *Journal of cheminformatics*, 12(1):70, 2020.

530 Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu,
531 Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms
532 via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025.

533 Taicheng Guo, Bozhao Nan, Zhenwen Liang, Zhichun Guo, Nitesh Chawla, Olaf Wiest, Xiangliang
534 Zhang, et al. What can large language models do in chemistry? a comprehensive benchmark on
535 eight tasks. *Advances in Neural Information Processing Systems*, 36:59662–59688, 2023.

540 Kan Hatakeyama-Sato, Naoki Yamane, Yasuhiko Igarashi, Yuta Nabee, and Teruaki Hayakawa.
 541 Prompt engineering of gpt-4 for chemical research: what can/cannot be done? *Science and*
 542 *Technology of Advanced Materials: Methods*, 3(1):2260300, 2023.

543

544 Esther Heid, Kevin P Greenman, Yunsie Chung, Shih-Cheng Li, David E Graff, Florence H Ver-
 545 meire, Haoyang Wu, William H Green, and Charles J McGill. Chemprop: a machine learning
 546 package for chemical property prediction. *Journal of Chemical Information and Modeling*, 64(1):
 547 9–17, 2023.

548

549 Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Os-
 550 trow, Akila Welihinda, Alan Hayes, Alec Radford, et al. Gpt-4o system card. *arXiv preprint*
 551 *arXiv:2410.21276*, 2024.

552

553 Ross Irwin, Spyridon Dimitriadis, Jiazen He, and Esben Jannik Bjerrum. Chemformer: a pre-
 554 trained transformer for computational chemistry. *Machine Learning: Science and Technology*, 3
 555 (1):015022, 2022.

556

557 Lei Jiang, Shuzhou Sun, Binqing Qi, Yuchen Fu, Xiaohua Xu, Yuqiang Li, Dongzhan Zhou, and
 558 Tianfan Fu. Chem3dilm: 3d multimodal large language models for chemistry. *arXiv preprint*
 559 *arXiv:2508.10696*, 2025.

560

561 Jan Hendrik Kirchner, Yining Chen, Harri Edwards, Jan Leike, Nat McAleese, and Yuri Burda.
 562 Prover-verifier games improve legibility of llm outputs. *arXiv preprint arXiv:2407.13692*, 2024.

563

564 Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. Large
 565 language models are zero-shot reasoners. *Advances in neural information processing systems*,
 566 35:22199–22213, 2022.

567

568 HJ Kuhn, SE Braslavsky, and R Schmidt. Chemical actinometry (iupac technical report). *Pure and*
 569 *Applied Chemistry*, 76(12):2105–2146, 2004.

570

571 Youngchun Kwon, Dongseon Lee, Youn-Suk Choi, and Seokho Kang. Uncertainty-aware prediction
 572 of chemical reaction yields with graph neural networks. *Journal of Cheminformatics*, 14(1):2,
 573 2022.

574

575 Hao Li, He Cao, Bin Feng, Yanjun Shao, Xiangru Tang, Zhiyuan Yan, Li Yuan, Yonghong Tian,
 576 and Yu Li. Beyond chemical qa: Evaluating llm’s chemical reasoning with modular chemical
 577 operations. *arXiv preprint arXiv:2505.21318*, 2025a.

578

579 Jiatong Li, Junxian Li, Yunqing Liu, Dongzhan Zhou, and Qing Li. Tomg-bench: Evaluating llms
 580 on text-based open molecule generation. *arXiv preprint arXiv:2412.14642*, 2024a.

581

582 Jiatong Li, Yunqing Liu, Wenqi Fan, Xiao-Yong Wei, Hui Liu, Jiliang Tang, and Qing Li. Empow-
 583 ering molecule discovery for molecule-caption translation with large language models: A chatgpt
 584 perspective. *IEEE transactions on knowledge and data engineering*, 2024b.

585

586 Jiatong Li, Weida Wang, Qinggang Zhang, Junxian Li, Di Zhang, Changmeng Zheng, Shufei Zhang,
 587 Xiaoyong Wei, and Qing Li. Mol-r1: Towards explicit long-cot reasoning in molecule discovery.
 588 *arXiv preprint arXiv:2508.08401*, 2025b.

589

590 Junxian Li, Di Zhang, Xunzhi Wang, Zeying Hao, Jingdi Lei, Qian Tan, Cai Zhou, Wei Liu, Yaotian
 591 Yang, Xinrui Xiong, et al. Chemvilm: Exploring the power of multimodal large language models
 592 in chemistry area. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39,
 593 pp. 415–423, 2025c.

594

595 Hunter Lightman, Vineet Kosaraju, Yuri Burda, Harrison Edwards, Bowen Baker, Teddy Lee, Jan
 596 Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. Let’s verify step by step. *The Twelfth*
 597 *International Conference on Learning Representations*, 2024.

598

599 Tianqi Liu, Yao Zhao, Rishabh Joshi, Misha Khalman, Mohammad Saleh, Peter J Liu, and
 600 Jialu Liu. Statistical rejection sampling improves preference optimization. *arXiv preprint*
 601 *arXiv:2309.06657*, 2023a.

594 Xinyi Liu, Lipeng Ma, Yixuan Li, Weidong Yang, Qingyuan Zhou, Jiayi Song, Shuhao Li, and
 595 Ben Fei. Chemau: Harness the reasoning of llms in chemical research with adaptive uncertainty
 596 estimation. *arXiv preprint arXiv:2506.01116*, 2025.

597 Zequn Liu, Wei Zhang, Yingce Xia, Lijun Wu, Shufang Xie, Tao Qin, Ming Zhang, and Tie-
 598 Yan Liu. Molxpt: Wrapping molecules with text for generative pre-training. *arXiv preprint*
 600 *arXiv:2305.10688*, 2023b.

601 Zhiyuan Liu, Sihang Li, Yanchen Luo, Hao Fei, Yixin Cao, Kenji Kawaguchi, Xiang Wang, and
 602 Tat-Seng Chua. Molca: Molecular graph-language modeling with cross-modal projector and uni-
 603 modal adapter. *arXiv preprint arXiv:2310.12798*, 2023c.

604 Renqian Luo, Lai Sun, Yingce Xia, Tao Qin, Sheng Zhang, Hoifung Poon, and Tie-Yan Liu. Biogpt:
 605 generative pre-trained transformer for biomedical text generation and mining. *Briefings in bioin-
 606 formatics*, 23(6):bbac409, 2022.

607 Pingchuan Ma, Tsun-Hsuan Wang, Minghao Guo, Zhiqing Sun, Joshua B Tenenbaum, Daniela Rus,
 608 Chuang Gan, and Wojciech Matusik. Llm and simulation as bilevel optimizers: A new paradigm
 609 to advance physical scientific discovery. *arXiv preprint arXiv:2405.09783*, 2024.

610 Siddharth M. Narayanan, James D. Braza, Ryan-Rhys Griffiths, Albert Bou, Geemi P. Wellawatte,
 611 Mayk Caldas Ramos, Ludovico Mitchener, Samuel G. Rodrigues, and Andrew D. White. Training
 612 a scientific reasoning model for chemistry. *arXiv preprint arXiv:2506.17238*, 2025.

613 Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong
 614 Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. Training language models to follow
 615 instructions with human feedback. *Advances in neural information processing systems*, 35, 2022.

616 Siru Ouyang, Zhuosheng Zhang, Bing Yan, Xuan Liu, Yéjin Choi, Jiawei Han, and Lianhui Qin.
 617 Structured chemistry reasoning with large language models. *arXiv preprint arXiv:2311.09656*,
 618 2023.

619 Richard Yuanzhe Pang, Weizhe Yuan, He He, Kyunghyun Cho, Sainbayar Sukhbaatar, and Jason
 620 Weston. Iterative reasoning preference optimization. *Advances in Neural Information Processing
 621 Systems*, 2024.

622 Qizhi Pei, Wei Zhang, Jinhua Zhu, Kehan Wu, Kaiyuan Gao, Lijun Wu, Yingce Xia, and Rui Yan.
 623 Biot5: Enriching cross-modal integration in biology with chemical knowledge and natural lan-
 624 guage associations. *arXiv preprint arXiv:2310.07276*, 2023.

625 Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea
 626 Finn. Direct preference optimization: Your language model is secretly a reward model. *Advances
 627 in neural information processing systems*, 2023.

628 Kohulan Rajan, Achim Zielesny, and Christoph Steinbeck. Stout: Smiles to iupac names using
 629 neural machine translation. *Journal of Cheminformatics*, 13(1):34, 2021.

630 Malik Sallam, Khaled Al-Salahat, Huda Eid, Jan Egger, and Behrus Puladi. Human versus artificial
 631 intelligence: Chatgpt-4 outperforming bing, bard, chatgpt-3.5 and humans in clinical chemistry
 632 multiple-choice questions. *Advances in Medical Education and Practice*, pp. 857–871, 2024.

633 Nadine Schneider, Nikolaus Stiefl, and Gregory A Landrum. What's what: The (nearly) definitive
 634 guide to reaction role assignment. *Journal of chemical information and modeling*, 56(12):2336–
 635 2346, 2016.

636 Parshin Shojaee, Kazem Meidani, Shashank Gupta, Amir Barati Farimani, and Chandan K Reddy.
 637 Llm-sr: Scientific equation discovery via programming with large language models. *arXiv
 638 preprint arXiv:2404.18400*, 2024.

639 Qian Tan, Dongzhan Zhou, Peng Xia, Wanhai Liu, Wanli Ouyang, Lei Bai, Yuqiang Li, and Tianfan
 640 Fu. Chemmlm: Chemical multimodal large language model. *arXiv preprint arXiv:2505.16326*,
 641 2025.

648 Ross Taylor, Marcin Kardas, Guillem Cucurull, Thomas Scialom, Anthony Hartshorn, Elvis Saravia,
 649 Andrew Poulton, Viktor Kerkez, and Robert Stojnic. Galactica: A large language model for
 650 science. *arXiv preprint arXiv:2211.09085*, 2022.

651 Yuxuan Tong, Xiwen Zhang, Rui Wang, Ruidong Wu, and Junxian He. Dart-math: Difficulty-aware
 652 rejection tuning for mathematical problem-solving. *Advances in Neural Information Processing
 653 Systems*, 2024.

654 Peiyi Wang, Lei Li, Zhihong Shao, RX Xu, Damai Dai, Yifei Li, Deli Chen, Yu Wu, and Zhifang Sui.
 655 Math-shepherd: Verify and reinforce llms step-by-step without human annotations. *Proceedings
 656 of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long
 657 Papers)*, 2024.

658 Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny
 659 Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in
 660 neural information processing systems*, 35:24824–24837, 2022.

661 David Weininger. Smiles, a chemical language and information system. 1. introduction to method-
 662 ology and encoding rules. *Journal of chemical information and computer sciences*, 28(1):31–36,
 663 1988.

664 Zhenqin Wu, Bharath Ramsundar, Evan N Feinberg, Joseph Gomes, Caleb Geniesse, Aneesh S
 665 Pappu, Karl Leswing, and Vijay Pande. Moleculenet: a benchmark for molecular machine learn-
 666 ing. *Chemical science*, 9(2):513–530, 2018.

667 Yingce Xia, Peiran Jin, Shufang Xie, Liang He, Chuan Cao, Renqian Luo, Guoqing Liu, Yue Wang,
 668 Zequn Liu, Yuan-Jyue Chen, et al. Nature language model: Deciphering the language of nature
 669 for scientific discovery. *arXiv preprint arXiv:2502.07527*, 2025.

670 An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu,
 671 Chang Gao, Chengan Huang, Chenxu Lv, et al. Qwen3 technical report. *arXiv preprint
 672 arXiv:2505.09388*, 2025.

673 Di Zhang, Xiaoshui Huang, Dongzhan Zhou, Yuqiang Li, and Wanli Ouyang. Accessing gpt-4 level
 674 mathematical olympiad solutions via monte carlo tree self-refine with llama-3 8b. *arXiv preprint
 675 arXiv:2406.07394*, 2024a.

676 Di Zhang, Wei Liu, Qian Tan, Jingdan Chen, Hang Yan, Yuliang Yan, Jiatong Li, Weiran Huang,
 677 Xiangyu Yue, Wanli Ouyang, et al. Chemllm: A chemical large language model. *arXiv preprint
 678 arXiv:2402.06852*, 2024b.

679 Di Zhang, Weida Wang, Junxian Li, Xunzhi Wang, Jiatong Li, Jianbo Wu, Jingdi Lei, Haonan He,
 680 Peng Ye, Shufei Zhang, et al. Control-r: Towards controllable test-time scaling. *arXiv preprint
 681 arXiv:2506.00189*, 2025a.

682 Lunjun Zhang, Arian Hosseini, Hritik Bansal, Mehran Kazemi, Aviral Kumar, and Rishabh Agarwal.
 683 Generative verifiers: Reward modeling as next-token prediction. *The Thirteenth International
 684 Conference on Learning Representations*, 2025b.

685 Yu Zhang, Yang Han, Shuai Chen, Ruijie Yu, Xin Zhao, Xianbin Liu, Kaipeng Zeng, Mengdi Yu,
 686 Jidong Tian, Feng Zhu, et al. Large language models to accelerate organic chemistry synthesis.
 687 *Nature Machine Intelligence*, pp. 1–13, 2025c.

688 Guojiang Zhao, Sihang Li, Zixiang Lu, Zheng Cheng, Haitao Lin, Lirong Wu, Hanchen Xia, Hengx-
 689 ing Cai, Wentao Guo, Hongshuai Wang, et al. Molreasoner: Toward effective and interpretable
 690 reasoning for molecular llms. *arXiv preprint arXiv:2508.02066*, 2025a.

691 Zihan Zhao, Bo Chen, Jingpiao Li, Lu Chen, Liyang Wen, Pengyu Wang, Zichen Zhu, Danyang
 692 Zhang, Yansi Li, Zhongyang Dai, Xin Chen, and Kai Yu. Chemdfm-x: towards large multimodal
 693 model for chemistry. *Science China Information Sciences*, 2024.

694 Zihan Zhao, Bo Chen, Ziping Wan, Lu Chen, Xuanze Lin, Shiyang Yu, Situo Zhang, Da Ma, Zichen
 695 Zhu, Danyang Zhang, et al. Chemdfm-r: An chemical reasoner llm enhanced with atomized
 696 chemical knowledge. *arXiv preprint arXiv:2507.21990*, 2025b.

702 Zihan Zhao, Da Ma, Lu Chen, Liangtai Sun, Zihao Li, Yi Xia, Bo Chen, Hongshen Xu, Zichen Zhu,
703 Su Zhu, et al. Developing chemdfm as a large language foundation model for chemistry. *Cell*
704 *Reports Physical Science*, 6(4), 2025c.

705

706 Zhiqiang Zhong, Kuangyu Zhou, and Davide Mottin. Benchmarking large language models for
707 molecule prediction tasks. *arXiv preprint arXiv:2403.05075*, 2024.

708 Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng
709 Zhang, and Guolin Ke. Uni-mol: A universal 3d molecular representation learning framework.
710 2023.

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756	Appendix	
757		
758		
759		
760		
761		
762		
763	A Overview of the Appendix	16
764		
765	B Task Description	16
766	B.1 Molecular Tasks	16
767	B.2 Reaction Tasks	18
768		
769		
770	C Implementation Details	20
771	C.1 Hyper-parameters	20
772	C.2 Data Statistics	21
773	C.3 CRP Extraction Prompt	23
774	C.4 Distillation Prompts	24
775	C.5 Analysis of CRPs	38
776	C.6 Human Evaluation	39
777		
778		
779		
780	D Experiment Result	40
781	D.1 Name Prediction	40
782	D.2 Property Prediction	41
783	D.3 Molecule Design	41
784	D.4 Molecule Captioning	42
785	D.5 Text-based Open Molecule Generation	43
786	D.6 Yield Prediction	45
787	D.7 Reagent Selection	45
788	D.8 Reaction Prediction	45
789	D.9 Retrosynthesis	46
790	D.10 Out of Domain	46
791	D.11 Effectiveness of CRP as a Prompting Strategy	47
792	D.12 Impact of Teacher Model Quality	47
793	D.13 Computational Efficiency Analysis	48
794		
795		
796		
797		
798		
799		
800		
801	E More Cases	49
802		
803		
804	F Use of LLMs	54
805		
806		
807		
808		
809		

810 A OVERVIEW OF THE APPENDIX
811812 **Section B (Task Description)** outlines the various tasks used to evaluate the model, which are
813 divided into two main categories. "Molecular Tasks" focus on individual molecules, including chal-
814 lenges like predicting chemical names and properties, designing new molecules, and generating
815 descriptions. "Reaction Tasks" involve chemical transformations, such as predicting reaction yields,
816 selecting reagents, and determining the products or starting materials of a reaction.817 **Section C (Implementation Details)** details the technical methodology of the study. It covers the
818 specific hyper-parameters for model training, the data statistics showing how datasets were divided,
819 [CRP extraction prompt](#), and the structured distillation prompts used to guide the model's reasoning
820 process. [It also presents an analysis of the Chemical Reasoning Protocols](#). Additionally, it explains
821 the criteria for the human evaluation performed by chemistry experts to assess the quality of the
822 model's logic.823 **Section D (Experiment Result)** presents the quantitative outcomes of the experiments. It contains
824 a series of tables with performance metrics that demonstrate the model's effectiveness on each of the
825 tasks described in Section B and on OOD tasks, providing a detailed, data-driven assessment of its
826 capabilities in chemical reasoning. [This section further includes supplementary analyses on the ef-
827 fectiveness of CRP as a prompting strategy, the impact of teacher model quality, and computational
828 efficiency.](#)829 **Section E (More Case)** provides additional, specific examples of the model's outputs.830 **Section F (Use of LLM)** clarifies the specific role and application of LLMs within this research.831 B TASK DESCRIPTION
832833 In this section, we describe the tasks used in our experiments, which span both molecu-
834 lar tasks and reaction tasks. We selected these tasks from four widely recognized chemistry
835 benchmarks—ChEBI-20, ChemLLMBench, USPTO, and TOMG-Bench—based on their practical
836 relevance and the diversity they offer in evaluating molecular and reaction-level capabilities. These
837 tasks cover a range of applications, from molecule generation and property prediction to reaction
838 prediction and retrosynthesis, providing a comprehensive evaluation of the model's performance in
839 chemical reasoning.840
841 B.1 MOLECULAR TASKS842 **(1) Name Prediction.** In this task, the goal is to convert between *IUPAC names* (International
843 Union of Pure and Applied Chemistry) and *SMILES* (Simplified Molecular Input Line Entry Sys-
844 tem) *strings*, which are two widely used methods for representing chemical structures. IUPAC
845 names provide a formal, systematic way to describe molecules based on their structure, while
846 SMILES is a textual representation that encodes molecular structure through a series of symbols
847 and characters. Converting between these formats requires reasoning about the chemical structure
848 itself—understanding the arrangement of atoms, bonds, functional groups, and molecular topology.
849 The model must interpret the chemical details embedded in IUPAC names or SMILES strings and
850 produce the corresponding representation, which involves complex reasoning about chemical con-
851 conventions and rules.852
853 Name Prediction (SMILES→IUPAC)854 **Question:** Provide the IUPAC name for the following molecule. SMILES:

855 CCOC(=O)C(C(C)=O)=C(C)N.

856 **Answer:** ethyl 2-acetyl-3-aminobut-2-enoate857
858 **(2) Property Prediction.** Property prediction involves classifying molecules into categories based
859 on their biological activity or toxicity (i.e., *BBBP*, *HIV*, *Tox21*, *ClinTox*, *BACE*). For this, the model
860 needs to reason about the underlying structure-property relationships. A molecule's structure in-
861 fluences its biological properties through factors like functional group interactions, molecular size,
862 and polarity. The model must learn these complex relationships from data and reason about which
863 aspects of the molecule's structure contribute to its biological effects. This makes property predic-

864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
tion a key task for reasoning models, as they must generalize these chemical insights across diverse molecular structures and predict their effects.

Property Prediction (BACE)

Question: Predict whether the following molecule can inhibit BACE1 (Yes/No). SMILES:

Clc1cc(cnc1)-c1cc2c(CC(CC23N=C(OC3)N)C)C)c1.

Answer: No

(3) Molecule Design. In this task, a model is given *a textual description of a molecule* and must generate its *corresponding SMILES representation*. The challenge here is that the model needs to map linguistic descriptions (which are often abstract) to concrete molecular structures. For instance, the description might mention the presence of a functional group, the molecule’s size, or other key features, which the model must translate into a valid SMILES string. This task requires the model to reason about the relationship between the described features and how they translate into a molecular structure. It tests the model’s ability to use abstract information to generate precise molecular representations, showcasing its reasoning in both language and chemistry.

Molecule Design

Question: Generate a molecule’s SMILES string that fits the following description. *Description: The molecule is a member of the class of cyclopentanols carrying 1,2,4-triazol-1-ylmethyl and 4-chlorobenzylidene and geminal dimethyl substituents at positions 1, 2 and 5 respectively. It is a member of triazoles, a member of monochlorobenzenes, a member of cyclopentanols, a tertiary alcohol and an olefinic compound.*

Answer: CC1(C)CC/C(=C)\c2ccc(Cl)cc2C1(O)Cn1cncn1

(4) Molecule Captioning. This task is the inverse problem of molecule design. Given a SMILES string, the model generates a natural language description of the molecule. Here, the model must reason about the structure encoded in the SMILES and generate a coherent description that accurately captures key molecular features, such as functional groups, chemical bonding, and overall molecular properties. The challenge lies in translating the complex, compact SMILES format into a human-readable description that captures both the structure and function of the molecule. It requires the model to interpret a structured representation and reason about how to explain it in a way that makes sense in natural language.

Molecule Captioning

Question: Provide a chemical description for the following molecule. SMILES:

CC1(C)CC/C(=C)\c2ccc(Cl)cc2C1(O)Cn1cncn1.

Answer: *The molecule is a member of the class of cyclopentanols carrying 1,2,4-triazol-1-ylmethyl and 4-chlorobenzylidene and geminal dimethyl substituents at positions 1, 2 and 5 respectively. It is a member of triazoles, a member of monochlorobenzenes, a member of cyclopentanols, a tertiary alcohol and an olefinic compound.*

(5) Text-based Open Molecule Generation. The TOMG-Bench benchmark focuses on text-based open molecule generation, evaluating a model’s ability to generate, modify, and optimize molecular structures based on textual descriptions or specified criteria. Tasks in this benchmark include Customized Molecule Generation, Molecule Editing, and Molecule Optimization. In *Customized Molecule Generation*, the model is tasked with creating molecules that meet specific constraints, such as a predefined number of atoms, bonds, or functional groups, while maintaining chemical validity. *Molecule Editing* requires the model to modify an existing molecule by adding, replacing, or removing functional groups, with the challenge of reasoning about how these changes affect the overall structure and properties of the molecule. *Molecule Optimization* involves optimizing molecules to improve specific properties like LogP (partition coefficient), QED (drug-likeness), and MR (molecular refractivity), where the model must navigate trade-offs between conflicting goals, such as balancing hydrophobicity to improve LogP without compromising QED. Together, these

918 tasks test a model’s ability to generate, edit, and optimize molecules, requiring reasoning about
 919 molecular structure, function, and the interdependencies between chemical properties.
 920

923 Text-based Open Molecule Generation (MolOpt-LogP)

925 **Question:** Modify the molecule

926 N#CCC1 (n2cc (-c3ccnc4 [nH] ccc34) cn2) C [NH+] (C2CCN (C (=O) Nc3cccc3C1)
 927 CC2) C1 to have a lower LogP value.

928 **Answer:** O=C (Nc1cccc1C1) N1CCC ([NH+] 2CC (CO) (n3cc (-c4ccnc5 [nH] ccc4
 929 5) cn3) C2) CC1

934 B.2 REACTION TASKS

937 **(6) Yield Prediction.** Yield prediction involves determining whether a given chemical reaction
 938 will result in *a high or low yield* based on the reactants and reaction conditions in the *Buchwald-*
 939 *Hartwig and Suzuki-coupling reactions*. Here, reasoning is necessary because predicting the yield
 940 requires the model to understand both the intrinsic properties of the reactants and the external factors
 941 that can influence the efficiency of the reaction. It requires the model to simulate the chemical
 942 behavior of the system, predict potential losses, and estimate the likelihood of a successful reaction
 943 based on prior examples. This is a classic task of predicting outcomes under uncertainty, demanding
 944 robust reasoning capabilities to account for various complex variables.

947 Yield Prediction (Buchwald-Hartwig reaction)

949 **Question:** Predict if the following Buchwald-Hartwig reaction is high-yielding (Yes for
 950 >70% yield, No otherwise). Reaction:

951 BrC1CCNc1.Cc1ccc (N) cc1 .
 952 O=S (=O) (O [Pd] 1c2cccc2-c2cccc2N-1) C (F) (F) F .
 953 COc1ccc (OC) c (P ([C@]23C [C@H] 4C [C@H] (C [C@H] (C4) C2) C3) [C@]23C [C@H]
 954 4C [C@H] (C [C@H] (C4) C2) C3) c1-c1c (C (C) C) cc (C (C) C) cc1C (C) C .
 955 CCN=P (N=P (N (C) C) (N (C) C) N (C) C) (N (C) C) N (C) C .
 956 CCOC (=O) c1cc (OC) no1>>Cc1ccc (Nc2ccnc2) cc1 .

957 **Answer:** No

962 **(7) Reagent Selection.** This task involves selecting the appropriate *reagents (reactants, solvents,*
 963 *and ligands)* from a predefined list for a given reaction. Reasoning is critical here, as the model must
 964 understand the chemical context of the reaction and predict which reagents will interact most effec-
 965 tively to drive the desired transformation. Importantly, we choose this task over USPTO-Condition
 966 because, in Reagent Selection, *each option comes with an associated yield value, making the reasoning more concrete and verifiable*. Additionally, this task more closely mimics real-world chemical
 967 practices, where chemists have to select reagents from a limited set of available chemicals, often due
 968 to budget, availability, or experimental constraints. In contrast, USPTO-Condition involves broader,
 969 less constrained reaction conditions that may not align with practical laboratory limitations, and its
 970 accuracy is harder to verify because it lacks specific yield values and focuses on a wider range of
 971 conditions. Thus, Reagent Selection provides a more focused and realistic task, better suited to
 evaluating a model’s ability to reason within the practical boundaries of chemical experimentation.

972
973**Reagent Selection**974
975**Question:** From the provided list, select the optimal reactant to maximize the yield for the following reaction setup.**Reaction Setup:**reactant: Ic1ccc2ncccc2c1catalyst: CC(=O)[O-].CC(=O)[O-].[Pd+2]

ligand: none

reagent: O=P([O-])([O-])[O-].[K+].[K+].[K+]solvent: C1CCOC1list of reactants for selection: ['B(O)O','Cc1ccc2c(cnn2C2CCCCO2)c1B1OC(C)(C)C(C)(C)O1','Cc1ccc2c(cnn2C2CCCCO2)c1[B-](F)(F)F.[K+]']**Answer:** Cc1ccc2c(cnn2C2CCCCO2)c1B1OC(C)(C)C(C)(C)O1

985

986

(8) Reaction Prediction. This task requires *predicting the products of a chemical reaction* based on the given reactants and reaction conditions. Reasoning is essential here because the model needs to understand the underlying chemistry, such as functional group reactivity, reaction mechanisms, and stereochemistry, in order to predict the correct products. Unlike simpler tasks that only involve pattern recognition, this task demands an ability to apply chemical principles (like how certain bonds break and form) and anticipate the reaction's outcome, which requires sophisticated reasoning beyond just memorization.

994

Reaction Prediction

995

Question: Predict the main product(s) for the following reaction. Reactants:CN(C)C=O.COc1cccc(N)c1.Cc1nc(CCl)cs1.C1.O=C([O-])[O-].[K+].[K+]**Answer:** COc1cccc(NCc2csc(C)n2)c1

999

1000

(9) Retrosynthesis. Retrosynthesis involves predicting the *starting materials (reactants)* required to synthesize a given target molecule. To perform reasoning, the model needs to deconstruct the target molecule into simpler components and reverse-engineer the chemical process. This requires understanding reaction pathways, identifying suitable reactions to break bonds, and selecting the appropriate reagents. It's a form of reverse reasoning, where the model must consider multiple potential pathways and choose the one that is most likely to lead to the desired product, based on its chemical structure and reactivity.

1007

1008

Retrosynthesis

1009

Question: Predict the necessary reactant(s) for the following product. Product:CNC(=O)c1c(-c2ccc(F)cc2)oc2cc(N(C)S(C)(=O)=O)c(-c3cccc([N+])(=O)[O-])c3)cc12**Answer:** CNC(=O)c1c(-c2ccc(F)cc2)oc2cc(N(C)S(C)(=O)=O)c(Br)cc12.O=[N+]([O-])c1cccc(B(O)O)c1

1014

1015

1016

1017

1018

1019

1020

1021

1022

1023

1024

1025

1026 **C IMPLEMENTATION DETAILS**
10271028 **C.1 HYPER-PARAMETERS**
10291030 In this section, we list the hyper-parameters used in different phases of training and inference. We
1031 used Llama-Factory to conduct SFT training (Phase 1 and Phase 2), and EasyRL for GRPO training
1032 (Phase 3). Table 3 provides the values for the hyper-parameters in Phase 1, Phase 2, and Phase 3, as
1033 well as for inference. The table includes settings such as the number of GPUs, learning rates, batch
1034 sizes, and the number of epochs for each phase.

Item	Value
<i>Phase 1</i>	
gpu_number (H100)	2
per_device_train_batch_size	1
gradient_accumulation_steps	4
learning_rate	1.0e-5
num_train_epochs	5
lr_scheduler_type	cosine
warmup_ratio	0.1
epoch	3
<i>Phase 2</i>	
gpu_number (H100)	2
per_device_train_batch_size	1
gradient_accumulation_steps	4
learning_rate	1.0e-5
num_train_epochs	5
lr_scheduler_type	cosine
warmup_ratio	0.1
epoch	5
<i>Inference</i>	
temperature	0.6
top_p	0.9
max_tokens	4096

1055 Table 3: Hyper-parameters for Different Phases of Training and Inference
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079

1080
1081

C.2 DATA STATISTICS

1082
1083
1084
1085

In this section, we outline the data splits used across different phases of training. The data partitioning follows the benchmark division strategies, ensuring that the test set is consistent with the evaluation criteria and standards. The amount of training data is strategically varied across the phases to match their distinct objectives:

1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098

- **Phase 1 (Chemical Foundation Training):** The goal is to build a comprehensive foundation of chemical knowledge. Therefore, this phase utilizes a *large volume* of question-answer pairs (e.g., 920k for Name Prediction) to ensure broad exposure to facts and patterns.
- **Phase 2 (CRP Distillation):** This phase focuses on teaching a structured reasoning method using high-quality synthetic CoT data. The strategy here is to provide a substantial and relatively balanced number of examples across different task categories, *generally targeting around 100k samples per major task*. For tasks identified as particularly difficult, such as Name Prediction, we ensure a higher volume of data to help the model master their complex reasoning protocols.
- **Phase 3 (Multi-task GRPO):** The objective is to refine the model’s reasoning skills. For this targeted alignment, the amount of training data for each task is not fixed but is calculated based on the model’s performance after Phase 2. This strategy, detailed in Section 3.3, allows us to concentrate the training effort on tasks the model finds more difficult, optimizing the refinement process.

1099
1100
1101
1102

Importantly, we ensure that no molecules appearing in the test set are included in the training sets. For a detailed overview of the data splits and their distribution, please refer to Table 4. It is particularly noteworthy that the TOMG tasks were intentionally excluded from the GRPO phase (Phase 3) due to their long evaluation times and better performance.

1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133

1134	Tasks	Phase 1	Phase 2	Phase 3	Train	Valid	Test
1135	<i>Name Prediction</i>						
1136	SMILES2IUPAC	920,734	100,000	6,978	828,661	500	100
1137	IUPAC2SMILES	920,734	100,000	7,821	828,661	500	100
1138	<i>Property Prediction</i>						
1139	BACE	1,413	20,000	3,489	1,413	50	100
1140	BBBP	1,950	20,000	2,286	1,950	50	100
1141	ClinTox	1,384	20,000	241	1,384	50	100
1142	HIV	41,027	20,000	0	41,027	50	100
1143	Tox21	7914	20,000	2,166	7,914	50	100
1144	<i>Molecule Design</i>						
1145	ChEBI-20	26,407	100,000	7,821	26,407	3,300	3,300
1146	<i>Molecule Captioning</i>						
1147	ChEBI-20	26,407	100,000	7,008	26,407	3,300	3,300
1148	<i>Text-based Open Molecule Generation</i>						
1149	MolCustom-AtomNum	133,334	33,333	—	133,334	—	5,000
1150	MolCustom-BondNum	133,334	33,333	—	133,334	—	5,000
1151	MolCustom-FunctionalGroup	133,334	33,333	—	133,334	—	5,000
1152	MolEdit-AddComponent	133,333	33,333	—	133,333	—	5,000
1153	MolEdit-DelComponent	133,333	33,333	—	133,333	—	5,000
1154	MolEdit-SubComponent	133,333	33,333	—	133,333	—	5,000
1155	MolOpt-LogP	133,333	33,333	—	133,333	—	5,000
1156	MolOpt-MR	133,333	33,333	—	133,333	—	5,000
1157	MolOpt-QED	133,333	33,333	—	133,333	—	5,000
1158	<i>Yield Prediction</i>						
1159	Buchwald-Hartwig	3,855	40,515	1,925	3,855	50	100
1160	Suzuki-Miyaura	5,660	58,485	1,685	5,660	50	100
1161	<i>Reagent Selection</i>						
1162	Reactant Selection	1,436	44,763	5,174	1,436	100	100
1163	Solvent Selection	1,340	41,770	6,497	1,340	100	100
1164	Ligand Selection	380	13,467	6,517	380	100	100
1165	<i>Reaction Prediction</i>						
1166	USPTO-Mixed	409,035	100,000	3,730	409,035	30,000	100
1167	<i>Retrosynthesis</i>						
1168	USPTO-50k	40,029	100,000	8,663	40,029	5,004	100
1169							
1170							

Table 4: Tasks and data splits across different phases. *Note* that the quantities listed for Phases 1, 2, and 3 refer to the total data volume, inclusive of any repeated samples.

1188
1189

C.3 CRP EXTRACTION PROMPT

1190 Prior to generating the protocol, we establish a reference dataset by sampling the teacher model
 1191 with a simple, generic task prompt. We evaluate the generated responses against ground truth to cat-
 1192 egorize them into two groups: **positive examples**, consisting of the input question and the correct
 1193 reasoning trace, and **negative examples**, which comprise the question, the flawed reasoning, and the
 1194 ground truth for correction. We then feed these curated examples into the CRP Extraction Prompt
 1195 (presented below), enabling the teacher model to synthesize a structured Chemical Reasoning Pro-
 1196 tocol defined by reasonable logical steps and hints for common mistakes.

1197
1198

CRP Extraction Prompt

1199
1200
1201
1202
1203
1204
1205
1206

You are an expert Chemical Methodology Analyst and Pedagogue. Your mission is to ana-
 lyze raw reasoning traces from specific chemical tasks and synthesize a robust, standardized
Chemical Reasoning Protocol (CRP). A CRP is a structured, algorithmic thinking blueprint
 designed to emulate the deliberate decision-making process of a professional chemist. It
 serves as a generalizable scaffold that breaks down complex chemical problems into a de-
 fined sequence of logical steps (derived from successful reasoning), while simultaneously
 integrating specific "Cautionary Guidance" (derived from failed reasoning) to preemptively
 block common hallucinations, syntactic errors, and logical fallacies.

1207
1208
1209

1. Task Context

Task Name: {{TASK_NAME}}
Description: {{TASK_DESCRIPTION}}

1210
1211
1212

2. Input Data

I will provide two sets of reasoning trajectories: (1) Positive Examples: 5 reasoning paths
 that logically lead to the correct answer. (2) Negative Examples: 5 reasoning paths that lead
 to incorrect answers or contain hallucinations.

1213
1214
1215

Positive Examples: {{POSITIVE_EXAMPLES}}
Negative Examples: {{NEGATIVE_EXAMPLES}}

1216
1217
1218

3. Analytical Objectives

Analyze these examples to construct a Generalizable Chemical Reasoning Protocol.

1219
1220
1221
1222
1223
1224

- **Synthesize Logical Steps:** Examine the **Positive Examples**. Abstract the common,
 underlying logical structure into 3-5 distinct steps (e.g., "Analyze Reactants", "Deduce
 Mechanism").
- **Extract Cautionary Hints:** Examine the **Negative Examples**. Identify the root causes of
 failure (e.g., ignoring stereochemistry, misidentifying functional groups). Convert these
 errors into *Cautionary Guidance* for the corresponding steps.
- **Generalization:** The protocol must be universal for this task type, avoiding references to
 specific molecules in the examples.

1225
1226
1227

4. Output Format

Please output the protocol strictly following the Markdown structure below. Do not include
 any other introductory text.

1228
1229
1230

1. **[Step 1 Title]:** [Detailed Description of what to do]
 * [Cautionary hint derived from negative examples]
 * [Strategic hint derived from positive examples]
2. **[Step 2 Title]:** [Detailed Description of what to do]
 * [Cautionary hint...]

1231
1232
1233

...
 (Continue for all necessary steps)

1234
1235
1236
1237
1238
1239
1240
1241

1242 C.4 DISTILLATION PROMPTS

1243

1244

1245

1246

Name Prediction (SMILES→IUPAC)1247 You are an expert chemist demonstrating how to determine the IUPAC name for a given
1248 molecule.

1249

You are given:

1250

1. A SMILES string representing the molecule.

1251

2. A list of functional groups present in the molecule (with representative SMILES patterns)
1252 to serve as hints.

1253

Your task is to generate a step-by-step reasoning process that logically deduces the systematic
1254 IUPAC name from the SMILES string. You MUST NOT mention ground-truth, given
1255 reactants, provided answer, or any similar phrases in your `<think>` block. You must write
1256 the reasoning as if you are deducing the answer from scratch, even though you already know
1257 the destination.

1258

Please structure your reasoning as follows:

1259

1. **Deconstruct & Analyze:** Thoroughly parse the SMILES string to visualize the molecular
1260 graph, identifying all functional groups, cyclic systems, points of unsaturation, and any
1261 specified stereocenters.

1262

* *Translate the linear SMILES notation into a clear mental or physical representation of the
molecular structure.*

1263

* *Systematically identify every functional group and ring, paying close attention to IUPAC
priority.*

1264

* *Note all stereochemical indicators (@, @@, /, \) for inclusion in the final name.*

1265

2. **Determine the Principal Chain or Parent Ring:** Apply IUPAC priority rules to identify
1266 the principal functional group, which in turn determines the parent structure (the longest
1267 chain or main ring system) and the suffix of the name.

1268

* *The parent structure must contain the highest-priority functional group.*

1269

* *If a choice exists, the parent is the structure with the most multiple bonds, then the longest
carbon chain.*

1270

3. **Number & Name the Parts:** Systematically number the atoms of the parent structure to
1271 assign the lowest possible locant to the principal functional group, then identify and name
1272 all substituent groups attached to the parent.

1273

* *The numbering must grant the lowest locant to the principal functional group above all
other considerations.*

1274

* *If a choice remains, assign lowest locants to multiple bonds, then to the first point of dif-
ference in substituent locants.*

1275

* *Correctly name each substituent (e.g., "methyl", "chloro") and note its locant.*

1276

4. **Assemble & Order:** Construct the complete IUPAC name by arranging the named sub-
1277 stituents alphabetically, prefixing them to the parent name, and incorporating all locants and
1278 stereochemical descriptors in their correct positions.

1279

* *Alphabetize substituents by name, ignoring multiplying prefixes (di, tri) but not iso or neo.*

1280

* *Assemble the final name in the standard order: Stereochemistry-Substituents-Parent-Suffix.*

1281

Input:

1282

SMILES: {instruction}

1283

FunctionalGroups: {functional_groups_str}

1284

Ground-truth IUPAC Name: {target}

1285

Output:

1286

<think>

1287

1. **Deconstruct & Analyze:** ...

1288

2. **Determine the Principal Chain or Parent Ring:** ...

1289

3. **Number & Name the Parts:** ...

1290

4. **Assemble & Order:** ...

1291

</think>

1292

<answer>{target}</answer>

1293

1294

1295

1296
1297**Name Prediction (IUPAC→SMILES)**1298
1299

You are an expert chemist demonstrating how to determine the IUPAC name for a given molecule.

1300

You are given:

1301

1. An IUPAC name representing the molecule.

1302
1303

2. A list of functional groups present in the molecule (with representative SMILES patterns) to serve as hints.

1304
1305
1306
1307

Your task is to generate a step-by-step reasoning process that logically deduces the SMILES string from the IUPAC name. You MUST NOT mention ground-truth, given reactants, provided answer, or any similar phrases in your `<think>` block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1308

Please structure your reasoning as follows:

1309
1310
1311

1. **Deconstruct IUPAC Name & Identify Components:** Break down the IUPAC name into its fundamental components, identifying the parent structure (chain or ring), principal functional group, all substituents, and their corresponding locants.

* *The parent name (e.g., "hexane", "cyclohexane") forms the SMILES backbone.*

* *Pay close attention to suffixes (-ol, -one, -oic acid) as they define the principal group.*

* *Isolate all stereochemical prefixes (R/S, E/Z, cis/trans) for later use.*

1312
1313
1314

2. **Formulate Conceptual Structure:** Translate the deconstructed components into a clear, verbal description of the molecule's connectivity, specifying how each substituent attaches to the parent structure at its designated locant.

* *This step acts as a "blueprint" before writing any SMILES code.*

* *Mentally or explicitly number the parent chain to map out attachment points.*

* *Clarify the precise atom-to-atom connections for all parts.*

1315
1316
1317

3. **Attempt SMILES Translation:** Individually convert each identified component—the parent structure, substituents, and functional groups—into its correct SMILES representation, treating them as separate fragments.

* *Use parentheses () for branches off the main chain.*

* *Represent double and triple bonds with '=' and '#', respectively.*

* *Keep track of numeric labels for defining ring closures.*

1318
1319
1320

4. **Construct Final SMILES String:** Systematically assemble the individual SMILES fragments into a single, valid string, ensuring correct connectivity, branching, ring closures, and the inclusion of stereochemical markers.

* *Start with the fragment for the principal functional group or a logical endpoint of the main chain.*

* *Insert substituent fragments inside parentheses at the correct atom of the main chain.*

* *Apply stereochemical symbols (@, @@, /, \) last to ensure correct placement relative to the final structure.*

1321

Input:

1322

IUPAC: {instruction}

1323

FunctionalGroups: {functional_groups_str}

1324

Ground-truth SMILES Name: {target}

1325

Output:

1326

`<think>`

1327

1. **Deconstruct Name & Identify Components:** ...

1328

2. **Formulate Conceptual Structure:** ...

1329

3. **Translate Components to SMILES:** ...

1330

4. **Assemble & Finalize String:** ...

1331

`</think>`

1332

`<answer>{target}</answer>`

1333

1334

1335

1336

1337

1338

1339

1340

1341

1342

1343

1344

1345

1346

1347

1348

1349

1350
1351**Property Prediction (BACE1)**1352
1353

You are an expert chemist demonstrating how to predict the BACE1 inhibition potential of a molecule from its chemical structure.

1354

You are given:

1355

1. A SMILES string representing the molecule.

1356

2. A list of functional groups present in the molecule (with representative SMILES patterns) to serve as hints.

1357

Your task is to generate a step-by-step reasoning process that logically deduces whether the molecule can inhibit (Yes) or cannot inhibit (No) the Beta-site Amyloid Precursor Protein Cleaving Enzyme 1 (BACE1). You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your `<think>` block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1363

Please structure your reasoning as follows:

1364

1. **Structural/Property Identification:** Analyze the SMILES to identify key pharmacophores, functional groups, and physicochemical properties relevant for BACE1 inhibition.

1365

* Identify key pharmacophoric features (H-bond donors/acceptors, aromatic rings).

1366

* Assess general drug-like properties such as MW, flexibility, and logP.

1367

* Check for structural motifs common to known inhibitors of this target class.

1368

2. **Property-Activity Correlation:** Correlate the identified features with general Structure-Activity Relationships (SAR) for BACE1 inhibition.

1369

* Assess how functional groups might form hydrogen bonds or electrostatic interactions in an active site.

1370

* Evaluate if hydrophobic regions are suitably sized for binding within lipophilic pockets.

1371

* Align the molecule's properties with general drug-likeness principles (e.g., Lipinski's Rules).

1372

3. **Holistic Assessment & Consideration:** Synthesize all positive and negative factors to form a holistic judgment of the molecule's inhibitory potential.

1373

* Weigh the combined evidence for and against binding, not just a simple count of features.

1374

* Consider potential liabilities that could prevent binding, such as steric hindrance or unfavorable charges.

1375

4. **Final Prediction Verdict:** Conclude with a definitive "Yes" or "No" prediction based on the integrated assessment.

1376

* The final verdict must be a direct logical consequence of the previous assessment.

1377

* State the primary molecular features that drive the final "Yes" or "No" decision.

1378

Wrap your entire reasoning process in `<think>` tags and output only the final answer (Yes or No) in `<answer>` tags.

1379

Input:

1380

SMILES: {instruction}

1381

FunctionalGroups: {functional_groups_str}

1382

Ground-truth Answer: {target}

1383

Output:

1384

`<think>`

1385

1. Structural/Property Identification: ...

1386

2. Property-Activity Correlation: ...

1387

3. Holistic Assessment & Consideration: ...

1388

4. Final Prediction Verdict: ...

1389

`</think>`

1390

`<answer>{target}</answer>`

1391

1392

1393

1394

1395

1396

1397

1398

1399

1400

1401

1402

1403

1404
1405**Property Prediction (BBBP)**1406
1407

You are an expert chemist demonstrating how to predict the Blood-Brain Barrier (BBB) penetration potential of a molecule from its chemical structure.

1408

You are given:

1409

1. A SMILES string representing the molecule.

1410

2. A list of functional groups present in the molecule (with representative SMILES patterns) to serve as hints.

1411

Your task is to generate a step-by-step reasoning process to predict the binary label (Yes for penetration, No for non-penetration) for the given molecule's ability to cross the Blood-Brain Barrier. You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your <think> block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1412

Please structure your reasoning as follows:

1413

1. **Structural/Property Identification:** Identify key physicochemical properties from the SMILES that govern membrane permeability.

1414

* Estimate properties like molecular weight (MW), lipophilicity (logP), and polar surface area (TPSA).

1415

* Count hydrogen bond donors/acceptors and identify any permanent charges.

1416

2. **Property-Activity Correlation:** Correlate the identified properties with principles favoring or hindering BBB penetration.

1417

* Assess if size, polarity, and lipophilicity fall within a favorable range for passive diffusion.

1418

* Recognize that high polarity or a permanent charge strongly hinders penetration.

1419

3. **Holistic Assessment & Consideration:** Synthesize all factors to form a balanced judgment on the molecule's overall profile for BBB penetration.

1420

* Weigh the combined impact of all properties; a single critical flaw can prevent penetration.

1421

* Consider the possibility of active efflux mechanisms (e.g., P-gp substrate).

1422

4. **Final Prediction Verdict:** Conclude with a definitive "Yes" or "No" prediction based on the holistic assessment.

1423

* The final verdict must be a logical conclusion of the prior analysis.

1424

* State the most influential properties driving the decision.

1425

Wrap your entire reasoning process in <think> tags and output only the final answer (Yes or No) in <answer> tags.

1426

Input:

1427

SMILES: {instruction}

1428

FunctionalGroups: {functional_groups_str}

1429

Ground-truth Answer: {target}

1430

Output:

1431

<think>

1432

1. **Structural/Property Identification:** ...

1433

2. **Property-Activity Correlation:** ...

1434

3. **Holistic Assessment & Consideration:** ...

1435

4. **Final Prediction Verdict:** ...

1436

</think>

1437

<answer>{target}</answer>

1438

1439

1440

1441

1442

1443

1444

1445

1446

1447

1448

1449

1450

1451

1452

1453

1454

1455

1456

1457

1458
1459**Property Prediction (ClinTox)**1460
1461

You are an expert chemist demonstrating how to predict the clinical trial toxicity of a molecule from its chemical structure.

1462

You are given:

1463

1. A SMILES string representing the molecule.

1464

2. A list of functional groups present in the molecule (with representative SMILES patterns) to serve as hints.

1465

3. The FDA approval status for clinical trials.

1466

Your task is to generate a step-by-step reasoning process to predict whether a molecule is Clinically-trial-Toxic (Yes) or Not Clinically-trial-toxic (No). The FDA-approved status is provided as an additional piece of evidence. You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your `<think>` block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1467

Please structure your reasoning as follows:

1468

1. **Structural/Property Identification:** Identify structural alerts, reactive functional groups, and other properties linked to toxicity from the SMILES.

1469

* Identify known toxicophores and chemically reactive motifs.

1470

* Assess if extreme physicochemical properties (e.g., high MW or logP) increase risk.

1471

2. **Property-Activity Correlation:** Correlate the identified structural risks with general toxicological principles.

1472

* Assess how reactive groups could potentially cause organ damage.

1473

* Note that a structural alert suggests risk but is not definitive proof of toxicity.

1474

3. **Holistic Assessment & Consideration:** Synthesize the structural risk analysis with the provided FDA approval status to form a comprehensive judgment.

1475

* Use the FDA approval status as powerful evidence to adjust the initial structural risk assessment.

1476

* An "Approved" status implies any identified structural risks were clinically acceptable.

1477

4. **Final Prediction Verdict:** Conclude with a definitive "Yes" or "No" prediction based on the integrated evidence.

1478

* The verdict must balance the structural analysis against the FDA approval status.

1479

* State the primary factors that drive the final "Yes" or "No" decision.

1480

Wrap your entire reasoning process in `<think>` tags and output only the final answer (Yes or No) in `<answer>` tags.

1481

Input:

1482

SMILES: {instruction}

1483

FunctionalGroups: {functional_groups_str}

1484

FDA_APPROVED: {FDA_APPROVED}

1485

Ground-truth Answer: {target}

1486

Output:

1487

`<think>`

1488

1. **Structural/Property Identification:** ...

1489

2. **Property-Activity Correlation:** ...

1490

3. **Holistic Assessment & Consideration:** ...

1491

4. **Final Prediction Verdict:** ...

1492

`</think>`

1493

`<answer>{target}</answer>`

1494

1495

1496

1497

1498

1499

1500

1501

1502

1503

1504

1505

1506

1507

1508

1509

1510

1511

1512
1513**Property Prediction (HIV)**1514
1515

You are an expert chemist demonstrating how to predict the HIV replication inhibition potential of a molecule from its chemical structure.

1516

You are given:

1517

1. A SMILES string representing the molecule.

1518

2. A list of functional groups present in the molecule (with representative SMILES patterns) to serve as hints.

1519

3. The result from an activity test, classified as CA (confirmed active), CM (confirmed moderately active), or CI (confirmed inactive).

1520

Your task is to generate a step-by-step reasoning process to predict the binary label for the molecule's ability to inhibit HIV replication (Yes or No), considering its structure and the provided activity test result. You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your `<think>` block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1521

Please structure your reasoning as follows:

1522

1. **Structural/Property Identification:** Analyze the SMILES for structural motifs and properties linked to antiretroviral activity.

1523

* Identify scaffolds similar to known HIV inhibitor classes (e.g., protease inhibitors).

1524

* Assess functional groups for key interactions (e.g., H-bonding) with viral targets.

1525

* Evaluate general properties like size and polarity relevant for cellular uptake.

1526

2. **Property-Activity Correlation:** Correlate the identified features with general mechanisms of HIV inhibition.

1527

* Assess the molecule's potential to mimic the natural substrates of viral enzymes.

1528

* Consider if the molecule's geometry is suitable for binding to key viral targets.

1529

3. **Holistic Assessment & Consideration:** Synthesize the structural analysis with the provided activity data to form a comprehensive judgment.

1530

* Use the activity result (CA/CM/CI) as key evidence to confirm or override the structural analysis.

1531

* A "CA" or "CM" result strongly supports a "Yes" prediction, even with subtle structural cues.

1532

* A "CI" result strongly supports a "No" prediction, regardless of favorable features.

1533

4. **Final Prediction Verdict:** Conclude with a definitive "Yes" or "No" prediction based on the integrated evidence.

1534

* The verdict must clearly integrate the structural analysis with the activity data.

1535

* State the primary factors (e.g., "Structurally promising and confirmed by 'CA' result") driving the decision.

1536

* The final verdict must explicitly state how the structural analysis and the activity data were combined.

1537

* Summarize the key factors (e.g., "Structurally promising and confirmed active by 'CA' result") driving the decision.

1538

Wrap your entire reasoning process in `<think>` tags and output only the final answer (Yes or No) in `<answer>` tags.

1539

Input:

1540

SMILES: {instruction}

1541

FunctionalGroups: {functional_groups_str}

1542

activity: {activity}

1543

Ground-truth Answer: {target}

1544

Output:

1545

`<think>`

1546

1. **Structural/Property Identification:** ...

1547

2. **Property-Activity Correlation:** ...

1548

3. **Holistic Assessment & Consideration:** ...

1549

4. **Final Prediction Verdict:** ...

1550

`</think>`

1551

`<answer>{target}</answer>`

1552

1566
1567**Property Prediction (Tox21)**1568
1569

You are an expert chemist demonstrating how to predict the toxicity of a molecule from its chemical structure.

1570

You are given:

1571

1. A SMILES string representing the molecule.

1572

2. A list of functional groups present in the molecule (with representative SMILES patterns) to serve as hints.

1573

Your task is to generate a step-by-step reasoning process to predict whether a molecule is toxic (Yes) or not toxic (No). You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your <think> block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1574

Please structure your reasoning as follows:

1575

1. **Structural/Property Identification:** Analyze the SMILES string for known toxicophores, structural alerts, and chemically reactive functional groups.

1576

* Identify highly reactive motifs (e.g., Michael acceptors, epoxides) or strained rings.

1577

* Search for groups that can be metabolically activated into toxic species (e.g., nitroaromatics).

1578

* Assess if extreme physicochemical properties (e.g., high logP) suggest non-specific toxicity.

1579

2. **Property-Activity Correlation:** Correlate the identified structural features with common mechanisms of chemical toxicity.

1580

* Evaluate the potential for reactive groups to cause cellular damage via covalent modification of proteins or DNA.

1581

* Consider the possibility of off-target pharmacology based on the molecule's overall profile.

1582

* Recognize that a structural alert increases toxicity risk but is not definitive proof.

1583

3. **Holistic Assessment & Consideration:** Synthesize all structural risk factors to form a comprehensive judgment on the molecule's likely toxicity.

1584

* Weigh the combined impact of all identified risks; the presence of multiple alerts is a strong indicator of toxicity.

1585

* Consider if the overall molecular context might mitigate the reactivity of a potential toxicophore (e.g., via steric hindrance).

1586

4. **Final Prediction Verdict:** Conclude with a definitive "Yes" (toxic) or "No" (not toxic) prediction based on the overall risk assessment.

1587

* The verdict must be a logical conclusion derived from the balance of identified structural risks.

1588

* State the primary structural reasons that drive the final "Yes" or "No" decision.

1589

Wrap your entire reasoning process in <think> tags and output only the final answer (Yes or No) in <answer> tags.

1590

Input:

1591

SMILES: {instruction}

1592

FunctionalGroups: {functional_groups_str}

1593

Ground-truth Answer: {target}

1594

Output:

1595

<think>

1596

1. **Structural/Property Identification:** ...

1597

2. **Property-Activity Correlation:** ...

1598

3. **Holistic Assessment & Consideration:** ...

1599

4. **Final Prediction Verdict:** ...

1600

</think>

1601

<answer>{target}</answer>

1602

1603

1604

1605

1606

1607

1608

1609

1610

1611

1612

1613

1614

1615

1616

1617

1618

1619

1620
1621**Molecule Design (Text→SMILES)**

1622

You are an expert chemist demonstrating how to solve a molecule design problem.

1623

You are given:

1624

1. A molecular description.

1625

2. A list of functional groups present in the molecule (with representative SMILES patterns).

1626

Your task is to generate a step-by-step reasoning process that logically deduces the SMILES from the description and the functional groups. You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your `<think>` block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1629

Please structure your reasoning as follows:

1630

1. **Identify the core structure:** Determine the central molecular scaffold or backbone by interpreting the root name in the chemical description.

1631

* Focus on identifying the parent compound, such as a main chain ('hexane') or a ring system ('benzene').

1632

* This core structure will serve as the primary chain for the SMILES string.

1633

2. **Summarize key elements:** Systematically list all substituents and functional groups from the description and identify their attachment points (locants) on the core.

1634

* Create a clear connectivity map, linking each group to its specific locant on the parent structure.

1635

* Use the provided functional group list to confirm the identity of complex groups.

1636

3. **Address stereochemistry:** Scan the description for any stereochemical descriptors and determine the specific configuration at each chiral center or double bond.

1637

* Identify specific prefixes such as (R)/(S) for chiral centers and (E)/(Z) for double bonds.

1638

* If no stereochemistry is mentioned, explicitly note this fact.

1639

4. **Construct SMILES:** Build the final SMILES string by translating the conceptual blueprint into the correct syntax, combining the core, substituents, and stereochemistry.

1640

* Use parentheses () for branches and numeric labels for ring closures.

1641

* Apply stereochemical markers (@, @@, /, \) at their correct positions after establishing atom connectivity.

1642

Wrap your reasoning in `<think>` tags and output the final SMILES in `<answer>` tags.

1643

Input:

1644

Description: {instruction}

1645

FunctionalGroups: {functional_groups_str}

1646

Ground-truth SMILES: {target}

1647

Output:

1648

`<think>`

1649

1. **Identify the core structure:** ...

1650

2. **Summarize key elements:** ...

1651

3. **Address stereochemistry:** ...

1652

4. **Construct SMILES:** ...

1653

`</think>`

1654

`<answer>{target}</answer>`

1655

1656

1657

1658

1659

1660

1661

1662

1663

1664

1665

1666

1667

1668

1669

1670

1671

1672

1673

1674
1675**Molecule Captioning (SMILES→Text)**

1676

You are an expert chemist demonstrating how to describe a molecule based on its structure.

1677

You are given:

1678

1. A SMILES string.

1679

2. A list of functional groups present in the molecule (with representative SMILES patterns). Your task is to generate a step-by-step reasoning process that logically deduces a detailed chemical description from the SMILES. You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your <think> block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1680

Please structure your reasoning as follows:

1681

1. Analyze Molecular Structure: Deconstruct the SMILES string by identifying the primary carbon backbone with its branching points, systematically locating all heteroatoms and specified functional groups, and characterizing any cyclic systems found via ring-closure numerals.

1682

* Account for all specified functional groups without omission.

1683

* Note aromaticity from lowercase letters to define the core scaffold.

1684

2. Summarize Key Elements: Integrate the individual components into a cohesive overview to characterize the core scaffold, such as an aliphatic chain or aromatic system, and determine the principal chemical class or classes to which the molecule belongs based on the combination of functional groups.

1685

* Synthesize findings into a chemical class (e.g., "a substituted ester"), not a simple list of parts.

1686

* Infer the molecule's dominant chemical character from its functional groups.

1687

* Consider the relative positions of substituents on the core structure.

1688

3. Address Stereochemistry: Scrutinize the SMILES string for stereochemical indicators, specifically '@' and '@@' for tetrahedral chirality and '/' and '\' for geometric isomerism, to explicitly determine the R/S or E/Z configuration for each identified stereocenter or double bond.

1689

* Explicitly state if stereochemistry is absent or unspecified.

1690

* Link each R/S or E/Z descriptor to its precise atom or bond.

1691

4. Predict Molecular Description: Construct a final, detailed chemical description by systematically assembling the preceding findings, starting with the core structure before adding the names and locations of all functional groups and incorporating stereochemical descriptors to form a complete narrative.

1692

* The final description must be a cohesive, flowing narrative.

1693

* Ensure all identified structural and stereochemical features are included.

1694

* Avoid a simple list of facts; the summary should be a comprehensive chemical story.

1695

Wrap your complete reasoning process in <think> tags and output the final description in <answer> tags.

1696

Input:

1697

SMILES: {instruction}

1698

FunctionalGroups: {functional_groups_str}

1699

Ground-truth description: {target}

1700

Output:

1701

<think>

1702

1. Analyze Molecular Structure: ...

1703

2. Summarize Key Elements: ...

1704

3. Address Stereochemistry: ...

1705

4. Predict Molecular Description: ...

1706

</think>

1707

<answer>{target}</answer>

1708

1709

1710

1711

1712

1713

1714

1715

1716

1717

1718

1719

1720

1721

1722

1723

1724

1725

1726

1727

1728
1729**TOMG (MolCustom)**1730
1731

You are an expert chemist demonstrating how to solve a molecule generation problem based on given constraints (Mol Custom).

1732

You are given:

1733

1. An instruction that specifies the constraints for the molecule to be generated.

1734

Your task is to generate a step-by-step reasoning process that logically constructs a valid molecule that satisfies all the given constraints. You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your <think> block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1738

Please structure your reasoning as follows:

1739

1. **Analyze Constraints:** Deconstruct the instruction to create a comprehensive checklist of all specified molecular requirements.

1740

* List all explicit constraints, such as required functional groups, atom counts, or ring systems.

1741

* Identify any implicit chemical rules, like ensuring correct valency for all atoms.

1742

* This checklist will serve as the guide and final validation tool.

1743

2. **Construct Core Backbone:** Begin molecular construction by creating the primary scaffold that satisfies the most significant structural constraint.

1744

* Start with the largest specified fragment, such as a named ring system or a carbon chain of a specific length.

1745

* If no core is defined, choose a simple starting point that facilitates adding other required features.

1746

3. **Assemble and Refine Structure:** Iteratively add the remaining functional groups and substituents to the core backbone to satisfy all remaining constraints.

1747

* Attach the required groups at chemically plausible positions on the backbone.

1748

* Adjust the structure as needed, ensuring all atoms adhere to standard valency rules during assembly.

1749

4. **Comprehensively Validate:** Perform a final, rigorous check of the generated molecule against the initial constraint checklist for accuracy and chemical validity.

1750

* Systematically verify that every single constraint from the initial list has been met.

1751

* Confirm the final SMILES string is syntactically correct and represents a chemically stable molecule.

1752

The final predicted molecule must be expressed as a SMILES string. Wrap your entire reasoning process in <think> tags and output only the final answer in <answer> tags.

Input:

1754

Instruction: {instruction}

1755

Ground-truth Molecule: {target}

Output:

1756

<think>

1757

1. **Analyze Constraints:** ...

1758

2. **Construct Core Backbone:** ...

1759

3. **Assemble and Refine Structure:** ...

1760

4. **Comprehensively Validate:** ...

1761

</think>

1762

<answer>{target}</answer>

1763

1764

1765

1766

1767

1768

1769

1770

1771

1772

1773

1774

1775

1776

1777

1778

1779

1780

1781

1782
1783 **TOMG (MolEdit)**1784 You are an expert chemist demonstrating how to solve a molecule editing problem.
1785

1786 You are given:

1787 1. A molecule editing problem.

1788 2. A list of key functional groups present in the SMILES, which can serve as hints.

1789 Your task is to generate a step-by-step reasoning process that logically deduces the final
1790 edited molecule's SMILES from the given starting molecule and editing instruction. You
1791 MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar
1792 phrases in your <think> block. You must write the reasoning as if you are deducing the
1793 answer from scratch, even though you already know the destination.

1794 Please structure your reasoning as follows:

1795 1. **Task Parsing:** Deconstruct the problem statement to identify the starting molecule, the
1796 specific editing operation, and the target substructure.1797 * *Clearly state the input molecule's SMILES and the intended action (e.g., 'add', 'remove',
1798 'replace').*1799 * *Identify the precise chemical group or site targeted for modification as described in the
1800 instruction.*1801 2. **Functional Group Analysis & Site Localization:** Analyze the starting molecule's
1802 SMILES to pinpoint the exact location for the edit.1803 * *Identify all functional groups within the starting molecule to create a structural map.*1804 * *Use this map to locate the exact atom(s) corresponding to the target site from the instruc-
1805 tion.*1806 3. **Edit Molecule:** Execute the specified modification on the SMILES string by performing
1807 the necessary bond-breaking and bond-forming operations.1808 * *Describe the transformation in terms of SMILES manipulation (e.g., "inserting a 'C' atom
1809 into the chain").*1810 * *Generate the new SMILES string that represents the post-edit molecular structure.*1811 4. **Structure Validation:** Verify that the resulting molecule accurately reflects the requested
1812 edit and is chemically plausible.1813 * *Confirm that the modification has been correctly applied to the intended site.*1814 * *Check the final SMILES for correct syntax and ensure all atoms have proper valency.*1815 The final predicted molecule must be expressed as a SMILES string. You must
1816 output your reasoning within <think>...</think> and the final answer within
1817 <answer>...</answer>.1818 **Input:**

1819 Problem: {problem}

1820 FunctionalGroups: {functional_groups_str}

1821 Ground-truth Edited Molecule: {target}

1822 **Output:**

1823 <think>

1824 1. **Task Parsing:** ...1825 2. **Functional Group Analysis & Site Localization:** ...1826 3. **Edit Molecule:** ...1827 4. **Structure Validation:** ...

1828 </think>

1829 <answer>{target}</answer>

1830

1831

1832

1833

1834

1835

1836
1837**TOMG (MolOptimization)**

1838

You are an expert chemist demonstrating how to solve a molecule optimization problem.

1839

You are given:

1840

1. A molecule optimization problem, which includes the starting molecule (SMILES) and the optimization goal.
2. A list of key functional groups present in the starting molecule, which can serve as hints. Your task is to generate a step-by-step reasoning process that logically deduces the final optimized molecule's SMILES. You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your `<think>` block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1841

Please structure your reasoning as follows:

1842

1. **Clarify Optimization Goal:** Interpret the optimization instruction to define the target property and the underlying chemical principle for its improvement.

1843

* *Clearly state the property to be optimized (e.g., solubility, binding affinity, metabolic stability).*

1844

* *Connect the goal to a specific chemical strategy (e.g., "increase solubility by adding polar groups").*

1845

2. **Molecular Analysis & Strategy Customization:** Analyze the starting molecule to identify a suitable site for modification that aligns with the optimization strategy.

1846

* *Identify regions of the molecule where modification is chemically feasible and unlikely to disrupt core activity.*

1847

* *Select a specific substituent or functional group to add/remove/replace based on the defined strategy.*

1848

3. **Modify Molecule:** Execute the planned structural modification on the starting molecule's SMILES to generate the new, optimized structure.

1849

* *Describe the specific edit performed on the SMILES string to implement the chemical change.*

1850

* *Generate the new SMILES string representing the final, optimized molecule.*

1851

4. **Validate Modification:** Confirm that the edit was performed correctly and that the resulting molecule plausibly achieves the optimization goal.

1852

* *Verify that the final structure incorporates the intended change at the correct location.*

1853

* *Briefly rationalize how the modification is expected to improve the target property (e.g., "the added hydroxyl group increases polarity").*

1854

The final predicted molecule must be expressed as a SMILES string. You must output your reasoning within `<think>...</think>` and the final answer within `<answer>...</answer>`.

1855

Input:

1856

Problem: {problem}

1857

FunctionalGroups: {functional_groups_str}

1858

Ground-truth Optimized Molecule: {target}

1859

Output:

1860

`<think>`

1861

1. **Clarify Optimization Goal:** ...

1862

2. **Molecular Analysis & Strategy Customization:** ...

1863

3. **Modify Molecule:** ...

1864

4. **Validate Modification:** ...

1865

`</think>`

1866

`<answer>{target}</answer>`

1867

1868

1869

1870

1871

1872

1873

1874

1875

1876

1877

1878

1879

1880

1881

1882

1883

1884

1885

1886

1887

1888

1889

1890

Reaction Prediction

1891

1892 You are an expert chemist demonstrating how to solve a reaction prediction problem.

1893

1894 You are given:

1895

1896 1. A set of chemical reactants represented by SMILES string.

1897

1898 2. A list of key functional groups present in the reactants, which serve as hints for discon-

1899

1900 nection.

1901 Your task is to generate a step-by-step reasoning process that logically predicts the product

1902 SMILES from the given reactant SMILES and its functional groups. You MUST NOT men-

1903 tion "ground-truth", "given reactants", "provided answer", or any similar phrases in your

1904 <think> block. You must write the reasoning as if you are deducing the answer from

1905 scratch, even though you already know the destination.

1906

1907 Please structure your reasoning as follows:

1908

1909 **1. Reactant Analysis:** Examine the reactants' structures to identify key features relevant to

1910 reactivity.

1911 * Assess the overall structure, including the carbon skeleton and stereochemistry.

1912 * Identify any notable structural motifs that could influence the reaction.

1913 **2. Functional Group Identification:** Locate and classify all functional groups to determine

1914 potential reaction sites.

1915 * Identify all key functional groups and classify them by reactivity (e.g., nucleophilic, elec-

1916 trophilic, acidic, basic).

1917 * Consider potential interactions between these groups or with external reagents.

1918 **3. Reaction Type Analysis:** Determine the most probable reaction pathway based on the

1919 reactants' structural and functional group properties.

1920 * Deduce the most plausible chemical transformation by correlating the functional groups

1921 present.

1922 * Propose a likely reaction type, such as a known named reaction or a fundamental class

1923 like substitution or addition.

1924 **4. Product Prediction:** Construct the final product by applying the proposed reaction mech-

1925 anism and generating the resultant SMILES string.

1926 * Detail the bond formations and breakages, considering regioselectivity and stereoselectiv-

1927 ity to determine the precise product structure.

1928 * Generate the new SMILES string that represents the final predicted molecule(s).

1929 The final predicted product must be expressed as a SMILES string. If multiple prod-

1930 ucts are predicted, they MUST be separated by a period '.' instead of commas. You

1931 must output your reasoning within <think>...</think> and the final answer within

1932 <answer>...</answer>.

Input:

1933 Reactants SMILES: {reactants}

1934 FunctionalGroups: {functional_groups_str}

1935 Ground-truth Product: {target}

Output:

1936 <think>

1937 1. **Reactant Analysis:** ...

1938 2. **Functional Group Identification:** ...

1939 3. **Reaction Type Analysis:** ...

1940 4. **Product Prediction:** ...

1941 </think>

1942 <answer>{target}</answer>

1943

1944

1945

1946

1947

1948

1949

1950

1951

1952

1953

1954

1955

1956

1957

1958

1959

1960

1961

1962

1963

1964

1965

1966

1967

1968

1969

1970

1971

1972

1973

1974

1975

1976

1977

1978

1979

1980

1981

1982

1983

1984

1985

1986

1987

1988

1989

1990

1991

1992

1993

1994

1995

1996

1997

1998

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1999

1944
1945**Retrosynthesis**

1946

You are an expert chemist demonstrating how to solve a retrosynthesis problem.

1947

You are given:

1948

1. A chemical product represented by a SMILES string.

1949

2. A list of key functional groups present in the product, which serve as hints for disconnection.

1950

Your task is to generate a step-by-step reasoning process that logically deduces the reactant SMILES from the given product SMILES and its functional groups. You MUST NOT mention "ground-truth", "given reactants", "provided answer", or any similar phrases in your <think> block. You must write the reasoning as if you are deducing the answer from scratch, even though you already know the destination.

1951

Please structure your reasoning as follows:

1952

1. Product Analysis: Examine the product's structure to identify strategic bonds for disconnection.

1953

** Assess the product's carbon skeleton, stereochemistry, and any notable structural motifs.*

1954

** Identify key bonds whose formation is synthetically plausible, often adjacent to functional groups.*

1955

2. Functional Group Identification: Analyze the functional groups to infer the last synthetic step.

1956

** Locate all key functional groups, recognizing them as the results of specific chemical reactions.*

1957

** Use these groups to guide the selection of a primary disconnection strategy.*

1958

3. Reaction Type Analysis: Propose a plausible forward reaction that could have formed a key bond in the product.

1959

** Based on the target bond and adjacent functionalities, suggest a reliable forward synthesis reaction (e.g., Wittig, Grignard, Diels-Alder).*

1960

** This proposed reaction dictates the nature of the precursor fragments (synthons).*

1961

4. Reactant Prediction: Perform the disconnection to derive the structures of the starting materials.

1962

** Break the target bond in the product to generate the corresponding synthons.*

1963

** Convert these synthons into stable, charge-neutral reactant molecules and generate their SMILES strings.*

1964

The final predicted reactants must be expressed as SMILES strings. If multiple reactants are predicted, they MUST be separated by a period '.' instead of commas. You must output your reasoning within <think>...</think> and the final answer within <answer>...</answer>.

1965

Input:

1966

Product SMILES: {products}

1967

FunctionalGroups: {functional_groups_str}

1968

Ground-truth Reactants: {target}

1969

Output:

1970

<think>

1971

1. Product Analysis: ...

1972

2. Functional Group Identification: ...

1973

3. Reaction Type Analysis: ...

1974

4. Reactant Prediction: ...

1975

</think>

1976

<answer>{target}</answer>

1977

1978

1979

1980

1981

1982

1983

1984

1985

1986

1987

1988

1989

1990

1991

1992

1993

1994

1995

1996

1997

1998
1999

C.5 ANALYSIS OF CRPs

2000
2001

We provide a detailed analysis regarding the quantification, generalization, and robustness of the Chemical Reasoning Protocols (CRP) distilled in Phase 2.

2002

Protocol Count and Generalization. Our cross-task merging process resulted in **9 core protocols** built around fundamental “reasonable steps” such as “Analyze Molecular Structure” or “Identify Functional Groups.” These steps are abstract and universal enough to cover virtually all chemical problem-solving workflows. The protocols generalize effectively for two reasons. First, the steps themselves remain consistent across tasks; what changes is the task-specific focus within each step. For example, “Analyze Molecular Structure” might focus on stereocenters for a property prediction task but on potential disconnection sites for a retrosynthesis task. Second, the “hints for common mistakes” derived from one task (e.g., misinterpreting a SMILES string for a reactant) are often relevant to others, creating a rich, shared knowledge base that enhances generalization.

2011

Robustness to Teacher Errors. Our protocol design is inherently robust to partially incorrect teacher traces. The goal of this phase is not to perfectly replicate every detail of a teacher’s trace, but to extract a high-level, structured workflow. We address teacher errors through two mechanisms: (1) If the final result is wrong, we analyze the reasoning chain to identify the failure mode, and this analysis directly informs the “hints for common mistakes” attached to the relevant protocol step. (2) If the final result is correct but the reasoning is partially flawed, our process remains effective because we only summarize the broad, high-level steps (e.g., “the model first identified reactants, then proposed a reaction type...”). Minor logical flaws within a step do not corrupt the extraction of this high-level structure. Therefore, the protocol remains sound, and the risk of propagating subtle errors is minimal.

2021

2022

2023

2024

2025

2026

2027

2028

2029

2030

2031

2032

2033

2034

2035

2036

2037

2038

2039

2040

2041

2042

2043

2044

2045

2046

2047

2048

2049

2050

2051

2052
2053

C.6 HUMAN EVALUATION

2054
2055
2056
2057
2058
2059
2060
2061
2062
2063

To provide a nuanced assessment of reasoning quality, we conducted a human evaluation with the help of chemistry experts. The experts were tasked with evaluating the generated CoT from Chem-R and several leading baseline models. We designed a comprehensive rubric consisting of six distinct, orthogonal metrics to capture different facets of a high-quality explanation: *Chemical Soundness* (the factual correctness of the chemistry), *Logical Coherence* (the step-by-step logical flow), *Step-by-Step Completeness* (whether crucial steps are missing), *Justification of the Conclusion* (the faithfulness of the reasoning to the final answer), *Clarity and Conciseness* (the quality of the language), and *Expert-level Insight* (the depth and nuance of the reasoning). Each metric was scored on a 0-5 scale, allowing for a detailed comparison of the models' ability to produce human-like, expert-level thought processes.

2064
2065
2066
2067
2068

Evaluation Limitations. A potential limitation is that CoT from Chem-R can be stylistically distinct, often more structured, due to its protocol-based training. To mitigate any resulting bias, the evaluation was conducted under *strictly blind conditions*. Experts assessed fully anonymized responses without any knowledge of the generating model, ensuring their ratings were based exclusively on the intrinsic quality of the reasoning rather than on stylistic patterns.

Score	Chemical Soundness	Logical Coherence	Step-by-Step Completeness	Justification of the Conclusion	Clarity and Conciseness	Expert-level Insight
5	Excellent: Flawless chemistry, all principles correctly applied.	Excellent: Perfectly logical and sequential flow.	Excellent: All critical steps are explicit and thorough.	Excellent: Correct conclusion is fully justified by the reasoning.	Excellent: Exceptionally clear, concise, and precise.	Excellent: Deep, nuanced understanding that mimics expert thinking.
4	Good: Minor, inconsequential chemical inaccuracy.	Good: Sound logic with a minor unclear transition.	Good: Omits a minor, obvious step.	Good: Correct conclusion, but the link to reasoning is slightly weak.	Good: Clear and easy to understand, but could be more polished.	Good: Strong competence, correctly identifies key factors.
3	Acceptable: Noticeable chemical error(s), but mostly correct.	Acceptable: Generally follows, but has some logical gaps.	Acceptable: Skips one or two important steps.	Acceptable: Reasoning is sound but leads to a minor error, OR the conclusion is not fully supported.	Acceptable: Generally understandable, but somewhat verbose or awkward.	Acceptable: Correct but superficial or "textbook" reasoning.
2	Poor: Significant misunderstanding of core concepts.	Poor: Disjointed and hard to follow.	Poor: Skips multiple crucial steps, making logical leaps.	Poor: Incorrect conclusion that is not supported by the reasoning.	Poor: Vague, confusing, and difficult to parse.	Poor: Misses key nuances and only addresses the obvious.
1	Very Poor: Built on fundamentally flawed chemical principles.	Very Poor: Largely illogical and chaotic.	Very Poor: A brief sketch missing almost all necessary logic.	Very Poor: Conclusion seems unrelated to the reasoning provided.	Very Poor: Nearly unintelligible.	Very Poor: Demonstrates a naive or simplistic understanding.
0	No Value: Completely incorrect chemistry.	No Value: No logical structure.	No Value: Provides only the final answer.	No Value: No connection between reasoning and answer.	No Value: Gibberish.	No Value: Shows a complete lack of understanding.

Figure 5: Human evaluation rubric for Chain-of-Thought quality. Experts are to score the generated reasoning on a 0-5 scale (0=worst, 5=best) across the six metrics provided: Chemical Soundness, Logical Coherence, Completeness, Justification, Clarity, and Expert-level Insight.

2096
2097
2098

Table 5: **Human evaluation of model-generated reasoning.** For each column the **best** and **second-best** models are highlighted.

2099
2100
2101
2102
2103
2104
2105

Metric	Chem-R	Gemini-2.5 Pro	DeepSeek-R1	ether0
Chemical Soundness	4.75	3.95	3.45	2.15
Logical Coherence	4.85	4.25	3.80	2.35
Step-by-Step Completeness	4.20	3.85	3.90	1.95
Justification of the Conclusion	4.28	4.10	3.55	2.05
Clarity and Conciseness	4.65	4.55	3.70	2.55
Expert-level Insight	4.55	3.75	3.20	1.85

2106 **D EXPERIEMT RESULT**
21072108 In this section, we will present the specific results of each subtask in the experiment to provide a
2109 better demonstration of the model’s performance.
21102111 **D.1 NAME PREDICTION**
21122113 In our name prediction task, we evaluated the model’s performance on two sub-tasks: SMILES
2114 to IUPAC name translation (SMILES2IUPAC) and IUPAC name to SMILES translation (IU-
2115 PAC2SMILES). For both sub-tasks, we use the **exact match accuracy** as the evaluation metric.
2116 [All SMILES comparisons are performed after canonicalization to ensure consistency.](#)
2117

2118 Model	2119 SMILES2IUPAC↑	2120 IUPAC2SMILES↑	2121 Average↑
<i>Task-specific specialist models</i>			
2120 STOUT (Rajan et al., 2021)	2121 0.55	2122 0.70	2123 0.63
<i>LLM-based generalist models</i>			
2123 GPT-4o	2124 0.00	2125 0.02	2126 0.01
2124 Gemini-2.5-Pro	2125 0.01	2126 0.33	2127 0.17
2125 DeepSeek-R1	2126 0.01	2127 0.09	2128 0.05
2126 ChemDFM-v1.5-8B	2127 0.06	2128 0.22	2129 0.14
2127 ether0-24B	2128 0.00	2129 0.29	2130 0.15
2128 Chem-R-8B	2129 0.51	2130 0.47	2131 0.49

2129 Table 6: Accuracy scores in name prediction tasks. The task-specific specialist models are sourced
2130 from (Guo et al., 2023).
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159

2160 D.2 PROPERTY PREDICTION
2161

2162 In the molecule property prediction task, we evaluated the model’s performance across a suite of
2163 benchmark datasets: BACE, BBBP, ClinTox, HIV, and Tox21. Each task is formulated as a binary
2164 classification problem. For evaluation, we uniformly use classification accuracy as the sole metric
2165 across all datasets. To further verify the competitiveness of our approach against specialized geo-
2166 metric deep learning models, we additionally evaluate a canonical Graph Neural Network (GNN)
2167 baseline, Chempool (D-MPNN) (Heid et al., 2023), using classification accuracy.
2168

2169 Table 7: AUC-ROC scores of different models in molecular property prediction tasks. **Avg.:** average.
2170 The task-specific specialist models are sourced from (Zhao et al., 2025c).
2171

Model	BACE↑	BBBP↑	ClinTox↑	HIV↑	Tox21↑	Avg.↑
<i>Task-specific specialist models</i>						
Uni-Mol (Zhou et al., 2023)	0.86	0.73	0.92	0.81	0.80	0.82
MolXPT (Liu et al., 2023b)	0.88	0.80	0.95	0.78	0.77	0.84
InstructMol (Cao et al., 2023)	0.86	0.64	-	0.74	-	-
<i>LLM-based generalist models</i>						
ChemDFM-v1.5-8B	0.78	0.75	0.69	0.88	0.74	0.77
ether0-24B	0.58	0.64	0.62	0.67	0.51	0.60
Chem-R-8B	0.78	0.81	0.95	1.00	0.71	0.85

2181
2182
2183 Table 8: Accuracy scores of different models in molecular property prediction tasks. We include
2184 Chemprop (Heid et al., 2023) as a specialized GNN baseline for comparison.
2185

Model	BACE↑	BBBP↑	ClinTox↑	HIV↑	Tox21↑	Avg.↑
<i>Specialized GNN Baseline</i>						
Chempool (D-MPNN)	0.66	0.72	0.84	0.85	0.83	0.78
<i>General Foundation Models</i>						
GPT-4o	0.34	0.61	0.98	1.00	0.49	0.68
Gemini-2.5-Pro	0.36	0.78	0.30	1.00	0.36	0.56
DeepSeek-R1	0.38	0.63	0.8	1.00	0.34	0.63
<i>Chemical LLMs</i>						
ChemDFM-v1.5-8B	0.77	0.57	0.69	0.85	0.81	0.74
ether0-24B	0.43	0.48	0.74	0.82	0.72	0.64
Chem-R-8B	0.74	0.82	0.98	1.00	0.80	0.87

2199 As shown in Table 8, Chem-R not only remains competitive but outperforms the strong Chempool
2200 baseline on average accuracy (0.87 vs. 0.78). This demonstrates the power of our reasoning-based
2201 approach, which matches or exceeds specialized geometric deep learning models while offering
2202 superior interpretability.
2203

2204 D.3 MOLECULE DESIGN

2205 In the text-based molecule design task, we evaluate the model’s ability to generate a correct SMILES
2206 string from a given textual description of a molecule, using the ChEBI-20 dataset. The evaluation
2207 is multi-faceted. First, we measure the fundamental **Validity** of the outputs, which is the percentage
2208 of generated SMILES that represent chemically valid molecules as verified by RDKit. To assess
2209 textual fidelity against the ground-truth SMILES, we employ several string-based metrics: **Exact**
2210 **Match (EM)** for identical strings, **Levenshtein Distance (Lev.)** to measure edit distance (lower is
2211 better), and the **BLEU score**, which quantifies n-gram overlap via the formula:
2212

$$BLEU = BP \cdot \exp \left(\sum_{n=1}^N w_n \log p_n \right) \quad (3)$$

where BP is the Brevity Penalty and p_n is the modified n-gram precision. Finally, to evaluate structural correctness, we calculate the **Tanimoto coefficient** between the molecular fingerprints (MACCS, RDKit, and Morgan) of the generated and ground-truth molecules, where a higher similarity score indicates a greater structural resemblance. We choose MolT5-large (Edwards et al., 2022), MolReGPT (Li et al., 2024b), Mol-Instruction (Fang et al., 2023), MolReasoner (Zhao et al., 2025a) and Mol-R1 (Li et al., 2025b) as the task-specific specialist models.

Model	BLEU↑	EM↑	Lev.↓	MACCS↑	RDK↑	Morgan↑	Validity↑
<i>Task-specific specialist models</i>							
MolT5-large	0.85	0.31	16.07	0.83	0.74	0.68	0.91
MolReGPT	0.86	0.28	17.14	0.90	0.80	0.74	0.90
Mol-Instruction	0.30	0.04	39.42	0.44	0.29	0.25	1.00
MolReasoner	0.78	0.08	26.93	0.68	0.44	0.36	0.97
Mol-R1	0.64	0.23	32.94	0.82	0.68	0.61	0.85
<i>LLM-based generalist models</i>							
GPT-4o	0.45	0.07	48.38	0.79	0.58	0.50	0.77
Gemini-2.5-Pro	0.69	0.29	130.26	0.95	0.88	0.82	0.91
DeepSeek-R1	0.51	0.22	169.36	0.92	0.82	0.75	0.78
ChemDFM-v1.5-8B	0.90	0.55	7.12	0.94	0.88	0.84	0.98
ether0-24B	0.39	0.27	860.99	0.82	0.70	0.64	0.73
Chem-R-8B	0.84	0.41	17.42	0.92	0.83	0.78	0.94

Table 9: Performance of different models on the text-based molecule generation task on the ChEBI-20 dataset. **BLEU**: Bilingual Evaluation Understudy, **EM**: Exact Match, **Lev.**: Levenshtein distance, **MACCS**: MACCS fingerprint similarity, **RDK**: RDKit fingerprint similarity, **Morgan**: Morgan fingerprint similarity, **Validity**: Percentage of valid molecules.

D.4 MOLECULE CAPTIONING

In the molecule captioning task, we evaluate the model’s ability to generate an accurate and fluent natural language description from a given molecular structure (SMILES string), using the ChEBI-20 dataset. To comprehensively assess the quality of the generated text, we employ a suite of standard metrics. We use the **BLEU** score to measure n-gram precision, specifically reporting **BLEU-2** ($N = 2$, $w_1 = w_2 = 0.5$) and **BLEU-4** ($N = 4$, $w_n = 0.25$), based on the general formula:

$$BLEU = BP \cdot \exp \left(\sum_{n=1}^N w_n \log p_n \right) \quad (4)$$

where BP is the Brevity Penalty and p_n is the modified n-gram precision. For a recall-oriented evaluation, we utilize the **ROUGE** family of metrics. We report the F1-scores for **ROUGE-1 (R-1)** and **ROUGE-2 (R-2)**, which measure unigram and bigram overlap, and **ROUGE-L (R-L)**, which is based on the longest common subsequence (LCS). The ROUGE-L F-score is computed as:

$$ROUGE-L_{f\text{-score}} = \frac{(1 + \beta^2)R_{lcs}P_{lcs}}{R_{lcs} + \beta^2P_{lcs}} \quad (5)$$

where R_{lcs} and P_{lcs} are the LCS-based recall and precision, and β is set to 1 to weigh recall and precision equally for the F1-score. Lastly, we incorporate the **METEOR** score, which enhances evaluation by considering synonymy and stemming. It is based on the harmonic mean of unigram precision (P) and recall (R), F_{mean} , which weights recall more than precision:

$$F_{mean} = \frac{10PR}{R + 9P} \quad (6)$$

The final score is calculated by applying a fragmentation penalty (Pen) to this value:

$$METEOR = F_{mean} \cdot (1 - Pen) \quad (7)$$

where Pen is a penalty for fragmentation based on the alignment of chunks between the generated and reference texts.

Model	BLEU-2↑	BLEU-4↑	R-1↑	R-2↑	R-L↑	METEOR↑
<i>Task-specific specialist models</i>						
MolT5-large	0.59	0.51	0.65	0.51	0.59	0.61
MolReGPT	0.61	0.53	0.63	0.48	0.56	0.61
MolReFlect	0.68	0.61	0.70	0.57	0.64	0.68
Mol-Instruction	0.10	0.07	0.28	0.18	0.26	0.19
MolReasoner	0.44	0.32	0.55	0.37	0.48	0.48
<i>LLM-based generalist models</i>						
GPT-4o	0.04	0.01	0.13	0.02	0.10	0.11
Gemini-2.5-Pro	0.09	0.04	0.20	0.06	0.13	0.31
DeepSeek-R1	0.10	0.04	0.14	0.03	0.11	0.15
ChemDFM-v1.5-8B	0.33	0.28	0.46	0.36	0.43	0.40
ether0-24B	0.01	0.00	0.05	0.01	0.04	0.05
Chem-R-8B	0.48	0.41	0.61	0.44	0.53	0.58

Table 10: Performance of different models on the molecule description task on the ChEBI-20 dataset. **R-1**: ROUGE-1 (Recall-Oriented Understudy for Gisting Evaluation-1), **R-2**: ROUGE-2, **R-L**: ROUGE-L (ROUGE-L stands for longest common subsequence), MTEOR: METEOR (Metric for Evaluation of Translation with Explicit ORdering).

D.5 TEXT-BASED OPEN MOLECULE GENERATION

In the text-based open molecule generation task, we evaluate the model’s ability to perform complex chemical reasoning and creative design, using the TOMG-Bench benchmark (Li et al., 2024a). The evaluation is structured around three distinct tasks designed to probe different capabilities: **Molecule Editing (MolEdit)**, **Molecule Optimization (MolOpt)**, and **Customized Molecule Generation (MolCustom)**. And we choose MolT5 (Edwards et al., 2022), BioT5 (Pei et al., 2023) and Open-MolIns (Li et al., 2024a) (which is trained on the full set of training data of TOMG-Bench) for the task-specific specialist models.

For the **MolEdit** and **MolOpt** tasks, which involve modifying an existing molecule, the assessment is threefold. First, we measure the **Success Rate (SR)** to verify if the model’s output correctly fulfills the textual instruction. Second, to ensure the modification is a rational and localized edit rather than a completely new structure, we calculate the Tanimoto **Similarity (Sim.)** between the Morgan fingerprints of the original and generated molecules.

For the **MolCustom** task, which requires generating a molecule from scratch (*de novo*), the metrics are adapted. The **Success Rate (SR)** evaluates if the generated molecule adheres to a set of specified structural constraints (e.g., atom counts, bond types). Instead of similarity, we measure **Novelty (Nov.)** to quantify the uniqueness of the generated molecule. The novelty n for a generated molecule m^g is calculated as:

$$n(m^g) = 1 - \frac{\sum_{m^k \in \text{Zinc}} \delta(m^g, m^k)}{|\text{Zinc}|} \quad (8)$$

where $\delta(m^g, m^k)$ is the Tanimoto similarity to a known molecule m^k in the Zinc database.

To provide a single, comprehensive ranking of model performance, the benchmark introduces a **Weighted Success Rate (WSR)**. This metric combines the core success rate with a quality metric relevant to each task—Similarity for MolEdit/MolOpt and Novelty for MolCustom. The WSR for a given subtask t is defined as:

$$WSR_t = \begin{cases} n_t \times SR_t, & t \in \{\text{MolCustom}\} \\ \delta_t \times SR_t, & t \in \{\text{MolEdit, MolOpt}\} \end{cases} \quad (9)$$

where n_t is the novelty score and δ_t is the similarity score for that subtask. The final performance is then the average WSR across all nine subtasks:

$$WSR = \frac{1}{9} \sum_t WSR_t \quad (10)$$

Finally, for all tasks, we measure the fundamental **Validity (Val.)** to ensure every generated SMILES string represents a chemically sound molecule.

MolEdit	AddComponent			DelComponent			SubComponent		
	SR↑	Sim.↑	Val.↑	SR↑	Sim.↑	Val.↑	SR↑	Sim.↑	Val.↑
<i>Task-specific specialist models</i>									
MolT5	0.28	0.11	0.93	0.22	0.12	0.92	0.17	0.09	0.94
BioT5	0.35	0.16	1.00	0.17	0.16	1.00	0.07	0.16	1.00
OpenMolIns	0.78	0.68	0.95	0.86	0.62	0.91	0.61	0.74	0.96
<i>LLM-based generalist models</i>									
GPT-4o	0.62	0.68	0.74	0.70	0.60	0.85	0.80	0.72	0.94
ChemDFM	0.29	0.65	0.91	0.21	0.69	0.93	0.22	0.64	0.87
Chem-R	0.89	0.69	0.98	0.89	0.62	0.93	0.62	0.75	0.99
MolOpt	LogP			MR			QED		
	SR↑	Sim.↑	Val.↑	SR↑	Sim.↑	Val.↑	SR↑	Sim.↑	Val.↑
<i>Task-specific specialist models</i>									
MolT5	0.42	0.10	0.82	0.45	0.11	0.87	0.47	0.12	0.92
BioT5	0.52	0.15	1.00	0.51	0.16	1.00	0.51	0.16	1.00
OpenMolIns	0.88	0.67	0.93	0.70	0.67	0.94	0.86	0.67	0.93
<i>LLM-based generalist models</i>									
GPT-4o	0.72	0.66	0.88	0.69	0.64	0.84	0.40	0.62	0.86
ChemDFM	0.30	0.66	0.76	0.25	0.68	0.82	0.31	0.66	0.81
Chem-R	0.90	0.66	0.97	0.92	0.68	0.97	0.74	0.68	0.98
MolCustom	AtomNum			BondNum			FunctionalGroup		
	SR↑	Nov.↑	Val.↑	SR↑	Nov.↑	Val.↑	SR↑	Nov.↑	Val.↑
<i>Task-specific specialist models</i>									
MolT5	0.02	0.71	0.84	0.01	0.56	0.89	0.04	0.61	0.94
BioT5	0.01	0.84	1.00	0.01	0.67	1.00	0.05	0.68	1.00
OpenMolIns	0.12	0.68	0.85	0.12	0.67	0.90	0.35	0.64	0.95
<i>LLM-based generalist models</i>									
GPT-4o	0.20	0.67	0.59	0.07	0.63	0.86	0.23	0.65	0.86
ChemDFM	0.01	0.68	0.67	0.02	0.61	0.90	0.03	0.60	0.81
Chem-R	0.22	0.69	0.85	0.10	0.70	0.70	0.31	0.64	0.93

Table 11: Detailed results on TOMG-Bench for different models. The indicators are: **SR** = Success Rate, **Sim.** = Similarity, **Nov.** = Novelty, **Val.** = Validity. MolT5 refers to MolT5-large, BioT5 refers to BioT5-base, OpenMolIns refers to the performance of the Llama-3.1-8B model trained on the largest instruction fine-tuning dataset OpenMolIns-xlarge proposed by TOMG-Bench, and ChemDFM refers to ChemDFM-v1.5-8B. The task-specific specialist models are sourced from (Li et al., 2024a).

Due to the extensive size of the TOMG-Bench (45,000 test instances across 9 tasks), evaluating the full set with closed-source, API-based models like Gemini and DeepSeek incurs prohibitive costs. To provide a direct comparison, we performed an evaluation on a 5% stratified random sample of the TOMG test set. On this subset, Chem-R achieves the highest overall weighted accuracy (wAcc), demonstrating superior performance against leading API-based models, particularly in complex editing and optimization tasks. The detailed results are presented in Table 12.

Table 12: Performance comparison on a 5% stratified random sample of the TOMG-Bench test set. The metric here is accuracy.

Models	MolCustom (Atom / Bond / Func)	MolEdit (Add / Del / Sub)	MolOpt (LogP / QED / MR)	wAcc
Gemini-2.5	0.420 / 0.250 / 0.550	0.780 / 0.820 / 0.680	0.850 / 0.680 / 0.880	0.4089
DeepSeek-R1	0.360 / 0.210 / 0.480	0.820 / 0.790 / 0.650	0.820 / 0.650 / 0.840	0.3921
Chem-R	0.260 / 0.260 / 0.472	0.912 / 0.916 / 0.652	0.912 / 0.728 / 0.932	0.4523

2376 D.6 YIELD PREDICTION
2377

2378 In the reaction yield prediction task, we evaluate the model’s performance on two reaction datasets:
2379 Buchwald-Hartwig and Suzuki-Miyaura. The task is framed as a binary classification problem to
2380 predict whether a reaction yield is high or low. We use **classification accuracy** as the sole evaluation
2381 metric, with the average accuracy across both datasets also reported (as shown in Table 13).

2383 Model	2384 Buchwald-Hartwig↑	2385 Suzuki-coupling↑	2386 Average↑
<i>Task-specific specialist models</i>			
2385 UAGNN (Kwon et al., 2022)	2386 0.97	2387 0.96	2388 0.96
<i>LLM-based generalist models</i>			
2389 GPT-4o	2390 0.20	2391 0.20	2392 0.20
2390 Gemini-2.5-Pro	2391 0.23	2392 0.47	2393 0.35
2391 DeepSeek-R1	2392 0.20	2393 0.45	2394 0.33
2392 ChemDFM-v1.5-8B	2393 0.35	2394 0.38	2395 0.37
2393 ether0-24B	2394 0.02	2395 0.03	2396 0.03
2394 Chem-R-8B	2395 0.87	2396 0.85	2397 0.86

2393 Table 13: Accuracy scores of different models in yield prediction tasks. The task-specific specialist
2394 models is sourced from (Zhao et al., 2025c).

2396 D.7 REAGENT SELECTION

2397 In the reagent selection task, we utilize the Suzuki High-Throughput Experimentation (HTE) dataset.
2398 This task is divided into three sub-tasks: predicting the correct reactant, solvent, and ligand for a
2399 given reaction. The evaluation metrics vary by sub-task. For reactant and solvent prediction, we
2400 report the **top-1 accuracy**. For ligand prediction, we report the **top-5 accuracy**, which considers a
2401 prediction correct if the ground-truth ligand is among the top five candidates proposed by the model.
2402 The ‘Avg.’ column in the table represents the average of these three accuracy scores.

2405 Model	2406 Reagent Selection				2407 Reaction	2408 Retrosyn-
	2409 Reactant	2410 Solvent	2411 Ligand	2412 Avg.	2413 Prediction	2414 thesis
<i>Task-specific specialist models</i>						
2409 Chemformer	2410 –	2411 –	2412 –	2413 –	2414 0.94	2415 0.54
2410 Mol-Instruction	2411 –	2412 –	2413 –	2414 –	2415 0.05	2416 –
2411 InstructMol	2412 –	2413 –	2414 –	2415 –	2416 0.54	2417 –
2412 ochem_predict_nn	2413 –	2414 –	2415 –	2416 –	2417 0.72	2418 –
2413 AiZynthFinder	2414 –	2415 –	2416 –	2417 –	2418 –	2419 0.47
<i>LLM-based generalist models</i>						
2414 GPT-4o	2415 0.48	2416 0.11	2417 0.19	2418 0.26	2419 0.04	2420 0.00
2415 Gemini-2.5-Pro	2416 0.17	2417 0.07	2418 0.56	2419 0.27	2420 0.35	2421 0.15
2416 DeepSeek-R1	2417 0.15	2418 0.17	2419 0.08	2420 0.13	2421 0.34	2422 0.13
2417 ChemDFM-v1.5-8B	2418 0.47	2419 0.10	2420 0.49	2421 0.35	2422 0.50	2423 0.07
2418 ether0-24B	2419 0.11	2420 0.35	2421 0.18	2422 0.21	2423 0.65	2424 0.04
2419 Chem-R-8B	2420 0.62	2421 0.57	2422 0.87	2423 0.69	2424 0.82	2425 0.39

2421 Table 14: Performance of task-specific specialist models and LLM-based generalist models on
2422 reagent selection, reaction prediction, and retrosynthesis tasks. The task-specific specialist
2423 models are sourced from (Zhao et al., 2025c). “–” means that the model was not designed for the task.
2424 We also include template-based (non-LLM based) baselines ochem_predict_nn (Coley et al., 2017)
2425 and AiZynthFinder (Genheden et al., 2020) for comparison.

2426 D.8 REACTION PREDICTION

2427 In the reaction prediction task, we evaluate the model’s ability to predict the major product of a
2428 chemical reaction, using the USPTO_Mixed dataset. Performance is measured using **exact match**

2430 **accuracy.** All SMILES comparisons are performed after canonicalization to ensure consistency.
 2431 We additionally benchmark against the specialized template-based model `ochem_predict_nn` (Coley
 2432 et al., 2017). And we choose Chemformer (Irwin et al., 2022), Mol-Instruction (Fang et al., 2023)
 2433 and InstructMol (Cao et al., 2023) for the task-specific specialist models. As shown in Table 14,
 2434 Chem-R outperforms this baseline (0.82 vs. 0.72), demonstrating that our reasoning-based approach
 2435 yields superior predictive accuracy compared to traditional template matching methods.
 2436

2437 D.9 RETROSYNTHESIS

2438 In the retrosynthesis task, the goal is to predict the reactants that form a given product molecule.
 2439 This is evaluated on the widely-used USPTO-50k dataset. Similar to reaction prediction, we em-
 2440 ploy **exact match accuracy** as the evaluation metric. All SMILES comparisons are performed
 2441 after canonicalization to ensure consistency. We further compare Chem-R with the search-based
 2442 tool `AiZynthFinder` (Genheden et al., 2020). And we choose Chemformer (Irwin et al., 2022) for
 2443 the task-specific specialist models. As shown in Table 14, while the specialized search algorithm
 2444 achieves higher accuracy (0.47 vs. 0.39), it operates as a black-box predictor. In contrast, Chem-R
 2445 provides step-by-step reasoning traces, delivering a competitive level of accuracy while ensuring the
 2446 interpretability and verifiability essential for scientific discovery.
 2447

2448 D.10 OUT OF DOMAIN

2449 To evaluate the out-of-domain (OOD) generalization of our model, Chem-R, we benchmarked its
 2450 performance on the challenging **Molecular Property Optimization** task from ChemCoTBench (Li
 2451 et al., 2025a). We selected four representative targets: Solubility, DRD2, JNK3, and GSK-3 β . The
 2452 evaluation, presented in Table 15, compares Chem-R against its base model, Llama-3.1-8B-Instruct,
 2453 and other powerful LLMs.
 2454

2455 The results clearly demonstrate the effectiveness of our training. The base Llama-3.1-8B-Instruct
 2456 model performs poorly, whereas Chem-R shows a dramatic improvement in both success rate (SR%)
 2457 and mean property improvement (Δ) across all tasks. This signifies that Chem-R has acquired robust
 2458 chemical reasoning skills that generalize effectively. Furthermore, Chem-R proves to be highly com-
 2459 petitive, significantly outperforming the much larger Llama-3.3-70B-Instruct model, which confirms
 2460 the strong OOD capabilities of our approach.
 2461

2462 Models	2463 Solubility		2464 DRD2		2465 JNK3		2466 GSK3- β	
	2467 Δ	2468 SR%	2469 Δ	2470 SR%	2471 Δ	2472 SR%	2473 Δ	2474 SR%
Qwen3-235B-A22B-think	0.27	42	0.03	31	-0.01	23	0.01	31
Qwen3-32B-think	0.11	23	0.00	6	-0.02	6	-0.02	5
Llama-3.1-8B-Instruct	0.01	10	-0.07	2	-0.10	4	-0.08	3
Chem-R	0.34	83	0.01	36	-0.02	24	-0.01	29

2469 Table 15: Performance of various models on different molecular optimization tasks. Evaluation was
 2470 conducted on one physicochemical property (Solubility) and three more challenging protein activity
 2471 targets (DRD2, JNK3, and GSK3- β). The mean improvement in a property is denoted by Δ ; a
 2472 negative Δ indicates degradation of the property. The success rate (SR%) represents the percentage
 2473 of optimizations that led to an increase in the target property.
 2474
 2475
 2476
 2477
 2478
 2479
 2480
 2481
 2482
 2483

2484 D.11 EFFECTIVENESS OF CRP AS A PROMPTING STRATEGY
2485

2486 To further decouple the contribution of our reasoning protocols from the model training, we eval-
2487 uated GPT-4o and Gemini-2.5-Pro by providing our task-specific Chemical Reasoning Protocols
2488 (CRPs) directly as part of their system prompt (denoted as “Prompted”). As shown in Table 16, the
2489 results yield two important findings. First, our CRPs are highly effective as a standalone prompting
2490 strategy, significantly improving the performance of generalist models (e.g., Prompted GPT-4o im-
2491 proves Name Prediction from 0.01 to 0.10), which validates the quality of our reasoning framework.
2492 Second, despite these gains, our much smaller Chem-R-8B still vastly outperforms these prompted
2493 state-of-the-art models. This proves that our superior performance stems from deeply internaliz-
2494 ing chemical reasoning through the three-phase training, a capability that in-context learning alone
2495 cannot replicate.

2496
2497 **Table 16:** Performance of Chem-R against baselines with and without protocol-guided (in-domain)
2498 prompting. “Prompted” indicates the model was given the CRP in the input context.

2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537	Molecule Tasks					Reaction Tasks			
	Model	Name	Prop.	Design	Capt.	TOMG	Yield	Reag.	React.
GPT-4o	0.01	0.68	0.07	0.01	0.32	0.20	0.26	0.04	0.00
GPT-4o (Prompted)	0.10	0.72	0.07	0.15	0.34	0.25	0.32	0.06	0.01
Gemini-2.5-Pro	0.17	0.56	0.29	0.04	–	0.35	0.27	0.35	0.15
Gemini-2.5-Pro (Prompted)	0.21	0.60	0.33	0.23	–	0.39	0.39	0.38	0.20
Chem-R-8B (Ours)	0.49	0.87	0.42	0.41	0.42	0.85	0.69	0.82	0.39

D.12 IMPACT OF TEACHER MODEL QUALITY

2509 To validate our framework’s robustness, we conducted the suggested experiment by distilling from
2510 a smaller Llama-3.1-8B teacher. The results lead to two insightful conclusions:

2511 (1) Our framework instills a reasoning structure, rather than just imitating. The most compelling evi-
2512 dence is that the Chem-R student model learns to solve complex problems that its own Llama-3.1-8B
2513 teacher could not. On tasks where the teacher scored 0 (e.g., Name Prediction, Reaction Prediction,
2514 and Retrosynthesis), the student achieved impressive scores of 0.39, 0.70, and 0.29, respectively.
2515 This demonstrates our framework unlocks latent capabilities by building a robust reasoning founda-
2516 tion.

2517 (2) The framework’s effectiveness scales with teacher quality. The results also show a strong corre-
2518 lation between teacher capability and the student’s performance ceiling. The Chem-R model distilled
2519 from the larger Llama-3.3-70B teacher consistently outperforms the student distilled from the 8B
2520 teacher. This confirms that while our framework is universally effective, a better teacher produces a
2521 better student. The detailed results are shown in Table 17.

2522
2523 **Table 17:** Performance of Chem-R-8B distilled from different teacher models (Llama-3.1-8B vs.
2524 Llama-3.3-70B) against various baselines. The results highlight that the Chem-R framework enables
2525 a student model to significantly outperform its teacher, especially on complex reasoning tasks where
the teacher scores zero.

2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537	Molecule Tasks					Reaction Tasks			
	Model	Name	Prop.	Design	Capt.	TOMG	Yield	Reag.	React.
Llama-3.1-8B-Instruct	0.00	0.47	0.00	0.01	0.07	0.26	0.26	0.00	0.00
Llama-3.3-70B	0.01	0.64	0.03	0.02	0.30	0.22	0.38	0.03	0.00
DeepSeek-R1	0.05	0.63	0.22	0.04	–	0.33	0.13	0.34	0.13
ether0-24B	0.15	0.64	0.30	0.03	0.03	0.03	0.21	0.65	0.04
Chem-R-8B (Distilled from Llama-3.1-8B)	0.39	0.75	0.31	0.36	0.34	0.78	0.50	0.70	0.29
Chem-R-8B (Ours, Distilled from Llama-3.3-70B)	0.49	0.87	0.42	0.41	0.42	0.85	0.69	0.82	0.39

2538 **D.13 COMPUTATIONAL EFFICIENCY ANALYSIS**
25392540 In this section, we provide a full computational cost analysis. However, we must first clarify that
2541 a direct comparison between a versatile LLM and traditional, specialized ML models is often mis-
2542 leading due to two key factors:2543

2544 - **Multi-Task Amortization:** Our single model is trained once to handle 25 distinct tasks. This
2545 amortizes the cost compared to the cumulative effort of developing and maintaining numerous
2546 separate, specialized models.
2547 - **The Cost of Interpretability:** Our model generates detailed reasoning chains, which are in-
2548 herently more computationally intensive than the single, black-box predictions from specialized
2549 models. This is a deliberate trade-off for the crucial scientific value of explainability and trust.

2550 **Training Cost.** The total training time for all three phases of Chem-R was approximately **158**
2551 hours on eight H200 GPUs (Phase 1: ~36h, Phase 2: ~81h, Phase 3: ~41h).2552 **Inference Cost.** A more meaningful comparison is against other LLMs. To provide a standard-
2553 ized measure of inference cost, we report the average number of generated tokens per response for
2554 each model across all tasks in Table 18. As the table shows, Chem-R generates responses with a
2555 token count that is comparable to other high-performing models and is significantly more concise
2556 than verbose reasoners like Gemini-2.5-Pro and DeepSeek-R1. This demonstrates that our frame-
2557 work achieves state-of-the-art accuracy without resorting to excessively long or inefficient reasoning
2558 chains, balancing performance with practical efficiency.2559
2560 **Table 18:** Inference cost comparison (average generated tokens per response). Chem-R maintains a
2561 reasonable token count, demonstrating efficient reasoning compared to other high-performing rea-
2562 soning models.2563

Model	Name	Prop.	Design	Capt.	TOMG	Yield	Reag.	React.	Retro
GPT-4o	572	339	499	489	413	522	562	465	518
Llama-3.1-8B-Instruct	1637	431	3031	984	1062	655	574	428	431
Llama-3.3-70B-Instruct	631	374	572	465	592	492	620	510	412
Gemini-2.5-Pro	1600	1348	1475	1926	1605	1658	1679	1422	1479
DeepSeek-R1	12265	5183	10568	10188	5581	4415	8915	8964	10096
QwQ-32B	8041	2390	7128	5392	5516	2986	3146	7669	7039
ChemLLM-20B-Chat	279	339	171	452	397	335	370	352	563
ChemDFM-v1.5-8B	38	1	58	58	94	1	15	29	42
ChemDFM-v1.0-13B	38	1	52	60	70	1	17	32	41
ether0	898	501	1295	782	700	607	850	811	982
Chem-R	659	478	560	675	526	573	619	505	552

2592 E MORE CASES
25932594 To provide a deeper qualitative analysis of Chem-R’s performance, we present additional case studies
2595 comparing its reasoning against ablation variants, accompanied by expert chemist comments. In
2596 the following examples, **green text** highlights correct, chemically sound reasoning, while **red text**
2597 indicates factual errors, hallucinations, or flawed logic.
25982599 **Reaction Prediction** (Chem-R generated response)
26002601 You should put your reasoning in `<think> </think>` tags. The final answer MUST BE
2602 put in `<answer> </answer>` tags. Please strictly follow the format.
26032604 Note: If multiple products are predicted, they MUST be separated by a period ‘.’ instead of
2605 commas.
26062607 Now predict the product for the following reaction:
26082609 Reactants: C1CCOC1.CC(C)(C)[O-].O=C(Cl)c1cccc([N+](=O)[O-])c1.[K+]
26102611 **Ground Truth:** CC(C)(C)OC(=O)c1cccc([N+](=O)[O-])c1
26122613 **LLM Answer:**
26142615 `<think>`2616 1. **Reactant analysis:** The reactant molecules are represented by their SMILES strings,
2617 indicating a complex mixture. One of the reactants is **an ether (C1CCOC1)**, suggest-
2618 ing the presence of an oxygen atom connected to two alkyl groups. Another reactant
2619 is a tertiary alkoxide (CC(C)(C)[O-]), which is **a strong nucleophile**. The presence of
2620 O=C(Cl)c1cccc([N+](=O)[O-])c1 indicates an aromatic ring with a nitro group and **an acyl**
2621 **chloride, which is highly reactive due to the chlorine atom attached to the carbonyl group**.
2622 The potassium ion [K+] suggests that there might be an ionic interaction or a potential for a
2623 nucleophilic substitution reaction.
26242625 2. **Functional group identification:** Key functional groups identified include the ether
2626 (COC), chlorine (Cl), and nitro (N(=O)[O-]) groups. These groups can serve as reaction
2627 sites for various chemical transformations. **The ether group might not be directly involved**
2628 **in the primary reaction due to its relatively stable nature**, but it could influence the reaction
2629 conditions or the solubility of the reactants. **The chlorine atom attached to the carbonyl group**
2630 **is highly reactive and susceptible to nucleophilic attack**, which could lead to the formation
2631 of a new bond.
26322633 3. **Reaction type analysis:** Given the presence of a strong nucleophile (the alkoxide ion)
2634 and a highly reactive acyl chloride, a **nucleophilic acyl substitution reaction is likely**. This
2635 type of reaction involves the **nucleophile attacking the carbonyl carbon of the acyl chloride**,
2636 **leading to the replacement of the chlorine atom with the nucleophile**. The nitro group on
2637 the aromatic ring might influence the electron density on the ring, potentially affecting the
2638 reaction rate or the stability of the product, but it is not directly involved in the primary
2639 reaction site.
26402641 4. **Product prediction:** Based on the identified reaction type, **the most likely product formed**
2642 **is the result of the nucleophilic substitution of the chlorine atom in the acyl chloride by the**
2643 **alkoxide ion**. This would lead to **the formation of an ester**. The predicted product SMILES
2644 string, considering the substitution of the chlorine with an alkoxide group (which would then
2645 be part of an ester linkage), is CC(C)(C)OC(=O)c1cccc([N+](=O)[O-])c1.
26462647 `</think>`
26482649 `<answer>CC(C)(C)OC(=O)c1cccc([N+](=O)[O-])c1</answer>`2650 **Comments by chemists:** *Spot on. If I were analyzing this, I would also immediately dismiss*
2651 *THF as the solvent and focus on the nucleophilic acyl substitution. The model correctly pairs*
2652 *the tert-butoxide with the acid chloride, leading straight to the correct ester product.*

2646
2647**Reaction Prediction (w/o CRP)**2648
2649

Problem: You are an expert chemist. Your task is to solve the given problem step by step. You should put your reasoning in `<think> </think>` tags. The final answer MUST BE put in `<answer> </answer>` tags. Please strictly follow the format.

2650
2651

Note: If multiple products are predicted, they MUST be separated by a period ‘.’ instead of commas.

2652
2653

Now predict the product for the following reaction:

2654

Reactants: C1CCOC1.CC(C)(C)[O-].O=C(Cl)c1cccc([N+](=O)[O-])c1.[K+]

2655
2656

Ground Truth: CC(C)(C)OC(=O)c1cccc([N+](=O)[O-])c1

2657
2658**LLM Answer:**

`<think>`

To solve this problem, we first need to identify the reactants and their possible roles in the reaction. The given reactants are C1CCOC1 (tetrahydrofuran, the solvent), CC(C)(C)[O-] with [K+] (potassium tert-butoxide), and O=C(Cl)c1cccc(N+[O-])c1 (a chloroformylated pyridine with a nitro group).

2659
2660

The key reactant is the potassium tert-butoxide. It is known as a **strong, non-nucleophilic, bulky base**. Its primary role in organic chemistry is to promote elimination reactions over substitution reactions. Given its strong basicity, the most likely first step is an **E2 elimination**. The tert-butoxide will abstract a proton from the aromatic ring, likely **ortho** to the carbonyl group, which is the most acidic position. This process will lead to the **expulsion of the chloride ion and the formation of a highly reactive benzyne intermediate**. This benzyne intermediate is then rapidly attacked by a molecule of tert-butoxide. This addition step, followed by protonation from trace amounts of water or other proton sources during workup, results in the final product. This mechanism ensures that the bulky base does not need to directly attack the sterically hindered carbonyl carbon. The final product forms by replacing the chloride with the tert-butoxy group.

2661
2662

`</think>`

`<answer>CC(C)(C)OC(=O)c1cccc([N+](=O)[O-])c1</answer>`

2663
2664

Comments by chemists: *The final answer is correct, but the reasoning is wild. If I were solving this, I would simply see a standard nucleophilic acyl substitution. The model hallucinates a benzyne mechanism that is chemically impossible here. It's a classic case of "right answer, wrong reason."*

2665
26662667
26682669
26702671
26722673
26742675
26762677
26782679
26802681
26822683
26842685
26862687
26882689
26902691
2692
2693
2694
2695
2696
2697
2698
2699

2700
2701**Reaction Prediction (w/o correct information)**2702
2703

Problem: You are an expert chemist. Your task is to solve the given problem step by step. You should put your reasoning in `<think> </think>` tags. The final answer MUST BE put in `<answer> </answer>` tags. Please strictly follow the format.

2704
2705

Note: If multiple products are predicted, they MUST be separated by a period ‘.’ instead of commas.

2706
2707

Now predict the product for the following reaction:

2708

Reactants: C1CCOC1.CC(C)(C)[O-].O=C(Cl)c1cccc([N+](=O)[O-])c1.[K+]

2709
2710

Ground Truth: CC(C)(C)OC(=O)c1cccc([N+](=O)[O-])c1

2711
2712**LLM Answer:**

`<think>`

2713
2714
2715
2716
2717

1. Reactant analysis: The reactants are represented by their SMILES strings. ‘C1CCOC1’ is tetrahydrofuran (THF), a common aprotic polar solvent. ‘CC(C)(C)[O-]’ and ‘[K+]’ together form potassium tert-butoxide, a strong base and a sterically hindered nucleophile. ‘O=C(Cl)c1cccc([N+](=O)[O-])c1’ is an aromatic compound with a nitro group and an acyl chloride group, acting as the electrophile.

2718
2719
2720
2721
2722
2723

2. Functional group identification: The key functional groups include the acyl chloride (-COCl) and the nitro group (-NO₂). Typically, the carbonyl carbon of an acyl chloride is a very strong electrophilic site. However, due to a misinterpretation of the SMILES string O=C(Cl)c1cccc([N+](=O)[O-])c1, the connectivity is incorrectly perceived. It is mistakenly assumed that the chlorine atom (Cl) is directly attached to the aromatic ring (c1) and the carbonyl group (C=O) is a separate ketone substituent. Aryl chlorides are significantly less reactive towards nucleophilic substitution than acyl chlorides.

2724
2725
2726
2727
2728
2729

3. Reaction type analysis: Based on the incorrect structural interpretation above, a nucleophilic acyl substitution will not occur. Since the molecule is now incorrectly viewed as an aryl chloride and a ketone, the most likely role for the strong base potassium tert-butoxide is as a base rather than a nucleophile. It will attempt to abstract a proton from the most acidic position on the aromatic ring (ortho to the nitro group) to initiate an elimination-addition (benzyne) mechanism.

2730
2731
2732
2733
2734

4. Product prediction: Following this incorrect mechanism, the potassium tert-butoxide acts as a strong base to deprotonate the carbon between the nitro and chloro groups, followed by the departure of the chloride ion to form a benzyne intermediate. The tert-butoxide anion then adds to this benzyne. This ultimately results in the replacement of the chlorine atom on the ring by the tert-butoxy group.

2735
2736

`</think>`
`<answer>CC(C)(C)Oc1cccc([N+](=O)[O-])c1</answer>`

2737
2738
2739
2740

Comments by chemists: This fails right at the start. The model completely misinterprets the SMILES, mistaking an acyl chloride for an aryl chloride. If I were reading this, the reactive -COCl group is obvious. This structural error forces the reasoning into a bizarre benzyne pathway, resulting in the wrong product entirely.

2741
2742
2743
2744
2745
2746
2747
2748
2749
2750
2751
2752
2753

2754
2755**Molecule Design** (Chem-R generated response)2756
2757

Problem: You are an expert chemist. Your task is to solve the given problem step by step. You should put your reasoning in `<think> </think>` tags. The final answer MUST BE put in `<answer> </answer>` tags. Please strictly follow the format.

2758

Now predict the SMILES representation for the following molecular design requirement:

2759

Description: The molecule is an N-carbamoylamino acid that is aspartic acid with one of its amino hydrogens replaced by a carbamoyl group. It has a role as a *Saccharomyces cerevisiae* metabolite, an *Escherichia coli* metabolite and a human metabolite. It is a N-carbamoyl-amino acid, an aspartic acid derivative and a C4-dicarboxylic acid. It is a conjugate acid of a N-carbamoylaspartate(2-).

2760
2761
2762
2763
2764

Ground Truth: C(C(C(=O)O)NC(=O)N)C(=O)O

2765

LLM Answer:

2766

`<think>`

2767
2768
2769
2770
2771
2772

1. **Identify the core structure:** The molecule is described as an **aspartic acid derivative**, which is a type of amino acid. **Aspartic acid** has a core structure that includes a **central carbon atom bonded to an amino group, a carboxylic acid group, and a side chain that also contains a carboxylic acid group**. Given that it's a derivative, we start with this basic amino acid structure.

2773
2774
2775
2776
2777
2778
2779

2. **Summarize key elements:** The molecule is an N-carbamoylaspartic acid, meaning **one of the hydrogens on the amino group of aspartic acid is replaced by a carbamoyl group (NC(=O)N)**. The key elements include: - A central carbon (C) that is part of the amino acid backbone. - An amino group (N) attached to this central carbon, which is then modified. - A carbamoyl group (NC(=O)N) replacing one of the hydrogens on the amino group. - Two carboxylic acid groups, one attached to the central carbon and the other as part of the side chain. - The side chain contains a carboxylic acid group (CC(=O)O).

2780
2781
2782
2783
2784
2785

3. **Address stereochemistry:** The **description does not specify any particular stereochemistry** for the molecule, suggesting that either the stereochemistry is not relevant for the identification of the molecule or it is assumed to be in a specific form by default (e.g., L-form for amino acids). However, since the task focuses on generating a SMILES string based on the functional groups and molecular description without specific stereochemical details, we will **not include explicit stereochemical indicators in our SMILES string**.

2786
2787
2788
2789
2790
2791

4. **Construct SMILES:** To construct the SMILES, we start with the **central carbon (C)** of the **amino acid**, which is **bonded to the amino group modified with the carbamoyl group (NC(=O)N)**, and **two carboxylic acid groups (C(=O)O)** - one directly attached to the central carbon and the other as part of the side chain. **Combining these elements, we get the SMILES string**. The carbamoyl group replaces one of the amino hydrogens, and the side chain contains a carboxylic acid group, leading to the structure: C(C(C(=O)O)NC(=O)N)C(=O)O.

2792
2793

`</think>`

`<answer>C(C(C(=O)O)NC(=O)N)C(=O)O</answer>`

2794
2795
2796

Comments by chemists: *A textbook approach. If I were solving this, I would start with the aspartic acid backbone and directly modify the amine. The model follows this exact logic, smartly filtering out biological distractions to build the correct SMILES.*

2797

2798

2799

2800

2801

2802

2803

2804

2805

2806

2807

2808
2809**Molecule Design (w/o CRP)**2810
2811

Problem: You are an expert chemist. Your task is to solve the given problem step by step. You should put your reasoning in `<think> </think>` tags. The final answer MUST BE put in `<answer> </answer>` tags. Please strictly follow the format.

2812

Now predict the SMILES representation for the following molecular design requirement:

2813

Description: The molecule is an N-carbamoylamino acid that is aspartic acid with one of its amino hydrogens replaced by a carbamoyl group. It has a role as a *Saccharomyces cerevisiae* metabolite, an *Escherichia coli* metabolite and a human metabolite. It is a N-carbamoyl-amino acid, an aspartic acid derivative and a C4-dicarboxylic acid. It is a conjugate acid of a N-carbamoylaspartate(2-).

2814
2815
2816
2817
2818

Ground Truth: C(C(C(=O)O)NC(=O)N)C(=O)O

2819
2820

LLM Answer:

2821

`<think>`

2822
2823
2824
2825
2826

To predict the SMILES representation of the molecule described, we first identify the base structure, which is aspartic acid. The SMILES for aspartic acid can be written by identifying its main chain and substituents. We can represent the backbone as 'C(C(C(=O)O))C(=O)O', with the amino group being a substituent on the alpha-carbon, which we would write as '(N)'.

2827
2828
2829
2830
2831

The description states that a carbamoyl group (-CONH₂) replaces one of the amino hydrogens. **In SMILES notation, when a new group is added, it is often treated as a new, separate substituent on the same parent atom.** Therefore, we will add the carbamoyl group, which can be written as (C(=O)N), as a second substituent on the alpha-carbon. This would initially give us the structure C(C(C(=O)O)(N)(C(=O)N))C(=O)O.

2832
2833
2834
2835
2836
2837
2838
2839
2840

However, this representation is not the most stable or conventional. **There is a widely used rule in SMILES generation for simplifying structures where two nitrogen-containing groups are attached to the same carbon atom. This rule involves an intramolecular rearrangement where the smaller nitrogen group (in this case, the amino group N) attacks the larger group (the carbamoyl group C(=O)N) to form a more stable, single linear substituent.** This "N-to-N bond fusion rule" merges the two parenthesized groups (N)(C(=O)N) into a single, more compact representation, NC(=O)N. By applying this simplification, we replace the two separate substituents with the newly formed single group, leading to the final, correct SMILES string.

2841
2842

`</think>`
`<answer>C(C(C(=O)O)NC(=O)N)C(=O)O</answer>`

2843
2844
2845
2846

Comments by chemists: *The answer is correct, but the reasoning is pure fantasy. If I were solving this, I would start with the aspartic acid backbone and directly modify the amine. Instead, the model oddly attaches substituents to the central carbon and invents a non-existent "fusion rule" to fix the structure.*

2847
2848
2849
2850
2851
2852
2853
2854
2855
2856
2857
2858
2859
2860
2861

2862
2863**Molecule Design (w/o correct information)**2864
2865

Problem: You are an expert chemist. Your task is to solve the given problem step by step. You should put your reasoning in `<think> </think>` tags. The final answer MUST BE put in `<answer> </answer>` tags. Please strictly follow the format.

2866

Now predict the SMILES representation for the following molecular design requirement:

2867

Description: The molecule is an N-carbamoylamino acid that is aspartic acid with one of its amino hydrogens replaced by a carbamoyl group. It has a role as a *Saccharomyces cerevisiae* metabolite, an *Escherichia coli* metabolite and a human metabolite. It is a N-carbamoyl-amino acid, an aspartic acid derivative and a C4-dicarboxylic acid. It is a conjugate acid of a N-carbamoylaspartate(2-).

2872

Ground Truth: C(C(C(=O)O)NC(=O)N)C(=O)O

2873

LLM Answer:

2874

`<think>`

2875

1. **Identify the core structure:** The molecule is a derivative of aspartic acid. The SMILES for the basic aspartic acid backbone can be represented as 'C(C(C(=O)O)O)N)C(=O)O', showing the alpha-carbon connected to a side-chain carbon, a carboxyl group, and an amino group.

2876

2. **Summarize key elements:** The carbamoyl group itself has the chemical structure - C(=O)N-. This group replaces one of the hydrogens on the aspartic acid's original amino group ('-NH₂').

2877

3. **Address stereochemistry:** We need to integrate this modification into the aspartic acid SMILES. The amino group 'N' in the original SMILES 'C(C(C(=O)O)O)N)C(=O)O' will be modified. The hydrogen is replaced by the carbamoyl group. This means the 'N' becomes part of a larger group, which we can write by attaching the SMILES for the carbamoyl group directly to it.

2878

4. **Construct SMILES:** Starting with the nitrogen 'N' of the amino acid, we attach the carbamoyl group '-C(=O)N' to it, resulting in the combined fragment 'NC(=O)N'. We then substitute this entire fragment back into the parenthesis where the original 'N' was located. This yields the final, complete SMILES string.

2879

`</think>`

2880

`<answer>C(C(C(=O)O)NC(=O)N)C(=O)O</answer>`

2881

Comments by chemists: *The answer is correct, but the terminology is sloppy. If I were solving this, I would start with the aspartic acid backbone and directly modify the amine. The model incorrectly defines the isolated carbamoyl group as -C(=O)N- instead of -C(=O)NH2. It arrives at the correct final structure, but lacks chemical rigor in its definitions.*

2882

2883

F USE OF LLMs

2884

During the preparation of this work, the author(s) used LLMs to improve the language and readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

2885

2886

2887

2888

2889

2890

2891

2892

2893

2894

2895

2896

2897

2898

2899

2900

2901