九元 机器学习的多相难熔合金原子间势 s

Jesper Byggm¨astar,[1,*] Tiago Lopes,[1] 樊哲勇,[2] 和 Tapio Ala-Nissila[3,4]
[1] *Department of Physics, University of Helsinki, Finland*
[2] *College of Physical Science and Technology, Bohai University, Jinzhou, China*
[3] *MSP group, Department of Applied Physics, P.O. Box 15600,Aalto University, FIN-00076 Aalto, Espoo, Finland*
[4] *Interdisciplinary Centre for Mathematical Modelling and Department of Mathematical Sciences, Loughborough University, Loughborough, Leicestershire LE11 3TU, United Kingdom*(日期：2026年3月5日)

人们不断设计和表征新型难熔合金，以满足对优异高温力学性能和稳定性的应用需求。基于原子模拟的计算设计受到原子间势的限制：这些势可能缺少关键元素、精度不足，或对于大规模模拟而言计算速度过慢。本文介绍了一个难熔合金数据库，以及两种计算高效、用途广泛的机器学习势函数（tabGAP 和 NEP）的开发。我们还设计了一种交叉采样策略，利用底层架构截然不同的两种势函数的预测结果，有效采样训练数据。这些势函数支持元素周期表第四至第六族元素（钛、锆、铪、钒、铌、钽、铬、钼、钨）的任意合金成分。该数据库内容丰富且针对多种目标，可用于模拟不同纯金属、固溶体、金属间化合物和玻璃态相中的难熔金属及合金。我们通过再现已知的压力、温度和溶质诱导相变、晶界偏析，以及模拟 WTaCrVHf金属玻璃中的辐照损伤，展示了这些势函数的实用性。

arXiv:2603.04147v1 [cond-mat.mtrl-sci] 4 Mar 2026

I. 引言

近几十年来，合金设计的范围不断扩大，涵盖了更多元素及其各种成分组合。这主要是由于人们对新型浓缩多组元合金（或中高熵合金）的兴趣迅速增长；这类合金将多种金属以显著浓度组合在一起 [1–4]。与由基体金属和少量合金元素组成的传统合金相比，浓缩合金的设计涉及广阔的化学空间，其中包含各种可能的合金成分。在这一化学空间中寻找新型合金，无论是在实验还是计算方面都面临挑战。定量原子级建模的一项关键挑战在于，开展模拟所需的模型开发既困难又复杂。具体而言，必须开发能够支持多种元素以各种组合和浓度进行模拟的原子间势，并且其精度理想情况下应足够高，从而实现真正可靠的模拟和预测。

机器学习原子间势（MLIP）的出现，在精度几乎没有损失的情况下，将计算成本高昂的量子力学计算与可随原子数线性扩展的经典原子模拟连接起来 [5–8]。如今，为各种关注的材料开发机器学习原子间势，已成为庞大的原子材料建模领域中常用的工具。针对单一元素或特定合金和化合物的机器学习原子间势，如今已可作为常规任务，以合理的工作量开发，并根据目标应用进行定制。近年来来，人们广泛致力于开发适用于元素周期表中各类材料的通用机器学习原子间势，以下称为基础势 [9–11]。越来越多的基础势不断涌现，其目标是实现对任意材料的模拟。然而，对于训练数据库未能充分涵盖的新型材料或微观结构特征（例如缺陷、表面、位错、相变），其准确性仍然较差或尚不明确。鉴于此，我们认为，针对特定对象的单元素或少元素机器学习原子间势与基础势之间仍存在巨大差距，值得关注。尽管可以通过微调方法解决基础势在目标应用中的精度不足问题 [12], ，但对于跨越较长时间尺度（纳秒）的大规模模拟（数百万个原子）而言，它们的计算成本仍然过高 [11]。近期的一个例外是 NEP，它旨在兼顾计算效率与对元素周期表中任意材料的迁移能力89 [13], ，但其在目标模拟中的准确性和可靠性仍需验证和微调。另一方面，要扩展现有高效且通用的单元素或少元素机器学习原子间势，使其涵盖更多元素，则需要付出大量努力。

由于高熵合金的重要应用，能够支持多种金属元素的原子间势受到强烈需求，但在关注特定合金成分时仍然稀缺。面向多元素合金的传统原子间势包括Zhou等人 [14], 的嵌入原子法（EAM）势函数、修正嵌入原子法（MEAM）势函数族 [15], 以及Starikov等人 [16]的角相关势等。近五年来，面向多元素合金的多种机器学习原子间势也相继出现，例如若干表格化高斯近似势（tabGAP） [17–20], 矩

* 通讯作者；jesper.byggmastar@helsinki.fi

张量势（MTP） [21], 、谱分析近邻势（SNAP） [22], 、神经进化势（NEP） [23, 24], 以及原子簇展开（ACE）势函数 [25]。这些势函数中的大多数仅支持 $4-6$ 种元素，这在实践中将它们限制于特定的高熵合金，并可能迁移应用于相应的低熵合金。

Zhou 等人提出的 EAM 势 [14] 支持含有 16 种不同元素的金属和合金。最近，研究人员基于 NEP 架构开发了一种支持相同 16 种元素的通用机器学习原子间势（MLIP），称为 UNEP-v1 [23]。两者结合，可高效计算模拟各种合金，灵活性高。然而，EAM 势和 UNEP-v1 均缺少难熔合金的关键元素（铌和铪），且尚不清楚它们能否适用于各种金属间相以及由压力或温度稳定的相的模拟。值得注意的是，铌是最初提出且至今仍受到广泛研究的两种 Senkov 合金 MoNbTaVW 和 HfNbTaTiZr 的组成元素 [3, 26]。近期研究还表明，铪是调控低活化抗辐射合金性能的关键元素 [27]。鉴于此，以及通用机器学习原子间势有必要弥补其与基础势之间的差距，我们在此建立了一个 RHEA 数据库，并为钛、锆、铪、钒、铌、钽、铬、钼、钨及其任意合金开发了两种机器学习原子间势。该 RHEA 数据库包含经过筛选的多样化结构，涵盖所有相关相和成分。我们通过模拟平衡及远离平衡条件下的各种合金和相，展示了这些机器学习原子间势的通用性，其中包括受辐照的百万原子金属玻璃。

II. 结果

A. RHEA 训练数据库

图 1 展示了所开发的 RHEA 数据库示意图。RHEA 训练数据库以我们此前工作中的纯钨、钼、钽、铌和钒数据库中的所有独特结构为基础，结合了 Mo-Nb-Ta-V-W 合金数据库 [28], 、W-Ta-Cr-V 数据库 [17, 18], 和 Hf-Nb-Ta-Ti-Zr 数据库 [20], 。为补充缺失的化学空间和构型空间数据，我们首先对体心立方相中的二元合金进行采样：针对所有元素对，选取三个不同体积和十种不同成分。三元合金也采用了类似的采样方式。理论上，对二元和三元体系进行采样足以填充 tabGAP 和 NEP 所用局部原子环境描述符所覆盖的特征空间，因为前者至多取决于三个原子，而 NEP 描述符本质上取决于原子对 [29], 。在液相采样中，我们利用了这一点；不过，对于晶态合金，我们还针对三种最相关的晶体结构（体心立方、六方密排结构、面心立方），加入了包含最多全部九种元素的等原子比成分 [23]，以确保模型对多元素合金具有良好的精度和迁移能力。

初始训练数据库已对其中包含的所有金属和合金的液相进行了充分采样。为覆盖缺失的化学空间，我们使用 tabGAP 的中间版本（以初始数据库中的所有液相数据训练）制备了压力为 0 吉帕和 30 吉帕的所有三元等原子比组成的熔融合金。此外，我们还制备了压力从 -60 到 100 吉帕、共十种不同压力下的九元熔融合金，以确保机器学习原子间势在极端压力和温度下的稳健性。

RHEA 数据库包含各种成分的已知以及假想金属间相。其中包括三种拉弗斯相 C15、C14、C36、B2 相，以及纯的和稀固溶的钛、锆和铪合金的 ω 相。我们还使用 icet 软件包 [30]枚举了含 $3-5$ 原子的体心立方晶格中所有对称性不同的二元有序相（1908 个独特结构）。Materials Project 数据库中的所有合金也被纳入 RHEA 数据库（少数人工一维结构除外） [31]。通过迭代训练确保对有限温度下的短程至长程有序的额外覆盖：我们使用机器学习原子间势的中间版本，在不同温度和体积下的混合蒙特卡罗分子动力学（MC/MD）模拟中使合金达到平衡。

由于我们希望 RHEA 数据库覆盖温度、压力以及辐照方面的极端条件，因此必须对短距离处的原子间排斥进行采样，并确保其准确且平滑。在极短距离处，两种机器学习原子间势所包含的屏蔽库仑势占主导（见补充信息），但除非有训练数据引导，向近平衡相互作用的过渡并不一定表现良好且准确。对于纯金属，这一点在初始数据库中通过纳入含有不稳定间隙原子、并与近邻原子处于较近（但不过近）接触的结构来实现 [32]。为覆盖缺失的排斥相互作用，我们设计了如下方法。在小型体心立方 16 原子随机九元合金中，将一个原子沿 $\langle 1\,1\,0\rangle$ 方向移动，同时保持其他原子固定，以诱发与近邻原子的强排斥作用。选择 $\langle 1\,1\,0\rangle$ 方向是因为在体心立方晶格中，该原子会接近并穿过两个近邻原子之间，经历一个鞍点处排斥作用达数十电子伏特的能量景观 [33]。这对应于定义阈位移能以及更一般地辐照过程中缺陷形成动力学的能量范围，因此是必须采样的关键能量区间。我们通过枚举并将被位移原子（移至鞍点）及其两个紧邻的近邻原子重命名为九种元素的所有元素三元组（405 个三元组）来创建训练结构。由此得到一套全面的结构，可对辐照诱发反冲事件早期阶段的所有原子对和三元组进行采样。

有关 RHEA 数据库中结构生成的更多详细信息，请参见补充信息。

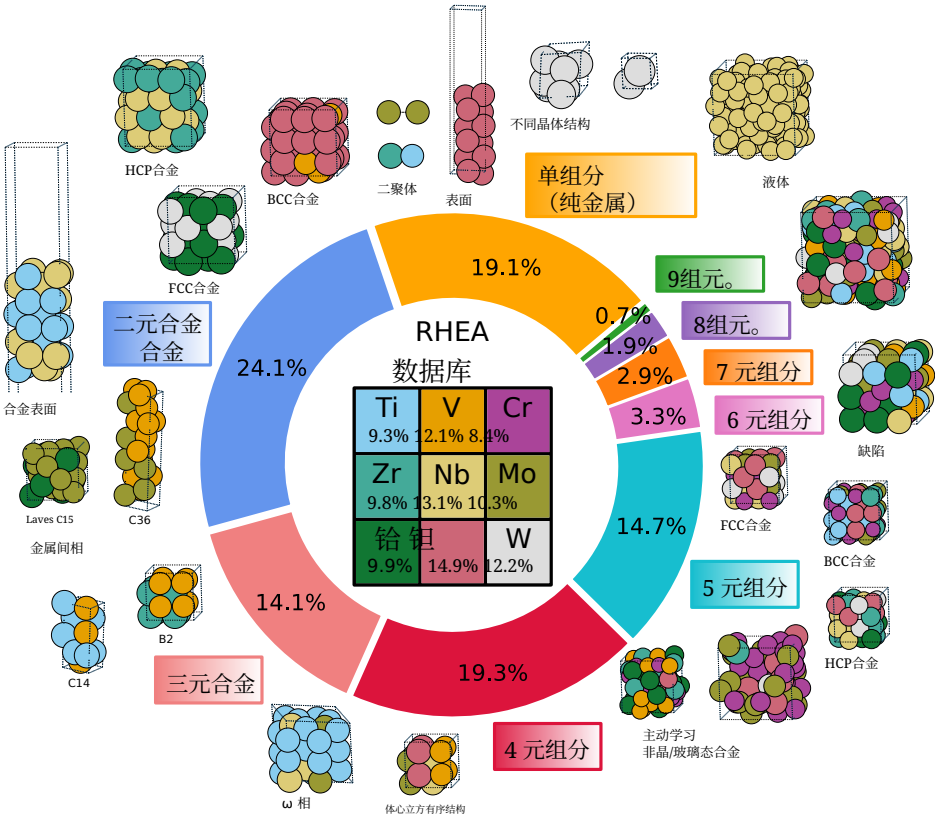

图1：**难熔高熵合金训练数据库**。训练数据库内容示意图，其中包含 22 477 个结构，共计 694 078 个原子。中央的元素网格显示了各元素的原子分数，外侧的环形图则显示了一至九元结构的原子分数，并附有各类结构的代表性快照。

B. 主动学习的交叉采样策略

除了上述训练数据生成方法外，我们还设计并采用了一种交叉采样主动学习策略，通过比较 tabGAP 和 NEP 的预测结果来开展主动学习。主动学习方法通常依赖单个模型提供的不确定性或外推度量（若可用），或者依赖从模型集成中估算出的不确定性。在这里，我们发现，一种有效的策略是利用两种完全不同的机器学习原子间势架构（本例中为 tabGAP 和 NEP）预测结果之间的差异来表征不确定性。

交叉采样的起点是用于采样的大型结构池。我们设计了六类合金结构池，这些结构在手动构建的数据库中缺失或采样不足：(1) 弹性畸变，(2) 点缺陷，(3) 熔体淬火结构（金属玻璃），(4) 熔融结构，(5) MC/MD优化结构，以及 (6) 表面。对于所有类别，我们都创建了包含数百至数千个结构的结构池。每个结构都是随机合金成分，其中包含随机数量的元素（两至九种），并施加所需的改造（例如，对晶胞进行弹性和随机畸变、引入空位或间隙原子），随后在适当的分子动力学模拟中演化（在 NVT 或 NPT 系综中弛豫、熔化并淬火、仅熔化，或进行 MC/MD 弛豫）。每个结构池中，一半结构使用 tabGAP 模拟，另一半使用 NEP 模拟。最后，使用两种机器学习原子间势计算所有结构的能量，并通过选取机器学习原子间势对总能量预测差异最大的 N 个结构来进行采样。随后，使用密度泛函理论计算交叉采样得到的结构，并将其加入训练数据库。使用交叉采样得到的结构重新训练机器学习原子间势后，以迭代方式重复这一过程，直到预测差异降至可接受的幅度（本研究中需要四次迭代）。补充信息中的图 S1 展示了机器学习原子间势之间的预测差异如何随迭代而减小。

在交叉采样过程中，我们监测了每个结构的真实密度泛函理论能量与两种机器学习原子间势能量之间的偏差。我们发现，通常情况下，密度泛函理论能量介于两种机器学习原子间势的预测值之间。通常，用于模拟该结构的机器学习原子间势会低估密度泛函理论能量，而交叉采样得到的结构则是另一种机器学习原子间势能够匹配或明显高估其能量的结构。这凸显了交叉采样策略的优势：它能同时剔除对应于虚假或过深局部极小值的异常结构，以及在任一种机器学习原子间势中能量虚高的异常结构。交叉采样也可以与成熟的模型或集成不确定性预测方法结合，但我们将其留待未来研究。

C. 机器学习原子间势 ls

在构建 RHEA 数据库的过程中，tabGAP 和 NEP 机器学习原子间势经过了迭代改进。最终版本的训练误差和测试误差如图2(a)所示，两种机器学习原子间势的精度相近。两种机器学习原子间势的计算速度与其他机器学习原子间势框架相当或更快；具体而言，它们比嵌入原子法慢一个数量级，与改进的嵌入原子法相当（基准测试结果见补充信息）。两种机器学习原子间势在训练集和测试集 A 上的误差都很低，尤其考虑到能量和力的取值范围很广。测试集 A 由 Mo-Nb-Ta-V-W测试集与 W-Mo训练数据组合而成，其中包含的主要是体心立方合金（有缺陷和无缺陷）以及液体。有趣的是，在测试集 B 上，NEP 的表现优于 tabGAP；测试集 B 中的结构来自 UNEP-v1 的训练数据 [23]。这表明，NEP 可能更擅长对远离训练数据的结构进行外推，也可能是因为这些结构最初由另一种 NEP（UNEP-v1 [23]）模拟生成，因此 NEP 的预测更准确。自然可以通过将测试集 B 中的结构或类似结构加入训练数据库来提升 tabGAP 在这些结构上的表现；但由于 tabGAP（以及 NEP）预测较差的结构能量很高且相当不自然，我们决定不采取这一步。

图2(b)展示了对各种合金能量-体积关系的验证测试，并将密度泛函理论计算结果与表格中的高斯近似势和 NEP 进行比较，涵盖所有二元合金的 B2 相、最稳定二元成分的 C15 拉弗斯相，以及材料项目数据库中的所有三元合金。根据能量-体积数据的极小值和曲率，我们计算了形成能和体积模量，并在图2(c)中将其与密度泛函理论计算结果进行比较。结果表明，机器学习原子间势具有若干令人期待的特性：在较宽的体积和能量范围内，能量-体积预测平滑且表现良好；两种机器学习原子间势都能很好地再现热力学最稳定相的排序；而且两种势预测的形成能和体积模量都相当准确。需要注意的是，虽然训练数据库中包含 B2 相、拉弗斯相和材料项目数据库中的合金，但其中仅包含一个或少数几个接近平衡态的体积。

对纯金属的稳健预测为复现任意合金的性质奠定了良好基础，这也是 RHEA 机器学习原子间势的主要目标。在图3(a)中，我们对机器学习原子间势预测的全部九种纯金属的性质进行了基准测试，并与密度泛函理论的结果以及可能情况下的实验数据进行了比较。图3(a)展示了六种晶相的能量-体积曲线。两种机器学习原子间势都预测出了平滑且表现良好的状态方程，并正确反映了晶相之间的相对稳定性。在钛和锆中，ω 相被正确预测为零温下的基态。最大的差异出现在铬中。这在一定程度上是有意为之。由于铬具有复杂的磁性基态，而训练数据和机器学习原子间势均未考虑这一点（因为缺乏显式的磁自由度），因此我们有意减少了纯铬的相关数据。因此，机器学习原子间势无法准确描述纯铬模拟，但对于含铬的典型浓缩合金，磁矩会受到抑制，此时机器学习原子间势是可靠的 [20, 35, 36]。

图3(b)总结了多种纯金属性质：体积模量、点缺陷形成能、最稳定取向的表面能和晶界能，以及熔化温度。总体而言，与密度泛函理论和实验相比，机器学习原子间势表现良好。一个异常值仍是纯铬的体积模量，相对于对应于铬真实磁性基态的实验值，该值被高估。磁性缺失对铬的其他性质（如熔点和表面能）影响较小。纯金属性质的更多验证见补充信息。

D. 相图与相变

图2-3表明，两种机器学习原子间势在预测零温合金稳定性、晶体相能量学以及纯金属的各种缺陷能和熔点方面均表现良好。我们的下一步目标是证明RHEA 数据库和机器学习原子间势也能再现随温度、压力和成分变化的相关相变。我们首先在图4中展示由tabGAP和NEP预测的纯钛压力-温度相图。该相图通过自由能计算和两相共存模拟绘制（见方法）。纯锆和纯铪的相图也以相同方式绘制，并展示于补充信息中。

图 4 显示，与参考文献 [37] 中以灰色虚线绘出的近似实验相变线相比，两种机器学习原子间势都预测出了定性上正确的相图。tabGAP 在一定程度上高估了三相点和六方密排稳定区的范围，而 NEP 则略微低估了二者，并且与实验线相比，ω-六方密排和 ω-体心立方相变线存在系统性偏移。在低压力下，两种机器学习原子间势都将六方密排-体心立方相变温度高估了 $100-150$ 开尔文。两种机器学习原子间势都预测熔点随压力升高而升高，这符合金属的典型规律。虽然有熔化线的实验数据，但数据差异极大，因此无法用于验证任一种机器学习原子间势 [38]。对于 Zr 和 Hf（见补充信息），相图在定性上也类似，唯一的例外是 NEP 预测在极高压力（>50吉帕）下 fcc 相是 Hf 最稳定的相。对于目标应用，可以通过在感兴趣的压力和温度下采集更多训练数据，进一步微调相图的精度。尽管如此，这里的结果表明，RHEA 数据库已能生成合理的相图；而要通过传统原子间势实现这一点，则需要仔细拟合 [39, 40]。更重要的是，这为再现化学成分函数下的相稳定性奠定了重要基础

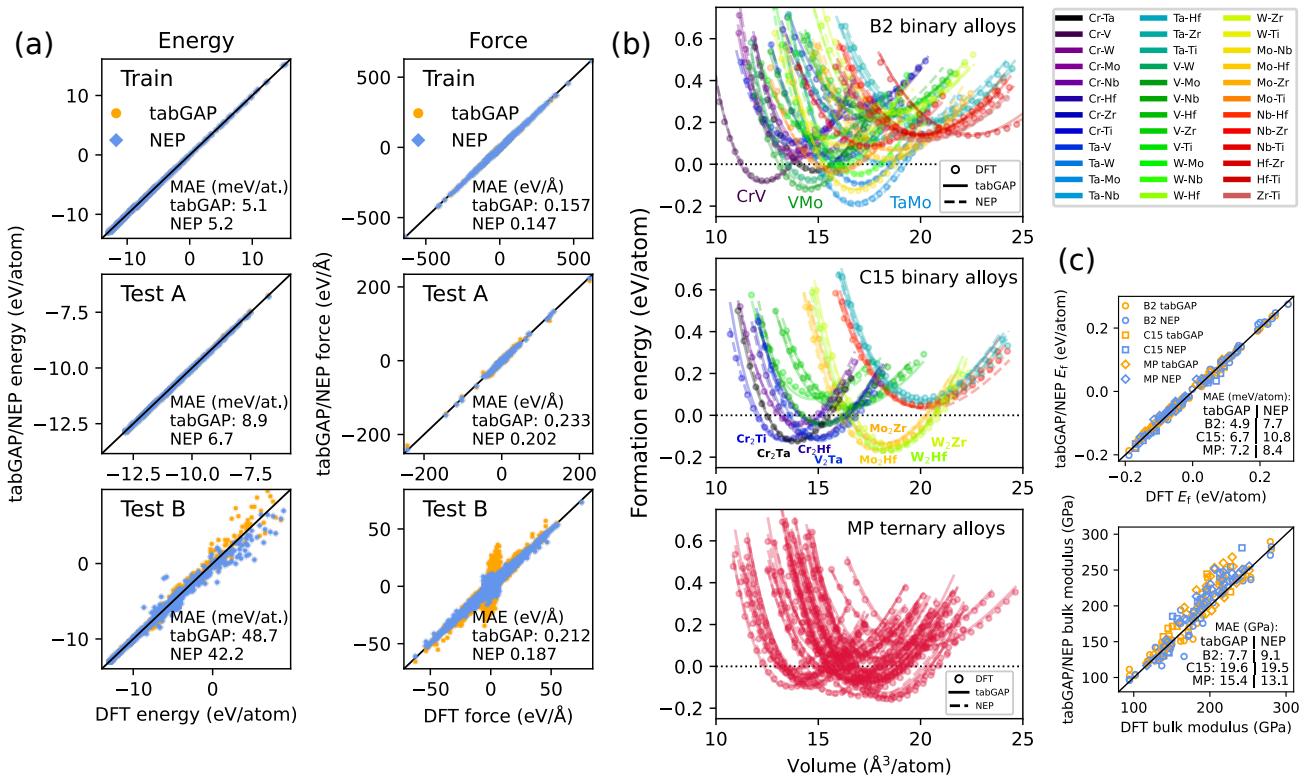

图2：**训练误差和验证误差**。列(a)展示了训练集、测试集 A（参考文献 [18] 中的 Mo-Nb-Ta-V-W 测试集与参考文献 [34] 中的 W-Mo 训练数据的组合）和测试集 B（参考文献 [23] 中训练数据的子集，且包含相同元素）的能量和力分量奇偶图及均方根误差。面板(b)展示了所有二元B2合金、最稳定的C15拉夫斯二元合金以及材料项目数据库中所有可用三元合金的能量-体积曲线，并将 tabGAP 和 NEP 的结果与密度泛函理论进行比较。面板(c)展示了根据(b)中的能量-体积数据提取的平衡形成能和体积模量的奇偶图及均方根误差。

成分。

为研究并验证机器学习原子间势预测的合金相稳定性，我们设计了六种不同的合金相，并对它们进行分子动力学加热-冷却模拟。结果总结于图 5。六种合金分别为：(1) 处于体心立方相的完整九元素体系；(2) 体心立方 MoNbTaVW Senkov合金；(3) 体心立方 HfNbTaTiZr Senkov合金；(4) 六方密排HfTiZr合金；(5) C15 拉维斯相 TaV_2，；(6) $Ti_{90}Nb_{10}$ 的 ω 相。除 TaV_2 拉维斯合金外，所有合金均为固溶体。从 100 开尔文开始，在零压力下，采用 NPT 系综，每种合金在 500 皮秒内加热至 3500 开尔文，随后在 1 纳秒内冷却回 100 开尔文。六种合金都使用两种机器学习原子间势进行模拟。按设计，这些合金在加热-冷却模拟过程中会经历若干次相变。前三种体心立方合金在熔化前一直保持体心立方相稳定。冷却过程中，九元素合金在两种机器学习原子间势下都过冷形成金属玻璃。两种 Senkov 合金都再结晶为体心立方相，或形成完美晶体，或带有残余缺陷。最初为六方密排结构的 HfTiZr 合金熔化后，在两种机器学习原子间势下均于冷却过程中再结晶为混合六方密排/体心立方/面心立方相。TaV_2 拉维斯合金熔化后，在冷却过程中凝固为金属玻璃。最后，最初处于 ω相的 $Ti_{90}Nb_{10}$ 合金未在两种机器学习原子间势中都经历了几次相变。首先，它在低温下便转变为体心立方相，随后一直保持稳定，直至熔化。在冷却过程中，两种机器学习原子间势都先使其再结晶为体心立方和六方密排混合相；而在 NEP 中，它最终又回到ω （但在 tabGAP 中并非如此，推测是因为 NEP 中 ω-Ti 比 tabGAP 中更稳定，见图4）。

以上模拟从多个方面展示了机器学习原子间势的稳健性和准确性。首先，所有观测到的相和相变都合理且符合预期。例如，已知 Senkov 合金是稳定的 BCC合金。$Ti_{90}Nb_{10}$ 转变为体心立方相，这一点已由实验相图证实 [41]。此外，在所有情况下、使用两种机器学习原子间势时，熔化都会导致势能（潜热）明显增加；所有最终形成的玻璃态相的能量也都高于初始晶体相，这符合预期。最后，为了进一步验证结果，我们从 tabGAP分子动力学轨迹中抽取结构，并使用密度泛函理论和 NEP 重新计算能量，以便直接比较。图5(c)显示了能量随时间和温度的变化，可以清楚地看出，对于所有观测到的相和温度下抽取的每个结构，密度泛函理论能量都接近 tabGAP 和 NEP 的能量。

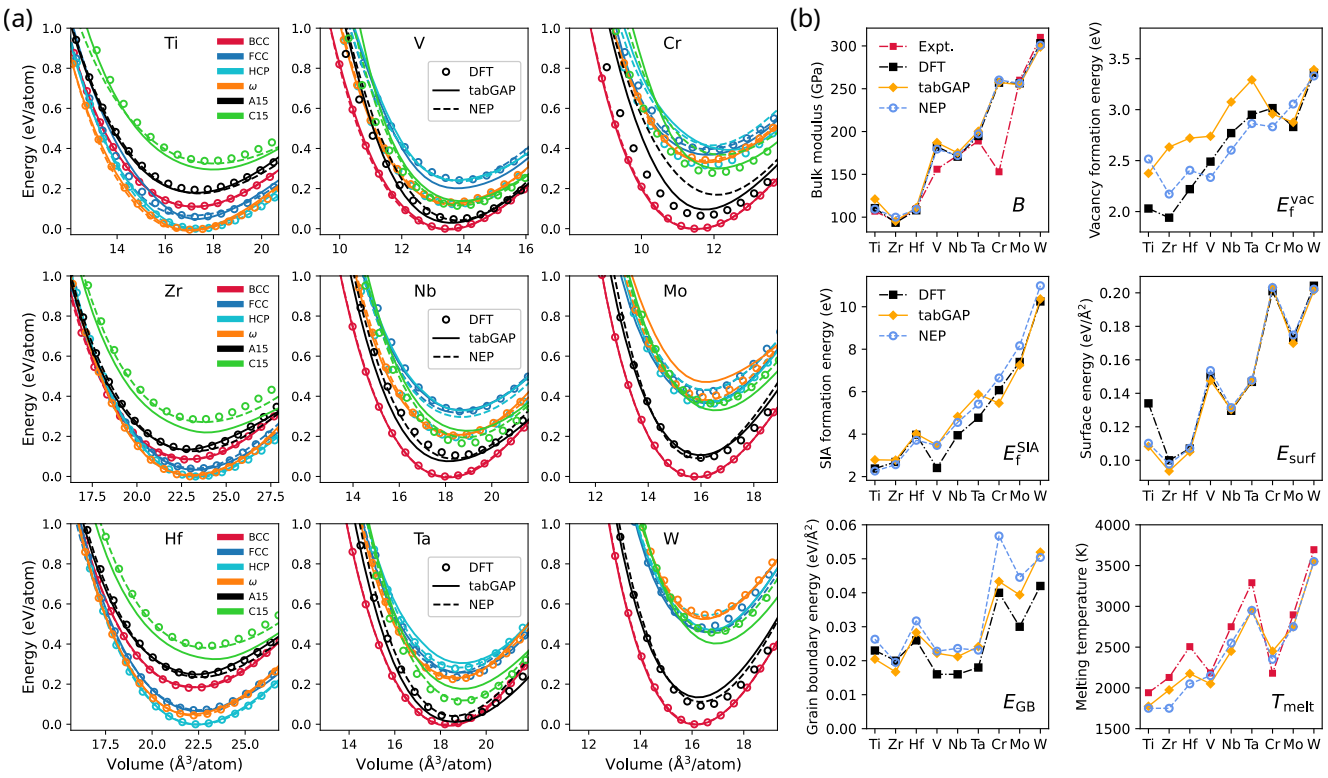

图3：**对九种纯金属的验证**。稳定和假想晶相各向同性体积缩放的能量-体积曲线，并将 tabGAP 和 NEP 与密度泛函理论进行比较。右栏展示了若干材料性质，并将其与密度泛函理论和实验数据进行比较。

E. Zr-Nb相稳定性

作为机器学习原子间势预测的成分依赖相变案例研究，我们聚焦于二元Zr-Nb合金体系。锆合金和钛合金常使用少量钼或铌等体心立方相稳定元素来改善性能或促进向体心立方相（β）转变。值得注意的是，Zr-Nb合金用作核燃料包壳材料；根据温度、成分、辐照和合成方法的不同，已观测到多种稳定相和亚稳相的混合物[43]。图6展示了根据参考文献[42]重新绘制的实验Zr-Nb相图中富锆部分。我们使用 tabGAP 进行了一系列恒温、恒定成分的分子动力学平衡模拟，以定性考察能否重现相图所预期的成分和温度诱导相变。对于恒温情况，我们选取 1000 K，并从铌含量逐渐增加的六方密排结构（α）Zr-Nb固溶体开始（图6中的橙色线）。实验相图表明，随着铌含量增加，α-Zr固溶体应转变为 β-Zr相，并经过一个两相 α-Zr + β-Zr区域。我们的模拟预测，1000 K时铌固溶体临界浓度为 10 at.%。当铌含量达到或超过这一数值时，晶格会变得动力学不稳定并转变为体心立方相，如图6中的快照所示。这在定性上与相图一致；相图显示，在 1000 K 下，超过 12 at.% 时，单相 β-锆为平衡相。

对于恒定浓度，我们选取 10 at.% 铌，并在 600 K 至 1200 K 的温度范围内进行模拟（图 6 中的红线）。这些模拟以α-锆-10铌固溶体为起始结构，并采用混合MC/MD方法，其中通过蒙特卡罗交换原子，促使可能发生热力学有利的偏析和第二相形核。在 10 at.% 铌且温度较低时，实验相图显示 α-锆和 β-铌共存。事实上，在 600 K 时，MC/MD 模拟产生了铌偏析，并形核形成富铌 β相析出物，其周围是几乎纯净的元素态 α-锆。随着温度升高，达到平衡的 β相区域逐渐增大，并包含更多锆，最终整个系统变为单相 α-锆固溶体，如图 6 的快照所示。这一趋势与实验相图一致。需要注意的是，这些分子动力学模拟的时间有限，且具有一定随机性，因此最终结果并不总是相同。例如，图 6 显示，在 800 K 时，α 相和 β 相大小相同；在 900 K 时，系统完全为 β相；而在 1000 K 时，模拟又产生了两相混合物。只有在 1100 K 及以上的温度下，模拟才稳定地呈现完全的 α到β 相变，这与相图所示相符。

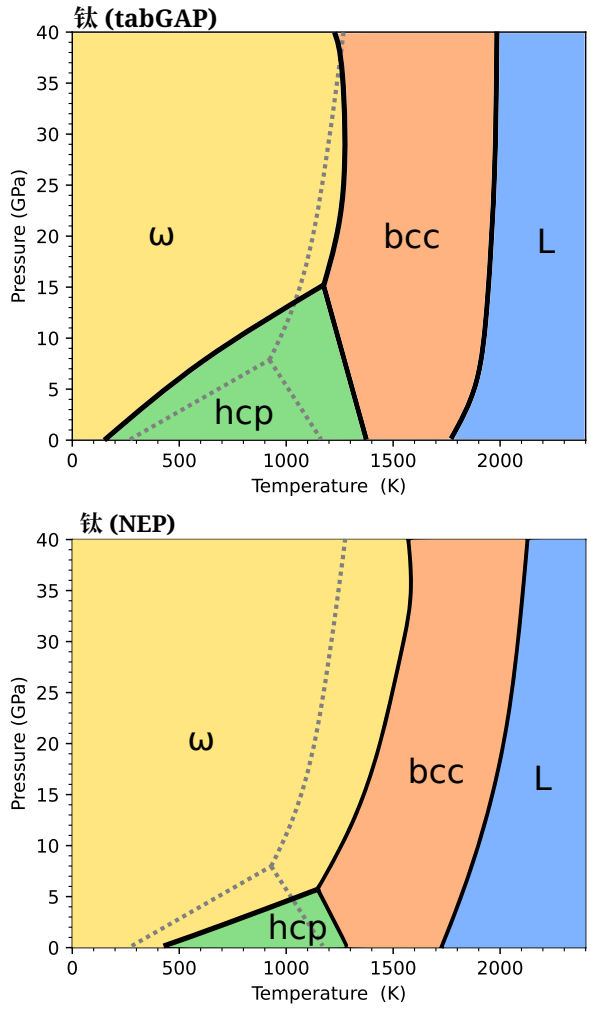

图 4：**钛的压力-温度相图**，基于使用 tabGAP 和 NEP 进行的自由能计算和固液模拟绘制。虚线表示根据实验 [37]绘制的相界。

F. 晶界偏析与实验对比

除了相稳定性和相变模拟外，准确且高效的机器学习原子间势还能用于模拟微观结构缺陷的作用。高浓度难熔合金块体样品的合成往往会导致某些元素偏析到晶界，或形成枝晶结构，使枝晶区域与枝晶间区域的成分存在差异。这可能是由元素间熔点差异较大所致，进而导致再结晶温度不同；也可能是晶界等微观结构特征的热力学驱动力所致 [3, 20, 44, 45]。我们从文献中选取了四种成分复杂的难熔合金，并将实验观察到的偏析趋势与机器学习原子间势的预测结果进行比较：$Mo_{89}Hf_{11}$ [46]，$W_{29.4}Ta_{42}Cr_5V_{16.1}Hf_{7.5}$ [27]，$Mo_{72.3}W_{12.8}Ta_{10}Ti_{2.5}Zr_{2.5}$ [47]，以及 $Hf_{15}Nb_{16}Ta_{17}Ti_{18}V_{20}Zr_{14}$ [48]。所有成分均以原子百分比表示。对于模拟，我们使用 tabGAP，为每种合金构建由取向随机的晶粒和晶界组成的多晶体系，然后在 1000 K 下通过 MC/MD模拟使其达到平衡。这使晶界偏析能够自然形成，并可与实验中报告的枝晶偏析进行定性比较。

图 7 展示了模拟中晶界区域内元素浓度随 MC步数的变化。在所有四种情况下都发生了偏析，平衡后的晶界成分与初始整体成分有明显差异。在二元 $Mo_{89}Hf_{11}$ 合金中，模拟显示铪明显优先偏析至晶界（$22-23$ at.%，而整体浓度为 11 at.%）。实验报告了铪含量为 $16-24$ at.% 的富铪通道，与模拟得到的晶界浓度一致。在 $W_{29.4}Ta_{42}Cr_5V_{16.1}Hf_{7.5}$中，实验也报告了铪偏析。模拟显示铪偏析最为显著，同时还预测钒和铬会偏析至晶界。

在 $Mo_{72.3}W_{12.8}Ta_{10}Ti_{2.5}Zr_{2.5}$中，实验报告显示，W 和 Ta 优先偏析至枝晶区域，Mo 和 Zr 偏析至枝晶间区域，而 Ti 则分布于各处，但略微偏向枝晶间区域。模拟也显示，Mo、Zr 以及部分 Ti 优先偏析至晶界，使晶粒内部保留了较多的 W 和 Ta。在 $Hf_{15}Nb_{16}Ta_{17}Ti_{18}V_{20}Zr_{14}$，中，实验报告显示，主要是 Ta、也有 Nb 偏析至枝晶区域；主要是 Zr、也有部分 Ti、V 和 Hf 偏析至枝晶间区域。同样，模拟预测 Ta 和 Nb 留在晶粒内部，而主要是 Zr、也有部分 V 和 Hf 偏析至晶界，Ti 则分布于各处。

在所有四种情况下，模拟预测的偏析趋势都与实验表征报告的结果一致，即使所考虑的微观结构并不相同（纳米晶粒与枝晶结构）。在现实长度尺度上模拟枝晶结构超出了本研究的范围，但我们认为偏析趋势仍具有可比性。此外，我们还确认，实验中观察到的 MoNbTaVW 样品 [49] 中 Nb 和 V 向晶界偏析的现象，也能通过我们的机器学习原子间势重现，正如先前研究 [20]所做的那样。

G. WTaCrVHf金属玻璃的辐照耐受性

高熵金属玻璃是一类新兴的非晶合金。Tunes等人制备了 $W_{24}Ta_{40}Cr_{18}V_5Hf_{13}$ 薄膜，并发现即使在高温（1173 K 和 1073 K）下经过长期退火和高剂量辐照，它们仍能保持稳定的玻璃相 [50]。为展示机器学习原子间势的多功能性，我们模拟了 $W_{24}Ta_{40}Cr_{18}V_5Hf_{13}$ 从熔体快速冷却过程中的玻璃形成，并

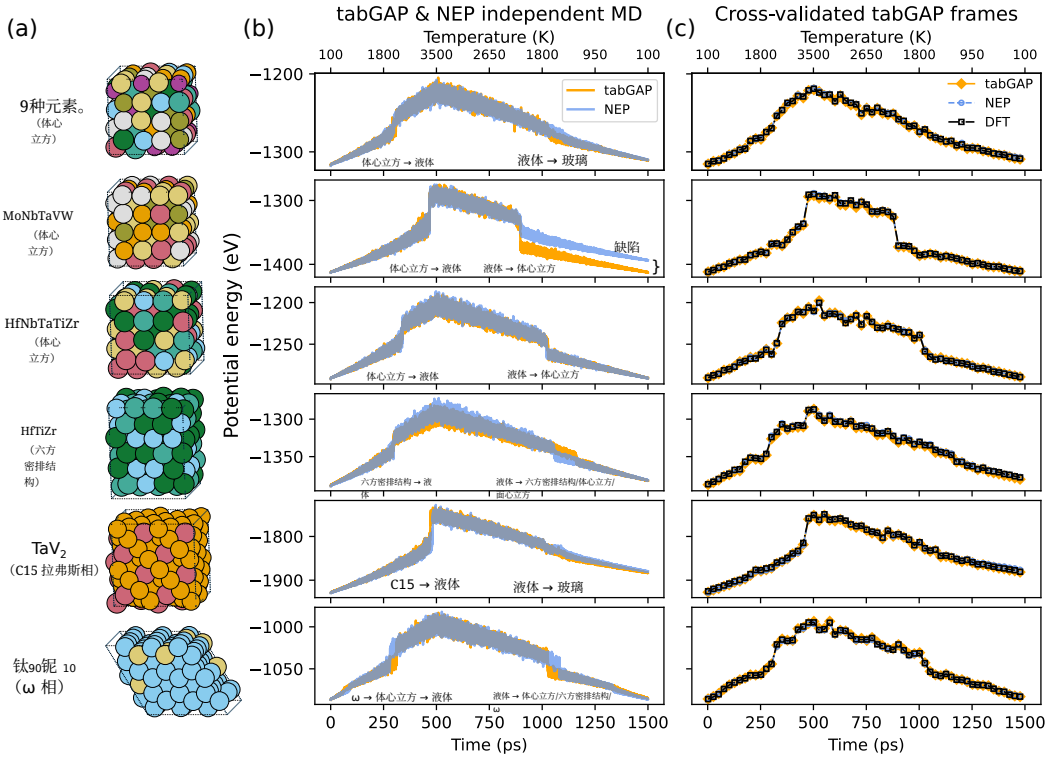

图 5：**加热和冷却过程中的相稳定性验证**。使用 tabGAP 和 NEP 将六种不同合金加热至熔点以上，再冷却至 100 K，期间的势能变化。图中标出了发生的各种相变。列 (c) 展示了对两种 MLIP 的验证：将从 tabGAP 轨迹中采样的结构的能量与密度泛函理论（DFT）结果进行比较。

在连续的碰撞级联模拟中对玻璃态系统进行强辐照。结果总结于图 8。

图 8(a) 展示了对包含 1.024 百万个原子的预制玻璃体系的多种表征。以 10^{12} K/s 的速率从 3100 K 快速冷却至 1100 K 时，势能随温度的变化呈现出明显的玻璃化转变，表现为斜率变化，但没有再结晶迹象。两种 MLIP 在冷却行为和玻璃的局部原子结构方面均给出了非常相似的结果。图 8(a) 还展示了径向分布函数、不同配位数原子数量的直方图（截断半径为 3.8 $^{\circ}$A），以及通过 OVITO [51] 中的公共近邻分析方法识别出的二十面体环境中原子的数量和成分。所有结果均为两种 MLIP 各自三种独立玻璃的平均值，这些玻璃由不同的随机合金起始构型制备。在两种 MLIP 中，配位数为 12 的原子在 $W_{24}Ta_{40}Cr_{18}V_5Hf_{13}$ 玻璃中最常见。二十面体原子的成分表明，与合金整体成分相比，铬和钨更常作为二十面体的中心原子，而以铪和钽为中心的二十面体较少，以钒为中心的二十面体则与整体成分相近。这些趋势在两种 MLIP 中一致。

我们使用 tabGAP 模拟了 50 次连续的 30 keV 碰撞级联，使制备好的玻璃态合金受到强烈辐照。每次碰撞级联均通过向靠近模拟体系中心且朝向中心的一个随机原子赋予 30 keV 的反冲能来启动。图 8(c) 展示了一次碰撞级联的示例，其中显示了某次模拟中最初 ≈ 0.44 ps 内所有动能高于 2 eV 的原子。金属玻璃中的级联在定性上与晶态金属和合金中的高能级联相似，会分裂成亚级联，并产生具有极热类液态核心的热尖峰 [52]。级联诱导熔化和无序会迅速部分恢复，但原子永久位移后仍会留下缺陷。

图 8(d) 将玻璃中的辐射诱导损伤量化为势能变化及不同配位原子数量的变化，并将其表示为连续 30 keV 级联次数的函数。辐照过程中未观察到再结晶，这与实验结果一致 [50]，玻璃结构仅发生细微变化。势能变化通过对每次模拟运行在恒温、零压力弛豫的最后 6 ps 内取平均值获得。配位数的变化则对三次相互独立的模拟运行取平均。图 8(d) 中的结果显示，强辐照使平均原子势能略微增加；经过 50 次级联后，增幅为 4 meV/原子。这相当于在含有百万个原子的玻璃中储存了 4000 eV 的缺陷能量。配位分析表明，辐照损伤主要表现为最受青睐的

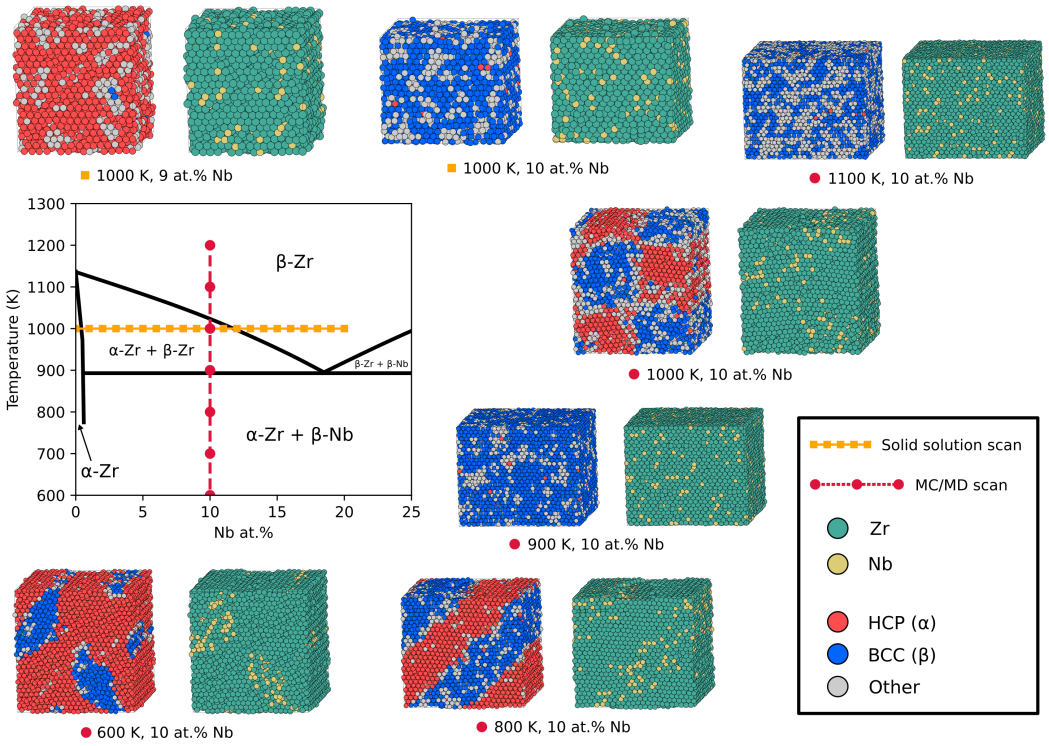

图 6：**Zr-Nb相变**。 Zr-Nb的实验相图（根据 [42]重新绘制）以及恒温（1000 K）、恒定浓度（10 at.% Nb）模拟所得平衡体系的快照。恒温模拟在 NPT系综中对固溶体进行（固溶体扫描），快照展示了hcp Zr-10Nb固溶体向bcc Zr-10Nb固溶体的转变。恒定浓度模拟采用混合MC/MD方法来获得热力学稳定态（MC/MD扫描）。图中展示了不同温度下的快照，说明随着bcc相稳定性增强，体系从两相（富Zr的六方密排结构和富Nb的体心立方相）转变为高温下的单相bcc固溶体。

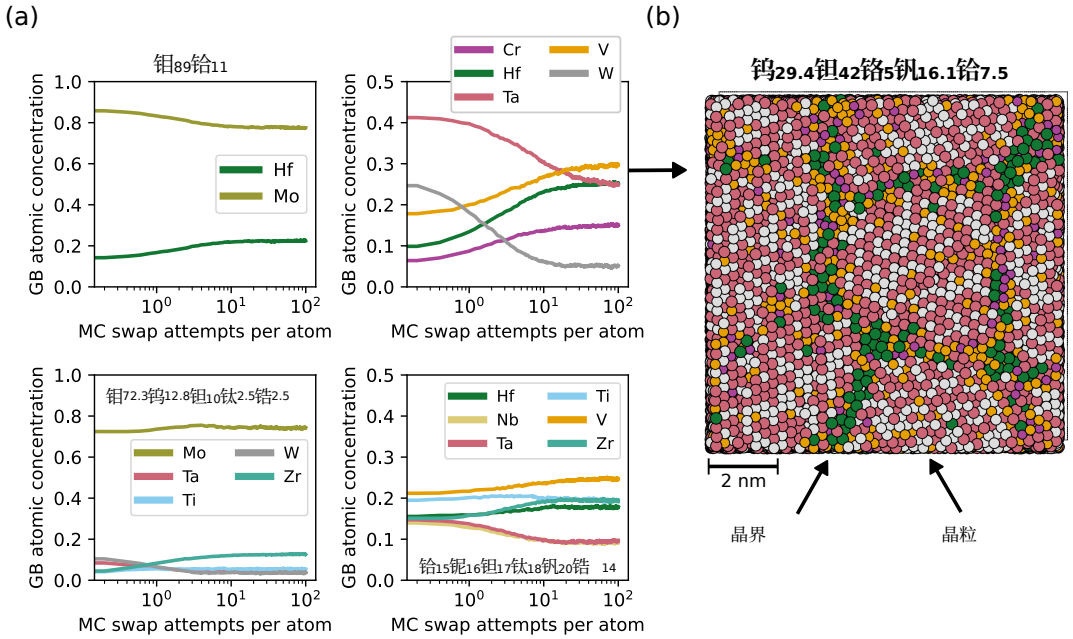

图 7：晶界偏析。 </style]面板 (a) 展示了四种经实验研究的合金中晶界原子的原子浓度随采用tabGAP进行的1000 K MC/MD模拟中MC步数的变化。面板 (b) 展示了最终平衡态$W_{29.4}Ta_{42}Cr_5V_{16.1}Hf_{7.5}$ 合金的快照，可以清晰看到晶界偏析。

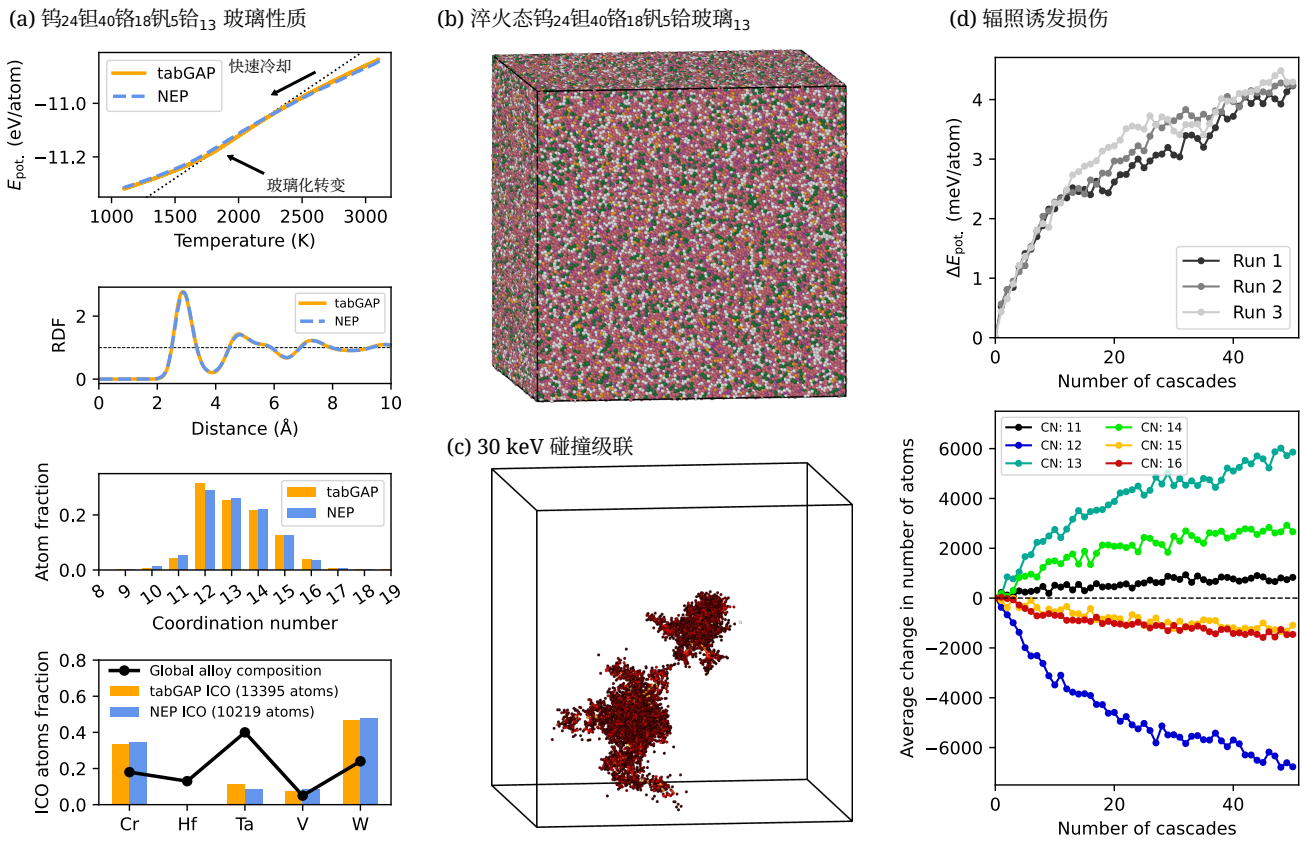

图 8：**一百万原子钨$_{24}$钽$_{40}$铬$_{18}$钒$_{5}$铪$_{13}$ 金属玻璃**。（a）显示从熔体冷却并形成玻璃态结构过程中的势能、最终玻璃的径向分布函数（RDF）、配位数，以及二十面体配位原子的数量和成分。每种机器学习原子间势（MLIP）的所有结果均为三次独立模拟的平均值。（b）显示在 1100 K 下制备的玻璃快照。（c）展示 30 keV 碰撞级联：图中显示级联模拟开始后 0.44 ps 内动能高于 2 eV 的原子，颜色表示动能。（d）显示辐照诱发损伤，即平均势能的变化和配位数平均变化随连续 30 keV 碰撞级联次数的变化。

12 配位原子以及 13 配位和 14 配位原子的相应增加。图 8(d) 中的结果还表明，随着辐照剂量增加，辐照损伤的增长速率逐渐减缓。这与晶态金属和合金中在计算和实验上广泛观察到的损伤饱和现象在定性上相似 [53, 54]。其主要原因是，新的级联会与先前产生的损伤发生重叠；当缺陷浓度达到临界水平后（该水平 strongly 取决于材料和辐照条件 [54, 55]），产生的新损伤会减少并最终不再产生。

III. 讨论

我们构建了一个难熔金属和合金结构多样化数据库，其中包含元素周期表第九族 $4-6$ 元素，并训练了两种计算效率较高的机器学习原子间势（tabGAP 和 NEP）。构建 RHEA 数据库的目标是在多样性与针对性结构之间取得平衡，使训练得到的机器学习原子间势既适用于通用用途，又能准确描述各种特定的金属间相。预计在例如对难熔合金的基础势进行微调，或开发包含更多元素的大规模合金机器学习原子间势时，RHEA 数据库将成为更大型数据库中稳健的一部分。通过复现基本材料性质，尤其是实验已知的、随压力、温度和合金成分变化的各种相变（包括液相和玻璃态相），我们展示了 RHEA 数据库和训练得到的机器学习原子间势的通用性与实用性。两种机器学习原子间势还通过屏蔽库仑势纳入了短距离处准确的排斥相互作用，从而能够模拟由这九种元素以任意合金成分组成的材料中的辐照损伤。作为示例，我们模拟了实验研究过的 $W_{24}Ta_{40}Cr_{18}V_5Hf_{13}$ 金属玻璃中的辐照损伤累积，观察到玻璃相具有很高的稳定性，这也与实验观察结果一致。

我们的工作还表明，在机器学习原子间势开发和应用的各个阶段——从生成训练数据，到验证并评估模拟结果的可靠性——训练和使用基于完全不同架构的两种机器学习原子间势，是一种强有力的策略。在生成训练数据时，我们广泛采用了交叉采样策略，对两种机器学习原子间势预测结果存在分歧的结构进行采样。这种方法可作为机器学习原子间势不确定性量化和主动学习的简单补充；后两者通常依赖于单一机器学习原子间势框架内的集成或外推度量。在模拟中应用时，可通过交叉核验两种机器学习原子间势是否得出相近结果，来确认结果的有效性。

我们认为，所开发的机器学习原子间势在通用性、准确性和计算效率之间实现了实用的平衡。这使它们处于一个广阔但尚未得到充分探索的领域：介于目前已常规开发的、功能有限的单元素或少元素机器学习原子间势，与数量日益增多的大型且极具通用性的机器学习原子间势之间。目的在于构建基础势。tabGAP 和 NEP 的计算速度都足够快，能够以具有竞争力的精度，低成本地进行涉及数百万个原子的模拟，时间尺度达纳秒。因此，这些机器学习原子间势为在近平衡和远离平衡条件下，对 Ti–Zr–Hf–V–Nb–Ta–Cr–Mo–W 成分丰富化学体系中的难熔合金进行模拟，提供了宝贵的工具。

IV. 方法

A. 密度泛函理论计算

用于训练和测试数据的密度泛函理论计算采用 VASP code [56–59]代码 [60, 61] 及投影缀加波势（Cr 和 Ta 使用 5.4 pv 版本，其他所有元素使用 sv 版本）。所有计算均采用一致的输入参数，包括 PBE GGA 交换-相关泛函、 [62], 500电子伏特的平面波截断能、15 $^{\circ}0.^{-1}$ Å 的最大 k点间距，并使用以 Γ为中心的 Monkhorst-Pack 网格 [63], 和 0.1 电子伏特的 Methfessel-Paxton 展宽 [64]。这些参数与我们先前工作中所用的相同 [17]。计算中未采用自旋极化，这意味着无法得到纯铬和稀铬合金的正确磁性基态。不过，纯铬和铬合金不同磁序之间弛豫后的能量差很小 [20]。

B. 表格化高斯近似势

正如我们在先前工作中所述 [17, 18], ，tabGAP 通过先使用 QUIP 代码 [6] 训练高斯近似势（GAP）来拟合，其中仅采用局部原子环境的低维描述符（二体、三体和 EAM密度 [18]），并将预先拟合的屏蔽库仑排斥势作为外部项（见补充信息）。低维特性使得 GAP 能够计算能量预测值，并将其制成一维和三维网格表，保存为预计算的 tabGAP 势函数文件。随后，通过针对 LAMMPS 的实现（https://gitlab.com/jezper/tabgap），利用三次样条插值 [17, 65, 66]即可高效完成评估。

用于生成 tabGAP 的主要 GAP 输入参数为：二体和 EAM 描述符的截断半径为 5 $^{\circ}$Å，三体描述符的截断半径为 4.1 $^{\circ}$Å。较短的三体截断半径提升了计算速度，且不会明显牺牲精度。GAP 使用稀疏高斯过程回归 [6, 67]。二体和 EAM 项的稀疏点数量为 20。对于三体描述符，我们为纯元素三元组使用了比混合元素三元组更多的稀疏点（400 对比 $80-148$），因为训练结构中纯元素三元组远多于混合元素三元组。需要注意的是，虽然稀疏点数量会直接影响 GAP 的计算速度，但由于 GAP 能量已预先计算并存储在文件中，因此不会影响 tabGAP 的速度。不过，在我们的案例中，稀疏点数量的上限实际上受到 QUIP 代码中数组最大尺寸的限制。

C. 神经进化势

NEP 使用 GPUMD 代码 [68]进行训练，训练期间采用与 tabGAP 相同的预拟合库仑排斥势（参见 SI） [69]。NEP 使用可分离自然进化策略优化人工前馈神经网络的参数，该网络具有一个隐藏层 [70]。局部原子环境的描述符分为径向和角向（三体、四体、五体）分量，在数学上与原子簇展开相似但并不相同 [8]。在迭代训练数据生成过程中，我们并行训练了大型和小型 NEP 模型，两者由 $n_{\max}$ 参数（小型为 4 4，大型为 8 8）区分。两种模型的 $l_{\max}$ 参数均为 4 2 1，截断半径 5 $^{\circ}$A，神经元数均为 80。在最终验证中，我们发现大型 NEP 模型在精度和速度之间取得了更好的平衡（精度显著高于小型模型，而速度并未显著降低）。因此，本文仅展示大型 NEP 模型，并舍弃小型模型。

D. 分子静力学与动力学模拟

使用机器学习原子间势进行的分子静力学与动力学模拟，采用 LAMMPS [71] （tabGAP 和 NEP）和 GPUMD [68]（仅 NEP）完成。所有 tabGAP 模拟和使用 NEP 进行的小规模计算均采用 LAMMPS。使用 GPUMD 在 GPU 上运行采用 NEP 的大规模（金属玻璃）模拟。

E. 自由能计算

我们通过自由能计算绘制纯钛、锆和铪的压力-温度相图。我们采用 Freitas 等 [72], 详细描述的方法，具体而言是 Frenkel-Ladd（FL）方法。FL 方法通过热力学积分和解析已知的参考相，计算亥姆霍兹自由能及其振动熵贡献。对于固相，我们遵循 Freitas 等 [72] 的做法，采用爱因斯坦晶体作为参考相。所需的模拟参数通过收敛性测试确定。切换时间为 25 ps，模拟盒大小为 $6000-7000$ 个原子（取决于相），所有模拟的时间步长均为 1 fs。得到亥姆霍兹自由能 F后，吉布斯自由能计算为 $G=F+PV$。

通过暴力采样（P, T）个点，并以 2 吉帕和 50 开尔文为增量，计算每个相关晶体相的吉布斯自由能，绘制出了相图。首先在所需的压力和温度下，使用 NPT系综使体系达到平衡；随后根据需要，在 NVT 系综中进行自由能计算。通过在不同压力下对整个模拟盒进行充分弛豫后，计算焓 $H=U+PV$，其中 U 为势能，从而准确确定了零温度相变。

固液相界面由不同压力下的传统固液两相界面模拟确定 [73] 。

F. 混合蒙特卡洛分子动力学 ics

使用 LAMMPS 对 Zr–Nb 相和晶界偏析进行了混合 MC/MD模拟。每 10 个分子动力学步交换的原子数设为原子总数的 1%。分子动力学积分在NPT系综中进行，持续 10^5 个时间步；结果证明，这足以完成偏析或相分解。对于晶界偏析模拟，我们以尺寸为 $10 \times 10 \times 10$ nm（67 000 个原子）的立方体心立方随机合金体系为起始结构，其中包含四个取向随机、尺寸相近的晶粒，这些晶粒使用 Atomsk [74] 创建。使用 OVITO [51] 中的晶粒分割算法识别晶界原子。

G. 辐照损伤模拟

通过使用 LAMMPS 模拟连续发生的 50 次碰撞级联，实现了金属玻璃中的辐照损伤。每次模拟的初级反冲能均为 30 keV，并赋予一个原子；该原子位于距盒子中心 60°A 处，沿随机方向朝盒子中心运动。这样，碰撞级联便可在盒子中心演化；该区域在 NVE 系综中模拟，而边界周围厚度为 8 °A 的区域则处于 NVT，以耗散热量。级联模拟采用自适应时间步长，以准确追踪高能原子轨迹。未考虑电子阻止造成的能量损失。每次级联的模拟总时间为 100 ps。随后，在 NPT系综中对整个系统弛豫 10 ps，以释放应力并使温度稳定在所需的 1100 K。之后，系统在三个维度上随机平移，超出周期性边界的原子会被包回模拟胞中。这样做是为了让下一次级联在新的随机位置启动，同时仍位于模拟胞的中心。

数据可用性

RHEA 结构数据库、密度泛函理论和势函数训练的输入文件，以及 tabGAP 和 NEP 势函数文件均可在参考文献 [75] 中公开获取。

致谢

J.B. 和 T.L. 获得了芬兰研究理事会通过 OCRAMLIP 项目提供的资助，项目资助编号为 354234。ZF 获得了辽宁省教育厅科学基金的资助（编号：LJ232510167001）。T.A-N. 获得了芬兰科学院通过欧盟“下一代欧盟”计划提供的部分资助，资助编号为 353298，以及编号为 370057 和 373647 的资助。谨此感谢 CSC 信息技术科学中心有限公司提供计算资源。

[1] M.-H. Tsai 和 J.-W. Yeh, 《材料研究快报》**2**, 107（2014）。[2] D. B. Miracle 和 O. N. Senkov, 《材料学报》**122**, 448（2017）。[3] O. N. Senkov、G. B. Wilks、D. B. Miracle、C. P. Chuang 和 P. K. Liaw, 《金属间化合物》**18**, 1758（2010）。[4] D. B. Miracle, 《MRS 公报》**50**, 1200（2025）。[5] J. Behler 和 M. Parrinello, 《物理评论快报》**98**, 10.1103/PhysRevLett.98.146401（2007）。[6] A. P. Bart´ok、M. C. Payne、R. Kondor 和 G. Cs´anyi, 《物理评论快报》**104**, 10.1103/Phys- RevLett.104.136403（2010）。[7] A. V. Shapeev, 《多尺度建模与仿真》**14**, 1153（2016）。[8] R. Drautz, 《物理评论B辑》**99**, 10.1103/Phys-RevB.99.014104（2019）。[9] C. Chen 和 S. P. Ong, 《自然计算科学》**2**, 718（2022）。[10] B. Deng、P. Zhong、K. Jun、J. Riebesell、K. Han、C. J. Bartel 和 G. Ceder, 《自然机器智能》**5**, 1031（2023）。[11] I. Batatia、P. Benner、Y. Chiang、A. M. Elena、D. P. Kov´acs、J. Riebesell、X. R. Advincula、M. Asta、M. Avaylon、W. J. Baldwin、F. Berger、N. Bernstein、A. Bhowmik、F. Bigi、S. M. Blau、V. Cˇarare、M. Ce- riotti、S. Chong、J. P. Darby、S. De、F. Della Pia、V. L. Deringer、R. Elijoˇsius、Z. El-Machachi、E. Fako、F. Falcioni、A. C. Ferrari、J. L. A. Gardner、M. J. Gawkowski、A. Genreith-Schriever、J. George、R. E. A. Goodall、J. Grandel、C. P. Grey、P. Grigorev、S. Han、W. Handley、H. H. Heenen、K. Hermansson、C. H. Ho、S. Hofmann, C. Holm, J. Jaafar, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, J. R. Ker- mode, P. Kourtis, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, C. Lin, J. T. Margraf, I.- B. Magdˇau, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. A. M. Ros- set, L. L. Schaaf, C. Schran, B. X. Shi, E. Sivonxay, T. K. Stenczel, C. Sutton, V. Svahn, T. D. Swinburne, J. Tilly, C. van der Oord, S. Vargas, E. Varga-Umbrich, T. Vegge, M. Vondr´ak, Y. Wang, W. C. Witt, T. Wolf, F. Zills, and G. Cs´anyi, 《化学物理杂志》**163**, 184110 (2025). [12] M. Radova, W. G. Stark, C. S. Allen, R. J. Maurer, and A. P. Bart´ok, npj计算材料**11**, 237 (2025). [13] T. Liang, K. Xu, E. Lindgren, Z. Chen, R. Zhao, J. Liu, E. Berger, B. Tang, B. Zhang, Y. Wang, K. Song, P. Ying, N. Xu, H. Dong, S. Chen, P. Erhart, Z. Fan, T. Ala-Nissila, and J. Xu, NEP89：适用于涵盖89种元素的无机和有机材料的通用神经进化势 (2025), arXiv:2504.21286 [凝聚态物质]。[14] X. W. Zhou, R. A. Johnson, and H. N. G. Wadley, 《物理评论B》**69**, 144113 (2004). [15] W.-M. Choi, Y. H. Jo, S. S. Sohn, S. Lee, and B.-J. Lee, npj计算材料**4**, 1 (2018). [16] S. Starikov, P. Grigorev, S.-H. Lee, Z. Xie, and P. A. T. Olsson, 《计算材料科学》**262**, 114369 (2026)。

[17] J. Byggm¨astar, K. Nordlund, and F. Djurabekova, 《物理评论B》104, 104101 (2021). [18] J. Byggm¨astar, K. Nordlund, and F. Djurabekova, 《物理评论：材料》6, 083801 (2022). [19] A. Fellman, J. Byggm¨astar, F. Granberg, F. Djurabekova, and K. Nordlund, 使用机器学习原子间势研究Al$ x$CrCuFeNi$ y$合金的辐照损伤和相稳定性 (2025), arXiv:2503.07344 [凝聚态物质]。 [20] J. Byggm¨astar, D. Sobieraj, J. S. Wr´obel, D. K. Schreiber, O. El-Atwani, E. Martinez, and D. Nguyen- Manh, 《材料学报》296, 121276 (2025). [21] S. Yin, Y. Zuo, A. Abu-Odeh, H. Zheng, X.-G. Li, J. Ding, S. P. Ong, M. Asta, and R. O. Ritchie, 《自然-通讯》12, 4873 (2021). [22] X.-G. Li, C. Chen, H. Zheng, Y. Zuo, and S. P. Ong, npj计算材料6, 1 (2020). [23] K. Song, R. Zhao, J. Liu, Y. Wang, E. Lindgren, Y. Wang, S. Chen, K. Xu, T. Liang, P. Ying, N. Xu, Z. Zhao, J. Shi, J. Wang, S. Lyu, Z. Zeng, S. Liang, H. Dong, L. Sun, Y. Chen, Z. Zhang, W. Guo, P. Qian, J. Sun, P. Erhart, T. Ala-Nissila, Y. Su, and Z. Fan, 《自然-通讯》15, 10208 (2024). [24] J. Liu, J. Byggm¨astar, Z. Fan, B. Bai, P. Qian, and Y. Su, 《核材料学杂志》616, 156004 (2025). [25] W. Wang, P. Kumar, D. H. Cook, F. Walsh, B. Zhang, P. P. P. O. Borges, D. Farkas, R. O. Ritchie, and M. Asta, npj计算材料11, 330 (2025). [26] O. N. Senkov, J. M. Scott, S. V. Senkova, D. B. Miracle, and C. F. Woodward, 《合金与化合物杂志》509, 6043 (2011). [27] O. El Atwani, H. T. Vo, M. A. Tunes, C. Lee, A. Al- varado, N. Krienke, J. D. Poplawsky, A. A. Kohnert, J. Gigax, W.-Y. Chen, M. Li, Y. Q. Wang, J. S. Wr´obel, D. Nguyen-Manh, J. K. S. Baldwin, O. U. Tukac, E. Aydogan, S. Fensin, and E. Martinez, 《自然-通讯》14, 2516 2023( )。 [28] J. Byggm¨astar, K. Nordlund, and F. Djurabekova, 《物理评论：材料》4, 093802 (2020). [29] T. Lopes and J. Byggm¨astar, 一种针对HfNbTaTiZr高熵合金体系的机器学习原子间势 (2026)，准备中。 [30] M. ˚Angqvist, W. A. Mu˜noz, J. M. Rahm, E. Fransson, C. Durniak, P. Rozyczko, T. H. Rod, and P. Erhart, 《先进理论与模拟》2, 1900015 (2019). [31] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, 《APL材料》1, 011002 (2013). [32] J. Byggm¨astar, A. Hamedani, K. Nordlund, and F. Djurabekova, 《物理评论B》100, 144105 (2019). [33] J. Byggm¨astar, F. Djurabekova, and K. Nordlund, 《物理评论：材料》8, 115406 (2024). [34] G. Nikoulis, J. Byggm¨astar, J. Kioseoglou, K. Nord- lund, and F. Djurabekova, 《物理学杂志：凝聚态物质》33, 315403 (2021). [35] S. Zhao, 《材料科学与技术杂志》44, 133 2020( )。 [36] O. El-Atwani, A. Alvarado, K. Unal, S. Fensin, J. Hinks, G. Greaves, J. K. S. Baldwin, S. Maloy, and E. Martinez, 《今日材料：能源》, 100599 (2020). [37] A. Dewaele, V. Stutzmann, J. Bouchet, F. Bottin, F. Occelli, and M. Mezouar, 《物理评论B》91, 134108 (2015). [38] V. Stutzmann, A. Dewaele, J. Bouchet, F. Bottin, and M. Mezouar, 《物理评论B》92, 224110 (2015). [39] R. G. Hennig, T. J. Lenosky, D. R. Trinkle, S. P. Rudin, and J. W. Wilkins, 《物理评论B》78，054121（2008）。[40] M. I. Mendelev, T. L. Underwood, and G. J. Ackland, 《化学物理学杂志》145，154102（2016）。[41] J. L. Murray, 《合金相图通报》2，55（1981）。[42] J. P. Abriata and J. C. Bolcich, 《合金相图通报》3，34（1982）。[43] V. Tuli, P. Burr, A. Claisse, and C. Cazorla, 《物理评论：材料》7, 113607（2023）。[44] L. Li, Z. Li, A. Kwiatkowski da Silva, Z. Peng, H. Zhao, B. Gault, and D. Raabe, 《材料学报》178, 1（2019）。[45] S. Ma, W. Liu, Q. Li, J. Zhang, S. Huang, Y. Xiong, B. Xu, T. Yang, and S. Zhao, 《材料学报》264, 119537（2024）。[46] O. N. Senkov, S. I. Rao, G. H. Balbus, and B. W. J. McArthur, 《材料科学与工程：A辑》922, 147625（2025）。[47] G. Ouyang, P. Singh, R. Su, D. D. Johnson, M. J. Kramer, J. H. Perepezko, O. N. Senkov, D. Miracle, and J. Cui，npj计算材料9，1（2023）。[48] M. C. Gao, B. Zhang, S. Yang, and S. M. Guo, 《冶金与材料学报A辑》47，3333（2016）。[49] A. Liski, M. Kemell, T. Vuorihei mo, K.-K. Tseng, T.-E. Shen, C.-W. Tsai, J.-W. Yeh, and K. Mizohata, 《核材料与能源》42，101870（2025）。[50] M. A. Tunes, H. T. Vo, J. K. S. Baldwin, T. A. Saleh, S. J. Fensin, and O. El-Atwani, 《今日应用材料》32，101796（2023）。[51] A. Stukowski, 《材料科学与工程建模与模拟》18, 015012（2010）。[52] K. Nordlund, S. J. Zinkle, A. E. Sand, F. Granberg, R. S. Averback, R. E. Stoller, T. Suzudo, L. Malerba, F. Banhart, W. J. Weber, F. Willaime, S. L. Dudarev, and D. Simeone, 《核材料学杂志》512，450（2018）。[53] F. Granberg, J. Byggm¨astar, and K. Nordlund, 《核材料学杂志》556，153158（2021）。[54] G. Wei, J. Byggm¨astar, J. Cui, K. Nordlund, J. Ren, and F. Djurabekova, 《材料学报》274，119991（2024）。[55] M. Boleininger, D. R. Mason, A. E. Sand, and S. L. Dudarev, 《科学报告》13，1684（2023）。[56] G. Kresse and J. Hafner, 《物理评论B》47，558（1993）。[57] G. Kresse and J. Hafner, 《物理评论B》49，14251（1994）。[58] G. Kresse and J. Furthm¨uller, 《计算材料科学》6，15（1996）。[59] G. Kresse and J. Furthm¨uller, 《物理评论B》54, 11169（1996）。[60] P. E. Bl¨ochl, 《物理评论B》50, 17953（1994）。[61] G. Kresse and D. Joubert, 《物理评论B辑》59，1758（1999）。[62] J. P. Perdew, K. Burke, and M. Ernzerhof, 《物理评论快报》77，3865（1996）。[63] H. J. Monkhorst and J. D. Pack, 《物理评论B》13，5188（1976）。[64] M. Methfessel and A. T. Paxton, 《物理评论B》40，3616（1989）。[65] A. Glielmo, C. Zeni, and A. De Vita, 《物理评论B辑》97，10.1103/PhysRevB.97.184307（2018）。[66] J. Vandermause, S. B. Torrisi, S. Batzner, Y. Xie, L. Sun, A. M. Kolpak, and B. Kozinsky，npj计算材料6，1（2020）。[67] A. P. Bart´ok and G. Cs´anyi, 《国际量子化学杂志》115, 1051（2015）。

[68] K. Xu, H. Bu, S. Pan, E. Lindgren, Y. Wu, Y. Wang, J. Liu, K. Song, B. Xu, Y. Li, T. Hainer, L. Svens- son, J. Wiktor, R. Zhao, H. Huang, C. Qian, S. Zhang, Z. Zeng, B. Zhang, B. Tang, Y. Xiao, Z. Yan, J. Shi, Z. Liang, J. Wang, T. Liang, S. Cao, Y. Wang, P. Ying, N. Xu, C. Chen, Y. Zhang, Z. Chen, X. Wu, W. Jiang, E. Berger, Y. Li, S. Chen, A. J. Gabourie, H. Dong, S. Xiong, N. Wei, Y. Chen, J. Xu, F. Ding, Z. Sun, T. Ala-Nissila, A. Harju, J. Zheng, P. Guan, P. Er- hart, J. Sun, W. Ouyang, Y. Su, and Z. Fan, 《材料基因组工程进展》**3**，e70028（2025）。[69] J. Liu, J. Byggm¨astar, Z. Fan, P. Qian, and Y. Su, 《物理评论B》**108**，054312（2023）。[70] Z. Fan, Y. Wang, P. Ying, K. Song, J. Wang, Y. Wang, Z. Zeng, K. Xu, E. Lindgren, J. M. Rahm, A. J. Gabourie, J. Liu, H. Dong, J. Wu, Y. Chen, Z. Zhong, J. Sun、P. Erhart、Y. Su 和 T. Ala-Nissila，《化学物理学杂志》**157**，114801（2022）。[71] A. P. Thompson、H. M. Aktulga、R. Berger、D. S. Bolin-tineanu、W. M. Brown、P. S. Crozier、P. J. in 't Veld、A. Kohlmeyer、S. G. Moore、T. D. Nguyen、R. Shan、M. J. Stevens、J. Tranchida、C. Trott 和 S. J. Plimp-ton，《计算物理通讯》**271**，108171（2022）。[72] R. Freitas、M. Asta 和 M. de Koning，《计算材料科学》**112**，333（2016）。[73] J. R. Morris、C. Z. Wang、K. M. Ho 和 C. T. Chan，《物理评论B》**49**，3109（1994）。[74] P. Hirel，《计算物理通讯》**197**，212（2015）。[75] J. Byggm¨astar，RHEA 数据库以及 tabGAP/NEP 势文件（2026）。

《九元机器学习原子间势：多相难熔合金》的补充信息

Jesper Byggm¨astar、[1]、*Tiago Lopes、[1]樊哲勇、[2]和 Tapio Ala-Nissila[3,4][1]*Department of Physics*、*University of Helsinki*、*Finland*[2]*College of Physical Science and Technology*、*Bohai University*、*Jinzhou*、*China*[3]*MSP group*、*Department of Applied Physics*、*P.O. Box 15600*、*Aalto University*、*FIN-00076 Aalto*、*Espoo*、*Finland*
[4]*Interdisciplinary Centre for Mathematical Modelling and Department of Mathematical Sciences*、*Loughborough University*、*Loughborough*、*Leicestershire LE11 3TU*、*United Kingdom*（日期：2026年3月5日）

目录

S1. 训练数据描述 S2. 交叉采样收敛 2
S3. 计算效率 S4. 排斥势 S5. 纯金属 4
性质 S6. 相图与自由能 参考文献 5
8
10
13
17

* 通讯作者；jesper.byggmastar@helsinki.fi

S1. 训练数据描述

正文对 RHEA 数据库的内容进行了说明和简要介绍。此处我们提供有关数据结构及其生成方式的更多细节。数据以扩展 .xyz 文件的形式存储在代码仓库 [1], 中，可直接用作 (tab)GAP 和 NEP 训练的输入。为方便使用，并允许在训练过程中为不同的结构组指定不同权重，我们将结构归入若干配置类型。下文以纯文本形式总结了配置类型关键词（列于每个结构的标题行中）、每种配置类型的结构数量、每个结构的平均原子数，以及每种配置类型中原子总数和各元素的原子数。

```
config_type              N_boxes    avg N_ats_in_box     N_atoms_total   # per-element atom Counter
-------------------------------------------------------------------------------------------------------
isolated_atom            9          1.00                 9               # {'Mo': 1, 'Nb': 1, 'Ta': 1, 'V': 1, 'W': 1, 'Cr': 1, 'Zr': 1, 'Ti': 1, 'Hf': 1})
dimer                    659        2.00                 1318            # {'Ti': 177, 'Zr': 167, 'Nb': 158, 'Hf': 151, 'Mo': 150, 'V': 146, 'Ta': 128, 'Cr': 123, 'W': 118})
pure_elem_lowE           4330       18.39                79622           # {'W': 13259, 'V': 13157, 'Mo': 13020, 'Nb': 13020, 'Ta': 12966, 'Cr': 4924, 'Hf': 3094, 'Ti': 3091, 'Zr': 3091})
pure_elem_highE          2777       9.21                 25583           # {'V': 4354, 'Ta': 4338, 'W': 4338, 'Mo': 4264, 'Nb': 4212, 'Ti': 1062, 'Zr': 1062, 'Hf': 1059, 'Cr': 894})
liquid                   599        114.04               68310           # {'Ta': 10663, 'W': 10328, 'V': 10201, 'Nb': 8906, 'Mo': 8611, 'Cr': 6727, 'Hf': 4443, 'Zr': 4246, 'Ti': 4185})
short_range              784        30.59                23985           # {'Ta': 4392, 'W': 4334, 'V': 4199, 'Mo': 2891, 'Nb': 2877, 'Cr': 2543, 'Zr': 934, 'Ti': 928, 'Hf': 887})
bcc_binary_alloys        1608       54.00                86832           # {'Hf': 13164, 'Ti': 12858, 'Zr': 12756, 'Nb': 11067, 'Ta': 11067, 'V': 6480, 'Mo': 6480, 'W': 6480, 'Cr': 6480})
bcc_alloys               2150       54.00                116100          # {'Ta': 20614, 'W': 16497, 'V': 16481, 'Nb': 15792, 'Mo': 11733, 'Cr': 9840, 'Ti': 8394, 'Hf': 8378, 'Zr': 8371})
defects                  430        114.51               49238           # {'Ta': 9258, 'W': 8904, 'V': 8393, 'Nb': 6268, 'Mo': 4901, 'Cr': 4323, 'Hf': 3247, 'Zr': 1974, 'Ti': 1970})
bcc_alloys_surface       204        57.45                11720           # {'V': 1877, 'W': 1791, 'Ta': 1534, 'Nb': 1387, 'Mo': 1362, 'Cr': 1321, 'Zr': 1021, 'Ti': 722, 'Hf': 705})
bcc_alloys_ordered       4750       12.05                57258           # {'W': 7418, 'Ta': 7321, 'V': 7192, 'Cr': 6825, 'Nb': 6076, 'Mo': 5966, 'Zr': 5488, 'Hf': 5488, 'Ti': 5484})
hcp_alloys               910        50.69                46128           # {'Ti': 7108, 'Ta': 7085, 'Nb': 7083, 'Hf': 7079, 'Zr': 7079, 'Mo': 2679, 'V': 2672, 'W': 2672, 'Cr': 2671})
hcp_binary_alloys        528        54.00                28512           # {'Hf': 6684, 'Ti': 6378, 'Zr': 6276, 'Nb': 4587, 'Ta': 4587})
fcc_alloys               518        32.00                16576           # {'Zr': 1900, 'Ta': 1896, 'Nb': 1889, 'Hf': 1887, 'Ti': 1877, 'V': 1789, 'Cr': 1781, 'Mo': 1779, 'W': 1778})
fcc_binary_alloys        180        32.00                5760            # {'Zr': 1320, 'Ti': 1300, 'Hf': 1240, 'Nb': 950, 'Ta': 950})
misc_alloys              1324       50.30                66592           # {'Zr': 10427, 'Hf': 9208, 'Cr': 9066, 'Ti': 7258, 'Mo': 6649, 'V': 6253, 'W': 6105, 'Ta': 6006, 'Nb': 5620})
intermetallics           570        16.93                9648            # {'Zr': 1625, 'Hf': 1620, 'Ti': 1549, 'Cr': 880, 'V': 876, 'Ta': 796, 'Nb': 795, 'W': 756, 'Mo': 751})
materials_project        147        6.03                 887             # {'Cr': 199, 'Zr': 125, 'Mo': 94, 'Ti': 90, 'Hf': 89, 'Nb': 84, 'V': 81, 'W': 77, 'Ta': 48})

All                      22477      30.88                694078          # {'Ta': 103650, 'Nb': 90772, 'W': 84856, 'V': 84152, 'Mo': 71331, 'Hf': 68424, 'Zr': 67863, 'Ti': 64432, 'Cr': 58598})
```

以下是每种配置类型中结构的描述。

- 孤立原子：真空中的孤立原子。在 VASP 中通过自旋极化计算得到，以获得正确的参考能量，从而得到正确的密度泛函理论预测的二聚体键能和体相内聚能。
- 二聚体：所有元素对的二聚体。主要包括排斥部分，距离下限高于 VASP 被认为可靠的距离；该距离通过与全电子密度泛函理论数据 [2] 比较来评估。还包括二聚体能量极小值，但不包括完整的解离曲线，因为这些密度泛函理论计算未考虑自旋极化，而孤立原子具有自旋极化，这会导致不一致。
- 纯元素低能结构：来自先前工作的纯元素结构，以及本工作中生成的一些额外结构。低能结构，主要是有或无应变、热位移或点缺陷的基态晶体结构。
- 纯元素高能结构：来自先前工作的纯元素结构，以及本工作中生成的一些额外结构。高能结构、不稳定晶体结构和扩展缺陷（表面），或应变极大的基态晶体。
- 液体：不同密度和温度下的液体结构，包括纯元素以及各种成分的合金。通过 tabGAP 的迭代训练（即使用先前版本的 tabGAP 进行分子动力学）以及交叉采样（见正文）生成。
- 短程：含有不稳定自间隙原子或弗伦克尔对的小型体心立方晶体，包括纯金属和合金，因此有一个或多个原子彼此非常接近。此类构型用于确保近距离下排斥作用的准确性，并平滑过渡到外部排斥势（见 S4 节）。需注意避免原子距离过近，以免 VASP 的计算结果不可靠（ 1 0 1 1 5 –.5 °A，具体取决于元素；对于 Cr-Cr 等小原子键为 1 °A，对于 Hf-Hf 键为 . A）。
- 体心立方二元合金：对体心立方结构中的所有二元合金进行采样，每种元素对取 10 种成分（5、10、20、.....、90、95 at.%），每种成分取三个体积（一个依据维加德定律确定，另两个约为该体积的 ±20 %）。使用标准差为 0.08 °A 的正态分布随机位移原子，以生成独特的局部环境和非零力。
- 体心立方合金：所有超出二元体系的 BCC合金。大多数合金的构建方式与上述二元合金类似。部分合金使用先前版本的 tabGAP 进行分子动力学弛豫，以生成真实的热位移。成分采样涵盖了由 3 至 9 种元素组成的所有合金，并略微偏向三元合金，包括高浓度和稀释成分。有些结构经过弹性畸变。部分经分子动力学弛豫的结构发生坍塌或转变为另一种相（例如，体心立方相不稳定时，这主要发生在以六方密排金属为主的成分中）。

- 缺陷：含有缺陷（空位、自间隙原子或两者兼有）的各种合金和一些纯金属。大多数结构通过迭代训练或交叉采样生成，即使用某个先前版本的 tabGAP 或 NEP 对其进行弛豫。

- 体心立方合金表面：少量随机 BCC 合金的 (1 0 0)、(1 1 0) 和 (11 1) 表面。通过手动方式和交叉采样生成。

- 有序体心立方合金：根据正文所述，使用自编脚本和 icet 软件包生成的各种有序 BCC 合金。

- 六方密排结构合金：对所有等原子比合金成分进行采样，晶格常数遵循维加德定律；其余生成方式与 BCC 二元合金相同。还包括先前研究中的 HCP 合金。
- 六方密排结构二元合金：先前研究中的二元 HCP 合金 [3]。
- 面心立方合金：与上述 HCP 合金类似。
- 面心立方二元合金：与上述 HCP 二元合金类似。
- 其他合金：不适合归入任何其他类别的合金集合。主要包括坍塌或转变为其他相的合金（无论是否有意）。大多数结构来自交叉采样迭代，其中许多结构坍塌为玻璃态或混合相。

- 《金属间化合物》：各种已知和假设的金属间相，包括拉弗斯 C14 相、C15 和 C36、A15 相，以及 Ti/Zr/Hf 合金的 ω 相。
- 材料项目数据库：Materials Project 数据库中包含所列元素的所有合金，但排除了少数明显为人工构造且能量较高的结构。

S2. 交叉采样收敛

图 S1 展示了主文所述交叉采样策略中使用的三组不同结构集合里，tabGAP 与 NEP 之间能量差的分布。选取能量差最大的结构，使用密度泛函理论进行计算，并将其加入训练数据库，随后对两个机器学习原子间势进行重新训练，以开展下一轮迭代。该图显示，随着每次迭代的进行，分布逐渐变窄，异常值也变得更少且极端程度更低（请注意对数刻度）。在第四轮迭代中，我们认为两个机器学习原子间势已达到足够的精度一致性，并选取最终的结构集合。

代码仓库 [1]中提供了实现交叉采样的示例脚本。

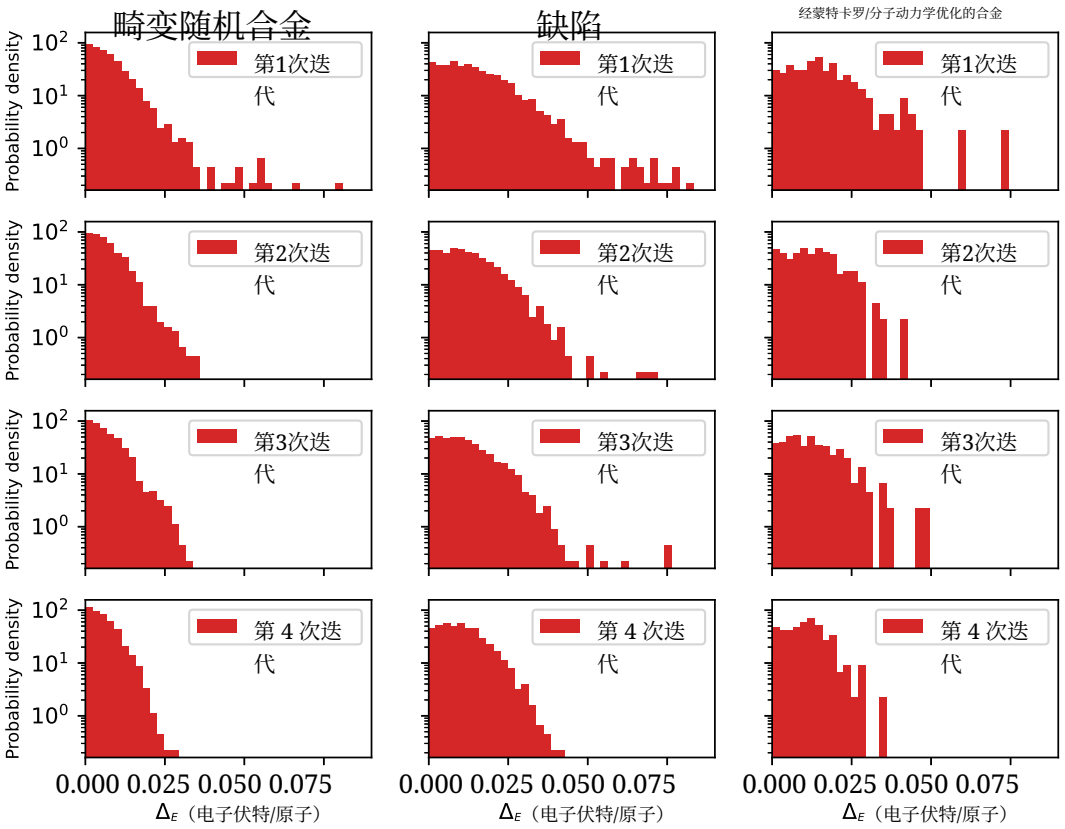

图 S1：正文所述迭代交叉采样策略中，三个不同结构池的 tabGAP 与 NEP 之间能量差的分布。

S3. 计算效率

图 S2 展示了在元素数量逐渐增加的随机合金模拟中，tabGAP 和 NEP 相对于嵌入原子法（Zhou 等人 [4]）及 UNEP-v1 [5]的计算速度。CPU 基准测试模拟使用 LAMMPS 在单个 AMD Ryzen 7 PRO 5750G 核心上完成。GPU 模拟中，tabGAP 和嵌入原子法使用 LAMMPS-KOKKOS，NEP 使用 GPUMD，均在单个 Nvidia Volta V100 GPU 上运行。所有基准测试均在初始温度为 300 K 的 NVE 系综中运行 1000 个时间步。所有测试均重复三次，以避免并消除异常计时结果。

图 S2 显示，tabGAP 比 NEP 和 UNEP-v1 更高效，但其速度会随元素数量增加而降低。这是因为 tabGAP 在三维三次样条插值过程中，依赖于存储和访问大型预计算能量数组。由于无法将所有能量数组都存储在快速缓存内存中，元素数量增加（能量数组也随之增多）会导致缓存未命中次数增加，从而使模拟速度变慢。此前在 LAMMPS-KOKKOS tabGAP 实现中采用了邻居列表排序程序 [6]，本研究也在 CPU 版本中采用了该程序，从而减轻了这一影响，但速度下降仍然十分显著。我们还注意到，tabGAP 效率随元素数量增加而下降的程度取决于许多因素（缓存内存大小、tabGAP 网格大小、合金成分及其排列、截断半径、模拟中的原子数量等），因此在其他模拟和势函数中，结果可能与图 S2 不同。EAM 势也会受到相同影响，但由于能量数组较小，程度稍轻。另一方面，NEP 的速度几乎不受模拟中元素数量的影响。

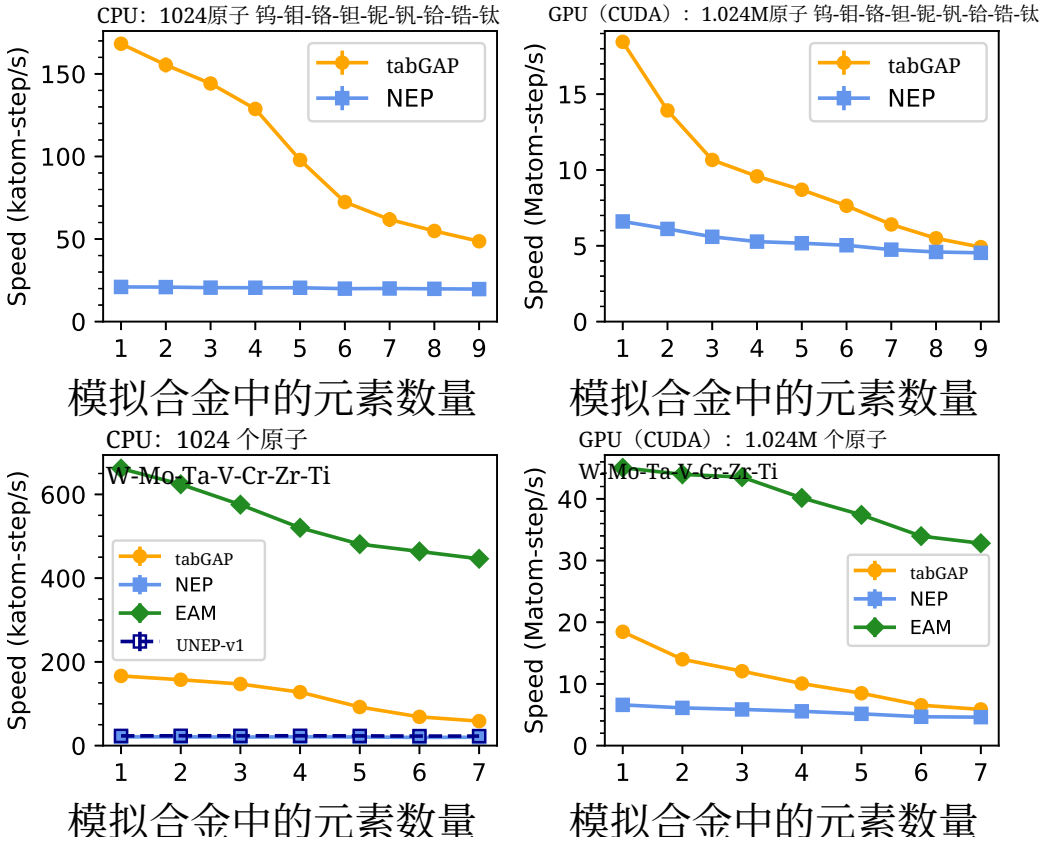

图 S2：tabGAP、NEP、UNEP-v1 和 EAM（Zhou 等人）的 CPU 和 GPU 计算速度基准测试，展示了随机合金中元素数量对计算速度的影响。CPU 模拟使用 LAMMPS，GPU 模拟则对 tabGAP/EAM 使用 LAMMPS-KOKKOS，对 NEP 使用 GPUMD。结果表明，随着合金中模拟元素数量增加，tabGAP（以及 EAM）的性能会下降，这是由缓存内存限制和缓存未命中造成的。

图S3展示了tabGAP、NEP以及标准LAMMPS编译中可直接使用的若干其他势函数在CPU和GPU上的速度基准测试。这些势函数被分为三类：机器学习原子间势、具有三体依赖的传统势函数（Tersoff势、修正嵌入原子法、角度依赖势）以及嵌入原子法。这些测试

在图S3中均针对纯金属进行，以便与许多其他势函数进行比较。但需注意，势函数的速度显然取决于截断半径内近邻原子的数量。这一点通过展示铌（原子密度较低，即截断半径内近邻原子较少）和铬（原子密度较高）的tabGAP与NEP结果加以说明。不同嵌入原子法和Tersoff势呈现出截然不同的速度，也清楚地表明了这一点。

在纯元素模拟中，tabGAP 的速度比 NEP 快 6–8 倍，具体取决于截断半径内近邻原子的数量。在 GPU 上，tabGAP 的速度仅比 NEP 快 2–3 倍，这表明 NEP 最适合搭配其原生的 GPUMD 代码使用。与其他势函数相比，tabGAP 在 CPU 和 GPU 上都显著快于所有受测的机器学习原子间势（其中仅包括效率最高的机器学习原子间势）。NEP 在 CPU 上的表现与其他机器学习原子间势相当，但在 GPU 上显著更快。

图S3还显示，势函数在 CPU 和 GPU 上的相对速度有所不同。例如，角依赖势效率很高，在 CPU 上比 tabGAP 快三倍，但在 GPU 上速度相近。与 tabGAP 相比，Tersoff势则呈现相反的趋势。

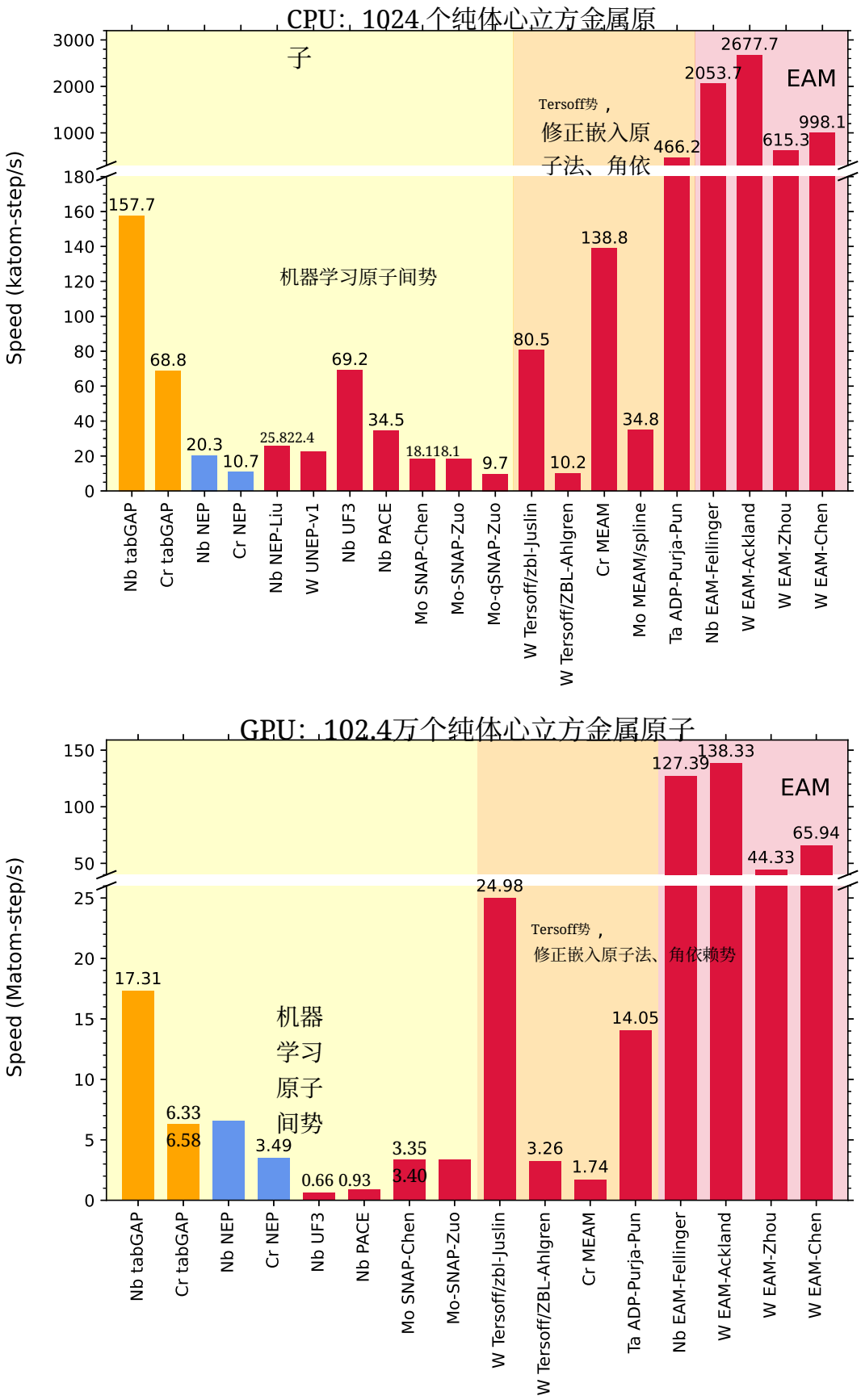

图 S3：针对具有现成原子间势的纯体心立方金属，在 CPU（上）和 GPU（下）上的计算速度基准测试。CPU 模拟使用 LAMMPS 在单个 AMD Ryzen 7 PRO 5750G 核心上完成。GPU 模拟在单个 Nvidia Volta V100 GPU 上使用 LAMMPS-KOKKOS（若可用）完成，但 NEP 除外，其使用原生 GPUMD 代码运行。请注意，速度强烈依赖于截断半径内的近邻原子数目（不同势函数之间有所不同），tabGAP 和 NEP 对铌（原子密度较低，即截断半径内近邻较少）和铬（原子密度较高）的结果说明了这一点。

S4. 排斥势

在训练 tabGAP 和 NEP 时，通过将外部纯排斥对势作为基线纳入其中，确保在极短原子间距离下具有物理上正确的排斥作用。随后，每个训练结构的目标能量变为 DFT 能量与该基线排斥能之差。为避免排斥势干扰近平衡原子间距离，排斥势通过截断函数平滑地衰减至零。该方法遵循参考文献 [7] 中所述的详细方法。

排斥势是针对每个元素对重新拟合的齐格勒-比尔萨克-利特马克（ZBL）势 [8]：

$$E_{\text{rep.}} = \sum_{i<j}^{N} \frac{1}{4\pi\varepsilon_0} \frac{Z_i Z_j e^2}{r_{ij}} \phi_{ij}(r_{ij}/a) f_{\text{cut}}(r_{ij}), \tag{1}$$

其中，$a = 0.46848/(Z_i^{0.23} + Z_j^{0.23})$ 与通用ZBL势中的定义相同。元素对 AB 的屏蔽函数为 $\phi_{\text{AB}}(x) = \sum_{i=1}^{m} \zeta_i e^{-\eta_i x}$。我们采用了 $m = 3$，并针对每个元素对，利用参考文献 [2] 中的全电子密度泛函理论数据拟合六个参数 ζ_i 和 η_i。截断函数 $f_{\text{cut}}(r_{ij})$ 与参考文献 [7] 中的相同，并在预定义范围内平滑地将能量强制降至零。

截断范围和拟合参数以原始文本形式列于下方。各列依次为：截断起点（Å）、截断终点（Å）、ζ_1, η_1, ζ_2, η_2, ζ_3, η_3, ζ_4, η_4。每行对应一个元素对。各行（元素对）的顺序为 1-1、1-2、1-3、……、1-9、2-2、2-3、……、8-8、8-9、9-9，其中 1、2、……、9 按字母顺序对应以下元素：铬 铪 钼 铌 钽 钛 钒 钨 锆。此格式以及下方粘贴的内容，与 NEP 训练所用 zbl.in 输入文件的格式完全相同。

```
1.0 2.0 0.62432709 0.52724872 0.17985857 1.59251784 0.19978689 2.32817531 0 0 1.0 2.2 0.46761483
0.88909152 0.21297607 2.98287067 0.31261612 0.40388501 0 0 0.8 2.2 0.20671801 9.44078152
0.49965657 1.26900977 0.428884 0.42626681 0 0 0.7 2.2 0.49811837 1.25665452 0.42783267
0.42607765 0.12694216 7.74046431 0 0 1.0 2.2 0.46700439 0.9213475 0.3270954 0.40805846
0.20402522 3.19206108 0 0 1.0 2.2 0.37440734 1.9039036 0.56883919 0.51978103 0.0507448 0.519783
0 0 1.0 2.0 0.1782988 2.29448513 0.19729121 1.64264918 0.62264504 0.52545345 0 0 1.0 2.2
0.32946711 0.40698282 0.19904867 3.3947085 0.47489525 0.93742154 0 0 1.0 2.2 0.50402738
1.22228461 0.10189284 6.41672406 0.41584423 0.42062546 0 0 1.0 2.2 0.59592609 0.56463409
0.3378634 2.39617239 0.06001117 0.25930197 0 0 1.0 2.2 0.17213842 3.69808253 0.31485428
0.37832137 0.51401042 0.97270952 0 0 1.0 2.2 0.51165313 0.90861154 0.29162012 0.36876844
0.18523058 3.23038014 0 0 1.0 2.2 0.5930524 0.56857818 0.33473184 2.39575589 0.06449729
0.26571147 0 0 1.0 2.2 0.25651787 0.37440487 0.21047215 2.9771018 0.5219775 0.84177329 0 0 1.0 2.2
0.21165315 2.979274 0.49082842 0.86742557 0.28862723 0.39247999 0 0 1.0 2.2 0.0704484 0.27305632
0.33301887 2.42079332 0.59005438 0.57393834 0 0 1.0 2.2 0.27621236 0.36224181 0.19530828
3.22049908 0.52075578 0.88265633 0 0 1.0 2.2 0.48796853 0.98194874 0.31579653 0.35741207
0.18798549 2.71874806 0 0 1.0 2.2 0.19020878 2.65056725 0.32076908 0.36142452 0.47789479
0.98036749 0 0 1.0 2.2 0.51344146 0.98119957 0.16477545 3.69293833 0.31816071 0.37914886 0 0 1.0
2.2 0.40970569 0.41443049 0.14048879 7.33390014 0.50897037 1.21591244 0 0 1.0 2.2 0.18396472
8.70291804 0.49897344 1.25387083 0.42609115 0.4247833 0 0 1.0 2.2 0.51684534 0.96288639
0.30904771 0.37485806 0.16779276 3.59037981 0 0 1.0 2.2 0.23908233 2.31575175 0.2905037
0.35143156 0.45403733 0.88232297 0 0 1.0 2.2 0.18049477 2.92908224 0.49502099 0.99695761
0.3227105 0.36116294 0 0 1.0 2.2 0.50494081 0.91672469 0.29923796 0.3728018 0.1826209 3.21613753
0 0 1.0 2.2 0.51944791 1.16631231 0.12504908 6.21792914 0.39038215 0.40495968 0 0 1.0 2.2
0.5026924 1.22895263 0.13951036 7.50424076 0.41849001 0.42091243 0 0 1.0 2.2 0.5085639
0.91518537 0.29670954 0.37129392 0.18666422 3.28455076 0 0 1.0 2.2 0.25827487 0.33608329
0.25427726 2.28636879 0.47267247 0.82744486 0 0 1.0 2.2 0.57289615 0.5763908 0.32394661
2.46358374 0.0739523 0.27722376 0 0 1.0 2.2 0.2028207 3.14399822 0.28344678 0.38591376
0.50676484 0.87847293 0 0 1.0 2.2 0.47408861 0.92079566 0.19860103 3.19599153 0.32249351
0.40575131 0 0 1.0 2.2 0.07833469 0.27911064 0.3294006 2.4745664 0.58841741 0.58199911 0 0 1.0 2.2
0.19519067 3.26708518 0.51520991 0.89414783 0.28453263 0.36642181 0 0 1.0 2.2 0.27696675
2.08564077 0.2225363 1.04949527 0.50214739 0.46974138 0 0 1.0 2.2 0.57377437 0.5002702
0.19290733 1.34666721 0.22937086 2.14239509 0 0 1.0 2.2 0.31065254 0.3962899 0.19280246
3.45766337 0.49736024 0.92468663 0 0 1.0 2.2 0.53564508 1.10903088 0.36371112 0.39206439
0.1189747 5.34000005 0 0 1.0 2.2 0.23582493 2.09349355 0.15563716 1.560913 0.60926898 0.51701385
0 0 1.0 2.2 0.47543791 0.94241702 0.19384435 3.38236532 0.33063109 0.40745887 0 0 1.0 2.2
0.10859677 6.0488931 0.51659195 1.17672301 0.39619262 0.41006409 0 0 1.0 2.2 0.32825 2.54931
0.09219 0.29182 0.5811 0.59231 0 0 1.0 2.2 0.29483836 0.37094151 0.19149279 3.38060865 0.51192616
0.91347455 0 0 1.0 2.2 0.25612795 2.34118951 0.49644544 0.8056594 0.23709107 0.32564545 0 0
```

拟合得到的排斥性基准势绘于图 S4 中（标记为“SC”，即屏蔽库仑势）。图中还显示了 tabGAP 和 NEP 的能量；它们与 SC 并不完全相同，因为机器学习能量贡献也不为零。不过，图 S4 表明，在高能量下，由于所有势函数呈现的能量相近，这一贡献可以忽略不计。该图还表明，即使原子间距离趋近于零，tabGAP 和 NEP 仍保持平滑且表现良好。

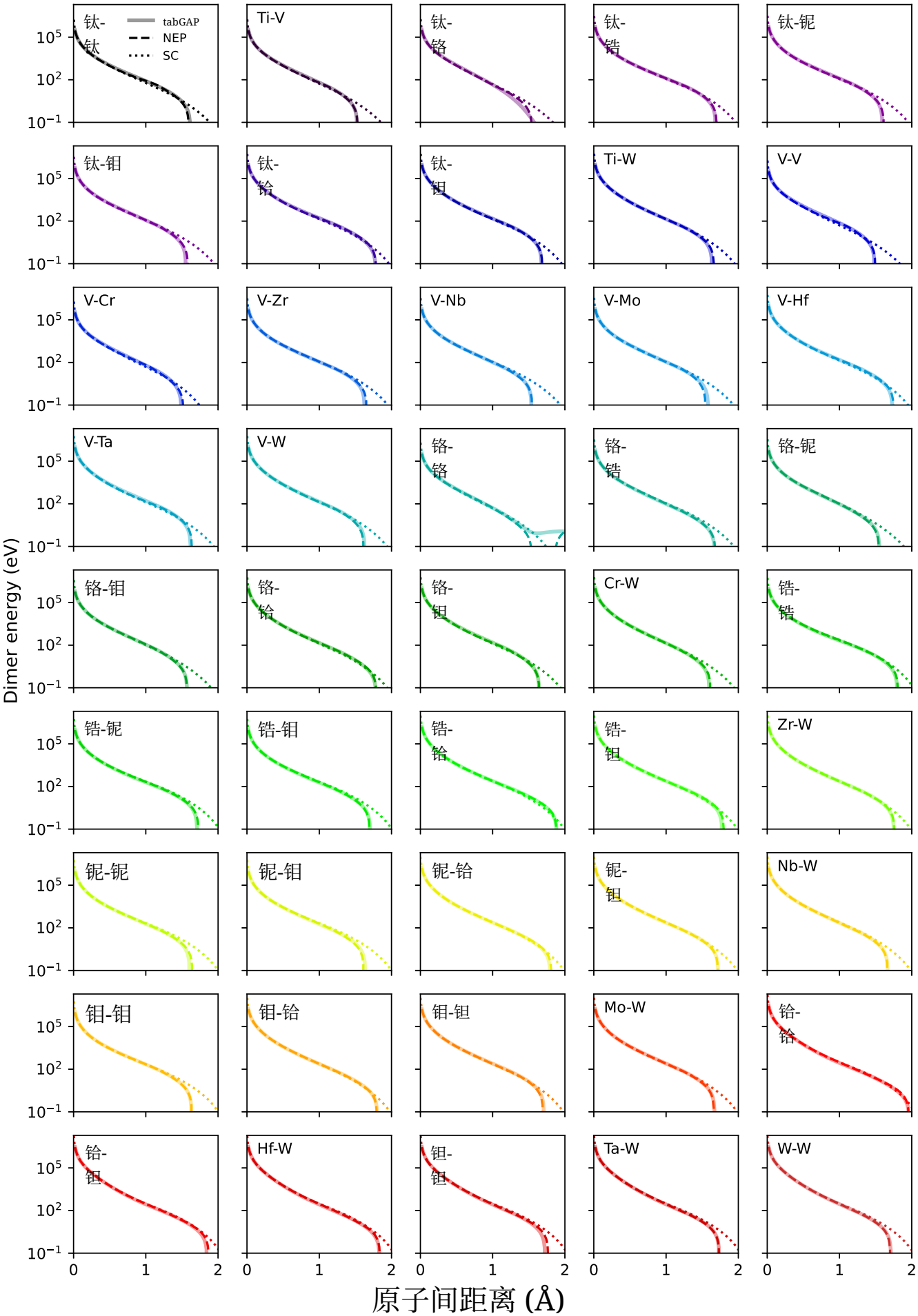

图 S4：二聚体曲线的排斥部分 tabGAP、NEP 和拟合全电子密度泛函理论数据的屏蔽库仑势（SC）。

S5. 纯金属性质

表 S1–S5 和图 S5–S7 展示了使用 tabGAP 和 NEP 计算的纯金属各种性质，并将其与实验和密度泛函理论数据进行了比较。关于 tabGAP 和 NEP 性能的一些观察如下：

•两种机器学习原子间势的弹性常数总体上较为准确，但也有一些例外；其中 NEP 的准确性略高。体心立方金属的弹性常数总体上比六方密排金属更准确。铬的弹性常数与密度泛函理论和实验结果存在偏差，主要是由于训练数据中缺少自旋极化。

• 与实验结果相比，两种机器学习原子间势的热膨胀系数要么准确，要么略微高估。值得注意的是，tabGAP 能够捕捉钛在低温下沿 c轴的特殊负热膨胀现象（tabGAP 中低于 300 K，实验和密度泛函理论中低于 170 K [9]），而 NEP 则不能。平均线性热膨胀近似为体积膨胀的三分之一，钛和所有其他金属的平均线性热膨胀仍为正值。

• 两种机器学习原子间势对空位迁移能的预测都相当准确。势垒形状有时有所不同，有时在中点处包含一个局部极小值；密度泛函理论对六方密排金属也有类似报道 [10]。两种机器学习原子间势都正确预测，在六方密排金属中，基面路径比非基面路径更有利。

表 S1：正文图3所示纯金属的性质。各项性质的数值按密度泛函理论（或 B 和 T_{melt}的实验）/tabGAP/NEP 的顺序给出。

	B (吉帕)	E_f^{vac} (电子伏特)	E_f^{SIA} (电子伏特)	E_{surf} (电子伏特/Å^2)	E_{GB}(毫电子伏特/Å^2)	T_{melt} (开尔文)
Ti	107/121/109	2.03/2.38/2.52	2.38/2.79/2.26	0.134/0.108/0.110	23/20/26	1941/1775/1750
Zr	95/95/100	1.94/2.63/2.17	2.69/2.78/2.56	0.100/0.094/0.098	20/17/20	2128/1975/1750
Hf	110/111/109	2.22/2.72/2.41	3.97/4.01/3.71	0.107/0.105/0.107	26/28/32	2506/2175/2050
V	156/187/179	2.49/2.74/2.33	2.41/3.48/3.47	0.150/0.148/0.154	16/22/23	2183/2050/2150
Nb	172/175/172	2.77/3.08/2.60	3.95/4.83/4.55	0.130/0.130/0.131	16/21/24	2750/2450/2550
Ta	189/200/197	2.95/3.29/2.86	4.77/5.88/5.41	0.147/0.147/0.148	18/24/23	3290/2950/2950
Cr	153/257/260	3.02/2.96/2.83	6.07/5.46/6.65	0.201/0.203/0.203	40/43/57	2180/2450/2350
Mo	260/255/256	2.83/2.88/3.05	7.40/7.26/8.16	0.173/0.170/0.175	30/39/44	2896/2750/2750
W	310/299/301	3.36/3.39/3.33	10.25/10.37/10.99	0.204/0.201/0.202	42/52/50	3695/3550/3550

表 S2：体心立方金属的弹性常数 在 0 K 时与密度泛函理论和室温实验数据的比较 a.

实验 [11]	密度泛函理论 [12]	tabGAP	NEP		实验 [11]	密度泛函理论 [12, 13]	tabGAP	NEP
V					Cr			
C_{11} 229	269	300	280		C_{11} 340	448	444	527
C_{12} 119	146	157	130		C_{12} 59	62	127	132
C_{44} 43	22	26	22		C_{44} 99	102	83	113
Nb					Mo			
C_{11} 247	237	257	257		C_{11} 464	468	462	475
C_{12} 135	138	135	133		C_{12} 158	155	150	155
C_{44} 29	11	17	13		C_{44} 109	100	97	110
Ta					W			
C_{11} 260	266	267	260		C_{11} 522	521	519	516
C_{12} 154	161	170	161		C_{12} 204	195	184	191
C_{44} 83	77	68	60		C_{44} 161	147	138	139

表 S3：0 K 时六方密排金属的弹性常数，与密度泛函理论计算结果和室温实验数据的比较。

	实验 [11]	密度泛函理论 [14]	tabGAP	NEP
Ti				
C_{11}	162	170	178	164
C_{12}	92	86	141	50
C_{13}	69	75	90	47
C_{33}	181	186	230	234
C_{44}	47	43	25	59
Zr				
C_{11}	143	141	136	141
C_{12}	73	68	78	82
C_{13}	65	66	71	75
C_{33}	165	164	192	177
C_{44}	32	26	26	32
Hf				
C_{11}	188	182	193	173
C_{12}	77	72	79	78
C_{13}	66	68	81	66
C_{33}	197	197	291	211
C_{44}	56	54	54	52

表 S4：室温下的线性热膨胀系数，单位为 10^{-6} K^{-1} 。实验值引自参考文献 [11]。

	实验	tabGAP	NEP		实验	tabGAP	NEP		实验	tabGAP	NEP
Ti	8.6	11.7	11.5	钒	8.4	8.4	11.6	铬	4.9	8.3	7.6
Zr	5.7	12.5	7.5	铌	7.3	7.0	9.4	钼	4.8	7.1	6.3
Hf	5.9	8.8	7.1	钽	6.3	5.6	7.4	W	4.5	5.6	5.5

表 S5：单空位迁移能，单位为电子伏特。六方密排金属对应的两个数值分别表示基面和非基面迁移路径。

	密度泛函理论 [10]	tabGAP	NEP		密度泛函理论 [15]	tabGAP	NEP		密度泛函理论 [15]	tabGAP	NEP
钛	0.43/0.57	0.40/0.46	0.28/0.42	钒	0.65	0.35	0.51	铬	1.11	1.25	1.25
锆	0.51/0.67	0.63/0.67	0.64/0.80	铌	0.65	0.52	0.60	Mo	1.21	1.41	1.36
铪	0.79/0.91	0.90/1.01	1.07/1.31	钽	0.76	0.61	0.71	W	1.73	1.90	1.95

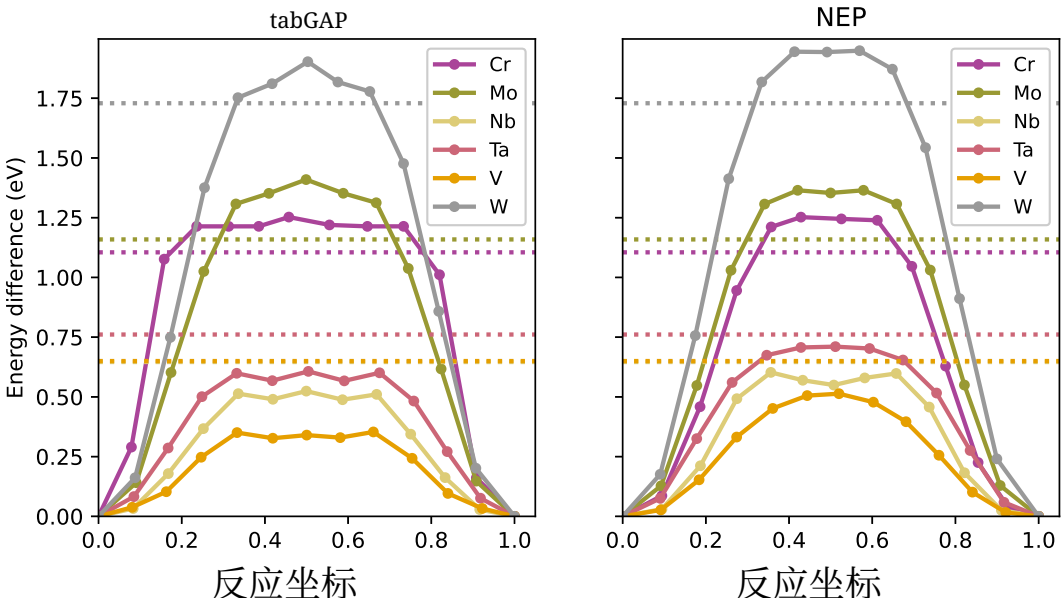

图 S5：用弹性带方法计算的纯体心立方金属中的空位迁移能垒。虚线为参考文献 [15] 中的密度泛函理论迁移能。

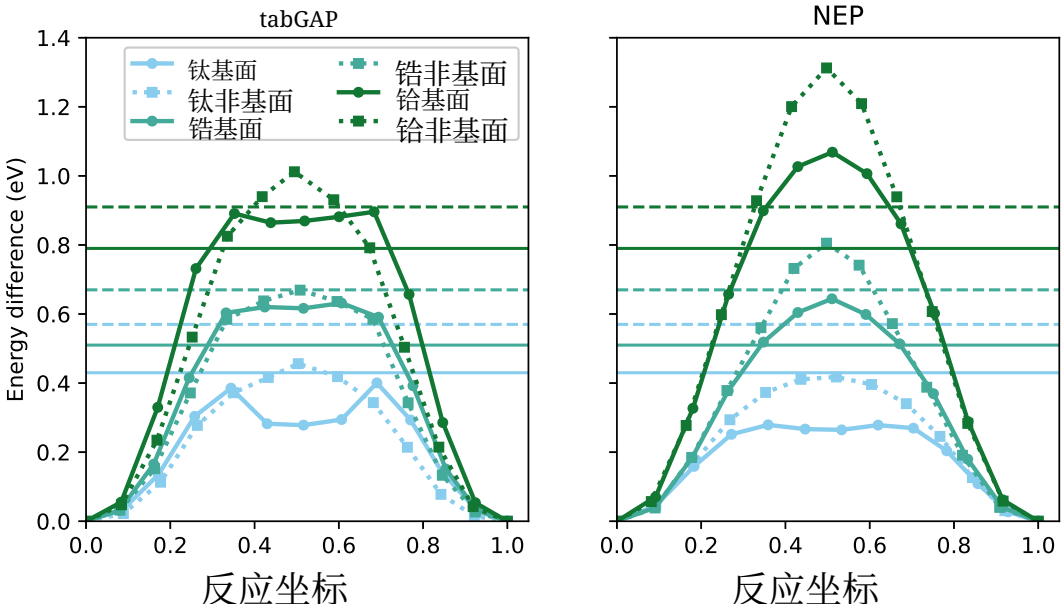

图 S6：用弹性带方法计算的纯六方密排金属中的空位迁移能垒，包括基面内迁移（基面）和垂直于基面的迁移（非基面）。实线和虚线分别为参考文献 [10] 中基面路径和非基面路径的密度泛函理论迁移能。

S6. 相图与自由能

图 S7 展示了使用 tabGAP 和 NEP 模拟得到的 Ti（正文中也有展示）、Zr 和 Hf 的相图。按照正文所述，通过蛮力采样，并利用非平衡分子动力学模拟计算 (P,T) 点处的吉布斯自由能，绘制相图。作为示例，图 S8-S10 展示了所考察的相的模拟和计算得到的 $G(P,T=0)$（即焓 H）曲线和 $G(T,P=0)$（即亥姆霍兹自由能 F）曲线。这些图还展示了相对于六方密排相的自由能差，从而揭示了相图中所示的零温度和零压相变。

自由能差通过对自由能数据进行三次样条插值计算得出。请注意，对于某些相，自由能曲线，尤其是能量差曲线，会出现人为的波动。这些波动并非真实现象，只是由于不稳定相在分子动力学模拟过程中部分转变为更稳定的相所导致的。因此，这种情况只会出现在高度不稳定的相中，不会影响平衡相和相变的判定。图 S8-S10 中仅在 $T=0$ K 曲线中纳入了 fcc 相。最初完全没有考虑 fcc 相，直到使用 NEP 对 Hf 进行的高压分子动力学模拟显示，其他相有转变为 fcc 相的迹象。静态（零温度）焓计算表明，fcc 相在高压下稳定。此后，我们对 (P,T) 点进行了稀疏采样，以绘制 Hf NEP 中的 fcc 相区域。在其他所有情况下，焓曲线均表明，fcc 相始终远非最稳定相。

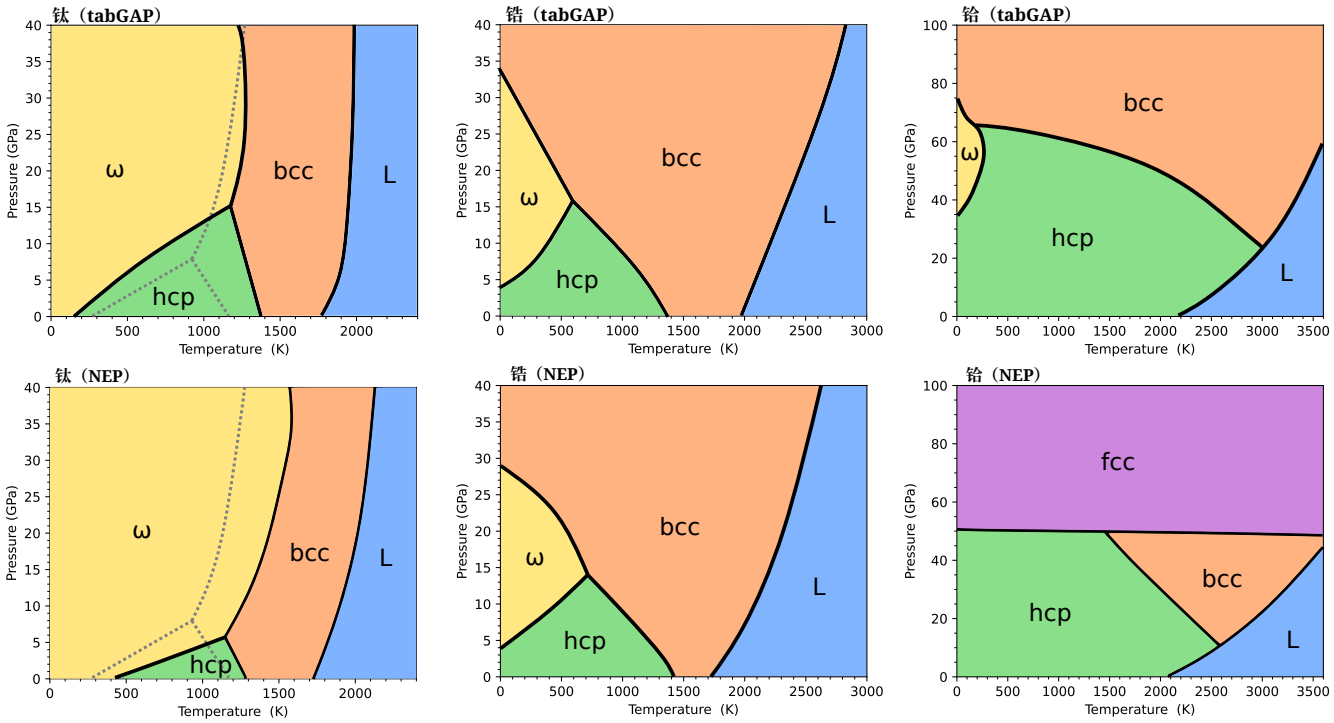

图 S7：tabGAP 和 NEP 预测的全部三种六方密排金属的相图。方法详见正文。

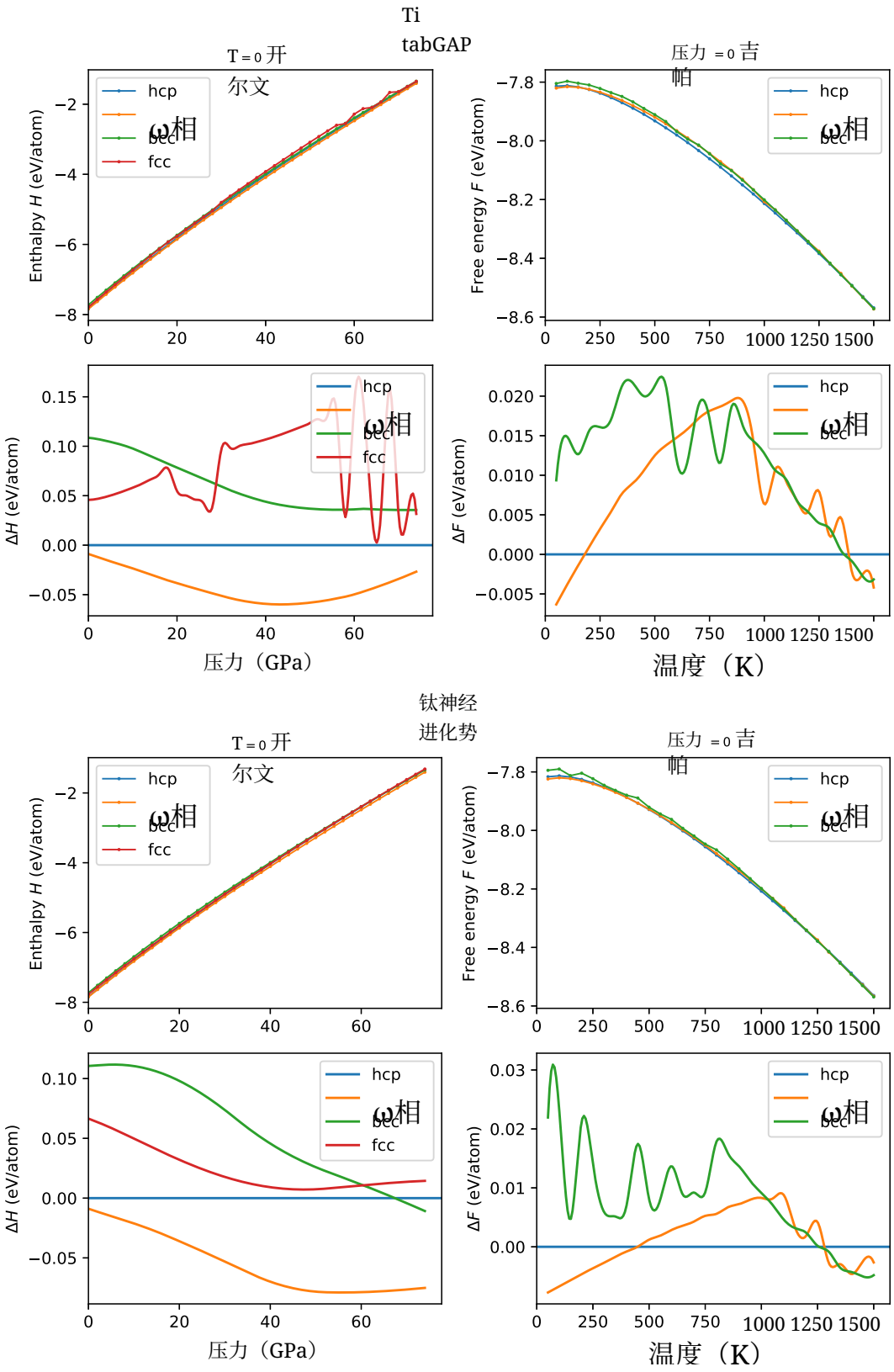

图 S8：模拟的自由能量及相对于钛六方密排相的自由能差，分别由 tabGAP 和 NEP. 焓 H（$T = 0$ 开尔文）和亥姆霍兹自由能 F（$P = 0$ 吉帕）。

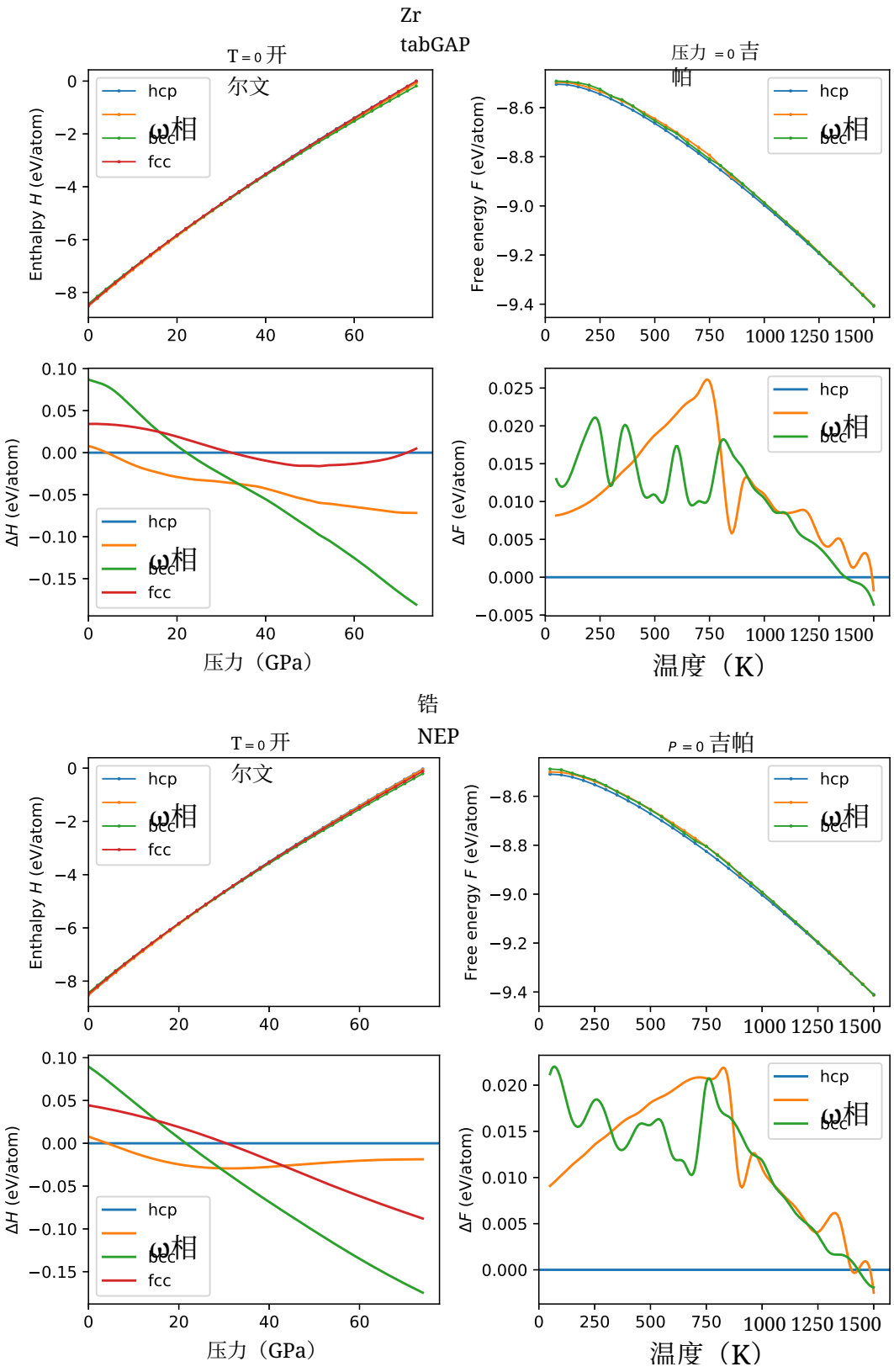

图 S9：模拟的自由 能量，以及相对于六方密排相的锆的自由能差，基于 tabGAP 和 NEP. 焓 H（$T = 0$ 开尔文）和亥姆霍兹自由能 F（$P = 0$ 吉帕）。

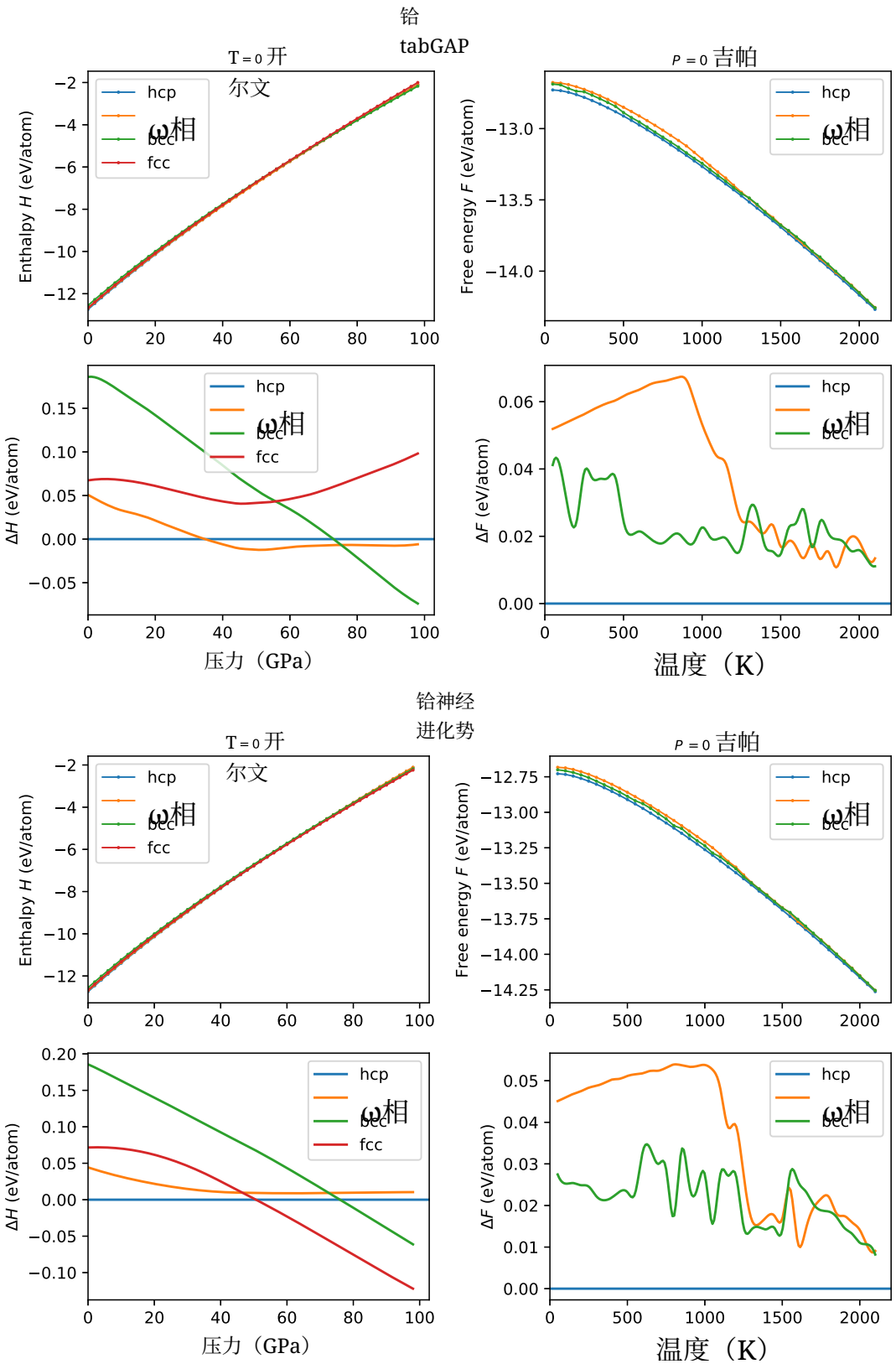

图 S10：模拟的自由能 自由能以及相对于铪的六方密排相的自由能差，来自tabGAP和NEP 相比，铪相对于六方密排相的自由能以及自由能差。焓 H（$T = 0$ 开尔文）和亥姆霍兹自由能 F（$P = 0$ 吉帕）。

[1] J. Byggm¨astar，RHEA 数据库和 tabGAP/NEP势函数文件（2026）。[2] K. Nordlund、S. Lehtola 和 G. Hobler，《物理评论A》**111**，032818（2025）。[3] T. Lopes 和 J. Byggm¨astar，《面向 HfNbTaTiZr高熵合金体系的机器学习原子间势》（2026），准备中。[4] X. W. Zhou、R. A. Johnson 和 H. N. G. Wadley，《物理评论B》**69**，144113（2004）。[5] K. Song、R. Zhao、J. Liu、Y. Wang、E. Lindgren、Y. Wang、S. Chen、K. Xu、T. Liang、P. Ying、N. Xu、Z. Zhao、J. Shi、J. Wang、S. Lyu、Z. Zeng、S. Liang、H. Dong、L. Sun、Y. Chen、Z. Zhang、W. Guo、P. Qian、J. Sun、P. Erhart、T. Ala-Nissila、Y. Su 和 Z. Fan，《自然-通讯》**15**，10208（2024）。[6] J. Byggm¨astar、V.-M. Yli-Suutala、A. Fellman、J. Astr¨om、J. Westerholm 和 F. Granberg，钨中初级辐照损伤的四种状态（2025），arXiv:2503.02710 [凝聚态物质]。[7] J. Byggm¨astar、A. Hamedani、K. Nordlund 和 F. Djurabekova，《物理评论B》**100**，144105（2019）。[8] J. F. Ziegler、J. P. Biersack 和 U. Littmarck，载于 *Treatise on Heavy-Ion Science*（Pergamon，纽约，1985），第93–129页。[9] P. Souvatzis、O. Eriksson 和 M. I. Katsnelson，《物理评论快报》**99**，015901（2007）。[10] G. V´erit´e、F. Willaime 和 C. C. Fu，《固体现象》**129**，75（2007）。[11] J. Rumble，编者，*CRC Handbook of Chemistry and Physics, 106th Edition*，第106版（CRC出版社，博卡拉顿，佛罗里达州，2025）。[12] J. Byggm¨astar、K. Nordlund 和 F. Djurabekova，《物理评论：材料》**4**，093802（2020）。[13] P.-W. Ma 和 S. L. Dudarev，《物理评论：材料》**3**，10.1103/PhysRevMaterials.3.013605（2019）。[14] A. R. Warwick、P.-W. Ma 和 S. L. Dudarev，《物理评论：材料》**9**，093605（2025）。[15] P.-W. Ma 和 S. L. Dudarev，《物理评论：材料》**3**，063601（2019）。