

## 3.12 First principles crystal structure prediction

Lewis J. Conway<sup>a</sup>, Chris J. Pickard<sup>a,b</sup>, and Andreas Hermann<sup>c</sup>, <sup>a</sup> Department of Materials Science & Metallurgy, University of Cambridge, Cambridge, United Kingdom; <sup>b</sup> Advanced Institute for Materials Research, Tohoku University, Sendai, Japan; and <sup>c</sup> School of Physics and Astronomy, The University of Edinburgh, Edinburgh, United Kingdom

© 2023 Elsevier Ltd. All rights reserved.

<b>3.12.1</b>	<b>Introduction</b>	<b>393</b>
<b>3.12.2</b>	<b>Crystalline configuration space</b>	<b>394</b>
<b>3.12.3</b>	<b>First principles calculations in solids</b>	<b>396</b>
<b>3.12.4</b>	<b>Modern approaches</b>	<b>397</b>
3.12.4.1	Random structure searching	397
3.12.4.2	Particle swarm optimization	398
3.12.4.3	Genetic and evolutionary algorithms	398
3.12.4.4	Database informed methods	400
3.12.4.5	Local approaches	401
3.12.4.5.1	Metadynamics	401
3.12.4.5.2	Minima hopping	402
3.12.4.5.3	Basin hopping	402
3.12.4.5.4	Following soft modes	403
<b>3.12.5</b>	<b>Fitness functions</b>	<b>403</b>
3.12.5.1	Compositional stability	404
3.12.5.2	Structural and mechanical properties	405
3.12.5.3	Electronic properties	405
<b>3.12.6</b>	<b>Accelerating the search process</b>	<b>406</b>
3.12.6.1	Hard sphere potentials and volumes	406
3.12.6.2	Modular decomposition	407
3.12.6.3	Orbital free DFT	407
3.12.6.4	Machine learned potentials	407
<b>3.12.7</b>	<b>Visualizing the PES</b>	<b>408</b>
<b>3.12.8</b>	<b>Examples</b>	<b>409</b>
3.12.8.1	Compositional complexity	409
3.12.8.2	High pressure	410
3.12.8.2.1	Elements	410
3.12.8.2.2	Hydrogen bonding	410
3.12.8.2.3	Polyhydride superconductors	411
3.12.8.3	Target properties	411
<b>3.12.9</b>	<b>Challenges and conclusions</b>	<b>411</b>
<b>Acknowledgments</b>		<b>413</b>
<b>References</b>		<b>413</b>

### Abstract

First principles calculations, usually in the shape of density functional theory, are an indispensable component of modern day solid-state inorganic chemistry. Over the past 15 years, those types of calculations have successfully evolved from explanatory to predictive tools. At the heart of this development is the now widely recognized ability of density functional theory to predict, without any input from experiment or recourse to chemical intuition, the crystal structures of inorganic compounds. In this chapter, we begin by presenting arguments why crystal structure prediction works, seemingly against the combinatorial odds, due to the shapes of the potential energy landscapes in crystalline configuration space. We then review the modern approaches used in the field that sample this space, either on a global scale aimed at finding global and metastable minima, or on a local scale aimed at finding nearby minima and transition states. We include an overview of attempts to tailor searches toward desired materials properties, and discuss ongoing work to make structure prediction faster and more efficient. Finally, we give examples from fields where crystal structure prediction is, or could be, a successful driving force for new research avenues, and conclude by pointing toward some of the remaining challenges.

### 3.12.1 Introduction

Determining the structure of a material has been a problem faced by computational scientists for decades. Without this information, it is difficult to predict or explain material properties. As the rise in computational resources matched the development of accurate

first principles calculation methods, high-throughput approaches have become more feasible routes to material design and can be used to study a wider variety of systems. This approach allows efficient screening of material properties, particularly in unusual chemical environments.

The study of a material by experiment is a labor intensive process. To use a trial-and-error approach experimentally to design and study new materials is prohibitively slow. Whereas, a computer — by its nature — can perform repetitive tasks and screen large numbers of potential compounds quite straightforwardly. Scientists who previously spent years studying a handful of materials, can now uncover insights into novel materials with relative ease using crystal structure prediction (CSP). That is not to say that high-throughput experiments are not possible. For instance, brilliant x-ray radiation from x-ray free electron lasers can be used in ‘high-throughput’ approaches to crystal structure determination.<sup>1</sup>

Even if such high-throughput experiments were more common, the results are not as definitive as those obtained in calculations. Typically, the crystal structure is inferred from diffraction based methods. Due to restrictions caused by noise, low symmetry, and sample synthesis, diffraction may only provide partial or inconclusive information about the crystal structure. This process can be helped significantly if guided by initial calculations which can provide likely stable stoichiometries, material properties, and diffraction patterns.

It would have been difficult to conceive of a novel high- $T_c$  superconducting sulfur hydride as a decomposition product of  $H_2S$ , without CSP.<sup>2</sup> Recent machine learned high-throughput methods can be used to quickly estimate superconducting temperatures<sup>3</sup> of a wide variety of hydrides, extending the predictive power of CSP; promising metastable polymorphs with desirable  $T_c$ 's can be quickly identified.

It would also be difficult to sample a wide range of complex stoichiometric environments without CSP. Cathode materials can consist of a wide variety of quaternary or even quinary stoichiometries. The space associated with conceivable structures of these types is very large, but CSP can help identify low energy polymorphs<sup>4,5</sup> with promising properties.

However, the methods used to obtain total energies and material properties are always approximate solutions to the problem, typically a trade-off between speed and accuracy. For example, in some cases, an ionic pair potential or force field is sufficient to calculate energies and is computationally very fast. In other cases, e.g. where electronic properties are relevant, such a potential might be insufficient to describe the behavior. The latter typically employs a density functional based approach, which in turn might be problematic for instance for strongly correlated materials. In either case, the search results should not be considered the final answer; temperature effects are difficult to consider in structure searches, and the choice of level of theory can influence the results. For these reasons, both experimental and theoretical methods should be used together to uncover the accurate picture; CSP should motivate experiments and vice versa.

This chapter will introduce the methods used to perform effective structure prediction from first principles. To appreciate the power of CSP, we begin by discussing historical approaches to determine molecular and crystal structures before introducing modern approaches to the global minimization problem.

An early example of theoretical structure prediction preceding an experimental discovery is that of van't Hoff, who postulated — to account for the optical activity of carbon-containing compounds — that a central carbon atom should be bonded to its neighbors such as to form a tetrahedron.<sup>6</sup> Van't Hoff's idea of this asymmetric carbon atom led to the idea that an allene — which takes the form of  $RR'C=C=CRR'$  where R represents a functional group — should be chiral. Many years later, this was confirmed experimentally.<sup>7</sup>

Following the development of x-ray diffraction and the discovery of atom-level periodicity in crystals, more systematic predictions were sought. One approach is to predict the likely coordination number for a given stoichiometry based on ionic radii ratios; a concept first applied to extended lattices by Victor Goldschmidt in 1926<sup>8</sup> and 1927.<sup>9</sup> For example, a 1:1 ionic binary compound with an ionic ratio of around 0.3 will likely be tetrahedrally coordinated in a cubic sphalerite structure or hexagonal wurtzite structure. For ratios around 0.6, it will likely form octahedrally coordinated structures such as rock-salt, and if it is close to 1 (i.e. the ions are similar in size), it will likely be 8-fold coordinated like CsCl. Similar arguments can be made for the stability of perovskite structures ( $ABO_3$ ). Intuition about the expected coordination of the atoms and the stoichiometry — in other words, to appeal to conventional wisdom — may be enough to generate some likely candidates from previously known structures. However, there are many ways to stack structures tetrahedrally or to stack perovskite units, and the conventional behavior of chemical bonding is different when pressure is applied. Additionally, stoichiometric variability is a concern. The ionic radii rules will not generate stable structures of the type  $Na_3Cl$ ,  $Na_2Cl$ ,  $Na_3Cl_2$ ,  $NaCl_3$  forming under pressure,<sup>2</sup> or the aforementioned  $H_3S$ <sup>10</sup> forming from the decomposition of  $H_2S$ .

With the advancement of density functional theory techniques, it became feasible to calculate the stable configurations of crystal structures, including the predictions mentioned in the previous paragraph. While computational limitations would mean initially limiting the candidates to small primitive unit cells with high symmetry, computational resources sufficient for unrestricted (and importantly, unsupervised) structure searches have been available for close to 20 years now.

While the determination of the global energy minima is a daunting task, the nature of the potential energy surface means it is surmountable. There are many crystal structure prediction methods, including evolutionary algorithms<sup>11,12</sup> and machine learning methods,<sup>13</sup> all of which have been successful in predicting novel crystal structures, particularly at high pressures<sup>14–21</sup> including the structure of phase III of solid hydrogen,<sup>16</sup> a transparent phase of sodium,<sup>22</sup> and superhard tungsten nitride.<sup>23,24</sup>

### 3.12.2 Crystalline configuration space

We begin by making a case for why crystal structure prediction works, despite the apparently overwhelming number of degrees of freedom on offer.

Long-range translational symmetry sets crystals apart from localized or extended amorphous systems. This symmetry is captured by defining three *lattice vectors* ( $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ ) (in three-dimensional space) such that the crystal is invariant under any translation vector  $\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$  that is a linear combination of the lattice vectors. The *atomic basis*, the set of atomic positions within the parallelepiped spanned by the lattice vectors (the *unit cell*), completes the definition of a specific crystal structure. The unit cell is not uniquely defined; an infinite number of choices of lattice vectors are possible that describe the same crystal. A *primitive unit cell* is the smallest possible unit cell (by volume or number of atoms), but is also not uniquely defined; several approaches exist that construct uniquely defined reduced cells for a given lattice, which involve finding the shortest possible primitive unit cell edges.<sup>25–28</sup>

Every lattice (in real space) has a corresponding *reciprocal lattice* (in reciprocal space), with lattice vectors ( $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ ) defined via the identities  $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$ . The reciprocal lattice (with units of inverse length) is important whenever phenomena are studied that are modulated by the crystal lattice, such as harmonic lattice vibrations, phonons, with specific wave vectors  $\mathbf{q}$  that reside in the reciprocal unit cell; see Section 3.12.4.5.4. More importantly, Bloch's theorem shows that single-electron states in a periodic potential have wave functions of the type  $\psi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r}) \exp i\mathbf{k}\mathbf{r}$ , with a lattice-periodic component  $u_{\mathbf{k}}(\mathbf{r})$  and a phase factor determined by wave vector  $\mathbf{k}$ , which resides in the first Brillouin zone, the region of reciprocal space closer to the origin than to any other reciprocal lattice vector; this is important in any electronic structure calculation, see Section 3.12.3.

The configurational space for an  $N$ -atom primitive unit cell is  $3N + 3$  dimensional ( $3N - 3$  dimensions for atomic positions, 6 dimensions for the unit cell). A specific unit cell volume  $V = NV_{\text{at}}$  or, equivalently, number density  $n = 1/V_{\text{at}}$ , which introduce a characteristic atomic volume  $V_{\text{at}}$ , reduces the dimension by one. A combinatorial approach shows<sup>29,30</sup> that the remaining space for atomic positions alone is incredibly vast. If atoms are placed within the unit cell on a discretized grid with spacing  $\delta$ , i.e.  $S = V_{\text{at}}/\delta^3$  sites per atom, the number of configurations  $C$  is

$$C = \binom{NS}{N} \approx (2\pi N)^{-1/2} e^N S^N,$$

utilizing Stirling's approximation. The exponential growth of the number of possible configurations makes exhaustive exploration of the space impossible. For compounds, where not all atom types are equal, an additional binomial factor increases  $C$  even further, but the largest contribution remains the freedom of placing atoms in space.

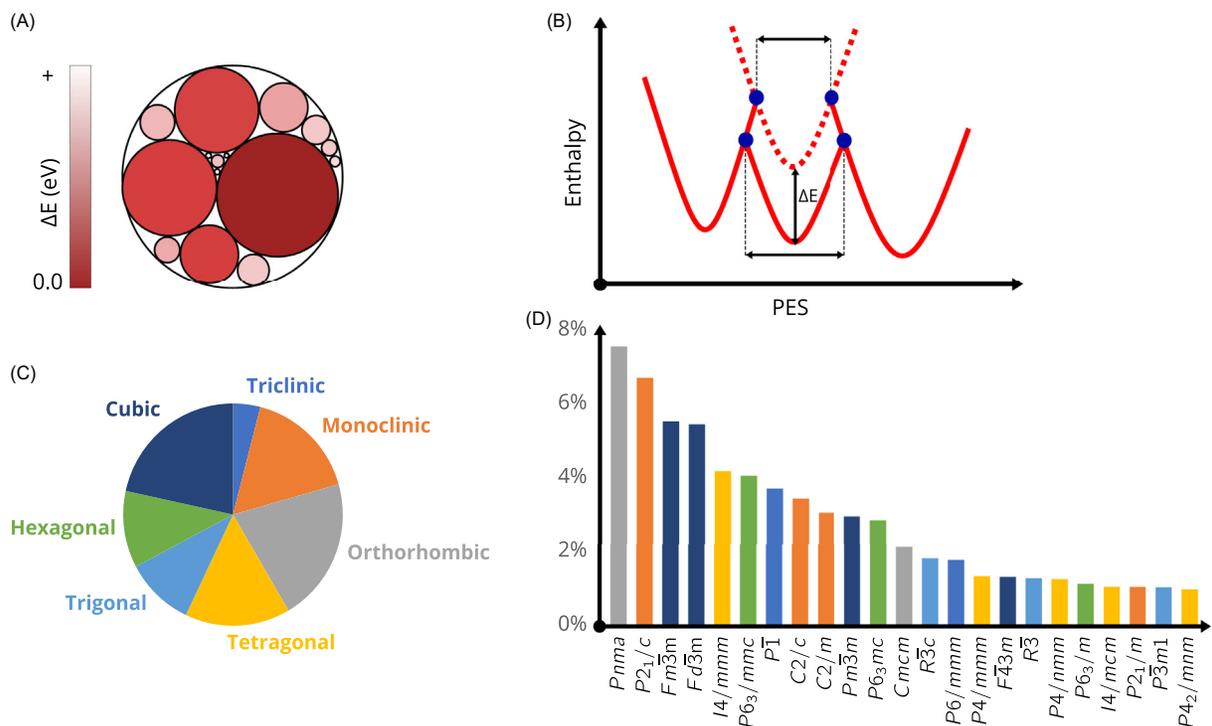
However, since the potential energy surface (PES), the energy associated with any point in the configurational space, is the result of physical interactions, it is, in most cases, smooth. Unphysical configurations, such as those with overlapping atoms, can be considered irrelevant for materials synthesis processes and ruled out altogether from detailed exploration. Most other configurations do not correspond to energy minima. Nonetheless, the number of local minima  $C_{\text{min}}$  still scales exponentially with system size,<sup>31,32</sup>  $C_{\text{min}} \sim e^{\alpha N}$ . However, the topology of the potential energy surface and the energy spectrum of these local minima have been subject to several studies that yield promising insights.

Doye and co-workers noted that lower energy minima on the PES tend to have larger basins of attraction.<sup>33,34</sup> They show that topologically, the PES resembles an Apollonian packing of spheres, in that the largest (and deepest) basins have the highest number of connections to adjacent basins, just like the largest Apollonian spheres have the most touching neighbors,<sup>35</sup> see Fig. 1. One consequence is that by performing geometry optimizations to find the *local* minimum from a given starting configuration, the likelihood of finding the global minimum is significantly increased. In fact, taken to its extreme, this basin size distribution implies that the global minimum has the highest probability of *any* minimum to be found with a random starting structure; however, due to the sheer number of local minima the *absolute* probability can still be very low, requiring many initial attempts. Another consequence of the expected high connectivity of the global minimum basin is that local exploration of the PES from some initial local minimum is a reasonable strategy to find more favorable minima.

The Bell-Evans-Polanyi principle provides some arguments for the basin size distribution and correlation with the minimum energy. It states that the smaller the energy barrier from one basin to a neighboring basin, the lower the neighboring basin's energy. In other words, low energy minima are close together. In a similar manner to the Marcus equation for electron transfer, this principle can be shown by assuming identical quadratic forms for neighboring basins with minima A and B, along a reaction coordinate  $x$ , at  $x = 0$  and  $x = 1$ . The result is that the transition coordinate is  $x = \frac{1}{2}$  for a thermoneutral reaction (that is to say both minima are equal) and that the transition coordinate and energy barrier is lower if minimum B is lower in energy than minimum A.<sup>36</sup> An alternative schematic is shown in Fig. 1: lowering the minimum energy of the central basin will both lower the transition barriers to access it from its neighbors, and increase its size.

The Bell-Evans-Polanyi principle relies on the assumption that basins have comparable shape; we assume equal curvatures in the previous paragraph. This is not unreasonable: a reaction coordinate connecting two local minima across a transition state represents coupled atomic displacements and lattice strains, and the restoring forces and stresses, in the harmonic/linear elastic regime, can be expected to be similar, if comparable bond breaking and bond formation processes occur around both minima.

Another assist for CSP is symmetry. Crystal structures can be highly symmetric, and high symmetry atomic configurations are often very low or very high energy minima.<sup>37</sup> In organic crystals, the uneven distribution of space groups (the 230 symmetry groups for three-dimensional crystals) has been documented for a long time<sup>38,39</sup> and is related to close-packing of unevenly shaped objects.<sup>40–42</sup> That restriction means space groups that combine inversion centers, 2-fold screw axes, and glide planes, are preferred<sup>43</sup> and indeed the six most dominant space groups in the Cambridge Structural Database (CSD), which are of that type, make up almost 83% of the over 1.1 m organic crystal structures in the CSD.<sup>44</sup> (Note that Kitaigorodskii<sup>41,42</sup> built and described a mechanical



**Fig. 1** Crystalline PES characteristics. (A) Basin size distribution as Apollonian sphere packing, energy axis indicated; (B) Bell-Evans-Polanyi principle connecting a basin's energy minimum to its size (black dotted bars) and transition barriers to adjacent basins (blue dots); (C) crystal system distribution in the ICSD database; (D) space group frequency of the most prominent space groups in the ICSD database.

'structure seeker' to explore and in effect predict close packed molecular crystals). In inorganic systems, the symmetry requirements and therefore their statistical distribution are markedly different. Fig. 1 shows crystal system and space group frequencies extracted from the over 200,000 entries of the Inorganic Crystal Structure Database (ICSD).<sup>45</sup> The different crystal systems are relatively evenly distributed, but within them particular space groups stand out. Those tend to be of the highest symmetries that each crystal system offers and suggest a symmetry principle in the formation of inorganic crystals.<sup>39</sup> The 20 most prominent space group account for just over 63% of all ICSD crystal structures. For structure prediction, this could mean that symmetry-constraint searches, or at least initial guesses, could be a fruitful avenue to explore. But note that four of the 20 most relevant space groups are triclinic or monoclinic; insisting on high symmetry for a particular compound might lead only to metastable states.

### 3.12.3 First principles calculations in solids

The exploration of the crystalline configurational space can, in principle, be done with any type of interatomic potential that attaches total energies to specific configurations and generates forces to optimize said configurations. Using off-the-shelf interatomic potentials or force fields is problematic if one aims to explore different conditions, e.g. high pressure, but even at ambient conditions every structure prediction approach (see next section for details) involves creating structural candidates that are initially far from structural equilibrium (database mining approaches would be an exception), where force field parameterization might be poor. Even so, in organic crystal structure prediction the use of force fields remains common, due to otherwise prohibitive computational cost,<sup>46</sup> but is most often combined with very sophisticated and specialized algorithms to generate structural candidates that avoid unphysical configurations. For inorganic systems, an additional challenge is their potentially non-trivial electronic structure, so metallic and insulating configurations need to be described with similar accuracy. Taken together, using parameter free ab initio methods has therefore become the default approach in inorganic crystal structure prediction, helped by advances in computational power that have made high-throughput first-principles calculations feasible about 20 years ago. However, note that recent advances in machine learning of interatomic potentials has meant that computationally cheap, transferrable potentials can be generated for a wide variety of systems, and we will discuss further below how this is used in crystal structure prediction.

Here, we give a very brief overview of total energy calculations for crystalline systems in conjunction with density functional theory, the standard method for first-principles calculations in extended systems. A wide selection of textbooks is available for more detailed reading in this area, see for example Refs. 47–50.

For any independent particle approach to electronic structure in crystals, Bloch's theorem means that the many-electron problem only needs to be solved for the number of electrons  $N_e$  per unit cell, instead of the whole crystal. For every wave

vector  $\mathbf{k}$ , the single-electron wavefunctions  $\psi_{n\mathbf{k}}$  and their eigenvalues  $\epsilon_{n\mathbf{k}}$  are labelled by a *band index*  $n$ ; in an insulator, the lowest  $N_e/2$  bands are occupied, all others unoccupied. In a macroscopic crystal, electrons occupy states on a dense mesh of  $\mathbf{k}$ -vectors; the mutual orthogonality of states at two wave vectors  $\mathbf{k}_1$  and  $\mathbf{k}_2$  satisfies Pauli's exclusion principle.

Many properties, such as total electronic energy, or electron density, need to be added up or averaged over all occupied electronic states. It turns out that the dense mesh of  $\mathbf{k}$ -vectors (with on the order of Avogadro's number of so-called  $\mathbf{k}$ -points) can be replaced for those purposes by a much coarser regular grid of  $\mathbf{k}$ -vectors, with on the order of 10–1000  $\mathbf{k}$ -points.<sup>51</sup> This, then, makes electronic structure calculations for crystals feasible: a relatively small number (10–1000) of eigenvalue problems of order  $N_e$  needs to be solved to obtain the electronic structure and hence information about total energy, Coulombic forces on the atoms, etc.

In standard density functional theory (DFT), the eigenvalue problems can be solved independently for different  $\mathbf{k}$ -vectors, which is ideal for numerical parallelization. In those flavors of DFT, the exchange-correlation energy is approximated by local or semilocal density functionals.<sup>52,53</sup> The Hamiltonian, as in standard SCF (Hartree-Fock) theory, depends via the density on its own solutions, and is solved self-consistently.

Due to Bloch's theorem it is advantageous to use plane waves as basis functions to expand all electronic states. The lattice-periodic component  $u_{\mathbf{k}}(\mathbf{r})$  of any Bloch state can be expanded in a Fourier series of plane waves with periodic lattice wave vectors,  $u_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) \exp i\mathbf{G}\mathbf{r}$ , where all  $\mathbf{G} = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3$ . The basis set quality can be captured by a single quantity, the cutoff radius  $\|\mathbf{G}_{\max}\|$  of the  $\mathbf{G}$ -vector expansion; most commonly expressed in terms of the "plane wave [kinetic] energy cutoff",  $E_c \equiv \frac{1}{2}\|\mathbf{G}_{\max}\|^2$ . In plane wave representation, solving the DFT eigenvalue equations then becomes a matrix diagonalization problem for the coefficients  $c_{n\mathbf{k}}(\mathbf{G})$ , replicated over the finite set of  $\mathbf{k}$ -vectors.

However, the choice of a plane-wave basis set is not well suited to the core regions of atoms where the spatial variation in the wavefunction is rapid and would require a very high value of  $E_c$  — and hence a prohibitively slow calculation — to give accurate results. However, core electron orbitals are unlikely to depend on the atomic environment and so electrons within a given cut-off radius of the nucleus are pre-calculated such that electronic structure calculations need only consider the remaining valence electrons. This is akin to the "effective core potential" approach, which in plane-wave DFT is extended to the whole periodic table.

The pseudopotential replaces the hard divergent potential close to the core with a softer potential. This is typically done by solving the radial Schrödinger equation as an 'all-electron' problem and generating a pseudo-wavefunction that satisfies a number of constraints.<sup>54,55</sup> First principles structure prediction relies on robust pseudopotentials that deliver accurate results with relatively low values of  $E_c$ .

Note that including the Fock operator explicitly in the exchange-correlation energy expression severely increases computational cost of plane wave calculations (unlike those based on localized basis sets). Hartree-Fock or hybrid functional DFT calculations for crystal structure prediction are therefore prohibitively expensive, as are any other wave-function based approaches for CSP. This can be a challenge in systems where so-called exact exchange is crucial to obtain the correct electronic structure.

Once the electronic structure is obtained, forces on all atoms and stresses on the unit cell are obtained using the Hellmann-Feynman approach. Those forces and stresses are then minimized using standard iterative techniques, until residuals are below acceptable thresholds; at which stage the final crystal structure (lattice and atomic positions) together with its energy and, optionally, other observables, are available for further processing. An efficient implementation of the geometry optimization is key for CSP, and the DFT packages typically chosen for CSP are those that have fast yet robust algorithms to minimize forces and stresses.

### 3.12.4 Modern approaches

A variety of different algorithms have been implemented in modern CSP approaches, and are described in detail below. They universally employ local geometry optimizations as described above. This distinguishes them from one of the most intuitive early structure searching approaches, simulated annealing.<sup>56</sup> In that method, Monte Carlo or molecular dynamics steps are applied to modify an initial structure, with an initially very high temperature that is continuously reduced during the simulation, leading ideally to the global minimum. No local geometry optimizations are involved during the individual simulation steps, which arguably restricts the method's efficiency.

#### 3.12.4.1 Random structure searching

As we will discuss, all approaches in *ab initio* structure searching necessarily utilize some level of randomness. The AIRSS code (Ab initio Random Structure Searching)<sup>57</sup> uses random structure search exclusively. While this approach may seem simplistic, it is remarkably effective thanks to the generation of *sensible* random structures, an emphasis on trivial parallelism and effective post-processing tools.

The speed of *ab initio* crystal structure searches depends on how many geometry optimization steps are required. This can be reduced by generating 'sensible' starting structures. This means the geometry optimization spends less time in unphysical configurations, far from the local minima. AIRSS incorporates a powerful set of keywords to generate sensible random structures by invoking a useful program, *buildcell*.

Important structure features such as the target volume and minimum atomic separations can be chosen manually or automatically based on initial searches. Symmetry can be imposed by specifying a range of symmetry operations or a specific space group.

For molecular crystals, buildcell can place molecules in random positions in the cell while applying symmetry operations. Constraints can be placed on the unit cell, and on positions of sets of atoms. This is useful when only partial information about the crystal structure is available.

A technique used in AIRSS is ‘relax and shake’ (RASH), which generates new structures by shaking structures around a local minimum. This can result in new structures with lower symmetry, or similar structures in distorted super-cells. This process can be performed recursively as a minimal input ‘learning’ approach.

Unlike more sophisticated approaches, random structure searching does not require any communication between different computations. This means calculations can be trivially parallelized so that thousands of structures can be generated and calculated simultaneously. For the same reason, AIRSS does not require special treatment for variable stoichiometry searching. The user can specify an arbitrarily complex composition space over which AIRSS can sample.

AIRSS uses the SHELX<sup>58</sup> .res file format for crystal structures, supplemented with metadata. These files can be read by visualization tools such as Vesta<sup>59</sup> and directly by AIRSS’ analysis tool, cryan, which can rank structures, build convex hulls and provide geometrical analyses.

### 3.12.4.2 Particle swarm optimization

A general approach to go beyond purely random structure generation is to use information from already completed geometry optimizations to map the configurational space and build new structures. Particle swarm optimization (PSO) is a global minimization technique used in a wide range of applications with non-convex minimization problems, such as modelling the electrical power market.<sup>60</sup>

The technique allows individual agents to move around configurational space (where particle  $i$  has coordinate  $x_{ij}^t$  at timestep  $t$  with velocity  $v_{ij}^t$ ). The positions and velocities are initialized randomly and allowed to relax to a global minimum with position  $y_{ij}^t$ . The new velocities are calculated from a weighted sum of;

- the difference between the initial and relaxed geometry,
- the difference between the starting configuration,
- the best known minimum energy configuration,  $z_{ij}^t$ ,
- and the velocity at the previous timestep.

The update equation used by the CALYPSO code<sup>61</sup> is given by,

$$v_{ij}^{t+1} = A_1 (y_{ij}^t - x_{ij}^t) + A_2 (z_{ij}^t - x_{ij}^t) + \omega v_{ij}^t. \quad (1)$$

Where  $A_1$  and  $A_2$  are randomly generated numbers between 0 and some value  $A_1^{\max}$  and  $A_2^{\max}$ . The ‘memory’ factor  $\omega$  decreases linearly between generations. A schematic of the algorithm and its traversal of the potential energy surface is shown in Fig. 2.

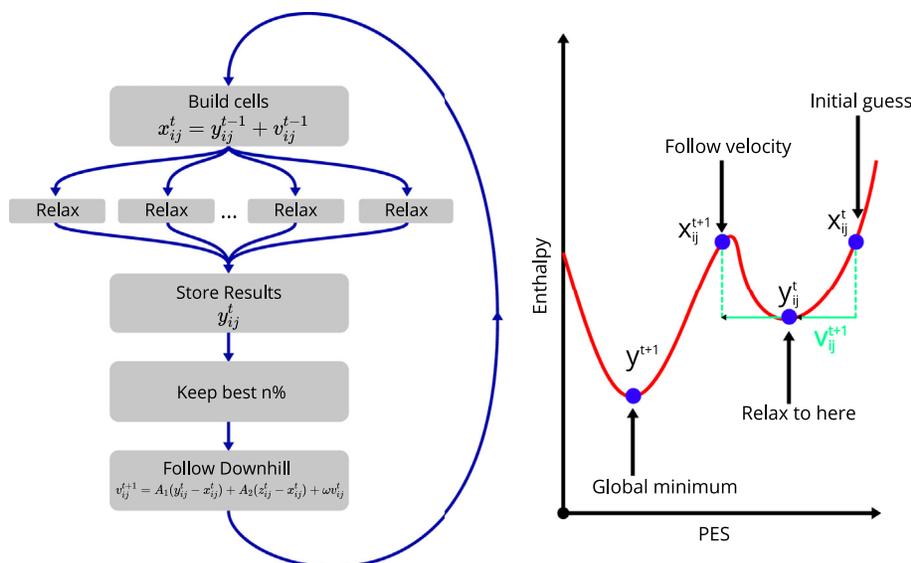
Several extra constraints are required to avoid stagnating in local minima: a penalty function is applied such that high energy structures are not used in the PSO, further exploiting the Bell-Evans-Polanyi principle; and each generation includes a certain (user-specified) percentage of new randomly generated structures to retain the population diversity.

### 3.12.4.3 Genetic and evolutionary algorithms

An alternative approach to exploit information during a structure search is to use genetic algorithms (GA’s) or evolutionary algorithms (EA’s). These borrow ideas from biological evolution, of mutation, breeding and natural selection, to build upon promising intermediate structural candidates. Ideally, that allows the search to focus on the relevant sections of the PES, generate more competitive guesses faster, and find the global minimum with less attempts. Their use for structure prediction was pioneered in the early 1990s for clusters and nanoparticles<sup>62–64</sup> and shortly thereafter for inorganic solids.<sup>65</sup>

GA’s encode all properties of a particular structure in a one-dimensional alphanumeric or binary string, the “genome”. All modifications (described below) are done by direct manipulation of “genes”. The projection from structure to genome (and vice versa) removes GA’s from the structure prediction realm, into the wider area where GA’s are applied. The advantage is that a large set of GA operators developed in different research fields can be applied straightforwardly; the disadvantage, that any links to crystalline configurational space, and to physico-chemical intuition as relates to structure formation, are severed. EA’s on the other hand describe algorithms where structural information (lattice vectors and atomic positions) is stored and manipulated directly, typically with the use of evolutionary operators that are designed for structural modifications. The majority of software packages in the field (see Table 1) are of the EA type.

Evolution of the Darwinian type implies that changes to the genome only occur in the procreation process, combining parental features with a small amount of random changes (mutations). For structure prediction this implies generating a new structure and determining the enthalpy for its given (fixed) geometry. This often has a very low success rate. However, allowing local optimization of the structure (i.e., relaxing to the bottom of its local basin in the PES) results in a much more relevant structural candidate for further consideration. This local optimization corresponds to changes to the genome during the “lifetime” of the structure, and allowing those changes to be passed on to later generations follows the Lamarckian view of evolution, that traits acquired during



**Fig. 2** Flow chart (left) outlining the particle swarm optimization technique (equation is specific to CALYPSO) to find the global minimum of a potential energy surface (right).

**Table 1** A selection of CSP packages for inorganic systems available for academic use (possibly following registration), and the methodology they are based on. Equivalent packages that specialize on organic CSP include DMACRYS, GAtor, GRACE, MOLPACK, and UPACK.<sup>66–70</sup>

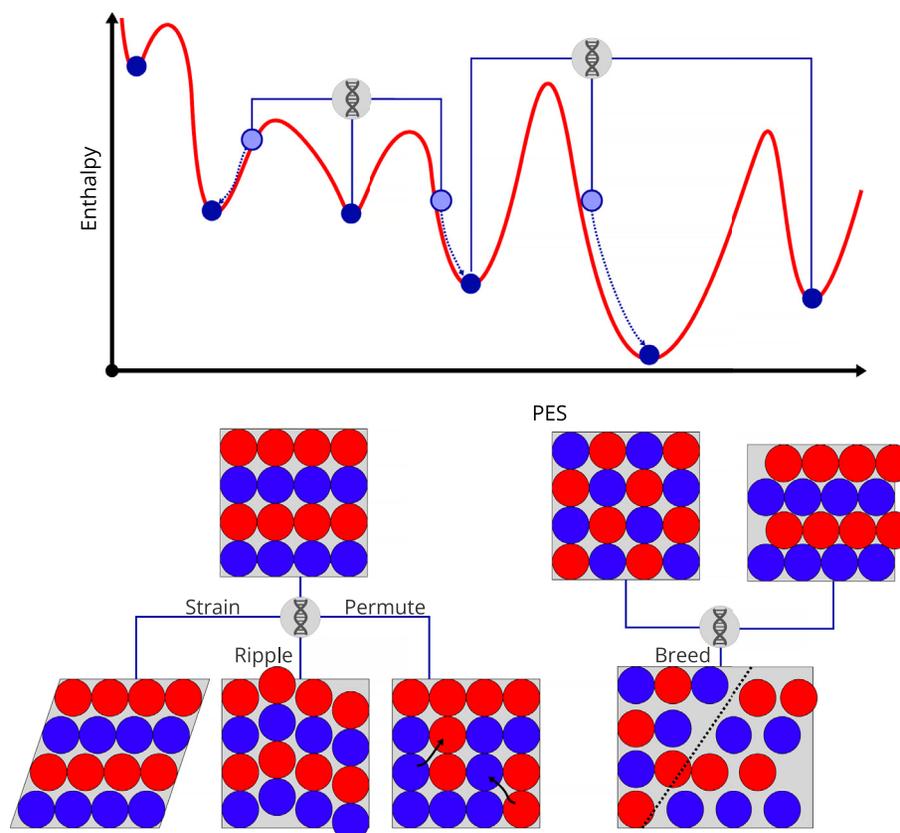
Package	Method	References
AIRSS	Random structure searching	17,57
CALYPSO	Particle Swarm Optimization	61
CRYSKY	Bayesian Optimization	71
GASP	Genetic Algorithm	72,73
MAISE	Evolutionary Algorithm	74
USPEX	Evolutionary Algorithm	11
XTALOPT	Evolutionary Algorithm	75

an organism's lifetime can be passed on to offspring. Such is the success of this approach, that local structural optimization has been identified as a key ingredient of GA or EA structure prediction.<sup>11</sup>

The typical workflow of a GA/EA structure search is shown in Fig. 3. Often a "generation" of structures is optimized, followed by a selection process of the best candidates, which influence the next generation of structures. The latter features various genetic or evolutionary operators. In another (incidental) nod to biology, EA implementations show great diversity in choice and weighting of particular evolutionary operators, but the following operations are commonly used:

- Mutation: manipulating lattice vectors and/or atomic positions of a single structure. Can include lattice strains, atom exchanges between sites, individual or collective atomic displacements.
- Breeding: combining two "parent" structures into one "offspring". Most often a real-space 'cut and splice' operation.
- Random generation: adding completely random new structures to the dataset to retain diversity of the "gene pool" and avoid premature focus on particular parts of the PES.

Structural candidates, whether created at random or via specific operations, might occupy the same basin within the PES, and are then (within numerical noise) identical following the geometry optimizations. This is not an issue for random structure searching – re-discovering the same structure several times in fact contains information about relative basin sizes – but is a challenge for algorithms with built-in memory: GA/EA and also PSO implementations need to be able to identify identical structures to remove duplicates from the pool of structures considered for the next generation. Given that crystal lattices and atomic positions are not uniquely defined, this is a non-trivial task. Duplicates are usually identified by having very similar enthalpies, and by either matching fingerprints based on local atomic connectivity,<sup>76</sup> pairwise separations of atoms<sup>77</sup> or successful mapping of suitably normalized unit cells.<sup>75</sup> Lyakhov et al.<sup>78</sup> show that fingerprints, and the 'distance' between these fingerprints, can be used to perform more effective EA structure prediction. To maintain diversity, structures with 'nearby' and hence similar fingerprints are identified and only the



**Fig. 3** Top: sketch outlining the evolutionary algorithms technique to find competitive minima of a potential energy surface. Bottom: visualizations of typical EA operations: mutations that involve lattice strain, collective displacements, or atom swapping; and hereditary ‘cut and splice’ operation.

lower energy structure is carried forward. They also note that structures which are very dissimilar should not ‘breed’ as the ‘‘offspring’’ structures will often be high energy.

#### 3.12.4.4 Database informed methods

Yet another approach is closer to the aforementioned ‘appeal to conventional wisdom’ approach. With high-throughput calculations easily accessible, it is feasible to build all candidate structures using many geometrically valid structures or structures from an extensive database (typically 1000s of structures) of known geometries.

The ICSD has over 200,000 experimental inorganic structures. The majority of these fall into a given structure type; a set of structures with identical lattice shapes and (relative) atomic positions.<sup>79</sup> For example, the cubic ZnS wurtzite structure type (labelled B4 in Strukturbericht notation) is shared by AgI, BeO, ZnO, GaN and many others. This suggests that a lot can be learned from these known structure types with different stoichiometries. As more new structures are discovered, new structure types may also emerge.

Before global optimization-based CSP methods, some effort had been applied to combine *ab initio* calculations with high-throughput approaches, building new structures from known structure types<sup>80</sup> and generating databases of electronic structure information. Since then, tools have been developed to assist these types of calculations such as high-throughput pseudopotentials<sup>81</sup> and band structures.<sup>82</sup> This technique has been applied to otherwise compositionally complex systems such as ternary and quaternary oxide compounds.<sup>83,84</sup> This method is similar to, but distinct from, database screening methods which aim to *screen* for material properties based on data from a database, without input from *ab initio* calculations.<sup>85,86</sup>

A related approach has an emphasis on local connectivity, or bond network topology, achieved by representing crystal structures as graphs. A graph is a set of nodes connected by edges. Just as a crystal, a graph can be infinite. Placing nodes in space (such as a plane or 3-dimensional space) is called embedding. A crystal can be considered a three-dimensional embedding of a graph where nodes are atoms and edges are bonds (or connections between nearest neighbors). The infinite graph of a crystal will have the same 3-periodicity. This can be simplified to a finite labelled graph by a process called the vector method.<sup>87</sup> In short, this involves drawing  $N$  nodes for the  $N$  atoms in the unit cell. An edge can be drawn between the nodes for bonded atoms where both are within the unit cell. Where atom  $i$  is inside the unit cell and is bonded to an atom  $j$  in the neighboring unit cell, an edge can be drawn from atom  $i$  to atom  $j$ 's image in the unit cell and labelled with the lattice vector describing the shift in unit cell.

Some examples of this method include considering all possible three-fold coordinated structures to find  $sp^2$  carbon polymorphs (restricted by the number of atoms in the unit cell),<sup>88,89</sup> or tetrahedral networks to find  $sp^3$  carbon polymorphs.<sup>90,91</sup> Further  $sp^3$  carbon polymorphs have been generated by using databases of zeolite structures.<sup>92</sup> A similar approach has been used to great effect to reproduce a priori all known phases of ice.<sup>93</sup> Structures generated using these methods may have much larger primitive unit cells and require more computational resources per structure, however each optimization starts from a chemically sensible configuration.

Graph theoretical methods have also been invoked in the study of aluminosilicates<sup>94,95</sup> where particular care can be taken to build structures obeying the Löwenstein avoidance rule,<sup>94,96</sup> which suggests aluminosilicates will be unlikely to have Al–O–Al bonds, an assertion that has kinetic<sup>97</sup> and thermodynamic arguments.<sup>98</sup>

Topological methods can be extremely efficient, if the underlying assumptions about relevant bond network topologies hold true. Conversely, they quickly fail if those assumptions break down. For instance, studying silica polymorphs based on connected  $SiO_4$  tetrahedra will fail to identify the high-pressure stishovite phase, which is based on  $SiO_6$  octahedra.<sup>99</sup>

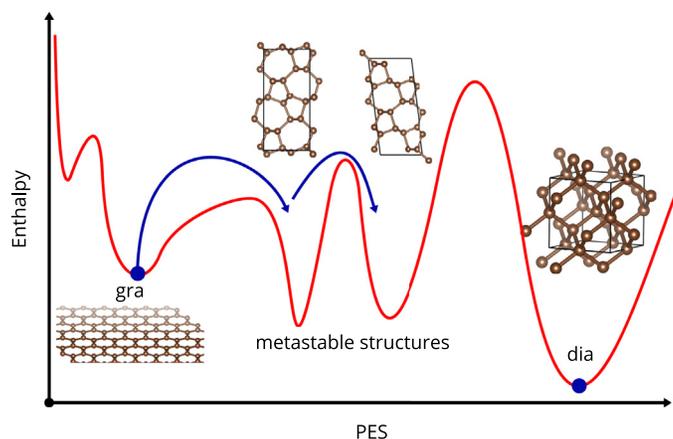
### 3.12.4.5 Local approaches

The prediction methods discussed so far are “global” approaches in the sense that they survey, at least initially, the full crystalline configuration space. Approaches that retain memories during the searches, such as the PSO and GA/EA algorithms, should eventually focus on the most relevant sections of this space. However, the location of this section can not be directed by the user, nor do the results give any indications about the local connectivity of the PES, i.e. which local minima are connected by transition states and the size of transition barriers. Further below we shall introduce recent approaches to extract such connectivity information from global structure prediction surveys, but in this section want to briefly discuss “local” approaches to predict crystal structures. By this we mean approaches that explore the local topology and connectivity of the PES, usually starting from a single user-defined configuration, and not always concerned with identifying the global minimum. There are a multitude of scenarios where such information is relevant. Fig. 4 depicts an idealized potential energy surface of carbon, which is dominated by graphite and diamond but features a myriad of other local minima that can be seen as intermediates in the sense that they combine  $sp^2$ - and  $sp^3$ -hybridized carbon (the x-axis in the figure should not be taken too literally; there are multitudes of purely  $sp^2$ - and  $sp^3$ -bonded carbon polymorphs other than graphite and diamond<sup>100</sup>). At ambient conditions graphite has the lowest free energy of all carbon polymorphs<sup>101</sup> but diamond is stabilized at moderate pressures. However, the cold (i.e., non-heated) compression of graphite does not result in diamond<sup>102–107</sup> because the potential barrier toward its formation can not be overcome, at least not on the experiments’ time scale. Instead, another metastable carbon allotrope forms that is configurationally “closer” to graphite but features  $sp^3$ -bonded carbon. Various crystal structure prediction approaches have yielded a variety of candidate allotropes very close in enthalpy, such as M-, bct- $C_4$ -, W-, and Z-carbon.<sup>29,108–112</sup> Of those, M-carbon seems the most likely candidate due to the lowest kinetic barriers to transformation from graphite.<sup>113</sup> Hence, knowledge of the local topology of the crystalline PES can be invaluable to interpret experimental findings or explore whether new materials are accessible from specific precursors.

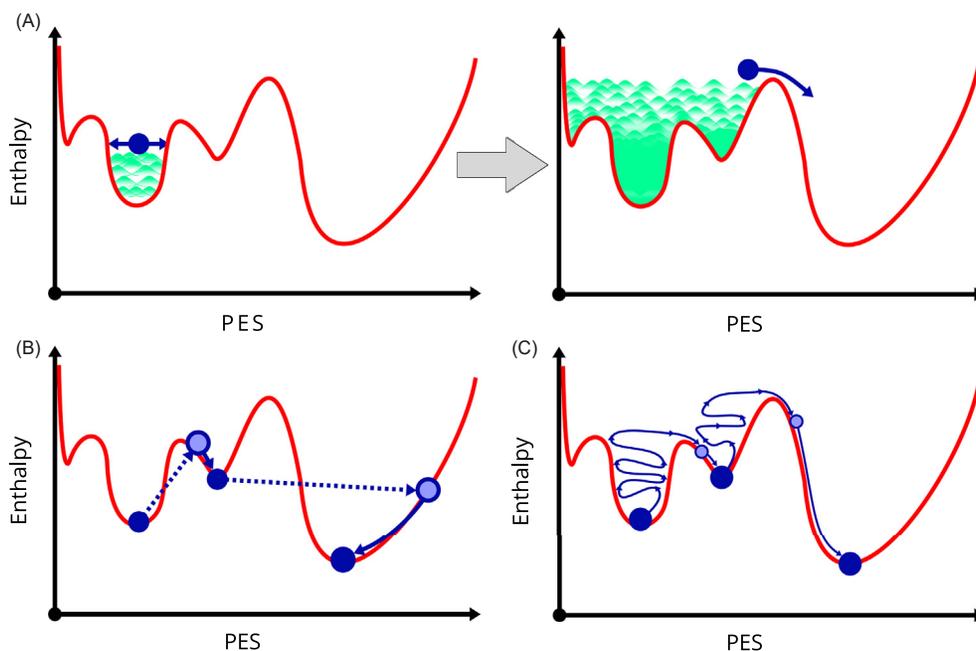
Most of the approaches described below are not exclusive or tailored for crystal structure prediction, they instead represent quite general methods for sampling free energy surfaces, often aiming to accelerate rare events such as phase transitions where substantial energy barriers need to be overcome.

#### 3.12.4.5.1 Metadynamics

In metadynamics,<sup>114–116</sup> a molecular dynamics simulation is performed, beginning from a local minimum on the PES. During the simulation, a bias potential is constructed that penalizes previously visited configurations, thus driving the system away from the initial and toward new metastable or stable configurations, see Fig. 5. One key ingredient is the concept of “collective variables”



**Fig. 4** Schematic of carbon’s potential energy surface at some moderate pressure where diamond (“dia”) has a lower enthalpy than graphite (“gra”) but other metastable structures are more readily accessed.



**Fig. 5** Local approaches' explorations of potential energy surfaces (red) with structural candidates (blue). (A) Metadynamics simulation early on (left) and at a later stage (right). Green indicates the bias potential. (B) Basin hopping. (C) Minima hopping.

(CV's)  $\{\mathbf{s}(\mathbf{R}, \mathbf{a})\}$ , a set of coarse-grained order parameters much smaller than the full set of atomic positions  $\mathbf{R}$  and lattice vectors  $\mathbf{a}$ , that can distinguish the different metastable states. The bias potential  $V(\mathbf{s})$  is built successively as a series of Gaussians  $G(\mathbf{s}, \mathbf{s}') = W \exp -(\mathbf{s} - \mathbf{s}')^2$  in  $\mathbf{s}$ -space that are centered around the CV's present at regular intervals during the simulation,  $\mathbf{s}_i$

$$V_n(\mathbf{s}) = V_{n-1}(\mathbf{s}) + G\left(\mathbf{s}, \mathbf{s}_n\right) \exp\left[-\frac{1}{\gamma-1} \beta V_{n-1}(\mathbf{s}_n)\right].$$

The exponential scaling term, introduced as well-tempered metadynamics,<sup>115</sup> ensures that the changes to the bias potential decay over time, and features the inverse temperature  $\beta = 1/k_B T$  and a positive constant  $\gamma > 1$ . Those, together with the width of the Gaussians  $W$  are the main computational variables in a metadynamics simulation. The  $x$ -axis of the metadynamics schematics in Fig. 5A should therefore be considered as one of these CV's.

Ongoing efforts in developing CV-based approaches are often centered around building better iterative bias potentials, to further accelerate rates of transitions between metastable minima.<sup>117-121</sup> However, the main physical choice is around appropriate CV's themselves. For proteins, these might be certain torsional angles that distinguish relevant foldings. For crystal structures (and many other systems) there are no obvious choices. Moreover, in exploratory searches, where the target structure is not known, it is nigh on impossible to guess the most efficient CV's that link initial and target structure. However, noting that space group symmetry and more generally the crystal lattice often sufficiently distinguish different polymorphs within a given system, metadynamics simulations in inorganic solids most often use the lattice vectors as CV's,  $\mathbf{s} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ .<sup>122,123</sup> Strictly speaking, the six independent lattice lengths and angles are used.

### 3.12.4.5.2 Minima hopping

Minima hopping<sup>124,125</sup> is another approach to survey metastable minima around an initial configuration that are accessible via increasingly larger energy barriers. It also uses molecular dynamics to explore the PES, but uses more explicit bookkeeping to retain memory of the explorations. After a relatively short MD run a full geometry optimization is performed, see Fig. 5. If the system has reached a new configuration, it will be added to the list of known configurations if it is less than a threshold  $E_{\text{diff}}$  higher in energy than the previous configuration. The threshold  $E_{\text{diff}}$  and the MD temperature  $T$  are then decreased. On the other hand, if the energy of a newly found configuration is more than  $E_{\text{diff}}$  above the previous configuration,  $E_{\text{diff}}$  will be increased; and if the optimization ended in a previously visited configuration,  $T$  will be raised. With those adjustments, repeated visits to known configurations are penalized, the exploration will always escape from local trappings, while locating new relevant minima slows the exploration down.

### 3.12.4.5.3 Basin hopping

The basin hopping approach was first introduced for clusters and is still mostly applied to localized systems<sup>126</sup>; it is similar to a prior "Monte Carlo minimization method" suggested to explore protein folding.<sup>127</sup> In this method, changes to the current configuration are made discontinuously, using random update steps such as moving or swapping atoms. Importantly, those updates are followed by full geometry optimizations, and the optimized configuration is accepted or rejected based on a Metropolis criterion, see Fig. 5.

By attempting to sample the space of local energy minima rather than the full PES, a high-dimensional problem is vastly reduced. But not only is the information from each PES basin reduced to its minimum energy value, this approach also loses information about barriers between adjacent basins. By scaling the overall update step size the method can be tuned to explore configurationally close to or farther away from the initial structure; if it is too large, the method essentially becomes random structure searching.

#### 3.12.4.5.4 Following soft modes

A method specific for crystal structure prediction is to “follow” soft phonons. Phonons are the elementary quasiparticles of lattice vibrations and span the space of all possible atomic displacements within a crystalline system. They are harmonic and monochromatic, characterized by a wave vector  $\mathbf{q}$ , frequency  $\omega_i(\mathbf{q})$ , and displacement vector  $\mathbf{e}_i(\mathbf{q})$  (index  $i$  labels the phonon “branch”, of which there are  $3N - 3$  for  $N$  atoms in the unit cell). Phonon frequencies and displacement vectors are the eigenvalues and -vectors of the dynamical matrix  $D(\mathbf{q})$ , which itself is the Fourier transform of the real-space force constant matrix. The appearance of non-zero  $\mathbf{q}$  is due to translational symmetry in the crystal; in molecular systems, the equivalent of the dynamical matrix is the Hessian. Solving for phonon frequencies is equivalent to diagonalizing the Hessian many times over, at each  $\mathbf{q}$ -vector of interest.

More precisely, the eigenvalues of  $D(\mathbf{q})$  are the squared phonon frequencies  $\omega_i^2(\mathbf{q})$ , proportional to the curvature of the PES along the displacement vector  $\mathbf{e}_i(\mathbf{q})$ . A dynamical instability is characterized by a negative curvature, and therefore  $\omega_i^2(\mathbf{q}) < 0$ . This is referred to as an “imaginary phonon mode”, though it isn’t actually a phonon (there is no lattice vibration without a restoring force), rather a saddle point on the PES. When plotting the phonon dispersion relations  $\omega_i(\mathbf{q})$  along special directions of  $\mathbf{q}$ , imaginary modes are commonly drawn as negative frequencies, as sketched in Fig. 6. The corresponding atomic displacement vector  $\mathbf{e}_i(\mathbf{q})$  points toward a lower-energy structure on the PES, and this can be exploited to find low-lying minima in configurational space. This method is similar in principle to the RASH technique used in AIRSS.

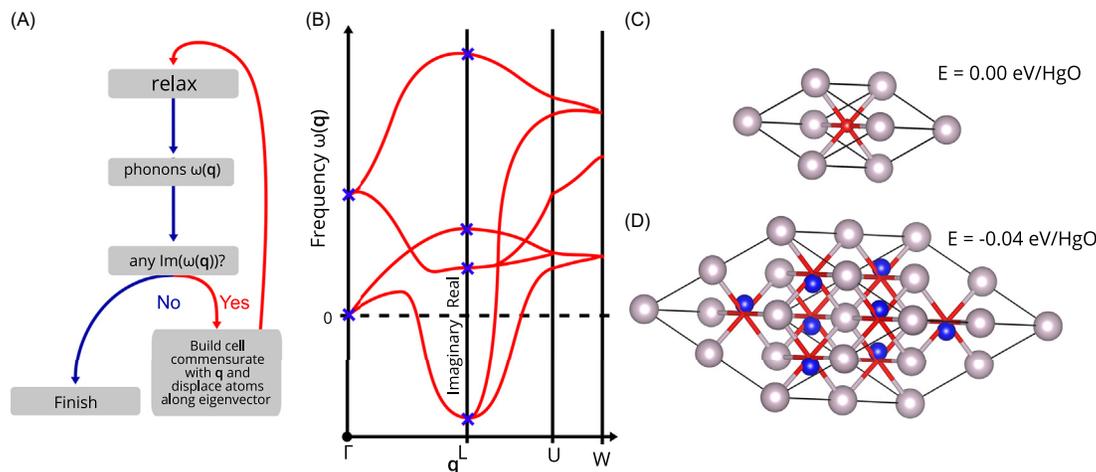
The entire procedure is sketched in Fig. 6. Starting from simple, high-symmetry initial structures for the system of interest (e.g., known structure types), geometries are relaxed and phonon dispersions calculated. If clear imaginary modes appear, the displacement vector is added to the atomic positions. Note that for non-zero  $\mathbf{q}$  vector, the displacement always has a wave length that spans multiple crystalline unit cells; it needs to be applied to the atomic positions in a suitable supercell. That can make the resulting structures very large, and it is usually advisable to follow imaginary modes at  $\mathbf{q} = 0$  (the  $\Gamma$ -point in the Brillouin zone), or otherwise modes at special points at the Brillouin zone boundary. Once a suitably scaled displacement is added, structures are relaxed, and the process repeated until dynamically stable structures have been generated. The absence of any imaginary modes means a minimum on the PES has been reached.

This approach has been used to predict new perovskite compounds<sup>129–131</sup> (starting from the ideal cubic perovskite structure), and new transition metal monoxides<sup>128,132</sup> (starting from the rocksalt structure). There are no limits to its applicability, but it will only discover structures that are connected to the initial guess via a series of saddle points on the PES.

A somewhat flipped approach that acknowledges the spirit of dynamical instabilities has been used in random structure searching to explicitly discover high-symmetry structures with one or two imaginary modes.<sup>133</sup> Just like many perovskites have a high-temperature phase with the cubic structure,<sup>134</sup> those saddle point structures are candidates for high-temperature phases of materials, where thermal excitations result in the averaged high-symmetry structure.

### 3.12.5 Fitness functions

So far, discussions have implicitly focused on the target to minimize the total *energy* (or *enthalpy*) during the structure prediction process. This is reasonable, as the global minimum on the potential energy surface for a given chemical composition is surely



**Fig. 6** (A) Illustration of following soft phonon modes. (B) Phonon dispersion curve of rocksalt HgO, as per reference 128. Points indicated by crosses are explicitly calculated phonon frequencies, red lines are Fourier interpolations. (C) Primitive unit cell of HgO in a rock-salt structure. (D) Imaginary phonon mode with wave vector  $\mathbf{q} = (0.5, 0.5, 0.5)$  of rock-salt HgO described in a  $2 \times 2 \times 2$  supercell of the primitive cell. Blue atoms are moved along the displacement vector. The perturbed structure is lower in energy.

the most likely to be synthesized. However, local search approaches have already been motivated by the need to identify low-barrier transitions from given initial structures to competitive metastable alternatives, rather than the global minimum. Furthermore, it is often desirable to identify and rank crystal structures by specific physical or electronic *properties*. Finally, compounds are often compared across composition ranges, so the total energy needs to be evaluated against all possible ‘escape’ or decomposition reactions. All these considerations have led to the development of more generalized ‘fitness functions’  $F$  to attach to a given crystal structure  $S$ ,

$$F(S) = (1 - \alpha)E(S) + \alpha \sum_i \beta_i P_i(S),$$

where  $P_i$  are appropriately normalized properties other than the compound’s total energy/enthalpy  $E$ , with weights  $\sum_i \beta_i = 1$  chosen to reflect the user’s priorities. The properties  $P_i$  can reflect a desirable absolute maximization or minimization of a property (e.g. hardness or mass density) or closeness to a target value (e.g. optical band gap or lattice constant). Once  $F$  is chosen, structural candidates can be ranked according to this more general criterion, which influences (in search approaches with memory) choice of future structural candidates accordingly.

To be fully utilizable, the properties  $P_i$  must be calculable with little effort; any property calculations that are computationally more demanding than a structure’s geometry optimization are unsuitable for unsupervised structure prediction. Some properties are trivially derived from a crystal structure, such as density or atomic coordination numbers; for others, for example Vickers hardness, estimators exist based on atomic configurations; yet others, such as electronic band gap or electronegativity, come at the cost of a single SCF calculation – which is a defensible cost if the structure prediction is done at DFT level of theory.

For structure-property relationships that are hard or impossible to calculate, correlations can be explored using machine-learning techniques.<sup>135</sup> Originally used to mine materials database information, machine-learned models can also be used in crystal structure prediction, provided the structural candidates are not too far removed from the model’s training dataset.

Pareto analysis is an economics-inspired alternative to specifying a fixed multi-component fitness function. It produces solutions that as best as possible optimize all desired properties ( $E, P_i$ ) simultaneously, ranking as best structures that are optimal in at least one property, and where every alternative includes a trade-off of properties.<sup>136–138</sup>

The Pareto analysis is an example of how desirable properties can be included in a structure prediction without designing a specific fitness function. A possible workflow could be to determine a full set of structural candidates, e.g. by random structure searching, then determine all properties of interest either for the full set or a sub-set of low-enthalpy candidates, depending on computational cost. The latter, in particular, is a very popular approach to find materials with specific properties among low-enthalpy (and therefore presumably synthesizable) structural candidates.

Below, we discuss some of the property estimators that can be employed directly as part of a fitness function. Before that, we introduce the Maxwell stability criteria for multi-component systems, also known as convex hull or tieline constructions.

### 3.12.5.1 Compositional stability

Where the system of interest contains a range of possible compositions or stoichiometries, the thermodynamic stability of each structure against a decomposition into a mixture of other stoichiometries should be calculated. A Maxwell, or convex hull, or tieline, construction can determine this.

If  $H(F)$  is the enthalpy of a structure with composition  $F$  comprising a set of species,  $\{A^i\} = \{A, B, \dots, X\}$ , the relative enthalpy of an  $n$ -species mixture containing elements A, B to X compared against decomposition is

$$\Delta H(\mathbf{x}) = H(A_{x_1} B_{x_2} \dots X_{x_n}^n) - \sum_{i=1}^n H(A^i) x_i. \quad (2)$$

An  $n$ -species composition space can be described by normalized barycentric coordinates in  $n - 1$  dimensions by imposing the constraint

$$\sum_{i=1}^n x_i = 1, \quad (3)$$

where  $x_i$  are the composition weights. The mechanism for calculating stability against decomposition to any composition is to construct the  $n$ -dimensional convex hull of points in a composition-enthalpy space. Points that form the (lower portion of the) convex hull are stable against decomposition.

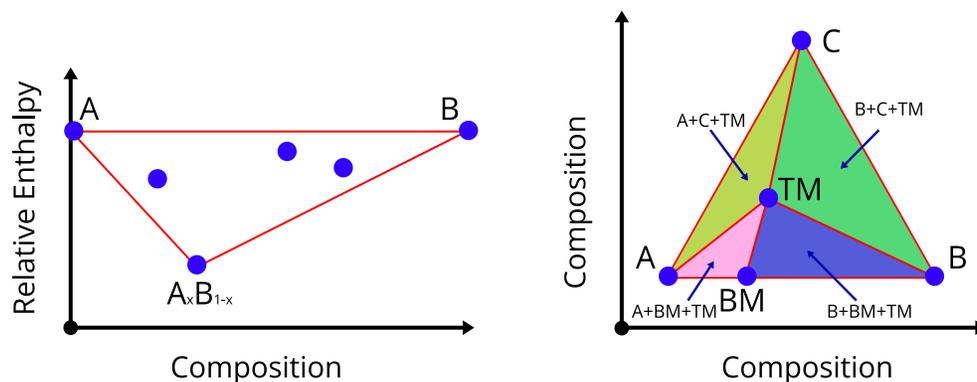
In a binary system this looks like

$$\Delta H(x) = H(A_x B_{1-x}) - H(A)x - H(B)(1-x). \quad (4)$$

The composition can be described by one parameter,  $x$ , and the convex hull is a polygon where each corner represents a stable composition, as in Fig. 7A).

In a ternary system this looks like

$$\Delta H(\mathbf{x}) = H(A_{x_1} B_{x_2} C_{1-x_1-x_2}) - H(A)x_1 - H(B)x_2 - H(C)(1-x_1-x_2) \quad (5)$$



**Fig. 7** Examples of binary and ternary convex hulls with A, B and C end members, binary material (BM =  $A_x B_{1-x}$ ) and ternary material (TM =  $A_{x_1} B_{x_2} C_{1-x_1-x_2}$ ). In the binary case, the data is represented in two dimensions; one for relative enthalpy and one for composition. In the ternary case, two dimensions are used to represent composition and the convex hull is projected onto the surface, dividing composition space into different (triangular) regions labeled by the relevant decomposition reactions.

where the pure A, B, C end members are corners of an equilateral triangle. Explicitly describing the ternary convex hull in Fig. 7B, the composition space is divided into irregular triangles, each of which represents a linear mixture of the surrounding stable structures. Points on the ridges connecting points represent binary mixtures of the two endpoints.

In a quaternary system, with a fourth species, D, this looks like

$$\begin{aligned} \Delta H(\mathbf{x}) = & H(A_{x_1} B_{x_2} C_{x_3} D_{1-x_1-x_2-x_3}) \\ & - H(A)x_1 - H(B)x_2 - H(C)x_3 - H(D)(1-x_1-x_2-x_3). \end{aligned} \quad (6)$$

where the end members are corners of a regular tetrahedron. In this case, the composition space is represented in three dimensions. The composition space is divided into irregular tetrahedra which contain compositions that form linear mixtures of the four surrounding compounds. Points on the faces of the tetrahedra represent ternary mixtures of the face corners and points on the edges of the tetrahedra represent binary mixtures of the two endpoints.

### 3.12.5.2 Structural and mechanical properties

Often, partial structural information about an unknown compound exists. In particular, partial powder x-ray diffraction (XRD) data might have been collected, but in insufficient detail or quality to fully solve the compound's structure. In such a case, the fitness function could be augmented by a term  $P \sim \|1 - d_{\text{fig}}\|$  that contains the difference  $d_{\text{fig}}$  between the measured and simulated XRD patterns for predicted structures.<sup>139-141</sup> This approach resembles a new type of XRD refinement procedure, based on first-principles structure optimizations. In practice, the use case for this approach has so far been limited. More often, the partial information from XRD (such as unit cell size or space group) is used to constrain conventional structure searches to the relevant regions of configurational space<sup>142,143</sup>; the interactions between experimental and computational collaborators tend to beat explicit fitness functions in those cases.

For a given structure, the response to external stresses via collective distortions away from equilibrium determines its mechanical properties. The mechanical properties of materials are extremely important for potential technological applications. Superhard and incompressible materials, for example, are sought after in various industrial and aerospace settings. While the bulk modulus, the inverse of compressibility, determines a material's resistance against compression, its shear modulus is a better descriptor for its hardness, the resistance against indentation.<sup>144,145</sup> Conventionally, elastic moduli of crystals are determined from a series of finite strain calculations, but this is too computationally demanding for high-throughput calculations. The bulk modulus can be estimated via an empirical relation to local bond lengths.<sup>146,147</sup> For the hardness, several different microscopic models have been developed<sup>148-151</sup> that (intrinsically or with small modifications) rely only on the atomic configurations and no additional electronic structure information. These computationally cheap empirical models have been implemented in several crystal structure prediction packages, and allow to include hardness in the fitness function.<sup>78,152</sup>

As an alternative,<sup>153</sup> the shear modulus of the predicted structures has been estimated by a machine-learned model based on the AFLOW materials database,<sup>85</sup> from which the hardness itself can be estimated.<sup>145</sup> We expect that machine-learned structure-property relationships will become increasingly more important to quickly assess newly predicted structures with respect to mechanical or other properties.

### 3.12.5.3 Electronic properties

For many applications of inorganic materials, their electronic properties are at least as important as mechanical or elastic properties – are they insulators, semiconductors, or conductors? Is their fundamental band gap (the solid-state equivalent of the

HOMO-LUMO gap) suitable for matter-light interaction, e.g. sensing or photovoltaics? Are their transport properties amenable to thermoelectric applications? Are they promising superconductors? Those properties are accessible in electronic structure calculations, but usually at (comparatively) large computational cost. For high-throughput scenarios such as crystal structure prediction, fast approximations are required.

The **optical band gap** of materials determines their ability to absorb light of given wave lengths. For solar cell materials, the optimal gap value is around 1.4 eV.<sup>154</sup> Hence, a fitness function that uses the property  $P \sim (E_g - 1.4 \text{ eV})^2$  would be straightforward to implement in a search process, where  $E_g$  is the band gap obtained from DFT. There are several issues that need to be considered in such an approach. Firstly, the optical gap is different from the single-particle energy gap; it is the smallest direct band gap including excitonic states, which are bound electron-hole pairs.<sup>155,156</sup> Calculations of exciton binding energies are feasible but computationally very demanding.<sup>157</sup> Secondly, all DFT implementations systematically underestimate even the fundamental band gap.<sup>158</sup> Exchange-correlation functionals can be chosen to closely mimic experimental data<sup>159</sup> but again can be computationally expensive (usually if they include exact exchange) and their appropriateness for new structure types discovered in a search is not always clear. Finally, for photovoltaic applications the band gap needs to be direct, and that assessment requires full band structure calculations instead of relying on the electronic density of states (DOS), which would be readily available within the structure search.

For metals, the potential of **superconductivity** is of interest. In fact, the prediction of conventional high-pressure polyhydride superconductors is one of the most impressive examples of the successes of crystal structure prediction. The superconducting transition temperature  $T_c$  can be calculated via the electron-phonon coupling strength  $\lambda$ , which is again computationally expensive.<sup>160,161</sup> Hence, the preferred *modus operandi* has been to sample a given system, and determine  $T_c$  for the most stable or metastable structural candidates only. However, the electron-phonon coupling can be re-written in terms of single-site Hopfield parameters  $\eta_i$ ,<sup>162</sup> which in turn can be estimated from scattering theory using only single-particle wave function information around each atom.<sup>163,164</sup> A fitness function to maximize the Hopfield parameters,  $P = \max_i \eta_i$ , could be used for prediction of high- $T_c$  candidate materials. In a recent alternative,<sup>165</sup> the Hopfield parameters were used as a secondary filter (after closeness to the convex hulls) to identify high- $T_c$  candidates, for which full calculations of  $\lambda$  and  $T_c$  were then performed; the amount of structural candidates that could be tested (in this approximate way) for superconducting behavior was several orders of magnitude higher than before.

Materials where electrons localize in the interstitial spaces between atoms are called **electrides**; they are stoichiometric pseudo-ionic compounds where the localized electrons take up the role of the anions.<sup>166</sup> The electron sites can be occupied by electron pairs or single electrons; the latter giving rise to interesting magnetic properties.<sup>167–169</sup> Either way, the high electron mobility and low work function make electrides valuable catalysts and reducing agents.<sup>170,171</sup> They are readily identified from electronic structure calculations by topological analyses of the electron density or derived measures, such as the electron localization function (ELF)<sup>172,173</sup>: ELF maxima not associated with ionic cores or covalent bonds, or density maxima away from atomic positions, are indicators for electride character that are obtained at very little computational cost during structure searching. A fitness function that maximizes such regions in the unit cell,  $P = V_{\text{inter}}/V_{\text{unit}}$ , has been used successfully to predict new electride materials.<sup>174</sup>

**Thermoelectrics** are energy materials where the thermoelectric effect, a voltage difference due to a thermal gradient, is exploited, for instance for refrigeration or waste heat recovery. Thermoelectrics are most often characterized by their Seebeck coefficient  $S = \Delta V/\Delta T$  and their figure of merit  $ZT = \sigma ST/\kappa$ , which combines  $S$  with their electrical conductivity  $\sigma$  and thermal conductivity  $\kappa$ . Accurately determining those transport coefficients is computationally expensive but reasonable approximations, such as the constant relaxation time approximation, have been implemented<sup>175,176</sup> that allow estimates of  $S$  and  $ZT$  based on standard DFT band structure information. This allows for high-throughput screening for thermoelectric properties.<sup>177,178</sup> For actual structure prediction, both  $S$  and  $ZT$  have been used (next to total energy) as part of multi-objective optimization searches.<sup>179,180</sup>

### 3.12.6 Accelerating the search process

The techniques described in Section 3.12.4 all require a series of first principles geometry optimisations, each of which consists of tens of energy and force calculations. This presents a bottleneck for CSP. Here, we describe some techniques that can reduce the number of expensive calculations and speed up the structure search.

#### 3.12.6.1 Hard sphere potentials and volumes

Considering the vast size of the potential energy surface, it should be appreciated that a significant portion of the potential energy landscape consists of unphysically dispersed or dense structures. Care should be taken to generate structures that are physically reasonable before invoking the chosen energy calculator. Therefore, candidate structures should be generated based on realistic estimates of the density and typical bond lengths. Any initial structures that deviate significantly from those should be discounted or adjusted prior to calculation. A related, albeit extreme case is that pseudopotentials have a finite cut-off radius separating the frozen core region and the Kohn-Sham valence electrons. When the core regions overlap too much, the ab initio calculation will often be nonsense and should be avoided.

Reasonable atomic separations should be considered pair-wise in multispecies systems; for example,  $\text{LiCoO}_2$ , which is used as a cathode material, has crystal structures which are likely to have close Li–O contacts but not Li–Li or Co–Li.<sup>5</sup> Each pair

combination can have a different minimal allowed separation. This is sometimes input as a matrix or list of pair distances. In Fig. 8, an example is shown for BaTiO<sub>3</sub>.

One approach, as implemented in AIRSS, is to apply hard-sphere pre-optimizations — with sphere radii determined by the expected bond length — to all proposed structures prior to first principles calculation. With a determined unit cell chosen, the atoms can be placed inside and optimized with the simple potential.

In some cases, the ideal density and bond lengths are unknown a priori. In this case, while the search can begin with no restrictions, after some time ‘sensible’ low energy structures will emerge. From those, hard sphere radii and volumes can be determined and used throughout the remaining searches. In a sense, when used to construct hard-sphere potentials, this is a basic — and perhaps trivial — example of training a potential on ab initio data.

### 3.12.6.2 Modular decomposition

Complex crystal structures can often be decomposed into recognizable building blocks. A trivial example is that a random search of stoichiometric H<sub>2</sub>O would find the low energy structures are networks of H<sub>2</sub>O molecules. The conclusion from this would be that the dimensionality of the search space could be reduced by building structures with H<sub>2</sub>O molecules from the beginning; the composition space of OH<sup>−</sup> and H<sup>+</sup> is less relevant and should be explored less. On a larger scale, complex structures like the allotropes of Boron can be decomposed into building blocks of clusters. Alpha Boron, a 12 atom rhombohedral structure is known to be formed of connected B<sub>12</sub> icosahedra. Gamma Boron, a 28 atom orthorhombic structure, found at high pressure, consists of B<sub>12</sub> icosahedra and B<sub>2</sub> dimers.<sup>181</sup> Ahnert et al. used this concept, conceiving a smart decomposition of structures with an emphasis on modularity, to perform structure searches for other allotropes of Boron.<sup>182</sup> The configuration space is reduced by building structures only containing the expected ‘building blocks’. This modular decomposition was also used to accelerate searches to rediscover complex fibrous Phosphorus crystal structures with a machine learned potential.<sup>183</sup>

In the Gamma Boron example, decomposing the structure into 12 atom icosahedra, the  $3N$  degrees of freedom reduce to just  $\frac{5N}{12}$  degrees of freedom containing  $\frac{3N}{12}$  translational degrees of freedom and  $\frac{2N}{12}$  rotational degrees of freedom.

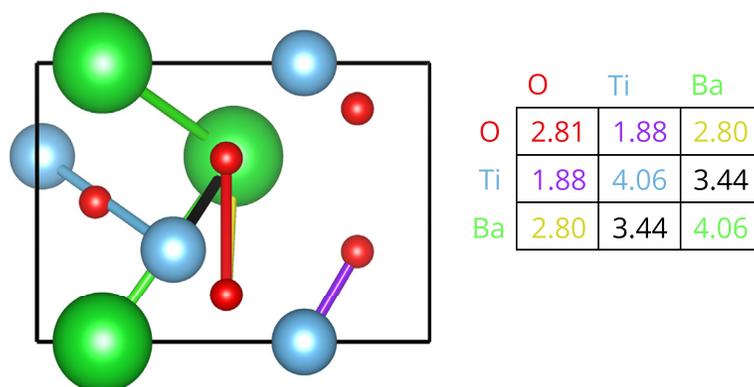
### 3.12.6.3 Orbital free DFT

Another approach is to use faster implementations of DFT to perform the geometry optimizations. The Kohn-Sham formalism of DFT described in Section 3.12.3 uses orbitals to calculate the kinetic energy and the electron density, which is in turn used to calculate the remaining energy contributions. In practice, this process requires a mesh sampling of the Brillouin zone to capture the orbitals as they vary with wave vector,  $\mathbf{k}$ . There is a trade-off between computational cost and accuracy depending on this sampling. In the case of metals, this becomes problematic as poor sampling leads to overly rugged landscapes, and as such the smoothness of the PES and hence surmountability of CSP described in Section 3.12.2 becomes less applicable.

Orbital free methods use a *kinetic energy* functional to determine the total energy using electronic density alone and as a result should scale in theory linearly with the number of atoms in the unit cell. This has been demonstrated to recover low energy structures of Li, Na, Mg, and Al<sup>184</sup> but requires a good functional to estimate the kinetic energy. Development of suitable kinetic energy functionals is ongoing.<sup>185</sup>

### 3.12.6.4 Machine learned potentials

Interatomic potentials have been used widely for a range of purposes since calculations are fast and can be used to study large systems. However, they will not describe well the novel environments seen in the candidate crystal structures. Hence, there is



**Fig. 8** The crystal structure of BaTiO<sub>3</sub>. Pair distances are shown in the matrix. Sensible low-energy structures will have similar pair separations. In AIRSS, this is given with the command, #MINSEP = 1.0 O-O = 2.81 O-Ti = 1.88 O-Ba = 2.80 Ti-Ti = 4.06 Ti-Ba = 3.44 Ba-Ba = 4.06.

a practical upper limit to the number of atoms used in first principles CSP. Machine learned (ML) potentials aim to describe the ab initio PES but calculated at a fraction of the cost.

The concept of fitting interatomic potentials to ab initio data is not new, but historically restricted to datasets of a few atoms in a small number of configurations.<sup>186</sup> The increased availability of high-throughput ab initio data and the renewed interest in non-linear fitting techniques such as artificial neural networks<sup>187</sup> and Gaussian regression<sup>188</sup> has garnered much interest in recent years.

Typically, ML potentials are effective because the model can “learn” local energies  $\epsilon_i$  for each atomic environment, taking advantage of the ‘shortsightedness’ of atomic interactions. The total energy is then a sum of each atoms contribution,  $E = \sum_i \epsilon_i$ . This means the ab initio training data can use relatively low numbers of atoms, remaining computationally manageable but still capturing a diverse set of local environments. The potential can then be used on much larger systems, previously inaccessible with DFT.

ML potentials generally consist of three components; a reference dataset, a descriptor, and a regression tool. The reference dataset consists of several atomic configurations with known energies and forces calculated with high accuracy such as with DFT. The transferability of the resulting potential depends heavily on this dataset.

A descriptor is selected to accurately represent each atomic environment, and eventually return a local energy and forces. Hence, the descriptor must be invariant to permutations, translations and rotations of atoms.<sup>189</sup> Meaning, swapping two identical atoms, should give an identical descriptor and hence an identical energy. For practicality, the descriptor only describes the environment within a sphere with radius of a few Angstrom.

Finally, the regression tool is trained to predict total energies for previously unseen crystal structures. This can be a linear regression,<sup>190</sup> artificial neural network<sup>187</sup> or kernel methods.<sup>188</sup>

One example descriptor is the ‘Smooth Overlap of Atomic Positions’, or SOAP,<sup>191</sup> which consists of an atomic density field formed of gaussians placed at each neighboring atom site, expanded in a basis of spherical harmonics and radial basis functions. This descriptor lends itself well to the Gaussian Approximation Potential (GAP) framework,<sup>188</sup> which uses Gaussian process regression to estimate the atomic energy and forces of a given environment by comparison with known energies for environments present in the reference dataset.

It can take a long time to generate potentials that are sufficiently transferable and accurate, often as a result of intuitive curation of reference datasets. These potentials can then be used on their own going forward in a variety of applications. An alternative approach is to routinely generate new potentials ‘on-the-fly’ with minimal human input. The latter potentials will often be less accurate but specific to the desired use. As an example, while silicon has been well studied and has very accurate machine learned interatomic potentials,<sup>187,192</sup> these would be of little use to study a less ubiquitous silicon-containing compound such as SiH<sub>4</sub> or SiC for which new potentials are required.

Crystal structure prediction, accelerated by MLPs, uses the ‘on-the-fly’ approach.<sup>193–195</sup> The potential need only be accurate enough to generate low energy structures, which can ultimately be refined using DFT. Generating potentials suitable for global optimization requires a reference dataset that spans low energy and high energy configurations to capture the atomic environments seen throughout the optimization. Training data for these potentials can be generated from single point energy calculations on randomly generated crystal structures. It can also be effective to iteratively train the model by using the potential to optimize structures to local minima and calculating ab initio energies for structures close to the predicted local minima. This enables structure discovery with significantly fewer explicit ab initio calculations and can search unit cells with significantly more atoms. Recent applications of MLPs with unit cells containing large numbers of atoms include a priori ‘rediscovering’ allotropes of Boron,<sup>13196</sup> prediction of allotropes of phosphorus<sup>197</sup> and prediction of a novel 60 atom phase of SiH<sub>4</sub> at high pressure.<sup>196</sup>

### 3.12.7 Visualizing the PES

The PES exists in a high dimensional space of all atomic configurations and so is inherently difficult to visualize. This represents a well known task in data science, to project the high dimensional data into three dimensions or less. A few approaches have been introduced to do this. However, as a first step the crystal structures must be ‘vectorized’ to define a consistent high dimensional basis. This means rather than choosing a *local* descriptor to describe atomic environments, we seek a *global* descriptor that describes the entire crystal structure. This is similar to the concept of collective variables discussed in [Section 3.12.4.5.1](#). However, these global descriptors are not constrained to just a few variables.

Earlier, we discussed the crystal fingerprints — or global descriptors — suggested by Valle and Oganov,<sup>77</sup> which are histograms of pair-wise atomic separations. These histograms are a  $1 \times N$  vector where  $N$  is the number of histogram bins. It is useful to define a measure of how separated the structures are from each other based on this descriptor. Valle and Oganov use a cosine distance measure — meaning the dot product of the descriptor vectors — and find that plotting energy against this distance measure gives information about the potential energy landscape. In particular they show how local minima can appear as ‘funnels’ toward the lower energy minima.<sup>198,199</sup>

An alternative vectorization is to, in some way, combine the local atomic environment descriptors in the unit cell to give a global descriptor. The aforementioned SOAP descriptor vectors, for example, can be simply averaged over atoms in the unit cell to give a global descriptor vector.<sup>200</sup> Of course, averaging the descriptors in this way loses some information; two distinct crystal structures could give similar averaged global fingerprints. Alternative but more complex global descriptors calculated from local environments are described by De et al.<sup>200</sup>

Several approaches have been used to reduce the dimensionality of this vector space, such as sketch maps<sup>201</sup> and disconnectivity graphs.<sup>202</sup> However, the specific task relevant to crystal structure prediction of representing a range of local minima — as opposed to a series of molecular dynamics snapshots — involves capturing information about the basin sizes, connectivity and separation. A recent approach by Shires and Pickard<sup>203</sup> does this by applying state-of-the-art dimension reduction techniques, such as t-SNE and UMAP,<sup>204,205</sup> to project the vectors on to three or fewer dimensions. The size of the basins are considered by combining similar structures, where more repetitions are represented by larger symbols. The connectivity of the basins is retained by imposing hard-sphere potentials to points in the reduced dimension projection. An example of this is shown in Fig. 9.

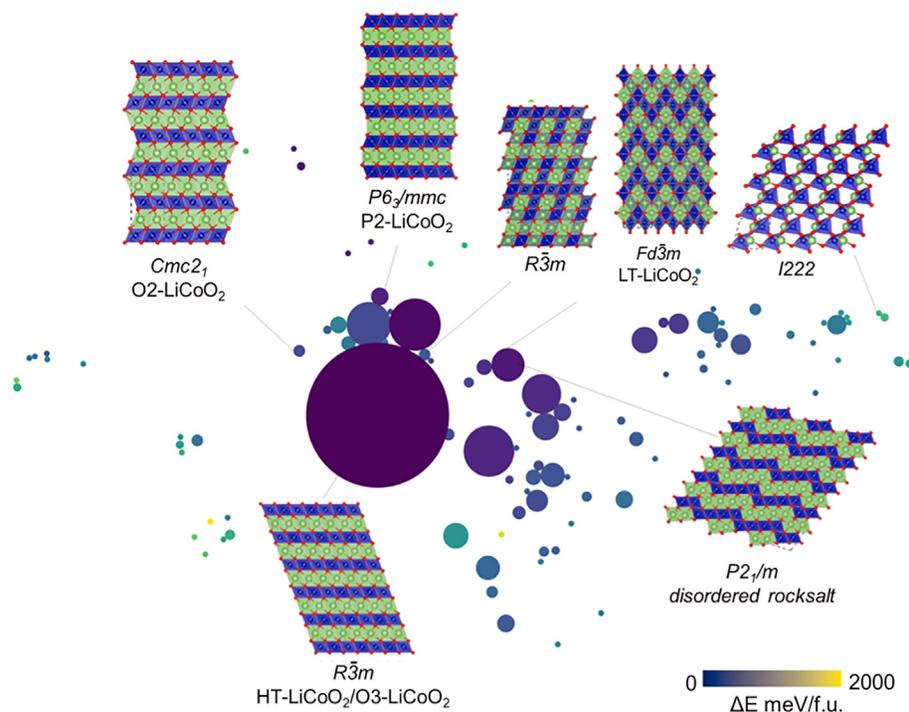
### 3.12.8 Examples

Thanks to the ubiquity of high-performance computing resources that can efficiently run multiple parallel instances of DFT codes, CSP has become a burgeoning field. Below we collect examples from the literature from different fields where first principles CSP is (or could be) a successful driving force of research.

#### 3.12.8.1 Compositional complexity

In chapter 2, we introduced a combinatorial estimation of the number of atomic configurations for a single element system. Naturally, introducing multi-element variation increases this number dramatically, making CSP more difficult. It is a challenge that is of interest, since superconducting cuprate materials typically contain between four and five element types,<sup>206</sup> and — as will be discussed — ternary superconducting hydrides are showing theoretical and experimental evidence for high temperature superconductivity.<sup>207–211</sup>

An application where compositional complexity dominates is in cathode materials. Particularly, with a desire to design next-generation Lithium-ion batteries for electric vehicles or ubiquitous energy storage, featuring longer lifespans and storage capacities.<sup>212</sup> A prominent material and the first commercialized Lithium-ion battery is  $\text{LiCoO}_2$ .<sup>213</sup> Since then, efforts have been made to improve the capacity by substituting transition metals into the  $\text{LiMO}_2$  framework, again presenting a combinatorial problem. This was explored initially through *ab-initio* calculations by Ceder et al.,<sup>214</sup> with manual construction of  $\text{LiMAlO}_2$  derivatives (where  $M = \text{Ti, V, Fe, Co}$ ). Then, more ambitiously, using high-throughput screening of ternary oxide compounds using known crystal structures from a database,<sup>83</sup> and calculations of  $A_xM(\text{YO}_3)(\text{XO}_4)$  (with  $A = \text{Na, Li}$ ;  $X = \text{Si, As, P}$ ;  $Y = \text{C, B}$ ;  $M =$  a redox active metal;) in the sidorenkite crystal structure.<sup>84</sup>



**Fig. 9** SHEAP map of  $\text{LiCoO}_2$  structures generated from a random structure search. Circle sizes correspond to the number of structure repetitions (hence the size of the local basins), colors represent energy scale from purple (most stable) to yellow (most unstable). Reproduced with permission from Lu et al. Lu, Z.; Zhu, B.; Shires, B.W.B.; Scanlon, D.O.; Pickard, C.J. Ab Initio Random Structure Searching for Battery Cathode Materials. *J. Chem. Phys.* **2021**, 154(17): 174111, doi:10.1063/5.0049309.

However, perhaps the next-generation material does not conform to a known structure type. Recently, calculations have attempted to generate novel structures for a wide range of stoichiometries using AIRSS,<sup>4,5</sup> reproducing known polymorphs and suggesting new stable and promising new materials in  $\text{LiCoO}_2$ ,  $\text{LiFePO}_4$ , and the  $\text{Li}_{x_1}\text{Cu}_{x_2}\text{F}_{x_3}$  ternary. Due to the sheer scale of the computational problem, these efforts may have only scratched the surface and perhaps the best material is a quaternary or quinary structure in an unknown structure type. Cathode materials will be at the cutting-edge of CSP in the coming years.

An entirely different proof-of-concept structure search by Pickard provides an unbiased sampling of the quaternary H–C–N–O chemical space at 1 GPa pressure.<sup>215</sup> While the pressure removes the dominance of dispersion interactions in the structures formed, the dataset reflects an ambitious survey of the most dominant organic chemical space (the PubChem database lists close to 45 million compounds within this composition space<sup>216</sup>). The dataset contains many known thermodynamic sinks, and the quaternary nature of the space is well reflected if visualized using the SHEAP mapping introduced above<sup>203</sup>; it remains available for further investigations. Unrelated studies of planetary ice mixtures, motivated by the unknown makeup of deep planetary interiors, presented quaternary H–C–N–O phase diagram at several Mbar pressures,<sup>217,218</sup> implying that compositionally complex systems can now be studied using first principles CSP.

### 3.12.8.2 High pressure

Energy landscapes of elements and compounds at high pressures often deviate from chemical intuition attuned to ambient conditions.<sup>219,220</sup> Due to the non-intuitive effects, the exploration of high pressure energy landscapes with crystal structure prediction methods has proven extremely useful to reveal fundamental new physics and chemistry that governs materials' stability and properties at high densities.<sup>221</sup>

#### 3.12.8.2.1 Elements

The lightest element, **hydrogen**, retains immense interest due to its complex high-pressure phase diagram and the promise of high-temperature superconductivity following its eventual metallization.<sup>222</sup> One of the most intriguing effects is the appearance of a strongly infrared-active vibron mode in its phase III above 150 GPa, implying the presence of polarized  $\text{H}_2$  molecules in an anisotropic matrix of their peers.<sup>223</sup> An early success of random structure searching was the discovery of a structure that fits this description<sup>16</sup>; while free energy calculations of the same dataset revealed a candidate structure for the high-temperature phase IV.<sup>224</sup> Computational work on hydrogen continues, and its unique challenges due to the low nuclear mass drive methodological developments in structure prediction.<sup>133</sup>

**Boron** tends to form very complex crystal structures, with polyhedral motifs borne out of an electron deficiency that favors multicenter bonding. This tests the ability of structure prediction methods to successfully perform searches with large numbers of atoms. The  $\gamma$ - $\text{B}_{28}$  phase was found in evolutionary algorithm searches by exploiting experimental unit cell information.<sup>181</sup> Boron provides motivation for continued methodology development to scale up structure predictions, such as automated detection of atomic motifs (making polyhedra instead of atoms the units of the search),<sup>182</sup> or on-the-fly training of interatomic potentials to vastly increase structure throughput.<sup>13196</sup>

**Sodium** is a textbook simple metal at ambient conditions but is also exemplar for non-trivial electronic transitions at high pressure, most strikingly by taking up an insulating phase above 200 GPa that was found in evolutionary algorithm structure searches.<sup>22</sup> The appearance of bad metals or insulators among most group I and II elements correlates with their tendencies to form high-pressure electrides, including in complex incommensurate host-guest phases.<sup>225</sup> Random structure searches predict that trivalent **aluminium** will also take up such host-guest phases in the TPa pressure range.<sup>226</sup>

#### 3.12.8.2.2 Hydrogen bonding

The evolution of hydrogen bonds under pressure is relevant to many fields beyond inorganic chemistry, including biology (boundaries of life at extremes), geosciences (water storage inside minerals) and planetary sciences (icy planet interiors).

Among hydrogen-bonded **ices** the pressure response varies significantly. **Ammonia** was predicted from random structure searching to self-ionize into  $(\text{NH}_2)^-(\text{NH}_4)^+$ ,<sup>15</sup> which was later confirmed in experiment.<sup>227,228</sup> Water is predicted to undergo a series of phase transitions beyond hydrogen bond symmetrisation,<sup>21,229–231</sup> ultimately decomposing in the TPa pressure range.<sup>232</sup> Experiments have not yet reached those pressure conditions, instead new ice phases are found at high temperatures, in the superionic regime of hot dense water.<sup>233–235</sup> **Ammonia-water mixtures** are again predicted to form ionic high pressure compounds,<sup>20,236</sup> which has been confirmed by experiment.<sup>237</sup> A cautionary tale is **hydrogen chloride**, where (post hydrogen bond symmetrization) a series of low-symmetry ground state phases was predicted,<sup>238–240</sup> whereas experiments and molecular dynamics simulations found a high-symmetry, room-temperature superionic phase.<sup>241</sup> The emergence of high-symmetry phases at elevated temperatures, whether caused by quantum or entropic effects, continues to present challenges for crystal structure prediction methods that rely on full structural optimizations toward local minima.

**Metal hydroxides** of formula  $\text{MOH}$  (for group I metals) or  $\text{M}(\text{OH})_2$  (for group II metals) are another class of hydrogen-bonded materials with rich structural variety under pressure. Crystal structure prediction was instrumental in both *enabling* and *correcting* previous experimental structure assignments in dense alkali hydroxides.<sup>242–244</sup>

**Hydrous minerals** are compounds that store hydrogen at deep Earth conditions; the hydrogen bonding influences individual phases' stabilities and elastic properties. Computational predictions have successfully preceded experimental syntheses in this

area,<sup>245,246</sup> while crystal structure prediction has been applied to scenarios that range from Earth's mantle transition zone to the primordial core.<sup>247–250</sup>

### 3.12.8.2.3 Polyhydride superconductors

The theory of metallic and superconducting dense hydrogen evolved toward the concept of 'chemically precompressed' hydrogen in hydride compounds.<sup>252</sup> The very first published random structure search results were in this area, on silane,<sup>17</sup> and since then the field has seen an extraordinary sequence of crystal structure prediction followed by experimental synthesis and characterization. The structure predictions usually follow the sequence of identifying low-enthalpy structures, nowadays across variable composition ranges, before calculating superconducting properties for the best candidates. As a general theme, high pressure favors the formation of polyhydrides with higher hydrogen content than at ambient or low pressure conditions, hence searches need to screen a wide range of hydrogen compositions. Polyhydrides that have been predicted (all with superconducting  $T_c$  above 200 K) and subsequently synthesized (all above 150 GPa) include  $\text{CaH}_6$ ,<sup>253–255</sup>  $\text{H}_3\text{S}$ ,<sup>18,256</sup>  $\text{LaH}_{10}$ ,<sup>257–260</sup> and  $\text{YH}_6 - 9$ .<sup>261–263</sup> Full screening of ternary composition spaces is now possible, and ternary polyhydrides have been predicted with even higher  $T_c$  or with lower formation pressures.<sup>207–210</sup>

### 3.12.8.3 Target properties

In Section 3.12.5 we discussed different types of fitness functions (other than total energy or enthalpy) that can be used to direct structure searches toward specific properties, using multi-objective optimizations. In truth, those targeted approaches have not been as prominent as one might expect – despite the need for new materials with specific properties. That is not to say that applications do not exist: multi-objective searches have yielded superhard materials,<sup>264–266</sup> electrides,<sup>174</sup> and thermoelectrics.<sup>179,180</sup> However, a much more common approach is to search for low-energy or low-enthalpy structures first, and then investigate those for properties of interest. This often includes expensive follow-up calculations, and is therefore restricted to a few candidate structures. There is therefore merit in developing or adapting computationally cheaper predictive models, to enable pre-screening for interesting properties across a wider range of candidates – ideally the entire dataset of structures that emerge from a CSP search. The surrogate model for superconducting  $T_c$ , based on single-electron wave functions, is an example, as it allows rapid qualitative statements on prospects of superconductivity.<sup>165</sup> Across materials informatics, a plethora of similar structure-property relationships have been codified for all kinds of materials properties in recent years, often using machine learning techniques, which could be used in conjunction with CSP.<sup>135,267–269</sup>

### 3.12.9 Challenges and conclusions

While first principles crystal structure prediction has been shown to be successful in materials discovery, challenges remain. Those challenges can stem from deficiencies in the DFT calculations to describe particular compounds; or from inadequacy of total energy calculations to deduce stability. The latter relates for example to the fact that DFT calculations of a material require well defined atomic positions; disorder is not usually handled by CSP.

However, real materials can be intrinsically **disordered**. In fact, a large number of structures on the ICSD contain partial site occupancy (SOF < 1), implying some disorder. This could be a structural disorder such as in  $\text{H}_2\text{O}$ , where the molecules appear to be tetrahedra with 0.5H atoms on each corner (see Fig. 10A). Crystal structure prediction will only discover the ordered approximation of these structures. In this example, it can discover the ordered structure, ice XI, but not the isostructural disordered ice Ih.

Alloys can have well defined crystal structures with non-integer occupancies that would not be described well using DFT. The stoichiometries accessible with CSP are limited by the number of atoms used in the search. Fig. 10B) shows an experimental determination of the crystal structure of an  $\text{Mn}_{0.95 - \delta}\text{Fe}_{1.05 - x + \delta}\text{Co}_x\text{B}$  alloy. However, the exact stoichiometry is not easily accessed in calculations. Furthermore, it is not clear that CSP is the right tool for alloys: an analysis of metallic compounds within the AFLOW materials library suggested that their formation is dominated by entropy rather than energy gains if they contain four or more elements.<sup>270,271</sup>

Configurational entropy is but one contribution to a compound's **Gibbs free energy** that DFT total energy calculations struggle to quantify. The aim of first principles CSP, to find the global minimum of the potential energy landscape, does not give the complete (or often, sufficient) picture. At finite temperatures, dynamic effects can change relative stabilities of different structures, which has two major consequences. Firstly, CSP methods need to find and catalog the spectrum of metastable states, in addition to the global minimum, and do this in an ergodic way, i.e. representing ensemble averages. For search methods with 'memory' this is not a priori clear. For the 'local' approaches discussed in Section 3.12.4.5 the challenge is the opposite, as they need to ensure to sample the local environment of the PES wide enough to encounter all relevant metastable states. For the 'global' survey approaches introduced in Section 3.12.4 additional benefit would emerge from information on the connectivity of different basins within the PES. Some progress is currently being made by visualizing the PES based on structural similarities, but identifying reaction pathways and transition states remains a big challenge.

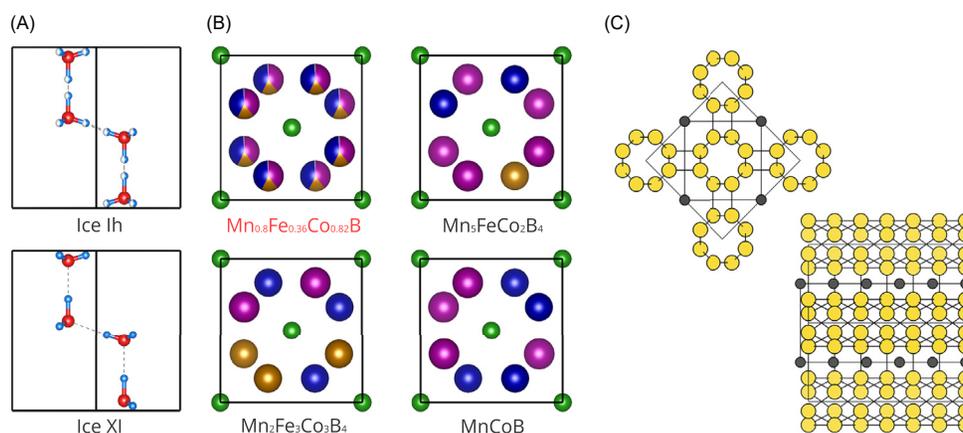
The second consequence of dynamic effects is the stabilization of phases that are not even local minima on the PES, as they are intrinsically stabilized by entropy. The prototypical cubic perovskite structure will not usually emerge from CSP of  $\text{ABO}_3$  compounds, because it often features a dynamical instability toward lower symmetry structures with rotated and tilted octahedra.

Molecular crystals may have rotational ('plastic') states at elevated temperatures,<sup>272</sup> which will not be uncovered from CSP alone. Some materials display superionicity, ionic diffusion, or chain melting (in each case a subset of atoms becomes mobile against a fixed lattice of the remaining atoms). Far below the characteristic temperature scales for these phenomena, free energies can be estimated using vibrational entropies alone. This has led to a candidate structure for hydrogen's phase IV.<sup>224</sup> Those calculations require expensive phonon calculations and are therefore not routinely accessible to CSP at the moment – but may be via structure-property relationships or efficient interatomic potentials, if those can readily deliver phonon densities of states. An entirely different issue with hydrogen, due to its very low mass, is the large contribution of **nuclear quantum effects**, which arguably require much more computationally intensive approaches than standard DFT to capture phase stabilities and transitions properly.

For some systems the periodic setup of DFT calculations for solids is insufficient. Some compounds contain periodicities that are **incommensurate** with the lattice. This can manifest as a host-guest structure, where the periodicity of the 'host atom' lattice is incommensurate with the 'guest atom' lattice. Known examples include the alkali metals, the heavier pnictogens, and predictions for Al, all under pressure,<sup>225,226,273</sup> while  $\text{LiB}_x$  and  $\text{Hg}_{3-\delta}\text{AsF}_6$  have linear incommensurate chains at ambient conditions.<sup>274–276</sup> Incommensurate structures can also manifest as charge density waves (CDW) or modulated crystal structures, where atoms are displaced along a wave vector incommensurate with the average crystal structure. Examples of these include uranium at ambient and the chalcogenides S–Te under pressure.<sup>277</sup> Finally, there are compounds, such as  $\text{NaNO}_2$  and CaAu-II that display *occupational* modulation where the site occupancies vary along an incommensurate wave vector. Those phases cannot be modelled in periodic DFT, nor discovered directly using CSP. However, commensurate or ordered approximants of host-guest and CDW structures can be constructed around a feasible superstructure that captures the two periodicities.<sup>226,278</sup> Incommensurate structures may have an ordered variant, stable above a transition temperature, where either the CDW amplitude vanishes (as in the high pressure S–IV to S–V transition), or the periodicity becomes commensurate. An excellent review of incommensurate structures and the related nomenclature has been given by Sander Van Smaalen.<sup>279</sup>

First principles CSP is also undermined in cases where DFT fails to capture significant portions of the **electronic interactions**. This used to include weakly bound systems dominated by long-range dispersion interactions but the development of dispersion-corrected exchange-correlation functionals in the last 10–15 years now allows accurate calculations of those. However, strongly correlated electron systems still suffer from the inadequate formalism of electron correlation within DFT. This includes open-shell systems (where DFT can struggle to reproduce magnetism) and Mott insulators (which tend to be metallic within vanilla DFT), among others. In addition, the systematic underestimation of electronic band gaps within DFT means that even if the electronic structure is qualitatively correct, quantitative predictions, e.g. of metallization pressures or optical band gaps, remain difficult. Corrections for these issues exist for specific systems, for example on-site Hubbard repulsion terms that penalize metallic states or material-specific functionals to reproduce experimental band gaps, but introducing system-specific parameters terms is not ideal in the spirit of automated CSP.

One main purpose of CSP is to uncover materials with **desirable properties** in silico, thereby bypassing expensive trial-and-error synthesis efforts. As pointed out in [Section 3.12.8.3](#), the direct search for specific properties has probably not reached its full potential. Partly, this might be due to the difficulty in capturing relevant physicochemical properties based on structural information alone: catalytic activity, for instance, often relates to defects and surfaces of materials, rather than their bulk crystal structure. Nonetheless, ongoing exploits in machine-learned structure-property relationships can hopefully be interfaced with CSP more often and thus lead to more relevant predictions in materials science.



**Fig. 10** Crystal structure prediction can be challenging on systems with partial occupancy (A, B) or incommensurate structures (C). (A) Ice Ih with partial H occupancy and ordered ice XI. O/H atoms in red/blue. Partial occupancies given by partially filled spheres. (B)  $\text{Mn}_{0.95-\delta}\text{Fe}_{1.05-x+\delta}\text{Co}_x\text{B}$  alloy with occupancies measured experimentally by XRD, and three ordered approximations in the same unit cell that may be found by CSP. Mn/Fe/Co/B atoms in purple/bronze/blue/green. Partial occupancies given by partially filled spheres. (C) Incommensurate host-guest phase of K-IIIa at high pressure, along the [001] (left) and the [110] directions. Reproduced with permission from Woolman, G.; Naden Robinson, V.; Marqués, M.; Loa, I.; Ackland, G.J.; Hermann, A. Structural and Electronic Properties of the Alkali Metal Incommensurate Phases. *Phys. Rev. Mater.* **2018**, 2(5), 053604, doi:10.1103/PhysRevMaterials.2.053604.

A final fundamental challenge is **the combinatorial problem** stated at the outset of this chapter. CSP (whether from first principles or not) will always be limited in unit cell size and compositional variety that can be searched by the computational resources necessary to sample configurational space. Even if the ground state minimum is the most likely structure to be found, in absolute terms it might still take many thousands or millions of attempts to locate it. The modern approaches in CSP have had remarkable success yet ongoing methodology developments aimed to speed up the searches, as discussed in this chapter, remain necessary to allow even more relevant searches to take place.

## Acknowledgments

We are grateful for support by the UK's Engineering and Physical Sciences Research Council over many years, in particular for financial support for L.J.C. (EP/L015110/1 and EP/S021981/1).

## References

- Schriber, E. A.; Paley, D. W.; Bolotovskiy, R.; Rosenberg, D. J.; Sierra, R. G.; Aquila, A.; Mendez, D.; Poitevin, F.; Blaschke, J. P.; Bhowmick, A.; Kelly, R. P.; Hunter, M.; Hayes, B.; Popple, D. C.; Yeung, M.; Pareja-Rivera, C.; Lisova, S.; Tono, K.; Sugahara, M.; Owada, S.; Kuykendall, T.; Yao, K.; Schuck, P. J.; Solis-Ibarra, D.; Sauter, N. K.; Brewster, A. S.; Hohman, J. N. Chemical Crystallography by Serial Femtosecond X-Ray Diffraction. *Nature* **2022**, *601* (7893), 360–365. <https://doi.org/10.1038/s41586-021-04218-3>.
- Zhang, W.; Oganov, A. R.; Goncharov, A. F.; Zhu, Q.; Bouffelfel, S. E.; Lyakhov, A. O.; Stavrou, E.; Somayazulu, M.; Prakapenka, V. B.; Konôpková, Z. Unexpected Stable Stoichiometries of Sodium Chlorides. *Science* **2013**, *342* (6165), 1502–1505. <https://doi.org/10.1126/science.1244989>.
- Hutcheon, M. J.; Shipley, A. M.; Needs, R. J. Predicting Novel Superconducting Hydrides Using Machine Learning Approaches. *Phys. Rev. B* **2020**, *101* (14), 144505. <https://doi.org/10.1103/PhysRevB.101.144505>.
- Zhu, B.; Lu, Z.; Pickard, C. J.; Scanlon, D. O. Accelerating Cathode Material Discovery through *Ab Initio* Random Structure Searching. *APL Mater.* **2021**, *9* (12), 121111. <https://doi.org/10.1063/5.0076220>.
- Lu, Z.; Zhu, B.; Shires, B. W. B.; Scanlon, D. O.; Pickard, C. J. *Ab Initio* Random Structure Searching for Battery Cathode Materials. *J. Chem. Phys.* **2021**, *154* (17), 174111. <https://doi.org/10.1063/5.0049309>.
- van't Hoff, J. H. A Suggestion Looking to the Extension into Space of the Structural Formulas at Present Used in Chemistry. And a Note upon the Relation between the Optical Activity and the Chemical Constitution of Organic Compounds. *Arch. Neerl. des Sci. Exactes Nat.* **1874**, *9*, 445–454.
- Maitland, P.; Mills, W. H. Experimental Demonstration of the Allene Asymmetry. *Nature* **1935**, *135* (3424), 994. <https://doi.org/10.1038/135994a0>.
- Goldschmidt, V. M. Die Gesetze der Krystallochemie. *Naturwissenschaften* **1926**, *14* (21), 477–485. <https://doi.org/10.1007/BF01507527>.
- Goldschmidt, V. M. Krystalbau und chemische Zusammensetzung. *Berichte der Dtsch. Chem. Gesellschaft* **1927**, *60* (5), 1263–1296. <https://doi.org/10.1002/cber.19270600550>.
- Errea, I.; Calandra, M.; Pickard, C. J.; Nelson, J.; Needs, R. J.; Li, Y.; Liu, H.; Zhang, Y.; Ma, Y.; Mauri, F. High-Pressure Hydrogen Sulfide from First Principles: A Strongly Anharmonic Phonon-Mediated Super-Conductor. *Phys. Rev. Lett.* **2015**, *114* (15), 157004. <https://doi.org/10.1103/PhysRevLett.114.157004>.
- Glass, C. W.; Oganov, A. R.; Hansen, N. USPEX — Evolutionary Crystal Structure Prediction. *Comput. Phys. Commun.* **2006**, *175* (11–12), 713–720. <https://doi.org/10.1016/j.cpc.2006.07.020>.
- Ishikawa, T.; Miyake, T. Evolutionary Construction of a Formation-Energy Convex Hull: Practical Scheme and Application to a Carbon-Hydrogen Binary System. *Phys. Rev. B* **2020**, *101* (21), 214106. <https://doi.org/10.1103/PhysRevB.101.214106>.
- Podryabinin, E. V.; Tikhonov, E. V.; Shapeev, A. V.; Oganov, A. R. Accelerating Crystal Structure Prediction by Machine-Learning Inter-Atomic Potentials with Active Learning. *Phys. Rev. B* **2019**, *99* (6), 064114. <https://doi.org/10.1103/PhysRevB.99.064114>.
- Gao, G.; Oganov, A. R.; Ma, Y.; Wang, H.; Li, P.; Li, Y.; Iitaka, T.; Zou, G. Dissociation of Methane under High Pressure. *J. Chem. Phys.* **2010**, *133* (14), 144508. <https://doi.org/10.1063/1.3488102>.
- Pickard, C. J.; Needs, R. J. Highly Compressed Ammonia Forms an Ionic Crystal. *Nat. Mater.* **2008**, *7* (10), 775–779. <https://doi.org/10.1038/nmat2261>.
- Pickard, C. J.; Needs, R. J. Structure of Phase III of Solid Hydrogen. *Nat. Phys.* **2007**, *3* (7), 473–476. <https://doi.org/10.1038/nphys625>.
- Pickard, C. J.; Needs, R. J. High-Pressure Phases of Silane. *Phys. Rev. Lett.* **2006**, *97* (4), 045504. <https://doi.org/10.1103/PhysRevLett.97.045504>.
- Duan, D.; Huang, X.; Tian, F.; Li, D.; Yu, H.; Liu, Y.; Ma, Y.; Liu, B.; Cui, T. Pressure-Induced Decomposition of Solid Hydrogen Sulfide. *Phys. Rev. B* **2015**, *91* (18), 180502. <https://doi.org/10.1103/PhysRevB.91.180502>.
- Li, Y.; Hao, J.; Liu, H.; Li, Y.; Ma, Y. The Metallization and Superconductivity of Dense Hydrogen Sulfide. *J. Chem. Phys.* **2014**, *140* (17). <https://doi.org/10.1063/1.4874158>.
- Naden Robinson, V.; Wang, Y.; Ma, Y.; Hermann, A. Stabilization of Ammonia-Rich Hydrate inside Icy Planets. *Proc. Natl. Acad. Sci.* **2017**, *114* (34), 201706244. <https://doi.org/10.1073/pnas.1706244114>.
- Hermann, A.; Ashcroft, N. W.; Hoffmann, R. High Pressure Ices. *Proc. Natl. Acad. Sci.* **2012**, *109* (3), 745–750. <https://doi.org/10.1073/pnas.1118694109>.
- Ma, Y.; Eremets, M.; Oganov, A. R.; Xie, Y.; Trojan, I.; Medvedev, S.; Lyakhov, A. O.; Valle, M.; Prakapenka, V. Transparent Dense Sodium. *Nature* **2009**, *458* (7235), 182–185. <https://doi.org/10.1038/nature07786>.
- Xia, K.; Gao, H.; Liu, C.; Yuan, J.; Sun, J.; Wang, H. T.; Xing, D. A Novel Superhard Tungsten Nitride Predicted by Machine-Learning Accelerated Crystal Structure Search. *Sci. Bull.* **2018**, *63* (13), 817–824. <https://doi.org/10.1016/j.scib.2018.05.027>.
- Salke, N. P.; Xia, K.; Fu, S.; Zhang, Y.; Greenberg, E.; Prakapenka, V. B.; Liu, J.; Sun, J.; Lin, J. F. Tungsten Hexanitride with Single-Bonded Armchairlike Hexazine Structure at High Pressure. *Phys. Rev. Lett.* **2021**, *126* (6), 065702. <https://doi.org/10.1103/PhysRevLett.126.065702>.
- Niggli, P. Handbuch der Experimentalphysik. In *Part 1. Kristallo-graphische und Strukturtheoretische Grundbegriffe*, vol. 7; Akademische Verlagsgesellschaft: Leipzig, 1928, p 317.
- Delaunay, B. Neue Darstellung der geometrischen Kristallographie. *Zeitschrift für Krist. - Cryst. Mater.* **1933**, *84* (1–6), 109–149. <https://doi.org/10.1524/zkri.1933.84.1.109>.
- Grosse-Kunstleve, R. W.; Sauter, N. K.; Adams, P. D. Numerically Stable Algorithms for the Computation of Reduced Unit Cells. *Acta Crystallogr. Sect. A Found. Crystallogr.* **2004**, *60* (1), 1–6. <https://doi.org/10.1107/S010876730302186X>.
- Andrews, L. C.; Bernstein, H. J.; Sauter, N. K. Selling Reduction Versus Niggli Reduction for Crystallographic Lattices. *Acta Crystallogr. Sect. A Found. Adv.* **2019**, *75* (1), 115–120. <https://doi.org/10.1107/S2053273318015413>.
- Oganov, A. R.; Glass, C. W. Crystal Structure Prediction Using *Ab Initio* Evolutionary Techniques: Principles and Applications. *J. Chem. Phys.* **2006**, *124* (24), 244704. <https://doi.org/10.1063/1.2210932>.

30. Oganov, A. R.; Lyakhov, A. O.; Valle, M. How Evolutionary Crystal Structure Prediction Works—And why. *Acc. Chem. Res.* **2011**, *44* (3), 227–237. <https://doi.org/10.1021/ar1001318>.
31. Stillinger, F. H.; Weber, T. A. Packing Structures and Transitions in Liquids and Solids. *Science* **1984**, *225* (4666), 983–989. <https://doi.org/10.1126/science.225.4666.983>.
32. Stillinger, F. H. Exponential Multiplicity of Inherent Structures. *Phys. Rev. E* **1999**, *59* (1), 48–51. <https://doi.org/10.1103/PhysRevE.59.48>.
33. Doye, J. P. K. Network Topology of a Potential Energy Land-scape: A Static Scale-Free Network. *Phys. Rev. Lett.* **2002**, *88* (23), 238701. <https://doi.org/10.1103/PhysRevLett.88.238701>.
34. Doye, J. P. K.; Massen, C. P. Characterizing the Network Topology of the Energy Landscapes of Atomic Clusters. *J. Chem. Phys.* **2005**, *122* (8), 084105. <https://doi.org/10.1063/1.1850468>.
35. Massen, C. P.; Doye, J. P. K.; Nash, R. W. Exploring the Origins of the Power-Law Properties of Energy Landscapes: An Egg-Box Model. *Phys. A: Stat. Mech. Appl.* **2007**, *382* (2), 683–692. <https://doi.org/10.1016/j.physa.2007.04.054>.
36. Jensen, F. *Introduction to Computational Chemistry*, 3rd ed.; John Wiley and Sons: Chichester, 2017.
37. Wales, D. J. Symmetry, Near-Symmetry and Energetics. *Chem. Phys. Lett.* **1998**, *285* (5–6), 330–336. [https://doi.org/10.1016/S0009-2614\(98\)00044-X](https://doi.org/10.1016/S0009-2614(98)00044-X).
38. Donohue, J. Revised Space-Group Frequencies for Organic Compounds. *Acta Crystallogr. Sect. A* **1985**, *41* (2), 203–204. <https://doi.org/10.1107/S010876738500040X>.
39. Baur, W. H.; Kassner, D. The Perils of Cc: Comparing the Frequencies of Falsely Assigned Space Groups with their General Population. *Acta Crystallogr. Sect. B Struct. Sci.* **1992**, *48* (4), 356–369. <https://doi.org/10.1107/S0108768191014726>.
40. Kitaigorodskii, A. I. The Close-Packing of Molecules in Crystals of Organic Compounds. *J. Phys.* **1945**, *9*, 351–352.
41. Kitaigorodskii, A. I. *Organic Chemical Crystallography*, Consultants Bureau: New York, 1961.
42. Kitaigorodskii, A. I. *Molecular Crystals and Molecules*, Academic Press: New York and London, 1973; p 555.
43. Gavezotti, A. Are Crystal Structures Predictable? *Acc. Chem. Res.* **1994**, *27* (10), 309–314. <https://doi.org/10.1021/ar00046a004>.
44. CCDC. *Documentation and Resources*. <https://www.ccdc.cam.ac.uk/support-and-resources/ccdcresources/Accessed>, 2022, 29/04/2022.
45. Hellenbrandt, M. The Inorganic Crystal Structure Database (ICSD)—Present and Future. *Crystallogr. Rev.* **2004**, *10* (1), 17–22. <https://doi.org/10.1080/08893110410001664882>.
46. Reilly, A. M.; Cooper, R. I.; Adjiman, C. S.; Bhattacharya, S.; Boese, A. D.; Brandenburg, J. G.; Bygrave, P. J.; Bylsma, R.; Campbell, J. E.; Car, R.; Case, D. H.; Chadha, R.; Cole, J. C.; Cosburn, K.; Cuppen, H. M.; Curtis, F.; Day, G. M.; DiStasio, R. A., Jr.; Dzyabchenko, A.; van Eijck, B. P.; Elkington, D. M.; van den Ende, J. A.; Facelli, J. C.; Ferraro, M. B.; Fusti-Molnar, L.; Gatsiou, C.-A.; Gee, T. S.; de Gelder, R.; Ghiringhelli, L. M.; Goto, H.; Grimme, S.; Guo, R.; Hofmann, D. W. M.; Hoja, J.; Hylton, R. K.; Iuzzolino, L.; Jankiewicz, W.; de Jong, D. T.; Kendrick, J.; de Klerk, N. J. J.; Ko, H.-Y.; Kuleshova, L. N.; Li, X.; Lohani, S.; Leusen, F. J. J.; Lund, A. M.; Lv, J.; Ma, Y.; Marom, N.; Masunov, A. E.; McCabe, P.; McMahon, D. P.; Meekes, H.; Metz, M. P.; Misquitta, A. J.; Mohamed, S.; Monserrat, B.; Needs, R. J.; Neumann, M. A.; Nyman, J.; Obata, S.; Oberhofer, H.; Oganov, A. R.; Orendt, A. M.; Pagola, G. I.; Pantelides, C. C.; Pickard, C. J.; Podeszwa, R.; Price, L. S.; Price, S. L.; Pulido, A.; Read, M. G.; Reuter, K.; Schneider, E.; Schober, C.; Shields, G. P.; Singh, P.; Sugden, I. J.; Szalewicz, K.; Taylor, C. R.; Tkatchenko, A.; Tucker-man, M. E.; Vacarro, F.; Vasileiadis, M.; Vazquez-Mayagoitia, A.; Vogt, L.; Wang, Y.; Watson, R. E.; de Wijs, G. A.; Yang, J.; Zhu, Q.; Groom, C. R. Report on the Sixth Blind Test of Organic Crystal Structure Prediction Methods. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2016**, *72* (4), 439–459. <https://doi.org/10.1107/S2052520616007447>.
47. Martin, R. M. *Electronic Structure*, Cambridge University Press: Cambridge, 2004.
48. Kaxiras, E. *Atomic and Electronic Structure of Solids*, Cambridge University Press: Cambridge, 2003.
49. Sholl, D. S.; Steckel, J. A. *Density Functional Theory: A Practical Introduction*, WILEY-VCH Verlag, 2011; p 256. <https://www.wiley.com/en-gb/Density+Functional+Theory%3A+A+Practical+Introduction-p-9781118211045>.
50. Giustino, F. *Materials Modelling Using Density Functional Theory*, Oxford University Press, 2014; p 304.
51. Monkhorst, H. J.; Pack, J. D. Special Points for Brillouin-Zone Integrations. *Phys. Rev. B* **1976**, *13* (12), 5188–5192. <https://doi.org/10.1103/PhysRevB.13.5188>.
52. Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), 864–871. <https://doi.org/10.1103/PhysRev.136.B864>.
53. Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A), 1133–1138. <https://doi.org/10.1103/PhysRev.140.A1133>.
54. Rappe, A. M.; Rabe, K. M.; Kaxiras, E.; Joannopoulos, J. D. Optimized Pseudopotentials. *Phys. Rev. B* **1990**, *41* (2), 1227–1230. <https://doi.org/10.1103/PhysRevB.41.1227>.
55. Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. *Phys. Rev. B* **1991**, *43* (3), 1993–2006. <https://doi.org/10.1103/PhysRevB.43.1993>.
56. Schön, J. C.; Jansen, M. First Step Towards Planning of Syntheses in Solid-State Chemistry: Determination of Promising Structure Candidates by Global Optimization. *Angew. Chem. Int. Ed.* **1996**, *35*, 1286. <https://doi.org/10.1002/anie.199612861>.
57. Pickard, C. J.; Needs, R. J. Ab Initio Random Structure Searching. *J. Phys. Condens. Matter* **2011**, *23* (5), 053201. <https://doi.org/10.1088/0953-8984/23/5/053201>.
58. Sheldrick, G. M. A Short History of SHELX. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2008**, *64* (1), 112–122. <https://doi.org/10.1107/S0108767307043930>.
59. Momma, K.; Izumi, F. VESTA 3 for Three-Dimensional Visualization of Crystal, Volumetric and Morphology Data. *J. Appl. Cryst.* **2011**, *44* (6), 1272–1276. <https://doi.org/10.1107/S0021889811038970>.
60. Mahor, A.; Prasad, V.; Rangnekar, S. Economic Dispatch Using Particle Swarm Optimization: A Review. *Renew. Sustain. Energy Rev.* **2009**, *13* (8), 2134–2141. <https://doi.org/10.1016/j.rser.2009.03.007>.
61. Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. Crystal Structure Prediction Via Particle-Swarm Optimization. *Phys. Rev. B* **2010**, *82* (9), 094116. <https://doi.org/10.1103/PhysRevB.82.094116>.
62. Hartke, B. Global Geometry Optimization of Clusters Using Genetic Algorithms. *J. Phys. Chem.* **1993**, *97* (39), 9973–9976. <https://doi.org/10.1021/j100141a013>.
63. Xiao, Y.; Williams, D. E. Genetic Algorithm: A New Approach to the Prediction of the Structure of Molecular Clusters. *Chem. Phys. Lett.* **1993**, *215* (1–3), 17–24. [https://doi.org/10.1016/0009-2614\(93\)89256-H](https://doi.org/10.1016/0009-2614(93)89256-H).
64. Johnston, R. L. Evolving Better Nanoparticles: Genetic Algorithms for Optimising Cluster Geometries. *Dalton Trans.* **2003**, *22*, 4193–4207. <https://doi.org/10.1039/b305686d>.
65. Bush, T. S.; Catlow, C. R. A.; Battle, P. D. Evolutionary Programming Techniques for Predicting Inorganic Crystal Structures. *J. Mater. Chem.* **1995**, *5* (8), 1269–1272. <https://doi.org/10.1039/jm9950501269>.
66. Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. Modelling Organic Crystal Structures Using Distributed Multipole and Polarizability-Based Model Intermolecular Potentials. *Phys. Chem. Chem. Phys.* **2010**, *12* (30), 8478. <https://doi.org/10.1039/c004164e>.
67. Curtis, F.; Li, X.; Rose, T.; Vázquez-Mayagoitia, A.; Bhattacharya, S.; Ghiringhelli, L. M.; Marom, N. GATOR: A First-Principles Genetic Algorithm for Molecular Crystal Structure Prediction. *J. Chem. Theory Comput.* **2018**, *14* (4), 2246–2264. <https://doi.org/10.1021/acs.jctc.7b01152>.
68. Neumann, M. A.; Leusen, F. J. J.; Kendrick, J. A Major Advance in Crystal Structure Prediction. *Angew. Chem. Int. Ed.* **2008**, *47* (13), 2427–2430. <https://doi.org/10.1002/anie.200704247>.
69. Holden, J. R.; Du, Z.; Ammon, H. L. Prediction of Possible Crystal Structures for C-, H-, N-, O-, and F-Containing Organic Compounds. *J. Comput. Chem.* **1993**, *14* (4), 422–437. <https://doi.org/10.1002/jcc.540140406>.
70. van Eijck, B. P.; Kroon, J. Structure Predictions Allowing More than One Molecule in the Asymmetric Unit. *Acta Crystallogr. Sect. B Struct. Sci.* **2000**, *56* (3), 535–542. <https://doi.org/10.1107/S0108768100000276>.
71. Yamashita, T.; Sato, N.; Kino, H.; Miyake, T.; Tsuda, K.; Oguchi, T. Crystal Structure Prediction Accelerated by Bayesian Optimization. *Phys. Rev. Mater.* **2018**, *2* (1), 013803. <https://doi.org/10.1103/PhysRevMaterials.2.013803>.

72. Tipton, W. W.; Hennig, R. G. A Grand Canonical Genetic Algorithm for the Prediction of Multi-Component Phase Diagrams and Testing of Empirical Potentials. *J. Phys. Condens. Matter* **2013**, *25* (49), 495401. <https://doi.org/10.1088/0953-8984/25/49/495401>.
73. Revard, B. C.; Tipton, W. W.; Hennig, R. G. *Structure and Stability Prediction of Compounds with Evolutionary Algorithms, Prediction and Calculation of Crystal Structures; Topics in Current Chemistry*, 345, Atahan-Evrenk, S., Aspuru-Guzik, A., Eds.; Springer, 2014; pp 181–222.
74. Hajinazar, S.; Thorn, A.; Sandoval, E. D.; Kharabazde, S.; Kolmogorov, A. N. MAISE: Construction of neural network interatomic models and evolutionary structure optimization. *Comput. Phys. Commun.* **2021**, *259*, 107679. <https://doi.org/10.1016/j.cpc.2020.107679>.
75. Lonie, D. C.; Zurek, E. Identifying Duplicate Crystal Structures: Xtal-Comp, an Open-Source Solution. *Comput. Phys. Commun.* **2012**, *183* (3), 690–697. <https://doi.org/10.1016/j.cpc.2011.11.007>.
76. Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. CALYPSO: A Method for Crystal Structure Prediction. *Comput. Phys. Commun.* **2012**, *183* (10), 2063–2070. <https://doi.org/10.1016/j.cpc.2012.05.008>.
77. Valle, M.; Oganov, A. R. Crystal Fingerprint Space – A Novel Paradigm for Studying Crystal-Structure Sets. *Acta Cryst. A: Found. Cryst.* **2010**, *66* (5), 507–517. <https://doi.org/10.1107/S0108767310026395>.
78. Lyakhov, A. O.; Oganov, A. R. Evolutionary Search for Superhard Materials: Methodology and Applications to Forms of Carbon and TiO<sub>2</sub>. *Phys. Rev. B* **2011**, *84* (9), 092103. <https://doi.org/10.1103/PhysRevB.84.092103>.
79. Allmann, R.; Hinek, R. The Introduction of Structure Types into the Inorganic Crystal Structure Database ICSD. *Acta Crystallogr. Sect. A Found. Crystallogr.* **2007**, *63* (5), 412–417. <https://doi.org/10.1107/S0108767307038081>.
80. Morgan, D.; Ceder, G.; Curtarolo, S. High-Throughput and Data Mining with Ab Initio Methods. *Meas. Sci. Technol.* **2005**, *16* (1), 296–301. <https://doi.org/10.1088/0957-0233/16/1/039>.
81. Garrity, K. F.; Bennett, J. W.; Rabe, K. M.; Vanderbilt, D. Pseudopotentials for High-Throughput DFT Calculations. *Comput. Mater. Sci.* **2014**, *81*, 446–452. <https://doi.org/10.1016/j.commatsci.2013.08.053>.
82. Setyawan, W.; Curtarolo, S. High-Throughput Electronic Band Structure Calculations: Challenges and Tools. *Comput. Mater. Sci.* **2010**, *49* (2), 299–312. <https://doi.org/10.1016/j.commatsci.2010.05.010>.
83. Hautier, G.; Fischer, C. C.; Jain, A.; Mueller, T.; Ceder, G. Finding Nature's Missing Ternary Oxide Compounds Using Machine Learning and Density Functional Theory. *Chem. Mater.* **2010**, *22* (12), 3762–3767. <https://doi.org/10.1021/cm100795d>.
84. Hautier, G.; Fischer, C.; Ehrlicher, V.; Jain, A.; Ceder, G. Data Mined Ionic Substitutions for the Discovery of New Compounds. *Inorg. Chem.* **2011**, *50* (2), 656–663. <https://doi.org/10.1021/ic102031h>.
85. Isayev, O.; Oses, C.; Toher, C.; Gossett, E.; Curtarolo, S.; Tropsha, A. Universal Fragment Descriptors for Predicting Properties of Inorganic Crystals. *Nat. Commun.* **2017**, *8*, 15679. <https://doi.org/10.1038/ncomms15679>.
86. Curtarolo, S.; Hart, G. L. W.; Nardelli, M. B.; Mingo, N.; Sanvito, S.; Levy, O. The High-Throughput Highway to Computational Materials Design. *Nat. Mater.* **2013**, *12* (3), 191–201. <https://doi.org/10.1038/nmat3568>.
87. Chung, S. J.; Hahn, T.; Klee, W. E. Nomenclature and Generation of Three-Periodic Nets: The Vector Method. *Acta Crystallogr. Sect. A Found. Crystallogr.* **1984**, *40* (1), 42–50. <https://doi.org/10.1107/S0108767384000088>.
88. Winkler, B.; Pickard, C. J.; Milman, V.; Klee, W. E.; Thimm, G. Prediction of a Nanoporous sp<sup>2</sup>-Carbon Framework Structure by Combining Graph Theory with Quantum Mechanics. *Chem. Phys. Lett.* **1999**, *312* (5–6), 536–541. [https://doi.org/10.1016/S0009-2614\(99\)00943-4](https://doi.org/10.1016/S0009-2614(99)00943-4).
89. Winkler, B.; Pickard, C. J.; Milman, V.; Thimm, G. Systematic Prediction of Crystal Structures. *Chem. Phys. Lett.* **2001**, *337* (1–3), 36–42. [https://doi.org/10.1016/S0009-2614\(01\)00126-9](https://doi.org/10.1016/S0009-2614(01)00126-9).
90. Strong, R. T.; Pickard, C. J.; Milman, V.; Thimm, G.; Winkler, B. Systematic Prediction of Crystal Structures: An Application to sp<sup>3</sup>-Hybridized Carbon Polymorphs. *Phys. Rev. B* **2004**, *70* (4), 045101. <https://doi.org/10.1103/PhysRevB.70.045101>.
91. Shi, X.; He, C.; Pickard, C. J.; Tang, C.; Zhong, J. Stochastic Generation of Complex Crystal Structures Combining Group and Graph Theory with Application to Carbon. *Phys. Rev. B* **2018**, *97* (1), 014104. <https://doi.org/10.1103/PhysRevB.97.014104>.
92. Baburin, I. A.; Proserpio, D. M.; Saleev, V. A.; Shipilova, A. V. From Zeolite Nets to sp<sup>3</sup> Carbon Allotropes: A Topology-Based Multiscale Theoretical Study. *Phys. Chem. Chem. Phys.* **2015**, *17* (2), 1332–1338. <https://doi.org/10.1039/C4CP04569F>.
93. Engel, E. A.; Anelli, A.; Ceriotti, M.; Pickard, C. J.; Needs, R. J. Mapping Uncharted Territory in Ice from Zeolite Networks to Ice Structures. *Nat. Commun.* **2018**, *9* (1), 2173. <https://doi.org/10.1038/s41467-018-04618-6>.
94. Klee, W. Al/Si Distributions in Tectosilicates: A Graph-Theoretical Approach. *Zeitschrift für Krist. - Cryst. Mater.* **1974**, *140* (3–4), 154–162. <https://doi.org/10.1524/zkri-1974-3-403>.
95. Eon, J.-G. Cation Ordering in Aluminophosphates and Aluminosilicates: A Combinatorial and Geometrical Analysis of the Avoidance Rule. *Struct. Chem.* **2016**, *27* (6), 1613–1621. <https://doi.org/10.1007/s11224-016-0770-5>.
96. Laves, F.; Goldsmith, J. R. The Effect of Temperature and Composition on the Al-Si Distribution in Anorthite. *Zeitschrift für Krist. - Cryst. Mater.* **1954**, *106* (1–6), 227–235. <https://doi.org/10.1524/zkri.1954.106.16.227>.
97. Catlow, C. R. A.; George, A. R.; Freeman, C. M. Ab Initio and Molecular-Mechanics Studies of Aluminosilicate Fragments, and the Origin of Lowenstein's Rule. *Chem. Commun.* **1996**, *11*, 1311. <https://doi.org/10.1039/cc960001311>.
98. Bell, R. G.; Jackson, R. A.; Catlow, C. R. A. Löwenstein's Rule in Zeolite a: A Computational Study. *Zeolites* **1992**, *12* (7), 870–871. [https://doi.org/10.1016/0144-2449\(92\)90065-W](https://doi.org/10.1016/0144-2449(92)90065-W).
99. Stishov, S. M. Memoir on the Discovery of High-Density Silica. *High Press. Res.* **1995**, *13* (5), 245–280. <https://doi.org/10.1080/08957959508200890>.
100. Hoffmann, R.; Kabanov, A. A.; Golov, A. A.; Proserpio, D. M. Homo Citans and Carbon Allotropes: For an Ethics of Citation. *Angew. Chem. Int. Ed.* **2016**, *55* (37), 10962–10976. <https://doi.org/10.1002/anie.201600655>.
101. Grochala, W. Diamond: Electronic Ground State of Carbon at Temperatures Approaching 0 K. *Angew. Chem. Int. Ed.* **2014**, *53* (14), 3680–3683. <https://doi.org/10.1002/anie.201400131>.
102. Aust, R. B.; Drickamer, H. G. Carbon: A New Crystalline Phase. *Science* **1963**, *140* (3568), 817–819. <https://doi.org/10.1126/science.140.3568.817>.
103. Bundy, F. P.; Kasper, J. S. Hexagonal Diamond—A New Form of Carbon. *J. Chem. Phys.* **1967**, *46* (9), 3437–3446. <https://doi.org/10.1063/1.1841236>.
104. Hanfland, M.; Syassen, K.; Sonnenschein, R. Optical Reflectivity of Graphite under Pressure. *Phys. Rev. B* **1989**, *40* (3), 1951–1954. <https://doi.org/10.1103/PhysRevB.40.1951>.
105. Zhao, Y. X.; Spain, I. L. X-Ray Diffraction Data for Graphite to 20 GPa. *Phys. Rev. B* **1989**, *40* (2), 993–997. <https://doi.org/10.1103/PhysRevB.40.993>.
106. Mao, W. L.; Mao, H. K.; Eng, P. J.; Trainor, T. P.; Newville, M.; Kao, C. C.; Heinz, D. L.; Shu, J.; Meng, Y.; Hemley, R. J. Bonding Changes in Compressed Superhard Graphite. *Science* **2003**, *302* (5644), 425–427. <https://doi.org/10.1126/science.1089713>.
107. Sundqvist, B. Carbon Under Pressure. *Phys. Rep.* **2021**, *909*, 1–73. <https://doi.org/10.1016/j.physrep.2020.12.007>.
108. Li, Q.; Ma, Y.; Oganov, A.; Wang, H.; Wang, H.; Xu, Y.; Cui, T.; Mao, H.-K.; Zou, G. Superhard Monoclinic Polymorph of Carbon. *Phys. Rev. Lett.* **2009**, *102* (17), 175506. <https://doi.org/10.1103/PhysRevLett.102.175506>.
109. Umemoto, K.; Wentzcovitch, R. M.; Saito, S.; Miyake, T. Body-Centered Tetragonal C<sub>4</sub>: Viable sp<sup>3</sup> Carbon Allotrope. *Phys. Rev. Lett.* **2010**, *104* (12), 125504. <https://doi.org/10.1103/PhysRevLett.104.125504>.
110. Wang, J.-T.; Chen, C.; Kawazoe, Y. Low-Temperature Phase Transformation from Graphite to sp<sup>3</sup> Orthorhombic Carbon. *Phys. Rev. Lett.* **2011**, *106* (7), 075501. <https://doi.org/10.1103/PhysRevLett.106.075501>.

111. Selli, D.; Baburin, I. A.; Martoňák, R.; Leoni, S. Superhard  $sp^3$  Carbon Allotropes with Odd and Even Ring Topologies. *Phys. Rev. B* **2011**, *84* (16), 161411. <https://doi.org/10.1103/PhysRevB.84.161411>.
112. Zhao, Z.; Xu, B.; Zhou, X.-F.; Wang, L.-M.; Wen, B.; He, J.; Liu, Z.; Wang, H.-T.; Tian, Y. Novel Superhard Carbon: C-Centered Orthorhombic  $C_8$ . *Phys. Rev. Lett.* **2011**, *107* (21), 215502. <https://doi.org/10.1103/PhysRevLett.107.215502>.
113. Bouffeffel, S. E.; Oganov, A. R.; Leoni, S. Understanding the Nature of "Superhard Graphite". *Sci. Rep.* **2012**, *2* (1), 471 <https://doi.org/10.1038/srep00471>.
114. Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (20), 12562–12566. <https://doi.org/10.1073/pnas.202427399>.
115. Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100* (2), 020603. <https://doi.org/10.1103/PhysRevLett.100.020603>.
116. Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. *Annu. Rev. Phys. Chem.* **2016**, *67*, 159–184. <https://doi.org/10.1146/annurev-physchem-040215-112229>.
117. Yang, Y. I.; Shao, Q.; Zhang, J.; Yang, L.; Gao, Y. Q. Enhanced Sampling in Molecular Dynamics. *J. Chem. Phys.* **2019**, *151* (7), 070902. <https://doi.org/10.1063/1.5109531>.
118. Valsson, O.; Parrinello, M. Variational Approach to Enhanced Sampling and Free Energy Calculations. *Phys. Rev. Lett.* **2014**, *113* (9), 090601. <https://doi.org/10.1103/PhysRevLett.113.090601>.
119. Bonati, L.; Zhang, Y. Y.; Parrinello, M. Neural Networks-Based Variationally Enhanced Sampling. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (36), 17641–17647. <https://doi.org/10.1073/pnas.1907975116>.
120. Invernizzi, M.; Parrinello, M. Rethinking Metadynamics: From Bias Potentials to Probability Distributions. *J. Phys. Chem. Lett.* **2020**, *11* (7), 2731–2736. <https://doi.org/10.1021/acs.jpclett.0c00497>.
121. Invernizzi, M.; Piaggi, P. M.; Parrinello, M. Unified Approach to Enhanced Sampling. *Phys. Rev. X* **2020**, *10* (4), 041034. <https://doi.org/10.1103/PhysRevX.10.041034>.
122. Martoňák, R.; Donadio, D.; Oganov, A. R.; Parrinello, M. Crystal Structure Transformations in  $SiO_2$  from Classical and Ab Initio Metadynamics. *Nat. Mater.* **2006**, *5* (8), 623–626. <https://doi.org/10.1038/nmat1696>.
123. Behler, J.; Martoňák, R.; Donadio, D.; Parrinello, M. Metadynamics Simulations of the High-Pressure Phases of Silicon Employing a High-Dimensional Neural Network Potential. *Phys. Rev. Lett.* **2008**, *100* (18), 185501. <https://doi.org/10.1103/PhysRevLett.100.185501>.
124. Goedecker, S. Minima Hopping: An Efficient Search Method for the Global Minimum of the Potential Energy Surface of Complex Molecular Systems. *J. Chem. Phys.* **2004**, *120* (21), 9911–9917. <https://doi.org/10.1063/1.1724816>.
125. Amsler, M.; Goedecker, S. Crystal Structure Prediction Using the Minima Hopping Method. *J. Chem. Phys.* **2010**, *133* (22), 224104. <https://doi.org/10.1063/1.3512900>.
126. Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101* (28), 5111–5116. <https://doi.org/10.1021/jp970984n>.
127. Li, Z.; Scheraga, H. A. Monte Carlo-Minimization Approach to the Multiple-Minima Problem in Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84* (19), 6611–6615. <https://doi.org/10.1073/pnas.84.19.6611>.
128. Derzsi, M.; Piekarz, P.; Grochala, W. Structures of Late Transition Metal Monoxides from Jahn-Teller Instabilities in the Rock Salt Lattice. *Phys. Rev. Lett.* **2014**, *113* (2), 025505. <https://doi.org/10.1103/PhysRevLett.113.025505>.
129. Ghosez, P.; Cockayne, E.; Waghmare, U. V.; Rabe, K. M. Lattice Dynamics of  $BaTiO_3$ ,  $PbTiO_3$  and  $PbZrO_3$ : A Comparative First-Principles Study. *Phys. Rev. B* **1999**, *60* (2), 836–843. <https://doi.org/10.1103/PhysRevB.60.836>.
130. Parlinski, K.; Kawazoe, Y. Ab Initio Study of Phonons and Structural Stabilities of the Perovskite-Type  $MgSiO_3$ . *Eur. Phys. J. B* **2000**, *16* (1), 49–58. <https://doi.org/10.1007/s100510070248>.
131. Baettig, P.; Schelle, C. F.; LeSar, R.; Waghmare, U. V.; Spaldin, N. A. Theoretical Prediction of New High-Performance Lead-Free Piezo-Electrics. *Chem. Mater.* **2005**, *17* (6), 1376–1380. <https://doi.org/10.1021/cm0480418>.
132. Derzsi, M.; Hermann, A.; Hoffmann, R.; Grochala, W. The Close Relationships between the Crystal Structures of  $MO$  and  $MSO_4$  ( $M =$  Group 10, 11, or 12 Metal), and the Predicted Structures of  $AuO$  and  $PtSO_4$ . *Eur. J. Inorg. Chem.* **2013**, *29*, 5094–5102. <https://doi.org/10.1002/ejic.201300769>.
133. Monserrat, B.; Drummond, N. D.; Dalladay-Simpson, P.; Howie, R. T.; López Ríos, P.; Gregoryanz, E.; Pickard, C. J.; Needs, R. J. Structure and Metallicity of Phase  $v$  of Hydrogen. *Phys. Rev. Lett.* **2018**, *120* (25), 255701. <https://doi.org/10.1103/PhysRevLett.120.255701>.
134. Lines, M. E.; Glass, A. M. *Principles and Applications of Ferroelectrics and Related Materials*, Oxford University Press: Oxford, UK, 2001. <https://doi.org/10.1093/acprof:oso/9780198507789.001.0001>.
135. Ramprasad, R.; Batra, R.; Pilania, G.; Mannodi-Kanakkithodi, A.; Kim, C. Machine Learning in Materials Informatics: Recent Applications and Prospects. *npj Comput. Mater.* **2017**, *3* (1), 54. <https://doi.org/10.1038/s41524-017-0056-5>.
136. Bligaard, T.; Jóhannesson, G. H.; Ruban, A. V.; Skriver, H. L.; Jacobsen, K. W.; Nørskov, J. K. Pareto-Optimal Alloys. *Appl. Phys. Lett.* **2003**, *83* (22), 4527–4529. <https://doi.org/10.1063/1.1631051>.
137. Greeley, J.; Jaramillo, T. F.; Bonde, J.; Chorkendorff, I.; Nørskov, J. K. Computational High-Throughput Screening of Electrocatalytic Materials for Hydrogen Evolution. *Nat. Mater.* **2006**, *5* (11), 909–913. <https://doi.org/10.1038/nmat1752>.
138. Lejaeghere, K.; Cottener, S.; Van Speybroeck, V. Ranking the Stars: A Refined Pareto Approach to Computational Materials Design. *Phys. Rev. Lett.* **2013**, *111* (7), 075501. <https://doi.org/10.1103/PhysRevLett.111.075501>.
139. Ward, L.; Michel, K.; Wolverton, C. Automated Crystal Structure Solution from Powder Diffraction Data: Validation of the First-Principles-Assisted Structure Solution Method. *Phys. Rev. Mater.* **2017**, *1* (6), 063802. <https://doi.org/10.1103/PhysRevMaterials.1.063802>.
140. Gao, P.; Tong, Q.; Lv, J.; Wang, Y.; Ma, Y. X-Ray Diffraction Data-Assisted Structure Searches. *Comput. Phys. Commun.* **2017**, *213*, 40–45. <https://doi.org/10.1016/j.cpc.2016.11.007>.
141. Tsujimoto, N.; Adachi, D.; Akashi, R.; Todo, S.; Tsuneyuki, S. Crystal Structure Prediction Supported by Incomplete Experimental Data. *Phys. Rev. Mater.* **2018**, *2* (5), 053801. <https://doi.org/10.1103/PhysRevMaterials.2.053801>.
142. Fortes, A. D.; Suard, E.; Lemée-Cailleau, M. H.; Pickard, C. J.; Needs, R. J. Crystal Structure of Ammonia Monohydrate Phase II. *J. Am. Chem. Soc.* **2009**, *131* (37), 13508–13515. <https://doi.org/10.1021/ja9052569>.
143. Dewaele, A.; Worth, N.; Pickard, C. J.; Needs, R. J.; Pascarelli, S.; Mathon, O.; Mezouar, M.; Irifune, T. Synthesis and Stability of Xenon Oxides  $Xe_2O_5$  and  $Xe_3O_2$  under Pressure. *Nat. Chem.* **2016**, *8* (8), 784–790. <https://doi.org/10.1038/nchem.2528>.
144. Teter, D. M. Computational Alchemy: The Search for New Superhard Materials. *MRS Bull.* **1998**, *23* (1), 22–27. <https://doi.org/10.1557/S0883769400031420>.
145. Chen, X. Q.; Niu, H.; Li, D.; Li, Y. Modeling Hardness of Polycrystalline Materials and Bulk Metallic Glasses. *Intermetallics* **2011**, *19* (9), 1275–1281. <https://doi.org/10.1016/j.intermet.2011.03.026>.
146. Phillips, J. C. Ionicity of the Chemical Bond in Crystals. *Rev. Mod. Phys.* **1970**, *42* (3), 317–356. <https://doi.org/10.1103/RevModPhys.42.317>.
147. Cohen, M. L. Calculation of Bulk Moduli of Diamond and Zinc-Blende Solids. *Phys. Rev. B* **1985**, *32* (12), 7988–7991. <https://doi.org/10.1103/PhysRevB.32.7988>.
148. Gao, F.; He, J.; Wu, E.; Liu, S.; Yu, D.; Li, D.; Zhang, S.; Tian, Y. Hardness of Covalent Crystals. *Phys. Rev. Lett.* **2003**, *91* (1), 015502. <https://doi.org/10.1103/PhysRevLett.91.015502>.
149. Šimunek, A.; Vackář, J. Hardness of Covalent and Ionic Crystals: First-Principle Calculations. *Phys. Rev. Lett.* **2006**, *96* (8), 085501. <https://doi.org/10.1103/PhysRevLett.96.085501>.
150. Šimunek, A. How to Estimate Hardness of Crystals on a Pocket Calculator. *Phys. Rev. B* **2007**, *75* (17), 172108. <https://doi.org/10.1103/PhysRevB.75.172108>.

151. Li, K.; Wang, X.; Zhang, F.; Xue, D. Electronegativity Identification of Novel Superhard Materials. *Phys. Rev. Lett.* **2008**, *100* (23), 235504. <https://doi.org/10.1103/PhysRevLett.100.235504>.
152. Zhang, X.; Wang, Y.; Lv, J.; Li, Q.; Zhu, C.; Li, Q.; Zhang, M.; Ma, Y. First-Principles Structural Design of Superhard Materials. *J. Chem. Phys.* **2013**, *138* (11), 114101. <https://doi.org/10.1063/1.4794424>.
153. Avery, P.; Toher, C.; Curtarolo, S.; Zurek, E. XTALOPT Version r12: An Open-Source Evolutionary Algorithm for Crystal Structure Prediction. *Comput. Phys. Commun.* **2019**, *237*, 274–275. <https://doi.org/10.1016/j.cpc.2018.11.016>.
154. Shockley, W.; Queisser, H. J. Detailed Balance Limit of Efficiency of P-N Junction Solar Cells. *J. Appl. Phys.* **1961**, *32* (3), 510–519. <https://doi.org/10.1063/1.1736034>.
155. Glutsch, S. *Excitons in Low-Dimensional Semiconductors. Springer Series in Solid-State Sciences*, vol. 141; Springer: Berlin, Heidelberg, 2004; p 298. <https://doi.org/10.1007/978-3-662-07150-2>.
156. Mueller, T.; Malic, E. Exciton Physics and Device Application of Two-Dimensional Transition Metal Dichalcogenide Semiconductors. *npj 2D Mater. Appl.* **2018**, *2* (1), 29. <https://doi.org/10.1038/s41699-018-0074-2>.
157. Onida, G.; Reining, L.; Rubio, A. Electronic Excitations: Density-Functional Versus Many-Body Green's-Function Approaches. *Rev. Mod. Phys.* **2002**, *74* (2), 601–659. <https://doi.org/10.1103/RevModPhys.74.601>.
158. Godby, R. W.; Schlüter, M.; Sham, L. J. Accurate Exchange-Correlation Potential for Silicon and Its Discontinuity on Addition of an Electron. *Phys. Rev. Lett.* **1986**, *56*, 2415–2418. <https://doi.org/10.1103/PhysRevLett.56.2415>.
159. Borlido, P.; Aull, T.; Huran, A. W.; Tran, F.; Marques, M. A. L.; Botti, S. Large-Scale Benchmark of Exchange-Correlation Functionals for the Determination of Electronic Band Gaps of Solids. *J. Chem. Theory Comput.* **2019**, *15* (9), 5069–5079. <https://doi.org/10.1021/acs.jctc.9b00322>.
160. Marsiglio, F.; Carbotte, J. P. Electron-Phonon Superconductivity. In *Superconductivity*; Benemann, K. H., Ketterson, J. B., Eds.; Springer: Berlin, Heidelberg, 2008; pp 73–162. [https://doi.org/10.1007/978-3-540-73253-2\\_3](https://doi.org/10.1007/978-3-540-73253-2_3). [http://link.springer.com/10.1007/978-3-540-73253-2\\_3](http://link.springer.com/10.1007/978-3-540-73253-2_3).
161. Flores-Livas, J. A.; Boeri, L.; Sanna, A.; Profeta, G.; Arita, R.; Eremets, M. A perspective on conventional high-temperature superconductors at high pressure: Methods and materials. *Phys. Rep.* **2020**, 1–78. <https://doi.org/10.1016/j.physrep.2020.02.003>.
162. Hopfield, J. J. Angular Momentum and Transition-Metal Superconductivity. *Phys. Rev.* **1969**, *186* (2), 443–451. <https://doi.org/10.1103/PhysRev.186.443>.
163. Gaspari, G. D.; Gyorffy, B. L. Electron-Phonon Interactions, *d* Resonances, and Superconductivity in Transition Metals. *Phys. Rev. Lett.* **1972**, *28* (13), 801–805. <https://doi.org/10.1103/PhysRevLett.28.801>.
164. Papaconstantopoulos, D. A.; Klein, B. M.; Mehl, M. J.; Pickett, W. E. Cubic H<sub>3</sub>S around 200 GPa: An Atomic Hydrogen Superconductor Stabilized by Sulfur. *Phys. Rev. B* **2015**, *91* (18), 1–5. <https://doi.org/10.1103/PhysRevB.91.184511>.
165. Shipley, A. M.; Hutcheon, M. J.; Needs, R. J.; Pickard, C. J. High-Throughput Discovery of High-Temperature Conventional Superconductors. *Phys. Rev. B* **2021**, *104* (5), 054501. <https://doi.org/10.1103/PhysRevB.104.054501>.
166. Dye, J. L. Electrides: Early Examples of Quantum Confinement. *Acc. Chem. Res.* **2009**, *42* (10), 1564–1572. <https://doi.org/10.1021/ar9000857>.
167. Pickard, C. J.; Needs, R. J. Predicted Pressure-Induced *s*-Band Ferromagnetism in Alkali Metals. *Phys. Rev. Lett.* **2011**, *107* (8), 087201. <https://doi.org/10.1103/PhysRevLett.107.087201>.
168. Inoshita, T.; Hamada, N.; Hosono, H. Ferromagnetic Instability of Interlayer Floating Electrons in the Quasi-Two-Dimensional Electride Y<sub>2</sub>C. *Phys. Rev. B* **2015**, *92* (20), 201109. <https://doi.org/10.1103/PhysRevB.92.201109>.
169. Lee, S. Y.; Hwang, J. Y.; Park, J.; Nandadasa, C. N.; Kim, Y.; Bang, J.; Lee, K.; Lee, K. H.; Zhang, Y.; Ma, Y.; Hosono, H.; Lee, Y. H.; Kim, S. G.; Kim, S. W. Ferromagnetic Quasi-Atomic Electrons in Two-Dimensional Electride. *Nat. Commun.* **2020**, *11* (1), 1526. <https://doi.org/10.1038/s41467-020-15253-5>.
170. Kitano, M.; Inoue, Y.; Yamazaki, Y.; Hayashi, F.; Kanbara, S.; Matsuishi, S.; Yokoyama, T.; Kim, S. W.; Hara, M.; Hosono, H. Ammonia Synthesis Using a Stable Electride as an Electron Donor and Reversible Hydrogen Store. *Nat. Chem.* **2012**, *4* (11), 934–940. <https://doi.org/10.1038/nchem.1476>.
171. Kim, Y. J.; Kim, S. M.; Cho, E. J.; Hosono, H.; Yang, J. W.; Kim, S. W. Two Dimensional Inorganic Electride-Promoted Electron Transfer Efficiency in Transfer Hydrogenation of Alkynes and Alkenes. *Chem. Sci.* **2015**, *6* (6), 3577–3581. <https://doi.org/10.1039/c5sc00933b>.
172. Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*, Oxford University Press, 1994.
173. Savin, A.; Jepsen, O.; Flad, J.; Andersen, O. K.; Preuss, H.; von Schnering, H. G. Electron Localization in Solid-State Structures of the Elements: The Diamond Structure. *Angew. Chem. Int. Ed. Engl.* **1992**, *31* (2), 187–188. <https://doi.org/10.1002/anie.199201871>.
174. Zhang, Y.; Wang, H.; Wang, Y.; Zhang, L.; Ma, Y. Computer-Assisted Inverse Design of Inorganic Electrides. *Phys. Rev. X* **2017**, *7* (1), 011017. <https://doi.org/10.1103/PhysRevX.7.011017>.
175. Madsen, G. K. H.; Singh, D. J. BoltzTraP. A code for calculating band-structure dependent quantities. *Comput. Phys. Commun.* **2006**, *175* (1), 67–71. <https://doi.org/10.1016/j.cpc.2006.03.007>.
176. Yan, J.; Gorai, P.; Ortiz, B.; Miller, S.; Barnett, S. A.; Mason, T.; Stevanović, V.; Toberer, E. S. Material Descriptors for Predicting Thermoelectric Performance. *Energ. Environ. Sci.* **2015**, *8* (3), 983–994. <https://doi.org/10.1039/c4ee03157a>.
177. Wang, S.; Wang, Z.; Setyawati, W.; Mingo, N.; Curtarolo, S. Assessing the Thermoelectric Properties of Sintered Compounds Via High-Throughput Ab-Initio Calculations. *Phys. Rev. X* **2011**, *1* (2), 021012. <https://doi.org/10.1103/PhysRevX.1.021012>.
178. Gorai, P.; Parrilla, P.; Toberer, E. S.; Stevanović, V. Computational Exploration of the Binary A<sub>3</sub>B<sub>3</sub> Chemical Space for Thermoelectric Performance. *Chem. Mater.* **2015**, *27* (18), 6213–6221. <https://doi.org/10.1021/acs.chemmater.5b01179>.
179. Núñez-Valdez, M.; Allahyari, Z.; Fan, T.; Oganov, A. R. Efficient Technique for Computational Design of Thermoelectric Materials. *Comput. Phys. Commun.* **2018**, *222*, 152–157. <https://doi.org/10.1016/j.cpc.2017.10.001>.
180. Yan, S.; Wang, Y.; Gao, Z.; Long, Y.; Ren, J. Directional Design of Materials Based on Multi-Objective Optimization: A Case Study of Two-Dimensional Thermoelectric SnSe. *Chinese Phys. Lett.* **2021**, *38* (2), 027301. <https://doi.org/10.1088/0256-307X/38/2/027301>.
181. Oganov, A. R.; Chen, J.; Gattai, C.; Ma, Y.; Ma, Y.; Glass, C. W.; Liu, Z.; Yu, T.; Kurakevich, O. O.; Solozhenko, V. L. Ionic High-Pressure Form of Elemental Boron. *Nature* **2009**, *457* (7231), 863–867. <https://doi.org/10.1038/nature07736>.
182. Ahnert, S. E.; Grant, W. P.; Pickard, C. J. Revealing and Exploiting Hierarchical Material Structure through Complex Atomic Networks. *npj Comput. Mater.* **2017**, *3* (1), 35. <https://doi.org/10.1038/s41524-017-0035-x>.
183. Deringer, V. L.; Proserpio, D. M.; Csányi, G.; Pickard, C. J. Data-Driven Learning and Prediction of Inorganic Crystal Structures. *Faraday Discuss.* **2018**, *211*, 45–59. <https://doi.org/10.1039/C8FD00034D>.
184. Witt, W. C.; Shires, B. W. B.; Tan, C. W.; Jankowski, W. J.; Pickard, C. J. Random Structure Searching with Orbital-Free Density Functional Theory. *J. Phys. Chem. A* **2021**, *125* (7), 1650–1660. <https://doi.org/10.1021/acs.jpca.0c11030>.
185. Xu, Q.; Ma, C.; Mi, W.; Wang, Y.; Ma, Y. Nonlocal Pseudopotential Energy Density Functional for Orbital-Free Density Functional Theory. *Nat. Commun.* **2022**, *13* (1), 1385. <https://doi.org/10.1038/s41467-022-29002-3>.
186. Biswas, R.; Hamann, D. R. Interatomic Potentials for Silicon Structural Energies. *Phys. Rev. Lett.* **1985**, *55* (19), 2001–2004. <https://doi.org/10.1103/PhysRevLett.55.2001>.
187. Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 146401. <https://doi.org/10.1103/PhysRevLett.98.146401>.
188. Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104* (13), 136403. <https://doi.org/10.1103/PhysRevLett.104.136403>.
189. Musil, F.; Grisafi, A.; Bartók, A. P.; Ortner, C.; Csányi, G.; Ceriotti, M. Physics-Inspired Structural Representations for Molecules and Materials. *Chem. Rev.* **2021**, *121* (16), 9759–9815. <https://doi.org/10.1021/acs.chemrev.1c00021>.

190. Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Model. Simul.* **2016**, *14* (3), 1153–1173. <https://doi.org/10.1137/15M1054183>.
191. Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B* **2013**, *87* (18), 184115. <https://doi.org/10.1103/PhysRevB.87.184115>.
192. Bartók, A. P.; Kermode, J.; Bernstein, N.; Csányi, G. Machine Learning a General-Purpose Interatomic Potential for Silicon. *Phys. Rev. X* **2018**, *8* (4), 041048. <https://doi.org/10.1103/PhysRevX.8.041048>.
193. Deringer, V. L.; Caro, M. A.; Csányi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv. Mater.* **2019**, *31* (46), 1902765. <https://doi.org/10.1002/adma.201902765>.
194. Tong, Q.; Xue, L.; Lv, J.; Wang, Y.; Ma, Y. Accelerating CALYPSO Structure Prediction by Data-Driven Learning of a Potential Energy Surface. *Faraday Discuss.* **2018**, *211*, 31–43. <https://doi.org/10.1039/C8FD00055G>.
195. Dolgirev, P. E.; Kruglov, I. A.; Oganov, A. R. Machine Learning Scheme for Fast Extraction of Chemically Interpretable Interatomic Potentials. *AIP Adv.* **2016**, *6* (8), 085318. <https://doi.org/10.1063/1.4961886>.
196. Pickard, C. J. Ephemeral Data Derived Potentials for Random Structure Search. *Phys. Rev. B* **2022**, *106* (1), 014102. <https://doi.org/10.1103/PhysRevB.106.014102>.
197. Deringer, V. L.; Csányi, G.; Proserpio, D. M. Extracting Crystal Chemistry from Amorphous Carbon Structures. *ChemPhysChem* **2017**, *18* (8), 873–877. <https://doi.org/10.1002/cphc.201700151>.
198. Oganov, A. R.; Valle, M. How to Quantify Energy Landscapes of Solids. *J. Chem. Phys.* **2009**, *130* (10), 104504. <https://doi.org/10.1063/1.3079326>.
199. Lyakhov, A. O.; Oganov, A. R.; Valle, M. How to Predict Very Large and Complex Crystal Structures. *Comput. Phys. Commun.* **2010**, *181* (9), 1623–1632. <https://doi.org/10.1016/j.cpc.2010.06.007>.
200. De, S.; Bartók, A. P.; Csányi, G.; Ceriotti, M. Comparing Molecules and Solids across Structural and Alchemical Space. *Phys. Chem. Chem. Phys.* **2016**, *18* (20), 13754–13769. <https://doi.org/10.1039/C6CP00415F>.
201. Ceriotti, M.; Tribello, G. A.; Parrinello, M. Simplifying the Representation of Complex Free-Energy Landscapes Using Sketch-Map. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (32), 13023–13028. <https://doi.org/10.1073/pnas.1108486108>.
202. Smeeton, L. C.; Oakley, M. T.; Johnston, R. L. Visualizing Energy Landscapes with Metric Disconnectivity Graphs. *J. Comput. Chem.* **2014**, *35* (20), 1481–1490. <https://doi.org/10.1002/jcc.23643>.
203. Shires, B. W. B.; Pickard, C. J. Visualising Energy Landscapes through Manifold Learning. *Phys. Rev. X* **2021**, *11* (4), 041026. <https://doi.org/10.1103/PhysRevX.11.041026>.
204. van der Maaten, L.; Hinton, G. Visualizing Data Using t-SNE. *J. Mach. Learn. Res.* **2008**, *9* (86), 2579–2605.
205. McInnes, L.; Healy, J.; Saul, N.; Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **2018**, *3* (29), 861. <https://doi.org/10.21105/joss.00861>.
206. Anderson, P. W.; Lee, P. A.; Randeria, M.; Rice, T. M.; Trivedi, N.; Zhang, F. C. The Physics behind High-Temperature Superconducting Cuprates: The Plain Vanilla Version of RVB. *J. Phys. Condens. Matter* **2004**, *16* (24), 755–769. <https://doi.org/10.1088/0953-8984/16/24/R02>.
207. Sun, Y.; Lv, J.; Xie, Y.; Liu, H.; Ma, Y. Route to a Superconducting Phase above Room Temperature in Electron-Doped Hydride Compounds under High Pressure. *Phys. Rev. Lett.* **2019**, *123* (9), 097001. <https://doi.org/10.1103/PhysRevLett.123.097001>.
208. Di Cataldo, S.; Heil, C.; von der Linden, W.; Boeri, L. LaBHg: Towards High- $T_c$  Low-Pressure Superconductivity in Ternary Superhydrides. *Phys. Rev. B* **2021**, *104* (2), L020511. <https://doi.org/10.1103/PhysRevB.104.L020511>.
209. Zhang, Z.; Cui, T.; Hutcheon, M. J.; Shipley, A. M.; Song, H.; Du, M.; Kresin, V. Z.; Duan, D.; Pickard, C. J.; Yao, Y. Design Principles for High-Temperature Superconductors with a Hydrogen-Based Alloy Back-Bone at Moderate Pressure. *Phys. Rev. Lett.* **2022**, *128* (4), 047001. <https://doi.org/10.1103/physrevlett.128.047001>.
210. Lucrezi, R.; Di Cataldo, S.; von der Linden, W.; Boeri, L.; Heil, C. In-Silico Synthesis of Lowest-Pressure High- $T_c$  Ternary Superhydrides. *npj Comput. Mater.* **2022**, *8* (1), 119. <https://doi.org/10.1038/s41524-022-00801-y>.
211. Semenok, D. V.; et al. Superconductivity at 253K in Lanthanum–Yttrium Ternary Hydrides. *Mater. Today* **2020**, *48*, 18–28. <https://doi.org/10.1016/j.mattod.2021.03.025>.
212. Booth, S. G.; Nedoma, A. J.; Anthonisamy, N. N.; Baker, P. J.; Boston, R.; Bronstein, H.; Clarke, S. J.; Cussen, E. J.; Daramalla, V.; De Volder, M.; Dutton, S. E.; Falkowski, V.; Fleck, N. A.; Geddes, H. S.; Gollapally, N.; Goodwin, A. L.; Griffin, J. M.; Haworth, A. R.; Hayward, M. A.; Hull, S.; Inkson, B. J.; Johnston, B. J.; Lu, Z.; MacManus-Driscoll, J. L.; Martínez De Luján Labalde, X.; McClelland, I.; McCombie, K.; Murdoch, B.; Nayak, D.; Park, S.; Pérez, G. E.; Pickard, C. J.; Piper, L. F. J.; Playford, H. Y.; Price, S.; Scanlon, D. O.; Stallard, J. C.; Tapia-Ruiz, N.; West, A. R.; Wheatcroft, L.; Wilson, M.; Zhang, L.; Zhi, X.; Zhu, B.; Cussen, S. A. Perspectives for Next Generation Lithium-Ion Battery Cathode Materials. *APL Mater.* **2021**, *9* (10), 109201. <https://doi.org/10.1063/5.0051092>.
213. Mizushima, K.; Jones, P.; Wiseman, P.; Goodenough, J. Li<sub>x</sub>CoO<sub>2</sub> (0 < x ≤ 1): A new cathode material for batteries of high energy density. *Solid State Ion.* **1981**, *3–4*, 171–174. [https://doi.org/10.1016/0167-2738\(81\)90077-1](https://doi.org/10.1016/0167-2738(81)90077-1).
214. Ceder, G.; Chiang, Y.-M.; Sadoway, D. R.; Aydinol, M. K.; Jang, Y.-I.; Huang, B. Identification of Cathode Materials for Lithium Batteries Guided by First-Principles Calculations. *Nature* **1998**, *392* (6677), 694–696. <https://doi.org/10.1038/33647>.
215. Pickard, C. J. AIRSS Data for Carbon at 10GPa and the C+N+H+O System at 1GPa, 2020; <https://doi.org/10.24435/materialscloud:2020.0026/v1>. <https://archive.materialscloud.org/record/2020.0026/v1>.
216. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem 2019 Update: Improved Access to Chemical Data. *Nucleic Acids Res.* **2019**, *47* (D1), 1102–1109. <https://doi.org/10.1093/nar/gky1033>.
217. Conway, L. J.; Pickard, C. J.; Hermann, A. Rules of Formation of H–C–N–O Compounds at High Pressure and the Fates of Planetary Ices. *Proc. Natl. Acad. Sci.* **2021**, *118* (19), 2026360118. <https://doi.org/10.1073/pnas.2026360118>.
218. Naumova, A. S.; Lepeshkin, S. V.; Bushlanov, P. V.; Oganov, A. R. Unusual Chemistry of the C–H–N–O System under Pressure and Implications for Giant Planets. *J. Phys. Chem. A* **2021**, *125* (18), 3936–3942. <https://doi.org/10.1021/acs.jpca.1c00591>.
219. Grochala, W.; Hoffmann, R.; Feng, J.; Ashcroft, N. W. The Chemical Imagination at Work in Very Tight Places. *Angew. Chem. Int. Ed.* **2007**, *46* (20), 3620–3642. <https://doi.org/10.1002/anie.200602485>.
220. Hermann, A. Chemical Bonding at High Pressure. In *Reviews in Computational Chemistry*; Parrill, A. L., Lipkowitz, K. B., Eds., John Wiley & Sons Inc.: Hoboken, NJ, 2017; pp 1–41. <https://doi.org/10.1002/9781119356059.ch1>.
221. Zhang, L.; Wang, Y.; Lv, J.; Ma, Y. Materials Discovery at High Pressures. *Nat. Rev. Mater.* **2017**, *2* (4), 17005. <https://doi.org/10.1038/natrevmats.2017.5>.
222. Ashcroft, N. W. Metallic Hydrogen: A High-Temperature Superconductor? *Phys. Rev. Lett.* **1968**, *21* (26), 1748–1749. <https://doi.org/10.1103/PhysRevLett.21.1748>.
223. Mao, H.-K.; Hemley, R. J. Ultrahigh-Pressure Transitions in Solid Hydrogen. *Rev. Mod. Phys.* **1994**, *66* (2), 671–692. <https://doi.org/10.1103/RevModPhys.66.671>.
224. Pickard, C. J.; Martínez-Canales, M.; Needs, R. J. Density Functional Theory Study of Phase IV of Solid Hydrogen. *Phys. Rev. B* **2012**, *85* (21), 214114. <https://doi.org/10.1103/PhysRevB.85.214114>.
225. Woolman, G.; Naden Robinson, V.; Marqués, M.; Loa, I.; Ackland, G. J.; Hermann, A. Structural and Electronic Properties of the Alkali Metal Incommensurate Phases. *Phys. Rev. Mater.* **2018**, *2* (5), 053604. <https://doi.org/10.1103/PhysRevMaterials.2.053604>.
226. Pickard, C. J.; Needs, R. J. Aluminium at Terapascal Pressures. *Nat. Mater.* **2010**, *9* (8), 624–627. <https://doi.org/10.1038/nmat2796>.
227. Ninet, S.; Datchi, F.; Dumas, P.; Mezour, M.; Garbarino, G.; Mafey, A.; Pickard, C. J.; Needs, R. J.; Saitta, A. M. Experimental and Theoretical Evidence for an Ionic Crystal of Ammonia at High Pressure. *Phys. Rev. B* **2014**, *89* (17), 174103. <https://doi.org/10.1103/PhysRevB.89.174103>.
228. Palasyuk, T.; Troyan, I.; Eremets, M.; Drozd, V.; Medvedev, S.; Zaleski-Egjiard, P.; Magos-Palasyuk, E.; Wang, H.; Bonev, S. A.; Dudenko, D.; Naumov, P. Ammonia as a Case Study for the Spontaneous Ionization of a Simple Hydrogen-Bonded Compound. *Nat. Commun.* **2014**, *5* (1), 3460. <https://doi.org/10.1038/ncomms4460>.
229. Ji, M.; Umemoto, K.; Wang, C.-Z.; Ho, K.-M.; Wentzcovitch, R. Ultrahigh-Pressure Phases of H<sub>2</sub>O Ice Predicted Using an Adaptive Genetic Algorithm. *Phys. Rev. B* **2011**, *84* (22), 220105. <https://doi.org/10.1103/PhysRevB.84.220105>.

230. Wang, Y.; Liu, H.; Lv, J.; Zhu, L.; Wang, H.; Ma, Y. High Pressure Partially Ionic Phase of Water Ice. *Nat. Commun.* **2011**, *2*, 563. <https://doi.org/10.1038/ncomms1566>.
231. McMahon, J. Ground-State Structures of Ice at High Pressures from Ab Initio Random Structure Searching. *Phys. Rev. B* **2011**, *84* (22), 220104. <https://doi.org/10.1103/PhysRevB.84.220104>.
232. Pickard, C. J.; Martinez-Canales, M.; Needs, R. J. Decomposition and Terapascal Phases of Water Ice. *Phys. Rev. Lett.* **2013**, *110* (24), 245701. <https://doi.org/10.1103/PhysRevLett.110.245701>.
233. Millot, M.; Coppari, F.; Bygg, J. R.; Correa Barrios, A.; Hamel, S.; Swift, D. C.; Eggert, J. H. Nanosecond X-Ray Diffraction of Shock-Compressed Superionic Water Ice. *Nature* **2019**, *569* (7755), 251–255. <https://doi.org/10.1038/s41586-019-1114-6>.
234. Prakapenka, V. B.; Holtgrewe, N.; Lobanov, S. S.; Goncharov, A. F. Structure and Properties of Two Superionic Ice Phases. *Nat. Phys.* **2021**, *17* (11), 1233–1238. <https://doi.org/10.1038/s41567-021-01351-8>.
235. Weck, G.; Queyroux, J.-A.; Ninet, S.; Datchi, F.; Mezouar, M.; Loubeyre, P. Evidence and Stability Field of Fcc Superionic Water Ice Using Static Compression. *Phys. Rev. Lett.* **2022**, *128* (16), 165701. <https://doi.org/10.1103/PhysRevLett.128.165701>.
236. Naden Robinson, V.; Marqués, M.; Wang, Y.; Ma, Y.; Hermann, A. Novel Phases in Ammonia-Water Mixtures under Pressure. *J. Chem. Phys.* **2018**, *149* (23), 234501. <https://doi.org/10.1063/1.5063569>.
237. Xu, W.; Robinson, V. N.; Zhang, X.; Zhang, H.-C.; Donnelly, M.-E.; Dalladay-Simpson, P.; Hermann, A.; Liu, X.-D.; Gregoryanz, E. Ionic Phases of Ammonia-Rich Hydrate at High Densities. *Phys. Rev. Lett.* **2021**, *126* (1), 015702. <https://doi.org/10.1103/PhysRevLett.126.015702>.
238. Zhang, L.; Wang, Y.; Zhang, X.; Ma, Y. High-Pressure Phase Transitions of Solid HF, HCl, and HBr: An Ab Initio Evolutionary Study. *Phys. Rev. B* **2010**, *82* (1), 014108. <https://doi.org/10.1103/PhysRevB.82.014108>.
239. Duan, D.; Tian, F.; He, Z.; Meng, X.; Wang, L.; Chen, C.; Zhao, X.; Liu, B.; Cui, T. Hydrogen Bond Symmetrization and Superconducting Phase of HBr and HCl under High Pressure: An Ab Initio Study. *J. Chem. Phys.* **2010**, *133* (7), 074509. <https://doi.org/10.1063/1.3471446>.
240. Chen, C.; Xu, Y.; Sun, X.; Wang, S. Novel Superconducting Phases of HCl and HBr under High Pressure: An Ab Initio Study. *J. Phys. Chem. C* **2015**, *119* (30), 17039–17043. <https://doi.org/10.1021/acs.jpcc.5b01653>.
241. Binns, J.; Hermann, A.; Peña-Alvarez, M.; Donnelly, M.-E.; Wang, M.; Kawaguchi, S. I.; Gregoryanz, E.; Howie, R. T.; Dalladay-Simpson, P. Superionicity, Disorder, and Bandgap Closure in Dense Hydrogen Chloride. *Sci. Adv.* **2021**, *7* (36), eabi9507. <https://doi.org/10.1126/sciadv.abi9507>.
242. Hermann, A.; Ashcroft, N. W.; Hoffmann, R. Lithium Hydroxide, LiOH, at Elevated Densities. *J. Chem. Phys.* **2014**, *141* (2), 024505. <https://doi.org/10.1063/1.4886335>.
243. Hermann, A.; Guthrie, M.; Nelves, R. J.; Loveday, J. S. Pressure-Induced Localisation of the Hydrogen-Bond Network in KOH-VI. *J. Chem. Phys.* **2015**, *143* (24), 244706. <https://doi.org/10.1063/1.4938260>.
244. Hermann, A. High-Pressure Phase Transitions in Rubidium and Caesium Hydroxides. *Phys. Chem. Chem. Phys.* **2016**, *18* (24), 16527–16534. <https://doi.org/10.1039/C6CP03203F>.
245. Tsuchiya, J. First Principles Prediction of a New High-Pressure Phase of Dense Hydrous Magnesium Silicates in the Lower Mantle. *Geophys. Res. Lett.* **2013**, *40* (17), 4570–4573. <https://doi.org/10.1002/grl.50875>.
246. Nishi, M.; Irifune, T.; Tsuchiya, J.; Tange, Y.; Nishihara, Y.; Fujino, K.; Higo, Y. Stability of Hydrous Silicate at High Pressures and Water Transport to the Deep Lower Mantle. *Nat. Geosci.* **2014**, *7* (3), 224–227. <https://doi.org/10.1038/ngeo2074>.
247. Hermann, A.; Mookherjee, M. High-Pressure Phase of Brucite Stable at Earth's Mantle Transition Zone and Lower Mantle Conditions. *Proc. Natl. Acad. Sci.* **2016**, *113* (49), 13971–13976. <https://doi.org/10.1073/pnas.1611571113>.
248. Zhong, X.; Hermann, A.; Wang, Y.; Ma, Y. Monoclinic High-Pressure Polymorph of AlOOH Predicted from First Principles. *Phys. Rev. B* **2016**, *94* (22), 224110. <https://doi.org/10.1103/PhysRevB.94.224110>.
249. Gao, H.; Liu, C.; Shi, J.; Pan, S.; Huang, T.; Lu, X.; Wang, H.-T.; Xing, D.; Sun, J. Superionic Silica-Water and Silica-Hydrogen Compounds in the Deep Interiors of Uranus and Neptune. *Phys. Rev. Lett.* **2022**, *128* (3), 035702. <https://doi.org/10.1103/physrevlett.128.035702>.
250. Li, H.-F.; Oganov, A. R.; Cui, H.; Zhou, X.-F.; Dong, X.; Wang, H.-T. Ultrahigh-Pressure Magnesium Hydroxides as Reservoirs of Water in Early Earth. *Phys. Rev. Lett.* **2022**, *128* (3), 035703. <https://doi.org/10.1103/physrevlett.128.035703>.
251. Hermann, A. Geoscience Material Structures Prediction via CALYPSO Methodology. *Chin. Phys. B* **2019**, *28* (10), 106107. <https://doi.org/10.1088/1674-1056/ab43bc>.
252. Ashcroft, N. W. Hydrogen Dominant Metallic Alloys: High Temperature Superconductors? *Phys. Rev. Lett.* **2004**, *92* (18), 187002. <https://doi.org/10.1103/PhysRevLett.92.187002>.
253. Wang, H.; Tse, J. S.; Tanaka, K.; Iitaka, T.; Ma, Y. Superconductive Sodalite-Like Clathrate Calcium Hydride at High Pressures. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (17), 6463–6466. <https://doi.org/10.1073/pnas.1118168109>.
254. Ma, L.; Wang, K.; Xie, Y.; Yang, X.; Wang, Y.; Zhou, M.; Liu, H.; Yu, X.; Zhao, Y.; Wang, H.; Liu, G.; Ma, Y. High-Temperature Super-Conducting Phase in Clathrate Calcium Hydride CaH<sub>6</sub> up to 215 K at a Pressure of 172 GPa. *Phys. Rev. Lett.* **2022**, *128* (16), 167001. <https://doi.org/10.1103/physrevlett.128.167001>.
255. Li, Z.; He, X.; Zhang, C.; Wang, X.; Zhang, S.; Jia, Y.; Feng, S.; Lu, K.; Zhao, J.; Zhang, J.; Min, B.; Long, Y.; Yu, R.; Wang, L.; Ye, M.; Zhang, Z.; Prakapenka, V.; Chariton, S.; Ginsberg, P. A.; Bass, J.; Yuan, S.; Liu, H.; Jin, C. Superconductivity above 200 K Discovered in Superhydrides of Calcium. *Nat. Commun.* **2022**, *13* (1), 2863. <https://doi.org/10.1038/s41467-022-30454-w>.
256. Drozdov, A. P.; Erements, M. I.; Troyan, I. A.; Ksenofontov, V.; Shylin, S. I. Conventional Superconductivity at 203 Kelvin at High Pressures in the Sulfur Hydride System. *Nature* **2015**, *525* (7567), 73–76. <https://doi.org/10.1038/nature14964>.
257. Liu, H.; Naumov, I. I.; Hoffmann, R.; Ashcroft, N. W.; Hemley, R. J. Potential High- $T_c$  Superconducting Lanthanum and Yttrium Hydrides at High Pressure. *Proc. Natl. Acad. Sci.* **2017**, *114* (27), 6990–6995. <https://doi.org/10.1073/pnas.1704505114>.
258. Geballe, Z. M.; Liu, H.; Mishra, A. K.; Ahart, M.; Somayazulu, M.; Meng, Y.; Baldini, M.; Hemley, R. J. Synthesis and Stability of Lanthanum Superhydrides. *Angew. Chem. Int. Ed.* **2018**, *57* (3), 688–692. <https://doi.org/10.1002/anie.201709970>.
259. Drozdov, A. P.; Kong, P. P.; Minkov, V. S.; Besedin, S. P.; Kuzovnikov, M. A.; Mozaffari, S.; Balicas, L.; Balakirev, F. F.; Graf, D. E.; Prakapenka, V. B.; Greenberg, E.; Knyazev, D. A.; Tkacz, M.; Erements, M. I. Superconductivity at 250 K in Lanthanum Hydride under High Pressures. *Nature* **2019**, *569* (7757), 528–531. <https://doi.org/10.1038/s41586-019-1201-8>.
260. Somayazulu, M.; Ahart, M.; Mishra, A. K.; Geballe, Z. M.; Baldini, M.; Meng, Y.; Struzhkin, V. V.; Hemley, R. J. Evidence for Superconductivity above 260 K in Lanthanum Superhydride at Megabar Pressures. *Phys. Rev. Lett.* **2019**, *122* (2), 027001. <https://doi.org/10.1103/PhysRevLett.122.027001>.
261. Peng, F.; Sun, Y.; Pickard, C. J.; Needs, R. J.; Wu, Q.; Ma, Y. Hydrogen Clathrate Structures in Rare Earth Hydrides at High Pressures: Possible Route to Room-Temperature Superconductivity. *Phys. Rev. Lett.* **2017**, *119* (10), 107001. <https://doi.org/10.1103/PhysRevLett.119.107001>.
262. Snider, E.; Dasenbrock-gammon, N.; McBride, R.; Wang, X.; Meyers, N.; Lawler, K. V.; Zurek, E.; Salamat, A.; Dias, R. P. Synthesis of Yttrium Superhydride Superconductor with a Transition Temperature up to 262 K by Catalytic Hydrogenation at High Pressures. *Phys. Rev. Lett.* **2021**, *126* (11), 117003. <https://doi.org/10.1103/PhysRevLett.126.117003>.
263. Kong, P.; Minkov, V. S.; Kuzovnikov, M. A.; Drozdov, A. P.; Besedin, S. P.; Mozaffari, S.; Balicas, L.; Balakirev, F. F.; Prakapenka, V. B.; Chariton, S.; Knyazev, D. A.; Greenberg, E.; Erements, M. I. Superconductivity up to 243 K in the Yttrium-Hydrogen System under High Pressure. *Nat. Commun.* **2021**, *12* (1), 5075. <https://doi.org/10.1038/s41467-021-25372-2>.
264. Avery, P.; Wang, X.; Oses, C.; Gossett, E.; Proserpio, D. M.; Toher, C.; Curtarolo, S.; Zurek, E. Predicting Superhard Materials via a Machine Learning Informed Evolutionary Structure Search. *npj Comput. Mater.* **2019**, *5* (1), 89. <https://doi.org/10.1038/s41524-019-0226-8>.
265. Allahyari, Z.; Oganov, A. R. Coevolutionary Search for Optimal Materials in the Space of all Possible Compounds. *npj Comput. Mater.* **2020**, *6* (1), 55. <https://doi.org/10.1038/s41524-020-0322-9>.

266. Wang, X.; Proserpio, D. M.; Oses, C.; Toher, C.; Curtarolo, S.; Zurek, E. The Microscopic Diamond Anvil Cell: Stabilization of Superhard, Superconducting Carbon Allotropes at Ambient Pressure. *Angew. Chem. Int. Ed.* **2022**, *61* (32), e202205129. <https://doi.org/10.1002/anie.202205129>.
267. Lee, J.; Seko, A.; Shitara, K.; Nakayama, K.; Tanaka, I. Prediction Model of Band Gap for Inorganic Compounds by Combination of Density Functional Theory Calculations and Machine Learning Techniques. *Phys. Rev. B* **2016**, *93* (11), 115104. <https://doi.org/10.1103/PhysRevB.93.115104>.
268. Zhuo, Y.; Mansouri Tehrani, A.; Brgoch, J. Predicting the Band Gaps of Inorganic Solids by Machine Learning. *J. Phys. Chem. Lett.* **2018**, *9* (7), 1668–1673. <https://doi.org/10.1021/acs.jpclett.8b00124>.
269. Schmidt, J.; Marques, M. R. G.; Botti, S.; Marques, M. A. L. Recent Advances and Applications of Machine Learning in Solid-State Materials Science. *npj Comput. Mater.* **2019**, *5* (1), 83. <https://doi.org/10.1038/s41524-019-0221-0>.
270. Sarker, P.; Harrington, T.; Toher, C.; Oses, C.; Samiee, M.; Maria, J.-P.; Brenner, D. W.; Vecchio, K. S.; Curtarolo, S. High-Entropy High-Hardness Metal Carbides Discovered by Entropy Descriptors. *Nat. Commun.* **2018**, *9* (1), 4980. <https://doi.org/10.1038/s41467-018-07160-7>.
271. Toher, C.; Oses, C.; Hicks, D.; Curtarolo, S. Unavoidable Disorder and Entropy in Multi-Component Systems. *npj Comput. Mater.* **2019**, *5* (1), 69. <https://doi.org/10.1038/s41524-019-0206-z>.
272. Schaack, S.; Ranieri, U.; Depondt, P.; Gaal, R.; Kuhs, W. F.; Falenty, A.; Gillet, P.; Finocchi, F.; Bove, L. E. Orientational Ordering, Locking-in, and Distortion of CH<sub>4</sub> Molecules in Methane Hydrate III under High Pressure. *J. Phys. Chem. C* **2018**, *122* (20), 11159–11166. <https://doi.org/10.1021/acs.jpcc.8b02783>.
273. Degtyareva, O. Crystal Structure of Simple Metals at High Pressures. *High Press. Res.* **2010**, *30* (3), 343–371. <https://doi.org/10.1080/08957959.2010.508877>.
274. Wörle, M.; Nesper, R.; Chatterji, T. K. LiBx (0.82 < x ≤ 1.0) – an Incommensurate Composite Structure below 150 K. *Z. Anorg. Allg. Chem.* **2006**, *632* (10–11), 1737–1742. <https://doi.org/10.1002/zaac.200600117>.
275. Hastings, J. M.; Pouget, J. P.; Shirane, G.; Heeger, A. J.; Miro, N. D.; MacDiarmid, A. G. One-Dimensional Phonons and “Phase-Ordering” Phase Transition in Hg<sub>3-8</sub>AsF<sub>6</sub>. *Phys. Rev. Lett.* **1977**, *39* (23), 1484–1487. <https://doi.org/10.1103/PhysRevLett.39.1484>.
276. Hermann, A.; Suarez-Alcubilla, A.; Gurtubay, I. G.; Yang, L.-M.; Bergara, A.; Ashcroft, N. W.; Hoffmann, R. LiB and its Boron-Deficient Variants under Pressure. *Phys. Rev. B* **2012**, *86*, 144110. <https://doi.org/10.1103/PhysRevB.86.144110>.
277. McMahon, M. I.; Nelmes, R. J. High-Pressure Structures and Phase Transformations in Elemental Metals. *Chem. Soc. Rev.* **2006**, *35* (10), 943. <https://doi.org/10.1039/b517777b>.
278. Whaley-Baldwin, J.; Hutcheon, M.; Pickard, C. J. Superconducting Incommensurate Host-Guest Phases in Compressed Elemental Sulfur. *Phys. Rev. B* **2021**, *103* (21), 214111. <https://doi.org/10.1103/PhysRevB.103.214111>.
279. Van Smaalen, S. Incommensurate crystal structures. *Crystallography Rev.* **1995**, *4* (2), 79–202. <https://doi.org/10.1080/08893119508039920>.