

Evolutionary crystal structure prediction as a tool in materials design

Artem R Oganov^{1,2,3} and Colin W Glass¹

¹ Laboratory of Crystallography, Department of Materials, ETH Zurich, Wolfgang-Pauli-Strasse 10, CH-8093 Zurich, Switzerland

² Geology Department, Moscow State University, 119992 Moscow, Russia

E-mail: a.oganov@mat.ethz.ch

Received 30 August 2007

Published 24 January 2008

Online at stacks.iop.org/JPhysCM/20/064210

Abstract

Ab initio methods allow a more or less straightforward prediction of numerous physical properties of solids, but require the knowledge of their crystal structure. The evolutionary algorithm USPEX, developed by us in 2004–2006, enables reliable prediction of the stable crystal structure without relying on any experimental data. Numerous tests (mostly for systems with up to 28 atoms in the unit cell, and a few tests with up to 128 atoms/cell) showed a success rate of nearly 100%. USPEX has resulted in a number of predictions of hitherto unknown stable structures. We give a short overview of the method, introducing some new developments and results, and discuss a few alternative approaches. The method is illustrated by a test on an 80 atom supercell of MgSiO₃ and by the search for new materials with compositions Al₁₃K and Al₁₂C.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The search for the stable crystal structure from the chemical composition has long been regarded as a difficult problem [1], central for materials design and for understanding the behaviour of materials at extreme conditions. Many approaches have been used to solve this problem. The simplest one is to explore thermodynamic properties of a list of candidate structures (which can be made of known structures of analogous materials, or new structures guessed using chemical intuition), and this is still perhaps the most widely used approach. Problems arise surprisingly often—almost every time when a totally unexpected and hitherto unknown structure is actually stable (only sometimes can a new structure be obtained from previously known ones, e.g. if there is a low-energy transition pathway between the structures). Some such cases will be illustrated below. A somewhat similar, but much more advanced, approach involves data mining [2], which derives rules of stability of crystal structures from a large set of *ab initio* calculations. A number of intuitive schemes have been developed (e.g. ideas of structure diagrams, polyhedral clusters—[3]), but their application usually requires a large experimental data set, and/or good understanding of the compound.

From the materials design perspective, it may be more desirable to have a method that requires no prior knowledge or assumptions about the system. Below we briefly discuss our personal views on a few such algorithms. Simulated annealing [4–7] is at first appealing, because this algorithm was inspired by crystallization through annealing—but its applications remain rather limited. Minima hopping [8] and metadynamics [9–11] show good promise. Recently it was shown that even relaxing randomly produced structures can deliver the stable structure [12] in a feasible time, but only for systems containing a small number of atoms in the unit cell. Finally, we will discuss our evolutionary algorithm USPEX (Universal Structure Predictor: Evolutionary Xtallography) [13–15] and illustrate it with a few examples. Successes of all these different methods show that non-empirical prediction of stable crystal structures from chemical composition is possible and is very promising for materials design.

2. Simulated annealing

Simulated annealing, being a modified version of the stochastic hill-climber, takes a random step from the current position and accepts it with a certain probability. This probability p is derived from the change in fitness (ΔE) and the ‘temperature’

³ Author to whom any correspondence should be addressed.

T , e.g. as

$$p = \exp\left(-\frac{\Delta E}{k_B T}\right), \quad (1)$$

where k_B is the Boltzmann constant. The temperature starts at a high value, making the search more or less a random walk (depending on exact implementation), and decreases during the run, coming ever closer to the hill-climber (where only better solutions get accepted). Simulated annealing is often used as a benchmark, due to its fast implementation, or when the goal is not to find the very best, but just a good solution. For application to crystal structure prediction the main drawback is that the algorithm easily gets stuck in local minima. The barriers between minima are usually very large compared to the energy differences—thus, to overcome a barrier (not just jumping over it, but overcoming it), steps with large decrease in fitness need to be made. Since the acceptance probability decays exponentially with decrease of fitness between steps and temperature (equation (1)), a high starting temperature and extremely slow cooling are required. This makes simulations expensive, forcing one to use cheap fitness functions, rather than *ab initio* energies. Furthermore, while for every optimization problem a careful balance between exploitation of information obtained during the run and exploration of new information is crucial, simulated annealing barely exploits any information obtained (except the current position and fitness), which seems to be unfavourable for crystal structure prediction. Successful early showcases of simulated annealing include the determinations of structures of framework materials [4] and NbF_4 [5] using experimental cell parameters and heuristic fitness functions, extensive structure search for silica polymorphs [6] and several test systems [7] using no information on the unit cell and more or less realistic potentials for the calculation of fitness.

3. Minima hopping

A more promising approach is minima hopping [8]. The random steps of simulated annealing are replaced by local optimization to find the local minimum and molecular dynamics to escape it. This is in itself clever, for it more efficiently finds the local minimum (using physical gradients) and most likely transitions via a low-energy saddle-point (such transitions are likeliest in molecular dynamics). Furthermore, it includes a tabu, i.e. a list of all previously visited minima that the system should (in order to maximize the efficiency of sampling the energy landscape) be discouraged from visiting again. All visited minima are kept in memory. Whether or not the current minimum has been visited before feeds back to the temperature in the molecular dynamics runs. This allows the algorithm to eventually escape even deep energy funnels. A problem with this approach is that the number of steps to the global minimum can vary (possibly by orders of magnitude) depending on the starting point. Starting from a good initial guess of the structure (good in the sense of close by in search space), this may be the most efficient method available. Starting from a random structure, it can become very expensive. This algorithm has been applied, e.g., to Lennard-Jones clusters and a silicon crystal with fixed cell parameters at two system sizes (64 and 216 atoms/cell) [8].

4. Metadynamics

Another very interesting method is metadynamics [9–11]. Its input includes a starting structure and the relevant order parameters (usually, the lattice vectors matrix); search is performed in the order parameter space. At every step, a Gaussian is added to the free energy surface at the current position; this discourages the system from visiting the same state again and therefore acts as a tabu. The next step is taken in the direction of the steepest descent of the modified free energy. After sufficiently many steps, the free energy well is filled with Gaussians and a transition via the lowest (with very high probability) saddle-point occurs. Not only does this approach systematically search for the lowest saddle-point, the path taken in the run directly gives an idealized transition path. However, metadynamics cannot perform a global search without a good starting point. The performance of metadynamics depends on the choice of Gaussian height and width parameters (their optimal choice is discussed in [10]); it was found [11] to be advantageous to use anisotropic Gaussians adapted to the shape of the landscape. In the usual case of a 6D order parameter (lattice vectors matrix) metadynamics possesses two interesting properties: (i) the dimensionality of the problem does not depend on the system size, which renders relatively large systems (up to several hundred atoms in the (super)cell) affordable—however, such a simplification of the problem decreases the success rate, especially for large systems (where the coupling between the cell and internal degrees of freedom is weak) so that the method often produces only amorphous structures, (ii) the use of cell vectors as order parameters allows one to interpret predicted transition paths as likely mechanisms of plastic deformation. Metadynamics has successfully reproduced the pressure-induced transformation in Si from the diamond to the simple hexagonal structure [9], found all known structures of benzene [16], established mechanisms of sluggish transitions in silica [11, 18], predicted a family of low-enthalpy polymorphs of MgSiO_3 at high pressure and an unexpected mechanism of plastic deformation of MgSiO_3 post-perovskite, the dominant mineral of the Earth's core–mantle boundary layer [17].

5. Evolutionary algorithm USPEX

Evolutionary algorithms are well known in computational science and, due to their attractiveness for global optimization problems, have been applied to the prediction of structures of crystals [19–22], colloids [23] and clusters [24]. The algorithm developed by Deaven and Ho [24] is, in our view, particularly interesting as some of its features (real-space representation of structures, local optimization and spatial heredity) are similar to our method. Their algorithm has successfully produced the structure of the C_{60} buckminsterfullerene, but to our knowledge has not been tested on hetero-atomic clusters, nor adapted to crystals. The algorithm of Bush and Woodley [19–21] was originally developed for crystals and successfully produced a starting model for solving the structure of Li_3RuO_4 [19]. However, systematic tests [20, 21]

showed that this algorithm very often fails even for rather simple systems with ~ 10 atoms/cell. Other drawbacks are that this algorithm requires experimental lattice parameters and simulations are very expensive, unless a cheap and crude heuristic expression is used for fitness. Unlike the Deaven–Ho algorithm and USPEX, in this method structures are represented by binary ‘0/1’ strings and there is no local optimization.

Our algorithm USPEX takes into account previous experiences with evolutionary algorithms, both successes and failures. Structures are represented by floating-point numbers, with fractional coordinates for the atoms and lattice vectors matrix describing the periodicity of the structure. USPEX operates with populations of structures. From a population, parent structures are selected. The selection probability of each structure is a function of its fitness rank. A new candidate structure is produced from parent structures using one of three operators: (i) Heredity combines spatially coherent slabs (in terms of fractional coordinates) of two parent structures, while the lattice vectors matrices are weighted averages of the two parent lattice vectors matrices. (ii) Permutation (as in [20, 21]) swaps chemical identities in randomly selected pairs of atoms. (iii) Mutation distorts the cell shape. All generated candidate structures are tested against three constraints—first, all interatomic distances must be above the specified minimal values; second, cell angles must be between 60° and 120° ; third, all cell lengths must be larger than a specified value (e.g. diameter of the largest atom). These constraints help to ensure stability of energy calculations and local optimization, and remove only redundant and infeasible regions of configuration space—thus the search is physically unconstrained. If in violation of these constraints, the candidate structure is discarded; otherwise, it is locally optimized (relaxed). The locally optimal structure is recorded (and used for producing structures of the next generation); the negative of its energy (or any other relevant thermodynamic potential) is used as fitness. Local optimization and energy calculations are done by external codes (currently, USPEX is interfaced with VASP [25], SIESTA [26] and GULP [27]). Once a sufficient number of new structures are produced, a new population is made of (one or more) lowest-enthalpy structures from the previous population and the new structures produced using variation operators. The above procedure is repeated in a loop.

The first generation usually consists of locally optimised random structures, but it is possible to include user-specified structures. If lattice parameters are known, runs can be done in the fixed cell, but this is not required and in most cases we do simulations with variable cell shape.

One of the reasons for the success of USPEX is that local arrangements of atoms (spatially coherent pieces of structures) are partly preserved and combined. This respects the predominant short-ranged interactions in crystals and exploits information from the current population. Combination, within a single heredity-produced structure, of slabs from two parent structures is reminiscent of the ‘two-phase’ simulation method used for calculations of melting temperatures. For large systems it may be advantageous to combine slabs of several

structures. Note, however, that for systems with very few atoms (or molecules) in the unit cell heredity becomes obsolete (in the limit of 1 atom/unit cell it is completely useless). The method has been described in [14, 15]; two new aspects (specific permutation and cell transformations) are described below.

The permutation operator now allows specific swaps. More precisely, the user can specify which types of atoms are to be interchanged. This is useful for systems with a large range in degree of chemical similarity between different atom types (e.g. in aluminosilicates Al–Si exchange is probably more relevant than Si–O exchange). The original method [14, 15] swapped all atomic types at random—this allows a more global search and this capability is retained as an option.

Furthermore, we have implemented a cell transformation, which optionally is applied to every candidate structure. The rationale for such a transformation is to further remove redundancies in the search space. The constraint that the cell angles be between 60° and 120° (in simulations we use a slightly wider interval) does not remove all redundancies and still allows inconvenient redundant cell shapes to be produced (e.g. cells with $\alpha = \beta = \gamma \sim 120^\circ$ are practically flat). Certain advantages would be gained by transforming such cells to a cell shape with shorter cell vectors. Such a transformation can be done if there is at least one lattice vector whose projection onto any other cell vector (or onto the diagonal vector of the opposite face of the unit cell) is greater (by absolute value) than half the length of that vector, i.e. for pairs \mathbf{a} and \mathbf{b} , or \mathbf{c} and $(\mathbf{a} + \mathbf{b})$ these criteria are

$$\left| \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|} \right| > \frac{|\mathbf{b}|}{2} \quad (2a)$$

$$\left| \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|} \right| > \frac{|\mathbf{a}|}{2} \quad (2b)$$

$$\left| \frac{\mathbf{c} \cdot (\mathbf{a} + \mathbf{b})}{|\mathbf{c}|} \right| > \frac{|\mathbf{c}|}{2} \quad (2c)$$

$$\left| \frac{\mathbf{c} \cdot (\mathbf{a} + \mathbf{b})}{|\mathbf{a} + \mathbf{b}|} \right| > \frac{|\mathbf{a} + \mathbf{b}|}{2} \quad (2d)$$

e.g. for the criterion (2a) the new vector \mathbf{a}_{new} is given by

$$\mathbf{a}_{\text{new}} = \mathbf{a} - \text{ceil} \left(\left| \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|} \right| \right) (\text{sign}(\mathbf{a} \cdot \mathbf{b})) \mathbf{b}. \quad (3)$$

During this transformation atomic fractional coordinates are transformed so that the original and the transformed structures are identical (during the transformation Cartesian coordinates of the atoms remain invariant).

6. Some illustrations of USPEX

A number of test cases have been presented in [14, 15, 28]. The first application of USPEX to finding new stable structures was published in [13], where two high-pressure phases of CaCO_3 were found. Both belong to new structure types (i.e. hitherto unknown structure topologies), one with triangular CO_3^{2-} ions (stable at pressures in the range 42–137 GPa and confirmed

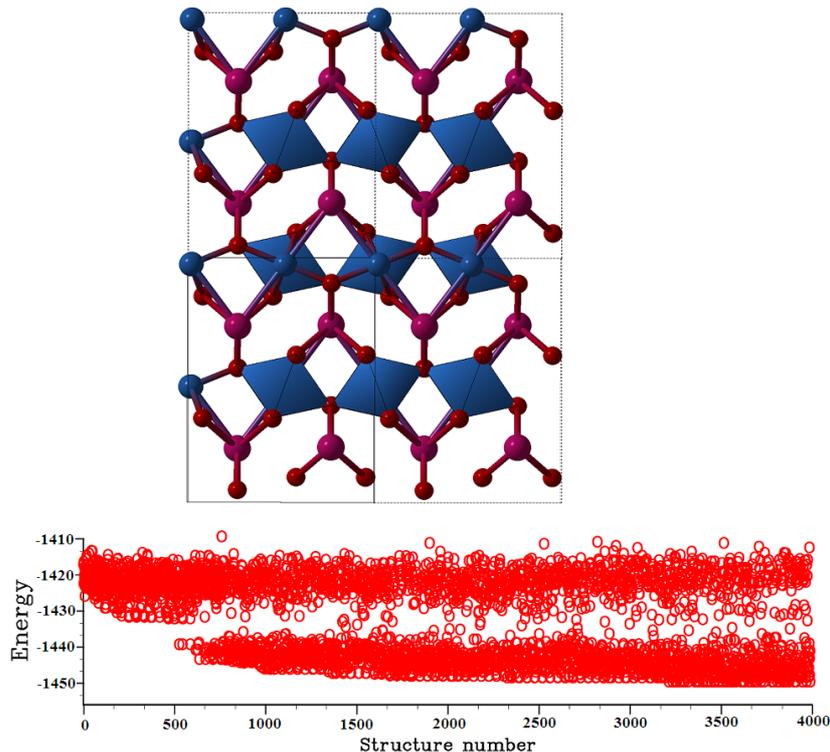


Figure 1. Evolutionary prediction of the structure of MgSiO_3 post-perovskite using the experimental cell parameters for an 80 atom supercell. Top: structure of post-perovskite (showing SiO_6 octahedra) obtained within ~ 3200 local optimizations. In this simulation, each generation consisted of 41 structures. Bottom: energies of structures along the evolutionary trajectory.

in the same paper [13]), and the other with chains of CO_4 tetrahedra (stable above 137 GPa and confirmed in [29]). Subsequently, we predicted new structures for high-pressure phases of oxygen [30], boron [31] and FeS [32]. Here, we would like to present one non-trivial test and two applications.

The test is the prediction of the structure of MgSiO_3 post-perovskite [33, 34] using a relatively large 80 atom supercell (with fixed supercell parameters) and an empirical potential [34] describing interatomic interactions within a partially ionic model. Local optimization and energy calculations were done using the GULP code [27]. Previously [28], we have shown that already in a 40 atom supercell this test is unfeasible using the simple random sampling (with local optimization): the correct structure was not produced even after 1.2×10^5 random attempts, but was found with 600–950 attempts of USPEX. With 80 atoms/cell the problem becomes much more complicated (one expects an exponential increase of complexity with increasing system size), but even in this case we correctly produced the post-perovskite structure in a reasonable number (~ 3200) of local optimizations—see figure 1. For an even larger test case, a Lennard-Jones crystal with 128 atoms in the (super)cell, we performed variable-cell structure search, which has correctly identified the hcp structure as the ground state within three generations (each consisting of only 10 structures). For larger Lennard-Jones systems (256 and 512 atoms/cell) we found the energetically very slightly less favourable fcc structure.

An interesting possibility has been proposed [35–37] that a new class of solids can exist, which are made of Al clusters—

ionic insulator $(\text{Al}_{13})\text{K}$ with the CsCl-type structure and icosahedral [35] or cuboctahedral [36] Al_{13}^- ions, and molecular Al_{12}C with icosahedral [37] Al_{12}C clusters bound together by weak intermolecular interactions. The rationale for this idea comes from the observation of icosahedral nanoclusters formed by boron and aluminium atoms and from the jellium model, which predicts particularly high stability for clusters with 40 valence electrons and suggests that the icosahedral 40-electron Al_{12}C and Al_{13}^- groups may have closed-shell electronic configurations and produce insulating structures. If correct, this idea could lead to the synthesis of technologically very interesting materials. Here, we explore these systems (with the only exception that for Al_{12}C we study only the 13 atom system, rather than the previously [37] suggested 104 atom cell) using USPEX. These calculations were performed within the generalized gradient approximation [38] and the PAW method [39, 40], using VASP [25] for local optimization and total energy calculations. In both cases we find very different structures, which do not contain any well defined clusters.

For Al_{13}K we found that when full cell relaxation is performed the icosahedral structure (figure 2(a)) spontaneously transforms into the cuboctahedral one (figure 2(b)). This is consistent with the conclusion [36] that the cuboctahedral structure is energetically much more favourable. The cuboctahedral clusters can be thought of as fragments of the fcc structure; however, these fragments are not isolated in Al_{13}K , but form bonds with each other (Al–Al distances within the cluster are 2.78, and 2.72 Å between the clusters), which makes this structure metallic and indicates limitations of the jellium

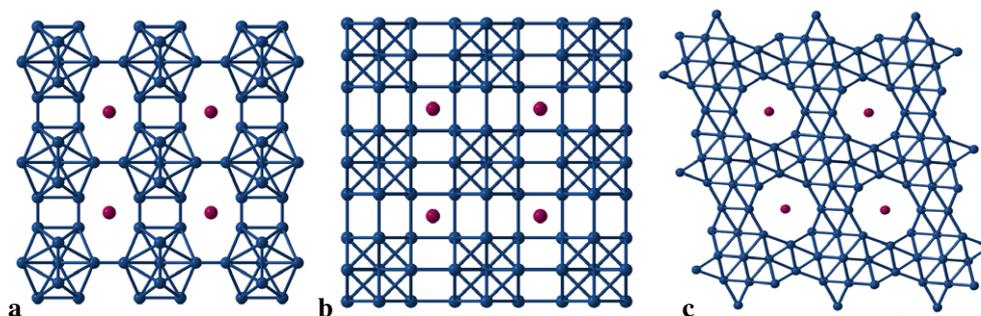


Figure 2. Structures of Al_{13}K with 14 atoms/cell: (a) with icosahedral Al_{13} clusters; (b) with cuboctahedral Al_{13} clusters; (c) structure produced by USPEX.

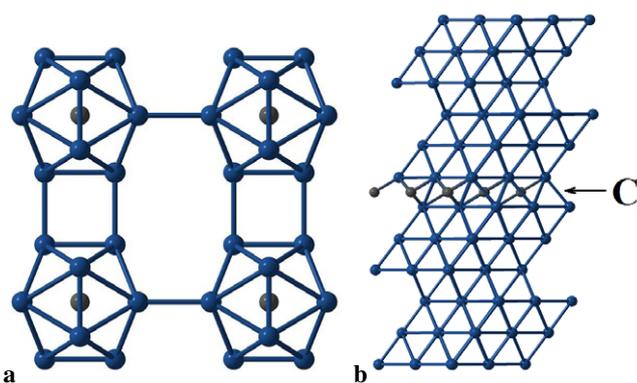


Figure 3. Structures of Al_{12}C with 13 atoms/cell: (a) with icosahedral clusters; (b) structure produces by USPEX.

model for real systems. The structure found with USPEX (figure 2(c)) is much lower in energy (by 78 meV/atom), has no clusters and shows a great perturbation that K atoms exert on the structure of Al. So far, no Al–K compounds are known to be stable with respect to decomposition into the elements—indeed, we find that even the USPEX-produced structure is 72 meV/atom higher in energy than the mechanical mixture of pure Al and K.

For Al_{12}C , the optimized icosahedral structure (figure 3(a)) also has bonds between the clusters: Al–Al distances are 2.67–2.72 Å within the cluster and 2.65 Å between the clusters. Thus, the expectation of a molecular structure made on the basis of the jellium model is again incorrect. Amazingly, however, this structure is a semiconductor (DFT band gap is 0.3 eV), but a very unstable one: with USPEX, we found a much more stable (by 557 meV/atom) structure shown in figure 3(b). This structure is nothing other than a close-packed structure, very similar to the fcc structure of pure Al, but with stacking faults at which C atoms occupy the octahedral voids in the aluminium close packing. This implies that C impurities in Al may be correlated with the presence of stacking faults. The same was found in our evolutionary simulations of Al_{16}C_4 . We conclude that the icosahedral coordination (coordination number 12) is probably too high for C atoms; the octahedral coordination is more stable. Indeed, C atoms in the stable compound Al_4C_3 occupy octahedrally and tetrahedrally coordinated sites. The USPEX-produced structure of Al_{12}C is metallic and only

39 meV/atom less stable than the mechanical mixture of pure aluminium and graphite.

The examples of Al_{12}C and Al_{13}K illustrate that evolutionary crystal structure prediction can result in lower energy structures than conventional chemical intuition. It allows thermodynamically feasible and potentially interesting materials to be predicted with a high degree of certainty. Results of such simulations can significantly improve our intuition and guide materials design in promising directions.

7. Conclusions

We have discussed a few different approaches (though a discussion of all existing or possible methods would not be practical within the limits of this article), including our evolutionary algorithm USPEX. They clearly show that, after many years of research in the area, methods have emerged that can deal with the problem of crystal structure prediction. One should keep in mind that the right approach depends on the system (its size and energy landscape) and the available information (partial structural information, symmetry etc). We are suggesting USPEX as the method of choice for crystal structure prediction of systems with ~6–40 atoms/cell, where no information (or just the lattice parameters) is available. Above 40 atoms/cell runs become expensive (although still feasible), due to the ‘curse of dimensionality’ eventually necessitating the use of other ideas within USPEX or another approach. Fortunately, many (or most) systems of practical or fundamental interest fall in the size range accessible to USPEX, opening a wide field of potential applications of this method in computational materials design.

Acknowledgments

Calculations were performed at the Joint Russian Supercomputer Centre (Russian Academy of Sciences, Moscow), ETH Zurich and CSCS (Manno). We thank D Vicente (Barcelona Supercomputer Centre), T Racic and G Sigut (ETH Zurich) for computational assistance, and M Valle for developing the STM3 library for the visualization of our results. This work is supported by the Swiss National Science Foundation under grant 200021-111847/1.

References

- [1] Maddox J 1988 Crystals from first principles *Nature* **335** 201
- [2] Curtarolo S, Morgan D, Persson K, Rodgers J and Ceder G 2003 Predicting crystal structures with data mining of quantum calculations *Phys. Rev. Lett.* **91** 135503
- [3] Urusov V S, Dubrovinskaya N A and Dubrovinsky L S 1990 *Generation of Likely Crystal Structures of Minerals* (Moscow: Moscow State University Press)
- [4] Deem M W and Newsam J M 1989 Determination of 4-connected framework crystal structures by simulated annealing *Nature* **342** 260–2
- [5] Pannetier J, Bassasalsina J, Rodriguez-Carvajal J and Caignaert V 1990 Prediction of crystal structures from crystal chemistry rules by simulated annealing *Nature* **346** 343–5
- [6] Boisen M B, Gibbs G V and Bukowinski M S T 1994 Framework silica structures generated using simulated annealing with a potential energy function based on an $\text{H}_6\text{Si}_2\text{O}_7$ molecule *Phys. Chem. Minerals* **21** 269–84
- [7] Schön J C and Jansen M 1996 First step towards planning of syntheses in solid-state chemistry: determination of promising structure candidates by global optimization *Angew. Chem. Int. Edn* **35** 1287–304
- [8] Gödecker S 2004 Minima hopping: an efficient search method for the global minimum of the potential energy surface of complex molecular systems *J. Chem. Phys.* **120** 9911–7
- [9] Martonak R, Laio A and Parrinello M 2003 Predicting crystal structures: the Parrinello–Rahman method revisited *Phys. Rev. Lett.* **90** 075503
- [10] Martonak R, Laio A, Bernasconi M, Ceriani C, Raiteri P, Zipoli F and Parrinello M 2005 Simulation of structural phase transitions by metadynamics *Z. Kristallogr.* **220** 489–98
- [11] Martoňák R, Donadio D, Oganov A R and Parrinello M 2006 Crystal structure transformations in SiO_2 from classical and *ab initio* metadynamics *Nat. Mater.* **5** 623–6
- [12] Pickard C J and Needs R J 2006 High-pressure phases of silane *Phys. Rev. Lett.* **97** 045504
- [13] Oganov A R, Glass C W and Ono S 2006 High-pressure phases of CaCO_3 : crystal structure prediction and experiment *Earth Planet. Sci. Lett.* **241** 95–103
- [14] Oganov A R and Glass C W 2006 Crystal structure prediction using *ab initio* evolutionary techniques: principles and applications *J. Chem. Phys.* **124** 244704
- [15] Glass C W, Oganov A R and Hansen N 2006 USPEX—evolutionary crystal structure prediction *Comput. Phys. Commun.* **175** 713–20
- [16] Raiteri P, Martonak R and Parrinello M 2005 Exploring polymorphism: the case of benzene *Angew. Chem. Int. Edn* **44** 3769–73
- [17] Oganov A R, Martoňák R, Laio A, Raiteri P and Parrinello M 2005 Anisotropy of Earth's D'' layer and stacking faults in the MgSiO_3 post-perovskite phase *Nature* **438** 1142–4
- [18] Martoňák R, Donadio D, Oganov A R and Parrinello M 2007 4- to 6-coordinated silica: transformation pathways from metadynamics *Phys. Rev. B* **76** 014120
- [19] Bush T S, Catlow C R A and Battle P D 1995 Evolutionary programming techniques for predicting inorganic crystal structures *J. Mater. Chem.* **5** 1269–72
- [20] Woodley S M 2004 Prediction of crystal structures using evolutionary algorithms and related techniques *Struct. Bonding* **110** 95–132
- [21] Woodley S M, Battle P D, Gale J D and Catlow C R A 1999 The prediction of inorganic crystal structures using a genetic algorithm and energy minimization *Phys. Chem. Chem. Phys.* **1** 2535–42
- [22] Bazterra V E, Ferraro M B and Facelli J C 2002 Modified genetic algorithm to model crystal structures. I. Benzene, naphthalene and anthracene *J. Chem. Phys.* **116** 5984–91
- [23] Gottwald D, Kahl G and Likos C N 2005 Predicting equilibrium structures in freezing processes *J. Chem. Phys.* **122** 204503
- [24] Deaven D M and Ho K M 1995 Molecular geometry optimization with a genetic algorithm *Phys. Rev. Lett.* **75** 288–91
- [25] Kresse G and Furthmüller J 1996 Efficient iterative schemes for *ab initio* total-energy calculations using a plane wave basis set *Phys. Rev. B* **54** 11169–86
- [26] Soler J M, Artacho E, Gale J D, Garcia A, Junquera J, Ordejon P and Sanchez-Portal D 2002 The SIESTA method for *ab initio* order-N materials simulation *J. Phys.: Condens. Matter* **14** 2745–79
- [27] Gale J D 2005 GULP: capabilities and prospects *Z. Kristallogr.* **220** 552–4
- [28] Martonak R, Oganov A R and Glass C W 2007 Crystal structure prediction and simulations of structural transformations: metadynamics and evolutionary algorithms *Phase Transit.* **80** 277–98
- [29] Ono S, Kikegawa T and Ohishi Y 2007 High-pressure transition of CaCO_3 *Am. Mineral.* **92** 1246–9
- [30] Ma Y-M, Oganov A R and Glass C W 2007 Structure of the metallic ζ -phase of oxygen and isosymmetric nature of the ε - ζ phase transition: *ab initio* simulations *Phys. Rev. B* **76** 064101
- [31] Oganov A R, Glass C W, Ma Y-Z, Ma Y-M and Chen J 2007 Ionic high-pressure form of elemental boron *Nature* submitted
- [32] Ono S, Oganov A R, Brodholt J P, Vocadlo L, Wood I G, Glass C W, Côté A S and Price G D 2007 High-pressure phase transformations of FeS : novel phases at conditions of planetary cores *Earth Planet. Sci. Lett.* submitted
- [33] Oganov A R and Ono S 2004 Theoretical and experimental evidence for a post-perovskite phase of MgSiO_3 in Earth's D'' layer *Nature* **430** 445–8
- [34] Murakami M, Hirose K, Kawamura K, Sata N and Ohishi Y 2004 Post-perovskite phase transition in MgSiO_3 *Science* **307** 855–8
- [35] Khanna S N and Jena P 1994 Designing ionic solids from metallic clusters *Chem. Phys. Lett.* **219** 479–83
- [36] Liu F, Mostoller M, Kaplan T, Khanna S N and Jena P 1996 Evidence for a new class of solids. First-principles study of $\text{K}(\text{Al})_{13}$ *Chem. Phys. Lett.* **248** 213–7
- [37] Gong X G 1997 Structure and stability of cluster-assembled solid $\text{Al}_{12}\text{C}(\text{Si})$: a first-principles study *Phys. Rev. B* **56** 1091–4
- [38] Perdew J P, Burke K and Ernzerhof M 1996 Generalized gradient approximation made simple *Phys. Rev. Lett.* **77** 3865–8
- [39] Blöchl P E 1994 Projector augmented-wave method *Phys. Rev. B* **50** 17953–79
- [40] Kresse G and Joubert D 1999 From ultrasoft pseudopotentials to the projector augmented-wave method *Phys. Rev. B* **59** 1758–75