Finding the lowest-energy crystal structure starting from randomly selected lattice vectors and atomic positions: first-principles evolutionary study of the Au–Pd, Cd–Pt, Al–Sc, Cu–Pd, Pd–Ti, and Ir–N binary systems

2008 J. Phys.: Condens. Matter 20 295212

(http://iopscience.iop.org/0953-8984/20/29/295212)

IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **20** (2008) 295212 (11pp) doi:10.1088/0953-8984/20/29/295212

Finding the lowest-energy crystal structure starting from randomly selected lattice vectors and atomic positions: first-principles evolutionary study of the Au–Pd, Cd–Pt, Al–Sc, Cu–Pd, Pd–Ti, and Ir–N binary systems

Giancarlo Trimarchi and Alex Zunger

National Renewable Energy Laboratory, Golden, CO 80401, USA

E-mail: Alex_Zunger@nrel.gov

Received 13 February 2008, in final form 27 April 2008
Published 1 July 2008
Online at stacks.iop.org/JPhysCM/20/295212

Abstract

Two types of global space-group optimization (GSGO) problems can be recognized in binary metallic alloys A_qB_{1-q}: (i) *configuration search* problems, where the underlying crystal lattice is known and the aim is finding the most favorable decoration of the lattice by A and B atoms and (ii) *lattice-type search* problems, where neither the lattice type nor the decorations are given and the aim is finding energetically favorable lattice vectors and atomic occupations. Here, we address the second, lattice-type search problem in binary A_qB_{1-q} metallic alloys, where the constituent solids A and B have different lattice types. We tackle this GSGO problem using an evolutionary algorithm, where a set of crystal structures with randomly selected lattice vectors and site occupations is evolved through a sequence of generations in which a given number of structures of highest LDA energy are replaced by new ones obtained by the generational operations of mutation or mating. Each new structure is locally relaxed to the nearest total-energy minimum by using the *ab initio* atomic forces and stresses. We applied this first-principles evolutionary GSGO scheme to metallic alloy systems where the nature of the intermediate A–B compounds is difficult to guess either because pure A and pure B have different lattice types and the (i) intermediate compound has the structure of *one* end-point (Al_3Sc, $AlSc_3$, $CdPt_3$), or (ii) *none* of them (CuPd, AlSc), or (iii) when the intermediate compound has lattice sites belonging simultaneously to a few types (fcc, bcc) ($PdTi_3$). The method found the correct structures, $L1_2$ type for Al_3Sc, $D0_{19}$ type for $AlSc_3$, '$CdPt_3$' type for $CdPt_3$, B2 type for CuPd and AlSc, and A15 type for $PdTi_3$. However, in such stochastic methods, success is not guaranteed, since many independently started evolutionary sequences produce at the end different final structures: one has to select the lowest-energy result from a set of such independently started sequences. Interestingly, we also predict a hitherto unknown ($P2/m$) structure of the hard compound IrN_2 with energy lower than all previous predictions.

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

1. Introduction

The close relationship between the atomic arrangement in a solid system and its electronic properties [1, 2], makes the theoretical determination of its crystal structure an important step towards predicting its physical properties. In predicting crystal structures there are in general two classes of structural degrees of freedom to consider: the decoration of each of the N lattice sites by atoms of type A, B, etc (requiring a 'configuration search'), and the selection of lattice type (requiring a 'lattice-vector search'). There are thus two corresponding search problems:

The *configuration search* problem corresponds to the classic generalized Ising problem on a given Bravais lattice [3, 4]. It is an 'NP-hard' problem [5] in that the computational labor to solve it increases exponentially with the number N of lattice sites. In the context of first-principles calculations, the energy versus configuration relationship could be calculated directly 'on the fly' (as in the Born–Oppenheimer *ab initio* dynamics [6]) or addressed via a 'cluster-expansion' method [7, 8] (i.e. mapping of first-principles total energies of a few configurations onto a multi-site generalized Ising Hamiltonian [8]). The configuration search of this cluster-expansion Hamiltonian can be done either via exhaustive enumeration [9] (often limited to $N \lesssim 20$ sites), or via genetic-algorithm approaches [10–12]. The combination of first-principles cluster expansion with such efficient configurational search methods has been applied in the past to numerous systems with *given underlying crystal lattice*, revealing in many cases interesting and unsuspected new ground-state configurations (see, e.g., [13–18]). Recent systematic searches of the cluster-expansion Hamiltonian have demonstrated the NP-hard character of this configuration search problem and offered improved (reciprocal-space) genetic algorithms [11] as well as Lamarckian 'virtual-atom' approaches [12] to efficiently address such a decoration optimization problem.

The *lattice-type search* problem has been addressed previously, e.g. by calculating the total-energy versus volume curves of a handful of candidate structure types, and selecting the lowest one from this limited group [19–21], as demonstrated by Yin and Cohen for silicon [21]. Variable-cell molecular dynamics [22, 23] extended by the meta-dynamics algorithm [24, 25] has enabled simulating changes of the cell shape and lattice type characterized by large energy barriers. Yet, the predictive power of the method is limited by the small number of neighboring local minima of the potential energy surface which it identifies and by the slowness of transforming a configuration (e.g. swapping a pair of non-neighboring atoms) via a number of concerted atom swaps.

In general, one can recognize material systems whose crystal structure search is dominated either by the decoration (i.e. configuration) search, or by the lattice-type search. Systems for which structure determination is largely *decoration-search limited* correspond to cases A_pB_q where A-on-B cross-substitutions ('anti-site' defect) represent the difficult-to-resolve *low-energy* (structural) excitations, whereas different lattice types are well separated energetically and are thus more easily resolved *high-energy* structural excitations. Examples include alloys whose constituent elemental solids A and B are chemically/electronically similar, and have the same underlying lattice as the ensuing intermediate compounds A_pB_q, e.g. the fcc-based Cu–Au [26, 27] and Au–Pd [16] or the bcc Mo–Ta [15, 28] intermetallics. Such decoration-limited searches are treated efficiently via cluster-expansion techniques. On the other hand, systems whose structure determination involves largely a *lattice-type-limited* search correspond to cases where different decorations are easily resolved high-energy excitations, but the choice of the lattice type is less obvious. Examples include covalent 'octet' compounds (III–V GaAs or II–VI ZnSe), where A-on-B (anion-on-cation or cation-on-anion) cross-substitution represents easily resolved octet-violating [1] high-energy structural excitations. Such cases of lattice-type-dominated structure search for octet compounds, originally performed by selecting among the total-energy versus volume curves for a handful of candidate structure types [21], were recently replaced by global space-group optimization (GSGO) approaches (see, e.g., [29–32]), in which one starts from randomly selected lattice vectors and randomly selected site occupations and determines the structure via an evolutionary algorithm progression. Applying this evolutionary algorithm selection of structures to Si [32] and GaAs [32] performed considerably better (i.e. required fewer evaluations of the relaxed total energy) than iterative random generation of crystal structures.

In the present paper we apply the procedure for GSGO described in [32] both to Au–Pd, representing a decoration-limited search problem, and to a set of metallic systems representing a lattice-type-limited search that are prone to adopt non-intuitive lattice types, and are thus difficult to treat by cluster expansion. We next explain the types of binary systems belonging to the lattice-type-limited class that we have selected here.

2. Systems selected for lattice-type-limited searches

One can recognize a few types of material systems for which it is difficult to guess the lattice type. A first such group of systems is cases where the elemental constituent solids A and B have different lattice types and the A_pB_q compounds have an underlying lattice which corresponds to that of one of the two constituents. For example, whereas solid Al is fcc and solid Sc is hcp, the compound Al_3Sc has the fcc-based $L1_2$ structure, whereas $AlSc_3$ has been recently predicted [33] to have the hexagonal-based $D0_{19}$ structure. Similarly, whereas solid Pt is fcc and solid Cd is hcp, the intermetallic $CdPt_3$ has been predicted [33] to have the fcc-based '$CdPt_3$' [34] structure. However, such predictions [33] were based on selecting the lattice type from a fixed list of 176 possible types. In this paper we apply our approach to GSGO, starting from randomly selected lattice vectors and randomly selected decoration to Al–Sc and Cd–Pt, in an attempt to determine the crystal structure in an unbiased way.

A second group of systems with difficult to guess lattice type is cases where the elemental constituent solids A and B have different lattice types and the intermetallic A_pB_q compound has a lattice type which differs altogether from those

of the end-point constituents. For example, whereas both Cu and Pd are fcc, their equimolar intermetallic Cu–Pd has the bcc-based B2 structure [33, 35, 36]. Similarly, whereas solid Al is fcc and Sc is hcp, their equimolar intermetallic Al–Sc has the bcc-based B2 structure [33, 35, 36]. Both in the first and second types of systems the compounds mentioned are based on decoration of a Bravais lattice or of a hcp lattice by A atoms and B atoms. Thus, such systems can be described in principle by a cluster-expansion approach, however for a few lattice types, one at a time (fcc for Al_3Sc and $CdPt_3$; hcp for $AlSc_3$; bcc for CuPd and AlSc).

A third group of systems with difficult to guess lattice type is the intermetallic compounds A_pB_q whose structure cannot be described as decoration of a simple parent Bravais lattice. Such is the A15-type structure of $PdTi_3$ [33], where the Pd atoms are located on the sites of a bcc sublattice, while the Ti atoms form pairs located at the faces of the conventional cubic cell with Ti pairs located on opposite faces of the cell parallel to the Cartesian axes.

3. Evolutionary procedure for predicting crystal structures

3.1. General scheme

Our approach to the GSGO problem of solids has been described previously [32]. It builds upon developments of recent works [29–31, 37] but represents an independent implementation and code. While the present paper emphasizes applications, for the sake of completeness, we give a brief account of the algorithm, depicted in figure 1 as a flowchart diagram. The search method uses a sequence of *ab initio* evaluations of the total energy of locally relaxed trial structures so as to seek the optimal lattice decoration and lattice vectors via a genetic-algorithm selection used widely in the past [29–31].

(1) A population of $N_{\rm pop}$ candidate structures $\mathcal{S}_i$, where $i = 1, \ldots, N_{\rm pop}$, is evolved through a sequence of generations. A structure is specified by the lattice vectors and the internal atomic positions. The structures of the initial population are randomly generated. At each new generation, the $N_{\rm rep}$ highest total-energy structures out of $N_{\rm pop}$ are replaced by new structures which are produced from the structures of the current population by performing the operations of mating and mutation.

(2) Since it is not obvious how to mate two structures which differ in lattice type, a similarity transformation [30, 31, 37] is first applied to the structures before mating. This maps the atomic positions onto fractional coordinates and the cell shape onto a cubic cell of unitary lattice parameter. After the similarity transformation,

(3) two structures are mated directly in real space via a cut-and-splice operation [29–32]. In this procedure, first one defines a plane perpendicular to one of the Cartesian axes which cuts the cubic unit cells of both parent structures into two slabs. Newly produced structures are immediately discarded and a new mating operation is done if they contain the wrong number of atoms or their composition is different than that

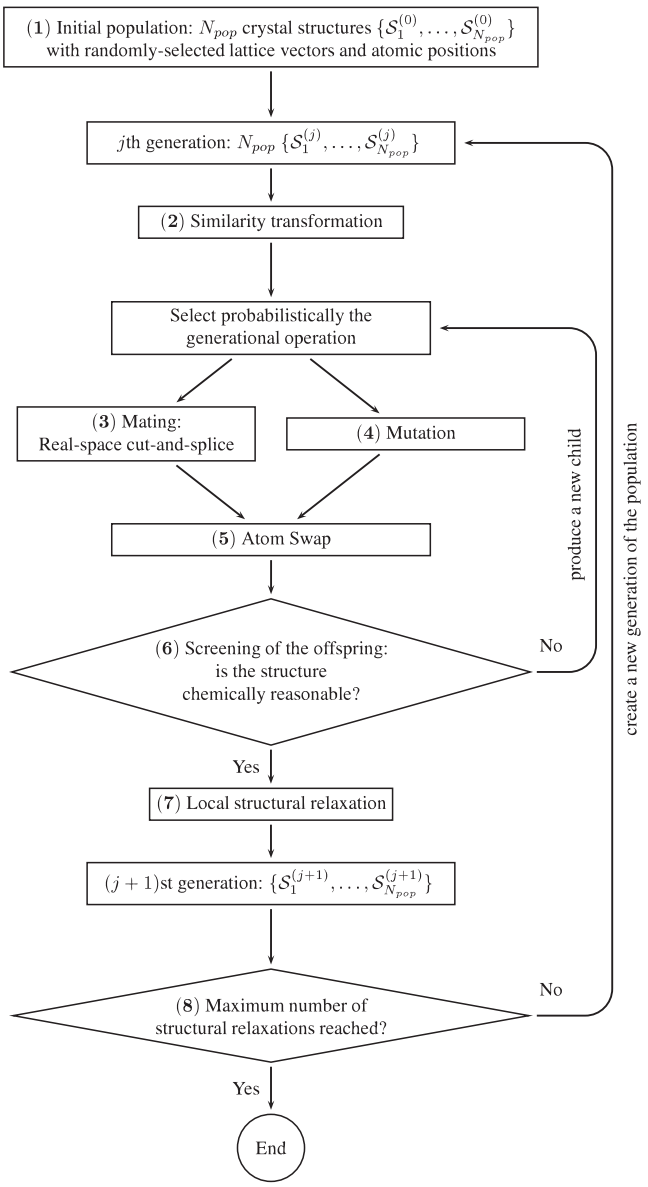

Figure 1. Flow chart of the evolutionary GSGO algorithm.

fixed for the compound which is being optimized. Then, corresponding slabs in the two parents are swapped, producing two new child structures. One of the two new structures is chosen according to a coin toss to be included in the new generation of the population.

(4) The mutation of a crystal structure consists in subjecting each atom to a displacement defined by a randomly selected vector $\boldsymbol{\delta}_i$ of maximal length $\delta^{\rm mut}$.

(5) The crystal structure obtained through the operation of mating or mutation is subjected to a sequence of swaps of pairs of unlike atoms in order to alter its atomic configuration. Such an operation allows us to explore the set of possible lattice site occupations for the newly generated crystal structures. After this step the child structure is assigned a triad of unit-cell vectors.

(6) In order to rapidly converge the evolutionary algorithm to chemically reasonable crystal structures, both at the stage of producing the initial population and of generating the $(j+1)$th

population from the jth one, a structure is discarded if the atoms are too close to each other, the unit-cell angles are too small, or the lattice vectors are shorter than the typical bond length in the system of interest.

(7) After N_{rep} chemically acceptable child crystal structures are obtained, they are fully relaxed to the nearest local total-energy minimum and added to the population.

(8) The whole procedure is repeated iteratively, evolving the population through a series of new generations until a given stopping criterion is satisfied. In the following, we provide details of the practice we follow in performing a GSGO search and analyzing its history and results.

3.2. Accuracy of the ab initio structural relaxation

The crystal structure found by GSGO depends on the accuracy with which the underlying *ab initio* energy is evaluated for a given structure. This accuracy decides if the algorithm is able to resolve energetically similar structures. Thus, each GSGO run is specified here also by the parameters used in the *ab initio* electronic structure calculations, which are done via the VASP code [38, 39]. Converging the ionic forces and the stress tensor with respect to the kinetic energy cutoff of the plane-wave basis and the resolution of the k-point mesh is crucial to obtain accurate lattice vectors and atomic positions from the local structural optimization. A sufficiently high-energy cutoff is also needed for it to be possible to compare the total energies of the crystal structures of the evolving population, that are described by supercells which, although of similar volume per atom, have different shapes. The parameters of the VASP calculations performed for each of the systems studied in this paper are listed in appendix A.

3.3. Stopping criteria for the evolutionary process and success rate

Defining the stopping criterion for the optimization of a function defined on a high-dimensional space and having many local minima is in general not obvious. This holds also for the potential energy surface of a solid that is a function of $3N + 3$ variables, where N is the number of atoms per unit cell. Indeed, the search for the global minimum proceeds visiting local minima of the potential energy surface and there is no way to tell *a priori* whether the local minimum of lowest energy among those that have been visited is also the global one. Indeed, the evolutionary GSGO algorithm used here is stochastic in nature, so it is unlikely that each independently started evolutionary sequence will retrieve the global minimum within a given number M of (relaxed) total-energy evaluations (where a single relaxed total-energy evaluation may include numerous relaxation steps). Rather, a number of restarts of the search from uncorrelated initial random populations are needed in order to effectively find the global minimum. Thus, success is not guaranteed. Indeed, the efficiency of the evolutionary GSGO scheme at finding the lowest-energy structure should be assessed statistically through a study of the average number $\overline{M}$ of (relaxed) total-energy evaluations needed to retrieve the global minimum with given confidence. Obviously, for $\overline{M}$ to be significant it should be calculated averaging over a sufficiently high number of independent runs. Such statistics is costly to obtain unless the total-energy functional is fast to calculate. This is the case when the functional is an already constructed cluster expansion. For the case of Au–Pd alloys, an accurate functional was constructed by Barabash *et al* [16] and used by d'Avezac and Zunger [11] to study the statistics for a pure configurational search. A supercell containing $(2\times2\times3)$, $(2\times2\times4)$, or $(2\times2\times6)$ sites required 580, 350, and 5600 (relaxed) total-energy evaluations to achieve in a real-space GA search the absolute lowest-energy structure with 95% confidence. Other, faster evolutionary algorithms, were also studied, and required fewer evaluations than real-space GA [11].

Using a plane-wave basis pseudopotential LDA energy functional is far more costly than a cluster-expansion functional, so such detailed statistics cannot be conveniently collected. Thus, here we performed just a few independently started GA sequences, each starting from *independent* random populations to get an idea on success rate. Each independent GSGO is evolved including at least 80 (relaxed) total-energy evaluations. A set of three independent runs of the above mentioned minimal length is hardly sufficient for distilling a meaningful statistics. However, from the GSGO calculations performed on binary systems following such a prescription, we obtain indications of the success rate and number of structure evaluations needed for the optimization of systems with moderate number of atoms per unit cell.

3.4. Crystallographic analysis

Once an evolutionary algorithm solution is obtained by fully relaxing an individual of the evolving population, we determine its space-group symmetry by using the ADDSYM module of the crystallographic analysis package PLATON [40–42]. ADDSYM is able to identify the space-group symmetry operations of a crystal structure even if the atoms show a small displacement ϵ from the high-symmetry Wyckoff positions, where usually $\epsilon < 0.02$ Å. Such small offsets of the atoms from ideal Wyckoff positions are to be expected in the crystal structures obtained during the evolutionary GSGO search relaxing structures created by the mating and mutation operations in which no symmetry constraint is imposed. To determine whether a crystal structure obtained by GSGO has fcc, bcc, or simple cubic parent lattice we identify the shells of nearest neighbors taking as the reference lattice site in turn all atoms in the unit cell. Then, the distance of the nth shell from the reference lattice site is compared with that of the nth nearest-neighbor shell in fcc, bcc, and simple cubic lattices of same volume per unit cell as the structure being analyzed: this comparison shows how far its underlying crystal lattice is from that of an fcc, bcc, and simple cubic lattice.

4. Results

4.1. The Au_8Pd_4 system: the appearance of numerous nearly degenerate configurations

4.1.1. Cluster expansion versus GSGO. The Au–Pd metallic alloy has elemental solid constituents Au and Pd

with the same fcc lattice type. The intermediate Au–Pd compounds also have fcc lattices like the constituents. Hence, predicting the most stable Au_pPd_q crystal structure is a configuration search problem that corresponds to finding the lowest-energy decoration of the given fcc lattice at this composition. The ground-state structures of the Au–Pd system have been investigated by Barabash *et al* [16] via an *ab initio* cluster-expansion (CE) approach, followed by an exhaustive enumeration [9] of all Au_pPd_q decorations of the fcc lattice by Au atoms and Pd atoms with up to 20 sites per unit cell. The fact that the lowest-energy structures of Au–Pd are known from such an exhaustive evaluation based on a CE with *ab initio* accuracy makes this binary alloy a good test system for the first-principles evolutionary GSGO. Thus, although such a configurational optimization is restricted to a fixed underlying lattice (as opposed to the evolutionary GSGO algorithm, which is not), it allows us to search for the ground-state structure scanning all concentrations $0 \leqslant x \leqslant 1$. Note that a 'ground-state structure' refers to first finding the lowest-energy configuration at each composition, and then removing from this list those configurations that are less stable than a linear combination of two structures with neighboring concentrations. For Au_8Pd_4, Barabash *et al* [16] found from the exhaustive enumeration of the LDA fitted CE energies two ground-state structures referred to as '4557' and '4905', having 12 atoms per unit cell each, and $C2/m$ and $Immm$ space-group symmetry, respectively, and whose LDA total energies differ by only 0.3 meV/atom.

Figure 2(a) shows the history plot of the present GSGO search for Au_8Pd_4 starting from random lattice vectors and random atomic positions, and figure 2(b) depicts the final structure found by GSGO after 28 structure evaluations. The crystallographic analysis shows that this structure has an fcc underlying lattice (see figure 2(c)) and the $C2/m$ space-group symmetry, like structure '4557' (see appendix B for description of the Au–Pd structures discussed here). Both the original CE and the current GSGO runs were performed with similarly large plane-wave basis sets and dense k-point meshes[1]. Thus, the GSGO search retrieved one of the degenerate lowest-energy structures predicted earlier by the *ab initio* cluster expansion [16] (provided that both methods use similar LDA convergence parameters).

4.1.2. LDA convergence parameters can affect the GSGO structure. When two or more configurations have similar energies, the convergence parameters deciding the precision of the underlying total-energy calculation can be very important. For example, our earlier [32] GSGO run of Au_8Pd_4 retrieved a structure with fcc lattice and $Cmcm$ space-group symmetry, called structure '4820', that was 2 meV/atom higher than the cluster-expansion structures '4557' and '4905'. The main difference between our present and earlier [32] GSGO runs of Au_8Pd_4 was in the parameters of the LDA Brillouin-zone sampling: in the present calculations we used k-point meshes that were as uniform as possible and of a given density,

[1] The difference in total energy between '4557' and '4905' is 0.3 meV/atom using the k-point mesh reported in appendix A, and 0.1 meV/atom using the k-point mesh used in the CE paper [16].

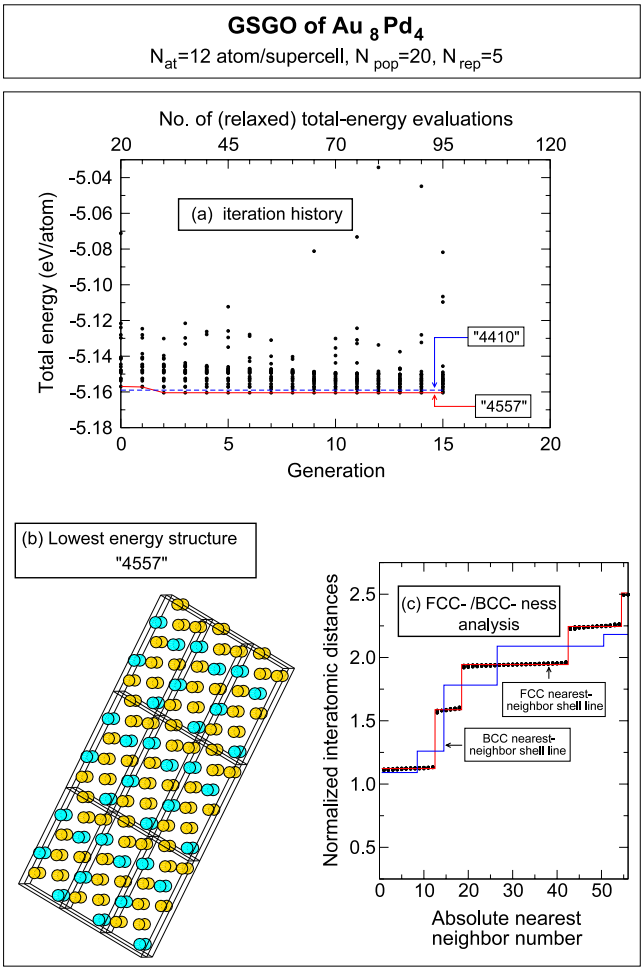

Figure 2. (a) History of the evolutionary optimization of Au_8Pd_4 ($N_{pop} = 20$ and $N_{rep} = 5$). (b) The lowest-energy structure that was found after 30 structure evaluations (space group No 12, $C2/m$). This structure has an fcc underlying lattice as shown by the fcc-/bcc-ness analysis of (c).

whereas in [32] we defined the k-point meshes for all structures assigning a fixed number of divisions of the unit vectors of the reciprocal cell. Not surprisingly, different LDA convergence parameters cannot distinguish structures whose energies are within the error range of the convergence parameters. We recalculated the LDA total energy of '4820' employing the computational parameters used in the present work, and we found that this structure is energetically degenerate within 0.3 < meV/atom with '4905' and '4557' found in the CE.

Recently, Oganov and Glass [43] repeated the calculation of Au_8Pd_4 using their independent implementation of the evolutionary GSGO (and similar convergence for the LDA total-energy calculations). They predicted a structure with fcc underlying lattice, and $P2_1/m$ space-group symmetry. We identified this structure as '4410' in our list of fcc structures [9, 16]. They thought that this structure represented a previously undetected lower-energy structure, because it had, in their calculation, an energy lower by 4 meV/atom than structure '4820' calculated with lower LDA convergence parameters. We found that structure '4410' is practically degenerate with '4905' and '4557', with a difference in total

energy of about 1 meV/atom. (We find that using a k-point grid with a resolution of $2\pi \times 0.02$ Å^{-1} or higher makes the structures '4557' and '4410' have an energy difference less than 1.0 meV/atom.) Thus, this is not a new lower-energy structure as alleged by Oganov and Glass [43].

We conclude from the present GSGO calculation, as well as from the earlier GSGO and CE studies, that the Au_8Pd_4 system has a few energetically quasi-degenerate ground-state structures. The presence of several quasi-degenerate structures is a well known phenomenon ('adaptive structures') observed in a few binary metallic systems [27], and poses a serious limitation to stochastic GA-based approaches, as independent restarts of the evolutionary GSGO of Au_8Pd_4 will often converge to one of these degenerate structures.

4.2. $CdPt_3$: the emergence of a rare fcc structure type

The experimental phase diagram [36] of Cd–Pt shows several stable compounds but only the crystal structures of the compounds at 50% and 75% Pt have been assigned, i.e. $L1_0$ and $L1_2$, respectively. Recently, Curtarolo *et al* [33] performed an *ab initio* data-mining study of Cd–Pt to predict its ground-state structures. Interestingly, for composition $CdPt_3$ the *ab initio* data-mining algorithm [33, 44] predicted a ground-state structure different from the usually expected $L1_2$. The structure predicted by data mining has orthorhombic $Cmmm$ space group, four atoms per unit cell, and represents a not-yet-observed structure prototype in binary intermetallics; here, it will be referred to as '$CdPt_3$' [34] (see appendix B and [33] for the details of the crystal structure). The data-mining prediction of '$CdPt_3$' was confirmed [33] by direct calculation of the total energy of all the 27 structure prototypes with AB_3 composition included in the data-mining library.

We performed an *ab initio* evolutionary GSGO of $CdPt_3$ in order to check whether the lowest-energy structure is '$CdPt_3$' or a structure type not included in the data-mining prediction library. The GSGO calculations were performed on Cd_2Pt_6 supercells. Out of the three independent GSGO calculations, two found as lowest-energy structure '$CdPt_3$' while the third one found a structure of higher total energy. Figure 3(a) shows the history plot of one of the GSGO calculations that found '$CdPt_3$'. The solution of the optimization is shown in figure 3(b) and was found after 32 (relaxed) total-energy evaluations: this structure has an fcc underlying lattice (see figure 3(c)), space-group symmetry $Cmmm$ and represents a realization of '$CdPt_3$' in a eight-atom supercell. The GSGO restart that did not find '$CdPt_3$' retrieved as solution an fcc-based structure with $I4/mmm$ space-group symmetry.

4.3. The Al–Sc system: the emergence of fcc-like $L1_2$ structure type for Al_3Sc, hexagonal $D0_{19}$ structure type for $AlSc_3$, and bcc-like B2 structure type for AlSc

The stable phases of Al–Sc known from experiment [36] are Al_3Sc in the $L1_2$ structure (fcc based), Al_2Sc in the C15 structure, AlSc in the B2 (bcc based), and $AlSc_2$ in the $B8_2$ structure (hcp based). The *ab initio* data-mining study of Curtarolo *et al* [33] confirms all the experimental

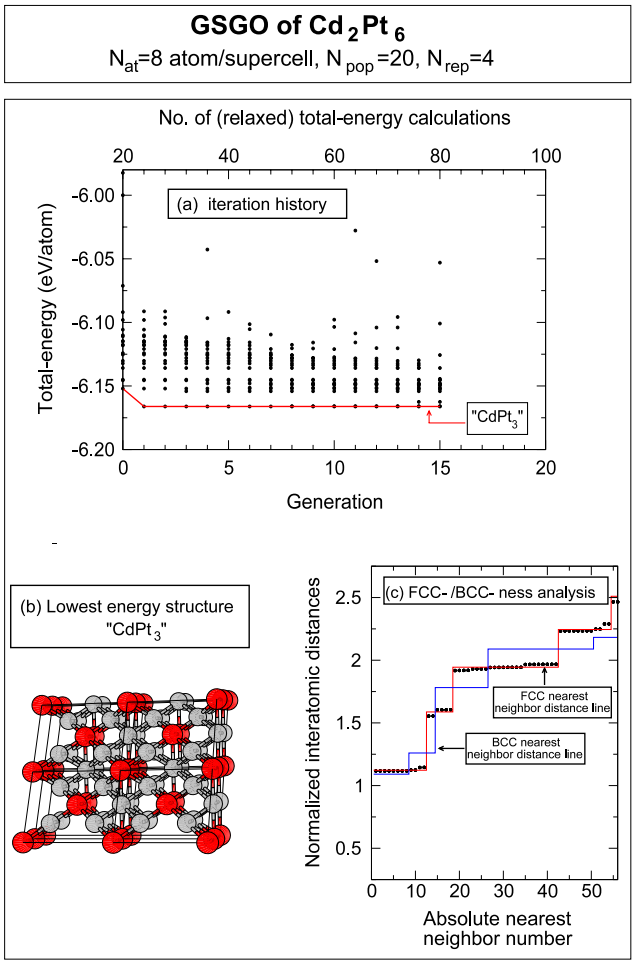

Figure 3. (a) GSGO of Cd_2Pt_6 ($N_{pop} = 20$, $N_{rep} = 4$). (b) The model of the ground-state structure found after 24 structure relaxations is $L1_3$ (space group No 65, $Cmmm$): the underlying crystal lattice is fcc as evidenced by the fcc/bcc analysis shown in (c).

ground-state structures and predicts also a stable compound at composition $AlSc_3$ with the hcp-based $D0_{19}$ structure, where the experimental phase diagram indicates, instead, coexistence of the Al and $AlSc_2$–$B8_2$ solid phases at $T > 0$ K. Here, we perform GSGO calculations for Al_6Sc_2, Al_2Sc_6, and Al_4Sc_4 supercells. We set the population size $N_{pop} = 20$ and replace the $N_{rep} = 4$ highest-energy structures at each generation. Figure 4(a) displays the evolution of one of the GSGO calculations for Al_6Sc_2. All the three independently started global optimization calculations performed for this system retrieved as lowest-energy structure $L1_2$ (space-group symmetry $Pm\bar{3}m$), which is also experimentally the most stable one at this composition. Figure 5 summarizes one of the GSGO searches performed on Al_2Sc_6. For this system in all the independent GSGO calculations that were performed $D0_{19}$ was retrieved as lowest-energy structure. In particular, the GSGO search reported in figure 5(a) attained $D0_{19}$ after 76 (relaxed) total-energy evaluations. The model of the final structure of the GSGO search is depicted in figure 5(b): $D0_{19}$ has $P6_3/mmc$ space-group symmetry and eight atoms per unit cell. Figure 6 shows the result of one of the GSGO calculations for Al_4Sc_4: for this material all the GSGO restarts found B2 as lowest-energy structure. The evolutionary search

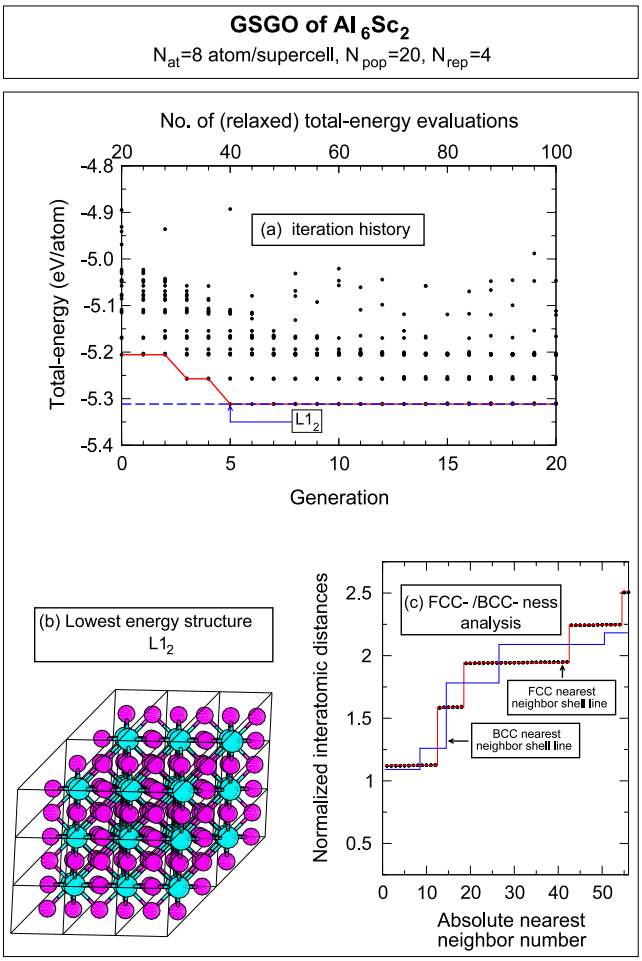

Figure 4. (a) History of the GSGO of Al_6Sc_2 ($N_{pop} = 20$ and $N_{rep} = 4$). (b) Model of the structure found by the evolutionary search, i.e. $L1_2$ (space group $Pm\bar{3}m$). This structure has fcc underlying lattice as shown by the fcc-/bcc-ness analysis (c).

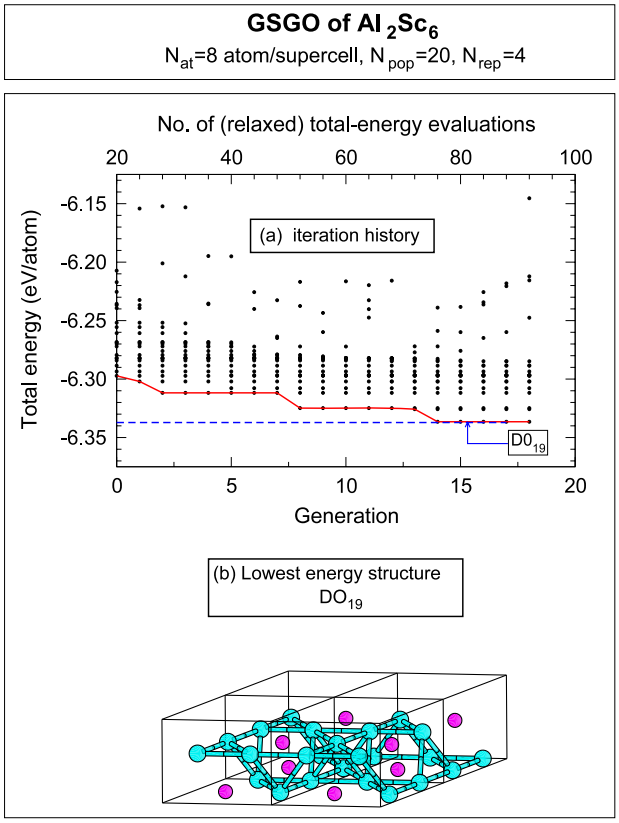

Figure 5. (a) GSGO of Al_2Sc_6 ($N_{pop} = 20$ and $N_{rep} = 4$). (b) The optimal structure obtained after 76 structure evaluations is $D0_{19}$ (space group No 194, $P6_3/mmc$).

depicted in (a) retrieved B2 as lowest-energy structure after six generations.

4.4. The Cu–Pd system: the emergence of a bcc intermetallic from fcc constituent solid elements

The Cu–Pd alloy, where both solid Cu and Pd have fcc crystal structure, exhibits a stable compound at 50%–50% concentration with the B2 ('CsCl') structure, which has an underlying bcc lattice. We performed a series of independent GSGO calculations on Cu_4Pd_4 supercells. All independent GSGO restarts retrieved the B2 structure. Figure 7 shows the evolution of one of these GSGO calculations: interestingly, in this calculation B2 was identified already in the initial, randomly generated population. Therefore, the GSGO search is able to resolve the bcc-based B2 from the fcc-based $L1_0$ structure located at energy 14 meV/atom above B2.

4.5. The $PdTi_3$ system: the emergence of the mixed fcc/bcc A15 structure-type from fcc-Pd and hcp-Ti

For the Pd–Ti alloy system we focused our attention on Ti concentration 75%. At this concentration the experimental phase diagram reports a compound with the A15 structure (see appendix A for the lattice vectors and the atomic positions of this structure). The *ab initio* study of Curtarolo *et al* [33] obtained A15 as lowest-energy structure at this composition. We performed three independent GSGO calculations taking Pd_2Ti_6 supercells. Out of these three GSGO calculations one retrieved A15. This required ∼100 (relaxed) total-energy evaluations. The history plot and the final A15 structure are shown in figure 8. The other two GSGO calculations found as lowest-energy structure the (100) superlattice (SL) with bcc parent lattice and $(Pd)_1/(Ti)_3$ layer sequence: this structure has tetragonal Bravais lattice with space group $P4/mmm$ and is higher in energy by 12 meV/atom than A15. Interestingly, the (100) SL $(Pd)_1/(Ti)_3$ was found as second-lowest-energy structure also in the optimization calculation which converged to A15.

4.6. Success rate of the evolutionary search

Because of the expense of the repeated LDA calculations, here each evolutionary sequence (part (a) of figures 2–8) is independently restarted just three or four times. Although this is not a strong basis for statistics, we comment that at least one out of the three or four independent evolutionary sequences produced the correct lowest total-energy crystal structure. For $CdPt_3$ and $PdTi_3$, the lowest total-energy structure was retrieved respectively in two out of three and one out of three

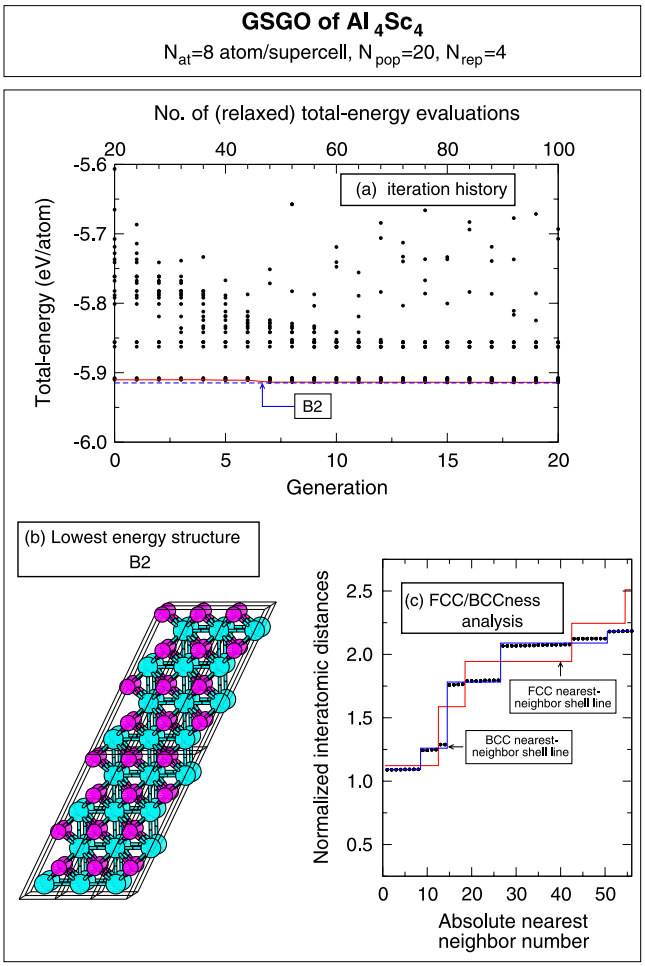

Figure 6. (a) Iteration history of the GSGO of Al_4Sc_4 ($N_{pop} = 20$ and $N_{rep} = 4$). The ground-state structure (b) is B2 (space group $Pm\bar{3}m$) obtained after 48 structure evaluations. The fcc-/bcc-ness analysis of (c) shows that the underlying lattice of the lowest-energy structure found by the GSGO search is bcc.

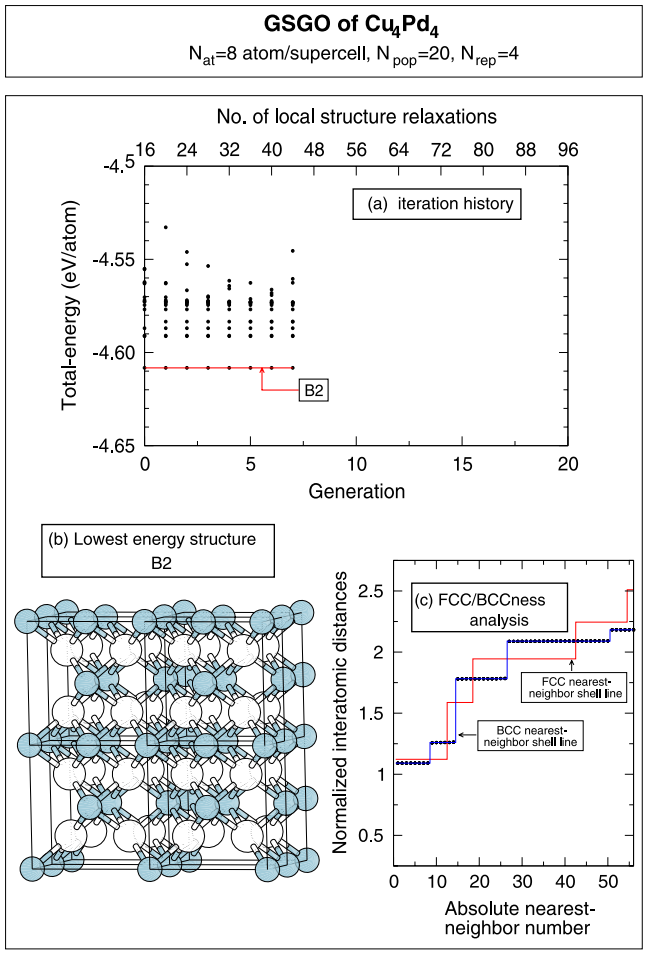

Figure 7. (a) Iteration history of the GSGO of Cu_4Pd_4 ($N_{pop} = 16$ and $N_{rep} = 4$). The ground-state structure (b) is B2 (space group $Pm\bar{3}m$) obtained already in the initial random population. The fcc-/bcc-ness analysis of (c) shows that the underlying lattice of the lowest-energy structure found by the GSGO search is bcc.

independent restarts. For all other systems the lowest total-energy structure was obtained in all restarts of the evolutionary search.

5. Discussion

We have performed a global space-group optimization of the crystal structure of the Au–Pd, Al–Sc, Cd–Pt, and Pd–Ti metallic alloys at selected compositions using an evolutionary algorithm procedure. The strength of the evolutionary-algorithm-based GSGO approach to crystal structure prediction is that no assumptions are needed for the lattice type or atomic configuration. The weakness of the evolutionary-algorithm-based GSGO is that due to its stochastic nature the search for the lowest-energy structure must proceed through a series of restarts of the evolutionary sequence (each from an independent random initial population), and success is not guaranteed. In this work only a limited number of restarts of the evolutionary algorithm were practically feasible, due to the high computational cost of performing at least 100 *ab initio* relaxed total-energy calculations per GSGO sequence. (Furthermore, each structural relaxation can be rather cumbersome due to the absence of symmetry operations, resulting in the need to include k-points over the full Brillouin zone.) Despite the limited number of restarts that were affordable, some indications can be obtained of the rate of success of the present GSGO procedure applied to binary alloys. For Au_8Pd_4 alloy, we observe several degenerate lowest-energy structures consisting of different decorations of the fcc lattice: in this case, a few independent GSGO restarts find different degenerate structures as global minimum all degenerate within the energy uncertainty of the underlying k-point and basis-set convergence of the LDA approach ($\leqslant$1 meV/atom). For materials where the lowest-energy structure has an underlying lattice of a Bravais type and has at most 10 atoms/cell, all independent restarts of the evolutionary search attained the global minimum within 80 (relaxed) total-energy evaluations. This suggests that the GSGO procedure might be able to predict the correct crystal structure with about 90% confidence requiring just 80 structure evaluations. The overall picture

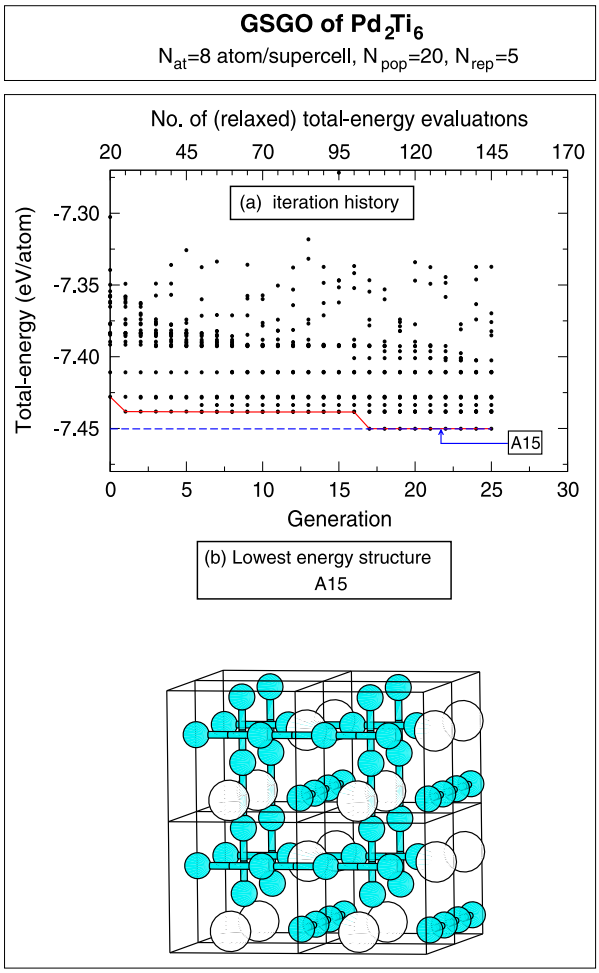

Figure 8. (a) History plot of the GSGO of Pd_2Ti_6 ($N_{pop} = 20$ and $N_{rep} = 5$). The ground-state structure (b) is A15 (space group $Pm\bar{3}n$, No 223) obtained after about 100 structure evaluations.

that emerges from the GSGO calculations performed in this work is that the first-principles evolutionary GSGO scheme can effectively address the problem of finding the lowest-energy structure of metallic binary systems where the elemental constituents have different lattice type, and for which the choice of the underlying lattice and ensuing structure type is highly non-trivial.

Acknowledgments

GT thanks Juarez Da Silva for several suggestions on the use of the VASP code. We thank Mayeul d'Avezac and Alberto Franceschetti for reading the manuscript, and Zhe Liu and Hannes Raebiger for useful discussions. This work is funded by the US Department of Energy, Office of Science, under Contract No DE-AC36-99GO10337 to NREL.

Appendix A. Parameters of the *ab initio* calculations

Here, we report the setup of the *ab initio* calculations performed on the systems studied in this work. To treat the interaction between valence electrons and core electrons and nucleus we used ultra-soft pseudopotentials. The local density approximation [45] and the generalized gradient approximation [46] to the exchange and correlation functional were used, the former for Au–Pd, Al–Sc, and Cd–Pt, while the latter for Cu–Pd and Pd–Ti.

Au–Pd. The kinetic energy cutoff of the plane-wave expansion of the wavefunctions was set equal to 20.21 Ryd. Structural relaxation and calculation of the total energy of the relaxed structure were performed using k-point meshes with a resolution of $2\pi \times 0.055$ Å^{-1} and $2\pi \times 0.025$ Å^{-1}, respectively.

Al–Sc. The pseudopotential used for Sc included the 2p electrons in the manifold of valence states. A kinetic energy cutoff of 16.4 Ryd was used for the plane-wave expansion of the wavefunctions. For the structural relaxation and the calculation of the total energy of the relaxed structure we used k-point meshes with a resolution of $2\pi \times 0.042$ Å^{-1} and $2\pi \times 0.030$ Å^{-1}, respectively.

Cd–Pt. The basis-set kinetic energy cutoff was set equal to 19.1 Ryd for the plane-wave expansion of the wavefunctions. For each individual of the population, the structural relaxation and the calculation of the total energy of the fully relaxed structure were performed using k-point meshes with a resolution of $2\pi \times 0.050$ Å^{-1} and $2\pi \times 0.030$ Å^{-1}, respectively.

Cu–Pd. In the first-principles structural relaxations for the Cu–Pd systems we used a kinetic energy cutoff of 23.9 Ryd for the plane-wave expansion of the wavefunctions. The structural relaxation and the calculation of the total energy of the final relaxed structure are carried out using k-point meshes with a resolution of $2\pi \times 0.050$ Å^{-1} and $2\pi \times 0.030$ Å^{-1}, respectively.

Pd–Ti. The ultra-soft pseudopotential for Ti treated also the 3p electrons as valence ones. The kinetic energy cutoff for the plane-wave expansion of the wavefunctions was set to 21.31 Ryd. Structural relaxation and calculation of the total energy of the relaxed structures were calculated using k-point meshes with a resolution of $2\pi \times 0.055$ Å^{-1} and $2\pi \times 0.030$ Å^{-1}, respectively.

Appendix B. Structure types of the compounds studied via GSGO

$L1_2$. Space group $Pm\bar{3}m$ (No 221 in the International Tables for Crystallography [47]); primitive vectors $\mathbf{a}_1 = a(1, 0, 0)$, $\mathbf{a}_2 = a(0, 1, 0)$, $\mathbf{a}_3 = a(0, 0, 1)$. Atomic positions (Wyckoff positions and fractional coordinates): A(1), (1a) (0, 0, 0); B(1), (3c) (0, 1/2, 1/2); B(2), (3c) (1/2, 0, 1/2); B(3), (3c)(1/2, 1/2, 0).

'$CdPt_3$'. Space group $Cmmm$ (No 65 in the International Tables for Crystallography [47]); primitive vectors $\mathbf{a}_1 = a(1, -1/2, 1/2)$, $\mathbf{a}_2 = a(-1, -1/2, 1/2)$, $\mathbf{a}_3 = a(0, -1/2, -1/2)$. Atomic positions (Wyckoff positions and fractional coordinates): A(1), (2a) (0, 0, 0); B(1), (4f) (0, 1/2, 1/2); B(2), (4f) (1/2, 0, 1/2); B(3), (2b) (1/2, 1/2, 0).

$D0_{19}$. Space group $P6_3/mmc$ (No 194 in the International Tables for Crystallography [47]); primitive vectors $\mathbf{a}_1 = a(1/2, -\sqrt{3}/2, 0)$, $\mathbf{a}_2 = a(1/2, \sqrt{3}/2, 0)$,

$\mathbf{a}_3 = a(0, 0, c/a)$. Atomic positions (Wyckoff positions and fractional coordinates): A(1), (2c) (1/3, 2/3, 1/4); A(2), (2c) (2/3, 1/3, 3/4); B(1), (6h) $(x, 2x, 1/4)$; B(2), (6h) $(-2x, -x, 1/4)$; B(3), (6h) $(x, -x, 1/4)$; B(4), (6h) $(-x, -2x, 3/4)$; B(5), (6h) $(2x, x, 3/4)$; B(6), (6h) $(-x, x, 3/4)$.

B2. Space group $Pm\bar{3}m$ (No 221 in the International Tables for Crystallography [47]); primitive vectors $\mathbf{a}_1 = a(1, 0, 0)$, $\mathbf{a}_2 = a(0, 1, 0)$, $\mathbf{a}_3 = a(0, 0, 1)$. Atomic positions (Wyckoff positions and fractional coordinates): A(1), (1a) (0, 0, 0); B(1), (1b) (1/2, 1/2, 1/2).

A15. Space group $Pm\bar{3}n$ (No 223 in the International Tables for Crystallography [47]); primitive vectors $\mathbf{a}_1 = a(1, 0, 0)$, $\mathbf{a}_2 = a(0, 1, 0)$, $\mathbf{a}_3 = a(0, 0, 1)$. Atomic positions (Wyckoff positions and fractional coordinates): A(1), (2a) (0, 0, 0); A(2), (2a)(1/2, 1/2, 1/2); B(1), (6c) (1/4, 1/2, 0); B(2), (6c) (3/4, 1/2, 0); B(3), (6c) (0, 1/4, 1/2); B(4), (6c) (0, 3/4, 1/2); B(5), (6c) (1/2, 0, 1/4); B(6), (6c) (1/2, 0, 3/4).

Structure '4557' space group $C2/m$ (No 12 in the International Tables for Crystallography [47]); primitive vectors $\mathbf{a}_1 = a(1.0, 0.0, 0.0)$, $\mathbf{a}_2 = a(0.5, 1.0, 0.5)$, $\mathbf{a}_3 = a(0.5, -1.0, 2.5)$. Atomic positions (Cartesian coordinates): A(1), (1.5, 0.0, 2.5); A(2), (0.5, −0.5, 2.0); A(3), (1.0, 0.0, 2.0); A(4), (0.5, 0.0, 1.5); A(5), (1.0, 0.5, 1.5); A(6), (0.5, 0.5, 1.0); A(7), (0.5, 0.0, 0.5); A(8), (1.0, 0.5, 0.5); B(1), (0.0, 0.0, 0.0); B(2), (1.0, −0.5, 2.5); B(3), (1.0, −0.5, 1.5); B(4), (1.0, 0.0, 1.0).

Structure '4905' space group $Immm$ (No 71 in the International Tables for Crystallography [47]); primitive vectors $\mathbf{a}_1 = a(1.0, 0.0, 0.0)$, $\mathbf{a}_2 = a(0.5, 1.5, 1.0)$, $\mathbf{a}_3 = a(-0.5, -1.5, 1.0)$. Atomic positions (Cartesian coordinates): A(1), (0.0, −1.0, 1.0); A(2), (0.0, 0.0, 1.0); A(3), (0.5, 0.5, 1.0); A(4), (1.0, 1.0, 1.0); A(5), (0.5, 0.0, 0.5); A(6), (0.0, −0.5, 1.5); A(7), (0.5, 0.0, 1.5); A(8), (0.0, −0.5, 0.5); B(1), (0.0, 0.0, 0.0); B(2), (0.5, −0.5, 1.0); B(3), (1.0, 0.5, 1.5); B(4), (1.0, 0.5, 0.5).

Structure '4820' space group $Cmcm$ (No 63 in the International Tables for Crystallography [47]); primitive vectors $\mathbf{a}_1 = a(1.0, 0.0, 0.0)$, $\mathbf{a}_2 = a(0.5, 1.5, 0.0)$, $\mathbf{a}_3 = a(0.0, 0.0, 2.0)$. Atomic positions (Cartesian coordinates): A(1), (0.5, 0.5, 0.0); A(2), (1.0, 1.0, 0.0); A(3), (0.5, 0.0, 0.5); A(4), (1.0, 0.5, 0.5); A(5), (0.5, 0.5, 1.0); A(6), (1.0, 1.0, 1.0); A(7), (0.5, 0.0, 1.5); A(8), (0.5, 1.0, 1.5); B(1), (0.0, 0.0, 0.0); B(2), (0.5, 1.0, 0.5); B(3), (0.0, 0.0, 1.0); B(4), (1.0, 0.5, 1.5).

Structure '4410' space group $P2_1/m$ (No 11 in the International Tables for Crystallography [47]); primitive vectors $\mathbf{a}_1 = a(1.0, 0.0, 0.0)$, $\mathbf{a}_2 = a(0.0, 1.5, 0.5)$, $\mathbf{a}_3 = a(0.0, 0.0, 2.0)$. Atomic positions (Cartesian coordinates): A(1), (0.5, 0.5, 2.0); A(2), (0.5, 0.0, 1.5); A(3), (0.0, 0.5, 1.5); A(4), (0.0, 0.0, 1.0); A(5), (0.0, 1.0, 1.0); A(6), (0.5, 0.0, 0.5); A(7), (0.0, 0.5, 0.5); A(8), (0.5, 1.0, 0.5); B(1), (0.0, 0.0, 0.0); B(2), (0.0, 1.0, 2.0); B(3), (0.5, 1.0, 1.5); B(4), (0.5, 0.5, 1.0).

Note added in proof. Recently, Aberg *et al* [48] have proposed a zero-pressure lowest-energy structure for IrN_2, which they refer to as ST_{AA}, based on static LDA calculations. Here, we performed an evolutionary search for Ir_2N_4 to determine the lowest-energy structure of this material at zero pressure. Figure 9(a) shows the history of the evolutionary sequence. This search run revealed a new low total-energy structure for IrN_2 displayed in figure 9(b); this structure has space group $P2/m$ and it is different from all structures previously proposed for IrN_2. Taking as reference the energy of the structure ST_{AA} this new GSGO structure has an energy of −0.057 eV/f.u. and −0.278 eV/f.u. respectively in LDA and GGA: clearly the structure obtained here by GSGO of lower energy than the lowest energy structure previously found for IrN_2.

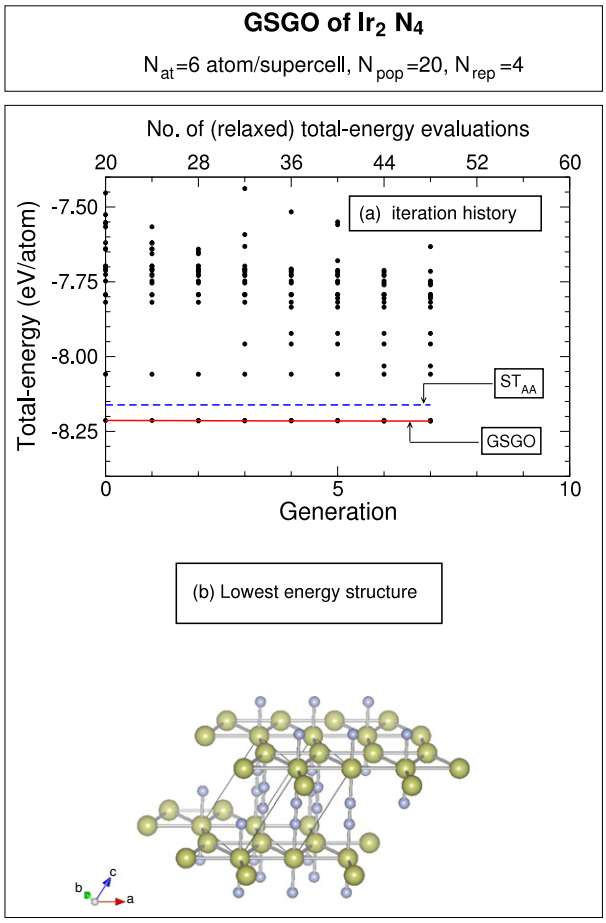

Figure 9. (a) Iteration history of the GSGO of Ir_2N_4 ($N_{pop} = 20$, $N_{rep} = 24$). The ground-state structure (b) has space group $P2/m$.

References

[1] Pauling L 1960 *The Nature of the Chemical Bond* (Ithaca: Cornell University Press)

[2] Hume-Rothery W and Rainor G 1954 *The Structure of Metals and Alloys* (London: Institute of Metals)

[3] Ising E 1925 *Z. Phys.* **31** 253

[4] Bragg W L and Williams E J 1934 *Proc. R. Soc.* A **145** 699

[5] Barahona F 1982 *J. Phys. A: Math. Gen.* **15** 3241

[6] Kresse G and Hafner J 1993 *Phys. Rev.* B **47** R558

[7] Sanchez J M, Ducastelle F and Gratias D 1984 *Physica* A **128** 334

[8] Zunger A 1994 *Statistics and Dynamics of Alloys Phase Transformations* (*NATO ASI*) ed P E A Turchi and A Gonis (New York: Plenum) p 361

[9] Ferreira L G, Wei S-H and Zunger A 1991 *Int. J. Supercomput. Appl.* **5** 34

[10] Trimarchi G, Graf P and Zunger A 2006 *Phys. Rev.* B **74** 014204

[11] d'Avezac M and Zunger A 2007 *J. Phys.: Condens. Matter* **19** 402201
[12] d'Avezac M and Zunger A 2008 submitted
[13] Wolverton C and Zunger A 1998 *Phys. Rev. Lett.* **81** 606
[14] Hart G L W and Zunger A 2001 *Phys. Rev. Lett.* **87** 275508
[15] Blum V and Zunger A 2004 *Phys. Rev.* B **70** 155108
[16] Barabash S V, Blum V, Muller S and Zunger A 2006 *Phys. Rev.* B **74** 035108
[17] Drautz R, Díaz-Ortiz A, Fähnle M and Dosch H 2004 *Phys. Rev. Lett.* **93** 067202
[18] Kolb B and Hart G L W 2005 *Phys. Rev.* B **72** 224207
[19] Froyen S and Cohen M L 1983 *Phys. Rev.* B **28** 3258
[20] Froyen S and Cohen M L 1984 *Phys. Rev. B (R)* **29** 3770
[21] Yin M T and Cohen M L 1984 *Phys. Rev. Lett.* **50** 2006
[22] Parrinello M and Rahman A 1980 *Phys. Rev. Lett.* **45** 1196
[23] Wentzcovitch R 1991 *Phys. Rev.* B **44** 2358
[24] Laio A and Parrinello M 2002 *Proc. Natl Acad. Sci. USA* **99** 12562
[25] Martonak R, Laio A and Parrinello M 2003 *Phys. Rev. Lett.* **90** 075503
[26] Ozoliņš V, Wolverton C and Zunger A 1998 *Phys. Rev.* B **57** 6427
[27] Sanati M, Wang L G and Zunger A 2003 *Phys. Rev. Lett.* **90** 045502
[28] Blum V and Zunger A 2004 *Phys. Rev.* B **69** 020103(R)
[29] Deaven D M and Ho K M 1995 *Phys. Rev. Lett.* **75** 288
[30] Oganov A R and Glass C W 2006 *J. Chem. Phys.* **124** 244704
[31] Abraham N L and Probert M I J 2006 *Phys. Rev.* B **73** 224104
[32] Trimarchi G and Zunger A 2007 *Phys. Rev.* B **75** 104113
[33] Curtarolo S, Morgan D and Ceder G 2005 *CALPHAD* **29** 163
[34] Hart G L W 2007 *Nat. Mater.* **6** 941
[35] Hultgren R R 1963 *Selected Values of Thermodynamic Properties of Metals and Alloys* (New York: Wiley–Interscience)
[36] Massalski T B (ed) 1992 *Binary alloy Phase Diagrams* (Metals Park, OH: ASM International)
[37] Oganov A R, Glass C W and Ono S 2006 *Earth Planet Sci. Lett.* **241** 95
[38] Kresse G and Hafner J 1993 *Phys. Rev.* B **47** 558
[39] Kresse G and Furthmuller J 1996 *Phys. Rev.* B **54** 11169
[40] Spek A L 2003 *J. Appl. Crystallogr.* **36** 7
[41] Spek A L 2003 *PLATON, A Multipurpose Crystallographic Tool* Utrecht University Utrecht, The Netherlands (1)
[42] http://www.cryst.chem.uu.nl/platon/
[43] Oganov A R, Ma Y, Glass C W and Valle M 2007 *Psi-k Newslett.* **84** 142 http://www.psi-k.org/newsletters/News_84/Highlight_84.pdf
[44] Curtarolo S, Morgan D, Persson K, Rodgers J and Ceder G 2003 *Phys. Rev. Lett.* **91** 135503
[45] Perdew J P and Zunger A 1981 *Phys. Rev.* B **23** 5048
[46] Perdew J P and Wang Y 1992 *Phys. Rev.* B **45** 13244
[47] Hahn T (ed) 1983 *International Tables for Crystallography* vol A (Dordrecht: Reidel)
[48] Aberg D, Sadigh B, Crowhurst J and Goncharov A F 2008 *Phys. Rev. Lett.* **100** 095501