PCCP

PAPER

Cite this: *Phys. Chem. Chem. Phys.*, 2014, **16**, 17845

Stochastic surface walking method for crystal structure and phase transition pathway prediction†

Cheng Shang, Xiao-Jie Zhang and Zhi-Pan Liu*

Received 5th April 2014, Accepted 11th June 2014

DOI: 10.1039/c4cp01485e

www.rsc.org/pccp

The determination of crystal structures and the solid-to-solid phase transition mechanisms are two important and related subjects in material science. Here we develop an unbiased general-purpose potential energy surface (PES) searching method, namely, SSW-crystal method, for prediction of both the crystal structure and the crystal phase transition pathway. The SSW-crystal method features with stochastic surface walking (SSW) *via* repeated small structural perturbation by taking into account the second derivative information on both the lattice and the atom degrees of freedom. The SSW-crystal method is capable of overcoming the high barrier of phase transition and identifying the desirable phase transition reaction coordinates. By applying the SSW-crystal method to a set of examples, including SiO_2 crystal up to 162 atoms per cell, Lennard-Jones model crystals up to 256 atoms, ternary $SrTiO_3$ crystal of 50 atoms and the rutile-to-anatase TiO_2 phase transition, we show that the SSW-crystal method can efficiently locate the global minimum (GM) from random initial structures without *a priori* knowledge of the system, and also allows for exhaustive sampling of the phase transition pathways, from which the lowest energy pathway can be obtained.

1. Introduction

The exploration of the PES of crystals has been a hot subject in theoretical modelling for the design of new materials.[1–12] Theoretical methods for the PES exploration of crystals can be generally divided into two categories, one targeting the structure search and the other aiming to provide kinetics information on the solid-to-solid transition pathway. The structure search algorithms are often assessed by their efficiency to identify important minima, *e.g.* GM, on PES,[13–21] which can be utilized to understand the thermodynamic properties of materials. For the kinetics involved in the phase transition, there is much less focus on theoretical modelling, not least because multiple pathways are present between the minima but also because the pathway sampling is theoretically much more challenging as compared to the minima sampling.[22,23] The identification of the correct reaction coordinate[24] among many degrees of freedom is the problem most concerned for pathway sampling but often neglected in structure search. It would be highly desirable if automated sampling on both the minima and the pathway can be achieved simultaneously in one single simulation.

The current theoretical methods for structure search generally involve aggressive structure change on either the atom or the lattice degree of freedom, which helps to bypass the transition region of PES and prevent trapping in local minima. For example, the GM searching methods starting from one single entry (one structure) on PES, such as annealing simulation,[5,7–10] basin hopping[14,15] and minimum hopping,[16,17] which were developed originally for finite systems (*e.g.* clusters) have been extended to periodic crystal systems, although their applications to material prediction are still limited. On the other hand, the methods starting from multiple entries (a group of different initial structures on PES), such as evolutionary algorithm (EA) as represented by USPEX[18,19] and the particle-swarm-optimization (PSO) algorithm[20,21] have shown great success in predicting unknown structures of materials. These methods do not yield a continuous trajectory during PES searching and thus the pathway information cannot be readily gleaned to connect the interested phases.

To provide insights into the kinetics, information on the transition region such as the saddle point on PES needs to be taken into account. One simple approach is to directly locate the transition state (TS) connecting two known crystal phases, *i.e.* the initial state (IS) and the final state (FS), as represented by some saddle point searching methods such as the ART

Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science (Ministry of Education), Department of Chemistry, Fudan University, Shanghai 200433, China. E-mail: zpliu@fudan.edu.cn

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c4cp01485e

method[25] and G-SSNEB.[26] These methods, however, do not sample the likely pathways between phases and thus may well miss the lowest energy pathway. For crystal systems, it is particularly difficult to guess the TS or FS (with the knowledge on IS) due to the flexibility of the lattice representation. The enhance sampling methods based on MD techniques,[27–31] such as metadynamics involving the constraints (bias potential)[13,32–34] and adiabatic free energy dynamics (crystal-AFED) with temperature acceleration,[35] are promising tools for sampling the PES of crystals. These methods have gentle structural perturbation and continuous trajectory as required in MD, which can be utilized to identify low energy phase transition pathways provided that the timescale of simulation is long enough to capture the transition event. For example, by imposing the collective variable on the lattice degrees of freedom, metadynamics can simulate the phase transition process between graphite and diamond. If the minima are known on PES, the local density of states and the barrier height of phase transition can also be estimated by random walking using, for example, the threshold method.[36–38] Overall, for these techniques, their efficiency to overcome the high barrier of phase transition is a general concern owing to the exponential dependence of the rate on the barrier.

In this work, we develop an unbiased general-purpose theoretical method for exploring the PES of crystals, namely the SSW-crystal method, without requiring *a priori* knowledge of the system. The method is designed on the basis of our recently-developed stochastic surface walking (SSW) method,[39,40] featuring the random direction generation, soft mode following and smooth surface walking. The SSW-crystal method can achieve the following: (i) explore the PES starting from one single entry (minimum) *via* stochastic surface walking; (ii) overcome high barriers separating the minima; and (iii) generate a continuous searching trajectory on the PES for pathway sampling. In the following sections, the methodology of the algorithm is first described and the applications of the method for structure search and pathway sampling are then detailed. A comparison with existing methods is also made, demonstrating good efficiency of the SSW-crystal method for crystal PES exploration.

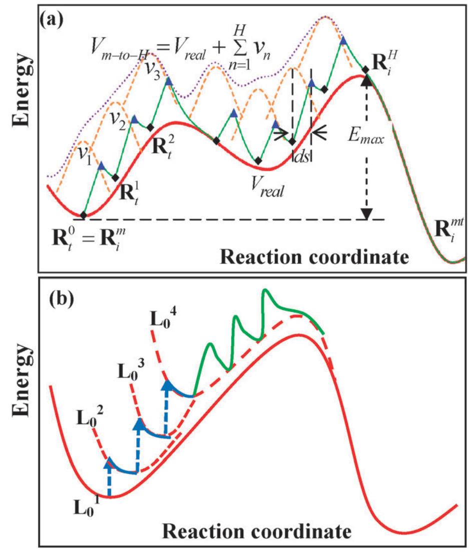


Fig. 1 Illustration of (a) the SSW method and (b) the SSW-crystal method. The real PES is described by the red curve showing separated minima. (a) One SSW step from one minimum to another, where the bias Gaussian potentials are imposed only on the atom degrees of freedom;[39,40] (b) One combined lattice and atom displacement step of the SSW-crystal method. The blue arrows and curves represent the lattice displacement and the local relaxation, respectively, as governed by the CBD-cell module, which gradually lifts up the potential energy of the crystal. The green curve represents the subsequent atom displacement (at the fixed lattice) as governed by the SSW-module, being a simplified view of (a).

2. Algorithm of SSW-crystal method

The SSW-crystal method is designed to explore the PES of periodic crystals based on the idea of the recently-developed SSW method.[39,40] The SSW method can explore molecular systems of complex PES by imposing constraints on the atom degrees of freedom *via* adding bias potentials. To extend the SSW method for crystals, one needs to couple effectively the lattice and atom movement during the crystal phase transition. In the following, we will present the algorithm of the SSW-crystal method in detail, focusing on how the new method can treat the lattice degrees of freedom together with those of the atoms.

In brief, the SSW-crystal method is constituted of three parts, an atom displacement module based on the SSW method (SSW module), a lattice displacement module based on the constrained-Broyden-dimer method for cell vectors (CBD-cell module) and a structure selection module, typically with a Metropolis Monte Carlo (MC) scheme, to judge whether a new minimum is accepted or refused. The SSW module and the CBD-cell module are responsible for the displacement of atom and lattice coordinates, respectively. By coupling these two modules, the SSW-crystal can mimic the collective movement in crystal phase transition which involves generally both the atom and the lattice degrees of freedom. Shown schematically in Fig. 1b is such a combined CBD-cell (blue curve) and SSW (green curve) displacement to manipulate one crystal structure from one minimum to another. We name such a process from one minimum to another as one SSW-crystal step. One SSW-crystal step contains typically 300–400 steps of energy and gradient calculations depending on the efficiency of local relaxation (similar to the SSW method). The SSW module and the CBD-cell module are elaborated below and the overall scheme of the SSW-crystal method is summarized at the end of the section.

2.1 SSW method

The SSW algorithm[39,40] features an automated climbing mechanism to manipulate one minimum to a high energy configuration along one random direction, which is inherited from the bias-potential driven constrained-Broyden-dimer (BP-CBD) method for TS location developed in the group.[41] The method has been successfully utilized for predicting the structure of finite clusters with complex PES, such as C_{100} fullerene, and the detail of the algorithm can be found in our previous papers.[39,40] For completeness, here we briefly outline the central idea of the

SSW method and an illustration of the SSW method is shown in Fig. 1a.

In one particular SSW step, labeled as i, a modified PES $V_{m\text{-to-}H}$, as shown in eqn (1), is utilized for moving from the current minimum, $\mathbf{R}_i^m$ to a high energy configuration $\mathbf{R}_i^H$, in which a series of bias Gaussian potential v_n (n is the index of the bias potential, n = 1,2...H) is added one by one consecutively along the direction $\mathbf{N}_i^n$.

$$\begin{aligned} V_{m\text{-to-}H} &= V_{\text{real}} + \sum_{n=1}^{H} v_n \\ &= V_{\text{real}} + \sum_{n=1}^{H} w_n \cdot \exp\left[-\left((\mathbf{R}^t - \mathbf{R}_t^{n-1}) \cdot \mathbf{N}_i^n\right)^2 \Big/ \left(2 \times \mathrm{ds}^2\right)\right] \end{aligned} \tag{1}$$

$$\begin{aligned} \mathbf{F}_{\text{tot}} &= \mathbf{F}_{\text{real}} + \sum_{n} w_n \cdot \exp\left[-\frac{\left((\mathbf{R}^t - \mathbf{R}_t^{n-1}) \cdot \mathbf{N}_i^n\right)^2}{2 \times \mathrm{ds}^2}\right] \\ &\quad \cdot \frac{(\mathbf{R}^t - \mathbf{R}_t^{n-1}) \cdot \mathbf{N}_i^n}{\mathrm{ds}^2} \cdot \mathbf{N}_i^n \end{aligned} \tag{2}$$

where $\mathbf{R}$ is the coordination vector of the structure and V_{real} represents the unmodified PES; $\mathbf{R}_t^n$ are the n^{th} local minima along the movement trajectory on the modified PES which is created after adding n Gaussian functions (see Fig. 1a). The PES at $\mathbf{R}_t^n$ is thus defined by $V_{m\text{-to-}n} = V_{\text{real}} + \sum_{k=1}^{n} v_k$. The Gaussian function is controlled by its height w and its width ds, and is always added along one particular walking direction as defined by $\mathbf{N}^n$.

$\mathbf{N}^n$ should desirably be one eigenstate of the Hessian matrix (normal mode) with low eigenvalues. Since the computation of the Hessian is expensive, SSW method optimizes $\mathbf{N}^n$ iteratively from the initial random direction $\mathbf{N}^0$ using the biased CBD rotation method,[39,41] which is utilized practically to fast locate one soft eigenvector of the Hessian matrix. It should be noted that as inherited from the CBD rotation,[42] the biased CBD rotation does not compute the whole Hessian matrix but aims to find one vector close to a soft eigenvector of the Hessian using the numerical approach. Owing to the extra bias added in the CBD rotation,[30,32] the biased CBD rotation distinguishes from other existing eigenvector updating schemes as utilized in the single-ended TS searching method like dimer,[43] CBD,[42] hybrid eigenvector-following[44] and GAD,[45,46] the rotation does not necessarily converge to the softest eigenvector or even any eigenvector of the Hessian exactly, but to a vector that mimics the initial random direction $\mathbf{N}^0$ and in the meantime is close to a soft eigenvector of the Hessian. In short, the biased CBD method is able to follow up one particular moving direction to escape the basin region, which is not always the softest mode direction (*e.g.* frustrated translational/rotational modes). More details on the random mode generation and optimization can be found in our previous works.[39,41]

Overall, the movement from $\mathbf{R}_i^m$ to $\mathbf{R}_i^H$ (see Fig. 1a) is a repeated procedure containing (i) the update of the direction $\mathbf{N}_i^n$ at $\mathbf{R}_t^{n-1}$; (ii) the adding of a new Gaussian function v_n and the displacement of $\mathbf{R}_t^{n-1}$ along the direction $\mathbf{N}_i^n$ by a magnitude of ds ($\mathbf{R}_t^{n-1}$ + $\mathbf{N}_i^n$·ds); and (iii) the local relaxation to $\mathbf{R}_t^n$ on the modified PES (the energy minimization is constrained by the added potentials). The force for the local optimization on the modified PES can be evaluated according to eqn (2). After reaching the high energy structure $\mathbf{R}_i^H$, we remove all added bias potentials and relax fully until a new minimum is found.

It might be noticed that the purpose of adding a bias potential in the SSW (also SSW-crystal) method is quite different from the MD-based metadynamics[32] method. In the SSW-crystal method, the bias potential is for driving the structure across the TS and will be removed before the full energy minimization. In metadynamics, the bias potentials are kept during the simulation to gradually fill the basins. A special reaction coordinate, *i.e.* collective variable, needs to be designed in order to fill only those related basins and drive the simulation to the target state. In contrast, the bias potential in the SSW-crystal method is imposed simply using the Cartesian coordinate, the direction of which is obtained from the biased CBD rotation. There is no need for guessing the reaction coordinate.

2.2 CBD-cell method

In the SSW method, only the atom degrees of freedom are constrained *via* adding Gaussian potentials. This is not sufficient to explore the PES of the crystal efficiently, where a significant distortion in the lattice might be essential. Following the key idea of the SSW method, we design a CBD-cell module for lattice displacement, where a few cycles are performed repeatedly to displace smoothly the lattice along one particular lattice mode direction. Each CBD-cell cycle contains three steps, namely, the soft lattice mode identification, the lattice displacement and the local relaxation of atoms in the fixed lattice. Maximally, 5 CBD-cell cycles, H_{Cell} = 5 are performed in this work. The CBD-cell cycle is schematically depicted in Fig. 1b using the blue arrows/curves.

Specifically, lattice displacement is achieved by displacing the current lattice $\mathbf{L}_0^n$ (n is the index of the current CBD-cell cycle) along a soft lattice mode direction $\mathbf{N}_{\text{cell}}$ to yield a new lattice $\mathbf{L}_0^{n+1}$, as expressed in eqn (3). The lattice $\mathbf{L}$ is a nine-dimensional (3 × 3) cell vector. After the lattice displacement, the local energy minimization (*e.g.* maximally 25 steps as utilized in this work) is then performed on the fixed lattice.

$$\mathbf{L}_0^{n+1} = \mathbf{L}_0^n + \Delta L \cdot \mathbf{N}_{\text{cell}} \tag{3}$$

It should be emphasized that a small step size, ΔL in eqn (3), is utilized for the lattice displacement, which is implemented by imposing constraints on the maximum change of the lattice length (*e.g.* a, b and c) or the maximum volume change. Here we utilize a lattice length criterion, *i.e.* $\Delta L = 0.15 \times \sqrt{\sum l_{ij}^2}, l_{ij} \in L_0^n$ (15% change for the cell vector). The gentle lattice displacement is analogous to atom displacement in the SSW method where the width of the added Gaussian (ds) is small. From eqn (3), it is obvious that the quality of $\mathbf{N}_{\text{cell}}$ determines largely how efficient the PES exploration will be.

To generate a physically meaningful $\mathbf{N}_{cell}$, we utilize the constrained-Broyden-dimer (CBD) rotation method[42] to identify one soft mode of the lattice Hessian, starting from a random vector $\mathbf{N}$ (a nine-component vector, three rotational modes are projected out in the CBD rotation). This is similar to that utilized in the SSW method for generating $\mathbf{N}_i^n$ in eqn (1). The CBD rotation algorithm for the lattice is summarized in eqn (4)–(6).

$$\mathbf{L}_1 = \mathbf{L}_0 + \mathrm{dL}\cdot\mathbf{N}_{\mathrm{cell}}^t \tag{4}$$

$$\mathbf{F}_{\mathrm{rot}}^{\mathbf{L}} = 2(\dot{\mathbf{L}}_0 - \dot{\mathbf{L}}_1) - 2[(\dot{\mathbf{L}}_0 - \dot{\mathbf{L}}_1)\cdot\mathbf{N}_{\mathrm{cell}}^t]\cdot\mathbf{N}_{\mathrm{cell}}^t \tag{5}$$

$$\dot{L}_{ij} = \partial H/\partial L_{ij} = -\Omega\sum_k(\sigma_{ik} - p_{\mathrm{ext}})\left[(\mathbf{L}^t)^{-1}\right]_{kj} \tag{6}$$

where $\dot{L}$ denotes the first derivative of the enthalpy (H) with respect to the lattice; σ is the stress tensor, Ω is the volume and p_{ext} is the external pressure (set as zero in this work); $\mathbf{L}_1$ and $\mathbf{L}_0$ are two lattice configurations, namely a dimer, separated by a fixed Euclidean distance of dL (*e.g.* 0.005 Å) along $\mathbf{N}_{\mathrm{cell}}^t$. In CBD rotation, the dimer rotates to minimize the rotation force ($\mathbf{F}_{\mathrm{rot}}^{\mathbf{L}}$)[42,47] until a preset criterion, *e.g.* 0.1 eV Å^{-1}, is reached. $\mathbf{N}_{\mathrm{cell}}$ will in principle converge to the softest eigenstate of the lattice Hessian matrix, if no extra constraints are applied (*e.g.* rotation steps). The curvature (eigenvalue) of $\mathbf{N}_{\mathrm{cell}}$ can then be derived using the numerical finite difference as:

$$C_{\mathrm{cell}} = (\dot{\mathbf{L}}_0 - \dot{\mathbf{L}}_1)\cdot\mathbf{N}_{\mathrm{cell}}^t/\mathrm{dL} \tag{7}$$

A key feature of the CBD rotation method is its ability to fast identify one soft lattice mode starting from a random vector, typically within a few steps of energy and force evaluations. Taking the α-quartz $(SiO_2)_{12}$ crystal as an example (a = 9.90, b = 9.93, c = 5.50 Å, $\alpha = \beta = 90°$, $\gamma = 120°$, the structure of which is shown in Fig. 2 inset), we have examined the efficiency of the CBD rotation to identify the lattice mode. The potential energy surface of SiO_2 is described by using the BKS potential.[48] The C_{cell} of the lattice modes of quartz are numerically calculated, being 4.98, 5.81, 8.47, 20.1, 21.7, 35.0 ($\times10^{-2}$ a.u.). The calculated softest (N_{cell}^1) and hardest (N_{cell}^6) lattice modes are as follows.

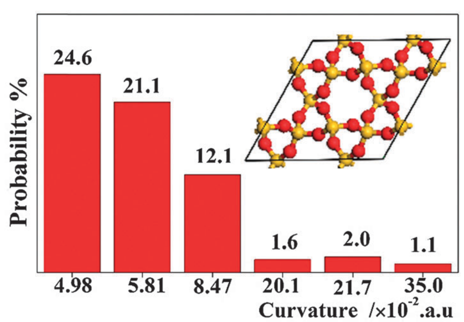


Fig. 2 Probability of identifying the lattice mode starting from random vectors using the CBD rotation method within six energy/force evaluation steps, as illustrated using α-quartz SiO_2 crystal. The structure of α-quartz SiO_2 is shown in the inset. The probability for converging to three soft lattice modes is much higher than that for converging to three hard lattice modes.

$$N_{\mathrm{cell}}^1 = \begin{pmatrix} -0.16 & 0.08 & -0.03 \\ 0.52 & -0.74 & -0.02 \\ -0.04 & -0.03 & 0.38 \end{pmatrix},$$

$$N_{\mathrm{cell}}^6 = \begin{pmatrix} 0.33 & 0.24 & 0.02 \\ 0.08 & 0.35 & 0.00 \\ 0.03 & 0.00 & 0.83 \end{pmatrix}$$

Obviously, the softest lattice mode represents the expansion along the c axis and the compression along the a and b axes, while the hardest lattice mode represents the compression along all three directions.

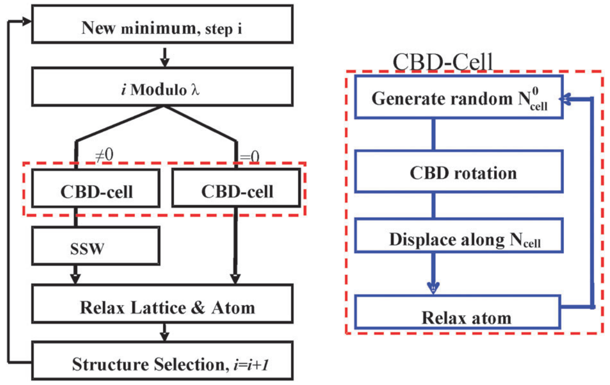
The efficiency of the CBD rotation is tested by randomly generating an initial 9-dimensional vector and then using the CBD rotation to identify the lattice mode with the maximum energy/force evaluation step being set to six. Our results are shown in Fig. 2. Averaged over 5000 independent runs, we found that the CBD rotation prefers clearly soft lattice modes: the probability for converging to three soft modes and to three hard modes are 57.8% and 4.7%, respectively. The rest 37.5% converges to the hybrid modes.

We emphasize that the displacement along the soft lattice modes, instead of a random direction, is crucial for achieving high efficiency of the crystal PES exploration. In Section 3.1 we will show that the efficiency of structure search by following soft lattice mode directions is much higher than that by following hard mode directions, as has been noticed previously that the soft modes are generally the preferred direction leading to more stable minima.[17,49,50] One major feature of the current SSW-crystal method is that both the lattice and the atom modes are explicitly calculated using the numerical biased-CBD method, which makes the SSW-crystal method transferable and robust for different PES.

Since the SSW-crystal method combines both the atom-move transformation as utilized in SSW, and the cell-move transformation as utilized in the CBD-cell module, it can be a useful tool for analyzing what type of displacement is critical for one particular phase transition. In Section 3.2 (Lennard-Jones crystal), we will show that the combination of both atom and cell movement can achieve the highest efficiency for structure search. The basic idea of coupling the atom and cell movement in the SSW-crystal method is similar to the atom and cell perturbation implemented in the USPEX[18] structure search method using the evolutional algorithm. However, in the literature, there are other methods by enforcing only the cell or the atom move for crystal structure prediction. For example, the cell-move transformation is often implemented *via* imposing external pressure[32] and the atom-move transformation can have different schemes such as those utilized in the original basin-hopping,[14] minimum-hopping[16] and PSO.[21]

2.3 Overall scheme of the SSW-crystal method

Since the SSW module and the CBD-cell module can function independently, it is flexible to combine these two modules for displacement of the crystal structure. For example, the structure



Scheme 1 Flow diagram of the SSW-crystal algorithm.

change can be driven by lattice displacement (CBD-cell only), atom displacement (SSW only), or by combining lattice and atom displacement (*i.e.* CBD-cell plus SSW, as shown in Fig. 1b). After a few preliminary trials in the example systems, we finally choose a hybrid scheme for the lattice and atom displacement in the SSW-crystal, where a parameter, λ, is utilized to partition the usage of SSW and CBD-cell modules. The flow diagram of the SSW-crystal algorithm is shown in Scheme 1 and the procedure is described below from step (1) to (8).

Step (1): Choose the structure displacement method at the current minimum $\mathbf{R}_i^m$. If the current SSW-crystal step i can be integer divided by the parameter λ (λ = 2 in this work), the combined lattice and atom displacement is performed, *i.e.* CBD-cell plus SSW. Otherwise, only the CBD-cell module is utilized for the structure displacement.

Step (2): Start the CBD-cell module by setting the current CBD-cell cycle as 1 ($n_{\text{Cell}} = 1$);

Step (3): utilize the CBD method to identify a soft mode $\mathbf{N}_{\text{cell}}$; Displace the lattice along $\mathbf{N}_{\text{cell}}$ (eqn (3)); and relax the atoms with the fixed lattice.

Step (4): Increment $n_{\text{Cell}} = n_{\text{Cell}} + 1$; if n_{Cell} reaches the maximum CBD-cell cycle (H_{Cell}), continue with step (5); else go back to step (3).

Step (5): Utilize the SSW method to displace the atom at the fixed lattice if the SSW-module is required as determined at step (1). If not, go to step (6).

Step (6): Relax all the degrees of freedom (including both atoms and lattice) to a new minimum, $\mathbf{R}_i^{mt}$ (all the added bias potentials and lattice constraints are removed, if present).

Step (7): Utilize the structure selection module to accept/refuse the new minimum. For the purpose of structure search, the Metropolis Monte Carlo scheme is utilized where the probability P of accepting the new configuration $\mathbf{R}_i^{mt}$ by making $\mathbf{R}_{i+1}^m = \mathbf{R}_i^{mt}$ is

$$P = \begin{cases} \exp\{[E(\mathbf{R}_i^m) - E(\mathbf{R}_i^{mt})]/RT\}, & \text{when } E(\mathbf{R}_i^{mt}) > E(\mathbf{R}_i^m) \\ 1, & \text{otherwise} \end{cases} \tag{8}$$

where E(X) is the energy of the structure X. For the purpose of pathway sampling, specific rules need to be designed, which will be discussed in Section 4.1 below.

Step (8): Increment $i = i + 1$ and repeat steps (1)–(8).

The SSW-crystal method for structure search is controlled by five main parameters, namely, ds: Gaussian width; H: maximum number of Gaussian potentials; ΔL: lattice displacement magnitude per CBD cycle; H_{Cell}: maximum CBD-cell cycles; and MC temperature. Because the central idea of the SSW-crystal method is to smoothly perturb the structure from one minimum to another by following one particular soft mode direction, the optimum values for the first four parameters, ds, H, ΔL and H_{Cell}, are in fact quite similar for different systems, *e.g.* ds = 0.6; H = 10; $\Delta L = 0.15 \times \sqrt{\sum l_{ij}^2}, l_{ij} \in L_0^n$; H_{Cell} = 5 utilized for most systems in this work. These values control the overall displacement distance (the Euclidean distance of the coordinates) from one minimum to another, which is in general less than 10 Å, and thus the neighboring minima in the SSW-crystal trajectory are geometrically closely connected on PES.

3. Structure search

3.1 SiO_2 crystal

We first tested the performance of the SSW-crystal method for finding the GM starting from an ordered crystal structure (less stable than GM). The SiO_2 crystal as described by the BKS potential[48] is utilized as an example where the efficiency for conversion of the α-quartz ($P3_121$, #152) to the GM, an anatase-type structure ($I4_1/amd$, #141) is examined. The BKS potential has been utilized previously for SiO_2 structure search using metadynamics and GA method, and it was noticed that the BKS potential predicts incorrectly the GM of SiO_2 at ambient pressure.[13,49] In the α-quartz form, Si is four-coordinated with the lattice constants being $a = b = 5.959$, $c = 5.500$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$, while in the anatase-type structure, Si becomes six-coordinated with the lattice constants being $a = b = 3.349$, $c = 8.510$ Å, $\alpha = \beta = \gamma = 90°$ (the coordinates of the structures are detailed in the ESI†). In order to convert the α-quartz to the GM, the algorithm needs to displace the coordinates of both the lattice and the atom.

We carried out 50 independent SSW-crystal runs for different-sized $(SiO_2)_n$ crystals with n being 12, 24, 36 or 54 (*i.e.* 162 atoms per cell by maximum), and each run was allowed for maximally 1000 SSW-crystal steps. Table 1 lists our results averaged over the 50 trajectories and Fig. 3 plots the probability of the GM being found against the total SSW-crystal steps. The key SSW-crystal parameters utilized in the simulation are as follows: ds = 0.6; H = 10; $\Delta L = 0.15 \times \sqrt{\sum l_{ij}^2}, l_{ij} \in L_0^n$; H_{Cell} = 5; MC temperature 4000 K. The probability of locating the GM is close to 100% within 200 SSW-crystal steps for n = 12 and it drops to 94 and 64% within 1000 steps for n = 36 and 54, respectively. The average SSW-crystal steps to find the GM (or the number of minima visited before the GM is reached, N_{GM}, see Table 1) is identified as 59.3, 61.2, 161.2 and 274.2 for the four systems. It is obvious that with the increase of the system size (number of atoms), the complexity of the PES increases rapidly and thus the GM identification becomes more difficult. For medium-sized crystals, *i.e.* n below 36, the SSW-crystal method exhibits high efficiency.

Table 1 GM search efficiency of the SSW-crystal method for SiO_2, LJ and $SrTiO_3$ systems averaged over 50 independent runs.[a] For SiO_2, the initial structure is an α-quartz phase. For LJ and $SrTiO_3$, the initial structure is randomly generated[b]

System	N_{atom}	N_{max}	N_{GM}	Succ. ratio (%)	Ave. force calc./step
SiO_2	36	1000	59.3	100	326
	36[b]	1000	74.74	100	416
	36[c]	1000	423.1	32	309
	72	1000	79.7	100	354
	108	1000	161.2	94	368
	162	1000	274.2	64	380
LJ	128	50	22.8	88	377
	128[d]	50	25.8	32	283
	128[e]	50	27	6	253
	256	50	36.8	34	403
	256	300	112.8	100	415
$SrTiO_3$	50	10 000	2797.9	100	496

[a] The listed data contain the number of atoms per cell (N_{atom}), the maximum SSW-crystal steps allowed for each run (N_{max}), the average number of minima visited before the GM is identified (N_{GM}), the success ratio and the average force/energy evaluation steps per SSW-crystal step. [b] Starting from random structure in a cube with periodic boundary condition, the volume of which equals to the total volume of all atoms (assuming 1.75 Å radius for atom and 0.65σ for LJ particle). [c] Only the hard lattice modes are utilized for the displacement (see eqn (3)). [d] Only the cell move (in CBD-cell module) is allowed. [e] Only the atom move (in SSW module) is allowed.

By comparing with existing methods, we found that the SSW-crystal method achieves similar or higher efficiency for this binary SiO_2 system, where the data for n = 24 are available for comparison: N_{GM} = 880 (22 generations with 40 structures of each generation) using the evolutionary metadynamics method with external pressure.[49]

It is interesting to examine the importance of the soft lattice mode as for the structure search. For this purpose, we have additionally carried out 50 independent SSW-crystal runs for n = 12 (36 atoms per cell), where the CBD rotation of the CBD-cell module is modified so that the rotation tends to converge to the hard modes (by simply reversing the sign of the rotation force in eqn (4)). The results are compared in Table 1 and also shown in Fig. 4, which demonstrate clearly that the soft mode following (red curve) outperforms the hard mode following (black curve). The efficiency of the structure search within 200 SSW-crystal steps differs by 7 times as measured by N_{GM}. Since the soft modes are generally the preferred direction leading to more stable minima,[49] the structure search on high dimensional PES can be effectively sped up by focusing on the soft modes related to only a few degrees of freedom. This also explains why the performance of the SSW-crystal method scales well (*i.e.* not exponentially) with the increase of the system size up to 162 atoms per cell (in total 489 degrees of freedom).

The SSW-crystal trajectories are also collected to see how the SSW-crystal method converts the α-quartz to anatase-type GM. Shown in Fig. 3b are the results of n = 54 containing 32 independent trajectories, where only minima that are exothermic along the trajectory are recorded for the purpose of clarity. Because the total SSW-crystal step for each trajectory varies significantly, from 35 to 968, we utilize a rescaled step

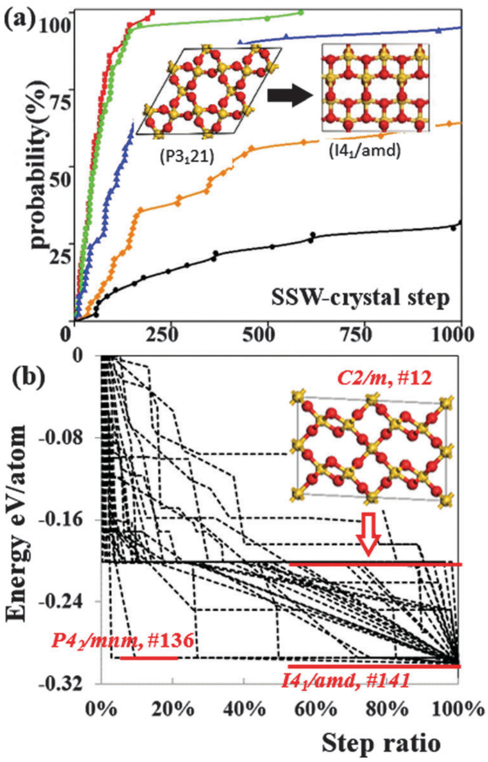


Fig. 3 (a): Probability of locating the GM of SiO_2 crystal starting from a quartz α-phase (BKS potential[48]) using the SSW-crystal method averaged over 50 independent runs. Red: 36 atoms per cell; green: 72 atoms per cell; blue: 108 atoms per cell; orange: 162 atoms per cell; black: 36 atoms per cell with the hard lattice modes utilized for displacement. (b): Trajectories of locating the GM of SiO_2 crystal starting from the α-quartz phase (162 atoms per cell). The y axis is the relative energy per atom with respect to the α-quartz phase; the x axis is the step ratio, being the step number divided by the total steps. The inset shows the structure of the major intermediate phase ($C2/m$, #12) from the searching trajectories. Red atom: oxygen; yellow atom: Si.

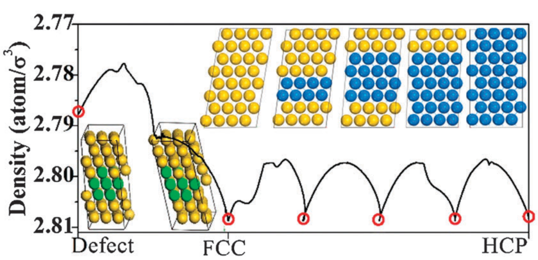


Fig. 4 Density profile in the phase transition pathway starting from a fcc crystal with a point vacancy to a perfect fcc crystal and finally to a hcp crystal of LJ_{128} crystal. The pathway is taken from the SSW-crystal searching trajectories and the TSs along the pathway are determined using the DESW method.[51] The red circles indicate the minima and the selected atoms are colored to highlight the movement (green: atoms surrounding the point defect; blue: hcp phase).

ratio, being the step number divided by the total steps, as the x axis to show the gradual energy decrease of the transformation process (see Fig. 3b). It is noticed that most of these trajectories (24 out of 32) reach a common intermediate phase ($C2/m$, #12; −0.20 eV per atom with respect to the α-quartz phase) within only 25% of the total steps and thus the computational cost is dominated by the transition from this intermediate to the GM ($I4_1/amd$, #141; −0.30 eV per atom with respect to the α-quartz phase). The other major minimum visited (9 out of 32 trajectories) is the stishovite phase ($P4_2/mnm$, #136), which is energetically very close to the GM (−0.29 eV per atom with respect to α-quartz phase).

3.2 Lennard-Jones crystal

The second example we tested is large size LJ crystals up to 256 atoms, where the interaction potential is described by $V_{\mathrm{LJ}} = 4\varepsilon \sum_{i<j} \left[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^{6} \right]$ (ε and σ are the depth and separation at the equilibrium well, respectively. We set $\varepsilon = 1$ and $\sigma = 1$). In this example, we start from the random initial structure in a cube with the periodic boundary condition (see Table 1 note). The GM of the LJ crystal is hcp packing, while the fcc packing structure is only $6.51 \times 10^{-4}\ \varepsilon$ less stable per atom as compared to the GM. Due to the large size of the crystal and the low energy separation between different phases, it is conceivable that the structural defects such as point vacancy and phase boundaries may well be present during the structure search, especially starting from random initial structures. It is thus a challenge for the algorithm to find the GM without being trapped in these local defected minima.

We carried out the SSW-crystal structure search to locate the GM of LJ crystals with 128 and 256 atoms per unit cell. The initial structure is randomly generated within a cubic unit cell with the lattice being 5.553 σ for LJ_{128} and 11.107 σ for LJ_{256} (the lattice estimated from the bond length). For both systems, 50 separated runs are performed with a maximum of 50 SSW-crystal steps. The key parameters utilized in the simulation are as follows: ds = 0.22 σ; $H = 10$; $\Delta L = 0.15 \times \sqrt{\sum l_{ij}^{2}}, l_{ij} \in L_0^n$; $H_{\mathrm{Cell}} = 5$; MC temperature 0.8. The results are summarized in Table 1.

Our results show that for LJ_{128} 44 out of 50 runs identify the GM, a success ratio of 88%, and the average N_{GM} is 22.8. For larger LJ_{256} the success ratio drops to 34% (17/50). Indeed, we found that for those trajectories that the GM is not identified, the majority of SSW-crystal runs are trapped at the stage of the fcc packing crystal and the minority stay either at the mixed fcc/hcp packing phases or at those phases with a point vacancy in the bulk. By further increasing the simulation steps, we found that simulation can eventually escape these local minima. Our test shows that as the maximum SSW-crystal steps are 300, the success ratio for LJ_{256} can finally reach 100% (also 50 independent runs).

The EA method using USPEX has been utilized for LJ crystals starting from a random structure.[19] The N_{GM} is reported to be 30 for LJ_{128} (3 generation and 10 structures per generation), which is comparable to our results, 22.8 (success ratio 88%); For larger systems, *e.g.* 256 atom, USPEX reaches the less favorable fcc structure, indicating the EA is also sensitive to the system size.[19] From our results, the efficiency of the SSW-crystal method scales reasonably with the increase of the system size and can identify the GM with a medium success ratio (34% within 50 steps and 100% within 300 steps).

It is of significance to ask why the SSW-crystal method can identify the GM and how it leaves the local minima for these large LJ crystals. To answer this, an additional SSW-crystal search for LJ_{128} was conducted. First, we tested the efficiency of the SSW-crystal method by using only the SSW module or the CBD-cell module starting from random structures, *i.e.* allowing only an atom or cell move in the structure displacement. In both cases the efficiency of identifying the GM drops dramatically, as shown in Table 1. From these results, it is obvious that the combination of both the atom and the cell move is more effective for crystal structure exploration.

Second, we performed SSW-crystal runs by deliberately setting the initial structure at a local minimum, the fcc-packing crystal with one point vacancy (fcc-with-vacancy). In total, 250 SSW-crystal runs were carried out with the maximum step being limited to 250. We found that 28 of them successfully identify the hcp GM, while most trajectories (214 out of 250) end with the fcc-packing structure. These trajectories have been compared carefully to retrieve the common features of these structure searches. We found that, interestingly, the SSW-crystal trajectories that finally visit the hcp GM generally follow a conversion sequence by transforming first the fcc-with-vacancy structure to the perfect fcc-packing structure and finally to the hcp-packing GM. The fcc-packing structure is therefore an important intermediate that locates most likely at the lowest energy conversion channel. In order to confirm this picture, we further utilize the TS searching method to identify the TS between the connecting phases, which are obtained from the SSW-crystal trajectories, and determine the pathway between them (note that no extensive pathway sampling is performed here since we are focusing on the common features of the SSW-crystal trajectories; and more details on the rigorous way of pathway sampling will be discussed in Section 4 below). The TS is located using our recently developed double-ended surface walking (DESW) method[51] and the pathways connecting two minima are then generated by steepest–descent optimization starting from located TSs.

Fig. 4 shows the lowest energy phase transformation pathway that we determined from these 250 trajectories, where the density of the crystal during phase transformation is plotted against the reaction coordinate. It is interesting to notice that the one-vacancy crystal needs to experience both expansion and densification in order to arrive at the perfect fcc crystal (triclinic lattice, $a = 2.89$, $b = 7.47$, $c = 5.56\ \sigma$, $\alpha = 94.9°$, $\beta = 85.8°$, $\gamma = 80.5°$). This fcc lattice will then transform to another triclinic lattice ($a = 7.32$, $b = 4.36$, $c = 3.93\ \sigma$, $\alpha = 106.1°$, $\beta = 78.1°$, $\gamma = 98.6°$) before it starts the transformation to the hcp phase. The fcc-to-hcp transformation involves a sequential expansion–compression process, which exhibits a layer-by-layer atomic gliding. From the pathway, we can conclude that the crystal phase transition occurs *via* collective movement of both the atom and lattice coordinates, as shown clearly in the layer gliding. The hybrid structure displacement (CBD-cell plus SSW) utilized in the SSW-crystal method is the key to achieve high efficiency of the crystal PES exploration.

The landscape of a two-dimensional LJ crystal (16 atoms per cell) has been analyzed by Schön and Jansen.[36,37,52] Interestingly, it was shown that the one-vacancy minimum is separated from the perfect crystal by a two-vacancy minimum. The crystal expansion followed by densification is required to remove the vacancy, demonstrating the importance of the hybrid way of structure displacement (move-class in their work). This is similar to what is observed here for the 3-D large system.

3.3 $SrTiO_3$ crystal

We then investigated a large ternary crystal, $SrTiO_3$, by using the SSW-crystal method for finding the GM starting from a random initial structure. PES exploration becomes much more challenging as compared to the single or binary crystals, in particular starting from random structures. The structure search of this system has been studied previous by EA methods as implemented in XtalOpt[53] and USPEX.[54] The potential proposed by Benedek *et al.*[55] is utilized to describe the interaction of crystals.

We have carried out 50 independent SSW-crystals with the maximum steps being 10 000, each starting from randomly placed atoms (50 atoms per cell) in a cubic cell with each lattice axis being 20 Å. These initial structures are highly disordered, *e.g.* with close Ti–Ti and O–O bonds, which are not present in the $SrTiO_3$ perovskite, the GM, which is a bcc packing unit cell with Sr occupying the center of the cube surrounded by Ti–O framework. The key parameters utilized in simulation are as follows: ds = 0.6; $H = 15$; $\Delta L = 0.15 \times \sqrt{\sum l_{ij}^2}, l_{ij} \in L_0^n$; $H_{\rm Cell} = 5$; MC temperature 1200 K. We found that the success ratio of the 50 SSW-crystal runs is 22% within 1000 SSW-crystal steps and reaches 100% within 10 000 SSW-crystal steps. The shortest trajectory takes 194 SSW-crystal steps to locate the GM. The results are also summarized in Table 1.

We also collected the SSW-crystal trajectories to examine how the SSW-crystal method converts the random structure to a perovskite phase. The results with 50 independent trajectories are shown in Fig. 5, where the x axis is the step ratio (the same as defined in Fig. 3b) and the y axis is the relative energy per atom with respect to the GM. Only the minima that are exothermic along the trajectory are recorded for the purpose of clarity. Interestingly, we found that unlike in SiO_2 (Fig. 3b), the 50 trajectories of $SrTiO_3$ are quite scattered and show few common features. Basically, each trajectory reaches the GM *via* its own way. This indicates that this ternary system has a highly complex PES and 50 independent runs are statistically not enough to capture the common feature in the phase conversion. Nevertheless, we identify two ordered intermediate phases which are common in several trajectories, as highlighted in Fig. 5 insets, with the energy being 4–8 eV above the GM. They belong to the defected crystal with wrongly-placed Ti/Sr atoms.

By comparing to previous results, we notice that the unbiased SSW-crystal method can achieve a reasonable efficiency (22% within 1000 minima visited, also see Table 1) for this complex system, higher than the EA method as implemented in XtalOpt code, *i.e.* 7–12% identifying the GM within 1000 minima visited. With the symmetry constraint and soft mutation techniques, the EA method implemented in USPEX code can achieve a higher success ratio for this system, namely 94% identifying the GM within 1000 minima visited.[53]

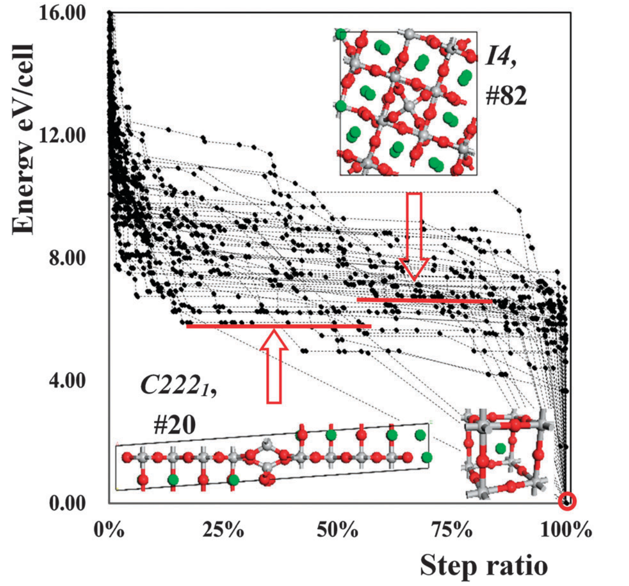


Fig. 5 SSW-crystal trajectories (50 runs) for finding the GM of $SrTiO_3$ crystal starting from a random initial structure (50 atoms per cell). The x axis is the step ratio and the y axis is the relative energy per atom with respect to the GM, namely perovskite structure. The insets show two intermediate structures from the trajectories, being defected crystal with wrongly-placed Ti/Sr atoms. The bcc packing unit cell of GM is shown at the bottom-right corner (Sr atom at the cubic center being surrounded by Ti–O framework).

Because the SSW-crystal method is a single-entry approach starting from a single initial structure (seed), different from the multiple-entry approaches such as EA and PSO methods, we expect that the efficiency of the current SSW-crystal method for structure search can be enhanced greatly when the massive parallel computation is utilized where the replica exchange between different seeds can be implemented. One possible strategy is to allow for multiple SSW-crystal runs at the current most stable replica to increase the probability of identifying better minima nearby, and simultaneously introduce other structure seeds (increase structure variety) to prevent the too long trapping at the local minimum. Our previous work has shown that for complex PES such as C_{100} cluster, the parallel replica exchange is essential for finding the fullerene GM among numerous energy nearly degenerate IPR isomers.[40] The parallel exchange and symmetry constraints could be a future direction to improve the efficiency of the SSW-crystal method.

4. Pathway sampling

The above results show that the SSW-crystal can be utilized for the crystal structure search and the SSW-crystal trajectory contains useful information on the pathway linking different minima. The overall structure change between two connecting minima obtained from SSW-crystal trajectory is generally not dramatic, *e.g.* within 10 Å (controlled by ds, H, ΔL and $H_{\rm Cell}$). Therefore, it is possible to match these two structures to identify the pathway between them by using the standard structure match algorithm (*e.g. via* minimizing the distance between two minima with the periodic boundary condition). The only problem is that such a connection may still not be the lowest energy pathway between the two connecting crystal phases. Nevertheless, it is likely to exploit the feature of the SSW-crystal method to sample exhaustively the pathways linking different crystal phases. Compared to the traditional approach for pathway sampling based on the MD, the trapping at deep local minima could be much less severe in the SSW-crystal method, as demonstrated above in the GM search of various systems, which increases the overall efficiency.

4.1 Method for pathway sampling and lowest energy pathway determination

In order to sample exhaustively all the likely pathways between two phases, we need to utilize a different structure selection module in the SSW-crystal, in which the new identified structure is accepted/refused based on the criterion whether the new phase is identical to the previous phase, instead of the Metropolis Monte-Carlo criterion as utilized in the structure search. The procedure for the pathway sampling and the lowest energy pathway determination is described as follows.

(i) We start from one single phase and utilize the SSW-crystal method to explore all the likely phases nearby this phase. If a new phase different from the starting phase is visited, we record the initial state (IS) and the final state (FS) of the current SSW-crystal step. Then, the SSW-crystal simulation will return to the IS (by rejecting the new phase) to continue the sampling. On the other hand, if the new minimum is the same as the starting phase (*e.g.* a lattice-permutation isomer), the SSW-crystal will accept the new isomeric phase and continue the sampling. The energy and crystal symmetry[56,57] are utilized as the criterion to distinguish the phases.

(ii) The pathway sampling stops until the accumulated IS–FS pairs reach a certain maximum number.

(iii) Based on the IS–FS pair, the DESW method[51] is utilized to identify the TSs explicitly.[41,42] The pathway connecting the IS and the FS is then determined by extrapolating the TS towards the IS or FS using the steepest-descendent method. The barrier is the energy difference between IS and TS.

(iv) The lowest energy pathway is identified by comparing the barrier of the pathways.

It should be mentioned that the above procedure for pathway sampling is only a one-end approach, namely, to identify all the pathways initiated from the starting phase. It is common, however, that the problem in question is associated with the lowest energy pathway linking two known phases (two-end) and the two phases may be well separated on PES without a short connection. For solving such problems, one needs to complete the whole connecting network by performing repeatedly the one-end pathway sampling in a recursive manner, which will be illustrated in the following example. The path optimization methods can also be used to analyze the database after the network is built.[22,58,59]

4.2 Phase transition pathway between TiO_2 rutile and anatase

The example we chose for the crystal phase transition pathway sampling is TiO_2, which has two common phases at ambient conditions, rutile and anatase. The conversion of the two phases has been widely exploited in experiments for synthesizing different morphologies of TiO_2[60] and is of great importance in applications, for example, of photocatalytic water splitting.[61–65] Despite extensive studies on these two phases of TiO_2,[66–69] the phase transition mechanism between rutile and anatase remains not clear at the atomic level.

In this work, the rutile-to-anatase pathway sampling is first performed to generate a database of possible pathways linking two phases. To speed up the process, we utilize the MA potential to describe a TiO_2 crystal, which has been utilized previously for investigating the phase diagram of TiO_2 at elevated pressures.[70,71] The TSs of all the pathways are determined explicitly and the low energy pathways are collected. The density functional theory (DFT) calculations are then utilized to verify the low energy pathways, typically the five lowest ones obtained from the MA potential search, where the TS is re-located in the first principles framework. By comparing the barrier of all these pathways, the lowest energy pathway can be finally obtained.

All DFT calculations are performed using the SIESTA package[72] with an optimized numerical double-ζ polarization basis set[73,74] at the GGA-PBE exchange–correlation functional level. The energy cutoff for the real space grid used to represent the density was set to 150 Ry. An energy shift of 0.01 eV was used to determine the orbital-confining cutoff radii. For all the lattices studied, the k-point mesh utilized has been Monkhorst–Pack ($4 \times 4 \times 4$) set, which is verified to be sufficient for obtaining good energetics.

We have carried out the SSW-crystal pathway sampling starting from rutile (12 atom per cell) with the maximum SSW-crystal step being set to 10 000. We identified 8413 minima that lie close to the rutile phase, among them 819 belonging to the TiO_2II phase, 797 being $Pn2_1m$ (#31), 430 being anatase, and so on. The results are summarized in Fig. 6 by plotting the averaged O–Ti bond length of the crystal against the sum of the lattice length. The structure coordinates are detailed in the ESI.† It shows that rutile, anatase and TiO_2II have very close O–Ti bond lengths; and compared to rutile, anatase and TiO_2II are more flexible in the lattice, showing a wide distribution of the lattice length.

Since the lowest energy pathway for rutile-to-anatase may not be a one-step reaction, we continue to conduct the one-end pathway sampling starting from TiO_2II and anatase. For each linkage (IS–FS pair) identified from the SSW-crystal trajectory, say, *e.g.* rutile-to-anatase, rutile-to-TiO_2II and TiO_2II-to-anatase, we have located the TS of the pathway and the low energy pathways are then verified using DFT.

Our DFT calculation results for the lowest energy pathway determined are highlighted in Fig. 7. The pathway is *via* two well-known high pressure phases, MI (Baddeleyite, $P2_1/c$, #14)

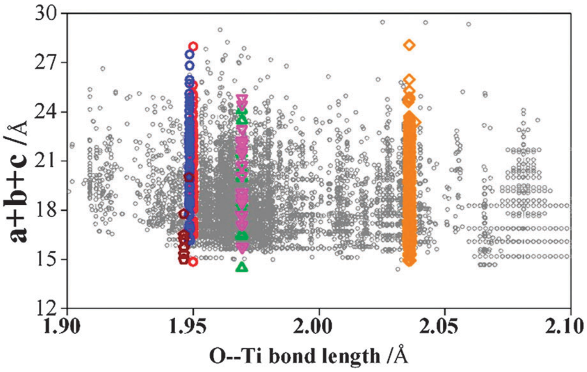


Fig. 6 The 8413 minima nearby the TiO_2 rutile phase (MA potential) as found with SSW-crystal pathway sampling. The x axis is the averaged O–Ti bond length of the crystal, which represents the strength of local bonding (the O–Ti bond is defined to be below 2.5 Å); the y axis is the sum of the lattice length, which is a measure of the lattice flexibility of the phase. Brown: rutile; red: TiO_2II; blue: anatase; pink: 88#; green: pyrite; orange: 31#.

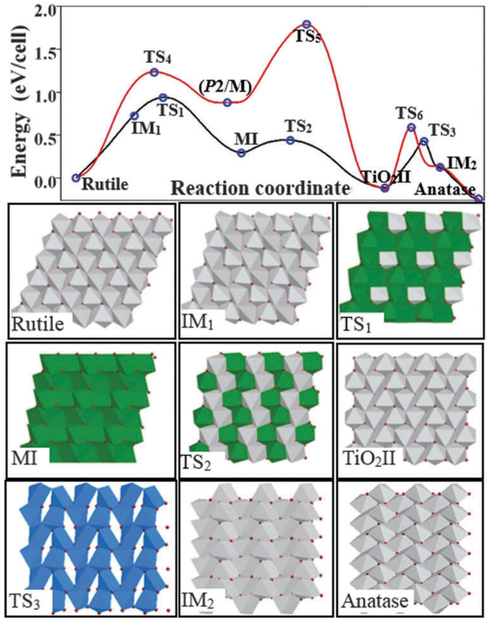


Fig. 7 (top) DFT energy profile for the lowest energy pathway from rutile to anatase. Black: the lowest energy pathway; red: the higher energy pathways connecting rutile to anatase. (bottom) Polyhedron structure snapshots from rutile to anatase in the lowest energy transition pathway (all views are taken along the c axis of each lattice, which is [100] direction in rutile and TiO_2II, and [110] direction in anatase). Polyhedrons are colored according to the coordination number of Ti ions. Blue: 5-coordinated; grey: 6-coordinated; green: 7-coordinated.

and TiO_2II (columbite, α-PdO_2 type, *Pbcn*, #60), ending with the anatase phase. The highest barrier is 0.84 eV per cell from DFT, related to the TS_1 in between rutile and TiO_2II. Since MI is significantly unstable at ambient pressure, TiO_2II is the key intermediate between anatase and rutile. Specifically, the rutile structure to initiate phase transition is in a monoclinic lattice ($a = b = 5.51$, $c = 4.65$ Å, $\alpha = \beta = 90°$, $\gamma = 65.61°$). This lattice is clearly different from that based on the Bravais conventional lattice of rutile (tetragonal, $a = b = 4.64$, $c = 5.95$ Å, $\alpha = \beta = \gamma = 90°$, a double cell along the c axis). The intermediate MI phase is in a monoclinic lattice ($a = 5.04$, $b = 4.84$, $c = 4.92$ Å, $\alpha = \beta = 90°$, $\gamma = 80.13°$) and the intermediate TiO_2II is in an orthorhombic lattice ($a = 5.58$, $b = 4.92$, $c = 4.57$ Å, $\alpha = \beta = \gamma = 90°$). The ending anatase phase is in a monoclinic lattice ($a = 5.55$, $b = 5.38$, $c = 5.37$ Å, $\alpha = \gamma = 90°$, $\beta = 61.46°$), again not the Bravais lattice of anatase (tetragonal, $a = b = 3.80$, $c = 9.76$ Å, $\alpha = \beta = \gamma = 90°$).

It should be emphasized that the Bravais lattice represents the smallest unit cell that contains the full symmetry of the crystal and is not the only possible lattice for a crystal. Because of the periodic constraint in simulation from one phase to another, the lowest energy phase transition pathway must require a good match for both the lattice and the atom coordinates between two connecting crystal minima. Consequently, the pathway does not generally initiate from or end with the Bravais cell. Our results show that the SSW-crystal method through the random pathway sampling can reveal desirable lattice/atom coordinates with a good structure match.

The snapshot of the transition pathway is shown in Fig. 7, highlighting changes in the coordination number of Ti in the process, namely, going from 6 of rutile to 7 of MI, back to 6 of TiO_2II, and having the 5 coordination transit from TiO_2II to anatase. In the transition, the density of the crystal first increases by 8.8% to reach the maximum at MI (a high pressure phase) and is then reduced by 6.1% to TiO_2II. Unlike that in rutile-to-II transformation, the lower coordinated O_{2c} and Ti_{5c} emerge in the II-to-anatase, as represented by the TS_3 in the pathway. The density at TS3 is 5.7% larger as compared to anatase but 5.7% lower as compared to TiO_2II. The pathway from anatase to II phase transition has a comparable but lower barrier (0.70 eV) than that of the pathway from rutile to II (0.84 eV). This indicates that TiO_2II is an important intermediate phase in between anatase–rutile and may be observed under experimental conditions. Indeed, TiO_2II has been found to grow at the grain boundaries of rutile ores in nature.[75]

It might be mentioned that there are many possible pathways for the rutile-to-anatase phase transition but with the higher barriers because there are many possible lattices for rutile and anatase (as shown in Fig. 6). For example, we also identify a two-step transition path (red curve shown in Fig. 7), initiating from the doubled Bravais lattice of rutile (tetragonal lattice, $a = c = 4.64$, $b = 5.95$ Å, $\alpha = \beta = \gamma = 90°$) to TiO_2II (orthorhombic lattice, $a = 4.56$, $b = 5.57$, $c = 4.94$ Å, $\alpha = \beta = \gamma = 90°$), which has a barrier of 1.82 eV per cell and passes a highly unstable phase, *P*2/*M*, #10. For TiO_2II to anatase, the second lowest energy transition path has a barrier of 0.1 eV (TS_6) higher than the lowest one (TS_3).

We also compared our rutile-to-anatase pathway with the one reported previously using the GSS-NEB method.[69] We noticed that the GSS-NEB method identifies a one-step pathway with an energy barrier of 9.74 eV (12 atoms per cell), where the IS and the FS utilized are based on the conventional Bravais cell of rutile and anatase. No pathway sampling was performed. Obviously, the rutile-to-anatase pathway from the SSW-crystal method, a three-step pathway, has a much lower barrier, 0.84 eV (12 atom per cell), which should therefore be kinetically more relevant for the phase transition. Our results, although performed here in a small unit cell of bulk, do suggest that the pathway sampling can be theoretically done and is essential to understanding solid-to-solid phase transition. More detailed analyses on the rutile-to-anatase phase transition pathway in large bulk cells and at the surfaces will be addressed in our future publications.

Finally, it is worthwhile to outline the main features of the SSW-crystal method and highlight its possible applications. First, similar to the other non-MD based methods for structure search, the current SSW-crystal explores the potential energy surface, not the free energy surface. This is largely because the energy minimization is always performed to locate the minima and the vibration contribution to entropy is removed.[76] This speeds up effectively PES searching without spending too much effort in visiting high density liquid-like phases. Therefore, the SSW-crystal method could be utilized as a first tool for exploring PES to sample important minima with large configuration entropy. Second, the SSW-crystal method involves only smooth structural displacement that is similar to MD-based techniques and thus the pathway information is retained during PES searching. The method can naturally identify the neighboring minima with close geometry. This leads to two interesting applications, *i.e.* pathway sampling, as demonstrated above, and investigation of crystals with rigid bonds, *e.g.* molecular crystals[13,32,33] where a large number of chemical bonds need to be conserved during PES exploration. Our ongoing work with an urea crystal as an example shows that the SSW-crystal method with reduced ds and ΔL parameters can be applied to study the phase transition of molecular crystals without destroying the chemical bond of the molecule.

5. Conclusion

The SSW-crystal method developed in this work represents a new theoretical method for exploring the PES of periodic crystals, which is a natural extension of the SSW method developed recently for finite systems. The SSW-crystal method involves two independent modules to treat the lattice degrees of freedom, the CBD-cell module, and the atom degrees of freedom, the SSW module. The effective coupling of these two modules helps the algorithm to mimic the collective lattice and atom movement during crystal phase transition. The new method can thus be used for both structure search and pathway sampling of a crystal system without *a priori* knowledge of the system (*e.g.* symmetry, bonding motif). Although the SSW-crystal method requires only the energy and the first derivative (stress and force) as input, it utilizes the numerical approach based on the biased CBD method to identify one soft lattice or an atom mode, starting from a random vector. By exploiting the second derivative information of PES and imposing repeatedly small structure perturbation along the soft mode, the SSW-crystal method can explore crystal PES efficiently, and simultaneously retain key information on the phase transition pathway.

We apply the new method for finding the GM of a SiO_2 crystal, Lennard-Jones crystal, $SrTiO_3$ crystal, and also for mapping out the phase transition pathway between TiO_2 rutile and anatase phases. We show that the SSW-crystal method is highly efficient for structure search, even starting from the random initial structure, which is comparable to that of existing algorithms. In addition, the SSW-crystal method in combination with the DESW method can sample exhaustively the pathways connecting different phases, and thus provide a unique opportunity to identify the lowest energy phase transition pathway and deduce the connection network between different crystal phases. Since the method is an unbiased general-purpose tool for material structure and pathway prediction, we expect that it could have wide application in the design of new materials in the future.

Acknowledgements

This work is financially supported by the National Science Foundation of China (21173051, 21361130019), the 973 program (2011CB808500, 2013CB834603), the Science and Technology Commission of Shanghai Municipality (08DZ2270500), and the Program for Professors of Special Appointment (Eastern Scholars) at the Shanghai Institute of Higher Learning.

References

1 J. Maddox, *Nature*, 1988, **335**, 201.
2 A. R. Oganov, *Modern Methods of Crystal Structure Prediction*, Wiley-VCH Verlag GmbH & Co. KGaA, 2010.
3 A. Gavezzotti, *Acc. Chem. Res.*, 1994, **27**, 309–314.
4 P. J. Spencer, *CLAPHAD*, 2008, **32**, 1–8.
5 J. Pannetier, J. Bassas-Alsina, J. Rodriguez-Carvajal and V. Caignaert, *Nature*, 1990, **346**, 343–345.
6 C. M. Freeman, J. M. Newsam, S. M. Levine and C. R. A. Catlow, *J. Mater. Chem.*, 1993, **3**, 531–535.
7 S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, *Science*, 1983, **220**, 671–680.
8 J. C. Schön and M. Jansen, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 1286–1304.
9 J. C. Schon and M. Jansen, *Z. Kristallogr.*, 2001, **216**, 307–325.
10 J. C. Schon and M. Jansen, *Z. Kristallogr.*, 2001, **216**, 361–383.
11 S. M. Woodley, P. D. Battle, J. D. Gale and C. R. A. Catlow, *Phys. Chem. Chem. Phys.*, 1999, **1**, 2535–2542.
12 S. M. Woodley and R. Catlow, *Nat. Mater.*, 2008, **7**, 937–946.
13 R. Martonak, D. Donadio, A. R. Oganov and M. Parrinello, *Nat. Mater.*, 2006, **5**, 623–626.
14 D. J. Wales and J. P. K. Doye, *J. Phys. Chem. A*, 1997, **101**, 5111–5116.
15 T. F. Middleton, J. Hernández-Rojas, P. N. Mortenson and D. J. Wales, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2001, **64**, 184201.
16 S. Goedecker, *J. Chem. Phys.*, 2004, **120**, 9911–9917.
17 M. Amsler and S. Goedecker, *J. Chem. Phys.*, 2010, **133**, 224104.

18 A. R. Oganov and C. W. Glass, *J. Chem. Phys.*, 2006, **124**, 244704.
19 A. R. Oganov and C. W. Glass, *J. Phys.: Condens. Matter*, 2008, **20**, 064210.
20 Y. C. Wang, J. A. Lv, L. Zhu and Y. M. Ma, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2010, **82**, 094116.
21 Y. Wang, J. Lv, L. Zhu and Y. Ma, *Comput. Phys. Commun.*, 2012, **183**, 2063–2070.
22 D. J. Wales, *Mol. Phys.*, 2002, **100**, 3285–3305.
23 J. M. Carr, S. A. Trygubenko and D. J. Wales, *J. Chem. Phys.*, 2005, **122**, 234903.
24 A. R. Oganov, A. O. Lyakhov and M. Valle, *Acc. Chem. Res.*, 2011, **44**, 227–237.
25 E. Machado-Charry, L. K. Béland, D. Caliste, L. Genovese, T. Deutsch, N. Mousseau and P. Pochet, *J. Chem. Phys.*, 2011, **135**, 034102.
26 D. Sheppard, P. Xiao, W. Chemelewski, D. D. Johnson and G. Henkelman, *J. Chem. Phys.*, 2012, **136**, 074103.
27 L. Rosso, P. Minary, Z. W. Zhu and M. E. Tuckerman, *J. Chem. Phys.*, 2002, **116**, 4389–4402.
28 L. Maragliano and E. Vanden-Eijnden, *Chem. Phys. Lett.*, 2006, **426**, 168–175.
29 J. I. Siepmann and D. Frenkel, *Mol. Phys.*, 1992, **75**, 59–70.
30 A. F. Voter, *Phys. Rev. Lett.*, 1997, **78**, 3908–3911.
31 C. J. Woods, J. W. Essex and M. A. King, *J. Phys. Chem. B*, 2003, **107**, 13703–13710.
32 A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12562–12566.
33 P. Raiteri, R. Martoňák and M. Parrinello, *Angew. Chem., Int. Ed.*, 2005, **44**, 3769–3773.
34 T. Huber, A. Torda and W. Gunsteren, *J. Comput.-Aided Mol. Des.*, 1994, **8**, 695–708.
35 T.-Q. Yu and M. E. Tuckerman, *Phys. Rev. Lett.*, 2011, **107**, 015701.
36 J. C. Schön and M. Jansen, *Ber. Bunsen-Ges.*, 1994, **98**, 1541–1544.
37 J. C. Schön, H. Putz and M. Jansen, *J. Phys.: Condens. Matter*, 1996, **8**, 143–156.
38 P. Sibani, J. C. Schön, P. Salamon and J. O. Andersson, *Europhys. Lett.*, 1993, **22**, 479.
39 C. Shang and Z. P. Liu, *J. Chem. Theory Comput.*, 2013, **9**, 1838–1845.
40 X.-J. Zhang, C. Shang and Z.-P. Liu, *J. Chem. Theory Comput.*, 2013, **9**, 3252–3260.
41 C. Shang and Z. P. Liu, *J. Chem. Theory Comput.*, 2012, **8**, 2215–2222.
42 C. Shang and Z.-P. Liu, *J. Chem. Theory Comput.*, 2010, **6**, 1136–1144.
43 G. Henkelman and H. Jonsson, *J. Chem. Phys.*, 1999, **111**, 7010–7022.
44 L. J. Munro and D. J. Wales, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **59**, 3969–3980.
45 W. N. E and X. Zhou, *Nonlinearity*, 2011, **24**, 1831–1842.
46 A. Samanta and W. N. E, *J. Chem. Phys.*, 2012, **136**, 124104.
47 H.-F. Wang and Z.-P. Liu, *J. Am. Chem. Soc.*, 2008, **130**, 10996–11004.
48 B. W. H. van Beest, G. J. Kramer and R. A. van Santen, *Phys. Rev. Lett.*, 1990, **64**, 1955–1958.
49 Q. Zhu, A. R. Oganov and A. O. Lyakhov, *CrystEngComm*, 2012, **14**, 3596–3601.
50 D. Wales, *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*, Cambridge University Press, 2003.
51 X.-J. Zhang, C. Shang and Z.-P. Liu, *J. Chem. Theory Comput.*, 2013, **9**, 5745–5753.
52 H. Putz, J. C. Schön and M. Jansen, *Ber. Bunsen-Ges.*, 1995, **99**, 1148–1153.
53 D. C. Lonie and E. Zurek, *Comput. Phys. Commun.*, 2011, **182**, 372–387.
54 A. O. Lyakhov, A. R. Oganov, H. T. Stokes and Q. Zhu, *Comput. Phys. Commun.*, 2013, **184**, 1172–1182.
55 N. A. Benedek, A. L. S. Chua, C. Elsässer, A. P. Sutton and M. W. Finnis, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2008, **78**, 064110.
56 R. Grosse-Kunstleve, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 1999, **55**, 383–395.
57 R. W. Grosse-Kunstleve and P. D. Adams, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 2002, **58**, 60–65.
58 D. Zagorac, J. C. Schoen and M. Jansen, *J. Phys. Chem. C*, 2012, **116**, 16726–16739.
59 C. Grebner, L. P. Pason and B. Engels, *J. Comput. Chem.*, 2013, **34**, 1810–1818.
60 A. L. Linsebigler, G. Lu and J. T. Yates, *Chem. Rev.*, 1995, **95**, 735–758.
61 R. I. Bickley, T. Gonzalez-Carreno, J. S. Lees, L. Palmisano and R. J. D. Tilley, *J. Solid State Chem.*, 1991, **92**, 178–190.
62 T. Ohno, K. Sarukawa, K. Tokieda and M. Matsumura, *J. Catal.*, 2001, **203**, 82–86.
63 J. Zhang, Q. Xu, Z. Feng, M. Li and C. Li, *Angew. Chem., Int. Ed.*, 2008, **47**, 1766–1769.
64 Y.-F. Li, Z.-P. Liu, L. Liu and W. Gao, *J. Am. Chem. Soc.*, 2010, **132**, 13008–13015.
65 Y.-F. Li and Z.-P. Liu, *J. Am. Chem. Soc.*, 2011, **133**, 15743–15752.
66 R. L. Penn and J. F. Banfield, *Am. Mineral.*, 1998, **83**, 1077–1082.
67 R. L. Penn and J. F. Banfield, *Science*, 1998, **281**, 969–971.
68 P. C. Ricci, C. M. Carbonaro, L. Stagi, M. Salis, A. Casu, S. Enzo and F. Delogu, *J. Phys. Chem. C*, 2013, **117**, 7850–7857.
69 N. H. Vu, H. V. Le, T. M. Cao, V. V. Pham, H. M. Le and D. Nguyen-Manh, *J. Phys.: Condens. Matter*, 2012, **24**, 405501.
70 M. Matsui and M. Akaogi, *Mol. Simul.*, 1991, **6**, 239–244.
71 N. A. Dubrovinskaia, L. S. Dubrovinsky, R. Ahuja, V. B. Prokopenko, V. Dmitriev, H. P. Weber, J. M. Osorio-Guillen and B. Johansson, *Phys. Rev. Lett.*, 2001, **87**, 275501.
72 J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejon and D. Sanchez-Portal, *J. Phys.: Condens. Matter*, 2002, **14**, 2745.
73 J. Junquera, O. Paz, D. Sanchez-Portal and E. Artacho, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2001, **64**, 235111.
74 E. Anglada, J. M. Soler, J. Junquera and E. Artacho, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2002, **66**, 205101.
75 A. El Goresy, M. Chen, P. Gillet, L. Dubrovinsky, G. Graup and R. Ahuja, *Earth Planet. Sci. Lett.*, 2001, **192**, 485–495.
76 J. P. K. Doye and D. J. Wales, *Phys. Rev. Lett.*, 1998, **80**, 1357–1360.

