PAPER

Data-driven discovery of ultraincompressible crystals from a universal correlation between bulk modulus and volumetric cohesive energy

To cite this article: Xiaoang Yuan and Enlai Gao 2024 *J. Phys.: Condens. Matter* **36** 105702

IOP Publishing Journal of Physics: Condensed Matter
J. Phys.: Condens. Matter **36** (2024) 105702 (7pp) https://doi.org/10.1088/1361-648X/ad0d2a

Data-driven discovery of ultraincompressible crystals from a universal correlation between bulk modulus and volumetric cohesive energy

Xiaoang Yuan and Enlai Gao*

Department of Engineering Mechanics, School of Civil Engineering, Wuhan University, Wuhan, Hubei 430072, People's Republic of China

E-mail: enlaigao@whu.edu.cn

Received 5 October 2023, revised 31 October 2023
Accepted for publication 16 November 2023
Published 7 December 2023

Abstract

Bulk modulus and cohesive energy are two important quantities of condensed matter. From the interatomic energy landscape, we here derived a correlation between the bulk modulus (B) and the volumetric cohesive energy (ρ_e), i.e. $B = 2(\ln 2)^2\rho_e/9{\varepsilon_s}^2 = k\rho_e$, where ε_s and k are the strain-to-failure of interatomic bonds and the factor of proportionality, respectively. By analyzing numerous crystals from first principles calculations, it was shown that this correlation is universally applicable to various crystals including simple substances and compounds. Most interestingly, it was found that ε_s of crystals with a similar structure are almost a constant, resulting in a linear relationship between B and ρ_e. Furthermore, we found that the value of k for any compound can be determined based on the rule of mixtures, i.e. $k = \sum x_i k_i$, where x_i and k_i are the atomic fraction and the factor of proportionality for each element in this compound, respectively. Finally, this correlation was used to predict the bulk moduli for a vast number of crystals with known ρ_e in databases. After first principles verification of the top 50 crystals with the highest predicted bulk modulus, 25 ultraincompressible crystals with a bulk modulus greater than 400 GPa that can rival diamond (436 GPa) were discovered.

Supplementary material for this article is available online

Keywords: data-driven discovery, universal correlation, bulk modulus, cohesive energy, ultraincompressible crystals

1. Introduction

Bulk modulus not only characterizes the resistance of a material to elastic volume deformation, but also closely correlates to other important properties, such as toughness and hardness [1–5]. Ultraincompressible crystals have been widely used in many industrial fields including abrasives, cutting tools, and coatings [6, 7]. To discover ultraincompressible crystals, several approaches have been developed to determine the bulk modulus of a material. Experimental measures and first principles calculations are high-fidelity but costly methods,

* Author to whom any correspondence should be addressed.

especially for a vast number of crystals. To obtain the bulk modulus more efficiently, many attempts have been made to predict the bulk modulus from structural parameters or other physical quantities in the past decades [8–15]. Cohen [8–10] proposed semiempirical formulae for predicting the bulk moduli of tetrahedrally bonded crystals from their bond length and coordination number. Li *et al* [11] reported an electronegativity-related model for estimating the bulk moduli of crystals. Karman *et al* [12] established semiempirical formulae for bulk moduli and shear moduli of diamondlike and zinc-blende covalent crystals in terms of their bond length and ionicity.

More recently, the development of computational power has driven the establishment of first principles crystal databases. The materials project (MP) [16] and the open quantum materials database (OQMD) [17] are two of the largest crystal databases. By August 2023, there are over one million crystal structures that have been collected in MP and OQMD. However, the bulk moduli for more than 99% of crystals in these databases remain unknown. The high cost of experimental measures and first principles calculations and the weak universality of empirical or semiempirical formulae limit their uses for data-driven discovery of ultraincompressible crystals from these crystal databases. Although machine learning developed in recent years is a promising approach to this end [7, 13], machine learning-directed identified ultraincompressible crystals with a bulk modulus greater than 400 GPa are still very rare. For example, only one crystal with a bulk modulus greater than 400 GPa (bulk modulus of 401 GPa for Re_2C) was found in a machine learning search of 18 493 crystals [13]. This is because there are still critical limitations, such as a lack of powerful descriptors, and insufficient training, in the past machine learning models [18]. Fortunately, the cohesive energies of these crystals can be conveniently accessed from the provided ground state energies of crystals and the energies of isolated atoms. Intuitively, it would be more accurate and efficient to predict the bulk modulus, if the valuable information of cohesive energy can be rationally utilized.

Bulk modulus and cohesive energy are two important quantities of condensed matter [19]. The bulk modulus is determined by the interatomic interaction of condensed matter in the equilibrium state, while the cohesive energy is determined by the interatomic interaction of condensed matter in both the equilibrium and dissociation states. Since the correlation between the bulk modulus and the cohesive energy is of fundamental importance for condensed matter, considerable effort has been devoted to establishing the correlation between the bulk modulus and the cohesive energy [20–23]. A dimensional analysis shows that, instead of the cohesive energy, volumetric cohesive energy (volume density of the cohesive energy) has the same dimension as the bulk modulus, and thus one should search for the correlation between the bulk modulus and the volumetric cohesive energy. Plendl *et al* [20] proposed a formula that relates compressibility (the reciprocal of the bulk modulus) with the volumetric cohesive energy for dielectric solids. Wacke *et al* [21] found that the bulk modulus is proportional to the volumetric cohesive energy for metals by analyzing the experimental data available in the literature. However, these attempts are restricted to a narrow range of substances.

In this work, we derived a correlation between the bulk modulus (B) and the volumetric cohesive energy (ρ_e), i.e. $B = 2(\ln 2)^2\rho_e/9{\varepsilon_s}^2 = k\rho_e$, where ε_s is the strain-to-failure of interatomic bonds, from Morse potential that well describes the interatomic energy landscape for a wide range of condensed matter in both the equilibrium and dissociation states, and k is the factor of proportionality. By analyzing a large number of first principles calculated crystals, we found that this correlation is universally applicable to various crystals. For crystals with a similar structure, B is almost linear to ρ_e, since ε_s is almost a constant. Finally, we demonstrated that this universal correlation can be used for data-driven identification of ultraincompressible crystals, which directs the discovery of 25 ultraincompressible crystals with a bulk modulus greater than 400 GPa.

2. Results and discussion

Volumetric cohesive energy (ρ_e) is defined as the ratio of the atomic cohesive energy [cohesive energy per atom (E_c)] to the atomic volume [equilibrium volume per atom (V)], which is related to the interatomic interaction of condensed matter in both the equilibrium and dissociation states. The bulk modulus is defined as $B = V\mathrm{d}^2E/\mathrm{d}V^2$, where E is the atomic cohesive energy function, which is related to the interatomic interaction of condensed matter in the equilibrium states. To correlate these two properties, a classical potential, i.e. Morse potential, that well describes the interatomic interaction for a wide range of condensed matter in both the equilibrium and dissociation states [24–26], is adopted to map the potential energy landscape of chemical bonds. The Morse potential energy function is of the form:

$$E_b(d) = D\left[1 - e^{-\alpha(d-d_0)}\right]^2, \tag{1}$$

where D determines the depth of the potential well, α is a stiffness parameter that controls the width of the potential [$\alpha = \ln 2/(d_m - d_0)$, where d_m is the bond length corresponding to the peak stretching force] [27], d_0 is the equilibrium bond length. For any crystal, the atomic cohesive energy, atomic cohesive energy function, and atomic volume can be expressed as $E_c = (N_c/2)D$, $E = (N_c/2)E_b$, and $V = m{d_0}^3$, respectively, where N_c is the coordination number and m is the ratio of the atomic volume to the cube of equilibrium bond length. Combining the expressions of E, V, and the definition of bulk modulus, the bulk modulus is derived as (see supplementary data for details)

$$B = V\frac{\mathrm{d}^2E}{\mathrm{d}V^2} = \frac{N_c}{9m}\frac{\alpha^2 D}{d_0}. \tag{2}$$

The volumetric cohesive energy can be derived as

$$\rho_e = \frac{E_c}{V} = \frac{N_c}{2m}\frac{D}{{d_0}^3}. \tag{3}$$

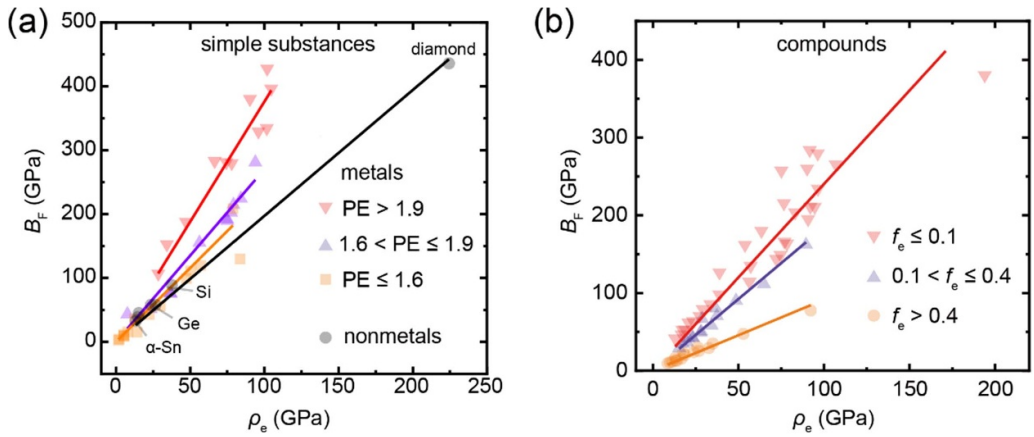


Figure 1. (a) Correlation between first principles calculated bulk modulus (B_F) and volumetric cohesive energy (ρ_e) for simple substances. The metals are divided into three groups with PE $>$ 1.9, 1.6 $<$ PE $\leqslant$ 1.9, and PE $\leqslant$ 1.6. (b) Correlation between B_F and ρ_e for compounds, which are divided into three groups with $f_e \leqslant 0.1$, $0.1 < f_e \leqslant 0.4$, and $f_e > 0.4$.

Therefore, the bulk modulus can be correlated with the volumetric cohesive energy as

$$B = k\rho_e, \tag{4}$$

where the factor of proportionality (k) is derived as

$$k = \frac{B}{\rho_e} = \frac{2\alpha^2 {d_0}^2}{9} = \frac{2(\ln 2)^2}{9} \cdot \frac{1}{{\varepsilon_s}^2}. \tag{5}$$

Here $\varepsilon_s = (d_m - d_0)/d_0$ represents the strain-to-failure of interatomic bonds. Equation (5) indicates that k solely depends on ε_s. From these derivations, it can be concluded that the correlation between bulk modulus and volumetric cohesive energy is applicable to any condensed matter on the condition that the potential energy landscape of chemical bonds can be approximated by the Morse potential. Fortunately, it has been demonstrated that the Morse potential can well describe the interatomic interaction for various condensed matter in both the equilibrium and dissociation states [28–30]. Therefore, a strong universality of the correlation between bulk modulus and volumetric cohesive energy can be expected. Based on this correlation, we can predict the bulk modulus using the volumetric cohesive energy. Since the volumetric cohesive energy captures the essence, such prediction of bulk modulus is rational in diverse scenarios. For example, even for α-iron under pressure, this correlation still remains valid and thus the predicted bulk moduli agree well with first principles calculations (figure S1). Moreover, considering that the volumetric cohesive energy encompasses the contributions of the outermost shell electrons and the inner shell electrons, the prediction of the bulk modulus using this correlation implicitly includes the influence of both the outermost shell electrons and the inner shell electrons, while it avoids the controversy of distinguishing the contributions of the outermost shell electrons and the inner shell electrons.

The following investigation supports the above expectation and further demonstrates that ε_s and k are near constants for crystals with a similar structure. The investigated crystals include simple substances (nonmetals and metals) and compounds (rock-salt and zinc-blende compounds, and transition-metal carbides and nitrides), and the bulk moduli and volumetric cohesive energies of crystals were obtained by first principles calculations. For nonmetals (figure 1(a)), the fitted value of k is 1.97, and the correlation coefficient is as high as 0.99. For metals (figure 1(a)), the diverse behaviors are more interesting. Based on the Pauling electronegativity (PE) of atoms [31], we divided metals into three groups with PE $>$ 1.9, 1.6 $<$ PE $\leqslant$ 1.9, and PE $\leqslant$ 1.6, respectively. The fitted values of k are 3.76, 2.70, and 2.31, respectively, and the corresponding correlation coefficients are as high as 0.99, 0.99, and 0.98. Since PE describes the power of an atom to attract electrons, the increase of PE is expected to improve the stiffness and reduce the strain-to-failure of metallic bonds, which accounts for the increase of k with the increase of PE (equation (5)). For compounds, based on the ionicity (f_e) with the definition of $f_e = \frac{(\chi_b - \chi_a)}{Z_a(\chi_a + \chi_b)}$ (χ_a and χ_b are the PE of the atoms, and Z_a is the valence state of cation) [11], we divided the compounds into three groups with $f_e > 0.4$, $0.1 < f_e \leqslant 0.4$, and $f_e \leqslant 0.1$, respectively (figure 1(b)). For these groups, the fitted values of k are 0.92, 1.85, and 2.40, respectively, and the corresponding correlation coefficients are as high as 0.99, 0.99, and 0.97. As shown in figure 1(b), it can be found that k decreases with the increase of f_e. This is because the increase of ionicity would induce asymmetry distributions of valence electrons in chemical bonds, which is expected to reduce the stiffness and increase the strain-to-failure of interatomic bonds, resulting in the decrease of k (equation (5)). In summary, the high correlation coefficients ($\geqslant$0.97) indicate the strong correlation between the bulk modulus and the volumetric cohesive energy for simple substances and compounds.

To further provide support for this correlation, we calculated the values of k and ε_s for numerous crystals. The values of ε_s were extracted from the Morse potential that was fitted with first principles calculated volumetric energy–strain curves of crystals (table S1). These energy–strain curves were calculated by applying equiaxial strain (volumetric strain) to these crystals. As shown in figure 2(a), we plotted the values

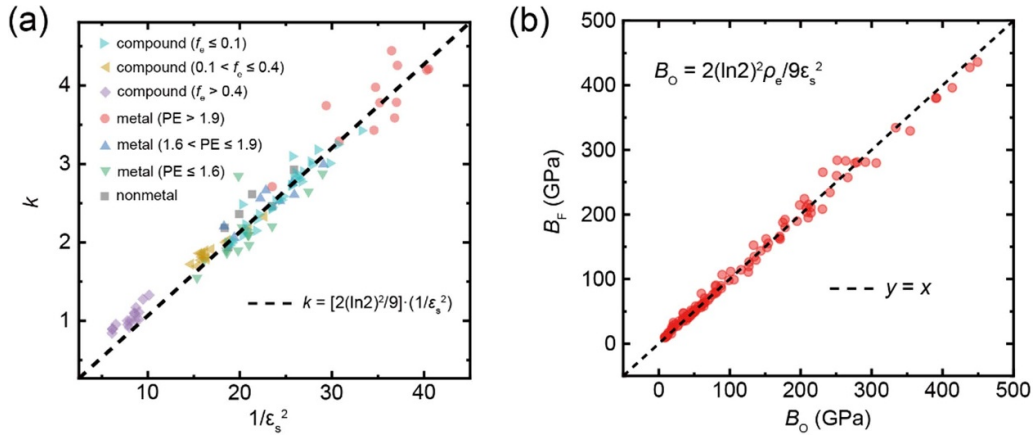


Figure 2. (a) k of simple substances and compounds as a function of $1/\varepsilon_s^2$. The dashed line represents the correlation between k and ε_s of equation (5). (b) Comparison of the first principles calculated bulk moduli (B_F) and the predicted bulk moduli from our formula (B_O) by equations (4) and (5) for simple substances and compounds.

of k as a function of $1/\varepsilon_s^2$ for these crystals, which is in good agreement with the theory (equation (5)). As these crystals are categorized into the groups with a similar structure (i.e. subgroups of simple substances, and compounds), ε_s and k fall within narrow sub-ranges. These results further indicate this theory is universally applicable to a wide range of crystals. On the other hand, the bulk moduli of these crystals can be predicted by substituting the known values of ρ_e and k into equation (4). The predicted bulk moduli from our formula are well consistent with the first principles calculated bulk moduli (figure 2(b)). These results indicate that this correlation can be used for fast and accurate prediction of the bulk moduli of crystals.

To predict B for a crystal (simple substances and compounds) with known ρ_e using the correlation ($B = k\rho_e$), the most important procedure is to determine the value of k for this crystal. However, the above procedure is still rough and complex, since it depends on how to categorize the crystals for determining the values of k. To solve this issue, we here proposed a simplified procedure to determine the value of k for a crystal, i.e. $k = \sum x_i k_i$, where x_i and k_i are the atomic fraction and the factor of proportionality for each element in this crystal, respectively, based on the rule of mixtures [32]. k_i for each element within the periodic table was calculated by substituting the α and d_0 of the corresponding simple substance into equation (5) (figure S2). Compared to the procedure of categorizing crystals for determining the values of k, this procedure uses the information of each element therein in a fine manner, which improves the predictive accuracy. To confirm the accuracy of this method, we calculated the k-values for 3765 crystals having first principles calculated bulk modulus in MP database by this method ($k = \sum x_i k_i$) and first-principles calculations ($k = B/\rho_e$). The mean absolute relative error (MARE) of k-values between our method and first-principles calculations is 17.0%, indicating the accuracy of this method. Meanwhile, we note that some flaws also exist in the current method. For example, we parameterized one value of k for each element, and for crystals having identical chemical compositions but different structures or phases, the determined values of k are identical using this method while they might be affected by structures or phases. For example, the determined values of k using this method for cubic-C_3N_4 (OQMD-14 925) and β-C_3N_4 (OQMD-14 924) are both 2.66, while the first principles values are 2.72 and 2.65, respectively. To address this issue, a possible solution in the future is to divide each element into distinct atomic subtypes based on local chemical environments (e.g. coordination numbers, valence states) and parameterize an individual value of k for each atomic subtype. Based on this procedure, the bulk modulus for a crystal can be predicted by substituting the known value of ρ_e and the corresponding value of k into equation (4).

We here provided additional evaluations on the performance of our formula as compared with previous approaches (i.e. empirical/semiempirical formulae, machine learning models, and first principles calculations). Among these approaches, first principles calculation is the most trustworthy yet high-cost approach for predicting the bulk modulus. Herein, we predicted bulk moduli for 3765 crystals in MP database that have a first principles calculated bulk modulus using previous empirical/semiempirical formulae (Cohen's formula [9], Li's formula [11]), our formula, and machine learning model. It should be noted that the previous empirical/semiempirical formulae were extended to these 3765 crystals for comparison, and the machine learning model was built by referring to previous works [7, 13] (see supplementary data for details). Compared to first principles calculations, the MARE of Cohen's formula [9], Li's formula [11], our formula, and the machine learning model for all 3765 crystals are 41.7%, 41.9%, 17.0%, and 14.6%, respectively (figure S3). This indicates the higher accuracy of our formula than previous empirical/semiempirical formulae. This is because previous empirical/semiempirical formulae are only accurate for a narrow range of crystals (e.g. diamondlike, zinc-blende, and chalcopyrite crystals), and their accuracies break down as extended to these 3765 crystals having diverse structures and compositions (figure S4). Since high-bulk-modulus crystals are just what we are looking for, the performance in the range of high bulk modulus is more important than for ranges of low

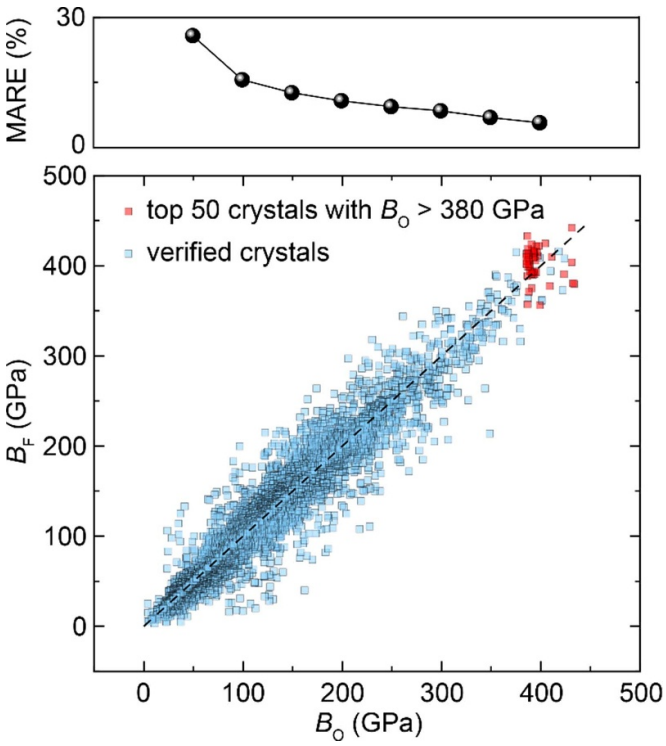


Figure 3. Predicted bulk moduli from our formula (B_O) for 3765 crystals in MP database that have a first principles calculated bulk modulus, and top 50 crystals with the highest predicted bulk moduli in MP and OQMD databases, as compared with first principles values (B_F). Each point of MARE is counted for crystals within the range of ±50 GPa.

or moderate bulk modulus. The MARE of Cohen's formula [9], Li's formula [11], our formula, and the machine learning model for all compounds having a first principles calculated bulk modulus greater than 350 GPa are 22.8%, 38.7%, 5.1%, and 10.1%, respectively. This indicates that our formula is more accurate than other approaches in the range of high bulk modulus. This is the key to directing the discovery of ultraincompressible crystals in the following investigation. As a comparison, both literature and our results show that machine learning models unusually underestimate the values in the range of high bulk modulus, and the machine learning-directed identified ultraincompressible crystals with a bulk modulus greater than 400 GPa are very rare [7, 13]. For example, there is only one crystal with a bulk modulus greater than 400 GPa (bulk modulus of 401 GPa for Re_2C) in the machine-learning search of 18 493 crystals [13]. This is because the past machine learning models have limitations, such as insufficient accuracy in the range of high bulk modulus due to sparse training data, and lack of critical physical descriptors [18]. These results indicate that our theoretical formula has clear advantages of high accuracy in the range of high bulk modulus, high interpretability, and great ease of use, as compared to machine learning models.

Finally, over one million crystals in MP and OQMD databases whose bulk moduli remain unknown were screened in the following procedures: firstly, the oxides, halides, and crystals that contain radioactive elements were filtered out (see supplementary data for details). Secondly, the crystals with a value of energy-above-hull >1.0 eV atom^{-1} or >20 atoms in the primitive cell were filtered out. Thus, the bulk moduli for the retained 461 285 crystals were predicted by equation (4). After a duplicate check, the top 50 crystals with a predicted bulk modulus greater than 380 GPa were further verified by first principles calculations (figure 3). Compared with first principles verifications, the MARE of predicted bulk moduli from our formula for these top 50 crystals is 4.5%, which further indicates the accuracy of our formula. Most importantly, 25 ultraincompressible crystals among them were identified to have a bulk modulus greater than 400 GPa that can rival diamond (table 1) [12]. It should be noted that the ideal values of bulk moduli for defect-free single crystals were explored in this work. For finite-size bulk counterparts with defects such as vacancies and grain boundaries, defects and size effects are expected to arise, which consequently would reduce the bulk moduli to be lower than those ideal values. These results indicate the power of our formula for high-throughput screening of ultraincompressible crystals.

Additionally, similar to the correlation between the bulk modulus and the volumetric cohesive energy, other important correlations can be derived. For instance, the sound speed of the crystal is

$$v \propto \sqrt{\frac{B}{\rho}} = \sqrt{\frac{kE_c}{m}}, \tag{6}$$

where $\rho = m/V$ is the mass density, and m is the atomic mass. Hence,

$$v^2 \propto k\frac{E_c}{m}. \tag{7}$$

Table 1. 25 crystals with a first principles calculated bulk modulus greater than 400 GPa. B_O, B_F, and B_L are the bulk moduli from our formula predictions, first principles calculations, and reported values in literature, respectively.

ID	Formula	k	B_O (GPa)	B_F (GPa)	B_L (GPa)
OQMD-14 925	C_3N_4	2.66	431	442	496 [33]
OQMD-21 515	ReC	2.98	387	436	446 [34]
MP-974 437	Re_2C	3.30	391	428	401 [13]
MP-867 141	$ReOs_3$	4.21	398	425	390 [13]
OQMD-1625 530	$ReIrOs_6$	4.26	395	422	394 [15]
OQMD-22 377	Re_2N	3.68	395	422	411 [35]
OQMD-22 307	ReN	3.55	387	419	453 [36]
OQMD-319 530	$IrOs_3$	4.31	391	419	394 [15]
OQMD-347 040	$ReOs_3$	4.21	392	417	395 [15]
OQMD-22 378	Re_3N	3.75	393	417	403 [35]
OQMD-1435 176	Re_3C	3.46	387	416	379 [15]
OQMD-301 434	$ReOs_3$	4.21	396	416	400 [15]
OQMD-1230 666	ReOs	4.13	397	415	388 [15]
OQMD-304 348	$IrOs_3$	4.31	386	414	394 [15]
OQMD-14 924	C_3N_4	2.66	411	411	451 [33]
OQMD-337 952	ReOs	4.13	385	410	388 [15]
OQMD-1625 528	$IrOs_6W$	4.23	388	410	385 [15]
OQMD-1625 537	$ReOs_6W$	4.18	393	409	388 [15]
MP-1102 681	CN_2	2.76	392	409	407 [37]
OQMD-327 494	ReOs	4.13	388	409	387 [15]
MP-1104 073	$C_{11}N_4$	2.31	431	405	460 [38]
OQMD-1625 541	$MoReOs_6$	4.21	386	405	383 [15]
MP-867 264	Re_3Os	4.04	386	402	371 [13]
OQMD-1625 866	Re_6IrOs	4.04	388	401	373 [15]
OQMD-301 470	Re_3Os	4.04	389	401	382 [15]

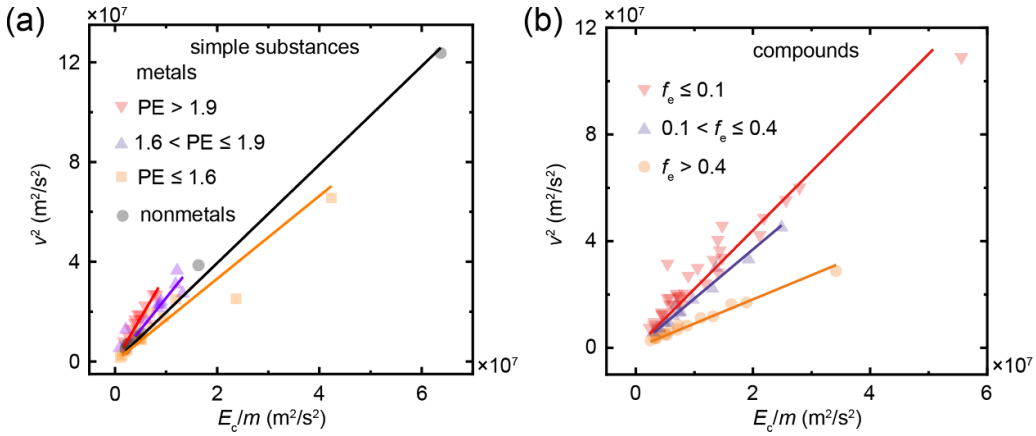


Figure 4. Correlation between the square of the sound speed (v^2) and massic cohesive energy (E_c/m) for (a) simple substances and (b) compounds.

Equation (7) indicates the correlation between the sound speed and the massic cohesive energy (E_c/m). To verify equation (7), the sound speed and the massic cohesive energy for simple substances and compounds are collected in figure 4. The high correlation coefficients ($\geqslant$0.97) support the strong correlation between the sound speed and the massic cohesive energy (equation (7)).

3. Conclusion

In conclusion, we derived a correlation between bulk modulus and volumetric cohesive energy from the interatomic energy landscape. The analyses of first principles calculated data support that this correlation is universally applicable to a wide range of crystals. Most interestingly, B is almost linear to ρ_e for crystals with a similar structure, since their ε_s fall within a narrow range. Finally, this universal correlation was used for high-throughput screening of ultraincompressible crystals. Directed by the prediction from our theoretical formula, 25 ultraincompressible crystals with a bulk modulus greater than 400 GPa were identified. This work demonstrates a universal correlation between important quantities of condensed matter, which not only provides a pathway to predict the bulk modulus from the volumetric cohesive energy, but also helps

to understand the microscopic origin of the ultrahigh bulk modulus.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

This work is supported by the National Natural Science Foundation of China (12172261). The numerical calculations in this work are done on the supercomputing system in the Supercomputing Center of Wuhan University.

ORCID iD

Enlai Gao https://orcid.org/0000-0003-1960-0260

References

[1] Kumar V, Shrivastava A K and Jha V 2010 Bulk modulus and microhardness of tetrahedral semiconductors *J. Phys. Chem. Solids* **71** 1513–20
[2] Mazhnik E and Oganov A R 2019 A model of hardness and fracture toughness of solids *J. Appl. Phys.* **126** 125109
[3] Gilman J J 2016 Chemical and physical "hardness" *Mater. Res. Innov.* **1** 71–76
[4] Anand Kumar P, Sanjay Kumar N R and Chandra Shekar N V 2022 Structure–compressibility correlation among MoB_X *J. Phys. Chem. Solids* **163** 110599
[5] Zhang Y, Sun H and Chen C 2004 Superhard cubic BC_2N compared to diamond *Phys. Rev. Lett.* **93** 195504
[6] Thornton A G and Wilks J 1978 Clean surface reactions between diamond and steel *Nature* **274** 792–3
[7] Mansouri Tehrani A, Oliynyk A O, Parry M, Rizvi Z, Couper S, Lin F, Miyagi L, Sparks T D and Brgoch J 2018 Machine learning directed search for ultraincompressible, superhard materials *J. Am. Chem. Soc.* **140** 9844–53
[8] Cohen M L 2000 The theory of real materials *Annu. Rev. Mater. Sci.* **30** 1–26
[9] Cohen M L 1993 Predicting useful materials *Science* **261** 307–8
[10] Cohen M L 1985 Calculation of bulk moduli of diamond and zinc-blende solids *Phys. Rev.* B **32** 7988–91
[11] Li K, Ding Z and Xue D 2011 Electronegativity-related bulk moduli of crystal materials *Phys. Status Solidi* b **248** 1227–36
[12] Kamran S, Chen K and Chen L 2008 Semiempirical formulae for elastic moduli and brittleness of diamondlike and zinc-blende covalent crystals *Phys. Rev.* B **77** 094109
[13] Zeng S, Li G, Zhao Y, Wang R and Ni J 2019 Machine learning-aided design of materials with target elastic properties *J. Phys. Chem.* C **123** 5042–7
[14] Ojih J, Al-Fahdi M, Rodriguez A D, Choudhary K and Hu M 2022 Efficiently searching extreme mechanical properties via boundless objective-free exploration and minimal first-principles calculations *npj Comput. Mater.* **8** 143
[15] Jin R, Yuan X and Gao E 2023 Atomic stiffness for bulk modulus prediction and high-throughput screening of ultraincompressible crystals *Nat. Commun.* **14** 4258
[16] Jain A *et al* 2013 Commentary: the materials project: a materials genome approach to accelerating materials innovation *APL Mater.* **1** 011002
[17] Saal J E, Kirklin S, Aykol M, Meredig B and Wolverton C 2013 Materials design and discovery with high-throughput density functional theory: the open quantum materials database (OQMD) *JOM* **65** 1501–9
[18] Ghiringhelli L M, Vybiral J, Levchenko S V, Draxl C and Scheffler M 2015 Big data of materials science: critical role of the descriptor *Phys. Rev. Lett.* **114** 105503
[19] Economou E N 2010 *The Physics of Solids* (Springer) (https://doi.org/10.1007/978-3-642-02069-8)
[20] Plendl J N, Mitra S S and Gielisse P J 1965 Compressibility, cohesive energy, and hardness of non-metallic solids *Phys. Status Solidi* b **12** 367–74
[21] Wacke S, Górecki T, Górecki C and Książek K 2011 Relations between the cohesive energy, atomic volume, bulk modulus and sound velocity in metals *J. Phys.: Conf. Ser.* **289** 012020
[22] Plendl J N and Gielisse P J M 1969 Compressibility and polymorphism of solids *Phys. Status Solidi* b **35** K151–K6
[23] Srivastava G P and Weaire D 1987 The theory of the cohesive energies of solids *Adv. Phys.* **36** 463–517
[24] Morse P M 1929 Diatomic molecules according to the wave mechanics. II. Vibrational levels *Phys. Rev.* **34** 57–64
[25] Costa Filho R N, Alencar G, Skagerstam B-S and Andrade J S 2013 Morse potential derived from first principles *Europhys. Lett.* **101** 10009
[26] Bétermin L 2019 Minimizing lattice structures for Morse potential energy in two and three dimensions *J. Math. Phys.* **60** 102901
[27] Charan H, Hansen A, Hentschel H G E and Procaccia I 2021 Aging and failure of a polymer chain under tension *Phys. Rev. Lett.* **126** 085501
[28] Lv B, Chen C, Zhang F, Poletaev G M and Rakitin R Y 2022 Potentials for describing interatomic interactions in γFe-Mn-C-N system *Metals* **12** 982
[29] Nickabadi S, Ansari R, Rouhi S and Aghdasi P 2021 On the derivation of coefficient of Morse potential function for the silicene: a DFT investigation *J. Mol. Model.* **27** 190
[30] Benassi E 2018 The zero point position in Morse's potential and accurate prediction of thermal expansion in metals *Chem. Phys.* **515** 323–35
[31] Pauling L 2002 The nature of the chemical bond. IV. The energy of single bonds and the relative electronegativity of atoms *J. Am. Chem. Soc.* **54** 3570–82
[32] Fan Z, Tsakiropoulos P and Miodownik A P 1994 A generalized law of mixtures *J. Mater. Sci.* **29** 141–50
[33] Teter D M and Hemley R J 1996 Low-compressibility carbon nitrides *Science* **271** 53–55
[34] Yang J and Gao F 2010 Hardness calculations of 5d transition metal monocarbides with tungsten carbide structure *Phys. Status Solidi* b **247** 2161–7
[35] Bannikov V V, Shein I R and Ivanovskii A L 2011 Elastic and electronic properties of hexagonal rhenium sub-nitrides Re_3N and Re_2N in comparison with *hcp*-Re and wurtzite-like rhenium mononitride ReN *Phys. Status Solidi* b **248** 1369–74
[36] Asvini Meenaatci A T, Rajeswarapalanichamy R and Iyakutti K 2011 First-principles study of electronic structure of transition metal nitride: ReN under normal and high pressure *Physica* B **406** 3303–7
[37] Li Q, Liu H, Zhou D, Zheng W, Wu Z and Ma Y 2012 A novel low compressible and superhard carbon nitride: body-centered tetragonal CN_2 *Phys. Chem. Chem. Phys.* **14** 13081–7
[38] Mattesini M and Matar S F 2002 Density-functional theory investigation of hardness, stability, and electron-energy-loss spectra of carbon nitrides with $C_{11}N_4$ stoichiometry *Phys. Rev.* B **65** 075110

Supplementary data for

Data-driven discovery of ultraincompressible crystals from a universal correlation between bulk modulus and volumetric cohesive energy

Xiaoang Yuan and Enlai Gao*

Department of Engineering Mechanics, School of Civil Engineering, Wuhan University, Wuhan, Hubei 430072, China.

*Corresponding author. Email: enlaigao@whu.edu.cn

Note 1: Derivation of the expression of bulk modulus

For a crystal, the atomic cohesive energy function (E) and atomic volume (V) can be expressed as $E = (N_c/2)E_b$, and $V = md_0^3$, respectively, where N_c is the coordination number, and m is the ratio of the atomic volume to the cube of equilibrium bond length. The first derivative of E with respect to V can be written as

$$\frac{\mathrm{d}E}{\mathrm{d}V} = \frac{\mathrm{d}E}{\mathrm{d}d_0}\frac{\mathrm{d}d_0}{\mathrm{d}V}, \tag{S1}$$

and the second derivative of E with respect to V can be written as

$$\frac{\mathrm{d}^2E}{\mathrm{d}V^2} = \frac{\mathrm{d}^2E}{\mathrm{d}d_0^{\,2}}\left(\frac{\mathrm{d}d_0}{\mathrm{d}V}\right)^2 + \frac{\mathrm{d}E}{\mathrm{d}d_0}\frac{\mathrm{d}^2d_0}{\mathrm{d}V^2}. \tag{S2}$$

From **Eq. 1**, the first derivative of E with respect to d_0 can be derived as

$$E'(d) = N_c\alpha De^{-\alpha(d-d_0)}[1-e^{-\alpha(d-d_0)}], \tag{S3}$$

By substituting d_0 into d,

$$\frac{\mathrm{d}E}{\mathrm{d}d_0} = E'(d_0) = 0. \tag{S4}$$

Meanwhile, the second derivative of E with respect to d_0 can be expressed as

$$E''(d) = -N_c\alpha^2De^{-\alpha(d-d_0)}[1-e^{-\alpha(d-d_0)}] + N_c\alpha^2D(e^{-\alpha(d-d_0)})^2, \tag{S5}$$

By substituting d_0 into d,

$$\frac{\mathrm{d}^2E}{\mathrm{d}d_0^{\,2}} = E''(d_0) = N_c\alpha D. \tag{S6}$$

Combining **Eqs. S2, S4, S6**, and $V = md_0^3$, the bulk modulus can be derived as

$$B = V\frac{\mathrm{d}^2E}{\mathrm{d}V^2} = \frac{N_c\alpha^2D}{9md_0}. \tag{S7}$$

Note 2: Parameterization of the Morse potential

To parametrize the Morse potential from first principles calculations, the following procedure was adopted. First, D was parameterized from first principles calculated ρ_e for each crystal. Second, α was parameterized by fitting the first principles energy-strain curves of each crystal to **Eq. 1.** These curves were calculated by applying equiaxial strain (volumetric strain). Substituting α and the known equilibrium bond length into **Eq. 5**, the values of ε_s can be calculated accordingly. **Table S1** summarizes relevant parameters.

Note 3: Prediction from empirical/semiempirical formulae and machine learning model

We extended Cohen's formula [1] and Li's formula [2] for predicting 3765 compounds in MP database. For the formula of Cohen [1], the bulk modulus for a crystal is calculated by

$$B = (N_c/4)(1972\text{-}220I)d^{-3.5}, \quad \text{(S8)}$$

where N_c, I, and d are the average coordination number, empirical ionicity parameter ($I = 2.9+0.6\ln(f_e)$, where f_e is the ionicity of the crystal), and the average bond length of the crystal, respectively. For the formula of Li [2], the bulk modulus for a crystal is calculated by

$$B = \frac{(1944.8\sum w_{ab}x_{ab} / v + 13.5)}{\exp[(1/\sum w_{ab} / f_{e(ab)})^2]}, \quad \text{(S9)}$$

where w_{ab}, x_{ab}, v, and $f_{e(ab)}$ are the weighting factor of *a*-*b* bond, bond electronegativity, average bond volume, and effective ionicity, respectively.

The machine learning model is developed as follows: Firstly, referring to the descriptors adopted in the previous machine learning models [3, 4], we imported 140 descriptors for characterizing the structural and compositional features of each crystal from Matminer, i.e., an open source toolkit for materials data mining [5]. Afterward, the Random forest [6] is used to describe the relationship between these descriptors and the bulk moduli of crystals. By adopting a 2-fold cross-validation scheme, we took one-half of these 3765 crystals as the training set, and then predicted the bulk moduli for the other half of these 3765 crystals from the trained model. Finally, the predicted bulk moduli for all 3765 crystals were obtained by this procedure.

Note 4: Prediction of ultraincompressible crystals from our formula

There are some notes on the prediction of ultraincompressible crystals from our formula. Firstly, since the k of a crystal is calculated based on the k of the simple substance of constitutive elements, the simple substances with ultrahigh bulk moduli (i.e., C, Os, and Re) are not taken into consideration in the prediction of ultraincompressible crystals. Therefore, the identified 25 ultraincompressible crystals with a bulk modulus greater than 400 GPa are all compounds. Secondly, since some elements are radioactive, the compounds that contain these elements are omitted in the prediction of ultraincompressible crystals. Thirdly, the oxides and halides were reported to have a low bulk modulus [4], and the top 50 ultraincompressible crystals predicted by the previous machine learning model [4] didn't contain oxides and halides, indicating that the low possibility of these crystals to be ultraincompressible crystals. Therefore, the oxides and halides are also omitted in this prediction.

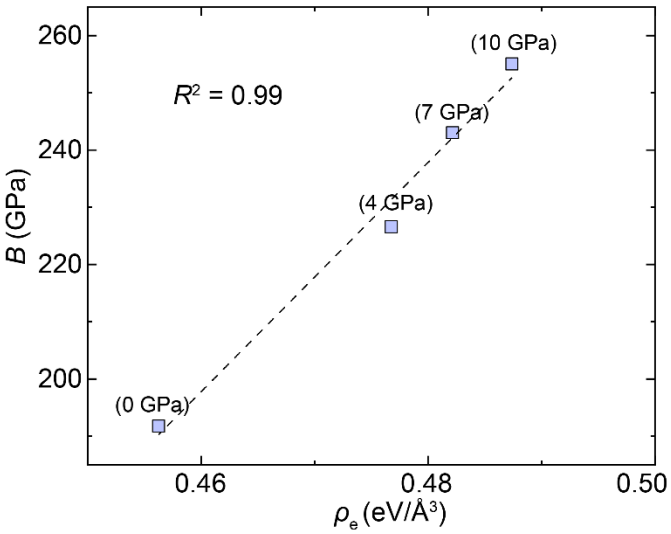


Figure S1. Correlation between first principles calculated bulk modulus (B) and volumetric cohesive energy (ρ_e) of α-iron under pressures of 0 GPa, 4 GPa, 7 GPa, and 10 GPa.

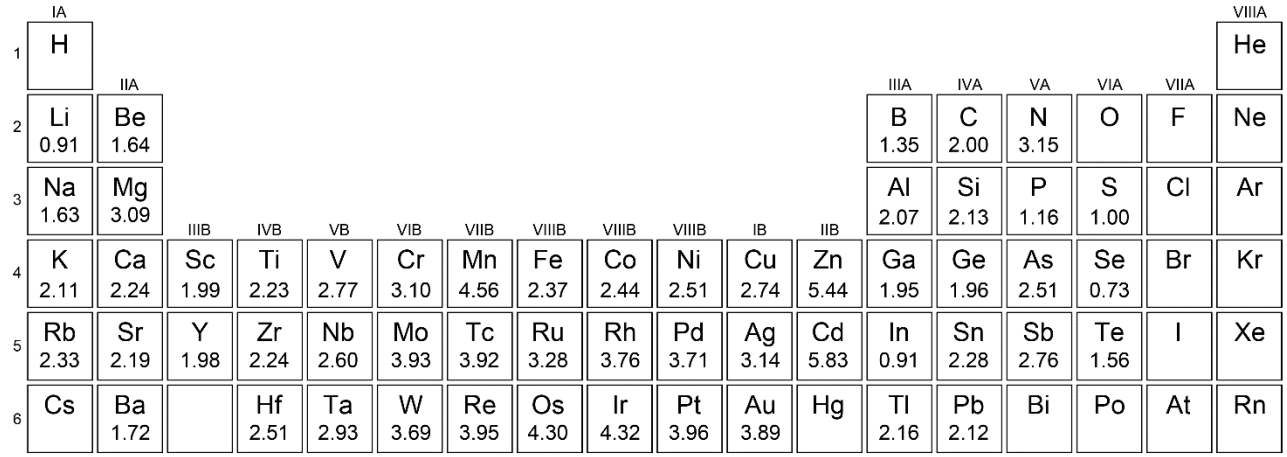


Figure S2. Factors of proportionality (k) for elements within the periodic table. The value of k for each element is determined in the following procedures: Firstly, α and d_0 of the simple substance of each element are parametrized. Afterward, the value of k is obtained by substituting the corresponding α and d_0 into **Eq. 5**. It should be noted that the value of k for nitrogen is parametrized from its compounds (C_3N_4, BC_2N, BN, CN_2, and VN), since nitrogen is gas at standard temperature and pressure.

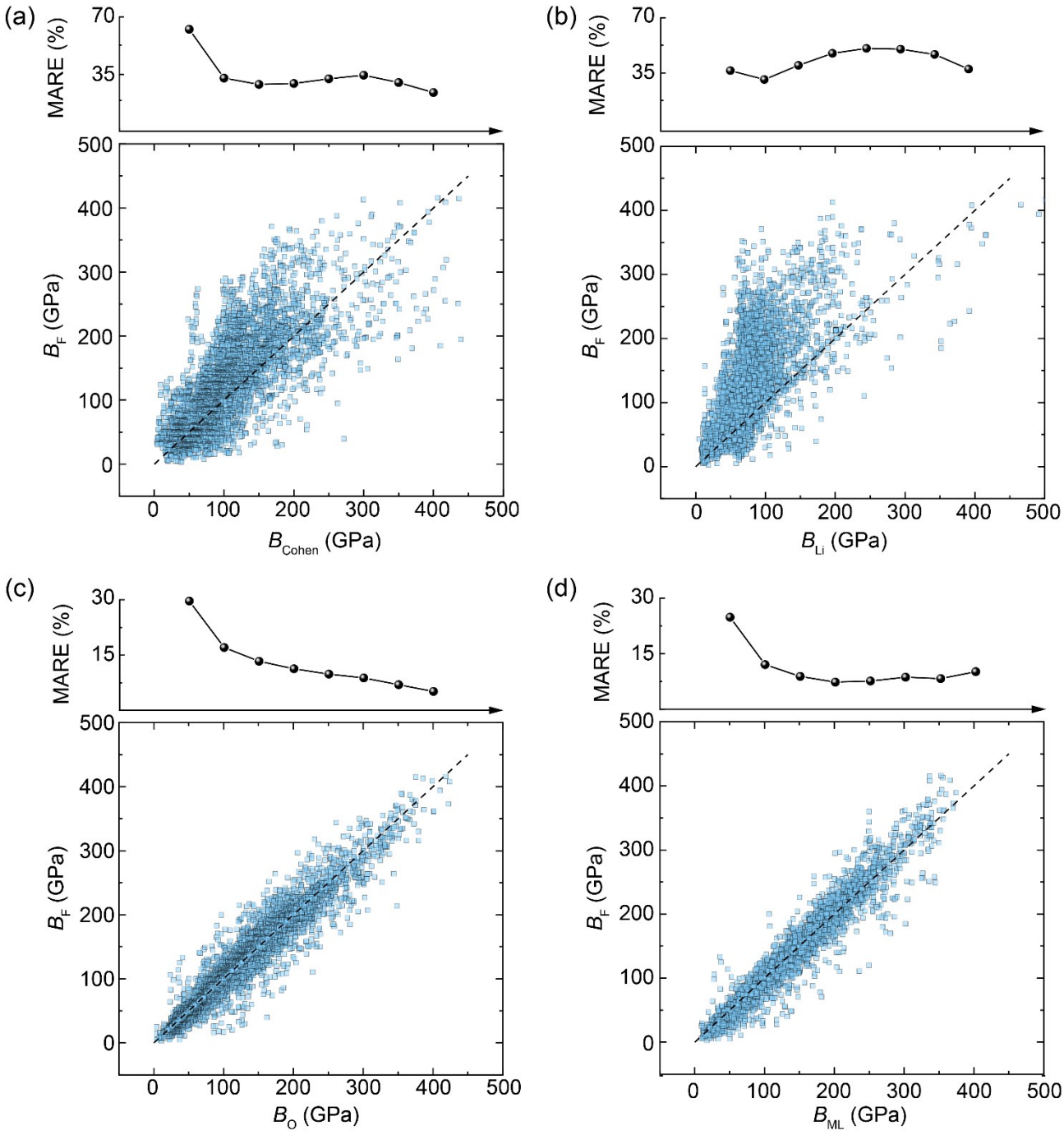


Figure S3. Predicted bulk moduli for 3765 crystals in MP database by (a) Cohen's formula (B_{Cohen}), (b) Li's formula (B_{Li}), (c) our formula (B_O), and (d) machine learning model (B_{ML}), as compared with first principles values (B_F). Each point of MARE is counted for crystals within the range of ±50 GPa.

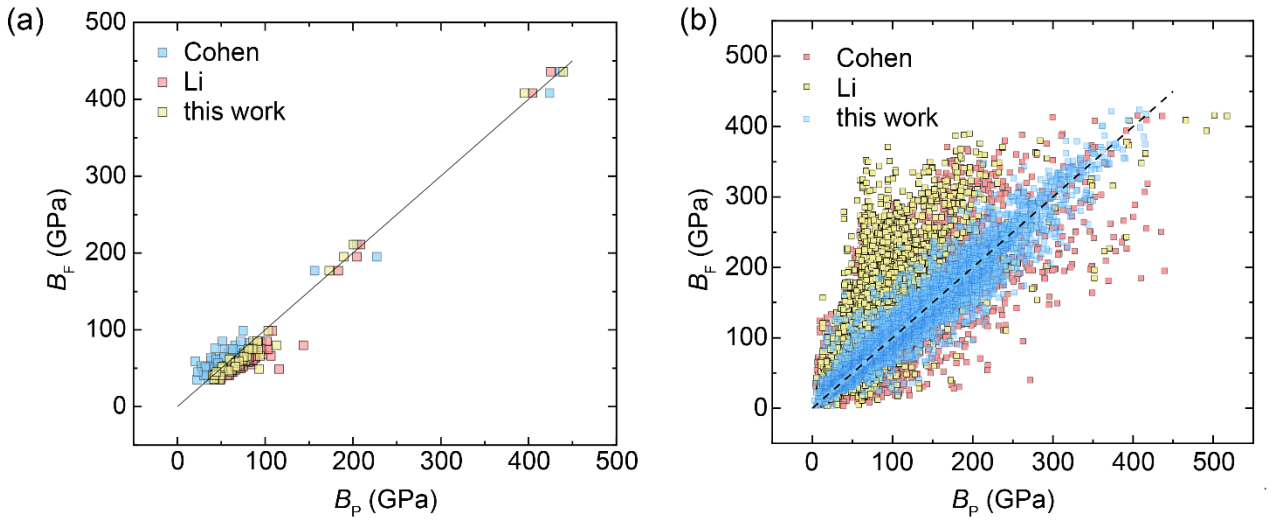


Figure S4. Predicted bulk moduli (B_P) for (a) typical zinc-blende and chalcopyrite crystals, and (b) 3765 crystals in MP database that have a first principles calculated bulk modulus by Cohen's formula, Li's formula, and our formula, as compared with first principles values (B_F). It is seen that Cohen's formula and Li's formula are only accurate for a narrow range of crystals.

Table S1. Parameters of Morse potential (D, α), the equilibrium bond length (d_0), the strain-to-failure of interatomic bonds (ε_s), and the predicted bulk modulus by our formula (B_O) and first principles calculated bulk modulus (B_F) for simple substances and compounds (rocksalt compounds and zincblende-structured compounds).

Formula	D (eV)	α (Å^{-1})	d_0 (Å)	ε_s	B_O (GPa)	B_F (GPa)
Diamond	3.97	1.94	1.54	0.23	449.27	435.67
Si	2.38	1.31	2.36	0.22	80.01	88.83
Ge	1.96	1.19	2.48	0.23	52.09	58.13
α-Sn	1.68	1.12	2.86	0.22	33.91	38.92
As	4.75	1.07	1.19	0.21	59.08	57.96
Sb	3.87	0.99	1.12	0.20	42.55	45.08
Os	1.48	1.63	2.70	0.16	438.30	427.59
Re	1.59	1.54	2.75	0.16	413.53	396.11
Ru	1.42	1.44	2.67	0.18	333.99	334.36
Ir	1.33	1.62	2.71	0.16	390.26	380.04
W	2.36	1.49	2.74	0.17	354.24	329.26
Mo	1.88	1.55	2.71	0.16	306.42	279.63
Pd	0.72	1.49	2.75	0.17	174.95	187.65
Ag	0.50	1.30	2.88	0.18	89.42	106.75
Ni	0.86	1.37	2.45	0.21	198.61	214.64
Rh	1.06	1.53	2.68	0.17	279.49	281.18
Fe	1.29	1.34	2.43	0.21	177.85	191.74
Nb	2.04	1.21	2.83	0.20	194.97	189.66
Co	0.92	1.35	2.46	0.21	205.32	224.42
V	1.57	1.38	2.56	0.20	214.62	202.40
Cr	1.64	1.53	2.44	0.19	291.37	281.10
Cu	0.67	1.39	2.52	0.20	154.46	155.37
Ti	0.98	1.11	2.86	0.22	127.06	119.55
Y	0.74	0.84	3.55	0.23	44.43	42.64
Sc	0.74	0.93	3.22	0.23	59.50	55.89
Zr	1.11	1.00	3.17	0.22	105.98	99.06
Ta	2.18	1.28	2.84	0.19	230.71	208.36
Hf	1.14	1.07	3.13	0.21	126.23	110.93
Pt	1.04	1.52	2.77	0.16	263.50	283.00
Au	0.62	1.44	2.90	0.17	133.61	152.41
Ga	0.83	1.09	2.71	0.23	45.49	51.41
Al	0.62	1.08	2.84	0.23	76.13	75.84
Mg	0.29	1.19	3.15	0.19	38.48	35.75
Ca	0.34	0.82	3.88	0.22	17.78	15.58
Sr	0.29	0.74	4.22	0.22	11.56	11.35

Ba	0.50	0.65	4.29	0.25	9.03	9.32
Be	0.66	1.23	2.20	0.26	137.10	129.60
Tl	0.57	0.92	3.38	0.22	26.62	27.00
Pb	0.53	0.88	3.52	0.22	34.97	46.96
LiF	1.53	0.86	2.00	0.40	60.19	77.67
LiCl	1.20	0.76	2.54	0.36	29.06	35.09
LiBr	1.09	0.75	2.72	0.34	24.18	30.41
LiI	0.96	0.74	2.97	0.31	19.10	23.45
NaF	1.37	0.74	2.32	0.40	34.78	47.47
NaCl	1.12	0.70	2.83	0.35	20.61	24.00
NaBr	1.02	0.68	2.98	0.34	17.05	20.31
NaI	0.91	0.67	3.21	0.32	13.37	16.86
KF	1.33	0.64	2.69	0.40	21.41	29.14
KCl	1.13	0.61	3.19	0.36	13.88	15.84
KBr	1.04	0.61	3.35	0.34	12.26	13.45
KI	0.94	0.57	3.55	0.34	9.13	11.06
RbF	1.30	0.62	2.85	0.39	18.65	25.63
RbCl	1.11	0.58	3.36	0.35	11.93	12.83
RbBr	1.03	0.59	3.52	0.33	10.89	11.53
RbI	0.93	0.56	3.73	0.33	8.32	9.57
MgO	1.74	1.31	2.10	0.25	151.34	162.27
MgS	1.34	1.22	2.59	0.22	82.22	80.31
MgSe	1.21	1.21	2.73	0.21	69.33	66.93
CaO	1.87	1.10	2.40	0.26	101.43	111.72
CaS	1.63	0.96	2.84	0.26	56.18	58.25
CaSe	1.52	0.95	2.96	0.25	49.30	50.50
CaTe	1.37	0.94	3.17	0.23	40.95	41.66
SrO	1.74	1.06	2.58	0.25	80.68	90.43
SrS	1.58	0.91	3.01	0.25	46.66	49.23
SrSe	1.48	0.90	3.13	0.25	40.75	42.24
SrTe	1.34	0.89	3.33	0.23	34.31	34.87
BaO	1.69	1.00	2.79	0.25	64.60	70.57
BaS	1.59	0.86	3.20	0.25	39.17	43.14
BaSe	1.51	0.84	3.31	0.25	33.88	36.75
BaTe	1.39	0.81	3.51	0.24	27.67	29.38
BN	3.58	1.92	1.57	0.23	391.56	380.20
BP	2.78	1.61	1.96	0.22	170.45	166.12
BAs	2.45	1.57	2.07	0.21	134.72	134.71
BSb	2.19	1.49	2.27	0.20	98.86	98.14
AlN	2.97	1.71	1.89	0.21	210.61	195.00

AlP	2.16	1.45	2.37	0.20	88.13	85.57
AlAs	1.97	1.40	2.47	0.20	71.89	70.47
AlSb	1.74	1.33	2.68	0.19	53.13	52.79
GaN	2.29	1.81	1.96	0.20	178.13	179.84
GaP	1.86	1.48	2.37	0.20	79.47	79.37
GaAs	1.70	1.42	2.47	0.20	64.06	63.81
GaSb	1.55	1.34	2.67	0.19	47.95	46.67
InN	1.93	1.76	2.18	0.18	127.32	126.31
InP	1.68	1.43	2.56	0.19	61.71	62.00
InAs	1.56	1.37	2.66	0.19	51.27	52.48
InSb	1.45	1.30	2.85	0.19	39.57	41.41
IrC	3.70	1.80	2.00	0.19	276.51	279.59
PtC	3.11	1.73	2.04	0.20	211.22	215.48
RuC	3.89	1.59	1.96	0.22	231.24	265.40
NbC	4.01	1.55	2.08	0.21	214.51	210.93
CoN	2.65	1.93	1.83	0.20	250.74	260.03
IrN	2.89	2.00	2.00	0.17	266.38	257.24
FeN	2.68	1.93	1.82	0.20	251.52	283.86
CuN	1.80	1.98	1.91	0.18	171.10	161.90
TaC	4.19	1.61	2.08	0.21	241.30	234.07
ZrC	3.79	1.38	2.19	0.23	152.54	148.95
HfC	3.86	1.44	2.17	0.22	169.77	163.54
HfN	3.77	1.59	2.11	0.21	208.65	203.63
TiN	3.51	1.59	1.98	0.22	208.12	210.77
ScN	3.24	1.42	2.11	0.23	143.77	143.79
YN	3.18	1.34	2.27	0.23	116.04	114.34

Structural information of 25 crystals with a bulk modulus greater than 400 GPa.

1. C_3N_4

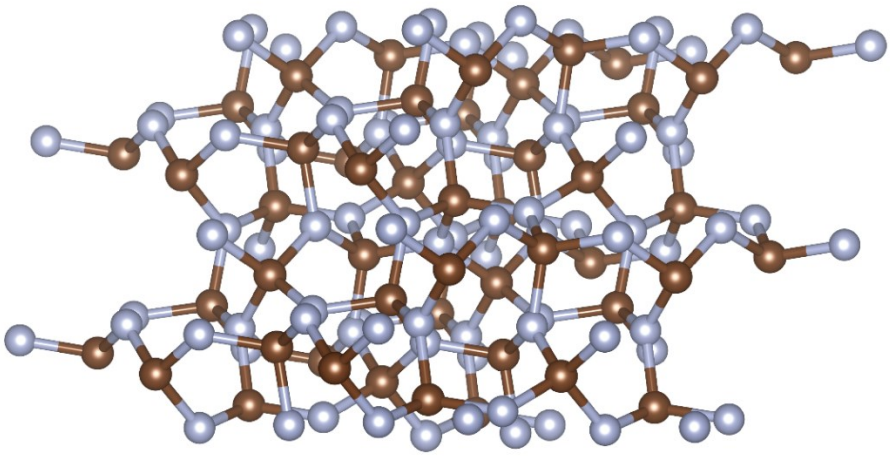
Formula	C_3N_4	**ID**	OQMD-14925
Density (g/cm^3)	3.859	**Symbol**	I-43d
Formation Energy / Atom (eV)	0.589	**Band Gap (eV)**	3.0

Crystal structure

Structural parameters: unit cell and atomic positions of C_3N_4 in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	2.70600000	2.70600000	-2.70600000
a_2	2.70600000	-2.70600000	2.70600000
a_3	-2.70600000	-2.70600000	-2.70600000

	x (Å)	*y* (Å)	*z* (Å)
C	0.25000000	0.37500000	0.12500000
C	0.87500000	0.62500000	0.25000000
C	0.75000000	0.12500000	0.37500000
C	0.37500000	0.25000000	0.62500000
C	0.62500000	0.87500000	0.75000000
C	0.12500000	0.75000000	0.87500000
N	0.50000000	0.43300000	0.00000000
N	0.50000000	0.93300000	0.00000000
N	0.06720000	0.06720000	0.06720000
N	0.00000000	0.00000000	0.43300000
N	0.93300000	0.50000000	0.50000000
N	0.43300000	0.50000000	0.50000000
N	0.56700000	0.06720000	0.56700000
N	0.00000000	0.00000000	0.93300000



2. ReC

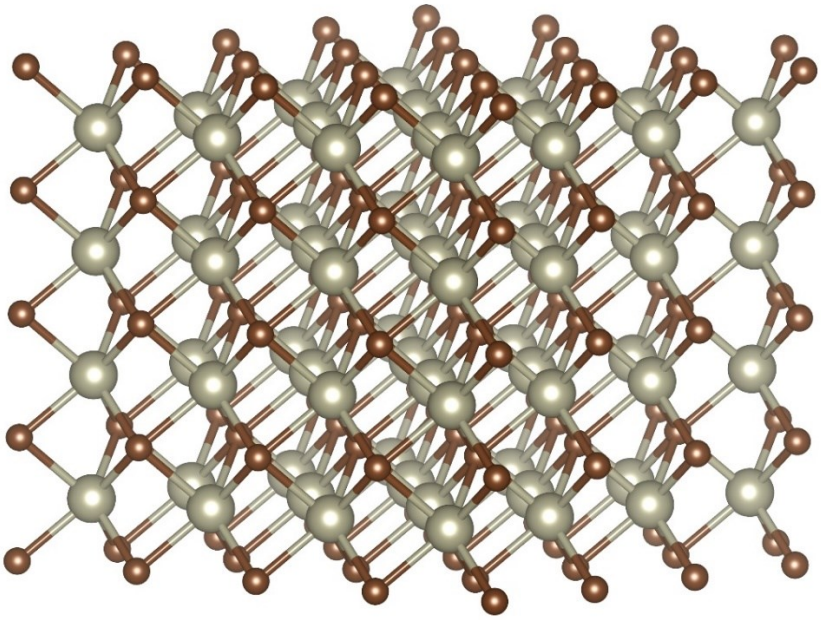
Formula	ReC	**ID**	OQMD-21515
Density (g/cm^3)	16.241	**Symbol**	P-6m2
Formation Energy / Atom (eV)	0.246	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of ReC in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	2.88100000	0.00000000	0.00000000
a_2	-1.44100000	2.49500000	0.00000000
a_3	0.00000000	0.00000000	2.82000000

	x (Å)	*y* (Å)	*z* (Å)
C	0.33333333	0.66666667	0.50000000
Re	0.00000000	0.00000000	0.00000000



3. Re_2C

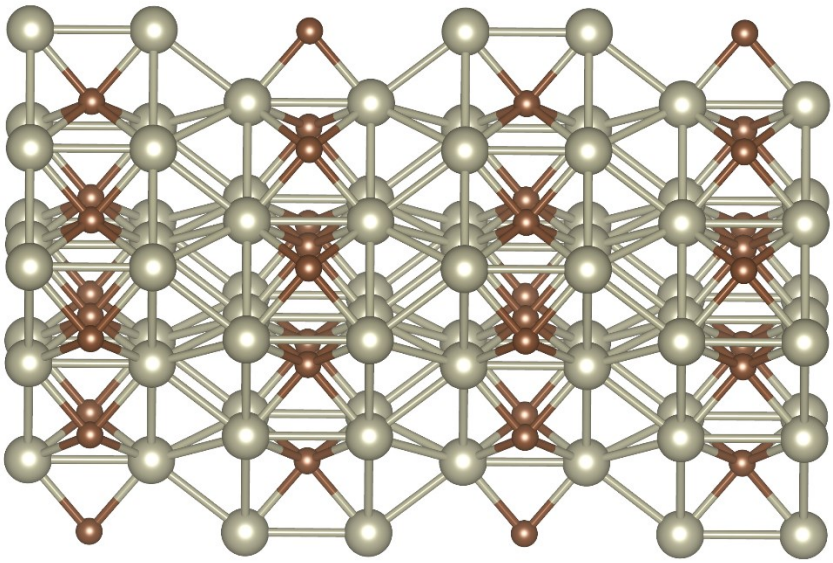
Formula	Re_2C	**ID**	mp-974437
Density (g/cm^3)	18.101	**Symbol**	P63/mmc
Formation Energy / Atom (eV)	-0.030	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of Re_2C in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	2.86600000	0.00000000	0.00000000
a_2	-1.43300000	2.48200000	0.00000000
a_3	0.00000000	0.00000000	9.91500000

	x (Å)	*y* (Å)	*z* (Å)
Re	0.33331876	0.66667211	0.60838437
Re	0.66667211	0.33331876	0.10839000
Re	0.66667211	0.33331876	0.39161428
Re	0.33331876	0.66667211	0.89161866
C	0.33331876	0.66667211	0.25000220
C	0.66667211	0.33331876	0.74999648



4. $ReOs_3$

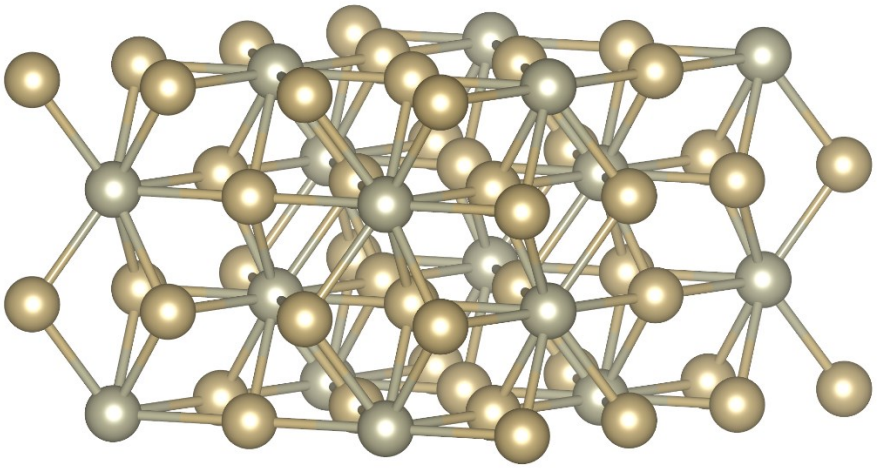
Formula	$ReOs_3$	**ID**	mp-867141
Density (g/cm^3)	21.667	**Symbol**	P63/mmc
Formation Energy / Atom (eV)	-0.085	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of $ReOs_3$ in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	5.53710794	0.00000000	0.00000000
a_2	-2.76855397	4.79527614	0.00000000
a_3	0.00000000	0.00000000	4.36940717

	x (Å)	*y* (Å)	*z* (Å)
Re	0.33334428	0.66667234	0.75001025
Re	0.66667234	0.33334428	0.25001102
Os	0.16728803	0.33459851	0.25001102
Os	0.66540008	0.83271056	0.25001102
Os	0.16730587	0.83271074	0.25001102
Os	0.83271056	0.66540008	0.75001025
Os	0.33459851	0.16728803	0.75001025
Os	0.83271074	0.16730587	0.75001025



5. $ReIrOs_6$

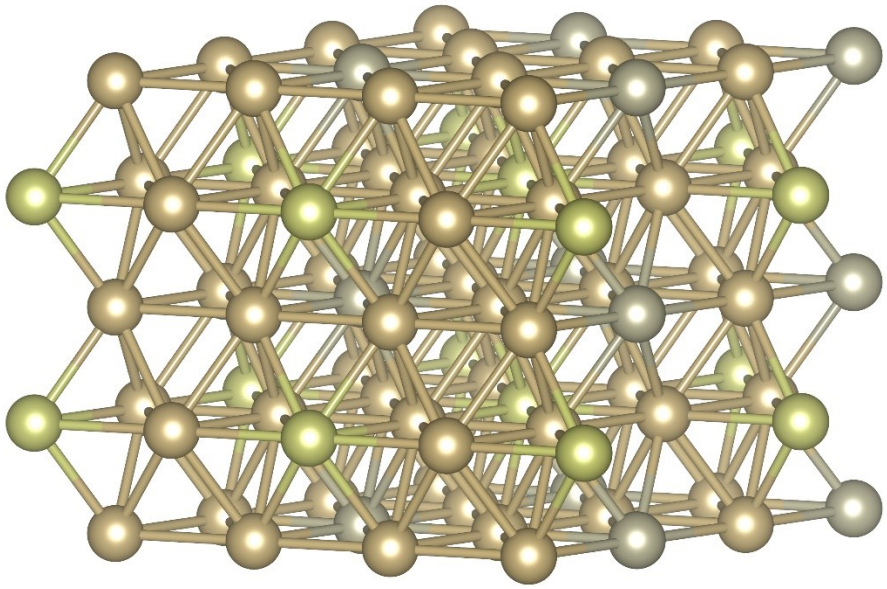
Formula	$ReIrOs_6$	**ID**	OQMD-1625530
Density (g/cm^3)	22.050	**Symbol**	P-6m2
Formation Energy / Atom (eV)	-0.046	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of $ReIrOs_6$ in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	5.51200000	0.00000000	0.00000000
a_2	-2.75600000	4.77300000	0.00000000
a_3	0.00000000	0.00000000	4.35200000

	x (Å)	*y* (Å)	*z* (Å)
Ir	0.00000000	0.00000000	0.50000000
Os	0.33100000	0.16500000	0.00000000
Os	0.83500000	0.16500000	0.00000000
Os	0.83500000	0.66900000	0.00000000
Os	0.50200000	0.00342000	0.50000000
Os	0.99700000	0.49800000	0.50000000
Os	0.50200000	0.49800000	0.50000000
Re	0.33300000	0.66700000	0.00000000



6. ReN

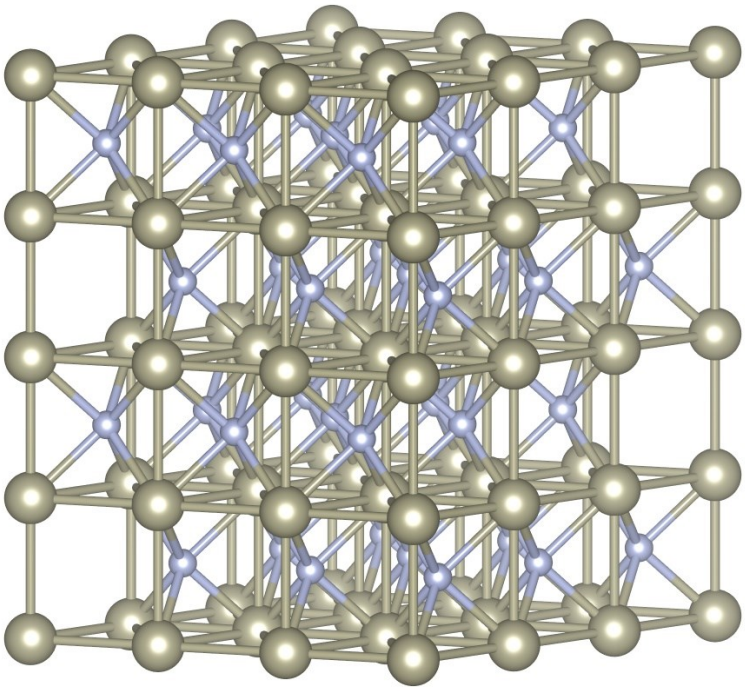
Formula	ReN	**ID**	OQMD-22307
Density (g/cm^3)	16.827	**Symbol**	P63/mmc
Formation Energy / Atom (eV)	0.317	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of ReN in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	2.78100000	0.00000000	0.00000000
a_2	-1.39100000	2.40900000	0.00000000
a_3	0.00000000	0.00000000	5.90100000

	x (Å)	*y* (Å)	*z* (Å)
N	0.33333333	0.66666667	0.25000000
N	0.66666667	0.33333333	0.75000000
Re	0.00000000	0.00000000	0.00000000
Re	0.00000000	0.00000000	0.50000000



7. $IrOs_3$

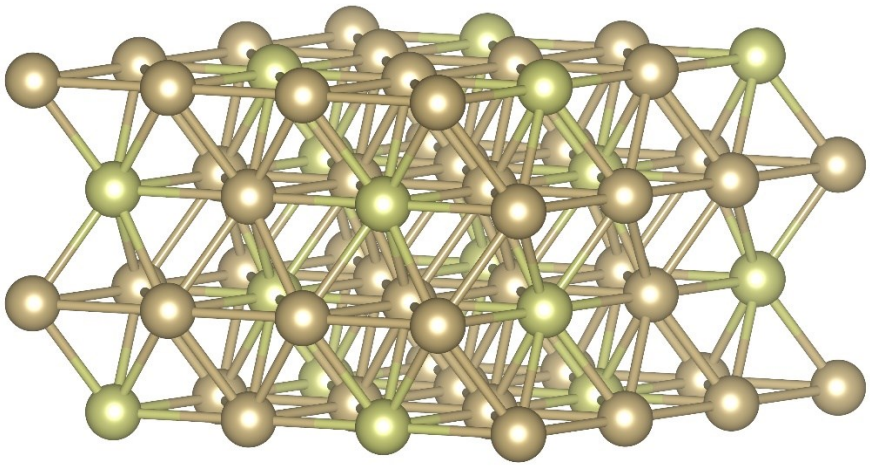
Formula	Os_3Ir	**ID**	OQMD-319530
Density (g/cm^3)	22.344	**Symbol**	P63/mmc
Formation Energy / Atom (eV)	-0.002	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of Os_3Ir in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	5.48900000	0.00000000	0.00000000
a_2	-2.74400000	4.75300000	0.00000000
a_3	0.00000000	0.00000000	4.34800000

	x (Å)	*y* (Å)	*z* (Å)
Ir	0.33300000	0.66700000	0.75000000
Ir	0.66700000	0.33300000	0.25000000
Os	0.16600000	0.33300000	0.25000000
Os	0.66700000	0.83400000	0.25000000
Os	0.16600000	0.83400000	0.25000000
Os	0.83400000	0.66700000	0.75000000
Os	0.33300000	0.16600000	0.75000000
Os	0.83400000	0.16600000	0.75000000



8. Re_2N

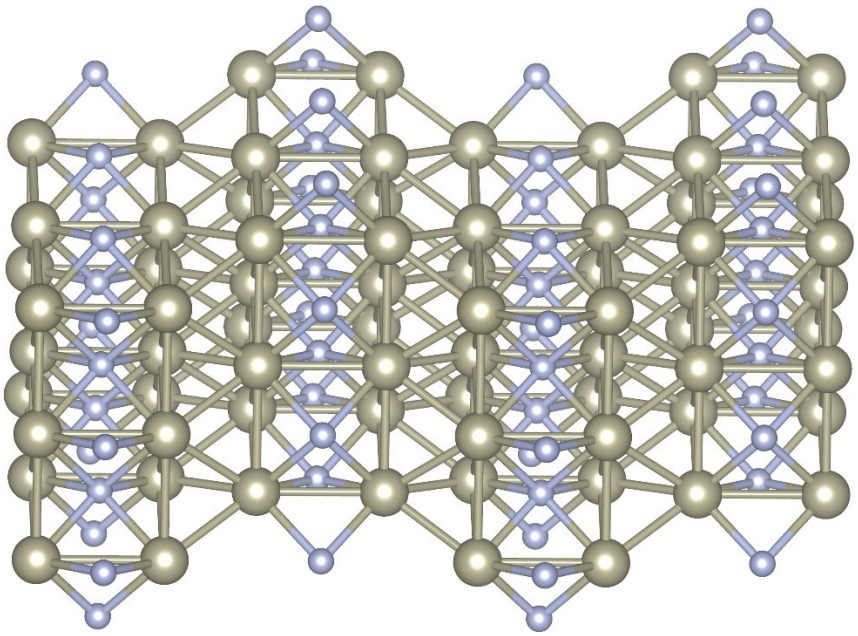
Formula	Re_2N	**ID**	OQMD-22377
Density (g/cm^3)	18.535	**Symbol**	P63/mmc
Formation Energy / Atom (eV)	-0.036	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of Re_2N in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	2.84800000	0.00000000	0.00000000
a_2	-1.42400000	2.46700000	0.00000000
a_3	0.00000000	0.00000000	9.85800000

	x (Å)	*y* (Å)	*z* (Å)
N	0.66700000	0.33300000	0.25000000
N	0.66700000	0.33300000	0.75000000
Re	0.33300000	0.66700000	0.10500000
Re	0.33300000	0.66700000	0.39500000
Re	0.66700000	0.33300000	0.60500000
Re	0.66700000	0.33300000	0.89500000



9. $ReOs_3$

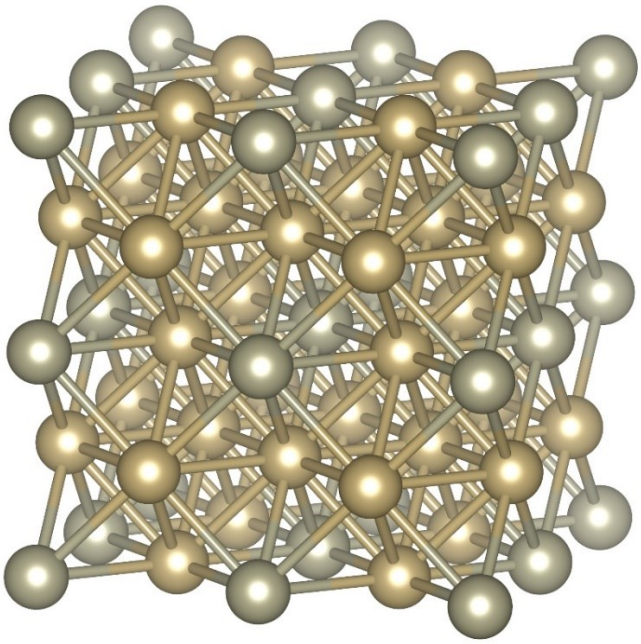
Formula	$ReOs_3$	**ID**	OQMD-347040
Density (g/cm^3)	21.788	**Symbol**	Pm-3m
Formation Energy / Atom (eV)	0.072	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of $ReOs_3$ in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	3.86400000	0.00000000	0.00000000
a_2	0.00000000	3.86400000	0.00000000
a_3	0.00000000	0.00000000	3.86400000

	x (Å)	*y* (Å)	*z* (Å)
Os	0.50000000	0.50000000	0.00000000
Os	0.50000000	0.00000000	0.50000000
Os	0.00000000	0.50000000	0.50000000
Re	0.00000000	0.00000000	0.00000000



10. $ReOs_3$

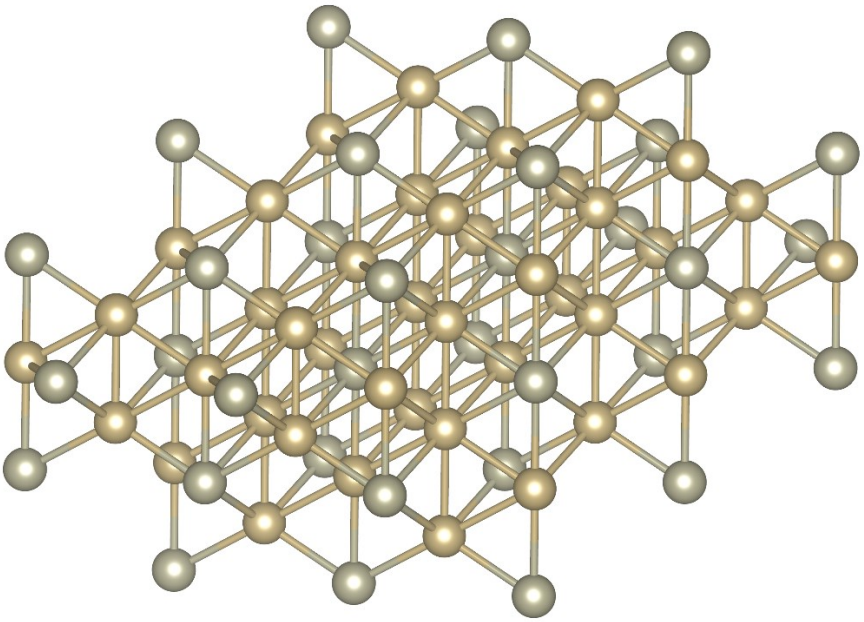
Formula	$ReOs_3$	**ID**	OQMD-301434
Density (g/cm^3)	22.029	**Symbol**	I4/mmm
Formation Energy / Atom (eV)	0.085	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of $ReOs_3$ in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	-1.91900000	1.91900000	3.87500000
a_2	1.91900000	-1.91900000	3.87500000
a_3	1.91900000	1.91900000	-3.87500000

	x (Å)	*y* (Å)	*z* (Å)
Os	0.75000000	0.25000000	0.50000000
Os	0.25000000	0.75000000	0.50000000
Os	0.50000000	0.50000000	0.00000000
Re	0.00000000	0.00000000	0.00000000



11. Re_3N

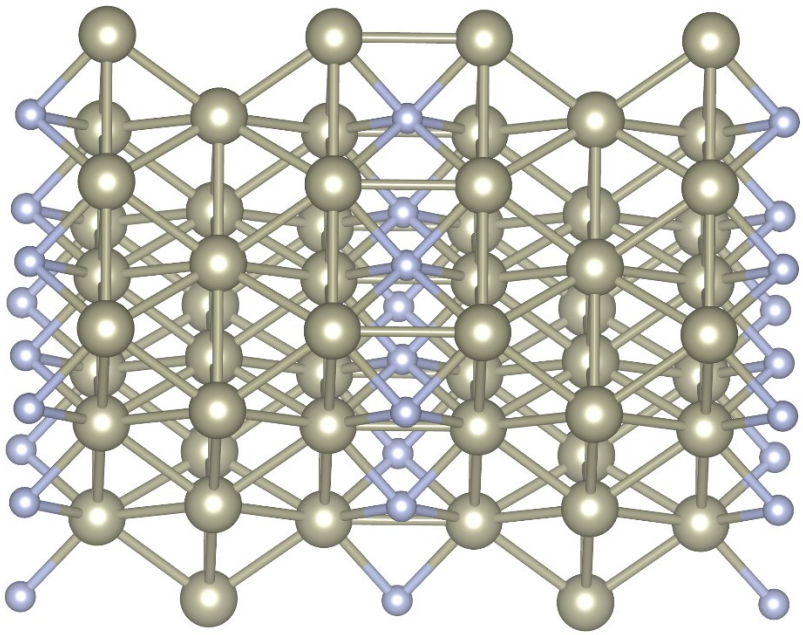
Formula	Re_3N	**ID**	OQMD-22378
Density (g/cm^3)	19.228	**Symbol**	P-6m2
Formation Energy / Atom (eV)	-0.087	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of Re_3N in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	2.82500000	0.00000000	0.00000000
a_2	-1.4130000	2.44700000	0.00000000
a_3	0.00000000	0.00000000	7.15800000

	x (Å)	*y* (Å)	*z* (Å)
N	0.66700000	0.33300000	0.00000000
Re	0.33300000	0.66700000	0.19900000
Re	0.66700000	0.33300000	0.50000000
Re	0.33300000	0.66700000	0.80100000



12. Re_3C

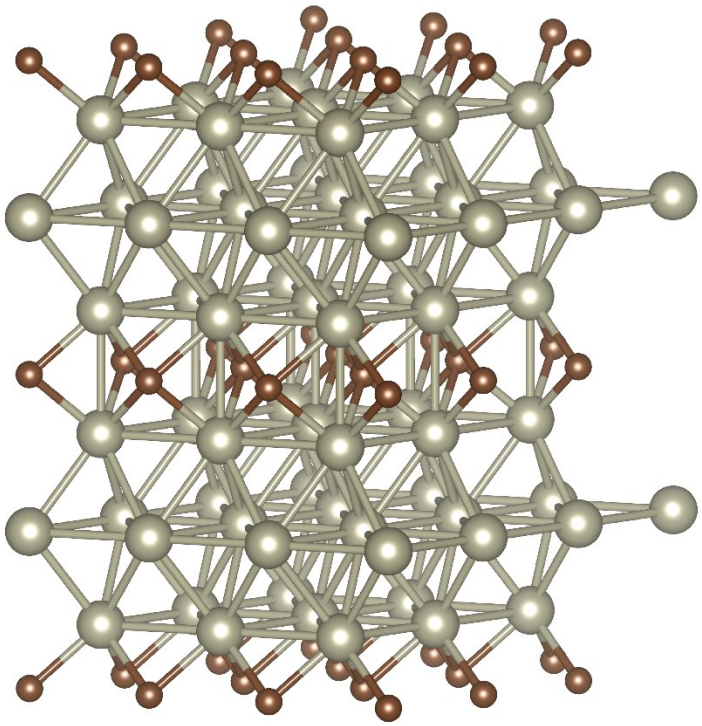
Formula	Re_3C	**ID**	OQMD-1435176
Density (g/cm^3)	18.923	**Symbol**	P-6m2
Formation Energy / Atom (eV)	-0.017	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of Re_3C in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	2.84300000	0.00000000	0.00000000
a_2	-1.4220000	2.46200000	0.00000000
a_3	0.00000000	0.00000000	7.15500000

	x (Å)	*y* (Å)	*z* (Å)
C	0.00000000	0.00000000	0.00000000
Re	0.66700000	0.33300000	0.19600000
Re	0.00000000	0.00000000	0.50000000
Re	0.66700000	0.33300000	0.80400000



13. $IrOs_3$

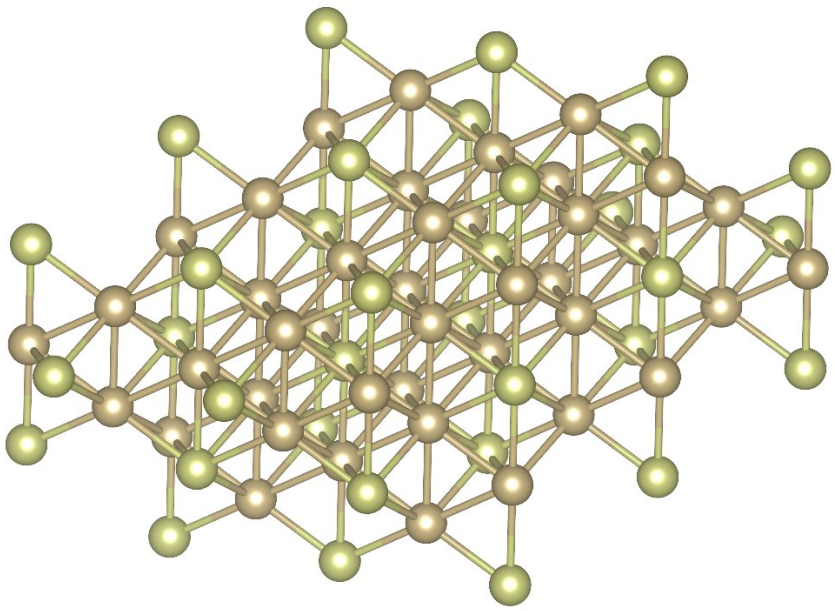
Formula	Os_3Ir	**ID**	OQMD-304348
Density (g/cm^3)	22.318	**Symbol**	I4/mmm
Formation Energy / Atom (eV)	0.091	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of Os_3Ir in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	-1.91800000	1.91800000	3.85800000
a_2	1.91800000	-1.91800000	3.85800000
a_3	1.91800000	1.91800000	-3.85800000

	x (Å)	*y* (Å)	*z* (Å)
Ir	0.00000000	0.00000000	0.00000000
Os	0.75000000	0.25000000	0.50000000
Os	0.25000000	0.75000000	0.50000000
Os	0.50000000	0.50000000	0.00000000



14. ReOs

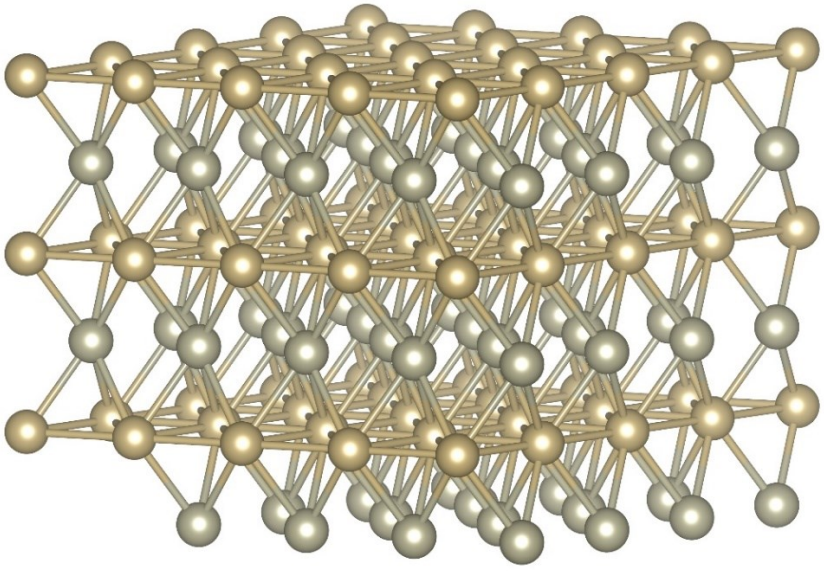
Formula	ReOs	**ID**	OQMD-1230666
Density (g/cm^3)	21.541	**Symbol**	P-6m2
Formation Energy / Atom (eV)	-0.083	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of ReOs in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	2.76400000	0.00000000	0.00000000
a_2	-1.3820000	2.39300000	0.00000000
a_3	0.00000000	0.00000000	4.38900000

	x (Å)	*y* (Å)	*z* (Å)
Os	0.66700000	0.33300000	0.75000000
Re	0.33300000	0.66700000	0.25000000



15. C_3N_4

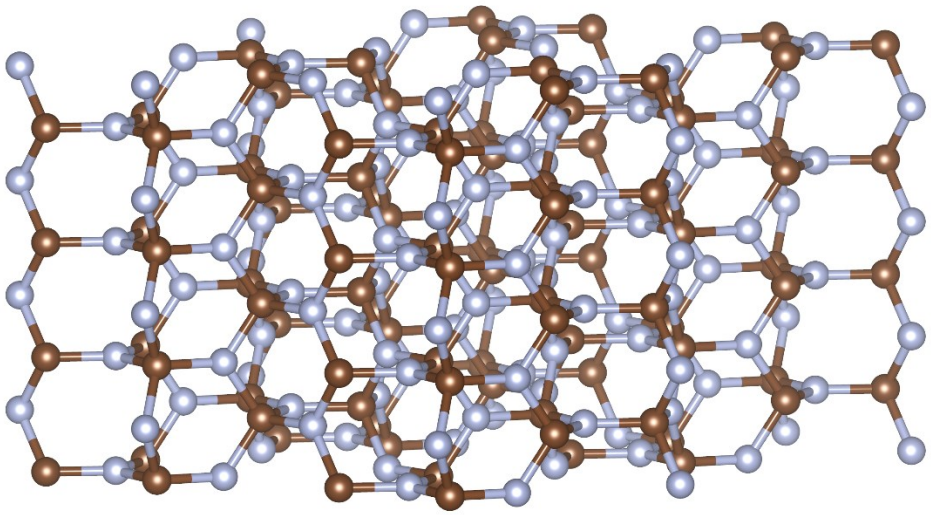
Formula	C_3N_4	**ID**	OQMD-14924
Density (g/cm^3)	3.572	**Symbol**	P-3
Formation Energy / Atom (eV)	0.421	**Band Gap (eV)**	3.4

Crystal structure

Structural parameters: unit cell and atomic positions of C_3N_4 in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	6.40700000	0.00000000	0.00000000
a_2	-3.20400000	5.54900000	0.00000000
a_3	0.00000000	0.00000000	2.40800000

	x (Å)	*y* (Å)	*z* (Å)
C	0.82200000	0.59500000	0.25000000
C	0.40500000	0.22700000	0.25000000
C	0.77300000	0.17800000	0.25000000
C	0.22700000	0.82200000	0.75000000
C	0.59500000	0.77300000	0.75000000
C	0.17800000	0.40500000	0.75000000
N	0.66700000	0.33300000	0.25000000
N	0.29700000	0.96700000	0.25000000
N	0.67000000	0.70300000	0.25000000
N	0.03310000	0.33000000	0.25000000
N	0.96700000	0.67000000	0.75000000
N	0.33000000	0.29700000	0.75000000
N	0.70300000	0.03310000	0.75000000
N	0.33300000	0.66700000	0.75000000



16. $ReOs_6W$

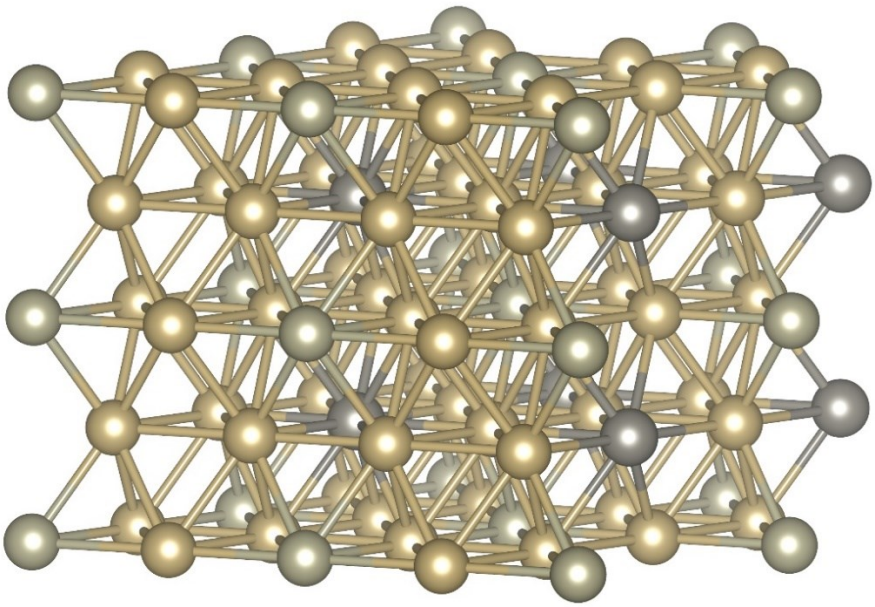
Formula	$ReOs_6W$	**ID**	OQMD-1625537
Density (g/cm^3)	21.643	**Symbol**	P-6m2
Formation Energy / Atom (eV)	-0.072	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of $ReOs_6W$ in fractional coordinates.

	x (Å)	y (Å)	z (Å)
a_1	5.51200000	0.00000000	0.00000000
a_2	-2.75600000	4.77300000	0.00000000
a_3	0.00000000	0.00000000	4.35200000

	x (Å)	y (Å)	z (Å)
Os	0.00211000	0.50100000	0.00000000
Os	0.49900000	0.50100000	0.00000000
Os	0.49900000	0.99800000	0.00000000
Os	0.16600000	0.33100000	0.50000000
Os	0.16600000	0.83400000	0.50000000
Os	0.66900000	0.83400000	0.50000000
Re	0.00000000	0.00000000	0.00000000
W	0.66700000	0.33300000	0.50000000



17. ReOs

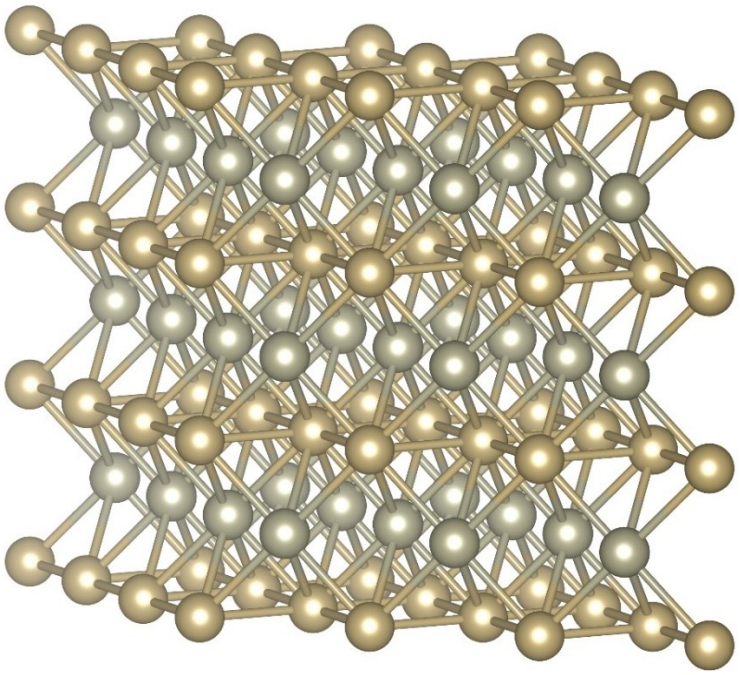
Formula	ReOs	**ID**	OQMD-337952
Density (g/cm^3)	21.523	**Symbol**	P4/mmm
Formation Energy / Atom (eV)	0.166	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of ReOs in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	1.93600000	1.93600000	0.00000000
a_2	0.00000000	3.87300000	0.00000000
a_3	0.00000000	0.00000000	3.87500000

	x (Å)	*y* (Å)	*z* (Å)
Os	0.00000000	0.00000000	0.00000000
Re	1.00000000	0.50000000	0.50000000



18. CN_2

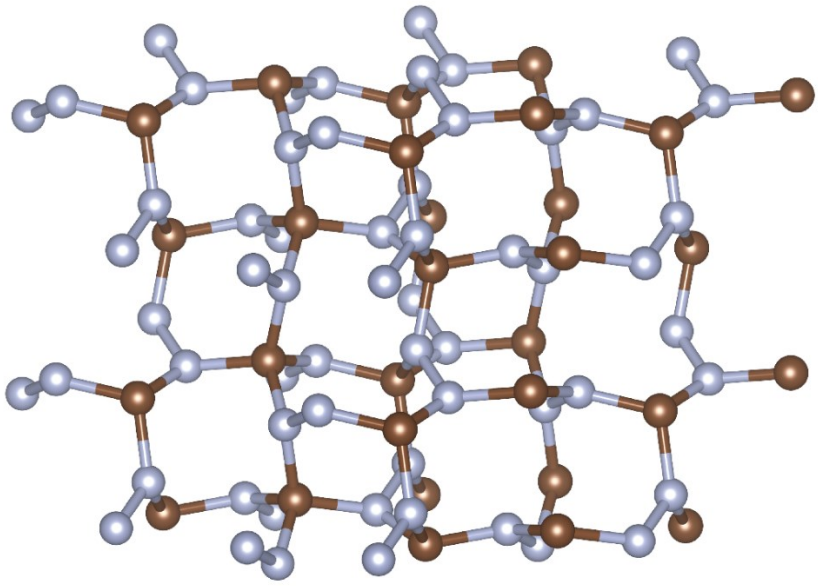
Formula	CN_2	**ID**	mp-1102681
Density (g/cm^3)	3.693	**Symbol**	I-42d
Formation Energy / Atom (eV)	0.554	**Band Gap (eV)**	3.7

Crystal structure

Structural parameters: unit cell and atomic positions of CN_2 in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	4.88940000	0.00000000	0.00000000
a_2	-3.68240625	3.21653797	0.00000000
a_3	-0.60349703	-1.60826940	4.57771703

	x (Å)	*y* (Å)	*z* (Å)
C	0.12498677	0.93209290	0.30709606
C	0.62499678	0.81789535	0.69290859
C	0.06789652	0.37499258	0.19291398
C	0.18207860	0.87498718	0.80709070
N	0.84676605	0.46637538	0.24245156
N	0.40324119	0.14569816	0.11960389
N	0.22390994	0.60430062	0.75753760
N	0.02608042	0.28362343	0.88038528
N	0.53361243	0.77606398	0.38039070
N	0.71636575	0.59676182	0.74244618
N	0.39568853	0.15323697	0.61959851
N	0.85431874	0.97392261	0.25754303



19. $IrOs_6W$

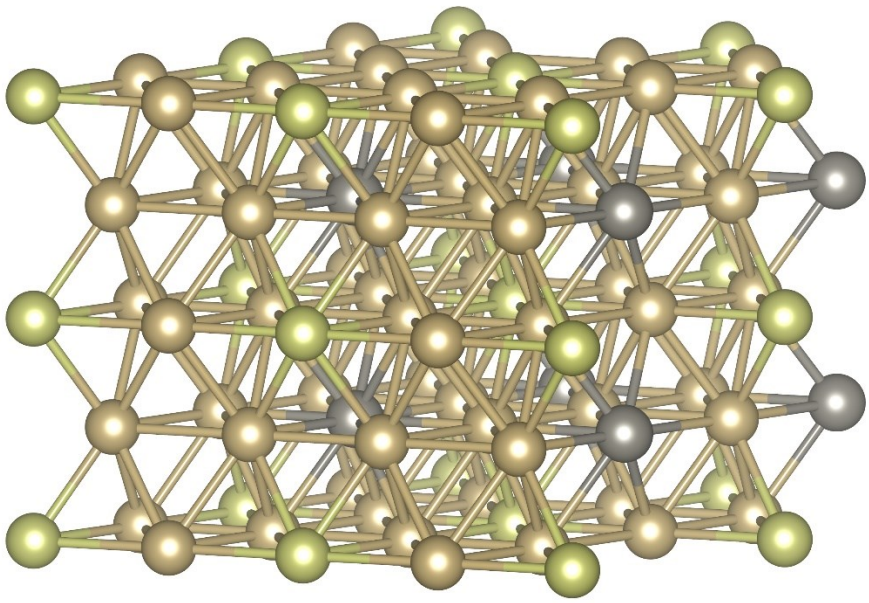
Formula	$IrOs_6W$	**ID**	OQMD-1625528
Density (g/cm^3)	21.782	**Symbol**	P-6m2
Formation Energy / Atom (eV)	-0.053	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of $IrOs_6W$ in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	5.52700000	0.00000000	0.00000000
a_2	-2.76300000	4.78600000	0.00000000
a_3	0.00000000	0.00000000	4.37500000

	x (Å)	*y* (Å)	*z* (Å)
Ir	0.00000000	0.00000000	0.00000000
Os	0.50300000	0.00594000	0.00000000
Os	0.99400000	0.49700000	0.00000000
Os	0.50300000	0.49700000	0.00000000
Os	0.83600000	0.16400000	0.50000000
Os	0.32800000	0.16400000	0.50000000
Os	0.83600000	0.67200000	0.50000000
W	0.33300000	0.66700000	0.50000000



20. ReOs

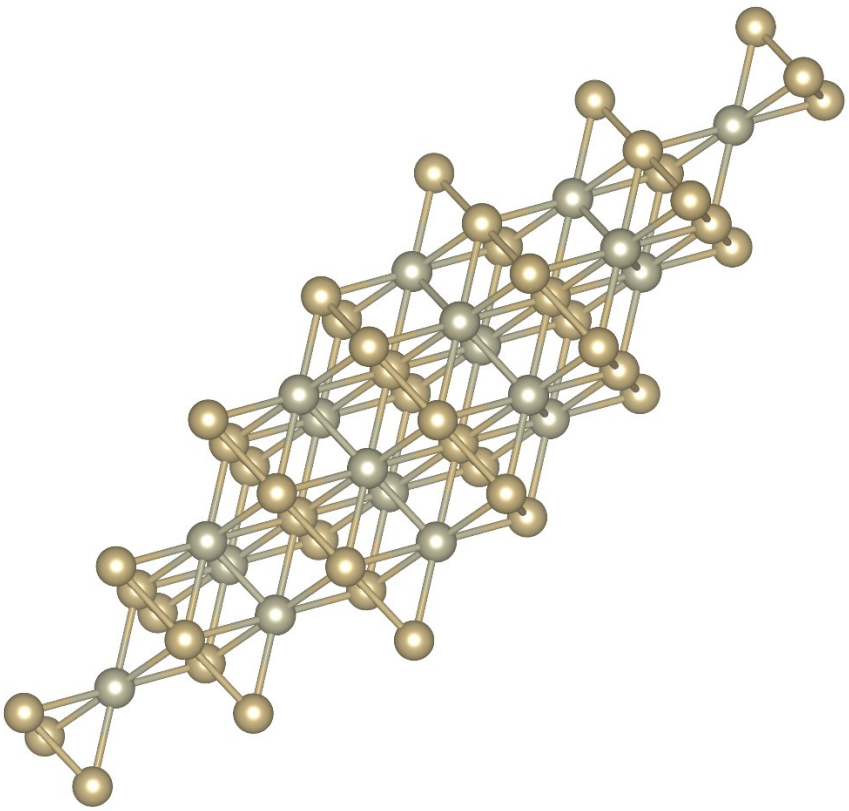
Formula	ReOs	**ID**	OQMD-327494
Density (g/cm^3)	21.453	**Symbol**	R-3m
Formation Energy / Atom (eV)	0.070	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of ReOs in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	1.37400000	0.79300000	4.46000000
a_2	-1.37400000	0.79300000	4.46000000
a_3	0.00000000	-1.58600000	4.46000000

	x (Å)	*y* (Å)	*z* (Å)
Os	0.00000000	0.00000000	0.00000000
Re	0.50000000	0.50000000	0.50000000



21. Re_3Os

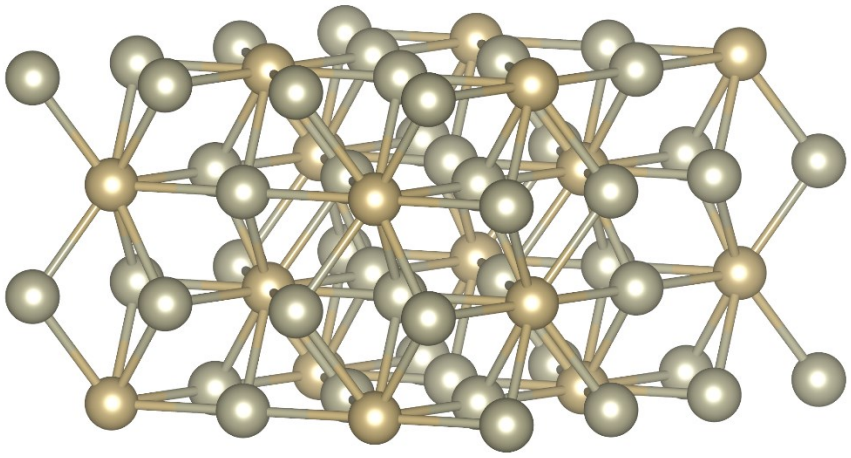
Formula	Re_3Os	**ID**	mp-867264
Density (g/cm^3)	20.894	**Symbol**	P63/mmc
Formation Energy / Atom (eV)	-0.059	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of Re_3Os in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	5.56494093	0.00000000	0.00000000
a_2	-2.78246982	4.81938059	0.00000000
a_3	0.00000000	0.00000000	4.43810320

	x (Å)	*y* (Å)	*z* (Å)
Re	0.83443618	0.16557422	0.75000513
Re	0.33112076	0.16556005	0.75000513
Re	0.83443236	0.66887164	0.75000513
Re	0.16557422	0.83443618	0.75000513
Re	0.66887164	0.83443236	0.75000513
Re	0.16556005	0.33112076	0.75000513
Os	0.66667604	0.33333439	0.75000513
Os	0.33333439	0.66667604	0.75000513



22. $C_{11}N_4$

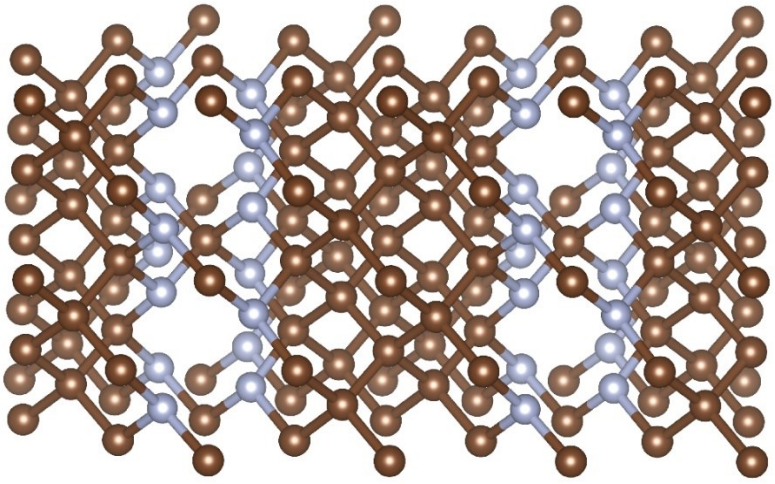
Formula	$C_{11}N_4$	**ID**	mp-1104073
Density (g/cm^3)	3.562	**Symbol**	P-42m
Formation Energy / Atom (eV)	0.303	**Band Gap (eV)**	2.8

Crystal structure

Structural parameters: unit cell and atomic positions of $C_{11}N_4$ in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	3.54172897	0.00000000	0.00000000
a_2	0.00000000	3.54172897	0.00000000
a_3	0.00000000	0.00000000	6.99318409

	x (Å)	*y* (Å)	*z* (Å)
C	0.50001001	0.00000000	0.25700739
C	0.50001001	0.00000000	0.74299490
C	0.00000000	0.50001001	0.74299490
C	0.00000000	0.50001001	0.25700739
C	0.00000000	0.00000000	0.00000000
C	0.50001001	0.50001001	0.00000000
C	0.74480575	0.74480575	0.12819625
C	0.25518611	0.25518611	0.12819625
C	0.25518611	0.74480575	0.87180608
C	0.74480575	0.25518611	0.87180608
C	0.00000000	0.00000000	0.50000113
N	0.76310182	0.76310181	0.62459385
N	0.23688994	0.23688994	0.62459385
N	0.23688994	0.76310182	0.37540838
N	0.76310182	0.23688994	0.37540838



23. $MoReOs_6$

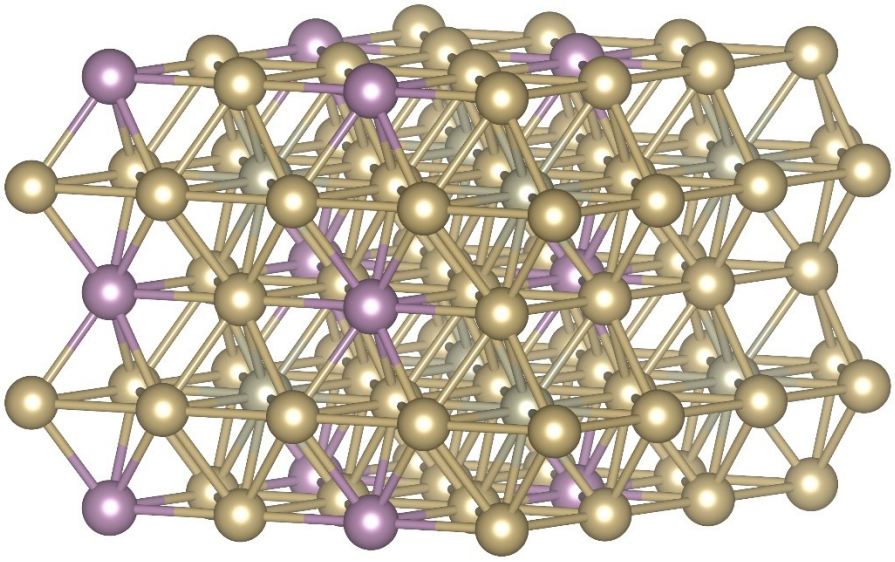
Formula	$ReMoOs_6$	**ID**	OQMD-1625541
Density (g/cm^3)	20.438	**Symbol**	P-6m2
Formation Energy / Atom (eV)	-0.069	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of $ReMoOs_6$ in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	5.52500000	0.00000000	0.00000000
a_2	-2.76200000	4.78400000	0.00000000
a_3	0.00000000	0.00000000	4.37700000

	x (Å)	*y* (Å)	*z* (Å)
Mo	0.66700000	0.33300000	0.00000000
Os	0.16700000	0.33300000	0.00000000
Os	0.16700000	0.83300000	0.00000000
Os	0.66700000	0.83300000	0.00000000
Os	0.33300000	0.16600000	0.50000000
Os	0.83400000	0.16600000	0.50000000
Os	0.83400000	0.66700000	0.50000000
Re	0.33300000	0.66700000	0.50000000



24. Re_6IrOs

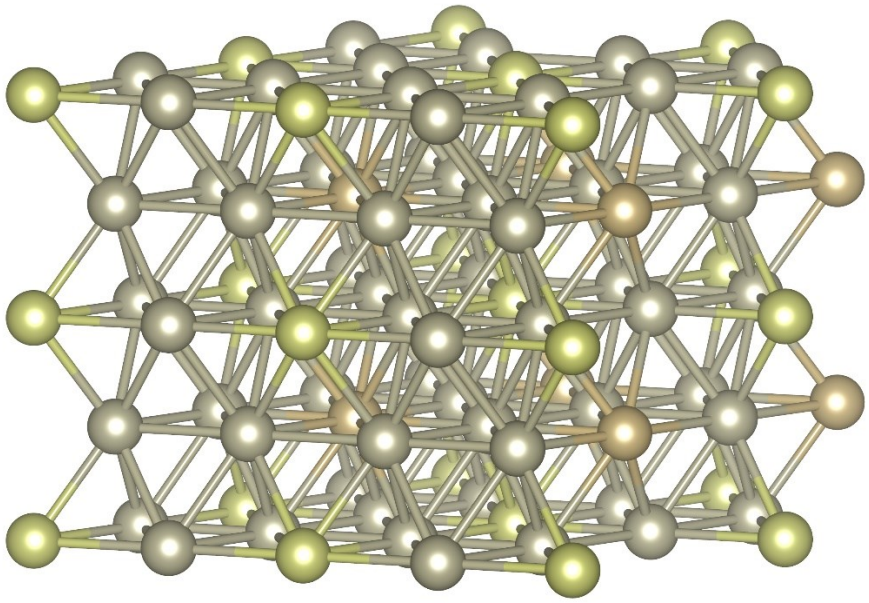
Formula	Re_6IrOs	**ID**	OQMD-1625866
Density (g/cm^3)	21.209	**Symbol**	P-6m2
Formation Energy / Atom (eV)	-0.135	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of Re_6IrOs in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	5.54600000	0.00000000	0.00000000
a_2	-2.77300000	4.80300000	0.00000000
a_3	0.00000000	0.00000000	4.40900000

	x (Å)	*y* (Å)	*z* (Å)
Ir	0.00000000	0.00000000	0.00000000
Os	0.33300000	0.66700000	0.50000000
Re	0.50300000	0.00539000	0.00000000
Re	0.50300000	0.49700000	0.00000000
Re	0.99500000	0.49700000	0.00000000
Re	0.33800000	0.16900000	0.50000000
Re	0.83100000	0.16900000	0.50000000
Re	0.83100000	0.66200000	0.50000000



25. Re_3Os

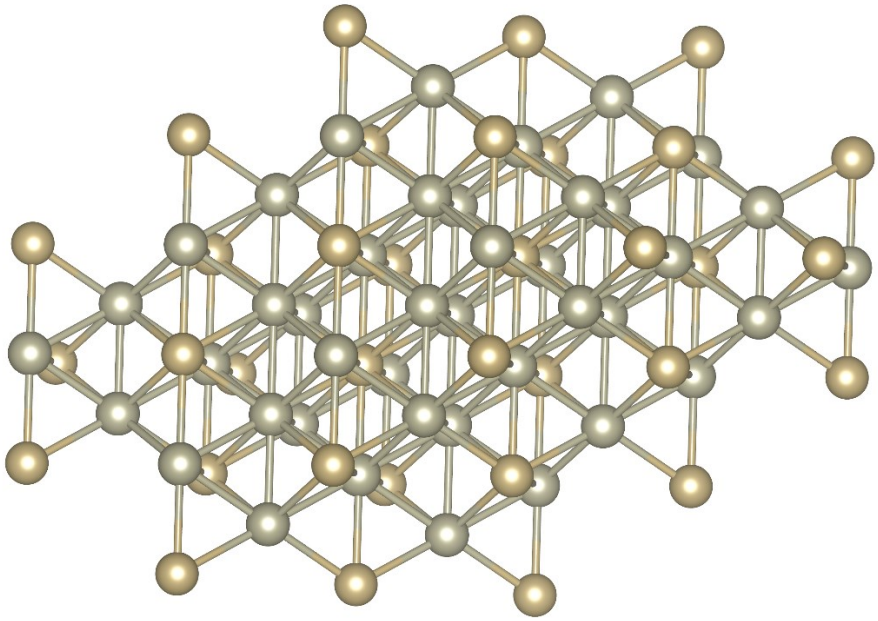
Formula	Re_3Os	**ID**	OQMD-301470
Density (g/cm^3)	21.356	**Symbol**	I4/mmm
Formation Energy / Atom (eV)	0.056	**Band Gap (eV)**	0.0

Crystal structure

Structural parameters: unit cell and atomic positions of Re_3Os in fractional coordinates.

	x (Å)	*y* (Å)	*z* (Å)
a_1	-1.93000000	1.93000000	3.90800000
a_2	1.93000000	-1.93000000	3.90800000
a_3	1.93000000	1.93000000	-3.90800000

	x (Å)	*y* (Å)	*z* (Å)
Os	0.00000000	0.00000000	0.00000000
Re	0.75000000	0.25000000	0.50000000
Re	0.25000000	0.75000000	0.50000000
Re	0.50000000	0.50000000	0.00000000



References

[1] Cohen M L 1993 Predicting useful materials *Science* **261** 307-8

[2] Li K, Ding Z and Xue D 2011 Electronegativity-related bulk moduli of crystal materials *Phys. Status Solidi B* **248** 1227-36

[3] Zeng S, Li G, Zhao Y, Wang R and Ni J 2019 Machine learning-aided design of materials with target elastic properties *J. Phys. Chem. C* **123** 5042-7

[4] Mansouri Tehrani A, Oliynyk A O, Parry M, Rizvi Z, Couper S, Lin F, Miyagi L, Sparks T D and Brgoch J 2018 Machine learning directed search for ultraincompressible, superhard materials *J. Am. Chem. Soc.* **140** 9844-53

[5] Ward L, Dunn A, Faghaninia A, Zimmermann N E R, Bajaj S, Wang Q, Montoya J, Chen J, Bystrom K, Dylla M, Chard K, Asta M, Persson K A, Snyder G J, Foster I and Jain A 2018 Matminer: An open source toolkit for materials data mining *Comput. Mater. Sci.* **152** 60-9

[6] Breiman L 2001 Random forests *Mach. Learn.* **45** 5-32