Sup.1 Featurization

Five types of features that include ElementProperty, IonProperty, Stochiometry, ValenceOrbital, and TMetalFraction were extracted from each of 2D materials and MXenes. These features not only entail information about the constituent elements of materials, such as their average atomic weight, the proportion of transition metals, and the number of unfilled states but also can be readily obtained just from materials' chemical formulas, without the need for additional DFT calculations.

1.1 ElementProperty (132 features)

The ElementProperty extracts information about 22 different elemental properties for each element present in the material and calculates six statistical measures, including mean, minimum, maximum, range, mean absolute deviation, and mode of the attribute values, thereby resulting total of 132 features. 22 different elemental properties are listed as below.

- Atomic number
- Mendeleev number
- Atomic weight
- Melting temperature
- Column
- Row
- Covalent radius
- Electronegativity
- Number of *s* valence electrons
- Number of *p* valence electrons
- Number of *d* valence electrons
- Number of *f* valence electrons
- Total number of valence electrons
- Number of unfilled *s* states
- Number of unfilled *p* states
- Number of unfilled *d* states
- Number of unfilled *f* states
- Total number of unfilled states
- Specific volume of 0 K ground state
- Band gap energy of 0 K ground state
- Magnetic moment (per atom) of 0 K ground state
- Space group number of 0 K ground state

1.2 IonProperty (2 features)

The IonProperty is related to an ionic character of a bond between elements which is computed from the electronegativity difference between two elements by using the relation below where I is the ionic character and X_A and X_B are electronegativities of element A and B, respectively.

$$I(X_A, X_B) = 1 - \exp\left(-0.25(X_A - X_B)^2\right)$$

In a substance with various elements, pairs of elements are made to calculate the ionic characters between them. From these values, the maximum and average are used as features.

1.3 Stoichiometry (6 features)

The Stoichiometric features refer to a set of properties that illustrate the elemental proportions within a material. These characteristics derive from the L^p norms (with p values of 0, 1, 2, 3, 7, 10) of a vector indicating the elemental fraction of the material for each constituent. The selected value of p in the L^p norm influences how sensitive the property is to variances in the elemental fractions. With an increasing value of p in the L^p norm, components with greater concentrations are prioritized more heavily. For an instance, the L^3 norm and L^{10} norm for MoS_2 is:

$$\|x\|_3 = \left(\left(\frac{1}{3}\right)^3 + \left(\frac{2}{3}\right)^3\right)^{1/3} \cong 0.693$$

$$\|x\|_{10} = \left(\left(\frac{1}{3}\right)^{10} + \left(\frac{2}{3}\right)^{10}\right)^{1/10} \cong 0.667$$

1.4 ValenceOrbital (8 features)

The ValenceOrbital attributes are the fraction-weighted average value of the number of valence electrons in each of four orbitals *(s, p, d, f)* and the same value divided by the fraction-weighted average value of the total number of valence electrons, resulting in eight features. As an example, for MoS_2, two ValenceOrbital features corresponding to a *p* orbital can be defined as below

$$p_1 = \frac{1}{3}(0) + \frac{2}{3}(4) = \frac{8}{3} \cong 2.667$$

$$p_2 = \frac{\frac{1}{3}(0) + \frac{2}{3}(4)}{\frac{1}{3}(6) + \frac{2}{3}(6)} = \frac{8}{18} \cong 0.444$$

1.5 TMetalFraction (1 feature)

The TMetalFraction is a feature that indicates the fraction of transition metals among constituent elements. For example, TMetalFraction of MoS_2 is given by

$$T = \frac{1}{3} \cong 0.333$$

Sup.2 Relationship between the Data in C2DB and the aNANt Database

Principal component analysis (PCA) was utilized to visualize and observe the relationship between the data in C2DB and the aNANt database. PCA can effectively reduce the dimensionality of the data by projecting them onto a smaller subspace by identifying the directions, known as principal components, where the variability of the data is most represented. Thus, each data point in both C2DB and aNANt database which was originally represented by 149 dimensional features was reduced to two-dimensions using PCA for the visualization. This method has been widely used in other studies to observe relationships between datasets [1-3]. The result, available in the figure below, shows that the data distribution of the 2D materials from C2DB used for training and the data distribution of MXenes from the aNANt database are noticeably closely located, indicating that the training and testing data are indeed related. The reason for this result, despite structural differences between the two datasets, is that the features used in this study were chemical features extracted from the constituent elements of 2D materials in C2DB and the MXenes in the aNANt database, respectively. Since the constituent elements of MXenes, such as transition metals, halogen atoms, carbon, and nitrogen, have been commonly used in previous 2D materials that were included in C2DB, the data distribution in the MXenes was able to be situated in a space close to that of the 2D materials in the C2DB. This similarity in the data distribution in two-dimensional space validates the relevance of the training and testing datasets.

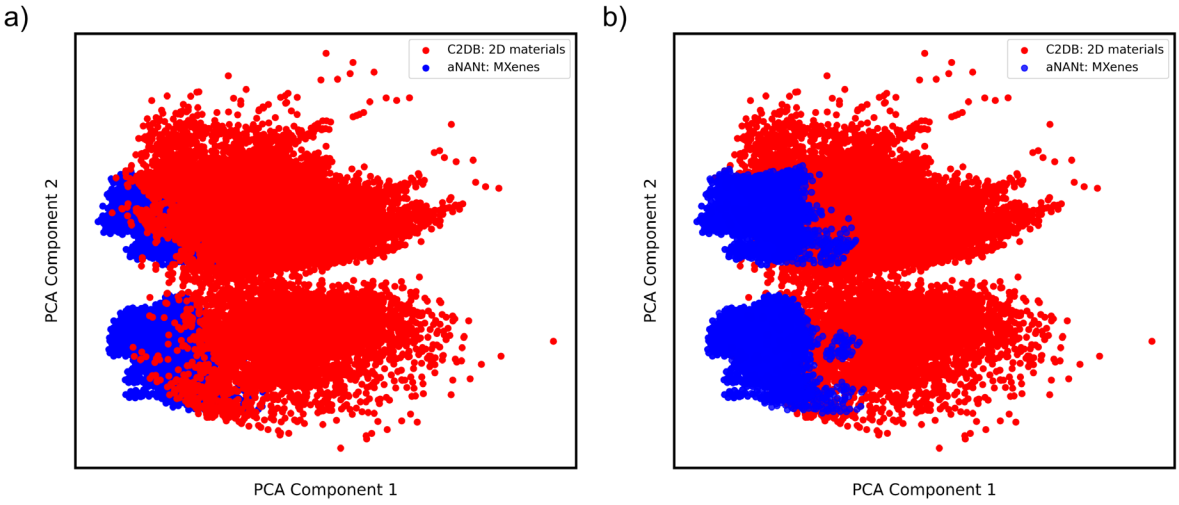

Fig.S1 Data distribution of the 2D materials from C2DB used for training and data distribution of MXenes from the aNANt database are represented in two-dimensional space; red dots indicate 2D materials from C2DB, and blue dots indicate MXenes from the aNANt database. To address the issue of blue dots being obscured by red dots, two cases are presented: a) where blue dots are scattered first, and b) where blue dots are scattered later

Sup.3 Details of Regression Methods on ΔH Predictions

The ML regressors tested in this study are RF, LGBM, ET, DT, and KNN, with DT playing a central role in RF, LGBM, and ET. First, DT begins by splitting the training data into multiple branches based on feature values, forming a tree-like structure. The algorithm selects the most effective feature for splitting and continues this process until a predetermined stopping condition, such as a maximum tree depth, is met. The completed tree then predicts outcomes by averaging the target values of training data segregated into the same branch as the new data. Second, RF constructs a forest of DTs, each built from a random subset of the training data. It trains multiple DTs independently and combines their predictions to produce the final regression outcome. Third, LGBM also employs multiple DTs but constructs them sequentially. It constructs the first DT, makes predictions, and uses the residuals (*i.e.*, the differences between the actual values and the predictions) as inputs for the subsequent DT. As more DTs are added, the residuals diminish, leading to predictions that closely match the actual values. Fourth, ET is also an ensemble method, but its biggest difference from other methods lies in the selection of features for splitting. While DT and RF select the most effective feature for splitting the training data, ET randomly selects this feature. Last, KNN makes predictions for a new data by considering the 'k' closest data (*i.e.*, neighbors) in the training dataset. It measures the similarity or distance between the feature values of the new data and those of its 'k' nearest neighbors, typically using Euclidean distance. KNN performs regression by taking an average of the target values of these neighbors once they are identified.

Sup.4 Regression Method on ΔH_{hull} Predictions

Regressor	R^2	MAE (eV/atom)	RMSE (eV/atom)
RF	0.602±0.017	0.102±0.002	0.145±0.003
LGBM	**0.672±0.020**	**0.086±0.002**	**0.131±0.004**
ET	0.598±0.021	0.103±0.003	0.146±0.004
Catboost	0.659±0.021	0.089±0.002	0.134±0.004

Sup.5 Hyperparameters of ML Algorithms

The hyperparameters of ML algorithms used in this study were tuned with the Bayesian optimization method available in the scikit-optimize package. A total of 150 iterations of optimization were performed for each of ML algorithms and resulted hyperparameters are as follows:

ET Regressor *
bootstrap = False, ccp_alpha = 0.0, max_depth = None, max_features = 'auto', max_leaf_nodes = None, max_samples = None, min_impurity_decrease = 0.0, min_samples_leaf = 1, min_samples_split = 2, min_weight_fraction_leaf = 0.0, n_estimators = 100, n_jobs = − 1, oob_score = False, random_state = 123, verbose = 0, warm_start = False

* : For the ET regressor, after 150 iterations of optimization, the performance was found to be lower than that of the pre-tuned parameters. Therefore, we used the original parameters in this study.

LGBM Classifier #
bagging_fraction = 0.7915053846417077, learning_rate = 0.055119720251062806, min_child_samples = 1, n_estimators = 278, num_leaves = 176, random_state = 123, reg_alpha = 8.613554304482821e−05, reg_lambda = 2.981904181205781e−06

CatBoost Classifier [#]
iterations = 236, l2_leaf_reg = 10, random_strength = 0, boost_from_average = False, subsample = 0.800000011920929, use_best_model = False, class_names = [0,1], random_seed = 123, depth = 10, learning_rate = 0.0745200514793396, bootstrap_type = 'MVS', max_leaves = 1024

RF Classifier [#]
bootstrap = False, max_features = 0.4, min_impurity_decrease = 4.091960728716533e−06, min_samples_leaf = 2, n_estimators = 300, n_jobs = − 1, random_state = 123

[#] : Hyperparameters not mentioned here are set to their default values.

Sup.6 Additional 15 Thermodynamically Stable MXenes

MXene	ΔH (eV/atom)	ΔH_{hull} (eV/atom)	MXene Type	Thermodynamic Stability	Dynamic Stability
Hf_2CBr_2	− 1.22	0.00	Iso-stoichiometric	Stable	Stable
Hf_2CH_2	− 0.63	0.09	Iso-stoichiometric	Metastable	Stable
Nb_2CBr_2	− 0.75	0.00	Iso-stoichiometric	Stable	Stable
Sc_2CH_2	− 0.56	0.08	Iso-stoichiometric	Metastable	Stable
Ta_2CBr_2	− 0.68	0.08	Iso-stoichiometric	Metastable	Stable
Ti_2CCl_2	− 1.60	0.00	Iso-stoichiometric	Stable	Stable
Ti_2NBr_2	− 1.44	0.01	Iso-stoichiometric	Metastable	Stable
Ti_2NCl_2	− 1.96	0.00	Iso-stoichiometric	Stable	Stable
V_2CBr_2	− 0.63	0.07	Iso-stoichiometric	Metastable	Stable
V_2NCl_2	− 1.39	0.08	Iso-stoichiometric	Metastable	Stable
Y_2CBr_2	− 1.31	0.00	Iso-stoichiometric	Stable	Stable
Zr_2CBr_2	− 1.25	0.00	Iso-stoichiometric	Stable	Stable
Zr_2CCl_2	− 1.71	0.00	Iso-stoichiometric	Stable	Stable
Zr_2CH_2	− 0.61	0.10	Iso-stoichiometric	Metastable	Stable
Zr_2NBr_2	− 1.53	0.00	Iso-stoichiometric	Stable	Stable

Sup.7 Energy of the Convex Hull (ΔH_{hull}) Values of 120 MXenes

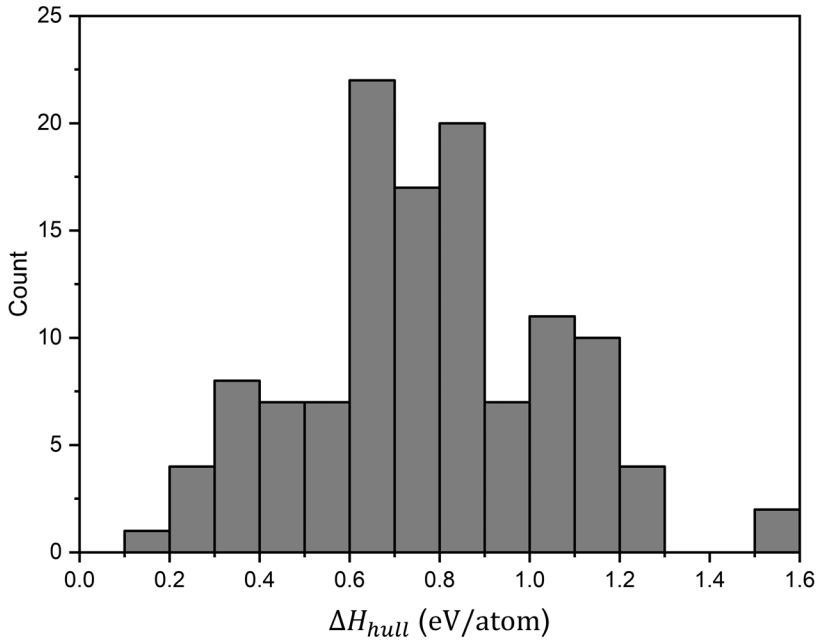

Fig.S2 Distribution of ΔH_{hull} values of randomly selected 120 MXenes

Sup.8 Phonon Dispersions of Thermodynamically Stable MXenes

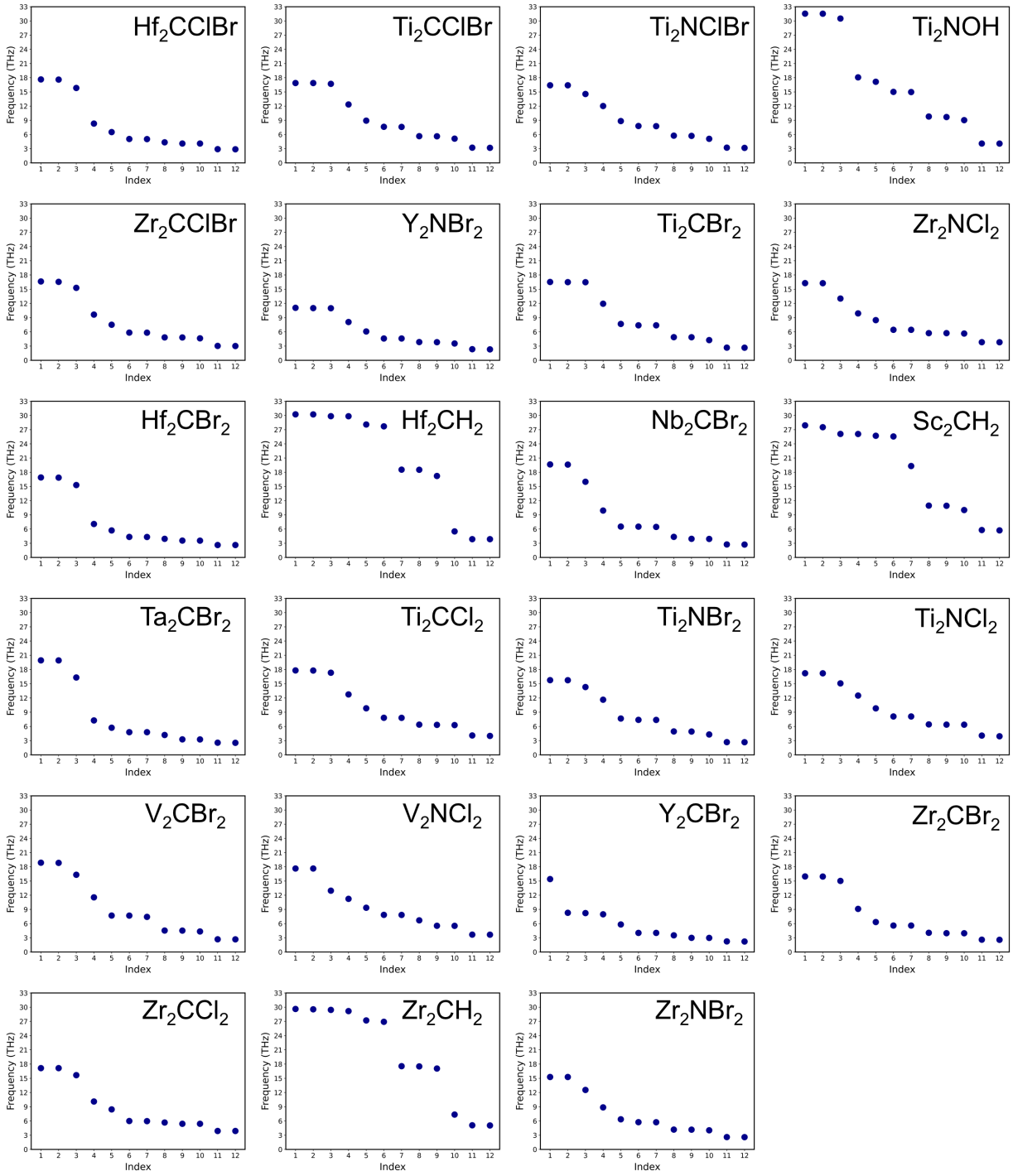

Fig.S3 Phonon information at the Γ-point for the 23 MXenes identified as stable for the first time in this study; horizontal axis represents the phonon modes without translational modes and vertical axis represents eigenfrequencies

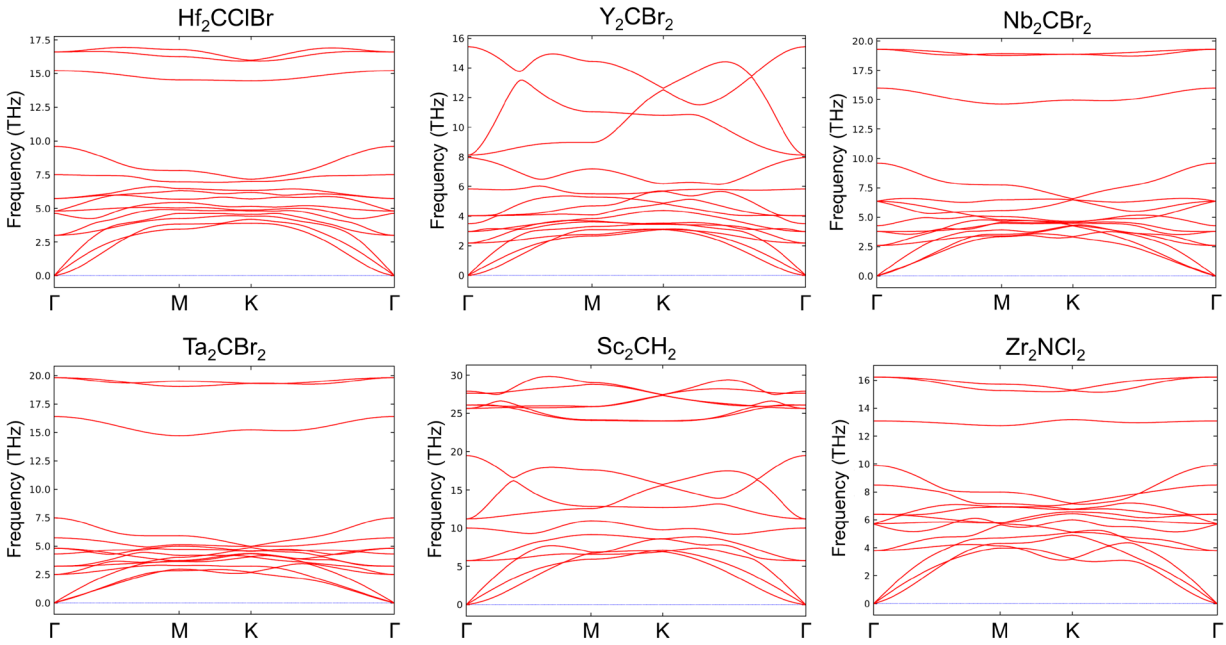

Fig.S4 Phonon band structures of six thermodynamically stable MXenes along high symmetry lines

Sup.9 Rationale for Exploring MXenes in aNANt Database

Despite the differences between the 23,857 MXenes and those successfully synthesized, the rationale for exploring this chemical space in this study is as follows:

1) The thermodynamic stability of these MXenes has not yet been fully assessed, either through DFT calculations or experimentally, making their exploration worthwhile. Investigating these MXenes, especially the ones that have not been successfully synthesized before, and selecting those with a high likelihood of synthesis can pave the way for developing new MXenes with novel chemical compositions and stoichiometries. Moreover, some MXenes, that include $Ti_3C_2Br_2$ and Nb_2CS_2, have been successfully synthesized despite having constituent elements that were not previously considered, underscoring the potential in investigating this unexplored chemical space [4, 5].
2) Janus MXenes, MXenes with distinct surface terminations and transition metals, hybrid transition metal MXenes, and iso-stoichiometric MXenes with uniform surface terminations are extensively studied with DFT calculations [6-11]. Particularly in the context of applications of MXenes, surface

termination plays a crucial role and MXenes with a uniform surface termination are preferred, especially when utilizing them as electrodes for Li-S batteries [12].

3) MXenes are known to have the broadest chemical space among 2D materials when considering theoretical materials as well, and if only the chemical space of successfully synthesized MXenes is considered, that space becomes limited. This study aimed to fully leverage the wide chemical space of MXenes, considering a diverse range of MXenes both in terms of constituent elements and stoichiometries. Hence, the MXenes in aNANt database, diverse in these two aspects, were chosen as the target materials to explore.

Sup.10 Mechanical Properties of 43 Thermodynamically Stable MXenes

MXene	E_x^{2D} (N/m)	E_y^{2D} (N/m)	G_{xy}^{2D} (N/m)	v_{xy}^{2D}
Hf_2CBr_2	157.4	158.6	123.2	0.28
Hf_2CClBr	151.5	155.7	118.8	0.29
Hf_2CCl_2	147.5	150.7	108.2	0.32
Hf_2CF_2	98.9	100	64	0.55
Hf_2CH_2	197.5	198.0	163.4	0.2
Hf_2NCl_2	136.2	136.5	102.4	0.34
Nb_2CBr_2	125.6	124.1	88.0	0.45
Sc_2CBr_2	131.6	131.2	105.4	0.25
Sc_2CCl_2	137.0	136.5	110.8	0.24
Sc_2CF_2	160.4	160.3	131.0	0.22
Sc_2CH_2	147.8	147.4	122.6	0.21
Sc_2NBr_2	136.6	136.5	109.0	0.26
Sc_2NCl_2	142.5	141.6	114.6	0.24
Sc_2NF_2	159.4	159.3	128.6	0.24
Ta_2CBr_2	118.3	118.6	83.2	0.42
Ta_2CCl_2	90.1	92.3	74.0	0.2

$TiTaCCl_2$	173.0	172.0	141.6	0.22
Ti_2CBr_2	175.9	176.5	143.8	0.23
Ti_2CClBr	178.6	178.6	146.0	0.23
Ti_2CCl_2	179.0	179.5	146.6	0.22
Ti_2CH_2	177.0	178.6	148.2	0.19
Ti_2NBr_2	156.8	156.2	120.2	0.3
Ti_2NClBr	159.5	159.0	123.8	0.29
Ti_2NCl_2	161.9	161.8	125.8	0.28
Ti_2NF_2	117.5	117.6	86.6	0.35
Ti_2NOH	184.2	185.1	141.2	0.31
V_2CBr_2	133.2	133.2	93.6	0.43
V_2CCl_2	127.2	127.0	85.8	0.48
V_2NCl_2	101.7	102.9	64.2	0.58
Y_2CBr_2	109.2	107.4	86.4	0.25
Y_2CCl_2	112.7	113.1	90.6	0.24
Y_2CF_2	138.6	138.8	113.0	0.23
Y_2NBr_2	112.3	112.2	88.4	0.26
Y_2NCl_2	115.4	116.1	90.8	0.27
Y_2NF_2	128.8	128.9	101.8	0.26
Zr_2CBr_2	144.7	145.0	113.0	0.29
Zr_2CClBr	140.6	142.2	109.0	0.29
Zr_2CCl_2	141.1	144.4	111.4	0.29
Zr_2CF_2	117.7	120.2	80.2	0.48
Zr_2CH_2	171.6	173.8	141.0	0.22
Zr_2NBr_2	136.2	135.4	104.2	0.3
Zr_2NCl_2	138.1	138.0	108.0	0.28
Zr_2NF_2	127.0	128.8	94.8	0.35

Sup.11 Expandability of the Proposed Framework

The proposed framework can be adjusted and expanded for practical applications through the three methods presented below.

i) A new ML model capable of predicting specific properties important to a practical application can be constructed and integrated with the proposed framework. For instance, to utilize MXenes as catalysts for the hydrogen evolution reaction (HER), the Gibbs free energy of hydrogen adsorption (ΔG_H) serves as an important descriptor indicating their performance. In this case, the proposed framework can be expanded by adding an ML model and using it to predict ΔG_H of thermodynamically stable MXenes. Consequently, it will be possible to expeditiously identify MXenes that are both thermodynamically stable and applicable as HER catalysts.

ii) Another method to expand this framework is to concatenate parameters related to the applications they are involved in with the 149 chemical features used in the proposed framework, thereby increasing the dimensionality of the features. This can help in accurately predicting both thermodynamic properties and specific properties related to practical applications. For example, since the Bader charge of transition metals, X site atoms (*i.e.*, carbon and nitrogen), and surface terminations, as well as bond lengths between them are known descriptors that can affect HER performance of MXenes [13], adding these to the feature dimension can expand the proposed framework towards the field of HER catalysts.

iii) Lastly, the proposed framework can be tuned to operate as a multi-output model. That is, ML models can be made to predict both thermodynamic properties and specific properties that are relevant to practical applications. In this way, it will be able to predict whether MXene meets the requirements for thermodynamic stability and the applications at the same time.

REFERENCES

1. See, X. Y.; Wen, X.; Wheeler, T. A.; Klein, C. K.; Goodpaster, J. D.; Reiner, B. R.; Tonks, I. A., Iterative supervised principal component analysis driven ligand design for regioselective Ti-catalyzed pyrrole synthesis. ACS catalysis 2020, 10 (22), 13504-13517.
2. Tranås, R.; Løvvik, O. M.; Tomic, O.; Berland, K., Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: Enhancing predictivity by active sampling with principal component analysis. Computational Materials Science 2022, 202, 110938.
3. Gorgannejad, S.; Gahrooei, M. R.; Paynabar, K.; Neu, R., Quantitative prediction of the aged state of Ni-base superalloys using PCA and tensor regression. Acta Materialia 2019, 165, 259-269.
4. Kamysbayev, V.; Filatov, A. S.; Hu, H.; Rui, X.; Lagunas, F.; Wang, D.; Klie, R. F.; Talapin, D. V., Covalent surface modifications and superconductivity of two-dimensional metal carbide MXenes. Science 2020, 369 (6506), 979-983.
5. Li, M.; Li, X.; Qin, G.; Luo, K.; Lu, J.; Li, Y.; Liang, G.; Huang, Z.; Zhou, J.; Hultman, L., Halogenated Ti3C2 MXenes with electrochemically active terminals for high-performance zinc ion batteries. ACS nano 2021, 15 (1), 1077-1085.
6. Özcan, S.; Biel, B., Exploring a novel class of Janus MXenes by first principles calculations: structural, electronic and magnetic properties of Sc 2 CXT, X= O, F, OH; T= C, S, N. Physical Chemistry Chemical Physics 2023, 25 (3), 1881-1888.
7. Zhao, X.; Wang, P.; Lv, E.; Wu, C.; Ma, K.; Gao, Z.; Gates, I. D.; Yang, W., Screening MXenes for novel anode material of lithium-ion batteries with high capacity and stability: A DFT calculation. Applied Surface Science 2021, 569, 151050.
8. Zeng, Z.; Chen, X.; Weng, K.; Wu, Y.; Zhang, P.; Jiang, J.; Li, N., Computational screening study of double transition metal carbonitrides M′ 2M ″CNO2-MXene as catalysts for hydrogen evolution reaction. npj Computational Materials 2021, 7 (1), 80.
9. Pandey, M.; Thygesen, K. S., Two-dimensional MXenes as catalysts for electrochemical hydrogen evolution: A computational screening study. The Journal of Physical Chemistry C 2017, 121 (25), 13593-13598.
10. Jin, W.; Wu, S.; Wang, Z., Structural, electronic and mechanical properties of two-dimensional Janus transition metal carbides and nitrides. Physica E: Low-dimensional Systems and Nanostructures 2018, 103, 307-313.
11. Zhan, C.; Sun, W.; Kent, P. R.; Naguib, M.; Gogotsi, Y.; Jiang, D.-e., Computational screening of MXene electrodes for pseudocapacitive energy storage. The Journal of Physical Chemistry C 2018, 123 (1), 315-321.
12. Anasori, B.; Gogotsi, Û. G., 2D metal carbides and nitrides (MXenes). Springer: 2019; Vol. 416.

13. Zheng, J.; Sun, X.; Qiu, C.; Yan, Y.; Yao, Z.; Deng, S.; Zhong, X.; Zhuang, G.; Wei, Z.; Wang, J., High-throughput screening of hydrogen evolution reaction catalysts in MXene materials. The Journal of Physical Chemistry C 2020, 124 (25), 13695-13705.