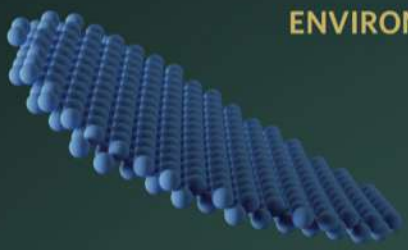
MARCH 2018 VOL 13 NO 3
www.nature.com/naturenanotechnology

nature nanotechnology

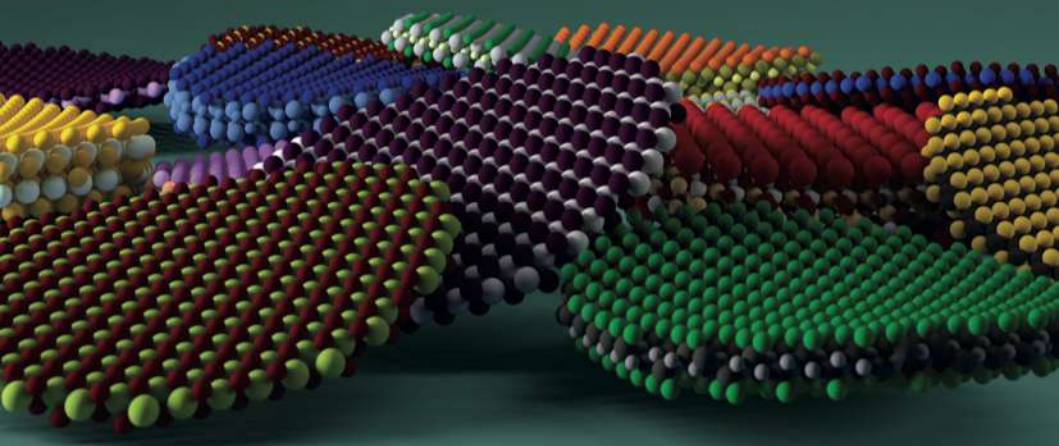
Computational quest for 2D materials



ENVIRONMENTAL NANOTECHNOLOGY
Interacting with the community

ACHROMATIC METALENSES
Visible images

NEUROMODULATION
Wireless excitement



2D MATERIALS

Searching for materials with reduced dimension

High-throughput electronic structure calculations, together with structural data-mining algorithms, allow the identification of new two-dimensional materials.

Olle Eriksson

Progress in materials science has been primarily driven by experimental exploration closely accompanied by theoretical work. In practice, apart from a few rare exceptions, theory seldom leads the way forward in finding new materials with desired functionality. Therefore, new tools are urgently needed to improve the rate at which new materials are being identified, and in this process theory is expected to play an important role. Now, writing in *Nature Nanotechnology*, N. Mounet and colleagues[1] have applied novel theoretical techniques, based on data-mining algorithms and advanced, high-throughput electronic structure theory, to identify a group of compounds that can serve as the basic platform for the synthesis of two-dimensional crystals for applications in nanotechnology.

The observed tunnelling magnetoresistance[2] in Fe/MgO/Fe sandwich structures, where a few layers of insulating MgO are inserted between thin films of Fe, was an experimental realization that followed theoretical predictions, based on calculations of the electronic structure[3,4]. From a practical point of view, this was a significant development because applications in nanotechnology, and particularly in sensing, could follow. Graphene — the celebrated two-dimensional material — where many of the observed exotic properties are the result of a linear dispersion of the electron states around the Fermi level[5], is another example where theory of the electronic structure preceded experimental measurements[6]. Yet the list of examples where theory actually makes predictions of new materials or new properties is short.

Mounet and co-workers suggest a way to theoretically identify new materials with layered structures, specifically two-dimensional functional materials. The approach can be briefly described as data-mining or data-filtering, as schematically shown in Fig. 1, where designed screening parameters are used to sift through thousands of compounds, in order to identify desired material characteristics.

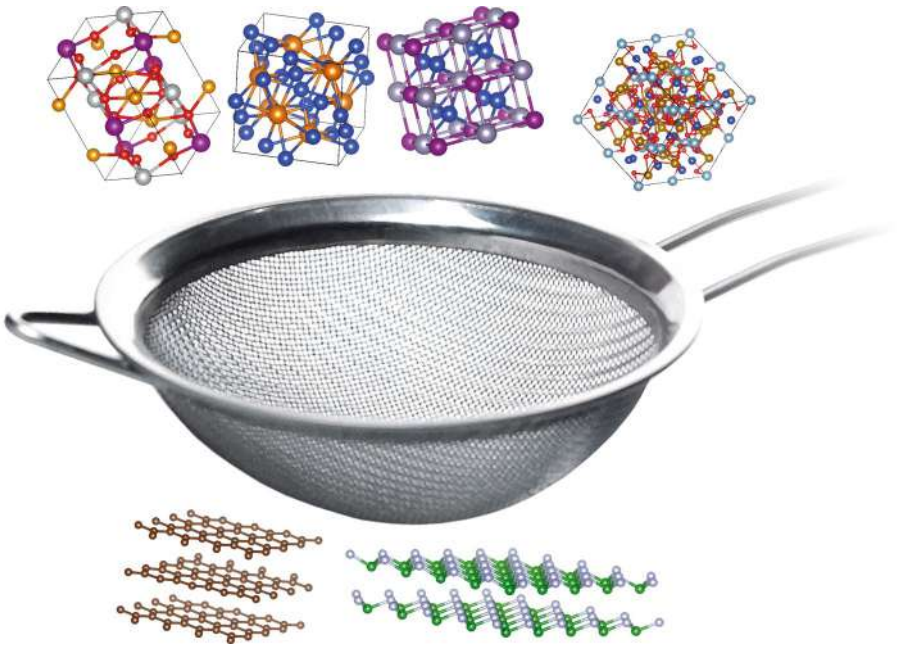


Fig. 1 | Schematic illustration of the method used by Mounet and co-workers. Here, high-throughput electronic structure calculations, together with structural data-mining algorithms, take the place of the sieve. Allowing for over 100,000 compounds (exemplified in the top part of the figure) to be filtered through this theoretical sieve, it has identified 258 prime candidates that may be synthesized in two-dimensional form (exemplified in the bottom part of the figure). Credit: Sieve: Liudmyla Lysenko / Alamy Stock Photo.

Starting from 108,423 experimentally known three-dimensional compounds, the authors distinguish 5,619 materials with layered crystal geometry. The list of potential two-dimensional materials is further reduced by means of high-throughput electronic structure calculations that help identify compounds with sufficiently weak interlayer binding. The final set of two-dimensional materials involves 258 entries, primarily distributed over ten structure types.

In addition, the researchers use accurate electronic structure theory to examine the mechanical stability, vibrational properties, electronic structure and possible magnetic ordering of these potentially exfoliable two-dimensional materials. Most of the proposed materials are found to be semiconducting (166). Furthermore, 56 of them are suggested to be magnetic and two were found to be quantum spin Hall insulators. The possibility of identifying materials where the functionality is primarily determined by the dimensionality is now significantly improved, and new materials in true two-dimensional form may be identified and studied. The properties that are likely to stand out from such investigations concern the electronic structure, where both practical aspects — such as, for example, light harvesting applications — and fundamental issues connected to correlations or screening are of interest. Collective phenomena associated with lattice dynamics and possible magnetic properties are also highly interesting, since they are expected to be uniquely coupled to the lattice dimension.

The hope is now that several of the suggested two-dimensional materials can be realized experimentally. However, both the synthesis and characterization may pose unexpected challenges. The pay-off for such studies is nevertheless high, since it would further add to the playground of exotic properties that is often associated with materials of reduced dimensionality, where the physical properties are expected to be dramatically different from those of the bulk counterparts. Therefore, this approach can prove even more fruitful than previous less-extensive theoretical methods[7] used to identify novel two-dimensional systems[8,9].

The approach used by Mounet and co-workers is naturally not limited to finding materials with reduced dimension, but can be used to identify functional materials in general, once a suitable filtering algorithm is designed, or an appropriate 'sieve' for a material with desired functionality is constructed. ❐

Olle Eriksson[1,2]

[1]Department of Physics and Astronomy, Uppsala University, Uppsala, Sweden. [2]School of Science and Technology, Örebro University, Örebro, Sweden.

e-mail: olle.eriksson@physics.uu.se

Published online: 05 February 2018

https://doi.org/10.1038/s41565-017-0060-4

References

1. Mounet, N. et al. *Nat. Nanotech.* https://doi.org/10.1038/s41565-017-0035-5 (2018).
2. Bowen, M. et al. *Appl. Phys. Lett.* **79**, 1655–1657 (2001).
3. Butler, W. H., Zhang, X.-G., Schulthess, T. C. & MacLaren, J. M. *Phys. Rev. B* **63**, 054416 (2001).
4. Mathon, J. & Umerski, A. *Phys. Rev. B* **63**, 220403 (R) (2001).
5. Wallace, P. R. *Phys. Rev.* **71**, 622–634 (1947).
6. Novoselov, K. S. et al. *Science* **306**, 666–669 (2004).
7. Lebègue, S., Björkman, T., Klintenberg, M., Nieminen, R. M. & Eriksson, O. *Phys. Rev. X* **3**, 031002 (2013).
8. Lin, M.-W. et al. *J. Mater. Chem. C* **4**, 315–322 (2016).
9. Gong, C. et al. *Nature* **546**, 265–269 (2017).

nature nanotechnology

ARTICLES

https://doi.org/10.1038/s41565-017-0035-5

Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds

Nicolas Mounet[1*], Marco Gibertini[1], Philippe Schwaller[1], Davide Campi[1], Andrius Merkys[1,2], Antimo Marrazzo[1], Thibault Sohier[1], Ivano Eligio Castelli[1], Andrea Cepellotti[1], Giovanni Pizzi[1] and Nicola Marzari[1*]

Two-dimensional (2D) materials have emerged as promising candidates for next-generation electronic and optoelectronic applications. Yet, only a few dozen 2D materials have been successfully synthesized or exfoliated. Here, we search for 2D materials that can be easily exfoliated from their parent compounds. Starting from 108,423 unique, experimentally known 3D compounds, we identify a subset of 5,619 compounds that appear layered according to robust geometric and bonding criteria. High-throughput calculations using van der Waals density functional theory, validated against experimental structural data and calculated random phase approximation binding energies, further allowed the identification of 1,825 compounds that are either easily or potentially exfoliable. In particular, the subset of 1,036 easily exfoliable cases provides novel structural prototypes and simple ternary compounds as well as a large portfolio of materials to search from for optimal properties. For a subset of 258 compounds, we explore vibrational, electronic, magnetic and topological properties, identifying 56 ferromagnetic and antiferromagnetic systems, including half-metals and half-semiconductors.

Two-dimensional (2D) materials provide opportunities to venture into largely unexplored regions of the materials space. On the one hand, their ultimate thinness makes them extremely promising for applications in electronics[1,2]. On the other, the physical properties of monolayers often change dramatically from those of their parent 3D materials, providing a new degree of freedom[3] for applications while also unveiling novel physics (for example, the valley Hall effect and composite excitations such as trions). Moreover, van der Waals (vdW) heterostructures[4] have recently emerged as an additional avenue to engineer new properties by stacking 2D materials in a desired fashion.

Progress in this area would be strongly accelerated by the availability of a broad portfolio of 2D candidate materials. To illustrate this point, we can compare the current situation for known 3D crystals, for which the knowledge accumulated in the past century (both crystal structures and measured physical properties) has been collected in databases such as the Pauling file[5], the Inorganic Crystal Structure Database[6] (ICSD) or the Crystallographic Open Database[7] (COD) (the latter two combined contain, to date, over half a million entries). In comparison, 2D materials databases are still scarce and limited in size: a first scan of the ICSD identified 92 2D compounds[8] (including Cu_2S, subsequently synthesized[9]). This was followed by 103 compounds selected among specific classes[10], while a recent study focused on transition-metal dichalcogenides and oxides, identifying 51 of them as stable[11]. More extensive efforts[12,13] have also been put forward to expand the set of prospective 2D materials by screening crystal structures from the Materials Project[14]. In fact, high-throughput computational methods represent a powerful tool[15] for exploring materials space and for screening materials without having to synthesize them first[16–20]. For instance, these techniques have been successfully employed in the search for materials for Li–air and Li-ion batteries[21,22], for hydrogen storage[23], scintillators[24], electrocatalysts[25], or to accelerate the discovery of light-absorbing materials[26].

Here, we systematically explore experimentally known compounds for possible exfoliation, paying particular attention to the mechanical stability of the exfoliated layers and the changes in the electronic structure that take place in reducing dimensionality—from the emergence of magnetic order to charge-density-wave instabilities. We perform this search starting from the inorganic compounds recorded in the ICSD and the COD, and then use vdW density functional theory (DFT) simulations to test these 3D parents for possible exfoliation. In particular, we compute the binding energy of 2,662 prospective layered structures and identify those that are held together by weak interactions and are ready for mechanical[27] or liquid-phase[28] exfoliation. This results in a portfolio of 1,825 materials that can be exfoliated in monolayers or multilayers. To showcase their potential, we explore the electronic, vibrational, magnetic and topological properties of 258 of the most promising systems, disclosing a number of functional materials that can be studied experimentally, notably including 56 magnetically ordered monolayers.

The reproducibility of all results is ensured by the deployment of the AiiDA[29] materials' informatics infrastructure, which keeps track of the full provenance of each calculation and result.

Identification of layered compounds

The search protocol starts from a comprehensive initial set of bulk 3D crystal structures extracted from the ICSD[6] and COD[7] databases,

[1]Theory and Simulation of Materials (THEOS) and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. [2]Vilnius University Institute of Biotechnology, Vilnius, Lithuania. *e-mail: nicolas.mounet@polytechnique.org; nicola.marzari@epfl.ch

Table 1 | Database statistics

	Unique to the ICSD	Unique to the COD	Common to both	Total sum
Experimental data				
CIF inputs	99,212	87,070		186,282
Unique 3D structures (set A)	34,548	60,354	13,521	108,423
Layered 3D structures (set B)	3,257	1,180	1,182	5,619
DFT calculations				
Layered 3D, relaxed (set C)	2,165	175	870	3,210
Binding energies (set D)	1,795	126	741	2,662
2D easily exfoliable (EE)	663	79	294	1,036
2D potentially exfoliable (PE)	524	34	231	789
Total	1,187	113	525	1,825

Experimental data: number of structures imported from the two databases (ICSD or COD, uniqueness not tested), number of unique 3D structures in each imported set or common to both (set A), and number of layered 3D structures identified using the geometrical criteria discussed in the text (set B). DFT calculations: number of structures that were relaxed (set C), number of structures that remain classified as layered after relaxation and for which binding energies were computed (set D), and number of easily or potentially (see text) exfoliable compounds.

containing, respectively, 177,343 (ICSD 2015.1) and 351,589 entries (COD revision 171462). Entries containing at most 6 different species are retrieved in the form of Crystallographic Information Files (CIF), correcting typical syntax errors[30]. We exclude structures with partial occupancies in the unit cell, those that are incompletely defined or are obviously incorrect, as well as theoretical structures that do not have an experimental counterpart. We then parse the CIF files using pymatgen[31], refine the atomic positions in the primitive cell with the spglib library[32], and finally remove duplicate structures (see Methods).

This procedure provides a starting point of 108,423 unique 3D crystal structures (set A, see Table 1), that can then be analysed to find promising candidates for exfoliation, that is, composed of chemically bonded manifolds held together by weak vdW interactions. Generalizing the pioneering effort of Lebègue and co-authors[8], we implement a protocol shown schematically in Fig. 1a and detailed in the Methods. Chemically bonded subunits are identified heuristically by comparing interatomic distances to the vdW radii[33] of the atoms involved. The dimensionality of each connected manifold is then obtained from the rank of the 3D lattice vectors embedded into the manifold (see Methods).

This approach is, in particular, able to deal with the following realistic situations: layered structures belonging to any crystal system (excluding cubic systems that cannot display a preferred exfoliation plane); arbitrary orientation of the stacking direction for the individual layers (Fig. 1b); the connectivity of the layers extending outside the 3D primitive unit cell, requiring the use of sufficiently large supercells (Fig. 1c); layers that are interpenetrating, that is, with overlapping projections along the stacking direction (Fig. 1d); composite structures that include manifolds with different dimensionalities (Fig. 1e); different chemical environments affecting the bond distances between atoms, requiring some tolerance in the parameters chosen to identify connected manifolds; and other low-dimensional units, such as 1D chains or 0D clusters.

By applying the full protocol to all 108,423 3D structures of set A, we find 5,619 materials that are classified as layered (set B in Table 1) and that are considered for further analysis.

Selection of exfoliable materials

Next, we need to determine if the interlayer interactions are weak enough for the parent compound to be a good candidate for exfoliation. The first step is the optimization of the 3D geometry (cell vectors and atomic positions) in set B, with a preference for the simplest structures, both for their higher relevance and for computational efficiency. As a result, at least all unary compounds containing less than 100 atoms in the unit cell, all binaries and ternaries with less than 40 atoms, and all other quaternary, quinary and senary compounds with less than 32 atoms are considered; in total, 3,210 layered compounds (set C in Table 1). These structures are relaxed using two different non-local vdW functionals: the vdW-DF2[34] with C09 exchange[35,36] (DF2-C09) and the revised Vydrov–Van Voorhis[37–39] (rVV10) functionals (see DFT computational details in the Methods). These functionals are validated against experimental structural parameters on a set of layered structures (Supplementary Fig. 1).

Then, we calculate the overall binding energy, E_b, as the difference between the ground-state energy of the relaxed 3D bulk structure and all of its unrelaxed substructures (of any dimensionality, including 1D chains or 0D clusters)[40]. To validate the quality of the vdW-DF2 and rVV10 functionals, we show in Fig. 2a the binding energies per layer and per unit surface calculated for well-known layered materials including graphite, hexagonal boron nitride and several transition-metal dichalcogenides. It can be seen that the values obtained with the two vdW functionals are in good agreement with accurate reference calculations performed in ref. [40] using the random phase approximation (RPA). We show in Fig. 2b the distribution of binding energies for the 2,662 structures that are still flagged as layered after vdW-DFT relaxation of set C; we label these as set D in Table 1. For both vdW functionals, a majority of the compounds (74% for DF2-C09 and 71% for rVV10) exhibit a binding energy smaller than 100 meV Å^{-2}, giving rise to the first peak in the distribution (note the logarithmic scale). A second broader peak appears between 200 and 400 meV Å^{-2} and is associated with structures containing lower-dimensionality units (that is, 0D or 1D) in addition to the 2D layer(s). Larger binding energies are to be expected for structures with mixed dimensionality (typically 0D and 2D) where charge transfer between intercalated units and 2D layers can take place, giving rise to ionic bonding. This means that while these additional 0D or 1D substructures have been separated from the 2D layers in the geometrical screening, the parent compounds might require different strategies (for example, infiltration) for exfoliation.

To further identify exfoliable materials, we correlate the binding energy with the relative discrepancy in equilibrium interlayer distance obtained in the structural relaxation of the 3D parents performed with vdW and non-vdW functionals. The non-vdW functional used was the revised Perdew–Burke–Ernzerhof functional (revPBE)[41], which, contrary to standard PBE[42], does not seem to display bonding interactions in the exchange energy[43]. We plot in Fig. 2c binding energies against this interlayer-distance difference for 1,535 compounds randomly chosen from set D that have been studied with the revPBE and DF2-C09 functionals, and in Supplementary Fig. 2 for the 1,482 compounds that have been studied with revPBE and rVV10 functionals. In both cases, we observe that most compounds with low binding energies exhibit a large relative discrepancy in interlayer distances when calculated with the vdW and non-vdW functionals. This is expected, since interlayer interactions are very sensitive to vdW forces, and it allows us to identify more clearly which layered compounds are characterized by dispersive interactions between layers. We thus choose thresholds of 30 meV Å^{-2} for binding energies computed with the DF2-C09 functional, and 35 meV Å^{-2} for those computed using rVV10, and we classify parent compounds falling below one or both of these thresholds as 'easily exfoliable' (EE; in blue); the label is also

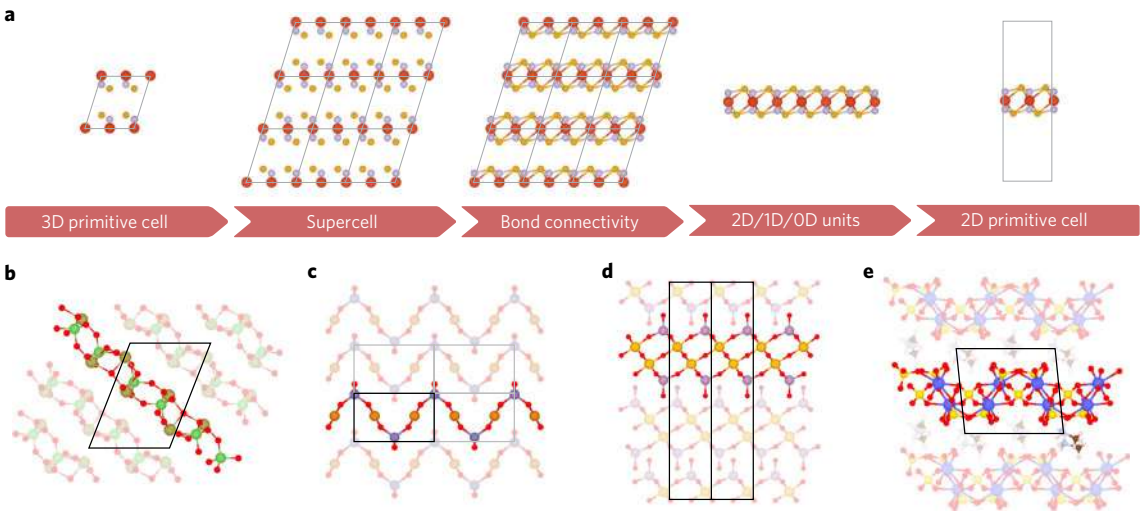


Fig. 1 | Screening for low-dimensional manifolds in a parent 3D crystal. a, Schematic showing the fundamental steps needed to find low-dimensional subunits of a parent 3D crystal (here $MgPS_3$). From left to right: a 3×3×3 supercell of the bulk primitive cell is constructed, chemical bonds are identified through the comparison of interatomic distances with the sum of vdW radii of the corresponding atoms minus a tunable parameter, low-dimensional subunits are identified from the rank of the lattice vectors embedded into chemically bonded units, and finally a 2D primitive cell is built. **b**–**e**, Examples illustrating non-trivial layered structures that can be identified: triclinic or monoclinic structures that are not layered along a standard crystallographic direction ($As_2Te_3O_{11}$, from ICSD #425299; **b**), layered compounds whose constitutive layers extend over multiple unit cells and thus require the use of supercells to be identified ($CuGeO_3$, from COD #1520902; **c**), layers that have partial overlap of the atomic projections along the stacking direction, with no manifest separation between them ($Mo_2Ta_2O_{11}$, from ICSD #247163; **d**), composite structures that contain subunits with different dimensionalities ($((CH_6N)_2(UO_2)_2(SO_4)_3$, from ICSD #192165, where 2D layers of uranyl sulfate are intercalated with 0D methylammonium molecules; **e**).

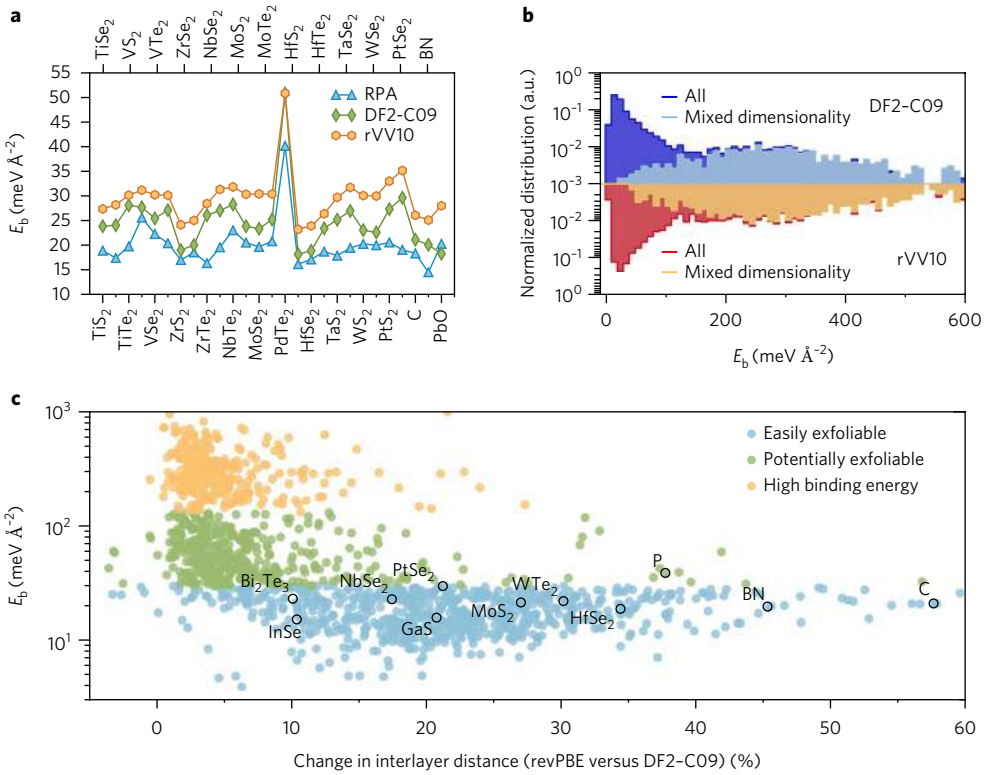


Fig. 2 | Binding energies of the bulk 3D compounds identified as geometrically layered. a, Binding energy, E_b, for a selection of layered materials identified in ref. [40], comparing results calculated with the random phase approximation (RPA) (ref. [40]) and with the DF2-C09 and rVV10 functionals (this work). **b**, Normalized distribution of all the 2,662 binding energies computed in this work (set D in Table 1). Results collected with the two different vdW functionals (DF2-C09 and rVV10) are reported in a specular fashion. Lighter colours highlight the distribution for the subset of structures that contain manifolds with different dimensionality. **c**, Binding energy versus change in interlayer distance for the 3D parent relaxed using the revPBE and the DF2-C09 functionals (1,535 structures; see Supplementary Fig. 2 for the 1,482 structures studied with revPBE and rVV10). Materials classified as easily exfoliable, potentially exfoliable, or with high binding energies are reported in different colours. Well-known 2D materials are highlighted.

applied to the resulting monolayers. As apparent from Fig. 2c and Supplementary Fig. 2, this choice identifies 2D materials commonly exfoliated in experiments[28], validating the approach.

In the top left corner of Fig. 2c and in Supplementary Fig. 2, a number of compounds exhibit high binding energies and a similar interlayer distance when comparing the revPBE and vdW functionals; many also contain substructures of mixed dimensionality, as discussed before. This group, shown in yellow, can be clearly separated from the other groups, the boundary being set at 130 meV $Å^{-2}$ for both vdW functionals. Above this value, compounds are not considered exfoliable and are excluded from the database.

Between these two regions, the remaining compounds (shown in green) exhibit relatively weak, possibly non-vdW, bonding. As an example, the compound $PdTe_2$ belongs to this region, and is metallic out-of-plane. We classify parents belonging to this group and their resulting monolayers as 'potentially exfoliable' (PE; in green). Last, materials for which the binding energy has been computed with both vdW functionals are classified according to the most optimistic prediction.

We note that 53 parent compounds can be exfoliated into two or more different monolayers.

The 2D database

Using these criteria, and after removal of duplicates (that is, identical 2D materials that come from different parents, see Methods), we obtain 1,036 easily exfoliable and 789 potentially exfoliable compounds, for a total of 1,825 candidates. In Fig. 3, we plot some statistics for this portfolio of promising materials, as well as for their 3D parents. We first show the distribution of packing ratios in the ICSD and COD structures (set A), in the group of 3D layered compounds (set B), and in the EE and PE compounds, using covalent radii[44] to compute the atomic volume. While the initial databases contain many closely packed structures with packing ratios of 0.7 or more, the layered structures peak around low packing ratios (0.3) but with a spread extending up to 0.8, even in the final set of exfoliable compounds. These considerations highlight how packing ratio alone is not necessarily a good criterion to filter 2D structures.

Figure 3b shows the distribution of point groups, arranged by crystal system, in all 3D structures, compared to that of the exfoliable 3D compounds. Most point groups are represented similarly in both sets, with the notable exception of all cubic point groups that are absent from the set of layered compounds. Moreover, the 222 point group is much rarer in the layered compounds than in the ICSD and COD databases, while the $3m$, $\bar{3}$, $\bar{3}m$ and $6mm$ point groups are more common among the 3D exfoliable compounds. It is interesting to note that while most of the 2D materials studied in the last ten years belong to the hexagonal crystal system, these structures represent only a minority of the full database of exfoliable materials.

Finally, in Fig. 3c we show how the layered and 2D compounds are distributed in terms of the number of species and number of atoms in the unit cell. As apparent from the graph, there are hundreds of 2D structures with few atoms in the unit cell (258 EE and 116 PE structures with at most 6 atoms, and 591 EE and 324 PE structures with at most 12 atoms), that can be used as a starting set for screening 2D materials with the desired properties. We find 18 unary monolayers (that is, containing a single species), including all known exfoliable 2D unaries such as graphene and phosphorene. Previously unreported unaries belong to the class of PE materials and mostly arise from intercalation layers in the parent 3D structures, although they are typically not mechanically stable. Interestingly, ternary 2D materials, which have been so far largely unexplored, represent a significant fraction of structures with less than 12 atoms in the unit cell, demonstrating that it may be worth inspecting these more versatile compounds beyond the current realm of unary/binary 2D materials.

We further classify the 2D materials of the EE group into different prototypes, according to their space groups and their structural similarities, considering all species undistinguishable and focusing only on the stoichiometry, using the pymatgen structure matcher (see Methods). 562 prototypes are found, among which the ten most common ones, representing a total of 214 structures, are shown in Fig. 4. The most common structural prototype is that of CdI_2, which groups together 64 similar structures, including many transition-metal dichalcogenides and dihalides. Contrary to the general distribution of point groups in Fig. 3b, most of these prototypes have high-order rotation axes, with the only exception being rectangular FeOCl. Although some of these structures have

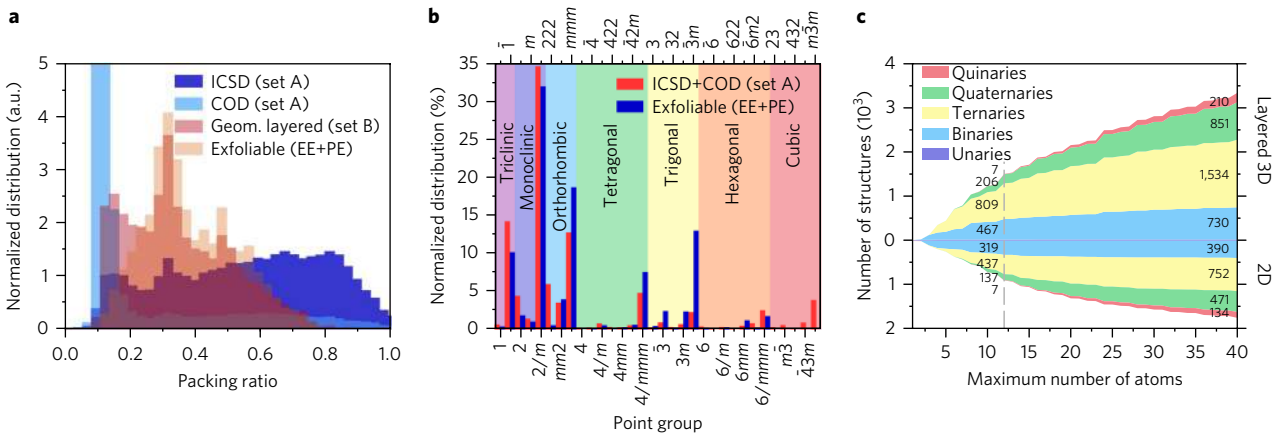


Fig. 3 | Statistics on the 2D and 3D databases. a, Packing ratios for the 3D structures considered according to their ICSD or COD provenance, layered compounds (set B) and exfoliable materials (EE+PE). Note that the peak of the COD materials at a packing ratio of 0.15—which extends beyond the upper limit of the graph—is due to the fact that COD is populated with a large number of molecular crystals, microporous networks, metal-organic frameworks and zeolites. **b**, Distribution of point groups in the set of structures exported from the ICSD and COD (set A) and in the set of easily and potentially exfoliable layered materials (EE+PE). The different colours mark the different crystal systems. **c**, Number of structures as a function of the number of atoms in the primitive cell (top, layered 3D structures; bottom, 2D structures in the EE+PE groups). The number of structures with at least two different species is explicitly reported for a maximum number of atoms in the primitive cell equal to 12 and 40. The number of unary structures with less than 12 (40) atoms in the primitive cell is 13 (15) for 3D layered compounds and 15 (15) for 2D EE+PE materials.

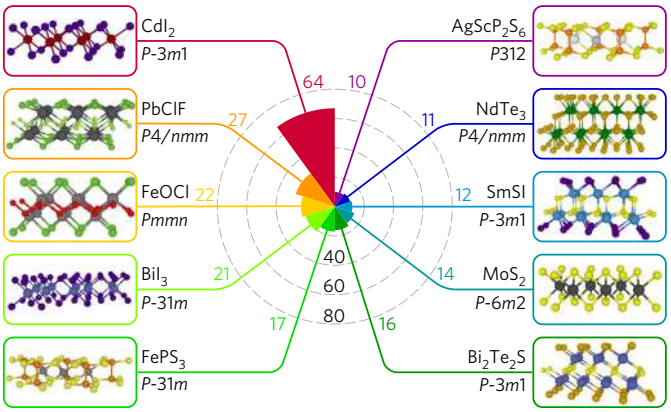


Fig. 4 | The most common 2D structural prototypes. Polar histogram showing the number of structures belonging to the ten most common 2D structural prototypes in the set of 1,036 easily exfoliable 2D materials. A graphical representation of each prototype is shown, together with the structure-type formula and the space group of the 2D systems.

already been suggested as exfoliable[8], the abundance of some relatively complex prototypes involving more than two species is new and compelling. Moreover, we emphasize that the $NdTe_3$ prototype has not been reported before and it is common to many rare-earth tritellurides.

Buoyed by these results, we extensively characterize here the most promising 2D structures, that is, the 258 easily exfoliable compounds with at most 6 atoms per unit cell. For this set, we study the mechanical stability and vibrational properties together with their electronic structure, exploring multiple ferromagnetic or antiferromagnetic states, to obtain the mechanical and electronic ground state of the exfoliated monolayers.

We start by assessing structural stability, as the existence of 2D subunits in a 3D parent compound does not guarantee that the isolated monolayer would still be stable. First, we compute phonon frequencies at the Γ point in the Brillouin zone and, if we find imaginary-frequency modes (this can happen if atoms are constrained by symmetries to be in a saddle point), we displace them along the modes' eigenvectors first (see Methods), and then fully relax the structure. The procedure is iterated until all phonon energies at the Γ point become real and positive. We then compute the full phonon dispersions in the correct 2D (out-of-plane open-boundary conditions) framework[45] (see Methods). This is necessary to correctly describe optical phonons in polar materials. All computed phonon dispersions are displayed in the Supplementary Information (we stress that all the phonon dispersions and relaxations are computed on the correct magnetic ground state, as discussed below).

More than 80% of the systems studied exhibited positive phonon dispersions, supporting the robustness of the exfoliation protocol; those systems that have soft phonons could be stabilized in supercells (although we did not attempt those calculations), but also display emerging phenomena, such as charge-density-wave instabilities as in the case of CoI_2. To further characterize these 2D materials, we also report their electronic band structure along high-symmetry lines, as calculated at the PBE[42] level. Most of the structures (166) are semiconducting, with 78 having bandgaps ranging from 0 to 1.5 eV, while 92 are metallic (see Supplementary Information).

To expand the family of known magnetic 2D materials[46–48], we fully characterized the electronic ground state of all 258 monolayers by comparing the energy of several magnetic configurations (including both ferromagnetic and antiferromagnetic supercell ordering) and hence identified the magnetic ground state of each 2D material (see Methods). A total of 56 compounds, reported in Table 2, are found to support non-trivial magnetic ordering, including magnetic insulators, half-semiconductors (for example, $CoCl_2$ and CrOBr) and half-metals (for example, $FeBr_2$), with possible applications in spintronics and data storage. The lowest-energy magnetic configuration and its corresponding magnetization for the 258 monolayers are reported in Supplementary Tables 1–3. We note that the recently synthesized ferromagnetic semiconductors CrI_3 (ref. [47]) and $CrGeTe_3$ (ref. [48]) are not present in this list as they contain more than 6 atoms per cell, although the database contains many more prospective magnetic 2D materials with the same structural prototype (for example, BiI_3 and $FePS_3$; Fig. 4).

For the insulating materials, we follow the evolution of the centres of the hybrid Wannier functions (see Methods), to look for $\mathbb{Z}_2$ topological order; excluding compounds with lanthanides, we identify only two quantum spin Hall insulators (QSHI) in the subset of 258 EE monolayers. The first QSHI found is pure bismuth, with space group 164 ($P\bar{3}m1$) and 2 atoms per unit cell, also known as bismuth bilayer or bismuthene[49]. The second one is TiNI, a ternary compound with space group 59 (*Pmmn*) and 6 atoms per unit cell, exhibiting a band inversion between the iodine *p* orbitals and titanium *d* orbitals gapped by spin–orbit coupling and was recently proposed as a QSHI[50]. In addition, we found systems (such as WTe_2 or $MoTe_2$) that are metallic at the DFT–PBE level, but whose band structure could in principle be adiabatically connected to a QSHI, for example, via a small strain or the interaction with a substrate.

In summary, we found a number of promising materials, including 37 ferromagnets (of which 13 are insulators), 19 antiferromagnets (of which 12 are insulators), 6 systems exhibiting sharp unstable phonon modes that could correspond to charge density waves, 2 quantum spin Hall insulators and 4 compensated semimetals (Supplementary Information). Such results hint at the usefulness of high-throughput studies in the search for novel high-performance 2D materials.

Conclusions

We identified an extensive database of exfoliable materials by systematic analysis and screening of experimental structures extracted from both the ICSD and the COD structural databases. We developed a geometrical algorithm to select potentially layered materials that is general enough to recognize substructures independently of their crystallographic orientation, dimensionality and environment. The list of 2D materials obtained by geometrical considerations was tested against accurate, validated vdW-DFT computations of the binding energies for all layers identified. Compounds with sufficiently small binding energies and where dispersion relations affect interlayer distances were deemed exfoliable. The materials identified are classified into groups of easily (1,036) or potentially (789)

Table 2 | Easily exfoliable magnetic compounds

	Ferromagnetic	Antiferromagnetic
Metals	$Co(OH)_2$, CoO_2, ErHCl, ErSeI, EuOBr, EuOI, $FeBr_2$, FeI_2, FeTe, LaCl, NdOBr, PrOBr, ScCl, SmOBr, SmSI, TbBr, TmI_2, TmOI, VS_2, VSe_2, VTe_2, YCl, YbOBr, YbOCl	CoI_2, $CrSe_2$, FeO_2, FeOCl, FeSe, PrOI, VOBr
Semiconductors	CdOCl, $CoBr_2$, $CoCl_2$, CrOBr, CrOCl, CrSBr, $CuCl_2$, ErSCl, HoSI, $LaBr_2$, $NiBr_2$, $NiCl_2$, NiI_2	$CrBr_2$, CrI_2, LaBr, $Mn(OH)_2$, $MnBr_2$, $MnCl_2$, MnI_2, VBr_2, VCl_2, VI_2, $VOBr_2$, $VOCl_2$

Chemical formulas of the easily exfoliable magnetic monolayers with at most 6 atoms in the unit cell. Compounds are grouped according to the type of magnetic ordering and the presence of an energy bandgap in the electronic structure. See also Supplementary Fig. 4.

exfoliable compounds, showing that only a very small fraction of possible 2D materials has been considered up to now. In particular, many opportunities might arise from materials with reduced symmetry or involving more than two atomic species but still with relatively simple structures; the identification of the most common prototypes could also allow to further expand the list of 2D materials by chemical substitutions and alternative site decorations.

For easily exfoliable compounds with up to 6 atoms per cell, we comprehensively characterized their vibrational, electronic, magnetic and topological properties, revealing a wealth of magnetic systems (56 magnetically ordered, including 37 ferromagnets, 19 antiferromagnets, 14 half metals and 6 half semiconductors) and highlighting a relative scarcity of insulators with $\mathbb{Z}_2$ topological order. All results can be fully reproduced thanks to the deployment of the AiiDA infrastructure that tracks the provenance of each data entry. To date, this is the largest available database of 2D compounds and it is available online on the Materials Cloud platform[51].

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41565-017-0035-5.

Received: 28 November 2016; Accepted: 20 November 2017;
Published online: 06 February 2018

References

1. Radisavljevic, B., Radenovic, A., Brivio, J., Giacometti, V. & Kis, A. Single-layer MoS_2 transistors. *Nat. Nanotechnol.* **6**, 147–150 (2011).
2. Chhowalla, M., Jena, D. & Zhang, H. Two-dimensional semiconductors for transistors. *Nat. Rev. Mater.* **1**, 16052 (2016).
3. Butler, S. Z. et al. Progress, challenges, and opportunities in two-dimensional materials beyond graphene. *ACS Nano* **7**, 2898–2926 (2013).
4. Geim, A. & Grigorieva, I. Van der Waals heterostructures. *Nature* **499**, 419–425 (2013).
5. Villars, P., Onodera, N. & Iwata, S. The Linus Pauling file (LPF) and its application to materials design. *J. Alloys Compd.* **279**, 1–7 (1998).
6. Inorganic Crystal Structure Database (ICSD); http://www.fiz-karlsruhe.com/icsd.html
7. Gražulis, S. et al. Crystallography open database (COD): an open-access collection of crystal structures and platform for world-wide collaboration. *Nucleic. Acids. Res.* **40**, D420–D427 (2012).
8. Lebègue, S., Björkman, T., Klintenberg, M., Nieminen, R. M. & Eriksson, O. Two-dimensional materials from data filtering and ab initio calculations. *Phys. Rev. X* **3**, 031002 (2013).
9. Romdhane, F. B. et al. Quasi-2D Cu_2S crystals on graphene: in-situ growth and ab-initio calculations. *Small* **11**, 1253–1257 (2015).
10. Miró, P., Audiffred, M. & Heine, T. An atlas of two-dimensional materials. *Chem. Soc. Rev.* **43**, 6537–6554 (2014).
11. Rasmussen, F. A. & Thygesen, K. S. Computational 2D materials database: electronic structure of transition-metal dichalcogenides and oxides. *J. Phys. Chem. C* **119**, 13169–13183 (2015).
12. Ashton, M., Paul, J., Sinnott, S. B. & Hennig, R. G. Topology-scaling identification of layered solids and stable exfoliated 2D materials. *Phys. Rev. Lett.* **118**, 106101 (2017).
13. Cheon, G. et al. Data mining for new two- and one-dimensional weakly bonded solids and lattice-commensurate heterostructures. *Nano Lett.* **17**, 1915–1923 (2017).
14. Jain, A. et al. Commentary: the materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
15. Gould, T., Lebègue, S., Björkman, T. & Dobson, J. in *Semiconductors and Semimetals 2D Materials* Vol. 95 (eds Iacopi, F. et al.) Ch. 1, 1–33 (Elsevier, 2016).
16. Franceschetti, A. & Zunger, A. The inverse band-structure problem of finding an atomic configuration with given electronic properties. *Nature* **402**, 60–63 (1999).
17. Johannesson, G. H. et al. Combined electronic structure and evolutionary search approach to materials design. *Phys. Rev. Lett.* **88**, 255506 (2002).
18. Curtarolo, S., Morgan, D., Persson, K., Rodgers, J. & Ceder, G. Predicting crystal structures with data mining of quantum calculations. *Phys. Rev. Lett.* **91**, 135503 (2003).
19. Curtarolo, S. et al. The high-throughput highway to computational materials design. *Nat. Mater.* **12**, 191–201 (2013).
20. Jain, A., Shin, Y. & Persson, K. A. Computational predictions of energy materials using density functional theory. *Nat. Rev. Mater.* **1**, 15004 (2016).
21. Mueller, T., Hautier, G., Jain, A. & Ceder, G. Evaluation of tavorite-structured cathode materials for lithium-ion batteries using high-throughput computing. *Chem. Mater.* **23**, 3854–3862 (2011).
22. Saal, J., Kirklin, S., Aykol, M., Meredig, B. & Wolverton, C. Materials design and discovery with high-throughput density functional theory: the Open Quantum Materials Database (OQMD). *JOM* **65**, 1501–1509 (2013).
23. Ozolins, V., Majzoub, E. H. & Wolverton, C. First-principles prediction of thermodynamically reversible hydrogen storage reactions in the Li-Mg-Ca-B-H system. *J. Am. Chem. Soc.* **131**, 230–237 (2009).
24. Ortiz, C., Eriksson, O. & Klintenberg, M. Data mining and accelerated electronic structure theory as a tool in the search for new functional materials. *Comput. Mater. Sci.* **44**, 1042–1049 (2009).
25. Greeley, J., Jaramillo, T. F., Bonde, J., Chorkendorff, I. & Norskov, J. K. Computational high-throughput screening of electrocatalytic materials for hydrogen evolution. *Nat. Mater.* **5**, 909–913 (2006).
26. Yu, L. & Zunger, A. Identification of potential photovoltaic absorbers based on first-principles spectroscopic screening of materials. *Phys. Rev. Lett.* **108**, 068701 (2012).
27. Novoselov, K. S. et al. Electric field effect in atomically thin carbon films. *Science* **306**, 666–669 (2004).
28. Nicolosi, V., Chhowalla, M., Kanatzidis, M. G., Strano, M. S. & Coleman, J. N. Liquid exfoliation of layered materials. *Science* **340**, 1226419 (2013).
29. Pizzi, G., Cepellotti, A., Sabatini, R., Marzari, N. & Kozinsky, B. AiiDA: automated interactive infrastructure and database for computational science. *Comput. Mater. Sci.* **111**, 218–230 (2016).
30. Merkys, A. et al. COD::CIF::Parser: an error-correcting CIF parser for the Perl language. *J. Appl. Crystallogr.* **49**, 292–301 (2016).
31. Ong, S. P. et al. Python Materials Genomics (pymatgen): a robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
32. Togo, A. spglib; https://atztogo.github.io/spglib/
33. Alvarez, S. A cartography of the van der Waals territories. *Dalton Trans.* **42**, 8617–8636 (2013).
34. Lee, K., Murray, É. D., Kong, L., Lundqvist, B. I. & Langreth, D. C. Higher-accuracy van der Waals density functional. *Phys. Rev. B* **82**, 081101 (2010).
35. Cooper, V. R. Van der Waals density functional: an appropriate exchange functional. *Phys. Rev. B* **81**, 161104 (2010).
36. Hamada, I. & Otani, M. Comparative van der Waals density-functional study of graphene on metal surfaces. *Phys. Rev. B* **82**, 153412 (2010).
37. Vydrov, O. A. & Van Voorhis, T. Nonlocal van der Waals density functional made simple. *Phys. Rev. Lett.* **103**, 063004 (2009).
38. Vydrov, O. A. & Van Voorhis, T. Nonlocal van der Waals density functional: the simpler the better. *J. Chem. Phys.* **133**, 244103 (2010).
39. Sabatini, R., Gorni, T. & de Gironcoli, S. Nonlocal van der Waals density functional made simple and efficient. *Phys. Rev. B* **87**, 041108 (2013).
40. Björkman, T., Gulans, A., Krasheninnikov, A. V. & Nieminen, R. M. Van der Waals bonding in layered compounds from advanced density-functional first-principles calculations. *Phys. Rev. Lett.* **108**, 235502 (2012).
41. Zhang, Y. & Yang, W. Comment on "generalized gradient approximation made simple". *Phys. Rev. Lett.* **80**, 890–890 (1998).
42. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
43. Dion, M., Rydberg, H., Schröder, E., Langreth, D. C. & Lundqvist, B. I. Van der Waals density functional for general geometries. *Phys. Rev. Lett.* **92**, 246401 (2004).
44. Cordero, B. et al. Covalent radii revisited. *Dalton Trans.* 2832–2838 (2008).
45. Sohier, T., Gibertini, M., Calandra, M., Mauri, F. & Marzari, N. Breakdown of optical phonons' splitting in two-dimensional materials. *Nano Lett.* **17**, 3758–3763 (2017).
46. Samarth, N. Condensed-matter physics: magnetism in flatland. *Nature* **546**, 216–218 (2017).
47. Huang, B. et al. Layer-dependent ferromagnetism in a van der Waals crystal down to the monolayer limit. *Nature* **546**, 270–273 (2017).
48. Gong, C. et al. Discovery of intrinsic ferromagnetism in two-dimensional van der Waals crystals. *Nature* **546**, 265–269 (2017).
49. Drozdov, I. K. et al. One-dimensional topological edge states of bismuth bilayers. *Nat. Phys.* **10**, 664–669 (2014).
50. Wang, A., Wang, Z., Du, A. & Zhao, M. Band inversion and topological aspects in a TiNI monolayer. *Phys. Chem. Chem. Phys.* **18**, 22154–22159 (2016).
51. Mounet, N. et al. Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds (data download). *Materials Cloud Archive* (2017); https://doi.org/10.24435/materialscloud:2017.0008/v1

Acknowledgements

This work was supported by the MARVEL National Centre of Competence in Research of the Swiss National Science Foundation. Simulation time was provided by the Swiss National

Supercomputing Centre (CSCS) under project IDs s580, mr0 and ch3, amounting to 60,000 DFT calculations and 5 million core hours. D.C., A.Ma. and N.Ma. gratefully acknowledge support from the EU Centre of Excellence MaX 'MAterials design at the eXascale' (grant no. 676598). D.C. acknowledges support from the 'EPFL Fellows' fellowship programme co-funded by Marie Skłodowska-Curie, Horizon 2020 grant agreement no. 665667. The authors would also like to acknowledge useful discussions with F. Ambrosio, and thank M. Giantomassi, M. J. van Setten and G. M. Rignanese for providing their fully relativistic ONCV pseudopotentials (https://github.com/abinit/pseudo_dojo).

Author contributions

M.G., G.P., N.Mo. and N.Ma. conceived the project. N.Mo., A.C., G.P., A.Me., T.S. and I.E.C. provided the necessary input, software tools and AiiDA workflows. N.Mo., P.S. and M.G. extracted and refined the structures from the source databases. P.S., N.Mo. and M.G. performed the geometrical screening of layered materials. N.Mo., D.C. and A.Ma. performed all first-principles simulations. N.Mo., M.G., G.P., D.C., A.Ma. and N.Ma. analysed the data. All authors contributed to the redaction of the manuscript.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41565-017-0035-5.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to N.M. or N.M.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Reproducibility and provenance. It is an often repeated tenet that results of scientific research must be reproducible. This objective is, however, challenging, especially in high-throughput research, due to the large number of simulations involved and the complex sequence of logical steps needed in any study. To ensure reproducibility, we use AiiDA[29] as an open-source materials' informatics infrastructure to implement the ADES model of automation, data, environment and sharing, as discussed in ref. [29]. All input information for each calculation is stored in the AiiDA repository before calculations are launched and actual code inputs are created by AiiDA using only data already stored in the repository. However, even the ability to replicate single calculations is not sufficient, since the majority of results are obtained via long sequences of simulations. It is thus essential to store the full provenance of each data item. This means, for instance, that for any 2D material identified in this work, one can track from which 3D parent structure it originated (including its source database), how the structure was processed to obtain its coordinates from the Wyckoff positions, how the low-dimensional units were identified, and which parameters were used to obtain the DFT results.

CIF files reformatting. The CIF entries are retrieved from crystallographic databases (ICSD[6], COD[7]) using routines distributed within AiiDA[29]. These entries sometimes contain syntax errors or unnecessary information that needs to be corrected and/or removed before parsing. To do so, we use cod-tools[30], in particular cif_filter and cif_select, to fix common CIF syntax errors, parse and reformat the summary chemical formula (from the tag _chemical_formula_sum), set explicit 90 degrees cell angles if they are not given (following the standard CIF specifications), remove empty non-loop tags, remove CIF data blocks without coordinates, and convert non-standard CIF tag name capitalization into the capitalization specified by the CIF dictionary; for example _atom_site_cartn_x is replaced by _atom_site_Cartn_x.

We also remove the tags _publ_author_name and _citation_journal_abbrev that often contain problematic characters for the pymatgen CIF parser (this information is still stored in the AiiDA database and therefore fully searchable).

As already mentioned in the main text, in this step we consider only structures coming from experimental measurements; nevertheless, the source databases are partially populated by purely theoretical structures. A selection is made using the flags set on the database entries by their curators. We also implemented some heuristics to detect clearly incorrect CIF files. For example, we discard structures where the chemical formula provided in the file is inconsistent with the elements in the unit cell. Regardless of all these efforts, it is possible that some incomplete or incorrect structures are still not filtered out from the original databases.

Crystal structure refinement. The crystal structures parsed from the CIF files are often subject to round-off errors in the atomic positions and/or cell parameters that artificially lower the number of symmetries. We use the spglib software[32] (version 1.8.2.2) to refine each structure and recover the maximum number of symmetries, using the algorithm described in ref. [52]. Interatomic distances and cell parameters are modified by at most 5×10^{-3} Å with respect to the initially parsed structure. At the same time (and with the same software), a standardized crystallographic primitive cell is extracted[53].

Removal of duplicates and prototyping. For the duplicate filtering procedure of both 3D and 2D structures, as well as for the prototyping of 2D structures, we have used extensively the structure matcher of pymatgen[31], which, thanks to the algorithm described in ref. [54], compares crystal structures. The goal is to (1) reduce the two structures to be compared to their primitive cells; (2) rescale the volumes (if 3D), or the in-plane areas (if 2D); (3) find, when it exists, an affine map between the two cells, within certain tolerances on the cell lengths and angles; and (4) compute the maximum distance between atomic sites when the two lattices are mapped onto a common average lattice. If this maximum is lower than a certain fraction of the average free length per atom, then the two structures are classified as similar.

When comparing 3D structures, the tolerances on cell lengths, cell angles and distances between paired atomic sites are chosen to be 20%, 5° and 30% of the average free length per atom, defined as the cubic root of the volume per atom. For 2D structures, due the artificial vacuum added to the cell in the out-of-plane direction the relative tolerances need to be set tighter, and we chose 10%, 1° and 10%, respectively.

To filter out duplicates of a set of structures, all compounds with the same composition (that is, the same species in the same relative proportion) are compared using the structure matcher algorithm, forming groups of matching structures. One unique structure is then chosen in each group as a representative for further calculations. For the initial 3D bulk compounds, the filtering of duplicates is performed first, independently, on each of the ICSD and COD sets; then, every structure of ICSD for which there exists a similar structure inside COD is also removed.

The same tool is used to prototype 2D structures, but with a number of modifications: (1) two structures must have the same space group to be considered as members of the same prototypical class (we used spglib, version 1.9.9, with a precision parameter of 0.5 Å to get the space group, see above); (2) the comparison is made irrespective of the specific species on each atomic site (for example, MoS_2 and WTe_2 are both considered as AB_2); (3) when the two primitive cells do not contain the same number of sites, one also attempts to generate a supercell of the smaller lattice before the mapping; and (4) the tolerances on the cell lengths, cell angles and maximum distance between sites are set to 10%, 5° and 20%, respectively. Again, all structures of the same kind of composition (for example, AB_2, A_2B_3, and so on) are compared in pairs, to classify them into distinct prototypical classes.

Identification of compounds containing low-dimensional units. The protocol is shown schematically in Fig. 1a. First, from the bulk primitive cell of the 3D candidate a $3\times3\times3$ supercell is created. All interatomic distances are evaluated and chemical bonds are heuristically identified as those for which

$$d_{i,j} < r_i^{\mathrm{vdW}} + r_j^{\mathrm{vdW}} - \Delta \quad (1)$$

where $d_{i,j}$ is the distance between two atoms i and j, r_i^{vdW} is the vdW radius of atom i according to the prescriptions of ref. [33] and Δ is a tunable parameter. The criterion in equation (1) is based on the statistical analysis of the distribution of over 5 million interatomic distances, according to which it was pointed out in ref. [33] that covalent bonds are likely to form only when the distance between atoms is shorter than the sum of the vdW radii by more than a quantity Δ, on average equal to 1.3 Å. Thus, if the interatomic distance is smaller than $r_i^{\mathrm{vdW}} + r_j^{\mathrm{vdW}} - \Delta$ the two atoms should be considered chemically connected. On the contrary, if the interatomic distance is larger, then strong chemical bonding is absent and only vdW interactions remain. Some special cases involving hydrogen or metallic bonding are still possible and discussed in ref. [33], but we will not consider them here; we have considered instead five equally spaced values for Δ between 1.1 Å and 1.5 Å to account for possible uncertainties in the values of the vdW radii and of Δ itself, and to maximize the number of candidates that would undergo the ultimate electronic-structure screening. We believe this criterion to be more robust and versatile than one based on covalent radii, as it allows to take into account the bond-length fluctuations associated with different chemical environments.

Once all bonds are identified, chemically connected groups are constructed. We start from $3\times3\times3$ supercells, as illustrated in Fig. 1c (even such supercells might not be large enough in principle), and identify the dimensionality of each connected manifold from the rank of the matrix formed by all the vectors linking an atom to all of its chemically connected periodic images (periodic according to the 3D lattice vectors).

This approach is able to identify subunits of any dimensionality: 2D layers, 1D chains and 0D clusters. In the case of 2D layers, the algorithm does not assume any specific orientation of the 2D plane with respect to the crystal or Cartesian axes.

DFT computational details. We use the Quantum ESPRESSO distribution[55] with the SSSP efficiency pseudopotentials' library (version 0.7)[56]. This library of extensively tested pseudopotentials from various sources[57–62] provides to date the best overall agreement with respect to all-electron calculations[63,64]. Wavefunction and charge-density cutoffs are chosen according to convergence tests with respect to cohesive energies, stresses and phonon frequencies performed for each individual element[56].

Two different vdW functionals are employed, namely the vdW-DF2-C09[34–36] and the revised VV10[37–39] that have been proven to perform well in layered systems[36,40,65,66], in addition to the PBE[42] and revPBE[41] functionals. These vdW functionals have a complex functional form[34,38,67] with a non-local dependence on density and parameters determined to reproduce specific functional dependencies or very accurate theoretical results. Sampling of the Brillouin zone is set using a Γ-centred Monkhorst-Pack grid, with an even number of k points in each direction of the reciprocal lattice such that the spacing between two consecutive k points along each direction is as close as possible to 0.2 $Å^{-1}$. All relaxations and binding energies of 3D structures are computed using a Marzari–Vanderbilt cold smearing[68] of 0.02 Ry, and without considering spin polarization or spin–orbit coupling (note that the binding energy of magnetic structures is not significantly affected by spin polarization; see Supplementary Fig. 3). When computing the ground-state energy of isolated 2D units, a vacuum space of 20 Å is set along the orthogonal direction to remove any fictitious interaction between periodic images of the 2D layers, while the atomic positions and in-plane cell parameters are kept fixed at the bulk values. In this particular case, because the energy of the 2D layer is directly compared with the energy of the parent bulk, it is important to use the same computational framework in 3D and 2D calculations.

All other electronic and vibrational properties of the 258 2D materials in the portfolio are computed in a different set-up. Namely, the following aspects are changed with respect to the previous set-up: (1) the functional, now set to PBE; (2) smearing is used only for metals while either totally suppressed or decreased by two orders of magnitude for insulators; (3) spin polarization is used for all the systems identified as magnetic (see below); and (4) the newly developed[69] 2D Coulomb cutoff for DFT and density-functional perturbation theory within Quantum ESPRESSO is used. This latter approach suppresses all interactions between periodic images, leading to the correct bidimensional framework

without the need for large supercell dimensions in the non-periodic direction, which reduces the computational costs, and, most importantly, allows to describe correctly the response to long-wavelength perturbations[45,70], such as optical phonons in polar materials.

The structural stabilization procedure follows the algorithm described in ref. [71]. In particular, phonons at Γ are computed (using the 2D Coulomb cutoff described above); one group of degenerate imaginary-frequency phonons is selected (the one with largest magnitude of the imaginary frequency), and a new structure is built where atoms are displaced following the eigenvectors of the unstable mode, choosing a maximum atomic displacement of 0.11 Å. If the mode is degenerate, a search is performed on all possible linear combinations of the two modes; the displacements with the largest number of symmetries are chosen. The cell is standardized using spglib[32,53] (with a symmetry precision of 0.05 Å), rotated if needed to have the non-periodic direction along z, and finally fully relaxed. The whole algorithm is repeated until all phonons at Γ are zero or positive real (this is needed when more than one group of degenerate unstable phonons is present).

The magnetic ground state of 2D compounds is identified thanks to an automated workflow that proceeds at first by screening purely ferromagnetic states, setting random initial spin states on each atomic species (identical for all the sites occupied by the same species). If after relaxation with spin-polarized DFT the compound ends up in a non-trivial magnetic state, this step is followed by a more complete screening of all possible antiferromagnetic states, performing calculations on supercells of twice the size of the unit cell when needed, while avoiding symmetry-redundant configurations (we use the algorithm of refs [72–74], through its interface with pymatgen[31]). Among all the configurations tested, the one with minimal energy per formula unit is selected.

QSHIs are identified by first-principles calculations of the $\mathbb{Z}_2$ invariant using Z2Pack[75,76], Wannier90[77] and Quantum ESPRESSO[55]. In the latter case, calculations are performed with the PBE functional and fully relativistic optimized norm-conserving Vanderbilt (ONCV) pseudopotentials[78] from the PseudoDojo library[63,79]. The screening for QSHIs is not performed on compounds containing lanthanides, due to the limited accuracy of standard DFT band structures in such cases and the presence of multiple magnetic minima.

Data availability. All the datasets generated and analysed during the current study for the structures with 6 atoms or less are included in this article and its Supplementary Information files, and are also available in the Materials Cloud archive repository[51]. Some of the original crystal structures that support the findings of this study are available from the ICSD but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are, however, available to licensed ICSD customers by using the ICSD ID number that we report in our data.

References

52. Grosse-Kunstleve, R. W. & Adams, P. D. Algorithms for deriving crystallographic space-group information. II. Treatment of special positions. *Acta. Crystallogr.* **A58**, 60–65 (2002).
53. Hinuma, Y., Togo, A., Hayashi, H. & Tanaka, I. Choice of basis vectors for conventional unit cells revisited. Preprint at http://arXiv.org/abs/1506.01455 (2015).
54. Hundt, R., Schön, J. C. & Jansen, M. CMPZ—an algorithm for the efficient comparison of periodic structures. *J. Appl. Crystallogr.* **39**, 6–16 (2006).
55. Giannozzi, P. et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).
56. Standard solid-state pseudopotentials (SSSP); http://www.materialscloud.org/sssp/
57. Garrity, K. F., Bennett, J. W., Rabe, K. M. & Vanderbilt, D. Pseudopotentials for high-throughput DFT calculations. *Comput. Mater. Sci.* **81**, 446–452 (2014).
58. Kucukbenli, E. et al. Projector augmented-wave and all-electron calculations across the periodic table: a comparison of structural and energetic properties. Preprint at http://arXiv.org/abs/1404.3015 (2014).
59. Dal Corso, A. Pseudopotentials periodic table: from H to Pu. *Comput. Mater. Sci.* **95**, 337–350 (2014).
60. Schlipf, M. & Gygi, F. Optimization algorithm for the generation of ONCV pseudopotentials. *Comput. Phys. Commun.* **196**, 36–44 (2015).
61. Willand, A. et al. Norm-conserving pseudopotentials with chemical accuracy compared to all-electron calculations. *J. Chem. Phys.* **138**, 104109 (2013).
62. Topsakal, M. & Wentzcovitch, R. Accurate projected augmented wave (PAW) datasets for rare-earth elements (RE = La-Lu). *Comput. Mater. Sci.* **95**, 263–270 (2014).
63. Lejaeghere, K. et al. Reproducibility in density functional theory calculations of solids. *Science* **351**, aad3000 (2016).
64. Lejaeghere, K., Van Speybroeck, V., Van Oost, G. & Cottenier, S. Error estimates for solid-state density-functional theory predictions: an overview by means of the ground-state elemental crystals. *Crit. Rev. Solid State Mater. Sci.* **39**, 1–24 (2014).
65. Björkman, T. Van der Waals density functional for solids. *Phys. Rev. B* **86**, 165109 (2012).
66. Björkman, T. Testing several recent van der Waals density functionals for layered structures. *J. Chem. Phys.* **141**, 074708 (2014).
67. Berland, K. et al. Van der Waals forces in density functional theory: a review of the vdW-DF method. *Rep. Prog. Phys.* **78**, 066501 (2015).
68. Marzari, N., Vanderbilt, D., De Vita, A. & Payne, M. C. Thermal contraction and disordering of the Al(110) surface. *Phys. Rev. Lett.* **82**, 3296 (1999).
69. Sohier, T., Calandra, M. & Mauri, F. Density functional perturbation theory for gated two-dimensional heterostructures: theoretical developments and application to flexural phonons in graphene. *Phys. Rev. B* **96**, 075448 (2017).
70. Sohier, T., Calandra, M. & Mauri, F. Two-dimensional Fröhlich interaction in transition-metal dichalcogenide monolayers: theoretical modeling and first-principles calculations. *Phys. Rev. B* **94**, 085415 (2016).
71. Togo, A. & Tanaka, I. Evolution of crystal structures in metallic elements. *Phys. Rev. B* **87**, 184104 (2013).
72. Hart, G. L. W. & Forcade, R. W. Algorithm for generating derivative structures. *Phys. Rev. B* **77**, 224115 (2008).
73. Hart, G. L. W. & Forcade, R. W. Generating derivative structures from multilattices: algorithm and application to hcp alloys. *Phys. Rev. B* **80**, 014120 (2009).
74. Hart, G. L., Nelson, L. J. & Forcade, R. W. Generating derivative structures at a fixed concentration. *Comput. Mater. Sci.* **59**, 101–107 (2012).
75. Soluyanov, A. A. & Vanderbilt, D. Computing topological invariants without inversion symmetry. *Phys. Rev. B* **83**, 235401 (2011).
76. Gresch, D. et al. Z2Pack: numerical implementation of hybrid Wannier centers for identifying topological materials. *Phys. Rev. B* **95**, 075146 (2017).
77. Mostofi, A. A. et al. An updated version of wannier90: a tool for obtaining maximally-localised Wannier functions. *Comput. Phys. Commun.* **185**, 2309–2310 (2014).
78. Hamann, D. R. Optimized norm-conserving Vanderbilt pseudopotentials. *Phys. Rev. B* **88**, 085117 (2013).
79. Pseudo-Dojo library; http://www.pseudo-dojo.org/