Emergence of a few distinct structures from a single formal structure type during high-throughput screening for stable compounds: The case of RbCuS and RbCuSe

Giancarlo Trimarchi, Xiuwen Zhang, Michael J. DeVries Vermeer, Jacqueline Cantwell, Kenneth R. Poeppelmeier, and Alex Zunger

Phys. Rev. B **92**, 165103 — Published 2 October 2015

DOI: 10.1103/PhysRevB.92.165103

Emergence of a few distinct structures from a single formal structure type during high-throughput screening for stable compounds: the case of RbCuS and RbCuSe

Giancarlo Trimarchi,[1,*] Xiuwen Zhang,[2] Michael J. DeVries Vermeer,[3]
Jacqueline Cantwell,[3] Kenneth R. Poeppelmeier,[3] and Alex Zunger[2]

[1]*Department of Physics and Astronomy, Northwestern University*

[2]*University of Colorado, Boulder*

[3]*Department of Chemistry, Northwestern University*

(Dated: August 31, 2015)

Abstract

Theoretical sorting of stable and synthesizable "missing compounds" from those that are unstable is a crucial step in the discovery of previously unknown functional materials. This active research area often involves high throughput (HT) examination of the total energy of a given compound in a list of candidate formal structure types (FSTs), searching for those with the lowest energy within that list. While it is well appreciated that local relaxation methods based on a fixed list of structure types can lead to inaccurate geometries, this approach is widely used in HT studies because it produces answers faster than global optimization methods (that vary lattice vectors and atomic positions without local restrictions). We find, however a different failure mode of the HT protocol: specific crystallographic classes of formal structure types correspond each to a series of chemically distinct "daughter structure types" (DSTs) that have the same space group, but posses totally different local bonding configurations, including coordination types. Failure to include such DSTs in the fixed list of examined candidate structures used in contemporary high throughput approaches can lead to qualitative misidentification of the stable bonding pattern, not just quantitative inaccuracies. In this work we (i) clarify the understanding of the general DST-FST relationship, thus improving current discovery HT approaches, (ii) illustrate this failure mode for RbCuS and RbCuSe (a newly predicted compound) by developing a synthesis method and accelerated crystal structure determination and, (iii) apply the genetic-algorithm based Global Space Group Optimization (GSGO) approach which is not vulnerable to the failure-mode of HT searches of fixed lists, demonstrating a correct identification of the stable DST. The broad impact of items (i)-(iii) lies in the demonstrated understanding of a new strategy–use HT as preliminary broad screening, followed by unbiased GSGO of the final candidates.

I. INTRODUCTION

The quest for new materials with important functionalities has stimulated the application of *ab initio* thermodynamics to predict the stability and structure of series of compounds given their chemical composition. This is illustrated in Fig. 1(a) for the I-I-VI ABX compounds showing that along with documented compounds (indicated by check marks) there are many unreported ones (indicated by question marks). Finding out whether a compound exists with a given composition, e.g., $A_pB_qX_r$ requires first to predict the lowest-energy crystal structure at that composition; second, one has to verify if the $A_pB_qX_r$ lowest energy structure is energetically more favorable than any of the possible phase decompositions[1–3] into the pure A, B, and X elements and the binary or ternary compounds they form.

To illustrate this we show in Fig. 1(b) the *ab initio* total-energy predictions of the stable and unstable I-I-VI ABXs assuming just the $C1_b$ structure as a possible competing phase. The "missing compounds" of Fig. 1(a) are thus sorted out in Fig. 1(b) into "missing-found-stable" (denoted by a "+" sign) or "missing-found-unstable" (denoted by a "-" sign). We see that using a single structure types as in Ref. 4 and 5 suggests that most previously experimentally missing compounds are unstable, thus providing a rational explanation for their absence.

The obvious generalization of this procedure is to inspect $N \gg 1$ different candidate structure prototypes, i.e., a longer, yet fixed list of structure types. The application of this more general screening method to the I-I-VI ABX compounds is illustrated in Fig. 1(c) where we examine a series of candidate structure (a subset of which is depicted in Fig. 2) the lowest energy structure of the ternary phase (as illustrated by the ladder diagram in Fig. 3) in parallel with assuring it is also stable with respect to decomposition into any combination of unary, binary and ternary competing phases. When this protocol is applied to ABX compounds that are previously synthesized we invariably predict stable structures in the observed phase.[6] We see from Fig. 1(c) that the vast majority of missing compounds predicted to be "unstable" when a single structure type was considered (Fig. 1(b)) are in fact false negative predictions; these compounds turn out to be mostly stable when more structure types are allowed (Fig. 1(c)). Specifically, RbCuS and RbCuSe are predicted stable in Fig. 1(c).

It is well appreciated that local relaxation methods based on a fixed list of structure types can lead to inaccurate geometries. Yet, precisely this approach is widely used in high-throughput studies such as those summarized in Fig. 1(b) and 1(c), the reason often being that it produces

answers faster than global optimization methods which vary the lattice vectors and atomic positions without local restrictions. Here, we find, a different failure mode of the HT protocol: specific crystallographic classes of "formal structure types" (FSTs) correspond each to a series of chemically distinct "daughter structure types" (DSTs) that have the same space group, but possess totally different local bonding configurations, including coordination types. Failure to include such DSTs in the fixed list of examined candidate structures used in contemporary HT approaches can lead to qualitative misidentification of the stable bonding pattern, not just quantitative inaccuracies. This significant "navigation error" is avoided if one starts with randomly guessed lattice vectors and atom positions and applies a global search method not confined to the local neighborhood of the initial guesses. This type of search produces for RbCuS and RbCuSe structures (Fig. 4(c)) that agree with those experimentally observed also in the local bonding configurations (Fig. 4(a)).

The significant findings of this work are: (i) we clarify the understanding of the general DST-FST relationship, thus improving current discovery HT approaches. The realization that the same "formal" structure type can conceal multiple daughter structures with different local atomic coordination and environments is important because many interesting physical properties are determined mainly by the bonding configurations and local symmetry, e.g., topological insulators, ferroelectricity, magnetism, etc. (ii) We illustrate this failure for RbCuS and RbCuSe (a newly predicted compound) by developing a synthesis method and accelerated crystal structure determination, and (iii) we apply the evolutionary-algorithm based Global Space Group Optimization (GSGO) approach which is not vulnerable to the failure mode of HT searches of fixed lists, demonstrating a correct identification of the stable DST. The broad impact of items (i)-(iii) lies in the demonstrated understanding of a new strategyuse HT as preliminary broad screening, followed by unbiased global Space Group optimization of the final candidates.

II. SCREENING MATERIALS FROM A FIXED LIST OF FORMAL STRUCTURE TYPES (FSTS) IN "HIGH THROUGHPUT" APPROACHES

A structure type is formally defined[7] by its space group symmetry, set of occupied Wyckoff positions, and chemical formula. Two structures that coincide with respect to these descriptors are deemed isotypical.[7,8] This is the classic definition used by Ewald and Hermann in the *Strukturbericht* to classify the crystal structures of solids, and we will refer to the so-defined structure type as to a "formal structure type" (FST). The constructs of a FST provide the information needed

to specify the topology of the periodic atom packing and bonding network in the crystal only if the coordinates of the occupied Wyckoff positions in the FST are constant numbers mandated by the symmetry operations of each Wyckoff orbit. The conventional fixed-list high-throughput screening approach[9–11] consists of three steps:

(i) In step 1 one selects a pool of candidate formal structure types. The common hypothesis made is that a new compound in a chemical family of systems of interest, e.g., ABX, will likely adopt one of the structure prototypes previously observed in compounds belonging to that chemical family. This hypothesis stems from the empirical observation that the compounds observed in a given chemical family (e.g., ABX, ABX_2, ABX_3, etc.) cluster in a finite and usually small number of subgroups each corresponding to a distinct structure prototype. Now, in some structures (as in the case of $C1_b$, see Tab. I) the Cartesian coordinates of a Wyckoff site are given by pure numbers mandated by space-group operations. However, in other structure types the Wyckoff sites are specified by (u, v, w) parameters some of which are not decided by symmetry (as in the compounds RbAuS and MgSrSi, see Tab. I).

(ii) In step 2, the initial numerical values of the atomic positions in the case of symmetry-unconstrained Wyckoff coordinates are taken from an experimentally known, (but often chemically different) compound. The atoms in such representative solids are replaced by the atoms of the target compound of interest. For instance, the atomic coordinates and lattice parameters of RbAuS are used as guess for the hypothetical compound RbCuS even though the lattice parameters and distances can be very different in Cu vs. Au compounds.

(iii) Finally, in step 3 total energy and force minimization (see Ihm *et al.*, Ref. 12) is applied to the target compound in each structure, searching for the minimum in the Born-Oppenheimer (BO) surface starting from the realization given by whatever similar compound is experimentally known. In doing so one hopes that the final structure predicted will depend neither on the initial choice of the FSTs in step (i) nor on the initial atomic parameters guessed in step (ii) above.

However, as will be shown here, the BO surface of a FST may have several local minima (daughter structure types or DSTs) different from each other by significant crystal structure features, (e.g., the local atomic coordination around atoms), yet all belonging to the same FST. For example Liu and Corbett[13] identified different local crystal coordination in CaSrSi and $CoSi_2$ that have the same structure type. In the present study we identify the general circumstances that cause such a splitting of a formal structure type into different daughter structures and use this understanding to point potential pitfalls in the high throughput protocol for material discovery which

uses only a list of FSTs. We point out that when at least some of the atom coordinates (u, v, w) in the unit cells are not decided by symmetry, then using values taken from a different material as initial guess of the local energy minimization may land in a different DST then the correct one. Reports of isostructural transitions in crystalline solids, i.e., where not only the space group but also the set of occupied Wyckoff positions does not change, are in fact rare. One of the few cases we were able to find is that of the iso-symmetric transition in $BiFeO_3$ under biaxial strain (see Schlom *et al.* in Ref. 14). Other cases are reviewed by Christy[15]. So, RbCuS and RbCuSe provide new examples of rarely observed iso-symmetrical poly-types.

Here, we examine the theoretical predictions of the stable structures for RbCuS and RbCuSe, which are thus far unreported compounds. Using the standard high throughput three-step protocol explained above with 40 structure types and initial values for the symmetry-unconstrained coordinates taken from prototype materials listed in the ICSD[16] (see Fig. 3) we predict that the lowest energy structure of RbCuS and RbCuSe belongs to a FST with same space group ($Cmcm$) and set of Wyckoff positions as RbAuS (shown in Fig. 2(d)). Subsequent synthesis of these never before made compounds described in Sec. V below confirmed the formation of RbCuS and RbCuSe in the predicted FST RbAuS-type space-group symmetry ($Cmcm$) and Wyckoff atomic positions (i.e., $4c$ and $4a$) shown in Fig. 4(b) but the experimental structures shown in Fig. 4(a) has qualitatively different local coordination environments than in the structure predicted by local minimization starting from the RbAuS lattice. The differences are apparent by comparing Fig. 4(a) and 4(b). The correct prediction (Fig. 4(c)) is provided by the Global Space Group Optimization (GSGO) approaches that start from randomly selected lattice structures and guide the exploration by evolutionary (e.g., genetic) algorithms.

III. THE GROUP OF MATERIALS CONSIDERED AND ITS "MISSING MEMBERS"

The I-I-VI compounds (see Fig.1(a)) based on two monovalent and one exavalent species form a rich group of solids that can be thought of as (formally) derived from the well-known octet[17,18] binary II-VI semiconductors with the zinc blend structure by splitting the column II cation into two monovalent cations, either from column IA or IB of the periodic table. The zinc blend structure in principle can accommodate both cations by stuffing the extra cation in the $(1/2, 1/2, 1/2)$ vacancy site[1,19–22], after the $(0, 0, 0)$ and $(1/4, 1/4, 1/4)$ sites along the diagonal of the cubic cell, thus obtaining one the so-called filled tetrahedral structures,[17] i.e., the half Heuseler alloy ($C1_b$)

structure depicted in Fig. 2(a).

$C1_b$ is only one of the numerous structure types in which the I-I-VI compounds can form and some of the frequently recurring ones are depicted in Fig. 2. Most notably, while several I-I-VI compounds have been previously synthesized (check marks in Fig. 1(a)) there are numerous "missing compounds" that are unreported members in this group (question marks in Fig. 1(a)) as well as in several other *ABX* groups. We focus on the RbCu$X^{\rm VI}$ compounds (shown in the upper left corner of each square in Fig. 1) and specifically on those with $X^{\rm VI}$=S and Se which are not reported in the ICSD[23], while RbCuO has been synthesized and forms[24] in a structure analogous to that of CuLiO.

IV. HIGH THROUGHPUT PREDICTION OF THE STRUCTURES OF RbCuS AND RbCuSe

In Ref. 1 the structure of the missing I-I-VI compounds were predicted by the procedure described in the previous section screening 41 structure prototypes. The compounds predicted stable are summarized in Fig. 1(c). Here, we revisit (and revise) the predictions of the lowest-energy structure for RbCuS and RbCuSe by performing *ab initio* local total-energy minimization calculations with stricter convergence criteria (see the Supporting Information for details). We consider a subset of the 41 prototypes used in Ref. 1 (namely a selection of the lowest-energy structures that were predicted in that work for the I-I-VI systems) as starting trial structures and relax them to the local energy minimum. The ladder diagram of Fig. 3 shows the total energies of RbCuS and RbCuSe for the fully relaxed structures of the considered prototypes. Each level is associated to a local energy-minimum structure and is labeled after the name of the known compound whose experimental structure was used as starting point of the local minimization. For instance the CuLiO-type minimum is obtained substituting Rb on Li and S on the O sites in the experimental structure of CuLiO and then doing a full relaxation.

We find that the RbAuS-type structure (see Tab. I) is the lowest-energy one both for RbCuS and RbCuSe. CuLiO-type is the prototype next in energy at about 0.06 eV/f.u. from the RbAuS-type one, followed by the MgSrSi-, LiCaN-, and ZrBeSi-type structures. MgSrSi and LiCaN have structures with the same space group symmetry and set of Wyckoff sites, therefore correspond to the same formal structure type. However, they have different numerical values of the Wyckoff coordinates and lattice parameters thus exhibiting qualitatively different bonding and coordination environments and making them two different starting configurations for the local energy minimiza-

tion of RbCuS and RbCuSe. Interestingly, in the case of RbCuSe the LiCaN-type and MgSrSi-type structures are degenerate after full local relaxation. Instead, in the case of RbCuS, the MgSrSi-type and LiCaN-type relaxed structures correspond to two different and non-degenerate atom arrangements which are therefore distinct daughter structure types of the same FST. The structure we identify as the lowest energy one both for RbCuS and RbCuSe is described for RbCuS in Fig. 4(b) and in Tables II and III, and with analogous tables and figure in the Supplementary Material for RbCuSe.

V. SYNTHESIS AND CRYSTAL STRUCTURE REFINEMENT OF RbCuS AND RbCuSe: REALIZING THAT THE EXPERIMENTAL AND HIGH-THROUGHPUT-PREDICTED STRUCTURES ARE DIFFERENT "DAUGHTERS" OF THE SAME "FORMAL STRUCTURE TYPE"

A synthesis of RbCuS was initially reported in Ref. 25 with details on the preparation and observed crystal structure reported by Boller in Ref. 26. At the best of our knowledge RbCuSe has not been previously reported[23]. Here, RbCuS and RbCuSe were synthesized via vacuum annealing of the reagents inside fused silica ampoules. Due to the air-sensitive nature of reagents and products, all preparation and characterization was done under a nitrogen atmosphere (see Supplementary Information for the details on the synthesis and structure characterization methods). The syntheses produced RbCuS and RbCuSe ternary compounds with the predicted $1{:}1{:}1$ composition. Refinement of the measured single-crystal XRD patterns of RbCuS and RbCuSe showed that both compounds crystallize in a structure with the orthorhombic $Cmcm$ space-group symmetry with the Cu atoms at the $4a$ sites and the Rb and S (Se) atoms at the $4c$ Wyckoff. The measured unit cell parameters and Wyckoff coordinates for RbCuS are reported respectively in Table II and III (see the Supplementary Material for the analogous structural informaiton for RbCuSe)

The refined crystal structure of RbCuS is drawn in Fig. 4(a). The Cu and S atoms form zigzag planar chains running along the a direction with the S atoms at the elbow sites in these chains. The Rb atoms lie in the plane defined by the Cu-S-Cu chains and are positioned 3.278 Å away from the S elbow sites. Fig. 4(b) displays RbCuS in the RbAuS-type structure which is the lowest-energy structure predicted by HT screening and obtained replacing Au with Cu in RbAuS and then fully relaxing to the local minimum (see Fig. 3). The experimental and the HT-predicted RbAuS-type structures are isotypical: both have $Cmcm$ space group symmetry with the Rb and S atoms occupying the $4c$ sites while Cu occupying the $4a$ sites (see Table I). However, RbCuS

in the RbAuS-type structure has remarkably different lattice parameters than in the experimental structure (see Tab. II). Indeed, the a parameter along the chain directions is 6.61 Å in the RbAuS-type structure vs. 4.90 Å in the experimental structure; the b parameter, i.e., the distance between coplanar chains, is 11.71 Å in the RbAuS-type structure, which is much shorter than 14.47 Å, the b value in the experimental structure. Furthermore, in the RbAuS-type structure of RbCuS the Rb-S distance is 6.456 Å, which is 3.178 Å longer than the 3.278 Å Rb-S distance in the experimental structure, thus indicating a qualitative different arrangement for the Rb atoms in the two structures. An analogous difference is found between the experimental and RbAuS-type structure predicted by HT screening for RbCuSe.

The experimental and RbAuS-type structures are different "daughter" structures of the same formal structure type and correspond to distinct minima on the Born-Oppenheimer. Indeed, the minimum found by *ab initio* relaxation of the RbCuS experimental structure is 0.04 eV/f.u. below the RbAuS-type minimum. In order to investigate the profile of the BO surface in the vicinity of these two local minina and in the region that separates them we applied solid-state nudged-elastic band method.[27] Fig. 5 shows the minimum energy reaction path connecting the two minima and consisting of a concerted switch in the height of the Rb atoms with respect to the S elbow atoms occurring along with a change in the cell parameters. We find that the two local minima following this reaction path are separated by a barrier of 0.4 eV/f.u. When the local energy optimization of RbCuS starts from the atom positions copied from the experimental structure of RbAuS one reaches the higher energy RbAuS-type daugher structure type instead of the correct ground state structure which is experimentally observed.

The present results on RbCuS and RbCuSe highlight the possibility that even though the correct FST is in the list of candidate prototypes examined by HT screening, the correct DST might be missed because of the unrealistic guess of the starting point for the local minimization. This issue originates from the fact that the crystal structure prediction problem is inherently a *global optimization* problem that requires the application an unbiased search method with a global view of the configuration space of the system to optimize. The HT search protocol is instead biased by the choice of the initial trial structures that in fact lock the screening into a few basins of attraction without the flexibility of exploring other possible basins on the BO surface of the solid of interest. In the next section we will apply an unbiased, global optimization method to predict the crystal structure of RbCuS and RbCuSe.

VI. APPLICATION OF THE GLOBAL SPACE GROUP OPTIMIZATION (GSGO) METHOD TO RbCuS AND RbCuSe: FINDING THE CORRECT DAUGHTER STRUCTURE TYPE

The prediction of the stable structure of an $A_pB_qX_r$ system requires to search for the lattice vectors and atom positions of the global minimum of the total energy $E(\mathbf{x})$ as a function of the the atom coordinates $\mathbf{x} = (\mathbf{r}_1, \ldots, \mathbf{r}_{N_{at}})$. The total-energy $E(\mathbf{x})$ of a solid is a highly non-linear function[28,29] and as a result of this in general it features multiple local minima corresponding to which often correspond structures that might be quite different than that of the global energy minimum. Several global optimization methods are currently widely used to search for the lowest energy structure of a solid by exploring in an unbiased fashion the Born-Oppenheimer surface of a solid calculated by DFT. These computational appraoches include evolutionary algorithms[30–34], particle swarm optimization,[35] methods based on molecular dynamics, such as simulated annealing[36], minima hopping,[37] and metadynamics[38]. Also a search approach based on the local energy minimization of randomly generated structures[39] has been successfully applied to numerous materials.

Here, we apply the global space-group optimization (GSGO)[33,34] implementation of the evolutionary algorithm for crystal structure prediction. The evolutionary search starts from a set of candidate structures generated by randomly selecting the lattice vectors and atom positions. The population of candidate structures is then evolved through a series of generations in which the higher energy structures (e.g., 25% of the whole population) are discarded and substituted with new structures produced by the evolutionary operators of mutation and heredity, here consisting of a cut-and-splice crossover[32]. Each newly generated structure is fully relaxed to the closest local minimum of the total-energy of the periodic system that is calculated by density functional theory. The fully relaxed structures are finally included in the next generation of the population (see, e.g., Refs.[34] for some applications). Since previous tests[1,3] of crystal structure prediction by the two approaches of GSGO automatic search and screening of a range of formal structure types gave the same answers, the computationally more costly GSGO is not always used.

The GSGO algorithm predicts for RbCuS the structure shown in Fig. 4(c) (see the Supplementary Information for the structure predicted by GSGO for RbCuSe). In order to determine the space group and set of Wyckoff positions in the GSGO structures we used the FINDSYM[40] symmetry analysis algorithm. We find that the GSGO-predicted structures for RbCuS and RbCuSe have the $Cmcm$ space group symmetry with the lattice parameters and Wyckoff positions reported

in Tables II and III and in the Supporting Information, respectively for RbCuS and RbCuSe. The GSGO predicted and the experimental structures have virtually the same total energies when these are fully relaxed by DFT and lower in energy than the local minimum RbAuS-type structure (see Fig. 3) found by prototype screening and shown in Fig. 4(c). The Rb, Cu, and X^{VI} atoms occupy the same Wyckoff orbits in the GSGO structures as in the experimental structures, and a visual inspection of the two structures confirms a perfect match between them.

VII. CONSEQUENCES OF THE EXISTENCE OF MULTIPLE DAUGHTER STRUCTURE TYPES ON MATERIALS DESIGN AND DISCOVERY

Our theoretical results on the RbCuS and RbCuSe compounds obtained from HT and GSGO calculations along with the synthesis and experimental observations we performed, point to the fact that the same "formal" structure type can conceal "daughter" structures with different local atomic coordination and environments. This is an important fact because many interesting physical properties, e.g., topological insulators, ferroelectricity, non-linear optical response, etc., are indeed determined mainly by local symmetry. As pointed out above, the HT screening is locked in a few basins of attractions and is thus ill suited to explore the several basins of attractions on the BO surface that might correspond to distinct daughter structures. The exploration of the possible daughter structures requires instead the application of an unbiased and unconstrained global minimizer to search for the lowest-energy minimum on the Born-Oppenheimer surface of a solid system. It is interesting to observe that the HT method might miss the correct minimum even in "simple" systems in which there are no unconstrained internal degrees of freedom and that might show multiples total energy minima as a function of the lattice parameter, as in the case of Ce.

The realization that a formal structure type does not uniquely define a local minimum (daughter) structure is also important in the context of the application of machine learning methods to predict the stable structure prototype of a candidate compound out a finite set of candidate formal structure prototypes. Examples of such machine learning predictions of binary and ternary compounds are given by the application of a Bayesian prediction algorithm (see Refs. 41 and 42) and principal component analysis to predict the most likely formal structure type for the stable compounds in binary and ternary systems. Ref. 43 shows the application of probabilistic modeling to predict the likely chemical substitution on a given structure type to generate a stable compound All of these prediction by machine-learned artificial intelligence do not take into account the pos-

sibility of daughter structure types, either because consider only formal structure types or because they focus on a given realization of a formal structure type itself.

Our study instead points to the need of exploring the BO surface to identify possible daughter structure types. In the context of machine learning compound prediction this can be accomplished with machine learning methods in which the candidate structures are represented by descriptors that are sensitive to the chemical environments and local coordination (see, e.g., Ref. 44). Such descriptors could be directly the atom coordinates and lattice vectors, as in the generalized neural networks applied by Behler *et al.* in Refs. 45 and 46 to model the BO energy function and that are functions of the atom coordinates.

VIII. DISCUSSION AND CONCLUSIONS

It is an interesting point to assess whether global optimization algorithms such as evolutionary algorithms and random search could be applied to predict the stable structure of large set of materials. With the current petaflop high-performance computing (HPC) platforms, calculation campaigns using evolutionary method or the random search approach (for instance in its Ab Initio Random Structure Searching (AIRSS)[39] formulation) on 10^2, or even 10^3, compounds are possible. On the exascale HPC facilities that will come on line in the next few years this type of campaigns using global optimization methods several hundreds of compounds will become routine routine and the boundaries can be pushed to perform campaigns on up to 10^4 different compounds. However, we do not think that only applying global optimization methods will be an efficient solution to the crystal structure prediction problem even when exascale HPC facilities will become available. We instead envisage a multi-step approach in which high-throughput screening will be the first stage of the search for stable structure types. These solutions will then be refined by global optimization algorithms, for instance including them as part of the initial population in an evolutionary algorithm search or by applying the random change operators that in AIRSS are used to explore new parts of the Born-Oppenheimer surface.

Based on the present study we propose a multi-step approach to predict new materials in which high-throughput screening is be the first step of the search for stable structure types. High-throughput screening should then be followed by global optimization searches, for instance based on evolutionary algorithms and or random search, to refine the solutions found by screening of a given set of structure types. The global optimization step will allow the survey of parts of the

Born-Oppenheimer surface that would not be otherwise explored only following the structural analogies with known materials on which high-throughput methods are based.

Acknowledgments. Work of MJDV and KRP was supported in part by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract No. DE-AC02-06CH11357. Work of XZ and AZ was supported by Office of Science, Basic Energy Science, MSE division under grant DE-FG02-13ER46959. GT used in this research resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. An award of computer time was provided by the ALCC program. As part of this ALCC award this research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. Single crystal x-ray data were collected at Northwestern University's Integrated Molecular Structure Education and Research Center (IMSERC) at Northwestern University, which is supported by grants from NSF-NSEC, NSF-MRSEC, the KECK Foundation, the State of Illinois, and Northwestern University.

---

* g-trimarchi@northwestern.edu

[1] X. Zhang, L. Yu, A. Zakutayev, and A. Zunger, Advanced Functional Materials **22**, 1425 (2012).

[2] X. Zhang, V. Stevanović, M. d'Avezac, S. Lany, and A. Zunger, Phys. Rev. B **86**, 014109 (2012).

[3] R. Gautier, X. Zhang, L. Hu, L. Yu, Y. Lin, T. O. L. Sunde, D. Chon, K. R. Poeppelmeier, and A. Zunger, Nature Chemistry **7**, 308 (2015).

[4] S. Chadov, X. Qi, J. Kübler, G. H. Fecher, C. Felser, and S. C. Zhang, Nat Mater **9**, 541 (2010).

[5] J. Carrete, W. Li, N. Mingo, S. Wang, and S. Curtarolo, Physical Review X **4**, 011019 (2014).

[6] (), Interestingly, when applying our T=0 protocol to 44 compounds reported in the ICSD (Mn_2SiO_4, Sr_2TiO_4, Al_2ZnS_4, Ba_2TiS_4, Ca_2SiS_4, Sc_2MgSe_4, In_2MgTe_4, BaZnSi, BaZnSn, CaZnGe, AgKO, KCaBi, CuKSe, KMgAs, KMgP, KZnSb, LiAlGe, LiAlSi, LiBeN, CuLiO, LiInGe, LiMgN, LiSrSb, LiYGe, NaAlSi, RbCaAs, SrZnSi, TiPtGe, VCoSi, VCoGe, NbCoSi, NbCoGe, NbCoSn, NbRhSi, NbRhGe, NbRhSn, NbIrSi, NbIrGe, NbIrSn, TaCoGe, TaRhGe, TaCoSi, TaRhSi, and TaIrSi), we find in all cases that we correctly predict their stability and in the correct (observed) structure.

[7] H. Burzlaff and Y. Malinovsky, Acta Cryst. A **53**, 217 (1997).

[8] P. P. Ewald and C. Hermann, *Strukturbericht*, Vol. 1 (1931) pp. 7–12.

[9] S. Curtarolo, D. Morgan, and G. Ceder, Calphad **29**, 163 (2005).

[10] W. Setyawan and S. Curtarolo, Computational Materials Science **49**, 299 (2010).

[11] S. Curtarolo, G. L. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, Nature materials **12**, 191 (2013).

[12] J. Ihm, A. Zunger, and M. L. Cohen, Journal of Physics C: Solid State Physics **12**, 4409 (1979).

[13] S. Liu and J. D. Corbett, Journal of Solid State Chemistry **179**, 830 (2006).

[14] D. G. Schlom, L.-Q. Chen, C. J. Fennie, V. Gopalan, D. A. Muller, X. Pan, R. Ramesh, and R. Uecker, MRS Bulletin **39**, 118 (2014).

[15] A. G. Christy, Acta Crystallographica Section B **51**, 753 (1995).

[16] Inorganic Crystal Structure Database, ICSD, Fachinformationszentrum Karlsruhe, Karlsruhe, Germany, 2015.

[17] E. Parthe', *Crystal Chemistry of Tetrahedral Structures* (Gordon and Breach Science Publishers, New York, 1964).

[18] L. I. Berger, *Semiconductor Materials* (CRC Press, Boca Raton, Florida, 1997).

[19] D. M. Wood, A. Zunger, and R. deGroot, Phys. Rev. B **31**, 2570 (1985).

[20] A. E. Carlsson, A. Zunger, and D. M. Wood, Phys. Rev. B **32**, 1386 (1985).

[21] S.-H. Wei and A. Zunger, Physical Review Letters **56**, 528 (1986).

[22] A. Zakutayev, X. Zhang, A. Nagaraja, L. Yu, S. Lany, T. O. Mason, D. S. Ginley, and A. Zunger, Journal of the American Chemical Society **135**, 10048 (2013).

[23] (), the structure of RbCuS reported in Ref. 25 and 26 was not posted to the ICSD. As of the time of submission of the present work (August 2015), RbCuS and its structure remain unreported in the current release of the ICSD.

[24] W. Losert and R. Hoppe, Zeitschrift für Anorganische und Allgemeine Chemie **524**, 7 (1985).

[25] M. Sing, Dissertation, Universität Linz, Austria (1993).

[26] H. Boller, Journal of Alloys and Compounds **442**, 3 (2007).

[27] D. Sheppard, P. Xiao, W. Chemelewski, D. D. Johnson, and G. Henkelman, J. of Chem. Phys. **136**, 074103 (2012).

[28] D. Wales, *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses* (Cambridge University Press, 2004).

[29] D. J. Wales, The Journal of Chemical Physics **142**, 130901 (2015).

[30] A. R. Oganov and C. W. Glass, The Journal of Chemical Physics **124**, 244704 (2006).

[31] A. R. Oganov, C. W. Glass, and S. Ono, Earth Panet. Sci. Lett. **241**, 95 (2006).

[32] N. L. Abraham and M. I. J. Probert, Phys. Rev. B **73**, 224104 (2006).

[33] G. Trimarchi and A. Zunger, Phys. Rev. B **75**, 104113 (2007).

[34] G. Trimarchi and A. Zunger, J. Phys.: Condens. Matter **20**, 295212 (2008).

[35] Y. Wang, J. Lv, L. Zhu, and Y. Ma, Phys. Rev. B **82**, 094116 (2010).

[36] K. Doll, J. C. Schön, and M. Jansen, Phys. Rev. B **78**, 144110 (2008).

[37] M. Amsler and S. Goedecker, The Journal of Chemical Physics **133**, 224104 (2010).

[38] R. Martoňák, A. Laio, and M. Parrinello, Phys. Rev. Lett. **90**, 075503 (2003).

[39] C. J. Pickard and R. J. Needs, Journal of Physics: Condensed Matter **23**, 053201 (2011).

[40] H. T. Stokes and D. M. Hatch, Journal of Applied Crystallography **38**, 237 (2005).

[41] C. C. Fischer, K. J. Tibbetts, D. Morgan, and G. Ceder, Nature Materials **5**, 641 (2006).

[42] G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, Chem. Mater. **22**, 3762 (2010).

[43] G. Hautier, C. Fischer, V. Ehrlacher, A. Jain, and G. Ceder, Inorganic Chemistry **50**, 656 (2011).

[44] K. T. Schütt, H. Glawe, F. Brockherde, A. Sanna, K. R. Müller, and E. K. U. Gross, Physical Review B **89**, 205118 (2014).

[45] J. Behler and M. Parrinello, Physical Review Letters **98**, 146401 (2007).

[46] J. Behler, Journal of Physics: Condensed Matter **26**, 183001 (2014).

[47] X. Zhang and A. Zunger, Advanced Functional Materials **20**, 1944 (2010).

[48] D. Kieven, A. Grimm, A. Beleanu, C. Blum, J. Schmidt, T. Rissom, I. Lauermann, T. Gruhn, C. Felser, and R. Klenk, Thin Solid Films **519**, 1866 (2011).

[49] W. Bronger and H. Kathage, Journal of the Less Common Metals **160**, 181 (1990).

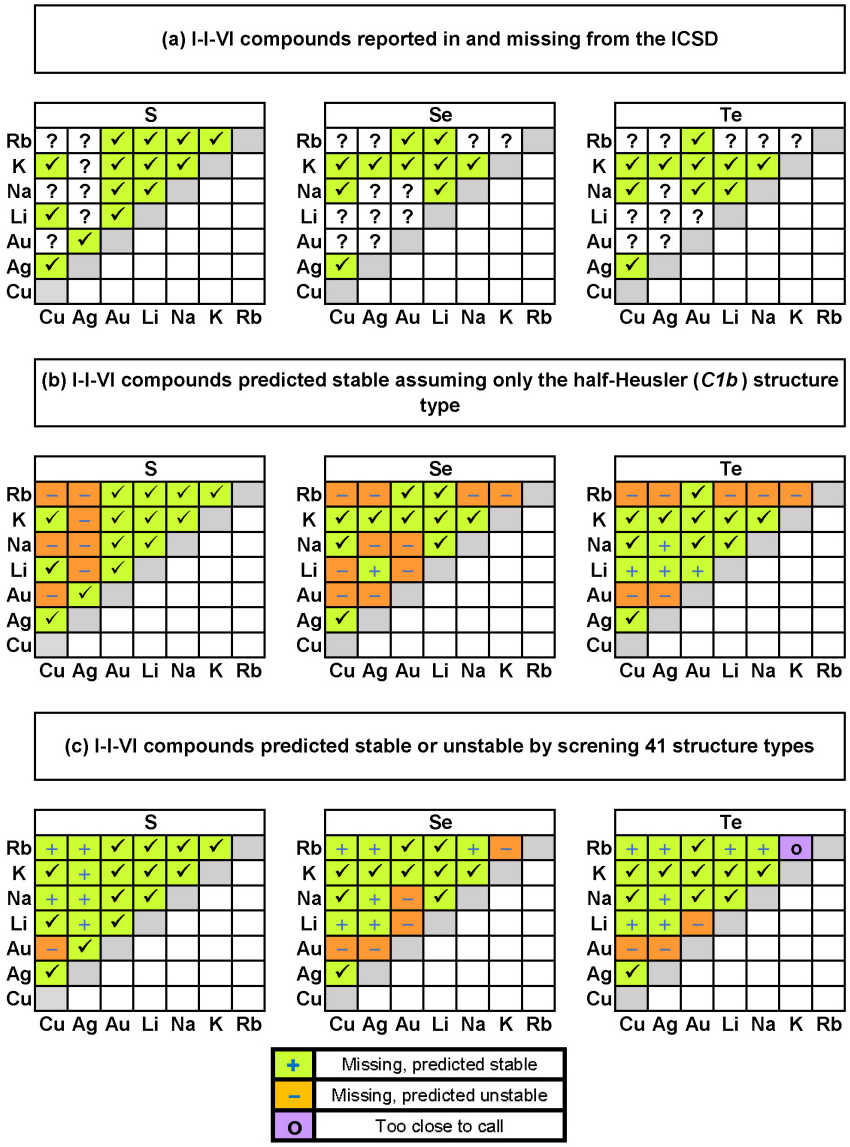

FIG. 1. (a) The I-I-VI ABX compounds that are reported in the ICSD[16] are shown by check marks, while the question marks indicate the compounds in this family that are missing from the ICSD. Note that RbCuS is missing from the current release of the ICSD[23] but was initially reported in Ref. 25 and 26. (b) I-I-VI compounds predicted by GGA formation energy calculations to be stable or unstable assuming only the $C1_b$ structure (see the Supporting Information for details). The "+" and "-" signs indicate missing materials predicted to be, respectively, thermodynamically stable and unstable in the $C1_b$ structure. (c) I-I-VI ABX compounds predicted stable or unstable in Ref. 1. In this work, the lowest energy structure for a given I-I-VI combination was determined by screening 41 ABX structure types. As discussed in Ref. 47, LiCuS[48] and KAgTe[49] were synthesized before, which are not listed in the ICSD database[16], and were initially added as plus signs in Fig. 4 of Ref. 1. The results in Ref. 1 on LiCuS and KAgTe agree with experiments[48,49] and thus are not theoretical predictions.

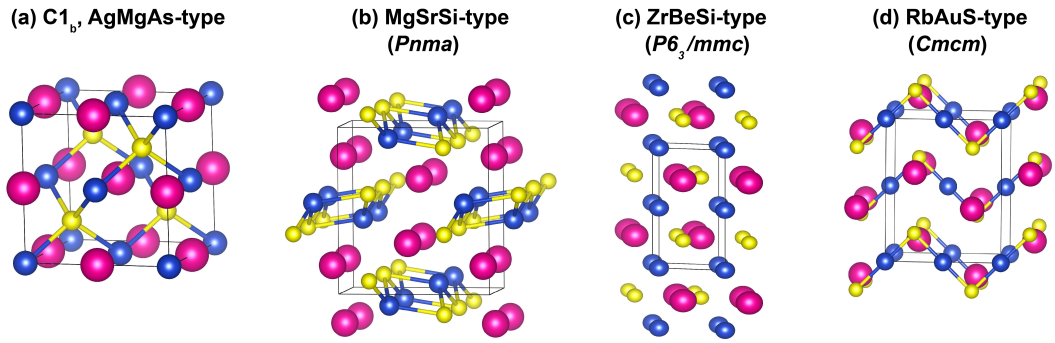

FIG. 2. Crystal structure model of the formal structure types whose total energy was calculated by full local structure relaxation and shown in Figure 3.

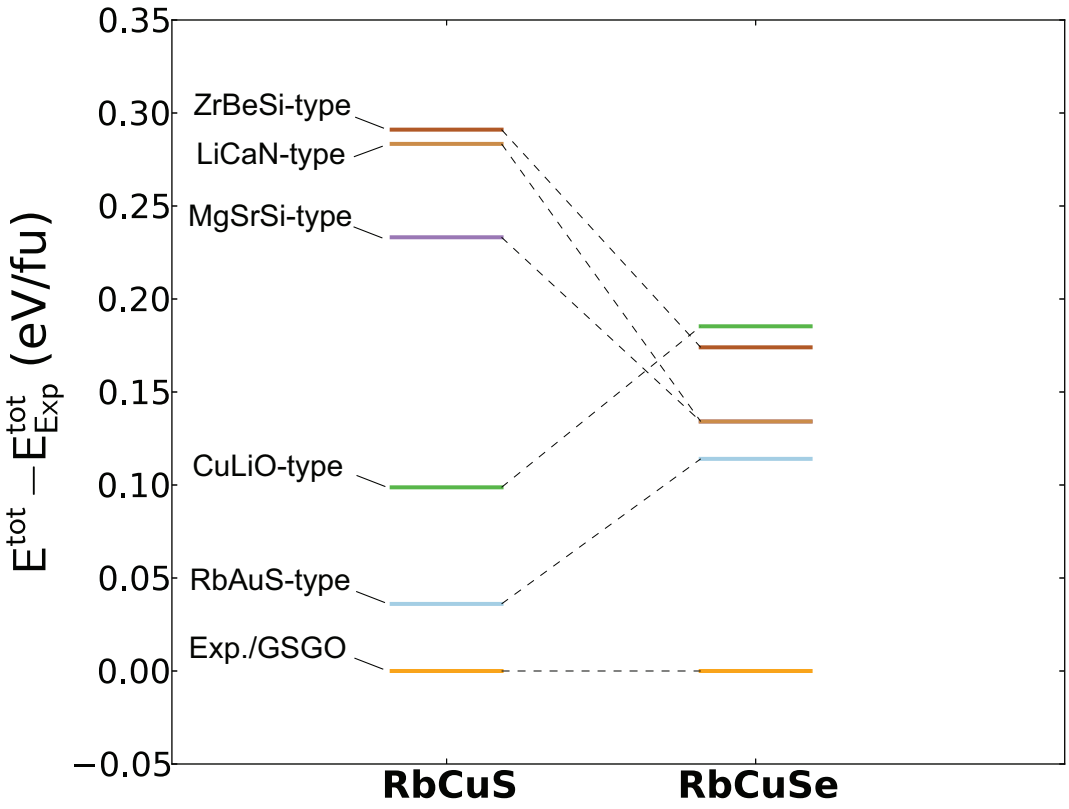

FIG. 3. Results of the high-throughput calculations on RbCuS and RbCuSe performed assuming the pool of structure types predicted for the I-I-VI compounds in Ref. 1. The cell-internal and lattice vector degrees of freedom were fully relaxed to the nearest local total energy minimum. As starting point for each local total energy minimization we took the atom coordinates of the prototype solid associated after which each structure type is named. The total energies are referred to the experimentally observed structures, which for both RbCuS and RbCuSe correspond to the structure predicted by global space-group optimization (GSGO).

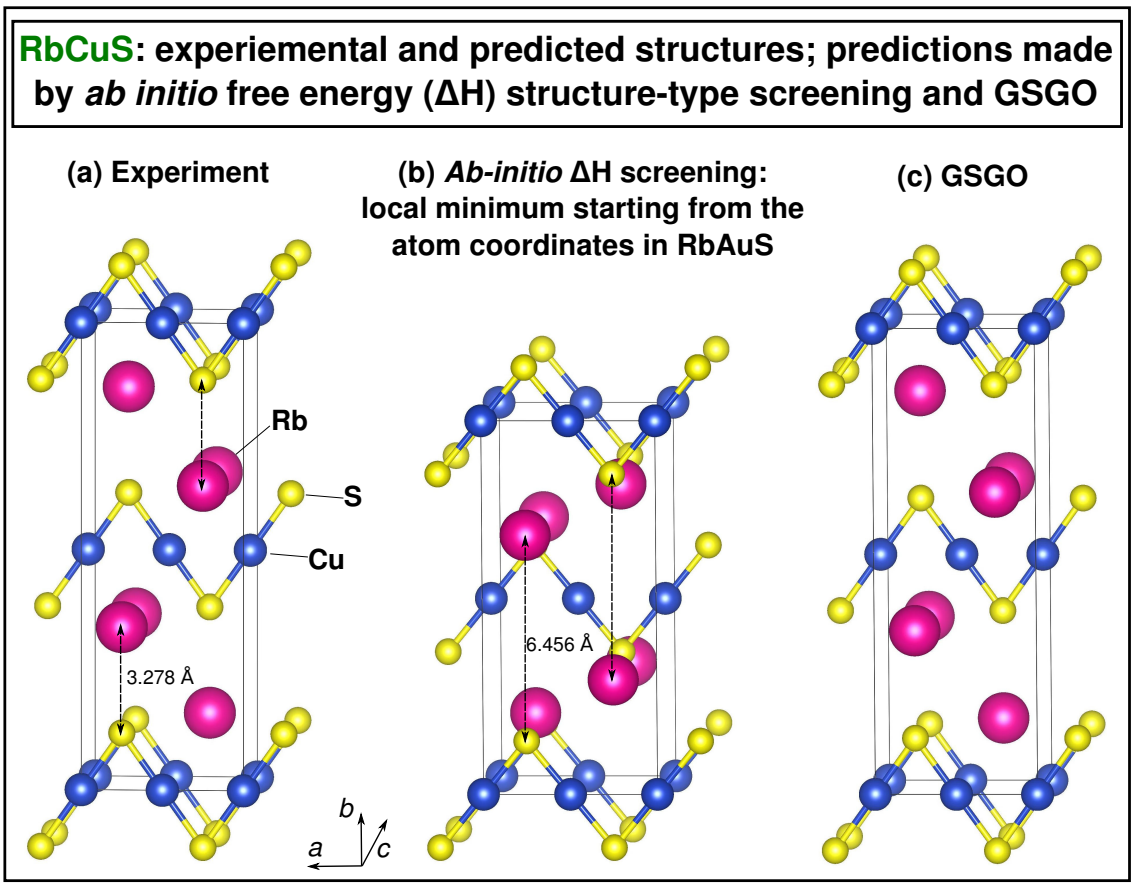

FIG. 4. Three-dimensional models of depicting for RbCuS the (a) experimental structure, (b) the local-minimum structure found starting the structural relaxation from the atom coordinated of RbAuS, and (c) the GSGO structure. The crystallographic parameters for these three structures are reported in the Tables II III.

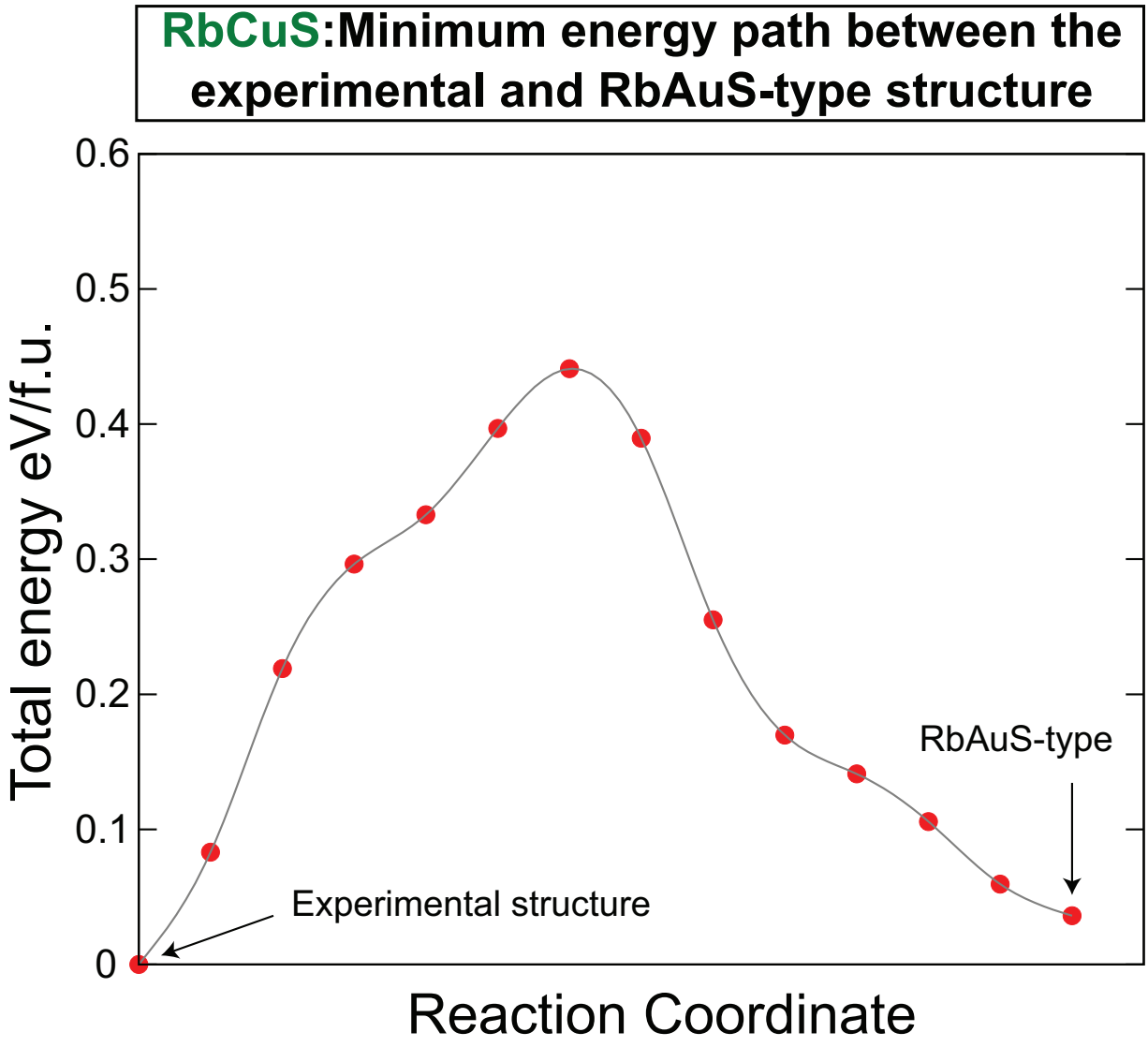

FIG. 5. Minimum energy path connecting the experimental structure of RbCuS and the local minimum energy structure found for this compound by local energy minimization starting from the atom positions in RbAuS and substituting Au with Cu. The zero of the energy is set equal to the total energy of the *ab-initio* relaxed experimental structure of RbCuS.

Selected ABX structure prototypes			
Formal Structure Type (Space Group)	**Species**	**Wyckoff site**	**Coordinate**
$C1_b$	A	$4a$	$(0,0,0)$
($F\bar{4}3m$)	B	$4c$	$(\frac{1}{4},\frac{1}{4},\frac{1}{4})$
	X	$4b$	$(\frac{1}{2},\frac{1}{2},\frac{1}{2})$
RbAuS-type	A	$4c$	$(0,v_{\mathrm{A}},\frac{1}{4})$
($Cmcm$)	B	$4a$	$(0,0,0)$
	X	$4c$	$(0,v_{\mathrm{X}},\frac{1}{4})$
MgSrSi-type	A	$4c$	$(u_{\mathrm{A}},\frac{1}{4},w_{\mathrm{A}})$
($Pnma$)	B	$4c$	$(u_{\mathrm{B}},\frac{1}{4},w_{\mathrm{B}})$
	X	$4c$	$(u_{\mathrm{X}},\frac{1}{4},w_{\mathrm{X}})$

TABLE I. Definition of the formal structure types (FSTs) that have RbAuS and MgSrSi as representative compounds. The values listed for the symmetry-unconstrained fractional coordinates are those experimentally observed for the representative materials and used as starting point for the local energy minimization.

	Single-Crystal X-Ray	Relaxed RbAuS-type structure	GSGO
Space group (No.)	$Cmcm$ (63)	$Cmcm$ (63)	$Cmcm$ (63)
Pearson symbol	$oC12$	$oC12$	$oC12$
a(Å)	4.9007	6.6106	5.1080
b(Å)	14.4778	11.7182	14.5751
c(Å)	5.0067	5.4684	5.0577

TABLE II. Crystallographic parameters, i.e., space group symmetry and lattice parameters, of RbCuS for the following structures: (a) the experimental structure measured by single-crystal X-ray diffraction; (b) the RbAuS-type local minimum structure obtained performing a full relaxation of RbCuS starting from the RbAuS atom positions and lattice vectors and substituting Au with Cu; (c) the structure found by GSGO and shown in Figure 4(c).

	Species	Wyckoff site	x	y	z
Experimental	Rb	$4c$	0	0.65172	$1/4$
	Cu	$4a$	0	0	0
	S	$4c$	0	0.87818	$1/4$
Relaxed RbAuS-type	Rb	$4c$	0	0.30641	$1/4$
	Cu	$4a$	0	0	0
	S	$4c$	0	0.85737	$1/4$
GSGO	Rb	$4c$	0	0.64873	$1/4$
	Cu	$4a$	0	0	0
	S	$4c$	0	0.879340	$1/4$

TABLE III. Atom coordinates of RbCuS for the following structures (see Tab. II for the corresponding lattice parameters): (a) the experimental structure measured by single-crystal X-ray diffraction; (b) the RbAuS-type local minimum structure obtained performing a full relaxation of RbCuS starting from the RbAuS atom positions and lattice vectors and substituting Au with Cu; (c) the structure found by GSGO and shown in Figure 4(c).