Thermodynamically limited Cu-Zn order in Cu_2ZnSnS_4 (CZTS) from Monte Carlo simulations

Suzanne K. Wallace,[a,b] **Jarvist Moore Frost,**[b] **and Aron Walsh**[*b,c]

[a] *Department of Chemistry, Centre for Sustainable Chemical Technologies, University of Bath, Claverton Down, Bath, BA2 7AY, UK*

[b] *Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, UK. Email: a.walsh@imperial.ac.uk*

[c] *Global E*^{3} *Institute and Department of Materials Science and Engineering, Yonsei University, Seoul 03722, Korea*

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

First published on the web Xth XXXXXXXXXX 200X

DOI: 10.1039/b000000x

Kesterite-structured Cu_2ZnSnS_4 (CZTS) is a semiconductor that is being studied for use as the absorber layer in thin-film solar cells. Currently the power-conversion efficiencies of this technology fall short of the requirements for commercialisation, despite the promising sunlight-matched optical band gap. Disorder in the Cu-Zn sub-lattice has been observed and is proposed as one possible explanation for the shortcomings of CZTS solar cells. Cation site disorder averaged over a macroscopic sample does not provide insights into the microscopic cation distribution that will interact with photogenerated electrons and holes. To provide atomistic insight into Cu-Zn disorder we have developed a Monte Carlo (MC) model based on pairwise interactions. We utilise two order parameters to relate Cu-Zn disorder to the processing temperature for stoichiometric systems: one based on cation site occupancies (the 'Q' order parameter) and the other based on cation pair correlation functions. Our model predicts that the order parameters reach a plateau at experimentally relevant low temperatures, indicating that Cu-Zn order in stoichiometric CZTS is thermodynamically limited. At room temperature we predict a minimum of 10% disorder in the cation site occupancy within (001) Cu-Zn planes.

Notes:

- Upload final stages of Eris data sets and pre-prints to Starcell OpenAIRE initiative: https://zenodo.org/communities/wmd-group – can do after submission
- Prepare final analysis notebooks for ErisCodes repo for open access (see Lucy's notebooks for e.g. https://github.com/WMD-group/hot-carrier-cooling) – can do after submission
- Open access ErisCodes git repo release of code? - Check with Aron and Jarv – release 'ErisCodes' repo because contains latest/ fixed version of code and without my excess data
- Revisit Scragg paper on cation ordering kinetics? (http://aip.scitation.org/doi/abs/10.1063/1.5010081)
- **Add citation for Arons dielectric constant calc/ details in SI?** - Check with Sunghyun, couldn't see value in killer defects paper
- Check SI with Aron and Jarv

1 Introduction

Amongst the semiconductors being developed for applications in thin-film photovoltaic (PV) devices, kesterite-structured Cu_2ZnSnS_4 (CZTS) stands out as being composed of low-cost, earth-abundant and non-toxic elements. While the material has many of the bulk properties required to be a high-efficiency photovoltaic absorber, such as a high absorption coefcient of $10^4 cm^{-1}$ and a direct band gap of 1.5 eV[1], the power-conversion efficiencies (PCEs) of solar cells are considerably less than the theoretical maximum of 28% as predicted by the Shockley-Queisser limit[2] based on its sunlight-matched optical band gap. The current confirmed record PCE for the kesterite-based alloy $Cu_2ZnSn(S_xSe_{1-x})_4$ (CZTSSe) is at 12.6%[3], while that of the pure sulfide material lags even further behind at 9.1%[4], both of which are far below that of

the similar PV technology $Cu(In_{1-y},Ga_y)Se_2$ (CIGSe) with a record PCE of 22.6%[5].

The low open-circuit voltage (compared to the optical band gap) limits achieved device efficiencies[6,7]. This is referred to as the V_{OC} deficit. It is possible that the efficiency of devices fabricated with absorber layers produced from different synthesis procedures may be limited by different factors, making it a difficult task to pinpoint a universal origin of the V_{OC} deficit in CZTS solar cells. Defects and bulk disorder in CZTS is one possible explanation for the V_{OC} deficit[8,9]. For record-efficiency devices, produced by the hydrazine-based solution method pioneered at the IBM T. J. Watson Research Center[3,10], this has been attributed to fluctuations in electrostatic potential due to Cu-Zn disorder, and associated band tailing[11]. The origin of the V_{OC} deficit is still an on-going debate[6].

Substitutional disorder within the cation sublattice of tetrahedrally bonded multinary semiconductors is a particularly likely form of disorder[12]. This can decisively alter the electronic properties of a material[13]. Substitutional disorder between Cu^+ and Zn^{2+} ions has a low enthalpic cost due to the similar ionic radii and chemical character of the two species. Density functional theory (DFT) predicts a low formation energy for the $[Cu^-_{Zn} + Zn^+_{Cu}]$ antisite defect pair[14] and there is a large body of evidence for the presence of disorder amongst Cu^+ and Zn^{2+} ions in CZTS[15–20]. Furthermore, Ref. 18–21 indicate a distinct order-disorder transition attributed to Cu-Zn substitution.

During the high-temperature synthesis of CZTS disorder can be 'frozen in' to the material as it cools to room temperature. Studies have been conducted to determine if low temperature post-deposition annealing could improve device performance and some improvements were observed from such treatments[22,23]. However, in the latter study the authors postulate that a high level of order amongst the Cu^+ and Zn^{2+} ions would require years of this treatment[23]. It is unclear if the disorder is due to slow kinetics, which could possibly be improved through optimising the processing conditions, or if the disorder is due to fundamental thermodynamic limitations for the material at room temperature[24]. In our study, we model only the thermodynamic equilibrium disorder as a function of temperature. Therefore, our model could be used to isolate disorder from equilibrium thermodynamics and kinetic limitations in experiments.

Devices made from an alloy of Cu_2ZnSnS_4 and $Cu_2ZnSnSe_4$, i.e. $Cu_2ZnSn(S_xSe_{1-x})_4$ make the highest performing devices[3,10]. In this study we focus on the pure sulfide. The V_{OC} deficit is worse in CZTS devices[6] and so potentially studying causes of the problem in this particular system could be more informative. Although ultimately the aim for this technology is to make thin-film devices from CZTS, in which the material is likely to be highly polycrystalline with grain boundaries, we focus on the bulk material. We are doing this for two reasons. Firstly, to improve the understanding of the fundamental material properties before attempting to understand a more complex system. Secondly, it has been proposed that the V_{OC} deficit in CZTSSe devices could be associated with properties of the bulk crystal[25]. It is believed that the most recent devices are not limited by interface recombination[26,27]. Furthermore, devices fabricated from single crystals have been demonstrated to have a V_{OC} deficit of 530 mV, which equals that of the record thin-film devices, indicating that the deficit could largely be due to bulk disorder[28].

Studies on various multinary semiconductors have indicated that it is not sufficient to consider only point defects to understand the defect physics of this type of compound due to the likely presence of structural disorder and extended antisite defects, which can dramatically lower the formation energy of defects[12,29]. Such system sizes would be beyond computationally feasible limits for density functional theory (DFT) calculations. However, some studies have investigated substitutional disorder through Metropolis Monte Carlo (MC) simulations of the redistribution of local structural motifs centred on the S-ions in CZTS, with energies calculated by DFT, out to nearest-neighbour interactions[29,30] and another study has been carried out using a cluster expansion of interacting dimers and trimers in CZTS[24], for system sizes of 1200 atoms and 512 atoms respectively. Prior to these studies there has been little work modelling disordered phases in CZTS, apart from one study where the choice of the disordered phase was arbitrary[31] and another investigating the configurational entropy of possible independent microstates in systems of up to 64 atoms[32].

In this study, we simulate substitutional disorder between Cu^+ and Zn^{2+} ions for system sizes of 10^5 atoms. We use on-lattice Metropolis MC simulation with a classical interatomic interaction model to calculate lattice energies, and perform Cu-Zn substitutions. We develop methods to ensure that the equilibrium disordered configuration has been reached at each simulation temperature and tools to quantify the structural disorder in the equilibrated configurations that are obtained. We assess finite size effects on the disorder process and parameterisation of the classical electrostatic interaction model to DFT calculations. We finish with the temperature dependence of thermodynamic Cu-Zn order in our model of CZTS.

2 Computational Methodology

2.1 Lattice model of Cu_2ZnSnS_4

The crystal structure of Cu_2ZnSnS_4 can be described by two inter-pentrating face-centred cubic (FCC) lattices: one of metal cations and one of sulfur anions. This is shown in Figure 1a, where green planes are a guide to the eye to distinguish the anion sublattice. We consider the sulfur sub-lattice to be

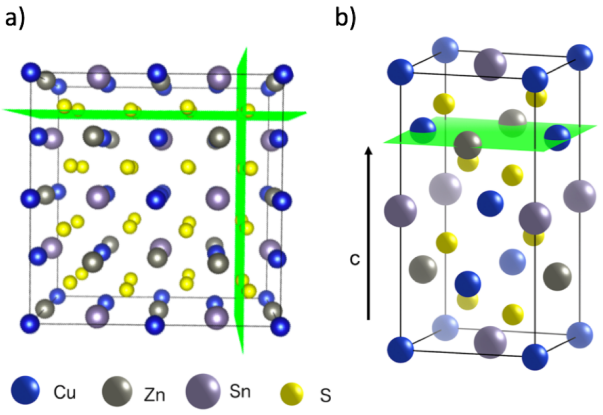

Fig. 1 Representations of the crystal structure of kesterite-structured Cu_2ZnSnS_4 where green planes are used as guides to the eye: a) supercell indicating the two inter-penetrating anion and cation sub-lattices, b) the conventional unit cell highlighting a Cu-Zn layer in the (001) planes along the *c*-axis. Visuals were produced using VESTA [33].

invariant as any substitution between ions in the cation sub-lattice and sulfur anion sub-lattice would be energetically infeasible. The sulfur sub-lattice is implicit during the MC simulations but incorporated later in calculations of lattice electrostatics. The cation lattice can be described by alternating layers of Cu-Sn and Cu-Zn in (001) planes, as shown in Fig. 1b. For computational convenience, we map this FCC lattice onto a simple cubic (SC) lattice for our simulations by introducing empty lattice sites.

The separation between lattice sites is our model is re-scaled using DFT-optimised lattice parameters of $a = b = 5.44\text{Å}$. Kesterite has a tetragonal lattice with a $\frac{c}{2a}$ ratio very close to 1 (0.998 from DFT-optimisation). We use a value of 1 in the MC simulations, which has a minor effect on the lattice energy as confirmed from explicit calculations using the General Utility Lattice Program (GULP) [34]. Lattice energies of an ordered 64 atom supercell with the exact DFT-optimised lattice parameters and an equivalent supercell with the approximated lattice parameters differed by less than 2%.

In our model we fix the position of Sn ions and simulate only two-dimensional disorder amongst Cu and Zn ions in Cu-Zn layers (Fig 1b). This is the most prevalent type of substitutional disorder observed experimentally for near stoichiometric samples [18]. To simulate only nearest-neighbour Cu-Zn disorder within the Cu-Zn layers, we use a cut-off radius of 2 lattice units (to account for the empty sites between each cation, shown in Fig. 4). Our model does not account for strain effects during Cu-Zn substitutions as it is fixed on-lattice. It has been reported that there is a small change in the *c* lattice parameter with increased disorder [35] which will not be included in our on-lattice model. However, due to the similar ionic radii of Cu and Zn we expect strain effects to be negligible.

2.2 Pair interaction model and Metropolis Monte Carlo simulation of cation disorder

The MC method can be used to calculate thermodynamic information about a system of interacting ions, which we represent on a 3D lattice as described above. We assume that the potential field of an ion is spherically symmetric and consider two-body forces acting between all pairs of ions in this system. If we know the positions of the N interacting ions on the lattice then the potential energy of the system can be calculated using equation 1, where V is the Coulombic potential between two ions and d_{ij} is the minimum distance between ions i and j [36].

$$E = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} V d_{ij} \tag{1}$$

To calculate the properties of the system, the canonical (NVT) ensemble is used where the temperature, number of ions and volume are all constant. In our model, the trial MC moves are swaps between nearest-neighbour Cu and Zn ions.

Using a standard MC method for our system would involve placing each of the N ions at random positions in the lattice to define a random point in the 3N-dimensional configuration space. However, many configurations are very improbable so performing this calculation for every possible configuration would be inefficient and unnecessary to sufficiently evaluate the ensemble. The custom MC code in this study makes use of the Metropolis modified MC scheme [36]. In this implementation of the MC method, instead of choosing configurations randomly and then weighting them, the Metropolis algorithm considers the relative probability of a system being in a new configuration, β, to that of being in the current configuration, α. This is shown in equation 2, where E_α is the energy of state α, E_β is the energy of state β and Q is the partition function. For most systems, calculating the value of the partition function requires the summation over a large number of states. However, within the Metropolis scheme, in the expression for the probability of the trial Q cancels out.

$$\frac{p_\beta}{p_\alpha} = \frac{e^{-\frac{E_\alpha}{k_b T}}}{Q}\frac{Q}{e^{-\frac{E_\alpha}{k_b T}}} = e^{-\frac{E_\beta - E_\alpha}{k_B T}} \tag{2}$$

The relative probabilities of the two states are completely determined by the energy difference, such that if:

$$\Delta E = E_\beta - E_\alpha \leq 0 \text{ then } \frac{p_\beta}{p_\alpha} \geq 1 \tag{3}$$

and if

$$\Delta E = E_\beta - E_\alpha > 0 \text{ then } \frac{p_\beta}{p_\alpha} < 1 \tag{4}$$

It is then decided if this new configuration should be added to the trajectory of the system (towards the minimum energy configuration), or not, based on the probability of the new configuration relative to the current configuration. If the relative probability is ≥ 1, as shown in equation 3, then the move is accepted and added to the trajectory. However, if the relative probability is < 1 then the move will only be accepted if $e^{-\frac{\Delta E}{k_B T}} \geq$ a random number generated between 0 and 1. In our simulations, we perform lattice energy summations out to a finite radius to obtain ΔE, but as we are using periodic boundary conditions the upper limit for the cut off radius is half the minimum dimension of the system. Details of the convergence in ΔE with respect to the cut off radius used in the lattice summations are given in the SI. Simulations are performed in parallel over different temperatures using GNU parallel[37].

3 Results and discussion

3.1 Equilibration

Due to the stochastic nature of the trajectory from an initial configuration in the MC method, we cannot draw any conclusions about the equilibrium thermodynamic properties of that system at the given simulation temperature until equilibrium has been reached. The simulation steps required to reach this point is the 'equilibration time'. Equilibration is often considered as the point at which the value of a quantity of interest, which initially changes by a large amount, eventually converges to fluctuating about a steady average value. This is dependent upon the principle that a system in equilibrium spends the overwhelming majority of its time in a small subset of states in which its properties take a narrow range of values[39].

Our MC model for Cu-Zn disorder is analogous to the Ising model of a ferromagnet and we describe the rationale for our equilibration procedure by referring to this common example. In the case of an Ising model, the trial moves in the Metropolis algorithm are spin flips, whereas in our model the trial moves are swaps between Cu and Zn ions. For the Ising model one MC step corresponds to attempting a trial spin-flip at all sites in the system once. Similarly, for our model one MC step corresponds to sweeping across the entire lattice and attempting a near-neighbour Cu-Zn swap at each Cu and Zn site. In the case of the Ising model it is usually the average magnetisation of the system, or internal energy, as a function of temperature that are the quantities of interest. For our system, we are interested in the configuration of the ions (and extent of thermodynamic disorder) and the corresponding distribution of the electrostatic potential across the system, as this can be related to the observed band tailing. We now explore two methods to gauge when the system has reached the equilibrium disordered configuration at each simulation temperature: the pair correlation function (PCF) for information on the structural disorder and also the variance of the distribution of on-site electrostatic potentials of species in the system.

3.1.1 Pair correlation functions from ordered and disordered initial lattices We first attempted to run two simulations for each temperature, one starting from an initial ordered lattice and one from an initial disordered lattice (produced by 'shuffling' Cu and Zn ions in the ordered lattice), until both simulations converged to the same equilibrium configuration. To gauge the point at which this had been reached, we compare the pair correlation functions (PCFs) for each configuration, an example of this analysis is given in Fig. 2.

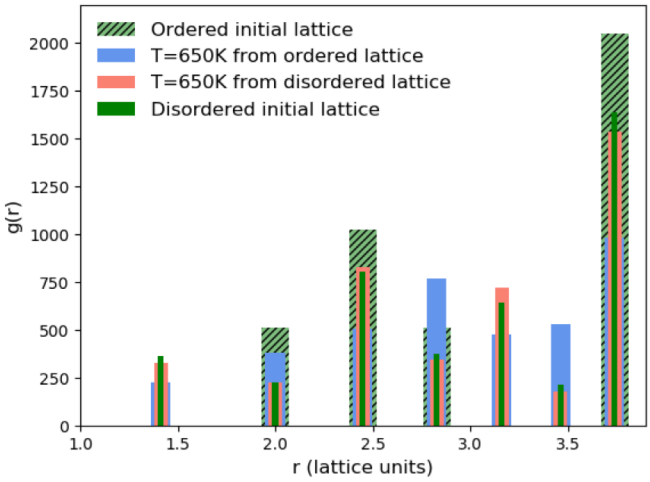

Fig. 2 The pair correlation function (PCF) between pairs of Zn ions in Cu_2ZnSnS_4. PCFs of an initial ordered lattice are plotted with that of a disordered initial lattice as reference points as well as systems that have been evolved from both of these initial configurations at T=650K. Widths of the bars plotted are arbitrarily chosen to ensure all data is visible.

We found that systems evolved from an initially disordered lattice required a substantially larger number of MC steps to evolve away from the initial configuration. This can be seen in Fig. 2 from the Zn-Zn PCFs. The most noticeable feature when comparing the PCF for an ordered initial lattice to that of a disordered lattice is the emergence of a new nearest-neighbour Zn-Zn peak at $\sqrt{2}$ due to the clustering of Zn ions once Cu and Zn ions have been allowed to substitute. This is discussed further in section 3.2 as an order parameter, but for now we just remark that the peak is largest for the disordered initial lattice and decreases for the system evolved from this initial configuration at moderate simulation temperatures. After a considerably large number of MC steps the peak for the two systems evolved from the ordered and disordered initial lattices were not of the same height. This observation may be explained by the entropic cost in going from a disordered to a

more ordered system, suggesting that this method may not be computationally efficient. We therefore adopted an alternative approach to check for equilibration, as outlined below.

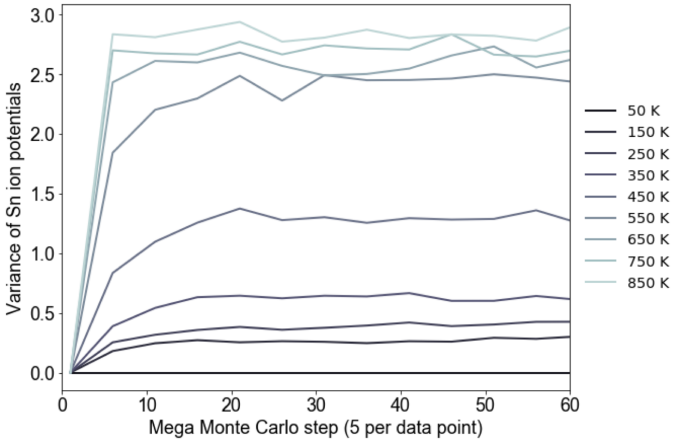

Fig. 3 Variance in the distribution of the on-site electrostatic potential of Sn ions in a $80\times80\times80$ Cu_2ZnSnS_4 system (containing 512,000 ions in total) across a range of simulation temperatures. Each mega Monte Carlo step corresponds to sweeping across the lattice and attempting 100 trial moves per lattice site.

3.1.2 Variance in the distribution of on-site electrostatic potentials Our second method is analogous to using the point at which the average magnetisation fluctuates about a steady value in the Ising model, as discussed earlier. For our system, we instead look for the number of MC steps required to reach a point after which the variance of the distribution of on-site electrostatic potentials of Sn ions fluctuates about a steady value. We use Sn ions because we have fixed the locations of Sn ions in our simulations, making them stationary reference points. There is one crystallographically distinct Sn ion in Cu_2ZnSnS_4. We start from an ordered lattice and as all ions are on their correct lattice sites, there is only one unique chemical environment for the Sn ions, regardless of system size. The variance in the distribution of electrostatic potentials for Sn ions in the ordered lattice is therefore zero. As the system evolves, and Cu and Zn ions are substituted, more unique chemical environments emerge for Sn ions in the system.

We calculate the on-site electrostatic potential out to a finite cut off radius of 10 lattice units. We check for a radius after which the value of the potential converges with increased cut-off. Details of our convergence test are included in the SI. An example of a test to determine a suitable number of MC steps for equilibration (i.e. steps to run before collecting data on the system) is shown in Fig 3. We perform this check for the largest system size and the whole simulation temperature range we study. A larger system may require a considerably larger number of MC steps to equilibrate and we perform the check for each temperature because if there is a phase transition (as suggested in several works [18–21]), there could be 'critical slowing down' close to the transition temperature.

From figure 3 we take 20 mega MC steps as a suitable number of equilibration steps before we start collecting data for our largest system size of $80\times80\times80$, which corresponds to 512,000 ions. One mega MC step corresponds to attempting on average 100 trial Cu-Zn substitutions per site per data point, i.e. 100 sweeps of the lattice per mega MC step. The absence of variance when the simulation temperature is at 0 K in this example is because at this simulation temperatures the system is ordered and therefore there is only one unique chemical environment for Sn, giving a variance in the on-site electrostatic potential of Sn ions of zero (with some numerical noise).

3.2 Order parameters

To quantify the extent of substitutional Cu-Zn disorder in our system, we consider two order parameters to enable us to investigate long- and short-ranged order in the system.

3.2.1 Pair correlation functions Pair correlation functions (PCFs) show the number of pairs of particular species with particular separations within the system. We generate reference PCFs of ordered and disordered systems (using equilibrated configurations at very low and very high temperatures respectively) of the same size. The most noticeable feature in the PCFs of the system was the emergence of a new nearest-neighbour peak in the Zn-Zn PCF at $\sqrt{2}$. This can be explained using Fig. 4. In the ordered lattice the shortest Zn-Zn spacing is 2 lattice units. Once Cu and Zn begin to substitute a new shortest Zn-Zn spacing of $\sqrt{2}$ lattice units becomes possible. An increase in the intensity of this $\sqrt{2}$ peak indicates more clustering of Zn ions and so provides insights into the extent of short-ranged disorder in the system. The same analysis is not possible using the Cu-Cu PCF because a $\sqrt{2}$ Cu-Cu separation is present between the (001) planes in the ordered lattice.

3.2.2 Cation site occupancy An order parameter used in experimental literature to quantify Cu-Zn disorder in kesterite-structured Cu_2ZnSnS_4, $Cu_2ZnSnSe_4$ and $Cu_2ZnSn(S_xSe_{1-x})_4$ is the 'Q order parameter', based on cation site occupancies [18]. In ordered CZTS, Cu ions occupy 2*c* sites in the Cu-Zn layers indicated in Fig. 1b and 2*a* sites in the Cu-Sn layers. Sn ions occupy the 2*b* sites and Zn ions occupy the 2*d* sites [21]. In completely disordered CZTS Cu and Zn are found evenly distributed over 2*c* and 2*d* sites, indicating disorder in the Cu-Zn layers. A measure of increasing order is when Cu shows a preference to occupy 2*c* sites and Zn to occupy 2*d* sites. For ordered CZTS, Q equals 1, corresponding to all Zn ions on 2*d*

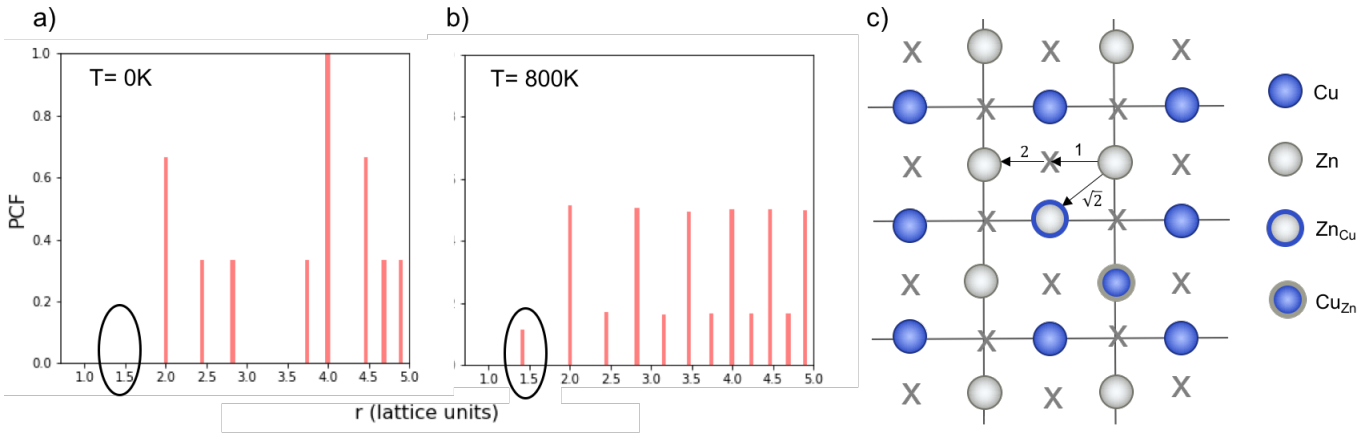

Fig. 4 Normalised Zn-Zn pair correlation functions (PCFs) at 0K (a) and 800K (b) for an a-b Cu-Zn plane in the cation sub-lattice of Cu_2ZnSnS_4 with structure shown in (c). Crosses denote the gap sites used in our lattice model to map an fcc lattice onto a sc lattice. Before a Cu-Zn swap, the nearest-neighbour Zn-Zn pair is 2 lattice units apart. After a Cu-Zn swap there is a Zn-Zn pair separated by r=$\sqrt{2}$. The 0K PCF is for the perfectly ordered lattice before any Cu-Zn substitutions have occurred and shows a Zn-Zn PCF peak intensity of zero at r=$\sqrt{2}$. The 800K PCF shows an increase in the peak intensity at r=$\sqrt{2}$, once Cu and Zn ions begin to substitute.

sites and all Cu ions on 2*c* sites. For completely disordered CZTS Q equals 0, corresponding to no preference for Cu or Zn to occupy their correct crystallographic site.

$$Q = \frac{[Cu_{2c} + Zn_{2d}] - [Zn_{2c} + Cu_{2d}]}{[Cu_{2c} + Zn_{2d}] + [Zn_{2c} + Cu_{2d}]} \quad (5)$$

However, it is possible that this metric may overestimate the extent of disorder in a system as locally ordered domains, displaced relative to the configuration of the initial lattice, would be considered as disordered. We therefore compare the extent of order at each simulation temperature inferred from our PCF analysis to that suggested by the Q parameter as we increase the system size to check for the formation of locally ordered domains. A decrease in the Q order parameter and an increase in Zn-Zn PCF $\sqrt{2}$ peak intensity correspond to a reduction in order in the system. For the case of locally ordered domains, a low Q parameter (suggesting large extents of Cu-Zn disorder) could coincide with a relatively small $\sqrt{2}$ Zn-Zn PCF peak, suggesting long-range disorder, but short range order within the Cu-Zn planes.

3.3 Finite size effects

To investigate finite size effects within our model, we perform simulations for system sizes ranging from 12×12×12 (= 1,728 ions) to 80×80×80 (= 512,000 ions). We investigate the disorder behaviour of the systems as a function of temperature using the Zn-Zn PCF analysis and Q order parameter outlined in the previous section in Fig. 5a and b respesctively. Fig. 5a shows the increase in the intensity of the Zn-Zn PCF at $r = \sqrt{2}$ with temperature. Fig. 5b shows the decrease in the Q order parameter from 1 to 0 with increased simulation temperature. Both order parameters show approximately the same disorder temperature.

Our model shows clear signs of finite size effects for the smallest systems, in the regime used in previous studies. We consider a 72×72×72 size system to give a converged disorder process with respect to system size. The high-temperature Zn-Zn nearest-neighbour PCF peak shown in Fig. 5a can be understood as the Cu-Zn planes beginning to melt. Within the cation sublattice, there are 12 sites $\sqrt{2}$ apart. In the Q=0 disorder regime, Cu shows no preference to occupy the *2c* sites and Zn to occupy the *2d* sites, we can therefore expect the density of Zn $\sqrt{2}$ away from every other Zn to converge towards 2/12 (approximately 0.167). However, we cannot expect the PCF peak at high-temperatures to reach a complete plateau at high-temperatures as is seen for the Q order parameter due to the lower stoichiometric ration of Zn relative to Cu in Cu_2ZnSnS_4.

3.4 Parameterisation of interaction energies

Equation 6 can be used to calculate the electrostatic interaction between a pair of ions, where q_1 and q_2 are the bare formal charges, r is the separation of the point charges, ε_r is the bulk static relative dielectric constant of the perfect crystal and ε_0 is the permittivity of free space. The interaction between all pairs of ions within the system are summed when calculating the change in the lattice energy of the system after performing a trial move in our MC simulations.

$$E_{electrostatic} = \frac{q_1 q_2}{4\pi\varepsilon_0\varepsilon_r} e^2 \frac{1}{r} = \frac{q_1 q_2}{r} I_{electrostatic} \quad (6)$$

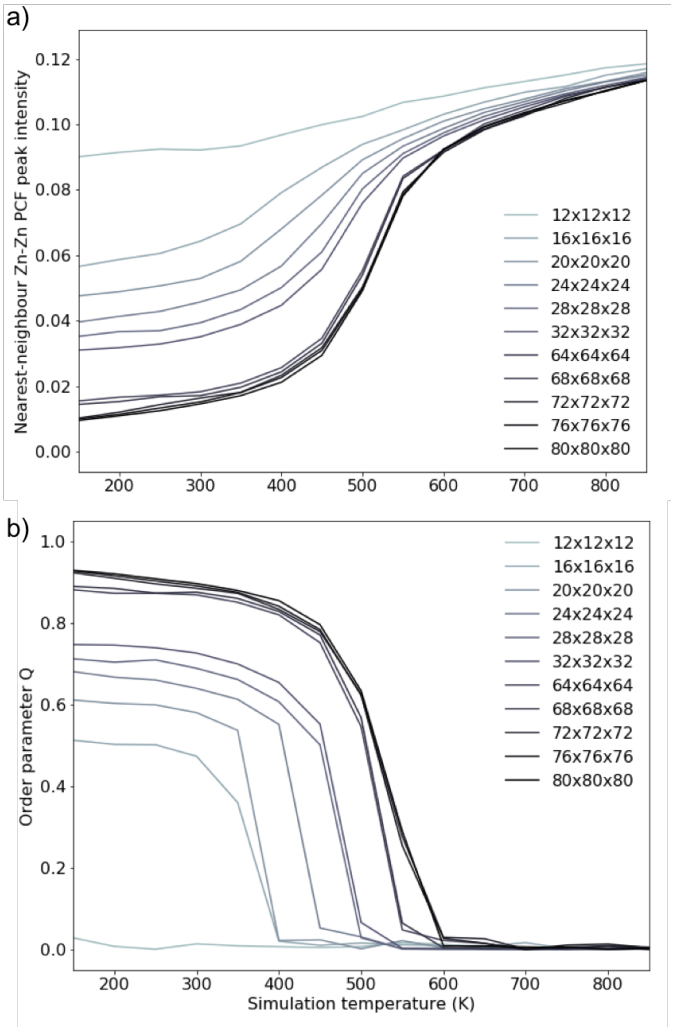

Fig. 5 To assess finite-size effects the Q order parameter based on Cu and Zn site occupancies in Cu_2ZnSnS_4 for systems of various sizes at thermodynamic equilibrium across simulation temperatures ranging from 150 to 850 K (a). Q=1 corresponds to a fully ordered lattice and Q=0 corresponds to complete Cu-Zn disorder within the (001) plane. The nearest-neighbour ($r = \sqrt{2}$) Zn-Zn pair correlation function peak intensity as a function of simulation temperature, indicating clustering of Zn ions and deviation from the perfectly ordered lattice with a $r = \sqrt{2}$ peak intensity of 0 (b).

We separate out the scaling factor $I_{electrostatic}$ in our simulations so that we can investigate the effect of this parameter on the disorder behaviour of the system. We first calculate the value of $I_{electrostatic}$ with the bulk, macroscopic dielectric constant of Cu_2ZnSnS_4 (both ionic and electronic components), ε_r, calculated from density functional theory (DFT) using the Heyd-Scuseria-Ernzerhof (HSE06) hybrid-functional[40] and the VASP code[41,42]. We use a value of 13.2 for ε_r and 3.8 $\AA$ for r, which is the separation of nearest-neighbour Cu-Zn ions. This gives a value of -0.284 eV for $I_{electrostatic}$. The limitation of this treatment is that it only accounts for the Coulombic interaction between the point charges and neglects any changes in the local electronic structure during defect formation.

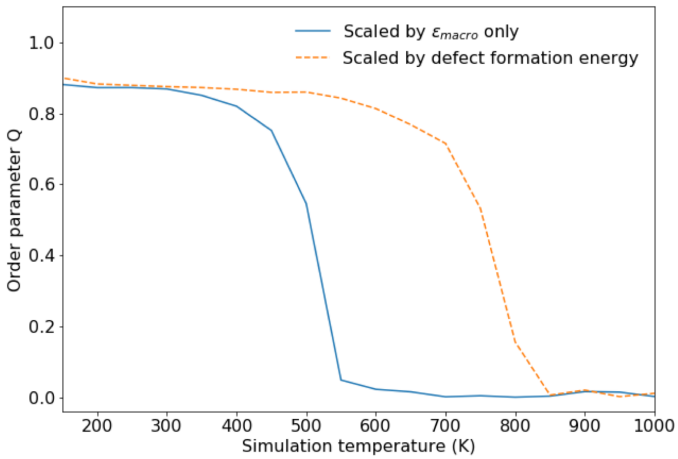

Fig. 6 The Q order parameter based on the occupancy of Cu 2*c* and Zn 2*d* sites in a system of 64×64×64 Cu_2ZnSnS_4 ions with pair interaction energies scaled by the macroscopic static dielectric constant of Cu_2ZnSnS_4 calculated by density functional theory (DFT) only (solid line) or also by DFT calculations of $[Cu_{Zn}^{-} + Zn_{Cu}^{+}]$ antisite pair formation energy (dashed line).

We also investigated parameterising the interaction energy, $I_{electrostatic}$, using DFT formation energies of Cu-on-Zn and Zn-on-Cu antisite defects. We compare the formation energies of a nearest-neighbour $[Cu_{Zn}^{-} + Zn_{Cu}^{+}]$ antisite defect pair calculated using classical interatomic potentials with the GULP program[34] to that obtained using quantum mechanical hybrid-density functional theory with HSE06, where effects of the electronic structure during defect formation are accounted for to scale $I_{electrostatic}$. Values obtained for the formation energy were 0.20 eV and 0.30 eV respectively. Further details of the DFT calculation is included in the SI. The defect formation energy obtained using HSE06 is approximately 1.5 times that obtained with GULP, we therefore trialled scaling the interaction energy by the same amount and so $I_{DFT} = 1.5 \times I_{electrostatic} = -0.425$ eV. Through doing this we aimed to scale the macroscopic dielectric constant to include the microscopic screening effects of the electronic structure during defect formation. Although our DFT $[Cu_{Zn}^{-} + Zn_{Cu}^{+}]$ antisite defect calculations still do not account for all of the physics of the system. These calculations are performed in the absence of temperature and at the dilute defect limit using the supercell method.

The resulting Q parameter disorder profiles with and without the additional scaling of the interaction energy from DFT defect formation energies are shown in Fig. 6. Experimen-

tally, the Cu-Zn order-disorder transition is at ~ 200°C for the pure selenide ($Cu_2SnSnSe_4$) and ~ 270°C for the pure sulfide (Cu_2SnSnS_4)[43]. Fig. 6 shows that further scaling of the interaction energy with DFT defect formation energies takes the order-disorder transition of the model away from the transition temperature observed experimentally. An explanation for this behaviour could be that the large local strain effects that are included in the DFT calculation are over-shadowed over larger distances where it is the bare formal charges and macroscopic dielectric response that determines the behaviour of the system. However, note that the only effect of the scaling of the interaction parameter is to shift the temperature-dependence of the order-disorder transition. The shape of the disorder profiles and maximum Q values are insensitive to the scaling of the interaction parameter as shown in Fig. 6. The final results we present in the next section are based on electrostatics scaled using the static dielectric constant calculated by DFT only.

3.5 Thermodynamically limited Cu-Zn order

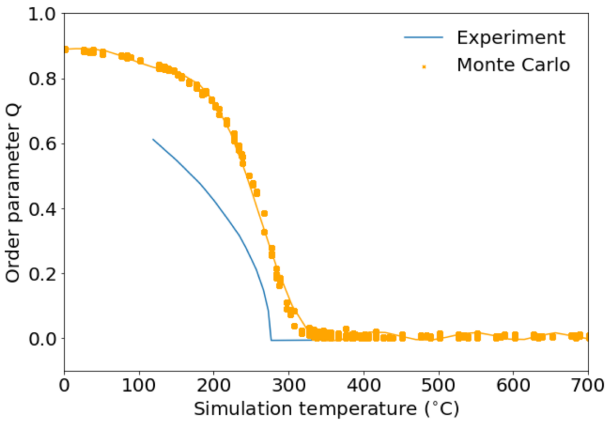

Fig. 7 The Q order parameter based on Cu and Zn site occupancies in Cu_2ZnSnS_4 for $72\times72\times72$ (= 373,248 ions) averaged over 20 independent Monte Carlo simulations and anomalous X-ray powder diffraction data for $Cu_2ZnSnSe_4$ from Ref. 18. Experimental data has been shifted by 70°C to account for the difference in the order-disorder transition temperature for the pure sulfide and pure selenide reported in Ref. 43.

For our production runs to probe Cu-Zn disorder in stoichiometric CZTS (Cu_2ZnSnS_4), we use a system size of $72\times72\times72$ (= 373,248 ions). This setup was found to give order-disorder behaviour that is converged with respect to system size (see Fig. 5). To improve our statistics, we perform 20 independent simulations, each using different random number seeds. There is a plateau in the maximum obtainable Cu-Zn order as described by the Q order parameter with decreasing temperature, where Q = 1.0 corresponds to perfect order in the Cu-Zn (001) planes and Q =8 0.0 corresponds to no preference for Cu or Zn to occupy their correct crystallographic site. Even for a system that is fully equilibrated at room temperature, overcoming any kinetic barriers, a value of Q=0.9 is obtained. This finding fits well with experimental works that have been unable to achieve high Cu-Zn order, even with long and low temperature anneals[23,44].

The thermodynamically limited disorder in the Cu-Zn sublattice arises due to the low energy for antisite formation as Cu and Zn have similar charge and ionic radii. The gain in configurational entropy, even at low temperatures, is sufficient to cause a disordered sublattice as described by the Monte Carlo simulations. Experimentally, the situation can be further complicated by kinetic barriers, whicSh give rise to metastable configurations. Our model suggests, however, that no improvement beyond a Q value of approximately 0.9 will be achievable at room temperature. We also note that there is no clear difference in the disorder behaviour described by our PCF analysis and the Q order parameter (Fig. 5), suggesting that locally ordered domains are not present in disordered configurations of our model.

4 Summary and further work

In summary, we have simulated substitutional Cu-Zn disorder in CZTS (Cu_2ZnSnS_4) using a MC model based on electrostatic pair interactions that have been scaled using the static dielectric constant of the material obtained from DFT calculations. Our model predicts a Cu-Zn order-disorder transition temperature in pure sulfide CZTS at approximately 250°C (Fig. 7), which is in good agreement with transition temperatures reported experimentally[19,43]. The value of the Q parameter as a function of temperature calculated for disordered lattice configurations from our simulations is also in good agreement with values reported experimentally[18].

We are able to simulate at temperatures much lower than those typically accessed during the annealing treatments of CZTS and our model suggests that at experimentally relevant low temperatures (of 0°C or larger) the Q order parameter based on the occupancy of Cu 2*c* and Zn 2*d* crystallographic sites plateaus with decreasing temperature. As our model considers only thermodynamic disorder, this suggests that CZTS is thermodynamically limited to not achieve Cu-Zn order corresponding to a Q parameter larger than 0.9, where 1.0 corresponds to fully ordered CZTS. Our model suggests that some Cu-Zn disorder will always be present in the material. The disorder present in our model of stoichiometric CZTS appears to be conistent with experimental studies[18]. However, it is possible that disorder may be suppressed in off-stoichiometric CZTS types that have been synthesised and structurally validated using neutron scattering[46]. Extending our model to

off-stoichiometric CZTS systems could provide insights into disorder suppression in compounds with slightly different stoichiometries. However, the exact role of Cu-Zn disorder on device performance is still not well understood.

5 Data access statement

Git repo for Eris (ErisCodes) and analysis jupyter notebooks (tidy up!). DFT Cu-Zn antisite defect pair calculation + dielectric constant calculation.

6 Acknowledgements

Funding + Jonathan (some of the defect calculations performed on Balena + early convergence tests w.r.t k-points and ENCUT). Use of cx1: Imperial College High Performance Computing Service, doi: 10.14469/hpc/2232

Mark Weller and Benjamin Morgan for discussions.

References

1 X. Liu, Y. Feng, H. Cui, F. Liu, X. Hao, G. Conibeer, D. B. Mitzi and M. Green, *Progress in Photovoltaics: Research and Applications*, 2016, **24**, 879–898.

2 W. Shockley and H. J. Queisser, *Journal of Applied Physics*, 1961, **32**, 510.

3 W. Wang, M. T. Winkler, O. Gunawan, T. Gokmen, T. K. Todorov, Y. Zhu and D. B. Mitzi, *Advanced Energy Materials*, 2013, **4**, 1301465.

4 S. Tajima, M. Umehara, M. Hasegawa, T. Mise and T. Itoh, *Progress in Photovoltaics: Research and Applications*, 2016, **25**, 14–22.

5 P. Jackson, R. Wuerz, D. Hariskos, E. Lotter, W. Witte and M. Powalla, *physica status solidi (RRL) - Rapid Research Letters*, 2016, **10**, 583–586.

6 S. Bourdais, C. Choné, B. Delatouche, A. Jacob, G. Larramona, C. Moisan, A. Lafond, F. Donatini, G. Rey, S. Siebentritt, A. Walsh and G. Dennler, *Advanced Energy Materials*, 2016, **6**, 1502276.

7 R. Aninat, L.-E. Quesada-Rubio, E. Sanchez-Cortezon and J.-M. Delgado-Sanchez, *Thin Solid Films*, 2017, **633**, 146 – 150.

8 S. K. Wallace, D. B. Mitzi and A. Walsh, *ACS Energy Letters*, 2017, **2**, 776–779.

9 J. Li, D. Wang, X. Li, Y. Zeng and Y. Zhang, *Advanced Science*, 2018, 1700744.

10 T. K. Todorov, K. B. Reuter and D. B. Mitzi, *Advanced Materials*, 2010, **22**, E156–E159.

11 T. Gokmen, O. Gunawan, T. K. Todorov and D. B. Mitzi, *Applied Physics Letters*, 2013, **103**, 103506.

12 L. L. Baranowski, P. Zawadzki, S. Lany, E. S. Toberer and A. Zakutayev, *Semiconductor Science and Technology*, 2016, **31**, 123004.

13 S. Lany, A. N. Fioretti, P. P. Zawadzki, L. T. Schelhas, E. S. Toberer, A. Zakutayev and A. C. Tamboli, *Phys. Rev. Materials*, 2017, **1**, 035401.

14 S. Chen, J.-H. Yang, X. G. Gong, A. Walsh and S.-H. Wei, *Physical Review B*, 2010, **81**, 245204.

15 S. Schorr, *Solar Energy Materials and Solar Cells*, 2011, **95**, 1482–1488.

16 T. Washio, H. Nozaki, T. Fukano, T. Motohiro, K. Jimbo and H. Katagiri, *Journal of Applied Physics*, 2011, **110**, 074511.

17 B. G. Mendis, M. D. Shannon, M. C. Goodman, J. D. Major, R. Claridge, D. P. Halliday and K. Durose, *Progress in Photovoltaics: Research and Applications*, 2012, **22**, 24–34.

18 D. M. Többens, G. Gurieva, S. Levcenko, T. Unold and S. Schorr, *physica status solidi (b)*, 2016, **253**, 1890–1897.

19 J. J. S. Scragg, L. Choubrac, A. Lafond, T. Ericson and C. Platzer-Björkman, *Applied Physics Letters*, 2014, **104**, 041911.

20 G. Rey, A. Redinger, J. Sendler, T. P. Weiss, M. Thevenin, M. Guennou, B. E. Adib and S. Siebentritt, *Applied Physics Letters*, 2014, **105**, 112106.

21 A. Ritscher, M. Hoelzel and M. Lerch, *Journal of Solid State Chemistry*, 2016, **238**, 68–73.

22 G. Rey, T. Weiss, J. Sendler, A. Finger, C. Spindler, F. Werner, M. Melchiorre, M. Hla, M. Guennou and S. Siebentritt, *Solar Energy Materials and Solar Cells*, 2016, **151**, 131 – 138.

23 K. Rudisch, Y. Ren, C. Platzer-Björkman and J. Scragg, *Applied Physics Letters*, 2016, **108**, 231902.

24 K. Yu and E. A. Carter, *Chemistry of Materials*, 2016, **28**, 864–869.

25 A. Polizzotti, I. L. Repins, R. Noufi, S.-H. Wei and D. B. Mitzi, *Energy & Environmental Science*, 2013, **6**, 3171.

26 S. Siebentritt, *Nature Energy*, 2017, **2**, 840–841.

27 K. Sun, C. Yan, F. Liu, J. Huang, F. Zhou, J. A. Stride, M. Green and X. Hao, *Advanced Energy Materials*, 2016, **6**, 1600046–n/a.

28 M. A. Lloyd, D. Bishop, O. Gunawan and B. McCandless, 2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC), 2016.

29 P. Zawadzki, A. Zakutayev and S. Lany, *Physical Review B*, 2015, **92**, 201204.

30 P. Zawadzki, A. Zakutayev and S. Lany, *Physical Review Applied*, 2015, **3**, 034007.

31 J. J. S. Scragg, J. K. Larsen, M. Kumar, C. Persson, J. Sendler, S. Siebentritt and C. P. Björkman, *physica status solidi (b)*, 2015, **253**, 247–254.

32 S. Shang, Y. Wang, G. Lindwall, N. R. Kelly, T. J. Anderson and Z.-K. Liu, *The Journal of Physical Chemistry C*, 2014, **118**, 24884–24889.

33 K. Momma and F. Izumi, *Journal of Applied Crystallography*, 2011, **44**, 1272–1276.

34 J. D. Gale and A. L. Rohl, *Molecular Simulation*, 2003, **29**, 291–341.

35 M. Quennet, A. Ritscher, M. Lerch and B. Paulus, *Journal of Solid State Chemistry*, 2017, **250**, 140 – 144.

36 N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.*, 1953, **21**, 1087–1092.

37 O. Tange, *The USENIX Magazine*, 2011, **36**, 42–47.

38 D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, 1987.

39 M. E. J. Newman and G. T. Barkema, *The Ising Model and the Metropolis Algorithm*, Oxford University Press, 1999.

40 J. Heyd, G. E. Scuseria and M. Ernzerhof, *Journal of Chemical Physics*, 2003, **118**, 8207–8215.

41 G. Kresse and J. Furthmüller, *Computational Materials Science*, 1996, **6**, 15–50.

42 G. Kresse and J. Furthmüller, *Physical Review B*, 1996, **54**, 11169–11186.

43 T. Gershon, D. Bishop, P. Antunez, S. Singh, K. W. Brew, Y. S. Lee, O. Gunawan, T. Gokmen, T. Todorov and R. Haight, *Current Opinion in Green and Sustainable Chemistry*, 2017, **4**, 29–36.

44 P. Bais, M. T. Caldes, M. Paris, C. Guillot-Deudon, P. Fertey, B. Domengs and A. Lafond, *Inorganic Chemistry*, 2017, **56**, 1177911786.

45 K. Biswas, S. Lany and A. Zunger, *Applied Physics Letters*, 2010, **96**, 201902.

46 A. Ritscher, A. Franz, S. Schorr and M. Lerch, *Journal of Alloys and Compounds*, 2016, **689**, 271–277.