

AMBER 中炼金术自由能模拟技术的发展历史及其应用

Pengfei Li

September 22, 2023

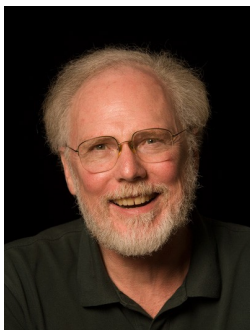
1 背景介绍

结合自由能的计算，尤其是针对蛋白-配体复合物的体系，是计算化学领域一个重要的研究方向。[炼金术自由能模拟](#)(Alchemical Free Energy Simulations) 技术被业内认为是计算结合自由能最精确的方法。该技术通过在始末热力学态之间增加一系列虚拟的中间态，从而实现缓慢的演变。这些演变过程可以包括将一个配体转变成另一个配体或者蛋白的某个残基转变成另一类残基(相对结合自由能模拟—RB-FEP)，以及将配体分子转变成不会被周围环境所感知的虚拟粒子(绝对结合自由能模拟—AB-FEP)等等。因为这些演变过程中的中间态往往并不对应真实存在的物理状态，所以将该技术称之为“炼金术”方法。

提及炼金术自由能模拟技术，就不得不提到自由能微扰理论 (FEP)。1954 年，美国理论物理学家和化学家[Robert Zwanzig](#)推导出 FEP 主方程 (也常被称为 Zwanzig 方程)[1], 奠定了自由能微扰理论 (FEP) 在物理与化学领域的基础地位。直到 20 世纪 80 年代，耶鲁大学教授[William L. Jorgensen](#)教授 (计算化学领域的先驱之一) 首次将 FEP 应用到在水中乙烷转变成甲醇的化学反应研究中 [2]; 美国物理化学家、加利福尼亚大学圣迭戈分校教授[J. Andrew McCammon](#)首次将 FEP 方法应用到了计算机辅助药物设计 (CADD) 领域 [3]; 已故的加州大学旧金山分校化学和药物化学[Peter Kollman](#)教授利用 FEP 对三肽类嗜热菌蛋白酶抑制剂的活性进行了预测 [4]。Peter Kollman 教授是当前仍十分流行的分子动力学模拟软件[AMBER](#)[5] 的早期开创者之一。当时，FEP 模拟技术已经存在于 AMBER 模拟软件中 [6]。



(a) Robert Zwanzig



(b) J. Andrew McCammon



(c) William L. Jorgensen



(d) Peter Kollman

由于采样算法的不断完善、蛋白和小分子力场的不断提升以及现代算力极大地提高，在工业界，薛定谔公司于 2014 年发布了商业版 FEP 模拟计算平台 FEP+。随后，薛定谔公司于 2015 年在 JACS 期刊上发表了关于 FEP+ 预测蛋白-配体亲和力的文章 [7]。该项工作中选取了 8 个靶标蛋白共有 199 个小分子配体，进行了结合能力预测，并将预测结果与实验结果进行了比较。对比结果显示 FEP+ 的预测值与实验值具有非常高的相关性，误差大概在 1.0 kcal/mol 左右。这也就意味着在实验活性上 5-10 倍的差距是可以很准确地预测出来的。这也是业内把 FEP 视为最为准确地预测结合自由能的方法的原因。同时，该项工作公布的 8 个蛋白-配体体系也被认为是业内验证亲和力计算方法准确度的标准体系。

2 近十年来，炼金术自由能模拟技术在 AMBER 中的发展

自 Peter Kollman 时期以来，炼金术自由能模拟技术主要实现在 AMBER 的 sander 模拟引擎中。2012 年，Joseph W. Kaus 和 J. Andrew McCammon 等人将该技术升级改进到 AMBER12 的 pmemd 高级模拟引擎中，通过更好地设计并行计算和减少对冗余能量项的计算等手段，实现的 pmemd 的计算效率提升至少是 sander 的 2.5 倍 [8]。而且，在模拟中无论是否引入“软核势”(softcore potential)，此改进的算法使得在 $\lambda = 0$ 和 $\lambda = 1$ 两端态附近都可以直接模拟而无需再通过外推算法得到该处哈密顿的能量值。请详见图1。

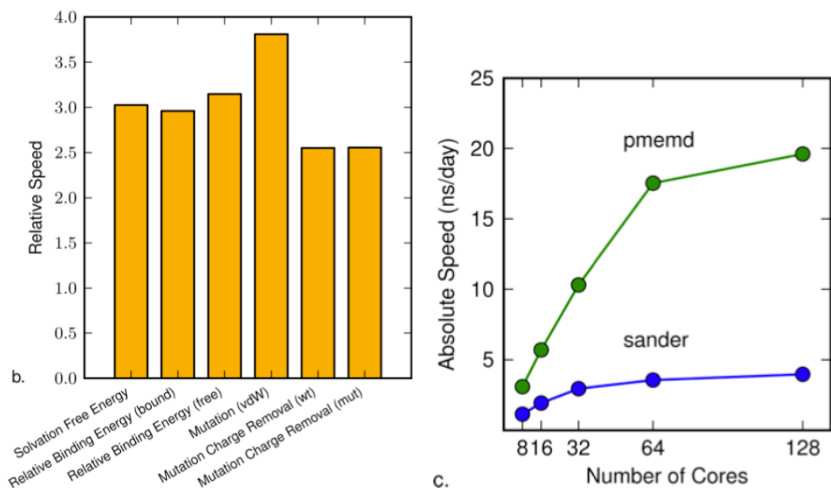


Figure 1: pmemd 较 sander 效率提升对比 (左图); pmemd 和 sander 的绝对模拟速度对比: 相对结合自由能计算体系 (右图)。

- 最初的 sander 模块,在没有引入软核势时,构造体系拓扑结构需要引入额外的虚原子 (dummy atoms) 以保证初末态分子具有相同的原子数。而且在当时还没有出现软核势时,通过将中间态哈密顿量 $V(\lambda) = (1-\lambda)^k V_0 + [1 - (1-\lambda)^k] V_1$ 的混合因子 λ 加上 k 次幂来帮助改善端态模拟的收敛性。当 $k=1$ 时,中间态哈密顿量也就是一般的线性混合 $V(\lambda) = (1-\lambda)V_0 + \lambda V_1$; 当 $k \geq 4$ 时,端态模拟才有可能收敛,甚至到 $k=6$ 时,可以做到类似色氨酸分子大小的变化。直至后来软核势的出现 [9],可以很好地解决当有原子消失或者出现所引起的端态难收敛的情况 (我们常称之为“端态灾难”问题)[10]。请详见图2。我们可以知道,“端态灾难”问题本质是由于消失或者出现的原子与环境原子靠近时所带来的范德华相互作用有着较大的波动引起的。[11]

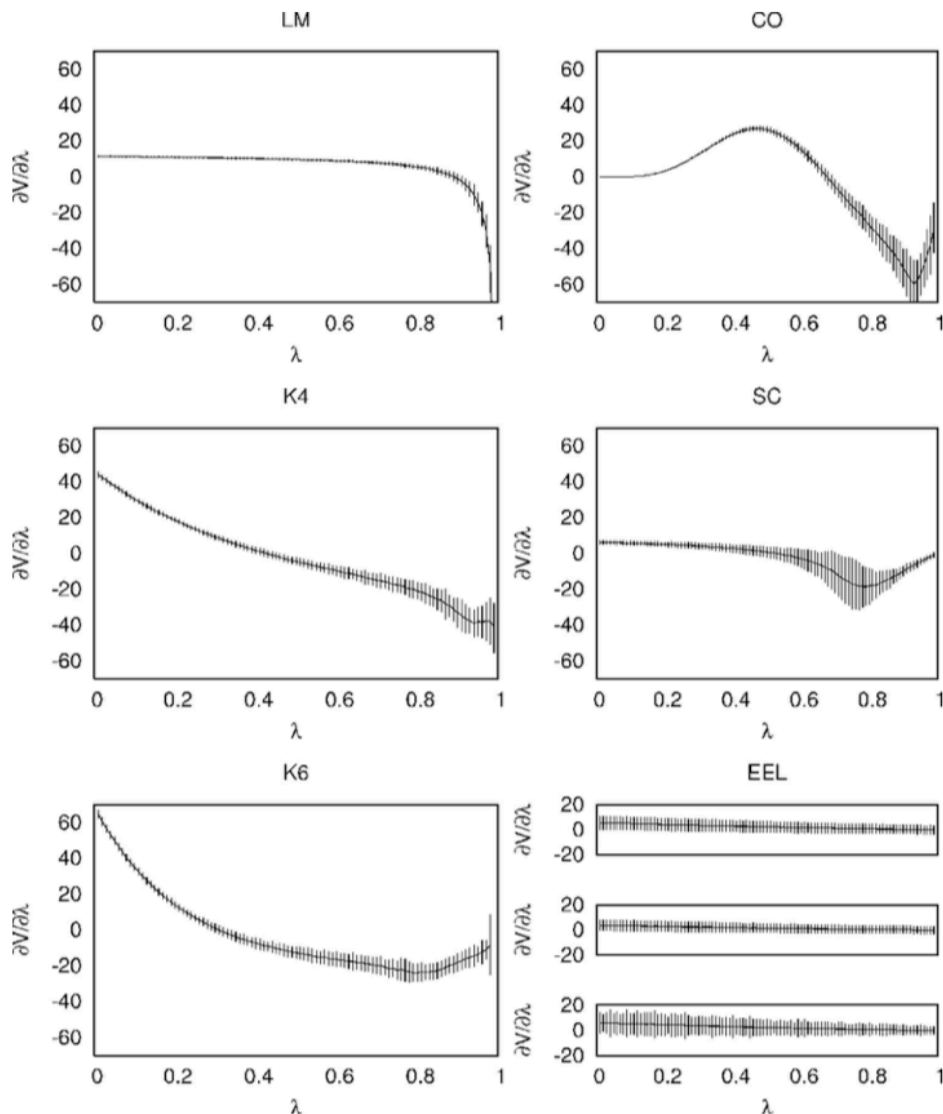


Figure 2: 几种不同方案对于自由能变化曲线的影响: 线性 LM、常用方法 CO、非线性 K4 和 K6 以及软核势 SC 方案。

- 当前广泛使用的 pmemd 模块,在使用软核势后,不需要引入额外的虚原子,只需要把要转变的原子分别放在 scmask1 和 scmask2 中即可,同时将初末态相同的原子部分保持相同,而

把分子中要转变的原子全部分别放在 timask1 和 timask2 中。

- 根据此时势能函数公式的定义以及相应拓扑结构的定义，AMBER pmemd TI 模块采取的是“混合拓扑” (hybrid topology) 方式，具体定义如图3所示。

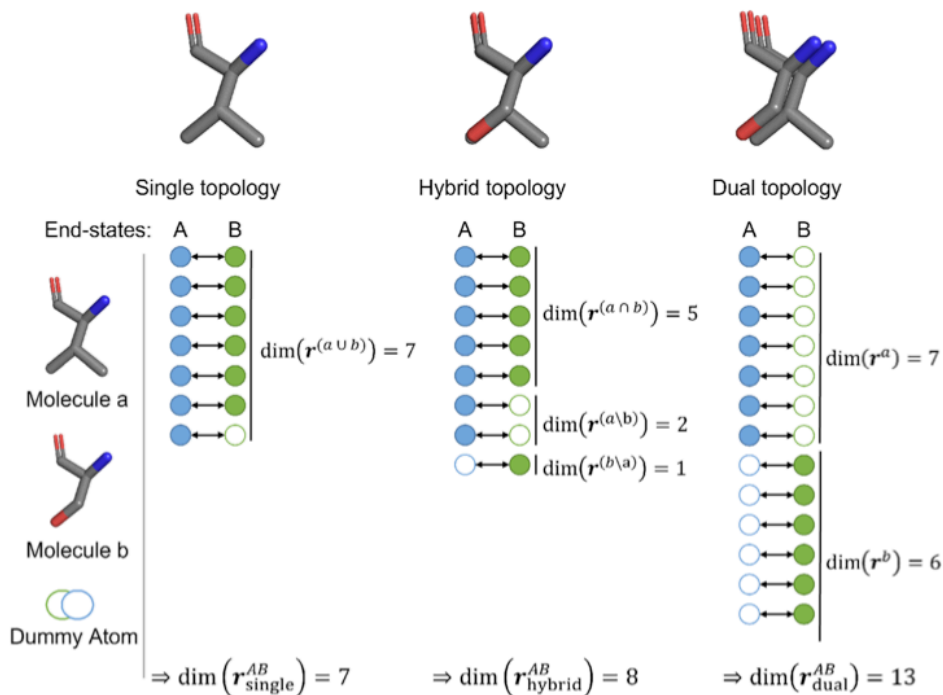


Figure 3: 单拓扑、混合拓扑以及双拓扑形式的定义 [12]。混合拓扑形式是结合了针对分子中定义为相同的原子部分采取单拓扑形式和针对要转变的原子部分采取双拓扑形式，这种方式也出现在 NAMD 炼金术自由能模拟技术的最新实现中 [13]。

- 关于三种不同拓扑形式的优缺点，请详见图4。小编认为，选取不同的拓扑形式是实现炼金术自由能模拟技术的底层框架基础，同时决定了在做该模拟时如何构造原子对匹配 (atom-to-atom mapping) 方案或在分子之间如何加约束的方案。以混合拓扑形式为例，通常根据这对分子的最大相似子结构 (MCS) 来定义要做原子对匹配的区域，具体 atom-to-atom 如何匹配有不同的实现方案 (在这里我们不做展开)。

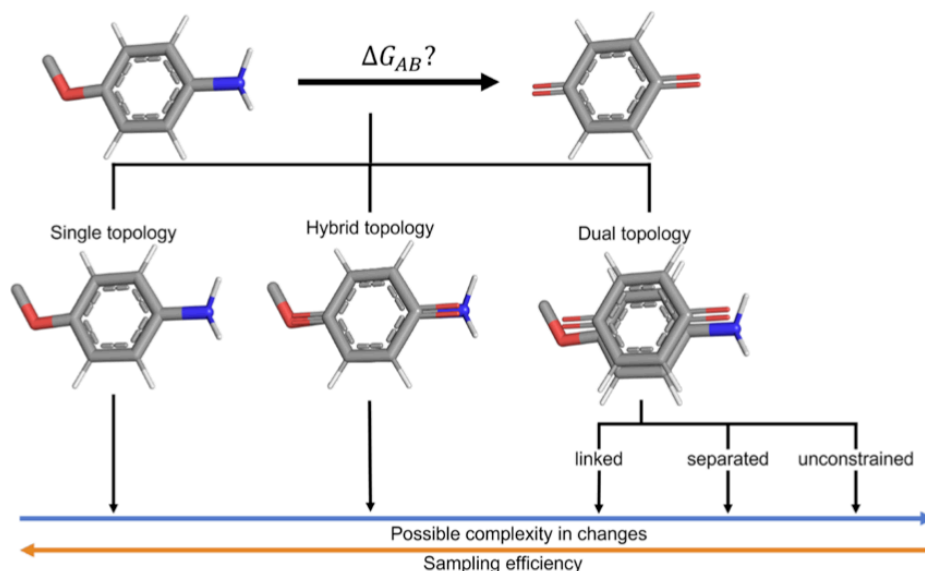


Figure 4: 单拓扑、混合拓扑以及双拓扑形式，在采样效率上依次下降，但在所能处理的微扰变化多样性上依次增强 [12]。在此坐标轴的最右端，相对结合自由能模拟 (RB-FEP) 近似等同于同时进行两个绝对结合自由能模拟 (AB-FEP)。

- 原子对的匹配及其匹配的原子区域 (common atoms) 的定义应该尽可能保证虚原子 (dummy/softcore atoms) 所形成的构象对于匹配区域结构的能量贡献在热力学循环的结合过程和未结合过程能够较好地抵消掉，否则会导致错误的预测结果 [14]。
- 在单拓扑和混合拓扑中，如何处理 common-softcore 原子之间的成键相互作用对结果可靠性的影响非常显著。严格合理的推荐方案是保留单个包含一个 softcore 原子的成键作用、单个包含一个 softcore 原子的键角作用、单个包含两个 softcore 原子的二面角作用 [15, 16, 17]。但是在 AMBER 中，默认的处理方案是保留所有 common-softcore 原子之间的成键相互作用。当然在目前最新的版本中也提供了其它定制化方案，可供用户选择。AMBER 软件手册中有详细的介绍。Stefan Boresch 等人最近的一项工作中详细做了这一方面的研究 [14]。该研究表明，即使是类似 AMBER 的默认处理方案，在大多数情况 (terminal junction/dual junction/planar triple junction) 中，可以得到相对可靠的结合自由能预测结果。尽管有的时候 softcore 原子对于自由能是有贡献的，但是可以在热力学循环的结合过程和未结合过程能够较好地抵消掉。但是对于 nonplanar triple junction 情况，类似 AMBER 的默认处理方案会带来一些较大误差。但是，小编认为可以通过改变 common 原子的范围尽可能避免产生 nonplanar triple junction 等情况。同时，该文章也指出研究总结的规律并不适用于有环键断裂/形成的情况，但是可以将涉及到环键断裂/形成的环放在 softcore 区域以形成 terminal junction 等情况。因为在这些情况下，预测结果相对可靠。其实，这正是利用了混合拓扑方

案的优势——针对要转变的原子部分采取双拓扑形式，不过这样也会带来模拟窗口和模拟时间需要增加的可能性。

- 此时，AMBER12 版本的 sander 和 pmemd 模拟引擎只支持 CPU 来做炼金术自由能模拟计算。

直到 2017 年，Darrin M. York 和 Tai-Sung Lee 等人基于 AMBER16 pmemd 中炼金术自由能模拟的 CPU 代码，进一步发展到可以在显卡上进行运算的 GPU 代码，将其命名为 pmemdGTI 模块，并在 AMBER18 版本中正式发布。这使得自由能模拟计算在单块 GPU 上的速度较单核 CPU 速度有了超过 100 倍以上的提升 [18]。请详见图5。

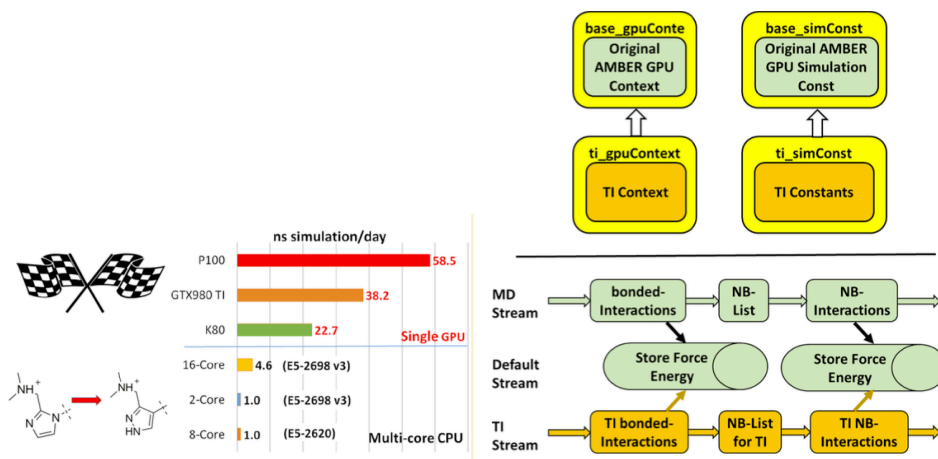


Figure 5: 在 AMBER18 中，多核 CPU 与单卡 GPU 的炼金术自由能模拟计算的效率对比 (左图); pmemdGTI 的 GPU 代码实现流程 (右图)。

2018 年，针对 AMBER18 这次重要的发布，AMBER 社区总结了其在炼金术自由能模拟计算模块的主要提升，包括有诸如 H-REMD、aMD、GaMD 等增强采样方法；常见的 NVE、NVT、NPT 等系综方法；TI 和 MBAR 后处理分析的数据产生；针对金属离子的 12-6-4 势能函数的支持 [19]。请详见图6。

Free energy compatible features		Amber16		Amber18	
Category	Functionality	CPU	GPU	CPU	GPU
Ensemble	NVE	✓	✗	✓	✓
	NVT	✓	✗	✓	✓
	NPT	✓	✗	✓	✓
	semi-iso P	✗	✗	✗	✗
	CpHMD	✓	✗	✓	✓
Free Energy Analysis	TI	✓	✗	✓	✓
	MBAR	✓	✗	✓	✓
Enhanced Sampling	H-REMD	✓	✗	✓	✓
	AMD	✓	✗	✓	✓
	SGLD	✓	✗	✓	✗
	GAMD	✓	✗	✓	✓
Potentials	12-6-4	✓	✗	✓	✓

Figure 6: 在炼金术自由能模拟方面，与 AMBER16 相比，AMBER18 主要发展了在 GPU 上的运算模块，几乎能够支持所有 CPU 支持的模拟功能。

在 AMBER18 发布之后, Darrin M. York 课题组作为 AMBER GPU-TI 的主要发展和维护者, 与 Kenneth M. Merz, Jr. 课题组合作, 系统地验证了 AMBER GPU-TI 在 RB-FEP 的表现 [20]。他们选用了前面提到的由薛定谔公司提供的测试集, 包括 8 个靶标蛋白 199 个分子组成的 330 个分子对。同时选用了 FF14SB+GAFF1.8 的力场组合。整体的表现 (直接由 GPU-TI 模拟分析得到的 $\Delta\Delta G$ 的结果): MUE 为 1.25 kcal/mol, RMSE 为 1.64 kcal/mol。同时, Junmei Wang 课题组也与 Darrin M. York 课题组合作, 同样选用了薛定谔测试集中的 4 个靶标蛋白 (BACE, CDK2, MCL1, PTP1B) 结合 FF14SB+GAFF1.8 的力场组合, 测试了一些详细的模拟协议细节。该研究结果表明在 $N_{run} \geq 4$ 且 $t > 4ns$ 时, 预测的结合自由能值能达到最好的表现 [21]。这里整体的表现结果, 如图7所示, GPU-TI 的预测能力与 FEP+ 的预测能力较为接近, 如图8所示。

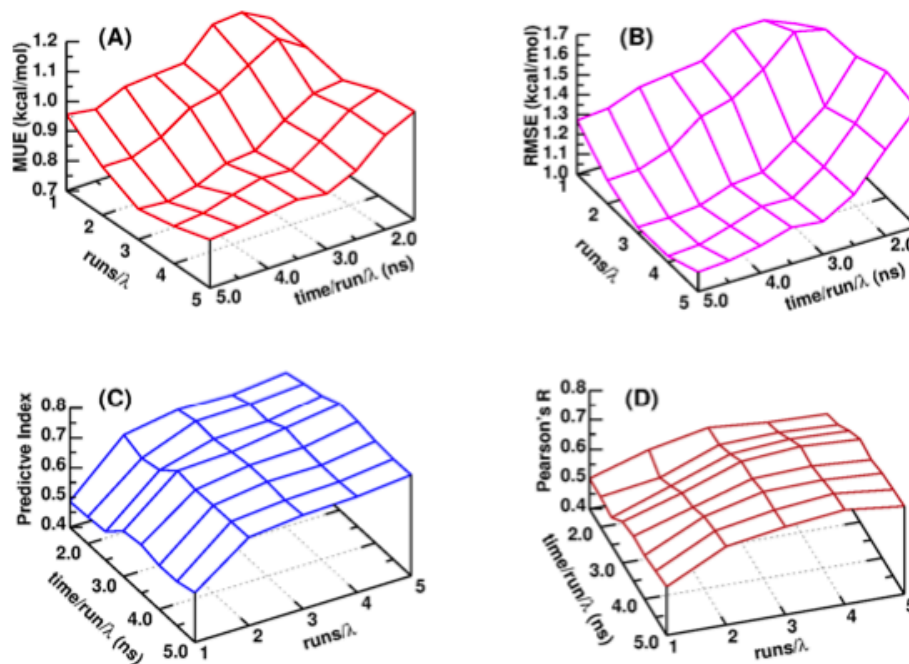


Figure 7: 随着模拟时间 t 与重复模拟次数 N_{run} 的增加, 针对 PTP1B 体系的 23 个分子, 其 ΔG 值的 MUE、RMSE、PI 以及 R 的变化情况。

systems	BACE ³²		CDK2 ³³		MCL1 ³⁴		PTB1B ³⁵	
protocols	FEP+	GPU-TI	FEP+	GPU-TI	FEP+	GPU-TI	FEP+	GPU-TI
#compounds	36	41	16	22	42	44	23	27
$\Delta\Delta G$ MUE (kcal/mol)	0.84	0.93	0.91	0.94	1.16	0.82	0.89	0.71
$\Delta\Delta G$ RMSE (kcal/mol)	1.03	1.22	1.11	1.16	1.41	1.01	1.22	0.91
Pearson's R for ΔG	0.78	0.61	0.48	0.64	0.77	0.65	0.80	0.75

Figure 8: 针对 BACE, CDK2, MCL1, PTP1B 四个体系, GPU-TI 的预测能力与 FEP+ 的预测能力的对比情况。

在 AMBER20 发布之前, Darrin M. York 课题组进一步验证测试了 AMBER 的 GPU-TI 模块 [22], 发现在处理 common 原子与 softcore 原子之间相互作用的时候, 1-4 范德华和 1-4 静电相互作用并没有随着 λ 的改变而改变, 这会导致较大的系统误差。如图9所示, 我们都知道乙酸分子在气相下, 倾向于 cis 结构, 而在水中, 倾向于 trans 结构。AMBER18 和 AMBER20 (文章里标记为 AMBER18*) 也给出了正确的自由能曲线 (黑色曲线和绿色曲线)。在 $\lambda = 1$ 窗口 (dummy 状态) 时, 在水中的乙酸分子和环境已经不存在相互作用, 在气相下的乙酸分子也同样和环境不存在相互作用。所以两者应该是相同的状态 (即红色曲线和蓝色曲线是相同的)。但是在 $\lambda = 0$ 和 $\lambda = 1$ 窗口下, 气相中的乙酸分子自由能曲线应该是不同的。因为随着 λ 的改变, common 与 softcore 原子之间的 1-4 范德华和 1-4 静电相互作用也应该随之消失。这一点在 AMBER20 版本中, 的确是正确的 (绿色曲线不同于蓝色曲线)。但是在 AMBER18 版本中, 绿色曲线是和蓝色曲线重合的, 这一点体现出原先的 common 与 softcore 原子之间的 1-4 范德华和 1-4 静电相互作用并没有随着 λ 的改变而改变, 而是保持和 $\lambda = 0$ 状态下一致的相互作用, 从而使得蓝色曲线与绿色曲线重合。最终, 这个问题在 AMBER20 发布版本中得以解决。

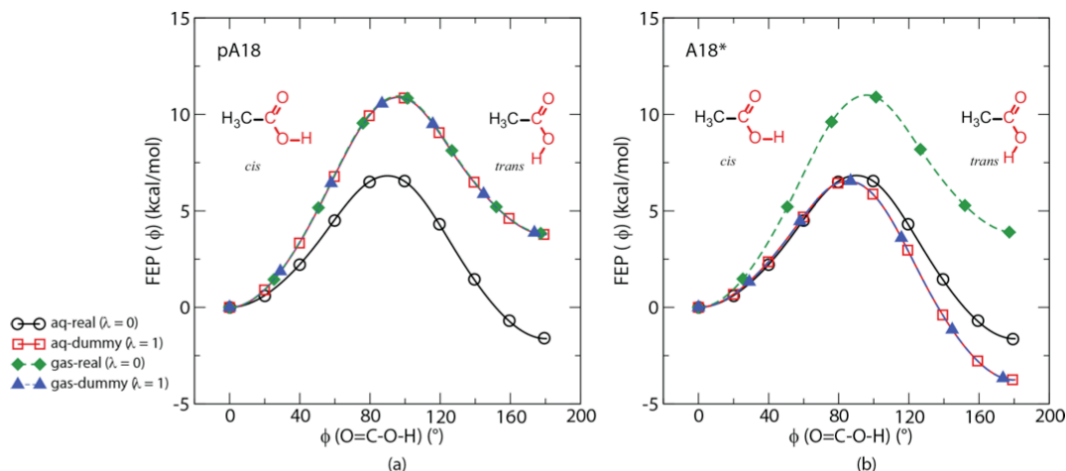


Figure 9: 乙酸分子 $\text{O}=\text{C}-\text{O}-\text{H}$ 二面角在水中 (aq) 和气相中 (gas) 分别在 $\lambda = 0$ 和 $\lambda = 1$ 窗口下, 相应的自由能变化曲线, a) 使用之前版本的 AMBER18(pA18); b) 使用更新版本的 AMBER18(A18* 即 AMBER20)。

前面我们重点解释过 softcore 函数通过软化近距离时的范德华相互作用可以解决“端态灾难”问题。除此之外，还有“粒子坍缩” (particle collapse) 问题，即在中间态窗口下有可能出现的情况：存在软核作用的原子之间会出现彼此重叠在一起的可能，主要由于软核范德华排斥相互作用不足以抵消软核静电吸引相互作用所引起的，这个可以通过调节软核势参数比 α/β 来调节，因此往往是依赖于体系变化的。还有在靠近端态处“势能梯度曲线陡变” (large gradient-jump) 问题，即自由能变化对于软核 β 参数取值较大比较敏感而导致，而往往 β 参数取值较大是为了平衡软核静电相互作用和软核范德华相互作用。因为其导致在靠近端态处的自由能曲线变陡，产生较大的能量波动，我们可以通过光滑的软核函数 SSC(Smoothstep Softcore) 形式来解决，其导数在端态处为零。分析两个问题的本质原因，不难发现另外一个方案，即只要把静电相互作用和范德华相互作用分开在不同的阶段来做，自然可以避免这两个问题的发生。但是将两种相互作用分开来处理 (通常被称为 stepwise 方案)，首先在关闭静电作用阶段，有时会遇到较大的净电荷变化，这在周期性模拟中往往是不希望遇到的；其次会增加模拟设置难度同时也增加了一定的计算量；最后这样的策略会增加与增强采样方法结合以及 λ 配置方案优化的难度。为了避免这几个问题，我们还是希望有一个能够同时变化两种相互作用的方案 (通常被称为 concerted 方案)。所以，在 AMBER20 版本中，针对 concerted 方案，Darrin M. York 课题组进一步发展了软核势函数 [23]。具体表现如图10所示，可以很明显地看出，新型软核势函数 SSC(2) 结合软核势参数 ($\alpha = 0.2; \beta = 50$) 有效地解决了“粒子坍缩”问题和“势能梯度曲线陡变”问题。

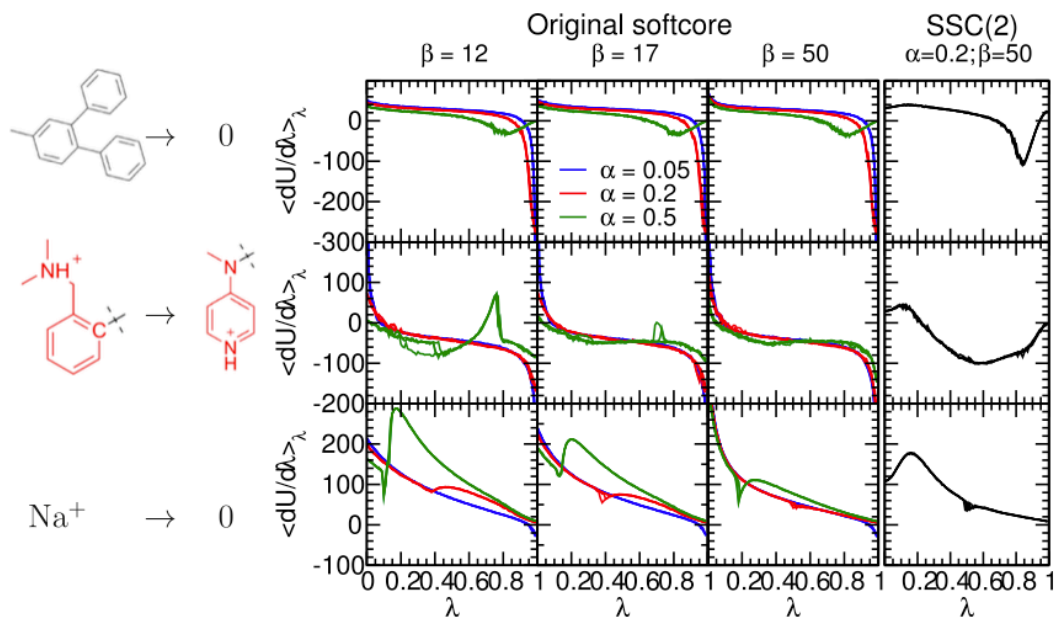


Figure 10: 新型软核势函数 (SSC(2) AMBER20) 和原始软核势函数 (Original softcore AMBER18) 与不同软核参数 (α, β) 的组合表现对比。

随后, Darrin M. York 课题组与 Silicon Therapeutics 公司合作, 系统性阐述和测试了 AMBER20 新增加的功能, 给出了如图11所示的标准测试结果 [24]。感兴趣的小伙伴, 可以详细读一下参考文章 [24], 这篇文章详细阐述了 AMBER 当时在对于炼金术自由能模拟的回顾、实现及展望, 并做了丰富的数据验证, 给出了一些切实的实践经验。

Table 1. Mean Unsigned Error (MUE) in kcal/mol for 8 Protein Targets^a

method	BACE	CDK2	JNK1	MCL1	P38	PTP1B	thrombin	TYK2	mean
AMBER20 (this work)	0.88	0.93	0.73	1.30	0.79	0.79	0.39	0.93	0.84
FEP+ (Wang et al. ¹²¹)	0.84	0.91	0.78	1.16	0.80	0.89	0.76	0.75	0.86
Cresset (Kuhn et al. ¹¹⁶)	0.95	0.95	0.78	1.36	1.18	1.04	0.23	0.71	0.90
PMX (Gapsys et al. ³⁴)	0.84	0.68	0.80	1.23	0.77	0.90	0.77	1.01	0.88

^aResults are presented for the work here and other recent relative binding free energy (RBFE) publications on the same data set.¹²¹ The AMBER20 results are an average of 10 independent runs using the smoothstep softcore SSC(2)⁵⁸ and the GAFF2 force field. Results for each individual perturbation, statistical errors across the 10 runs, correlation coefficients, and null model results can be found in [Supporting Information](#).

Figure 11: 针对同一套标准测试体系, 基于 AMBER20 搭建的 FEP 平台与其它主流 FEP 平台表现的对比。

最后, 还需提到的是, Darrin M. York 课题组基于 AMBER20 搭建了一套 FEP 模拟计算平台, 被称之为 **AMBER-DD** (AMBER Drug Discovery Boost)[25]。AMBER-DD 所包含的功能以及主要流程, 如图12所示。

Category	description	input/enable	Status
REMD	Odd number REMD		A20
REMD	Targeted Volume	NPT=4	DD
Output	Detailed TI component output	gti_output	A20
Fix	Fix inconsistency of SC cutoff	gti_cut	A20
Fix	Charge neutralization	gti_chg_keep	A20
Fix	Fix cross NB problem	gti_add_sc (default 0 to 1)	A20
New Feature	User-control intra-SC NB terms	gti_add_sc	A20
New Feature	User-control cross NB terms	gti_add_sc	A20
New Feature	SSC(P)	different SC functional form	A20
New Feature	Lambda scheduling	different SC timing	A20
New Feature	cross bonded term correction	gti_bat_sc	A20
New Feature	User-control cross bonded terms	gti_bat_sc	A20
New Feature	12-6-4 potential TI/GPU		A20
New Feature	12-6-4 potential MD/GPU		A20
New Feature	Different m and n for SC		DD
New Feature	scaled NMR restraint for TS regions		A20
New Feature	ACES enhanced sampling	gti_add_sc/REMD	DD
New Feature	Self-adjusted mixture sampling (SAMS)		DDX
New Feature	REST-type enhanced sampling		DDX
New Feature	BARnet		DDi
New Feature	Workflow		DDs

A20: already in AMBER20 pmemd.cuda. DD: only in AMBER-DD pmemd.cuda. DDX: only in AMBER-DD pmemd.cuda, under α -phase tests. DDi: only in AMBER-DD, as a set of separate programs. DDs: only in AMBER-DD, as a set of scripts.

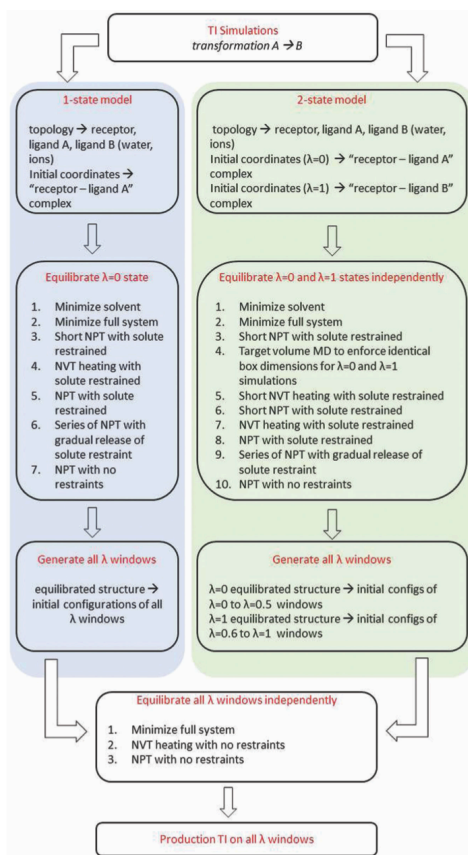


Figure 12: AMBER-DD 所包含的功能 (左图); AMBER-DD 的主要流程图 (右图)。

并且，AMBER-DD 也测试了 ACES(AlChemical Enhanced Sampling) 增强采样方法在自由能模拟计算中的作用表现，如图13所示。

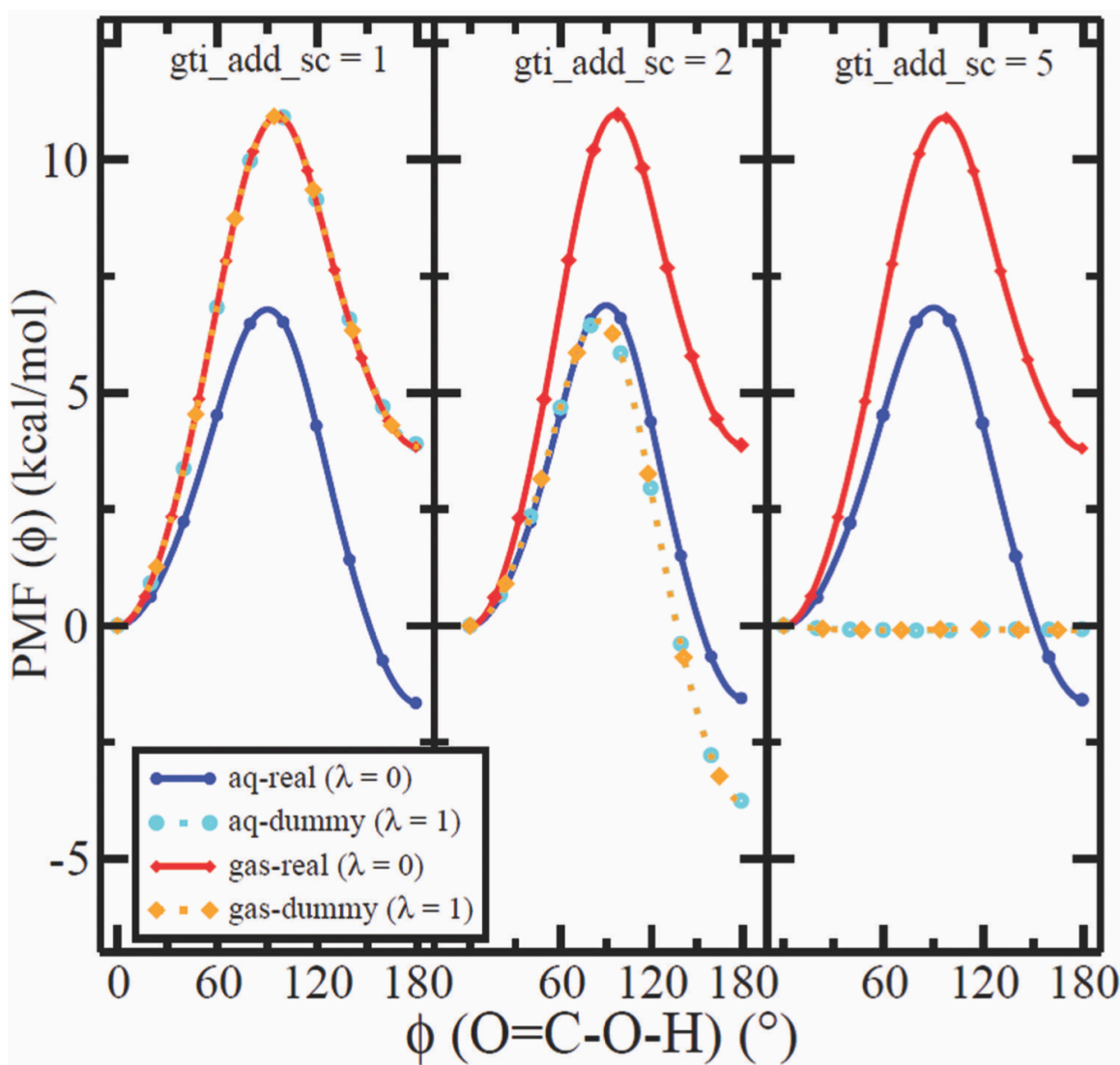


Figure 13: 乙酸分子 O=C—O—H 二面角在水中 (aq) 和气相中 (gas) 分别在 $\lambda = 0$ 和 $\lambda = 1$ 窗口下，相应的自由能变化曲线。由此可以看出，在 $gti_add_sc=5$ 即软化了该二面角时，在 $\lambda = 1$ 窗口下该自由度不再存在较大能垒，可以实现二面角的自由旋转从而实现了增强采样。

将该增强采样方案应用到 CDK2 靶标蛋白的 1h1r 和 1h1q 这对分子的变化中。选取该例子的原因是：1h1r 中存在一个苯环的连接处有一个二面角，它有着明显的两个角度分布：syn 式角度为-8.81 度和 anti 式角度为 150.75 度。在 REST2 增强采样方法发展的工作中，也是选用了这个例子，的确看到了不使用增强采样方法时，syn 和 anti 两个角度不会同时被采样到 [26]。ACES 增强采样方案 (SC5) 结合 H-REMD 方案和 REST2 方案类似，都能同时采样到 syn 和 anti 两个角度，同时从不同角度的构象出发，最终都能得到收敛到相同值的结合自由能结果，如图14和图15所示。

感兴趣的小伙伴，可以免费下载这个软件包，体验一下 AMBER-DD 的功能。

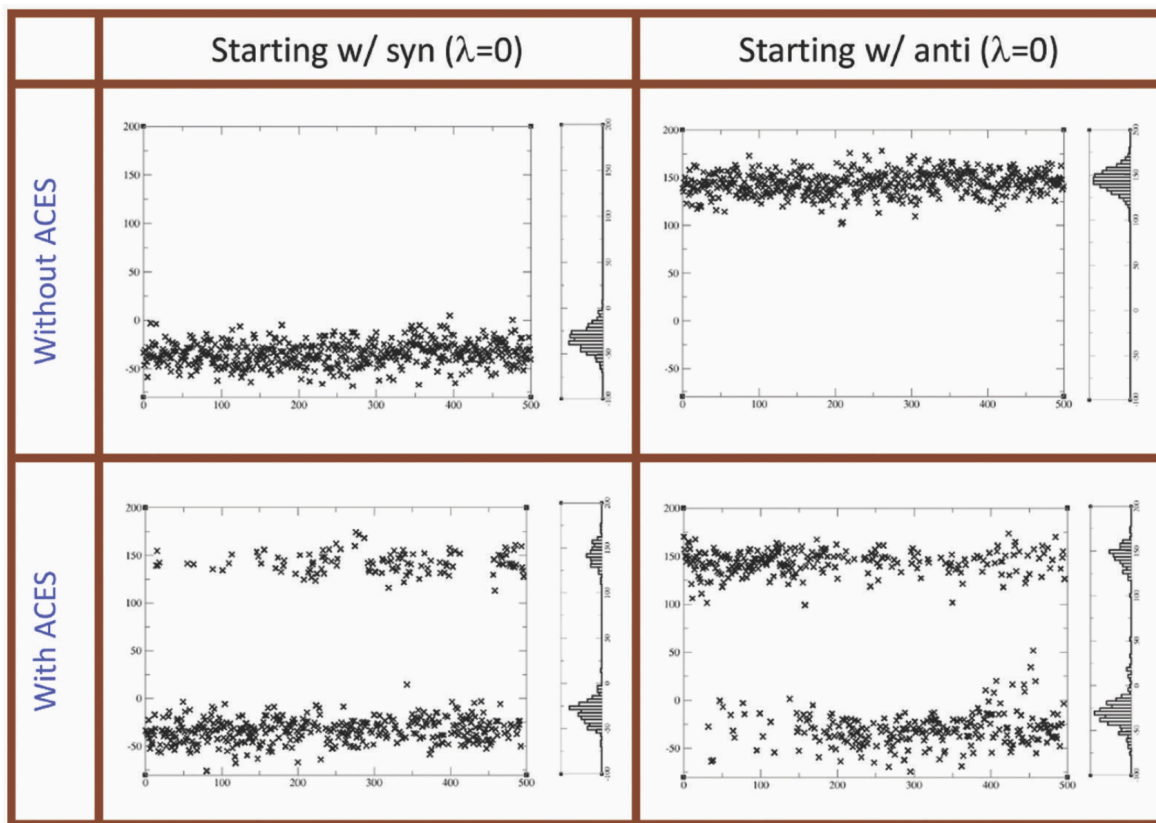


Figure 14: 在使用 ACES 增强采样方案时，分别从 syn 和 anti 两个角度构象出发，可以同时采样到两种不同构象。而不使用 ACES 方案时，则只能采样到单一构象。

		<i>Syn</i>	<i>Anti</i>
no REMD	SC2	0.41(27)	0.63(28)
	SC5	0.46(32)	0.55(22)
REMD	SC2	0.33(20)	0.50(13)
	SC5	0.50(16)	0.48(12)

Syn: Simulation starting with the syn conformation. Anti: Simulation starting with the anti conformation.
 SC2: Simulation with gti_add_sc=2. SC5: Simulation with gti_add_sc=5.

Figure 15: 在使用 ACES 增强采样方案时，分别从 syn 和 anti 两个角度构象出发的模拟计算，可以得到收敛到相同值的结合自由能结果，而且同时使用 H-REMD 方案，可以帮助加速模拟计算的收敛。

3 FEP Spell 的发展及其表现

FEP Spell 作为 SingleParticle 公司的 FEP 计算平台，基于 AMBER22 中的炼金术自由能模拟技术实现了完全自动化的操作流程部署在公司内部的 HPC 集群上，可以方便快捷地执行 FEP 流程。我们选用上述标准测试体系 [7] 来对 FEP Spell 平台进行标准化测试，具体测试结果—相对结合自由能值—比较如下表所示。

Platform	BACE	CDK2	JNK1	MCL1	P38	PTP1B	THROMBIN	TYK2	Wt. MUE ^a
FEP SPell (Ours)	0.88	1.02	0.75	1.17	1.02	1.05	0.43	0.84	0.97
AMBER20 (Lee et al.[24])	0.88	0.93	0.73	1.30	0.79	0.79	0.39	0.93	0.91
FEP+ (Wang et al.[7])	0.84	0.91	0.78	1.16	0.80	0.89	0.76	0.75	0.90
Cresset (Kuhn et al.[27])	0.95	0.95	0.78	1.36	1.18	1.04	0.23	0.71	1.02
PMX (Gapsys et al.[28])	0.84	0.68	0.80	1.23	0.77	0.90	0.77	1.01	0.91

^a Wt. MUE (Weighted Mean Unsigned Error: kcal/mol);

由此可以看出，我们搭建的 FEP Spell 平台与包括 FEP+ 平台在内的其他几个平台表现相近，可以在药物设计项目中投入使用，期待与大家合作。

References

- [1] Robert W Zwanzig. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *The Journal of Chemical Physics*, 22(8):1420–1426, 1954.
- [2] William L. Jorgensen and C. Ravimohan. Monte carlo simulation of differences in free energies of hydration. *The Journal of Chemical Physics*, 83(6):3050–3054, 1985.
- [3] Chung F. Wong and J. Andrew. McCammon. Dynamics and design of enzymes and inhibitors. *Journal of the American Chemical Society*, 108(13):3830–3832, 1986.
- [4] Kenneth M. Merz and Peter A. Kollman. Free energy perturbation simulations of the inhibition of thermolysin: prediction of the free energy of binding of a new inhibitor. *Journal of the American Chemical Society*, 111(15):5649–5658, 1989.
- [5] D. A. Case, H. M. Aktulga, K. Belfon, I. Y. Ben-Shalom, J. T. Berryman, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, , M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, K. Kovalenko, R. Krasny, T. Kurtzman, Tai-Sung Lee, S. Le-Grand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, V. Man, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, , H. Nguyen, A. Onufriev, F. Pan, S. Pantano, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wilson, R. M. Wolf, X. Wu, Y. Xiong, Y. Xue, D. M. York, S. Zhao, and P. A. Kollman. Amber 22. 2022.
- [6] Peter. Kollman. Free energy calculations: Applications to chemical and biochemical phenomena. *Chemical Reviews*, 93(7):2395–2417, 1993.
- [7] Lingle Wang, Yujie Wu, Yuqing Deng, Byungchan Kim, Levi Pierce, Goran Krilov, Dmitry Lupyan, Shaughnessy Robinson, Markus K. Dahlgren, Jeremy Greenwood, Donna L. Romero, Craig Masse, Jennifer L. Knight, Thomas Steinbrecher, Thijs Beuming, Wolfgang Damm, Ed Harder, Woody Sherman, Mark Brewer, Ron Wester, Mark Murcko, Leah Frye, Ramy Farid, Teng Lin, David L. Mobley, William L. Jorgensen, Bruce J. Berne, Richard A. Friesner, and Robert Abel. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *Journal of the American Chemical Society*, 137(7):2695–2703, 2015.
- [8] Joseph W. Kaus, Levi T. Pierce, Ross C. Walker, and J. Andrew McCammon. Improving the efficiency of free energy calculations in the amber molecular dynamics package. *Journal of Chemical Theory and Computation*, 9(9):4131–4139, 2013.
- [9] Thomas C. Beutler, Alan E. Mark, René C. van Schaik, Paul R. Gerber, and Wilfred F. van Gunsteren. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chemical Physics Letters*, 222(6):529–539, 1994.

- [10] Thomas Steinbrecher, David L. Mobley, and David A. Case. Nonlinear scaling schemes for lennard-jones interactions in free energy calculations. *The Journal of Chemical Physics*, 127(21):214108, 2007.
- [11] Thomas Steinbrecher, InSuk Joung, and David A. Case. Soft-core potentials in thermodynamic integration: Comparing one- and two-step transformations. *Journal of Computational Chemistry*, 32(15):3253–3263, 2011.
- [12] Benjamin Ries, Salome Rieder, Clemens Rhiner, Philippe H. Hunenberger, and Sereina Riniker. Restraintmaker: A graph-based approach to select distance restraints in free-energy calculations with dual topology. *Journal of Computer-Aided Molecular Design*, 36:175–192, 2022.
- [13] Wei Jiang, Christophe Chipot, and Benoît Roux. Computing relative binding affinity of ligands to receptor: An effective hybrid single-dual-topology free-energy perturbation approach in namd. *Journal of Chemical Information and Modeling*, 59(9):3794–3802, 2019.
- [14] Markus Fleck, Marcus Wieder, and Stefan Boresch. Dummy atoms in alchemical free energy calculations. *Journal of Chemical Theory and Computation*, 17(7):4403–4419, 2021.
- [15] Stefan Boresch and Martin Karplus. The role of bonded terms in free energy simulations: 1. theoretical analysis. *The Journal of Physical Chemistry A*, 103(1):103–118, 1999.
- [16] Stefan Boresch and Martin Karplus. The role of bonded terms in free energy simulations. 2. calculation of their influence on free energy differences of solvation. *The Journal of Physical Chemistry A*, 103(1):119–136, 1999.
- [17] Stefan Boresch. The role of bonded energy terms in free energy simulations - insights from analytical results. *Molecular Simulation*, 28(1-2):13–37, 2002.
- [18] Tai-Sung Lee, Yuan Hu, Brad Sherborne, Zhuyan Guo, and Darrin M. York. Toward fast and accurate binding affinity prediction with pmemdgti: An efficient implementation of gpu-accelerated thermodynamic integration. *Journal of Chemical Theory and Computation*, 13(7):3077–3084, 2017.
- [19] Tai-Sung Lee, David S. Cerutti, Dan Mermelstein, Charles Lin, Scott LeGrand, Timothy J. Giese, Adrian Roitberg, David A. Case, Ross C. Walker, and Darrin M. York. Gpu-accelerated molecular dynamics and free energy methods in amber18: Performance enhancements and new features. *Journal of Chemical Information and Modeling*, 58(10):2043–2050, 2018.
- [20] Lin Frank Song, Tai-Sung Lee, Chun Zhu, Darrin M. York, and Kenneth M. Merz. Using amber18 for relative free energy calculations. *Journal of Chemical Information and Modeling*, 59(7):3128–3135, 2019.
- [21] Xibing He, Shuhan Liu, Tai-Sung Lee, Beihong Ji, Viet H. Man, Darrin M. York, and Junmei Wang. Fast, accurate, and reliable protocols for routine calculations of protein–

- ligand binding affinities in drug design projects using amber gpu-ti with ff14sb/gaff. *ACS Omega*, 5(9):4611–4619, 2020.
- [22] Hsu-Chun Tsai, Yujun Tao, Tai-Sung Lee, Kenneth M. Merz, and Darrin M. York. Validation of free energy methods in amber. *Journal of Chemical Information and Modeling*, 60(11):5296–5300, 2020.
- [23] Tai-Sung Lee, Zhixiong Lin, Bryce K. Allen, Charles Lin, Brian K. Radak, Yujun Tao, Hsu-Chun Tsai, Woody Sherman, and Darrin M. York. Improved alchemical free energy calculations with optimized smoothstep softcore potentials. *Journal of Chemical Theory and Computation*, 16(9):5512–5525, 2020.
- [24] Tai-Sung Lee, Bryce K. Allen, Timothy J. Giese, Zhenyu Guo, Pengfei Li, Charles Lin, T. Dwight McGee, David A. Pearlman, Brian K. Radak, Yujun Tao, Hsu-Chun Tsai, Huafeng Xu, Woody Sherman, and Darrin M. York. Alchemical binding free energy calculations in amber20: Advances and best practices for drug discovery. *Journal of Chemical Information and Modeling*, 60(11):5595–5623, 2020.
- [25] Tai-Sung Lee, Hsu-Chun Tsai, Abir Ganguly, Timothy J. Giese, and Darrin M. York. *Robust, Efficient and Automated Methods for Accurate Prediction of Protein-Ligand Binding Affinities in AMBER Drug Discovery Boost*, chapter 7, pages 161–204.
- [26] Lingle Wang, Yuqing Deng, Jennifer L. Knight, Yujie Wu, Byungchan Kim, Woody Sherman, John C. Shelley, Teng Lin, and Robert Abel. Modeling local structural rearrangements using fep/rest: Application to relative binding affinity predictions of cdk2 inhibitors. *Journal of Chemical Theory and Computation*, 9(2):1282–1293, 2013.
- [27] Maximilian Kuhn, Stuart Firth-Clark, Paolo Tosco, Antonia S. J. S. Mey, Mark Mackey, and Julien Michel. Assessment of binding affinity via alchemical free-energy calculations. *Journal of Chemical Information and Modeling*, 60(6):3120–3130, 2020.
- [28] Vytautas Gapsys, Laura Pérez-Benito, Matteo Aldeghi, Daniel Seeliger, Herman van Vlijmen, Gary Tresadern, and Bert L. de Groot. Large scale relative protein ligand binding affinities using non-equilibrium alchemy. *Chemical Science*, 11:1140–1152, 2020.