Accelerated Simulations of Xylene Adsorption Using a Simplified Molecular Model and Advanced Monte Carlo Techniques

Zhao Li, Haoyuan Chen, Arun Gopalan, and Randall Q. Snurr\*

Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208, United States

# ABSTRACT

Xylene separation using Metal-Organic Frameworks (MOFs) were studied and analyzed using two different methodologies. In the first approach, a hybrid move combining energy-bias Monte Carlo and configuration bias Monte Carlo was implemented to speed up the fully atomistic simulations. The center of aromatic ring is inserted into energetically favorable sites using energy-bias Monte Carlo and the orientations are refined using the orientational bias scheme from configuration bias Monte Carlo. In the second simpler approach, three-site models of xylene isomers were developed and deployed to capture the packing and selectivity of xylene isomers in MOF pores with similar accuracy as fully atomistic simulations, but at a much-reduced computational cost. The final configurations of three-site models can then be used as excellent starting points for the fully atomistic model, resulting in significant speed up of the overall simulations.

# 1. INTRODUCTION

Xylene isomer separations are one of the most important technologies in chemical industries in the last century1. Para-xylene (pX) is the one with higher economical value because it can be synthesized into polyethylene terephthalate (PET), which is the building block for plastics1. Because of the similar physical properties of the xylene isomers, separating pX is extremely difficult. Current methods of separating xylene isomers include fractional crystallization, which relies on the fact that the isomers differ in freezing point. However, this method is not energetically efficient since low temperature separations often require large energy inputs. Zeolites are also applied for pX separation by means of selective adsorption. Among them, barium-X (BaX) is widely used. Lachet et al.2 used molecular simulations to investigate effects on selectivity of pX in a pX/mX binary mixture of different cations and zeolite X and Y. They discovered that because barium comparing to sodium cation is larger in size, and the steric effect destabilizes site II and guarantees pX selectivity. Recently, researchers discovered that MOFs, a class of porous crystalline materials, can separate xylene isomers with high purities. Numerous works have been performed to screen for structures and to generate mechanisms and design rules for selectivities3–31. One method leading to high selectivity is to use the MOFs as molecular sieves. In this path, MOFs that have specific pore diameter only allow the passage for pX, while blocking oX and mX. Wu et al.13 developed a MOF membrane, MIL-160, that can only allow pX to pass through while oX and mX cannot. They found that pX is adsorbed over oX because of stronger host-guest interactions and pX diffuses faster than mX because mX has a larger kinetic diameter and will cause steric hindrance. Another method is commensurate stacking in 1D channels proposed by Torres-Knoop et al.27 They made a vivid analogy that commensurate stacking was like “books on a bookshelf”32,33. They found that MAF-X827, a MOF with square shape one-dimensional channels, can separate pX from a mixture of isomers. The pX molecules stack face-to-face inside the MAF-X8 channels. Among all the mechanisms for high pX-selectivity, commensurate stacking seems to be the most promising because it guarantees both high pX loading and high pX selectivity simultaneously.

However, challenges arise when one needs to screen through databases for MOFs that can enable this mechanism. According to the hypothesis by Torres-Knoop et al.27,32,33, commensurate stacking requires liquid phase adsorption inside of 1D channel structures. For liquid phase adsorption, normal grand canonical Monte Carlo (GCMC) insertions are extremely inefficient. Works by Torres-Knoop et al.3,27 tried to overcome this issue by developing a new advanced Monte Carlo (MC) technique that can accelerate insertion of adsorbate molecules at high density. The new move, called configurational-bias/continuous fractional component3,34 (CB/CFC) MC, combines the merits of both configurational-bias35 (CBMC), which tries to probe different positions and orientations of inserted molecule, and continuous fractional component36 (CFCMC), which tries to insert a molecule continuously like inflating a balloon. They found that this hybrid CB/CFC move can accelerate the GCMC simulation of xylene adsorption in MTW zeolite comparing to GCMC simulations which only used CBMC or CFCMC.

Recently, Zhang et al. 37 used energy biasing Monte Carlo (EBMC) for water adsorption in ZIF-8 and observed speed-up over normal GCMC simulations. In their work, the interaction energies of water with the framework is used to bias the insertion of water molecules in energy favorable sites within ZIF-8. In the light of the work by Zhang et al.37, we propose a new MC move to bias the insertions of xylene molecules at energetically favorable sites inside MOFs. EBMC combined with the CB/CFC move developed by Torres-Knoop et al.3 showed great potential of accelerating the systems that we studied. At high density, when xylene molecules pack, the positions and orientations of the molecules are highly localized. Thus, accelerations can be obtained when we bias the insertions at these localized positions. We found that simulations are much faster when energy-biased while studying xylene adsorption in MTW zeolite.

In contrast to accelerating simulations using state-of-the-art MC moves for GCMC, which is complicated and requires mathematical derivations, a simpler but equally powerful path is coarse graining. Thus, we propose a coarse-grained model for xylene isomers that captures their essential characteristics. We have observed significant speed up when the coarse-grained model was applied in MTW zeolite.

The origins of the two methods differ drastically. EBMC emerges from chemical intuition and mathematical derivations, while the coarse-grained model comes from a critical thinking and thought experiments. However, from a philosophical standpoint, the two methods share a common origin. They are the two branches that prospers from the same stem38. In the following sections, we will explain this from a statistical mechanical point of view.

# 2. METHODOLOGY

## 2.1 ENERGY-BIAS MONTE CARLO

We start from the grand canonical partition function, using fractional coordinates39

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

where represents the temperature of the system, represents the volume of the system, *N* represents number of particles, , which is the thermal de Broglie wavelength is the mass of the particle, is the Boltzmann’s constant, is the chemical potential, , and is the configurational space.

Probability density for a configuration with N particles, using convention of Allen and Tildesley39, is

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

Probability density for a configuration with (N+1) particles is

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

Acceptance rule of insertion move can be derived from the ratio of the two probability densities:

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

By using fugacity , we can get:

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

Similarly, we can also get the acceptance rule for deletion move:

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

In EB insertion moves, we bias the insertion towards energy favorable sites. We first generate an energy grid with equal spacing. Each grid point is the center of a small cubelet. Each cubelet is assigned a weight based on the energy calculated at its grid point. The weights are assigned using a Boltzmann distribution.

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

During an insertion, we first randomly choose a cubelet based on the Boltzmann’s weights, then the molecule can be placed anywhere inside of the cubelet.

The probability of attempting such an insertion move is:

|  |  |  |
| --- | --- | --- |
|  |  | (8) |

To satisfy detailed balance, we need to consider the probability of attempting a deletion move, since the deletion move is not biased:

|  |  |  |
| --- | --- | --- |
|  |  | (9) |

Thus, the acceptance rule for insertion can be written as:

|  |  |  |
| --- | --- | --- |
|  |  | (10) |

Similarly, the acceptance rule for deletion move is:

|  |  |  |
| --- | --- | --- |
|  |  | (11) |

In RASPA40, CBMC35 is implemented and deeply coupled to every move, for example, re-insertion and identity swap move to boost the efficiency of the move. In CBMC, the algorithm grows the molecule by first placing its first bead. In RASPA, this is done by randomly selecting 10 points in the simulation box, compute the energies of the first bead at these locations, assigning a Boltzmann’s weight and choose from them. To make this process more efficient, we implement an algorithm that combines CBMC with EBMC. Now the insertion of the first bead is energy biased.

In normal CBMC, from the paper by Esselink et al.41, when placing the first bead, the total Rosenbluth weight of the first bead is:

|  |  |  |
| --- | --- | --- |
|  |  | (12) |

and is the number of trial positions for placing the first bead. In RASPA, is set to 10 by default. So,

|  |  |  |
| --- | --- | --- |
|  |  | (13) |

Probability of one of the positions being selected is following a Boltzmann’s distribution:

|  |  |  |
| --- | --- | --- |
|  |  | (14) |

The Rosenbluth weight of growing the rest of the molecule is , so the Rosenbluth weight of growing the whole molecule is:

|  |  |  |
| --- | --- | --- |
|  |  | (15) |

When energy biasing is combined with CBMC, instead of having 10 trial positions, the algorithm randomly selects 1 point from all the grid points. Each grid point is the center of a cubelet, so the point to place the first bead is randomly selected within the volume of the cubelet. Each grid has a weight, . To maintain a detailed balance, from Equation (10), we replaced the Rosenbluth weight of the first bead from original CBMC with the part from Equation (10) that accounts for the energy biasing. The Rosenbluth weight of the new EB-CBMC is:

|  |  |  |
| --- | --- | --- |
|  |  | (16) |

Growing the rest of the molecule will be the same as the normal CBMC algorithm, so the total Rosenbluth weight is:

|  |  |  |
| --- | --- | --- |
|  |  | (17) |

For MC moves that requires the Rosenbluth weight of the old configuration, such as deletion, identity change, and reinsertion, we just need to get . We first determine the cubelet to which the first bead of the old configuration belongs, then we calculate using the energy at that grid point pre-tabulated in the energy grid. Finally, the Rosenbluth weight of the old molecule is calculated as where .

For GCMC simulations using CBMC, the acceptance rules of insertions and deletions are:

|  |  |  |
| --- | --- | --- |
|  |  | (18) |
|  |  | (19) |

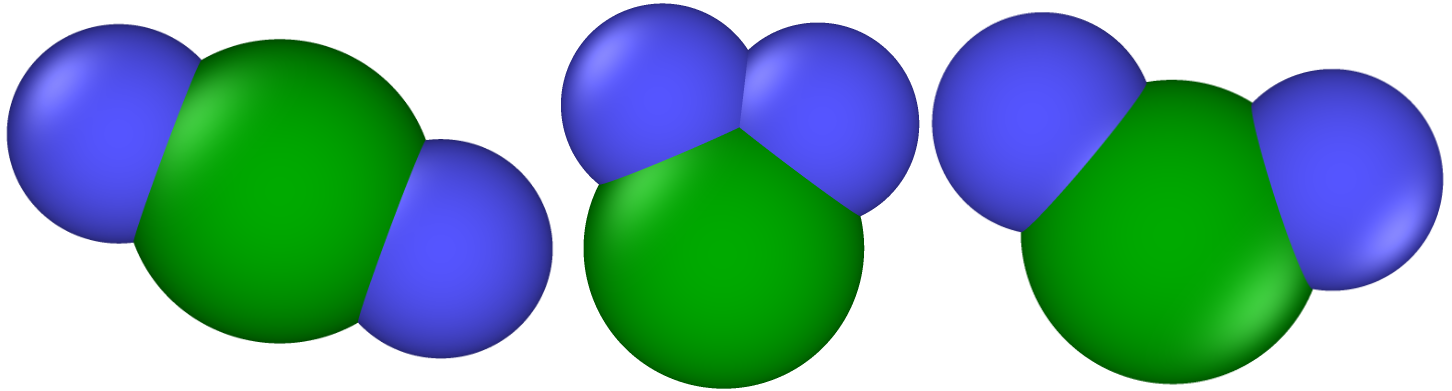
where is the averaged ideal Rosenbluth weight of the molecule. For a rigid molecule, the ideal Rosenbluth weight is set to 1.42–44

Later in this article, more swap moves will be used (CFCMC and CB/CFC), we are calling their EB counterparts EB-CFCMC and EB-CB/CFC. When energy-bias MC is mentioned in general, we will use EBMC. The new move combines the merits of both EBMC and CBMC. EBMC will try to insert the first bead in an energy-favorable location and then CBMC will sample different orientations for the molecule. Because of this, EB-CBMC has a higher chance of inserting molecules at the correct location than trying random sites in original CBMC. To implement this EBMC algorithm, we tailored the xylene molecules to incorporate fictional, approximate benzene (AB) site. Please refer to Figure S1 for the implementation of the xylene model. We made approximate benzene the first bead of the xylene isomers. The AB pseudo-atom is used to for energy-biasing and determines where the ring center is placed. The interaction of the AB pseudo-atom is ignored when computing the energies.

The grids were generated using the LJ parameters of AB pseudo-atoms. However, we modified the source code of RASPA40 so that the energy of the AB is not accounted when calculating the potential energies. The AB energy grid is only used for biasing. The workflow of using EBMC in RASPA40 is summarized in Figure S2.

## 2.2 THE THREE-SITE XYLENE (TSX) MODEL

From hypothesis of Torres-Knoop et al.32,45, it can be implied, that commensurate stacking is promoting , face-to-face interactions between planar xylene molecules. Although we can treat benzene-ring molecules as discs, for xylene molecules the situation is more complicated. A xylene molecule is composed of an aromatic plane and two methyl groups. A question can be proposed: Which part of a xylene molecule has more impact on xylene selectivity? If we can proof that a part of the xylene molecule model is not essential towards its selectivity, by the law of the razor of Occam, that part can be reduced or simplified. To put our thought to test, we proposed xylene models that tried to emphasize the methyl groups rather than the benzene ring. The new three-site xylene (TSX) models composed of two united methyl groups and a central Lennard-Jones site of approximate benzene38. The Lennard-Jones parameters of Approximate Benzene were taken from Snurr et al38. We ignored the partial charges on the pseudo-atoms of the TSX models because electrostatics calculations will slow the simulations down and that the close packing at saturation is determined primarily by the Van der Waals interactions. The parameters for pseudo-atoms of the three-site models were summarized in Table S1. Figure 1 shows the examples of TSX models. Because the TSX isomers have no planar characteristics, it can test whether methyl groups or aromatic rings are more important. Thus, if we can reproduce commensurate stacking with three-site models, we can say that methyl groups play a more important role, and *vice versa*.



**Figure 1.** Illustrations of 3-site models for para-, ortho- and meta-xylene. The central green pseudo-atom is approximate benzene, the two blue pseudo-atoms are methyl groups.

## 2.3 MOLECULAR SIMULATION DETAILS

To validate our TSX model and EBMC method, we first conducted simulations in MTW zeolite and MOF MIL-47 because they are studied in previous literatures3,19. We then applied the TSX model and EBMC on all 1D channels structures in the MOF database generated by ToBaCCo 1.046. Channel dimensionalities of the structures in ToBaCCo 1.046 were calculated using Zeo++47. We chose three probe radii for this calculation. The helium probe (σ=2.64 Å48) yielded the least number of 1D channel MOFs, while the AB probe (σ=5.27 Å38) gave the most and the methyl probe (σ=3.80 Å49) in between. To maximize our selection of 1D channel MOFs, we chose to consider all MOFs that were classified as 1D-channel-only structures by either one of the probe radii. This eventually yielded 640 structures. A full list of the structures can be found in the Supplemental Materials. We realized that the 640 selected MOFs cannot be rigorously defined as 1D channel MOFs, for example, MOF #199 from Figure S9. However, we can see from the future sections that the dimensionality of the channels does not have an effect of selectivity. The textural properties histograms of the full ToBaCCo 1.046 database and the 640 selected MOFs are shown in Figure S3. We see that the distributions of the selected structures are narrower than the distributions of the full database. Also, the selected structures generally have a smaller gravimetric surface area. Based on our assumption of the effect of 1D channels and steric effect on selectivity, as the size of the pores increase within a structure, the selectivity will be less because it is harder for the xylene isomers to form patterns and xylene mixtures will act towards the bulk phase. It then justifies our purpose of using Zeo++47 for 1D channel selections: it helps us narrow down our scope and eliminate structures that are not suitable for our application. For reference, we included the names, linkers information, nodes information, and the MOFkeys. From the work of Bucior et al.50, we realized that a few ToBaCCo 1.046 MOFs cannot yield correct SMILES string51 and will cause inconsistencies when processed by the MOF-ID code.50 For example, MOF 2974 contains nodes with cobalt that will form additional Co-Co bond50 in the process of identifying the MOF. Thus, we included the node and linker information from the ToBaCCo 1.0 paper46 to help readers better locate the CIF files.

Regarding the forcefield used for ToBaCCo 1.046 MOFs, the LJ interactions of framework atoms were modeled using the Universal Force Field (UFF), and we ignored partial charges from the framework. For zeolites, the LJ parameters were from the DREIDING forcefield (DFF). For MIL-47, we chose to use DFF except for the vanadium atom (UFF was used). For zeolite MTW and MIL-47, since the partial charges are provided, we used the partial charges given in the CIF files. The partial charges are listed in Table S1. The frameworks atoms were held fixed during the GCMC simulations. Lorentz-Berthelot mixing rule was used for cross-interactions. A cutoff of 12 Å was applied.

The parameters for the TSX isomers were shown in Table S1. The AB pseudo-atoms were placed in the center of the aromatic rings of the xylene isomers. The TSX molecules were held rigid during the GCMC simulations. Full model xylenes are modeled with OPLS49 and contains partial charges49. The parameters are summarized in Table S1.

GCMC simulations were performed using CB/CFC at 433 K and 103 bar fugacity. For mixture simulations, the bulk phase fugacity of pX, oX, and mX were set to be equal. We used 30,000 equilibration + 30,000 production + 5,000 initialization cycles for the TSX simulations. In RASPA40, each cycle contains steps, where equals to the maximum of 20 and the number of adsorbate molecules in the simulation box. We used translation, rotation, reinsertion, swap move, CFC swap move, CB/CFC swap move, and identity swap move. The moves were set to have equal probabilities. A swap move is made of insertion and deletion moves executed with equal probabilities. For a CFC and CB/CFC move, a fractional molecule is used for each species. Each fractional molecule has a parameter , ranges from 0 to 1, that controls its interaction strength with other atoms. The modified VdW and electrostatic interaction is given by:

|  |  |  |
| --- | --- | --- |
|  |  | (20) |
|  |  | (21) |

A CFC or CBCFC swap move randomly perturbs and if , the move makes the original fractional molecule whole and inserts a new fractional molecule with if < 0, the moves delete the original fractional molecule and makes a randomly selected molecule fractional. More details about a CB/CFC swap move can be found in Torres-Knoop et al3.

EB-CB/CFC simulations were performed at the same conditions. We generated energy grid with a spacing of 0.1Å for AB pseudo atom. We also generated and used energy grids at the same grid spacing for pseudo-atoms of xylene: carbon on the ring (C\_xyl), hydrogen on the ring (H\_xyl) and methyl group (CH3\_xyl) to accelerate the simulation. In RASPA40, a grid file of the interaction between a pseudo-atom with a framework includes the local potential energy , and its first, second, and third derivatives at the grid points3,52. For a GCMC that uses grids, RASPA40 applies cubic interpolation52 to approximate the interaction energy between the pseudo-atom with the framework at locations between grid points. EB-CB/CFC simulations used 50K initialization + 200K equilibration and 200K production cycles. When reporting the snapshots, loadings, and selectivities, we only report the loadings of the full molecules. The selectivities of pX are calculated as the ratio between GCMC loadings of pX against the sum of oX and mX loadings:

|  |  |  |
| --- | --- | --- |
|  |  | (22) |

# 3. RESULTS AND DISCUSSIONS

## 3.1 EBMC VALIDATION

To test EBMC simulation on a system that has known results, we chose MIL-4753–57. MIL-47 is constructed by vanadium nodes and di-carboxylate linkers. The structure forms 1D channels with a diamond shape along its x axis and is known to have a high oX-selectivity. Recently, Castillo et al.19 successfully predicted the selectivity of xylene isomers in MIL-47 and discovered that the packing effects accounts for the oX selectivity. oX and pX can stack 8 molecules within a channel while mX can only have in 4.

However, instead of starting with xylene isomer systems, which is hard, we conducted ethane adsorption isotherm in MOF MIL-47. Ethane was modeled as a rigid molecule using TraPPE58 force field. The ethane forcefield parameters are shown in Table S1. Our EB-CBMC algorithm works with the TraPPE58 ethane because ethane was defined as a rigid molecule in TraPPE58. Similar to approximate benzene for EBMC calculation for xylene molecules, we added a fictional approximate ethane (AE) bead in the middle of the bond of ethane molecule as the first bead for the EB-CBMC simulation. The LJ parameters of AE are shown below, calculated from empirical relation between the critical constants of ethane59 and the estimated LJ parameters60,61.

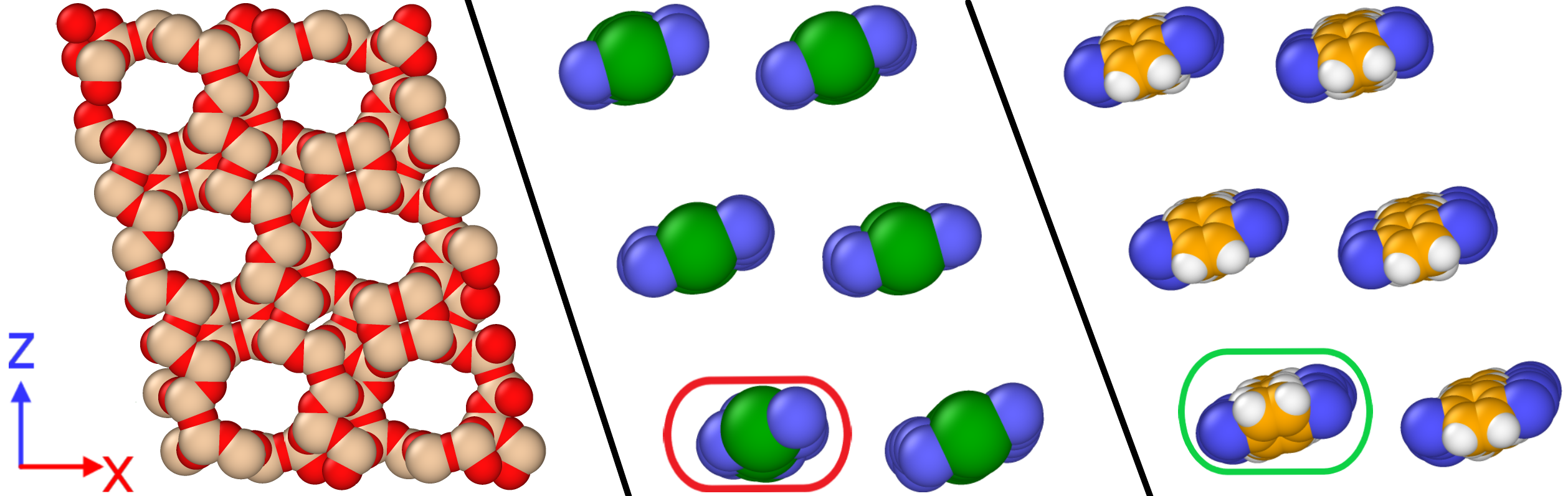
|  |  |  |  |
| --- | --- | --- | --- |
|  |  | | (23) |
|  | |  | (24) |

The resulting comparison of isotherms generated by normal GCMC with CBMC and EB-CBMC is shown in Figure S4. The results show that EB-CBMC simulation (orange) can generate the correct isotherm as the normal CBMC simulation (blue). It also shows that the error bars of EBMC simulations at higher pressures (100 Bar, 101 Bar, 102 Bar) are larger than those of CBMC. This indicates that EBMC increased the insertion and deletion success rate, thus creating larger fluctuations for sampling. This validates the conclusion of Snurr et al.38 that EBMC increases the acceptance rate of insertions and deletions.

After the successful validation of ethane adsorption, we performed equimolar xylene isomer (pX/oX/mX) adsorption isotherm at 343 K in MIL-47. Peng-Robinson62 equation of state was used to convert pressures into fugacities. We used CB/CFC swap, CFC swap and CBMC swap moves for the GCMC simulations and their EB counterparts for the EBMC simulations. The isotherm comparison can be found in Figure S5. We can barely distinguish the EB-CB/CFC simulations from the CB/CFC simulations. This shows our EBMC implementation can generate reliable results.

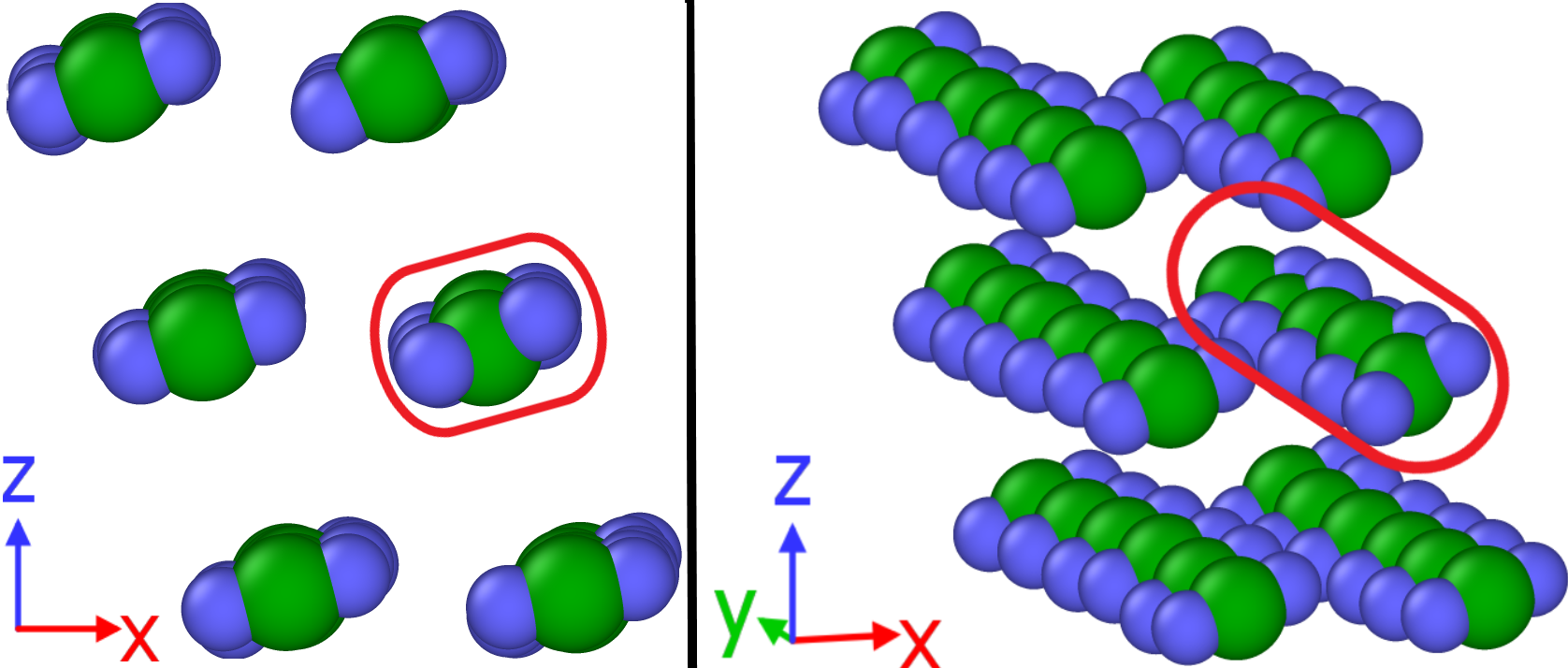
## 3.2 MTW COMPARISON

Our next testing structure is the MTW zeolite. We simulated pX using the TSX model in MTW at 433 K and 103 Bar fugacity. For comparison, we simulated the same system using the full model for pX and EB-CB/CFC. The comparison snapshot is shown in Figure 2. It indicates that the TSX can reproduce the orientations of pX molecules in the full model simulation with a few defects. However, the channel with wrong-oriented pX molecules (red-circled in Figure 2, middle) will never reach its saturation loading, which is six pX molecules per channel. Another thing we need to point out is that for the full model, although the packing or pX is extremely tight, there is one channel with the opposite orientation from the rest (green-circled in Figure 2, right). This indicates that there is a rotational degree of freedom (RDOF) for the pX molecules inside the channels around the methyl axis. Although the selection of orientation of a channel is random and the two orientations are equivalent, during a full-model simulation, this RDOF requires the simulation to take longer time and more cycles to turn all pX molecules in one channel to have the same orientation. The TSX model ignores this RDOF by representing the benzene ring with an AB pseudo-atom. This allows the simulation to converge faster because three-site model simulations will only need to equilibrate the methyl group orientations.



**Figure 2.** Left: Empty MTW framework (red: oxygen, gold: silicon). Middle: snapshot of TSX model pX in MTW at 433 K and 103 Bar fugacity, simulated using CB/CFC. Right: snapshot of full model pX within MTW at the same condition, simulated using EB-CB/CFC. For the middle and right plot, the structure was deleted to show the pX molecules clearly. All the sub-snapshots share the same point of view.

We then tested equimolar mixture of pX/oX/mX simulation for MTW zeolite at the same condition (433 K, 103 Bar fugacity) using CB/CFC for TSX model and EB-CB/CFC for full model. The full model snapshot is shown in Figure S6. The full model simulation snapshot yielded invidivual loadings for pX/oX/mX = 33/1/0, which corresponds to 1.09/0.03/0.0 mol/kg. The full-model averaged loadings for pX/oX/mX were 1.136/0.022/0.0 with error bars of 0.03/0.022/0.028 mol/kg. The TSX resulting snapshots are shown in Figure 3. The individual loadings of pX/oX/mX = 29/0/5 molecules, which corresponds to 0.99/0.0/0.17 mol/kg. The TSX model averaged loadings for pX/oX/mX were 0.984/0.0002/0.197 with error bars of 0.001/0.0003/0.0001 mol/kg. Because of the fractional molecule introduced by CB/CFC algorithm, the fluctuation caused a channel of pX (the channel above the circled one) has only 5 molecules. There is also one channel, circled in Figure 3, with mX stacking, thus reduced the pX loading and reduced pX-selectivity. However, while the other channels filled with 6 pX molecules, it is hard for the channel with mX molecules to fill up. This is due to the difference in positions of methyl groups in pX and mX. If mX wants its methyl groups in the favorable positions, it needs to tilt its orientation a little bit, which eventually takes up the space for a sixth molecule.

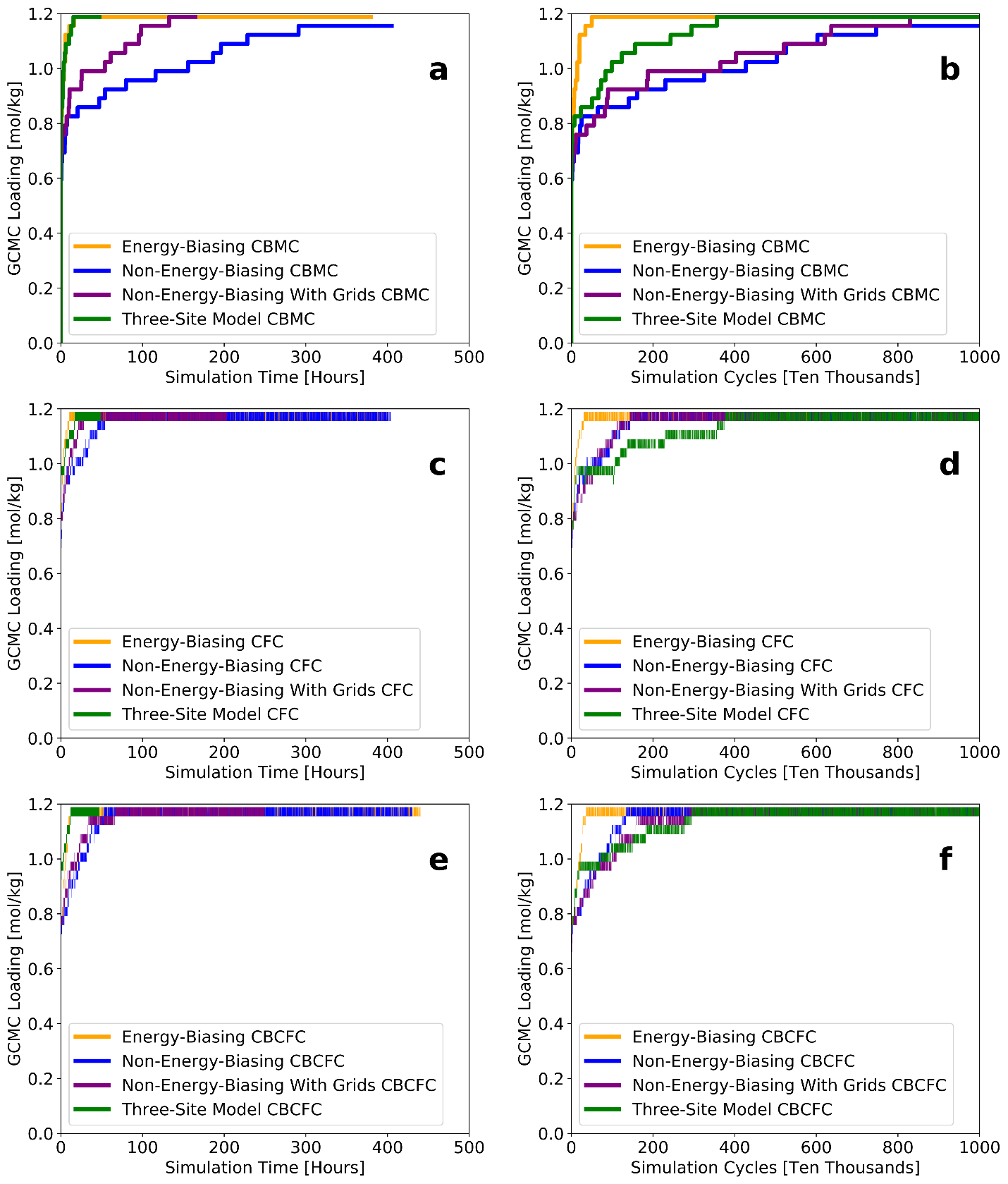


**Figure 3.** Equimolar xylene isomers adsorption using the TSX model at 433 K and 103 Bar fugacity in MTW zeolite. To show the packing directly, we deleted the zeolite atoms from the snapshot. For the figure on the right, we tilted the box to show the packing style of all the molecules.

We ran a further test on zeolite JSW. This is a zeolite with narrow (largest cavity diameter = 5.38 Å63) 1D channels (Figure S7, left). In mixture simulation, the TSX model yielded pX selectivity (Figure S7, right) with averaged individual loadings for pX/oX/mX = 1.38/0.0/0.0 with error bars equal 0.005/0.00008/0.0015 mol/kg. The results for EB-CB/CFC in Figure S7 and the fact that the averaged individual loadings for pX/oX/mX = 1.361/0.0/0.0 with error bars equal 0.0003/0.0/0.00006 mol/kg show that the selectivity was accurately reproduced. It also displays the same orientations as those in the full model simulation.

## 3.2 SPEED COMPARISON

We also tested the convergence speed of full model simulations in terms of time and cycles and made comparison with the TSX model. We still took pX in MTW zeolite at 433 K and 103 Bar fugacity as the example. pX adsorption in MTW zeolite is a perfect example to test convergence speed for new methods. We ran 10,000,000 cycles for the full model. We chose to simulate the system at three “levels of method” in terms of moves used. The results are shown in Figure 4. The first set of simulations only used CBMC swap move (Figure 4a and b). The second one used CFC swap move in addition to CBMC swap move (Figure 4c and d). The last set used CBCFC swap move on top of the moves mentioned before (Figure 4e and f). We also simulated the EBMC moves at the different levels of method. Finally, the TSX simulations without EBMC at different level of theories were performed at the same conditions and same number of cycles.



**Figure 4.** Time and cycle comparison between simulations using the TSX model and the full model simulated at 433 K and 103 Bar fugacity. First row (a and b): CBMC simulations. Second row (c and d): CFC simulations. Right (e and f): CB/CFC simulations. Orange line: Energy Biased simulations using full model. Blue line: Non-energy biased simulations using full model. Purple line: Non-energy biased simulations with grids of full model xylene pseudo atoms. Green line: TSX simulations. TSX simulations are not energy biased.

From all three sets of comparisons between the full model results, we see that CFC (Figure 4c and d) and CB/CFC (Figure 4e and f) are the fastest among all methods. The CFC and CB/CFC simulations all converged within 100 hours. In terms of cycles, CFC and CB/CFC simulations, for the full model, converged within 2,000,000 cycles. CBMC (Figure 6a and 6b) sets of simulations are much slower. This verifies the speed relationship that CB/CFC ~ CFC >> CBMC.

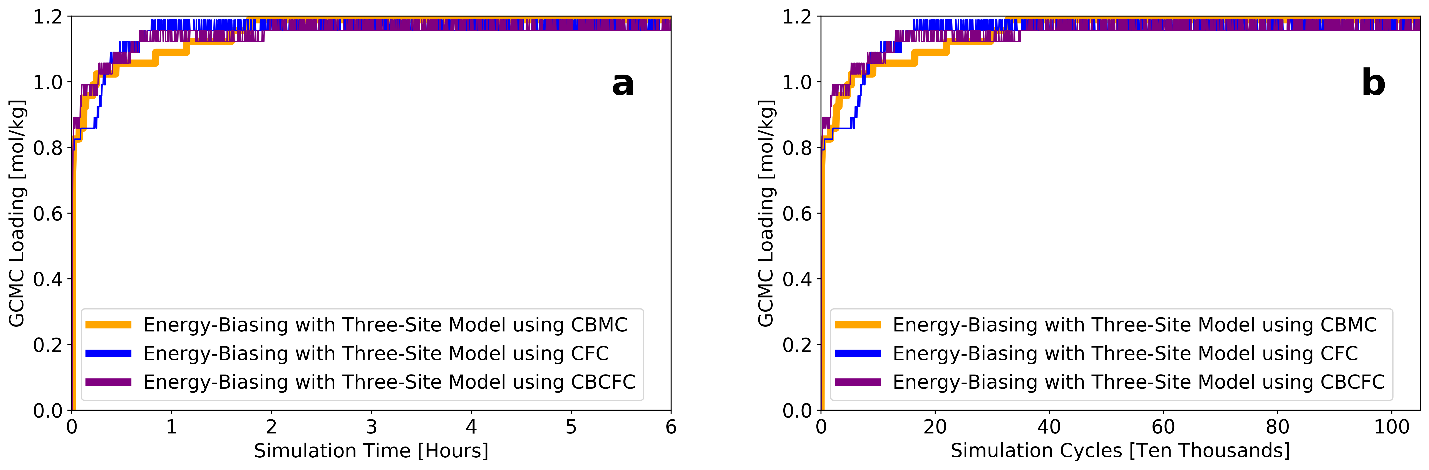
For a given level of method with full model pX, take CFC (Figure 4c and d) as the example, EB-CFC (orange line) is the fastest both in terms of cycles and time. CFC with grids (purple line) finishes two times faster than CFC without grids (blue line). Thus, for full model simulations, we see the following speed hierarchy: EBMC with grids >> GCMC with grids but not energy-biased >> GCMC without grids and no energy bias.

Now that we compared all simulations for the full model, we can add the three-site model results into the comparison. We see that TSX simulations for CFC and CB/CFC converged within 40 hours. For a given level of theory, take CFC as example, TSX result (green line) is slower than EB-CFC with full model (orange line) in terms of time, faster than non-energy-biased CFC simulations using the full model.

Although we cannot fairly compare simulations with different models, from a practical point of view, TSX is slower than EBMC with full model. However, full-model-EBMC requires energy grids at sufficient accuracies, which is 0.1 Å spacing for our case, while TSX has no storage or memory requirement. The storage requirement for EBMC for this system is 4.9 gigabytes and the EBMC simulation requires about 15G memory to store everything during the simulation. TSX is still much faster than normal GCMC simulation, either with or without grids. It converges much faster because of the reduction of LJ sites and no electrostatic interactions.

Every simulation reaches 0.8 mol/kg immediately. However, there seems to be a huge energy barrier at 1.0 mol/kg. At 1.0 mol/kg, which correspond to 30 molecules in the simulation cell, there are on average 5 out of 6 channels filled up. MC algorithm will not fill channels one by one; all channels are filled in simultaneously. At 2M cycles for CFC (Figure 4d) and CB/CFC (Figure 4f) , the TSX model reached 1.0 mol/kg (Figure S8, left). At this point, only one out of the six channels are filled with 6 pX molecules. We see that although the methyl groups are still in the favorable regions of the channels, the axial positions are not aligned. Same situation can be observed from the snapshot of the full model (Figure S8, right). Thus, we can conclude that the TSX model converge to the correct loading with a shorter time because there are less pairwise calculations.

Finally, a well-educated question arises: what if we combine the three-site model with EBMC? This time, we simulated the same system with the EB-CBMC, EB-CFC, and EB-CB/CFC, running 1,050,000 cycles. The grids for methyl groups and central benzene pseudo-atom of the three-site model are also provided. The results are shown in Figure 5. The results indicate that all the three EBMC+TSX simulations converged within 2 hours. In terms of cycles, EBMC+TSX simulation converged within 400,000 cycles, which is much faster than their non-energy-biasing counterparts in **Figure 4**, green lines. The synergy of an efficient MC method and a simplified xylene model leads to orders of magnitudes acceleration comparing to the normal GCMC simulations using full model.



**Figure 5**. Time and cycle versus loading for simulations using the TSX model with EBMC simulated at 433 K and 103 Bar fugacity.

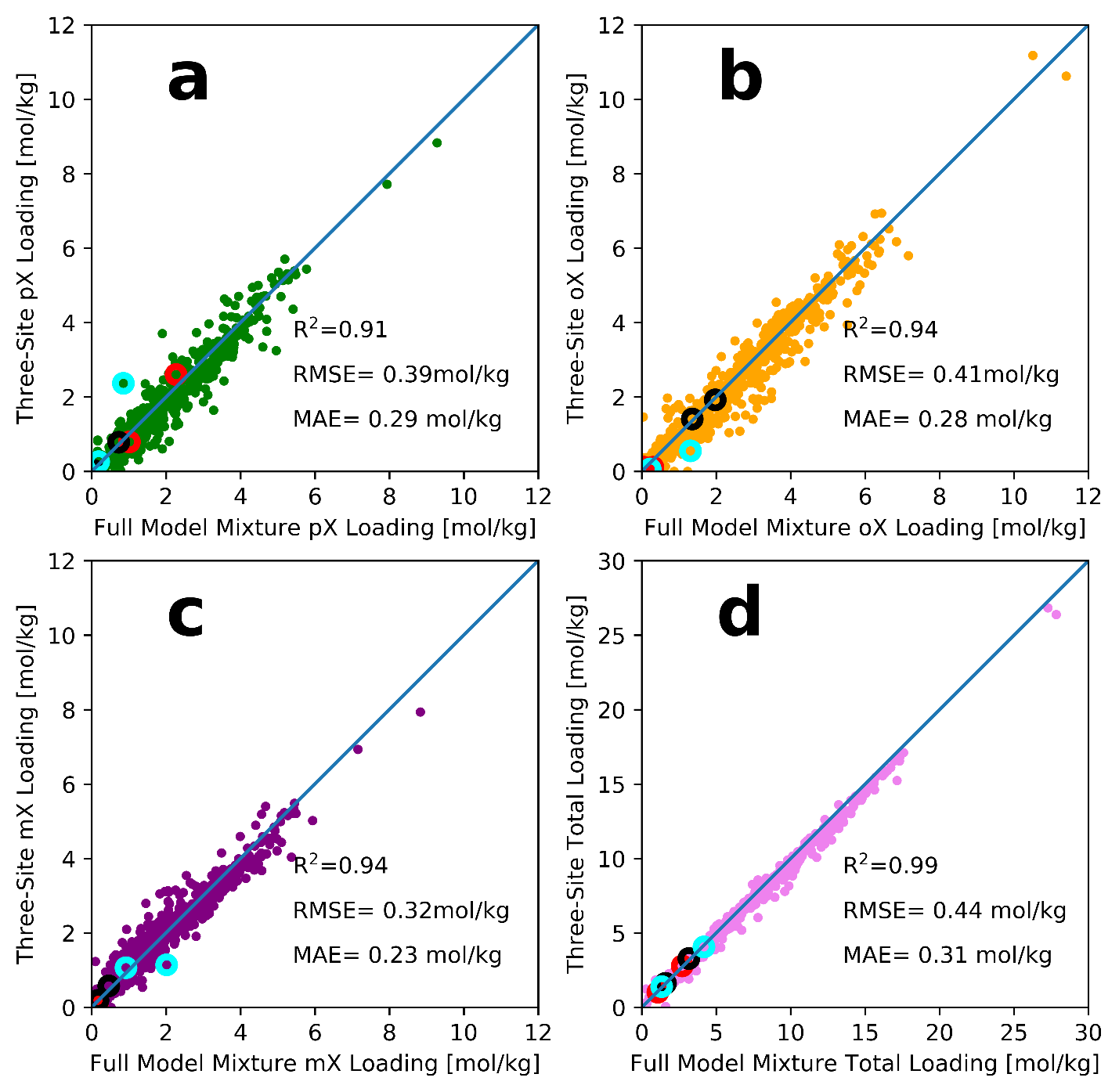
From this thorough and practical comparison, we see that EBMC is the fastest for full model simulations. However, EBMC has huge storage and memory requirement. In optimization problems, we always encounter the “space-time tradeoff”: one cannot save time and space simultaneously for a task. EBMC is a perfect example of this tradeoff. Comparing to EBMC, TSX can avoid it by simplifying the simulation. Thus, the TSX model gives the most economical, though not the fastest, solution while EB-CB/CFC gives the fastest.

## 3.3 OVERALL PERFORMANCE OF TSX

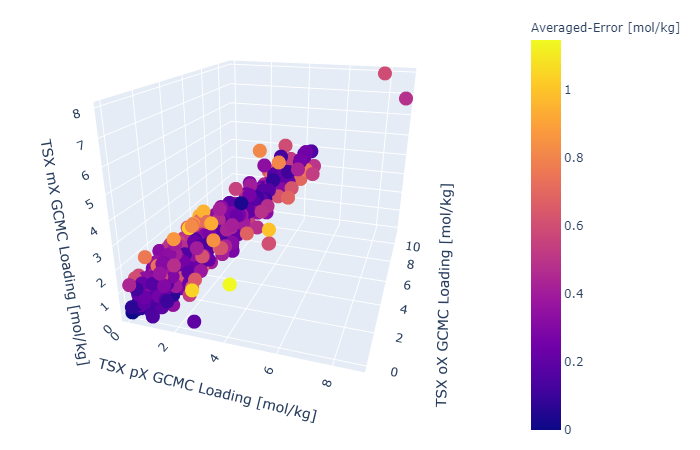
Finally, we applied the TSX model for all 640 MOFs selected from ToBaCCo 1.046 database. We compared the individual loadings of the isomers and the total loadings from the TSX models simulated using CB/CFC to the GCMC results generated by EB-CB/CFC using the full model. The results are shown in Figure 6. We also highlighted two selective structures for each species. A table summarizing the results of the highlighted structures are shown in Table S2. The results indicate that the TSX model can predict full model individual loadings of isomers with high accuracies. The TSX model can get the saturation loading or capacity with extremely high accuracy. The highlighted structures also display high fidelity for individual loadings (Figure 6a, b, and c) and very high fidelity for total loadings (Figure 6d).

We also tried to display the errors of individual points and formed a 3D plot in Figure 7. We calculated the absolute errors between the three-site model loading and the full model loading for each point and averaged across the species of the isomers. We discovered that although some points are having high errors (greater than 1 mol/kg, and look yellow on the graph), the rest of the data set still has small errors.

|  |  |  |
| --- | --- | --- |
|  |  | (25) |



**Figure 6.** Parity plots of loadings of equimolar mixture of xylene isomers (a: pX, b: oX, c: mX) and the total loading (d) at 433 K and 103 Bar fugacity in 640 ToBaCCo 1.0 MOFs with 1D channels with the three-site models using CB/CFC and the full models using EB-CB/CFC. We chose to highlight two pX-selective (red circles), two oX-selective (black circles), and two mX-selective (cyan circles) structures.



**Figure 7.** 3D plot of the individual loadings of pX/oX/mX of the three-site model. The color bar represents the average error of the three-site model individual loadings against the full model individual loadings calculated by Equation (25).

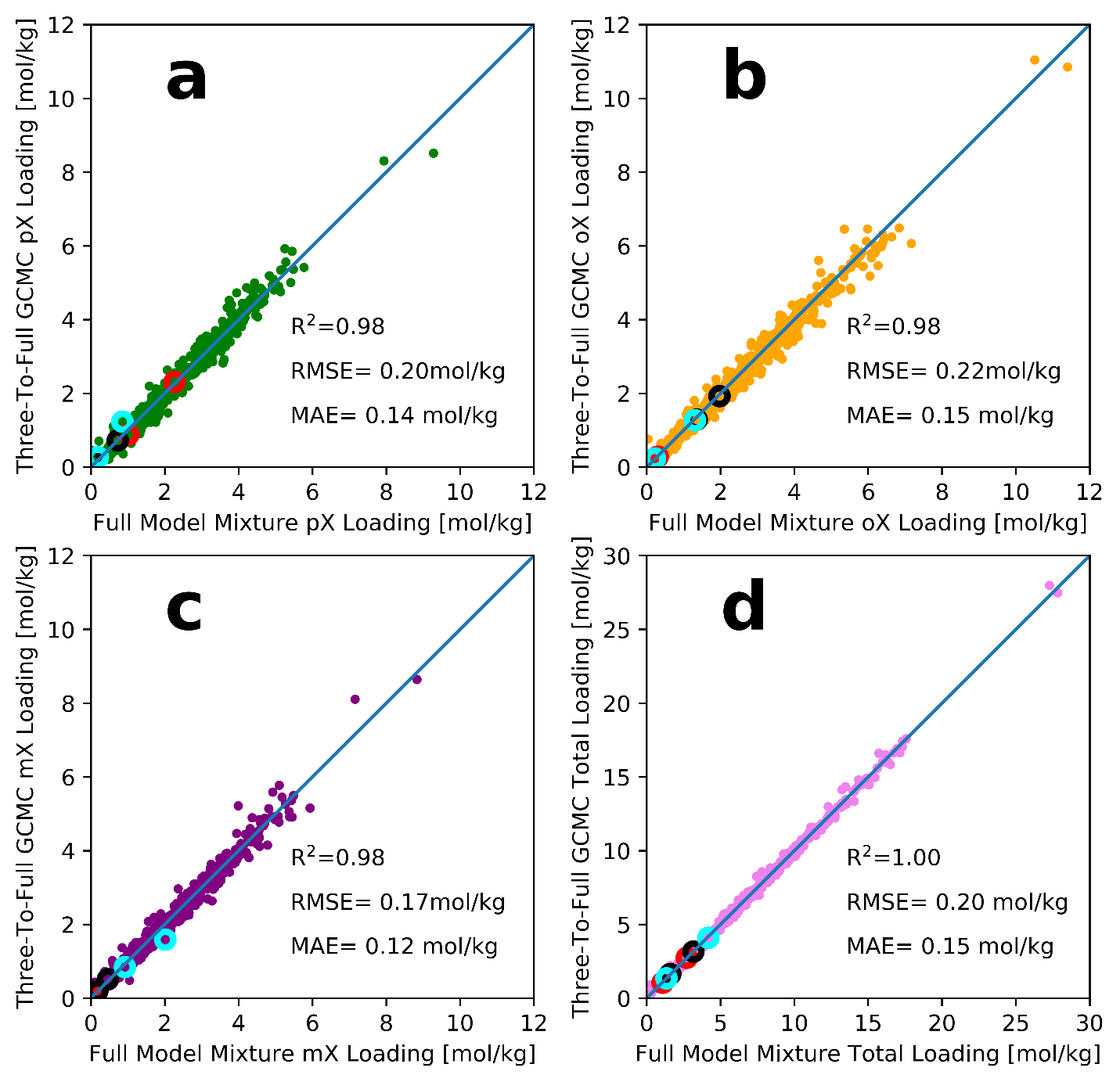
We selected to demonstrate a comparison of results in example MOFs using full model and the three-site model. MOF 199 has interconnected channels in three directions. It was identified as 1D channel MOF by the methyl probe. MOF 199 enhances perpendicular stacking of pX inside of the channels. Figure S9 shows that three-site model reproduced the results of the full-model in terms of stacking pattern, loading and selectivity. MOF 8995 allows for parallel packing inside of the 1D channels (Figure S10). The TSX model reproduced the pattern and xylene loading. However, we saw defects in the packing of the three-site model (circled in red). There was a mX molecule. However, the channel with mX cannot fit a sixth molecule. Another parallel stacking MOF, 9273, also displayed good agreement between results from the three-site model and the full model (Figure S11). The small channels in MOF 9273 provides accurate space and landscape for pX to form parallel stacking, while the big channels do not exhibit a pX-selectivity because of its big size.

In Figure S12, we included the comparison between three-site selectivities versus full model selectivities. The results indicate that there is a considerable number of points that the selectivites do not agree numerically. A possible reason is that the three-site model lacks the planar shape of the full model. For pX molecules, the three-site model has one less rotational degree of freedom around the methyl-methyl axis than the full model. Also, the approximate benzene pseudo-atom () is thicker than the benzene ring (thickness of benzene ring is the diameter of the carbon atom on the benzene ring, ). Take MOF 4721 as an example. From Figure S13, which are the snapshots of the simulation using the full model, we see that the pyrine linker (starred region of Figure S13) of this structure can fit a xylene ring and prefers pX. The stacked pX molecule then can stabilize the pX molecules next to it, generating interactions between benzene ring and the methyl groups. Figure S14 summarized the calculated probability of pX in the regions of MOF 4721 from snapshots. This also validates that the parallel slit has the highest selectivity for pX. We did not observe this phenomenon using the TSX model because the size of spherical benzene is too big comparing to the size of the slit. Another reason is the calculation of selectivity. We see from the equation 22 that selectivity is defined as a ratio of pX loading and the sum of oX and mX loading. This indicates that highly pX-selective structures can have a highly sensitive selectivity. We take MOF 199 as an example. The TSX model yielded a result of pX/oX/mX = 2.51/0.19/0.03 with error bars of 0.039/0.025/0.032 while the full model had 2.27/0.29/0.15 with error bars of 0.018/0.028/0.021, all in mol/kg. We see that the three-site model correctly predicted that 199 is a highly pX-selective structure, however, the selectivities differ by 100%. This is because the mX + oX loadings for the three-site model (0.217 mol/kg), although small, is half of that for the full model (0.448 mol/kg).

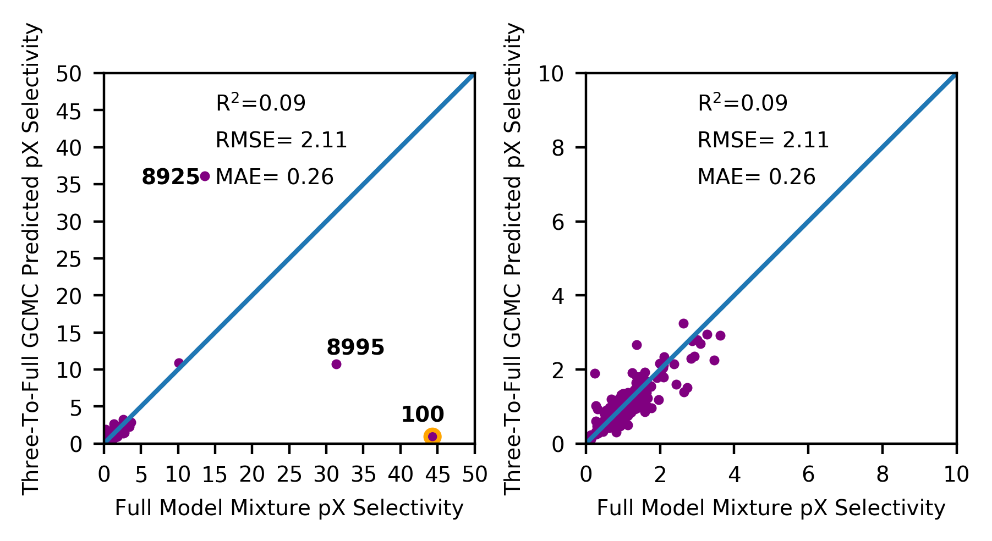
## 3.4 THREE-TO-FULL WORKFLOW

Although the three-site model has difficulties when predicting the exact number of selectivities, we can circumvent this issue by taking advantages of its merits. Since the individual and total loadings from the three-site models are highly accurate, the predictions from the three-site model are not far away from the full model prediction. Thus, we can use the three-site model results as starting points for the full model simulation and hopefully bypass some free energy barrier of insertions. A full description of the details and process of the “three-to-full workflow” is summarized in the SI and in Figure S15.

We applied the three-to-full workflow for all 640 1D channels structures in the ToBaCCo 1.0 database46. We ran 4,000 cycles for the semi-grand simulation and 20,000 for the final GCMC simulation. The results for loadings and selectivities are summarized in Figure 8 and Figure 9. We see that after the workflow, the selectivities are much closer to the full model values. This shows that the three-site model can be used to generate snapshots as the starting point of full model simulations. Also, the three-to-full workflow does not require energy grids to be generated, thus freeing up storage spaces.



**Figure 8.** Parity plots of loadings of equimolar mixture of xylene isomers (a: pX, b: oX, c: mX) and the total loading (d) at 433 K and 103 Bar fugacity in 640 ToBaCCo 1.0 MOFs with 1D channels with the three-site models using CB/CFC and the full models using EB-CB/CFC. The highlighted structures were the same from Figure 6.



**Figure 9.** Comparison between result after the three-to-full workflow versus the full model result. On the right we tried to zoom in. The outlier is labeled and highlighted with an orange circle. Other two structures with deviations are also labeled.

However, we see that there is 1 outlier in Figure 9. This one is MOF 100. Figure S16 shows the snapshot and iso-surface of the MOF. It has a low capacity while most of the cavities only allows for pX. According the EB-CB/CFC simulation, the MOF can take 3 pX molecules per channel. EBMC and normal GCMC all confirmed that this structure is pX-selective. The full model simulation yielded individual loadings at pX/oX/mX = 0.871/0.039/0.00 with error bars equal to 0.035/0.011/0.0 mol/kg. We compared the full model snapshots with the one generated from the TSX model, which is shown in Figure S17. The TSX snapshot shows MOF 100 is an oX-selective structure. We can see TSX-oX forms efficient packing within the channels. However, when converted to the full model, this packing will be eliminated with long enough cycles. As we can see from Figure S17, because the packing of TSX-oX molecules are very tight, when the snapshot gets converted to using full model, there are severe overlaps between methyl groups and hydrogen atoms of the packed oX molecules. This is because of the sizes of TSX oX and full model oX. The TSX-oX is shorter and bulkier comparing to full model oX, which has a big and flat aromatic ring. This caused the newly grown hydrogen atoms overlap with the methyl groups of another oX next to it.

We ran the final step of the three-to-full workflow for 1,050,000 cycles. Figure S18 summarized the cycle versus loading plot for this structure after running the GCMC process of the three-to-full workflow for 1,050,000 cycles. The snapshot of this GCMC simulation is shown in Figure S16e. Each channel can take 3 pX molecules despite an oX molecule in one of the channels. The pX loading goes up as more cycles used while the oX loading goes down. The averaged loadings for the three-to-full workflow is pX/oX/mX = 0.846/0.048/0.021 with error bars of 0.064/0.058/0.026 mol/kg. This is very close to the full model simulation result. The figure also indicates that total loading goes down quickly and shows that the structure may have multiple metastable states. Because the initial guess from TSX lands far away from zero loading, it might take a long time for the system to relax, comparing to starting from empty framework.

We also saw there was a data point that deviates from the y = x line with a very high pX-selectivity. This MOF is numbered 8925. The structures and adsorption snapshots are shown in Figure S19. The averaged loadings from EB-CB/CFC and the three-to-full workflow for pX/oX/mX were 0.526/0.00/0.077 with error bars equal to 0.065/0.0/0.031 and 0.636/0.00/0.077 with error bars equal to 0.032/0.001/0.025 mol/kg, respectively. We compared the difference between snapshots and saw there were almost no difference. The EB-CB/CFC snapshot does have 1 more mX molecules (highlighted in red circles) than the snapshot from the three-to-full workflow. Thus, we do not consider MOF #8925 as an outlier. Similarly, MOF #8995 displayed a same packing style in Figure S20 that also has very high selectivity and was not considered as an outlier as well. The averaged loadings for pX/oX/mX from EB-CB/CFC the three-to-full workflow for #8995 were 1.016/0.043/0.022 with error bars of 0.105/0.028/0.032 mol/kg, and 0.881/0.112/0.052 with error bars of 0.064/0.028/0.051 mol/kg, respectively.

After solving for the issues of the outlier MOF #100 from the three-to-full workflow, the first thing we need to resolve is whether the GCMC in our three-to-full process has reached convergence. For the case of MOF 100, if we use more cycles for the GCMC in the three-to-full workflow, we may see the loadings reach agreement with the ones generated from EB-CB/CFC. To form a systematic way of re-running the simulations until it converges to the correct loadings, we used the convergence criteria formulated by Gelman and Rubin64. The Gelman-Rubin (GR) diagnostic assumes that for several independent Markov chains, the variance within each chain should be the same if the chains are converged. The GR factor, denoted as R, is calculated below:

|  |  |  |
| --- | --- | --- |
|  | , | (26) |

where is the average of within block variance, B is the variance of block averages, and is the length of the Markov chain. Inspired by the work of Anstine et al.65, since purpose of block averages is to decorrelate the blocks, we replace the independent Markov chains with blocks from a single Monte Carlo simulation. For our case, the final GCMC in the three-to-full workflow uses 20,000 cycles. We printed the loadings every 200 cycles. When the simulation is finished, we extract all the printed loadings and discard the numbers from the first 10,000 cycles. Then, we divide the last 10,000 cycles into 5 blocks and calculate and for both pX loading and total loading. If the GR factor for either pX loading or total loading is greater than 1.15, we rerun the simulation with the latest snapshot. This completes a GR round. We cap it with a maximum of 5 GR rounds.

After this process, we compared the selectivities against the ones from EB-CB/CFC. The result is summarized in Figure 10. MOF #100 was able to reach a much closer selectivity comparing to the one from EB-CB/CFC. However, there appears to be another newly emerged outlier.

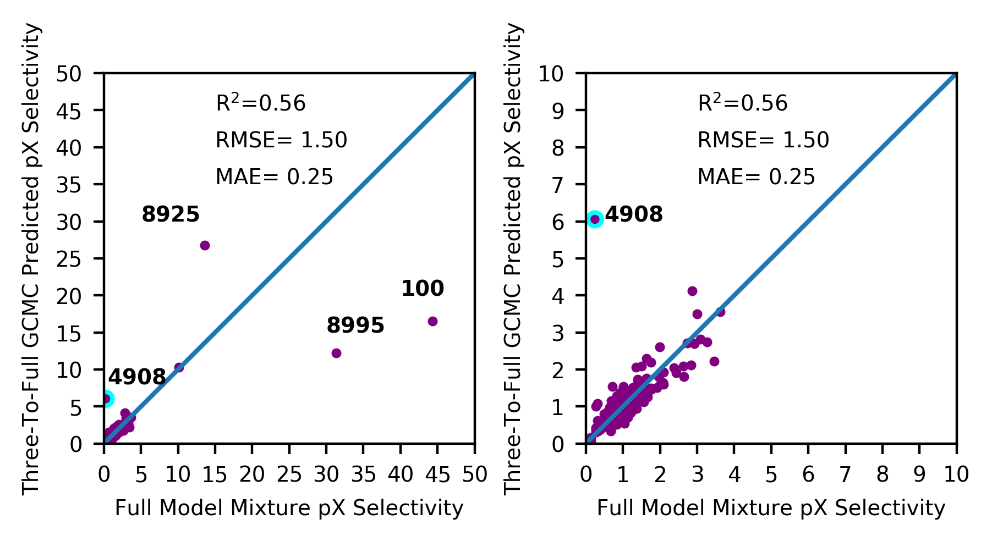


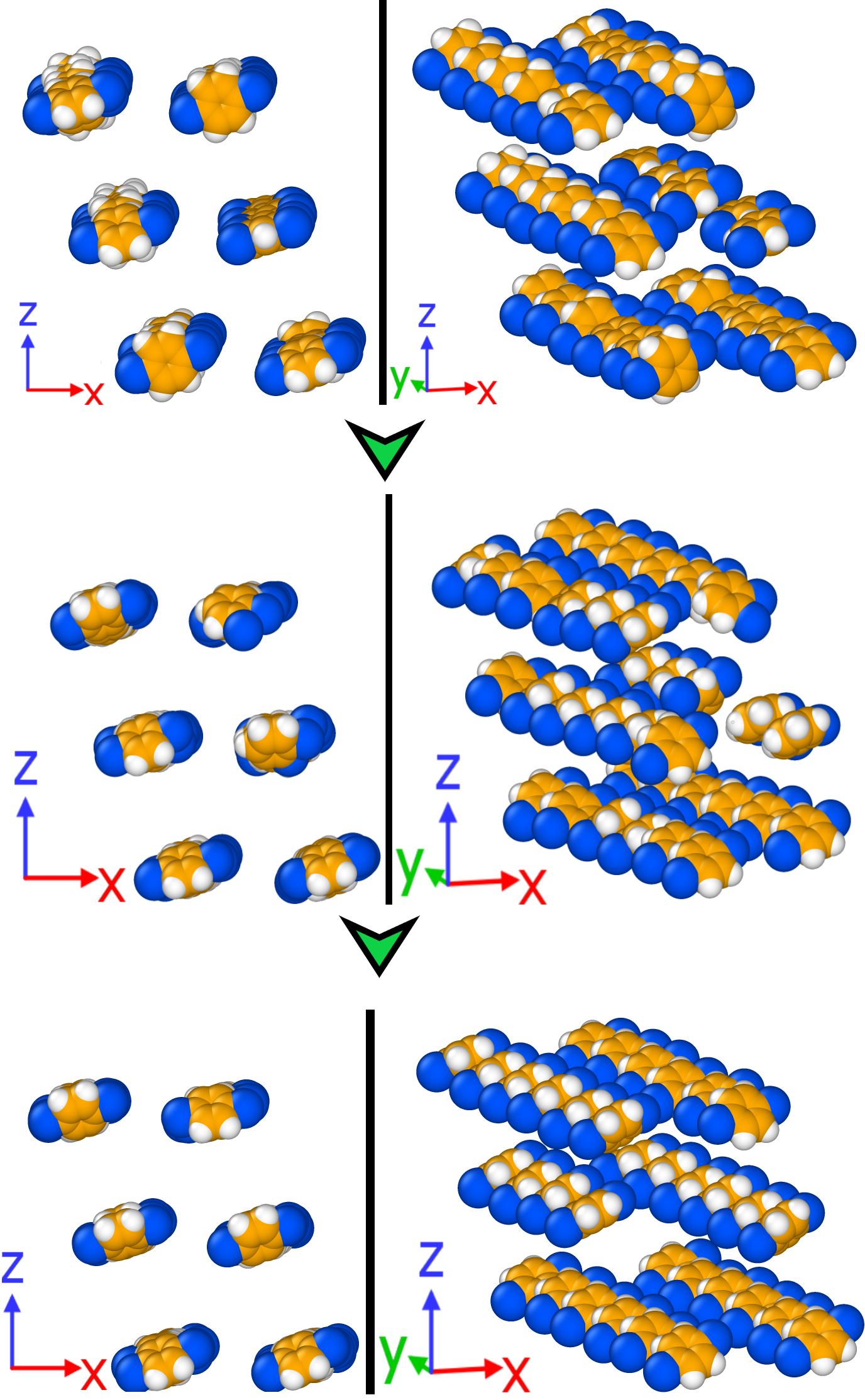
Figure 10**.** Comparison between result after the three-to-full workflow after GR process versus the full model result. On the right we tried to zoom in. The new outlier is labeled and highlighted with a cyan circle. Other structures with deviations are also labeled.

The new outlier was on the bottom-left corner of Figure 10 (outlier number 1 in Figure 10). The full model result generated by EB-CB/CFC has no selectivity while the result generated by three-to-full workflow + GR shows a pX-selectivity of around 5. This outlier MOF is #4908. The structure of #4908 is shown in Figure S21. We saw that from Figure S21a that the structure has three different visible regions possible for molecule insertions. We found that EB-CB/CFC tends to insert oX/mX molecules in the yellow region of Figure S21 while normal CB/CFC inserts pX molecules in the red region, as shown in Figure S22. The green region cannot fit any xylene isomers for both EB-CB/CFC and CB/CFC simulations.

To investigate the reason for this discrepancy, we first studied the distribution of energy favorable sites for approximate benzene. EBMC simulations try to insert xylene isomers at energy favorable sites of approximate benzene in a structure. If a specific site is disliked by AB, then the chance of that point being selected to insert the molecule is low. An easy way of doing this is simulating the adsorption of AB pseudo atom at the same condition (433 K, 103 Bar fugacity). The simulated snapshots are shown in Figure S23. We see that AB only goes into the region of MOF #4908 highlighted with yellow circles. AB never goes into the red ones. This is because of a steric effect. We saw from Figure S21 that the horizontal distance for the red region pore is 5.48 Å, while the vertical distance is only 3.44 Å. An AB pseudo atom has a diameter 5.27 Å. This means AB can never fit in the pores highlighted in red circles. For an overlapping region like this, the EBMC weight is very low. This explains why EBMC cannot fit pX into the red region. However, the height of the aromatic ring of pX is the diameter of the carbon pseudo-atom on the aromatic ring, which is 3.55 Å, while the diameter of the aromatic ring is 4.96 Å. These two sizes can very tightly fit the red region pore. This is the reason we saw that the normal CB/CFC inserts the pX molecules in the red region.

After realizing that the approximate benzene causes frequent overlaps in this MOF, we tried to do EBMC with a different probe. We used the LJ parameter of a methyl group to bias the insertion. The resulting cycles versus loading for this new EBMC with methyl probe and its comparison with other simulations are summarized in Figure S24. This time, we saw that there was no problem for EB-CB/CFC to agree with normal CB/CFC results. Also, EB-CB/CFC with methyl probe converges much faster than the normal CB/CFC simulation. Thus, we discovered a special case where EBMC with AB may not provide acceleration. With the help of a smaller pseudo-atom for biasing, the abnormality was resolved.

To then test our three-site model and the three-to-full workflow on a known structure, we still take xylene isomers in MTW zeolite as the example. The snapshots of steps of the three-to-full workflow are shown in Figure 11. We see that the snapshot generated from the three-site simulation has a lot of wrong orientations and overlaps. The loading of pX/oX/mX is 30/0/5. We modified the source code of RASPA to account for the overlap energies during the full model simulations for MTW. More discussion can be found in the SI. After the semi-grand simulation, we see that the pX molecules in the channels are oriented harmonically. One mX molecule (left middle channel) is changed to pX. This snapshot resolved almost all the overlaps and is good for the next GCMC step. The loadings of pX/oX/mX at this step is 31/3/1. After the GCMC step, we see that all the impurities were wiped out. Now the loadings are 34/0/0, and the selectivity matches the result from the full model simulation.

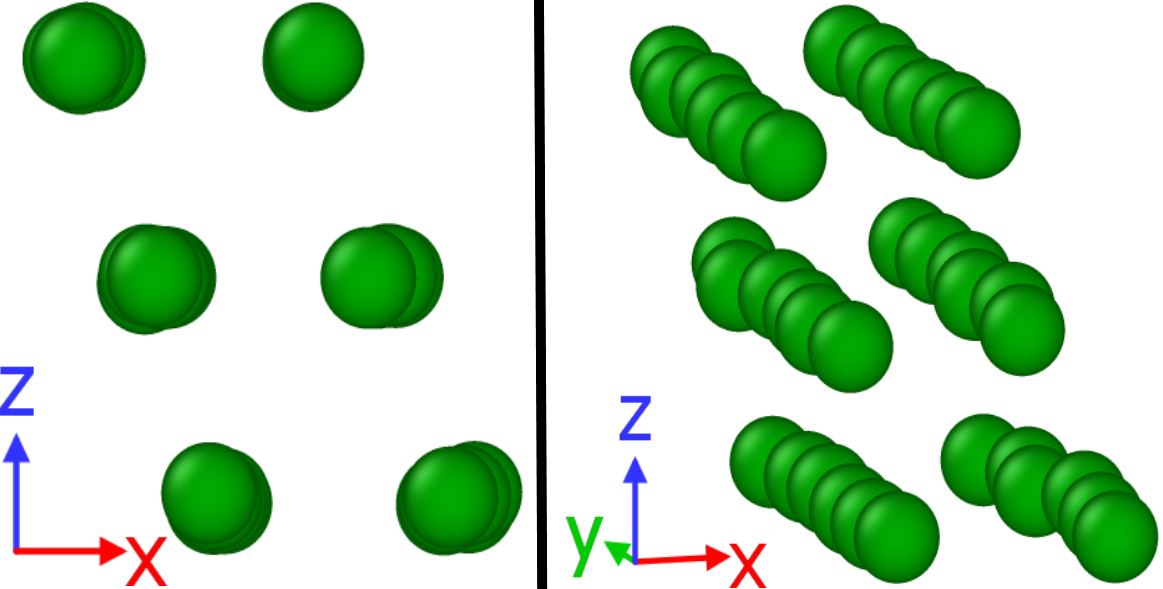


**Figure 11.** Snapshots of steps of the three-to-full workflow on xylene isomers adsorption in MTW zeolite at 433 K and 103 Bar fugacity. First row: snapshot after the three-site model simulation. Second row: snapshot after the semi-grand simulation. Third row: snapshot after the final GCMC simulation.

To test our three-site model and the three-to-full workflow on a real MOF, we selected MIL-47, which was the testing simulation for EBMC in the previous sections. We started by simulating equimolar tertiary mixture xylene adsorption at 343 K and 0.04 Bar using the three-site model. The simulation yielded a slight oX selectivity with the loadings of pX/oX/mX equals 0.31/1.33/0.87 with errors equal to 0.070/0.078/0.109 mol/kg. Snapshots (Figure S25) shows that the three-site model is too massive. Xylene cannot stack within the channels and each channel can only take 5 molecules (Figure S25, red circled region). We then tried to grow the full model xylene molecules from the three-site snapshot and ran a semi-grand simulation. This time we saw that although the total loading stays the same, since we enabled identity changes, the selectivities after the equilibration step was much closer to what Castillo et al.19 reported and the packing start to become correct with pX/oX/mX equals to 0.73/1.34/0.61 with errors equal to 0.207/0.272/0.088 mol/kg. From Figure S26, pX molecules started to replace mX and replicates the correct order of the abundance of the isomers: oX > pX > mX. We see that from the Front snapshot (Figure S26, right) that the oX molecules are using the space efficiently by forming parallel stacking within the channel though not at the correct capacity of the structure. The GCMC step filled the structure with oX and pX and the individual loadings became very close (pX/oX/mX = 1.19/2.41/0.26 with errors equal to 0.44/0.39/0.12 mol/kg) to that Castillo et al.19 reported. From Figure S27, we see that for the highlighted, circled channel, the eight xylene molecules are packing harmonically and form two decks of xylenes. Comparing to Figure S26, we see all channels start to form the double decks. This eventually shows that our three-to-full workflow can be applied to real MOF systems to yield accurate results.

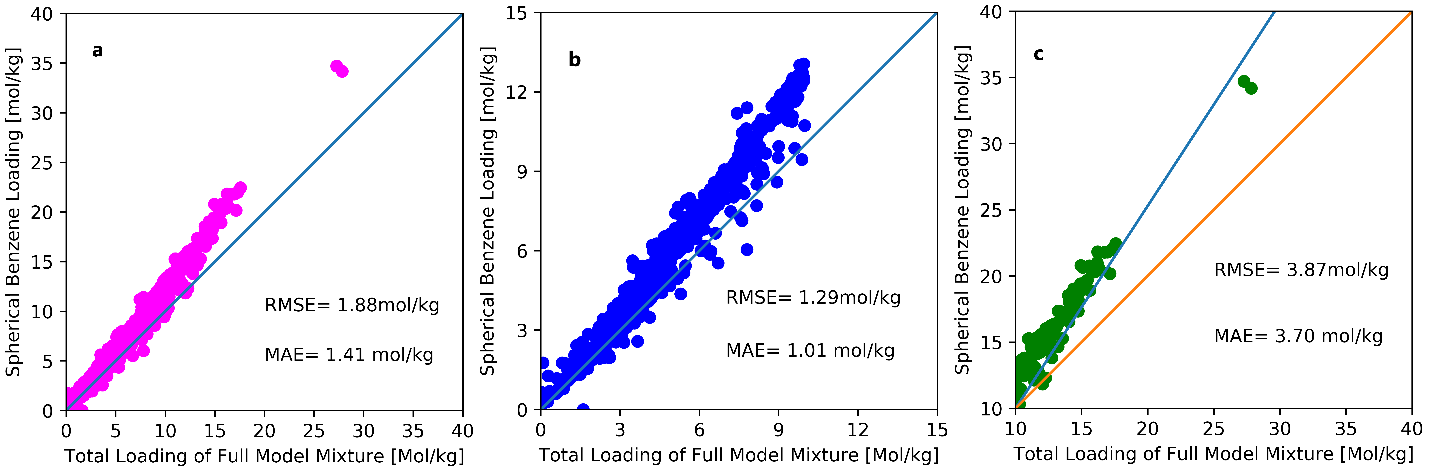
## 3.5 EFFECT OF APPROXIMATE BENZENE

Now that we finalized the three-to-full workflow, a side question arises on why EBMC simulations achieves speed-ups. Another related question on why TSX model reproduces the capacity of a structure with high accuracies. To solve for them, we used AB as a monatomic molecule and simulated single component adsorption GCMC at the same conditions in zeolite MTW. The saturation loading of AB shown in Figure 12 matched the full model mixture simulation for xylene isomers (Figure 2 and Figure 3), which is six molecules per channel. The reason for the acceleration of EBMC is that the locations of pX molecules correspond highly with the locations of AB. Thus, when performing EBMC biased with AB interaction energies, the algorithm will directly insert pX molecules in these locations. This also answered our question on TSX model. Because AB adsorption shows a loading that equals the full model capacity (6 molecules in 1 channel), TSX model that uses AB as a pseudo-atom inherits this merit.



**Figure 12.** Snapshot of AB adsorption in MTW zeolite at 433 K and 103 Bar fugacity using CBMC. For a clearer view, the framework atoms are deleted. Approximate benzene pseudo-atoms are colored in green.

We then calculated the loadings of AB for all 640 selected MOFs in ToBaCCo 1.0, the resulting parity plot is shown in Figure 13. The results indicate that the AB loadings are generally higher than the full model loadings. This is because AB does not consider the interactions of the methyl groups, so the AB molecule is smaller than xylene isomers. Despite the discrepancies, there is strong correlation between loadings of AB and the total loadings of full model mixture. We observed that there is a change of slope at full model capacity equals 10 mol/kg. For the low loading regions, the slope follows the y = x line, meaning that for this region the AB loading can accurately predict the total capacity for xylene isomers. However, when shifting to higher loadings (full model loading >= 10 mol/kg), we see that the slope deviates from the y = x line. To obtain a better understanding of the change of slope, we compared the sizes of the AB and the xylene isomers. Since the three-site model can accurately predict the total loading, we calculated the volume of three-site models and compared to the volume of AB. The volume of a three-site pX or mX (approximate benzene plus two methyl groups minus overlaps) is 117.5 Å3, while a AB pseudo-atom is 76.6 Å3. The size ratio of three-site pX/AB is 1.53. In Figure 13c, we plotted the points which has a full-model capacity greater than 10 mol/kg. We see that the parity line y = 1.53\*x corrected with the AB loadings well. This indicates that for lower loading structures, since there are confinements, the AB atoms stay at where the xylene molecules stay. However, when the confinements are reduced, the packing of AB become unrestricted. The ratio of between the loading of AB and the loading of xylene isomers indicates that, comparing to the effects of methyl groups, which determines the selectivities, spherical benzene determines the capacity of the structures, which is the combined loadings of the xylene isomers. Thus, when AB and methyl groups are combined, the three-site model can mimic the total loading and the selectivities.



**Figure 13.** Parity plot comparing total loading of full model xylene isomer simulations using EB-CB/CFC and loading of AB loading simulations using CBMC. a) Overall comparison between loadings of AB and loadings of full models of xylene isomers. b) Zoomed in comparison at the low loadings (full model loading ≤ 10 mol/kg). c) Zoomed in comparison at higher loadings, the blue line is y = 1.53\*x and the orange one is y = x.

# 4. CONCLUSION

We developed a new MC move that can accelerate the insertion of xylene molecules at high densities in a structure and a new xylene model that can mimic the behaviors of full model xylene at saturation inside 1D channel structures. EBMC places newly inserted molecules in energy favorable positions, thus faster than randomly placing them. The TSX model emerged from a thought experiment that intends to demystify the impact of different parts of a xylene molecule on its selectivity. The TSX model generated good overall agreements with the full model in terms of individual loadings of xylene isomers and the total capacity for xylene isomers of a structure. The success of the TSX model indicates that methyl groups play a more important role in determining selectivity than the aromatic ring. However, we also see outlier structures that were classified as 1D channel MOFs, such as MOF #4721, have motifs that prefers planar shape and thus enhance pX selectivity. The three-site model also had difficulties predicting the same selectivity as the full model since the selectivity is very sensitive to the loadings of oX and mX. We then developed a simple workflow that takes the three-site snapshot to grow the full model atoms and use this as a starting point for a short semi-grand and GCMC simulation. We found that this workflow greatly enhanced the quality of the selectivities. Two outliers, MOF #100 and MOF #4908, appear during the process of the three-to-full workflow. These outliers show the challenges that needs to be systematically addressed in the future. MOF #100 shows the case when the TSX model lands far away from the full model result. By combining convergence diagnostics with three-to-full workflow, the resulting selectivities gets better. And finally, during the Gelman-Rubin process, we found MOF #4908, on the other hand, suffers from a poor choice of probe for EBMC.

We also tested its performance in time for a difficult system, pX adsorption in MTW zeolite, and discovered that because of its fewer LJ sites and the absence of electrostatic interactions, the TSX model can perform more cycles and reach the correct loading in less time. Comparing to TSX and full model normal GCMC simulations, EBMC is the fastest method for this system. However, EBMC suffers from huge memory requirement and thus TSX is the most economical. We then discovered the effect of approximate benzene in the simulations of TSX models. We found that for low-capacity MOFs, the saturation loadings of spherical benzene are the capacities for xylene isomers. This is proved by our simulation of AB adsorption within MTW zeolite. For higher-capacity structures, the slope of the parity plot for this region is related to the volume ratio of the TSX model xylenes and the AB pseudo-atom, indicating that the methyl groups start to affect xylene isomer capacities of higher-capacity MOFs.

This study also compares different paths to cope with engineering problems. One can certainly develop new MC algorithms that requires huge amount of effort and time but is correct for all pressures. But one can also try to simplify the problem by utilizing the TSX model which can reproduce nothing of the full model but the saturation loadings. It shows that for specific applications, over-simplification can still capture the essential characteristics. Further research can be conducted along this path. For example, despite the lack of attractive forces from hard sphere models, one can simulate xylene separation at saturation condition by treating them as hard spheres, thus eliminating all the energy calculations.

This work also shows great potential of the combination of three-to-full workflow and the EBMC simulation. On one hand, EBMC simulations are shown to be much faster than normal GCMC simulations. However, the choice of biasing bead is a problem for structures like #4908. Then the use of three-to-full workflow provides a much less expensive approach of validating the EBMC simulation. On the other hand, for MOF #100, EBMC is much faster. From a philosophical point of view, the EBMC method and the TSX model are the dichotomy of the use of an approximate benzene bead. EBMC treats the bead as fictional and only use it for biasing. TSX model used the approximate bead for real atomistic simulations. Another way to understand the relation of EBMC and TSX model is the dichotomy of two classes of molecular simulations. EBMC tries to accelerate the atomistic simulation itself, while the TSX model tried to speed-up the process by coarse graining the guest molecules.

Another direction that is hard, but important for future work in this topic is to find smarter ways of implementing the three-to-full workflow. We know that because of the random nature of MC simulations, we cannot know what happens if we run the simulation 1 cycle further. This leads to the unsolvable question of MC convergence. Currently, we implemented the GR diagnostic for the three-to-full workflow. However, we still cannot say for sure what can happen if we run the simulation longer. Thus, a better diagnostic awaits us for the convergence of MC simulations in the molecular simulation community.

# ACKNOWLEDGEMENT

This work was supported by the U.S. Department of Energy under Award DE-FG02-08ER15967. Z.L. also acknowledges support from Data Science Fellowship via the Northwestern Institute on Complex Systems (NICO). This research was partly supported through the computational resources provided for the Quest high-performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology. Z.L. also gratefully thanks the resources provided by the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The authors thank Professor David Dubbeldam for providing simulation data on xylene in MTW zeolite and other helpful discussions on CBCFC simulations for us to reproduce their work. Zhao Li thanks Dr. Kaihang Shi for other helpful discussions. Z.L. also thanks Dr. Benjamin Sikora and Dr. Hongda Zhang for previous but crucial developments of algorithms in RASPA within the Snurr group. Finally, Z.L. thanks Mr. Andrew S. Rosen for helps and discussions on the MOFkey generations.

# ASSOCIATED CONTENT

Lennard-Jones parameters for the MOF from ToBaCCo 1.0 database, Lennard-Jones parameters, and partial charges for MIL-47, MTW zeolite, and xylene isomers. Validation of EBMC algorithm of ethane and xylene isomer separation in MIL-47. Textural properties distributions of the selected MOFs. Snapshots of selected MOFs and outliers in the three-to-full workflow including the loading versus cycle plots.

Excel spreadsheet that contains MOFkeys, linker and node information for 640 selected ToBaCCo 1.0 MOFs.

RASPA with EBMC implementation is available at: <https://github.com/snurr-group>

RASPA with energy cap removed is available at: <https://github.com/snurr-group/Three-to-Full-Xylene>

Three-to-full workflow python script is available at: <https://github.com/snurr-group/Three-to-Full-Xylene>

The 3D interactive plot of comparison between individual loadings generated by TSX and full model is available at: <https://github.com/snurr-group/Three-to-Full-Xylene/blob/master/TSX-Compare-3D.html>

# REFERENCES

(1) Sholl, D. S.; Lively, R. P. Seven Chemical Separations to Change the World. *Nature News* **2016**, *532* (7600), 435. https://doi.org/10.1038/532435a.

(2) Lachet, V.; Buttefey, S.; Boutin, A.; H. Fuchs, A. Molecular Simulation of Adsorption Equilibria of Xylene Isomer Mixtures in Faujasite Zeolites. A Study of the Cation Exchange Effect on Adsorption Selectivity. *Physical Chemistry Chemical Physics* **2001**, *3* (1), 80–86. https://doi.org/10.1039/B007281H.

(3) Torres-Knoop, A.; Balaji, S. P.; Vlugt, T. J. H.; Dubbeldam, D. A Comparison of Advanced Monte Carlo Methods for Open Systems: CFCMC vs CBMC. *J. Chem. Theory Comput.* **2014**, *10* (3), 942–952. https://doi.org/10.1021/ct4009766.

(4) Tarditi, A. M.; Horowitz, G. I.; Lombardo, E. A. A Durable ZSM-5/SS Composite Tubular Membrane for the Selective Separation of p-Xylene from Its Isomers. *Journal of Membrane Science* **2006**, *281* (1), 692–699. https://doi.org/10.1016/j.memsci.2006.04.041.

(5) Wei, J. A Mathematical Theory of Enhanced Para-Xylene Selectivity in Molecular Sieve Catalysts. *Journal of Catalysis* **1982**, *76* (2), 433–439. https://doi.org/10.1016/0021-9517(82)90272-X.

(6) Sun, N.; Wang, S.-Q.; Zou, R.; Cui, W.-G.; Zhang, A.; Zhang, T.; Li, Q.; Zhuang, Z.-Z.; Zhang, Y.-H.; Xu, J.; J. Zaworotko, M.; Bu, X.-H. Benchmark Selectivity p -Xylene Separation by a Non-Porous Molecular Solid through Liquid or Vapor Extraction. *Chemical Science* **2019**, *10* (38), 8850–8854. https://doi.org/10.1039/C9SC02621E.

(7) Smith, G. D.; Jaffe, R. L. Comparative Study of Force Fields for Benzene. *J. Phys. Chem.* **1996**, *100* (23), 9624–9630. https://doi.org/10.1021/jp9535194.

(8) Gee, J. A.; Zhang, K.; Bhattacharyya, S.; Bentley, J.; Rungta, M.; Abichandani, J. S.; Sholl, D. S.; Nair, S. Computational Identification and Experimental Evaluation of Metal–Organic Frameworks for Xylene Enrichment. *J. Phys. Chem. C* **2016**, *120* (22), 12075–12082. https://doi.org/10.1021/acs.jpcc.6b03349.

(9) Ma, Y.; Jue, M. L.; Zhang, F.; Mathias, R.; Jang, H. Y.; Lively, R. P. Creation of Well-Defined “Mid-Sized” Micropores in Carbon Molecular Sieve Membranes. *Angewandte Chemie* **2019**, *131* (38), 13393–13399. https://doi.org/10.1002/ange.201903105.

(10) O’Brien-Abraham, J.; Lin, Y. S. Effect of Isomorphous Metal Substitution in Zeolite Framework on Pervaporation Xylene-Separation Performance of MFI-Type Zeolite Membranes. *Ind. Eng. Chem. Res.* **2010**, *49* (2), 809–816. https://doi.org/10.1021/ie900926t.

(11) Ma, Y.; Zhang, F.; Yang, S.; Lively, R. P. Evidence for Entropic Diffusion Selection of Xylene Isomers in Carbon Molecular Sieve Membranes. *Journal of Membrane Science* **2018**, *564*, 404–414. https://doi.org/10.1016/j.memsci.2018.07.040.

(12) Yang, X.; Zhou, H.-L.; He, C.-T.; Mo, Z.-W.; Ye, J.-W.; Chen, X.-M.; Zhang, J.-P. Flexibility of Metal-Organic Framework Tunable by Crystal Size at the Micrometer to Submillimeter Scale for Efficient Xylene Isomer Separation https://spj.sciencemag.org/research/2019/9463719/?utm\_source=trendmd&utm\_medium=cpc&utm\_campaign=research\_trendmd\_1 (accessed Jun 23, 2020).

(13) Wu, X.; Wei, W.; Jiang, J.; Caro, J.; Huang, A. High-Flux High-Selectivity Metal–Organic Framework MIL-160 Membrane for Xylene Isomer Separation by Pervaporation. *Angewandte Chemie International Edition* **2018**, *57* (47), 15354–15358. https://doi.org/10.1002/anie.201807935.

(14) Wang, S.-Q.; Mukherjee, S.; Patyk‐Kaźmierczak, E.; Darwish, S.; Bajpai, A.; Yang, Q.-Y.; Zaworotko, M. J. Highly Selective, High-Capacity Separation of o-Xylene from C8 Aromatics by a Switching Adsorbent Layered Material. *Angewandte Chemie* **2019**, *131* (20), 6702–6706. https://doi.org/10.1002/ange.201901198.

(15) Agrawal, M.; Bhattacharyya, S.; Huang, Y.; Jayachandrababu, K. C.; Murdock, C. R.; Bentley, J. A.; Rivas-Cardona, A.; Mertens, M. M.; Walton, K. S.; Sholl, D. S.; Nair, S. Liquid-Phase Multicomponent Adsorption and Separation of Xylene Mixtures by Flexible MIL-53 Adsorbents. *J. Phys. Chem. C* **2018**, *122* (1), 386–397. https://doi.org/10.1021/acs.jpcc.7b09105.

(16) He, Z.; Yang, Y.; Bai, P.; Guo, X. Metal-Organic Framework MIL-53(Cr) as a Superior Adsorbent: Highly Efficient Separation of Xylene Isomers in Liquid Phase. *Journal of Industrial and Engineering Chemistry* **2019**, *77*, 262–272. https://doi.org/10.1016/j.jiec.2019.04.047.

(17) Sappidi, P.; Rabideau, B. D.; Turner, C. H. Molecular Simulation of the Separation of Toluene and P-Xylene with the Thermally-Robust Ionic Liquid Triphenyl-p-Phenyl Sulfonyl Phenyl Phosphonium. *Chemical Engineering Science* **2020**, *224*, 115790. https://doi.org/10.1016/j.ces.2020.115790.

(18) Caro-Ortiz, S.; Hens, R.; Zuidema, E.; Rigutto, M.; Dubbeldam, D.; Vlugt, T. J. H. Molecular Simulation of the Vapor-Liquid Equilibria of Xylene Mixtures: Force Field Performance, and Wolf vs. Ewald for Electrostatic Interactions. *Fluid Phase Equilibria* **2019**, *485*, 239–247. https://doi.org/10.1016/j.fluid.2018.12.006.

(19) Castillo, J. M.; Vlugt, T. J. H.; Calero, S. Molecular Simulation Study on the Separation of Xylene Isomers in MIL-47 Metal−Organic Frameworks. *J. Phys. Chem. C* **2009**, *113* (49), 20869–20874. https://doi.org/10.1021/jp908247w.

(20) Miyamoto, M.; Ono, S.; Oumi, Y.; Uemiya, S.; Van der Perre, S.; Virdis, T.; Baron, G. V.; Denayer, J. F. M. Nanoporous ZSM-5 Crystals Coated with Silicalite-1 for Enhanced p-Xylene Separation. *ACS Appl. Nano Mater.* **2019**, *2* (5), 2642–2650. https://doi.org/10.1021/acsanm.9b00037.

(21) Kumar, P.; Kim, D. W.; Rangnekar, N.; Xu, H.; Fetisov, E. O.; Ghosh, S.; Zhang, H.; Xiao, Q.; Shete, M.; Siepmann, J. I.; Dumitrica, T.; McCool, B.; Tsapatsis, M.; Mkhoyan, K. A. One-Dimensional Intergrowths in Two-Dimensional Zeolite Nanosheets and Their Effect on Ultra-Selective Transport. *Nature Materials* **2020**, *19* (4), 443–449. https://doi.org/10.1038/s41563-019-0581-3.

(22) Lee, S.; Kim, D.; Kim, J.; Jung, O.-S. Pseudo-2D Porous Networks via Interpenetration of 1D Zigzag Ladder-Type Coordination Polymers: Adsorption and Separation of Xylene Isomers. *Crystal Growth & Design* **2020**, *20* (6), 3601–3604. https://doi.org/10.1021/acs.cgd.0c00336.

(23) Khabzina, Y.; Laroche, C.; Pérez-Pellitero, J.; Farrusseng, D. Quantitative Structure–Property Relationship Approach to Predicting Xylene Separation with Diverse Exchanged Faujasites. *Physical Chemistry Chemical Physics* **2018**, *20* (36), 23773–23782. https://doi.org/10.1039/C8CP04042G.

(24) Mullen, R. G.; Maginn, E. J. Reaction Ensemble Monte Carlo Simulation of Xylene Isomerization in Bulk Phases and under Confinement. *J. Chem. Theory Comput.* **2017**, *13* (9), 4054–4062. https://doi.org/10.1021/acs.jctc.7b00498.

(25) du Plessis, M.; Nikolayenko, V. I.; Barbour, L. J. Record-Setting Selectivity for p-Xylene by an Intrinsically Porous Zero-Dimensional Metallocycle. *J. Am. Chem. Soc.* **2020**, *142* (10), 4529–4533. https://doi.org/10.1021/jacs.9b11314.

(26) Koh, D.-Y.; McCool, B. A.; Deckman, H. W.; Lively, R. P. Reverse Osmosis Molecular Differentiation of Organic Liquids Using Carbon Molecular Sieve Membranes. *Science* **2016**, *353* (6301), 804–807. https://doi.org/10.1126/science.aaf1343.

(27) Torres‐Knoop, A.; Krishna, R.; Dubbeldam, D. Separating Xylene Isomers by Commensurate Stacking of P-Xylene within Channels of MAF-X8. *Angewandte Chemie International Edition* **2014**, *53* (30), 7774–7778. https://doi.org/10.1002/anie.201402894.

(28) Kim, S. H.; Park, J. H.; Go, E. M.; Kim, W.-S.; Kwak, S. K. Separation Principle of Xylene Isomers and Ethylbenzene with Hydrogen-Bonded Host Frameworks via First-Principles Calculation. *Journal of Industrial and Engineering Chemistry* **2020**, *85*, 276–281. https://doi.org/10.1016/j.jiec.2020.02.010.

(29) Khabzina, Y.; Laroche, C.; Perez-Pellitero, J.; Farrusseng, D. Xylene Separation on a Diverse Library of Exchanged Faujasite Zeolites. *Microporous and Mesoporous Materials* **2017**, *247*, 52–59. https://doi.org/10.1016/j.micromeso.2017.03.026.

(30) Hartlieb, K. J.; Holcroft, J. M.; Moghadam, P. Z.; Vermeulen, N. A.; Algaradah, M. M.; Nassar, M. S.; Botros, Y. Y.; Snurr, R. Q.; Stoddart, J. F. CD-MOF: A Versatile Separation Medium. *J. Am. Chem. Soc.* **2016**, *138* (7), 2292–2301. https://doi.org/10.1021/jacs.5b12860.

(31) Holcroft, J. M.; Hartlieb, K. J.; Moghadam, P. Z.; Bell, J. G.; Barin, G.; Ferris, D. P.; Bloch, E. D.; Algaradah, M. M.; Nassar, M. S.; Botros, Y. Y.; Thomas, K. M.; Long, J. R.; Snurr, R. Q.; Stoddart, J. F. Carbohydrate-Mediated Purification of Petrochemicals. *J. Am. Chem. Soc.* **2015**, *137* (17), 5706–5719. https://doi.org/10.1021/ja511878b.

(32) Torres‐Knoop, A.; Balestra, S. R. G.; Krishna, R.; Calero, S.; Dubbeldam, D. Entropic Separations of Mixtures of Aromatics by Selective Face-to-Face Molecular Stacking in One-Dimensional Channels of Metal–Organic Frameworks and Zeolites. *ChemPhysChem* **2015**, *16* (3), 532–535. https://doi.org/10.1002/cphc.201402819.

(33) Torres-Knoop, A.; Dubbeldam, D. Exploiting Large-Pore Metal–Organic Frameworks for Separations through Entropic Molecular Mechanisms. *ChemPhysChem* **2015**, *16* (10), 2046–2067. https://doi.org/10.1002/cphc.201500195.

(34) Sikora, B. J.; Colón, Y. J.; Snurr, R. Q. Continuous Fractional Component Monte Carlo Simulations of High-Density Adsorption in Metal–Organic Frameworks. *Molecular Simulation* **2015**, *41* (16–17), 1339–1347. https://doi.org/10.1080/08927022.2015.1043629.

(35) Siepmann, J. I.; Frenkel, D. Configurational Bias Monte Carlo: A New Sampling Scheme for Flexible Chains. *Molecular Physics* **1992**, *75* (1), 59–70. https://doi.org/10.1080/00268979200100061.

(36) 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* (4), 1451–1463. https://doi.org/10.1021/ct7000039.

(37) Zhang, H.; Snurr, R. Q. Computational Study of Water Adsorption in the Hydrophobic Metal–Organic Framework ZIF-8: Adsorption Mechanism and Acceleration of the Simulations. *J. Phys. Chem. C* **2017**, *121* (43), 24000–24010. https://doi.org/10.1021/acs.jpcc.7b06405.

(38) Snurr, R. Q.; Bell, A. T.; Theodorou, D. N. Prediction of Adsorption of Aromatic Hydrocarbons in Silicalite from Grand Canonical Monte Carlo Simulations with Biased Insertions. *J. Phys. Chem.* **1993**, *97* (51), 13742–13752. https://doi.org/10.1021/j100153a051.

(39) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids: Second Edition*; Oxford University Press, 2017.

(40) Dubbeldam, D.; Calero, S.; Ellis, D. E.; Snurr, R. Q. RASPA: Molecular Simulation Software for Adsorption and Diffusion in Flexible Nanoporous Materials. *Molecular Simulation* **2016**, *42* (2), 81–101. https://doi.org/10.1080/08927022.2015.1010082.

(41) Esselink, K.; Loyens, L. D. J. C.; Smit, B. Parallel Monte Carlo Simulations. *Phys. Rev. E* **1995**, *51* (2), 1560–1568. https://doi.org/10.1103/PhysRevE.51.1560.

(42) Smit, B. Grand Canonical Monte Carlo Simulations of Chain Molecules: Adsorption Isotherms of Alkanes in Zeolites. *Molecular Physics* **1995**, *85* (1), 153–172. https://doi.org/10.1080/00268979500101011.

(43) MACEDONIA, M. D.; MAGINN, E. J. A Biased Grand Canonical Monte Carlo Method for Simulating Adsorption Using All-Atom and Branched United Atom Models. *Molecular Physics* **1999**, *96* (9), 1375–1390. https://doi.org/10.1080/00268979909483082.

(44) Gupta, A.; Clark, L. A.; Snurr, R. Q. Grand Canonical Monte Carlo Simulations of Nonrigid Molecules:  Siting and Segregation in Silicalite Zeolite. *Langmuir* **2000**, *16* (8), 3910–3919. https://doi.org/10.1021/la990756f.

(45) Torres‐Knoop, A.; Krishna, R.; Dubbeldam, D. Separating Xylene Isomers by Commensurate Stacking of P-Xylene within Channels of MAF-X8. *Angewandte Chemie International Edition* **2014**, *53* (30), 7774–7778. https://doi.org/10.1002/anie.201402894.

(46) Colón, Y. J.; Gómez-Gualdrón, D. A.; Snurr, R. Q. Topologically Guided, Automated Construction of Metal–Organic Frameworks and Their Evaluation for Energy-Related Applications. *Crystal Growth & Design* **2017**, *17* (11), 5801–5810. https://doi.org/10.1021/acs.cgd.7b00848.

(47) Willems, T. F.; Rycroft, C. H.; Kazi, M.; Meza, J. C.; Haranczyk, M. Algorithms and Tools for High-Throughput Geometry-Based Analysis of Crystalline Porous Materials. *Microporous and Mesoporous Materials* **2012**, *149* (1), 134–141. https://doi.org/10.1016/j.micromeso.2011.08.020.

(48) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114* (25), 10024–10035. https://doi.org/10.1021/ja00051a040.

(49) Jorgensen, W. L.; Laird, E. R.; Nguyen, T. B.; Tirado‐Rives, J. Monte Carlo Simulations of Pure Liquid Substituted Benzenes with OPLS Potential Functions. *Journal of Computational Chemistry* **1993**, *14* (2), 206–215. https://doi.org/10.1002/jcc.540140208.

(50) Bucior, B. J.; Rosen, A. S.; Haranczyk, M.; Yao, Z.; Ziebel, M. E.; Farha, O. K.; Hupp, J. T.; Siepmann, J. I.; Aspuru-Guzik, A.; Snurr, R. Q. Identification Schemes for Metal–Organic Frameworks To Enable Rapid Search and Cheminformatics Analysis. *Crystal Growth & Design* **2019**, *19* (11), 6682–6697. https://doi.org/10.1021/acs.cgd.9b01050.

(51) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31–36. https://doi.org/10.1021/ci00057a005.

(52) Lekien, F.; Marsden, J. Tricubic Interpolation in Three Dimensions. *International Journal for Numerical Methods in Engineering* **2005**, *63* (3), 455–471. https://doi.org/10.1002/nme.1296.

(53) Alaerts, L.; Kirschhock, C. E. A.; Maes, M.; van der Veen, M. A.; Finsy, V.; Depla, A.; Martens, J. A.; Baron, G. V.; Jacobs, P. A.; Denayer, J. F. M.; De Vos, D. E. Selective Adsorption and Separation of Xylene Isomers and Ethylbenzene with the Microporous Vanadium(IV) Terephthalate MIL-47. *Angewandte Chemie International Edition* **2007**, *46* (23), 4293–4297. https://doi.org/10.1002/anie.200700056.

(54) Alaerts, L.; Maes, M.; A. Jacobs, P.; M. Denayer, J. F.; Vos, D. E. D. Activation of the Metal–Organic Framework MIL-47 for Selective Adsorption of Xylenes and Other Difunctionalized Aromatics. *Physical Chemistry Chemical Physics* **2008**, *10* (20), 2979–2985. https://doi.org/10.1039/B719513C.

(55) Finsy, V.; Verelst, H.; Alaerts, L.; De Vos, D.; Jacobs, P. A.; Baron, G. V.; Denayer, J. F. M. Pore-Filling-Dependent Selectivity Effects in the Vapor-Phase Separation of Xylene Isomers on the Metal−Organic Framework MIL-47. *J. Am. Chem. Soc.* **2008**, *130* (22), 7110–7118. https://doi.org/10.1021/ja800686c.

(56) Finsy, V.; Kirschhock, C. E. A.; Vedts, G.; Maes, M.; Alaerts, L.; De Vos, D. E.; Baron, G. V.; Denayer, J. F. M. Framework Breathing in the Vapour-Phase Adsorption and Separation of Xylene Isomers with the Metal–Organic Framework MIL-53. *Chemistry – A European Journal* **2009**, *15* (31), 7724–7731. https://doi.org/10.1002/chem.200802672.

(57) Krishna, R.; Baten, J. M. van. A Molecular Simulation Study of Commensurate–Incommensurate Adsorption of n-Alkanes in Cobalt Formate Frameworks. *Molecular Simulation* **2009**, *35* (12–13), 1098–1104. https://doi.org/10.1080/08927020902744672.

(58) Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *J. Phys. Chem. B* **1998**, *102* (14), 2569–2577. https://doi.org/10.1021/jp972543+.

(59) Acree, Jr, W. E.; Chickos, J. S. Phase Transition Enthalpy Measurements of Organic and Organometallic Compounds. *NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899*. https://doi.org/10.18434/T4D303.

(60) Hirschfelder; Curtiss, C.; Bird, B. The Molecular Theory of Gases and Liquids https://www.wiley.com/en-us/The+Molecular+Theory+of+Gases+and+Liquids-p-9780471400653 (accessed Jul 17, 2019).

(61) Bird, R. B.; Stewart, W. E.; Lightfoot, E. N. *Transport Phenomena, Revised 2nd Edition*; John Wiley & Sons.

(62) Peng, D.-Y.; Robinson, D. B. A New Two-Constant Equation of State. *Ind. Eng. Chem. Fund.* **1976**, *15* (1), 59–64. https://doi.org/10.1021/i160057a011.

(63) JSW: Framework Type https://america.iza-structure.org/IZA-SC/framework.php?STC=JSW (accessed Sep 21, 2020).

(64) Gelman, A.; Rubin, D. B. Inference from Iterative Simulation Using Multiple Sequences. *Statist. Sci.* **1992**, *7* (4), 457–472. https://doi.org/10.1214/ss/1177011136.

(65) Anstine, D. M.; Demidov, A. G.; Mendez, N. F.; Morgan, W. J.; Colina, C. M. Screening PIM-1 Performance as a Membrane for Binary Mixture Separation of Gaseous Organic Compounds. *Journal of Membrane Science* **2020**, *599*, 117798. https://doi.org/10.1016/j.memsci.2019.117798.