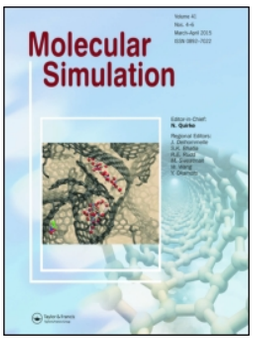

Molecular Simulation

ISSN: 0892-7022 (Print) 1029-0435 (Online) Journal homepage: http://www.tandfonline.com/loi/gmos20

Enhanced replica exchange reactive Monte Carlo simulations for constructing zeolite frameworks

Cecilia Bores, Scott M. Auerbach & Peter A. Monson

To cite this article: Cecilia Bores, Scott M. Auerbach & Peter A. Monson (2018) Enhanced replica exchange reactive Monte Carlo simulations for constructing zeolite frameworks, Molecular Simulation, 44:6, 453-462, DOI: 10.1080/08927022.2017.1399375

To link to this article: https://doi.org/10.1080/08927022.2017.1399375

Published online: 21 Nov 2017.

MOLECULAR SIMULATION, 2018
VOL. 44, NO. 6, 453–462
https://doi.org/10.1080/08927022.2017.1399375

Taylor & Francis
Taylor & Francis Group

Enhanced replica exchange reactive Monte Carlo simulations for constructing zeolite frameworks

Cecilia Bores[a], Scott M. Auerbach[a,b] and Peter A. Monson[a]

[a]Department of Chemical Engineering, University of Massachusetts Amherst, Amherst, MA, USA; [b]Department of Chemistry, University of Massachusetts Amherst, Amherst, MA, USA

ABSTRACT

We investigate the ability of a reactive model of silica to build crystalline zeolite frameworks with an enhanced sampling approach based on replica exchange Monte Carlo (REMC) simulations. In our implementation of REMC, the silica hydrolysis/condensation equilibrium constant controlling network formation is chosen as the index parameter characterising each replica. We show that improving the performance of the REMC method by increasing its efficiency in producing zeolite crystals allows zeolites with larger unit cells to be constructed. In particular, we perform an efficiency case study on the sodalite structure, containing 12 tetrahedra per unit cell and then apply the resulting enhancements to construct a significantly larger zeolite. We have improved our simulations in two ways; first, we have analysed the roles of the Monte Carlo parameters on the efficiency of this REMC method. Second, we have implemented an adaptive protocol that resets the values of the hydrolysis/condensation equilibrium constants among the replicas, 'on the fly' of each simulation, to optimise replica exchange for the purpose of constructing zeolite crystals. Finally, we show that by applying these enhancements, the REMC method can produce the crystal structure of zeolite AWW, a nanoporous material with 24 tetrahedra per unit cell.

ARTICLE HISTORY
Received 17 July 2017
Accepted 24 October 2017

KEYWORDS
Replica exchange Monte Carlo; Zeolites; Silica Polymerisation; Self Assembly

1. Introduction

Zeolites are crystalline microporous materials that consist of a network of corner-sharing tetrahedra, each with T-atoms (Si, Al, P or other metals) in the centre and four oxygens at the vertices. Zeolite materials enjoy substantial technological relevance, as shape-selective catalysts and as molecular sieves for green separations, and are scientifically important as platforms to study the dynamics of confined systems [1]. There is significant interest in understanding more fully how zeolites form microporous crystals, both for technological control of zeolite pore structures and crystalline architectures, and to understand the science of hierarchical structure formation [2]. Molecular modelling stands to play an important role in elucidating zeolite formation, because the sizes of zeolite nuclei likely fall into the nanoscale blindspot between IR/NMR and XRD characterisation methods. In previous work, we reported a replica exchange (RE), reaction ensemble (Rx) Monte Carlo (RE-RxMC) simulation procedure for building structures of zeolite crystals using a reactive model of silica polymerisation [3]. We note that RE-RxMC trajectories are optimised to equilibrate rapidly and thus do not follow physical trajectories, especially when barriers separate thermodynamically stable structures or phases. RE-RxMC nonetheless represents an important computational approach for materials discovery, by determining the crystalline phases consistent with our reactive model of silica. The initial study was limited to zeolites with small unit cells due to the restrictions of the methodology. In the present work, we extend this RE-RxMC method to larger zeolites by exploring enhancements in the efficiency of the RE-RxMC approach.

Computational scientists have made striking progress in modelling the structure of crystalline materials from the knowledge of their composition during the last 30 years [2]. In the field of zeolites, significant progress has been made in predicting the existence of new structures. Using graph theory and combinatorial tiling theory, millions of energetically feasible structures have been found during the last decades [4]: Deem et al., Treacy et al. and Yi et al. have generated databases of hypothetical zeolites [5–7]. However, a more complete molecular-level understanding of the zeolite synthesis process remains elusive. Computational simulations of silica polymerisation have been recently performed by Malani et al. [8,9]. These simulations accurately reproduced the evolution of silica network formation as determined experimentally by ^{29}Si NMR, which measures the Q_n molar fraction distribution as a function of time. The simulations of Malani et al. accomplished this agreement by performing Reaction Ensemble MC simulations (RxMC) in combination with the silica tetrahedron model proposed by Astala et al. [10]. This model was originally devised to reproduce the mechanical properties of crystalline silica solids and represents silicic acid molecules through flexible and unbreakable tetrahedra. Chien et al. later extended this silica polymerisation model to study silica nanoparticles [11] and zeolite crystallisation [3]. Because of the glassy nature of silica, applying this model with available computational resources to find zeolite crystals required a specialised sampling procedure: Replica Exchange Monte Carlo (RE-MC) [12].

RE-MC [12] is a useful technique for simulating systems with rugged energy landscapes like zeolites, where the system

CONTACT Scott M. Auerbach auerbach@chem.umass.edu; Peter A. Monson monson@ecs.umass.edu

may get stuck in metastable structures far from the equilibrium state. The key of this method is to simultaneously perform independent MC simulations of N parallel replicas of the system considering a slightly different index parameter (e.g. temperature, pressure, chemical potential) value for each replica. The maximum and minimum index parameter values are chosen to lead to different MC equilibration times and thus, by periodically exchanging configurations between the replicas, the slowest equilibrating replica will more rapidly reach the equilibrium state. Hence, this technique increases the efficiency of sampling near the equilibrium state, allowing the system to rapidly escape from metastable local minima. In this work, as we did in our previous contribution [3], we have adapted a Replica Exchange (RE) technique to Reactive Ensemble (Rx) Monte Carlo simulations (RE-RxMC) [13]. We have chosen the equilibrium constant controlling silica hydrolysis reactions (K_{hyd}) as the index parameter that varies from one replica to another. Large values of K_{hyd} promote disassembly of silica networks, important for avoiding glassy states, and small values of K_{hyd} drive the system to form silica networks. Thus, a suitably chosen grid of K_{hyd} values avoids glass formation and forms silica crystalline states.

The limitations of RE-MC methods for systems with many degrees of freedom are well known [14], especially when sampling phase changes that are separated by free energy barriers. Various optimisation approaches have been proposed for RE-MC to cross free energy barriers, particularly in the field of protein folding. Most of these methods emphasise the importance of the distribution of index parameter values among the replicas, yielding an efficient exchange protocol. While most other work on RE-MC uses the temperature as the index parameter, giving so-called 'parallel tempering' Monte Carlo [12], we aim for an optimal grid of K_{hyd} values. The simplest distribution of index parameters, proposed in Ref. [15], is a grid defined by a geometric progression (i.e. $K_{\text{hyd},i+1} = \text{constant} \times K_{\text{hyd},i}$). However, this approach has been shown to be inefficient when a phase transition is present within range of index parameter values [14]. Some authors, aware of the bottleneck around the phase transition, use an adaptive grid of index parameters yielding an *equal* acceptance probability of replica exchanges [16–18]. On the other hand, other authors emphasise the importance of maximising the diffusion of configurations near the phase transition [19–21]. Towards this end, they advocate using an adaptive grid that yields a *peaked* distribution (which we have parametrised with a Gaussian function) of replica exchange probabilities around the bottleneck. In the present work, we have compared the effectiveness of these three different distributions: (i) a fixed geometric distribution, (ii) an adaptive grid that yields a nearly constant probability of replica exchanges and (iii) an adaptive grid that yields a peaked probability of replica exchanges. Below we show that using adaptive grids of K_{hyd} is essential for enhancing the performance of RE-RxMC.

Zeolites we have assembled so far with RE-RxMC are among those with the smallest unit cells. The goal of this article is to enhance the performance of this RE-RxMC method to build zeolites with larger unit cells. First, we identify MC parameters that can be varied to speed up the simulations and to polymerise silica more efficiently. Second, we develop and test fixed and adaptive grids of K_{hyd} values to optimise replica exchange as discussed above. These enhancements were tested in a case study on the RE-RxMC construction of the sodalite (SOD) zeolite framework. With 12 tetrahedra per unit cell, SOD is the largest zeolite framework found in our previous work [3]. SOD is not a trivial system but its size allows us to perform a large number of test simulations. Finally, we show how enhancing the method with SOD makes possible the RE-RxMC construction of the AWW framework, with 24 tetrahedra per unit cell.

The present article is organised as follows. Section 2 describes the model and method we are using; Section 3 shows the results of the enhancements made to RE-RxMC and how they impact the probability of finding the crystalline states of SOD and AWW; finally, in Section 4 we summarise the main findings of this work and we discuss the outlook for additional progress with this approach.

2. Model and methods

Here we describe the model and methods behind our Replica Exchange Reactive Ensemble Monte Carlo (RE-RxMC) simulations, which produce crystalline structures of zeolites through the self-assembly of SiO_4 tetrahedra. The goal of this work is to improve, and if possible to optimise, the performance of the RE-RxMC method – to push this approach to the assembly of zeolites with larger unit cells than those studied previously [3]. In our previous work, we applied this method to produce nonporous and nanoporous crystals for systems with simulation cells containing between 5 and 16 SiO_4 tetrahedra. Below, we use the all-silica sodalite (SOD) system as the basis for a case study to improve the RE-RxMC approach. SOD contains 12 SiO_4 tetrahedral [22], and as such is both tractable and sufficiently challenging to test our method. We then apply the improved RE-RxMC to search for crystals of zeolites with 24 SiO_4 tetrahedra with attention on the AWW framework [22].

We have focused on improving two aspects of the method: the RxMC parameters controlling frequencies of attempted reactive and non-reactive moves, and the grid of hydrolysis equilibrium constants $\{K_{\text{hyd}}\}$ among the replicas in REMC. For computational efficiency, we have used the best (adaptive) $\{K_{\text{hyd}}\}$ grid when improving the RxMC parameters, and the best set of RxMC parameters when adapting the grid of $\{K_{\text{hyd}}\}$ values. Equilibration is not guaranteed due to the glassiness of the silica model; for this reason we have performed 20 statistically independent RE-RxMC runs (2 million steps each) in the case of SOD and 5 independent RE-RxMC runs (10 million steps each) for AWW. To quantify how improving RxMC parameters and adapting $\{K_{\text{hyd}}\}$ grids improve RE-RxMC performance, we have counted the fraction of simulations that yield crystal structures, denoted as $N_{\text{crystals}}/N_{\text{runs}}$ in Section 3.

2.1. Reactive model of silica polymerisation

The details of our 'spring-tetrahedron' model have been extensively described in our previous work [10]; as such, we outline only its main features here (see Figure 1). Silicon atoms are located in the centres of SiO_4 tetrahedra and are represented by hard spheres of diameter $\sigma_{Si-Si} = 2.0$ Å. Two oxygenic species are considered: hydroxyl groups (OH) if they are the terminal group, and bridging oxygens (BO) if they connect two

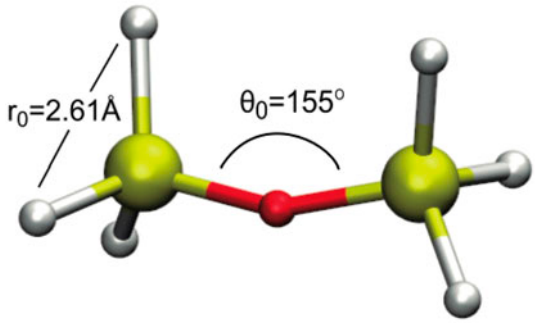

Figure 1. (colour online) Silica tetrahedral model: Silicon atoms (yellow) are in the centres of the tetrahedra. Oxygenic species are terminal hydroxyls (white) and bridging oxygens (red). Tetrahedral shape and flexibility are ensured by springs between all vertices.

tetrahedra. Harmonic springs between BO and/or OH groups ensure structural integrity and flexibility of each tetrahedron. The potential energy due to the distortion of each tetrahedron is given by $U_1 = \sum_{i=1}^{3}\sum_{j=i+1}^{4} k_S(|\boldsymbol{r}_i - \boldsymbol{r}_j| - r_0)^2/2$. Here k_S is the spring constant set to 851 kJ mol^{-1}Å^{-2} in Ref. [10], r_0 is the equilibrium distance between oxygenic species (r_0 = 2.61 Å based on the Si–O bond length of 1.6 Å and the tetrahedral SiO_4 geometry [10]), and $\boldsymbol{r}_i$ is the position of the ith OH/BO vertex.

Our model accounts for the flexibility of tetrahedral networks using an angular harmonic potential that restricts the Si–O–Si angles in the range of the observed values in silica materials: 130°-180° [23]. This angular term is expressed as follows: $U_2 = k_A(\cos\theta - \cos\theta_0)^2/2$, where k_A is the angular force constant (set to 226.74 kJ mol^{-1} Å^{-2} in Ref. [10]), θ_0 is the Si–O-Si reference angle (determined to be 155° by periodic DFT calculations in Ref. [23]), and θ is the Si–O–Si angle between each pair of connected tetrahedra. The model parametrisation used herein has been shown to reproduce structures and bulk moduli of many all-silica zeolites [10], as well as structures of nanoporous phosphates [3]. We note that a re-parametrisation of this model may be required when heteroatoms such as B or Ge with very different bond lengths and angles are present in the system [24].

2.2. Reactive Ensemble Monte Carlo

To represent the polymerisation process among tetrahedra, we have performed Monte Carlo simulations using the 'spring tetrahedron' model [8,9] in the Reactive Ensemble (RxMC) [25]. As with our spring-tetrahedron model, our implementation of RxMC has been described in detail in previous publications [8,9] (see in particular *Section 1.2* in Ref. [9]); here we briefly outline our RxMC approach. This sampling method includes tetrahedral rotations; translations of OH/BO species, tetrahedra centres-of-mass and entire clusters; and reactive moves that are essential for simulating network formation. Reactive moves reproduce the condensation and hydrolysis reactions that occur during the silica polymerisation process: $\equiv Si - OH + HO - Si \equiv \leftrightharpoons \equiv Si - BO - Si \equiv + H_2O$. Although explicit water molecules are not present in our simulations, water influences the energetics of silica condensation and hydrolysis reactions [26], which in turn establish physically meaningful K_{hyd} values [3].

Condensations may be inter- or intra-cluster, depending on whether the pair of OH/BO involved belong to different clusters or to the same cluster, respectively. Force bias Monte Carlo moves are performed to relax highly distorted structures that result after each intra-cluster condensation move. The force bias moves are described in Ref. [9]. The success of this model and sampling approach has been established by comparison with various structural data such as distance and angle data from detailed molecular dynamics simulations [27], ring-size distributions as determined by reverse Monte Carlo [28], and the evolution of network formation probed by the time-dependent Q_n distribution as measured by ^{29}Si NMR [29].

Tuning the attempt probabilities for the various kinds of MC moves in RxMC, and the various maximum displacement values, has proved to be essential for improving the efficiency of the RE-RxMC method for constructing zeolite crystals. In Table 1, *Section 3.1*, we show the set of RxMC parameters we have found to yield the best performance of the RE-RxMC method for generating crystals of SOD.

2.3. Replica exchange reactive ensemble Monte Carlo (RE-RxMC)

We showed in our previous article [3] how RxMC simulations find crystalline phases of alumino-silicate and alumino-phosphate zeolites using the 'spring-tetrahedron' model along with the Replica Exchange protocol, reported by Turner et al. [13], for traversing rough energy landscapes of silica systems. The protocol to exchange adjacent replicas and its probability is given in Appendix 1 of the present article. In particular, we perform Replica Exchange and Reactive Ensemble Monte Carlo moves with probabilities of 0.01 and 0.99, respectively.

The hydrolysis equilibrium constant, K_{hyd}, controlling network formation is taken as the index parameter characterising the replicas, as was done in Ref. [3]. The meaning of K_{hyd} as an index parameter can be conceptually seen as serving roughly like an effective temperature in this Replica Exchange approach: at high K_{hyd} values, the likelihood of silica polymerisation is low and the system is formed by disconnected tetrahedra; at low K_{hyd} values, the polymerisation probability is higher and the system forms connected networks. K_{hyd} is distributed among the m replicas as follows: $K_{hyd,1} < K_{hyd,2} < \cdots < K_{hyd,m}$. The minimum and maximum values in our simulations are fixed to the values used in our earlier work, chosen to drive network formation ($K_{hyd,min}$=10^{-6}) and silica hydrolysis ($K_{hyd,max}$=537), respectively.

Optimising the distribution of the intermediate K_{hyd} values among the replicas is essential to ensure efficient diffusion throughout the replica space. For this reason, we have implemented two adaptive methods for resetting the index parameters among the replicas, on the fly of the RE-RxMC simulations, to reproduce a certain replica exchange acceptance ratio as a function of the index parameter, denoted as $acc_{target}(K_{hyd})$. We have considered the following three K_{hyd} grids:

- Constant geometric grid of K_{hyd} along the replicas: $K_{hyd,i+1} = c \times K_{hyd,i}$. For consistency with previous work, we have taken c = 4 and i = 1, . . ., 16).
- Adaptive K_{hyd} grid leading to uniform replica exchange acceptance probabilities. In this case $acc_{target}(K_{hyd})$ was set to 0.4, constant for all K_{hyd} values.

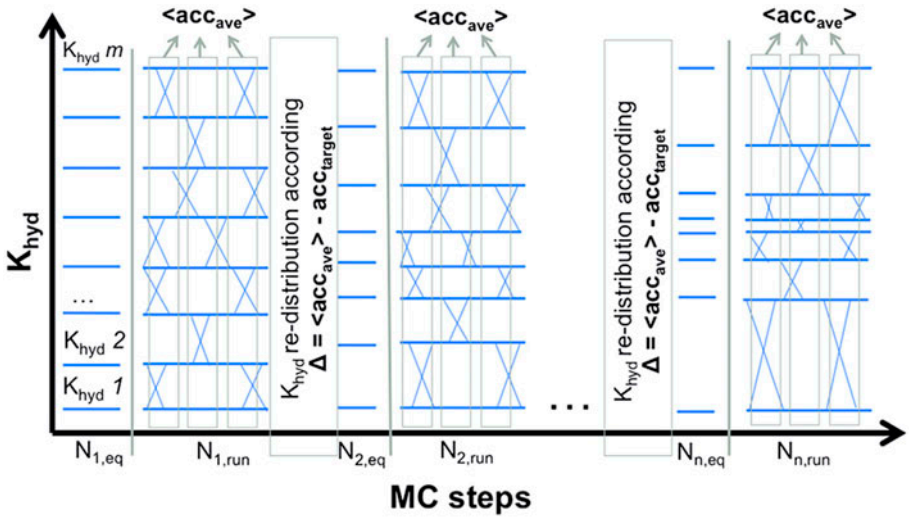

Figure 2. (colour online) Description of the adaptive RE-RxMC method. The system is allowed to equilibrate during N_{eq} Reactive Ensemble Monte Carlo steps at the corresponding K_{hyd} value of each replica. Afterwards, information will be transferred between the replicas and statistics will be collected during $N_{RE-RxMC}$ Replica exchange reactive ensemble MC steps. The index parameter, K_{hyd}, is redistributed along the replicas on the fly to generate the desired probability of exchange between each pair of replicas, $acc_{\text{target},ij}$.

- Adaptive K_{hyd} grid leading to a maximum of the replica exchange acceptance probabilities around the transition region determined. In this case, $acc_{\text{target}}(K_{\text{hyd}})$ is the following Gaussian function with parameters $a = 0.35$, $b = 0.01638$ and $c = 1.01$:

$$acc_{\text{target}}(K_{\text{hyd}}) = ae^{-\frac{(K_{\text{hyd}}-b)^2}{2c^2}} + 0.25 \qquad (1)$$

The adaptive method that resets the K_{hyd} values along the m replicas to reproduce a certain $acc_{\text{target}}(K_{\text{hyd}})$ works as follows (see also Figure 2):

(1) N_{eq} RxMC steps are performed to equilibrate the system at the corresponding K_{hyd} value of each replica. We have considered $N_{\text{eq}} = 10,000$ steps for SOD and $N_{\text{eq}} = 20,000$ for AWW.
(2) During N_{run} RE-RxMC steps, configurations are exchanged between the replicas, in addition to the various RxMC moves, and statistics are collected. In particular $acc_{\text{ave},ij}$, that is the probability of accept replica exchanges between replicas i and j. We have considered $N_{\text{run}} = 50,000$ steps for SOD and $N_{\text{run}} = 100,000$ for AWW.
(3) New K'_{hyd} values are determined as follows:

$$K'_{\text{hyd}}(i) = K'_{\text{hyd}}(j)\left(\frac{K_{\text{hyd}}(i)}{K_{\text{hyd}}(j)}\right)^{(1+\Delta)^2} \qquad (2)$$

where $\Delta = acc_{\text{ave},ij} - acc_{\text{target},ij}$, for each pair of replicas i and $j = i + 1$, and with the constraint that $K_{\text{hyd,min}}$=10^{-6} and $K_{\text{hyd,max}}$=537.
Thus, if the resulting $acc_{\text{ave},ij}$ is smaller than $acc_{\text{target},ij}$ then $\Delta < 0$ and replicas i and j come closer in K_{hyd} values; otherwise $\Delta > 0$ and replicas i and j become further separated in K_{hyd}-space.

(4) We repeat this process n times, until the end of the simulation.

All RE-RxMC simulations were initiated with random configurations and performed under three-dimensional periodic boundary conditions. The dimensions of each simulation box and the number of tetrahedra were fixed given the unit cell and the density of each target zeolite: SOD has 12 tetrahedra in a cubic unit cell (8.9561 Å × 8.9561 Å × 8.9561 Å), and AWW has 24 tetrahedra in a tetragonal unit cell (13.634 Å × 13.634 Å × 7.627 Å) [22].

RE-RxMC simulations were performed in parallel using a different number of replicas for each system, varying from 16 to 28 – large enough to yield the desired acc_{target}. The temperature was kept fixed at $T = 300$ K for all replicas. In future work, we will examine the effects of increasing temperature, which allows for more distorted silica tetrahedra, on the Monte Carlo construction of zeolites. Simulation length depends on the system under study: 2 million steps was found to be sufficient for the case of study SOD, and 10 million steps were required for AWW. Crystals found using RE-RxMC were relaxed using 5 million MC steps in the canonical ensemble at a lower temperature, $T = 50$ K, to reduce thermal distortions from ideal crystal structures before analysing the resulting structures. We have computed the X-ray diffraction patterns of the crystals we found, and compared them with those of the target zeolites obtained from the IZA database [22]. In both cases, we used Debyer software [30] on 25 × 25 × 25 periodic extensions of unit cells to compute the XRD patterns.

3. Results and discussion

Our aim in this article is to improve the efficiency of RE-RxMC simulations in constructing zeolite crystals. To do so, we have enhanced the parameters and RE-RxMC protocols used in Ref. [3]. First, we have tuned the RxMC parameters with the aim

Table 1. Probabilities of attempt for each of the RxMC moves: displacement of OH/BO, tetrahedra and clusters, rotation of tetrahedra, condensation and hydrolysis reactions. Left column corresponds to the set of parameters used in Ref. [3] and Right column to the set of parameters we have found to give the best results, i.e. leading to the formation of more crystals.

MC move	Chien et al. [11]	Improved
	Probability of Attempt	
OH/BO displacement	0.1515	0.05
Tetrahedra displacement	0.3030	0.075
Tetrahedra rotation	0.0	0.075
Cluster displacement	0.1515	0.05
Tetrahedra displacement in Cluster	0.3637	0.075
Tetrahedra rotation in Cluster	0.0	0.075
Condensation	0.01515	0.30
Hydrolysis	0.01515	0.30
	Maximum Displacement	
OH/BO displacement	0.01Å	0.025Å
Tetrahedra displacement	0.01Å	0.2Å
Tetrahedra rotation	0.00°	0.02°
Cluster displacement	0.01Å	0.1Å
Tetrahedra displacement in Cluster	0.01Å	0.025Å
Tetrahedra rotation in Cluster	0.00°	0.02°
	$N_{crystals}/N_{runs}$	
SOD (after 2M MC steps)	2/20	19/20

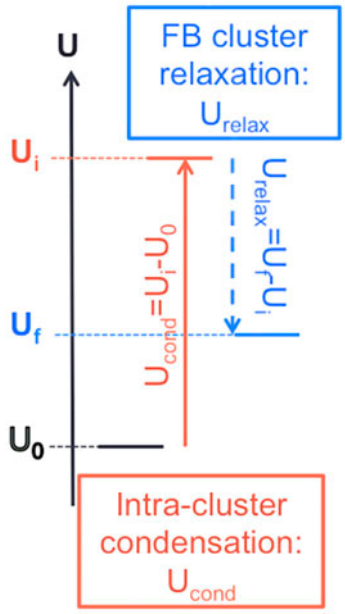

Figure 3. (colour online) Diagram showing the energy relaxation process through the Force Bias mechanism. U_0 is the initial energy of the system, U_i is the energy after the highly distorting intra-cluster condensation move, and U_f is the final energy after the FB relaxation process.

of achieving faster equilibration for each replica. Second, we have modified the replica exchange protocol to improve the transfer of information between the replicas. Finally, we show that only by implementing both enhancements are we able to find the emergence of a zeolite with a significantly bigger unit cell: AWW zeolite. As an aside, we note that other zeolites with 24 SiO_2 units per unit cell could also be studied; we will discuss others in a forthcoming article on the role of structure directing agents.

3.1. Reactive moves

Reactive moves (condensation and hydrolysis) are essential to achieve silica polymerisation, and to reach the equilibrium of the system at a certain K_{hyd} value. We increased the number of attempted condensation and hydrolysis reactions by multiplying by 20 the attempt probability of reaction moves used in our previous work [3]. As a result, the attempt probabilities were chosen as 60% for reactive moves and 40% for translation and rotations. In addition, the maximum allowed atomic displacements have been increased to more efficiently sample configuration space. Table 1 shows the parameters used by Chien et al. [3] and the new set of enhanced parameters we have implemented herein. We note that other parameters were studied, but for brevity we only indicate the best set we have found.

The resulting configuration after each intra-cluster condensation move may have a very high energy associated with stretched spring bonds ($U_{cond} = U_i - U_0$ is the energy change after an intra-cluster condensation). Force bias (FB) moves are essential for intra-cluster polymerisation and ring formation because they relax the distorted structure by draining the energy excess ($U_{relax} = U_f - U_i$ is the energy reduction after the FB cluster relaxation, see Figure 3). The problem is that FB moves are highly time-consuming. We have investigated the number (N_{FB}) and maximum displacement (dr_{FB}) of FB moves to obtain more efficient simulations. We seek to maximise the amount of energy drained per Force Bias move, U_{relax}/N_{FB}; to minimise CPU time required; and to maximise the probability of accepting the intra-cluster condensation move (see results in Table 2). After this analysis, we have concluded that 50 FB attempts and 0.2 Å maximum displacement are the best choices.

Table 2. Ability of the FB method to reduce the energy (U_{relax}) of the highly distorted configuration obtained after an intra-cluster condensation move for different number of Force Bias moves (N_{FB}) and different maximum displacement values (dr_{FB}). The CPU time spent and the probability to accept the condensation moves (acc_{cond}) are also useful to decide the best choices.

N_{FB}	dr_{FB} (Å)	CPU time (s)	U_{relax} (kJ/mol)/N_{FB}	acc_{cond}
500	0.01	25.1	0.83	0.025
200	0.1	13.0	1.95	0.039
200	0.2	8.0	1.86	0.092
50	0.1	3.0	6.79	0.034
50	**0.2**	**2.7**	**9.077**	**0.093**
20	0.1	0.78	21.6	0.0094
20	0.2	0.72	22.34	0.029

The RxMC parameters used in Ref. [3] were inspired by previous research on reproducing silica network formation kinetics [8,9] in comparison with NMR data [29]. Here we are concerned about speeding up the network formation to achieve the crystalline zeolite state in a minimum amount of time. We have found that performing more reactive moves during the simulation allows each replica to reach its equilibrium degree of polymerisation more easily. The bottom row of Table 1 shows that this feature is directly reflected in the probability of finding the crystal state. In particular, the number of RE-RxMC simulations that produce crystals was found to increase from 2, with the old parameters, to 19 with the new ones, an increase of nearly a factor of 10 in crystallising efficacy.

3.2. Optimisation of the K_{hyd} distribution

We have developed a feedback method for resetting the K_{hyd} values among the replicas during the RE-RxMC simulation to

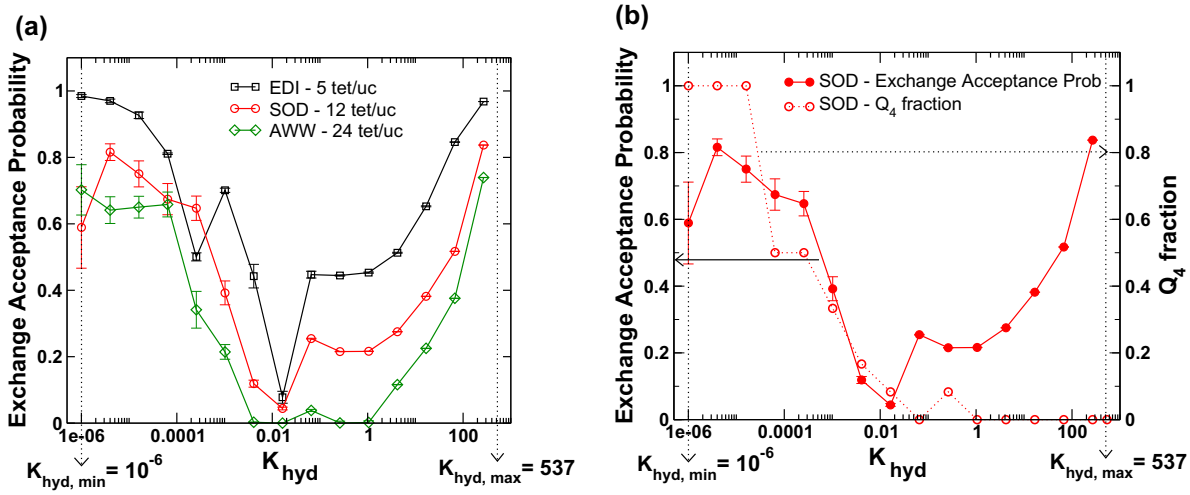

Figure 4. (colour online) Probability of replica exchange acceptance as a function of the index parameter K_{hyd} of SOD in comparison with (a) other zeolites of different size and (b) the final value of Q_4 molar fraction (see text for definition) for each replica. All results correspond to a constant geometric distribution of K_{hyd} values among 16 replicas.

Table 3. 20 statistically independent simulations of SOD were performed considering the parameters and method from Ref. [3] (first row), and for the three K_{hyd} distributions presented in this article (next three rows). $N_{replicas}$ is the number of replicas considered in each case; $N_{crystals}$ is the number of simulations leading to the SOD crystalline structure; and N_{trips} is the number of times a replica goes from $K_{hyd,max}$ to $K_{hyd,min}$ or vice versa, a measure of replica exchange efficiency.

	$N_{replicas}$	$N_{crystals}/N_{runs}$	N_{trips}
Chien et al. [3]	16	0/20	13.9
Constant geometric	16	11/20	31.5
Constant geometric	20	7/20	37.7
Adaptive Uniform fitting	16	15/20	35.3
Adaptive Gaussian fitting	20	19/20	51.9

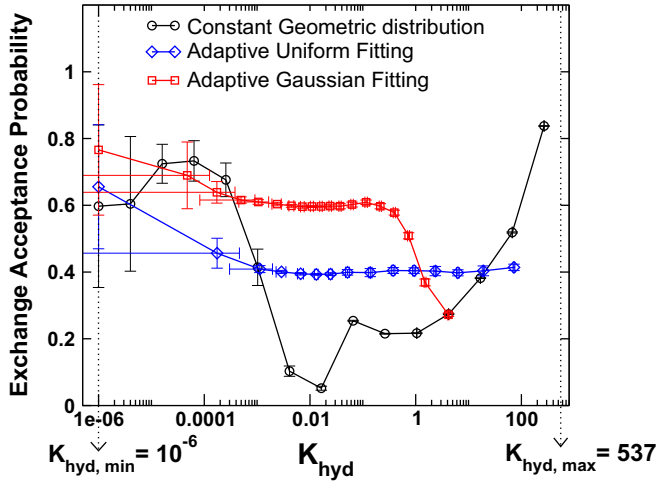

Figure 5. (colour online) Three different replica distributions considered in this paper and their correspondent replica exchange acceptance probabilities for zeolite SOD (average over 20 independent simulations).

obtain the desired acceptance ratios of replica exchanges. We have compared the ability of our method to find SOD crystals using three K_{hyd} distributions: constant geometric distribution, an adaptive grid leading to a uniform distribution of probabilities, and an adaptive grid giving a Gaussian distribution of probabilities.

The constant geometric distribution gives rise to a low replica exchange probability in the transition region. This behaviour is shown in Figure 4(a) where replica exchanges near K_{hyd} = 0.016 are quite rare, as indicated by the minimum in the replica exchange probability. Different lines in Figure 4 correspond to three different zeolites: EDI (5 tetrahedra per unit cell), SOD (12 tetrahedra per unit cell) and AWW (24 tetrahedra per unit cell), showing how this bottleneck in the replica exchanges became worse as the system size increases. Figure 4(b) shows that replicas at the lowest and highest K_{hyd} values correspond to two different phases characterised by the Q_4 molar fraction, which is the fraction of tetrahedra that are fully connected with all four oxygens as bridging oxygens. We see that at high K_{hyd} values there are no tetrahedra connected ($Q_4 = 0$), and at low K_{hyd} tetrahedra are highly connected, meaning network formation is occurring ($Q_4 > 0$).

Figure 5 compares RE-RxMC results from the constant geometric grid of K_{hyd} values, the adaptive constant probability approach, and the adaptive Gaussian probability method. The black line corresponds to a geometric distribution (ratio equal to 4) of K_{hyd} values; the red line shows replica exchange probabilities for the adaptive approach with a constant target probability of 0.4; and the blue line shows results for the adaptive method seeking a Gaussian distribution of exchange probabilities centred on K_{hyd} = 0.016. Figure 5 shows that both adaptive methods gather more replicas near the bottleneck at K_{hyd} = 0.016 to achieve more efficient replica exchanges. We note in Figure 5 that the replica exchange probabilities associated with $K_{hyd,min} = 10^{-6}$ exhibit the largest non-systematic errors explaining their deviations from target values.

Table 3 shows that the adaptive replica methods enhance the ability of RE-RxMC to build the crystal structure of SOD. The first row corresponds to the RxMC parameters and the RE protocol used in Ref. [3]. We note that no SOD crystal was found after 2 million RE-RxMC steps across 20 statistically

independent runs. In our previous publication [3], we needed as many as 3.3 million RE-RxMC steps to produce SOD crystals. The next four rows of Table 3 show results using the improved RxMC parameters given in Tables 1 and 2, along with the three replica exchange protocols described above. Comparing rows 2 and 3 in Table 3 shows that adding more replicas to the Constant geometric grid approach does not enhance the efficiency of the method. We can see in Table 3 that the number of independent runs yielding the SOD crystal structure increases to 7–11/20 with a constant grid, 15/20 with a constant probability adaptive grid and 19/20 with the Gaussian probability adaptive grid. This result suggests that the Gaussian probability is optimal among those studied herein.

The Gaussian adaptive grid causes better diffusion of replicas along the K_{hyd} range, which is reflected in how often a replica travels from $K_{hyd,max}$ to $K_{hyd,min}$ or vice versa. We have measured this feature by counting how many times the hydrolysis equilibrium constant characterising each replica goes from the highest to the lowest value or vice versa during the simulation (N_{trips}) and compared this number, averaged over the 20 statistically independent runs, in the last column of Table 3; this shows how the enhancements implemented here have improved the RE-RxMC diffusion efficiency. In addition, a visual representation of the diffusion of replicas along the K_{hyd} range is shown in Figure 6. We have plotted in this figure how the replica starting at the minimum value of K_{hyd} traverses the K_{hyd} space for all the 20 independent RE-RxMC runs. Panels (a), (b), (c) and (d) in Figure 6 correspond to rows 1, 2, 4 and 5, respectively, in Table 3. Figure 6 shows visually the progressive improvement in replica diffusion from (a) to (b) to (c) to (d), consistent with the progressive larger number of crystals in Table 3. In particular, the substantial gaps in Figure 6(a) explain why no crystal emerged in these simulations; while Figure 6(d) shows the most complete diffusion, consistent with generating the most crystals in Table 3.

In addition to the case study on SOD, we have reproduced the self-assembly of the following structures also found in Ref. [3]: EDI and ATT [22]. By implementing improved RxMC parameters and more efficient replica exchange, we have found more efficient crystallisation during RE-RxMC. In particular, the EDI structure was found in Ref. [3] after 3 million RE-RxMC steps; SOD required 3.3 million steps; and ATT required 50 million RE-RxMC steps. Using the improved RE-RxMC approach, these crystals emerge after 0.2, 1.0 and 1.5 million

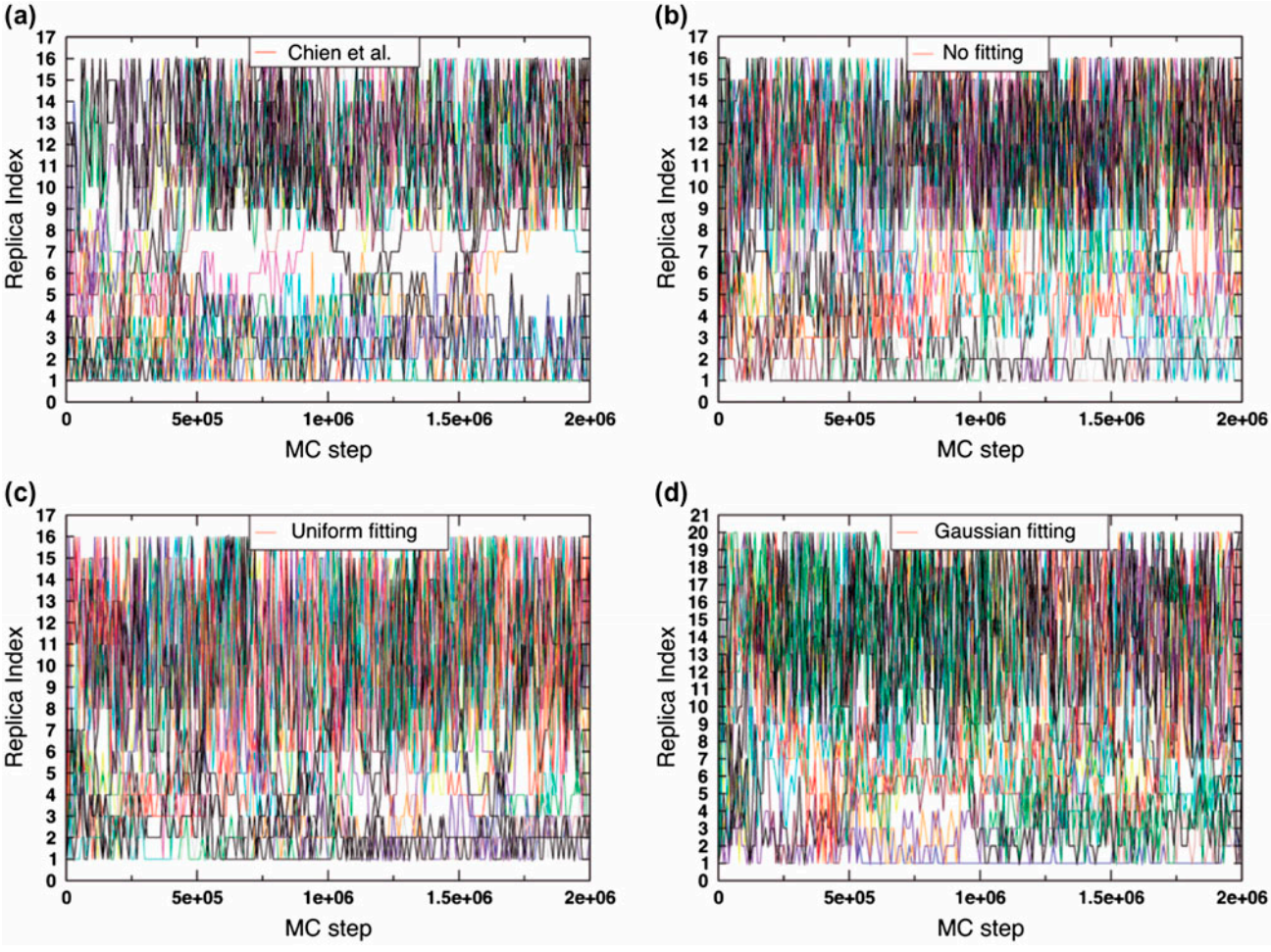

Figure 6. (colour online) Diffusion of the configuration that started at the lowest K_{hyd} value along the replica space for SOD: (a) Unoptimised simulations corresponding to our previous publication [3]; Best set of RxMC parameters so far plus (b) constant geometric K_{hyd} distribution; (c) uniform adaptive K_{hyd} grid; and (d) Gaussian adaptive K_{hyd} grid.

RE-RxMC steps, respectively, representing substantial speed-ups in the range of 3.3–33.3.

3.3. Application to bigger systems: AWW

We have applied the RE-RxMC enhancements found in the case study on SOD – the largest zeolite found in our previous work with a unit cell volume of 718 Å^3 – to AWW zeolite, a system with 24 tetrahedra in a unit cell of 1418 Å^3 (see Figure 7). In our study applying RE-RxMC to the AWW zeolite, we investigated the efficacies of all four approaches shown in Table 3 and Figure 6. We found that only the combination of improved RxMC parameters and the Gaussian adaptive grid was sufficient to generate the AWW crystal, running five statistically independent runs, each with 28 replicas for 10 million steps. Only one out of five RE-RxMC runs produced an AWW crystal, whose structure is shown in Figure 7. We note that comparison simulations on AWW using the constant geometric grid with 28 replicas produced no crystals, indicating that adding replicas alone does not improve the performance of the method in constructing crystals.

Figure 8(a) shows average replica exchange probabilities as a function of K_{hyd} for different K_{hyd} distributions: constant geometric (black), adaptive grid with uniform fitting (blue) and adaptive grid with Gaussian fitting (red). Figure 8(a) emphasises how the adaptive grid approach overcomes the bottleneck in the transition region by bunching replicas in this region of K_{hyd} space. Figure 8(b) shows the Q_4 molar fraction of the final structure at each replica, for all three K_{hyd} distributions; these data correspond to the simulation showing the highest Q_4 value for the lowest K_{hyd}, not the average value over all the independent simulations. Because the Q_4 molar fraction denotes the number of Si atoms bonded to four bridging oxygens divided by the total number of Si atoms, Q_4 equals 1 only if the structure is fully connected. Figure 8(b) shows that only the Gaussian adaptive grid can produce such a fully connected structure, the

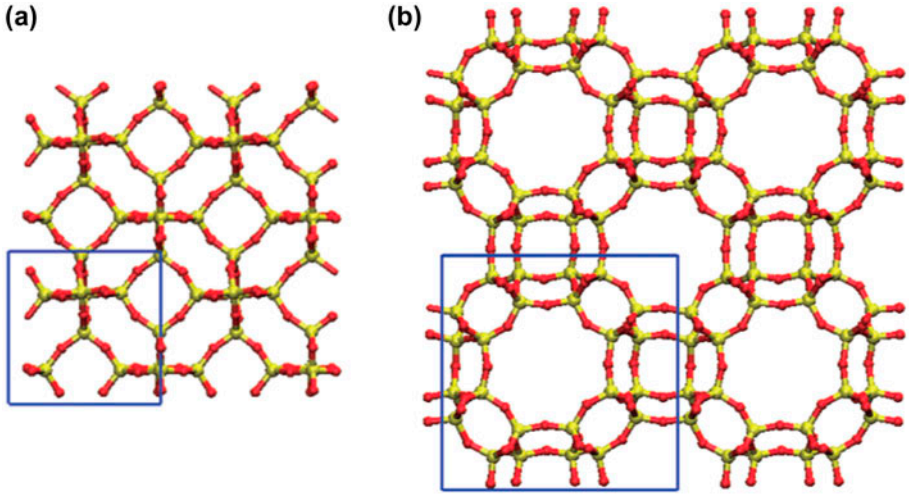

Figure 7. (colour online) Snapshots a 2 × 2 × 2 periodic extensions of (a) SOD and (b) AWW frameworks. Unit cell boundaries are drawn in blue.

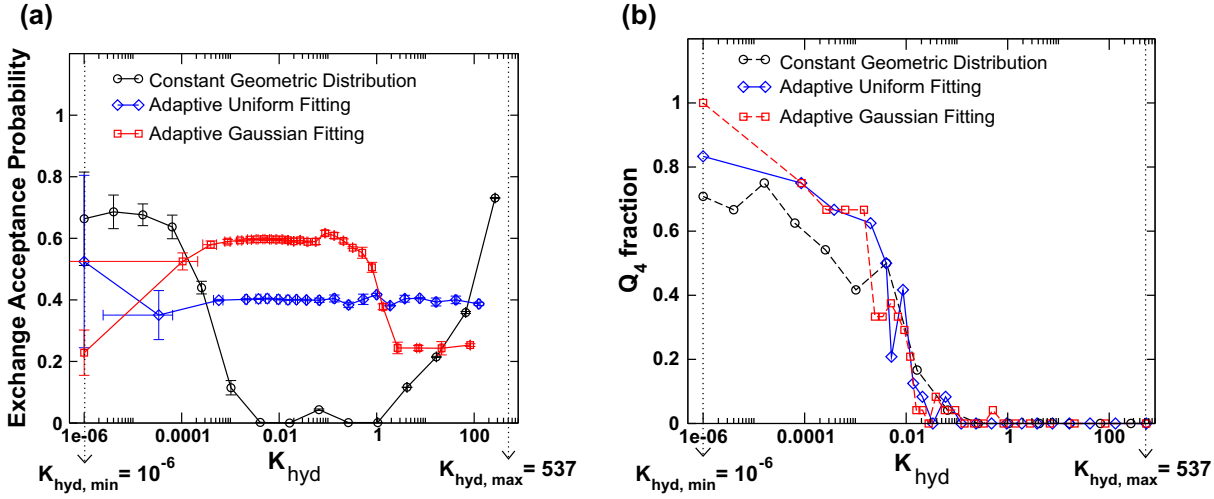

Figure 8. (colour online) (a) Replica exchange probabilities (averaging over the five independent simulations), and (b) Final value of Q_4 mole fraction as a function of K_{hyd} values, for AWW. Constant geometric distribution (black), adaptive uniform (blue), and adaptive Gaussian (red) data shown.

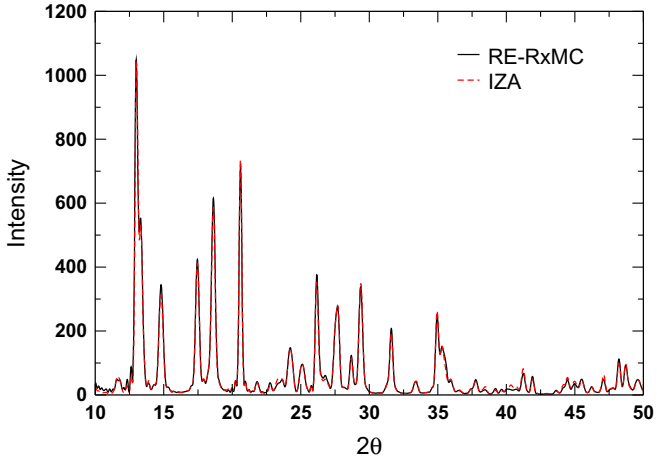

Figure 9. (colour online) XRD patterns for AWW structure. We compare the structure resulting from RE-RxMC (black solid line) with the experimental coordinates taken from IZA database (dashed red line).

AWW zeolite crystal, which appears only in the replica with the lowest K_{hyd} value.

Figure 9 plots the XRD patterns of the AWW crystal structure generated by RE-RxMC, compared with that of the experimental XRD found in the IZA database [22]. Figure 9 shows excellent, essentially perfect, agreement between the IZA structure and the simulated AWW framework using RE-RxMC.

4. Summary and conclusions

We have presented improvements to the efficiency of replica exchange reactive ensemble Monte Carlo (RE-RxMC) simulations, applied with our 'spring-tetrahedron' model, for constructing the crystalline structures of zeolites in a reactive model of silica polymerisation through the self-assembly of tetrahedral silica units. To investigate improvements to the RE-RxMC approach, we have engaged in a case study on crystallising the sodalite (SOD) zeolite, whose unit cell contains 12 tetrahedra. First, we have tuned the parameters of the reactive Monte Carlo moves to speed up the sampling of silica polymerisation. Second, we have devised an adaptive protocol that resets the values of the hydrolysis equilibrium constant, K_{hyd}, among the replicas to generate certain distributions of replica exchange acceptance probabilities. Such adaptive grids have proved useful for enhancing the transfer of configurations between the replicas and, in turn, the ability of the RE-RxMC method to find the crystalline states. By enhancing both the reactive Monte Carlo parameters and the replica exchange protocol, we have found speed-ups of the Monte Carlo crystallisation in the range of 3.3–33, depending on the system. In addition, we have extended the RE-RxMC method to crystallise AWW, a zeolite with a unit cell twice as large as that of SOD.

The 'spring-tetrahedron' model taken together with advanced sampling methods such as RE-RxMC has proved useful for simulating the construction of zeolite frameworks. Such simulations constitute an important initial step towards a deeper understanding of how zeolites actually crystallise in laboratory experiments. To generate such deeper understanding, the following next steps need to be considered: (i) the roles of structure-directing agents in zeolite crystallisation; (ii) more physically relevant crystallisation pathways obtained by rare event sampling approaches; and (iii) more industrially relevant zeolite topologies such as MFI (e.g. ZSM-5) and FAU (e.g. US-Y), which feature much larger unit cells – containing 96 and 192 tetrahedra, respectively. In addition, simulations performed in a constant pressure ensemble will remove the constraint of constant volume, and thus will bring our simulations even closer to providing information about the cell parameters and crystal structure simultaneously. Altogether this represents a promising roadmap towards a deeper understanding of zeolite formation.

Acknowledgements

The authors acknowledge Dr. Szu-Chia Chien for all the assistance with the RE-RxMC code. We also thank computational resources provided by the Massachusetts Green High-Performance Computing Center (MGHPCC).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the US Department of Energy [Contract No. DE-FG02-07ER46466].

References

[1] Auerbach SM, Karrado KA, Dutta PK. Handbook of zeolite science and technology. New York (NY): Marcel Dekker; 2003.
[2] Auerbach SM, Fan W, Monson PA. Modelling the assembly of nanoporous silica materials. Int Rev Phys Chem. 2015;34:35–70.
[3] Chien SC, Auerbach SM, Monson PA. Reactive ensemble Monte Carlo simulations of silica polymerization that yield zeolites and related crystalline microporous structures. J Phys Chem C. 2015;119:26682.
[4] Catlow CRA. Modelling and predicting crystal structures. Interdisciplinary Sci Rev. 2015;40:294–307.
[5] Foster MD, Treacy MMJ. A database of hypothetical zeolite structures. 2006. Available from:http://www.hypotheticalzeolites.net.
[6] Li Y, Yu J. Hypothetical zeolite database. 2003. Available from: http://mezeopor.jlu.edu.cn/hypo/
[7] Earl DJ, Deem MW. Toward a database of hypothetical zeolite structures. Ind Eng Chem Res. 2006;45:5449.
[8] Malani A, Auerbach SM, Monson PA. Probing the mechanism of silica polymerization at ambient temperatures using monte carlo simulations. J Phys Chem Lett. 2010;1:3219–3224.
[9] Malani A, Auerbach SM, Monson PA. Monte Carlo simulations of silica polymerisation and network formation. J Phys Chem C. 2011;115:15988–16000.
[10] Astala R, Auerbach SM, Monson PA. Normal mode approach for predicting the mechanical properties of solids from first principles: application to compressibility and thermal expansion of zeolites. Phys Rev B. 2005;71:014112.
[11] Chien SC, Auerbach SM, Monson PA. Modeling the self-assembly of silica-templated nanoparticles in the initial stages of zeolite formation. Langmuir. 2015;31:4940–4949.
[12] Earl DJ, Deem MW. Parallel tempering: Theory, applications and new perspectives. Phys Chem Chem Phys. 2005;7:3910–3916.
[13] Turner C, Brennan J, Lisal M. Replica exchange for reactive Monte Carlo simulations. J Phys Chem C. 2007 July;111:15706–15715.

[14] Predescu C, Predescu M, Ciobanu C. The imcomplete beta function law for parallel tempering sampling of classical canonical systems. J Chem Phys. 2004;120:4119.
[15] Kofke DA. On the acceptance probability of replica-exchange Monte Carlo trials. J Chem Phys. 2002;117:6911.
[16] Rathore N, Chopra M, de Pablo JJ. Optimal allocation of replicas in parallel tempering simulations. J Chem Phys. 2005;122:024111.
[17] Guidetti M, Rolando V, Tripiccione R. Efficient assignment of the temperature set for parallel tempering. J Comput Phys. 2012;231:1524.
[18] Bittner E, Nubaumer A, Janke W. Make life simple: unleash the full power of the parallel tempering algorithm. Phys Rev Lett. 2008;101:130603.
[19] Predescu C, Predescu M, Ciobanu CV. On the efficiency of exchange in parallel tempering Monte Carlo simulations. J Phys Chem B. 2005;109:4189.
[20] Katzgraber H, Trebst S, Huse D, et al. Feedback-optimized parallel tempering Monte Carlo. J Statistical Mech. Theory Exp. 2006;2006:P03018.
[21] Trebst S, Troyer M, Hansmann U. Optimized parallel tempering simulations of proteins. J Chem Phys. 2006 May;124:174903.
[22] Baerlocher C, McCusker LB, Olson D. Atlas of zeolite framework types. 6th ed. Elsevier; 2007. Available from: http://www.iza-structure.org/databases/
[23] Astala R, Auerbach SM, Monson PA. Density functional theory study of silica zeolite structures: stabilities and mechanical properties of SOD, LTA, CHA, MOR and MFI. J Phys Chem B. 2004;108:9208–9215.
[24] Jiang J, Yu J, Corma A. Extra-large-pore zeolites: bridging the gap between micro and mesoporous structures. Angew Chem Int Ed. 2010;49:3120–3145.
[25] Johnson JK, Panagiotopoulos AZ, Gubbins KE. Reactive canonical Monte Carlo. A new simulation technique for reacting or associating fluids. Mol Phys. 1994;81:717–733.
[26] Mora-Fonz MJ, Catlow CRA, Lewis DW. Oligomerization and cyclization processes in the nucleation of microporous silicas. Angew Chem. 2005;117:3142–3146.
[27] Garofalini S, Martin G. Molecular simulations of the polymerization of silicic acid molecules and network formation. J Phys Chem. 1994;98:1311.
[28] Wu MG, Deem MW. Monte Carlo study of the nucleation process during zeolite synthesis. J Chem Phys. 2002;116:2125.
[29] Devreux F, Boilot JP, Chaput F, et al. Sol-gel condensation of rapidly hydrolyzed silicon alkoxides: A joint ^{29}Si NMR and small-angle X-ray scattering study. Phys Rev A. 1990 Jun;41:6901–6909. doi:10.1103/PhysRevA.41.6901.
[30] Wojdyr M. 2011 [cited July 14th 2017]. Available from: http://debyer.readthedocs.org/

Appendix 1. Adaptive RE-RxMC protocol

We have used the following protocol to perform RE-RxMC simulations of a system with N_{tet} number of tetrahedra and m replicas:

- Randomly choose between a Replica Exchange (RE) move or a Reactive Ensemble (Rx) move. The best probabilities found so far are $P_{\text{RE}} = 0.01$ and $P_{\text{Rx}} = 0.99$, although other possibilities are shown in Table A1. A large number of exchanges are important to ensure good transfer of information, but exchanging replicas too frequently may prevent each replica from equilibrating, and hence hindering the emergence of crystals.
- If a Reactive move is chosen, N_{tet} RxMC moves (displacement of all species, rotations of tetrahedra, condensations or hydrolyses) are performed according to their specific attempt probabilities given in *Section* 3.1. The Monte Carlo probabilities of each of these moves are described and derived in [9].
- On the other hand, if a Replica Exchange is selected, we perform a two-step exchange protocol as in [19]. We try to exchange half $(m/2)$ of the replicas with their adjacent replica, randomly choosing between those replicas with odd or even indices. Finally, given the number of bridging oxygens N_{BO} and the hydrolysis equilibrium constant K_{hyd} characterising each replica, we calculate the probability (see Equation A1) and accept or reject replica exchange for each pair of replicas.

Table A1. Number of replicas leading to SOD crystal framework, out of 20 independent simulations, depending on the probability of attempting a Replica Exchange. Probabilities of attempting rotations, displacements, or reactive moves were 0.9, 0.99 and 0.999 in rows 1, 2 and 3, respectively.

P_{RE}	$N_{\text{crystals}}/N_{\text{runs}}$
0.1	18/20
0.01	19/20
0.001	13/20

The probability of exchanging adjacent replicas i and j can be found in [11] and is the following:

$$P_{i,j} = \min\left\{1, \left(\frac{K_{\text{hyd},j}}{K_{\text{hyd},i}}\right)^{N_{\text{BO}i}-N_{\text{BO}j}}\right\} \tag{A1}$$

where $N_{\text{BO}i}$ and $N_{\text{BO}j}$ are the numbers of bridging oxygens, and $K_{\text{hyd},i}$ and $K_{\text{hyd},i}$ the values of the hydrolysis equilibrium constant, in replicas i and j, respectively.