relevant for applications and limited measuring specificity

ChemBench

LabSafety Bench LabSafety Bench

MaScQA

MolPuzzle 🔀 👸

relevant for applications and specific measurement

LAB-Bench

SciAssess 🔀

SciKnowEval 🚰

CAMEL - Chemistry

ChemIQ 🚭

ChemLLMBench



limited applicability and measuring specificity

end-to-end runable



