7

Automated Computation of Materials Properties

Cormac Toher[1], Corey Oses[1], and Stefano Curtarolo[2]

[1]*Duke University, Department of Mechanical Engineering and Materials Science, 144 Hudson Hall, Durham, NC 27708, USA*
[2]*Duke University, Materials Science, Electrical Engineering, Physics and Chemistry, 144 Hudson Hall, Durham, NC 27708, USA*

7.1 Introduction

Materials informatics requires large repositories of materials data to identify trends in and correlations between materials properties, as well as for training machine learning models. Such patterns lead to the formulation of descriptors that guide rational materials design. Generating large databases of computational materials properties requires robust, integrated, automated frameworks [1]. Built-in error correction and standardized parameter sets enable the production and analysis of data without direct intervention from human researchers. Current examples of such frameworks include *A*utomatic *FLOW* (AFLOW) [2–10], Materials Project [11–14], *O*pen *Q*uantum *M*aterials *D*atabase (OQMD) [15–17], the Computational Materials Repository [18] and its associated scripting interface *A*tomic *S*imulation *E*nvironment (ASE) [19], *A*utomated *I*nteractive *I*nfrastructure and *Da*tabase for Computational Science (AiiDA) [20–22], and the Open Materials Database at `httk.openmaterialsdb.se` with its associated *H*igh-*T*hroughput *T*ool*k*it (HTTK). Other computational materials science resources include the aggregated repository maintained by the *No*vel *Ma*terials *D*iscovery (NoMaD) Laboratory [23], and the *T*heoretical *C*rystallography *O*pen *D*atabase (TCOD) [24]. For this data to be consumable by automated machine learning algorithms, it must be organized in programmatically accessible repositories [4, 5, 7, 11, 12, 15, 23]. These frameworks also contain modules that combine and analyze data from various calculations to predict complex thermomechanical phenomena, such as lattice thermal conductivity and mechanical stability.

Computational strategies have already had success in predicting materials for applications including photovoltaics [25], water splitters [26], carbon capture and gas storage [27, 28], nuclear detection and scintillators [29–32], topological insulators [33, 34], piezoelectrics [35, 36], thermoelectric materials [37–40], catalysis [41], and battery cathode materials [42–44]. More recently, computational

materials data has been combined with machine learning approaches to predict electronic and thermomechanical properties [45, 46] and to identify superconducting materials [47]. Descriptors are also being constructed to describe the formation of disordered materials and have recently been used to predict the glass-forming ability (GFA) of binary alloy systems [48]. These successes demonstrate that accelerated materials design can be achieved by combining structured data sets generated using autonomous computational methods with intelligently formulated descriptors and machine learning.

7.2 Automated Computational Materials Design Frameworks

Rapid generation of materials data relies on automated frameworks such as AFLOW [2–6], Materials Project's `pymatgen` [13] and `atomate` [14], OQMD [15–17], ASE [19], and AiiDA [21]. The general automated workflow is illustrated in Figure 7.1. These frameworks begin by creating the input files required by the electronic structure codes that perform the quantum mechanical level calculations, where the initial geometry is generated by decorating structural prototypes (Figure 7.1a,b). They execute and monitor these calculations, reading any error messages written to the output files and diagnosing calculation failures. Depending on the nature of the errors, these frameworks are equipped with a catalog of prescribed solutions – enabling them to adjust the appropriate parameters and restart the calculations (Figure 7.1c). At the end of a successful calculation, the frameworks parse the output files to extract the relevant materials data such as total energy, electronic bandgap, and relaxed cell volume. Finally, the calculated properties are organized and formatted for entry into machine-accessible, searchable, and sortable databases.

In addition to running and managing the quantum mechanical level calculations, the frameworks also maintain a broad selection of post-processing libraries for extracting additional properties, such as calculating X-ray diffraction (XRD) spectra from relaxed atomic coordinates, and the formation enthalpies for the convex hull analysis to identify stable compounds (Figure 7.1d). Results from calculations of distorted structures can be combined to calculate thermal and elastic properties [2, 49–51], and results from different compositions and structural phases can be amalgamated to generate thermodynamic phase diagrams.

7.2.1 Generating and Using Databases for Materials Discovery

A major aim of high-throughput computational materials science is to identify new, thermodynamically stable compounds. This requires the generation of new materials structures, which have not been previously reported in the literature, to populate the databases. The accuracy of analyses involving sets of structures, such as that used to determine thermodynamic stability, is contingent on sufficient exploration of the full range of possibilities. Therefore, autonomous materials design frameworks such as AFLOW use crystallographic prototypes to generate new materials entries consistently and reproducibly.
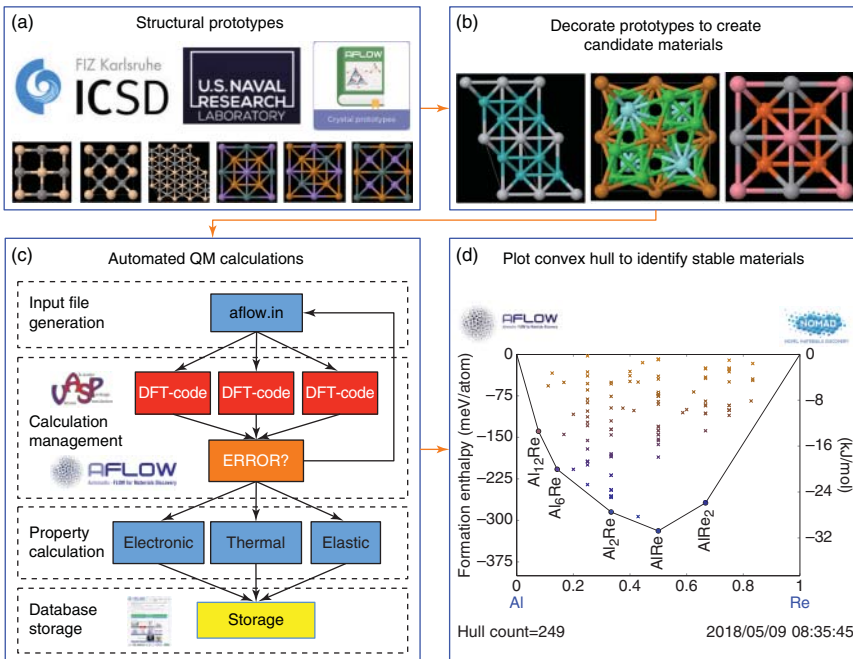
Figure 7.1 Computational materials data generation workflow. (a) Crystallographic prototypes are extracted from databases such as the ICSD or the NRL crystal structure library, or are generated by enumeration algorithms. The illustrated examples are for the rocksalt, zincblende, wurtzite, Heusler, inverse Heusler, and half-Heusler structures. (b) New candidate materials are generated by decorating the atomic sites with different elements. (c) Automated DFT calculations are used to optimize the geometric structure and calculate energetic, electronic, thermal, and elastic properties. Calculations are monitored to detect errors. The input parameters are adjusted to compensate for the problem, and the calculation is rerun. Results are formatted and added to an online data repository to facilitate programmatic access. (d) Calculated data is used to plot the convex hull phase diagrams for each alloy system to identify stable compounds.

Crystallographic prototypes are the basic building blocks used to generate the wide range of materials entries involved in computational materials discovery. These prototypes are based on (i) structures commonly observed in nature [52, 53], such as the rocksalt, zincblende, wurtzite, or Heusler structures illustrated in Figure 7.1b, as well as (ii) hypothetical structures, such as those enumerated by the methods described in Refs. [54, 55]. The AFLOW Library of Crystallographic Prototypes [53] is also available online at aflow.org/CrystalDatabase/, where users can choose from hundreds of crystal prototypes with adjustable parameters, which can be decorated to generate new input structures for materials science calculations.

New materials are then generated by decorating the various atomic sites in the crystallographic prototype with different elements. These decorated prototypes serve as the structural input for *ab initio* calculations. A full relaxation of the geometries and energy determination follows, from which phase diagrams for

stability analyses can be constructed. The resulting materials data are then stored in an online data repository for future consideration.

The phase diagram of a given alloy system can be approximated by considering the low-temperature limit in which the behavior of the system is dictated by the ground state [56, 57]. In compositional space, the lower-half convex hull defines the minimum energy surface and the ground-state configurations of the system. All non-ground-state stoichiometries are unstable, with the decomposition described by the hull facet directly below it. In the case of a binary system, the facet is a tie line as illustrated in Figure 7.2a. The energy gained from this decomposition is geometrically represented by the (vertical) distance of the compound from the facet and quantifies the excitation energy involved in forming this compound. While the minimum energy surface changes at finite temperature
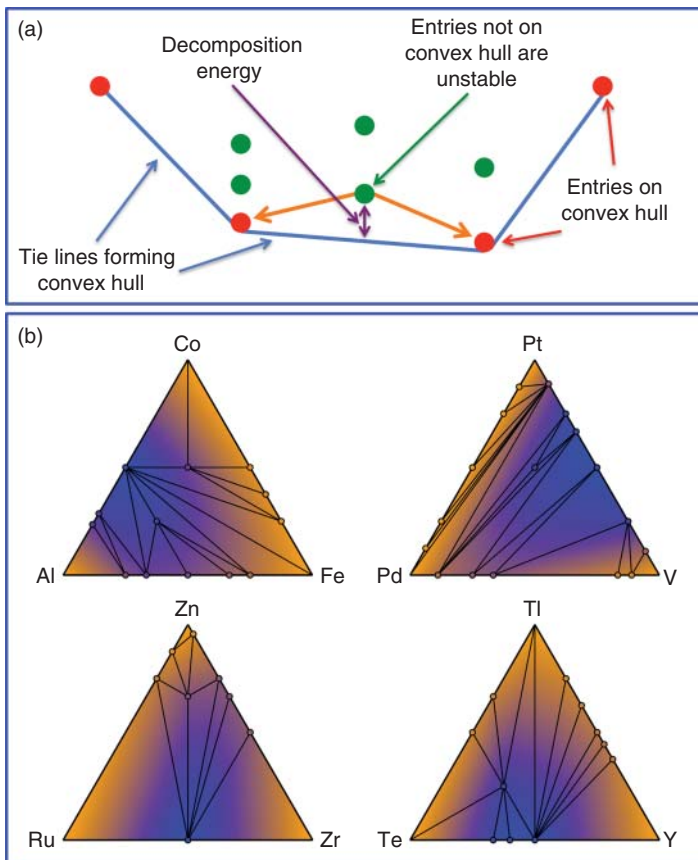

Figure 7.2 Convex hull phase diagrams for multicomponent alloys systems. (a) Schematic illustrating construction of convex hull for a general binary alloy system A_xB_{1-x}. Ground-state structures are depicted as red points, with the minimum energy surface outlined with blue lines. The minimum energy surface is formed by connecting the lowest-energy structures with tie lines that form a convex hull. Unstable structures are shown in green, with the decomposition reaction indicated by orange arrows and the decomposition energy indicated in purple. (b) Example ternary convex hulls as generated by AFLOW.

(favoring disordered structures), the $T = 0$ K excitation energy serves as a reasonable descriptor for relative thermodynamic stability [58]. This analysis generates valuable information such as ground-state structures, excitation energies, and phase coexistence for storage in the online data repository. This stability data can be visualized and displayed by online modules, such as those developed by AFLOW [58], the Materials Project [59], and the OQMD [16, 60]. An example visualization from AFLOW is shown in Figure 7.2b.

Convex hull phase diagrams have been used to discover new thermodynamically stable compounds in a wide range of alloy systems, including hafnium [61, 62], rhodium [63], rhenium [64], ruthenium [65], and technetium [66] with various transition metals, as well as the Co–Pt system [67]. Magnesium alloy systems such as the lightweight Li–Mg system [68] and 34 other Mg-based systems [69] have also been investigated. This approach has also been used to calculate the solubility of elements in titanium alloys [70], to study the effect of hydrogen on phase separation in iron–vanadium [71], and to find new superhard tungsten nitride compounds [72]. The data has been employed to generate structure maps for hcp metals [73], as well as to search for new stable compounds with the Pt_8Ti phase [74] and with the $L1_1$ and $L1_3$ crystal structures [75]. Note that even if a structure does not lie on the ground-state convex hull, this does not rule out its existence. It may be synthesizable under specific temperature and pressure conditions and then be metastable under ambient conditions.

7.2.2 Standardized Protocols for Automated Data Generation

Standard calculation protocols and parameters sets [6] are essential to the identification of trends and correlations among materials properties. The workhorse method for calculating quantum mechanically resolved materials properties is *d*ensity *f*unctional *t*heory (DFT). DFT is based on the Hohenberg–Kohn theorem [76], which proves that for a ground-state system, the potential energy is a unique functional of the density: $V(\vec{r}) = V(\rho(\vec{r}))$. This allows for the charge density $\rho(\vec{r})$ to be used as the central variable for the calculations rather than the many-body wavefunction $\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$, dramatically reducing the number of degrees of freedom in the calculation.

The Kohn–Sham equations [77] map the n coupled equations for the system of n interacting particles onto a system of n independent equations for n noninteracting particles:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_s(\vec{r})\right]\phi_i(\vec{r}) = \varepsilon_i\phi_i(\vec{r}) \tag{7.1}$$

where $\phi_i(\vec{r})$ are the noninteracting Kohn–Sham eigenfunctions and ε_i are their eigenenergies. $V_s(\vec{r})$ is the Kohn–Sham potential:

$$V_s(\vec{r}) = V(\vec{r}) + \int e^2 \frac{\rho_s(\vec{r}')}{|\vec{r} - \vec{r}'|}d^3\vec{r}' + V_{XC}[\rho_s(\vec{r})] \tag{7.2}$$

where $V(\vec{r})$ is the external potential (which includes influences of the nuclei, applied fields, and the core electrons when pseudopotentials are used),

the second term is the direct Coulomb potential, and $V_{\text{XC}}[\rho_s(\vec{r})]$ is the exchange–correlation term.

The mapping onto a system of n noninteracting particles comes at the cost of introducing the exchange–correlation potential $V_{\text{XC}}[\rho_s(\vec{r})]$, the exact form of which is unknown and must be approximated. The simplest approximation is the *l*ocal *d*ensity *a*pproximation (LDA) [78], in which the magnitude of the exchange–correlation energy at a particular point in space is assumed to be proportional to the magnitude of the density at that point in space. Despite its simplicity, LDA produces realistic results for atomic structure, elastic, and vibrational properties for a wide range of systems. However, it tends to overestimate the binding energies of materials, even putting crystal bulk phases in the wrong energetic order [79]. Beyond LDA is the *g*eneralized *g*radient *a*pproximation (GGA), in which the exchange–correlation term is a functional of the charge density and its gradient at each point in space. There are several forms of GGA including those developed by Perdew, Burke, and Ernzerhof (PBE, [80]) or by Lee, Yang, and Parr (LYP, [81]). A more recent development is the meta-GGA *s*trongly *c*onstrained and *a*ppropriately *n*ormed (SCAN) functional [82], which satisfies all 17 known exact constraints on exchange–correlation functionals.

The major limitations of LDA and GGA include their inability to adequately describe systems with strongly correlated or localized electrons, due to the local and semilocal nature of the functionals. Treatments include the Hubbard U corrections [83, 84], self-interaction corrections [78], and hybrid functionals such as Becke's three-parameter modification of LYP (B3LYP, [85]) and that of Heyd, Scuseria, and Ernzerhof (HSE, [86]).

Within the context of *ab initio* structure prediction calculations, GGA-PBE is the usual standard since it tends to produce accurate geometries and lattice constants [56]. For accounting for strong correlation effects, the DFT+U method [83, 84] is often favored in large-scale automated database generation due to its low computational overhead. However, the traditional DFT+U procedure requires the addition of an empirical factor to the potential [83, 84]. Recently, methods have been implemented to calculate the U parameter self-consistently from first principles, such as the ACBN0 functional [87].

DFT also suffers from an inadequate description of excited/unoccupied states, as the theory is fundamentally based on the ground state. Extensions for describing excited states include time-dependent density functional theory (TDDFT) [88] and the GW correction [89]. However, these methods are typically much more expensive than standard DFT and are not generally considered for large-scale database generation.

At the technical implementation level, there are many DFT software packages available, including VASP [90–93], QuantumESPRESSO [94, 95], ABINIT [96, 97], FHI–AIMS [98], SIESTA [99], and GAUSSIAN [100]. These codes are generally distinguished by the choice of basis set. There are two principle types of basis sets: plane waves, which take the form $\psi(\vec{r}) = \sum \text{e}^{i\vec{k}\cdot\vec{r}}$, and local orbitals, formed by a sum over functions $\phi_a(\vec{r})$ localized at particular points in space, such as Gaussians or numerical atomic orbitals [101]. Plane-wave-based packages include VASP, QuantumESPRESSO and ABINIT, and are generally better

suited to periodic systems such as bulk inorganic materials. Local orbital-based packages include FHI–AIMS, SIESTA and GAUSSIAN, and are generally better suited to nonperiodic systems such as organic molecules. In the field of automated computational materials science, plane-wave codes such as VASP are generally preferred: it is straightforward to automatically and systematically generate well-converged basis sets since there is only a single parameter to adjust, namely, the cutoff energy determining the number of plane waves in the basis set. Local orbital basis sets tend to have far more independently adjustable degrees of freedom, such as the number of basis orbitals per atomic orbital as well as their respective cutoff radii, making the automated generation of reliable basis sets more difficult. Therefore, a typical standardized protocol for automated materials science calculations [6] relies on the VASP software package with a basis set cutoff energy higher than that recommended by the VASP potential files, in combination with the PBE formulation of GGA.

Finally, it is necessary to automate the generation of the reciprocallattice k-point grid and pathways in reciprocal space used for the calculation of forces, energies, and the electronic band structure. In general, DFT codes use standardized methods such as the Monkhorst–Pack scheme [102] to generate k-point grids, although optimized grids have been calculated for different lattice types and are available online [103]. Optimizing k-point grid density is a computationally expensive process that is difficult to automate, so instead standardized grid densities based on the concept of "k-*p*oints *p*er *r*eciprocal *a*tom" (KPPRA) are used. The KPPRA value is chosen to be sufficiently large to ensure convergence for all systems. Typical recommended values used for KPPRA range from 6000 to 10 000 [6], so that a material with two atoms in the calculation cell will have a k-point mesh of at least 3000–5000 points. Standardized directions in reciprocal space have also been defined for the calculation of the band structure as illustrated in Figure 7.3 [3]. These paths are optimized to include all of the high-symmetry points of the lattice.

7.3 Integrated Calculation of Materials Properties

Automated frameworks such as AFLOW combine the computational analysis of properties including symmetry, electronic structure, elasticity, and thermal behavior into integrated workflows. Crystal symmetry information is used to find the primitive cell to reduce the size of DFT calculations, to determine the appropriate paths in reciprocal space for electronic band structure calculations (see Figure 7.3, [3]), and to determine the set of inequivalent distortions for phonon and elasticity calculations. Thermal and elastic properties of materials are important for predicting the thermodynamic and mechanical stability of structural phases [104–107] and assessing their importance for a variety of applications. Elastic properties such as the shear and bulk moduli are important for predicting the hardness of materials [108, 109] and thus their resistance to wear and distortion. Elasticity tensors can be used to predict the properties of composite materials [110, 111]. They are also important in geophysics for modeling
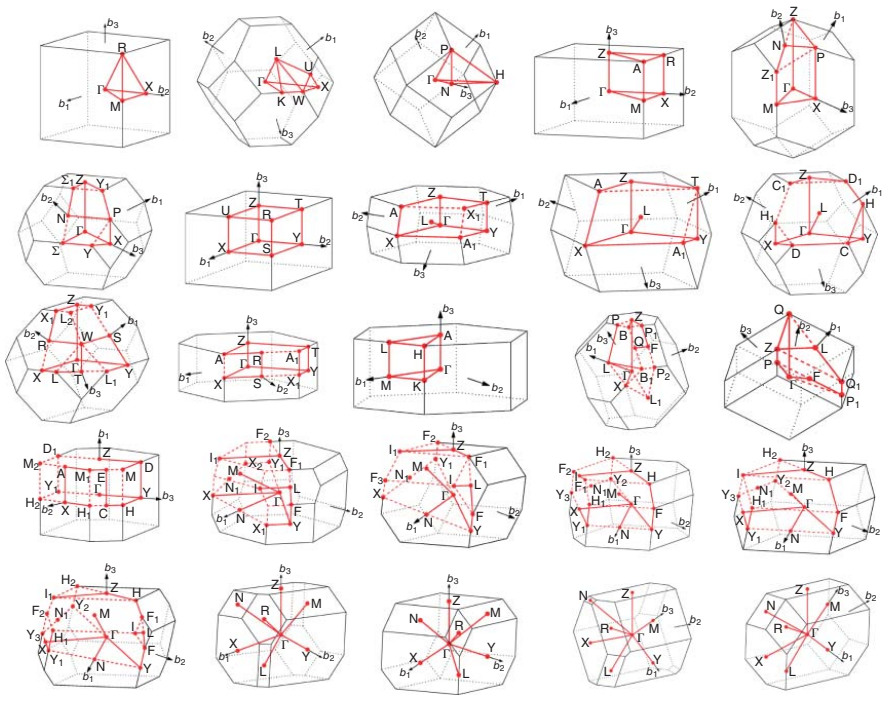
Figure 7.3 Standardized paths in reciprocal space for calculation of the electronic band structures for the 25 different lattice types [3]. Source: Setyawan and Curtarolo 2010 [3]. Reproduced with permission of Elsevier.

the propagation of seismic waves in order to investigate the mineral composition of geological formations [105, 112, 113]. The lattice thermal conductivity (κ_{L}) is a crucial design parameter in a wide range of important technologies, such as the development of new thermoelectric materials [39, 114, 115], heat sink materials for thermal management in electronic devices [116], and rewritable phase-change memories [117]. High thermal conductivity materials, which typically have a zincblende or diamond-like structure, are essential in micro-electronic and nanoelectronic devices for achieving efficient heat removal [118] and have been intensively studied for the past few decades [119]. Low thermal conductivity materials constitute the basis of a new generation of thermoelectric materials and thermal barrier coatings [120].

The calculation of thermal and elastic properties offers an excellent example of the power of integrated computational materials design frameworks. With a single input file, these frameworks can automatically set up and run calculations of different distorted cells, and combine the resulting energies and forces to calculate thermal and mechanical properties.

7.3.1 Autonomous Symmetry Analysis

Critical to any analysis of crystals is the accurate determination of the symmetry profile. For example, symmetry serves to (i) validate the forms of the elastic constants and compliance tensors, where the crystal symmetry dictates equivalence or absence of specific tensor elements [50, 106, 121], and (ii) reduce the number of *ab initio* calculations needed for phonon calculations, where, in the case of the finite displacement method, equivalent atoms and distortion directions are identified through factor group and site symmetry analyses [122].

Autonomous workflows for elasticity and vibrational characterizations therefore require a correspondingly robust symmetry analysis. Unfortunately, standard symmetry packages [123–126], catering to different objectives, depend on tolerance-tuning to overcome numerical instabilities and atypical data – emanating from finite temperature measurements and uncertainty in experimentally reported observations. These tolerances are responsible for validating mappings and identifying isometries, such as the *n*-fold operator depicted in Figure 7.4a. Some standard packages define separate tolerances for space, angle [126], and even operation type [123–125] (e.g. rotation vs. inversion). Each parameter introduces a factorial expansion of unique inputs, which can result in distinct symmetry profiles as illustrated in Figure 7.4b. By varying the spatial tolerance ϵ, four different space groups can be observed for AgBr (ICSD #56551 [www.aflow.org/material.php?id=56551]), if one is found at all. Gaps in the range, where no consistent symmetry profile can be resolved, are particularly problematic in automated frameworks, triggering critical failures in subsequent analyses.

Cell shape can also complicate mapping determinations. Anisotropies in the cell, such as skewness of lattice vectors, translate to distortions of fractional and reciprocal spaces. A uniform tolerance sphere in Cartesian space, inside which points are considered mapped, generally warps to a sheared spheroid, as depicted in Figure 7.4c. Hence, distances in these spaces are direction dependent,
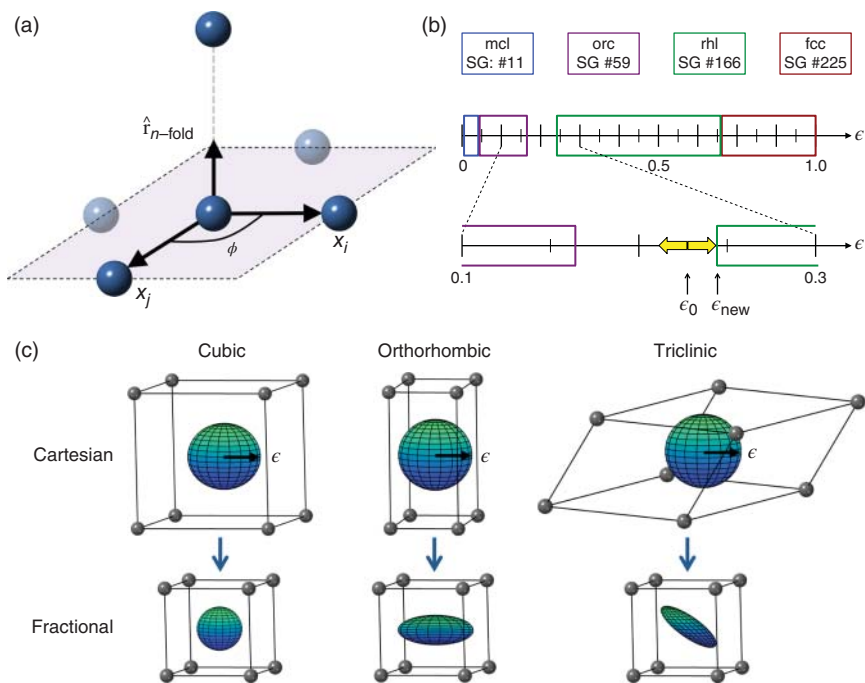
Figure 7.4 Challenges in autonomous symmetry analysis. (a) An illustration of a general *n*-fold symmetry operation. (b) Possible space group determinations with mapping tolerance ϵ for AgBr (ICSD #56551). (c) Warping of mapping tolerance sphere with a transformation from Cartesian to fractional basis.

compromising the integrity of rapid minimum-image determinations [127] and generally warranting prohibitively expensive algorithms [128]. Such failures can result in incommensurate symmetry profiles, where the real space lattice profile (e.g. bcc) does not match that of the reciprocal space (fcc).

The new AFLOW-SYM module [128] within AFLOW offers careful treatment of tolerances, with extensive validation schemes, to mitigate the aforementioned challenges. Although a user-defined tolerance input is still available, AFLOW defaults to one of two predefined tolerances, namely, `tight` (standard) and `loose`. Should any discrepancies occur, these defaults are the starting values of a large tolerance scan, as shown in Figure 7.4b. A number of validation schemes have been incorporated to catch such discrepancies. These checks are consistent with crystallographic group theory principles, validating operation types and cardinalities [129]. From considerations of different extreme cell shapes, a heuristic threshold has been defined to classify scenarios where mapping failures are likely to occur – based on skewness and mapping tolerance. When benchmarked against standard packages for over 54 000 structures in the Inorganic Crystal Structure Database, AFLOW-SYM consistently resolves the symmetry characterization most compatible with experimental observations [128].

Along with accuracy, AFLOW-SYM delivers a wealth of symmetry properties and representations to satisfy injection into any analysis or workflow. The full set of operators – including that of the point, factor, crystallographic point, space group, and site symmetries – are provided in matrix, axis–angle, matrix generator, and quaternion representations in both Cartesian and fractional coordinates. A span of characterizations, organized by degree of symmetry-breaking, are available, including those of the lattice, superlattice, crystal, and crystal spin. Space group and Wyckoff positions are also resolved. The full data set is made available in both plain text and JSON formats.

7.3.2 Elastic Constants

There are two main methods for calculating the elastic constants based on the response of either the stress tensor or the total energy to a set of applied strains [50, 51, 130–132]. Automated implementations of these methods are included in the AFLOW (referred to as the *A*utomatic *E*lasticity *L*ibrary, AEL [51]) and Materials Project frameworks [50].

To calculate the elastic tensor, several different normal and shear strains should be applied to the calculation cell in each independent direction [50, 51], as illustrated in Figure 7.5a. The resulting stress tensor elements σ_{ij}, obtained from the directional forces on the cell calculated with DFT, can then be fitted to the applied strains ϵ_{ij} to obtain the corresponding elastic constants c_{ij} in the form of the stiffness tensor:

$$\begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} \\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66} \end{pmatrix} \begin{pmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ 2\epsilon_{23} \\ 2\epsilon_{13} \\ 2\epsilon_{12} \end{pmatrix} \tag{7.3}$$

written in the 6×6 Voigt notation using the mapping [105] $11 \mapsto 1$, $22 \mapsto 2$, $33 \mapsto 3$, $23 \mapsto 4$, $13 \mapsto 5$, and $12 \mapsto 6$. Symmetry analysis such as that provided by AFLOW-SYM can be used to reduce the number of required calculations by up to a factor of 3 in the case of cubic systems, as well as for verification of the computed tensors [106].

The elastic constants can then be used in the Voigt or Reuss approximations, which for polycrystalline materials correspond to assuming uniform strain and uniform stress, respectively, and give the upper and lower bounds on the elastic moduli. In the Voigt approximation, the bulk modulus is given by

$$B_{\text{Voigt}} = \frac{1}{9}[(c_{11} + c_{22} + c_{33}) + 2(c_{12} + c_{23} + c_{13})] \tag{7.4}$$

and the shear modulus is given by

$$G_{\text{Voigt}} = \frac{1}{15}[(c_{11} + c_{22} + c_{33}) - (c_{12} + c_{23} + c_{13})]$$
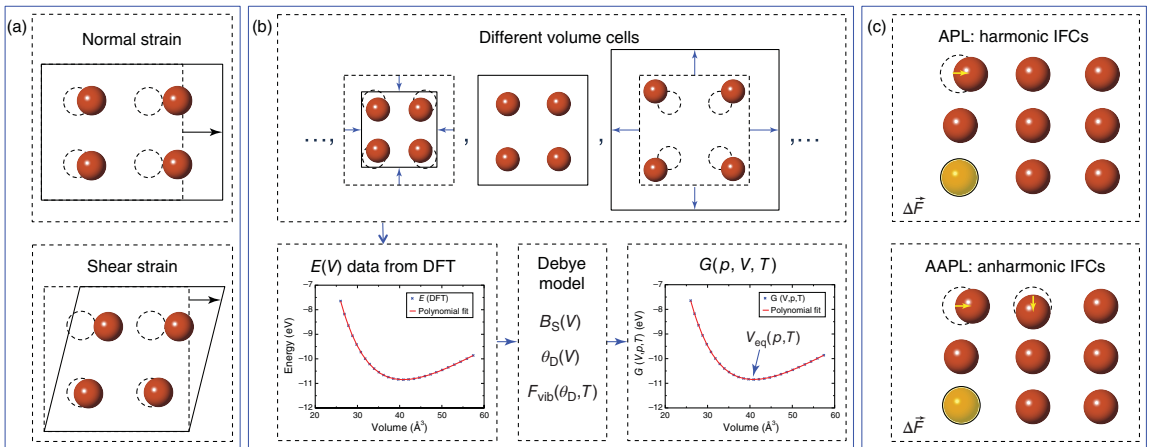$$+ \frac{1}{5}(c_{44} + c_{55} + c_{66}) \tag{7.5}$$

Figure 7.5 (a) AEL applies a set of independent normal and shear strains to the crystal structure to obtain the elastic constants. (b) AGL applies a set of isotropic strains to the unit cell to obtain energy vs. volume data, which is fitted by a polynomial in order to calculate the bulk modulus as a function of volume, $B_S(V)$. $B_S(V)$ is then used to calculate the Debye temperature as a function of volume and thus the vibrational free energy as a function of temperature. The Gibbs free energy as a function of volume is then minimized for each pressure and temperature point to obtain the equilibrium volume and other thermomechanical properties. (c) APL obtains the harmonic *i*nteratomic *f*orce *c*onstants (IFCs) from supercell calculations where inequivalent atoms are displaced in inequivalent directions, and then the changes in the forces on the other atoms are calculated. The IFCs are then used to construct the dynamical matrix, which is diagonalized to obtain the phonon eigenmodes. AAPL calculates three-phonon scattering effects by performing supercell calculations where pairs of inequivalent atoms are displaced in inequivalent directions, and the changes in the forces on the other atoms in the supercell are calculated to obtain the third-order anharmonic IFCs.

The Reuss approximation uses the elements of the compliance tensor s_{ij} (the inverse of the stiffness tensor) to calculate the bulk modulus

$$\frac{1}{B_{\mathrm{Reuss}}} = (s_{11} + s_{22} + s_{33}) + 2(s_{12} + s_{23} + s_{13}) \tag{7.6}$$

while the shear modulus is given by

$$\frac{15}{G_{\mathrm{Reuss}}} = 4(s_{11} + s_{22} + s_{33}) - 4(s_{12} + s_{23} + s_{13})$$
$$+ 3(s_{44} + s_{55} + s_{66}) \tag{7.7}$$

The two approximations are combined to obtain the *V*oigt–*R*euss–*H*ill (VRH) averages [133] for the bulk modulus

$$B_{\mathrm{VRH}} = \frac{B_{\mathrm{Voigt}} + B_{\mathrm{Reuss}}}{2} \tag{7.8}$$

and the shear modulus

$$G_{\mathrm{VRH}} = \frac{G_{\mathrm{Voigt}} + G_{\mathrm{Reuss}}}{2} \tag{7.9}$$

The Poisson ratio v is then given by

$$v = \frac{3B_{\mathrm{VRH}} - 2G_{\mathrm{VRH}}}{6B_{\mathrm{VRH}} + 2G_{\mathrm{VRH}}} \tag{7.10}$$

7.3.3 Quasi-harmonic Debye–Grüneisen Model

Thermal properties can be predicted by several different methods, such as the quasi-harmonic Debye–Grüneisen model, which uses volume as a proxy for temperature [49], and by calculating the phonon dispersion from the dynamical matrix of IFCs [122].

The energy vs. volume data from a set of simple static primitive cell calculations can be fitted to a quasi-harmonic Debye–Grüneisen model such as the "GIBBS" method [49, 51, 134] to obtain thermal properties, as demonstrated in Figure 7.5b. This method has been implemented in the AFLOW framework in the form of the *A*utomatic *GIBBS L*ibrary (AGL).

First, the adiabatic bulk modulus B_{S} as a function of cell volume V is obtained either (i) by fitting the $E_{\mathrm{DFT}}(V)$ data to an equation of state (EOS) or (ii) by taking the numerical second derivative of a polynomial fit of $E_{\mathrm{DFT}}(V)$, which gives the static bulk modulus B_{static}:

$$B_{\mathrm{S}}(V) \approx B_{\mathrm{static}}(\vec{x}) \approx B_{\mathrm{static}}(\vec{x}(V))$$
$$= V\left(\frac{\partial^2 E(\vec{x}(V))}{\partial V^2}\right) = V\left(\frac{\partial^2 E(V)}{\partial V^2}\right) \tag{7.11}$$

Three different empirical EOS have been implemented within AGL: the Birch–Murnaghan EOS [105, 134, 135], the Vinet EOS [134, 136], and the Baonza–Cáceres–Núñez spinodal EOS [134, 137]. However, these EOS often introduce an additional source of error into the results since they are calibrated

for specific sets of systems and pressure–temperature regimes. Recent studies have found the numerical calculation of B to be just as, if not more, reliable as the empirical EOS [51]. Therefore, the numerical method is the default for the automated generation of thermomechanical properties for the AFLOW database.

The bulk modulus can then be used to calculate the Debye temperature as a function of volume:

$$\theta_{\mathrm{D}}(V) = \frac{\hbar}{k_{\mathrm{B}}} [6\pi^2 V^{1/2} n]^{1/3} f(v) \sqrt{\frac{B_{\mathrm{S}}}{M}} \tag{7.12}$$

where M is the mass of the unit cell and $f(v)$ is a function of the Poisson ratio v:

$$f(v) = \left\{ 3 \left[2 \left(\frac{2}{3} \cdot \frac{1+v}{1-2v} \right)^{3/2} + \left(\frac{1}{3} \cdot \frac{1+v}{1-v} \right)^{3/2} \right]^{-1} \right\}^{\frac{1}{3}} \tag{7.13}$$

The integration offered by the AFLOW framework allows the value of v required by this expression to be obtained directly and automatically from the AEL calculation (Eq. (7.10)).

To obtain the equilibrium volume at a particular (p, T) point, the Gibbs free energy is minimized with respect to volume. In the quasi-harmonic approximation (QHA), the vibrational component of the free energy, $F_{\mathrm{vib}}(\vec{x}; T)$, is given by

$$F_{\mathrm{vib}}(\vec{x}; T) = \int_0^\infty \left[\frac{\hbar\omega}{2} + \frac{1}{\beta} \log(1 - e^{-\beta\hbar\omega}) \right] g(\vec{x}; \omega) d\omega \tag{7.14}$$

where $\beta = (k_{\mathrm{B}} T)^{-1}$ and $g(\vec{x}; \omega)$ is the phonon density of states, which depends on the system geometry $\vec{x}$. In the Debye–Grüneisen model, F_{vib} can be written as

$$F_{\mathrm{vib}}(\theta_{\mathrm{D}}; T) = \frac{n}{\beta} \left[\frac{9}{8} \frac{\theta_{\mathrm{D}}}{T} + 3 \log \left(1 - e^{-\theta_{\mathrm{D}}/T} \right) - D \left(\frac{\theta_{\mathrm{D}}}{T} \right) \right] \tag{7.15}$$

where $D(\theta_{\mathrm{D}}/T)$ is the Debye integral

$$D \left(\theta_{\mathrm{D}}/T \right) = 3 \left(\frac{T}{\theta_{\mathrm{D}}} \right)^3 \int_0^{\theta_{\mathrm{D}}/T} \frac{x^3}{e^x - 1} dx \tag{7.16}$$

Next, the full Gibbs free energy as a function of temperature and pressure is calculated by

$$\mathsf{G}(V; p, T) = E_{\mathrm{DFT}}(V) + F_{\mathrm{vib}}(\theta_{\mathrm{D}}(V); T) + pV \tag{7.17}$$

and fitted by a polynomial in V, the minimum of which gives the equilibrium volume, V_{eq}. Note that the symbol G is used for shear modulus, while G is used for the Gibbs free energy.

θ_{D} is then determined from its value at V_{eq}, while other thermal properties such as the Grüneisen parameter can be calculated using the expression

$$\gamma = -\frac{V}{\theta_{\mathrm{D}}} \frac{\partial \theta_{\mathrm{D}}(V)}{\partial V} \tag{7.18}$$

The specific heat capacity at constant volume can be obtained using the expression

$$C_{\text{V,vib}} = 3nk_{\text{B}} \left[4D\left(\frac{\theta_{\text{D}}}{T}\right) - \frac{3\theta_{\text{D}}/T}{e^{\theta_{\text{D}}/T} - 1} \right] \tag{7.19}$$

while the specific heat capacity at constant pressure is given by

$$C_{\text{p,vib}} = C_{\text{V,vib}}(1 + \alpha\gamma T) \tag{7.20}$$

where α is the coefficient of thermal expansion

$$\alpha = \frac{\gamma C_{\text{V,vib}}}{B_{\text{T}} V} \tag{7.21}$$

The lattice thermal conductivity can be calculated using the Leibfried–Schlömann equation [138–140] using the Debye temperature and the Grüneisen parameter:

$$\kappa_{\text{L}}(\theta_{\text{a}}) = \frac{0.849 \times 3\sqrt[3]{4}}{20\pi^3(1 - 0.514\gamma_{\text{a}}^{-1} + 0.228\gamma_{\text{a}}^{-2})}$$
$$\times \left(\frac{k_{\text{B}}\theta_{\text{a}}}{\hbar}\right)^2 \frac{k_{\text{B}}mV^{\frac{1}{3}}}{\hbar\gamma_{\text{a}}^2} \tag{7.22}$$

where V is the volume of the unit cell and m is the average atomic mass, while θ_{a} and γ_{a} are the acoustic Debye temperature and Grüneisen parameter obtained by only considering the acoustic modes, based on the assumption that the optical phonon modes in crystals do not contribute to heat transport [139, 140]. θ_{a} and γ_{a} can be derived directly from the phonon DOS by only considering the acoustic modes [139, 141]. θ_{a} can also be estimated from the traditional Debye temperature θ_{D} using the expression $\theta_{\text{a}} = \theta_{\text{D}}n^{-\frac{1}{3}}$ [139, 140]. There is no simple way to extract γ_{a} from the traditional Grüneisen parameter, so the approximation $\gamma_{\text{a}} = \gamma$ is used in the AEL–AGL approach to calculating the thermal conductivity.

The thermal conductivity at temperatures other than θ_{a} is estimated using the expression [139, 140, 142] $\kappa_{\text{L}}(T) = \kappa_{\text{L}}(\theta_{\text{a}})\theta_{\text{a}}/T$.

7.3.4 Harmonic Phonons

Thermal properties can also be obtained by directly calculating the phonon dispersion from the dynamical matrix of IFCs. The approach is implemented within the AFLOW *P*honon *Li*brary (APL) [2]. The IFCs are determined from a set of supercell calculations in which the atoms are displaced from their equilibrium positions [122] as shown in Figure 7.5c.

The IFCs derive from a Taylor expansion of the potential energy, V, of the crystal about the atoms' equilibrium positions:

$$V = V|_{\vec{r}(i,t)=0, \forall i} + \sum_{i,\alpha} \left.\frac{\partial V}{\partial r(i,t)^\alpha}\right|_{\vec{r}(i,t)=0, \forall i} r(i,t)^\alpha$$
$$+ \frac{1}{2} \sum_{\substack{i,\alpha, \\ j,\beta}} \left.\frac{\partial^2 V}{\partial r(i,t)^\alpha \partial r(j,t)^\beta}\right|_{\vec{r}(i,t)=0, \forall i} r(i,t)^\alpha r(j,t)^\beta$$
$$+ \cdots \tag{7.23}$$

where $r(i, t)^\alpha$ is the α-Cartesian component ($\alpha = x, y, z$) of the time-dependent atomic displacement $\vec{r}(t)$ of the ith atom about its equilibrium position, $V|_{\vec{r}(i,t)=0,\forall i}$ is the potential energy of the crystal in its equilibrium configuration, $\partial V / \partial r(i, t)^\alpha|_{\vec{r}(i,t)=0,\forall i}$ is the negative of the force acting in the α direction on atom i in the equilibrium configuration (zero by definition), and $\partial^2 V / \partial r(i, t)^\alpha \partial r(j, t)^\beta|_{\vec{r}(i,t)=0,\forall i}$ constitute the IFC $\phi(i, j)_{\alpha,\beta}$. To first approximation, $\phi(i, j)_{\alpha,\beta}$ is the negative of the force exerted in the α direction on atom i when atom j is displaced in the β direction with all other atoms maintaining their equilibrium positions, as shown in Figure 7.5c. All higher-order terms are neglected in the harmonic approximation.

Correspondingly, the equations of motion of the lattice are

$$M(i)\ddot{r}(i, t)^\alpha = - \sum_{j,\beta} \phi(i, j)_{\alpha,\beta} r(j, t)^\beta \quad \forall i, \alpha \tag{7.24}$$

and can be solved by a plane-wave solution of the form

$$r(i, t)^\alpha = \frac{v(i)^\alpha}{\sqrt{M(i)}} e^{i(\vec{q}\cdot\vec{R}_l - \omega t)} \tag{7.25}$$

where $v(i)^\alpha$ form the phonon eigenvectors (polarization vector), $M(i)$ is the mass of the ith atom, $\vec{q}$ is the wave vector, $\vec{R}_l$ is the position of lattice point l, and ω form the phonon eigenvalues (frequencies). The approach is nearly identical to that taken for electrons in a periodic potential (Bloch waves) [143]. Plugging this solution into the equations of motion (Eq. (7.24)) yields the following set of linear equations:

$$\omega^2 v(i)^\alpha = \sum_{j,\beta} D_{i,j}^{\alpha,\beta}(\vec{q}) v(j)^\beta \quad \forall i, \alpha \tag{7.26}$$

where the dynamical matrix $D_{i,j}^{\alpha,\beta}(\vec{q})$ is defined as

$$D_{i,j}^{\alpha,\beta}(\vec{q}) = \sum_l \frac{\phi(i, j)_{\alpha,\beta}}{\sqrt{M(i)M(j)}} e^{-i\vec{q}\cdot(\vec{R}_l - \vec{R}_0)} \tag{7.27}$$

The problem can be equivalently represented by a standard eigenvalue equation:

$$\omega^2 [\vec{v}] = [\mathbf{D}(\vec{q})] [\vec{v}] \tag{7.28}$$

where the dynamical matrix and phonon eigenvectors have dimensions ($3n_a \times 3n_a$) and ($3n_a \times 1$), respectively, and n_a is the number of atoms in the cell. Hence, Eq. (7.28) has $3n_a$ solutions/modes referred to as branches indexed by λ. In practice, Eq. (7.28) is solved for discrete sets of $\vec{q}$-points to compute the phonon density of states (grid over all possible $\vec{q}$) and dispersion (along the high-symmetry paths of the lattice [3]). Thus, the phonon eigenvalues and eigenvectors are appropriately denoted $\omega_\lambda(\vec{q})$ and $\vec{v}_\lambda(\vec{q})$, respectively.

Similar to the electronic Hamiltonian, the dynamical matrix is Hermitian, i.e. $\mathbf{D}(\vec{q}) = \mathbf{D}^*(\vec{q})$. Thus $\omega_\lambda^2(\vec{q})$ must also be real, so $\omega_\lambda(\vec{q})$ can either be real or purely imaginary. However, a purely imaginary frequency corresponds to vibrational motion of the lattice that increases exponentially in time. Therefore, imaginary frequencies, or those corresponding to soft modes, indicate the structure is

dynamically unstable. In the case of a symmetric, high-temperature phase, soft modes suggest there exists a lower symmetry structure stable at $T = 0$ K. Temperature effects on phonon frequencies can be modeled with

$$\tilde{\omega}_\lambda^2(\vec{q}, T) = \omega_\lambda^2(\vec{q}, T = 0) + \eta T \tag{7.29}$$

where η is positive in general. The two structures, the symmetric and the stable, differ by the distortion corresponding to this "frozen" (non-vibrating) mode. Upon heating, the temperature term increases until the frequency reaches zero, and a phase transition occurs from the stable structure to the symmetric [144].

In practice, soft modes [145] may indicate the following: (i) the structure is dynamically unstable at T; (ii) the symmetry of the structure is lower than that considered, perhaps due to magnetism; (iii) strong electronic correlations, or (iv) long-range interactions play a significant role, and a larger supercell should be considered.

With the phonon density of states computed, the following thermal properties can be calculated: the internal vibrational energy

$$U_{\text{vib}}(\vec{x}, T) = \int_0^\infty \left(\frac{1}{2} + \frac{1}{e^{(\beta \hbar \omega)} - 1} \right) \hbar \omega g(\vec{x}; \omega) \mathrm{d}\omega \tag{7.30}$$

the vibrational component of the free energy $F_{\text{vib}}(\vec{x}; T)$ (Eq. (7.14)), the vibrational entropy

$$S_{\text{vib}}(\vec{x}, T) = \frac{U_{\text{vib}}(\vec{x}, T) - F_{\text{vib}}(\vec{x}; T)}{T} \tag{7.31}$$

and the isochoric specific heat

$$C_{\text{V,vib}}(\vec{x}, T) = \int_0^\infty \frac{k_{\text{B}}(\beta \hbar \omega)^2 g(\vec{x}; \omega)}{(1 - e^{-(\beta \hbar \omega)})(e^{(\beta \hbar \omega)} - 1)} \mathrm{d}\omega \tag{7.32}$$

7.3.5 Quasi-harmonic Phonons

The harmonic approximation does not describe phonon–phonon scattering and so cannot be used to calculate properties such as thermal conductivity or thermal expansion. To obtain these properties, either the *Quasi-Harmonic Approximation* (QHA) can be used, or a full calculation of the higher-order anharmonic IFCs can be performed. QHA is the less computationally demanding of these two methods and compares harmonic calculations of phonon properties at different volumes to predict anharmonic properties. The different volume calculations can be in the form of harmonic phonon calculations as described above [146, 147] or simple static primitive cell calculations [49, 134]. QHA is implemented within APL and referred to as QHA-APL [49]. In the case of the quasi-harmonic phonon calculations, the anharmonicity of the system is described by the mode-resolved Grüneisen parameters, which are given by the change in the phonon frequencies as a function of volume:

$$\gamma_\lambda(\vec{q}) = -\frac{V}{\omega_\lambda(\vec{q})} \frac{\partial \omega_\lambda(\vec{q})}{\partial V} \tag{7.33}$$

where $\gamma_\lambda(\vec{q})$ is the parameter for the wave vector $\vec{q}$ and the λth mode of the phonon dispersion. The average of the $\gamma_\lambda(\vec{q})$ values, weighted by the specific heat capacity of each mode $C_{V,\lambda}(\vec{q})$, gives the average Gruneisen parameter:

$$\gamma = \frac{\sum_{\lambda,\vec{q}} \gamma_\lambda(\vec{q}) C_{V,\lambda}(\vec{q})}{C_V} \tag{7.34}$$

The specific heat capacity, Debye temperature, and Grüneisen parameter can then be combined to calculate other properties such as the specific heat capacity at constant pressure C_p, the thermal coefficient of expansion α, and the lattice thermal conductivity κ_L [147], using similar expressions to those described in Section 7.3.3.

7.3.6 Anharmonic Phonons

The full calculation of the anharmonic IFCs requires performing supercell calculations in which pairs of inequivalent atoms are displaced in all pairs of inequivalent directions [148–157] as illustrated in Figure 7.5c. The third-order anharmonic IFCs can then be obtained by calculating the change in the forces on all of the other atoms due to these displacements. This method has been implemented in the form of a fully automated integrated workflow in the AFLOW framework, where it is referred to as the *A*FLOW *A*nharmonic *P*honon *L*ibrary (AAPL) [157]. This approach can provide very accurate results for the lattice thermal conductivity when combined with accurate electronic structure methods [157] but quickly becomes very expensive for systems with multiple inequivalent atoms or low symmetry. Therefore, simpler methods such as the quasi-harmonic Debye model tend to be used for initial rapid screening [49, 51], while the more accurate and expensive methods are used for characterizing systems that are promising candidates for specific engineering applications.

7.4 Online Data Repositories

Rendering the massive quantities of data generated using automated *ab initio* frameworks available for other researchers requires going beyond the conventional methods for the dissemination of scientific results in the form of journal articles. Instead, this data is typically made available in online data repositories, which can usually be accessed both manually via interactive Web portals and programmatically via an *a*pplication *p*rogramming *i*nterface (API).

7.4.1 Computational Materials Data Web Portals

Most computational data repositories include an interactive Web portal front end that enables manual data access. These Web portals usually include online applications to facilitate data retrieval and analysis. The front page of the AFLOW data repository is displayed in Figure 7.6a. The main features include a search bar where information such as ICSD reference number, *A*FLOW *u*nique *id*entifier (AUID) or the chemical formula, can be entered in order to retrieve specific materials entries. Below are buttons linking to several different online applications
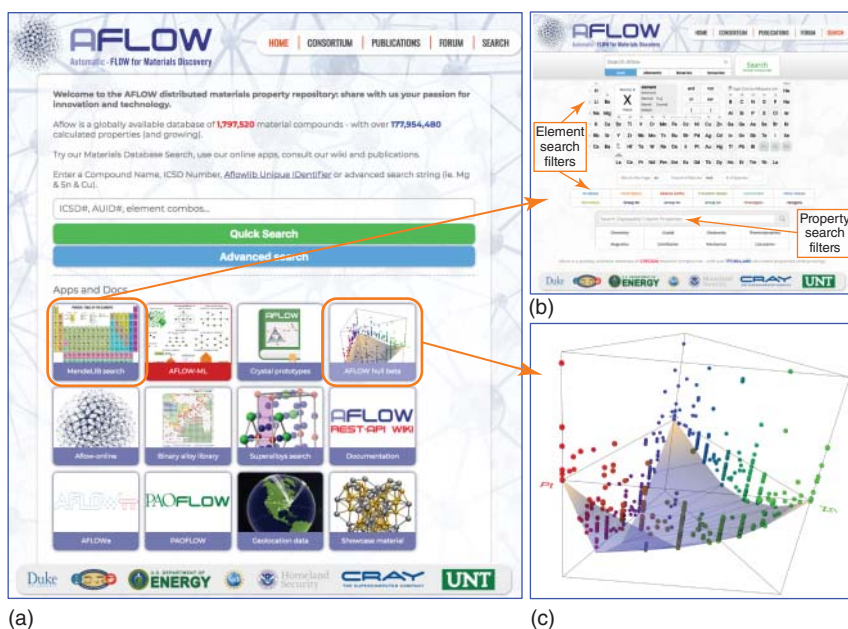
Element search filters

Property search filters

(a)

(b)

(c)

Figure 7.6 (a) Front page of the AFLOW online data repository, highlighting the link to (b) the AFLOW advanced search application, which facilitates complex search queries including filtering by chemical composition and materials properties, and (c) the AFLOW interactive convex hull generator, showing the 3D hull for the Pt–Sc–Zn ternary alloy system.

such as the advanced search functionality, convex hull phase diagram generators, machine learning applications [45, 158, 159], and AFLOW-online data analysis tools. The link to the advanced search application is highlighted by the orange square, and the application page is shown in Figure 7.6b. The advanced search application allows users to search for materials that contain (or exclude) specific elements or groups of elements, and also to filter and sort the results by properties such as electronic band structure energy gap (under the "Electronics" properties filter group) and bulk modulus (under the "Mechanical" properties filter group). This allows users to identify candidate materials with suitable properties for specific applications.

Another example online application available on the AFLOW Web portal is the convex hull phase diagram generator. This application can be accessed by clicking on the button highlighted by the orange square in Figure 7.6a, which will bring up a periodic table allowing users to select two or three elements for which they want to generate a convex hull. The application will then access the formation enthalpies and stoichiometries of the materials entries in the relevant alloy systems and use this data to generate a two- or three-dimensional convex hull phase diagram as depicted in Figure 7.6c. This application is fully interactive, allowing users to adjust the energy axis scale, rotate the diagram to view from different directions, and select specific points to obtain more information on the corresponding entries.

7.4.2 Programmatically Accessible Online Repositories of Computed Materials Properties

In order to use materials data in machine learning algorithms, it should be stored in a structured online database and made programmatically accessible via a *re*presentational *s*tate *t*ransfer API (REST API). Examples of online repositories of materials data include AFLOW [4, 5], Materials Project [11], and OQMD [15]. There are also repositories that aggregate results from multiple sources such as NoMaD [23] and Citrine [160].

REST APIs facilitate programmatic access to data repositories. Typical databases such as AFLOW are organized in layers, with the top layer corresponding to a project or catalog (e.g. binary alloys), the next layer corresponding to data sets (e.g. all of the entries for a particular alloy system), and then the bottom layer corresponding to specific materials entries, as illustrated in Figure 7.7a.

In the case of the AFLOW database, there are currently four different "projects," namely, the "ICSD," "LIB1," "LIB2," and "LIB3" projects, along with three more under construction: "LIB4," "LIB5," and "LIB6." The "ICSD" project contains calculated data for previously observed compounds [52], whereas the other three projects contain calculated data for single elements, binary alloys,
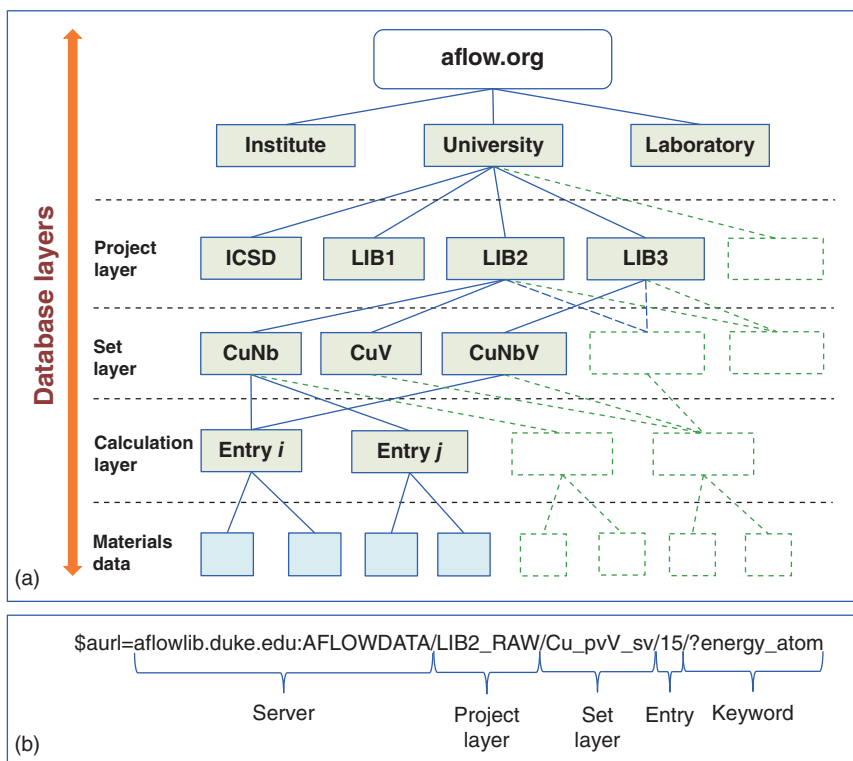

Figure 7.7 (a) The AFLOW database is organized as a multilayered system. (b) Example of an AURL that enables direct programmatic access to specific materials entry properties in the AFLOW database.

and ternary alloys, respectively, and are constructed by decorating prototype structures with combinations of different elements. Within "LIB2" and "LIB3," there are many different data sets, each corresponding to a specific binary or ternary alloy system. Each entry in the set corresponds to a specific prototype structure and stoichiometry. The materials properties values for each of these entries are encoded via keywords, and the data can be accessed via URLs constructed from the different layer names and the appropriate keywords. In the case of the AFLOW database, the location of each layer and entry is identified by an *AFLOW uniform resource locator* (AURL) [5], which can be converted to a URL providing the absolute path to a particular layer, entry, or property. The AURL takes the form server:AFLOWDATA/project/set/entry/? keywords, for example, aflowlib.duke.edu:AFLOWDATA/LIB2_RAW/Cu_pvV_sv/15/?energy_atom, where aflowlib.duke.edu is the Web address of the physical server where the data is located, LIB2_RAW is the binary alloy project layer, Cu_pvV_sv is the set containing the binary alloy system Cu–V, 15 is a specific entry with the composition Cu_3V in a tetragonal lattice, and energy_atom is the keyword corresponding to the property of energy per atom in units of eV, as shown in Figure 7.7b. Each AURL can be converted to a Web URL by changing the ":" after the server name to a "/," so that the AURL in Figure 7.7b would become the URL aflowlib.duke.edu/AFLOWDATA/LIB2_RAW/Cu_pvV_sv/15/?energy_atom. This URL, if queried via a Web browser or using a UNIX utility such as `wget`, returns the energy per atom in eV for entry 15 of the Cu–V binary alloy system.

In addition to the AURL, each entry in the AFLOW database is also associated with an AUID [5], which is a unique hexadecimal (base 16) number constructed from a checksum of the AFLOW output file for that entry. Since the AUID for a particular entry can always be reconstructed by applying the checksum procedure to the output file, it serves as a permanent, unique specifier for each calculation, irrespective of the current physical location of where the data are stored. This enables the retrieval of the results for a particular calculation from different servers, allowing for the construction of a truly distributed database that is robust against the failure or relocation of the physical hardware. Actual database versions can be identified from the version of AFLOW used to parse the calculation output files and post-process the results to generate the database entry. This information can be retrieved using the keyword aflowlib_version.

The search and sort functions of the front-end portals can be combined with the programmatic data access functionality of the REST API through the implementation of a Search-API. The AFLUX Search-API uses the LUX language to enable the embedding of logical operators within URL query strings [161]. For example, the energy per atom of every entry in the AFLOW repository containing the element Cu or V, but not the element Ti, with an electronic bandgap between 2 and 5 eV, can be retrieved using the command aflowlib.duke .edu/search/API/?species((Cu:V),(!Ti)),Egap(2*,*5),energy_atom. In this AFLUX search query, the comma "," represents the logical AND operation, the colon ":" the logical OR operation, the exclamation mark "!" the logical NOT operation, and the asterisk "*" the "loose" operation that defines a range of values to search within. Note that by default AFLUX returns only the first 64 entries matching the search query. The number and set of entries can be

controlled by appending the `paging` directive to the end of the search query as follows: aflowlib.duke.edu/search/API/?species((Cu:V),(!Ti)),Egap(2*,*5), energy_atom,paging(0), where calling the `paging` directive with the argument "0" instructs AFLUX to return all of the matching entries (note that this could potentially be a large amount of data, depending on the search query). The AFLUX Search-API allows users to construct and retrieve customized data sets, which they can feed into materials informatics machine learning packages to identify trends and correlations for use in rational materials design.

The use of APIs to provide programmatic access is being extended beyond materials data retrieval, to enable the remote use of pretrained machine learning algorithms. The AFLOW-ML API [159] facilitates access to the two machine learning models that are also available online at aflow.org/aflow-ml [45, 158]. The API allows users to submit structural data for the material of interest using a utility such as `cURL` and then returns the results of the model's predictions in JSON format. The programmatic access to machine learning predictions enables the incorporation of machine learning into materials design workflows, allowing for rapid prescreening to automatically select promising candidates for further investigation.

7.5 Materials Applications

The automated approach to computational materials science has been used to accelerate the design of materials for structural applications such as metallic glasses and superalloys, and for functional applications including thermoelectrics, magnets, catalysts, batteries, photovoltaics, and superconductors.

7.5.1 Disordered Materials

Section 7.2 describes how the thermodynamic stability of ordered compounds at zero temperature can be predicted from the convex hull phase diagrams generated using the formation enthalpies available in computational materials data repositories such as AFLOW [2, 4–6]. At finite temperature, however, entropic contributions due to thermally driven disorder play an important role and lead to the formation of disordered materials such as metallic glasses and solid solutions. The thermodynamically favored phase at a given temperature and pressure is the phase with the lowest Gibbs free energy. Since the entropy term in the Gibbs free energy is multiplied by the temperature T, the entropic contribution to the Gibbs free energy becomes increasingly important at higher temperatures. The entropy of materials has two main components: the vibrational entropy, S_{vib}, which can be calculated from the phonon dispersion or the Debye model as described in Section 7.3, and the configuration entropy, S_{config}, due to the disorder in the atomic positions or site occupations. Configurational entropy originates from chemical disorder as in the case of high entropy alloys in which all of the atoms are arranged on a regular lattice (but the specific lattice sites are randomly occupied by different chemical species) or structural disorder as in the

case of metallic glasses, where the atoms no longer occupy regular lattice sites, resulting in an amorphous material.

7.5.1.1 High Entropy Materials

High entropy materials display structural order (i.e. all of the atoms are arranged on a periodic crystal lattice) but chemical disorder (i.e. the actual occupation of these lattice sites is random) [162].

In the ideal entropy limit in which the occupation of the atomic sites is completely random, the configuration entropy per atom is given by $S_{\text{config}} = k_B \sum_i x_i \log_e(x_i)$ [162], where x_i is the fractional composition of each species component. Note that this expression increases with increasing numbers of species and is also maximized when all of the values of x_i are equal, i.e. for equimolar compositions.

The expression for the ideal entropy can be combined with calculations for *special quasirandom structures* (SQS) [163], which are special structural configurations where the radial correlation functions mimic those of a perfectly random structure, to estimate the Gibbs free energy for high entropy alloys. This can then be used in conjunction with the energies of the ordered phases obtained from computational materials data repositories such as AFLOW [2, 4–6] to generate structural phase diagrams as a function of temperature and composition, predicting the phase transition boundaries between ordered compounds, phase separation regions, and single-phase solid solutions [107, 164, 165]. The calculated ordered structure energies in AFLOW can also be used to train cluster expansion models [166] to predict the energies of large ensembles of configurations, which can be combined with thermodynamic descriptors to estimate the transition temperature and miscibility gaps for solid solutions and high entropy alloys [167].

The concept of entropy stabilization has recently been extended beyond metallic alloys to include multicomponent ceramics, such as high entropy oxides [168, 169]. High entropy oxides consist of an ordered anion sublattice occupied by oxygen ions, with a disordered cation sublattice randomly occupied by five different metal ions, such as Co, Cu, Mg, Ni, and Zn [168, 169]. The oxygen ions screen the metal ions from each other, reducing the energy cost associated with forming a random configuration of the metal ions, enabling the formation of a single-phase, entropy-stabilized ceramic.

7.5.1.2 Metallic Glasses

Metallic glasses are alloys in which the atoms do not occupy the sites of a regular periodic lattice, but instead form a structurally disordered amorphous phase. These materials are of great commercial and industrial interest due to their unique combination of superb mechanical properties [170] and plastic-like processability [171–173] for several potential applications [174–178].

Several different attempts have been made to understand the formation of metallic glasses and predict the GFA of different alloy compositions. Most of these efforts center around maximizing the packing density of the different atoms [179], which requires elements with a range of different atomic radii [180–184]. Other efforts have been made to use phase diagram data on liquidus

temperatures to predict GFA [185–187]. Work is also underway to use machine learning techniques to predict potential glass formers [188].

Much of the theoretical work described above relies on the use of experimental rather than *ab initio* computational data to predict new materials, due to the difficulty of modeling amorphous structures using first principles techniques. However, Perim et al. [48] recently demonstrated that the energies of different structural phases can be combined into a descriptor to predict the formation of metallic glasses. If there are many different structural phases with similar formation enthalpy, this will frustrate crystallization during solidification and thus promote glass formation [48]. This frustration can be quantified to formulate a spectral descriptor for GFA using the structural and energetic information available in computational materials data repositories such as AFLOW [2, 4–6]. The differences in the geometry between two structures are quantified by describing each structure in terms of its atomic environments [189–191], while the formation enthalpy differences between the respective structures are expressed in the form of Boltzmann factors. The energetic and structural descriptors can then be combined with appropriate normalization factors to formulate a spectral descriptor for GFA as function of composition x: GFA $(\{x\})$. Comparisons with known glass-forming compositions available in the literature can then be used to define a threshold, such that if GFA($\{x\}$) exceeds this threshold, then the composition x would be expected to be glass forming.

The GFA($\{x\}$) descriptor has been used to perform an automated analysis of the GFA of over 1400 binary alloy systems from the AFLOW data repository [48]. While over half of all binary alloy systems are predicted to have a GFA below that of the threshold, nevertheless some 17% of alloy systems display a maximum value of *GFA*($\{x\}$) greater than the maximum value for the Cu–Zr system, a well-known good glass former. These included several alloy systems for which glass formation had never previously been observed or sometimes even investigated, suggesting that there are many possible glass forming compositions that remain to be discovered. This success demonstrates the power of combining descriptors based on the easily calculated properties of periodic crystalline phases with large pre-calculated databases for predicting the synthesizability of complex disordered materials.

7.5.1.3 Modeling Off-Stoichiometry Materials

Incorporating the effects of disorder is a necessary, albeit difficult, step in materials modeling. Not only is disorder intrinsic to all materials, but it also offers a route to enhanced and even otherwise inaccessible functionality, as demonstrated by its ubiquity in technological applications. Prominent examples include fuel cells [192], high-temperature superconductors [193, 194], and low thermal conductivity thermoelectrics [195].

Specifically, chemical disorder can arise in the form of doping, vacancies, and even in the occupation of lattice sites themselves (random), which cannot inherently be modeled using periodic systems. One approach for modeling such effects includes SQS [163]. These quasirandom approximates are very computationally effective but only offer a single representation of the disordered states, i.e. that with the lowest site correlations. Instead of reducing down to a single

representation, AFLOW treats such systems as an ensemble of ordered supercells [196]. Properties are resolved through ensemble averages of the representative states, with opportunities to optimize computation (via supercell size/site error) and tune the level of disorder explored (via parameter T). AFLOW partial occupation module (AFLOW-POCC) has already resolved significant stoichiometric trends in wide-gap semiconductors and magnetic systems while offering additional insight into underlying physical mechanisms. Ultimately, the screening criteria and property predictions generated by these *bona fide* thermodynamic models and descriptors are accelerating design of new, technologically significant materials, including advanced ceramics [197] and metallic glasses [48].

7.5.2 Superalloys

Superalloys are characterized by their extraordinary mechanical properties, particularly at temperatures near their melting point. Such traits make them the ideal candidates for applications in the aerospace and power generation industries. Among the more common examples, many have a face-centered cubic structure with base elements nickel, cobalt, and iron, though nickel-based superalloys dominate the market. A novel cobalt-based superalloy, $Co_3(Al,W)$, was discovered in 2006 that exhibits mechanical properties better than many nickel-based superalloys. This inspired a thorough computational investigation with AFLOW of alloys containing 40 different elements, yielding over 2224 relevant ternary systems [58]. The search offered 102 systems shown to (i) be more stable than $Co_3[Al_{0.5},W_{0.5}]$, the $L1_2$-like random structure previously characterized thermodynamically [198] and very close to the compositions reported by experiments [199], (ii) have a relevant concentration ($X_3[A_xB_{1-x}]$) that is in two-phase equilibrium with the host matrix, and (iii) exhibit only small deviations from the host matrix lattice (within 5% relative mismatch).

For these 102 candidates, additional pertinent properties were extracted, including the density and bulk modulus (as a proxy for hardness). Low density materials are preferred to mitigate the stress on turbine components. Significant trends for the bulk modulus are elucidated when plotted with respect to component B on a Pettifor scale: Ni-based materials show a peak at or before Ni, whereas Co-based materials monotonically increase. Additionally, Co-based materials are generally more resistant to compression compared with Ni-based materials.

Of the 102 candidates, 37 materials have no reported phase diagrams in standard databases and are thus expected to be unexplored or new. Additional screening based on the toxicity and (low) melting temperature of components uncovered six priority candidates for experimental validation.

7.5.3 Thermoelectrics

Thermoelectric materials generate an electric voltage when subjected to a temperature gradient and can also generate a temperature gradient when a voltage is applied [200, 201]. Their lack of moving parts and resulting scalability means that they have potential applications in power generation for spacecraft, energy recovery from waste heat in automotive and industrial facilities [202, 203] and

in spot cooling for nanoelectronics using the Peltier cooling effect [202, 203]. However, most of the available thermoelectric materials have low efficiency, only converting a few percent of the available thermal energy into electricity. Therefore, a major goal of thermoelectrics research is to develop new materials that have higher thermoelectric efficiency.

The thermoelectric efficiency of a material is determined by the figure of merit zT, which is obtained from [200, 201]

$$zT = \frac{\sigma S^2 T}{\kappa_{\mathrm{L}} + \kappa_{\mathrm{e}}} \tag{7.35}$$

where S is the Seebeck coefficient, σ is the electrical conductivity, κ_{L} is the lattice thermal conductivity, and κ_{e} is the electronic thermal conductivity. The lattice thermal conductivity κ_{L} can be calculated using the methods described in Section 7.3. Most of the electronic thermal conductivity κ_{e} will depend directly on the electrical conductivity σ through the Wiedemann–Franz law [200]

$$\kappa_{\mathrm{e}} = L\sigma T \tag{7.36}$$

where L is the Lorenz factor, which has a value of 2.4×10^{-8} J^2/(K$^2 \cdot$ C^2) for free electrons. The Seebeck coefficient S is given by [200]

$$S = \frac{8\pi^2 k_{\mathrm{B}}^2}{3eh^2} m^* T \left(\frac{\pi}{3n} \right)^{\frac{2}{3}} \tag{7.37}$$

where n is the charge carrier concentration, e is the electronic charge, and m^* is the density of states effective mass of the charge carriers in the material. The effective mass tensor m_{ij} can be calculated from the curvature of electronic band structure dispersion $E(\vec{k})$:

$$m_{ij}^{-1} = \frac{1}{\hbar^2} \frac{\mathrm{d}^2 E}{\mathrm{d}k_i \mathrm{d}k_j} \tag{7.38}$$

where k_i and k_j are components of the wave vector $\vec{k}$. Larger curvature of the band structure implies a lower effective mass, while flat narrow bands tend to result in a large effective mass. Charge carrier mobility and thus electrical conductivity tend to reduce with increasing effective mass. However, as can be seen from Eq. (7.37), the Seebeck coefficient increases with effective mass, and κ_{e} also increases with σ. Therefore, a compromise should be found between high effective mass to maximize S and high charge carrier mobility to give high σ in order to optimize the thermoelectric efficiency of the device.

Several computational high-throughput searches have been performed for thermoelectric materials [37–40, 115, 204–206]. Many of the efforts toward developing more efficient thermoelectric materials have focused on either lowering the lattice thermal conductivity κ_{L} or finding materials in which the electronic properties are highly directional, allowing for a narrow energy band distribution while simultaneously having a low effective mass, thus increasing the power factor σS^2. High-throughput searches for materials with low lattice thermal conductivity have focused on materials such as half-Heusler structures [39, 40, 207], which have lower densities and thus lower thermal conductivities than the full Heusler structures. Other promising materials include structures such as

clathrates [208–212] and skutterudites [206, 213–215], which contain hollow voids that can be filled with "rattler" atoms to reduce the lattice thermal conductivity. Filled skutterudites in particular, such as $R_xCo_4Sb_{12}$, are excellent thermoelectric materials because of their combination of a high effective mass with high carrier mobility due to the existence of a secondary conduction band with 12 conducting charge carrier pockets [206].

Searches of large databases of inorganic materials to find new thermoelectrics include the study of 48 000 materials from the Materials Project database [204], where the power factor was calculated using the BoltzTraP code [216] and the thermal conductivity was estimated using the Clarke [217] and Cahill–Pohl [218] models. Almost 600 oxides, nitrides, and sulfides from the ICSD were investigated by Garrity [115], where the lattice thermal conductivity was calculated at the quasi-harmonic phonon level of approximation, with particular attention being paid to degeneracies in the conduction band minimum, or materials with strongly anisotropic conduction bands, that result in an effective low-dimensional conductor with a corresponding increase in the power factor. The thermoelectric material LiZnSb was proposed by an automated search of the calculated band structures of 1640 compounds in the ICSD containing Sb [38], although later experimental measurements did not find a high thermoelectric efficiency for this compound [219].

Other strategies to increase the power factor include engineering the band structure [220] through volume changes by alloying different materials to create solid solutions, such as antifluorite Mg_2Si and Mg_2Ge with Mg_2Sn or orthorhombic Ca_2Si and Ca_2Ge with Ca_2Sn [221]. Tuning the composition of alloys can also be used to converge the valence and conduction bands, enabling high valley degeneracy to be achieved in materials such as $PbTe_{1-x}Se_x$ alloys [222]. Solid solutions can also produce local anisotropic structural disorder, increasing phonon scattering and thus improving the thermoelectric efficiency [223, 224].

The exploitation of thermodynamic phenomena such as spinodal decomposition to self-assemble heterostructures with increased phonon scattering [225] has also been proposed to enhance the efficiency of thermoelectric devices. In this approach, materials such as PbSe and PbTe, which are miscible at high temperatures, undergo phase separation when the mixture is cooled slowly, creating a layered heterostructure with a network of boundaries between the different components, which scatter phonons and thus suppress the thermal conductivity. This concept has also been extended to other nanotechnology applications, e.g. as a means to embed a network of electrically conducting nanowires, in the form of topologically protected interface states, within an insulating matrix [226].

The combination of different competing materials properties that must be optimized to maximize the thermoelectric efficiency highlights the importance of integrated frameworks such as AFLOW, which can automatically calculate different types of materials properties such as thermal conductivity and electronic band structures. Having all of these electronic and thermal properties calculated and available in an integrated, searchable, sortable data repository such as AFLOW.org accelerates the design of new, high-efficiency thermoelectric materials.

7.5.4 Magnetic Materials

The search for new magnetic systems remains a long-standing challenge despite their ubiquity in modern technology [227]. Magnetism demonstrates remarkable sensitivity to a number of properties, including electronic configuration, bond length/angle, and magnetic ion valence, and thus its presence is rather uncommon and difficult to predict. In fact, only two percent of the known inorganic compounds [52] exhibit magnetic order of any kind. Consumer applications place additional practical restrictions for magnets, with the current global market effectively populated by only two dozen compounds. These obstacles motivated a large-scale computational search with AFLOW for new magnets among the Heusler structure family. Heusler structures are of particular interest for a number of reasons: (i) several are known high-performance magnets, (ii) the breadth of distinct compounds offers an excellent chance for discovery, (iii) the full set of materials will likely offer other types of interesting materials (aside from magnets), and (iv) they are metallic and thus well described by DFT. There are three types of Heusler structures, i.e. the regular-Heuslers X_2YZ (Cu_2MnAl-type), inverse-Heuslers $(XY)XZ$ (Hg_2CuTi-type), and the half-Heuslers XYZ (MgCuSb-type). By decorating these prototypes with ternary combinations of 55 elements, a total of 236 115 compounds were generated and added to the AFLOW.org repository.

As a first attempt, the analysis is limited to Heuslers containing elements of the 3d, 4d, and 5d periods, i.e. a subset of 36 540 compounds. Of this set, 248 are determined to be thermodynamically stable, and 22 have a magnetic ground state compatible with the unit cells considered. Among these 22 magnetic ground-state compounds, a few prominent classes can be identified, including Co_2YZ and Mn_2YZ. Upon further analysis of these classes, four materials were of particular interest. In the first class Co_2YZ, there already exists 25 known compounds all lying on the Slater–Pauling curve (magnetic moment per formula unit vs. number of valence electrons) [228]. The regression predicts Co_2MnTi to have the notably high Curie transition temperature T_C of 940 K – a feature shared by only two dozen known magnets. The second class Mn_2YZ is of interest because of their high T_C and potentially large magnetocrystalline anisotropy [229]. Two known examples from this class, Mn_2VAl and Mn_2VGa, show ferrimagnetic ordering, matching two candidates from the list of 22, Mn_2PtCo and Mn_2PtV. One more compound was highlighted for satisfying a stringent thermodynamic constraint. Mn_2PdPt is robustly stable by at least 30 meV, where the criterion derives from the distance of the stable phase from the pseudo-convex hull that neglects it. This criterion quantifies the impact of the structure on the minimum energy surface.

Following an attempt to synthesize these four candidates, two were successful (Co_2MnTi and Mn_2PtPd), and the other two decomposed into binary compounds. In fact, Co_2MnTi shows a T_C of 938 K, almost exactly as predicted by the Slater–Pauling curve. Surprisingly, Mn_2PdPt shows antiferromagnetic ordering and tetragonal distortion ($c/a \sim 1.8$), a result corroborated by calculation upon further analysis. Beyond the synthesis of these two systems, this investigation offers a new, accelerated pathway to materials discovery over traditional trial-and-error approaches.

7.6 Conclusion

Automated computational materials design frameworks have the capability to rapidly generate materials data without the need for laborious human intervention. They are being used to construct large repositories of programmatically accessible materials properties, calculated in a standardized, consistent fashion so as to facilitate the identification of trends and the training of machine learning models to predict electronic, thermal, and mechanical behavior. When combined with physical models and intelligently formulated descriptors, the data becomes a powerful tool to accelerate the discovery of new materials for applications ranging from high-temperature superalloys to thermoelectrics and magnets.

Acknowledgments

References

1 Curtarolo, S., Hart, G.L.W., Buongiorno Nardelli, M. et al. (2013). The high-throughput highway to computational materials design. *Nat. Mater.* 12: 191–201.

2 Curtarolo, S., Setyawan, W., Hart, G.L.W. et al. (2012). AFLOW: an automatic framework for high-throughput materials discovery. *Comput. Mater. Sci.* 58: 218–226.

3 Setyawan, W. and Curtarolo, S. (2010). High-throughput electronic band structure calculations: challenges and tools. *Comput. Mater. Sci.* 49: 299–312.

4 Curtarolo, S., Setyawan, W., Wang, S. et al. (2012). AFLOWLIB.ORG: a distributed materials properties repository from high-throughput ab initio calculations. *Comput. Mater. Sci.* 58: 227–235.

5 Taylor, R.H., Rose, F., Toher, C. et al. (2014). A RESTful API for exchanging materials data in the AFLOWLIB.org consortium. *Comput. Mater. Sci.* 93: 178–192.

6 Calderon, C.E., Plata, J.J., Toher, C. et al. (2015). The AFLOW standard for high-throughput materials science calculations. *Comput. Mater. Sci.* 108 Pt. A: 233–238.

7 Setyawan, W. and Curtarolo, S. (2011). AflowLib: Ab-initio electronic structure library database. http://www.aflow.org (accessed 17 April 2019).

8 Toher, C., Oses, C., Hicks, D. et al. (2018). The AFLOW Fleet for Materials Discovery. In: *Handbook of Materials Modeling* (ed. W. Andreoni and S. Yip), 1–28 Cham, Switzerland: Springer International Publishing. doi: 10.1007/978-3-319-42913-7_63-1.

9 Supka, A.R., Lyons, T.E., Liyanage, L.S.I. et al. (2017). AFLOWπ: a minimalist approach to high-throughput *ab initio* calculations including the generation of tight-binding hamiltonians. *Comput. Mater. Sci.* 136: 76–84.

10 Buongiorno Nardelli, M., Cerasoli, F.T., Costa, M. et al. (2017). PAOFLOW: a utility to construct and operate on ab initio Hamiltonians from the projections of electronic wavefunctions on atomic orbital bases, including characterization of topological materials. *Comput. Mater. Sci.* 143: 462–472.

11 Jain, A., Hautier, G., Moore, C.J. et al. (2011). A high-throughput infrastructure for density functional theory calculations. *Comput. Mater. Sci.* 50: 2295–2310.

12 Jain, A., Ong, S.P., Hautier, G. et al. (2013). Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* 1: 011002.

13 Ong, S.P., Richards, W.D., Jain, A. et al. (2013). Python Materials Genomics (pymatgen): a robust, open-source python library for materials analysis. *Comput. Mater. Sci.* 68: 314–319.

14 Mathew, K., Montoya, J.H., Faghaninia, A. et al. (2017). Atomate: a high-level interface to generate, execute, and analyze computational materials science workflows. *Comput. Mater. Sci.* 139: 140–152.

15 Saal, J.E., Kirklin, S., Aykol, M. et al. (2013). Materials design and discovery with high-throughput density functional theory: the Open Quantum Materials Database (OQMD). *JOM* 65: 1501–1509.

16 Kirklin, S., Meredig, B., and Wolverton, C. (2013). High-throughput computational screening of new Li-Ion battery anode materials. *Adv. Energy Mater.* 3: 252–262.

17 Kirklin, S., Saal, J.E., Hegde, V.I., and Wolverton, C. (2016). High-throughput computational search for strengthening precipitates in alloys. *Acta Mater.* 102: 125–135.

18 Landis, D.D., Hummelshøj, J.S., Nestorov, S. et al. (2012). The computational materials repository. *Comput. Sci. Eng.* 14: 51–57.

19 Bahn, S.R. and Jacobsen, K.W. (2002). An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.* 4: 56–66.

20 Pizzi, G., Cepellotti, A., Sabatini, R. et al. (2016). AiiDA. http://www.aiida .net (accessed 17 April 2019).

21 Pizzi, G., Cepellotti, A., Sabatini, R. et al. (2016). AiiDA: automated interactive infrastructure and database for computational science. *Comput. Mater. Sci.* 111: 218–230.

22 Mounet, N., Gibertini, M., Schwaller, P. et al. (2018). Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds. *Nat. Nanotechnol.* 13: 246–252.

23 Scheffler, M. and Draxl, C. (2014). *Computer Center of the Max-Planck Society*. Garching: The NoMaD Repository. http://nomad-repository.eu.

24 Merkys, A., Mounet, N., Cepellotti, A. et al. (2017). A posteriori metadata from automated provenance tracking: integration of AiiDA and TCOD. *J. Cheminform* 9: 56.

25 Yu, L. and Zunger, A. (2012). Identification of potential photovoltaic absorbers based on first-principles spectroscopic screening of materials. *Phys. Rev. Lett.* 108: 068701.

26 Castelli, I.E., Olsen, T., Datta, S. et al. (2012). Computational screening of perovskite metal oxides for optimal solar light capture. *Energy Environ. Sci.* 5: 5814–5819.

27 Lin, L.-C., Berger, A.H., Martin, R.L. et al. (2012). In silico screening of carbon-capture materials. *Nat. Mater.* 11: 633–641.

28 Alapati, S.V., Johnson, J.K., and Sholl, D.S. (2008). Large-scale screening of metal hydride mixtures for high-capacity hydrogen storage from first-principles calculations. *J. Phys. Chem. C* 112: 5258–5262.

29 Derenzo, S., Bizarri, G., Borade, R. et al. (2011). New scintillators discovered by high-throughput screening. *Nucl. Inst. Methods Phys. Res. A* 652: 247–250.

30 Ortiz, C., Eriksson, O., and Klintenberg, M. (2009). Data mining and accelerated electronic structure theory as a tool in the search for new functional materials. *Comput. Mater. Sci.* 44: 1042–1049.

31 Setyawan, W., Gaumé, R.M., Lam, S. et al. (2011). High-throughput combinatorial database of electronic band structures for inorganic scintillator materials. *ACS Comb. Sci.* 13: 382–390.

32 Setyawan, W., Gaumé, R.M., Feigelson, R.S., and Curtarolo, S. (2009). Comparative study of nonproportionality and electronic band structures features in scintillator materials. *IEEE Trans. Nucl. Sci.* 56: 2989–2996.

33 Yang, K., Setyawan, W., Wang, S. et al. (2012). A search model for topological insulators with high-throughput robustness descriptors. *Nat. Mater.* 11: 614–619.

34 Lin, H., Wray, L.A., Xia, Y. et al. (2010). Half-Heusler ternary compounds as new multifunctional experimental platforms for topological quantum phenomena. *Nat. Mater.* 9: 546–549.

35 Armiento, R., Kozinsky, B., Fornari, M., and Ceder, G. (2011). Screening for high-performance piezoelectrics using high-throughput density functional theory. *Phys. Rev. B* 84: 014103.

36 Roy, A., Bennett, J.W., Rabe, K.M., and Vanderbilt, D. (2012). Half-Heusler semiconductors as piezoelectrics. *Phys. Rev. Lett.* 109: 037602.

37 Wang, S., Wang, Z., Setyawan, W. et al. (2011). Assessing the thermo-electric properties of sintered compounds via high-throughput Ab-Initio calculations. *Phys. Rev. X* 1: 021012.

38 Madsen, G.K.H. (2006). Automated search for new thermoelectric materials: the case of LiZnSb. *J. Am. Chem. Soc.* 128: 12140–12146.

39 Carrete, J., Li, W., Mingo, N. et al. (2014). Finding unprecedentedly low-thermal-conductivity half-Heusler semiconductors via high-throughput materials modeling. *Phys. Rev. X* 4: 011019.

40 Carrete, J., Mingo, N., Wang, S., and Curtarolo, S. (2014). Nanograined half-Heusler semiconductors as advanced thermoelectrics: An Ab initio high-throughput statistical study. *Adv. Funct. Mater.* 24: 7427–7432.

41 Nørskov, J.K., Bligaard, T., Rossmeisel, J., and Christensen, C.H. (2009). Towards the computational design of solid catalysts. *Nat. Chem.* 1: 37–46.

42 Hautier, G., Jain, A., Chen, H. et al. (2011). Novel mixed polyanions lithium-ion batery cathode materials predicted by high-throughput. *ab initio* computations. *J. Mater. Chem.* 21: 17147–17153.

43 Hautier, G., Jain, A., Ong, S.P. et al. (2011). Phosphates as lithium-ion battery cathodes: an evaluation based on high-throughput *ab initio* calculations. *Chem. Mater.* 23: 3495–3508.

44 Mueller, T., Hautier, G., Jain, A., and Ceder, G. (2011). Evaluation of tavorite-structured cathode materials for lithium-ion batteries using high-throughput computing. *Chem. Mater.* 23: 3854–3862.

45 Isayev, O., Oses, C., Toher, C. et al. (2017). Universal fragment descriptors for predicting properties of inorganic crystals. *Nat. Commun.* 8: 15679.

46 de Jong, M., Chen, W., Notestine, R. et al. (2016). A statistical learning framework for materials science: application to elastic moduli of k-nary inorganic polycrystalline compounds. *Sci. Rep.* 6: 34256.

47 Isayev, O., Fourches, D., Muratov, E.N. et al. (2015). Materials cartography: representing and mining materials space using structural and electronic fingerprints. *Chem. Mater.* 27: 735–743.

48 Perim, E., Lee, D., Liu, Y. et al. (2016). Spectral descriptors for bulk metallic glasses based on the thermodynamics of competing crystalline phases. *Nat. Commun.* 7: 12315.

49 Toher, C., Plata, J.J., Levy, O. et al. (2014). High-throughput computational screening of thermal conductivity, Debye temperature, and Grüneisen parameter using a quasiharmonic Debye model. *Phys. Rev. B* 90: 174107.

50 de Jong, M., Chen, W., Angsten, T. et al. (2015). Charting the complete elastic properties of inorganic crystalline compounds. *Sci. Data* 2: 150009.

51 Toher, C., Oses, C., Plata, J.J. et al. (2017). Combining the AFLOW GIBBS and elastic libraries to efficiently and robustly screen thermomechanical properties of solids. *Phys. Rev. Mater.* 1: 015401.

52 Bergerhoff, G., Hundt, R., Sievers, R., and Brown, I.D. (1983). The inorganic crystal structure data base. *J. Chem. Inf. Comput. Sci.* 23: 66–69.

53 Mehl, M.J., Hicks, D., Toher, C. et al. (2017). The AFLOW Library of Crystallographic Prototypes: Part 1. *Comput. Mater. Sci.* 136: S1–S828.

54 Hart, G.L.W. and Forcade, R.W. (2008). Algorithm for generating derivative structures. *Phys. Rev. B* 77: 224115.

55 Hart, G.L.W. and Forcade, R.W. (2009). Generating derivative structures from multilattices: algorithm and application to HCP alloys. *Phys. Rev. B* 80: 014120.

56 Curtarolo, S., Morgan, D., and Ceder, G. (2005). Accuracy of *ab initio* methods in predicting the crystal structures of metals: a review of 80 binary alloys. *Calphad* 29: 163–211.

57 Hart, G.L.W., Curtarolo, S., Massalski, T.B., and Levy, O. (2013). Comprehensive search for new phases and compounds in binary alloy systems based

on Platinum-Group metals, using a computational first-principles approach. *Phys. Rev. X* 3: 041035.

58 Nyshadham, C., Oses, C., Hansen, J.E. et al. (2017). A computational high-throughput search for new ternary superalloys. *Acta Mater.* 122: 438–447.

59 Ong, S.P., Wang, L., Kang, B., and Ceder, G. (2008). Li-Fe-P-O$_2$ phase diagram from first principles calculations. *Chem. Mater.* 20: 1798–1807.

60 Akbarzadeh, A.R., Ozoliņš̆, V., and Wolverton, C. (2007). First-principles determination of multicomponent hydride phase diagrams: application to the Li-Mg-N-H system. *Adv. Mater.* 19: 3233–3239.

61 Levy, O., Hart, G.L.W., and Curtarolo, S. (2010). Uncovering compounds by synergy of cluster expansion and high-throughput methods. *J. Am. Chem. Soc.* 132: 4830–4833.

62 Levy, O., Hart, G.L.W., and Curtarolo, S. (2010). Hafnium binary alloys from experiments and first principles. *Acta Mater.* 58: 2887–2897.

63 Levy, O., Chepulskii, R.V., Hart, G.L.W., and Curtarolo, S. (2010). The new face of rhodium alloys: revealing ordered structures from first principles. *J. Am. Chem. Soc.* 132: 833–837.

64 Levy, O., Jahnátek, M., Chepulskii, R.V. et al. (2011). Ordered structures in rhenium binary alloys from first-principles calculations. *J. Am. Chem. Soc.* 133: 158–163.

65 Jahnátek, M., Levy, O., Hart, G.L.W. et al. (2011). Ordered phases in ruthenium binary alloys from high-throughput first-principles calculations. *Phys. Rev. B* 84: 214110.

66 Levy, O., Xue, J., Wang, S. et al. (2012). Stable ordered structures of binary technetium alloys from first principles. *Phys. Rev. B* 85: 012201.

67 Chepulskii, R.V. and Curtarolo, S. (2011). Revealing low-temperature atomic ordering in bulk Co-Pt with the high-throughput *ab-initio* method. *Appl. Phys. Lett.* 99: 261902.

68 Taylor, R.H., Curtarolo, S., and Hart, G.L.W. (2010). Ordered magnesium-lithium alloys: first-principles predictions. *Phys. Rev. B* 81: 024112.

69 Taylor, R.H., Curtarolo, S., and Hart, G.L.W. (2011). Guiding the experimental discovery of magnesium alloys. *Phys. Rev. B* 84: 084101.

70 Chepulskii, R.V. and Curtarolo, S. (2009). Calculation of solubility in titanium alloys from first principles. *Acta Mater.* 57: 5314–5323.

71 Bloch, J., Levy, O., Pejova, B. et al. (2012). Prediction and hydrogen acceleration of ordering in iron–vanadium alloys. *Phys. Rev. Lett.* 108: 215503.

72 Mehl, M.J., Finkenstadt, D., Dane, C. et al. (2015). Finding the stable structures of $N_{1-x}W_x$ with an *ab initio* high-throughput approach. *Phys. Rev. B* 91: 184110.

73 Levy, O., Hart, G.L.W., and Curtarolo, S. (2010). Structure maps for hcp metals from first-principles calculations. *Phys. Rev. B* 81: 174106.

74 Taylor, R.H., Curtarolo, S., and Hart, G.L.W. (2010). Predictions of the Pt_8Ti phase in unexpected systems. *J. Am. Chem. Soc.* 132: 6851–6854.

75 Nelson, L.J., Hart, G.L.W., and Curtarolo, S. (2012). Ground-state characterizations of systems predicted to exhibit $L1_1$ or $L1_3$ crystal structures. *Phys. Rev. B* 85: 054203.

76 Hohenberg, P. and Kohn, W. (1964). Inhomogeneous electron gas. *Phys. Rev.* 136: B864–B871.

77 Kohn, W. and Sham, L.J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev.* 140: A1133.

78 Perdew, J.P. and Zunger, A. (1981). Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* 23: 5048–5079.

79 Zupan, A., Blaha, P., Schwarz, K., and Perdew, J.P. (1998). Pressure-induced phase transitions in solid Si, SiO_2, and Fe: performance of local-spin-density and generalized-gradient-approximation density functionals. *Phys. Rev. B* 58: 11266.

80 Perdew, J.P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett.* 77: 3865–3868.

81 Lee, C., Yang, W., and Parr, R.G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37: 785.

82 Sun, J., Ruzsinszky, A., and Perdew, J.P. (2015). Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* 115: 036402.

83 Liechtenstein, A.I., Anisimov, V.I., and Zaanen, J. (1995). Density-functional theory and strong interactions: orbital ordering in Mott-Hubbard insulators. *Phys. Rev. B* 52: R5467.

84 Dudarev, S.L., Botton, G.A., Savrasov, S.Y. et al. (1998). Electron-energy-loss spectra and the structural stability of nickel oxide: an LSDA+U study. *Phys. Rev. B* 57: 1505–1509.

85 Becke, A.D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys* 98: 5648.

86 Heyd, J., Scuseria, G.E., and Ernzerhof, M. (2003). Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* 118: 8207–8215.

87 Agapito, L.A., Curtarolo, S., and Buongiorno Nardelli, M. (2015). Reformulation of DFT + U as a pseudohybrid Hubbard density functional for accelerated materials discovery. *Phys. Rev. X* 5: 011006.

88 Hedin, L. (1965). New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* 139: A796–A823.

89 Aryasetiawan, F. and Gunnarsson, O. (1998). The *GW* method. *Rep. Prog. Phys.* 61: 237.

90 Kresse, G. and Hafner, J. (1993). *Ab initio* molecular dynamics for liquid metals. *Phys. Rev. B* 47: 558–561.

91 Kresse, G. and Furthmüller, J. (1996). Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* 54: 11169–11186.

92 Kresse, G. and Furthmüller, J. (1996). Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* 6: 15–50.

93 Kresse, G. and Joubert, D. (1999). From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* 59: 1758–1775.

94 Giannozzi, P., Baroni, S., Bonini, N. et al. (2009). QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* 21: 395502.

95 Giannozzi, P., Andreussi, O., Brumme, T. et al. (2017). Advanced capabilities for materials modelling with QUANTUM ESPRESSO. *J. Phys. Condens. Matter* 29: 465901.

96 Gonze, X., Beuken, J.-M., Caracas, R. et al. (2002). First-principles computation of material properties: the ABINIT software project. *Comput. Mater. Sci.* 25: 478–492.

97 Gonze, X., Amadon, B., Anglade, P.-M. et al. (2009). ABINIT: first-principles approach to materials and nanosystem properties. *Comput. Phys. Commun.* 180: 2582–2615.

98 Blum, V., Gehrke, R., Hanke, F. et al. (2009). *Ab initio* molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* 180: 2175–2196.

99 Soler, J.M., Artacho, E., Gale, J.D. et al. (2002). The SIESTA method for ab initio order-N materials simulation. *J. Phys. Condens. Matter* 14: 2745.

100 Frisch, M.J., Trucks, G.W., Schlegel, H.B. et al. (2009). *Gaussian09 Revision D.01*. Wallingford, CT: Gaussian, Inc.

101 Hehre, W.J., Stewart, R.F., and Pople, J.A. (1969). Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of slater-type atomic orbitals. *J. Chem. Phys.* 51: 2657–2664.

102 Monkhorst, H.J. and Pack, J.D. (1976). Special points for Brillouin-zone integrations. *Phys. Rev. B* 13: 5188.

103 Wisesa, P., McGill, K.A., and Mueller, T. (2016). Efficient generation of generalized Monkhorst–Pack grids through the use of informatics. *Phys. Rev. B* 93: 155109.

104 Greaves, G.N., Greer, A.L., Lakes, R.S., and Rouxel, T. (2011). Poisson's ratio and modern materials. *Nat. Mater.* 10: 823–837.

105 Poirier, J.-P. (2000). *Introduction to the Physics of the Earth's Interior*, 2e. Cambridge University Press.

106 Mouhat, F. and Coudert, F.-X. (2014). Necessary and sufficient elastic stability conditions in various crystal systems. *Phys. Rev. B* 90: 224104.

107 Barzilai, S., Toher, C., Curtarolo, S., and Levy, O. (2016). Evaluation of the tantalum-titanium phase diagram from *ab-initio* calculations. *Acta Mater.* 120: 255–263.

108 Chen, X.-Q., Niu, H., Li, D., and Li, Y. (2011). Modeling hardness of polycrystalline materials and bulk metallic glasses. *Intermetallics* 19: 1275–1281.

109 Teter, D.M. (1998). Computational alchemy: the search for new superhard materials. *MRS Bull.* 23: 22–27.

110 Hashin, Z. and Shtrikman, S. (1963). A variational approach to the theory of the elastic behaviour of multiphase materials. *J. Mech. Phys. Solids* 11: 127–140.

111 Zohdi, T.I. and Wriggers, P. (2001). Aspects of the computational testing of the mechanical properties of microheterogeneous material samples. *Int. J. Numer. Methods Eng.* 50: 2573–2599.

112 Anderson, O.L., Schreiber, E., Liebermann, R.C., and Soga, N. (1968). Some elastic constant data on minerals relevant to geophysics. *Rev. Geophys.* 6: 491–524.

113 Karki, B.B., Stixrude, L., and Wentzcovitch, R.M. (2001). High-pressure elastic properties of major materials of Earth's mantle from first principles. *Rev. Geophys.* 39: 507–534.

114 Zebarjadi, M., Esfarjani, K., Dresselhaus, M.S. et al. (2012). Perspectives on thermoelectrics: from fundamentals to device applications. *Energy Environ. Sci.* 5: 5147–5162.

115 Garrity, K.F. (2016). First principles search for n-type oxide, nitride and sulfide thermoelectrics. *Phys. Rev. B* 94: 045122.

116 Yeh, L.-T. and Chu, R.C. (2002). *Thermal Management of Microloectronic Equipment: Heat Transfer Theory, Analysis Methods, and Design Practices*. ASME Press.

117 Wright, C.D., Wang, L., Shah, P. et al. (2011). The design of rewritable ultra-high density scanning-probe phase-change memories. *IEEE Trans. Nanotechnol.* 10: 900–912.

118 Watari, K. and Shinde, S.L. (2001). High thermal conductivity materials. *MRS Bull.* 26: 440–444.

119 Slack, G.A., Tanzilli, R.A., Pohl, R.O., and Vandersande, J.W. (1987). The intrinsic thermal conductivity of AlN. *J. Phys. Chem. Solids* 48: 641–647.

120 Toberer, E.S., Zevalkink, A., and Snyder, G.J. (2011). Phonon engineering through crystal chemistry. *J. Mater. Chem.* 21: 15843–15852.

121 Nye, J.F. (1985). *Physical Properties of Crystals: Their Representation by Tensors and Matrices*. Oxford Science Publications (Clarendon Press).

122 Maradudin, A.A., Montroll, E.W., Weiss, G.H., and Ipatova, I.P. (1971). *Theory of Lattice Dynamics in the Harmonic Approximation*. New York: Academic Press.

123 Stokes, H.T. and Hatch, D.M. (2005). FINDSYM: program for identifying the space group symmetry of a crystal. *J. Appl. Crystallogr.* 38: 237–238.

124 Stokes, H.T. (1995). Using symmetry in frozen phonon calculations. *Ferroelectrics* 164: 183–188.

125 Spek, A.L. (2003). Single-crystal structure validation with the program PLATON. *J. Appl. Crystallogr.* 36: 7–13.

126 Togo, A. and Tanaka, I. (2017). Spglib: a software library for crystal symmetry search. https://atztogo.github.io/spglib/ (accessed 17 April 2019).

127 Hloucha, M. and Deiters, U.K. (1998). Fast coding of the minimum image convention. *Mol. Simul.* 20: 239–244.

128 Hicks, D., Oses, C., Gossett, E. et al. (2018). *AFLOW-SYM*: platform for the complete, automatic and self-consistent symmetry analysis of crystals. *Acta Crystallogr., Sect. A: Found. Adv.* 74: 184–203.

129 Hahn, T. (ed.) (2002). *International Tables of Crystallography. Volume A: Space-Group Symmetry*. Chester, England: Kluwer Academic publishers, International Union of Crystallography.

130 Golesorkhtabar, R., Pavone, P., Spitaler, J. et al. (2013). ElaStic: a tool for calculating second-order elastic constants from first principles. *Comput. Phys. Commun.* 184: 1861–1873.

131 da Silveira, P.R.C., da Silva, C.R.S., and Wentzcovitch, R.M. (2008). Metadata management for distributed first principles calculations in VLab-A collaborative cyberinfrastructure for materials computation. *Comput. Phys. Commun.* 178: 186–198.

132 da Silva, C.R.S., da Silveira, P.R.C., Karki, B. et al. (2007). Virtual laboratory for planetary materials: system service architecture overview. *Phys. Earth Planet. Inter.* 163: 321–332.

133 Hill, R. (1952). The elastic behaviour of a crystalline aggregate. *Proc. Phys. Soc., Sect. A* 65: 349.

134 Blanco, M.A., Francisco, E., and Luaña, V. (2004). GIBBS: isothermal-isobaric thermodynamics of solids from energy curves using a quasi-harmonic Debye model. *Comput. Phys. Commun.* 158: 57–72.

135 Birch, F. (1938). The effect of pressure upon the elastic parameters of isotropic solids, according to Murnaghan's theory of finite strain. *J. Appl. Phys.* 9: 279.

136 Vinet, P., Rose, J.H., Ferrante, J., and Smith, J.R. (1989). Universal features of the equation of state of solids. *J. Phys. Condens. Matter* 1: 1941–1963.

137 Baonza, V.G., Cáceres, M., and Núñez, J. (1995). Universal compressibility behavior of dense phases. *Phys. Rev. B* 51: 28–37.

138 Leibfried, G. and Schlömann, E. (1954). *Wärmeleitung in elektrisch isolierenden Kristallen*, Nachrichten d.Akadd. Wiss. in Göttingen. Math.-physik. Kl. 2a. Math.-physik.-chem. Abt Vandenhoeck & Ruprecht.

139 Slack, G.A. (1979). The thermal conductivity of nonmetallic crystals. In: *Solid State Physics*, vol. 34 (ed. H. Ehrenreich, F. Seitz, and D. Turnbull), 1–71. New York: Academic Press.

140 Morelli, D.T. and Slack, G.A. (2006). High lattice thermal conductivity solids. In: *High Thermal Conductivity Materials* (ed. S.L. Shinde and J.S. Goela), 37–68. New York, NY: Springer.

141 Wee, D., Kozinsky, B., Pavan, B., and Fornari, M. (2012). Quasiharmonic vibrational properties of TiNiSn from *ab-initio* phonons. *J. Electron. Mater.* 41: 977–983.

142 Bjerg, L., Iversen, B.B., and Madsen, G.K.H. (2014). Modeling the thermal conductivities of the zinc antimonides ZnSb and Zn_4Sb_3. *Phys. Rev. B* 89: 024304.

143 Ashcroft, N.W. and Mermin, N.D. (1976). *Solid State Physics*. Philadelphia, PA: Holt-Saunders.

144 Dove, M.T. (1993). *Introduction to Lattice Dynamics*, Cambridge Topics in Mineral Physics and Chemistry. Cambridge University Press.

145 Parlinski, K. (2010). Computing for materials: phonon software. http://www.computingformaterials.com/phoncfm/3faq/100softmode1.html.

146 Nath, P., Plata, J.J., Usanmaz, D. et al. (2016). High-throughput prediction of finite-temperature properties using the quasi-harmonic approximation. *Comput. Mater. Sci.* 125: 82–91.

147 Nath, P., Plata, J.J., Usanmaz, D. et al. (2017). High throughput combinatorial method for fast and robust prediction of lattice thermal conductivity. *Scr. Mater.* 129: 88–93.

148 Broido, D.A., Malorny, M., Birner, G. et al. (2007). Intrinsic lattice thermal conductivity of semiconductors from first principles. *Appl. Phys. Lett.* 91: 231922.

149 Li, W., Mingo, N., Lindsay, L. et al. (2012). Thermal conductivity of diamond nanowires from first principles. *Phys. Rev. B* 85: 195436.

150 Ward, A., Broido, D.A., Stewart, D.A., and Deinzer, G. (2009). *Ab initio* theory of the lattice thermal conductivity in diamond. *Phys. Rev. B* 80: 125203.

151 Ward, A. and Broido, D.A. (2010). Intrinsic phonon relaxation times from first-principles studies of the thermal conductivities of Si and Ge. *Phys. Rev. B* 81: 085205.

152 Zhang, Q., Cao, F., Lukas, K. et al. (2012). Study of the thermoelectric properties of lead selenide doped with boron, gallium, indium, or thallium. *J. Am. Chem. Soc.* 134: 17731–17738.

153 Li, W., Lindsay, L., Broido, D.A. et al. (2012). Thermal conductivity of bulk and nanowire $Mg_2Si_xSn_{1-x}$ alloys from first principles. *Phys. Rev. B* 86: 174307.

154 Lindsay, L., Broido, D.A., and Reinecke, T.L. (2013). First-principles determination of ultrahigh thermal conductivity of boron arsenide: a competitor for diamond?. *Phys. Rev. Lett.* 111: 025901.

155 Lindsay, L., Broido, D.A., and Reinecke, T.L. (2013). Ab initio thermal transport in compound semiconductors. *Phys. Rev. B* 87: 165201.

156 Li, W., Carrete, J., Katcho, N.A., and Mingo, N. (2014). ShengBTE: a solver of the Boltzmann transport equation for phonons. *Comput. Phys. Commun.* 185: 1747–1758.

157 Plata, J.J., Nath, P., Usanmaz, D. et al. (2017). An efficient and accurate framework for calculating lattice thermal conductivity of solids: AFLOW-AAPL Automatic Anharmonic Phonon Library. *NPJ Comput. Mater.* 3: 45.

158 Legrain, F., Carrete, J., van Roekeghem, A. et al. (2017). How chemical composition alone can predict vibrational free energies and entropies of solids. *Chem. Mater.* 29: 6220–6227.

159 Gossett, E., Toher, C., Oses, C. et al. (2018). AFLOW-ML: a RESTful API for machine-learning predictions of materials properties. *Comput. Mater. Sci.* 152: 134–145.

160 Meredig, B. and Mulholland, G. (2015). Citrine informatics. http://www.citrine.io (accessed 17 April 2019).

161 Rose, F., Toher, C., Gossett, E. et al. (2017). AFLUX: the LUX materials search API for the AFLOW data repositories. *Comput. Mater. Sci.* 137: 362–370.

162 Widom, M. (2016). Prediction of structure and phase transformations. In: *High-Entropy Alloys: Fundamentals and Applications*, Chapter 8 (ed. M.C. Gao, J.-W. Yeh, P.K. Liaw, and Y. Zhang). Cham: Springer 267–298.

163 Zunger, A., Wei, S.-H., Ferreira, L.G., and Bernard, J.E. (1990). Special quasirandom structures. *Phys. Rev. Lett.* 65: 353–356.

164 Barzilai, S., Toher, C., Curtarolo, S., and Levy, O. (2017). The effect of lattice stability determination on the computational phase diagrams of intermetallic alloys. *J. Alloys Compd.* 728: 314–321.

165 Barzilai, S., Toher, C., Curtarolo, S., and Levy, O. (2017). Molybdenum-titanium phase diagram evaluated from *ab initio* calculations. *Phys. Rev. Mater.* 1: 023604.

166 van de Walle, A., Asta, M.D., and Ceder, G. (2002). The alloy theoretic automated toolkit: a user guide. *Calphad* 26: 539–553.

167 Lederer, Y., Toher, C., Vecchio, K.S., and Curtarolo, S. (2018). The search for high entropy alloys: a high-throughput ab-initio approach, *Acta Mater.* 159: 364–383.

168 Rost, C.M., Sachet, E., Borman, T. et al. (2015). Entropy-stabilized oxides. *Nat. Commun.* 6: 8485.

169 Rak, Z., Rost, C.M., Lim, M. et al. (2016). Charge compensation and electrostatic transferability in three entropy-stabilized oxides: results from density functional theory calculations. *J. Appl. Phys.* 120: 095105.

170 Chen, W., Ketkaew, J., Liu, Z. et al. (2015). Does the fracture toughness of bulk metallic glasses scatter? *Scr. Mater.* 107: 1–4.

171 Schroers, J. and Paton, N. (2006). Amorphous metal alloys form like plastics. *Adv. Mater. Processes* 164: 61.

172 Schroers, J., Hodges, T.M., Kumar, G. et al. (2011). Thermoplastic blow molding of metals. *Mater. Today* 14: 14–19.

173 Kaltenboeck, G., Demetriou, M.D., Roberts, S., and Johnson, W.L. (2016). Shaping metallic glasses by electromagnetic pulsing. *Nat. Commun.* 7: 10576.

174 Johnson, W.L. (1999). Bulk glass-forming metallic alloys: science and technology. *MRS Bull.* 24: 42–56.

175 Greer, A.L. (2009). Metallic glasses…on the threshold. *Mater. Today* 12: 14–22.

176 Schroers, J. (2010). Processing of bulk metallic glass. *Adv. Mater.* 22: 1566–1597.

177 Johnson, W.L., Na, J.H., and Demetriou, M.D. (2016). Quantifying the origin of metallic glass formation. *Nat. Commun.* 7: 10313.

178 Ashby, M.F. and Greer, A.L. (2006). Metallic glasses as structural materials. *Scr. Mater.* 54: 321–326.

179 Miracle, D.B. (2004). A structural model for metallic glasses. *Nat. Mater.* 3: 697–702.

180 Egami, T. and Waseda, Y. (1984). Atomic size effect on the formability of metallic glasses. *J. Non-Cryst. Solids* 64: 113–134.

181 Greer, A.L. (1993). Confusion by design. *Nature* 366: 303–304.

182 Egami, T. (2003). Atomistic mechanism of bulk metallic glass formation. *J. Non-Cryst. Solids* 317: 30–33.

183 Lee, H.-J., Cagin, T., Johnson, W.L., and Goddard, W.A. III (2003). Criteria for formation of metallic glasses: the role of atomic size ratio. *J. Chem. Phys.* 119: 9858–9870.

184 Zhang, K., Dice, B., Liu, Y. et al. (2015). On the origin of multi-component bulk metallic glasses: atomic size mismatches and de-mixing. *J. Chem. Phys.* 143: 054501.

185 Cheney, J. and Vecchio, K. (2009). Evaluation of glass-forming ability in metals using multi-model techniques. *J. Alloys Compd.* 471: 222–240.

186 Cheney, J. and Vecchio, K. (2007). Prediction of glass-forming compositions using liquidus temperature calculations. *Mater. Sci. Eng., A* 471: 135–143.

187 Lu, Z.P. and Liu, C.T. (2002). A new glass-forming ability criterion for bulk metallic glasses. *Acta Mater.* 50: 3501–3512.

188 Ward, L., Agrawal, A., Choudhary, A., and Wolverton, C. (2016). A general-purpose machine learning framework for predicting properties of inorganic materials. *NPJ Comput. Mater.* 2: 16028.

189 Villars, P. (2000). Factors governing crystal structures. In: *Crystal Structures of Intermetallic Compounds* (ed. J.H. Westbrook and R.L. Fleisher), 1–49. New York: Wiley.

190 Daams, J.L.C. (2000). Atomic environments in some related intermetallic structure types. In: *Crystal Structures of Intermetallic Compounds* (ed. J.H. Westbrook and R.L. Fleisher), 139–159. New York: Wiley.

191 Daams, J.L.C. and Villars, P. (2000). Atomic environments in relation to compound prediction. *Eng. Appl. Artif. Intell.* 13: 507–511.

192 Xie, L., Brault, P., Coutanceau, C. et al. (2015). Efficient amorphous platinum catalyst cluster growth on porous carbon: a combined molecular dynamics and experimental study. *Appl. Catal. B* 162: 21–26.

193 Bednorz, J.G. and Müller, K.A. (1986). Possible high T_c superconductivity in the Ba-La-Cu-O system. *Z. Phys. B: Condens. Matter* 64: 189–193.

194 Maeno, Y., Hashimoto, H., Yoshida, K. et al. (1994). Superconductivity in a layered perovskite without copper. *Nature* 372: 532–534.

195 Winter, M.R. and Clarke, D.R. (2007). Oxide materials with low thermal conductivity. *J. Am. Ceram. Soc.* 90: 533–540.

196 Yang, K., Oses, C., and Curtarolo, S. (2016). Modeling off-stoichiometry materials with a high-throughput *Ab-Initio* approach. *Chem. Mater.* 28: 6484–6492.

197 Rohrer, G.S., Affatigato, M., Backhaus, M. et al. (2012). Challenges in ceramic science: a report from the workshop on emerging research areas in ceramic science. *J. Am. Ceram. Soc.* 95: 3699–3712.

198 Saal, J.E. and Wolverton, C. (2013). Thermodynamic stability of Co-Al-W L1$_2$ γ'. *Acta Mater.* 61: 2330–2338.

199 Sato, J., Omori, T., Oikawa, K. et al. (2006). Cobalt-Base High-Temperature Alloys. *Science* 312: 90–91.

200 Snyder, G.J. and Toberer, E.S. (2008). Complex thermoelectric materials. *Nat. Mater.* 7: 105–114.

201 Nolas, G.S., Sharp, J., and Goldsmid, H.J. (2001). *Thermoelectrics: Basic Principles and New Materials Developments*. Springer-Verlag.

202 Bell, L.E. (2008). Cooling, heating, generating power, and recovering waste heat with thermoelectric systems. *Science* 321: 1457–1461.

203 DiSalvo, F.J. (1999). Thermoelectric cooling and power generation. *Science* 285: 703–706.

204 Chen, W., Pöhls, J.-H., Hautier, G. et al. (2016). Understanding thermoelectric properties from high-throughput calculations: trends, insights, and comparisons with experiment. *J. Mater. Chem. C* 4: 4414–4426.

205 Zhu, H., Hautier, G., Aydemir, U. et al. (2015). Computational and experimental investigation of TmAgTe$_2$ and *XYZ*$_2$ compounds, a new group

of thermoelectric materials identified by first-principles high-throughput screening. *J. Mater. Chem. C* 3: 10554–10565.

206 Tang, Y., Gibbs, Z.M., Agapito, L.A. et al. (2015). Convergence of multi-valley bands as the electronic origin of high thermoelectric performance in $CoSb_3$ skutterudites. *Nat. Mater.* 14: 1223–1228.

207 Zeier, W.G., Schmitt, J., Hautier, G. et al. (2016). Engineering half-Heusler thermoelectric materials using Zintl chemistry. *Nat. Rev. Mater.* 1: 16032.

208 Shi, X., Yang, J., Bai, S. et al. (2010). On the design of high-efficiency thermoelectric clathrates through a systematic cross-substitution of framework elements. *Adv. Func. Mater.* 20: 755–763.

209 Zhang, H., Borrmann, H., Oeschler, N. et al. (2011). Atomic interactions in the p-type clathrate I $Ba_8Au_{5.3}Ge_{40.7}$. *Inorg. Chem.* 50: 1250–1257.

210 Saiga, Y., Du, B., Deng, S.K. et al. (2012). Thermoelectric properties of type-VIII clathrate $Ba_8Ga_{16}Sn_{30}$ doped with Cu. *J. Alloys Compd.* 537: 303–307.

211 Christensen, M., Johnsen, S., and Iversen, B.B. (2010). Thermoelectric clathrates of type I. *Dalton Trans.* 39: 978–992.

212 Madsen, G.K.H., Katre, A., and Bera, C. (2016). Calculating the thermal conductivity of the silicon clathrates using the quasi-harmonic approximation. *Phys. Status Solidi A* 213: 802–807.

213 Sales, B.C., Mandrus, D., and Williams, R.K. (1996). Filled skutterudite antimonides: a new class of thermoelectric materials. *Science* 272: 1325–1328.

214 Bai, S.Q., Pei, Y.Z., Chen, L.D. et al. (2009). Enhanced thermoelectric performance of dual-element-filled skutterudites $Ba_xCe_yCo_4Sb_{12}$. *Acta Mater.* 57: 3135–3139.

215 Yang, J., Qiu, P., Liu, R. et al. (2011). Trends in electrical transport of *p*-type skutterudites RFe_4Sb_{12} (R=Na, K, Ca, Sr, Ba, La, Ce, Pr, Yb) from first-principles calculations and Boltzmann transport theory. *Phys. Rev. B* 84: 235205.

216 Madsen, G.K.H. and Singh, D.J. (2006). BoltzTraP. A code for calculating band-structure dependent quantities. *Comput. Phys. Commun.* 175: 67–71.

217 Clarke, D.R. (2003). Materials selection guidelines for low thermal conductivity thermal barrier coatings. *Surf. Coat. Technol.* 163–164: 67–74.

218 Cahill, D.G., Braun, P.V., Chen, G. et al. (2014). Nanoscale thermal transport. II. 2003-2012. *Appl. Phys. Rev.* 1: 011305.

219 Toberer, E.S., May, A.F., Scanlon, C.J., and Snyder, G.J. (2009). Thermoelectric properties of p-type LiZnSb: Assessment of *ab initio* calculations. *J. Appl. Phys.* 105: 063701.

220 Pei, Y., Wang, H., and Snyder, G.J. (2012). Band engineering of thermoelectric materials. *Adv. Mater.* 24: 6125–6135.

221 Bhattacharya, S. and Madsen, G.K.H. (2015). High-throughput exploration of alloying as design strategy for thermoelectrics. *Phys. Rev. B* 92: 085205.

222 Pei, Y., Shi, X., LaLonde, A. et al. (2011). Convergence of electronic bands for high performance bulk thermoelectrics. *Nature* 473: 66–69.

223 Zeier, W.G., LaLonde, A., Gibbs, Z.M. et al. (2012). Influence of a nano phase segregation on the thermoelectric properties of the p-type doped stannite compound $Cu_{2+x}Zn_{1-x}GeSe_4$. *J. Am. Chem. Soc.* 134: 7147–7154.

224 Zeier, W.G., Pei, Y., Pomrehn, G. et al. (2012). Phonon scattering through a local anisotropic structural disorder in the thermoelectric solid solution $Cu_2Zn_{1-x}Fe_xGeSe_4$. *J. Am. Chem. Soc.* 135: 726–732.

225 Usanmaz, D., Nath, P., Plata, J.J. et al. (2016). First principles thermodynamical modeling of the binodal and spinodal curves in lead chalcogenides. *Phys. Chem. Chem. Phys.* 18: 5005–5011.

226 Usanmaz, D., Nath, P., Toher, C. et al. (2018). Spinodal superlattices of topological insulators. *Chem. Mater.* 30: 2331–2340.

227 Sanvito, S., Oses, C., Xue, J. et al. (2017). Accelerated discovery of new magnets in the Heusler alloy family. *Sci. Adv.* 3: e1602241.

228 Graf, T., Felser, C., and Parkin, S.S.P. (2011). Simple rules for the understanding of Heusler compounds. *Prog. Solid State Chem.* 39: 1–50.

229 Kreiner, G., Kalache, A., Hausdorf, S. et al. (2014). New Mn_2-based Heusler compounds. *Z. Anorg. Allg. Chem.* 640: 738–752.