

基于神经网络势与增强采样的气相水团簇成核过程研究^{*}

徐森^a 吴丽铃^b 李震宇^{*,a,b}

(^a 中国科学技术大学 合肥微尺度物质科学国家研究中心 合肥 230026)

(^b 中国科学技术大学 化学物理系 合肥 230026)

摘要 由于气相分子密度低,对气相成核过程的理论模拟往往需要很大的计算量.为了提高模拟效率,本工作将神经网络势与增强采样技术结合,并以水团簇成核为例进行了研究.采用密度泛函理论方法对不同尺寸的水团簇进行了能量和力的计算,并由此训练出一套能较好描述水团簇体系相互作用的神经网络势.将这个势应用于蒙特卡洛模拟并结合多种增强采样方法,实现了在不同尺寸水团簇之间的随机行走,由此可得到水团簇在特定条件下的概率分布以及吉布斯自由能.通过后续的蒙特卡洛模拟结合伞形采样和变分过渡态理论,可以进一步计算出不同水团簇的水分子蒸发速率.观察到了四聚体到五聚体的自由能和蒸发速率的突变现象.结构分析表明虽然五聚体的最低能量构型是二维环状结构,但是在有限温度下五聚体中三维氢键网络已经开始形成.这导致了在四聚体过渡到五聚体时的异常.本工作提供的第一性原理精度下对气相水团簇成核进行研究的方法可以推广到更为复杂的多组分体系,为研究大气颗粒物形成机理奠定了基础.

关键词 神经网络; 蒙特卡洛; 水; 成核; 增强采样

Nucleation of Water Clusters in Gas Phase: A Computational Study Based on Neural Network Potential and Enhanced Sampling^{*}

Xu, Sen^a Wu, Liling^b Li, Zhenyu^{*,a,b}

(^a Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei 230026, China)

(^b Department of Chemical Physics, University of Science and Technology of China, Hefei 230026, China)

Abstract Due to their low density in atmosphere, theoretical simulations of the nucleation of gas-phase molecules are computationally very expensive. In this study, neural network potential (NNP) is combined with enhanced sampling techniques to effectively investigate the nucleation of water clusters in gas phase. The neural network potential is trained based on water-cluster energies and forces from density functional theory (DFT). The problem that the binding between water molecules is too weak in the previous empirical force field model has been solved in the NNP. This NNP potential is then applied to Monte Carlo simulations in grand canonical ensemble with enhanced sampling methods such as aggregation-volume-bias Monte Carlo (AVBMC) and transition-matrix Monte Carlo (TMMC) to realize a random walk among different cluster sizes. Probability distribution of water cluster sizes and the corresponding Gibbs free energies can then be obtained. Subsequently, the evaporation rates of water clusters can be calculated via umbrella sampling Monte Carlo simulations in canonical ensemble combined with variational transition state theory (VTST). We observe a big change of free energy and evaporation rate from tetramer to pentamer. A statistical analysis of the number of hydrogen bonds suggests that more hydrogen bonds are required to be broken in the evaporation reaction of tetramer compared to that of trimer and pentamer. Structure analysis indicates that, although the ground state of the pentamer has a two-dimensional ring structure, three-dimensional hydrogen bond network begins to form in pentamer at finite temperature. Therefore, it is a two-dimensional to three-dimensional transition from tetramer to pentamer. The fact that the most probable configuration of pentamer is different from the lowest energy configuration demonstrates the importance of molecular simulations. Simply finding the lowest energy configuration via global geometry optimization and then calculating the free energy within a harmonic approximation of vibrations are not a universal protocol for cluster systems. Methods used in this study are expected to be applicable for more complicated multi-component systems, which opens an avenue for the research of particulate matter formation in atmosphere.

Keywords neural network; Monte Carlo; water; nucleation; enhanced sampling

^{*} Dedicated to the 10th anniversary of the Youth Innovation Promotion Association, CAS.

^{*} E-mail: zyl@ustc.edu.cn

Received January 3, 2022; published February 21, 2022.

Supporting information for this article is available free of charge via the Internet at <http://sioc-journal.cn>.

Project supported by the National Natural Science Foundation of China (Nos. 21825302, 21573201).

^{*} 庆祝中国科学院青年创新促进会十年华诞.

项目受国家自然科学基金(Nos. 21825302, 21573201)资助.

1 引言

大气中的物理化学过程对于我们的日常生活有着重要的影响^[1]. 水分子是大气的重要组成部分. 大气中水团簇的成核问题长期以来受到广泛关注^[2-5]. 之前的实验研究^[6-8]通常关注较大的水团簇或者纳米粒子, 运用经典成核理论^[9]来解释实验结果. 经典成核理论将水团簇抽象成一个球体, 不适应于较小的团簇. 实验上准确获得水分子小团簇结构和氢键网络信息具有较大的挑战性. 最近, Zhang 等^[5]应用红外光谱对水分子小团簇开展了系统的研究. 他们的研究发现了一些非能量最低构型起着重要贡献. 因此, 在理论研究中仅仅通过全局几何优化得到能量最低结构, 并在此基础上运用简谐近似计算自由能往往是不够的. 通过分子模拟对多种可能构型进行充分采样对研究水分子小团簇十分必要.

通过分子模拟进行充分采样往往需要较大的计算量. 因此, 之前对水分子团簇的分子模拟研究大都采用 SPC/E^[10]、TIP3P^[11]、TIP4P^[11-13]、OPC^[14]等经验势. 这些经验势中的参数通常是通过液态水的宏观性质进行拟合而得到的, 经常被应用到液态水的分子动力学模拟^[15-16]中, 但是应用在水团簇研究上可能精度不够. 同时, 部分经验势把水分子处理成一个刚体, 无法考虑分子内部自由度的影响, 这对于水分子氢键网络的描述可能造成较大的偏差. 基于上述考虑, 对一些之前用刚性水分子力场模拟预测的结果^[17]需要用更高精度的模型进行确认. 考虑到直接进行第一性原理模拟往往在计算量上无法承受, 机器学习提供了一种解决计算量与计算精度之间矛盾的可能方案. 基于神经网络势可以得到第一性原理精度水平的计算结果, 同时计算速度比第一性原理计算有数十倍的提升. 从 2007 年 Behler 和 Parrinello^[18]提出神经网络势(NNP)模型开始, 人们发展了多种机器学习势模型, 例如 GAP^[19]、SchNet^[20]、ANI-1^[21]、AMP^[22]、LASP^[23-24]、DeePMD^[25]. 其中, 深度势分子动力学(DeePMD)是一个获得广泛应用的模型. 这些模型为准确且高效地研究水分子团簇问题提供了一条可能的途径.

在经典成核理论^[9], 水分子团簇被抽象成一个球体, 很多原子细节被忽略. 在原子尺度描述水分子团簇的成核过程可以采用动力学成核理论^[26]. 在该理论中, 成核生长过程被分解成一系列分子从团簇中蒸发以及分子在团簇表面贴附的过程. 基于正则变分过渡态理论, 从团簇亥姆霍兹自由能与团簇体积或者半径的函数关系可以确定分子单体蒸发速率常数. 在特定的大气环境中, 如果知道各个不同尺寸团簇的吉布斯自由能, 通过细致平衡原理可以得到贴附速率. 最终, 可以通过速率方程确定成核速率. 动力学成核理论摆脱了经典成核理论对宏观性质的依赖, 可以帮助我们系统地了解成核动力学, 包括成核瓶颈和生长路径等信息. 为了得到动力学成核理论需要的亥姆霍兹自由能随团簇半径的变

化以及随团簇内分子数的变化, 需要进行分子模拟. 通常可以采用蒙特卡洛算法^[27]. 为了克服自由能垒、加速收敛, 通常可以采用一些增强采样技术, 包括不同的偏倚采样方法^[28-29]和伞形采样^[30]等.

在本研究中, 我们将先针对水团簇体系构造训练集, 得到一套可移植性较好精度较高的神经网络势. 基于神经网络势, 对水团簇体系进行巨正则系综蒙特卡洛模拟. 模拟中允许水分子直接添加到一个团簇或者被从一个团簇移除, 从而加速对不同尺寸的团簇的采样. 得到平衡的团簇尺寸分布以后, 可以得到团簇吉布斯自由能随团簇内水分子个数的变化. 基于变分过渡态理论, 通过正则系综蒙特卡洛模拟得到团簇半径的分布函数, 从而获得不同团簇的水分子蒸发速率. 模拟发现五聚体容易形成 4+1 结构, 有较高的蒸发速率. 而四聚体形成的环状结构很稳定, 蒸发一个水分子需要打断较多的氢键, 蒸发速率较低. 本工作的研究结果为在原子分子尺度理解大气中的物理化学过程提供了基础.

2 理论背景和方法

2.1 对巨正则系综的偏倚采样

研究团簇的成核首先需要合理的团簇定义. 一个可能的定义^[31]是要求团簇中每个粒子至少有一个近邻粒子也属于该团簇. 而近邻粒子可以通过一个粒子间距的阈值来定义. 之前的研究^[17,32]表明阈值的选择对最终结果的影响不是很大. 对于水团簇体系, 我们采用氧-氧原子之间的距离小于 5 Å 作为水分子近邻的判据. 有了上述定义, 我们可以对一个气相水分子体系统计其团簇尺寸分布. 例如, 对巨正则系综在特定的化学势下, 可以通过改变体系中水分子数目改变团簇尺寸. 困难在于由于成核自由能垒的存在, 接近临界成核尺寸的团簇出现的概率很低^[33-34]. 这意味着需要很长的模拟时间去达到平衡. 为了提高模拟效率, 需要引入增强采样方法.

在巨正则系综蒙特卡洛模拟中, 可能的采样尝试操作有两类. 一类是插入或者删除分子, 另一类是移动系统中原子分子的位置. 本研究中我们关注自由能随团簇尺寸的变化, 所以主要关注对第一类操作的采样加速. 在通常的模拟中, 插入分子会随机地放入模拟单胞中的任意位置. 由于气体浓度很低, 这样添加的分子使得团簇尺寸增加的几率很小. 而减少分子数目时, 如果所选取的分子属于一个团簇, 则该操作有较大的能量惩罚, 接受率较低. 总之, 由于气相分子体系的势能面主要由深而窄的势阱组成, 聚集过程需要克服熵的障碍, 而蒸发过程需要克服能量的障碍, 采样效率较低. 为此, 我们采用聚集体积偏倚蒙特卡洛(AVBMC)方法加速采样. 在插入一个新的分子时, 我们总是要求该分子添加到团簇上. 这实际上相当于将插入分子可能放置的空间由原来的整个模拟单胞限制到一个可以导致团簇增长的区域.

巨正则系综 AVBMC 模拟^[28]步骤如下: (1)随机选择是插入还是删除一个分子, (2)随机选择一个分子 j 作为目标分子, (3)如果是插入操作, 则将插入分子放置在目标分子周围; 如果是删除操作, 则从分子 j 周围选择一个分子删除, (4)计算初始状态 o 和最终状态 n 的能量差, $\Delta E = E_n - E_o$, (5)按如下的几率接受该操作. 插入操作的接受率为

$$\text{acc}(o \rightarrow n) = \min\left[1, w(r^{N+1}) \times \frac{N}{N+1} \times \frac{V_{\text{in}} \times \exp(\mu/k_B T) \times \exp(-\Delta E/k_B T)}{(N_{\text{in}}+1) \times q(T)}\right] \quad (1)$$

而删除操作的接受率为

$$\text{acc}(o \rightarrow n) = \min\left[1, w(r^{N-1}) \times \frac{N}{N-1} \times \frac{N_{\text{in}} \times q(T) \times \exp(-\mu/k_B T) \times \exp(-\Delta E/k_B T)}{V_{\text{in}}}\right] \quad (2)$$

上两式中第一项来自于满足团簇定义的限制, $w(r^N)$ 当 N 个粒子满足团簇定义时候等于 1, 否则等于 0. 第二项来自于选择目标分子步骤导致的插入和删除的几率差异, 其中 N 是系统中总粒子数. 第三项是从受限空间插入或删除分子的巨正则系综蒙特卡洛模拟所需满足的接受率(见支持信息, 公式 S3), 其中 $q(T)$ 是包含动量积分项部分的分子配分函数(对于原子体系 $q(T) = \Lambda^{-3}$, Λ 是热德布罗意波长), 化学势 $\mu = k_B T \ln \rho - k_B T \ln q(T)$, k_B 是玻尔兹曼常数, T 是模拟的温度, ρ 代表模拟的粒子数密度. N_{in} 和 V_{in} 分别代表按照团簇的定义目标分子 j 的近邻分子数以及其近邻区域的体积.

为了进一步提高模拟效率, 在选取目标分子时可以计算 Rosenbluth 因子, 对采样施加能量偏倚^[35]. 同时, 考虑到水分子本身由多个原子组成, 插入过程中确定了氧原子位置之后还需要确定两个氢原子位置, 可以通过构型偏倚提高接受率^[36-38].

基于上述 AVBMC 算法, 原则上模拟的时间足够长以后我们可以得到一个关于团簇尺寸的几率分布 $\Pi(N)$, 其中 N 为团簇内水分子的个数. 困难在于对于某些 N , 几率非常小. 因此, 要使 $\Pi(N)$ 收敛非常困难. 为此, 可以定义转移矩阵来统计在不同粒子数之间变化的几率^[29]. 从 N 分子状态变到 $N+\delta$ 分子状态的转移矩阵元为

$$C_{N,\delta} = C_{N,\delta} + p_a \quad (3)$$

$$C_{N,0} = C_{N,0} + 1 - p_a \quad (4)$$

其中 $p_a = \min(1, f_{o,n})$ 为该蒙特卡洛采样尝试操作的接受率, 其具体表达式见公式(1)和(2). 根据细致平衡原理, 可以通过转移矩阵估计 $\Pi(N)$. 为了达到增强采样的目的, 在实际的模拟中我们使用偏倚的接受率

$$P_\eta = \min\left[1, \frac{\exp(\eta(N_n))}{\exp(\eta(N_o))} f_{o,n}\right] \quad (5)$$

其中偏倚因子 $\eta(N) = -\ln \Pi(N)$ 时, 可以在不同分子数 N 的状态间实现随机行走. 由于 $\Pi(N)$ 正是我们需要通过模拟得到的量, 没法用来直接设置偏倚因子. 可以通过迭代法来解决这一问题. 首先随机设置 $\eta(N)$ 的初始值, 经过一定的模拟步以后, 通过转移矩阵更新对 $\Pi(N)$ 的估计, 并重新设置偏倚因子. 迭代多次以后, 偏倚因子基本保持不变. 这时, 我们已经得到 $\Pi(N)$ 的一个较好的估计, 同时实现了在不同分子数 N 的状态间的随机行走.

2.2 动力学成核理论

在动力学成核理论(DNT)^[39-41]中, 整个成核生长过程被分解为一系列单体从团簇蒸发和单体贴附到团簇上的气相分解和聚合反应过程.



在单体浓度足够小的情况下, 忽略多个单体同时从团簇蒸发或者贴附到团簇是一个合理的假设. 这样, 如果能得到所有的蒸发和贴附速率常数, 通过写出不同尺寸团簇浓度随时间变化的方程, 最终可以得到成核速率. 一个简单的假设是每次碰撞都产生贴附, 这样贴附速率常数可以通过气体运动学理论得到. 然后, 如果知道不同尺寸团簇的自由能根据细致平衡原理可以得到所有的蒸发速率常数. DNT 理论采用了一种更严格的相反的策略: 先通过变分过渡态理论得到蒸发速率常数, 再由蒸发速率常数根据细致平衡原理得到贴附速率常数.

在一个分子从团簇中蒸发的过程中, 整个体系的半径越变越大. 因此, DNT 理论采用团簇半径作为反应坐标, 以团簇质心为中心的球面作为反应物与产物的分割面. 事实上, 以质心为中心的半径早已被广泛应用于团簇定义^[42-44]. 该定义的难点在于如何确定截断半径的具体数值. DNT 理论通过变分过渡态理论来确定截断半径. 假设分隔反应物和产物的球面的半径为 r_{cut} , 反应流正比于过渡态配分函数 $q^{\text{GT}}(r_{\text{cut}}, T)$. 变分过渡态理论认为应该选取反应流最小的 $r^{\#}$ 对应的球面作为最终的反应物-产物分界面, 该半径即为团簇的截断半径. 在具体计算时可以统计最外层分子距离团簇质心的半径的分布, 该径向分布的最小值点给出截断半径 $r^{\#}$ ^[45]. 这时, 反应速率常数可以写成

$$k_f^{\text{TST}} = - \sqrt{\frac{\beta}{2\pi m}} \frac{dA(r_{\text{cut}}, T)}{dr_{\text{cut}}} \Big|_{r_{\text{cut}}=r^{\#}} \quad (7)$$

其中 m 是分子质量, $\beta = 1/k_B T$. 亥姆霍兹自由能 $A(r_{\text{cut}}, T) = -k_B T \ln q^{\text{R}}(r_{\text{cut}}, T)$, q^{R} 是反应物配分函数.

原则上, 我们可以通过足够长时间的蒙特卡洛模拟得到一个收敛的半径分布函数以及亥姆霍兹自由能随半径的变化, 从而得到蒸发速率常数. 在实际的模拟中, 为了加速收敛, 我们采用了伞形采样方法. 同时运行多

个蒙特卡洛模拟, 每个蒙特卡洛链受到一个偏势的限制使得团簇半径只分布在一个较窄的范围. 这样比较容易保证在该范围内的采样是充分的. 最后, 分段的分布可以通过加权直方分析(WHAM)^[46]合并成一个完整的分布. 从该分布可以得到亥姆霍兹自由能随截断半径的变化, 从而得到蒸发速率与截断半径.

3 结果与讨论

3.1 神经网络势

对水分子团簇体系的神经网络势的训练我们同时优化能量和力, 采用多轮迭代的方式产生训练集. 我们首先通过势能面随机游走算法^[47]对包含 2~10 或者 32 个水分子的 10 种水分子团簇各产生 200 个结构. 这里加入 32 个水团簇训练集是希望在一定程度上增加可移植性, 帮助后续训练产生更大团簇体系的机器学习势. 用这 2000 个结构进行第一性原理密度泛函理论(DFT)计算作为初始训练集. 基于不同的随机初始参数训练产生 4 套不同的初始神经网络势. 用其中一种势进行蒙特卡洛模拟, 对产生的结构计算四种不同神经网络势计算力的相对误差, 对最大相对误差在 0.25~0.35 eV/Å 之间的结构重新进行第一性原理计算. 将误差确实较大的结构

加入训练集重新训练. 如此迭代, 直到能量误差最后到达了几个 meV/atom 的量级, 力的误差在 0.15 eV/Å 以下时, 停止训练. 最后训练集的水分子数目分布图展示在图 S12(见支持信息). 在后续的模拟中, 我们从四套神经网络势中随机取一套作为最终的机器学习势模型.

为了测试神经网络势的精度, 我们对含 2、5 和 10 个水分子的体系进行了测试. 对每个体系, 通过势能面随机游走算法产生 200 个结构作为测试集. 对 2 个水分子的体系, DFT 计算 200 个结构能量和力的标准差分别是 0.6 eV 和 1.4 eV/Å. 对比神经网络势的预测和 DFT 计算结果, 均方根误差(RMSE)分别是 0.032 eV 和 0.096 eV/Å, 相对误差分别是 5.3%和 6.9%. 对于 5 个水的体系, DFT 能量和力的标准差是 0.47 eV 和 0.79 eV/Å. RMSE 分别是 0.031 eV 和 0.051 eV/Å, 相对误差分别是 6.6%和 6.5%. 对于 10 个水分子的体系, DFT 能量和力的标准差分别是 0.85 eV 和 0.57 eV/Å. 能量和力的 RMSE 分别是 0.055 eV 和 0.045 eV/Å, 相对误差分别是 6.5%和 7.9%. 三个体系的具体数据散点图见图 S1、图 S2(见支持信息)和图 1. 我们对水分子团簇拟合的神经网络势的精度与之前对液态水的神经网络势模型^[25]的精度接近.

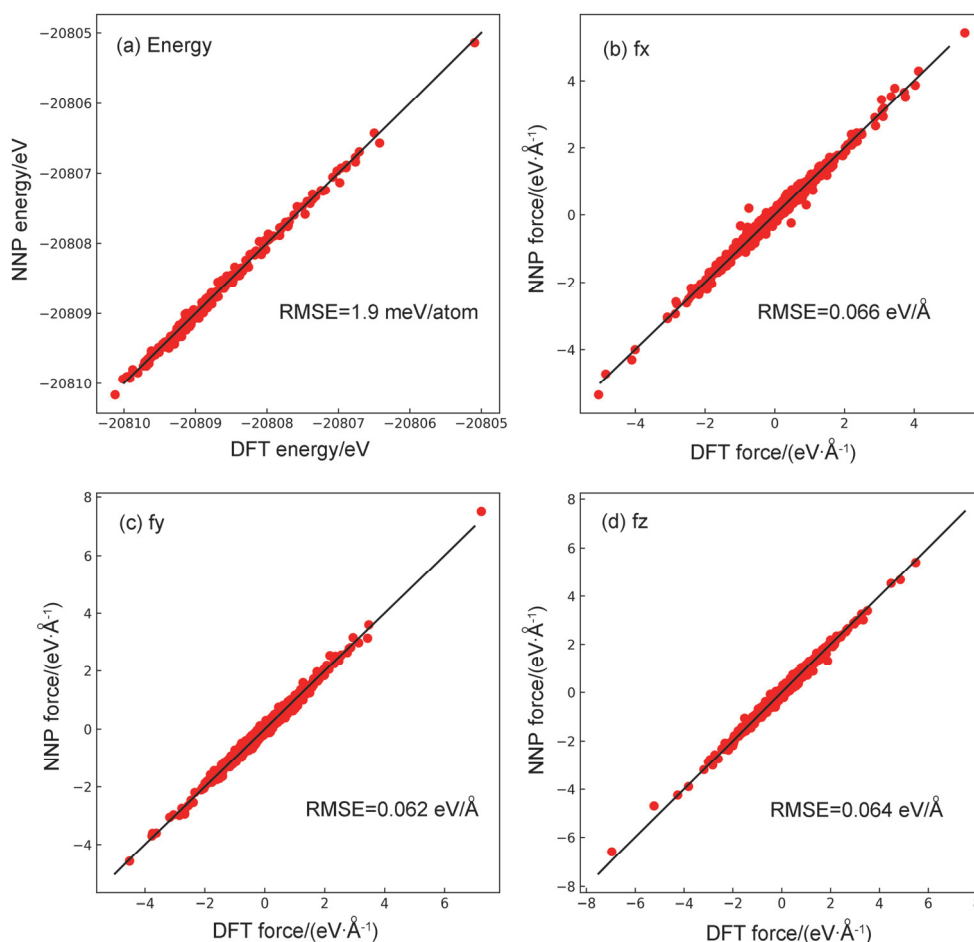


图 1 十个水分子组成的团簇神经网络势与第一性原理能量与力的对比图

Figure 1 Comparison of (a) energies and (b~d) forces from neural network potential (NNP) and DFT for the cluster with ten water molecules

作为一个额外的测试,我们对水分子团簇体系用神经网络势进行了一次巨正则系综蒙特卡洛模拟.模拟过程中团簇的尺寸遍历1~10的范围.对 1×10^7 蒙特卡洛步我们每1000步进行一次取样,然后用我们得到的四套不同的神经网络势进行能量和力的计算(见支持信息,图S3).结果表明不同势模型的能量均方根误差在3.5 meV/atom以下,力的均方根误差在0.1 eV/Å以下,最大力误差在0.20 eV/Å以下.

3.2 自由能与团簇尺寸的关系

首先,对经典力场,我们采用 AVBMC 结合转移矩阵蒙特卡洛方法重复了之前文献^[34]的结果(见支持信息,图S4),表明了该方法的可靠性.接下来我们用拟合的神经网络势进行了模拟.我们选择了5个不同的化学势进行模拟,分别对应于300 K温度下理想气体浓度为 1.0×10^{-5} 、 5.0×10^{-6} 、 2.5×10^{-6} 、 1.8×10^{-6} 和 1.0×10^{-6} Å⁻³时的化学势.我们将这5种条件分别命名为s1到s5,对应过饱和度越来越低.以s1为例,选择蒙特卡洛模拟最后 5×10^7 步每1000步取一次样.对应的团簇尺寸分布(见支持信息,图S5)显示我们实现了在不同团簇尺寸间的随机行走.尺寸为1~10的水分子团簇出现的概率都大概等于10%,其在模拟不同阶段的分布也已经足够均匀(见支持信息,图S6).

从团簇尺寸概率分布我们可以计算吉布斯自由能随团簇尺寸的变化(图2a).得到的结果在定性上与SPC/E经典力场结果是一致的.在s1~s4的条件下,临界核的尺寸分别为3、3、7和8.随着过饱和度的降低,临界核的尺寸越来越大.同时,在自由能随团簇尺寸变化的曲线中,我们看到水分子数目为4的团簇自由能有明显的降低趋势.为了更好地看清楚这一趋势,在图2b中,我们画出了蒙特卡洛模拟中转移矩阵对应不同尺寸团簇间转移的矩阵元的值.整体趋势是随着团簇尺寸增

加,水分子贴附的趋势逐渐增加.这与团簇变大以后碰撞几率增加的趋势是一致的.对蒸发而言,我们看到随着原子尺寸增加蒸发一般有减弱的趋势.而从四聚体到五聚体是一个例外.五聚体的蒸发概率显著地大于四聚体.之前经典力场的模拟中也看到了类似的规律^[17],我们的神经网络势研究表明这不是经典力场由于精度原因产生的一个虚假结果.因此,需要对这一异常形成的原因进行细致分析.

3.3 团簇蒸发速率

上面提到的转移矩阵的矩阵元虽然代表了热力学上水分子从团簇蒸发以及水分子贴附到团簇的趋势,但是定量地分析蒸发速率还需要借助变分过渡态理论.为此,我们对不同数目的水分子体系进行了蒙特卡洛模拟,统计半径分布函数.在正式模拟之前,我们先对两个水分子的体系进行测试.为了与之前文献一致,测试模拟中对水分子采用刚体模型,模拟温度为243 K.可以看到,对截断半径不同方法预测的结果差别不大,但是极化力场预测的蒸发速率常数^[26]比基于高精度量子化学计算(MP2)的模拟结果^[45]高出一个量级(见支持信息,表S1),表明其描述的水分子结合趋势明显偏弱.而采用神经网络势以后,预测的蒸发速率常数得到了较好的改善,优于Hartree-Fock以及有效片段势(EFP2)等基于第一性原理的结果^[48].完成测试后,我们在300 K下对2~6和10个水分子的体系分别进行了模拟.这时,水分子的内部自由度在模拟中完全放开.表1给出了水分子蒸发速率常数与截断半径随着分子数变化的关系.就像可以预期的一样,团簇半径随着团簇包含的水分子数目而增加.同时,我们观察到蒸发速率随着团簇内水分子数目的增加而减小.但是从四聚体到五聚体蒸发速率常数增加.这与前面巨正则系综转移矩阵模拟的结论(图2b)是一致的.

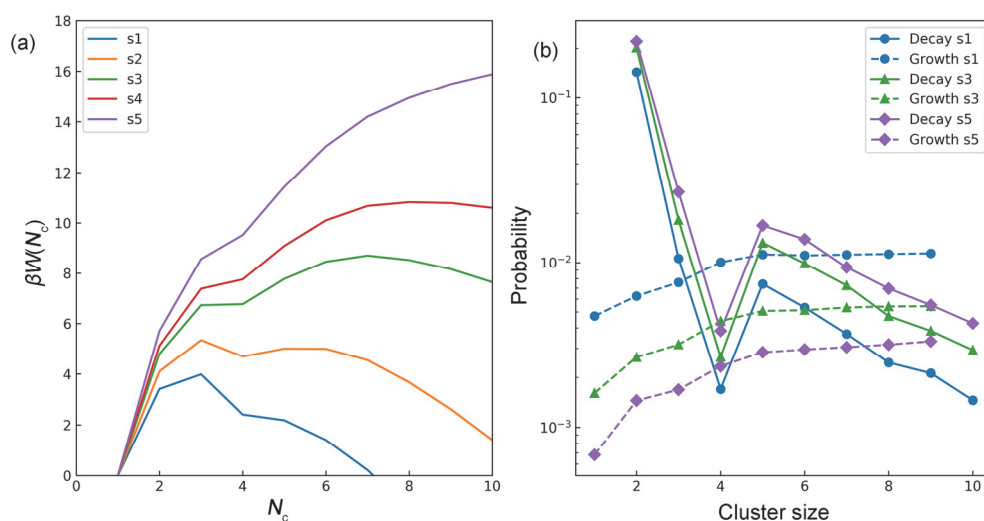


图2 (a) 自由能随着水团簇尺寸的变化. (b) 巨正则系综平均增长和衰减概率随团簇尺寸的变化
Figure 2 (a) Gibbs free-energy and (b) grand canonical average growth and decay probabilities as a function of cluster size

表1 300 K 下水团簇截断半径与蒸发速率常数

Table 1 Cutoff radii and evaporation rate constants of water clusters at 300 K

Cluster	(H ₂ O) ₂	(H ₂ O) ₃	(H ₂ O) ₄	(H ₂ O) ₅	(H ₂ O) ₆	(H ₂ O) ₁₀
$r_{\text{cut}}/\text{\AA}$	2.56	4.10	5.03	5.45	6.11	7.530
Rate constant/ ($\times 10^9 \text{ s}^{-1}$)	159.70	25.52	2.86	9.56	7.39	1.965

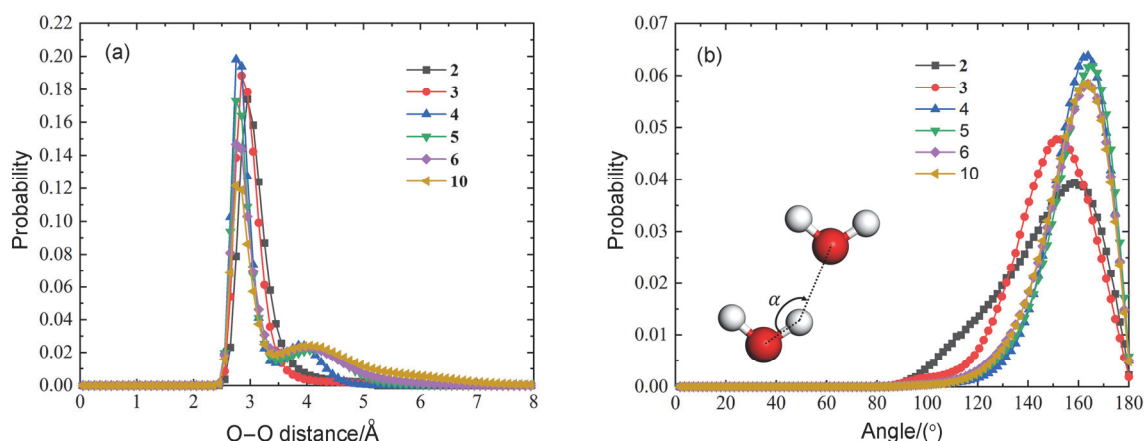
为了探究水团簇蒸发速率与结构之间的内在联系,我们将分子模拟得到的半径小于截断半径的团簇结构进行了统计分析. 首先分析 O—O 径向分布. 如图 3a 所示, (H₂O)₂ 团簇的 O—O 径向分布函数只有一个主峰位于 2.95 Å, 对应于氢键的 O—O 键长, 与文献报道的高精度量子化学计算(MP2)给出的 O—O 距离 2.94 Å^[49] 非常接近. 这也反映了神经网络势的高精度. 随着水分子数目从 2 个增加到 4 个, 可以看到 O—O 距离在逐步的缩小, 这是由于多体协同效应产生的氢键的增强^[50]. 这与我们观察到的蒸发速率常数从 2 到 4 个水分子的减弱趋势是一致的. 同时, 我们也观察到当团簇中水分子数目达到 4 的时候开始出现位于 3.5~5.5 Å 比较宽的次峰, 这对应于液态水的第一溶剂化层的特征半径^[51], 来源于三维氢键网络. 随着水分子数从 4 增加到 10, 次峰的比重不断增加, 表明三维氢键网络不断完善.

我们还对 OH—O 夹角进行了统计分析(如图 3b 所示). OH—O 夹角 α 定义为供体 H—O 键矢量和供体 H 指向受体 O 矢量的夹角(如图 3 的插图所示), 描述氢键的弯折程度. 180° 表示线性氢键. 在 (H₂O)₂ 团簇, α 角度分布在 160° 呈现极大值, 整体的分布相对比较弥散. 虽然 (H₂O)₃ 团簇主峰偏移到 151.2°, 但是相对二聚体来说 (H₂O)₃ 团簇 OH—O 角度分布变得更集中, 重度弯折的构型明显减少. 随着水分子数增加到 4 个, α 角度分布的峰值移到 165.0° 附近, 同时峰宽明显变窄. 水分子数从 4 进一步增长到 10 时, α 角度分布基本保持稳定. 上述结果表明有多个水分子时, 氢键形成网络, 互相制约, 给出较稳定的分布.

上面的分析表明氢键在水分子团簇结构中起着重

要作用. 如果要对氢键数目进行定量分析, 需要一个氢键的定义. 这里我们采用文献中常用的定义^[52]: 供体 O_d 到受体 O_a 距离小于 3.5 Å 且 $\angle \text{H—O}_d \cdots \text{O}_a < 30^\circ$. 基于该定义, 我们将不同团簇的统计分布按氢键个数分类. 对 (H₂O)_n ($n=2\sim 6, 10$) 团簇氢键数目分辨的能量分布如图 4 所示. (H₂O)₂ 团簇主要由 0 和 1 个氢键的构型组成, 1 个氢键的比重略大于 0 个氢键, 二者都接近 50%, 这反映了整个二聚体团簇的不稳定性非常高. 对于 (H₂O)₃ 团簇主要有 2~4 个氢键的构型, 其中 2 个氢键的分布最多, 达到 54%. 注意到 (H₂O)₃ 的能量最低构型是含三个氢键的三元环, 但所有含三个氢键构型的占比不足 20%. 这也表明了非能量最低构型的重要作用. 对于 (H₂O)₄ 团簇主要有 2~4 个氢键的构型, 4 个氢键的占比最高, 达到 75%. 对于 (H₂O)₅ 团簇主要有 3~6 个氢键的构型, 5 个氢键占比最高, 达到了 66%. 六聚体主要有 4~8 个氢键的构型, 其中以 6 个氢键为主, 达到 44%. 十聚体主要由 8~15 个氢键构型组成, 其中 12 个氢键占比最高, 达到 30%.

此外, 我们还统计了水分子蒸发反应过渡态的氢键组成. (H₂O)₂ 的蒸发的过渡态普遍不具有氢键; (H₂O)₃ 的蒸发过渡态 1 个氢键贡献比较大, 达到 61%; (H₂O)₄ 的蒸发过渡态两个氢键贡献为主达到 48%; (H₂O)₅ 的蒸发过渡态主要由 4 个氢键组成, 达到 68%; (H₂O)₆ 的蒸发过渡态主要由 5 氢键组成, 达到 60%; (H₂O)₁₀ 的蒸发过渡态主要由 10、11 和 9 个氢键组成, 占比分别达到 27%、25% 和 22%. 根据反应物和过渡态的各类氢键占比, 我们计算了反应物和过渡态的平均氢键数目以及蒸发所需打破的氢键个数(表 2). 可以看到, 需要破坏的氢键数目与我们模拟得到的蒸发速率变化趋势大体一致. (H₂O)₄ 中蒸发一个水分子需要打破的氢键数目比较多, 因此其蒸发速率较低. 当然, 需要破坏的氢键数目只能定性解释蒸发速率常数, 氢键强度以及其他多体协同效应也将影响蒸发速率常数. 事实上, 三聚体蒸发速率大于五聚体, 与这里氢键数目的预测并不一致.

图3 300 K 下水团簇(H₂O)_n ($n=2\sim 6, 10$)的(a) O—O 径向分布函数和(b) OH—O 角度分布函数Figure 3 (a) Oxygen-oxygen distance and (b) OH—O angle distributions of (H₂O)_n ($n=2\sim 6, 10$) at 300 K

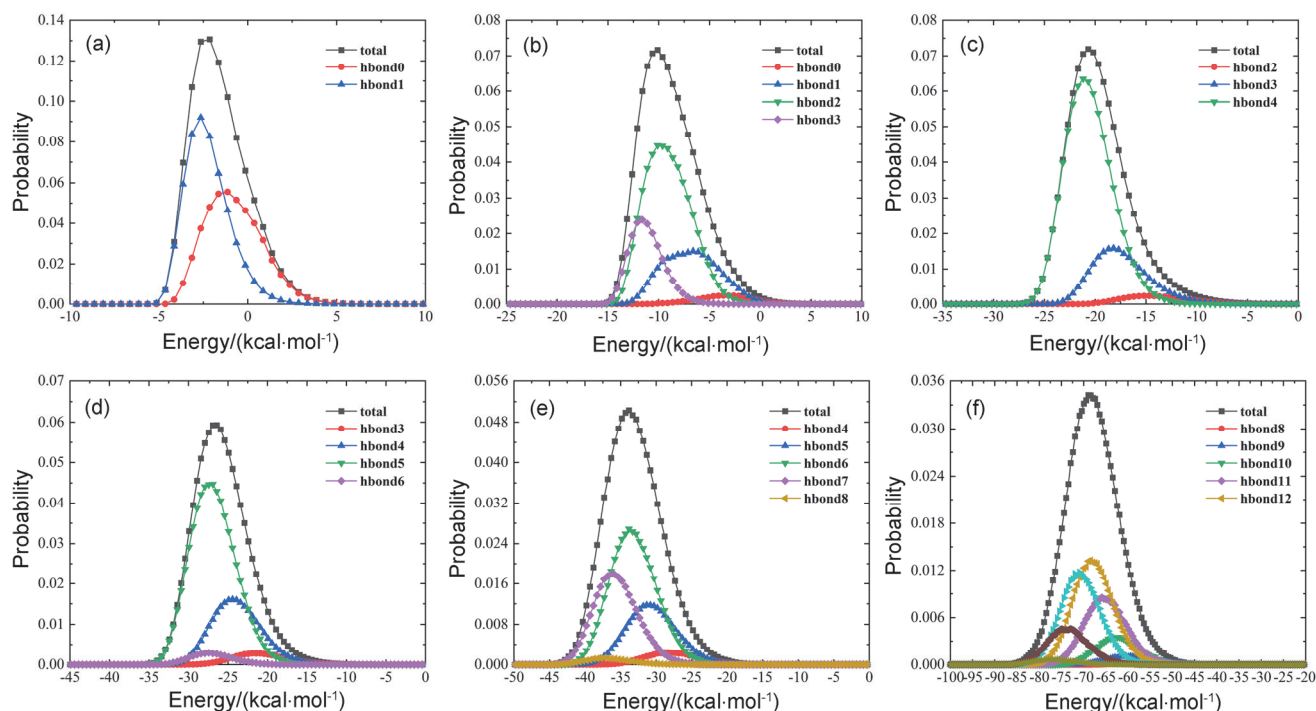


图4 300 K 下 $(\text{H}_2\text{O})_n$ ($n=2\sim6, 10$) 不同氢键数目构型能量分布

Figure 4 Energy distribution of $(\text{H}_2\text{O})_n$ ($n=2\sim6, 10$) configurations with different numbers of hydrogen bonds at 300 K

表2 300 K 下 $(\text{H}_2\text{O})_n$ ($n=2\sim6, 10$) 团簇蒸发反应物(IS)、过渡态(TS)氢键平均数目和蒸发过程打破氢键数目

Table 2 Average number of hydrogen bonds in reactants (IS), transition state (TS) and break hydrogen bond number of the evaporation reaction in $(\text{H}_2\text{O})_n$ ($n=2\sim6, 10$) at 300 K

Hbond number	2	3	4	5	6	10
IS	0.5	1.9	3.7	4.7	6.0	12.0
TS	0.0	0.6	1.8	3.6	4.6	10.0
Break Hbond	0.5	1.3	1.9	1.1	1.4	2.0

3.4 最低能量结构与最概然结构

为了获得一个直观的图像,我们对不同体系的最低能量结构和最概然结构进行了分析.从巨正则系综模拟和正则系综模拟得到的结构进行分析得到的结果是类似的.事实上,将巨正则系综产生的结构按团簇尺寸和半径进行统计,得到的结果和正则系综模拟得到的结果几乎一致(见支持信息,图 S11).唯一的区别在于前者得到的团簇结构全部落在反应物区间,过渡态附近几乎没有采样.因此,我们这部分的分析主要用正则系综的结果进行. $(\text{H}_2\text{O})_n$ ($n=2\sim6, 10$) 团簇的最低能量结构如图 S8(见支持信息)所示. $(\text{H}_2\text{O})_2$ 团簇的最低能量构型具有一个氢键, $\text{OH}\cdots\text{O}$ 角度(176.6°)非常接近线性氢键,键长为 $2.89 \text{ \AA}$. $(\text{H}_2\text{O})_3$ 能量最低的构型接近正三角形, $\text{O}\cdots\text{O}\cdots\text{O}$ 平均角度为 60.0° , $\text{O}\cdots\text{O}$ 平均键长为 $2.77 \text{ \AA}$, 氢键角度平均为 154.0° . 水分子中不参与形成氢键的自由氢原子相对于氢键网络形成的二维平面的指向为上

下. $(\text{H}_2\text{O})_4$ 能量最低的构型接近正方形, $\text{O}\cdots\text{O}\cdots\text{O}$ 平均角度为 89.8° , $\text{O}\cdots\text{O}$ 平均键长为 $2.82 \text{ \AA}$, 氢键角度平均为 165.7° . 水分子中不参与形成氢键的自由氢原子相对于氢键网络形成的二维平面的指向为上下上下排布. $(\text{H}_2\text{O})_5$ 能量最低的构型是具有五个氢键的手性同向旋转环状五聚体^[53], $\text{O}\cdots\text{O}\cdots\text{O}$ 平均角度为 106.9° , $\text{O}\cdots\text{O}$ 平均键长为 $2.72 \text{ \AA}$, 氢键角度平均为 170.6° . 这与 5 K 下分子束偏转实验^[54]得出的五聚体的结构($\text{O}\cdots\text{O}$ 平均实验距离为 $2.76 \text{ \AA}$, $\text{O}\cdots\text{O}\cdots\text{O}$ 平均角度为 108.0° , 氢键角度接近线性)吻合较好. $(\text{H}_2\text{O})_6$ 能量最低的构型是三棱柱. $(\text{H}_2\text{O})_{10}$ 能量最低的构型是直六棱柱.我们从分子模拟中得到的最低能量构型和之前进行全局结构优化得到的结果^[55]吻合较好,这也从一个侧面表明我们的采样充分、神经网络势精度高.

我们采用 molclu 软件^[56]对蒙特卡洛采样得到的构型进行了聚类.最概然构型为包含该构型结构数最多的类.同一类中我们展示能量最低的结构.3~6 个水分子组成的团簇的最概然构型如图 5 所示.对三聚体和四聚体,氧原子大致处于一个平面内,呈现一个环形结构.最概然结构与最低能量结构构型一致.对五聚体,最概然结构不再是一个平面构型,而是一个四元环结构加一个额外的水分子(4+1 结构).这就解释了为什么我们看到五聚体更加容易蒸发.从四聚体到五聚体,最概然结构从平面结构过渡到三维结构,因此我们的模拟发现在自由能、蒸发速率等特征上都在这一位置发生了趋势的变化.另外,对于五聚体最概然构型不同于最低能量构

型, 表明了进行分子模拟的必要性. 通过全局优化寻找最低能量构型, 然后通过简谐近似计算自由能的方法对团簇体系并不总是适用的.

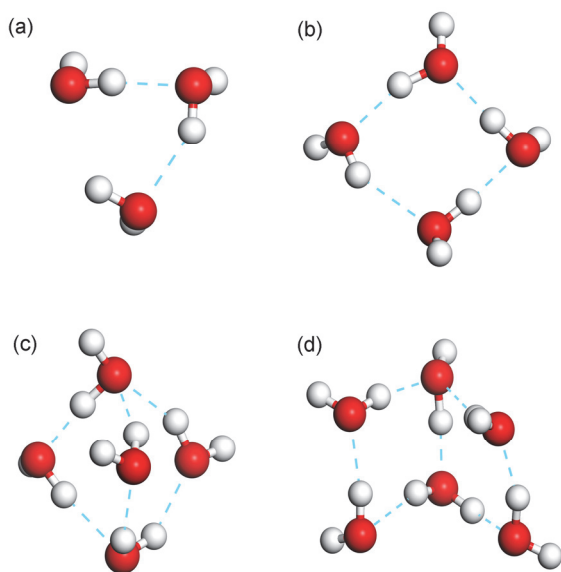


图5 300 K 下 $(\text{H}_2\text{O})_n$ 的最概然结构. (a) $n=3$, (b) $n=4$, (c) $n=5$, (d) $n=6$

Figure 5 Most probable conformers of $(\text{H}_2\text{O})_n$ cluster at 300 K. (a) $n=3$, (b) $n=4$, (c) $n=5$, (d) $n=6$

$(\text{H}_2\text{O})_n$ ($n=2\sim6, 10$) 团簇的最低能量过渡态结构如图 S9(见支持信息)所示. 水二聚体过渡态最低能量构型是由两个取向夹角为 59.3° 的水分子构成. 三聚体能量最低的过渡态是以一个由氢键连接的二聚体和另外一个水分子组成, 该水分子的 H 指向二聚体中一个水分子的 O. 四聚体的最低能量过渡态以三个氢键形成的三元环和一个环外的水分子组成, 该水分子的一个 H 指向三元环平面. 五聚体能量最低的过渡态是由一个四水环形团簇和一个额外水分子组成. 该水分子的两个 H 一个指向四元环中的一个水分子另一个接近平行于四元环. 六聚体能量最低的过渡态也是以五元环和一个水分子组成. 蒸发水分子的 OH 与五元环共面, 另一个 H 朝外与五元环面呈 107.0° 夹角. $(\text{H}_2\text{O})_{10}$ 最低能量过渡态构型是由 13 个氢键连接的九聚体最低能量结构(一个 D_{2d} 八面体中的一条边额外与一个水形成氢键形成了一种两个供体一个给体的三角结构)^[57] 和一个额外的水分子组成. 该水分子的两个 H 完全朝外侧不指向九聚体. 从上面结果可以看出最低能量过渡态的结构一般是一个水分子基本挣脱了氢键的束缚, 剩下的部分则尽量保持其能量最低环状或者氢键网络结构存在.

4 结论

我们拟合了一套具有第一性原理精度的机器学习势来准确描述水团簇相互作用, 修正了之前经验力场结

合能偏弱的趋势. 通过采用 AVBMC 结合转移矩阵蒙特卡洛方法, 我们实现了水团簇在水分子数反应坐标上的全局随机行走. 由此得到了在第一性原理精度下的水团簇吉布斯自由能随水分子数的变化. 我们的结果确认了在四聚体到五聚体处存在一个突变. 该突变需要通过分子模拟才能发现, 如果简单考虑能量最低构型, 五聚体像四聚体一样也倾向保持二维环状结构. 而对蒙特卡洛模拟采样的结构进行聚类分析可以发现, 最概然的五聚体结构由水四元环和一个通过氢键接在环上的水分子组成. 五聚体的这种“4+1”结构导致蒸发比较容易, 蒸发速率常数更大. 本工作采用的研究方案可以推广到更复杂的硫酸、水二元体系以及硫酸、氨、水三元体系等与大气颗粒物成核生长密切相关的体系.

5 计算细节

我们采用深度神经网络的平滑版(DeepPot-SE)模型^[58]进行训练. 在该模型中总势能等于所有原子能量的求和: $E=\sum E_i$. 每个原子的能量由该原子周围其他近邻原子的种类及相对位置所决定: $E_i=E(R_i, \{R_j\})$. 这里近邻列表半径设置为 8 Å, 在 7.8 Å 开始使用平滑函数. 输入端到输出端的大小分别为 50 和 100. 采用 5 个隐藏层, 每层 120 个节点. 学习率从 0.001 开始, 每 10000 步衰减率为前一步的 95%. 最终的训练集包含 17343 个构型. 这些构型的能量和力通过 Gaussian 16 软件包^[59]采用 B3LYP 杂化交换关联泛函加 DFT-D3 长程色散作用修正^[60]和 6-311++G(d, p) 基组计算. 为了提高拟合精度, 我们将长程库伦相互作用单独剥离^[61]. 其截断与 DeePMD 势能的截断保持一致都设为 8 Å. 氧原子和氢原子的电荷设定为 SPC/E 经典力场模型中值, 分别为 -0.8476 e 和 0.4238 e . 损失函数考虑了能量和力, 训练初始的能量和力的权重为 1 : 100, 训练结束后调整权重比到 1 : 1. 由于蒙特卡洛模拟中只有能量被直接用到, 我们以得到的模型参数为初始值将能量和力的权重设为 100 : 1 进行一次重复训练.

AVBMC 结合转移矩阵蒙特卡洛模拟的温度为 300 K. 水分子的数目截断到 10. 模拟的盒子大小为 $40\text{ Å} \times 40\text{ Å} \times 40\text{ Å}$. 总模拟步长不小于 10^8 步, 45% 的采样尝试操作是粒子数交换, 20% 是分子内部的原子移动, 20% 是分子整体的平移, 剩下的 15% 是单个分子的旋转. 每 5000 步调节一下平动和转动的步长, 使其接受率保持在 50% 左右. 每 10000 步更新一次转移矩阵.

DNT 模拟在 300 K 下 NVT 系综进行. 为了避免采样过程出现超出神经网络力场预测能力之外的构型, 我们限制了分子内键长不小于 0.85 Å 且不大于 1.15 Å, 键角不小于 73° 且不大于 120° . $(\text{H}_2\text{O})_2$ 、 $(\text{H}_2\text{O})_3$ 、 $(\text{H}_2\text{O})_4$ 、 $(\text{H}_2\text{O})_5$ 、 $(\text{H}_2\text{O})_6$ 和 $(\text{H}_2\text{O})_{10}$ 的最小采样半径分别为 0、1.25、1.25、1.25、1.25 和 2.25 Å, 最大采样半径分别为 4.00、6.25、8.25、8.25、8.25 和 10.25 Å. 在采样半径范围内

均匀放置 48 个伞势, 其力常数为 $0.02 \text{ Hartree}/\text{\AA}^2$. 每条蒙特卡洛链步数为 3×10^8 步.

致谢

数值计算得到了中国科学技术大学超级计算中心和鸿之微云超算的支持.

References

- [1] Elm, J.; Kubečka, J.; Besel, V.; Jääskeläinen, M. J.; Halonen, R.; Kurtén, T.; Vehkamäki, H. *J. Aerosol. Sci.* **2020**, *149*, 105621.
- [2] Jahl, L. G.; Brubaker, T. A.; Polen, M. J.; Jahn, L. G.; Cain, K. P.; Bowers, B. B.; Fahy, W. D.; Graves, S.; Sullivan, R. C. *Sci. Adv.* **2021**, *7*, eabd3440.
- [3] Zhu, J.; Penner, J. E.; Yu, F.; Sillman, S.; Andreae, M. O.; Coe, H. *Nat. Commun.* **2019**, *10*, 1.
- [4] He, X.-C.; Tham, Y. J.; Dada, L.; Wang, M.; Finkenzeller, H.; Stolzenburg, D.; Iyer, S.; Simon, M.; Kürten, A.; Shen, J. *Science* **2021**, *371*, 589.
- [5] Zhang, B.; Yu, Y.; Zhang, Y.-Y.; Jiang, S.; Li, Q.; Hu, H.-S.; Li, G.; Zhao, Z.; Wang, C.; Xie, H. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 15423.
- [6] Wyslouzil, B. E.; Wölk, J. *J. Chem. Phys.* **2016**, *145*, 211702.
- [7] Wölk, J.; Strey, R. *J. Phys. Chem. B* **2001**, *105*, 11683.
- [8] Viisanen, Y.; Strey, R.; Reiss, H. *J. Chem. Phys.* **1993**, *99*, 4680.
- [9] Laaksonen, A.; Talanquer, V.; Oxtoby, D. W. *Annu. Rev. Phys. Chem.* **1995**, *46*, 489.
- [10] Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
- [11] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- [12] Horn, H. W.; Swope, W. C.; Pitner, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665.
- [13] Abascal, J. L.; Vega, C. *J. Chem. Phys.* **2005**, *123*, 234505.
- [14] Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. *J. Phys. Chem. Lett.* **2014**, *5*, 3863.
- [15] Zhao, M. Y.; Yang, X. P.; Yang, X. N. *Acta Phys.-Chim. Sin.* **2015**, *31*, 1489 (in Chinese). (赵梦尧, 杨雪平, 杨晓宁, 物理化学学报, **2015**, *31*, 1489.)
- [16] Sun, Y. R.; Yu, F.; Ma, J. *Acta Phys.-Chim. Sin.* **2017**, *33*, 2173 (in Chinese). (孙怡然, 于飞, 马杰, 物理化学学报, **2017**, *33*, 2173.)
- [17] Merikanto, J.; Vehkamäki, H.; Zapadinsky, E. *J. Chem. Phys.* **2004**, *121*, 914.
- [18] Behler, J.; Parrinello, M. *Phys. Rev. Lett.* **2007**, *98*, 146401.
- [19] Bartók, A. P.; Csányi, G. *Int. J. Quantum Chem.* **2015**, *115*, 1051.
- [20] Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. *J. Chem. Phys.* **2018**, *148*, 241722.
- [21] Smith, J. S.; Isayev, O.; Roitberg, A. E. *Chem. Sci.* **2017**, *8*, 3192.
- [22] Khorshidi, A.; Peterson, A. A. *Comput. Phys. Commun.* **2016**, *207*, 310.
- [23] Huang, S. D.; Shang, C.; Kang, P. L.; Zhang, X. J.; Liu, Z. P. *WIREs Comput. Mol. Sci.* **2019**, *9*, e1415.
- [24] Kang, P.-L.; Shang, C.; Liu, Z.-P. *Chin. J. Chem. Phys.* **2021**, *34*, 583.
- [25] Wang, H.; Zhang, L.; Han, J.; Weinan, E. *Comput. Phys. Commun.* **2018**, *228*, 178.
- [26] Kathmann, S. M.; Schenter, G. K.; Garrett, B. C. *J. Chem. Phys.* **1999**, *111*, 4688.
- [27] Jorgensen, W. L.; Tirado-Rives, J. *J. Phys. Chem.* **1996**, *100*, 14508.
- [28] Chen, B.; Siepmann, J. I.; Oh, K. J.; Klein, M. L. *J. Chem. Phys.* **2001**, *115*, 10903.
- [29] Paluch, A. S.; Shen, V. K.; Errington, J. R. *Ind. Eng. Chem. Res.* **2008**, *47*, 4533.
- [30] Kästner, J. *Wiley Interdiscip. Rev.- Comput. Mol. Sci.* **2011**, *1*, 932.
- [31] Stillinger Jr., F. H. *J. Chem. Phys.* **1963**, *38*, 1486.
- [32] Chen, B.; Siepmann, J. I.; Klein, M. L. *J. Phys. Chem. A* **2005**, *109*, 1137.
- [33] Oh, K.; Zeng, X. C. *J. Chem. Phys.* **2000**, *112*, 294.
- [34] Kusaka, I.; Wang, Z.-G.; Seinfeld, J. H. *J. Chem. Phys.* **1998**, *108*, 3416.
- [35] Loeffler, T. D.; Sepehri, A.; Chen, B. *J. Chem. Theory Comput.* **2015**, *11*, 4023.
- [36] Wick, C. D.; Siepmann, J. I. *Macromolecules* **2000**, *33*, 7207.
- [37] Siepmann, J. I. *Mol. Phys.* **1990**, *70*, 1145.
- [38] Martin, M. G.; Frischknecht, A. L. *Mol. Phys.* **2006**, *104*, 2439.
- [39] Kathmann, S. M.; Schenter, G. K.; Garrett, B. C.; Chen, B.; Siepmann, J. I. *J. Phys. Chem. C* **2009**, *113*, 10354.
- [40] Schenter, G. K.; Kathmann, S. M.; Garrett, B. C. *Phys. Rev. Lett.* **1999**, *82*, 3484.
- [41] Schenter, G. K.; Kathmann, S. M.; Garrett, B. C. *J. Chem. Phys.* **1999**, *110*, 7951.
- [42] Reiss, H.; Katz, J.; Cohen, E. *J. Chem. Phys.* **1968**, *48*, 5553.
- [43] Reiss, H.; Tabazadeh, A.; Talbot, J. *J. Chem. Phys.* **1990**, *92*, 1266.
- [44] Lee, J. K.; Barker, J.; Abraham, F. F. *J. Chem. Phys.* **1973**, *58*, 3166.
- [45] Crosby, L. D.; Kathmann, S. M.; Windus, T. L. *J. Comput. Chem.* **2009**, *30*, 743.
- [46] Souaille, M.; Roux, B. *Comput. Phys. Commun.* **2001**, *135*, 40.
- [47] Shang, C.; Liu, Z.-P. *J. Chem. Theory Comput.* **2013**, *9*, 1838.
- [48] Devarajan, A.; Windus, T. L.; Gordon, M. S. *J. Phys. Chem. A* **2011**, *115*, 13987.
- [49] Scheiner, S. *Annu. Rev. Phys. Chem.* **1994**, *45*, 23.
- [50] Qian, P.; Song, W.; Lu, L.; Yang, Z. *Int. J. Quantum Chem.* **2010**, *110*, 1923.
- [51] Netzloff, H. M.; Gordon, M. S. *J. Chem. Phys.* **2004**, *121*, 2711.
- [52] Luzar, A.; Chandler, D. *Phys. Rev. Lett.* **1996**, *76*, 928.
- [53] Day, M. B.; Kirschner, K. N.; Shields, G. C. *J. Phys. Chem. A* **2005**, *109*, 6773.
- [54] Keutsch, F. N.; Saykally, R. J. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10533.
- [55] Qian, P.; Song, W.; Lu, L.; Yang, Z. *Int. J. Quantum Chem.* **2010**, *110*, 1923.
- [56] Lu, T. *Molclus Program*, Version 1.9.9.7, <http://www.keinsci.com/research/molclus.html>.
- [57] Rodriguez, J.; Moriena, G.; Laria, D. *Chem. Phys. Lett.* **2002**, *356*, 147.
- [58] Zhang, L.; Han, J.; Wang, H.; Saidi, W. A.; Car, R. arXiv preprint arXiv:1805.09003. **2018**.
- [59] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16*, Revision C.01, Gaussian, Inc., Wallingford CT, **2016**.
- [60] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
- [61] Yue, S.; Muniz, M. C.; Calegari Andrade, M. F.; Zhang, L.; Car, R.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **2021**, *154*, 034111.

(Cheng, B.)