MATERIALS SCIENCE

Simplifying inverse materials design problems for fixed lattices with alchemical chirality

Guido Falk von Rudorff[1,2] and O. Anatole von Lilienfeld[1,2]*

Brute-force compute campaigns relying on demanding ab initio calculations routinely search for previously unknown materials in chemical compound space (CCS), the vast set of all conceivable stable combinations of elements and structural configurations. Here, we demonstrate that four-dimensional chirality arising from antisymmetry of alchemical perturbations dissects CCS and defines approximate ranks, which reduce its formal dimensionality and break down its combinatorial scaling. The resulting "alchemical" enantiomers have the same electronic energy up to the third order, independent of respective covalent bond topology, imposing relevant constraints on chemical bonding. Alchemical chirality deepens our understanding of CCS and enables the establishment of trends without empiricism for any materials with fixed lattices. We demonstrate the efficacy for three cases: (i) new rules for electronic energy contributions to chemical bonding; (ii) analysis of the electron density of BN-doped benzene; and (iii) ranking over 2000 and 4 million BN-doped naphthalene and picene derivatives, respectively.

INTRODUCTION

The computational simulation of molecules and materials, performed to predict their physical, material, and chemical properties, has become a routine tool in the molecular and materials sciences. Current efforts geared towards computational materials and molecular design might, one day, enable the realization of the holy grail of automatized experimental design and discovery. Driven by the accelerating progress of computer hardware and statistical learning (artificial intelligence), the first seminal examples of integrating sophisticated software and robotics to perform experimental sequences and to establish rules and trends among properties and materials, as well as their synthesis, in realiter have recently been introduced (*1–5*). However, the lofty goal of "materials on demand" has still remained elusive, even when doing it just in silico.

The use of empirical trends to guide experimental design has had a long tradition in the chemical sciences. Popular examples include Mendeleev's discovery of the periodic table, Hammett's relationship, Pettifor's numbering scheme, the Bell-Evans-Polanyi principle, Hammond's postulate, or Pauling's covalent bond postulate (*6*). Modern systematic attempts to establish and exploit such rules in terms of quantitative structure-property relationships have led to computationally advanced bio-, chem-, and materials-informatics methodologies (*7*). Unfortunately, these methods are typically inherently limited to certain applicability domains and do not scale because of their empirical nature (*8*). To rigorously explore the high-dimensional chemical compound space (CCS) (*9*), i.e., the combinatorially scaling number of all conceivable molecules or materials (usually defined by composition, constitution, and conformation), the quantum mechanics of electrons ought to be invoked.

It is thus not unexpected that ab initio–based materials design approaches have been at the forefront for more than 20 years (*10–14*) and have played a major role in popularizing the use of efficient and accurate quantum methods, such as density functional theory (DFT)

(*15–19*). Sampling CCS from scratch, even when done within efficient optimization algorithms, is typically an encyclopedical endeavor by nature and ignores the many underlying relations among different properties and materials. Quantum machine learning models (*20*) statistically exploit these implicit correlations, hidden in the data, and have successfully accelerated CCS exploration campaigns (*21*). Machine learning efficiency and transferability demonstrably benefit greatly from explicitly enforcing known relationships (e.g., translational, rotational, or atom index invariances) directly in the model construction rather than having to learn them agnostically from data (*22*). Specific examples include explicitly imposing forces and curvatures in the loss function (*23–25*), spatial symmetry relations (*26, 27*), or arbitrary differential relations (*28*). However, even for the most efficient and transferable statistical models, e.g., the atom-in-molecule fragment–based approach (*29*), the acquisition of training data in sufficient quantity and quality requires considerable up-front investments.

Here, we introduce the fundamental notion of a new symmetry relation in CCS, which is fully consistent with the ab initio view of matter (*30*) and effectively enables us to solve the inverse materials design problem in a nonempirical and highly efficient manner. Spatial symmetry considerations have been crucial for the unraveling of some of the most fundamental laws of nature and are heavily used in many fields. In ab initio calculations, for example, symmetry group theory arguments are common to reduce computational complexity and load. Symmetry constraints on compositional degrees of freedom would be highly desirable to establish general rules among distinct materials and properties and to generally improve our understanding of CCS (*30*).

In analogy to conventional spatial chirality, we here define "alchemical chirality" as a reflection plane in the space spanned by nuclear charges at fixed atomic positions as they enter the electronic Schrödinger equation. An illustrative comparison is given in Fig. 1 (A and B) for conventional enantiomers consisting of a tetravalent carbon atom with four different substitutions and for alchemical enantiomers consisting of doubly boron nitride (BN)–doped carbon in the diamond crystal structure. Figure 1C compares and relates this newly described alchemical reflection σ_A with the conventional spatial reflection σ_S for the same dopant pattern as in Fig. 1B for a

[1]University of Vienna, Faculty of Physics, Kolingasse 14-16, 1090 Vienna, Austria. [2]Institute of Physical Chemistry and National Center for Computational Design and Discovery of Novel Materials (MARVEL), Department of Chemistry, University of Basel, 4056 Basel, Switzerland.
*Corresponding author. Email: anatole.vonlilienfeld@univie.ac.at.
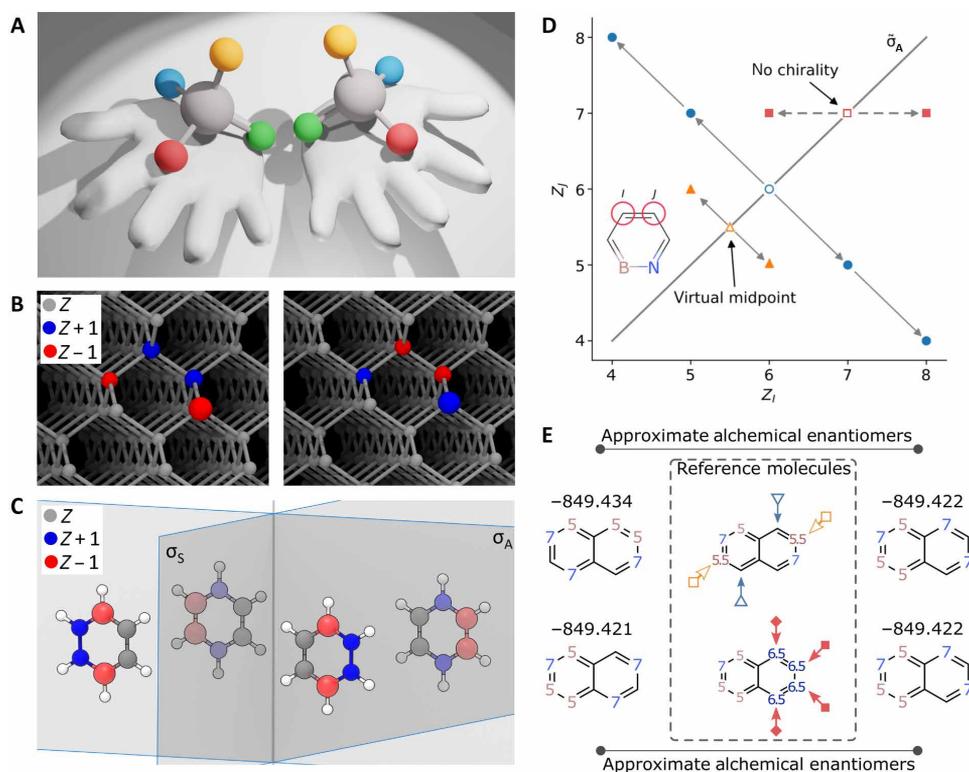
Fig. 1. Illustrations of alchemical chirality and spatial chirality. (**A**) Enantioselective catalysis enables the synthesis of either conventional enantiomer related to its counterpart through spatial reflection symmetry. (**B**) Alchemical (not spatial) chirality at four sites in a diamond cubic lattice relates BN-doped alchemical enantiomers through alchemical reflection (boron, nitrogen, and carbon in red, blue, and gray, respectively). (**C**) Alchemical reflection plane σ_A for the six carbon nuclei ($Z = 6$) in benzene connecting 1,2,3,6-tetrahydro-1,2,3,6-diazadiborinine (THDADB) to 1,2,3,4-tetrahydro-1,4,2,3-diazadiborinine compared to the spatial reflection plane σ_S. Boron, nitrogen, carbon, and hydrogen in red, blue, gray, and white, respectively. (**D**) Illustration of approximate alchemical reflection plane $\tilde{\sigma}_A$ (diagonal) in the space defined by the nuclear charges Z of two atoms I, J in three cyclic CBNC-containing reference molecules (empty symbols). Three reflections via virtual reference with fractional nuclear charges (orange triangle), real reference with two carbons (blue circle), and real reference with two nitrogens (red square) are indicated. The latter reflection does not yield alchemical enantiomers. (**E**) Chains of approximate alchemical enantiomers. The top left and top right molecules are connected via one reference molecule, while the bottom left and bottom right molecules are connected via another reference molecule. Both molecules on the right are identical, and therefore, the three distinct molecules form a chain of approximate alchemical enantiomers. Transmutating atoms and symmetry-equivalent sites are denoted with nuclear charges and identical arrows, respectively. Electronic energies are given in Hartrees.

single molecular skeleton. In this case, four subsequent reflections alternating between alchemical and spatial reflections return to the original molecule. While any spatial reflection leaves the molecule unchanged, an alchemical reflection affects the nuclear charges and therefore creates a different molecule as reflection image.

Exchange of the dopant atom sequence in Fig. 1B from NBBN to BNNB is an alchemical reflection around pristine diamond: For each site, the nuclear charge difference to diamond gets inverted. Hence, in the space spanned by all nuclear charges of the system, pristine diamond corresponds to a reflection center. All reflections that leave the total sum of all nuclear charges unchanged are defined by a hyperplane (cf. Fig. 1D), which we refer to as the nuclear charge reference plane. In other words, treating the change from pristine diamond to the two doped variants as a perturbation of the system Hamiltonian, there is an antisymmetry relation between these alchemical perturbations.

No other spatial symmetry operation (rotation, reflection, or inversion) can interconvert the constitutional isomers in Fig. 1B, thereby necessitating a fourth dimension, namely, the nuclear charges. This amounts to a four-dimensional (4D) alchemical chirality. Note that chirality crucially depends on dimensionality, e.g., the letter L is

chiral within two dimensions only. The chiral center of our chirality operation can be a compound in itself (diamond in Fig. 1B and benzene in Fig. 1C).

Alchemical enantiomers exist only if pairs of distinct atomic environments can be mapped onto each other under a symmetry operation, a consequence of the reflection in nuclear charge space that defines alchemical enantiomers. This implies that the total sum of all nuclear charges of alchemical enantiomers is identical (see Fig. 1D). If a compound has no spatial symmetry, then atom sites with similar electron density derivatives $\partial_Z^n \rho$ constitute these pairs just like strictly symmetry-equivalent atoms do. In any case, alchemical chirality requires a one-to-one correspondence of sites with opposite change in nuclear charge. At least two such pairs need to exist in a compound to obtain alchemical enantiomers, which are different chemical objects. For a single pair of symmetry-equivalent atoms, the alchemical reflection in nuclear charge space would trivially connect two spatially symmetric compounds, e.g. CO and OC.

In summary, exact alchemical enantiomers are defined as two spatially non-superimposable, alchemically coupled, and iso-electronic compounds with the same formal charge, where each transmutating atom is assigned to exactly one subset within each of which

averaging of nuclear charges results in identical chemical environments. Note that alchemical enantiomers can differ in chemical compositions, and that any given alchemical enantiomer can have as many alchemical mirror images as different valid averaged reference compounds can be defined. An alchemical enantiomer and all its possible mirror images are degenerate in the electronic energy up to the third order. Approximate alchemical enantiomers differ from their exact analog in that the pair-wise averaging results in similar, rather than identical, atoms within each pair. Alchemical enantiomers have approximately the same electronic energy (see the "Materials and Methods" section below). This symmetry is then broken by (i) the nuclear-nuclear repulsion and (ii) geometry relaxation into distinct total energy minima. The nuclear repulsion will typically dominate, implying that alchemical enantiomers with spatially closer atoms of higher nuclear charge will exhibit higher total energies than their respective mirror images.

While the restriction to fixed atomic positions might seem severe at first, we highlight in the following the importance and relevance of classes of fixed configurational frameworks for materials design applications. More specifically, we discuss the implications of alchemical chirality for system classes that are low dimensional in their structural degrees of freedom and where the problem is dominated by the combinatorial scaling due to varying chemical composition.

Examples abound and include any variation of graphitic motifs (studied below) prevalent in nanotechnological applications, in inorganic materials such as MAX-phases, or in organic electronics, e.g., polycyclic aromatic hydrocarbon derivatives and other rigid scaffolds, such as metal-organic frameworks or perovskites. All these systems would be directly amenable to alchemical chirality–based estimates (ACEs), i.e., approximating the electronic energies for alchemical enantiomers as degenerate and adding nuclear repulsions to obtain relative energies. This way, ACEs enable the ranking and grouping of large subsets of materials. Within any of these classes, the number of possible materials increases combinatorially with the number of building blocks. For real-world applications, defects often need to be included in the quantum chemistry model, which further increases the chemical space under consideration. For particularly rigid frameworks and extended periodic materials, only a local and low-dimensional geometric response to a change of composition is typically observed. Hence, ACEs amount to a complementary means to rigorously sample compositional ensembles within a given framework class. As shown below, ACEs become an enabling tool for the systematic identification of structure property relations within given compound classes that further the understanding of the respective impact of the compositional degrees of freedom in materials design. Here, we exemplify this by identifying design rules from analysis of more than 400 million alchemical estimates of BN-doped derivatives in the picene framework.

RESULTS
Estimating bond energies
The external potential energy differences between the reference compound (defining the reflection plane such as carbon atoms in diamond in Fig. 1B) and either alchemical enantiomer are exactly mirrored in magnitude (opposite sign). Hence, while the corresponding electronic Hamiltonians are alchemical mirror images of each other, their respective solutions to the electronic Schrödinger

equation are not necessarily equal. More specifically, for reflections around atoms with identical chemical environments, we show that the electronic energy of corresponding alchemical enantiomers must be degenerate up to the third order within alchemical perturbation DFT (APDFT) (31) (see the "Materials and Methods" section below), higher-order terms carrying different signs. Hence, approximate alchemical enantiomers are only approximately degenerate in their electronic energy.

It turns out, however, that this approximation is quite fair and that useful rules for chemical binding can be derived. For example, by enumerating all colored connected graphs that are subgraphs of hexagonal lattices, we obtain the following ACE of two-body interatomic bonding for the electronic energy

$$E_{QR} \approx E_{SR} + 0.5 \times (E_{QQ} - E_{SS}) \qquad (1)$$

where QRS correspond to three adjacent elements in the periodic table. Hence, this rule confirms the naive expectation that given two heteroatomic bonds—QR and SR—to the same reference atom R, that electronic contribution to bonding will be larger, which contains the heteroatom with the larger electronic homoatomic bond, i.e., QQ versus SS, respectively. We believe that this simple rule was hidden in the past because of the obfuscation caused by the contributing nuclear repulsion term.

While seemingly reminiscent of Pauling's electronegativity-based bond energy estimate (6), $E_{QS} = 0.5 \times (E_{QQ} + E_{SS}) + \Delta\chi$, we stress that this ACE relation is derived from pure symmetry and perturbation theory–based considerations that, for alchemical enantiomers, are exact up to the third order. Pauling, by contrast, merely postulated his Ansatz. Our rule relies on bonding information involving three distinct and adjacent chemical elements rather than just two, and maybe more importantly, it pertains to the electronic potential energy only, i.e., without the nuclear repulsion, which is trivial to add a posteriori. This rule is easily verified for the example of BC and CN bonds in tetrahydrodiazadiborine (THDADB) by simply comparing the sum of all bond energies in either alchemical enantiomer (resulting in $2E_{BC} + E_{NN} \approx 2E_{NC} + E_{BB}$).

Subtracting nuclear repulsion estimates from published DFT data for all single bonds between main-group elements from the second and third periods saturated by hydrogen (32), ACEs yield bond energies of elements QR, i.e., solely based on bonding information of elements SR, QQ, and SS for adjacent elements QRS in the same period. As shown in Fig. 2, predicting bond dissociation energies among elements R and Q or S (to the left or right in the period) generates remarkably accurate predictions for bonds among elements in either period with a mean absolute error of ~10 kJ/mol, not far from the highly coveted "chemical accuracy" of ~4 kJ/mol. A linear fit through all the diverse chemistries encountered yields a slope of 1.04, an offset of ~7 kJ/mol, and a correlation coefficient of 0.981. Successive daisy chaining across groups in any given period represents a straightforward extension, which markedly reduces the number of reference bonds required for ACE—but only at the cost of reduced accuracy: The mean absolute error (MAE) increases to ~22 kJ/mol, i.e., still better or on par with common generalized gradient–based approximations to the exchange correlation potential in DFT (33).

Furthermore, ACE rules emerge when increasing the alchemical nuclear charge radius ($\Delta Z = \pm 1, \pm 2$) with PQRST corresponding to five adjacent elements in the periodic table: $E_{PR} \approx E_{TR} + (E_{PP} - E_{TT})/2$,

$E_{PQ} + E_{QT} + 2E_{SR} \approx E_{PS} + E_{ST} + 2E_{QR}$, and $2E_{PR} + E_{QT} + E_{ST} \approx E_{PQ} + E_{PS} + 2E_{TR}$. These rules are identical for the 3D diamond lattice and the 2D hexagonal graphene structure. Note that other graph lattices could yield additional rules and that rules for interatomic 3- and n-body contributions to binding exist as well. For example, invoking $\Delta Z = \pm 1$ only, one finds the three-body rule that $E_{SQS} + 2E_{RSQ} + E_{RQR} \approx E_{QSQ} + 2E_{RQS} + E_{RSR}$. A systematic enumeration of ACE rules for two- and three-body terms for graphene and diamond lattices is given in the Supplementary Materials.
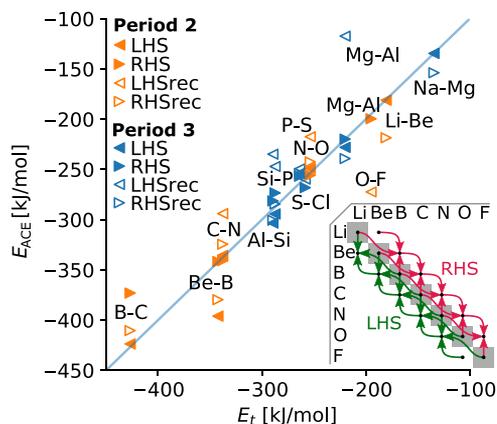

Fig. 2. Predicted (E_{ACE}) from ACEs versus true (E_t) single-bond dissociation energies. These are predicted according to the two-body rule for adjacent elements QRS using published DFT results for main-group elements in the second and third rows of the periodic table and assuming all bond distances to correspond to 2 Bohr. Predictions of bonding energies between QR and RS elements are denoted as LHS and RHS, respectively. Successive recursive application of the rule throughout either period is denoted by "rec". Inset: Each field in the matrix corresponds to a bond of the elements in the row and column headers. Arrows show which values, when combined together, yield a predicted bond energy. This shows the symmetry in going forward (RHS) and backward (LHS) through the groups in the periodic table. Using only the initial off-diagonal matrix element and then using one prediction to obtain the next following the arrows would correspond to rec in the main figure.

Ranking molecules

To apply ACE to inverse materials design problems, we showcase applications in well-defined subregions of CCS and solve three specific and increasingly challenging design tasks. All three use cases address the combinatorial design problem of how to dope planar hexagonal lattices to an increasing extent (6, 10, and 22 carbon atoms, respectively). Hexagonal lattices are archetypical scaffolds, e.g., relevant in the design of graphene-inspired materials for nanotechnological devices (34), catalytic surfaces (35–37), porous BN-doped–based nanofibers for battery materials (38), or 2D materials in general (39). Just doping with BN already results in a combinatorial explosion; already for the 77 smallest benzoid-like structures, the number of possible unique BN-doped derivatives exceeds 7 tera (40). Hence, we believe that it is warranted, and without loss of generality, to focus on constitutional isomers in rigid lattices for which thermal or geometrical distortions can be neglected. Note that relative offsets in total energies due to differences in stoichiometry are straightforward to estimate, and typically occur on different orders of magnitude, and that subsequent inclusion of configurational degrees of freedom within alchemical predictions is possible, as already demonstrated for small molecules (41, 42) and ionic, metallic, and semiconducting solids (43–45). Hence, the focus of the following applications lies on breaking down the combinatorially scaling problem of energy predictions throughout colored chemical bond connectivity graphs.

As a first use case, we provide an in-depth but intuitive illustration of alchemical chirality with benzene, its six equivalent carbon atoms as the mirroring reflection plane, and the alchemical enantiomers corresponding to carborazine [$C_2H_6B_2N_2$, THDADB]. Because second-order APDFT—resulting in alchemical normal modes to define a complete basis in certain subspaces of CCS—has already been applied to benzene (46), the occurrence of six degenerate estimates among the 11 constitutional isomers of THDADB ($C_2H_6B_2N_2$) can now be readily explained, thanks to alchemical chirality: They represent three pairs of alchemical enantiomers. For the select enantiomer pair BNNBCC/NBBNCC, a molecular planar analog to the crystalline example (Fig. 1B) is given in Fig. 1C, contrasting the
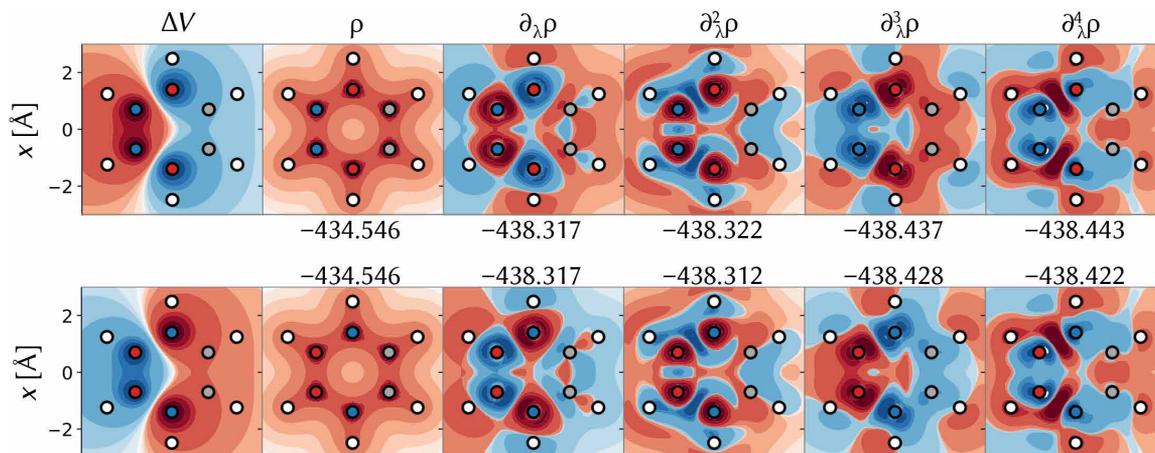

Fig. 3. Alchemical enantiomers in carborazine [$C_2H_6B_2N_2$, tetrahydrodiazadiborine (THDADB)] with the reference molecule benzene (C_6H_6). Left-hand column: the perturbing potential ΔV acting on benzene. Remaining columns correspond to electron density derivatives at CCSD/def2-TZVP level. Positive/negative values are shown in red/blue, respectively. All density contour lines are set at the same percentiles for each plot to render their shape comparable at different magnitudes. The percentiles are chosen to emphasize extremal values. Next to each electron density derivative, the corresponding electronic energy estimate by APDFT is given in Hartree for energy order 1 up to 5. The actual electronic energies are −438.413 Ha (top) and −438.401 Ha (bottom).

conventional spatial reflection plane σ_s with the alchemical reflection plane σ_A. While NBBNCC and BNNBCC are mutually achiral in the sense of conventional spatial 3D chirality, the alchemical chirality relation between the two becomes obvious with regular benzene corresponding to the reflection plane.

In Fig. 3, the corresponding perturbing potentials (exact mirror images of each other) for the two benzene enantiomers from Fig. 1C are shown, as well as the resulting electron density derivatives of benzene up to the fourth order. Corresponding figures for all other BN-doped benzene mutants are provided in the Supplementary Materials. As it becomes obvious already by visual inspection, the perturbed densities are nearly identical for even orders and anti-symmetric for odd orders. This indicates that the perturbational series in APDFT (31) extrapolates to an approximately equal amount, leading to the near degeneracy of the electronic energies of the two alchemical enantiomers. To quantify this effect, Fig. 3 also shows the corresponding APDFT-based electronic energy estimates up to the fifth order (within perturbation theory the energy order precedes the wave function order), numerically demonstrating that the degeneracy is exact in the second order and starts to deviate by ~0.01, 0.02, and 0.025 Ha for the third, fourth, and fifth orders, respectively. The fifth-order predictions deviate from the actual electronic energies by ~0.03 and 0.02 Ha, respectively. Addition of the nuclear repulsion terms typically lifts the approximate degeneracy, resulting in quantitative ACEs and simple inequality rules for total energies. Hence, $E_{BNNBCC} > E_{NBBNCC}$ because of the larger nuclear repulsion experienced when atoms with larger nuclear charges (nitrogens in this case) are closer in proximity to each other than atoms with lower nuclear charges (boron).

Such accuracy, while not on par with explicitly correlated quantum chemistry calculations in large basis sets, is on a similar scale as generalized gradient–approximated DFT or semiempirical quantum chemistry methods—typically sufficiently accurate for computational materials design studies as demonstrated for the many examples of successful computational discovery of heterogeneous catalysts (47) or the materials project dataset (48). The following two use cases explore this question for ACE-based ranking in a more systematic and comprehensive fashion.

Because the equivalence of the electronic energies is transitive, having a second reference that connects one alchemical enantiomer with a third one allows us to build chains of alchemical enantiomers that must have similar electronic energies, consequently enabling a ranking of molecules within one such group solely by the magnitude of their respective nuclear repulsion. Figure 4A illustrates the chaining of exact (σ_A) and approximate ($\tilde{\sigma}_A$) alchemical enantiomers for BN-doped naphthalene where four pairs of alchemical enantiomers could form such a chain. For reflections along the exact alchemical symmetry axis σ_A, the energy difference is 1 or 8 mHa, respectively. For cases with approximate alchemical symmetries $\tilde{\sigma}_A$, this energy difference is notably larger (237 mHa) but typically still much smaller than the nuclear repulsion energy contributions.

Overall, exhaustive charge-neutral and isoelectronic BN doping in naphthalene leads to 2285 unique derivatives for which we have used ACE to establish energetic ranks. Having identified all alchemical enantiomers through exhaustive scanning within one stoichiometry [stoichiometries as a whole can be ranked with existing relations (49–51)], we have ranked all possible molecules using ACE within groups of molecules approximately degenerate in electronic energy. These groups form the connected components of a
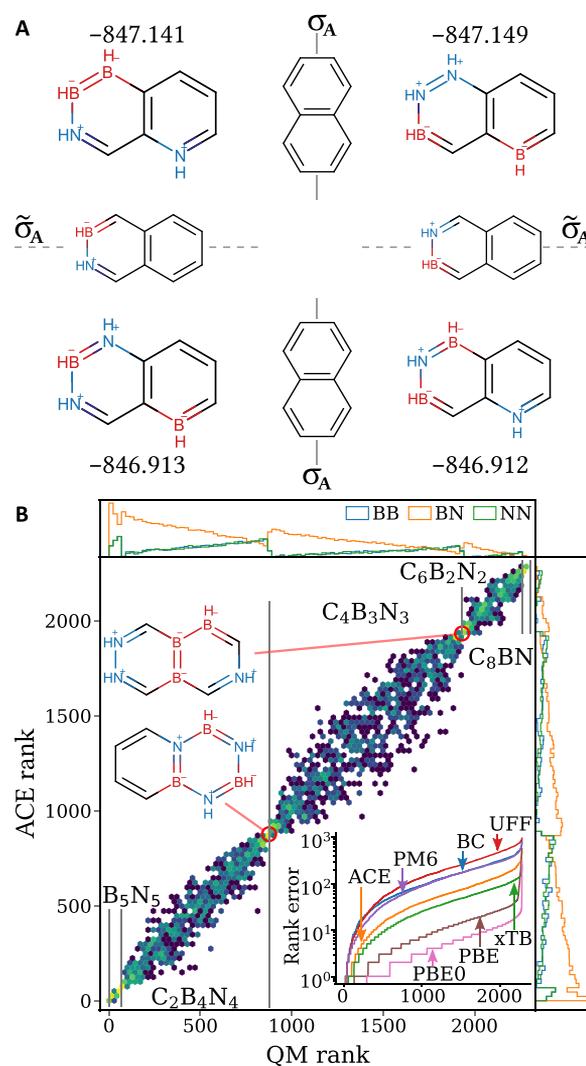

Fig. 4. Ranking molecules via chains of alchemical enantiomers. (**A**) Each alchemical symmetry plane is denoted by a gray line and the reference molecule, the chiral center. Electronic energies given in Hartrees. (**B**) All 2285 unique BN-doped naphthalene derivatives ranked by their total energy as obtained from quantum chemistry calculations compared to the ranking from ACEs. Histogram is colored by number of molecules in a given bin. Two representative molecules of same sum formula $C_4(BN)_3$. Inset shows the distribution of rank errors sorted in ascending order for different methods: the force field UFF, bond counting (BC), semiempirical PM6, ACE (this work), semiempirical xTB-GFN2 (xTB), and the DFT methods PBE and PBE0, both with density-fitted cc-pVDZ basis set. Side panels show bond-type frequencies.

graph where all molecules are nodes and only alchemical enantiomers are connected, thus allowing us to exploit the transitivity of the electronic energy degeneracy. Within these groups, the nuclear repulsion dominates the ranking. Among the groups, we rank on the basis of averaged bond-counting results for their bond energies, where bond energies have been obtained by fitting to the imposed energy degeneracies among all pairs.

Using Coupled cluster with singles and doubles (CCSD/cc-pVDZ) as a ground-truth Quantum Mechanics (QM) method for all naphthalene derivatives on preoptimized geometry (40), we have performed an exhaustive validation of alchemical ranking, resulting

in a Spearman correlation coefficient of 0.9899. Dissecting CCS in all the various possible stoichiometries of BN-doped naphthalene, the scatterplot of ACE rank versus QM rank results in a very reasonable correlation, as shown in Fig. 4B. To place ACE-based ranking in perspective, the inset of Fig. 4B also reports corresponding sorted ranking errors (wrt. CCSD) when using PBE0 (*52–54*), PBE (*16*), xTB (*55*), PM6 (*56*), bond counting, and UFF (*57*) with respective Spearman correlation coefficients of 0.9998, 0.9983, 0.9966, 0.9592, 0.9562, and 0.9021. In terms of computational cost invested to rank all naphthalene derivatives, ACE ranking is slightly more expensive than bond counting, UFF, or PM6. In terms of accuracy; however, it clearly outperforms all three—without any need for empirical knowledge or fitting to external data. It is worth noting that bond counting is the only method that has been reparametrized on this particular dataset, which explains its good performance compared to PM6. As also shown in Fig. 4B, ACE reproduces bond-type frequency trends as a function of rank. While the accuracy of ACE itself is not on par with more advanced semiempirical quantum chemistry (xTB), the sorted ranking error distribution suggests that alchemical ranking is closer to xTB than it is to PM6. However, in contrast to alchemical ranking, xTB relies on substantial empirical data for fitting. While PBE and PBE0 are obviously more accurate, their computational cost is also three orders of magnitude larger than xTB (~40 s). Hence, we conclude that alchemical ranking is superior to bond counting and PM6 and could represent a viable and less empirical alternative to xTB if accuracy thresholds are less stringent and computational load is high.

As a third use case, we have used ACE to explore and deepen our understanding of the CCS spanned by the 413,887,189 unique *k*-fold BN-doped derivatives of picene (see Fig. 5, with *k* < 7 and *k* > 9). Dissecting its CCS first by all stoichiometries, we can use the trends to map out obvious structural features as a function of rank to detect useful descriptors for structure-property relationships. Roughly speaking, results in Fig. 5 suggest that the energy will typically decrease within any given stoichiometry as the number of CC, BN, NN, and BB bonds increases, increases, decreases, and decreases, respectively. Differences in B and N counts in any ring (a measure of compositional homogeneity per ring), hardly affects the energy except for the heavily BN-doped stoichiometries (11 and 10 BN pairs). The degree of clustering (as measured by root mean squared distances) varies wildly with little correlation for BB, BN, and NN, as long as only few sites have been doped. As the degree of BN doping increases, the strong stabilizing effect of BN bonds subdivides the constitutional isomers into groups of identical BN bond count within which boron clusters are stabilizing while nitrogen clusters are destabilizing. Because the BB bond effect is visible for any fixed number of BN bonds, we can conclude that the impact of BN bonds on stability is larger than the one of BB bonds. Following this concept of conditional order and with the data shown in Fig. 5 and the extended version thereof in the Supplementary Materials, we can identify the following stabilizing design patterns in decreasing order of strength: (i) add BN pairs, (ii) maximize number of CC bonds, (iii) substitute sites shared between rings, (iv) maximize number of BN bonds, (v) avoid N substitutions on rings sharing a larger amount of bonds with other rings, and (vi) balance BN substitutions in each ring. Note how ACE has given us access to a complete ranking without a single quantum chemistry calculation. While the individual rank might not be completely accurate, as shown in Fig. 4B, the emerging pattern yields relevant and novel structure property trends as a direct consequence of alchemical chirality.

A common problem in materials design consists of identifying global optima. In Fig. 5, we also show the least and most stable BN-doped derivatives for each stoichiometry, as identified by ACE. As more and more carbon sites are BN doped, the sites interact more strongly, and patterns emerge that are in line with the aforementioned summary observations, e.g., the energetically unfavorable nature of nitrogen clusters. Thus, even if high accuracy is needed eventually, prefiltering with ACE can markedly accelerate the identification and discovery of candidate compound lists that are to be treated with higher level quantum methods subsequently.

DISCUSSION

We have introduced "alchemical" 4D chirality, resulting from a reflection in the nuclear charge space manifested by the external potentials in the electronic Hamiltonian. This symmetry relation is exact for the perturbing Hamiltonians of the corresponding enantiomers. The corresponding variational electronic energies are degenerate only up to the third order and reminiscent of parity violation (*58*), which also lifts the exact energy degeneracy between spatial enantiomers (albeit less by many orders of magnitude). Our numerical findings indicate that ACEs solely from symmetry considerations alone are sufficiently accurate to enable the exploration of substantial swaths of subdomains in CCS. From a practical point of view, both experiments and simulations have a resolution limit resulting from method uncertainty beyond which molecules or crystals are indistinguishable. Alchemical chirality offers a new way to find those symmetrically related enantiomers with practically identical energies from which only one needs to be considered in terms of measurements or calculations. In the current state, this could be applied to, e. g., surface adsorption problems such as surface catalysis or molecular sensing. Future work on bounds on approximate alchemical enantiomers for relaxed geometries might give access to weak interaction problems as they occur in molecular self-assembly.

Specifically, we have provided novel ACE-based bonding rules of chemical bonds and angles. Numerical evidence for single bonds of second- and third-row main-group elements even suggests that DFT level of accuracy can be reached with these rules. CCSD-based perturbation theory results for THDADB (and all other possible BN-doped benzene derivatives) have served the illustration of alchemical chirality, indicating near degeneracy for the electronic energies of alchemical enantiomers, deviating by roughly two orders of magnitude less than a covalent bond. Correspondingly, ACE can serve the energy ranking of more complex systems, as discussed for the two use cases of BN doping in naphthalene and in picene. Using CCSD/cc-pVDZ as a reference, the comparison to bond counting, semiempirical DFT, generalized gradient–approximated DFT, and hybrid DFT results for the more than 2000 naphthalene derivatives indicates that the alchemical chirality–based ranking outperforms bond counting and common semiempirical quantum chemistry (PM6), while approaching the performance of semiempirical DFT (xTB) in terms of fidelity—at negligible computational cost and without empiricism. For the BN doping of picene, an ACE of more than 400 million unique derivatives has enabled the establishment of structural trends and the identification of the least and most stable derivatives for each stoichiometry. We stress that for all the alchemical chirality–derived results presented, not a single quantum chemistry calculation was necessary.
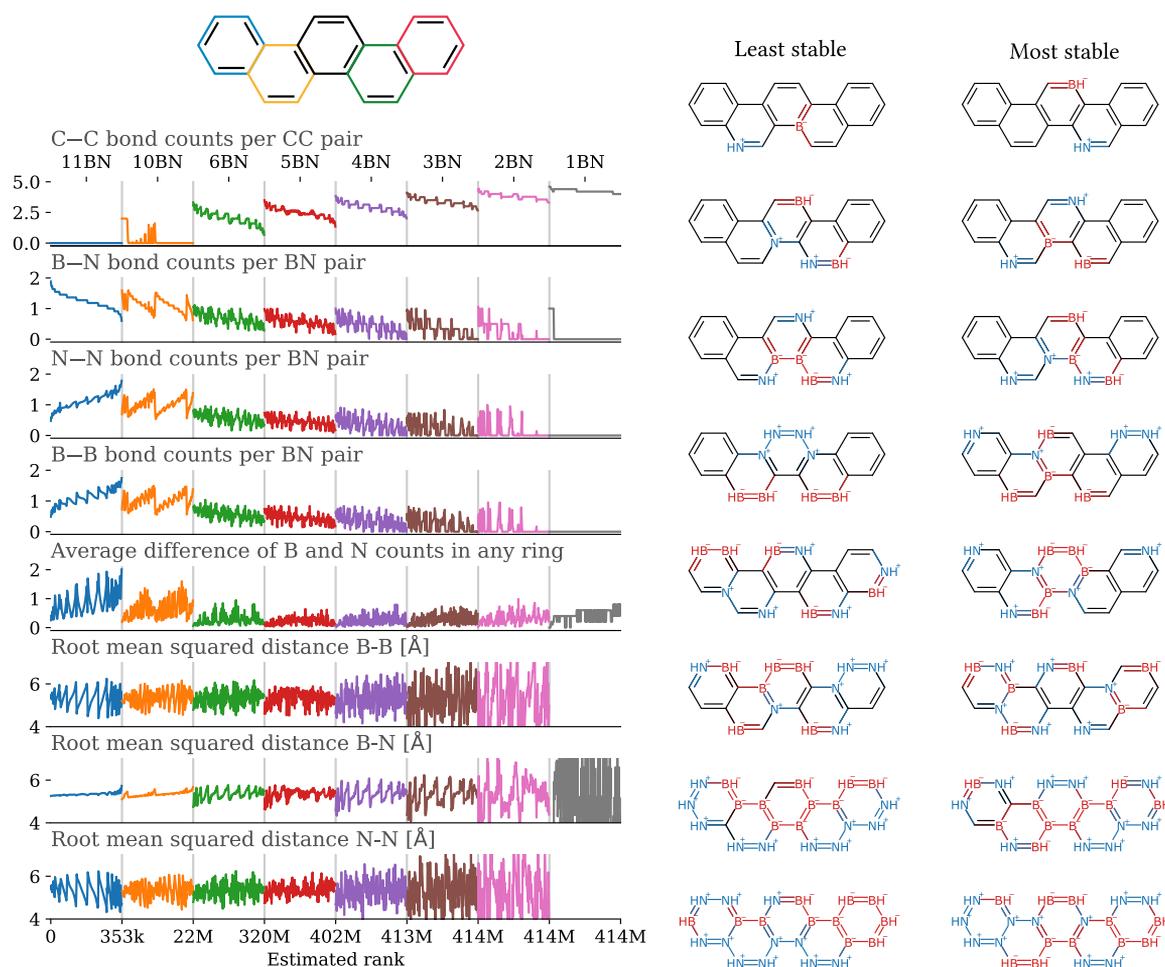
Fig. 5. Trends among all 414 million BN-doped picenes of the select stoichiometries (top) ranked by their total energy on the basis of ACE. Each row shows a structural feature for each stoichiometry as a function of the ACE rank and averaged over more than 200 bins. Those molecules that are the most or least stable for each stoichiometry are listed separately.

Overall, our arguments and numerical results indicate that the concept of alchemical chirality represents a novel, fundamental, and useful symmetry relation in CCS. Its power to dissect, group, and rank throughout the CCS of constitutional isomers holds great promise to further progress toward the overall goal of virtual computational materials discovery. Future work will deal with current limitations, such as the necessity to perform alchemical changes only in close vicinity in nuclear charge space, to restrict changes to be isoelectronic and neutral, or to rely on scaffold lattices with fixed nuclear positions. It would also be interesting to study how alchemical chirality can be exploited using quantum machine learning models.

MATERIALS AND METHODS
Previous methodological work tackling CCS from first principles through variable (alchemical) nuclear charges included 4D DFT (*59*), the use of thermodynamic integration (*60*), trends among vertical energies (*50*), entire potential energy surfaces (*61*), nuclear grand-canonical ensembles (*62*, *63*), linear combinations of atomic potentials (*64*), and APDFT (*31*). Starting in 1996 with stability of solid solutions (*65*), multiple promising applications, based on quantum alchemical changes, have been published over recent

years, including thermodynamic integrations (*66*), mixtures in metal clusters (*67*, *68*), reactivity estimates (*69*), water adsorption on BN-doped graphene (*70*), BN doping in fullerenes (*71*), and protonation energy predictions (*72*, *73*) However, apart from nearsightedness studies on chemical transferability (*74*), general rules rooted in the quantum mechanical framework of variable composition are mostly still lacking. Here, we describe alchemical chirality, defined by compositional reflection in the nuclear charge mirror plane of some reference system. Such a reflection defines alchemical enantiomers as distinct constitutional isomers with electronic energies being identical up to the third order. Alchemical chirality relates distinct molecules and materials in ways that, to the best of our knowledge, have not yet been discussed before.

Calculating the total potential energy of any compound U, most commonly obtained within the Born-Oppenheimer approximation and neglecting relativistic effects, represents the probably most crucial step in any atomistic simulation study. It consists of two contributions, the nuclear repulsion, which can be evaluated trivially, and the more complex electronic energy E, which, within the picture of DFT, sums up the electrons' kinetic contributions as well as their interactions with each other and with the nuclear charges. Hence, E is key and typically represents the principal goal of most modern

electronic structure theory developments, including improved DFT approximations. Furthermore, textbook discussions, such as the virial theorem in physical chemistry, deal with the discussion of the chemical bond in terms of the electronic energy.

The difference in electronic energy E between a reference reflection molecule, constituting a maximum in electronic energy, with electron density ρ_r and some "adjacent" isoelectronic alchemical enantiomer i, $\Delta E = E[\rho_i(\mathbf{r})] - E[\rho_r(\mathbf{r})]$, can be obtained through thermodynamic integration over the coupling constant $0 \leq \lambda_i \leq 1$, which linearly interpolates the nuclear charges between reflection molecule and alchemical enantiomer. According to Hellmann-Feynman, $\Delta E = \int_{-\infty}^{+\infty} d\mathbf{r}\, \Delta v_{ri}(\mathbf{r}) \int_0^1 d\lambda_i\, \rho(\mathbf{r}, \lambda_i)$, with $\Delta v_{ri}(\mathbf{r})$ as the difference in external potential between reflection molecule and alchemical enantiomer (31, 75).

Approximating this difference by a Taylor series expansion $\left(\approx \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{+\infty} d\mathbf{r}\, \Delta v_{ri}(\mathbf{r}) \partial_{\lambda_i}^{n-1} \rho(\mathbf{r}) \Big|_{\lambda=0} \right)$, subtracting the energy of the other alchemical enantiomer j (i.e., $\Delta v_{ri} = -\Delta v_{rj}$) and rearranging results in

$$\Delta E_{ij}^{\text{sym}} \approx \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{+\infty} d\mathbf{r}\, \Delta v_{ri}(\mathbf{r}) \left(\frac{\partial^{n-1} \rho_r(\mathbf{r})}{\partial \lambda_i^{n-1}} + \frac{\partial^{n-1} \rho_r(\mathbf{r})}{\partial \lambda_j^{n-1}} \right) \quad (2)$$

which is zero for all orders n as long as $\partial_{\lambda_i}^{n-1} \rho_r = -\partial_{\lambda_j}^{n-1} \rho_r$. Equation 2 must be exactly zero up to the third order because (i) the zeroth-order term $E[\rho_r]$ cancels, (ii) the first-order Hellmann-Feynman term $\int d\mathbf{r} \Delta v_{ri} \rho_r$ is zero for highly symmetric systems (as necessary to define an alchemical reflection plane) due to Δv and ρ_r being odd and even functions (46) and, (iii) due to the second-order term canceling because $\partial_\lambda \rho_r = \Sigma_I \Delta Z_I \partial_{Z_I} \rho_r$ differs for i and j only by the sign of ΔZ_I. Consequently, knowing all the higher-order contributions for one alchemical enantiomer also yields all higher-order terms for the other enantiomer without any further calculation. In other words, if a molecule contains a set of disjoint pairs of atoms of the same elements, which share nearly the same chemical environments, then it can be seen as the nuclear charge reflection plane, or chiral center, of all corresponding alchemical enantiomers. For example, the benzene molecule is the chiral center for three pairs of alchemical enantiomers (for $\Delta Z = \pm 1$, i.e., BN doping), all discussed in the Supplementary Materials and one highlighted in Figs. 3 and 1. For isoelectronic charge-neutral mutations, alchemical enantiomers differ from their chiral center only in their nuclear charges such that the net nuclear charge difference of each atom pair is zero. We used nauty (76) for graph enumeration and the Coulomb Matrix (77) as implemented in qmlcode (78) as similarity measure. Reference calculations were performed with Molpro (79) and MRCC (80, 81), in part using basis set extrapolation (82).

SUPPLEMENTARY MATERIALS

REFERENCES AND NOTES

1. P. Raccuglia, K. C. Elbert, P. D. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, A. J. Norquist, Machine-learning–assisted materials discovery using failed experiments. *Nature* **533**, 73–76 (2016).
2. L. Porwol, D. J. Kowalski, A. Henson, D.-L. Long, N. L. Bell, L. Cronin, An autonomous chemical robot discovers the rules of inorganic coordination chemistry without prior knowledge. *Angew. Chem. Int. Ed.* **59**, 11256–11261 (2020).
3. P. S. Gromski, J. M. Granda, L. Cronin, Universal chemical synthesis and discovery with the chemputer. *Trends Chem.* **2**, 4–12 (2020).
4. M. M. Flores-Leonar, L. M. Mejía-Mendoza, A. Aguilar-Granda, B. Sanchez-Lengeling, H. Tribukait, C. Amador-Bedolla, A. Aspuru-Guzik, Materials acceleration platforms: On the way to autonomous experimentation. *Curr. Opin. Green Sustain. Chem.* **25**, 100370 (2020).
5. B. P. MacLeod, F. G. Parlane, T. D. Morrissey, F. Häse, L. M. Roch, K. E. Dettelbach, R. Moreira, L. P. Yunker, M. B. Rooney, J. R. Deeth, V. Lai, G. J. Ng, H. Situ, R. H. Zhang, M. S. Elliott, T. H. Haley, D. J. Dvorak, A. Aspuru-Guzik, J. E. Hein, C. P. Berlinguette, Self-driving laboratory for accelerated discovery of thin-film materials. *Sci. Adv.* **6**, eaaz8867 (2020).
6. L. Pauling, D. M. Yost, The additivity of the energies of normal covalent bonds. *Proc. Natl. Acad. Sci. U.S.A.* **18**, 414–416 (1932).
7. E. N. Muratov, J. Bajorath, R. P. Sheridan, I. V. Tetko, D. Filimonov, V. Poroikov, T. I. Oprea, I. I. Baskin, A. Varnek, A. Roitberg, O. Isayev, S. Curtalolo, D. Fourches, Y. Cohen, A. Aspuru-Guzik, D. A. Winkler, D. Agrafiotis, A. Cherkasov, A. Tropsha, Qsar without borders. *Chem. Soc. Rev.* **49**, 3525–3564 (2020).
8. G. Schneider, Virtual screening: An endless staircase? *Nat. Rev.* **9**, 273–276 (2010).
9. A. Mullard, The drugmaker's guide to the galaxy. *Nature* **549**, 445–447 (2017).
10. S. R. Marder, D. N. Beratan, L.-T. Cheng, Approaches for optimizing the first electronic hyperpolarizability of conjugated organic molecules. *Science* **252**, 103–106 (1991).
11. G. Ceder, Computational materials science: Predicting properties from scratch. *Science* **280**, 1099–1100 (1998).
12. C. Wolverton, A. Zunger, First-principles prediction of vacancy order-disorder and intercalation battery voltages in Li$_x$CoO$_2$. *Phys. Rev. Lett.* **81**, 606–609 (1998).
13. A. Franceschetti, A. Zunger, The inverse band-structure problem of finding an atomic configuration with given electronic properties. *Nature* **402**, 60–63 (1999).
14. K. Alberi, M. B. Nardelli, A. Zakutayev, L. Mitas, S. Curtarolo, A. Jain, M. Fornari, N. Marzari, I. Takeuchi, M. L. Green, M. Kanatzidis, M. F. Toney, S. Butenko, B. Meredig, S. Lany, U. Kattner, A. Davydov, E. S. Toberer, V. Stevanovic, A. Walsh, N.-G. Park, A. Aspuru-Guzik, D. P. Tabor, J. Nelson, J. Murphy, A. Setlur, J. Gregoire, H. Li, R. Xiao, A. Ludwig, L. W. Martin, A. M. Rappe, S.-H. Wei, J. Perkins, The 2019 materials by design roadmap. *J. Phys. D Appl. Phys.* **52**, 013001 (2018).
15. P. Hohenberg, W. Kohn, Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964).
16. J. P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
17. J. Hafner, C. Wolverton, G. Ceder, Toward computational materials design: The impact of density functional theory on materials research. *MRS Bulletin* **31**, 659–668 (2006).
18. K. Burke, Perspective on density functional theory. *J. Chem. Phys.* **136**, 150901 (2012).
19. N. Marzari, Materials modelling: The frontiers and the challenges. *Nat. Mater.* **15**, 381–382 (2016).
20. O. A. von Lilienfeld, Quantum machine learning in chemical compound space. *Angew. Chem. Int. Ed.* **57**, 4164–4169 (2018).
21. O. A. von Lilienfeld, K.-R. Müller, A. Tkatchenko, Exploring chemical compound space with quantum-based machine learning. *Nat. Rev. Chem.* , 347–358 (2020).
22. K. Schütt, S. Chmiela, O. von Lilienfeld, A. Tkatchenko, K. Tsuda, K. Müller, *Machine Learning Meets Quantum Physics* (Lecture Notes in Physics, Springer International Publishing, 2020).
23. V. Botu, R. Ramprasad, Adaptive machine learning framework to accelerate ab initio molecular dynamics. *Int. J. Quantum Chem.* **115**, 1074–1083 (2015).
24. A. Glielmo, P. Sollich, A. De Vita, Accurate interatomic force fields via machine learning with covariant kernels. *Phys. Rev. B* **95**, 214302 (2017).
25. S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, K.-R. Müller, Machine learning of accurate energy-conserving molecular force fields. *Sci. Adv.* **3**, e1603015 (2017).
26. A. Grisa, D. M. Wilkins, G. Csányi, M. Ceriotti, Symmetry-adapted machine learning for tensorial properties of atomistic systems. *Phys. Rev. Lett.* **120**, 036002 (2018).
27. S. Chmiela, H. E. Sauceda, I. Poltavsky, K.-R. Müller, A. Tkatchenko, sgdml: Constructing accurate and data efficient molecular force fields using machine learning. *Comput. Phys. Commun.* **240**, 38–45 (2019).
28. A. S. Christensen, F. A. Faber, O. A. von Lilienfeld, Operators in quantum machine learning: Response properties in chemical space. *J. Chem. Phys.* **150**, 064105 (2019).
29. B. Huang, O. A. von Lilienfeld, Quantum machine learning using atom-in-molecule–based fragments selected on the fly. *Nat. Chem.* **12**, 945–951 (2020).
30. O. A. von Lilienfeld, First principles view on chemical compound space: Gaining rigorous atomistic control of molecular properties. *Int. J. Quantum Chem.* **113**, 1676–1689 (2013).
31. G. F. von Rudorff, O. A. von Lilienfeld, Alchemical perturbation density functional theory. *Phys. Rev. Res.* **2**, 023220 (2020).
32. O. Mó, M. Yáñez, M. Eckert-Maksić, Z. B. Maksić, I. Alkorta, J. Elguero, Periodic trends in bond dissociation energies. A theoretical study. *J. Phys. Chem. A* **109**, 4359–4365 (2005).
33. W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory* (Wiley-VCH, 2002).
34. K. Ryczko, P. Darancet, I. Tamblyn, Inverse design of a graphene-based quantum transducer via neuroevolution. *J. Phys. Chem. C* **124**, 26117–26123 (2020).

35. K. Saravanan, J. R. Kitchin, O. A. von Lilienfeld, J. A. Keith, Alchemical predictions for computational catalysis: Potential and limitations. *J. Phys. Chem. Lett.* **8**, 5002–5007 (2017).

36. C. D. Griego, K. Saravanan, J. A. Keith, Benchmarking computational alchemy for carbide, nitride, and oxide catalysts. *Adv. Theory Simulat.* **2**, 1800142 (2019).

37. C. D. Griego, J. R. Kitchin, J. A. Keith, Acceleration of catalyst discovery with easy, fast, and reproducible computational alchemy. *Int. J. Quantum Chem.* **121**, e26380 (2021).

38. H. Wang, J. Liang, Y. Wu, T. Kang, D. Shen, Z. Tong, R. Yang, Y. Jiang, D. Wu, X. Li, C.-S. Lee, Porous BN nanofibers enable long-cycling life sodium metal batteries. *Small* **16**, 2002671 (2020).

39. N. Mounet, M. Gibertini, P. Schwaller, D. Campi, A. Merkys, A. Marrazzo, T. Sohier, I. E. Castelli, A. Cepellotti, G. Pizzi, N. Marzari, Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds. *Nat. Nanotechnol.* **13**, 246–252 (2018).

40. S. Chakraborty, P. Kayastha, R. Ramakrishnan, The chemical space of B,N-substituted polycyclic aromatic hydrocarbons: Combinatorial enumeration and high-throughput first-principles modeling. *J. Chem. Phys.* **150**, 114106 (2019).

41. K. Y. Chang, O. A. von Lilienfeld, Quantum mechanical treatment of variable molecular composition: From 'Alchemical' Changes of state functions to rational compound design. *CHIMIA Int. J. Chem.* **68**, 602–608 (2014).

42. K. Y. S. Chang, S. Fias, R. Ramakrishnan, O. A. von Lilienfeld, Fast and accurate predictions of covalent bonds in chemical space. *J. Chem. Phys.* **144**, 174110 (2016).

43. M. to Baben, J. O. Achenbach, O. A. von Lilienfeld, Guiding ab initio calculations by alchemical derivatives. *J. Chem. Phys.* **144**, 104103 (2016).

44. A. Solovyeva, O. A. von Lilienfeld, Alchemical screening of ionic crystals. *Phys. Chem. Chem. Phys.* **18**, 31078–31091 (2016).

45. K. S. Chang, O. A. von Lilienfeld, $Al_xGa_{1-x}As$ crystals with direct 2 eV band gaps from computational alchemy. *Phys. Rev. Mater.* **2**, 073802 (2018).

46. S. Fias, K. S. Chang, O. A. von Lilienfeld, Alchemical normal modes unify chemical space. *J. Phys. Chem. Lett.* **10**, 30–39 (2019).

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

48. S. P. Ong, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, D. Bailey, D. Skinner, K. A. Persson, G. Ceder, *The Materials Project* (2011); http://materialsproject.org/.

49. P. G. Mezey, *J. Math. Chem.* **23**, 65–84 (1998).

50. P. G. Mezey, Electronic energy inequalities for isoelectronic molecular systems. *Theoretica Chimica Acta* **59**, 321–332 (1980).

51. E. E. Daza, A. Bernal, Energy bounds for isoelectronic molecular sets and the implicated order. *J. Math. Chem.* **38**, 247–263 (2005).

52. J. P. Perdew, M. Ernzerhof, K. Burke, Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **105**, 9982–9985 (1996).

53. M. Ernzerhof, G. E. Scuseria, Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional. *J. Chem. Phys.* **110**, 5029–5036 (1999).

54. C. Adamo, V. Barone, Toward reliable density functional methods without adjustable parameters: The pbe0 model. *J. Chem. Phys.* **110**, 6158–6170 (1999).

55. C. Bannwarth, S. Ehlert, S. Grimme, Gfn2-xtb an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).

56. J. J. P. Stewart, Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **13**, 1173–1213 (2007).

57. A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, W. M. Skid, UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **114**, 10024–10035 (1992).

58. M. Quack, Frontiers in spectroscopy. *Faraday Discuss.* **150**, 533–565 (2011).

59. E. B. Wilson Jr., Four dimensional electron density function. *J. Chem. Phys.* **36**, 2232–2233 (1962).

60. P. Politzer, R. G. Parr, Some new energy formulas for atoms and molecules. *J. Chem. Phys.* **61**, 4258–4262 (1974).

61. P. G. Mezey, New global constraints on electronic energy hypersurfaces. *Int. J. Quantum Chem.* **29**, 85–99 (1986).

62. O. A. von Lilienfeld, R. D. Lins, U. Rothlisberger, Variational particle number approach for rational compound design. *Phys. Rev. Lett.* **95**, 153002 (2005).

63. O. A. von Lilienfeld, M. E. Tuckerman, Molecular grand-canonical ensemble density functional theory and exploration of chemical space. *J. Chem. Phys.* **125**, 54104 (2006).

64. M. Wang, X. Hu, D. N. Beratan, W. Yang, Designing molecules by optimizing potentials. *J. Am. Chem. Soc.* **128**, 3228–3232 (2006).

65. N. Marzari, S. de Gironcoli, S. Baroni, Structure and phase stability of $Ga_xIn_{1-x}P$ solid solutions from computational alchemy. *Phys. Rev. Lett.* **72**, 4001–4004 (1994).

66. A. Beste, R. J. Harrison, T. Yanai, Direct computation of general chemical energy differences: Application to ionization potentials, excitation, and bond energies. *J. Phys. Chem.* **125**, 074101 (2006).

67. F. Weigend, C. Schrodt, R. Ahlrichs, Atom distributions in binary atom clusters: A perturbational approach and its validation in a case study. *J. Chem. Phys.* **121**, 10380–10384 (2004).

68. F. Weigend, Extending dft-based genetic algorithms by atom-to-place re-assignment via perturbation theory: A systematic and unbiased approach to structures of mixed-metallic clusters. *J. Chem. Phys.* **141**, 134103 (2014).

69. D. Sheppard, G. Henkelman, O. A. von Lilienfeld, Alchemical derivatives of reaction energetics. *J. Chem. Phys.* **133**, 084104 (2010).

70. Y. S. Al-Hamdani, A. Michaelides, O. A. von Lilienfeld, Exploring dissociative water adsorption on isoelectronically bn doped graphene using alchemical derivatives. *J. Chem. Phys.* **147**, 164113 (2017).

71. R. Balawender, M. Lesiuk, F. D. Proft, P. Geerlings, Exploring chemical space with alchemical derivatives: *BN*-simultaneous substitution patterns in c60. *J. Chem. Theory Comput.* **14**, 1154–1168 (2018).

72. G. F. von Rudorff, O. A. von Lilienfeld, Rapid and accurate molecular deprotonation energies from quantum alchemy. *Phys. Chem. Chem. Phys.* **22**, 10519–10525 (2020).

73. M. Muñoz, A. Robles-Navarro, P. Fuentealba, C. Cárdenas, Predicting deprotonation sites using alchemical derivatives. *Chem. A Eur. J.* **124**, 3754–3760 (2020).

74. S. Fias, F. Heidar-Zadeh, P. Geerlings, P. W. Ayers, Chemical transferability of functional groups follows from the nearsightedness of electronic matter. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11633–11638 (2017).

75. G. F. von Rudorff, O. A. von Lilienfeld, Atoms in molecules from alchemical perturbation density functional theory. *J. Phys. Chem. B.* **123**, 10073–10082 (2019).

76. B. D. McKay, A. Piperno, Practical graph isomorphism, II. *J. Symbolic Comput.* **60**, 94–112 (2014).

77. M. Rupp, A. Tkatchenko, K.-R. Müller, O. A. von Lilienfeld, Fast and accurate modeling of molecular atomization energies with machine learning. *Phys. Rev. Lett.* **108**, 058301 (2012).

78. A. S. Christensen, F. A. Faber, B. Huang, L. A. Bratholm, A. Tkatchenko, K.-R. Müller, O. A. von Lilienfeld, *Qml: A Python Toolkit for Quantum Machine Learning* (2017); https://github.com/qmlcode/qml.

79. H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, S. J. Bennie, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, S. J. R. Lee, Y. Liu, A. W. Lloyd, Q. Ma, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, T. F. Miller III, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, M. Welborn, Molpro, version 2018.1, a package of ab initio programs (2018).

80. P. R. Nagy, G. Samu, M. Kállay, Optimization of the linear-scaling local natural orbital CCSD(t) method: Improved algorithm and benchmark applications. *J. Chem. Theory Comput.* **14**, 4193–4215 (2018).

81. M. Kállay, P. R. Nagy, Z. Rolik, D. Mester, G. Samu, J. Csontos, J. Csóka, B. P. Szabó, L. Gyevi-Nagy, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei, B. Hégely, *Mrcc, A Quantum Chemical Program Suite* (2019); www.mrcc.hu.

82. R. M. Richard, M. S. Marshall, O. Dolgounitcheva, J. V. Ortiz, J.-L. Brédas, N. Marom, C. D. Sherrill, Accurate ionization potentials and electron affinities of acceptor molecules i. reference data at the CCSD(t) complete basis set limit. *J. Chem. Theory Comput.* **12**, 595–604 (2016).

83. G. F. von Rudorff, O. A. von Lilienfeld, *Solving the Inverse Materials Design Problem with Alchemical Chirality* (2020); https://zenodo.org/record/3994178#.YHXS8-gzaM8.

ScienceAdvances

Simplifying inverse materials design problems for fixed lattices with alchemical chirality

Guido Falk von Rudorff and O. Anatole von Lilienfeld