Strategic Optimization of the Electronic Transport Properties of Pseudo-Ternary Clathrates

Joakim Brorsson, Anders E. C. Palmqvist, and Paul Erhart*

While alloying is a powerful handle for materials engineering, it is an ongoing challenge to navigate the large and complex parameter space of these materials. This applies in particular for thermoelectrics and even more so clathrates. Here, a combination of density functional theory calculations, alloy cluster expansions, Monte Carlo simulations, and Boltzmann transport theory calculations is used to identify compositions that yield high power factors in the pseudo-ternary clathrates $Ba_8Al_xGa_yGe_{46-x-y}$ and $Ba_8Ga_xGe_ySi_{46-x-y}$, while accounting for weight and raw material costs. The results show how a cost-efficient performance can be achieved by reducing the number of Al and Ga atoms per unit cell, while compensating the resulting increase in the carrier concentration via an extrinsic dopant. The approach used in this study is transferable and can be a useful tool for mapping the thermodynamic and transport properties of other multinary systems.

1. Introduction

Computationally driven strategies have been demonstrated to accelerate the development of thermoelectrics[1–10] as well as other functional materials.[11–13] We have previously shown that atomic scale simulations can be used to estimate the physical properties for pseudo-binary type-I clathrates, in spite of the inherently large configuration space.[14,15] Here, a similar approach is applied to investigate two even more complex cases, namely the pseudo-ternary systems $Ba_8Al_xGa_yGe_{46-x-y}$ and $Ba_8Ga_xGe_ySi_{46-x-y}$. We focus on thermoelectric performance, since these materials exhibit extremely low thermal conductivities (κ) together with relatively high electrical conductivities (σ) and Seebeck coefficients (S).[16–18] In turn, this leads to a reasonably larger thermoelectric figure of merit $zT = \sigma S^2 T/\kappa$, where σS^2 is the power factor (PF).

J. Brorsson, A. E. C. Palmqvist
Department of Chemistry and Chemical Engineering
Chalmers University of Technology
Gothenburg SE-412 96, Sweden

P. Erhart
Department of Physics
Chalmers University of Technology
Gothenburg SE-412 96, Sweden
E-mail: erhart@chalmers.se

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aelm.202100756.

Based on our previous studies of $Ba_8Ga_{16}Ge_{30}$[14] and $Ba_8Al_xGa_{16-x}Ge30$,[19] we know that the electronic transport properties are strongly influenced by chemical ordering, which is a direct consequence of the unique structural motifs of clathrates. They feature a rigid host framework that contains cages which enclose loosely bound guest atoms. Like most studies, our investigation pertains to structures that belong to the type-I subgroup, which possesses $Pm\bar{3}n$ (International Tables of Crystallography number 223) symmetry and can be described by the chemical formula $A_8B_xC_{46-x}$. Here, A represents the guest atoms, which are found at Wyckoff sites $2a$ and $6d$, while the host atoms, B and C, occupy $6c$, $16i$, and $24k$ sites (see **Figure 1**a. According to the Zintl concept, it is expected that the guest atoms are not directly bonded to the host framework, but rather donate their electrons to the latter so as to ensure that all atoms have a full octet in their valence shell. In $Ba_8Al_xGa_{16-x}Ge30$ and $Ba_8Ga_{16}Ge_xSi_{16-x}$, the 8 pairs of electrons provided by the Ba atoms (+2) are compensated by the relative deficiency (−1) of the 16 trivalent elements (Al, Ga). While both compounds should thus be expected to behave as intrinsic semiconductors, actual samples tend to deviate from the stoichiometric composition, which leads to either a lack (p-type) or surplus (n-type) of electrons.

The Zintl concept alone is, however, not enough to explain the differences in the property measurements that are generally observed, even for samples that have the same nominal composition.[20–33] Detailed structural characterisations of different pseudo-binary clathrates have, moreover, revealed that different synthesis methods produce materials with remarkably different site occupation factors (SOFs).[16,34] In addition, our previous studies of the pseudo-binaries, $Ba_8Ga_xGe_{46-x}$, $Ba_8Al_xGe_{46-x}$, $Ba_8Ga_xSi_{46-x}$, and $Ba_8Al_xSi_{46-x}$[14,15] as well as $Ba_8Al_xGa_{16-x}Ge_{30}$[35] have revealed the existence of significant correlations between the electronic transport properties and the degree of chemical ordering. The fact that the latter varies strongly with both temperature and composition, indicates that there exist multiple levers that can be used to tune the thermoelectric performance, especially for pseudo-ternary systems.

It should be emphasized that while the figure of merit of $Ba_8Ga_xGe_{46-x}$ can reach values above 1,[36] this compound is among the most expensive of all known thermoelectric materials,[37,38] which severely limits its viability for practical applications. Since the market prices for Al and Si ($\approx 2\,kg^{-1}$)[39,40] are substantially lower than Ga ($\approx 200\,kg^{-1}$) and Ge ($\approx 1000\,kg^{-1}$),[40]
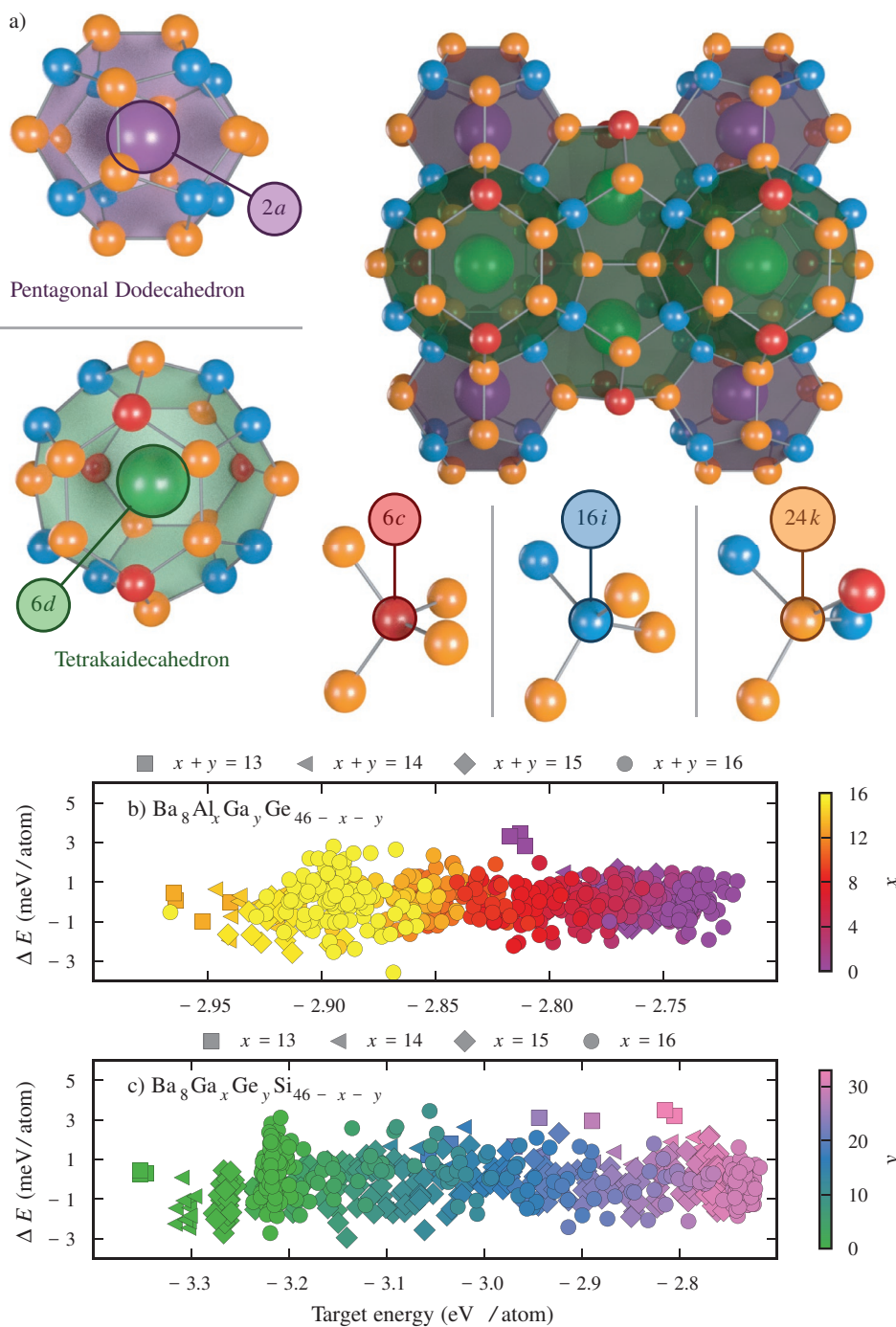
Figure 1. a) Crystal structure of a type-I clathrate. 6c (red), 16i (blue), and 24k (orange) host sites as well as the 2a (purple) and 6d (green) guest sites can be distinguished. (b,c) Prediction errors associated with the b) $Ba_8Al_xGa_yGe_{46-x-y}$ and c) $Ba_8Ga_xGe_ySi_{46-x-y}$ CE models.

a potential solution could be to alloy $Ba_8Ga_xGe_{46-x}$ with $Ba_8Al_xGe_{46-x}$ or $Ba_8Ga_xSi_{46-x}$. Another key issue that has been revealed in recent years is that $Ba_8Ga_xGe_{46-x}$ has a tendency to decompose, both in air and vacuum, if subjected to high temperatures for prolonged periods of time.[41] In this context, the addition of a third framework species could provide entropic stabilization, which is the basic principle behind high-entropy alloys.[42–44]

Given these considerations, here, we present a systematic computational study of the thermoelectric PF as a function of composition in $Ba_8Al_xGa_yGe_{46-x-y}$ and $Ba_8Ga_xGe_ySi_{46-x-y}$ in relation to the cost of these compounds. This is accomplished via a combination of density functional theory (DFT) calculations, Monte Carlo (MC) simulations based on cluster expansion (CE) models and Boltzmann transport theory (BTT) calculations. In the next section, we first summarize the CE models and then show evidence for chemical ordering, followed by a detailed discussion of the calculated PFs. Technical information related to the methodology can be found in the supplementary

information, which includes details regarding the DFT calculations (Note S1, Supporting Information),[45–49] CE construction (Note S2, Supporting Information)[50,51] and sampling via MC simulations, (Note S3, Supporting Information)[52–54] as well as the BTT calculations (see Note S4, Supporting Information and Figure S1, Supporting Information).[55,56]

2. Results and Discussion

2.1. Performance of the Cluster Expansions

Based on a thorough testing of three alternative fitting methods, we decided to use recursive feature elimination (RFE) with ordinary least squares (OLS) to train the CEs, since it results in low cross-validation (CV) root-mean-square error (RMSE) scores for relatively few non-zero parameters (Figure S2, Supporting Information). The final $Ba_8Al_xGa_yGe_{46-x-y}$ (Ba$_8Ga_xGe_ySi_{46-x-y}$) model gave an error of only 1.49 meV site^{-1} (0.54 meV site^{-1}), when 90% of the available structures were used for training, even though it included just 23 (29) pairs as well as 5 (6) triplets (Figure S3, Supporting Information). As a further measure of the performance of the CEs we calculated the predictive errors for all target energies, which fall between ±2 meV atom^{-1}. Both the $Ba_8Al_xGa_yGe_{46-x-y}$ (Figure 1b and $Ba_8Ga_xGe_ySi_{46-x-y}$ (Figure 1c CEs are thus comparable in performance to those that we have previously reported for $Ba_8Ga_xGe_{46-x}$, $Ba_8Ga_xSi_{46-x}$, $Ba_8Al_xGe_{46-x}$, and $Ba_8Al_xSi_{46-x}$.[15] To further validate the models, we compared the SOFs obtained from sampling CEs ensembles, for both pseudo-ternary and pseudo-binary clathrates, with experimental data[34,35,57–67] (see Note S5, Supporting Information and Figure S4, Supporting Information). Overall, good agreement is obtained with respect to the measured data as well as between the pseudo-ternary and pseudo-binary CEs.

2.2. Chemical Ordering

By performing MC simulations based on the CEs, we mapped out the SOFs at 700 K across a large portion of the pseudo-ternary compositional space. To this end, we employed a hybrid method, which involves randomly switching between a canonical and a variance constrained semi-grand canonical (VCSGC) ensemble, to sample all $Ba_8Al_xGa_yGe_{46-x-y}$ (**Figure 2**) and $Ba_8Ga_xGe_ySi_{46-x-y}$ (Figure S5, Supporting Information) configurations with between 6 and 20 trivalent elements (Al and Ga). This approach enabled us to sample a continuous range of compositions while at the same time ensuring that the end points are included (Figure S6, Supporting Information).

The SOF data show that Al has a much stronger tendency to occupy the 6c site, and to a lesser extent 24k, compared to Ga, which leads to a pronounced asymmetry in the $Ba_8Al_xGa_yGe_{46-x-y}$ system. Such clear trends are not observed in $Ba_8Ga_xGe_ySi_{46-x-y}$, although there are slight preferences for Ge and Si to occupy 24k and 16i sites, respectively. These observations can be partially explained by the differences between effective cluster interactions (ECI) singlets, which in the case of $Ba_8Al_xGa_yGe_{46-x-y}$ (Figure S3, Supporting Information a–c) are more positive for Ga compared to Al and only negative for Ge. For $Ba_8Ga_xGe_ySi_{46-x-y}$ (Figure S3, Supporting Information f–h), meanwhile, Si provides a larger negative contribution than Ge while the term associated with Ga is positive. As is verified by counting the number of nearest neighbors (Figure S7, Supporting Information), we can also conclude that the strong preference for the 6c site exhibited by Al confirms that
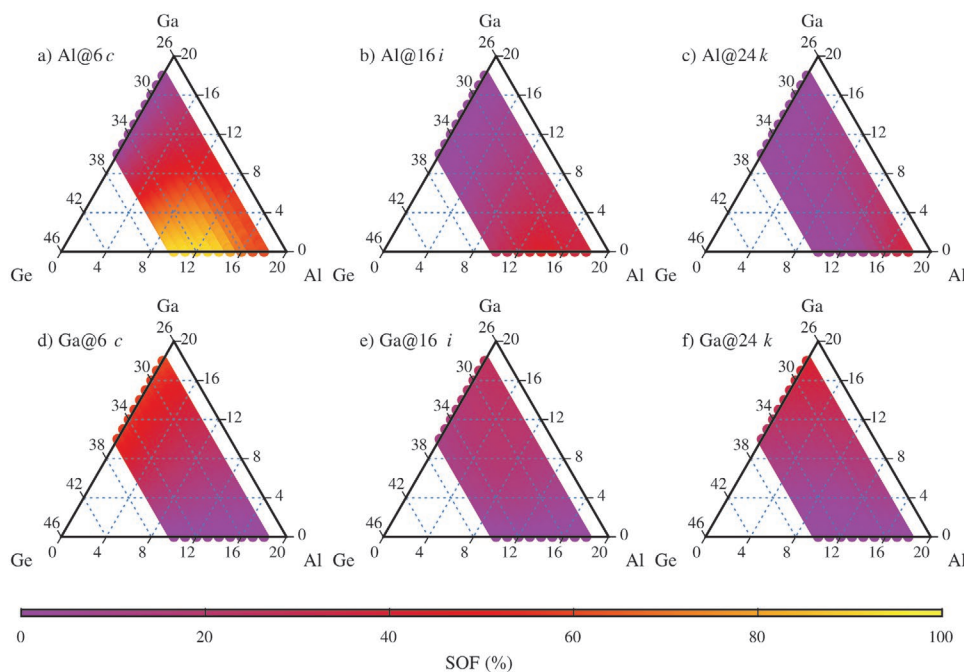
Figure 2. Maps of a–c) Al and d–f) Ga SOFs at the 6c (a, d), 16i (b, e), and 24k (c, f) sites for $Ba_8Al_xGa_yGe_{46-x-y}$, obtained from MC simulations at 700 K.

Al–Al and Al–Ga bonds are less favorable than Ga–Ga bonds as expected.[16]

2.3. Mapping of the Power Factor

While the previously described chemical ordering is interesting in its own right, and a key to understanding the complex chemistry of clathrate compounds, other physical properties are of interest depending on the intended application. Here, we focus on the thermoelectric performance and in particular the PF, which is the product of the electrical conductivity and the square of the Seebeck coefficient (S^q). We consider the PF rather than the thermoelectric figure-of-merit zT since the latter involves the lattice thermal conductivity. The lattice thermal conductivity is intrinsically low in these compounds[18] and is dominated by phonon-phonon scattering at the temperatures of interest, which means that further reductions are not only difficult to achieve, but also less effective.[68] As the calculation of the lattice thermal conductivity is also computationally orders of magnitude more expensive, we therefore limit ourselves to the PF.

Pseudo-ternary systems, such as $Ba_8Al_xGa_yGe_{46-x-y}$ and $Ba_8Ga_xGe_ySi_{46-x-y}$, are especially interesting with regard to tuning the electric transport properties since they offer the possibility to affect the band structure via the Al/Ga and Si/Ge ratios, respectively, while maintaining nearly the same carrier concentration, which is achieved by fixing the number of trivalent elements (Al and Ga). Our simulations are designed to determine if it is possible to retain or even enhance the PF when substituting Ga and Ge with more earth abundant and inexpensive Al and Si.

As a reference we have chosen the single crystal sample (disk 15) produced by Saramat et al.,[36] which displays a zT of 1.24 at 1000 K. Since this is one of the highest values reported for any clathrate, it seems likely that the PF must be close to the optimum. For this reason, and because one cannot expect perfect agreement between calculations and measurements, we deem all compositions that correspond to a PF higher than 80% to be of possible interest. Moreover, it is assumed that carrier concentration and relaxation time, which is assumed to be given by the expression $\tau_{eff} = \tau_{300K}(300K/T)^a$, are the same for all compositions (see ref. [14] for details). Specifically, the corresponding parameters were selected to achieve agreement with the experimental data reported by Saramat et al.[36] for a single crystalline sample (disk 15) of $Ba_8Ga_{16}Ge_{30}$ (see Note S4, Supporting Information and Figure S1, Supporting Information). In practice, this means that some type of extrinsic dopant would need to be introduced to compensate for the fact that deviations from the stoichiometric condition, that is 16 trivalent elements, lead to electron excess (Al/Ga < 16) or deficiency (Al/Ga > 16).
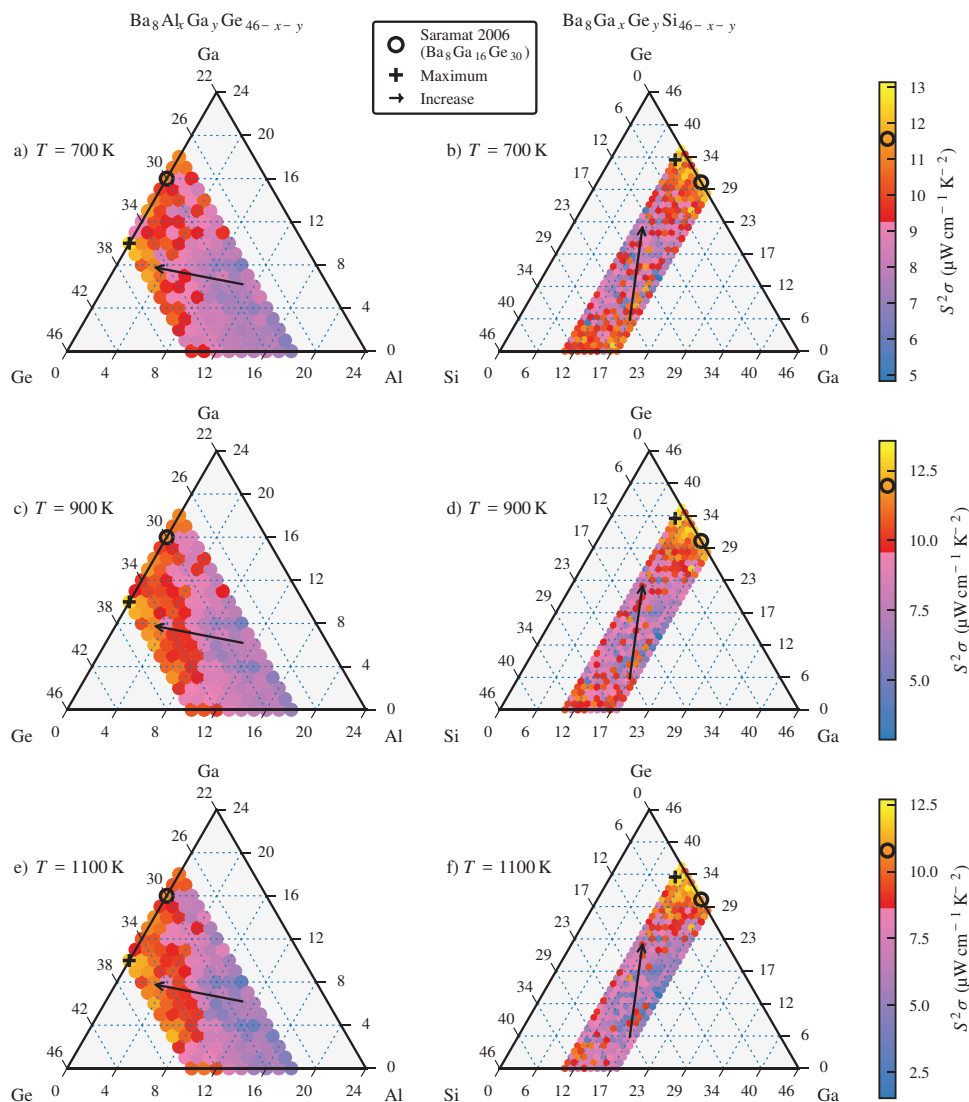
Our calculations show that a reduction in the number of Al and Ga atoms generally results in an increase in the PF (**Figure 3**). Moreover, it is clear that this trend is persistent even when the temperature changes, which means that in further studies it can be sufficient to consider a handful or even just a single temperature. Though mixing of $Ba_8Ga_xGe_{46-x}$ with either $Ba_8Al_xGe_{46-x}$ (Figure 3a,c,e or $Ba_8Ga_xSi_{46-x}$ (Figure 3b,d,f is in many cases detrimental for the performance, it is also evident that an enhancement can in principle be achieved. In this context, one should keep in mind that the experimental meas-urements reported so far clearly show that $Ba_8Ga_{16}Ge_{30}$, and in particular the sample selected as a reference,[36] tends to have superior properties compared to $Ba_8Ga_xSi_{46-x}$, $Ba_8Al_xGe_{46-x}$, and $Ba_8Al_xSi_{46-x}$.[18] This trend is the most obvious for $Ba_8Al_xGa_yGe_{46-x-y}$ (Figure 3a,c,e since the compositions close to $Ba_8Al_{16}Ge_{30}$ and $Ba_8Ga_{10}Ge_{36}$ exhibit the lowest and highest PFs, respectively. A possible explanation for this trend is the strong preference of Al for 6c sites (Figure 2) together with the fact that the PF of $Ba_8Ga_{16}Ge_{30}$, as we have previously shown,[14] is maximal for a hypothetical structure with no tri-valent (Ga) atoms on 6c sites. Still, a performance similar to that of $Ba_8Ga_{16}Ge_{30}$ is achieved for compositions with up to 8 Al per unit cell, indicating that it might be worthwhile to investigate at least this region in more detail. This especially applies to $Ba_8Al_xGa_yGe_{46-x-y}$ with $x \leqslant 4$ and $x + y \leqslant 11$. The situation is somewhat different for $Ba_8Ga_xGe_ySi_{46-x-y}$ because the minimum appears in the middle region, that is when $Ba_8Ga_xGe_{46-x}$ and $Ba_8Ga_xSi_{46-x}$ are mixed in equal proportions. Yet there exist multiple compositions across the entire range for which the PF is similar to or even higher than $Ba_8Ga_{16}Ge_{30}$, in particular at 700 K. More precisely, it would be interesting to explore $Ba_8Ga_xGe_ySi_{46-x-y}$ with $x + y < 16$ and either $x \geqslant 28$ or $x \leqslant 2$.

We have also performed more detailed test calculations for various $Ba_8Al_xGa_{16-x}Ge_{30}$ compositions ($0 \leqslant x \leqslant 16$) as a way of assessing the validity of the earlier observations (see Figure S8, Supporting Information and Figure S9, Supporting Information). These were performed using a previously reported procedure[35] and not only account for the existence of an order-disorder transition but also include a correction of the band gap. The results show that the structural transformation has a greater impact for some compositions than others and, especially for a higher Al content, can lead to a reduction in the PF at temperatures above the transition ($\approx 650K$. Another interesting conclusion is that even though the PF for the compositions with 4 and 8 Al atoms would be lower than for $Ba_8Ga_{16}Ge_{30}$ for the same carrier concentration, the values become equally high if the doping level is slightly adjusted. This is a key result since we do not account for this possibility when performing the screening calculations, which encompass a much wider composition range. One should also note that the use of a different reference sample, in the form of a different sample (disk 18) extracted from the same single crystal,[36] has a significant impact on the predictions (see Figure S8, Supporting Information and Figure S9, Supporting Information). It is evident that the carrier concentration is almost optimal for disk 18 (Figure S9, Supporting Information) but not disk 15 (Figure S8, Supporting Information). Moreover, the PF is generally larger for the high temperature configuration than the ground state above the transition ($\approx 650K$) for the former sample while the opposite is true for the latter. These results show that it is indeed wise to not be overly restrictive when analyzing the transport calculations. As was explained earlier, we accounted for this fact by introducing a threshold, which has here been set to 75% of the value for the reference, and treated all compositions with a higher or equal value to be of potential interest.

An important aspect to consider in this context is that the replacement of Ga with Al or Ge with Si confers additional benefits, mainly related to the fact that Al and Si are both lighter, significantly more abundant, and cheaper than Ga and Ge. For this reason we also examined the PF after weighting
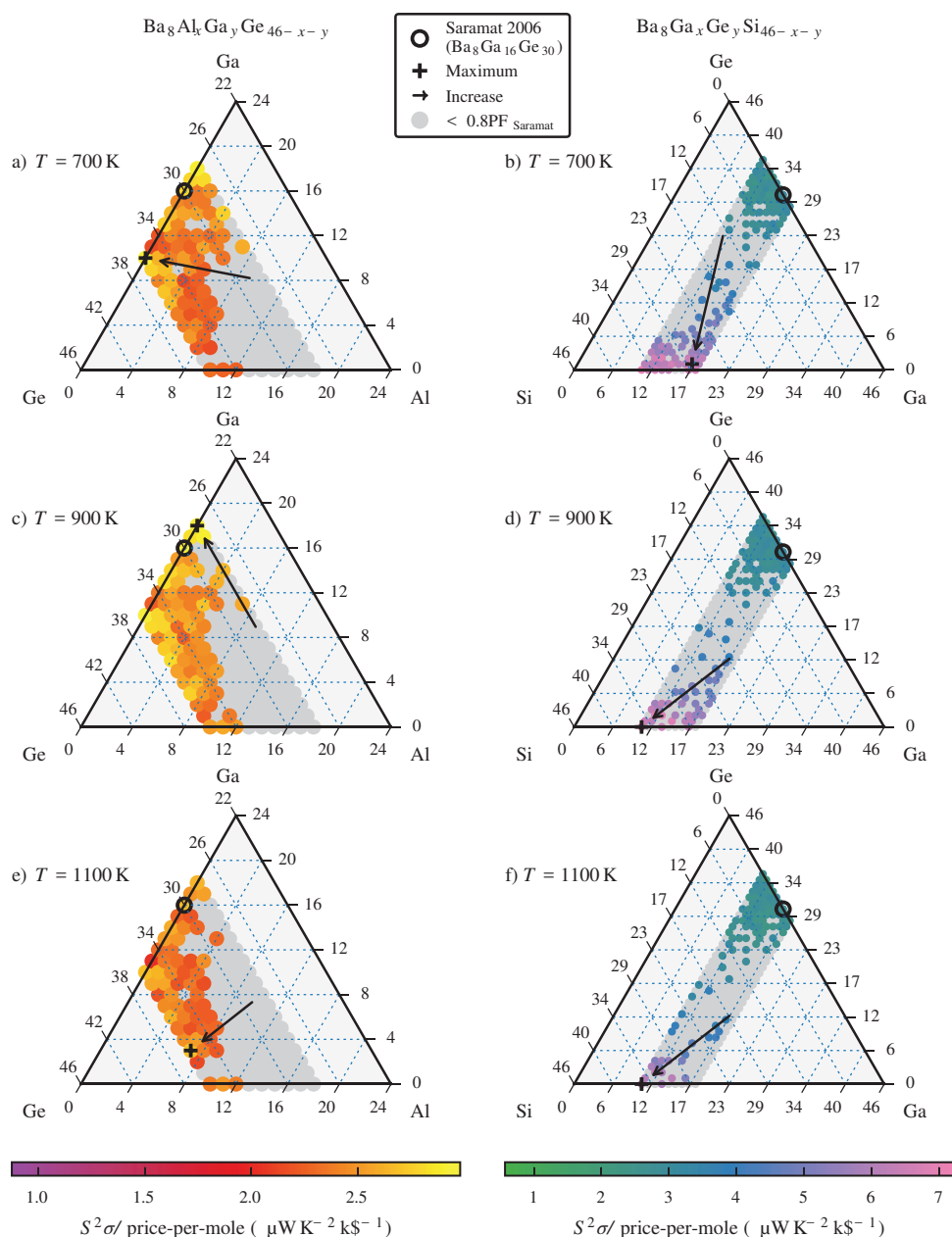
**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
ELECTRONIC
MATERIALS**

www.advelectronicmat.de

Figure 3. Mapping of the PF (σS^2) for a,c,e) $Ba_8Al_xGa_yGe_{46-x-y}$ and b,d,f) $Ba_8Ga_xGe_ySi_{46-x-y}$ at 700 K (a, b), 900 K (c, d), and 1100 K (e, f). The compositions yielding the maximum PFs (plus sign) as well as the composition of the reference ($Ba_8Ga_{16}Ge_{30}$, circle) are indicated. The color range for each temperature has been chosen so that the sharp shift (from pink to red) corresponds to 80% of the value predicted for $Ba_8Ga_{16}Ge_{30}$.

it by mass (Figure S10, Supporting Information) and cost (**Figure 4**) per mole. As should be expected, based on the large difference in mass and market price[69] for Al and Si compared with Ga and Ge (see Table S1, Supporting Information and Figure S11, Supporting Information), this shifts the location of the optimum composition, especially in the case of $Ba_8Ga_xGe_ySi_{46-x-y}$. This type of analysis further emphasizes that $Ba_8Ga_xSi_{46-x}$ is expected to be a better alternative than $Ba_8Ga_xGe_{46-x}$ for commercial applications. The effect is smaller for $Ba_8Al_xGa_yGe_{46-x-y}$, because Ge is even more expensive than Ga and, additionally, the number of trivalent elements is limited to about 18 atoms per unit cell or lower, which only corresponds to one third of the total number of atoms (54). One should note, however, that this is primarily true at 900 K and 1100 k. At 700 K, inclusion of the cost reduces the variations between the different compositions, which makes it more viable to substitute Ga in $Ba_8Ga_xGe_{46-x}$ with Al. It is crucial to emphasize that the

trends outlined above are more pronounced when considering mass (rather than cost), which further enhances the viability of compositions closer to $Ba_8Al_xGe_{46-x}$ and $Ba_8Ga_xSi_{46-x}$.

In summary, our calculations show that a potential strategy to achieve a high thermoelectric performance for these clathrate systems, is to reduce the number of trivalent elements while at the same time introducing a hole dopant, to compensate for the intrinsic n-doping. Possibly, this could be achieved via a variant of modulation doping, which we have shown to be applicable for the $Ba_8Al_xGa_yGe_{46-x-y}$ system in its original form,[70] where the charge carriers are transferred from the matrix to the doped phase instead of vice versa, which is the conventional idea behind this strategy. Another option is to reduce the number of electrons donated by the guest atoms by partially substituting Ba with an alkali metal such as K[71] or Rb. It is also important to note that our calculations of the mixing free energies indicate that it should be possible to mix

ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

ADVANCED
ELECTRONIC
MATERIALS
www.advelectronicmat.de

Figure 4. Mapping of the PF (σS^2) weighted by material cost (see Table S1, Supporting Information)[69] for a,c,e) $Ba_8Al_xGa_yGe_{46-x-y}$ and b,d,f) $Ba_8Ga_xGe_ySi_{46-x-y}$ at 700 k (a, b), 900 k (c, d), and 1100 k (e, f). All compositions that give a power factor lower than 80 % of the value for the reference sample have been uniformly colored (gray). The maximum value (plus) and the reference composition (circle) have also been indicated.

$Ba_8Ga_xGe_{46-x}$ with $Ba_8Ga_xSi_{46-x}$ or $Ba_8Al_xGe_{46-x}$ in any ratio, provided that the corresponding binaries can be synthesized. The contribution is in fact large enough that the free energy is significantly reduced, suggesting that the mixed system might be less prone to decomposition. Even so, one must keep in mind that all possible phases need to be considered in order to determine thermodynamic stability. Yet there exists substantial experimental evidence that it is indeed possible to produce $Ba_8Al_xGa_yGe_{46-x-y}$[35,66,70,72] and $Ba_8Ga_xGe_ySi_{46-x-y}$[64,73,74] for a wide range of different compositions using a variety of methods. Though these studies have been mainly focused on almost stoichiometric compositions, samples of the corresponding

binary compounds $Ba_8Ga_{16}Ge_{30}$[20,75,76], $Ba_8Ga_{16}Si_{30}$,[30,67] and $Ba8Al16Ge30$[23,34,77] have been successfully synthesized that contain either more (only $Ba_8Ga_{16}Ge_{30}$ and $Ba_8Al_{16}Ge_{30}$) or significantly less than 16 trivalent elements per unit cell. Consequently, it seems probable that significant portions of the compositional ranges explored in this study could be investigated experimentally.

3. Conclusions

In this study, we have used an efficient computational method to map the PFs and SOFs across a large portion of

the configuration space of two pseudo-ternary clathrates, $Ba_8Al_xGa_yGe_{46-x-y}$ and $Ba_8Ga_xGe_ySi_{46-x-y}$. These results show that the presence of Al in $Ba_8Al_xGa_yGe_{46-x-y}$ leads to a pronounced asymmetry in the SOFs, mainly because it has a much stronger preference for the $6c$ site compared to Ga, while the variations for Ge and Si in $Ba_8Ga_xGe_ySi_{46-x-y}$ are almost identical. In agreement with experimental measurements reported so far, our calculations indicate that while improving the electronic transport properties of $Ba_8Ga_{16}Ge_{30}$ is not an easy task, there exist multiple $Ba_8Al_xGa_yGe_{46-x-y}$ and $Ba_8Ga_xGe_ySi_{46-x-y}$ compositions that are predicted to exhibit a similar, and in a few cases even higher, performance. There are, moreover, indications that this applies in particular if the carrier concentration is optimized for each configuration. Even so, the general tendency is that the substitution of Ga with Al or Ge with Si decreases the PF. Precise control over the stoichiometry and extensive testing is, therefore, likely to be required in order to achieve an enhancement. The situation changes, however, when material weight and cost are taken into account. This is especially apparent in the case of $Ba_8Ga_xGe_ySi_{46-x-y}$, since compositions close to $Ba_8Ga_xSi_{46-x}$ are potentially more attractive than $Ba_8Ga_xGe_{46-x}$. Though this effect is less pronounced for $Ba_8Al_xGa_yGe_{46-x-y}$, it is deemed worthwhile to consider compositions with up to 8 Al atoms per unit cell. For both $Ba_8Al_xGa_yGe_{46-x-y}$ and $Ba_8Ga_xGe_ySi_{46-x-y}$, however, the maximal PF is achieved by reducing the number of trivalent elements (Al and Ga). Though it might prove difficult to reach the lower limit of the interval examined in this study (Al/Ga = 10), it should be possible to introduce an extrinsic dopant while at the same time decreasing the starting amount of Al and Ga when synthesizing these materials. Lastly, we want to stress that the procedure presented in this paper can, in principle, be applied to any multinary system that displays some degree of chemical ordering. Thus, it can serve as a guide for determining what range of compositions might be the most promising for a certain type of application, which need not be limited to thermoelectrics.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

[1] M. W. Gaultois, T. D. Sparks, C. K. H. Borg, R. Seshadri, W. D. Bonificio, D. R. Clarke, *Chem. Mater.* **2013**, *25*, 2911.
[2] J. Yang, L. Xi, W. Qiu, L. Wu, X. Shi, L. Chen, J. Yang, W. Zhang, C. Uher, D. J. Singh, *npj Comput. Mater.* **2016**, *2*, 15015.
[3] W. Chen, J.-H. Pöhls, G. Hautier, D. Broberg, S. Bajaj, U. Aydemir, Z. M. Gibbs, H. Zhu, M. Asta, G. J. Snyder, B. Meredig, M. A. White, K. Persson, A. Jain, *J. Mater. Chem. C* **2016**, *4*, 4414.
[4] P. Gorai, V. Stevanović, E. S. Toberer, *Nat. Rev. Mater.* **2017**, *2*, 17053.
[5] Q. Hao, D. Xu, H. Zhao, *ECS J. Solid State Sci. Technol.* **2017**, *6*, N3095.
[6] P. Ren, Y. Liu, J. He, T. Lv, J. Gao, G. Xu, *Inorg. Chem. Front.* **2018**, *5*, 2380.
[7] S. Hao, V. P. Dravid, M. G. Kanatzidis, C. Wolverton, *npj Comput. Mater.* **2019**, *5*, 58.
[8] D. Beretta, N. Neophytou, J. M. Hodges, M. G. Kanatzidis, D. Narducci, M. Martin- Gonzalez, M. Beekman, B. Balke, G. Cerretti, W. Tremel, A. Zevalkink, A. I. Hofmann, C. Möller, B. Dörling, M. Campoy-Quiles, M. Caironi, *Mater. Sci. Eng.: R: Rep.* **2019**, *138*, 100501.
[9] L. Fernandes, E. Fernández, P. Martins, N. Ferreira, P. Antunes, S. Lanceros-Mendez, *J. Mater. Sci.* **2020**, *55*, 925.
[10] J. Recatala-Gomez, A. Suwardi, I. Nandhakumar, A. Abutaha, K. Hippalgaonkar, *ACS Appl. Energy Mater.* **2020**, *3*, 2240.
[11] L. Ward, C. Wolverton, *Curr. Opin. Solid State Mater. Sci.* **2017**, *21*, 167.
[12] E. George, D. Raabe, R. Ritchie, *Nat. Rev. Mater.* **2019**, *4*, 515.
[13] R. R. Schnepf, J. J. Cordell, M. B. Tellekamp, C. L. Melamed, A. L. Greenaway, A. Mis, G. L. Brennecka, S. Christensen, G. J. Tucker, E. S. Toberer, S. Lany, A. C. Tamboli, *ACS Energy Lett.* **2020**, *5*, 2027.
[14] M. Ångqvist, D. O. Lindroth, P. Erhart, *Chem. Mater.* **2016**, *28*, 6877.
[15] M. Ångqvist, P. Erhart, *Chem. Mater.* **2017**, *29*, 7554.
[16] M. Christensen, S. Johnsen, B. B. Iversen, *Dalton Trans.* **2010**, *39*, 978.
[17] T. Takabatake, K. Suekuni, T. Nakayama, E. Kaneshita, *Rev. Modern Phys.* **2014**, *86*, 669.
[18] J.-A. Dolyniuk, B. Owens-Baird, J. Wang, J. V. Zaikina, K. Kovnir, *Mater. Sci. Eng.: R: Rep.* **2016**, *108*, 1.
[19] Y. Zhang, J. Brorsson, T. Kamiyama, T. Saito, P. Erhart, A. E. C. Palmqvist, **2021**.
[20] H. Anno, M. Hokazono, M. Kawamura, J. Nagao, K. Matsubara, in *Proc. ICT '02. Twenty-First Int. Conf. on Thermoelectrics, 2002*, **2002**, p. 77.
[21] J. D. Bryan, N. P. Blake, H. Metiu, G. D. Stucky, B. B. Iversen, R. D. Poulsen, A. Bentien, *J. Appl. Phys.* **2002**, *92*, 7281.
[22] M. Avila, K. Suekuni, K. Umeo, T. Takabatake, *Phys. B: Condens. Matter* **2006**, *383*, 124.
[23] T. Uemura, K. Akai, K. Koga, T. Tanaka, H. Kurisu, S. Yamamoto, K. Kishimoto, T. Koyanagi, M. Matsuura, *J. Appl. Phys.* **2008**, *104*, 013702.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
ELECTRONIC
MATERIALS**

www.advelectronicmat.de

[24] J. Martin, H. Wang, G. S. Nolas, *Appl. Phys. Lett.* **2008**, *92*, 222110.

[25] A. F. May, E. S. Toberer, A. Saramat, G. J. Snyder, *Phys. Rev. B* **2009**, *80*, 125205.

[26] S.-K. Deng, X.-F. Tang, R.-S. Tang, *Chin. Phys. B* **2009**, *18*, 3084.

[27] N. Tsujii, J. H. Roudebush, A. Zevalkink, C. A. Cox-Uvarov, G. J. Snyder, S. M. Kauzlarich, *J. Solid State Chem.* **2011**, *184*, 1293.

[28] J. Leszczynski, A. Kolezynski, K. Wojciechowski, *J. Solid State Chem.* **2012**, *193*, 114.

[29] Y. Nagatomo, N. Mugita, Y. Nakakohara, M. Saisho, M. Tajiri, R. Teranishi, S. Munetoh, *J. Phys.: Conf. Series* **2012**, *379*, 012008.

[30] H. Anno, H. Yamada, T. Nakabayashi, M. Hokazono, R. Shirataki, *J. Solid State Chem.* **2012**, *193*, 94.

[31] H. Anno, M. Hokazono, R. Shirataki, Y. Nagami, *J. Mater. Sci.* **2013**, *48*, 2846.

[32] H. Anno, M. Hokazono, R. Shirataki, Y. Nagami, *J. Electron. Mater.* **2013**, *42*, 2326.

[33] L.-H. Wang, L.-S. Chang, *J. Alloys Compounds* **2017**, *722*, 644.

[34] M. Christensen, B. B. Iversen, *Chem. Mater.* **2007**, *19*, 4896.

[35] J. Brorsson, Y. Zhang, A. E. C. Palmqvist, P. Erhart, *Chem. Mater.* **2021**, *33*, 4500.

[36] A. Saramat, G. Svensson, A. E. C. Palmqvist, C. Stiewe, E. Mueller, D. Platzek, S. G. K. Williams, D. M. Rowe, J. D. Bryan, G. D. Stucky, *J. Appl. Phys.* **2006**, *99*, 023708.

[37] G. G. Yadav, J. A. Susoreny, G. Zhang, H. Yang, Y. Wu, *Nanoscale* **2011**, *3*, 3555.

[38] S. LeBlanc, S. K. Yee, M. L. Scullin, C. Dames, K. E. Goodson, *Renewable Sustainable Energy Rev.* **2014**, *32*, 313.

[39] Institut für seltene Erden und Metalle (ISE), Base metals prices in December 2020, https://en.institut-seltene-erden.de/prices-for-base-metals-in-december/ (accessed: May 2021).

[40] Institut für seltene Erden und Metalle (ISE), Prices for non-ferrous metals in December 2020, https://en.institut-seltene-erden.de/strategic-metals-prices-in-december/ (accessed: May 2021).

[41] H. Reardon, A. B. Blichfeld, H. Kasai, H. Yin, E. D. Bojesen, B. B. Iversen, *Phys. Chem. Chem. Phys.* **2017**, *19*, 15734.

[42] J.-W. Yeh, S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, S.-Y. Chang, *Adv. Eng. Mater.* **2004**, *6*, 299.

[43] S. Shafeie, S. Guo, Q. Hu, H. Fahlquist, P. Erhart, A. Palmqvist, *J. Appl. Phys.* **2015**, *118*, 184905.

[44] S. Shafeie, S. Guo, P. Erhart, Q. Hu, A. Palmqvist, *Adv. Mater.* **2019**, *31*, 1805392.

[45] G. Kresse, J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169.

[46] P. E. Blöchl, *Phys. Rev. B* **1994**, *50*, 17953.

[47] G. Kresse, D. Joubert, *Phys. Rev. B* **1999**, *59*, 1758.

[48] K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard, B. I. Lundqvist, *Reports Progress Phys.* **2015**, *78*, 066501.

[49] K. Berland, P. Hyldgaard, *Phys. Review B* **2014**, *89*, 035412.

[50] M. Ångqvist, W. A. Muñoz, J. M. Rahm, E. Fransson, C. Durniak, P. Rozyczko, T. H. Rod, P. Erhart, *Adv. Theory Simul.* **2019**, *2*, 1900015.

[51] E. Fransson, F. Eriksson, P. Erhart, *npj Comput. Mater.* **2020**, *6*, 135.

[52] T. Mueller, G. Ceder, *Phys. Review B* **2009**, *80*, 024103.

[53] D. Foreman-Mackey, D. W. Hogg, D. Lang, J. Goodman, *Publ. Astron. Soci. Pac.* **2013**, *125*, 306.

[54] B. Sadigh, P. Erhart, *Phys. Review B* **2012**, *86*, 134204.

[55] A. van de Walle, P. Tiwary, M. de Jong, D. L. Olmsted, M. Asta, A. Dick, D. Shin, Y. Wang, L.-Q. Chen, Z.-K. Liu, *Calphad* **2013**, *42*, 13.

[56] G. K. Madsen, J. Carrete, M. J. Verstraete, *Comput. Phys. Commun.* **2018**, *231*, 140.

[57] M. Christensen, N. Lock, J. Overgaard, B. B. Iversen, *J. Am. Chem. Soc.* **2006**, *128*, 15657.

[58] B. Eisenmann, H. Schäfer, R. Zagler, *J. Less Common Metals* **1986**, *118*, 43.

[59] M. Christensen, G. J. Snyder, B. Iversen, in *2006 25th Int. Conf. Thermoelectrics.* **2006**, p. 40.

[60] L. Qiu, I. P. Swainson, G. S. Nolas, M. A. White, *Phys. Rev. B* **2004**, *70*, 035208.

[61] M. Bobnar, B. Böhme, M. Wedel, U. Burkhardt, A. Ormeci, Y. Prots, C. Drathen, Y. Liang, H. D. Nguyen, M. Baitinger, Y. Grin, *Dalton Trans.* **2015**, *44*, 12680.

[62] A. Bentien, B. Iversen, J. Bryan, G. Stucky, A. Palmvist, A. Schultz, R. Henning, *J. Appl. Phys.* **2002**, *91*, 5694.

[63] A. Bentien, E. Nishibori, S. Paschen, B. B. Iversen, *Phys. Rev. B* **2005**, *71*, 144107.

[64] M. Christensen, S. Johnsen, M. Søndergaard, J. Overgaard, H. Birkedal, B. B. Iversen, *Chem. Mater.* **2009**, *21*, 122.

[65] J. H. Roudebush, C. de la Cruz, B. C. Chakoumakos, S. M. Kauzlarich, *Inorg. Chem.* **2012**, *51*, 1805.

[66] W. R. Puspita, H. Takeya, T. Mochiku, Y. Ishikawa, S. Lee, S. Torii, M. Hagihala, T. Kamiyama, In *2019 2nd Int. Conf. on Applied Engieering (ICAE)*. IEEE, Piscataway, NJ **2019**, p. 1.

[67] D. Nataraj, J. Nagao, *J. Solid State Chem.* **2004**, *177*, 1905.

[68] D. O. Lindroth, J. Brorsson, E. Fransson, F. Eriksson, A. Palmqvist, P. Erhart, *Phys. Rev. B* **2019**, *100*, 045206.

[69] Institut für seltene Erden und Metalle (ISE), Prices for high purity metals in December 2020, https://en.institut-seltene-erden.de/prices-for-high-purity-metals-in-december-2020/ (accessed: May 2019).

[70] Y. Zhang, J. Brorsson, R. Qiu, A. E. C. Palmqvist, *Adv. Electron. Mater.* **2020**, *7*, 2000782.

[71] F. Sui, S. M. Kauzlarich, *Chem. Mater.* **2016**, *28*, 3099.

[72] S. Deng, X. Tang, P. Li, Q. Zhang, *J. Appl. Phys.* **2008**, *103*, 073503.

[73] J. Martin, S. Erickson, G. S. Nolas, P. Alboni, T. M. Tritt, J. Yang, *J. Appl. Phys.* **2006**, *99*, 044903.

[74] J. Martin, G. S. Nolas, H. Wang, J. Yang, *J. Appl. Phys.* **2007**, *102*, 103719.

[75] N. L. Okamoto, K. Kishida, K. Tanaka, H. Inui, *J. Appl. Phys.* **2006**, *100*, 073504.

[76] A. Saramat, E. S. Toberer, A. F. May, G. J. Snyder, *J. Electron. Mater.* **2009**, *38*, 1423.

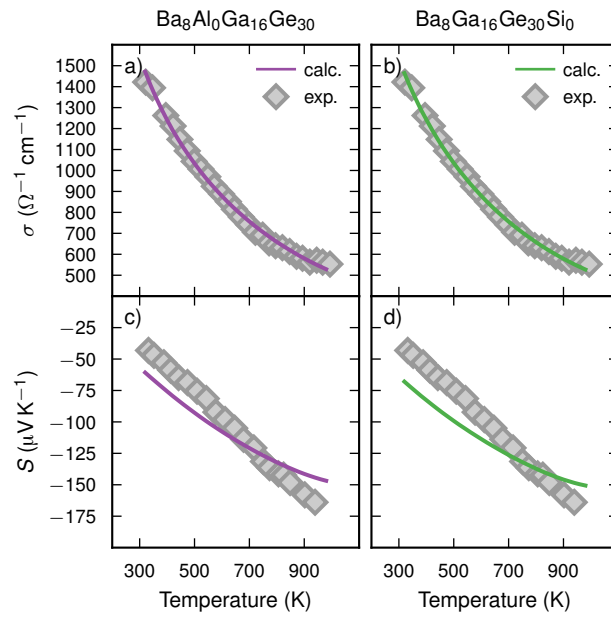[77] S. Y. Rodriguez, L. Saribaev, J. H. Ross, *Phys. Rev. B* **2010**, *82*, 064111.

Strategic optimization of the electronic transport properties of pseudo-ternary clathrates

Joakim Brorsson, Anders E. C. Palmqvist, and Paul Erhart

Contents

Supplementary Figures

Supplementary Figure S1: Fit of electronic transport properties. Electrical conductivities (a, b) and Seebeck coefficients (c, d) for $Ba_8Al_0Ga_{16}Ge_{30}$ (a, c) and $Ba_8Ga_{16}Ge_{30}Si_0$ (b, d) that have been fitted (solid lines) to the experimental data (gray diamonds) reported by Saramat *et al.* (1) for a single crystalline sample (disk 15) of $Ba_8Ga_{16}Ge_{30}$.
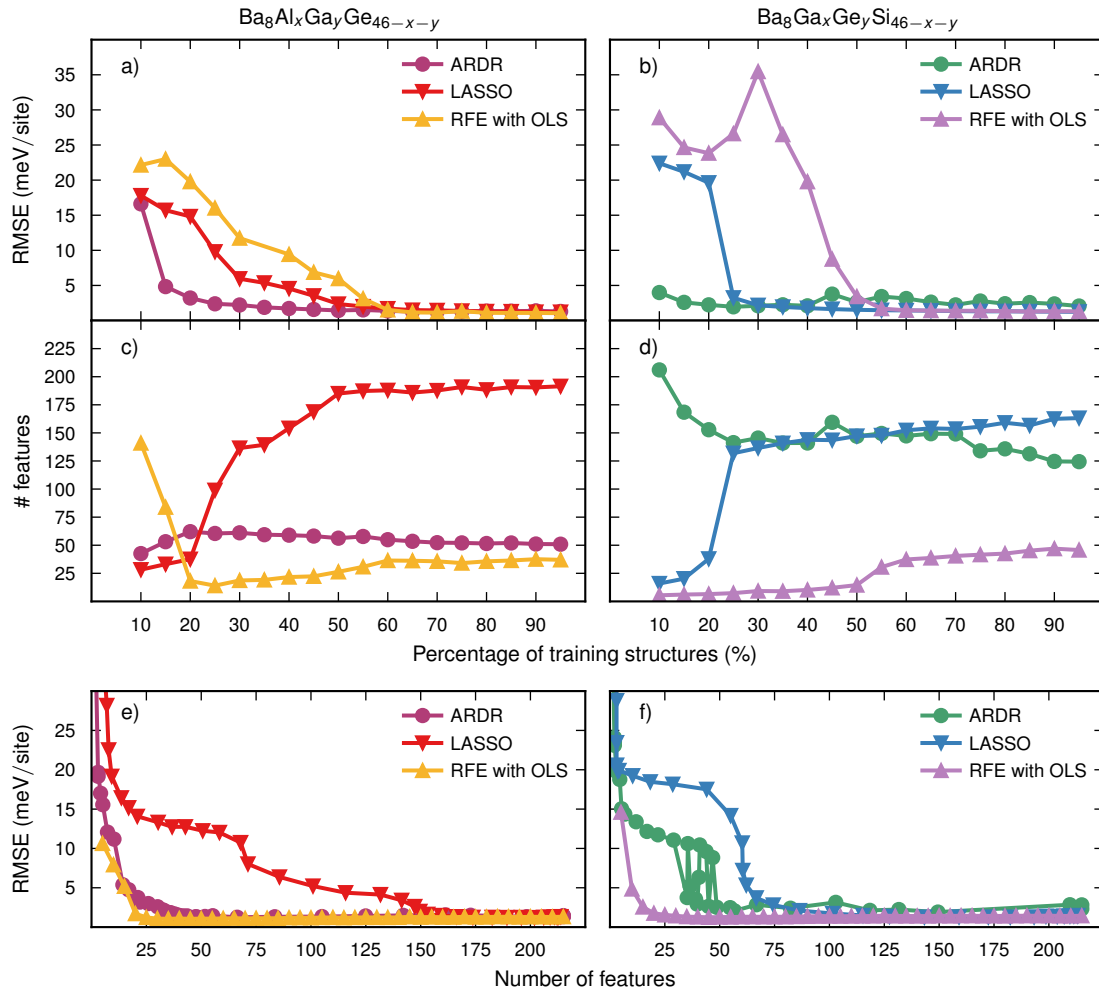
Supplementary Figure S2: Learning curves. Learning curves for $Ba_8Al_xGa_yGe_{46-x-y}$ (a, c, e) $Ba_8Ga_xGe_ySi_{46-x-y}$ (b, d, f), which include (a, b) the root-mean-square error (RMSE) calculated via cross-validation (CV) and (b, c) the number of features (non-zero parameters) versus the percentage of training structures as well as (d, e) the CV-RMSE as a function of the number of features when 90 % of the structures have been used for training.
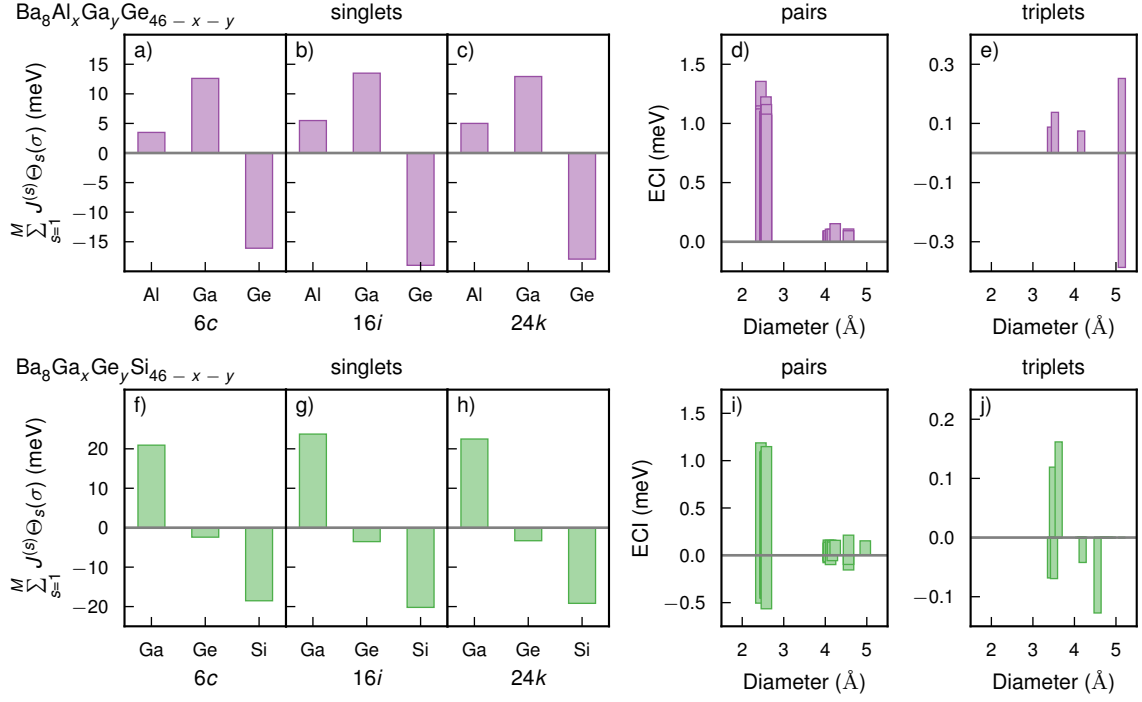
Supplementary Figure S3: Effective cluster interactions. Effective cluster interactions (ECIs) for the $Ba_8Al_xGa_yGe_{46-x-y}$ (a-e) and $Ba_8Ga_xGe_ySi_{46-x-y}$ (f-j) cluster expansions (CEs), including the sums of the contributions from singlets for each host atom and Wyckoff site, calculated as $\sum_{s=1}^{M} J^{(s)}\Theta_s(\sigma)$, (a-c, f-h) as well as the ECIs for pairs (d, i) and triplets (e, j).
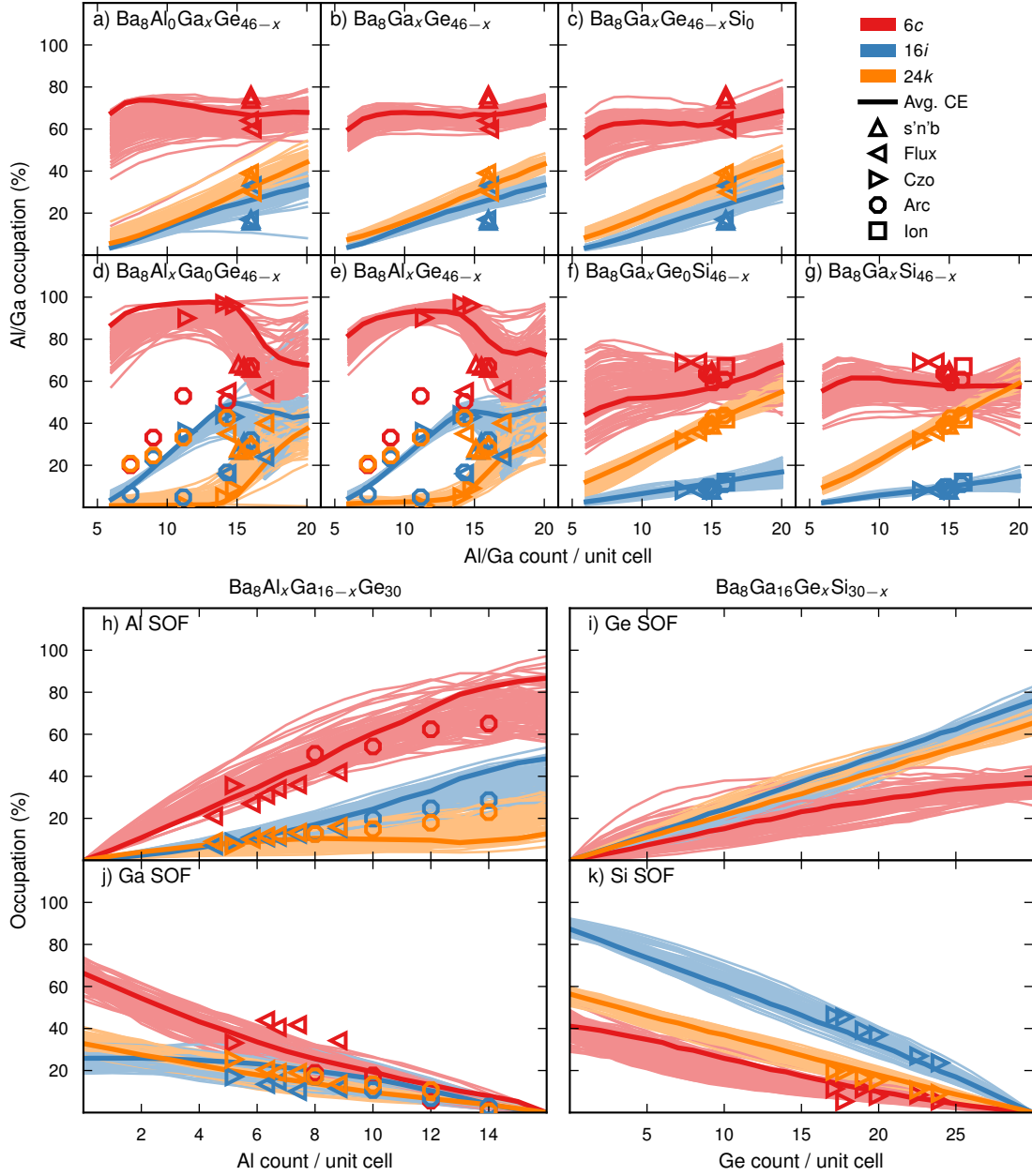
Supplementary Figure S4: Experimental and calculated SOFs. Site occupation factors (SOFs), at the $6c$ (red); $16i$ (blue), and $24k$ (orange) Wyckoff sites, from Monte Carlo (MC) simulations based on ensembles of psedo-binary (b,e,g) and pseudo-ternary (a,c,d,f) CEs together with experimental data for $Ba_8Ga_xGe_{46-x}$ (a-c), $Ba_8Al_xGe_{46-x}$ (d,e), and $Ba_8Ga_xSi_{46-x}$ (f,g). The occupations of Al (h) and Ga (j) in $Ba_8Al_xGa_{16-x}Ge_{30}$ as well as Ge (i) and Si (k) $Ba_8Ga_{16}Ge_xSi_{16-x}$ are also shown. Specifically, the data includes measurements on samples synthesised via solid state synthesis, also referred to as "shake and bake" (s'n'b) (2), (upward triangles) (3–8), flux-growth (3, 4, 6, 9–11) (left triangle) Czochralski-pulling (4, 6, 11, 12), arc-melting (13, 14), and Ar ion melting (15).

5
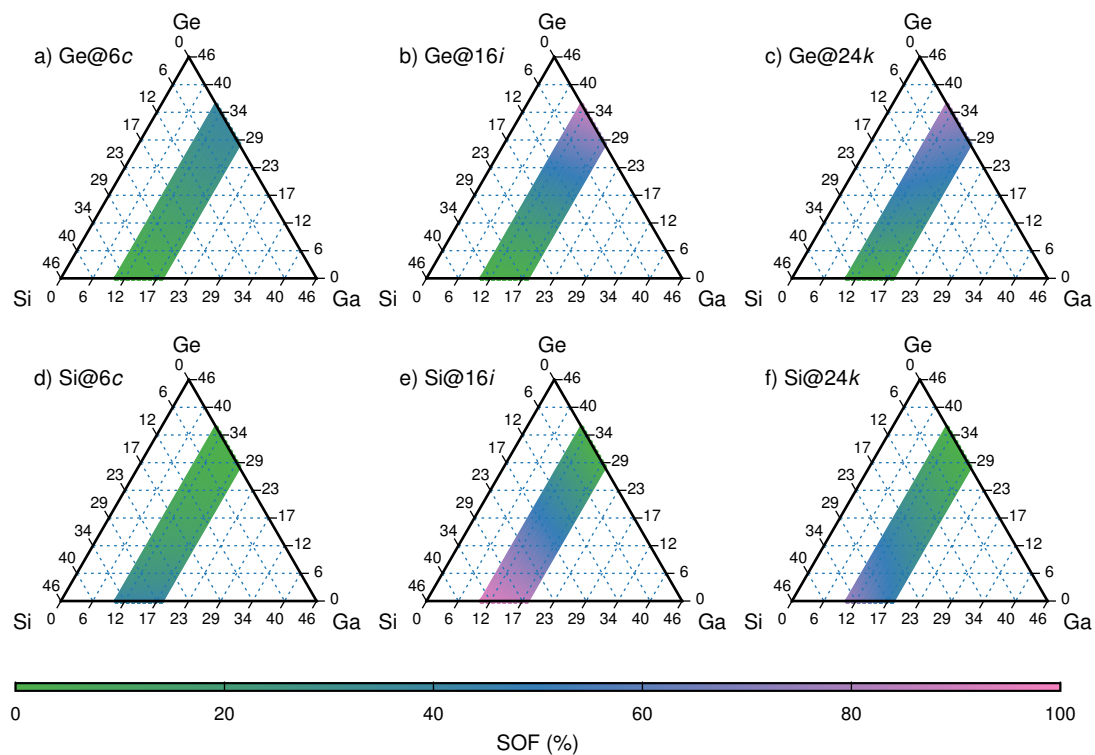
Supplementary Figure S5: Hybrid SOF maps for $Ba_8Ga_{16}Ge_xSi_{16-x}$. Maps with the Ge (a-c) and Si (d-f) SOFs at the $6c$ (a, d), $16i$ (b, e), and $24k$ (c, f) sites for $Ba_8Ga_xGe_ySi_{46-x-y}$, obtained from MC simulations, based on hybrid canonical/VSCGC ensembles, at $700\,K$.
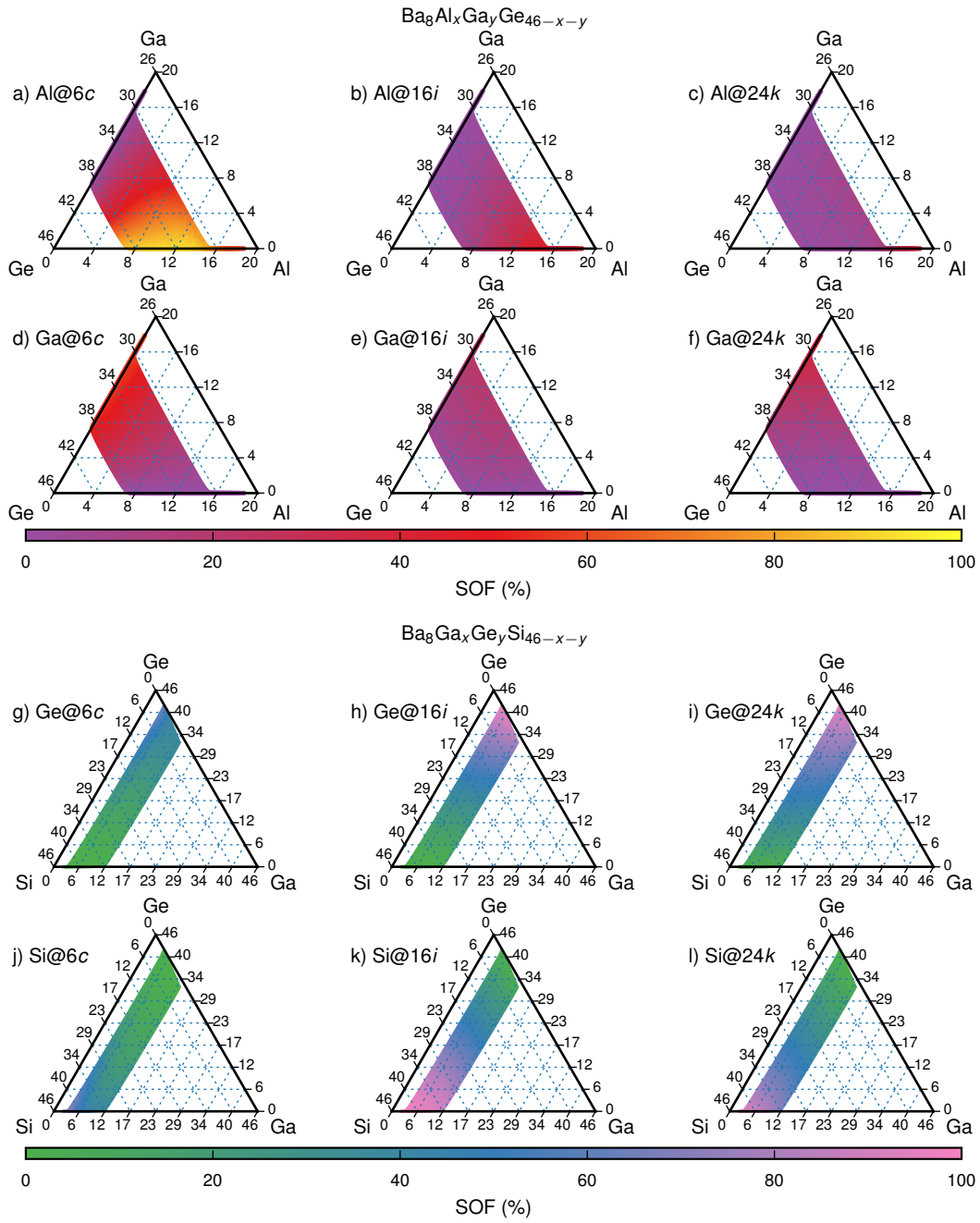
Supplementary Figure S6: VCSGC SOF maps. SOFs maps for Al (a-c) and Ga (d-f) in $Ba_8Ga_xGe_ySi_{46-x-y}$ and Ge (g-i) and Si (j-l) in $Ba_8Ga_xGe_ySi_{46-x-y}$ at the $6c$ (a, d, g, j), $16i$ (b, e, h, k), and $24k$ (c, f, i, l) sites, extracted from MC simulations, based on variance constrained semi-grand canonical (VCSGC) ensembles, at 700 K.
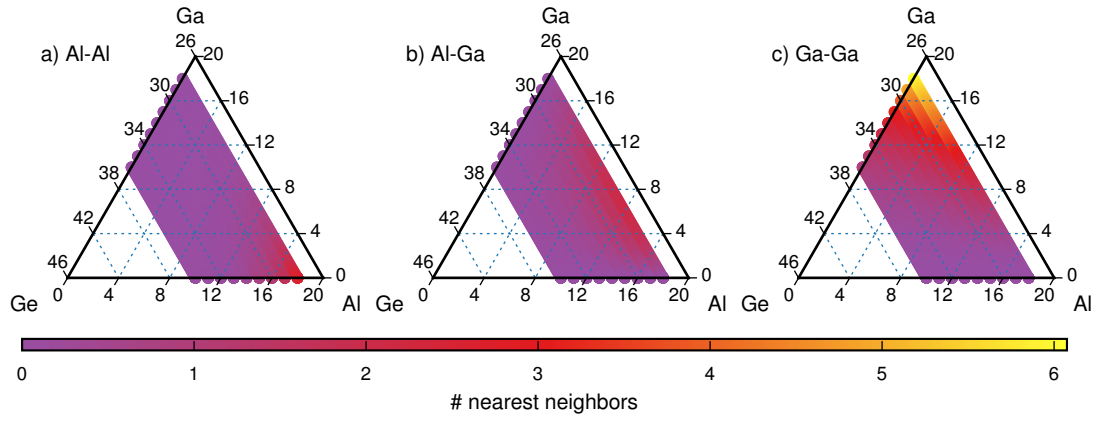
Supplementary Figure S7: Nearest neighbor maps for $Ba_8Al_xGa_{16-x}Ge_{30}$. Maps of the number of Al-Al (a), Al-Ga (b), and (Ga-Ga) (c) nearest neighbors for $Ba_8Al_xGa_yGe_{46-x-y}$, obtained from MC simulations, based on hybrid canonical/VSCGC ensembles, at 700 K.
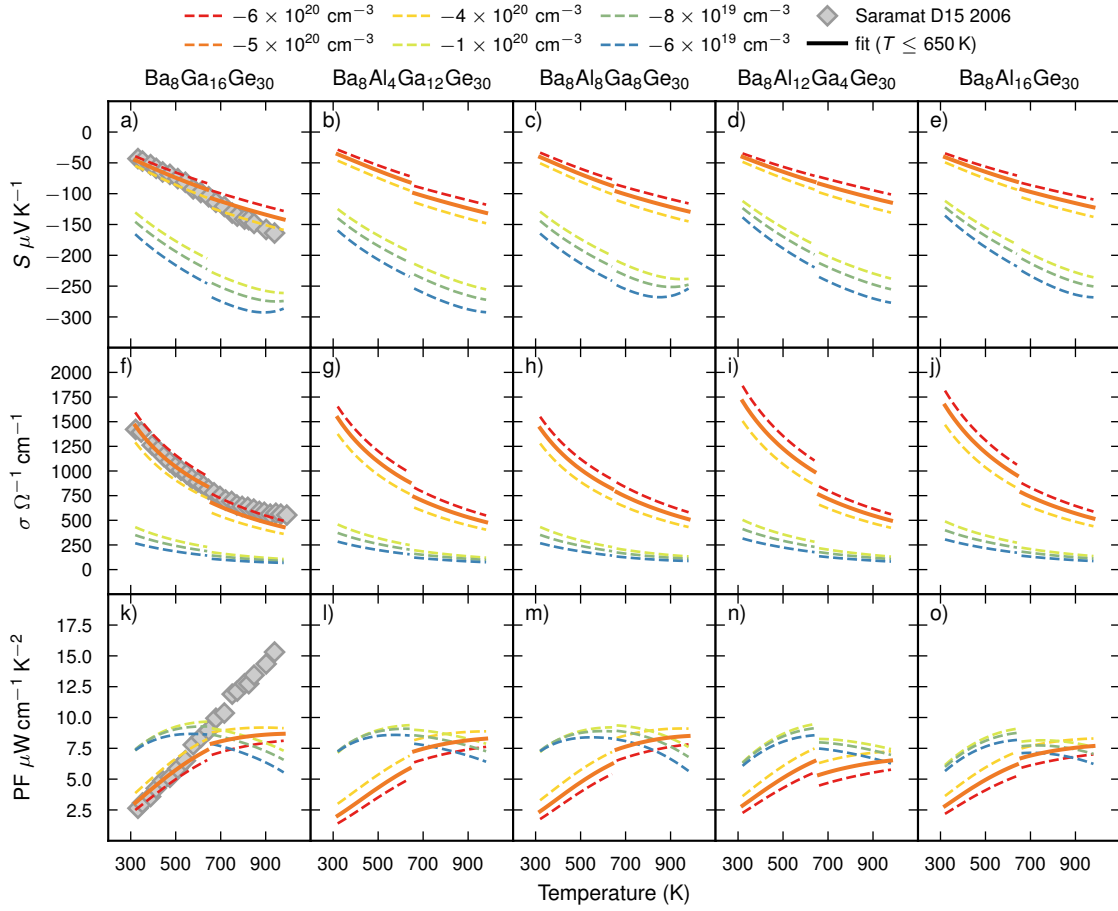
Supplementary Figure S8: Electronic transport in $Ba_8Al_xGa_{16-x}Ge_{30}$ for reference sample. Calculated Seebeck coefficient (a-e), electrical conductivity (f-j) and power factor (k-o) for $Ba_8Al_5Ga_{11}Ge_{30}$ with $x = 0$ (a,f,k); $x = 4$ (b,g,l); $x = 8$ (c,h,m); $x = 12$ (d,i,n); and $x = 16$ (e,j,o). The impact of the order-disorder transition has been taken into account by combining the predictions for the ground state and a representative high temperature (1200 K configuration, which was achieved by following the procedure described in our previous study of $Ba_8Al_xGa_yGe_{46-x-y}$ (11). It has also been assumed that the temperature dependence of the relaxation time follows a simple power law, $\tau_{\text{eff}} = \tau_{300\,\text{K}}(300\,\text{K}/T)^a$, and is the same for all compositions. More precisely, the corresponding parameters have been fitted to the electrical conductivity reported by Saramat *et al.* (1) for a single crystalline sample (disk 15) of $Ba_8Ga_{16}Ge_{30}$. The dashed lines, specifically, represent the results for different doping levels while the solid line gives the best fit between the calculated and measured Seebeck coefficient for $x = 0$.

Supplementary Figure S9: Electronic transport in $Ba_8Al_xGa_{16-x}Ge_{30}$ for alternative reference sample. Calculated Seebeck coefficient (a-e), electrical conductivity (f-j) and power factor (k-o) for $Ba_8Al_5Ga_{11}Ge_{30}$ with $x = 0$ (a,f,k); $x = 4$ (b,g,l); $x = 8$ (c,h,m); $x = 12$ (d,i,n); and $x = 16$ (e,j,o). The impact of the order-disorder transition has been taken into account by combining the predictions for the ground state and a representative high temperature (1200 K configuration, which was achieved by following the procedure described in our previous study of $Ba_8Al_xGa_yGe_{46-x-y}$ (11). It has also been assumed that the temperature dependence of the relaxation time follows a simple power law, $\tau_{\text{eff}} = \tau_{300\,\text{K}}(300\,\text{K}/T)^a$, and is the same for all compositions. More precisely, the corresponding parameters have been fitted to the electrical conductivity reported by Saramat *et al.* (1) for a single crystalline sample (disk 18) of $Ba_8Ga_{16}Ge_{30}$. The dashed lines, specifically, represent the results for different doping levels while the solid line gives the best fit between the calculated and measured Seebeck coefficient for $x = 0$.
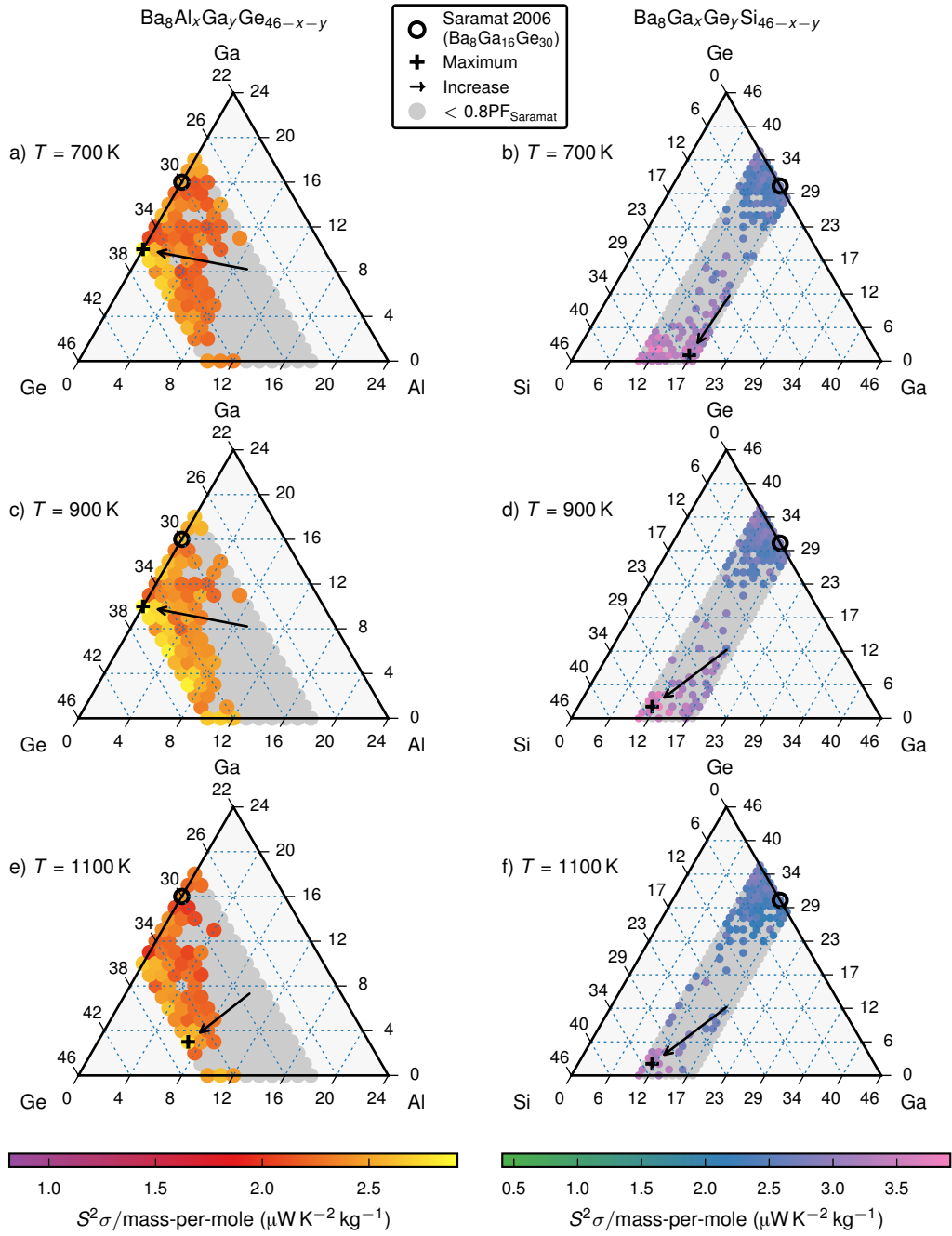
Supplementary Figure S10: PF map weighted by molar mass. Mapping of the power factor (PF) (σS^2) weighted by the molar mass (see Table S1) for $Ba_8Al_xGa_yGe_{46-x-y}$ (a, c, e) and $Ba_8Ga_xGe_ySi_{46-x-y}$ (b, d, f) at 700 K (a, b), 900 K (c, d), and 1100 K (e, f). All compositions that give a power factor lower than 80 % of the value for the reference sample have been uniformly colored (grey). The maximum value (plus) and the reference composition (circle) have also been indicated.
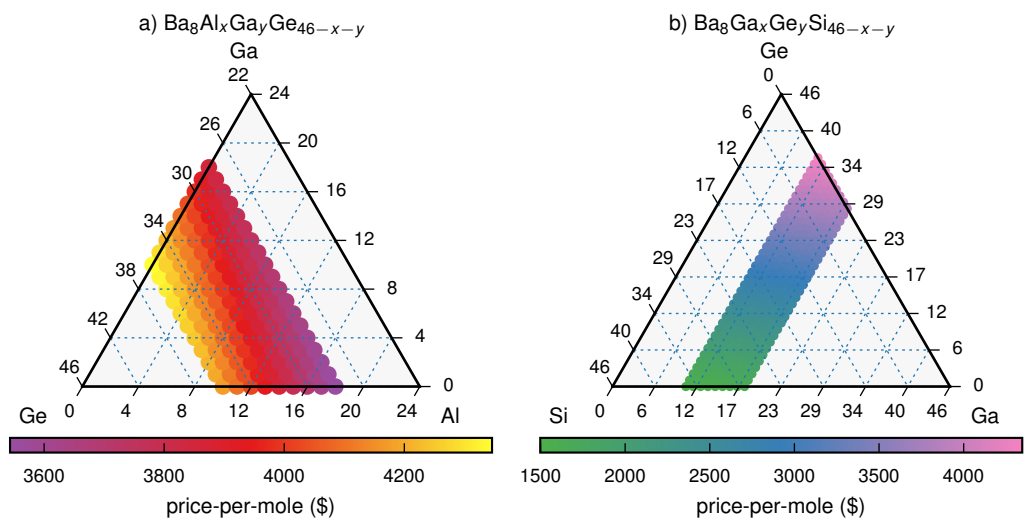
Supplementary Figure S11: Materials price map. Mapping of the materials price per mole (see Table S1) for $Ba_8Al_xGa_yGe_{46-x-y}$(a) and $Ba_8Ga_xGe_ySi_{46-x-y}$(b).

Supplementary Tables

chemical element	atomic weights (16) $(\mathrm{g\,mol^{-1}})$	material price	
		$(\$\,\mathrm{kg^{-1}})$	$(\$\,\mathrm{mol^{-1}})$
Ba	137.3	1206.3 (17)	165.7
Al	26.98	2.6475 (18)	0.07143
Ga	69.72	245.00 (19)	17.08
Ge	72.63	1090.0 (19)	79.17
Si	28.09	2.2672 (19)	0.06367

Supplementary Table S1: Elemental weights and market prices. Atomic weights and market prices for the chemical elements in $\mathrm{Ba_8Al_xGa_yGe_{46-x-y}}$ and $\mathrm{Ba_8Ga_xGe_ySi_{46-x-y}}$. Note that the fourth column has been calculated using the data from the second and third column.

Supplementary Notes

Supplementary Note S1: Electronic structure calculations.

The target data required for constructing the alloy CEs (Supplementary Note S2) consisted of energies for a total of 528 (585) $\mathrm{Ba_8Al_xGa_yGe_{46-x-y}}$ ($\mathrm{Ba_8Ga_xGe_ySi_{46-x-y}}$) randomly generated configurations, which included $132+132$ ($146+146$) $\mathrm{Ba_8Ga_xGe_{46-x}}$ + $\mathrm{Ba_8Al_xGe_{46-x}}$ ($\mathrm{Ba_8Ga_xGe_{46-x}}$ + $\mathrm{Ba_8Ga_xSi_{46-x}}$) structures. Density functional theory (DFT) based on the projector augmented wave (PAW) method (20, 21), as implemented in the Vienna ab initio simulation package (VASP) (22), was used for the relaxation of both the cell metric and ionic positions. Γ-centered $3 \times 3 \times 3$ $\boldsymbol{k}$-point meshes and $319\,\mathrm{eV}$ plane wave energy cutoffs were used for all calculations. Also, the width for the Gaussian smearing was set to $0.1\,\mathrm{eV}$ while the convergence criteria for the residual forces was kept at $5\,\mathrm{meV\,\mathring{A}}^{-1}$. Exchange-correlation effects, meanwhile, were accounted for by employing the van-der-Waals density functional method (23) with consistent exchange (vdW-DF-cx) (24).

Supplementary Note S2: Cluster expansion construction.

Alloy CEs can best be described as lattice models that represent particular chemical configurations through the population of the inherent sites. Though these possess many similarities with the generalized Ising model, a key difference, compared to the standard version of the latter, is that the interactions can extend beyond the nearest-neighbors and, additionally, involve more than two sites. Each such group, which is often referred to as a cluster, is associated with an ECI (J_α) that, in turn, corresponds to an *a priori* unknown parameter in the mathematical expression for the total energy, or indeed any other function $Q(\boldsymbol{\sigma})$, (25),

$$Q(\boldsymbol{\sigma}) = \sum_\alpha J_\alpha \Pi_\alpha(\sigma) \tag{S1}$$

Here, $\boldsymbol{\sigma}$ represents the configuration vector, which elements correspond to the chemical species that occupy each site, and $\Pi_\alpha(\sigma)$ the basis functions. The α, meanwhile, are the vector representations of the different clusters, consisting of a list of point functions that indicate if the corresponding site is included or not. In order to describe the energetics of a multi-component system, the expansion coefficients (J_α) are fitted using data obtained from first-principles calculations, which are typically based on DFT. While the model thus obtained operates on a rigid lattice, it will effectively include atomic relaxation given that the training structures have been properly relaxed. When alloy CEs are combined MC simulations, this gives rise to a versatile methodology that is sufficiently effective to allow for the sampling of multiple compositional degrees of freedom with a level of accuracy equivalent to that of first-principles calculations.

After having generated sets of relaxed $\mathrm{Ba_8Al_xGa_yGe_{46-x-y}}$ and $\mathrm{Ba_8Ga_xGe_ySi_{46-x-y}}$ structures using DFT (Supplementary Note S1), we employed the ICET software package (25) to train the alloy CEs. For these pseudo-ternary systems, we obtained a total of 215 symmetry inequivalent clusters, which included 6 singlets, 46 pairs and 162 triplets, when setting the cutoffs, for both pairs and triplets, just below one half of the lattice parameter ($5.4\,\mathrm{\mathring{A}}$). Before constructing the final CE, with help of ordinary least squares (OLS) with recursive feature elimination (RFE), we performed an extensive testing of various fitting methods, which also included least absolute shrinkage and selection operator (LASSO) and automatic relevance detection regression (ARDR) (see Reference (26) and (25) as well as Figure S2). As it turns out, the RMSEs, obtained via CV, for different training set sizes are very similar when more than half of the structures are used for training (Figure S2a,b). Even so, RFE with OLS gives the lowest score, followed by ARDR. The fact that the trend is exactly the opposite for training set sizes below $50\,\%$, is expected since LASSO and ARDR are known to give a better performance for underdetermined systems (25, 26). Another crucial factor to consider, however, is the number of non-zero parameters and in this respect RFE based on OLS is clearly the best choice, especially for $\mathrm{Ba_8Ga_xGe_ySi_{46-x-y}}$ (Figure S2d) while ARDR is almost as good in the case of $\mathrm{Ba_8Al_xGa_yGe_{46-x-y}}$ (Figure S2c).

As we have previously reported (27), the ECIs contain key information regarding the interactions between the atomic species found on the active sublattice. In that case, it was observed that the pair contributions were higher for pseudo-binary clathrates that contained Ga ($\mathrm{Ba_8Ga_xGe_{46-x}}$ and $\mathrm{Ba_8Ga_xSi_{46-x}}$) rather than Al ($\mathrm{Ba_8Al_xGe_{46-x}}$ and $\mathrm{Ba_8Al_xSi_{46-x}}$), which we deemed to be the foremost reason for the sharp changes in the SOFs as $x \to 16$ exhibited by the latter. One can also observe (see Figure 3 in reference 27) that the singlets are slightly more negative for $\mathrm{Ba_8Ga_xGe_{46-x}}$ ($\mathrm{Ba_8Ga_xSi_{46-x}}$) compared to $\mathrm{Ba_8Al_xGe_{46-x}}$ ($\mathrm{Ba_8Al_xSi_{46-x}}$). Though a direct comparison with the pseudo-ternary systems, which have six rather than three singlets, is problematic, some conclusions can be drawn by summing the contributions associated with each species and Wyckoff site as $\sum_{s=1}^{M} J^{(s)}\Theta_s(\sigma)$. Here, $\Theta_s(\sigma)$ is one of the point functions in the product that represents the basis functions in Equation S1 $\Pi_\alpha(\sigma) = \Theta_{n_1}(\sigma_1)\Theta_{n_2}(\sigma_2)...\Theta_{n_l}(\sigma_l)$. For $\mathrm{Ba_8Al_xGa_yGe_{46-x-y}}$, the values thus obtained vary only weakly depending on the Wyckoff site, while there are distinctive differences depending on the nature of the species (Figure S3a-c). In particular, the sum of the ECI is more positive for Ga compared to Al and

only negative for Ge, meaning that the incorporation of the trivalent elements makes the system less stable. The situation is similar for $Ba_8Ga_xGe_ySi_{46-x-y}$ (Figure S3f-h) in the sense that Si provides a larger negative contribution than Ge while the term associated with Ga is positive. Less can be said about the pairs and triplets other than the fact that the only positive pair interactions are observed for $Ba_8Al_xGa_yGe_{46-x-y}$ (Figure S3e,f) while some are negative in the case of $Ba_8Ga_xGe_ySi_{46-x-y}$ (Figure S3i,j).

Supplementary Note S3: Monte Carlo simulations.

Standard MC simulations, as implemented in the MCHAMMER module of ICET (25), was used to sample the CEs (Supplementary Note S2) in order to determine the chemical ordering, as measured by the SOFs, and mixing energies as well as to generate representative structures for the transport calculations. This included a type of simulated annealing, in the canonical ensemble, that were initiated at 1200 K whereafter the temperature was lowered in 100 K step, after performing 22 000 MC cycles, down to 0 K. These simulations were performed for a wide range of different stoichiometries, namely $Ba_8Al_xGa_yGe_{46-x-y}$ ($Ba_8Ga_xGe_ySi_{46-x-y}$) with $6 \leq x+y \leq 20$ ($6 \leq x \leq 20$), using $1 \times 1 \times 1$ and $2 \times 2 \times 2$ supercells. For the purpose of further testing the sensitivity of the models, ensembles consisting of 100 CEs were constructed and sampled in a similar fashion, although in this case the temperature was kept fixed, at 700 K.

We, furthermore, employed the VCSGC ensemble (28) at a single temperature (700 K) for the purpose of continuously varying the chemical composition. This method alone is, however, not sufficient to be able to integrate the free energies and, thereby, determine the mixing energies and entropies, since the clathrate phase is only thermodynamically stable for limited range of trivalent (Al, Ga) concentrations. For instance, several experimental studies of pseudo-binary clathrates, such as $Ba_8Ga_xGe_{46-x}$, has shown that the solubility limit is only slightly higher, or even lower than, the charge balanced composition ($x = 16$). In terms of the VCSGC calculations, this means that the end points that are required to calculated the mixing energies within the entire configurational space, which in the case of $Ba_8Al_xGa_yGe_{46-x-y}$ would correspond to $Ba_8\{Al,Ga,Ge\}_{46}$, are not available (Figure S6). The fact that the pair of parameters being varied represent the average and variance rather than the actual concentration, moreover means that some compositions outside the nominal range are being sampled, while some within are not (see the Supplementary Information). Consequently, it is not possible to calculate the energy associated with the mixing of two pseudo-binary compositions, such as $Ba_8Ga_{16}Ge_{30}$ and $Ba_8Al_{16}Ge_{30}$. This problem, can, however, be circumvented by employing a hybrid canonical/VCSGC ensemble. Taking $Ba_8Al_xGa_yGe_{46-x-y}$ as an example, this approach is based on randomly alternating between trial steps within the VCSGC and canonical ensemble. In particular, the former is only allowed to switch the occupation of a single site from Al to Ga, or the vice versa, while the latter involves swapping the species, which can be any of Al; Ga; or Ge, occupying two different sites. Consequently, it is possible to keep the the trivalent concentration, fixed and, therefore, ensure that the entire range of interest is covered.

Supplementary Note S4: Calculations of electronic properties.

To map the electronic transport properties across the entire composition range and, thereby, enable the identification of promising compositions, we extracted representative structures from MC simulations performed within the canonical ensemble, using $2 \times 2 \times 2$ supercells (Supplementary Note S3). These were, more precisely, obtained by constructing the equivalent of a primitive, special quasi-ordered structure (29), which was selected so as to match the average cluster vector, for a given chemical composition, at 700 K. As we have shown before (11), the alternative approach of considering multiple randomly selected structures and calculating the statistical averages yields similar results, although at a significantly higher computational cost.

After having extracted the structures as outline above, these were relaxed by using the same settings as before (Supplementary Note S1), with only a few exceptions, namely that a denser ($5 \times 5 \times 5$) k-points grid and more stringent convergence criteria (by a factor of 10) were employed. Next, a non-self consistent calculation was carried out after switching to an interpolated Γ-centered $25 \times 25 \times 25$ mesh. The final step involved the use of the Boltzmann transport theory (BTT) based BOLTZTRAP2 (30) software to calculate the transport coefficients, with both the interpolated eigenenergies and the group velocities as input, at the relaxation time approximation (RTA) level. We have previously applied a similar approach and shown that it provides a good agreement with experiments both in the case of $Ba_8Ga_{16}Ge_{30}$ (31) and $Ba_8Al_xGa_{16-x}Ge_{30}$ (11). In accordance with these studies, it was deemed sufficient to model the temperature dependence of the relaxation time using a simple power law, $\tau_{eff}(T) = \tau_{300\,K}(300\,K/T)^a$. To determine the parameters, in the form of the value at 300 K ($\tau_{300\,K}$) and the exponent (a), the calculated electrical conductivity was fitted to the experimental data reported by Saramat $et~al.$ (1) for a $Ba_8Ga_{16}Ge_{30}$ single crystal (see Figure S1). In addition, test calculations were performed (see Figure S8 and Figure S9) based on a procedure that we have previously used for studying $Ba_8Al_xGa_yGe_{46-x-y}$. More precisely, we thereby account for the existence of an order-disorder transition and, moreover, use corrected band gaps, which should be in better agreement with experiments.

Because the number of electrons varies strongly with the concentration of trivalent elements, this often leads to deviations from the charge balanced, stoichiometric, composition. It is, furthermore, a well-known fact that the PF is only appreciable for a relatively limited range of carrier concentrations, which means that an undoped inorganic type-I pseudo-binary clathrate, such as $Ba_8Ga_xGe_{46-x}$, can only be expected to display a decent performance if $x \approx 16$. As our aim is to single out the contribution from the band structure we have, thus, decided to fix the number of electrons so that it corresponds to the doping level that provides the best agreement between the calculated and experimental Seebeck coefficients for $Ba_8Ga_{16}Ge_{30}$. In essence, our calculations therefore indicate to what extent the replacement of the Ga or Ge atoms in $Ba_8Ga_{16}Ge_{30}$ with Al or Si, while ensuring that the carrier concentration is kept fixed via extrinsic doping, can lead to an increase in the PF. This can, in other words, be seen as a proof of concept for a computational approach that can be used to explore the possibilities of band structure engineering within a pseudo-ternary system.

Supplementary Note S5: Comparison between experimental and calculated SOFs.

To validate our calculations of the $Ba_8Al_xGa_yGe_{46-x-y}$ and $Ba_8Ga_xGe_ySi_{46-x-y}$ SOFs we have compared the results obtained by sampling ensembles of pseudo-ternary, and pesudo-binary, CEs with experimental data(3–15), not only for $Ba_8Al_xGa_{16-x}Ge_{30}$ (Figure S4h,j) and $Ba_8Ga_{16}Ge_xSi_{16-x}$ (Figure S4i,k) but also $Ba_8Ga_xGe_{46-x}$ (Figure S4a-c); $Ba_8Al_xGe_{46-x}$ (Figure S4d,e); and $Ba_8Ga_xSi_{46-x}$ (Figure S4f,g). Given the relatively large variations in the measured SOFs, it can be concluded that an overall good agreement is achieved. While the pseudo-binary and pseudo-ternary CEs yield very similar results, it also evident that the spread of the latter is slightly larger, possibly due to the higher number of non-zero parameters. Though large discrepancies with respect to the experimental measurements can be observed, especially for $Ba_8Al_xGe_{46-x}$ and $x \ll 16$, one must also keep in mind that these samples contain vacancies and, in addition, that we have calculated the SOFs at a temperature (700 K) that lies above the order-disorder transition (~ 650 K) (11). The fact that the predictions and measurements for both $Ba_8Al_xGa_{16-x}Ge_{30}$, which we have previously reported (32), and $Ba_8Ga_{16}Ge_xSi_{16-x}$ match well across significant portions of the composition range, can be seen as further evidence for this conclusion.

Supplementary References

[1] A. Saramat, G. Svensson, A. E. C. Palmqvist, C. Stiewe, E. Mueller, D. Platzek, S. G. K. Williams, D. M. Rowe, J. D. Bryan, and G. D. Stucky, *Large thermoelectric figure of merit at high temperature in Czochralski-grown clathrate $Ba_8Ga_{16}Ge_{30}$,* Journal of Applied Physics **99**, 023708 (2006). `doi:doi:10.1063/1.2163979`.

[2] M. Christensen, S. Johnsen, and B. B. Iversen, *Thermoelectric clathrates of type I,* Dalton Transactions **39**, 978 (2010). `doi:10.1039/B916400F`.

[3] M. Christensen, N. Lock, J. Overgaard, and B. B. Iversen, *Crystal Structures of Thermoelectric n- and p-type Ba8Ga16Ge30 Studied by Single Crystal, Multitemperature, Neutron Diffraction, Conventional X-ray Diffraction and Resonant Synchrotron X-ray Diffraction,* Journal of the American Chemical Society **128**, 15657 (2006). PMID: 17147375. `doi:10.1021/ja063695y`.

[4] M. Christensen and B. B. Iversen, *Host Structure Engineering in Thermoelectric Clathrates,* Chemistry of Materials **19**, 4896 (2007). `doi:10.1021/cm071435p`.

[5] B. Eisenmann, H. Schäfer, and R. Zagler, *Die verbindungen $AI_8^I B_{16}^{III} B_{30}^{IV}$ ($A^{II} \equiv$ Sr,Ba; $B^{III} \equiv$ Al,Ga; $B^{IV} \equiv$ Si,Ge,Sn) und ihre käfigstrukturen,* Journal of the Less Common Metals **118**, 43 (1986). `doi:http://dx.doi.org/10.1016/0022-5088(86)90609-0`.

[6] M. Christensen, G. J. Snyder, and B. Iversen, *High temperature thermoelectric properties of Czochralski-pulled $Ba_8Ga_{16}Ge_{30}$,* in *2006 25th International Conference on Thermoelectrics*, 40, 2006. `doi:10.1109/ICT.2006.331265`.

[7] L. Qiu, I. P. Swainson, G. S. Nolas, and M. A. White, *Structure, thermal, and transport properties of the clathrates $Sr_8Zn_8Ge_{38}$, $Sr_8Ga_{16}Ge_{30}$, and $Ba_8Ga_{16}Si_{30}$,* Physical Review B **70**, 035208 (2004). `doi:10.1103/PhysRevB.70.035208`.

[8] M. Bobnar, B. Böhme, M. Wedel, U. Burkhardt, A. Ormeci, Y. Prots, C. Drathen, Y. Liang, H. D. Nguyen, M. Baitinger, and Y. Grin, *Distribution of Al atoms in the clathrate-I phase $Ba_8Al_xSi_{46-x}$ at $x = 6.9$,* Dalton Transactions **44**, 12680 (2015). `doi:10.1039/C5DT01198A`.

[9] A. Bentien, B. Iversen, J. Bryan, G. Stucky, A. Palmqvist, A. Schultz, and R. Henning, *Maximum entropy method analysis of thermal motion and disorder in thermoelectric clathrate $Ba_8Ga_{16}Si_{30}$,* Journal of Applied Physics **91**, 5694 (2002). `doi:10.1063/1.1466531`.

[10] A. Bentien, E. Nishibori, S. Paschen, and B. B. Iversen, *Crystal structures, atomic vibration, and disorder of the type-I thermoelectric clathrates $Ba_8Ga_{16}Si_{30}$, $Ba_8Ga_{16}Ge_{30}$, $Ba_8In_{16}Ge_{30}$, and $Sr_8Ga_{16}Ge_{30}$,* Physical Review B **71**, 1 (2005). `doi:10.1103/PhysRevB.71.144107`.

[11] J. Brorsson, Y. Zhang, A. E. C. Palmqvist, and P. Erhart, *Order–Disorder Transition in Inorganic Clathrates Controls Electrical Transport Properties,* Chemistry of Materials , (2021). `doi:10.1021/acs.chemmater.1c00731`.

[12] M. Christensen, S. Johnsen, M. Søndergaard, J. Overgaard, H. Birkedal, and B. B. Iversen, *Fast Preparation and Characterization of Quarternary Thermoelectric Clathrates,* Chemistry of Materials **21**, 122 (2009). `doi:10.1021/cm802289n`.

[13] J. H. Roudebush, C. de la Cruz, B. C. Chakoumakos, and S. M. Kauzlarich, *Neutron Diffraction Study of the Type I Clathrate $Ba_8Al_xSi_{46-x}$: Site Occupancies, Cage Volumes, and the Interaction between the Guest and the Host Framework,* Inorganic Chemistry **51**, 1805 (2012). `doi:10.1021/ic202095e`.

[14] W. R. Puspita, H. Takeya, T. Mochiku, Y. Ishikawa, S. Lee, S. Torii, M. Hagihala, and T. Kamiyama, *Structural Study of Thermoelectric Clathrate $Ba_8Al_{16-x}Ga_xGe_{3}0$,* in *2019 2nd International Conference on Applied Engineering (ICAE)*, 1, IEEE, 2019.

[15] D. Nataraj and J. Nagao, *Structure and Raman scattering study on Ba8GaxSi46-x(x=10 and 16) type I clathrates,* Journal of Solid State Chemistry **177**, 1905 (2004). `doi:http://dx.doi.org/10.1016/j.jssc.2004.01.011`.

[16] J. Meija, T. B. Coplen, M. Berglund, W. A. Brand, P. D. Bièvre, M. Gröning, N. E. Holden, J. Irrgeher, R. D. Loss, T. Walczyk, and T. Prohaska, *Atomic weights of the elements 2013 (IUPAC Technical Report),* Pure and Applied Chemistry **88**, 265 (2016). `doi:doi:10.1515/pac-2015-0305`.

[17] Institut für seltene Erden und Metalle (ISE), *Prices for high purity metals in December 2020,* Accessed: 2021-05-19. `https://en.institut-seltene-erden.de/prices-for-high-purity-metals-in-december-2020/`.

[18] Institut für seltene Erden und Metalle (ISE), *Base metals prices in December 2020,* Accessed: 2021-05-19. `https://en.institut-seltene-erden.de/prices-for-base-metals-in-december/`.

[19] Institut für seltene Erden und Metalle (ISE), *Prices for non-ferrous metals in December 2020,* Accessed: 2021-05-19. `https://en.institut-seltene-erden.de/strategic-metals-prices-in-december/`.

[20] P. E. Blöchl, *Projector augmented-wave method,* Physical Review B **50**, 17953 (1994). `doi:10.1103/PhysRevB.54.11169`.

[21] G. Kresse and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method,* Physical Review B **59**, 1758 (1999). `doi:10.1103/PhysRevB.59.1758`.

[22] G. Kresse and J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set,* Physical Review B **54**, 11169 (1996). `doi:10.1103/PhysRevB.54.11169`.

[23] K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard, and B. I. Lundqvist, *Van Der Waals Forces in Density Functional Theory: A Review of the vdW-DF Method,* Reports on Progress in Physics **78**, 066501 (2015). `doi:10.1088/0034-4885/78/6/066501`.

[24] K. Berland and P. Hyldgaard, *Exchange functional that tests the robustness of the plasmon description of the van der Waals density functional,* Physical Review B **89**, 035412 (2014). `doi:10.1103/PhysRevB.89.035412`.

[25] M. Ångqvist, W. A. Muñoz, J. M. Rahm, E. Fransson, C. Durniak, P. Rozyczko, T. H. Rod, and P. Erhart, *ICET – A Python Library for Constructing and Sampling Alloy Cluster Expansions,* Advanced Theory and Simulations **2**, 1900015 (2019). `doi:10.1002/adts.201900015`.

[26] E. Fransson, F. Eriksson, and P. Erhart, *Efficient construction of linear models in materials modeling and applications to force constant expansions,* npj Computational Materials **6**, 135 (2020). `doi:10.1038/s41524-020-00404-5`.

[27] M. Ångqvist and P. Erhart, *Understanding chemical ordering in intermetallic clathrates from atomic scale simulations,* Chemistry of Materials **29**, 7554 (2017). `doi:10.1021/acs.chemmater.7b02686`.

[28] B. Sadigh and P. Erhart, *Calculation of excess free energies of precipitates via direct thermodynamic integration across phase boundaries,* Physical Review B **86**, 134204 (2012). `doi:10.1103/PhysRevB.86.134204`.

[29] A. van de Walle, P. Tiwary, D. L. O. de Jong, M., M. Asta, A. Dick, D. Shin, Y. Wang, L.-Q. Chen, and Z.-K. Liu, *Efficient stochastic generation of special quasirandom structures,* Calphad **42**, 13 (2013). `doi:10.1016/j.calphad.2013.06.006`.

[30] G. K. Madsen, J. Carrete, and M. J. Verstraete, *BoltzTraP2, a program for interpolating band structures and calculating semi-classical transport coefficients,* Computer Physics Communications **231**, 140 (2018). `doi:10.1016/j.cpc.2018.05.010`.

[31] M. Ångqvist, D. O. Lindroth, and P. Erhart, *Optimization of the thermoelectric power factor: Coupling between chemical order and transport properties,* Chemistry of Materials **28**, 6877 (2016). `doi:10.1021/acs.chemmater.6b02117`.

[32] Y. Zhang, J. Brorsson, T. Kamiyama, T. Saito, P. Erhart, and A. E. C. Palmqvist, *Investigating the chemical ordering in quaternary clathrate $Ba_8Al_xGa_{16-x}Ge_{30}$,* Manuscript (2021).