Supporting Information for: Brick-CFCMC: open source software for Monte Carlo simulations of phase and reaction equilibria using the Continuous Fractional Component method

Remco Hens,[†] Ahmadreza Rahbari,[†] Sebastián Caro-Ortiz,[†] Noura Dawass,[†] Máté Erdős,[†] Ali Poursaeidesfahani,[†] Hirad S. Salehi,[†] Alper T. Celebi,[†] Mahinder Ramdin,[†] Othonas A. Moultos,[†] David Dubbeldam,[‡] and Thijs J.H. Vlugt[*,†]

†*Engineering Thermodynamics, Process & Energy Department, Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Leeghwaterstraat 39, 2628CB Delft, The Netherlands*

‡*Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, Science Park 904, 1098XH Amsterdam, The Netherlands*

E-mail: t.j.h.vlugt@tudelft.nl

Brick-CFCMC is an open-source software package for force field-based Monte Carlo simulations in open ensembles (e.g. the Grand-Canonical Ensemble,[1,2] the Gibbs Ensemble,[3,4] and, the Reaction Ensemble[5]). It uses the Continuous Fractional Component method[6–10] to obtain phase and/or reaction equilibria. One can also calculate properties such as chemical potentials, partial molar enthalpies, and partial molar volumes directly from a single simulation.[6–8,11] Brick-CFCMC is written in Fortran 90/95, terminal based, and works most conveniently in combination with the Bash shell. This document contains a description of some of the features in Brick-CFCMC. For a more detailed overview of the possibilities of Brick-CFCMC, we refer the reader to the manual of Brick which can be found at `https://gitlab.com/ETh_TU_Delft/Brick-CFCMC`

Origin of the name Brick

This software package is called Brick after the character Brick Heck, from the TV series *The Middle*, whose parents gave him this name in the hope of that giving him a special name would make him do special things.

The Continuous Fractional Component Method

The Continuous Fractional Component (CFC) method uses expanded versions of thermodynamic ensembles.[7–10,12] Extra molecule(s) are added to the system, so-called fractional molecules, with an additional degree of freedom: the fractional parameter λ. Other molecules in the system, that are not the fractional molecules, are referred to as whole molecules. Interactions of fractional molecule(s) are scaled by the parameter λ. It ranges from 0, which means that the fractional molecule does not have any interaction with surrounding molecules, to 1, which means full (normal) interactions with the whole molecules. New Monte Carlo trial moves are introduced which facilitate the transfer of molecules in the system. The CFC method makes simulations more efficient because of higher acceptance probabilities for the transfer of molecules when λ is close to 0 and identity changes when λ is close to 1. It has been shown that the presence of a fractional molecule(s) has a negligible effect on the thermodynamic properties of the system.[13]

Ensembles and Partition Functions

In this section, we list the different ensembles that can be used in Brick-CFCMC (hereafter referred to as Brick). We also provide the partition functions for each of these ensembles. Here, we consider the partition functions for the expanded (CFC) ensembles with a single fractional molecule. Deriving the partition functions for ensembles with more than one fractional molecule is straightforward and left to the reader as an exercise.

NVT

The NVT ensemble has the following partition function:[14]

$$Q_{NVT} = \frac{V^N}{\Lambda^{3N} N!} \int \mathrm{d}\mathbf{s}^N \exp[-\beta U(\mathbf{s}^N)] \tag{S1}$$

N is the number of molecules, V is the volume of the system, Λ is the thermal wavelength, $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, and U is the potential energy.

CFC*NVT*

The CFC version of the NVT ensemble has the following partition function:[6]

$$Q^{\text{CFC}}_{NVT} = \frac{V^{N+1}}{\Lambda^{3(N+1)}N!}\int_0^1 \mathrm{d}\lambda \int \mathrm{d}\mathbf{s}^N \exp[-\beta U(\mathbf{s}^N)] \tag{S2}$$

$$\times \int \mathrm{d}\mathbf{s}_{\text{frac}} \exp[-\beta U_{\text{frac}}(\mathbf{s}_{\text{frac}}, \mathbf{s}^N, \lambda)] \tag{S3}$$

N is the number of whole molecules, V is the volume of the system, Λ is the thermal wavelength, $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, U is the potential energy (excluding the interactions of the fractional molecule), U_{frac} is the potential energy of the fractional molecule(s) (this term also includes the interactions between fractional molecules in case there is more than one fractional molecule in the system), $\lambda \in [0, 1]$ is the fractional parameter, and $\mathbf{s}_{\text{frac}}$ is the scaled coordinate vector of the fractional molecule.

NPT

The NPT ensemble has the following partition function:[14]

$$Q_{NPT} = \frac{\beta P}{\Lambda^{3N}N!}\int \mathrm{d}V V^N \exp[-\beta PV] \int \mathrm{d}\mathbf{s}^N \exp[-\beta U(\mathbf{s}^N)] \tag{S4}$$

N is the number of molecules, V is the volume of the system, P is the imposed pressure, Λ is the thermal wavelength, $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, and U is the potential energy.

CFC*NPT*

The CFC version of the NPT ensemble has the following partition function:[6]

$$Q_{NPT}^{\text{CFC}} = \frac{\beta P}{\Lambda^{3(N+1)} N!} \int_0^1 \mathrm{d}\lambda \int \mathrm{d}V V^{N+1} \exp[-\beta PV] \tag{S5}$$

$$\times \int \mathrm{d}\mathbf{s}^N \exp[-\beta U(\mathbf{s}^N)] \int \mathrm{d}\mathbf{s}_{\text{frac}} \exp[-\beta U_{\text{frac}}(\mathbf{s}_{\text{frac}}, \mathbf{s}^N, \lambda)] \tag{S6}$$

N is the number of whole molecules, V is the volume of the system, P is the imposed pressure, Λ is the thermal wavelength, $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, U is the potential energy (excluding the interactions of the fractional molecule), U_{frac} is the potential energy of the fractional molecule(s) (this term also includes the interactions between fractional molecules in case there is more than one fractional molecule in the system), $\lambda \in [0,1]$ is the fractional parameter, and $\mathbf{s}_{\text{frac}}$ is the scaled coordinate vector of the fractional molecule.

Gibbs Ensemble at constant total volume

The Gibbs Ensemble combined with the CFC method can be used to simulate phase equilibria.[7,12] There are two simulation boxes that can exchange molecules and volume. In this ensemble, the fractional molecule can be transferred between the two simulation boxes (denoted by $i = 1$ and $i = 2$). It has the following partition function when the total volume is constant:

$$Q^{\mathrm{CFC}}_{\mathrm{GE},NVT} = \frac{1}{\Lambda^{3(N_{\mathrm{T}}+1)} N_{\mathrm{T}}!} \sum_{i=1}^{2} \sum_{N_1=0}^{N_{\mathrm{T}}} \frac{N_{\mathrm{T}}!}{N_1!(N_{\mathrm{T}}-N_1)!} \int_0^1 \mathrm{d}\lambda \tag{S7}$$

$$\times \int_0^{V_{\mathrm{T}}} \mathrm{d}V_1 V_1^{N_1+\delta_{i1}} (V_{\mathrm{T}} - V_1)^{N_{\mathrm{T}}-N_1+\delta_{i2}} \tag{S8}$$

$$\times \int \mathrm{d}\mathbf{s}^{N_1} \exp[-\beta U_1(\mathbf{s}^{N_1})] \tag{S9}$$

$$\times \int \mathrm{d}\mathbf{s}^{N_{\mathrm{T}}-N_1} \exp[-\beta U_2(\mathbf{s}^{N_{\mathrm{T}}-N_1})] \tag{S10}$$

$$\times \Big(\delta_{i1} \int \mathrm{d}\mathbf{s}_{\mathrm{frac}} \exp[-\beta U_{\mathrm{frac},1}(\mathbf{s}_{\mathrm{frac}}, \mathbf{s}^{N_1}, \lambda)] \tag{S11}$$

$$+ \delta_{i2} \int \mathrm{d}\mathbf{s}_{\mathrm{frac}} \exp[-\beta U_{\mathrm{frac},2}(\mathbf{s}_{\mathrm{frac}}, \mathbf{s}^{N_{\mathrm{T}}-N_1}, \lambda)]\Big) \tag{S12}$$

N_j and V_j are the number of whole molecules and volume of simulation box j, N_{T} is the total number of whole molecules in the two simulation boxes, V_{T} is the total volume of the two simulation boxes, Λ is the thermal wavelength, $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, U is the potential energy, U_{frac} is the potential energy of the fractional molecule, $\lambda \in [0,1]$ is the fractional parameter, $\mathbf{s}_{\mathrm{frac}}$ is the scaled coordinate vector of the fractional molecule, and the δ-function is used to indicate if the fractional molecule is in simulation box i ($\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$)

Gibbs Ensemble at constant pressure

The Gibbs Ensemble combined with the CFC method can be used to simulate phase equilibria.[7] There are two simulation boxes that can exchange molecules and change volume. This ensemble can only be used for simulations of two or more components, otherwise the phase rule would be violated. Here, we consider the ensemble for a single component for simplicity. In this ensemble the fractional molecule can be transferred between the two simulation boxes. It has the following partition function at constant pressure:

$$Q^{\mathrm{CFC}}_{\mathrm{GE},NPT} = \frac{\beta P}{\Lambda^{3(N_T+1)} N_{\mathrm{T}}!} \sum_{i=1}^{2} \sum_{N_1=0}^{N_{\mathrm{T}}} \frac{N_{\mathrm{T}}!}{N_1!(N_{\mathrm{T}}-N_1)!} \int_0^1 \mathrm{d}\lambda \tag{S13}$$

$$\times \int \mathrm{d}V_1 V_1^{N_1+\delta_{i1}} \exp[-\beta P V_1] \tag{S14}$$

$$\times \int \mathrm{d}V_2 V_2^{N_{\mathrm{T}}-N_1+\delta_{i2}} \exp[-\beta P V_2] \tag{S15}$$

$$\times \int \mathrm{d}\mathbf{s}^{N_1} \exp[-\beta U_1(\mathbf{s}^{N_1})] \tag{S16}$$

$$\times \int \mathrm{d}\mathbf{s}^{N_{\mathrm{T}}-N_1} \exp[-\beta U_2(\mathbf{s}^{N_{\mathrm{T}}-N_1})] \tag{S17}$$

$$\times \Big(\delta_{i1} \int \mathrm{d}\mathbf{s}_{\mathrm{frac}} \exp[-\beta U_{\mathrm{frac},1}(\mathbf{s}_{\mathrm{frac}}, \mathbf{s}^{N_1}, \lambda)] \tag{S18}$$

$$+ \delta_{i2} \int \mathrm{d}\mathbf{s}_{\mathrm{frac}} \exp[-\beta U_{\mathrm{frac},2}(\mathbf{s}_{\mathrm{frac}}, \mathbf{s}^{N_{\mathrm{T}}-N_1}, \lambda)]\Big) \tag{S19}$$

N_j and V_j are the number of whole molecules and volume of simulation box j, N_{T} is the total number of whole molecules in the two simulation boxes, P is the imposed pressure, Λ is the thermal wavelength, $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, U is the potential energy, U_{frac} is the potential energy of the fractional molecule, $\lambda \in [0,1]$ is the fractional parameter, $\mathbf{s}_{\mathrm{frac}}$ is the scaled coordinate vector of the fractional molecule, and the δ-function is used to indicate if the fractional molecule is in simulation box i ($\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$)

Reaction Ensemble at constant volume

The Reaction Ensemble combined with the CFC method can be used to simulate reaction equilibria.[8] Fractional molecules are either reactants or products. The chemical reaction is mimicked as additional Monte Carlo trial moves that transform reactants into products and vice versa. It has the following partition function at constant volume:

$$Q^{\mathrm{CFC}}_{\mathrm{RE,V}} = \sum_{N_1=0}^{\infty} \cdots \sum_{N_{\mathrm{S}}=0}^{\infty} \sum_{\delta=0}^{1} \int_0^1 \mathrm{d}\lambda \tag{S20}$$

$$\times \exp\left[\sum_{i=1}^{\mathrm{R}} \beta\mu_i \left(N_i + \delta\nu_i\right) + \left(N_i + \delta\nu_i\right) \ln\left(\frac{Vq_i}{\Lambda_i^3}\right) - \ln N_i!\right] \tag{S21}$$

$$\times \exp\left[\sum_{i=R+1}^{\mathrm{S}} \beta\mu_i \left(N_i + \left(1-\delta\right)\nu_i\right) + \left(N_i + \left(1-\delta\right)\nu_i\right) \ln\left(\frac{Vq_i}{\Lambda_i^3}\right) - \ln N_i!\right] \tag{S22}$$

$$\times \int \mathrm{d}\mathbf{s}^N \exp[-\beta U(\mathbf{s}^N)] \int \mathrm{d}\mathbf{s}_{\mathrm{frac}}^{N_{\mathrm{frac}}} \exp[-\beta U(\mathbf{s}_{\mathrm{frac}}^{N_{\mathrm{frac}}}, \mathbf{s}^N, \lambda, \delta)] \tag{S23}$$

N_i is the number of whole molecules of component i, Λ_i is the thermal wavelength, q_i is the partition function of the isolated molecule excluding the translational part (see the Appendix of this document for details), μ_i is the chemical potential, ν_i is the stoichiometric coefficient, δ indicates the reaction step ($\delta = 0$ indicates that the fractional molecules are reactants and $\delta = 1$ indicates that the fractional molecules are products), $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, U is the potential energy (excluding the interactions of fractional molecules), U_{frac} is the potential energy of the fractional molecule (including interactions between fractional molecules), λ is the fractional parameter, and $\mathbf{s}_{\mathrm{frac}}$ is the scaled coordinate vector of the fractional molecule. The components are labeled such that components 1 to R are reactants, and components $R+1$ to S are products. Multiple reactions can be defined in the Reaction Ensemble.

Reaction Ensemble at constant pressure

The Reaction Ensemble combined with the CFC method can be used to simulate reaction equilibria.[8] Fractional molecules are either reactants or products. The chemical reaction is mimicked as additional Monte Carlo trial moves that transform reactants into products and vice versa. It has the following partition function at constant pressure:

$$Q_{\mathrm{RE,P}}^{\mathrm{CFC}} = \beta P \sum_{N_1=0}^{\infty} \cdots \sum_{N_{\mathrm{S}}=0}^{\infty} \sum_{\delta=0}^{1} \int_0^1 \mathrm{d}\lambda \int \mathrm{d}V \exp[-\beta PV] \tag{S24}$$

$$\times \exp\left[\sum_{i=1}^{\mathrm{R}} \beta\mu_i \left(N_i + \delta\nu_i\right) + \left(N_i + \delta\nu_i\right) \ln\left(\frac{Vq_i}{\Lambda_i^3}\right) - \ln N_i!\right] \tag{S25}$$

$$\times \exp\left[\sum_{i=R+1}^{\mathrm{S}} \beta\mu_i \left(N_i + (1-\delta)\,\nu_i\right) + \left(N_i + (1-\delta)\,\nu_i\right) \ln\left(\frac{Vq_i}{\Lambda_i^3}\right) - \ln N_i!\right] \tag{S26}$$

$$\times \int \mathrm{d}\mathbf{s}^N \exp[-\beta U(\mathbf{s}^N)] \int \mathrm{d}\mathbf{s}_{\mathrm{frac}}^{N_{\mathrm{frac}}} \exp[-\beta U(\mathbf{s}_{\mathrm{frac}}^{N_{\mathrm{frac}}}, \mathbf{s}^N, \lambda, \delta)] \tag{S27}$$

N_i is the number of whole molecules of component i, Λ_i is the thermal wavelength, q_i is the partition function of the isolated molecule excluding the translational part (see the Appendix of this document for details), μ_i is the chemical potential, ν_i is the stoichiometric coefficient, δ indicates the reaction step ($\delta = 0$ indicates that the fractional molecules are reactants and $\delta = 1$ indicates the fractional molecules are products), P is the imposed pressure, $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, U is the potential energy (excluding the interactions of fractional molecules), U_{frac} is the potential energy of the fractional molecule (including interactions between fractional molecules), λ is the fractional parameter, and $\mathbf{s}_{\mathrm{frac}}$ is the scaled coordinate vector of the fractional molecule. The components are labeled such that components 1 to R are reactants, and components $R+1$ to S are products. Multiple reactions can be defined in the Reaction Ensemble.

Grand-Canonical Ensemble

The Grand-Canonical Ensemble combined with the CFC method can be used to study absorption processes in liquids.[9] There is one simulation box that can exchange molecules with an ideal gas reservoir. It has the following partition function at constant volume:

$$Q_{\mathrm{GC}}^{\mathrm{CFC}} = \sum_{N=0}^{\infty} \frac{\exp\left[\beta\mu N\right] V^{N+1}}{\Lambda^{3(N+1)} N!} \int_0^1 \mathrm{d}\lambda \int \mathrm{d}\mathbf{s}^N \exp[-\beta U(\mathbf{s}^N)] \tag{S28}$$

$$\times \int \mathrm{d}\mathbf{s}_{\mathrm{frac}} \exp[-\beta U(\mathbf{s}_{\mathrm{frac}}, \mathbf{s}^N, \lambda)] \tag{S29}$$

N is the number of whole molecules, V is the volume of the system, Λ is the thermal wavelength, μ is the imposed chemical potential, $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, U is the potential energy (excluding the interactions of the fractional molecules), U_{frac} is the potential energy of the fractional molecule (including interactions between fractional molecules), $\lambda \in [0, 1]$ is the fractional parameter, and $\mathbf{s}_{\mathrm{frac}}$ is the scaled coordinate vector of the fractional molecule.

Osmotic Ensemble

The Osmotic Ensemble combined with the CFC method is a variant of the Grand-Canonical Ensemble with multiple components.[9] It imposes a constant pressure and fixes one extensive parameter. This ensemble has the following partition function:

$$Q_{\mathrm{Osmotic}}^{\mathrm{CFC}} = \sum_{N=0}^{\infty} \frac{\exp\left[\beta\mu N\right]}{\Lambda^{3(N+1)} N!} \int_0^1 \mathrm{d}\lambda \int \mathrm{d}V V^{N+1} \exp[-\beta P V] \tag{S30}$$

$$\times \int \mathrm{d}\mathbf{s}^N \exp[-\beta U(\mathbf{s}^N)] \int \mathrm{d}\mathbf{s}_{\mathrm{frac}} \exp[-\beta U(\mathbf{s}_{\mathrm{frac}}, \mathbf{s}^N, \lambda)] \tag{S31}$$

N is the number of whole molecules, V is the volume of the system, P is the imposed pressure, Λ is the thermal wavelength, μ is the imposed chemical potential, $\mathbf{s}^N$ is the scaled coordinate vector of all molecules, U is the potential energy (excluding the interactions of the fractional molecules), U_{frac} is the potential energy of the fractional molecule (including interactions between fractional molecules), $\lambda \in [0, 1]$ is the fractional parameter, and $\mathbf{s}_{\mathrm{frac}}$ is the scaled coordinate vector of the fractional molecule.

Fractional Molecules

All expanded ensembles make use of fractional molecules. In Brick, we can distinguish between four different types:

- NVT/NPT: these fractional molecules are used as probes to obtain partial molar properties.[6,11] Furthermore, these fractional molecules also enhance the sampling of other properties in dense systems.
- Gibbs Ensemble: these fractional molecules are used to facilitate the exchange of molecules between two simulation boxes to obtain phase equilibrium.[7]
- Reaction Ensemble: these fractional molecules are used to facilitate the transformation of reactants into products (and vice versa) to obtain reaction equilibrium.[8]
- Grand-Canonical Ensemble: these fractional molecules are used to facilitate the exchange of molecules between the simulation box and reservoir.[9] This also applies to the Osmotic Ensemble.

The fractional molecules from the NVT/NPT ensembles (probes) can be used in any other ensemble. The fractional molecules from the Gibbs Ensemble and Reaction Ensemble can also be combined. However, the fractional molecules from the Grand-Canonical Ensemble can not be combined with the Gibbs Ensemble at constant pressure or the Reaction Ensemble at constant pressure unless at least one extensive parameter is fixed.

The Fractional Parameter

The fractional parameter λ is used to scale interactions of the fractional molecules in the system. When $\lambda = 0$, the fractional molecule has no interactions with its surroundings.

For $\lambda = 1$, the interactions of the fractional molecule are the same as a whole molecule. Between these two state points, a path can be chosen how to "switch on" the interactions from $\lambda = 0$ to $\lambda = 1$. The LJ interactions and electrostatic interactions are scaled differently, using λ_{LJ} and λ_{EL}, respectively. The following procedure is used in Brick. Starting from $\lambda = 0$, with increasing λ we only increase the Lennard-Jones interactions of the fractional molecule i.e. increase λ_{LJ} and keep $\lambda_{\mathrm{EL}} = 0$. Then, at a certain chosen value of λ, which we refer to as λ_{s}, we reach $\lambda_{\mathrm{LJ}} = 1$. From that point (till $\lambda = 1$) we start increasing the electrostatic interactions of the fractional molecule i.e. increase λ_{EL} and keep $\lambda_{\mathrm{LJ}} = 1$. This way we gradually increase first the LJ interactions and then the electrostatic interactions which is more efficient in simulations than scaling both interactions at the same time.[15] Intramolecular interactions can be changed in the same way, but only for fractional molecules in the NVT/NPT ensembles and Gibbs Ensemble, and not in the Reaction Ensemble and Grand-Canonical/Osmotic Ensemble. Simulations of multicomponent systems often require multiple fractional molecules, each with a fractional parameter. This allows the calculation of chemical potentials and partial molar properties of each component in the mixture. It is also possible to define a fractional group consisting of several molecules. For example, in ionic systems, to maintain charge neutrality, a fractional group consists of an equal number of anions and cations. The strength of the interactions of the each fractional molecule in the fractional group is regulated by a single fractional parameter λ. Similarly, in the Reaction Ensemble, a fractional group is defined for the reactants or the products. All molecules in the fractional group share the same fractional parameter. For more details the reader is referred to the manual of Brick.[16]

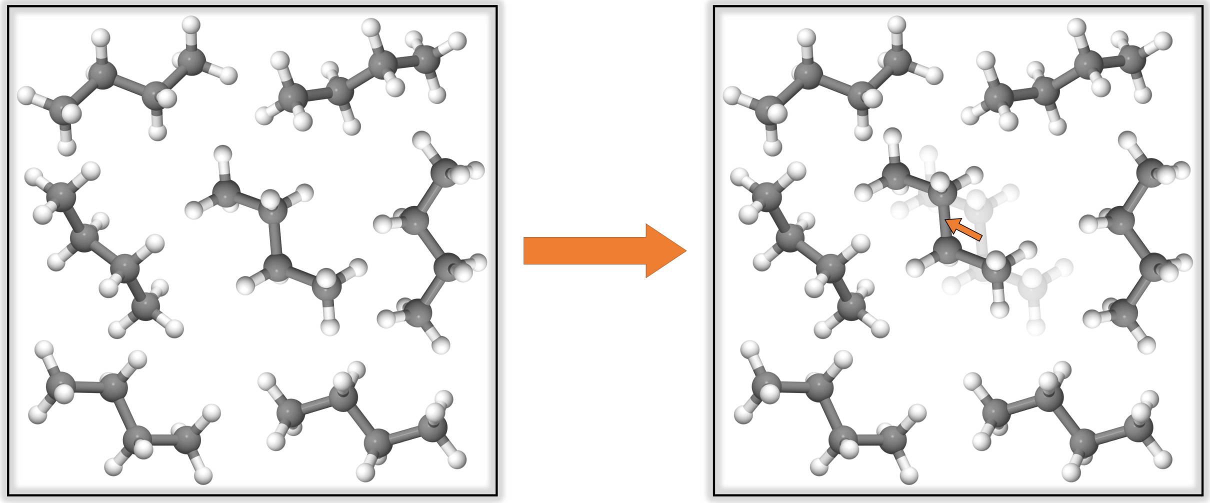

Figure S1: Translation Trial Move: a randomly selected molecule is displaced by a random distance in a random direction.

Monte Carlo Trial Moves and Acceptance Rules

In this section, we introduce the Monte Carlo trial moves and acceptance rules for the different ensembles in Brick. During a simulation, one trial move is randomly selected from the available trial moves at each MC step. For a more detailed derivation and description of the trial moves, we refer the reader to the manual of Brick.[16]

Translation Trial Moves

In a translation trial move, one molecule is selected at random and moved in a random direction by a random distance (Figure S1). The rest of the system remains unchanged. The acceptance rule for this trial move is:

$$\mathrm{acc}(\mathrm{o} \to \mathrm{n}) = \min(1, \exp[-\beta \Delta U]) \tag{S32}$$

where ΔU is the change in the intermolecular energy.

Pair and Cluster Translations

In some systems, molecules are strongly bonded to each other, and in this case, it can be more efficient to displace two or more molecules in one trial move. For this reason, in Brick, it is possible to use pair translation trial moves[17] and cluster translation trial moves. In the pair translation trial move, two molecules are displaced in a random direction by a random distance. In the cluster translation trial move a cluster of molecules is displaced in a random direction. For more details about these trial moves we refer the reader to the manual of Brick.[16]

Smart Translations

The efficiency of the translation trial moves can be increased using the information on the forces on the molecules in the system. Instead of generating random displacements, one can choose to generate the displacement based on the forces acting on the molecules.[18–20] Therefore, in Brick, there is the possibility to use so-called smart translation trial moves. This trial move displaces each molecule based on the force acting on them. For more details about this trial move we refer the reader to the manual of Brick.[16]

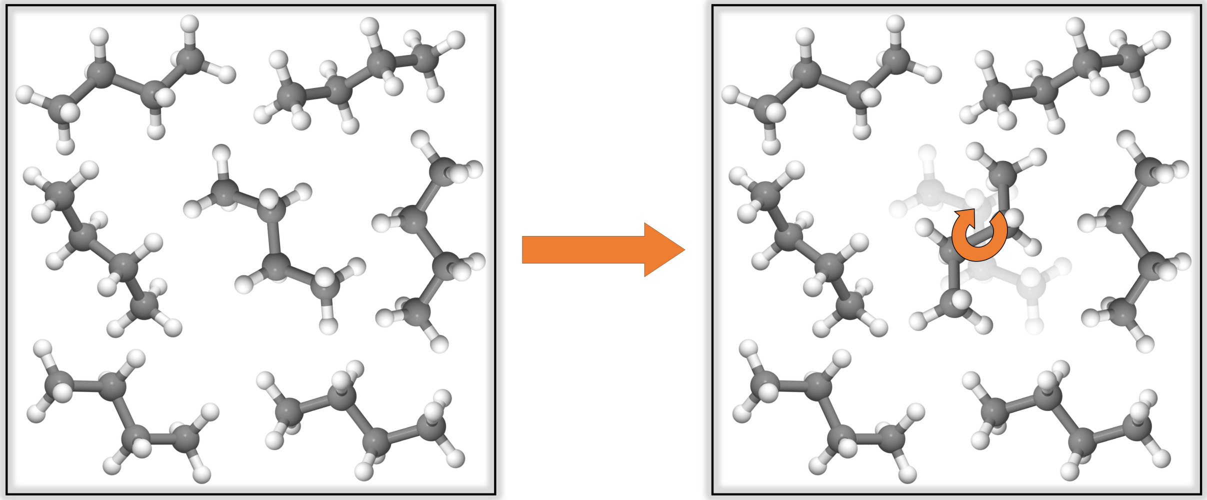

Figure S2: Rotation Trial Move: a randomly selected molecule is rotated by a random angle around a random axis.

Rotation Trial Moves

In a rotation trial move, a randomly selected molecule is rotated around a random axis by a random angle (Figure S2). The rest of the system remains unchanged. The acceptance rule for this trial move is:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min(1, \exp[-\beta \Delta U]) \tag{S33}$$

where ΔU is the change in the intermolecular energy.

Pair and Cluster Rotations

In some systems, molecules are strongly bonded to each other, and in this case, it can be more efficient to rotate two or more molecules in one trial move. For this reason, in Brick, it is possible to use pair rotation trial moves[17] and cluster rotation trial moves. In the pair rotation trial move, two molecules are rotated along one common axis (going through the geometric center) by a random angle. In the cluster rotation trial move a cluster of molecules is rotated along one common axis (going through the geometric center) by a random angle.

For more details about these trial moves we refer the reader to the manual of Brick.[16]

Smart Rotations

The efficiency of the rotation trial moves can be increased using the torques on the molecules in the system. Instead of generating random angles for rotation one can choose to generate the angles based on the torques acting on the molecules.[18–20] Therefore, in Brick, there is the possibility to use so-called smart rotation trial moves. This trial move rotates each molecule based on the torque acting on them. For more details about this trial move we refer the reader to the manual of Brick.[16]

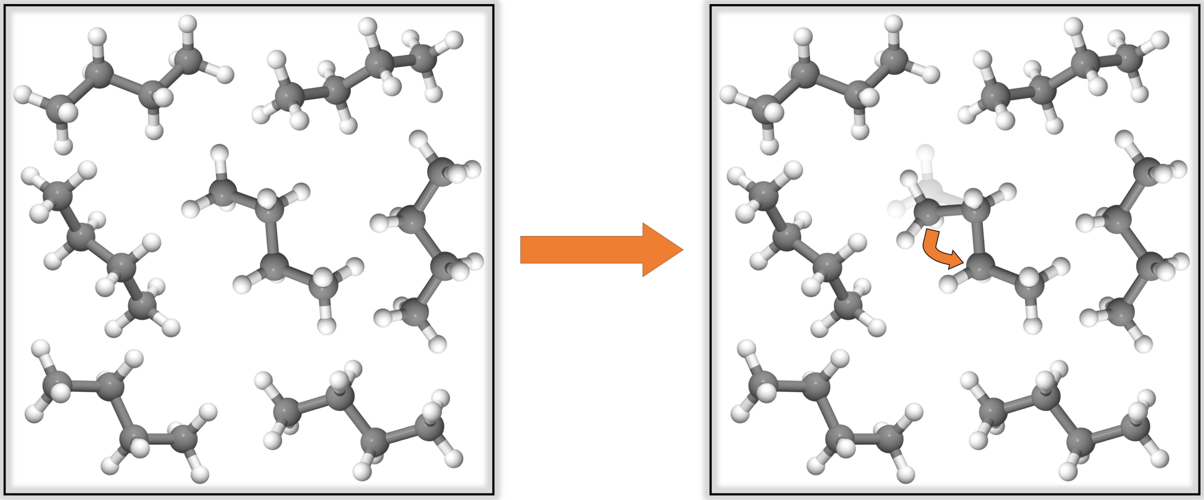

Figure S3: Bending Trial Move: a randomly selected bond angle is randomly changed in the same plane in a randomly selected molecule.

Bending Trial Moves

In a bending trial move, in a randomly selected molecule the bond angle between randomly selected atoms is changed by a random value, while keeping the atoms of the bending angle in the same plane (Figure S3). The rest of the system remains unchanged. The acceptance rule for this trial move is:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min\left(1, \frac{\sin\theta_\mathrm{n}}{\sin\theta_\mathrm{o}} \exp[-\beta\Delta U]\right) \tag{S34}$$

where θ_o and θ_n are the old and new bond angles, respectively, and ΔU is the change in the intermolecular and intramolecular energy.

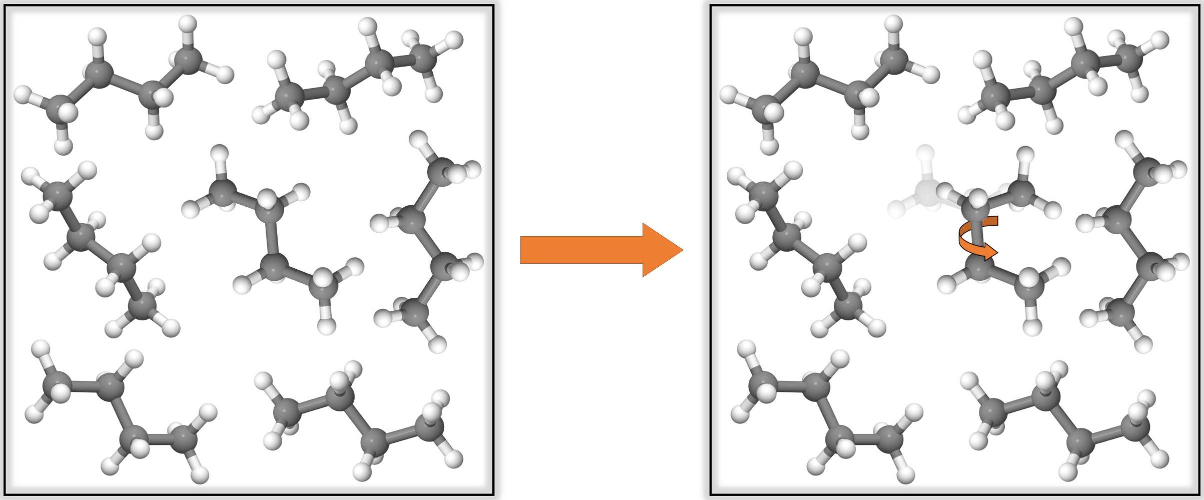

Figure S4: Torsion Trial Move: a molecule is selected at random and a randomly selected torsional angle in this molecule is changed.

Torsion Trial Moves

In a torsion trial move, in a randomly selected molecule a randomly selected torsional angle is changed by a random value (Figure S4). The rest of the system remains unchanged. The acceptance rule for this trial move is:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min(1, \exp[-\beta \Delta U]) \tag{S35}$$

where ΔU is the change in the intermolecular and intramolecular energy.

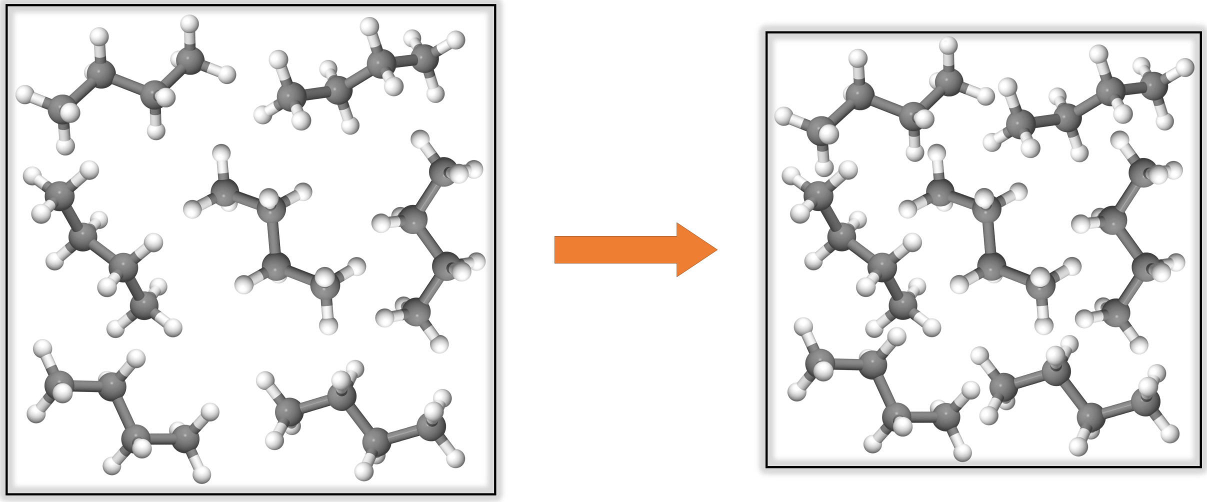

Figure S5: Volume Change Trial Move: the volume of the simulation box is changed by a random value.

Volume Change Trial Moves

In a volume change trial move, the volume of the simulation box is changed by a random value $V_\mathrm{n} = V_\mathrm{o} + \Delta V$ (Figure S5). The centers of mass of the molecules are scaled according to this volume change but the intramolecular distances for each molecule remain unchanged. The acceptance rule for this trial move is:

$$\mathrm{acc}(\mathrm{o} \to \mathrm{n}) = \min\left(1, \exp[-\beta P(V_\mathrm{n} - V_\mathrm{o}) - \beta \Delta U + N \ln\left(\frac{V_\mathrm{n}}{V_\mathrm{o}}\right)]\right) \tag{S36}$$

where P is the imposed pressure, ΔU is the change in the intermolecular energy only (since changing the volume does not change the intramolecular interactions), and N is the total number of molecules in the system (including the fractional molecules).

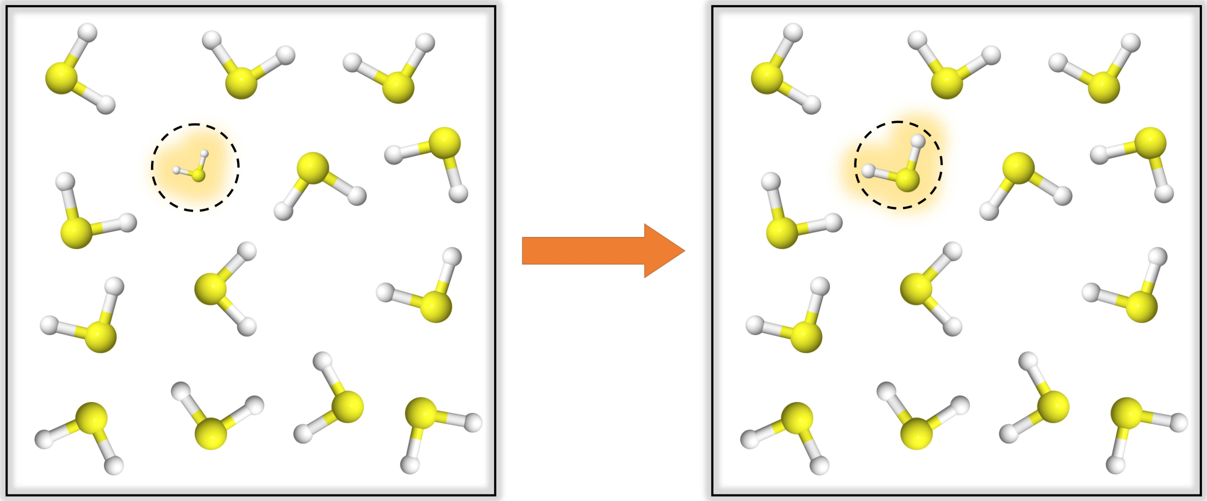

Figure S6: Lambda Trial Move: the value of the fractional parameter λ is changed by a random value which changes the interaction strength of the fractional molecule with its surroundings. The new value of λ should always be in the interval $[0, 1]$.

NVT/NPT Lambda Trial Moves

In this trial move, the value of the fractional parameter λ is changed by a random value (Figure S6). We have the following acceptance rule:

$$\mathrm{acc}(\mathrm{o} \to \mathrm{n}) = \min(1, \exp[-\beta \Delta U_{\mathrm{frac}}]) \tag{S37}$$

where ΔU_{frac} is the change in energy of the fractional molecule (including the intramolecular energy if those interactions are scaled). Note that $\lambda \in [0, 1]$. If λ is outside this interval the trial move is rejected immediately.

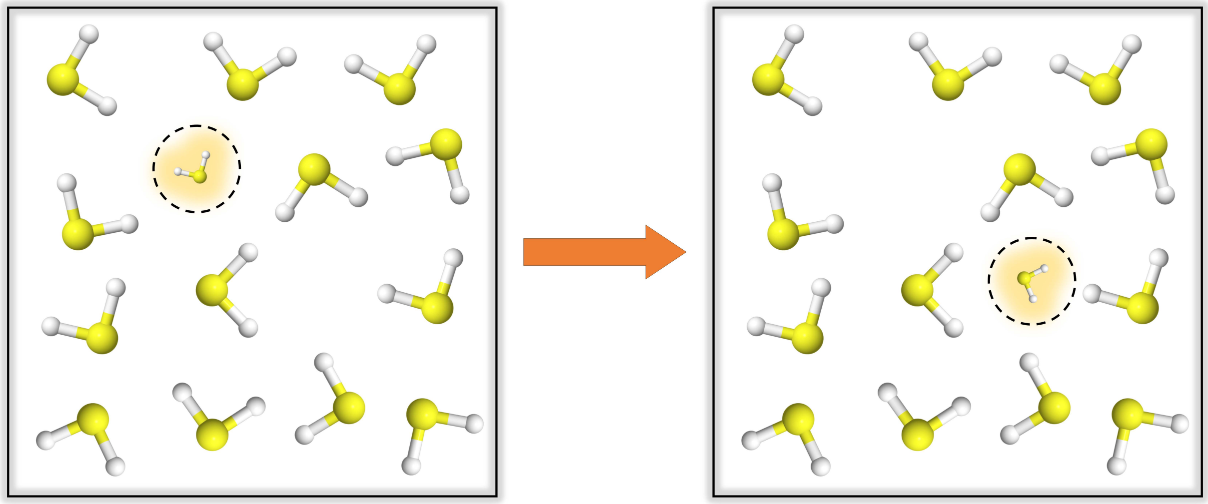

Figure S7: Reinsertion Trial Move: the fractional molecule is removed from the system and reinserted at a random position in the simulation box with a random orientation. The internal configuration does not change.

NVT/NPT Reinsertion Trial Moves

In this trial move, the fractional molecule is reinserted at a random position with a random orientation in the simulation box (Figure S7). We have the following acceptance rule:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min(1, \exp[-\beta \Delta U_{\mathrm{frac}}]) \tag{S38}$$

where ΔU_{frac} is the change in intermolecular energy of the fractional molecule. This trial move has a high acceptance probability for λ close to 0.

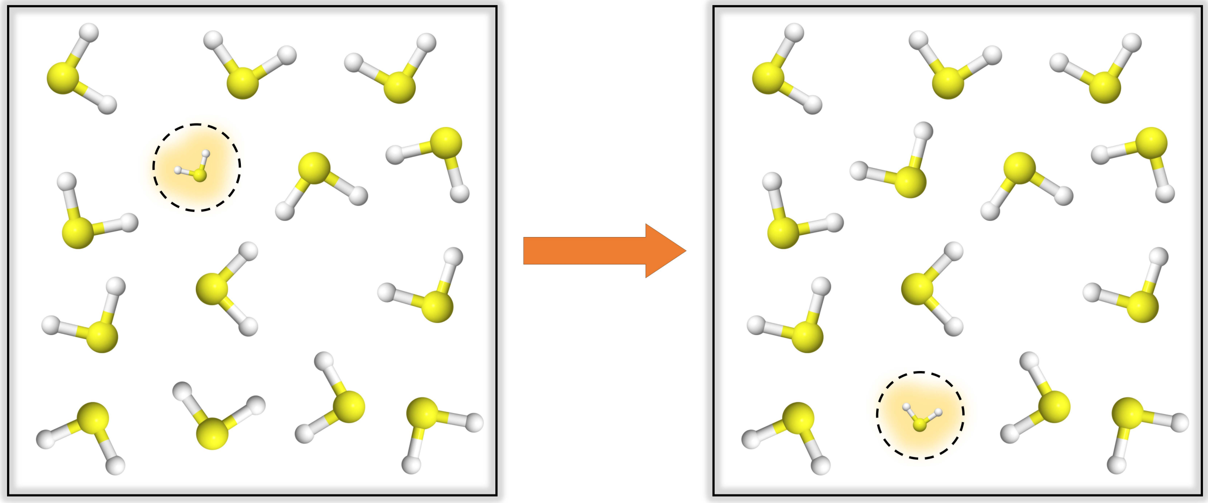

Figure S8: Identity Change Trial Move: the fractional molecule is transformed into a whole molecule and a randomly selected whole molecule is transformed into a fractional molecule with the same value of λ.

NVT/NPT Identity Change Trial Moves

In this trial move, the fractional molecule is transformed into a whole molecule while at the same time a randomly selected whole molecule of the same component is transformed into a fractional molecule with the same fractional parameter λ (Figure S8). We have the following acceptance rule:

$$\mathrm{acc}(\mathrm{o} \to \mathrm{n}) = \min(1, \exp[-\beta \Delta U_{\mathrm{frac}}]) \tag{S39}$$

where ΔU_{frac} is the change in energy of the fractional molecule and whole molecule (including the intramolecular energy if this energy is scaled). This trial move has a high acceptance probability for λ close to 1.

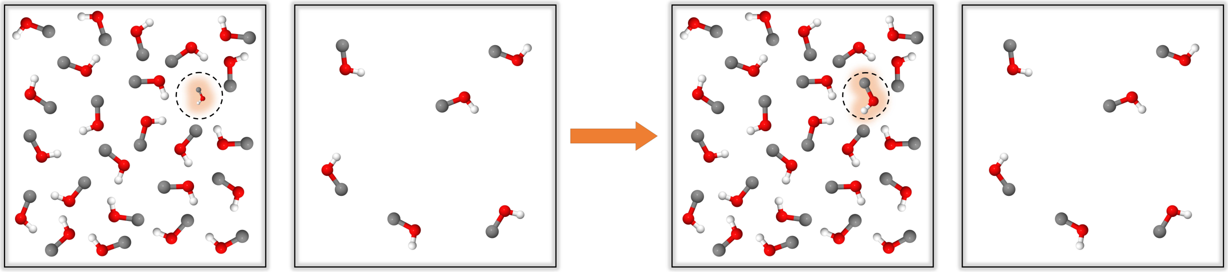

Figure S9: Gibbs Ensemble Lambda Trial Move: the value of the fractional parameter λ is changed by a random value which increases or decreases the interaction strength of the fractional molecule with its surroundings. The new value of λ should always be in the interval $[0, 1]$.

Gibbs Ensemble Lambda Trial Moves

In this trial move, the value of the fractional parameter λ is changed by a random value (Figure S9). We have the following acceptance rule:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min(1, \exp[-\beta \Delta U_{\mathrm{frac}}]) \tag{S40}$$

where ΔU_{frac} is the change in energy of the fractional molecule (including the intramolecular energy if those interactions are scaled). Note that $\lambda \in [0, 1]$. If λ is outside this interval the trial move is rejected immediately.

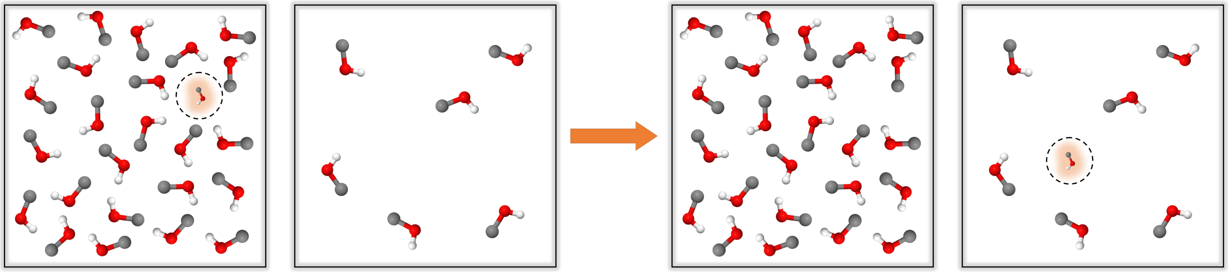

Figure S10: Gibbs Ensemble Swap Trial Move: the fractional molecule is transferred from one simulation box to the other. The position and orientation in the new simulation box are generated at random and λ remains unchanged.

Gibbs Ensemble Swap Trial Moves

In this trial move, the fractional molecule is removed from one simulation box and reinserted at a random position with a random orientation in the other simulation box leaving λ unchanged (Figure S10). Transferring the fractional molecule from simulation box i to simulation box $j \neq i$ results in the following acceptance rule:

$$\text{acc}(\text{o} \rightarrow \text{n}) = \min\left(1, \frac{V_j}{V_i} \exp[-\beta \Delta U]\right) \tag{S41}$$

where ΔU is the total energy change in both simulation boxes (including the intramolecular energy if those interactions are scaled), and V_i is the volume of simulation box i. This trial move has a high acceptance probability for λ close to 0.

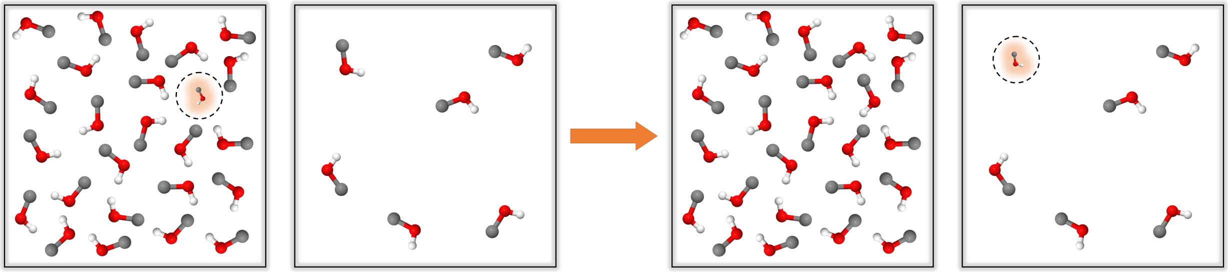

Figure S11: Gibbs Ensemble Identity Change Trial Move: the fractional molecule is transformed into a whole molecule and a randomly selected molecule in the other simulation box is transformed into a fractional molecule. The value of the fractional parameter remains unchanged.

Gibbs Ensemble Identity Change Trial Moves

In this trial move, the fractional molecule is transformed into a whole molecule while at the same time a randomly selected molecule in the other simulation box is transformed into a fractional molecule (Figure S11). Changing the fractional molecule in simulation box i into a whole molecule and changing a whole molecule in simulation box $j \neq i$ into a fractional molecule results in the following acceptance rule:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min\left(1, \frac{N_j}{N_i + 1} \exp[-\beta \Delta U]\right) \tag{S42}$$

where ΔU is the total energy change in both simulation boxes (including the intramolecular energy if this energy is scaled), and N_i is the number of whole molecules in box i. This trial move has a high acceptance probability for λ close to 1.

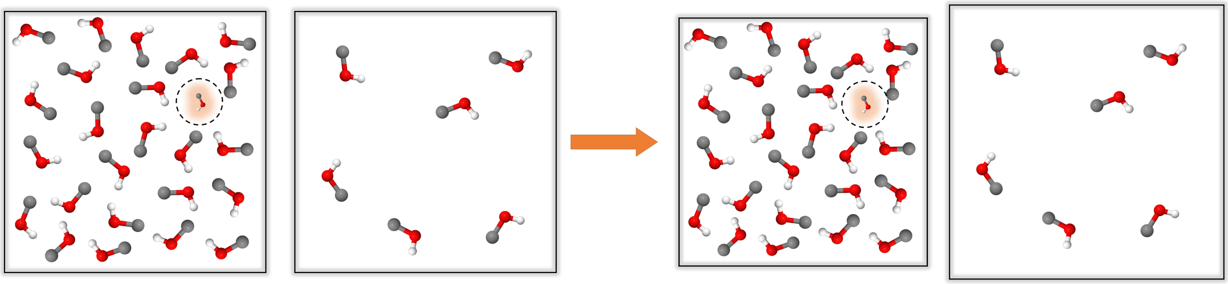

Figure S12: Gibbs Volume Change Trial Move: the volume of the simulation box(es) is randomly changed. For the Gibbs Ensemble at constant volume the total volume of the two simulation boxes is conserved. For the Gibbs Ensemble at constant pressure one simulation box is selected at random of which the volume is changed.

Gibbs Ensemble Volume Change Trial Moves

For a volume change (Figure S12), where the total volume of the two simulation boxes is constant we have $V_{1,\mathrm{n}} = V_{1,\mathrm{o}} \pm \Delta V$ and $V_{2,\mathrm{n}} = V_{2,\mathrm{o}} \mp \Delta V$, where o and n denote the old and new configuration. In each simulation box, the centers of mass of the molecules are scaled according to the volume change but the intramolecular distances for each molecule remain unchanged. The acceptance rule for this trial move is:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min\left(1, \frac{V_{1,\mathrm{n}}^{N_1+\delta_{i1}} V_{2,\mathrm{n}}^{N_2+\delta_{i2}}}{V_{1,\mathrm{o}}^{N_1+\delta_{i1}} V_{2,\mathrm{o}}^{N_2+\delta_{i2}}} \exp[-\beta \Delta U]\right) \tag{S43}$$

where i indicates the simulation box in which the fractional molecule is and ΔU is the change in the intermolecular energy only, since changing the volume does not change the intramolecular interactions. $N_{j,\mathrm{o/n}}$ is the total number of whole molecules in simulation box j, i indicates in which box the fractional molecule is, such that δ_{ij} indicates if the fractional molecule is in box j.

If the pressure is constant, the volume of both simulation boxes can be changed independently of each other and the acceptance rule reduces to that of the one for the NPT ensemble:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min\left(1, \exp[-\beta P(V_{i,\mathrm{n}} - V_{i,\mathrm{o}}) - \beta \Delta U + (N + \delta_{ij}) \ln\left(\frac{V_{i,\mathrm{n}}}{V_{i,\mathrm{o}}}\right)]\right) \tag{S44}$$

where P is the imposed pressure, i indicates the simulation box of which the volume is changed, $V_{i,\mathrm{o/n}}$ is the total number of whole molecules in simulation box i (o and n denote the old and new configuration), j indicates in which simulation box the fractional molecule is, and ΔU is the change in the intermolecular energy only since changing the volume does not change the intramolecular interactions. The δ-function is used to indicate if the fractional molecule is in simulation box i ($\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$)

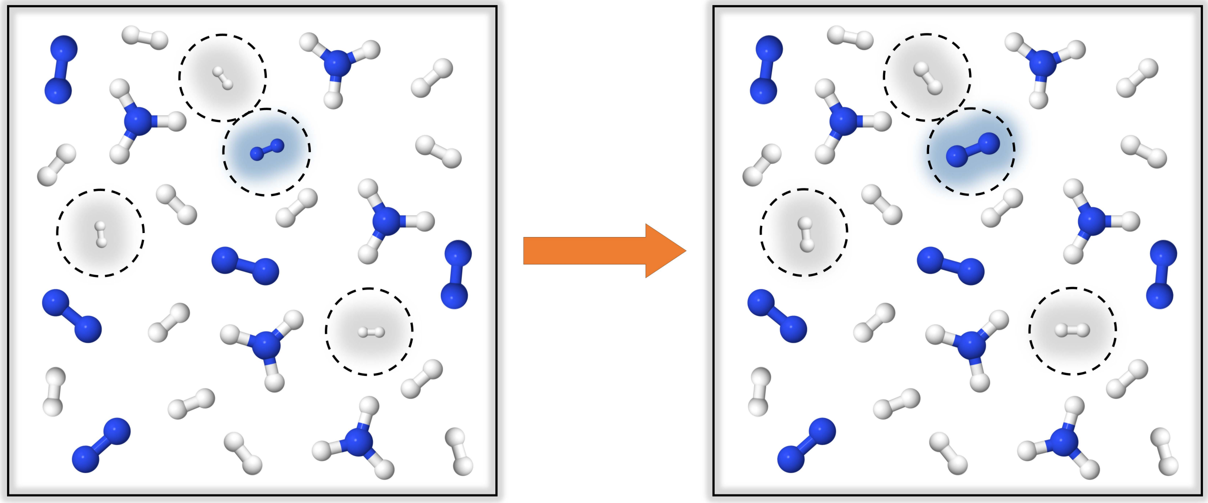

Figure S13: Reaction Lambda Trial Move: the value of the fractional parameter λ is changed by a random value which increases or decreases the interaction strength of the fractional molecule(s) with its surroundings. The new value of λ should always be in the interval $[0, 1]$.

Reaction Ensemble Lambda Trial Moves

In this trial move the value of the fractional parameter λ is changed by a random value (Figure S13). We have the following acceptance rule:

$$\mathrm{acc(o \to n)} = \min(1, \exp[-\beta \Delta U_{\mathrm{frac}}]) \tag{S45}$$

where ΔU_{frac} is the change in intermolecular energy of the fractional molecule(s) (intramolecular interactions can not be scaled in the Reaction Ensemble). Note that $\lambda \in [0, 1]$. If λ is outside this interval the trial move is rejected immediately.

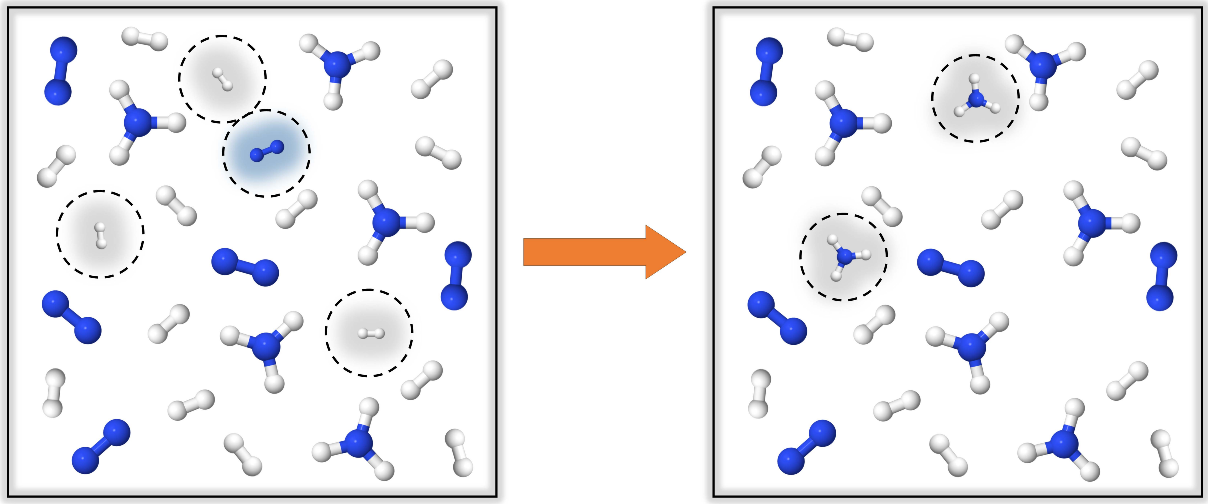

Figure S14: Reaction Swap Trial Move: fractional molecules of reactants/products are removed and fractional molecules of the products/reactants are inserted at random positions with random orientations. The fractional parameter λ remains unchanged.

Reaction Ensemble Swap Trial Moves

In this trial move, the fractional molecules of the reactants/products are removed from the simulation box and fractional molecules of the products/reactants are inserted at random positions, with a random orientation and a random internal configuration (Figure S14). For non-rigid molecules, the configurations of the molecules are generated in a short, separate, MC simulation of the isolated molecule where only bending and torsion trial moves are used i.e. the molecules are taken from an ideal gas reservoir. This ensures that their internal structure is equilibrated. Changing from fractional reactants to fractional products results in the following acceptance rule:

$$\text{acc}(\text{o} \rightarrow \text{n}) = \min\left(1, \prod_{i=1}^{\text{R}} \left(\frac{V q_i}{\Lambda_i^3}\right)^{-\nu_i} \prod_{i=R+1}^{\text{S}} \left(\frac{V q_i}{\Lambda_i^3}\right)^{\nu_i} \exp[-\beta \Delta U]\right) \tag{S46}$$

where ΔU is the total intermolecular energy change (intramolecular interactions can not be scaled in the Reaction Ensemble), V is the volume of the simulation box, Λ_i is the thermal wavelength of component i, q_i is the partition function of the isolated molecule excluding the translational part, and ν_i is the stoichiometric coefficient. The components are labeled such that components 1 to R are reactants, and components $R+1$ to S are products. This trial move has a high acceptance probability for λ close to 0.

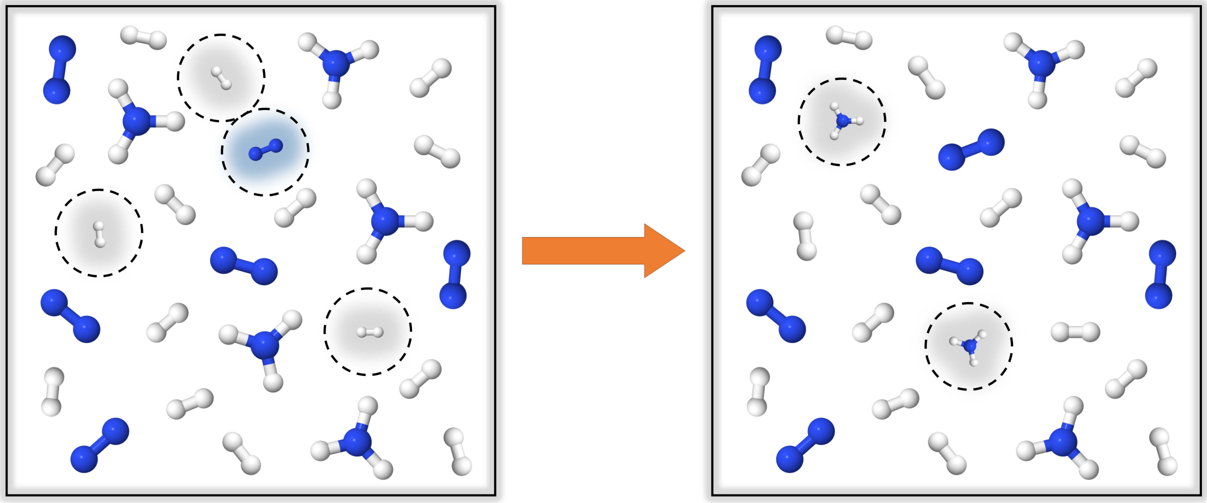

Figure S15: Reaction Identity Change Trial Move: fractional molecules of the reactants/products are transformed into whole molecules while randomly selected whole products/reactants are transformed into fractional molecules. The fractional parameter λ remains unchanged.

Reaction Ensemble Identity Change Trial Moves

In this trial move fractional molecules of the reactants/products are transformed into whole molecules and randomly selected molecules from the products/reactants are transformed into fractional molecules with the same fractional parameter (Figure S15). Changing fractional reactants into whole molecules and whole products molecules into fractional products results in the following acceptance rule:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min\left(1, \prod_{i=1}^{\mathrm{R}} \frac{N_i!}{(N_i+\nu_i)!} \prod_{i=R+1}^{\mathrm{S}} \frac{N_i}{(N_i-\nu_i)!} \exp[-\beta \Delta U]\right) \tag{S47}$$

Where ΔU is the total intermolecular energy change (intramolecular interactions can not be scaled in the Reaction Ensemble), N_i is the number of whole molecules of component i, and ν_i is the stoichiometric coefficient. The components are labeled such that components 1 to R are reactants, and components $R+1$ to S are products. For the reverse case, fractional products to fractional reactants, the only difference is the change in signs of the stoichiometric coefficients ν_i. For the reverse case, changing fractional products into whole molecules and whole reactant molecules into fractional reactants, the only difference is the change in signs of the stoichiometric coefficients ν_i. This trial move has a high acceptance probability for λ close to 1.

Grand-Canonical/Osmotic Ensemble Lambda Trial Moves

In the Grand-Canonical and Osmotic Ensemble,[9] the value of the fractional parameter is randomly changed. There are three cases to consider:

I. $\lambda < 0$. The fractional molecule is removed from the system and a randomly selected molecule is transformed into a fractional molecule with $\lambda \rightarrow \lambda + 1$ (Figure S16). The acceptance rule is:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min\left(1, \frac{N}{f\beta V}\exp[-\beta\Delta U]\right) \tag{S48}$$

II. $0 < \lambda < 1$. This trial move acts like a normal change in λ (Figure S17). The acceptance rule is:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min\left(1, \exp[-\beta\Delta U]\right) \tag{S49}$$

III. $1 < \lambda$. A new fractional molecule is added to the system at a random position, with a random orientation, and random internal configuration. The configuration of the molecule is generated in a short, separate, MC simulation of the isolated molecule where only bending and torsion trial moves are used i.e. the molecules are taken from a ideal gas reservoir. The new fractional parameter of this fractional molecule is $\lambda \rightarrow \lambda - 1$ (Figure S18). The acceptance rule is:

$$\mathrm{acc}(\mathrm{o} \rightarrow \mathrm{n}) = \min\left(1, \frac{f\beta V}{N+1}\exp[-\beta\Delta U]\right) \tag{S50}$$

where ΔU is the change in intermolecular energy (intramolecular energy can not be scaled in the Grand-Canonical/Osmotic Ensemble), N is the number of whole molecules, V is the volume of the simulation box and f is the fugacity which is related to the excess chemical potential: $f = \frac{N}{\beta V} \exp[\mu_{\mathrm{excess}}/k_{\mathrm{B}}T]$.[8] The fugacity is used as input in Brick and can be obtained from the excess chemical potential or an equation of state such as the one by Peng and Robinson.[21]

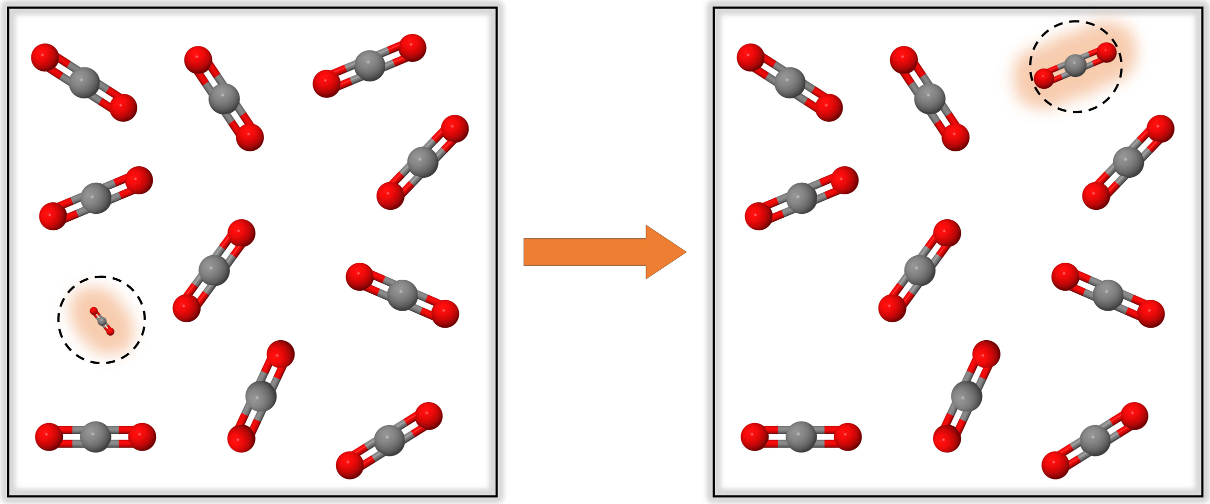

Figure S16: Grand-Canonical Deletion Trial Move: the fractional molecule is removed from the system and a randomly selected molecule is transformed into a fractional molecule.

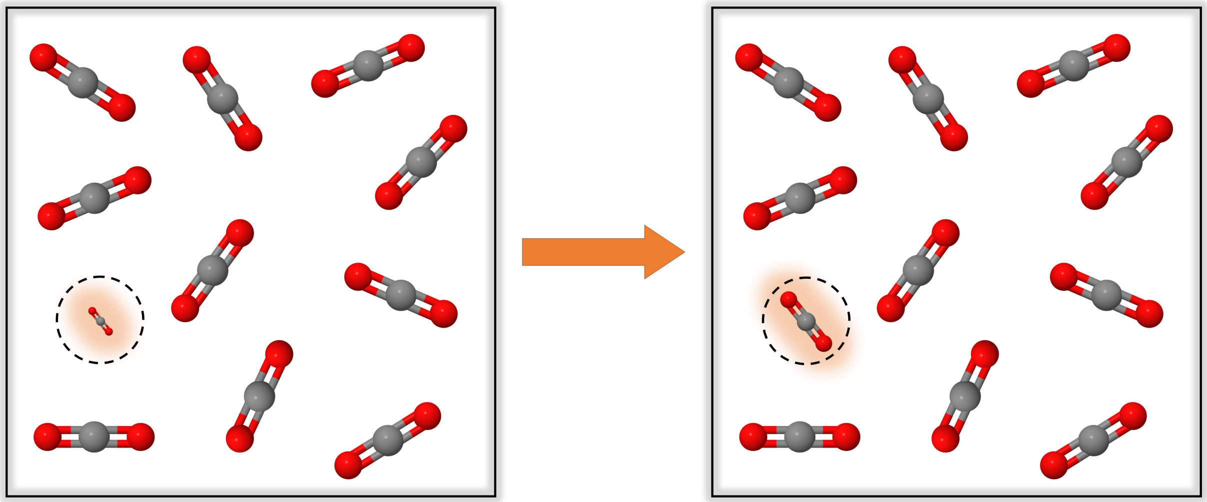

Figure S17: Grand-Canonical Lambda Trial Move: the fractional parameter of the fractional molecule is changed, increasing or decreasing the strength of its interactions.

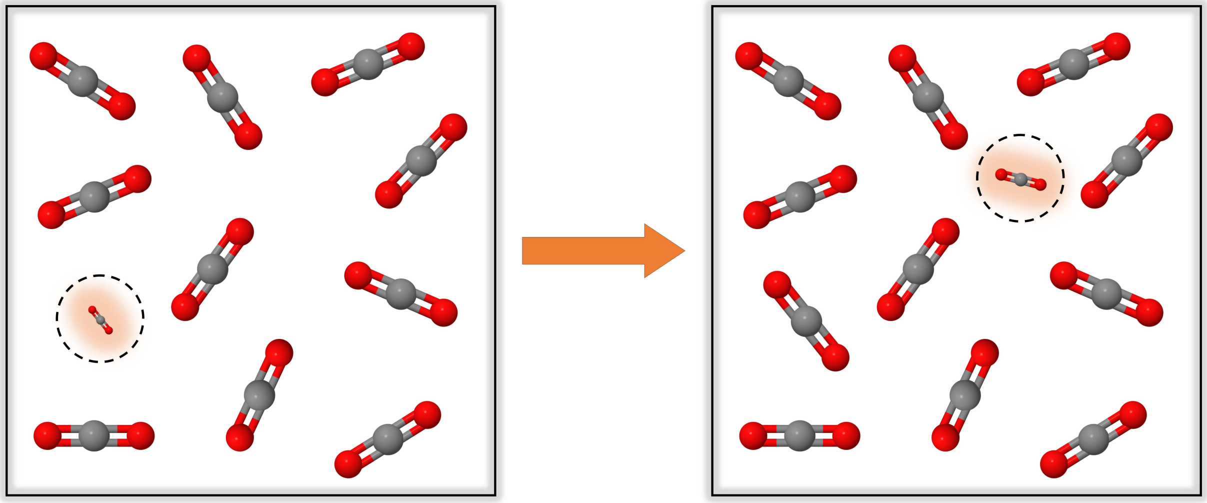

Figure S18: Grand-Canonical Insertion Trial Move: the fractional molecule is transformed into a whole molecule and a new fractional molecule is inserted at a random position with a random orientation.

Weight Functions

With the CFC method, additional degrees of freedom are introduced to the system: the fractional parameter λ. To make simulations efficient, this parameter should be used to add a bias to the system such that the observed probability distribution of λ during the simulation is flat.

- NVT/NPT: the weight function is one-dimensional and only depends on λ: $W = W(\lambda)$. It aims for equally probable values of λ.
- Gibbs Ensemble: the weight function is two-dimensional and depends on λ and the simulation box (i) the fractional molecule is in: $W = W(\lambda, i)$. It aims for equally probable values of λ and makes the fractional molecule equally likely to be found in one of the simulation boxes (50% in simulation box 1 an 50% in simulation box 2).
- Reaction Ensemble: the weight function is two-dimensional and depends on λ and the reaction step (fractional molecules are reactants: $\delta = 0$ or fractional molecules are products: $\delta = 1$): $W = W(\lambda, \delta)$. It aims for equally probable values of λ and makes the fractional molecule equally likely to be found in one of the reaction steps.
- Grand-Canonical Ensemble: the weight function is one-dimensional and only depends on λ: $W = W(\lambda)$. It aims for equally probable values of λ.

In a system with more than one fractional molecule the weight functions are independent,[22] i.e.:

$$W(\lambda_1, \lambda_2, ..., \lambda_n) = \sum_{i=1}^{n} W(\lambda_i). \tag{S51}$$

Interactions

In this section, we consider the potentials that can be used in Brick for intermolecular en intramolecular interactions.

Lennard-Jones Interactions

The Lennard-Jones potential has the functional form:

$$U_{\mathrm{LJ}}\left(r_{ij}\right)=4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{S52}$$

where ϵ_{ij} and σ_{ij} are the LJ-parameters between sites i and j. There are two commonly used methods for dealing with the truncation of the potential. One of them is shifting the potential to 0 at the cutoff radius R_{c}:

$$U_{\mathrm{LJ}}^{\mathrm{shifted}}\left(r_{ij}\right)=\begin{cases}U_{\mathrm{LJ}}\left(r_{ij}\right)-U_{\mathrm{LJ}}\left(R_{\mathrm{c}}\right) & r_{ij}\leq R_{\mathrm{c}}\\ 0 & r_{ij}>R_{\mathrm{c}}\end{cases} \tag{S53}$$

Another option is adding an average energy to the system that accounts for the interactions of the molecules beyond the cutoff radius. This term can be calculated analytically and is referred to as the analytic tail correction. It equals[19]

$$U_{\mathrm{LJ}}^{\mathrm{tailcorrection}}=\frac{1}{2}\sum_{i,j}\frac{16\pi N_iN_j\epsilon_{ij}}{V}\left(\frac{\sigma_{ij}^{12}}{9R_{\mathrm{c}}^{9}}-\frac{\sigma_{ij}^{6}}{3R_{\mathrm{c}}^{3}}\right) \tag{S54}$$

where the sum ranges over all atom types in the system, N_i is the number of atoms of type i (excluding the fractional molecules), and V is the volume of the simulation box. The factor $\frac{1}{2}$ corrects for counting interactions twice. For fractional molecules the Lennard-Jones interactions are scaled. The functional form for these interactions is[23]

$$U_{\mathrm{LJ}}\left(r,\lambda_{\mathrm{LJ}}\right)=4\lambda\epsilon\left[\frac{1}{\left(\alpha_{\mathrm{LJ}}\left(1-\lambda_{\mathrm{LJ}}\right)^{b}+\left(\frac{r}{\sigma}\right)^{c}\right)^{\frac{12}{c}}}-\frac{1}{\left(\alpha_{\mathrm{LJ}}\left(1-\lambda_{\mathrm{LJ}}\right)^{b}+\left(\frac{r}{\sigma}\right)^{c}\right)^{\frac{6}{c}}}\right] \tag{S55}$$

where typical values for α_{LJ}, b and c are 0.5, 1, and 6 respectively. The term $\alpha_{\mathrm{LJ}}(1-\lambda_{\mathrm{LJ}})^{b}$ prevents singularities for small values of r. Shifting this potential at the cutoff is straightforward. For tail corrections the substitution $N_i\rightarrow N_i+\lambda_{\mathrm{LJ}}$ is applied to Eq. S54.[8,24]

Intramolecular LJ interactions are calculated directly from Eq. S52 without truncating the interactions.

Electrostatic Interactions

There are three methods for handling electrostatic interactions in Brick. The most commonly used one is the Ewald method.[25,26] This method splits the electrostatic interactions in a short-range and long-range part and uses a Fourier Transform on the long-range part.

$$U_{\mathrm{electrostatic}}^{\mathrm{Ewald}} = \frac{1}{2} \underbrace{\sum_{i=1}^{N_{\mathrm{m}}} \sum_{a=1}^{N_{\mathrm{a}}^{i}} \sum_{\substack{j=1 \\ j \neq i}}^{N_{\mathrm{m}}} \sum_{b=1}^{N_{\mathrm{a}}^{j}}}_{r_{iajb} < R_{\mathrm{c}}} q_{ia} q_{jb} \frac{\operatorname{erfc}\left(\alpha r_{iajb}\right)}{r_{iajb}} \tag{S56}$$

$$+ \frac{1}{2} \sum_{\substack{i=1 \\ \\ r_{iaib} < R_{\mathrm{c}}}}^{N_{\mathrm{m}}} \sum_{a=1}^{N_{\mathrm{a}}^{i}} \sum_{\substack{b=1 \\ b \neq a}}^{N_{\mathrm{a}}^{j}} q_{ia} q_{ib} \frac{\operatorname{erfc}\left(\alpha r_{iaib}\right)}{r_{iaib}} \tag{S57}$$

$$- \frac{1}{2} \sum_{i=1}^{N_{\mathrm{m}}} \sum_{a=1}^{N_{\mathrm{a}}^{i}} \sum_{b=1}^{N_{\mathrm{a}}^{j}} q_{ia} q_{ib} \frac{1}{r_{iaib}} \tag{S58}$$

$$+ \frac{1}{2V} \sum_{\vec{k} \neq \vec{0}} \frac{4\pi}{k^2} \left| \sum_{i=1}^{N_{\mathrm{m}}} \sum_{a=1}^{N_{\mathrm{a}}^{i}} q_{ia} \exp\left[i\vec{k} \cdot \vec{r}_{ia} \right] \right|^2 \exp\left[-\frac{k^2}{4\alpha^2} \right] \tag{S59}$$

$$- \frac{\alpha}{\sqrt{\pi}} \sum_{i=1}^{N_{\mathrm{m}}} \sum_{a=1}^{N_{\mathrm{a}}^{i}} q_{ia}^{2} \tag{S60}$$

where N_{m} is the number of molecules, N_{a}^{i} is the number of atoms in molecule i, q_{ia} is the partial charge of atom a in molecule i, $\operatorname{erfc}(x)$ is the complementary error function, α is a damping parameter, $r_{iajb} = |\vec{r_{ia}} - \vec{r_{jb}}|$ is the distance between atom a in molecule i and atom b in molecule j, and R_{c} is the cutoff radius. Term S56 is the damped electrostatic potential for the short-ranged interactions. Terms S57 and S58 are corrections for intramolecular interactions (since they are calculated separately) and are referred to as the *exclusion* terms. One could say that term S57 completes the sum over all atoms in the simulation box and term S58 substracts all intramolecular interactions to obtain the proper intermolecular interactions. The sum in term S59 ranges over vectors $\vec{k} = \frac{2\pi}{L}(n_x, n_y, n_z)$ with $L = V^{\frac{1}{3}}$ (the length of one side of the simulation box) and integers $n_x, n_y, n_z \in \mathbb{N}$. Since this is a converging sum we can truncate it at a certain maximum vector k_{max} (or n_{max} such that $n_x, n_y, n_z \leq n_{\text{max}}$). The Fourier Transform (term S59) makes the Ewald method a computationally expensive method but can be optimized by storing the values of the double sum.[27] Term S60 is the self interaction.

Since the Ewald method involves the calculation of a Fourier Transform it is computationally expensive. Therefore, it is also possible to use the Wolf method[28], which works particularly well for dense systems (e.g. liquids). The Wolf method uses the following expression to calculate electrostatic interactions:

$$U_{\text{electrostatic}}^{\text{Wolf}} = \frac{1}{2} \sum_{i=1}^{N_{\text{m}}} \sum_{a=1}^{N_{\text{a}}^{i}} \sum_{\substack{j=1 \\ j \neq i \\ r_{iajb} < R_{\text{c}}}}^{N_{\text{m}}} \sum_{b=1}^{N_{\text{a}}^{j}} q_{ia} q_{jb} \left(\frac{\text{erfc}\,(\alpha r_{iajb})}{r_{iajb}} - \frac{\text{erfc}\,(\alpha R_{\text{c}})}{R_{\text{c}}} \right) \tag{S61}$$

$$+ \frac{1}{2} \sum_{i=1}^{N_{\text{m}}} \sum_{a=1}^{N_{\text{a}}^{i}} \sum_{\substack{b=1 \\ b \neq a \\ r_{iaib} < R_{\text{c}}}}^{N_{\text{a}}^{j}} q_{ia} q_{ib} \left(\frac{\text{erfc}\,(\alpha r_{iaib})}{r_{iaib}} - \frac{\text{erfc}\,(\alpha R_{\text{c}})}{R_{\text{c}}} \right) \tag{S62}$$

$$- \frac{1}{2} \sum_{i=1}^{N_{\text{m}}} \sum_{a=1}^{N_{\text{a}}^{i}} \sum_{b=1}^{N_{\text{a}}^{j}} q_{ia} q_{ib} \frac{1}{r_{iaib}} \tag{S63}$$

$$- \left(\frac{\text{erfc}\,(\alpha R_{\text{c}})}{2R_{\text{c}}} + \frac{\alpha}{\sqrt{\pi}} \right) \sum_{i=1}^{N_{\text{m}}} \sum_{a=1}^{N_{\text{a}}^{i}} q_{ia}^{2} \tag{S64}$$

where the variables and parameters have the same meaning as for the Ewald method. It should be kept in mind that typically the values for the damping parameters α are different for the Ewald and Wolf method.[24] Term S61 is the damped and shifted electrostatic potential. Terms S62 and S63 are corrections for intramolecular interactions (since they are calculated separately) and are referred to as the exclusion terms. One could say that term S62 completes the sum over all atoms in the simulation box and term S63 substracts all intramolecular interactions to obtain the proper intermolecular interactions. Term S64 is the self interaction.

Although the electrostatic interactions can be accurately calculated using the Wolf method, for some systems artificial structuring around the cutoff distance is a potential issue.[29] A modification of the Wolf method by Fenell and Gezelter[30] solves this issue. This is the third method that can be used in Brick.

$$
\begin{aligned}
U_{\text{electrostatic}}^{\text{Fennell-Gezelter}} = \frac{1}{2} & \underset{r_{iajb}<R_{\mathrm{c}}}{\sum_{i=1}^{N_{\mathrm{m}}} \sum_{a=1}^{N_{\mathrm{a}}^{i}} \sum_{\substack{j=1 \\ j\neq i}}^{N_{\mathrm{m}}} \sum_{b=1}^{N_{\mathrm{a}}^{j}}} q_{ia} q_{jb} \Bigg[\frac{\operatorname{erfc}\left(\alpha r_{iajb}\right)}{r_{iajb}} - \frac{\operatorname{erfc}\left(\alpha R_{\mathrm{c}}\right)}{R_{\mathrm{c}}} \\
& + \left(\frac{\operatorname{erfc}\left(\alpha R_{\mathrm{c}}\right)}{R_{\mathrm{c}}^{2}} + \frac{2\alpha}{\sqrt{\pi}} \frac{\exp\left[-\alpha^{2} R_{\mathrm{c}}^{2}\right]}{R_{\mathrm{c}}} \right) \left(r_{iajb} - R_{\mathrm{c}} \right) \Bigg] & \text{(S65)} \\
& + \frac{1}{2} \underset{r_{iaib}<R_{\mathrm{c}}}{\sum_{i=1}^{N_{\mathrm{m}}} \sum_{a=1}^{N_{\mathrm{a}}^{i}} \sum_{\substack{b=1 \\ b\neq a}}^{N_{\mathrm{a}}^{j}}} q_{ia} q_{ib} \left(\frac{\operatorname{erfc}\left(\alpha r_{iaib}\right)}{r_{iaib}} - \frac{\operatorname{erfc}\left(\alpha R_{\mathrm{c}}\right)}{R_{\mathrm{c}}} \right) & \text{(S66)} \\
& - \frac{1}{2} \sum_{i=1}^{N_{\mathrm{m}}} \sum_{a=1}^{N_{\mathrm{a}}^{i}} \sum_{b=1}^{N_{\mathrm{a}}^{j}} q_{ia} q_{ib} \frac{1}{r_{iaib}} & \text{(S67)} \\
& - \left(\frac{\operatorname{erfc}\left(\alpha R_{\mathrm{c}}\right)}{2 R_{\mathrm{c}}} + \frac{\alpha}{\sqrt{\pi}} \right) \sum_{i=1}^{N_{\mathrm{m}}} \sum_{a=1}^{N_{\mathrm{a}}^{i}} q_{ia}^{2} & \text{(S68)}
\end{aligned}
$$

Term S65 is the damped and shifted electrostatic potential. Terms S66 and S67 are corrections for intramolecular interactions (since they are calculated separately) and are referred to as the exclusion terms. One could say that term S66 completes the sum over all atoms in the simulation box and term S67 substracts all intramolecular interactions to obtain the proper intermolecular interactions. Term S68 is the self interaction.

The electrostatic interactions for fractional molecules are scaled similar to the Lennard-Jones interactions. This is achieved by substituting $r_{iajb} \rightarrow r_{iajb} + \alpha_{\mathrm{EL}}(1 - \lambda_{\mathrm{EL}})$ (which prevents singularities at small distances) and $q_{ia} \rightarrow \lambda_{\mathrm{EL}} q_{ia}$ where we typically chose $\alpha_{\mathrm{EL}} = 0.01$ Å. So far we only considered the intermolecular electrostatic interactions. The intramolecular interactions are calculated separately, directly from the Coulomb potential without truncating the interactions:

$$U_{\text{electrostatic}}^{\text{intramolecular}} = \gamma_{ij} \frac{q_i q j}{r_{ij}} \tag{S69}$$

where γ_{ij} is a scaling factor for the intramolecular electrostatic interactions between atoms i and j in a molecule.

Simulation Details

Simulations in the combined Gibbs Ensemble and Reaction Ensemble for the esterification were performed starting from an initial configuration of 900 molecules. The cutoff radius for Lennard-Jones and electrostatic interactions is 14 Å. Analytic tail corrections for the Lennard-Jones interactions are used, and the Ewald method is used for electrostatic interactions. Details about the force field parameters can be found in the Supporting Information. Partition functions of the isolated molecule were obtained from the Gibbs free energy of formation[31] and are listed in the Supporting Information. We consider the system at $T = 343$ K and $P = 1$ bar. For equilibrating the system, $5{\cdot}10^5$ cycles were used, where each cycle consists of N trial moves (N being the total number of molecules in the system). We study two cases: (1) all molecules are rigid; (2) all molecules (except water) are flexible (i.e. bond bending and torsion are taken into account). The trial moves are randomly selected with the following probabilities: 25% translations, 25% rotations, 1% volume changes, 25% λ changes, 16% Gibbs Ensemble swap/identity changes and, 8% Reaction Ensemble swap/identity changes for the case where molecules are treated as rigid. When flexibility of the molecules is taken into account, the trial moves are selected with the following probabilities: 20% translations, 20% rotations, 14% bond bendings, 7% torsions, 1% volume changes, 20% λ changes, 12% Gibbs Ensemble swap/identity changes and, 6% Reaction Ensemble swap/identity changes. The Gibbs Ensemble swap and identity change trial moves facilitate the particle transfer between the two simulation boxes.[7] The Reaction Ensemble swap and identity change trial moves facilitate the conversion of reactants into products (and vice versa) in each simulation box.[8] In the first 10^5 cycles, the Wang-Landau scheme was used to obtain weight functions for the fractional molecules. Then, every 10^5 cycles an iterative scheme was used to improve the weight function further. After equilibrating, 10^5 production cycles were performed to obtain average properties of the system such as the equilibrium composition. Additional simulations were run in the *NPT* ensemble at this equilibrium composition to obtain the chemical potentials using probe molecules of the four species involved in the reaction.

The thermodynamic activities are defined as:[32]

$$a_i = \exp\left[\frac{\mu_i - \mu_i^{\text{ref}}}{RT}\right] = \gamma_i x_i \tag{S70}$$

where a_i is the thermodynamic activity of component i, μ_i is the chemical potential, μ_i^{ref} is the reference chemical potential, γ_i is the activity coefficient, and x_i is the mole fraction. The reference chemical potentials μ_i^{ref} are defined as the chemical potentials of the pure components at the same temperature and pressure. Additional simulations were performed in the *NPT* ensemble to obtain the reference chemical potentials. All computed reference chemical potentials can be found in Table S1.

Tables S1, S2, S3, S4 and S5 list the input that was used for simulations of the esterification of methanol with acetic acid and calculations of chemical potentials.

				Rigid		Flexible	
Molecule	$\mu_i^\circ/[\mathrm{kJ \cdot mol^{-1}}]$	$q_i/\Lambda_i^3/[\text{Å}^{-3}]$	$\rho_i^{\text{Lit.}}/[\mathrm{kg{\cdot}m^{-3}}]$	$\rho_i/[\mathrm{kg{\cdot}m^{-3}}]$	$\mu_i^{\text{ref}}/[\mathrm{kJ \cdot mol^{-1}}]$	$\rho_i/[\mathrm{kg{\cdot}m^{-3}}]$	$\mu_i^{\text{ref}}/[\mathrm{kJ \cdot mol^{-1}}]$
CH_3OH	-125.67	$1.37 \cdot 10^{19}$	743 (Ref. [33])	732(3)	-30.4(3)	731(2)	-30.4(1)
CH_3COOH	-344.66	$3.07 \cdot 10^{52}$	972 (Ref. [34])	968(2)	-33.3(4)	978(3)	-34.0(7)
CH_3COOCH_3	-281.53	$7.46 \cdot 10^{42}$	840 (Ref. [35])	856(1)	-29.6(2)	857(2)	-29.8(2)
H_2O	-195.87	$6.73 \cdot 10^{29}$	969 (Ref. [33])	971(3)	-37.9(3)	n.a.	n.a.

Table S1: Chemical potentials of different molecules at $T = 343$ K calculated from Gibbs free energies of formation (calculated with Outotec HSC Chemistry Software[31]) $\mu_i^\circ = -RT\ln(q_i V_0/\Lambda_i^3)$, partition functions of the isolated molecule (excluding the translational part, $V_0 = 1$ Å^3), densities reported in literature, and densities and chemical potentials at the reference point (defined as the density/chemical potential of the pure component at the same temperature and pressure $P = 1$ bar). Densities and chemical potentials at the reference point were calculated from simulations in the NPT ensemble with the CFC method. The numbers between brackets denote the uncertainty (standard deviation) in the last digit.

Table S2: Force field parameters of methanol,[36] acetic acid,[37] methyl acetate,[35] and water.[38] CH_3 is described as united atom and M is a dummy site.

Molecule	Atom/site	ϵ/k_B/[K]	σ/[Å]	q/[e]
CH_3OH	CH_3	98.0	3.75	0.265
	O	93.0	3.02	-0.700
	H	0	0	0.435
CH_3COOH	CH_3	98.0	3.75	0.120
	C	41.0	3.90	0.420
	O(=C)	79.0	3.05	-0.450
	O(-H)	93.0	3.02	-0.460
	H	0	0	0.370
CH_3COOCH_3	CH_3(-C)	98.0	3.75	0.050
	C	41.0	3.90	0.550
	O(=C)	79.0	3.05	-0.450
	O	55.0	2.80	-0.400
	CH_3(-O)	98.0	3.75	0.050
H_2O	H	0	0	0.52422
	O	81.899	3.16435	0
	M	0	0	-1.04844

Table S3: Bond lengths in methanol,[36] acetic acid,[37] methyl acetate,[35] and water.[38,39]

Molecule	Bond	Length/[Å]
CH_3OH	CH_3-O	1.43
	O-H	0.945
CH_3COOH	CH_3-C	1.520
	C=O	1.214
	C-O	1.364
	O-H	0.970
CH_3COOCH_3	CH_3-C	1.520
	C=O	1.2
	C-O	1.344
	O-CH_3	1.41
H_2O	O-H	0.9572
	O-M	0.15

Table S4: Force field parameters for bond bendings in methanol,[36] acetic acid,[37] methyl acetate,[35] and water.[38,39] Bond bending is described by the potential: $U_{\text{bending}} = \frac{k_\theta}{2}\left(\theta - \theta_0\right)^2$.

Molecule	Bending	$k_\theta/k_\text{B}/[\text{K}]$	θ_0
CH_3OH	CH_3-O-H	55400	108.5°
CH_3COOH	CH_3-C=O	40300	126°
	CH_3-C-O	35300	111°
	O=C-O	40300	123°
	C-O-H	17600	107°
CH_3COOCH_3	CH_3-C=O	62500	125°
	CH_3-C-O	70600	110°
	O=C-O	62500	125°
	C-O-CH_3	62500	115°
H_2O	H-O-H	Rigid	104.52°
	H-O-M	Rigid	52.26°

Table S5: Force field parameters for torsions in acetic acid,[37] and methyl acetate.[35] Torsion is described by the potential: $U_{\text{torsion}} = \sum_{i=0}^{3} c_i \cos^i(\phi)$.

Molecule	Torsion	$c_0/k_\text{B}/[\text{K}]$	$c_1/k_\text{B}/[\text{K}]$	$c_2/k_\text{B}/[\text{K}]$	$c_3/k_\text{B}/[\text{K}]$
CH_3COOH	O=C-O-H	2192.4	-630.0	-1562.4	0
	CH_3-C-O-H	2192.4	630.0	-1562.4	0
CH_3COOCH_3	O=C-O-CH_3	10874.6	-2654.2	-4118.0	613.6
	CH_3-C-O-CH_3	6551.3	1566.1	-4196.0	789.2

Table S6: Sum of chemical potentials of reactants (CH_3OH + CH_3COOH) and products (CH_3COOCH_3 + H_2O) at reaction and phase equilibrium of the esterification of methanol with acetic acid at $T = 343$ K and $P = 1$ bar. Results for a system where all molecules are rigid and a system where molecule are flexible (i.e. bond bending and torsion are taken into account). The number between brackets indicate the undertainty (one standard deviation) in the last digit.

	Phase	$\sum_{\text{Reactants}} \nu_i\mu_i / [\text{kJ}\cdot\text{mol}^{-1}]$	$\sum_{\text{Products}} \nu_i\mu_i / [\text{kJ}\cdot\text{mol}^{-1}]$
Rigid	water-rich	-544(2)	-544(2)
	ester-rich	-545(2)	-544(2)
Flexible	water-rich	-545(2)	-545(2)
	ester-rich	-546(2)	-545(2)

Appendix

In this Appendix we explain how to obtain the partition functions of isolated molecules. These partition functions are required as input for simulations in Brick using the Reaction Ensemble.

Partition Functions of Isolated Molecules

Molecular partition functions are used to calculate thermochemical properties such as the internal energy, entropy, chemical potential, heat capacity, etc.[40,41] These quantities can be obtained from different sources including available thermochemistry data or quantum calculations, each with a well-defined choice of reference state for zero of energy. The main purpose of this section is to explain how reference states for energy calculations can be chosen consistently using different data sets or computer programs. It is assumed that the reader is partially familiar with statistical mechanics and basic concepts in computational chemistry, and this section should not be blindly used as a "cook book" for computation chemistry problems. The theoretical part in this section is based on the Physical Chemistry book by McQuarrie[40] and Essential Statistical Thermodynamics from Computational Chemistry Comparison and Benchmark Data Base (CCCBDB).[42] In this section, commonly used thermodynamic data sets (JANAF tables[43,44]) and the Gaussian09 software[41] are used to provide examples on how to calculate molecular partition functions for nitrogen, hydrogen and ammonia. The results are used to solve the chemical equilibrium in MC simulations of the Haber-Bosch process in the Reaction Ensemble.[6] To compute the molecular partition function, a complete set of molecular energy levels is required, which is almost never available.[42,45] As an approximation, the energy of a molecule can be estimated by decoupling translational, vibrational, rotational and electronic contributions, which means that different energy contributions are unaffected by each other. Using this approximation, the molecular partition function of an isolated molecule can be written as[40]

$$q^*(T) = q_{\mathrm{trans}}(T) q_{\mathrm{rot}}(T) q_{\mathrm{vib}}(T) q_{\mathrm{elec}}(T) \tag{S71}$$

in which the terms on the right hand side of Eq. S71 denote translational, rotational, vibrational and electronic partition functions. Here, it is explained briefly how different contributions in Eq. S71 are calculated to obtain $q^*(T)$.

Translational Partition Function: the translational contribution is obtained from[40]

$$q_{\mathrm{trans}}(T) = \frac{V_0}{\Lambda^3} \qquad \Lambda = \frac{h}{\sqrt{2\pi M k_{\mathrm{B}} T}} \tag{S72}$$

in which Λ is the thermal de Broglie wavelength of the molecule, V_0 is a reference volume, k_{B} is the Boltzmann constant, h is the plank constant, T is the temperature, and M is the sum of all atomic masses in the molecule. The choice of reference volume in Eq. S72 may be defined differently depending on the software. In Gaussian09,[41] the ideal gas law is used to calculate the volume at atmospheric pressure and the translational partition function is defined as:[41]

$$q_{\mathrm{trans}}^{\mathrm{Gaussian}}(T) = \frac{k_{\mathrm{B}} T}{P\Lambda^3} \tag{S73}$$

It is shown later in this section that only the temperature dependent part of the partition function is used to obtain the chemical potential. Therefore, the choice of volume V_0 in Eq. S72 does not affect the chemical potential. For Brick, we take the convention $V_0 = 1$ Å^3.

From here on, we take the convention to explicitly write out the translational part of the partition function in Eq. S71:

$$q^* (T) = q_{\text{trans}}(T)q_{\text{rot}}(T)q_{\text{vib}}(T)q_{\text{elec}}(T) = \frac{q(T)V_0}{\Lambda^3} \tag{S74}$$

$$\text{where } q(T) = q_{\text{rot}}(T)q_{\text{vib}}(T)q_{\text{elec}}(T) \tag{S75}$$

in which we use the reference volume $V_0 = 1$ Å^3 in Eq. S72, and $q(T)$ is the partition function excluding the translational part.

Electronic Partition Function: For a monotonic ideal gas, the electronic partition function is obtained using

$$q_{\text{elec}}(T) = \sum_i g_{\text{e}i} \exp[-\beta\varepsilon_{\text{e}i}] \tag{S76}$$

in which $g_{\text{e}i}$ and $\varepsilon_{\text{e}i}$ are the degeneracy and the energy of the i^{th} electronic level, respectively. The degeneracy of electronic levels is determined by the spin multiplicity,[41] $2S+1$, in which $S =$ is the net electron spin or total spin quantum number.[42] For a monoatomic ideal gas, the zero of electronic energy is fixed at the ground state ($\varepsilon_{\text{e}0} = 0$). Normally, only the first and the second term in the summation in Eq. S76 are considered for the electronic contribution to the partition function.[40] This is because the excited electronic energy levels at typical temperatures are around tens of thousands of wave numbers,[40] which means that the excited energy levels for most substances are nearly inaccessible even at temperatures up to $T = 1000$ K. As an illustrative example, hydrogen atom has a first excited state ${}^2P_{1/2}$ with the energy 82258 in units of cm^{-1}.[40] At $T = 1000$ K, the contribution of the second term in Eq. S76 is in order of 10^{-52}. When considering excited electronic states, the translation, vibrational and rotational contirbutions can be approximated as those in the electronic ground state if no other data is available.[45] For diatomic or polyatomic ideal gas, the arbitrary zero of electronic energy is taken to be the infinitely dissociated atoms at rest in their ground electronic state.[40] The ground electronic level for a diatomic ideal gas

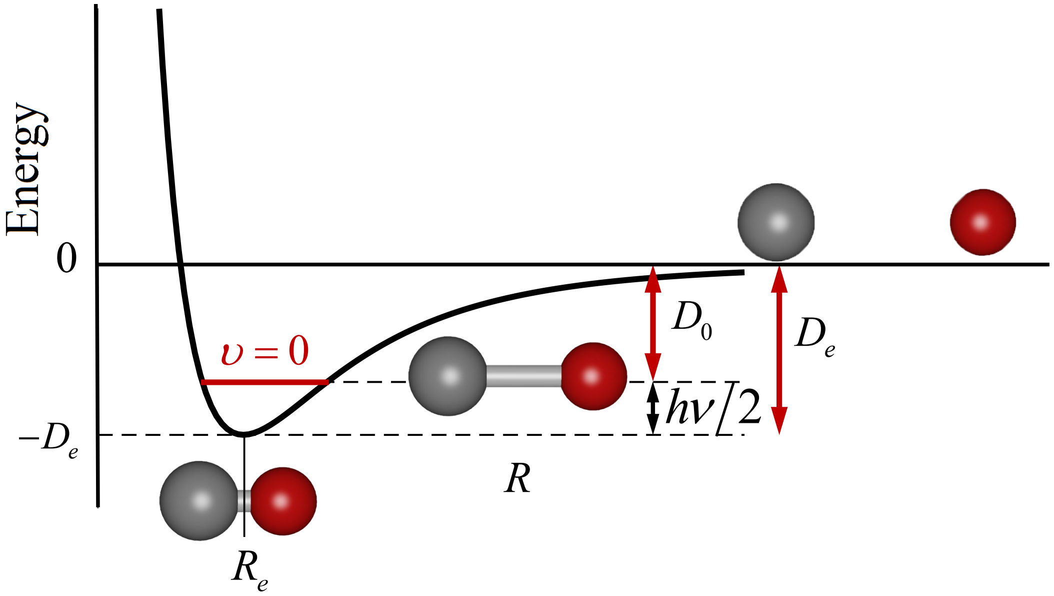

Figure S19: Electronic potential curve of a diatomic molecule as a function of internuclear distance R. $-D_\mathrm{e}$ is the energy of the molecule in the ground state (lowest value), relative to infinitely dissociated atoms. The vibrational energy of the ground state ($\upsilon = 0$) is called zero-point energy which equals $h\nu/2$. $-D_0$ is the corresponding dissociation (atomization) energy when the zero-point energy is selected as zero of energy.

is shown Eq. S19. It is shown in Eq. S19 that the energy difference between the minimum of the internuclear potential well fully dissociated limit (bare nuclei and free electrons) is denoted with D_e, which is the dissociation energy of the molecule. The energy of the ground electronic state is $\varepsilon_{\mathrm{e}1} = -D_\mathrm{e}$. Note the same definition of the molecular ground state holds for a polyatomic molecule. The electronic partition function for a diatomic or polyatomic ideal gas molecule is

$$q_\mathrm{elec}(T) = g_{\mathrm{e}1} \exp[\beta D_\mathrm{e}] + \cdots \tag{S77}$$

in which the contributions from exited electronic levels are not considered at ordinary temperatures. It should be noted that in Gaussian09, the electronic contribution only contains the degeneracy of the electronic ground state,[41] which means

$$q_\mathrm{elec}(T) = g_{\mathrm{e}1} \tag{S78}$$

The distinction between equations S77 and S78 is very important when obtaining the molecular partition function from Gaussian09, as the energy reference for the electronic contribution is different by D_e.

Vibrational partition function: Under the harmonic-oscillator approximation, the accessible vibrational energy levels of a diatomic molecule, relative to the bottom of the internuclear potential well, as shown in Fig. S19 are obtained from[40]

$$\varepsilon_\upsilon = (\upsilon + 1/2\ )\, h\nu \qquad \upsilon = 0, 1, 2, \cdots \tag{S79}$$

in which $\nu = (k/\mu^{1/2}\ )/2\pi$ is the frequency of vibration. k is the force constant of the molecule, υ is the quantum number and μ is the reduced mass. A zero of energy is also required for the vibrational energy levels. Two choices are common for zero of vibrational energy: *(1)* the minimum of the internuclear potential energy curve as shown in Fig. S19, which means $\varepsilon_0 = h\nu/2$. *(2)* The energy of the ground vibrational state is set to zero, which means $\varepsilon_0 = 0$. As shown in Fig. S19, the corresponding dissociation energy is $D_0 = D_\mathrm{e} - h\nu/2$ which means that the zero-point energy is taken to be the ground vibrational state. For a polyatomic molecule, the vibrational motion is described using an independent harmonic oscillator, in terms of normal coordinates. The vibrational energies are written as

$$\varepsilon_\upsilon = \sum_{j=1}^{\alpha} (\upsilon_j + 1/2\ )\, h\nu_j \qquad \upsilon_j = 0, 1, 2, \cdots \tag{S80}$$

in which j denotes the j^th normal mode, and α is the vibrational degree of freedom. For a linear molecule with n atoms, $\alpha = 3n - 5$, and a nonlinear molecule, $\alpha = 3n - 6$. Similar to a diatomic molecule. Considering the first choice for the zero of vibrational energy, the vibrational partition function of a diatomic molecule equals:[40]

$$q_{\mathrm{vib}}^{\mathrm{BOT}}(T) = \prod_{j=1}^{\alpha} \frac{\exp\left[-\Theta_{\mathrm{vib},j}/2T\ \right]}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\ \right]} \tag{S81}$$

in which $\Theta_j = h\nu_j/k_{\mathrm{B}}$ is the characteristic vibrational temperature corresponding to the j^{th} normal mode. The notation "BOT" refers to the zero of energy at the bottom of the internuclear potential well ($-D_{\mathrm{e}}$). This notation is used in Gaussian09 to report the vibrational partition function using Eq. S81. The exponential term in the nominator on the right hand side of Eq. S81, is the contribution of the ground vibrational state.[40] Note that the vibrational energy levels are non-degenerate. Considering the ground vibrational state as zero of vibrational energy, the vibrational partition function is obtained using[41]

$$q_{\mathrm{vib}}^{\mathrm{V=0}}(T) = \prod_{j=1}^{\alpha} \frac{1}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\ \right]} \tag{S82}$$

The notation "V=0" refers to the vibrational ground state is taken to be the arbitrary zero-point energy. This notation is also used in Gaussian09 to report the vibrational partition function using Eq. S82. When the vibrational ground state is taken to be zero of energy, the corresponding dissociation energy, for a general case of a polyatomic molecule relative to the is obtained from, see Fig. S19:

$$D_0 = D_{\mathrm{e}} - \sum_{j=1}^{\alpha} h\nu_j/2 \tag{S83}$$

The second term on the right hand side of Eq. S83 is the zero-point vibrational energy at zero Kelvin, usually abbreviated as ZPE or ZPVE.[41,42] Normally, empirical scaling factors are used to scale the vibrational frequencies obtained from ab inito calculations to correct for vibrational anharmonicity and incomplete electron correlations.[42,46] The empirical scaling factors are found in literature and reference data bases[42,46].

Rotational partition function: Under the rigid-rotator approximation, the rotational energy levels of a diatomic molecule are written as[40]

$$\varepsilon_J = \frac{\hbar^2 J(J+1)}{2I} \qquad J = 0, 1, 2, \cdots \tag{S84}$$

in which I is the moment of inertia of the rotor. The degeneracy of each rotational level equals $g_J = 2J + 1$. As shown in Eq. S84, rotation of a rigid molecule is quantized and only certain rotational energy levels can be occupied.[45] Based on Eq. S84, a convenient choice for the zero of rotational energy is the level $J = 0$. The energies and degeneracies of a linear polyatomic molecule are the same as a diatomic molecule.[40] The expression for the rotational partition function of a linear polyatomic molecule or a diatomic molecule is[40]

$$q_{\mathrm{rot}}(T) = \frac{T}{\sigma \Theta_{\mathrm{rot}}} \tag{S85}$$

in which $\Theta_{\mathrm{rot}} = h^2/8\pi^2 I k_{\mathrm{B}}$ is the characteristic rotational temperature, and σ is the symmetry number or rotational symmetry number of the molecule. The symmetry number is the number of unique orientations of the rigid molecule obtained by interchanging identical atoms.[40–42,45] For a heteronuclear diatomic or unsymmetrical molecule, $\sigma = 1$, and for a homonuclear diatomic or symmetrical molecule $\sigma = 2$. The rotational symmetry number can be identified using group theory if the point group of the molecule is known. The symmetry number corresponding to different point groups is provided in Table S7.[42] As a reference, a list of molecules and corresponding symmetry numbers is found on the website of CCCBDB.[42] It is also possible to obtain σ from counting manually the number of unique orientations of molecule. It is highly recommended to double-check the point group of the molecule from Gaussian09 output. For non-linear, polyatomic molecules, the partition

Table S7: Point groups and the corresponding symmetry number. The values in this table are obtained from the Computational Chemistry Comparison and Benchmark DataBase (CCCBDB) website.[42]

Group	σ
$C_1, C_i, C_s, C_{\infty\nu}$	1
$C_n, C_{n\nu}, C_{nh}$	n
$D_{\infty h}$	2
D_n, D_{nh}, D_{nd}	$2n$
T, T_d	12
S_n	$n/2$
O_h	24
I_h	60

function is obtained from[40]

$$q_{\mathrm{rot}}(T) = \frac{\pi^{1/2}}{\sigma}\left(\frac{T^3}{\Theta_{\mathrm{rot,A}}\Theta_{\mathrm{rot,B}}\Theta_{\mathrm{rot,C}}}\right)^{1/2} \tag{S86}$$

in which $\Theta_{\mathrm{rot,j}}$ is the characteristic rotational temperature corresponding to the three principal moments of inertia; A,B,C.[40] If $\Theta_{\mathrm{rot,A}} = \Theta_{\mathrm{rot,B}} = \Theta_{\mathrm{rot,C}}$ the molecule is a symmetric top, and if $\Theta_{\mathrm{rot,A}} \neq \Theta_{\mathrm{rot,B}} \neq \Theta_{\mathrm{rot,C}}$ the molecule is a asymmetric top. The molecule is called symmetric top if $\Theta_{\mathrm{rot,A}} = \Theta_{\mathrm{rot,B}} \neq \Theta_{\mathrm{rot,C}}$.[40]

Molecular partition function: Combining equations S71, S72, S77, S81, and S86, the molecular partition function for a polyatomic ideal gas equals

$$q\,(T) = \frac{\pi^{1/2}}{\sigma}\left(\frac{T^3}{\Theta_{\mathrm{rot,A}}\Theta_{\mathrm{rot,B}}\Theta_{\mathrm{rot,C}}}\right)^{1/2}\prod_{j=1}^{\alpha}\frac{\exp\left[-\Theta_{\mathrm{vib},j}/2T\right]}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\right]}\cdot g_{\mathrm{e1}}\exp[D_{\mathrm{e}}/k_{\mathrm{B}}T\,] \tag{S87}$$

in which the zero-point energy for the vibrational ground state is the bottom of the internuclear potential well, and the zero of electronic energy is the dissociated atoms in the ground state. This means that the ground state electronic energy equals $-D_{\mathrm{e}}$ as shown in Fig. S19. In the thermochemistry output of Gaussian09, the energy of the electronic ground state is taken to be zero, and the corresponding partition function denoted by "Q Total Bot" is the same as Eq. S87 except for the factor of $\exp[\beta D_\mathrm{e}]$.[41] It is important to consider this factor when evaluating equilibrium constant of reaction from molecular partition functions. This will be highlighted later in this section. The molecular partition function (excluding the translational part) is also obtained by combining equations S71, S72, S77, S81, S83, and S86.

$$q\left(T\right) = \frac{\pi^{1/2}}{\sigma}\left(\frac{T^3}{\Theta_{\mathrm{rot,A}}\Theta_{\mathrm{rot,B}}\Theta_{\mathrm{rot,C}}}\right)^{1/2}\prod_{j=1}^{\alpha}\frac{1}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\ \right]}\cdot g_{\mathrm{e1}}\exp[D_0/k_\mathrm{B}T] \tag{S88}$$

in which the zero-point for the vibrational energy is the ground state energy, and the corresponding the ground state electronic energy is $-D_0$ as shown in Fig. S19. Based on equations S81, S82, and S83, one can observe that equations S87 and S88 are identical. In thermodynamic tables, it is however common to take the ground state of molecule (vibrational and electronic) as zero of energy, instead of the dissociated atoms. This is performed by factoring out the contribution of the ground state energy of the molecule from the partition function excluding the translational part:

$$\begin{aligned} q\left(T\right) &= \sum_j \exp\left[-\varepsilon_j/k_\mathrm{B}T\ \right] \\ &= \exp\left[-\varepsilon_0/k_\mathrm{B}T\ \right]\sum_j \exp\left[-(\varepsilon_j-\varepsilon_0)/k_\mathrm{B}T\ \right] \\ &= \exp\left[-\varepsilon_0/k_\mathrm{B}T\ \right] q_0\left(T\right) \end{aligned} \tag{S89}$$

The notation $q_0(T)$ for the molecular partition functions highlights the fact that the ground state energy of the molecule is zero (instead of $-D_0$). By comparing equations S89 and S88, it is clear that $\varepsilon_0 = -D_0$. For a general case of a polyatomic molecule, we have

$$q_0\left(T\right) = \frac{\pi^{1/2}}{\sigma}\left(\frac{T^3}{\Theta_{\mathrm{rot,A}}\Theta_{\mathrm{rot,B}}\Theta_{\mathrm{rot,C}}}\right)^{1/2}\prod_{j=1}^{\alpha}\frac{1}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\ \right]}\cdot g_{\mathrm{e1}} \tag{S90}$$

In the thermochemistry output of Gaussian09, the value reported as "Q Total V=0"[41] is equal to $\frac{k_\mathrm{B}T}{P\Lambda^3}q_0\,(T)$.

Atomization energy (D_0) : The experimental values for the atomization energy of several molecules are reported in the NIST database[42] and other thermodynamic references.[40] The atomization energy can also be obtained using ab inito calculation of molecular energies. As shown in Fig. S19, for a diatomic molecule, D_0 is obtained from the difference between the molecular energy (vibrational and electronic) in the ground state and the dissociated atoms in their respective ground state. The same approach holds for a polyatomic molecule. Using the notation as in the Gaussian09 manual,[41] the atomization energy of a polyatomic molecule is obtained from

$$D_0 = \sum_{i=1}^{N} y_i(\varepsilon_{\mathrm{e},i}) - (\varepsilon_\mathrm{e} + \varepsilon_\mathrm{ZPE}) \tag{S91}$$

in which y_i indicates the number of atoms of kind i in the molecule. $\varepsilon_{\mathrm{e},i}$ is the electronic energy of the i^th atom (dissociated). The second term on the right hand side is the sum of electronic and ZPE of the molecule (vibrational energy in the ground state). It should be noted that computing the atomization energy to chemical accuracy (usually defined as 1 kcal/mol) is not trivial. Advanced methods (*e.g.* Gaussian-n composite methods[47,48] are recommended for accurate calculation of atomization energies. This is beyond the scope of this thesis. From JANAF tables, the atomization energy is obtained from the difference between the enthalpy of formation of the molecule and the dissociated atoms:

$$D_0 = \sum_{i=1}^{N} y_i \Delta_f H_i^\circ(0\ \mathrm{K}) - \Delta_f H^\circ(0\ \mathrm{K}) \tag{S92}$$

in which $\Delta_f H_i^\circ(0\ \mathrm{K})$ is the enthalpy of formation of the i^th atom (dissociated) and $\Delta_f H^\circ(0\ \mathrm{K})$ is the enthalpy of formation of the molecule.

Chemical Potentials

It is shown that the standard chemical potential and the molecular partition function are related:[40]

$$\mu^{\circ}(T) = -RT \ln \left[\frac{q\,(T)\,k_{\mathrm{B}}T}{P\Lambda^3} \right] \tag{S93}$$

$\mu^{\circ}(T)$ is the standard chemical potential of an ideal gas molecule and $q\,(T)$ is the partition function (excluding the translational part) from Eq. S87 or Eq. S88. To tabulate the chemical potential in Eq. S93, a zero of energy is required. In thermodynamic tables, it is common to take the ground state energy of the molecule to be zero. This leads to

$$\mu^{\circ}(T) - E_0 = -RT \ln \left[\frac{q_0\,(T)\,k_{\mathrm{B}}T}{P\Lambda^3} \right] \tag{S94}$$

in which $E_0 = -D_0 = N_{\mathrm{A}}\varepsilon_0$ (N_{A} being Avogadro's number) and $q^0\,(T)$ is the partition function with the ground state energy equal to zero (Eq. S89). With this energy reference, we show that the E_0 is the standard molar enthalpy of molecule at $T = 0$ K. For an ideal gas, the enthalpy can be written in terms of partition function:

$$\begin{aligned} H^{\circ}(T) &= U + RT \\ &= RT^2 \frac{\partial \ln q(T)}{\partial T} + RT \\ &= N_{\mathrm{A}}\varepsilon_0 + RT^2 \frac{\partial \ln q_0(T)}{\partial T} + RT \end{aligned} \tag{S95}$$

Evaluating the enthalpy in Eq. S95 at $T = 0$ shows that $H^{\circ}(0\text{ K}) = E_0 = N_{\mathrm{A}}\varepsilon_0$. The ideal gas partition function excluding the translational part $q_0(T)$ in Eq. S95 is the same as in Eq. S90 which can be obtained by rearranging Eq. S93 or Eq. S94. $q_0(T)$ can be obtained using rigid rotator-harmonic oscillator approximation which agrees reasonably well with experiments.[40] To improve the accuracy, experimental data may be used to complement the theoretical calculations.[40] The combination of experimentally determined thermodynamic

Table S8: An example of a JANAF table entry.[43,44] In the JANAF tables, the standard Gibbs energy (chemical potential) is referenced to the enthalpy at $T_\mathrm{r} = 298.15$ K. To obtain the chemical potential of a molecule in which the ground state energy is taken to be zero, the values for the Gibbs energy are shifted relative to the enthalpy at $T_\mathrm{r} = 0$ K, as described in Eq. S96

.

Nitrogen Oxide		(NO$_2$)					N$_1$O$_2$ (g)
		J.K^{-1}.mol^{-1}		**kJ.mol^{-1}**			
$\boldsymbol{T}$**/K**	$\boldsymbol{C_P^\circ}$	$\boldsymbol{S^\circ}$	$\boldsymbol{-[G^\circ - H^\circ(T_\mathrm{r})]/T}$	$\boldsymbol{H - H^\circ(T_\mathrm{r})}$	$\boldsymbol{\Delta_\mathrm{f}H^\circ}$	$\boldsymbol{\Delta_\mathrm{f}G^\circ}$	**log K$_\mathrm{f}$**
0	0	0	INFINITE	-10.186	35.927	35.927	INFINITE
100	33.276	202.563	271.168	-6.861	34.898	39.963	-20.874
200	34.385	225.852	243.325	-3.495	33.897	45.422	-11.863
250	35.593	233.649	240.634	-1.746	33.46	48.355	-10.103
298.15	36.974	240.034	240.034	0	33.095	51.258	-8.98

properties and theoretical calculations can be found in the Joint, Army, Navy Air Force (JANAF) tables.[43,44] Thermodynamic functions and parameters including the Gibbs free energy, enthalpy and heat capacity are extensively tabulated in JANAF tables.[43,44] A JANAF table entry for NO$_2$ is provided as an example in Fig. S8. Using JANAF tables, it is possible to calculate $q_0(T)$ in Eq. S94 without performing direct quantum mechanical calculations. For a pure component, we have $\mu^\circ = G^\circ$. For a pure substance, $\mu^\circ(T) - E_0^\circ$ in Eq. S94 is obtained from JANAF tables (denoted with $G^\circ(T) - H_0^\circ$). To obtain the chemical potential of a molecule in which the ground state energy is taken to be zero (see Eq. S94), the values for the Gibbs energy are shifted relative to the enthalpy at $T_\mathrm{r} = 0$ K. By rearranging the data in the fourth and fifth columns and multiplying by the temperature (see Table S8) we obtain the chemical potential as shown in Eq. S94:

$$\begin{aligned}\mu^\circ(T) - E_0 &= -RT \ln\left[\frac{q_0\,(T)\,k_\mathrm{B}T}{P\Lambda^3}\right] \\ &= G^\circ(T) - H^\circ(0\text{ K}) \\ &= [G^\circ(T) - H^\circ(298.15\text{ K})] - [H^\circ(0\text{ K}) - H^\circ(298.15\text{ K})]\end{aligned} \tag{S96}$$

For obtaining the $E_0 = -D_0$ the reader is refered to the section about the atomization energy.

Chemical Equilibrium

For a general case of homogeneous gas phase chemical reaction, it is well-known that the Gibbs energy of reaction and the equilibrium constant are related to chemical potentials of reactants and products (at P°). For a multicomponent reacting mixture of S distinguishable components we have[40,49]

$$\begin{aligned}\Delta G_r^{\circ}(T) &= \sum_{i=1}^{S} \nu_i \mu_i^{\circ}(T) \\ &= -RT \ln K(T) \\ &= -RT \ln \left[\prod_{i=1}^{S} \left(\frac{q(T)V_0}{\Lambda^3} \right) \right]^{\nu_i} \\ &= -RT \sum_{i=1}^{S} \nu_i \ln \left[\left(\frac{q(T)V_0}{\Lambda^3} \right) \right] \qquad \text{(S97)}\end{aligned}$$

ν_i is the stoichiometric coefficient of component i, and $K(T)$ is the equilibrium constant of the reaction.[40] The reaction enthalpy at standard pressure is calculated directly from the Gibbs-Helmholtz equation[50,51]

$$\left(\frac{\partial \Delta G_r^{\circ}/T}{\partial T} \right)_P = -\frac{\Delta H_r^{\circ}}{T^2} \qquad \text{(S98)}$$

Ammonia Synthesis Reaction

The ideal gas partition functions are calculated for nitrogen, hydrogen, and ammonia using experimental thermochemistry data,[40] JANAF tables[44] and quantum calculations using Gaussian09.[41] The frequency analysis is performed using two different basis sets: B3LYP level of theory with a 6-31G** basis set, and MP2 level of theory with a 6-311G** basis set.[41] It should be noted that other software packages are also available for performing frequency analysis, such as ADF,[52] Spartan[53] etc. For details about Gaussian09 input files, the reader is referred to the manual.[41] The characteristic vibrational and rotational temperatures of nitrogen, hydrogen, and ammonia obtained from experimental data and ab intio calculations are provided in Table S9. For atomization energies of components, only the experimental data provided by McQuarrie[40] are used and not the ones obtained from Gaussian09.[41] Atomization energies can also be obtained from JANAF tables using Eq. S92. The results are shown in Table S10. One can easily see that computation of the atomization energies using a single basis set results in energy differences well above chemical accuracy.[54] Advanced methods such as (*e.g.* Gaussian-n composite methods[47,48] are recommended for ab initio calculation of atomization energies which is beyond the scope of this thesis. The thermochemical data in Tables S9 and S10 are used to compute the partition functions of nitrogen, hydrogen and ammonia at temperatures between $T = 573$ K and $T = 873$ K. The results are presented in Table S11. Note that the corresponding vibrational and rotational partition functions for diatomic molecules were used for hydrogen and nitrogen, and for ammonia the corresponding vibrational and rotational partition functions were used.

Partition Functions from thermochemistry as input for Brick

To calculate molecular partition functions (for example from Gaussian09[41]) and use them as input for simulations in Brick one can go through the following steps to calculate the different contributions:

1. q_{trans}, the translational partition function. This term can be directly calculated from the thermal wavelength:

$$\Lambda = \frac{h}{\sqrt{2\pi M k_{\mathrm{B}} T}} \tag{S99}$$

$$q_{\text{trans}} = \frac{V_0}{\Lambda^3} \text{ with } V_0 = 1\ \text{Å}^3. \tag{S100}$$

   M is the mass of one molecule (in kg) and the thermal wavelength in Å.

2. q_{elec}, the electronic partition function where typically only the first electronic level is

considered (the last factor on the right-hand side in Eq. S88):

$$q_{\mathrm{elec}} = g_{\mathrm{e1}} \exp\left[\frac{D_0}{k_{\mathrm{B}}T}\right] \tag{S101}$$

where the degeneracy g_{e1}, and dissociation energy D_0 can also be obtained from literature. The partition function q_{elec} can also be directly obtained from Gaussian09.

3. q_{vib}, the vibrational partition function. This term can be obtained directly from Gaussian09 or computed manually:

$$q_{\mathrm{vib}} = \prod_{j=1}^{\alpha} \frac{1}{1 - \exp\left[-\Theta_{\mathrm{vib},j}/T\right]} \tag{S82}$$

   where the vibrational temperatures $\Theta_{\mathrm{vib},j}$ can be obtained from literature.[42]

4. q_{rot}, the rotational partition function. This term can be obtained directly from Gaussian09 or computed manually:

$$q_{\mathrm{rot}} = \frac{\pi^{1/2}}{\sigma}\left(\frac{T^3}{\Theta_{\mathrm{rot,A}}\Theta_{\mathrm{rot,B}}\Theta_{\mathrm{rot,C}}}\right)^{1/2} \tag{S86}$$

   where the symmetry number σ (Table S7) and rotational temperatures $\Theta_{\mathrm{rot},A/B/C}$ can be obtained from literature.[42]

5. Multiply all terms to obtain the molecular partition function (Eq. S71):

$$q^* = q_{\mathrm{trans}} q_{\mathrm{rot}} q_{\mathrm{vib}} q_{\mathrm{elec}} = \frac{q_{\mathrm{rot}} q_{\mathrm{vib}} q_{\mathrm{elec}} V_0}{\Lambda^3} = \frac{q V_0}{\Lambda^3} \tag{S71}$$

   with $V_0 = 1\ \mathrm{\mathring{A}}^3$.

6. Since typical numbers for $\frac{qV_0}{\Lambda^3}$ are large, in Brick the natural logarithm is used as input:

$$\textbf{Input for Brick:} \quad \ln\left(\frac{qV_0}{\Lambda^3}\right)$$

Example

We calculate the partition function of NH_3 at $T = 573$ K following the steps above as an example.

1. $M = 2.828 \cdot 10^{-26}$ kg so that $\Lambda = 0.1767$ Å and $q_{\mathrm{trans}} = 181.2$.
2. $D_0 = 1158$ kJ/mol and the degeneracy g_{e1}=1 so that $g_{\mathrm{elec}} = 3.641 \cdot 10^{105}$.
3. The vibrational temperatures are 1360 K, 4800 K, 2330 K, 2330 K, 4880 K, and 4880K so that $q_{\mathrm{vib}} = 1.142$.
4. The rotational temperatures are 13.60 K, 13.60 K, and 8.92 K (McQuarrie et al.[40] as reported in Table S9) and the symmetry number is 3, so that $q_{\mathrm{rot}} = 199.5$.
5. Multiplying all terms yields: $\frac{qV_0}{\Lambda^3} = 1.503 \cdot 10^{110}$.
6. Taking the logarithm of this number is the input that can be used in Brick:

$$\textbf{Input for Brick:} \quad \ln\left(\frac{qV_0}{\Lambda^3}\right) = \mathbf{253.7}$$

Partition Functions from JANAF tables as input for Brick

To calculate molecular partition functions, using JANAF tables, and use them as input for simulations in Brick one can go through the following steps:

1. Calculate the chemical potential:

$$\mu^{\circ}(T) - E_0 = [G^{\circ}(T) - H^{\circ}(298.15\text{ K})] - [H^{\circ}(0\text{ K}) - H^{\circ}(298.15\text{ K})] \tag{S96}$$

using the JANAF tables.

2. Calculate the partition function excluding the contribution from the atomization energy (S90):

$$\frac{q_0}{\Lambda^3} = \frac{P}{k_{\mathrm{B}}T} \exp\left[-\frac{\mu^\circ(T) - E_0}{RT}\right] \tag{S102}$$

   where P is the reference pressure (typically $P = 1$ atm).

3. Multiply the result by the contribution from the atomization energy:

$$\frac{q}{\Lambda^3} = \frac{q_0}{\Lambda^3} \exp\left[\frac{D_0}{RT}\right] \tag{S103}$$

   where the dissociation energy D_0 can be obtained from literature.[42,44]

4. Multiply the result by the reference volume used in Brick $V_0 = 1$ Å^3 and take the natural logarithm of the result:

$$\textbf{Input for Brick:} \quad \ln\left(\frac{qV_0}{\Lambda^3}\right)$$

Example (JANAF)

We calculate the partition function of NO_2 at $T = 298.15$ K following the steps above as an example from the JANAF tables (Table S8).

1. From Table S8 we calculate

$$\begin{aligned} \mu^\circ(T) - E_0 &= [G^\circ(298.15) - H^\circ(298.15\text{ K})] - [H^\circ(0\text{ K}) - H^\circ(298.15\text{ K})] && \text{(S104)} \\ &= -6.134 \cdot 10^4 \text{ J/mol} && \text{(S105)} \end{aligned}$$

   Be aware of the different units used in the JANAF tables for different columns.

2. In JANAF $P = 1.013 \cdot 10^5$ Pa which yields: $\frac{q_0}{\Lambda^3} = 4.407 \cdot 10^{-16}$ Å^{-3}.

3. The dissociation energy $D_0 = 927.7$ kJ/mol.[42] This yields: $\frac{q}{\Lambda^3} = 1.480 \cdot 10^{147}$ m^{-3}.

4. Multiplying the result by the reference volume, $V_0 = 1$ Å^3 and taking the logarithm of this number is the input that can be used in Brick:

$$\textbf{Input for Brick:} \qquad \ln\left(\frac{qV_0}{\Lambda^3}\right) = \mathbf{338.8}$$

Table S9: Characteristic vibrational and rotational temperatures of nitrogen, hydrogen and ammonia obtained from experimental data (McQuarrie et al.[40]) and ab initio calculations using Gaussian09.[41] In the table, Gaussian (1) denotes frequency calculations using B3LYP level of theory with a 6-31G** basis set, and Gaussian (2) denotes frequency calculations using MP2 level of theory with a 311G**basis set. To correct for vibrational anharmonicity, the vibrational frequencies obtained from Gaussian09 are scaled by empirical factors, 0.96 for the B3LYP/6-31G** and 0.95.[42,46] The numbers in parentheses indicate the degeneracy of the modes.

	McQuarrie		Gaussian (1)		Gaussian (2)	
	Θ_{vib}/[K]	Θ_{rot}/[K]	Θ_{vib}/[K]	Θ_{rot}/[K]	Θ_{vib}/[K]	Θ_{rot}/[K]
N_2	3374	2.88	3393	2.83	2979	2.76
H_2	6332	85.2	6161	87.16	6196	88.29
NH_3	1360	13.60	1509	14.10	1542	14.24
	4800	13.60	4781	14.10	4819	14.24
	2330(2)	8.92	2340(2)	9.14	2292(2)	9.18
	4880(2)		4958(2)		5015(2)	

Table S10: Experimental atomization energies (McQuarrie et al.[40]) and atomization energies computed for nitrogen, hydrogen, ammonia using Gaussian09 using the B3LYP level of theory with a 6-31G** basis set.[41]

Component	D_0 / [kJ/mol]	
	McQuarrie	Gaussian
N_2	941.6	917.6
H_2	432.1	432.1
NH_3	1158	1149.8

Table S11: Computed ideal gas partition functions of nitrogen, hydrogen, ammonia obtained as defined in this. The reported values are based on experimental data (McQuarrie[40] and JANAF tables[44]) and quantum computations using Gaussian09.[41] In the table, Gaussian (1) denotes frequency calculations using B3LYP level of theory with a 6-31G** basis set, and Gaussian (2) denotes frequency calculations using MP2 level of theory with a 311G** basis set. It is important to note that for the values obtained from Gaussian09, the experimental atomization energies provided by McQuarrie[40] are used (see Table S10). One can easily see that using the atomization energies from Gaussian09, listed in Table S10, results in large deviations.

	$q/\Lambda^3\ [\mathrm{N_2}]/[\mathrm{\AA}^{-3}]$		$q/\Lambda^3\ [\mathrm{H_2}]/[\mathrm{\AA}^{-3}]$		$q/\Lambda^3\ [\mathrm{NH_3}]/[\mathrm{\AA}^{-3}]$	
$T/[K]$	Gaussian (1)	Gaussian (2)	Gaussian (1)	Gaussian(2)	Gaussian (1)	Gaussian (2)
573	$2.65 \cdot 10^{90}$	$2.72 \cdot 10^{90}$	$5.94 \cdot 10^{40}$	$5.87 \cdot 10^{40}$	$1.40 \cdot 10^{110}$	$1.38 \cdot 10^{110}$
673	$7.00 \cdot 10^{77}$	$7.22 \cdot 10^{77}$	$1.25 \cdot 10^{35}$	$1.23 \cdot 10^{35}$	$5.01 \cdot 10^{94}$	$4.94 \cdot 10^{94}$
773	$3.50 \cdot 10^{68}$	$3.62 \cdot 10^{68}$	$8.09 \cdot 10^{30}$	$7.99 \cdot 10^{30}$	$1.94 \cdot 10^{83}$	$1.92 \cdot 10^{83}$
873	$2.46 \cdot 10^{61}$	$2.56 \cdot 10^{61}$	$4.96 \cdot 10^{27}$	$4.90 \cdot 10^{27}$	$3.33 \cdot 10^{74}$	$3.29 \cdot 10^{74}$
$T/[K]$	McQuarrie	JANAF	McQuarrie	JANAF	McQuarrie	JANAF
573	$2.60 \cdot 10^{90}$	$2.67 \cdot 10^{90}$	$6.08 \cdot 10^{40}$	$6.53 \cdot 10^{40}$	$1.50 \cdot 10^{110}$	$1.46 \cdot 10^{110}$
673	$6.89 \cdot 10^{77}$	$7.04 \cdot 10^{77}$	$1.27 \cdot 10^{35}$	$1.36 \cdot 10^{35}$	$5.42 \cdot 10^{94}$	$5.26 \cdot 10^{94}$
773	$3.44 \cdot 10^{68}$	$3.52 \cdot 10^{68}$	$8.27 \cdot 10^{30}$	$8.79 \cdot 10^{30}$	$2.12 \cdot 10^{83}$	$2.06 \cdot 10^{83}$
873	$2.42 \cdot 10^{61}$	$2.48 \cdot 10^{61}$	$5.07 \cdot 10^{27}$	$5.38 \cdot 10^{27}$	$3.65 \cdot 10^{74}$	$3.58 \cdot 10^{74}$

References

(1) Adams, D. Chemical Potential of Hard-Sphere Fluids by Monte Carlo Methods. *Mol. Phys.* **1974**, *28*, 1241–1252.

(2) Adams, D. Grand Canonical Ensemble Monte Carlo for a Lennard-Jones Fluid. *Mol. Phys.* **1975**, *29*, 307–311.

(3) Panagiotopoulos, A. Z. Direct Determination of Phase Coexistence Properties of Fluids by Monte Carlo Simulation in a New Ensemble. *Mol. Phys.* **1987**, *61*, 813–826.

(4) Panagiotopoulos, A.; Quirke, N.; Stapleton, M.; Tildesley, D. Phase Equilibria by Simulation in the Gibbs Ensemble. *Mol. Phys.* **1988**, *63*, 527–545.

(5) Johnson, J. K.; Panagiotopoulos, A. Z.; Gubbins, K. E. Reactive Canonical Monte Carlo. *Mol. Phys.* **1994**, *81*, 717–733.

(6) Rahbari, A.; Hens, R.; Nikolaidis, I. K.; Poursaeidesfahani, A.; Ramdin, M.; Economou, I. G.; Moultos, O. A.; Dubbeldam, D.; Vlugt, T. J. H. Computation of Partial Molar Properties using Continuous Fractional Component Monte Carlo. *Mol. Phys.* **2018**, *116*, 3331–3344.

(7) Poursaeidesfahani, A.; Torres-Knoop, A.; Dubbeldam, D.; Vlugt, T. J. H. Direct Free Energy Calculation in the Continuous Fractional Component Gibbs Ensemble. *J. Chem. Theory Comput.* **2016**, *12*, 1481–1490.

(8) Poursaeidesfahani, A.; Hens, R.; Rahbari, A.; Ramdin, M.; Dubbeldam, D.; Vlugt, T. J. H. Efficient Application of Continuous Fractional Component Monte Carlo in the Reaction Ensemble. *J. Chem. Theory Comput.* **2017**, *13*, 4452–4466.

(9) Shi, W.; Maginn, E. J. Continuous Fractional Component Monte Carlo: An Adaptive Biasing Method for Open System Atomistic Simulations. *J. Chem. Theory Comput.* **2007**, *3*, 1451–1463.

(10) Rosch, T. W.; Maginn, E. J. Reaction Ensemble Monte Carlo Simulation of Complex Molecular Systems. *J. Chem. Theory Comput.* **2011**, *7*, 269–279.

(11) Rahbari, A.; Hens, R.; Moultos, O. A.; Dubbeldam, D.; Vlugt, T. J. H. Multiple Free Energy Calculations from Single State Point Continuous Fractional Component Monte Carlo Simulation using Umbrella Sampling. *J. Chem. Theory Comput.* **2020**, *16*, 1757–1767.

(12) Shi, W.; Maginn, E. J. Improvement in Molecule Exchange Efficiency in Gibbs Ensemble Monte Carlo: Development and Implementation of the Continuous Fractional Component Move. *J. Comput. Chem.* **2008**, *29*, 2520–2530.

(13) Poursaeidesfahani, A.; Rahbari, A.; Torres-Knoop, A.; Dubbeldam, D.; Vlugt, T. J. H. Computation of Thermodynamic Properties in the Continuous Fractional Component Monte Carlo Gibbs Ensemble. *Mol. Simul.* **2017**, *43*, 189–195.

(14) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Academic Press: 525 B Street, Suite 1900, San Diego, California 92101-4495, United States of America, 2002.

(15) Naden, L. N.; Pham, T. T.; Shirts, M. R. Linear Basis Function Approach to Efficient Alchemical Free Energy Calculations. 1. Removal of Uncharged Atomic Sites. *J. Chem. Theory Comput.* **2014**, *10*, 1128–1149.

(16) Brick-CFCMC. `https://gitlab.com/ETh_TU_Delft/Brick-CFCMC`.

(17) Orkoulas, G.; Panagiotopoulos, A. Z. Free Energy and Phase Equilibria for the Restricted Primitive Model of Ionic Fluids from Monte Carlo Simulations. *J. Comput. Phys.* **1994**, *101*, 1452–1459.

(18) Rossky, P.; Doll, J.; Friedman, H. Brownian Dynamics as Smart Monte Carlo Simulation. *J. Comput. Phys.* **1978**, *69*, 4628–4633.

(19) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*, 2nd ed.; Oxford University Press: Great Clarendon Street, Oxford, OX2 6DP, United Kingdom, 2017.

(20) Moučka, F.; Nezbeda, I.; Smith, W. R. Computationally Efficient Monte Carlo Simulations for Polarisable Models: Multi-Particle Move Method for Water and Aqueous Electrolytes. *Mol. Simul.* **2013**, *39*, 1125–1134.

(21) Peng, D.-Y.; Robinson, D. B. A New Two-Constant Equation of State. *Ind. Eng. Chem. Fundam.* **1976**, *15*, 59–64.

(22) Rahbari, A.; Hens, R.; Dubbeldam, D.; Vlugt, T. J. H. Improving the Accuracy of Computing Chemical Potentials in CFCMC Simulations. *Mol. Phys.* **2019**, *117*, 3493–3508.

(23) Pham, T. T.; Shirts, M. R. Identifying Low Variance Pathways for Free Energy Calculations of Molecular Transformations in Solution Phase. *J. Comput. Phys.* **2011**, *135*, 034114.

(24) Hens, R.; Vlugt, T. J. H. Molecular Simulation of Vapor-Liquid Equilibria using the Wolf Method for Electrostatic Interactions. *J. Chem. Eng. Data* **2018**, *63*, 1096–1102.

(25) Ewald, P. P. Die Berechnung Optischer und Elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, *369*, 253–287.

(26) Skeel, R. D. An Alternative Construction of the Ewald Sum. *Mol. Phys.* **2016**, *114*, 3166–3170.

(27) Vlugt, T. J. H.; García-Pérez, E.; Dubbeldam, D.; Ban, S.; Calero, S. Computing the Heat of Adsorption using Molecular Simulations: The Effect of Strong Coulombic Interactions. *J. Chem. Theory Comput.* **2008**, *4*, 1107–1118.

(28) Wolf, D.; Keblinski, P.; Phillpot, S. R.; Eggebrecht, J. Exact Method for the Simulation

of Coulombic Systems by Spherically Truncated, Pairwise r^{-1} Summation. *J. Comput. Phys.* **1999**, *110*, 8254–8282.

(29) Rahbari, A.; Hens, R.; Jamali, S. H.; Ramdin, M.; Dubbeldam, D.; Vlugt, T. J. H. Effect of Truncating Electrostatic Interactions on Predicting Thermodynamic Properties of Water-Methanol Systems. *Mol. Simul.* **2019**, *45*, 336–350.

(30) Fennell, C. J.; Gezelter, J. D. Is the Ewald Summation still Necessary? Pairwise Alternatives to the Accepted Standard for Long-Range Electrostatics. *J. Comput. Phys.* **2006**, *124*, 234104.

(31) Outotec, HSC Chemistry Software 9. `https://www.outotec.com/products-and-services/technologies/digital-solutions/hsc-chemistry/` (visited on 2/20/2020).

(32) Renner, T. In *Quantities, Units and Symbols in Physical Chemistry*; Cohen, E. R., Cvitas, T., Frey, J. G., Holström, B., Kuchitsu, K., Marquardt, R., Mills, I., Pavese, F., Quack, M., Stohner, J., Strauss, H. L., Takami, M., Thor, A. J., Eds.; The Royal Society of Chemistry, 2007; pp P001–232.

(33) Dortmund Data Bank, 2020, `www.ddbst.com` (visited on 3/17/2020).

(34) Briggs, J. M.; Nguyen, T. B.; Jorgensen, W. L. Monte Carlo Simulations of Liquid Acetic Acid and Methyl Acetate with the OPLS Potential Functions. *J. Phys. Chem.* **1991**, *95*, 3315–3322.

(35) Kamath, G.; Robinson, J.; Potoff, J. J. Application of TraPPE-UA Force Field for Determination of Vapor-Liquid Equilibria of Carboxylate Esters. *Fluid Phase Equilib.* **2006**, *240*, 46 – 55.

(36) Chen, B.; Potoff, J. J.; Siepmann, J. I. Monte Carlo Calculations for Alcohols and Their Mixtures with Alkanes. Transferable Potentials for Phase Equilibria. 5. United-Atom

Description of Primary, Secondary, and Tertiary Alcohols. *J. Phys. Chem. B* **2001**, *105*, 3093–3104.

(37) Kamath, G.; Cao, F.; Potoff, J. J. An Improved Force Field for the Prediction of the Vapor-Liquid Equilibria for Carboxylic Acids. *J. Phys. Chem. B* **2004**, *108*, 14130–14136.

(38) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Comput. Phys.* **2004**, *120*, 9665–9678.

(39) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Comput. Phys.* **1983**, *79*, 926–935.

(40) McQuarrie, D. A.; Simon, J. D. *Physical Chemistry: a Molecular Approach*, 1st ed.; University Science Books: Sausalito, California, 1997.

(41) Frisch, M. J. et al. Gaussian 09, Revision C.02. Gaussian Inc. Wallingford CT, 2016.

(42) NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101 Release 20, August 2019, Editor: Russell D. Johnson III, `http://cccbdb.nist.gov`.

(43) Chase, M. W.; Curnutt, J.; Prophet, H.; McDonald, R.; Syverud, A. JANAF Thermochemical Tables, 1975 Supplement. *J. Phys. Chem. Ref. Data.* **1975**, *4*, 1–176.

(44) Chase, M. W. NIST-JANAF Themochemical Tables, Fourth Edition. *J. Phys. Chem. Ref. Data* **1998**, *4*, 1–1951.

(45) Irikura, K. K. *Computational Thermochemistry*; 1998; Chapter 22, pp 402–418.

(46) Irikura, K. K.; Johnson, R. D.; Kacker, R. N. Uncertainties in Scaling Factors for Ab Initio Vibrational Frequencies. *J. Phys. Chem. A* **2005**, *109*, 8430–8437.

(47) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 Theory. *J. Chem. Phys.* **2007**, *126*, 084108.

(48) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) Theory for Molecules Containing First and Second-Row Atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

(49) Gmehling, J.; Kolbe, B.; Kleiber, M.; Rarey, J. *Chemical Thermodynamics for Process Simulation*, 1st ed.; Wiley-VCH Verlag & GmbH Co. KGaA.: Weinheim, Germany, 2012.

(50) Sandler, S. I. *Chemical, Biochemical, and Engineering Thermodynamics*, 4th ed.; John Wiley & Sons: Hoboken, N.J., USA, 2006.

(51) Smith, J. M.; Van Ness, H. C.; Abbott, M. M. *Introduction to Chemical Engineering Thermodynamics*, 7th ed.; McGraw-Hill: New York, USA, 2005.

(52) Te Velde, G.; Bickelhaupt, F.; Baerends, E.; Fonseca Guerra, C.; Dr. van Gisbergen, S.; Snijders, J.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.

(53) Wavefunction Inc., Irvine, CA. `www.wavefun.com` (visited on 2/20/2020).

(54) Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical Accuracy in Ab Initio Thermochemistry and Spectroscopy: Current Strategies and Future Challenges. *Theor. Chem. Acc.* **2012**, *131*, 1079.