ROYAL SOCIETY
OF CHEMISTRY

PAPER

Metashooting: a novel tool for free energy reconstruction from polymorphic phase transition mechanisms

Samuel Alexander Jobbins,[a] Salah Eddine Boulfelfel [iD] [b]
and Stefano Leoni [iD] *[a]

We introduce a novel scheme for the mechanistic investigation of solid–solid phase transitions, which we dub 'metashooting'. Combining transition path sampling, molecular dynamics and metadynamics, this scheme allows for both a complete mechanistic analysis and detailed mapping of the free energy surface. This is illustrated by performing metashooting calculations on the pressure-induced B4/B3 → B1 phase transition in ZnO. The resulting free energy map helps to clarify the role of intermediate configurations along this activated process and the competition between different mechanistic regimes with superior accuracy. We argue that metashooting can be efficiently applied to a broader class of activated processes.

Introduction

Molecular dynamics (MD) has become the method of choice to investigate complex chemical phenomena at the atomic level. Thanks to both efficiently written algorithms and high-performance computational infrastructures, the scope of numerical simulations has markedly widened. Nonetheless, the presence of inherent kinetic bottlenecks means that entire regions of configuration space may remain unexplored. If such states are separated by large energy barriers, even the most efficient MD will end up spending considerable time within long-lived, metastable states. Furthermore, depending on the height of those barriers, the event of interest may fully elude observation over the time scale of the simulation.[1–3]

Protein folding, nucleation, crystallisation, phase transitions and chemical reactions are all important phenomena which are characterised by activated steps. Their detailed mechanistic investigation would benefit considerably from precise mapping of the underlying free energy landscape. However, despite

[a]School of Chemistry, Cardiff University, Cardiff, CF10 3AT, Wales, UK. E-mail: leonis@cf.ac.uk
[b]Georgia Institute of Technology, School of Chemical and Biomolecular Engineering, Atlanta, GA 30332-0100, USA

improving numerical algorithms and increasing computational power, their systematic investigation remains extremely demanding and often completely impractical. Methods to explore multi-minima energy landscapes exist[4,5] and recently there has been a pronounced effort to develop novel approaches to overcome this time-scale problem.[3,6] Historically, the use of an external bias aimed at enhancing the frequency of rare events in MD simulations has been realised in methods such as umbrella sampling (US).[7] This is achieved by systematically adding a bias along well-chosen coordinates, in order to sample only small but overlapping regions from which a free energy can be reconstructed. Another bias-oriented method, metadynamics (metaD),[3] exploits the concept of collective variables (CVs), which span a lower dimensional coordinate space. Within the space of these CVs, a history-dependent bias is deposited, which rapidly drives the system away from its initial probable state. The definition of a good set of CVs is therefore key to the success of the method.[8] Within a CV space, standard MD would linger around a minimum close to the initial state corresponding to some value of the CV. Overcoming free energy barriers by metadynamics inevitably leads to broader exploration of the CV values by crossing over to different states within a finite computational time.

Metadynamics takes place within the space of one or a set of CVs, which must be known at the beginning of the calculation. A valid set of collective variables optimally covers the entire space of the conformations of interest. This clearly requires that CVs be engineered with a global free energy map in mind. However, this is evidently problematic, as often the underlying free energy landscape is the desired final result of the simulation and it often cannot be deduced by intuition. Additionally, the requirement for such 'global' CVs directly clashes with the effort to keep their dimensionality as low as possible.

Recently, a novel approach has been introduced, which allows for a "bespoke bias" based on a variational approach.[9,10] Within this approach, the CV problem is partially overcome by reinterpreting its meaning as a local descriptor of meta-stable states. Enhancing static fluctuations will amplify the probability of a crossover to another metastable state. Importantly, the bias can be deposited in such a way that partial filling can be achieved in all the basins of interest. While the success of the method still depends on an appropriate choice of CVs, this represents a novel general scheme for enhanced-sampling MD simulations.

To tackle the intrinsic disparity between the time scales of activated processes and molecular dynamics simulations, the transition path sampling method (TPS)[6,11] was introduced. TPS is based upon the idea that, in a complex system, crossing a barrier between (meta-)stable states may happen in a multitude of different ways. In such a scenario, the potential energy surface of a system is unlikely to be represented by saddle points alone. Instead, numerous points on the potential energy surface become relevant, some of which may still be stationary points, but many others not. The implementation of the method consists of collecting such an ensemble of reactive trajectories, without previous knowledge of the details of the process (that is, no *a priori* information or guess about the reaction coordinates is required – such information should naturally result from the simulation approach). Within a transition ensemble, the relevance of a path is weighted by its probability. TPS collects true dynamical trajectories and can be plugged on top of different molecular dynamics schemes, within different simulation ensembles.[2]

TPS belongs to the class of pathway-orientated methods and allows for an exquisite level of detail in mechanistic investigations.[12–18] Unlike metadynamics, TPS does not deposit any bias potential; similarly to metadynamics, it deploys a low-dimensional order parameter, which distinguishes between the basins of interest. TPS is designed to enhance statistics in intermediate regions between long-lived basins, however it is generally not a method of choice for free energy reconstruction as the main statistical weights are still contributed by (meta-) stable basins.

Here, we present a novel combination of TPS and metadynamics techniques, which we dub 'metashooting'. This novel approach allows for reconstruction of the underlying free energy of a process from TPS trajectories. This combination is based on the propagation of CVs from TPS to metaD. While TPS controls the rapidity of CV variation (inherent to true dynamical trajectories), metadynamics deposits a bias around regions of the CV that vary comparatively slowly. In doing so, TPS keeps the path crossing probability ratios (A → B to B → A) unchanged, thus preventing unilateral over-biasing. By rapidly varying the controlling CVs in TPS, partial basin filling can be achieved while maintaining the intermediate regions bias-free for most of the calculations. Only in its final phases do the details of the intermediate regions appear.

In the following we illustrate the details of our approach by performing a complete metashooting analysis of the solid–solid reconstructive phase transition between the B3 zincblende phase and the B1 high-pressure rocksalt phase of zinc oxide (ZnO). ZnO has proven to be extremely challenging over the years,[19–23] with several attempts to conclusively shed light on competitive mechanisms and intermediates. Saitta and Decremps first discussed the competitive nature of the 'tetragonal' and 'hexagonal' routes in the B4–B1 phase transition, and their work concluded that semiconductors involving d-electrons, such as ZnO, preferred the 'tetragonal' path.[21] However, other work seems to contradict this – notably the experimental work of Liu *et al.*, which utilised high resolution angular dispersive X-ray diffraction, which seemed to indicate that the 'hexagonal' pathway was the more favoured for ZnO at lower pressures.[24] Second harmonic generation experiments pinpointed the role of the pressure medium (hydrostatic *vs.* non-hydrostatic) in promoting the formation of centrosymmetric *iH* or acentric tetragonal *iT*, respectively.[22] In previous work,[15] TPS was used to produce a detailed mechanistic pathway for the B4–B1 transition, and it was shown that the mechanism proceeds *via* a rich series of transition states and intermediates.[12] Additionally, the two 5-coordinate intermediate motifs were visited and were shown to compete with one another in the mechanism. The analysis, based on dynamical trajectories, concluded that a hexagonal intermediate could indeed be visited, however it was not essential to the transformation. In fact, the reconstruction was found to propagate over the interfaces of local tetragonal motifs (Fig. 1).

It is therefore apparent that the phase transitions between different polymorphs of ZnO are non-trivial. Indeed, much debate still rages over their precise nature and work is still being published to attempt to unravel the minutiae of these elusive mechanisms, with a particular emphasis on the B4–B1 transformation. In the following, after a summary of the implementation details, the results of metashooting on a challenging solid–solid phase transition in this material between the B3 and B1 phases are discussed.
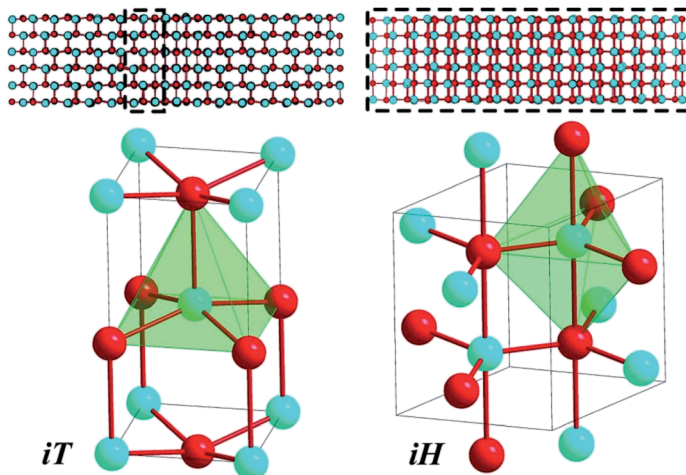
Fig. 1 Two 5-coordinate intermediates (tetragonal *iT* (*I4mm*, left) and hexagonal *iH* (*P6₃/mmc*, right)) were shown to compete against one another in the B4–B1 transformation. Whilst *iT* was only found at the interface between B4 and B1, the *iH* structure sometimes dominated the entire system under study. Despite this, the *iH* phase was shown not to be crucial to the transformation.[12]

Methods

Our complete analysis of the B3–B1 phase transition in ZnO proceeds over 2 steps: (i) TPS calculations, including the definition of a suitable order parameter and (ii) free energy mapping with metashooting, where TPS and metaD are combined.

Transition path sampling

The aim of TPS is to obtain the transition path ensemble,[6,11] and an efficient approach to this task involves the use of Monte Carlo techniques. A random walk is carried out in the space of the trajectories, whereby a new path $x^{(n)}(\tau)$ is generated from an old one, $x^{(o)}(\tau)$. The implementation of the random walk moves is based on the shooting algorithm,[2] whose general premise is to generate pathways by "shooting off" a new pathway from a randomly selected time slice $x_{t'}^{(o)}$ of the previous pathway. Modifications are introduced to $x_{t'}^{(o)}$, yielding $x_{t'}^{(n)}$. From this perturbed time slice at time t', two new path segments are generated forward to time τ and backward to time 0 in molecular dynamics runs within the chosen ensemble. A trajectory is rejected if the introduced modifications cause the new trajectory not to connect the basins of interest. Otherwise, it is accepted with a given probability.[2] The acceptance/rejection criteria thus depend only on the shooting point, and can easily be translated into an algorithm. The introduction of such a probability weighting scheme allows the system to cross intermediate activation barriers of the order of $k_B T$ or lower. Perturbations are introduced as momenta modifications with strict conservation of the total linear and angular momenta.

Order parameter

Transition path sampling calculations require the initial and final configurations of the system to be strictly distinguishable, such that there is no overlap of their

respective basins when projecting onto the order parameter.[25] A handy yet powerful order parameter used in previous path sampling calculations is the average coordination number of the system. The coordination sphere (CS) is a set of integers $n_1, n_2, \ldots n_i$, in which the ith term is the number of atoms in "shell" i that are bonded to atoms in "shell" $i - 1$. Shell 0 consists of a single atom, and the number of atoms in the 1st shell n_1 is the conventional coordination number (CN). This quantity can be calculated for all atoms in a system, collated, and averaged over the whole structure. The 1st three shells of the coordination spheres of the three phases of zinc oxide are: zincblende (B3) [4,12,24], wurtzite (B4) [4,12,25] and rocksalt (B1) [6,18,38]. The 1st coordination numbers (1st CNs) of zincblende and rocksalt are 4 and 6, respectively. During TPS, the coordination spheres of each atom were calculated up to the 3rd shell, but only the 1st CN was utilised to trace the progress of the path sampling calculations. The 2nd and 3rd CNs were recorded simply to gain further insight into the underlying details of the transition pathways.

Using only the 1st coordination shell means that there is no discrimination between the B3 and B4 configurations. As the system starts from a metastable B3 structure (see below), changes in the final regime can be expected. Furthermore, using solely the 1st coordination number as the order parameter exerts absolutely no bias on the calculation. The variation in the order parameter arises simply as a result of the application of the transition pressure to the system. Hence, monitoring the 1st coordination sphere does not force the system to evolve in any particular way over the course of the path sampling iterations. In TPS, both the overall mechanism and the trajectory endpoints are rectified.[26] Therefore, the order parameter is chosen with minimal constraints in mind, in order to allow the process to express the most probable outcome. A higher-order CS would drive the system into a specific transformation regime within a biased TPS approach, in order to study specific, less probable transformation pathways.

Metashooting

Metashooting combines TPS and metadynamics to efficiently reconstruct the underlying free energy landscape from TPS trajectories. This way, computing time will not be expended by filling other regions of the configuration space that have no relevance to the transition.

The general procedure of the metashooting approach for a transition A → B is as follows:

(1) TPS step 1: Starting from a randomly chosen snapshot along the TPS trajectories, shoot off a new trajectory and time propagate.

(2) MetaD step 1: Once basin A is reached, deposit a bias using a metadynamics scheme (in this case, using the well-tempered approach), in order to locally partially fill the as-yet uncharacterised energy well. At this stage, a small number of metasteps are chosen, allowing the CV to vary about the characteristic value of the basin.

(3) TPS step 2: Generate another trajectory (as in step 1) and propagate into the now-biased basin A, to generate a true MD velocity distribution on the biased potential.

(4) TPS step 3: Similarly to TPS step 1, shoot off a new trajectory from the same snapshot as above and propagate, only this time backwards in time, to basin B.

(5) MetaD step 2: In basin B, apply the same metadynamics scheme as above for a number of steps, in order to partially fill this second basin of attraction.

(6) TPS step 4: Re-shoot a novel trajectory from a random snapshot of the previous trajectory. Propagate into the now-biased basin B.

Repeat the above iterations of metadynamics and TPS moves until the underlying free energy profile is fully converged.

The number of metadynamics steps for each metashooting move can be chosen such that only moderate variations around the current CV value are allowed, as ΔCV_{Max}. Alternatively, as a better means to control bias deposition without imbalance, the TPS path acceptance probability in TPS can be chosen to stay the same, *i.e.* at no point can crossing over the activation energy barrier become easier for either direction A → B or B → A. This is realised within an adaptive scheme of rescaling momenta modifications (see below).

Calculation details

Molecular dynamics. Born–Oppenheimer molecular dynamics simulations were performed using the *cp2k* package within the *NpT* ensemble at $T = 300$ K and $p = 9.8$ GPa. At this pressure, the enthalpies of B1 and B3 are equal. Constant pressure and temperature were ensured by the Martyna–Tobias–Klein[27] algorithm, allowing for anisotropic shape changes of the simulation box.[28] Interatomic forces were calculated from a Buckingham pair potential,[29] which has been previously validated.[15] Long-range electrostatic effects were accounted for using an Ewald summation. Time propagation was performed with a time-reversible predictor–corrector algorithm. An integration timestep of 0.2 fs was used. The simulation box contained 1200 Zn–O atom pairs.

Transition path sampling. The sampling starts from an initial trajectory connecting the limiting phases. A new trajectory is generated by randomly selecting a configuration from an existing trajectory and slightly modifying the atomic momenta. The modifications δp are applied to randomly chosen pairs of atoms (i,j) according to: $\vec{p}_{new} = \vec{p}_{old} + \delta p(\vec{r}_j - \vec{r}_i)/|\vec{r}_j - \vec{r}_i|$ and $\vec{p}_{new} = \vec{p}_{old} - \delta p(\vec{r}_j - \vec{r}_i)/|\vec{r}_j - \vec{r}_i|$, conserving both momentum and angular momentum. The resulting atomic momenta are rescaled by a factor of $\sqrt{E_{kin}^{old} \Big/ \left(\sum_{i=1} |\vec{p}_i|^2 \Big/ 2m_i \right)}$, in order to conserve the total kinetic energy. The propagation of the modified configuration in both directions of time $(-t, +t)$ generates a new trajectory. Repeating this step provides a set of trajectories, with successful examples representing the transition ensemble. TPS requires an initial guess trajectory, which we obtained from a topological/geometric approach of interpolation between the two limiting crystal structures (B1 and B3), as described elsewhere.[13,14,26] Momenta modifications were introduced using an adaptive scheme, which entailed rescaling δp to larger or smaller values, based on whether this δp led to a successful or a failed trajectory, respectively. The acceptance ratio for the production runs was generally between 40–60%.

Metadynamics. Metadynamics calculations were carried out using the plumed plug-in,[30] in which a history-dependent Gaussian bias is added to the underlying potential. The well-tempered[31] prescription of metadynamics was used, in which the heights of the Gaussian hills are rescaled at each step. The collective variables corresponded to the 1st and 3rd coordination spheres of the zinc oxide

systems – analogous to the order parameter used in TPS. The coordination sphere in plumed is mapped onto a switching, differentiable function,

$$s(r) = \frac{1 - \left(\frac{r - d_0}{r_0}\right)^n}{1 - \left(\frac{r - d_0}{r_0}\right)^m}.$$ r is the Zn–O inter-atomic distance, while r_0, d_0, m and n are

adjustable parameters. For the 1st and 3rd coordination spheres the parameters were set to (2.6, 0.1, 6, 12) and (5.3, 0.1, 6, 12), respectively. These parameters reflect a one-to-one relationship between the geometrically-calculated coordination numbers and the values based on the switching function, despite resulting in slightly exaggerated values for the 3rd coordination sphere of B4/B3 (28/29 instead of 24/25). Gaussian functions were deposited every 500 simulation steps. On average, 5 metasteps were calculated on each side of the trajectory during each iteration of the metashooting procedure. The height of the Gaussians was set at 1000 kJ mol^{-1}. The Gaussian widths were chosen in accordance with the variance of the CV in an unbiased simulation – as such, these were set to 0.2 and 1.0 for the 1st and 3rd coordination numbers, respectively. The well-tempered 'bias-factor' was set at 10 000, which is large enough to escape the minima and re-scale the deposited Gaussians within a meaningful time-frame.

Results

Mechanistic analysis

TPS was initiated from a B3–B1 trajectory regime, obtained from the modelling approach described and used elsewhere.[13,32] One of the advantages of initiating the process from a plausibly designed trajectory is that it takes fewer TPS iterations for the system to move away from the initial, unfavourable regime to a more probable one. The process of moving towards a probable regime is known as trajectory decorrelation. In fewer than 50 TPS iterations, the procedure steered the trajectory away from the collective motion encoded by the geometric model towards a regime characterised by nucleation events. The trajectories continued to evolve until no further major mechanistic changes were observed, which occurred after several hundred iterations of the path sampling procedure.

Fascinatingly, the final 4-coordinated product was not a pure zincblende (B3) structure, but a mixed B4/B3 system dominated by the wurtzite (B4) form. In all the successful trajectories, the final material consisted of only 10–20% B3, with the more stable hexagonal form forming the remainder of the structure. The order parameter based on the 1st coordination number permits both the wurtzite and zincblende configurations. Whilst the initial, unfavorable regime was B1 → B3, TPS rectified the mechanism into B1 → B4/B3, which contains predominantly hexagonal components.

The complexity of the system leads to several types of trajectories, with some degrees of (dis)similarity in the details of the reconstruction.

A representative B1 → B4/B3 trajectory is shown in Fig. 2. A nucleus of 4-connected ZnO (Fig. 2, 0.6 ps) forms within B1, followed by the formation of regions of 5-fold coordination at the interface with B1 (Fig. 2, 0.7 ps), which are subsequently outgrown by 4-fold coordinated regions (Fig. 2, 0.9–4.7 ps). Regions of 5-fold coordination (mostly iH) are found between the expanding 4-connected
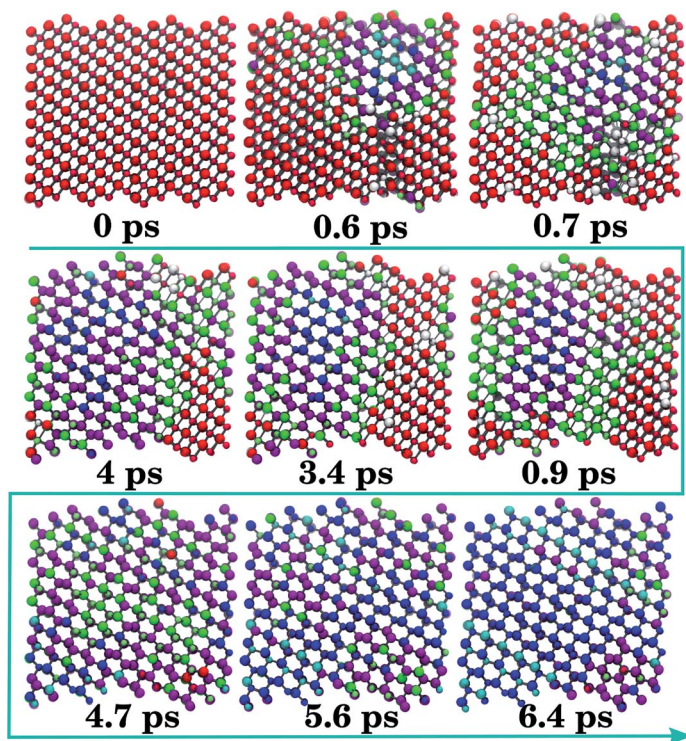
Fig. 2 Coordination-coloured snapshots from a representative trajectory, showing the sequence from rocksalt (B1) to a mixed wurtzite/zincblende (B4/B3) configuration. A nucleus forms in B1 within 500 fs, followed by shearing along $(111)_{B1}$ at 0.6–0.7 ps. The nucleus grows further over the intermediates characterised by 5- and 4-fold coordination (0.9–4.7 ps), until the system transforms into a 4-coordinate motif, which is composed primarily of hexagonal wurtzite (5.6–6.4 ps). Colour code: red: B1; green: *iH*, *iT*; dark blue: B4 (wurtzite); light blue: B3 (zincblende); purple: 4-coordinated, but not corresponding to B3 or B4.

region and B1. The transformation into mixed B4/B3 entails the complete resorption of areas of 5-fold coordination (Fig. 2, 4.7–6.4 ps). The transformation of the remaining B1 motifs (4 ps) is associated with a sharp volume expansion, which causes the box volume to fluctuate, under the appearance of local 5-fold motifs (Fig. 2, 4.0–4.7 ps).

From a configuration of the type seen in Fig. 2 at 0.6 ps, a less severe compression leads to the growth of a 5-fold region at the expense of any 4-connected one. This accompanies the formation of a mostly 5-coordinated intermediate containing residual B1, denoted B1–*iH*. This configuration persists for as long as 500 ps in subsequent molecular dynamics simulations. The presence of an intermediate, which is not vital to the transformation and corresponds to a 5-membered inter-mediate along the transition pathway, was noted in previous work.[12] The nature of this intermediate is the hexagonal *iH* form, whilst the tetragonal analogue *iT* only appears as a motif at interfaces during the reconstruction.

The B4/B3 → B1 mechanism slightly deviates from the reverse process (Fig. 3). In this trajectory type, the starting 4-coordinate material is first collectively distorted
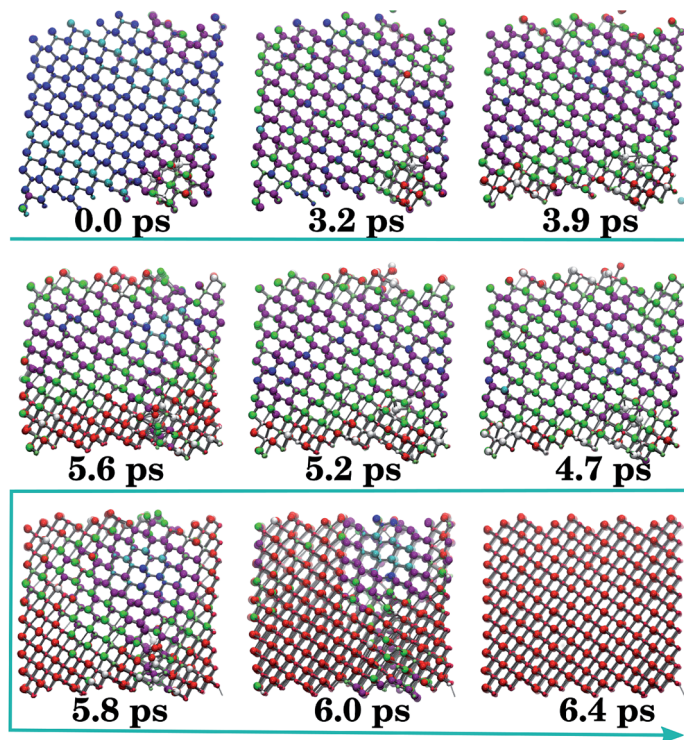
Fig. 3 Coordination-coloured snapshots of a representative B4/B3 → B1 trajectory. Mixed (B4/B3) wurtzite/zincblende (0.0 ps) is followed by a 4-coordinated distorted configuration (0.0–3.2 ps), purple atoms. Subsequent growth of 5-coordinated areas (green), mainly of type *iH*, is observed (3.2–5.2 ps), from which B1 (red) develops (5.2–6.0 ps) until the structure locks-in to 6-coordinate B1 (6.4 ps). Colours as in Fig. 2.

(Fig. 3, 0.0–3.2 ps, purple atoms). From the 5-coordinated areas of the *iH* motifs (Fig. 3, 3.9–5.2 ps), the B1 structure develops (Fig. 3, 5.2–6.4 ps). Thus, over half of the transformation time is spent slowly squeezing the system until it can transform into 5-coordinated *iH* regions, indicating the presence of a deep but fairly gently-sloping energy barrier separating the B4/B3 basin from the intermediate configurations. In all the configurations, the *iH* motifs are relatively long-lived. This is not the same as the reverse B1 → B4/B3 transition, in which the *iH* regions were less extensive and promptly eliminated at the interfaces (Fig. 2), with the majority of the simulation time being dominated by coexisting 4- and 6-coordinated motifs.

Consistent with previous calculations,[12] different intermediates are associated with different types of trajectories. The presence of B3 adds further structures to an already complex scenario. In previous work, low transition pressures (less than 10 GPa) were observed to favour the appearance of *iH*, which benefits from a less severe $[0001]_{B4}$ compression. The results presented here appear to corroborate that view.

Metashooting

In order to better understand the competitive nature of different mechanisms, the free energy (FE) was evaluated by coarse-graining the potential energy surface

onto a suitable set of collective variables. Free energy surfaces have found extensive uses for the description of biological systems (in particular peptide calculations[33]), but are less frequently used in the solid state.

In metashooting, mechanistic details acquired from the transition path sampling procedure remain unspoiled. As the process is driven by TPS trajectories, metastable basins are filled in a balanced way – starting from the two limiting configurations, and finishing with the characterisation of all the relevant intermediate configurations. Repeated iterations of the metashooting algorithm (see Methods) enable the free energy landscapes corresponding only to the courses of the TPS-derived trajectories to be built up in a stepwise and equally distributed manner. As a result, regions around the transition states are considered only at the end of the analysis, avoiding any potential bias that may scramble the mechanistic details.

The algorithm should first fill up the two starting basins of attraction from the final TPS trajectories (B4/B3 and B1). Visiting these regions should become discouraged as a result of the bias potential. The progress of metashooting can be monitored by visualising the Gaussian deposition. Eventually, the metashooting algorithm should fully characterise the central (around the metastable basins) and peripheral regions of the CV space over the course of the trajectory. Convergence will be seen in a similar way to a standard metadynamics scheme – that is, the underlying free energy landscape should no longer significantly change with further iterations of the procedure. At this point, meaningful analysis of the intermediates and energy barriers associated with each process can be carried out.

Harnessing the free energy in this way should have a number of advantages. The landscape can be thought of as the surface upon which the trajectories travel, and the individual characteristics of the trajectories are determined by which regions of the surface are visited. Thus, its elucidation allows for a complete quantitative analysis of the underlying thermodynamics and kinetics of the process; measurements of the activation energy and changes in the free energy between different steps of the procedure, as well as the starting and finishing products, can trivially be taken from the plot. Moreover, the intermediates and transition states that are vitally important to the transformation will appear obvious along the reaction scheme. Such insight into trajectories using only path-based analyses of transformations is not trivial.

The free energy profile was constructed from the converged TPS trajectories. After 250 iterations of the metashooting procedure, the free energy had fully converged into a detailed landscape.

Progression of metashooting

The evolution of the free energy landscape according to the metashooting algorithm is first discussed, followed by an analysis of the final profile generated by the procedure.

The first few iterations of the procedure (Fig. 4) accumulated Gaussian functions only within the two limiting basins (B4/B3 and B1). A small 'shoulder-peak' formed in the 4-coordinate basin as a result of oscillations between different forms of the low-coordination material, including B3, B4 and other forms.
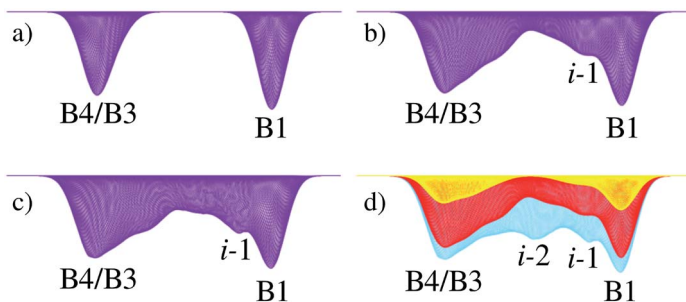
Fig. 4 Free energy landscapes at different stages of metashooting. (a) Filling of the limiting basins centred on the B4/B3 and B1 motifs (1st metashooting cycle), (b) appearance of a minimum *i*-1 after 50 cycles, (c) intermediate areas are explored after 200 cycles. (d) Stages 10 (yellow), 75 (red) and 225 (blue). In the latest cycles only the intermediate *i*-2 appears, which corresponds to B1–*iH*. The 3D free energy is plotted as a function of the 1st and 3rd coordination numbers. See main text for further details.

After 50 iterations (Fig. 4b), a notable 'shoulder-peak' adjacent to the B1 basin was observed. After 150 iterations, this was resolved into a *bona fide* minimum, *i*-1. By iteration 200, a second minimum centred halfway between the two limiting basins started to appear (Fig. 4c), which was fully resolved by iteration 225 (Fig. 4d), denoted *i*-2. By 250 iterations (Fig. 5), the plot was fully converged – the energy barriers and positions of the minima remained unchanged with any further iterations of the metashooting procedure.

Energy profile and intermediate configurations

There are at least seven points of interest on the free energy surface – the two initial basins of attraction (corresponding to B4/B3 and B1), two intermediate
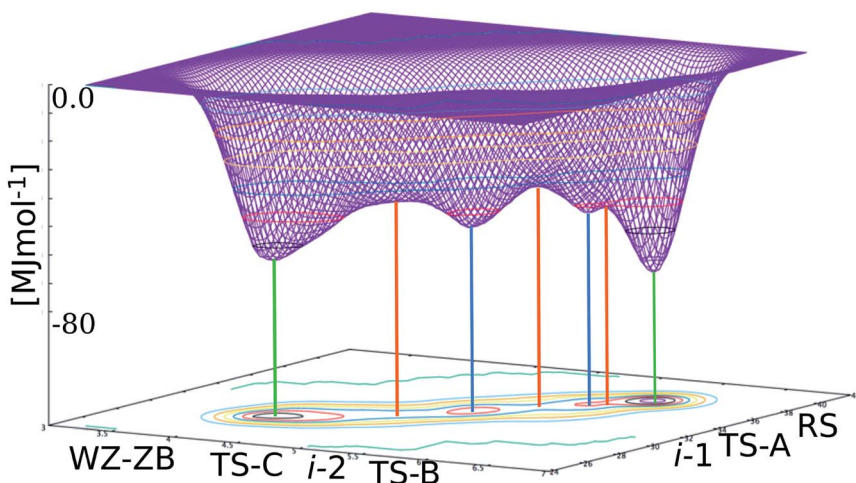

Fig. 5 Three-dimensional plot of the converged free energy surface for the transition after 250 iterations of the metashooting procedure. Two main basins, WZ-ZB (B4/B3) and RS (B1), two intermediate basins (*i*-1 and *i*-2), and three transition states (TS-A, TS-B, TS-C) are visible.

basins and three maxima, corresponding to the transition state regions. The structures and roles of these turning points in the transformation can be ascertained by obtaining the 1st and 3rd coordination numbers of the turning points on the graph and matching these up to structures along the hundreds of trajectories from both the 'plain' TPS runs and the metashooting trajectories (Table 1). It stands to reason that these regions must play important roles in the transformation mechanism.

Free energy intermediates and maxima

i-1 is the metastable basin adjacent to the B1 basin. Every trajectory congregates at *i*-1 before proceeding along a single pathway to or from the rocksalt structure. Despite the small and shallow dimensions of the basin, it clearly plays a central role in the transformation. It is separated from B1 by TS-A. The configuration associated with this maximum is an almost entirely B1 structure at the moment of deformation of the rocksalt motifs (Fig. 6). For the first few Zn–O sites this entails overcoming an activation energy, as they leave their 6-fold coordination.

In *i*-1 a portion of the pristine B1 fragments contains 4- and 5-coordinated motifs. With respect to the way the square pattern is opened, the locally 5-coordinated motifs correspond to *iH*. This small region is extremely short-lived but is present in every successful trajectory sampled, thus explaining its place on the free energy profile. Interestingly, this mechanistic pattern cannot propagate as it is. The next free energy barrier TS-B is overcome by including an additional $(111)_{B1}$ shearing mechanism, as seen in both the forward and reverse TPS trajectories. This configuration, which includes an enlarged seed region and the onset of shearing, corresponds to the free energy maximum TS-B. The TS-B configuration could easily be seen in the TPS trajectories, however the fact that this part of the transformation corresponds to two distinct events would have been completely missed without the meta-shooting analysis.

The second intermediate region, *i*-2, is a deep energy well placed halfway between the two limiting basins. Trajectories visiting configurations characterised by the 1st CN in the interval [5.0–5.3] and the 3rd CN within [32–34] can in principle contribute to this basin. Accordingly, three different types of trajectory can be distinguished:

(1) Successful B4/B3 → B1 trajectories, along which extended 5-coordinated regions are present (see Fig. 3, 3.9–5.6 ps).

Table 1   1st and 3rd coordination spheres (as defined by the switching function) at the centre of the regions defining the turning points

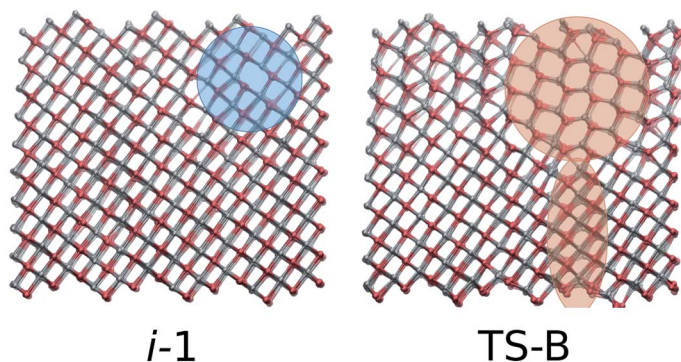| Turning point	1st CN	3rd CN
RS (B1)	5.98	37.90
TS-A	5.75	36.66
i-1	5.68	36.21
TS-B	5.43	34.86
i-2	5.15	33.13
TS-C	4.87	31.62
WZ-ZB (B4/B3)	4.10	29.22

i-1   TS-B

Fig. 6 (Left) Configuration corresponding to the small minimum *i*-1, which is just after the very first rocksalt motifs locally deform into hexagons. (Right) A representative configuration corresponding to the maximum TS-B, showing an enlarged seed accompanied by $(111)_{B1}$ layer shearing.

(2) Successful B1 → B4/B3 trajectories, in which Zn or O is 4-, 5- and 6-coordinated, giving an overall average 1st coordination sphere of five (see Fig. 2, 0.7–0.9 ps).

(3) Trajectories where the system visits and lingers in the predominantly 5-coordinate intermediate B1–*iH*. The average time spent in this basin is >200 ps.

Representative trajectories of the three types above are mapped on the free energy surface in Fig. 7. Successful trajectories tend to keep a higher 3rd coordination number while they approach the large *i*-2 basin. On the other hand, a lower 3rd CN is associated with a trajectory entering an intermediate basin. Closer to the B1 basin, the trajectories are more similar, while the B4/B3 basin can be escaped in different ways. Around TS-B, mechanistic competition (assisted by box size fluctuations[12,22]) may result in very different trajectories.

It seems likely, therefore, that the depth of the intermediate basin (and hence the energy barriers associated with it) cannot be taken as representing a proper intermediate along a particular type of trajectory. Instead, numerous possible configurations with similar average coordination number values all add weight to the free energy at these values of the collective variables, even though not every configuration is realisable in every trajectory.

Unlike TS-A and TS-B, TS-C corresponds to a very 'gentle' maximum, with a very shallow gradient either side of it. This again corroborates what was seen in the path sampling investigations – the system spends a great deal of time in a 4-coordinate state, or a mixture of 4- and 5-coordinate states, as the result of a slow, steady process, which is mapped here onto a gentle free energy gradient. Like its associated intermediate basin *i*-2, the height of TS-C is influenced by a number of configurations, all of which contain some combination of 4-coordinated ZnO, *iH* intermediate and residual B1. Crossing this maximum implies a significant coordination number (order parameter) change, and corresponds to mixed motif configurations as at 4 ps in Fig. 2 for the B1 → B4/B3 direction, or at 4.7 ps for B4/B3 → B1 (Fig. 3).

The appearance of mixed B1–*iH* agrees with previous studies on ionic compounds undergoing the same B4/B3 → B1 transition,[34–36] which may indicate the transferability of this mechanistic analysis.
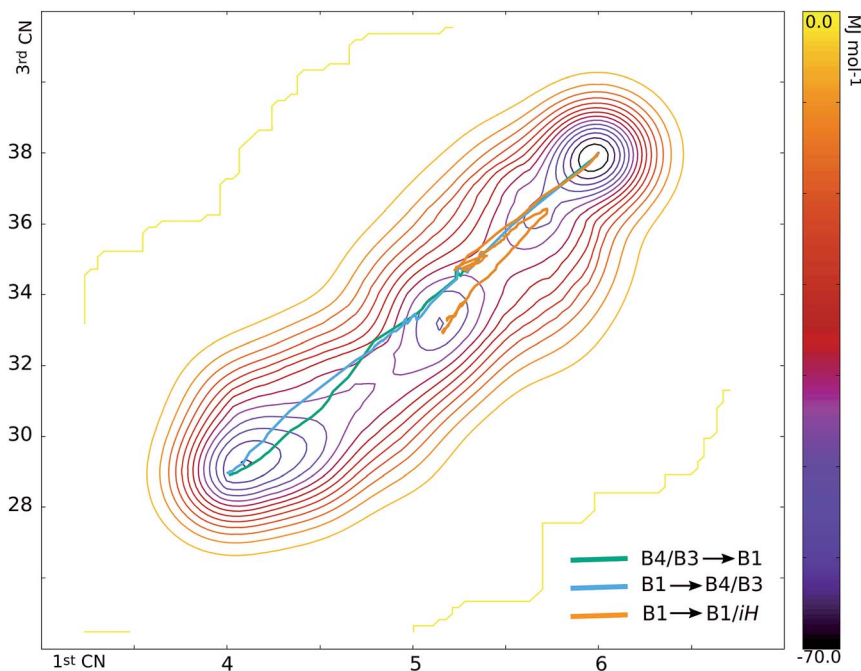
Fig. 7 Three representative TPS trajectories, mapped on the free energy surface. Successful B4/B3 → B1 trajectory (green), successful B1 → B4/B3 trajectory (light blue) and B1 → mixed B1–*iH* (orange). The latter is a long-lived intermediate configuration, which coincides with the intermediate *i*-2 basin.

Energy barriers

From the converged free energy surface, it is apparent that there are six distinguishable energy barriers associated with the pathway – three for B4/B3 → B1 (Fig. 8, E_{1-3}), and three for the reverse process (Fig. 8, E_{5-7}).

Trajectory variety has shown that the system may proceed along different pathways, therefore crossing different features of the free energy plot. This will
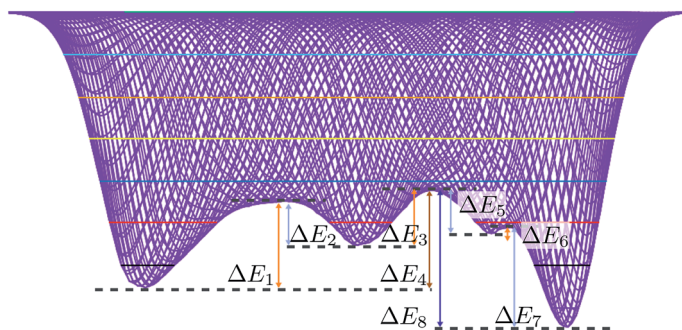

Fig. 8 Summary of energy barriers associated with the free energy profile. The activation energies are defined as the largest barriers from the starting basin. Note that these barriers are calculated as the difference from the centre of each of the turning points.

lead to different values for the total amount of energy overcome by the individual trajectories. The following quantitative analysis of the energy differences assumes measurement to and from the centre of each of the turning points, as described in Table 2, thereby indicating the energy associated with the minimum energy transition pathway that visits all maxima and minima. Every trajectory must surmount an energy barrier at least equivalent to the activation energy presented, as successful trajectories must overcome the highest energy barrier to cross to the other side, irrespective of the path taken.

It is also apparent from inspection that the energy barriers for each stage of the transformation are very different for the forward and reverse trajectories. This may go some way towards explaining why there are different pathways associated with the forward and reverse trajectories, and perhaps why there is such notable hysteresis present in such transformations.

The change in free energy is calculated as the difference between the energy values of the B1 and B4/B3 basin centres. The initial pressure was chosen based on the zero enthalpy difference between B3 and B1, which was dictated by the model used to commence the TPS cycles (see above).

The difference in free energy between the forward (B4/B3 → B1) and backward transformation is not zero at the chosen simulation pressure ($p = 9.8$ GPa). The inclusion of lattice dynamics[37] in first principles calculations was shown to shift the ($\Delta H = 0$) "equilibrium" pressure to a lower value. Consistent with those findings, Fig. 8 pinpoints the role of entropy, which is positive going from B4/B3 to B1. The disordered nature of B4/B3 reinforces this effect.

Discussion and conclusions

Using transition path sampling methods, the phase transition characterised by a sequence of nucleation and growth events between mixed wurtzite/zincblende (B4/B3) and rocksalt (B1) has been described. The forward and reverse trajectories were found to be different, each with different competing intermediates and preferences for one mechanism over the other for both transformations.

A typical B1 → B4/B3 trajectory was characterised by the formation of an initial seed, followed by the gradual transformation to a mixed 4-coordinate product via coexisting motifs. Trajectories could also end up trapped in a long-lived intermediate 5-coordinate basin, corresponding to a mixed rocksalt-5-coordinate

Table 2 Energy barriers associated with the forward and reverse trajectories. Energy differences are measured to and from the centres of the turning points on the energy surface, corresponding to the (unobserved) minimum energy pathway trajectory

Energy difference	kJ mol^{-1}	$k_{\mathrm{B}}T$/ZnO
ΔE_1	20 447.0	6.8
ΔE_2	10 483.8	3.5
ΔE_3	13 085.9	4.4
ΔE_5	10 376.6	3.5
ΔE_6	1357.0	0.5
ΔE_7	23 726.3	7.9
ΔE_4	35 069.9	11.6
ΔE_8	44 586.7	14.9

hexagonal structure denoted B1–*iH*. The B4/B3 → B1 transition entailed a more stepwise process, which involved a transition to a 4-coordinate mixture *via* local motifs of hexagonal *iH* before transforming to the final B1 product.

A novel method denoted 'metashooting' was developed as a combination of TPS and well-tempered metadynamics, using the 1st and 3rd average coordination spheres as the collective variables. TPS ensured rapid variations of the CVs, and balanced filling of the basins of interest. The metashooting procedure is able to decipher the underlying free energy landscape of the transformation, and structural information about the nature of appropriate maxima and minima can be extracted from the resultant plot. In addition, the energy barriers and changes in free energy can be ascertained and the successful and failed trajectories can be mapped onto the final free energy plot, in order to gain a true understanding of the energetics and structural changes inherent within each individual trajectory.

Metashooting can accommodate additional collective variables, for example higher coordination sequences, in order to further distinguish between the phases. Using the 1st and 3rd coordination spheres was convenient in this case as it was an identical measurement to the order parameter used in the path sampling calculations. In general, the CV set used on the metashooting layer can be more detailed then the order parameter utilised in the TPS steps.

It is believed that the metashooting procedure, which combines transition path sampling and metadynamics, could set a novel paradigm for future investigations into condensed matter phase transitions. Such a scheme could easily be transferred to another system, under any level of theory, to ascertain the underlying thermodynamics and kinetics of a temperature- or pressure-induced phase transformation.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 D. Chandler, *J. Chem. Phys.*, 1978, **68**, 2959–2970.
2 C. Dellago, P. G. Bolhuis and P. L. Geissler, *Lect. Notes Phys.*, 2006, **703**, 349–391.
3 A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12562–12566.
4 F. Glover, *Interfaces*, 1990, **20**, 74–94.

5 P. Sibani, J. C. Schön, P. Salamon and J. O. Andersson, *Europhys. Lett.*, 1993, **22**, 479–485.

6 C. Dellago, P. G. Bolhuis, F. S. Csajka and D. Chandler, *J. Chem. Phys.*, 1998, **108**, 1964–1977.

7 G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, 1977, **23**, 187–199.

8 C. Abrams and G. Bussi, *Entropy*, 2014, **16**, 163–199.

9 O. Valsson and M. Parrinello, *Phys. Rev. Lett.*, 2014, **113**, 090601.

10 O. Valsson and M. Parrinello, *J. Chem. Theory Comput.*, 2015, **11**, 1996–2002.

11 P. G. Bolhuis, C. Dellago and D. Chandler, *Faraday Discuss.*, 1998, **110**, 421–436.

12 S. E. Boulfelfel and S. Leoni, *Phys. Rev. B*, 2008, **78**, 125204.

13 S. Leoni and D. Zahn, *Z. für Kristallogr.*, 2004, **219**, 339–344.

14 D. Zahn and S. Leoni, *Phys. Rev. Lett.*, 2004, **92**, 250201.

15 S. E. Boulfelfel, D. Zahn, Y. Grin and S. Leoni, *Phys. Rev. Lett.*, 2007, **99**, 125505.

16 S. Leoni, S. E. Boulfelfel and I. A. Baburin, *Z. Anorg. Allg. Chem.*, 2011, **637**, 864–869.

17 S. E. Boulfelfel, G. Seifert, Y. Grin and S. Leoni, *Phys. Rev. B*, 2012, **85**, 014110.

18 S. E. Boulfelfel, A. R. Oganov and S. Leoni, *Sci. Rep.*, 2012, **2**, 471.

19 D. Zagorac, J. C. Schoen and M. Jansen, *J. Phys. Chem. C*, 2012, **116**, 16726–16739.

20 C. R. A. Catlow, S. A. French, A. A. Sokol, A. A. Al-Sunaidi and S. M. Woodley, *J. Comput. Chem.*, 2008, **29**, 2234–2249.

21 A. M. Saitta and F. Decremps, *Phys. Rev. B*, 2004, **70**, 035214.

22 L. Bayarjargal and B. Winkler, *Z. Kristallogr. - Cryst. Mater.*, 2014, **229**, 92–100.

23 J. Cai and N. Chen, *J. Phys. Condens. Matter.*, 2007, **19**, 266207.

24 H. Z. Liu, Y. Ding, M. Somayazulu, J. Qian, J. Shu, D. Häusermann and H. K. Mao, *Phys. Rev. B*, 2005, **71**, 212103.

25 P. G. Bolhuis, D. Chandler, C. Dellago and P. L. Geissler, *Annu. Rev. Phys. Chem.*, 2002, **53**, 291–318.

26 S. Leoni, R. Ramlau, K. Meier, M. Schmidt and U. Schwarz, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 19612–19616.

27 G. Martyna, D. Tobias and M. Klein, *J. Chem. Phys.*, 1994, **101**, 4177–4189.

28 G. Martyna, M. Tuckerman, D. Tobias and M. Klein, *Mol. Phys.*, 1996, **87**, 1117–1157.

29 D. J. Binks and R. W. Grimes, *J. Am. Ceram. Soc.*, 1993, **76**, 2370–2372.

30 G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni and G. Bussi, *Comput. Phys. Commun.*, 2014, **185**, 604–613.

31 A. Barducci, G. Bussi and M. Parrinello, *Phys. Rev. Lett.*, 2008, **100**, 020603.

32 S. Leoni and R. Nesper, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 2000, **56**, 383–393.

33 D. Granata, F. Baftizadeh, J. Habchi, C. Galvagnion, A. De Simone, C. Camilloni, A. Laio and M. Vendruscolo, *Sci. Rep.*, 2015, **5**, 15449.

34 J. C. Schön and M. Jansen, *Comput. Mater. Sci.*, 1995, **4**, 43–58.

35 J. C. Schön, *Z. Anorg. Allg. Chem.*, 2004, **630**, 2354–2366.

36 J. C. Schön, *Process. Appl. Ceram.*, 2015, **9**, 157–168.

37 A. Seko, F. Oba, A. Kuwabara and I. Tanaka, *Phys. Rev. B*, 2005, **72**, 485.