

分子对接与分子动力学模拟：经典力学视角与 GROMACS 裁剪实现的理论支撑

- 作者：GaoZheng
- 日期：2025-09-25
- 版本：v1.0.0

摘要

分子对接与分子动力学（MD）并非对立：在经典力学框架下，MD 在势能面上进行随时间演化的连续动力过程，而对接可视为在相同势能面上进行的“静态（或准静态）构型搜索与打分”。因此，可将“对接”理解为“MD 的静态切面”：在给定力场与环境假设下，对受体-配体构型空间进行采样、最小化与排序。基于此，GROMACS 现有的能量计算、邻域搜索、最小化器、能量分组与并行设施，为把“动力模拟退化为对接式搜索”提供了充分的理论与工程支撑。

1. 经典力学框架：分子系统的基本模型

- 状态描述：原子坐标 r 与动量 p ；哈密顿量 $H(r, p) = K(p) + U(r)$ 。
- 运动方程： $F = -\nabla U(r)$, $m \frac{d^2 r}{dt^2} = F$ ；约束（如键长/角度）可由拉格朗日乘子或 SHAKE/LINCS 实现。
- 势能分解： $U = U_{bond} + U_{angle} + U_{dihedral} + U_{vdw} + U_{coul} +$ (可选极化/特殊项)。
- 力场与环境：真空/隐式溶剂/显式溶剂、介电模型、截断与 PME 等共同决定 $U(r)$ 的形状。

结论：无论做 MD 还是对接，核心都是在同一 $U(r)$ 上评估能量与（必要时）梯度。

2. “分子对接是分子动力学的静态切面”

- 对接问题：给定受体 R 与配体 L ，寻求使相互作用能（或经验评分）最优的构型 x （含刚体 6 自由度与可转二面角）。
- 静态化假设：常将受体刚性化或弱柔性；以 $T \rightarrow 0$ 的极限，把“找到稳定结合构型”近似为“在势能面上做最小化/采样并取低能态”。
- 等价视角：
 - MD 视角：在 $U(r)$ 上以时间推进取样；

- 对接视角：在 $U(r)$ 上以（启发式/MC/网格）探索“静态”构型并以 U 或评分 S 排序。
- 评分近似： $S \approx E_{inter}(R, L) + E_{intra}(L) + \eta \cdot \text{惩罚/经验项}$ 。许多对接程序用栅格近似 E_{inter} 以加速评估，本质仍是在势能面上做静态搜索。

3. 统计力学视角：时间平均与构型系综

- MD 以微正则/正则系综采样，观测量取时间平均 \approx 系综平均；
- 对接多取 $T \rightarrow 0$ 的“能量主导”极限，聚焦局部极小及其相对排序；
- 自由能 $\Delta G \approx \Delta H - T\Delta S$ 。多数对接评分忽略显式熵与溶剂效应，或以经验项近似；若用力场直接评分，得到的是势能排序（近似 ΔH ），需谨慎解读与校准。

4. GROMACS 的内在支撑：为何能“退化为对接”

- 能量评估：高效非键相互作用核（nbnxm, CPU/GPU）、约束与 PME/截断等完备能量项；
- 最小化器：steep/L-BFGS，可作为“静态打分的局部细化器”；
- 邻域与并行：Verlet 列表、多模拟（-multidir）有助于“多姿势并行评估”；
- 选择/索引与能量分组：energygrps=Protein Ligand，可直接输出“Protein–Ligand”分项能量用于评分；
- 冻结/约束：freezegrps 或位置约束，使受体刚性；
- 模块化模拟器：允许新增轻量算法循环（候选生成→能量评估→记录）。

5. 可剪裁实现 A：无源码改动的“伪对接”流程

5.1 模型与假设

- 受体冻结（刚体近似），配体刚体 + 少量扭转；
- 评分采用 GROMACS 势能：优先比较 $E_{inter}(\text{Protein, Ligand})$ 与（可选） $E_{intra}(\text{Ligand})$ ；
- 每个候选做短最小化，避免几何重叠引出的虚高能。

5.2 最小 MDP 示例（真空/隐式近似）

```
integrator = steep
nsteps    = 200
emtol     = 100.0
nstlist   = 20
rlist     = 1.2
coulombtype = Cut-off
rcoulomb  = 1.2
vdwtype   = Cut-off
rvdw      = 1.2
constraints = h-bonds
pbc       = xyz
; 受体/配体能量分组
energygrps = Protein Ligand
; 冻结受体 (需 index.ndx 中有 Protein/Ligand 分组)
freezegrps = Protein
freezedim  = Y Y Y
```

5.3 管道示意（每个候选）

1. 构造随机刚体位姿: `editconf -translate/-rotate` 作用于配体;
2. 合并受体+配体, 生成 index: `select "resname ..."` 得 Protein/Ligand 分组;
3. `grompp -f em.mdp -c start.gro -p topol.top -n index.ndx` ;
4. `mdrun -v -deffnm em` ;
5. `gmx energy -f em.edr` 读取 `LJ-SR:Protein-Ligand` 与 `Coul-SR:Protein-Ligand` 求和评分;
6. 记录前 N 姿势 (`em.gro`→PDB) , 并并行化多候选 (`-multidir`) 。

该流程把“动态”转化为“静态打分+局部最小化”，即对接的核心思想。

6. 可剪裁实现 B：源码级最小改造（示意）

- 新增“dock”算法环：生成候选→（可选）最小化→能量评估→排行榜；
- MDP 新键: `dock-center/dock-size/dock-poses/dock-min-steps/dock-allow-torsions` ;
- 冻结受体、仅更新配体自由度；每次刚体移动后触发一次能量与（可选）梯度；
- 并行：以多模拟拆分候选批次，在主进程归并前 N；
- 代码落点: `src/gromacs/modularsimulator/*`（调度）、`src/gromacs/mdrun/*`（入口）、`mdtypes/inputrec*`（参数），能量评估复用现有 `nbnxm` 核，无需改物理核。

7. 局限与改进

- 小分子参数化必要 (GAFF/CGenFF 等) , 否则势能排序不可靠;
- 栅格加速可进一步“对接化”: 预计算受体势能栅格, 搜索阶段做插值 + 配体内能;
- 溶剂与熵: 真空/隐式近似会偏置评分, 可在筛选后用显式溶剂 MD 做再排序;
- 受体柔性: 可用多构象受体 (ensemble docking) 或对热点侧链施加弱约束;
- 数值稳健性: 初始重叠可用软核/位移截断, 或最小化前先做轻度排斥预处理。

8. 小结

从经典力学角度, 对接即在同一势能面上的“静态采样与排序”, 而 MD 是“时间驱动连续采样”。GROMACS 提供的能量评估、最小化、邻域与并行机制, 使“动力模拟退化为对接式搜索”在理论与工程上都可: 先以“伪对接”原型验证方法学, 再逐步演化为源码级集成或引入栅格评分, 以在速度与物理一致性之间取得平衡。

许可声明 (License)

Copyright (C) 2025 GaoZheng

本文档采用[知识共享-署名-非商业性使用-禁止演绎 4.0 国际许可协议 \(CC BY-NC-ND 4.0\)](#)进行许可。