Dear Editor and Reviewers,

Thank you for your valuable feedback. We have made substantial revisions to address all reviewer concerns. Below is a summary of our major changes:

Major Revisions

1. **Reorganized the entire framework** around four domain-specific challenges (Section 2.3, pages 2-3) instead of generic training paradigms, addressing concerns about lack of molecular discovery insights.
2. **Added extensive new content:**
    - Eight comprehensive appendices (A-H, pages 16-27) covering datasets, metrics, tools, and evaluation frameworks
    - Empirical analysis with qualitative comparison (Figure 6, page 24) and quantitative results (Table 2, page 24)
    - Distribution shift and OOD generalization discussion (Appendix G, page 25)
3. **Expanded coverage of missing works and tools:**
    - Included mentioned missing papers
    - Added external tools section (Section 5.3, page 7; Appendix D, pages 21-22)
    - Discussed 2D vs 3D molecular representations (Appendix A, pages 16-17)
4. **Provided concrete performance comparisons:**
    - Added Table 2 (page 24) with quantitative results
    - Included analysis showing SFT performance gains (Appendix F.2, pages 24-25)
    - Added Figure 6 (page 24) showing trade-offs between paradigms
5. **Improved readability:**
    - Restructured content with clear subsections throughout
    - Added visual elements (Figures 1-6)
6. **Clarified scope and future directions:**
    - Clearly defined inclusion criteria (Section 2.2, page 2)
    - Provided three concrete, chemistry-specific future research directions (Section 6, pages 8-9)

These revisions directly address the reviewers' main concerns about domain specificity, missing content, lack of performance comparisons, and readability issues. We believe the revised manuscript now provides a comprehensive, domain-focused resource for researchers at the intersection of LLMs and molecular discovery.

Thank you for your consideration.

Sincerely,

All authors