Self-validated Machine Learning Study of Graphdiyne-based Dual Atomic Catalyst

Mingzi Sun[1], Tong Wu[1], Alan William Dougherty[2], Maggie Lam[1], Bolong Huang[1], Yuliang Li[3], Chun-Hua Yan[4,5]*

M. Sun, T. Wu, M. Lam, Prof. B. Huang

Department of Applied Biology and Chemical Technology,

The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, China

E-mail: bhuang@polyu.edu.hk

Dr. A. W. Dougherty

Hong Kong Applied Science and Technology Research Institute, Shatin, New Territories, Hong Kong SAR, China

Prof. Y. Li

Institute of Chemistry, Chinese Academy of Sciences, Beijing, 100190, PR China

Prof. C.-H. Yan

Beijing National Laboratory for Molecular Sciences, State Key Laboratory of Rare Earth Materials, Chemistry and Applications, PKU-HKU Joint Laboratory in Rare Earth Materials and Bioinorganic Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

Prof. C.-H. Yan

College of Chemistry and Chemical Engineering, Research Center of Biomedical Nanotechnology, Lanzhou University, Lanzhou 730000, China

Keywords: dual-atomic catalyst, graphdiyne, *f-d* orbital coupling, machine-learning, self-validation

Although the atomic catalyst has attracted intensive attention in the past few years, the current progress of this field is still limited to the single atomic catalyst (SAC). With very few

successful cases of dual atomic catalysts (DACs), the most challenging part of experimental synthesis still lies in two main directions: the thermodynamic stability of synthesis and the optimal combination of metals. To address such challenges, we propose comprehensive theoretical investigations on graphdyine (GDY)-based DAC by considering both the formation stability and the *d*-band center modifications. Unexpectedly, we prove the introduction of selected lanthanide metals (Ln) to the transition metals (TM) contributes to the optimized stability and electroactivities. With further verification by machine-learning, we unravel the potential *f-d* orbital coupling as the pivotal factor in modulating the *d*-band center with enhanced stability by less orbital repulsive forces. These findings supply the delicate explanations of the atomic interactions and screen out the most promising DAC to surpass the limitations of conventional trial and error synthesis. This work has supplied an insightful understanding of DAC, which opens up a brand new direction to advance the research in atomic catalysts for broad applications.

Due to increasing demands for catalyst performances, nanoengineering based strategies have been proposed to design the electrocatalyst with superior performances. The advanced performances of the individual metal site rather than nanoparticles in electrocatalysts have been proposed by Flytzani-Stephanopoulos *et al.* in the water-gas shifting reactions and Bashyam & Zelenay in oxygen reduction reactions (ORR)[1]. With the development of the characterization technique, the successful realization of isolated Pt sites by Zhang *et al.* has been confirmed that plays a significant role in the electrocatalysis of CO oxidation[2]. Since then, the heterogeneous catalyst has entered a new stage. As one of the most effective solutions in the nanoengineering approaches, atomic catalysts (ACs) have become the most significant research topics due to their superior performances mastering in broad chemical reactions covering chemical

synthesis[3], green energy production[4], pollution gas conversion[2, 5]. Currently, the transition metals with well-known electroactivity have been considered as promising selections for AC, including Pt, Pd, Co, Ni, Fe, Ru, Ir, etc[6]. Although the unique structural properties of Ln metals at the atomic scale may further endow them with an unexpected performance, the rare-earth metal based AC are only reported by a few groups[7]. Such a phenomenon is not only induced by electroactivity limitations but also the synthesis challenges. Owing to the significant progress achieved in AC, the understanding of the electron transfer and interactions at the atomic level has been successfully achieved in previous theoretical explorations[5b, 8]. However, the in-depth understanding of the electronic structure in AC still requires further explorations from the advanced theoretical approaches. Therefore, regarding the impressive progress in the single atomic catalyst, it is the timing for a further step towards the DAC[9].

Although the SAC has achieved success in the past decade, the single atomic active site of one single metal also limits the modulation space for further performance improvements in many material systems. Through the introduction of the second metal atom, the electronic redistribution of surface active sites is able to further optimize the adsorption strength of the intermediates during the reactions, which further improves the electroactivity. Although there have been some theoretical investigations[10], the DAC has only been achieved in very few cases from experiments. Utilizing MOF with different metal precursors has been another potential approach to achieve DAC[11]. Some other approaches such as competitive complexation strategy, bottom-up atomic layer deposition[12]. Li et al. have successfully prepared the DAC based on the precursor-preselected" wet-chemistry strategy[12a]. Meanwhile, for the DACs without direct bonding between metals, several approaches have been introduced. Recently, Fan et al. have reported the W-Mo DAC for pH-universal hydrogen evolution reaction (HER),

which realizes the stable anchoring and fine distribution based on the formation of O-coordinated bonding environments[12b]. A similar strategy has also been introduced in the Co-Fe DAC by Hu et al. to stabilize the two metal atoms[12c]. To achieve the stable DACs, several main challenges existed in the synthesis including 1) the aggregation issues of the surface metal atoms, 2) the competition during the synthesis, and 3) The identification of valence states for surface metal atoms.

Considering the current developments of DACs, several significant questions have appeared – What is the real dual atomic catalyst? What are the correlations between two metal atomic sites? The DAC is constructed by two close SAC sites without direct bonding or the coupling of two neighboring metal atomic sites? These questions are revealed under three circumstances in **Figure 1**. For the first circumstances, owing to the strong repulsive force between *d*-orbitals, the synthesis of both homo-coupling and hetero-coupling TM-based DACs usually lead to the isolated metal sites on the surface. The individual SAC will be more favored during the SAC. For the second circumstance, additional assistances will be introduced to facilitate the synthesis of DAC such as O or N atoms to coordinate the two different metal atom sites. For the last circumstance, the introduction of Ln ions activates the *f-d* coupling to stabilize the DAC and achieve high electroactivity, which results in a preferred DAC formation than the Ln-SAC or TM-SAC. Although these questions have been pivotal for the future design strategy of DAC, related investigations and explanations are still significantly lacking.

Until now, the synthesis method, characterization techniques, electrocatalysis mechanism have achieved impressive progress[3c, 13]. However, the developments of the DAC still face significant challenges regarding the precise control of metal atom anchoring, the suppression

of metal agglomerations, and the accurate characterizations. Moreover, such promising strategies of DAC is challenged by a large number of possible combinations among all the TM and Ln, which requires the extremely high cost to verify the performances of these catalysts case by case. The best solution is to integrate the theoretical calculations and machine learning techniques to effectively explore and screen the most possible and potential candidates, which supplies the pivotal guidance for experiments.

The most significant foundation for a good AC largely relies on the strong interactions between the anchoring metal atom and the support material, which is the key factor to guarantee the stable isolation of surface metal atoms as active sites to achieve high electrocatalytic performances[8b]. The available support materials for AC are classified into carbon-based materials and oxides. The carbon-based support has offered a new type of substrate with unique low dimension structures and coordination environments. Among different carbon-based support materials, the GDY has become one of the most promising candidates. Most importantly, GDY is composed of both sp^2 and *sp* hybridized carbon atoms with a high level of π conjugations, where the unique *sp* hybridization optimizes the coordination of metal atoms and the adsorption of intermediates during the electrocatalysis. In addition, the naturally uniform pores in GDY have supplied the suitable anchoring sites for both the SAC and DAC. Our previous works also confirm the untapped potential of GDY-based AC in realizing novel zero-valence electrocatalysts, which propose the unique selection of anchoring metals for future experimental synthesis design[14]. Herein, we have performed the parallel theoretical investigations on the GDY based DACs by considering all the combinations of TM and Ln metals for the first time. Regarding the stability and the electroactivity of the DAC, we have screened out the most promising experimental synthesis directions to achieve highly stable and

electroactive catalysts. The absence of the homo-coupling and hetero-coupling DAC in the current SAC synthesis process is ascribed to the strong *d-d* repulsive force. Compared with the highly unstable TM based DACs, the introduction of Ln metals to SAC realizes the electronic self-balance effect in the Ln-TM DACs systems through the *f-d* coupling effect. This work has opened a new direction for searching the efficient DAC candidates with desired electrocatalytic properties, supplying the pivotal understanding and guidance to the future experimental synthesis of atomic catalyst developments.

Results and Discussions

As the most important factor, the thermodynamic trends of the formation should be one of the most important factors to determine the synthesis possibility in experiments. To clearly compare the formation preference, we have demonstrated the combinations between all TM and Ln metals. Considering the similar atomic radius and electronic configurations, anchoring two TM metals from the same row seems to be the most potential solution. The formation energies of such DACs have been demonstrated in **Figure 2a-c**. Interestingly, it is noted that such DACs are not generally favored, where the majority of the DACs show the positive formation energies. The only exception appears at the Sc based 3d-3d DACs due to the weak interactions between the nearly empty d-bands. The accommodation effect of the empty orbitals is more evident in the Ln metal with *f*-orbitals. These results support that the DACs constructed by similar TM might not be the optimal solution for the experimental synthesis induced by the nature of instability. Among the same row, the formation energies are not regularly changed with the periodic table. However, as the mass number of transition metal becomes larger, the formation energies follow the order as 3d-3d<4d-4d<5d-5d, which indicates the low formation preference of these DAC. These results explain that the absence of TM clusters in the previous

GDY-based SAC works is ascribed to the much lower thermodynamic preference[8a, 15]. The mixture of TM has significantly balanced the formation energies of DACs, showing a similar energy range (**Figure 2d-f**). Notably, we noticed the existence of a peak region near d^5 and a valley region near d^8 in the energies of all the transition metal combinations. The phenomenon is mostly regulated in 3d-4d and 3d-3d DAC. As the 4d and 5d TMs are introduced, such a phenomenon has been slightly perturbed. Such a trend demonstrates the existence of late transition metals leads to the most stabilized structure while the d^3-d^5 transition metals should not be considered as components of DAC. Unexpectedly, the introduction of Ln has shown a significant role in stabilizing the DAC system, which shows much-lowered formation energies (**Figure 2g-i**). For the Ln-3d DACs, most of the systems display thermodynamically favored trends, in which the experimental synthesis of these DACs are highly promising with lowe energy costs or even spontaneous. Moreover, the appearance of Ln has strongly alleviated the fluctuations of energy change. Although we notice some extreme points, the evident fluctuations with contrasting peak and valley regions are not evident. The formation energies of the homo-coupling DACs are shown in Figure S1. It is noticed that as the d-orbital increases, the overall formation energies increases from 3d-3d to 5d-5d homo-coupling DACs. In comparison, the formation of Ln-Ln homo-coupling DACs is much more favorable since most of the formation energies are exothermal. Meanwhile, the energy difference between the homo-coupling and hetero-coupling is not evident, supporting the existence of competition between DACs. Compared to the TM-TM homo-coupling DACs, the Ln-TM hetero-coupling DACs show much lower formation energies, supporting the preference of the Ln-TM hetero-coupling DACs during the formation competition. The introduced Ln perturb the regular energetic variations of the TM based DAC, enabling more flexibility in the thermodynamic modulations by the potential interactions between TM and Ln metals. Therefore, from the thermodynamic

perspective, the combination of transition and Ln metals is a promising solution to realize the stable DAC.

To demonstrate the structure, we have selected the representative structures (the most stable) for all the combinations (3d-3d, 3d-4d, 3d-5d, 4d-5d, 4d-4d, 5d-5d, Ln-3d, Ln-4d, Ln-5d, and Ln-Ln) of GDY-based DACs in Figure S2-S3 and the corresponding bond lengths have been summarized in Table S1. Notably, the minimum equilibrium bond lengths between the metal sites vary largely from 2.52 to 4.01 Å, which correspond to distinct structures of DACs. For most of the conditions, we notice the stabilization of metal atoms in the same plane within the pore of the GDY, such as Sc-Sc, Zr-Zr, and Hf-Hf DACs, in which the bond length changes are very similar. For the hetero-coupling DACs, different distortions are noted. Owing to the different ionic radius and electronic structures, the atomic distributions are not symmetrical, leading to different distorition of GDY bonding as well. For the 4d-5d and 5d-5d hetero-coupling DACs, we notice the strong repulsive force leads to the separation of metal from the GDY, forming the dangling metal sites on the GDY. Compared to the TM, Ln metals prefer to occupy the central place of the pores in GDY, in which the dual-metal sites are not at the same plane. Such a unique structure results in the alleviation of the formation energies in Ln-based DACs. These results suggest that different morphologies of the DAC are possible for different combinations.

To further explicitly compare the formation energy difference among the metal combinations, all the formation energies have been mapping in **Figure 3a**. Evidently, the data located below the zero line are dominated by the Ln-TM combinations, supporting their intrinsic high stability after formation, implying the *f-d* coupling effect is playing a pivotal role in realizing the

thermodynamic stability of anchoring the dual metal atom sites on the GDY. The TM-TM DACs are mostly unstable while the Ln-Ln DACs show the overall highest stability. Meanwhile, since our previous works have confirmed that GDY-Ln catalysts are active in the inter- and intra- atomic electron transfer, the strong stability of the Ln-Ln DACs are attributed to the *f-d* orbital coupling by different Ln sites to achieve the thermodyanmaic stablization. Through the Fourier Transfer of the energy, we also notice the gradual change of the amplitude from 3d to Ln series (**Figure 3b**). For 3d, there are three evident sharp peaks at 0.03, 0.05, and 0.09, respectively. For 4d and 5d, the peaks still exist but become much weaker and blurred. For Ln-based DACs, we cannot identify any obvious peak in the spectra, which supports the weakened trend of specific stability preference in the pure TMs combinations. More importantly, the box plot also clearly reveals the energy change variations (**Figure 3c**). We notice a volcano trend for 3d, 4d, and 5d based DACs, where the peak always near the middle TM at V, Mo, and W for 3d, 4d and 5d based DACs. In comparison, the Ln-based DACs demonstrate averagely lower energies, supporting the higher synthesis possibility. Distinct from the volcano plots of TM based DAC, we notice several small peaks at Ce, Pm and Lu. The middle Ln based DACs construct a peak region from Gd to Dy. More importantly, due to the couplings between *d* and *f* orbitals, more abnormally high energies are noticed for Dy-Cd, Dy-Hf and W-Nd based DACs, implying more complicated factors in the orbital couplings. Overall, compared to the TM-TM based DACs, the Ln-TM based DACs offer a subtle energy variation trend, which supplies higher controllability during the experimental synthesis.

Considering the thermodynamic stability, there are several perspectives needed to be considered regarding the stabilization energies, anchoring energies, and the comparison with the homogeneous structures. From the comprehensive mapping of the formation energies, it is

further confirmed that the thermodynamic preference of the Ln-TM combined DACs (**Figure 4a**). Only a few pure TM based DACs show the negative formation energies to support an efficient synthesis. The stabilization energies refer to the energetic cost of two SACs to form the DAC (**Figure 4b**). Notably, compared to two individual SACs, most of the metal atoms prefer to form the DAC, except for the Ln-based DACs. These results that the mixture of metal precursors in experimental synthesis leads to a stronger preference to form the DAC at one appropriate anchoring site rather than forming isolated SAC in different anchoring sites. Although the Ln-based DACs all show lower formation energies than TM-based DACs, they still show less selectivity than the Ln-based SACs, especially for the light Ln elements. These results further confirm that high thermodynamic stability cannot guarantee the synthesis of DAC. The selectivity competition between the SAC and DAC is another pivotal factor. Meanwhile, the anchoring energies further confirm the synthesis difficulties of the DAC (**Figure 4c**). It is noted that most of the DACs show relatively high anchoring energy costs for the second metal atom. Only limited combinations show a spontaneous reaction to anchor the second metal atom on the formed SAC. Our results have explained the difficulties of the continuous anchoring method to synthesis the DAC as previous works [11b, 12c, 16]. Meanwhile, based on our results, we are able to screen out the promising combinations to achieve the stable DAC. The strong energetic contrast between the stabilization and anchoring energies indicates that the simultaneous anchoring of two different metal atoms is the optimal synthesis method. Instead, introducing the second metal atom in the formed SAC faces much higher energy costs. The homo-coupling DAC has already been successfully achieved in previous research[12a], which shows more facile synthesis control than the hetero-coupling DAC (**Figure 4d**). For both TM-TM and Ln-Ln DAC, the stabilized homo-coupling DACs are slightly more favored than the hetero-coupling DACs. Meanwhile, the hetero-coupling Ln-TM DACs are slightly more preferred than the homo-coupling TM-TM DACs due to the relatively lower energy cost. Thus,

during the synthesis of DAC by mixing Ln and TM precursor, Ln-TM hetero-coupling and Ln-Ln homo-coupling DACs are promising products since they have shown superior stability to both homo-and hetero- coupling TM-TM DACs due to the *f-d* coupling between TM and Ln atoms.

Beyond the thermodynamic stability, the electroactivity is carefully studied regarding the classic *d*-band center approach. As shown in **Figure 5**, we have demonstrated the *d*-band centers of all the DAC combinations. For the 3d-(3d, 4d, 5d) DACs, the *d*-band demonstrates a regular downshifting as the d orbitals become more filled. From 3d to 5d, the overall *d*-band centers of the 3d-(3d, 4d, 5d) DACs display a decreasing trend, especially for 3d-5d DACs with an evident drop of the *d*-band center towards -8.0 eV below the Fermi level (**Supplementary Figure S4-S9**). A similar decreasing trend is also preserved in the 4d-(3d, 4d, 5d) and 5d-(3d, 4d, 5d) DACs, especially at the fully filled d^{10} orbitals (**Supplementary Figure S10-S15**). However, such trends will be interrupted by the Ln, which is similar to energy mapping results in **Figure 4**. Distinct from the continuous downshifting trend, we notice the "hump" *d*-band centers with two peaks (**Supplementary Figure S16-S23**). Instead of the light or heavy Ln based DACs, the middle Ln based DACs demonstrate the valley-like *d*-band center trend, supporting a low electroactivity. More importantly, the introduction of Ln metal has evidently lifted the *d*-band center even for the fully occupied d^{10} element, which indicates the evident modulations by the *f*-electrons from Ln metals. In comparison, the Ln-Ln DACs with fully occupied *d*-orbitals demonstrates a high position of *d*-band, crossing the Fermi level. Compared to the large *d*-band variation in TM-based DACs, the combination of between Ln elements only induces the limited fluctuation of *d*-band change from E_V+0.5 to E_V+2.5 eV (We choose the valence band edge E_V

at 0 eV as the reference line). This limited fluctuations near E_V in Ln-TM DACs supply more appropriate electronic structures to facilitate the electron transfer in different reactions.

To better reveal the electroactivity change, we have mapped out the *d*-band center positions in all the DAC combinations to unravel the evolution of electronic environments. From **Figure 6a**, the regular change trend of the *d*-band center becomes more evident. In particular, due to the deep *d*-band center of Zn, Cd and Hg, their combination with other TMs still demonstrate an evident drop, indicating an inert electroactivity depending on the on-site electron transfer between *d*-*d* orbitals. In comparison, the evident increases of the *d*-band center in TM-Ln DACs are noted, in which the combination of middle TM metals and Ln metals are able to display an appropriate *d*-band center with high electroactivity within the range of E_V-2.0 eV to E_V. The overall stable *d*-band models indicate the electron migration (EM) induced by the Ln metals. Through the long-range site-to-site *d*-*f* electronic migration (SEM) model from *d*-orbitals of TM → *p*-orbitals of GDY →*f*-orbitals of introduced Ln, the self-electronic balance of DAC is preserved, leading to the evident alleviation of *d*-band center decreasing in TM-TM based DAC. Moreover, all Ln-Ln DACs show the *d*-band center over the Fermi level, representing the high valence states of Ln metals on the DAC. The involvement of Ln metals significantly uplifts the *d*-band center position, which is a good selection to flexibly modify the electroactivity of DACs. Meanwhile, we also compare the *d*-band center of the DAC and SAC (**Figure 6b**). Interestingly, the TM-TM DACs show a distinct *d*-band center with the SAC, indicating the coupling effect even between similar TM metal atoms. Since the *d*-band centers of SAC show the regular downshifting trends, the *d*-band center shifting induced by the introduction of hetero TM metals still preserves a similar downshifting trend. The combination with early TM will upshift the *d*-band center while the late TM will lower the *d*-band center, which supplying an explicit

direction for us to predict the electroactivity of DACs based on the change of *d*-band center in TM-TM DACs. This strategy is also applicable to TM-Ln DACs, which is able to balance the *d*-band center position to achieve the optimal electronic environment for electrocatalysis. It is noted that the *d*-band center change induced by the interactions of Ln metals are very weak. These results confirm that the *d-d* and *f-d* couplings effect between anchoring metal atoms are much more feasible than the *f-f* coupling to achieve the desired modulations.

In previous works, the discussion of the *d*-band center and their correlations with the electroactivity has been well interpreted[17]. In addition, the *p*-band has also been proposed to explain the intrinsic electroactivity of lattice oxygen in many oxide-based materials[18]. Although the high electroactivity of Ln materials has been often attributed to their unique electronic environments, the detailed involvements of 4*f* orbitals have rarely been discussed. More importantly, owing to the lack of atomic catalyst works on the Ln metals, the corresponding *f*-band center concept has never been demonstrated. However, since the *f-d* coupling has indicated an evident modulation to the electronic environment, we also need to further consider the *f*-band center of the Ln-based DACs (**Figure 7a**). The *f*-band has shown an evident valley trend, in which the middle Ln metals show the deepest positions of the *f*-band center. This indicates the high electroactivity of light and heavy Ln-Ln DACs for promoting electron transfer. It is worth mentioning that Lu-Ln DACs has shown an abnormally deep *f*-band center, which is attributed to the strong repulsive force between the fully filled *f*-shell and other *f* orbitals. For the *f*-band center change, the homogenous combination of Ln shows limited influence due to the limited *f-f* orbital couplings between the same Ln metals (**Figure 7b**). Since the *f*-band center of Ln also demonstrate the valley trend, the Tb- and Gd-based DACs have displayed the upshifting of the *f*-band center while the Lu-based DACs have shown a

downshifting trend. The *f*-band center does not follow the same variation trend as the *d*-band center, which has rarely been discussed in previous literatures. Most importantly, the *f*-band center of Ln-Ln DACs show relatively rare sensitivity to the electronic modulation by the second Ln metals, confirming the limited modulation between *f-f* orbital interactions. Through the difference between the *d*-band and *f*-band has been quantified (**Figure 7c**). The *f-d* orbital coupling in DAC activates the electron transfer behaviors from d-orbitals of transition metals to 4f-orbitals of Ln metals, which realizes the electronic self-balance effect to optimize the valence and structural stability. Based on our theoretical calculations, we notice that the *d*-band-center of TM atoms has been uplifted to a closer position to the Fermi level, which represents the higher electroactivity with an improved electron transfer. The electrons in the *d*-orbitals of TM atoms transfer to *4f* orbitals of Ln atoms, which improves both the electroactivity of DAC and the stability of TM metals on the surfaces. It is worth mentioning that the *f-d* couplings are more preferred in light Ln (Ce to Eu) and heavy Ln metals (Er to Yb) due to their higher position of 4f orbitals. In comparison, the middle Ln elements (Gd to Ho) show deep 4f orbitals, which cannot activate the efficient electron transfer from TM to Ln, displaying very limited modulations on the electronic structures. Therefore, the *f-d* orbital coupling is highly significant to achieve a stable and efficient DAC in the future experimental synthesis. With the detailed investigation of *f*-band center and *f-d* couplings, we not only explain the challenges of synthesizing DAC but also propose the promising candidates of the hetero-coupling DAC for future experiment targets. Combined with thermodynamic mapping, this work has supplied important references to the future rational design of DACs with desired electroactivity and durability.

By considering the current thermodynamic results, we have identified Ln-based DACs are more preferred than TM-TM based DACs, in which Sc-Pr, Y-Pr, Au-Sm, and Nd-Sm are the most stable DACs for Ln-3d, Ln-4d, Ln-5d, and Ln-Ln DAC, respectively. Meanwhile, some other combinations also show relatively low formation energies to support the promising synthesis in the future. However, the most stable structures usually have to sacrifice the electroactivity of the DAC. To screen out the most promising candidates, we have mainly compared the electronic structures of all DACs To achieve a high electroactivity for electron transfer, the Ln-Ln combinations are not suitable, which all show a d-band-center crossing the Fermi level. In comparison, the Ln-TM DACs show a more appropriate *d*-band-center for electrocatalysis. In particular, we have identified that Ln-Ni, Ln-Rh, Ln-Rh, Ln-Pd, Ln-Os, Ln-Ir and Ln-Pt DACs have shown the suitable d-band-center, which locate within 1.0 eV to the Fermi level, to guarantee different electrocatalysis processes. If we take into account the formation energies of these DACs, the Co-Nd, Ni-Er, Ru-Pr, Rh-Pr, Pd-Nd, Os-Pr, Ir-Pr and Pt-Ho will be the most potential candidates to achieve stable synthesis as well as high electroactivity in future experiments. We think that this work supplies valuable references for future experiments in the selection of optimal element combinations in DACs.

Although the theoretical calculations have explored both the thermodynamic stability and electronic structures of the GDY-based DACs, the intrinsic origins that result in the difference still need further investigation and study. As a powerful tool, the machine learning algorithm is able to predict the potential properties as the rational design guidance for future experimental synthesis. To verify the theoretical simulated results, we further introduce the machine learning method based on the Gaussian Process Regression (GPR) algorithm to verify the results. The GPR models have been designed to directly capture the model uncertainty by supplying a distribution for the prediction value rather than one simple value. Meanwhile, the GPR method

performs matrices inversions with different kernel functions to define the prior knowledge and specifications about the shape of the model, which often leads to the extremely large computational time when accessing large training datasets. Therefore, for our relatively small dataset (e.g. 990 combinations in this work), GPR is a powerful algorithm to identify significant information and accurate prediction with limited input knowledge. In this work, we have selected the Radial Basis Function (RBF) in the expression form as below.

$$\hat{f}(x) = \sum_{x}^{N} \omega_k G(x - x_k) \quad (1)$$

Eq. (1) represents the interpolation of function $f(x)$ with the known dataset x_k, where the function G is a radially symmetric function of its argument in the form of $G(r) \coloneqq \phi(|r|)$. The x is the vector of joint angles or other parameters describing the current pose of the skeleton and x_k is the pose of the k^{th} example. w_k represent the different weights of each vertex coefficient.

As the input data source, we have considered the fundamental physicochemical properties of the metal atoms to predict the formation energies of the GDY-based DACs. We have compared the prediction results of three types of different data input for the formation energies (**Figure 8a-c**). It is noted that most of the methods have all shown a high accuracy of the formation energies, where the similar root-mean-square error (RMSE) near 0.16. The introduction of additional parameters and the *d*-band does not further enhance the prediction, supporting that the thermodynamic stability of DAC is more related to the fundamental parameters of the anchoring metals. The electronic related properties show a limited contribution to the accuracy of predictions. Notably, even the abnormal formation energies have shown highly close predictions, supporting the pivotal role of the essential physicochemical properties in the thermodynamic stability (**Figure 8d-f**). Compared with the TM-TM based DACs, most of the deviation of prediction appear at the Ln-based DACs. Meanwhile, for the electronic structures,

the predictions show reduced accuracies, in which the prediction difference becomes more scattering with increased RMSE (**Figure 8g-i**). It is noted that the different parameter input indeed results in varied accuracy. However, more input parameters do not generate the highest accuracy. The method with 7 parameters input shows the smallest RMSE, which demonstrates the consideration of thermodynamic stability is not required for realizing high accuracy predictions. The decreases in input parameters evidently enlarge the deviation of predictions. The predicted data of the *d*-band center has shown the general deviation in both TM- and Ln-based DACs, which is different from the thermodynamic predictions (**Figure 8j-l**). These results support that the overall prediction of the *d*-band center still needs further optimization on the algorithm. The pure machine learning method based on the essential parameters cannot fully reveal the interactions between *d-d* and *f-d* orbitals, which have displayed a significant effect on the electronic environments. Therefore, these results demonstrate that the *d*-band centers in DACs are not completely essential properties that rely on the intrinsic properties of anchoring metals, in which some of the orbital couplings of metal/metal and metal/substrate inevitably affect the *d*-band center as well as the electroactivity.

The GPR based machine learning results agree well with our DFT models, indicating that thermodynamic properties are easily predicted while the electronic properties show much more complicated contributions from different factors. We conclude that the deviated prediction of the electronic structure is induced by the varied orbital interactions between dual metal sites, in which the self-balanced SEM model plays a significant role in achieving the optimal electronic environment of the DAC. The machine-learning based GPR technique have enabled the feasibility of designing complicated multi-atomic catalysts with large amounts of combinations. However, reliable theoretical calculation guided investigations are still needed, which

compensate for the insufficiency of machine learning-driven studies, especially in revealing the underlying interactions between the orbitals of the anchoring metals.

Conclusion

While atomic catalysts have existed in different forms and varied reactions, the fast developments of experimental and characterization methods have sparked the recent research interest in the further step of current SAC. Relying on the parallel theoretical calculations of both the DFT and machine learning techniques, we propose the Ln-TM DAC systems as the promising electrocatalyst candidates with expected high electroactivity and durability in the long-term applications. Such performances are contributed by the electronic self-balance mechanism between the long-range site-to-site *f-d* orbital interactions. Accordingly, the hetero-coupling TM-TM DACs are not the optimal solutions to reach both the thermodynamic stability and the electroactivity due to the *d-d* couplings. The GPR based prediction further confirm the underlying *f-d* orbital interactions while the thermodynamic stability is dominated by the intrinsic physicochemical properties. To surpass the limitation of the current SAC, our works provide invaluable guidance and explicit direction towards the GDY-based DACs with high experimental synthesis possibility and optimal electroactivity, which will also open up a new avenue to investigate multi-atomic catalyst systems in future research.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The authors gratefully acknowledge the support of the Natural Science Foundation of China (Grant No.: NSFC 21771156), and the Early Career Scheme (ECS) fund (Grant No.: PolyU 253026/16P) from the Research Grant Council (RGC) in Hong Kong.

Received: ((will be filled in by the editorial staff))
Revised: ((will be filled in by the editorial staff))
Published online: ((will be filled in by the editorial staff))

References

[1] a) Q. Fu, H. Saltsburg, M. Flytzani-Stephanopoulos, *Science* **2003**, *301*, 935-938; b) R. Bashyam, P. Zelenay, *Nature* **2006**, *443*, 63-66.

[2] B. Qiao, A. Wang, X. Yang, L. F. Allard, Z. Jiang, Y. Cui, J. Liu, J. Li, T. Zhang, *Nat. Chem.* **2011**, *3*, 634-641.

[3] a) P. Liu, Y. Zhao, R. Qin, S. Mo, G. Chen, L. Gu, D. M. Chevrier, P. Zhang, Q. Guo, D. Zang, B. Wu, G. Fu, N. Zheng, *Science* **2016**, *352*, 797-801; b) H. Wei, X. Liu, A. Wang, L. Zhang, B. Qiao, X. Yang, Y. Huang, S. Miao, J. Liu, T. Zhang, *Nat. Commun.* **2014**, *5*, 5634; c) E. Jung, H. Shin, B. H. Lee, V. Efremov, S. Lee, H. S. Lee, J. Kim, W. Hooch Antink, S. Park, K. S. Lee, S. P. Cho, J. S. Yoo, Y. E. Sung, T. Hyeon, *Nat. Mater.* **2020**, *19*, 436-442.

[4] a) C. Li, X. Zhao, A. Wang, G. W. Huber, T. Zhang, *Chem. Rev.* **2015**, *115*, 11559-11624; b) C. Wu, A. C. Wang, W. Ding, H. Guo, Z. L. Wang, *Adv. Energy Mater.* **2019**, *9*; c) Z. Li, Y. Chen, S. Ji, Y. Tang, W. Chen, A. Li, J. Zhao, Y. Xiong, Y. Wu, Y. Gong, T. Yao, W. Liu, L. Zheng, J. Dong, Y. Wang, Z. Zhuang, W. Xing, C. T. He, C. Peng, W. C. Cheong, Q. Li, M. Zhang, Z. Chen, N. Fu, X. Gao, W. Zhu, J. Wan, J. Zhang, L. Gu, S. Wei, P. Hu, J. Luo, J. Li, C. Chen, Q. Peng, X. Duan, Y. Huang, X. M. Chen, D. Wang, Y. Li, *Nat. Chem.* **2020**, *12*, 764-772.

[5] a) W. Yang, Z. Gao, X. Liu, X. Li, X. Ding, W. Yan, *Catal. Sci. Technol.* **2018**, *8*, 4159-4168; b) X. F. Yang, A. Wang, B. Qiao, J. Li, J. Liu, T. Zhang, *Acc. Chem. Res.* **2013**, *46*, 1740-1748; c) Y. Chen, S. Ji, C. Chen, Q. Peng, D. Wang, Y. Li, *Joule* **2018**, *2*, 1242-1264.

[6] a) S. K. Kaiser, Z. Chen, D. Faust Akl, S. Mitchell, J. Perez-Ramirez, *Chem. Rev.* **2020**; b) S. Ji, Y. Chen, X. Wang, Z. Zhang, D. Wang, Y. Li, *Chem. Rev.* **2020**; c) X. Zhang, S. Zhang, Y. Yang, L. Wang, Z. Mu, H. Zhu, X. Zhu, H. Xing, H. Xia, B. Huang, J. Li, S. Guo, E. Wang, *Adv. Mater.* **2020**, *32*, e1906905.

[7] a) J. Liu, X. Kong, L. Zheng, X. Guo, X. Liu, J. Shui, *ACS Nano* **2020**, *14*, 1093-1101; b) S. Ji, Y. Qu, T. Wang, Y. Chen, G. Wang, X. Li, J. Dong, Q. Chen, W. Zhang, Z. Zhang, S. Liang, R. Yu, Y. Wang, D. Wang, Y. Li, *Angew. Chem.* **2020**, *132*, 10738-10744.

[8] a) Y. Xue, B. Huang, Y. Yi, Y. Guo, Z. Zuo, Y. Li, Z. Jia, H. Liu, Y. Li, *Nat Commun* **2018**, *9*, 1460; b) N. J. O'Connor, A. S. M. Jonayat, M. J. Janik, T. P. Senftle, *Nat. Catal.* **2018**, *1*, 531-539.

[9] Y. Pan, C. Zhang, Z. Liu, C. Chen, Y. Li, *Matter* **2020**, *2*, 78-110.

[10] a) J. Zhao, J. Zhao, F. Li, Z. Chen, *The J.Phys. Chem. C* **2018**, *122*, 19712-19721; b) X. Zhang, A. Chen, Z. Zhang, Z. Zhou, *J. Mater. Chem. A* **2018**, *6*, 18599-18604.

[11] a) M. Xiao, H. Zhang, Y. Chen, J. Zhu, L. Gao, Z. Jin, J. Ge, Z. Jiang, S. Chen, C. Liu, W. Xing, *Nano Energy* **2018**, *46*, 396-403; b) J. Wang, Z. Huang, W. Liu, C. Chang, H. Tang, Z. Li, W. Chen, C. Jia, T. Yao, S. Wei, Y. Wu, Y. Li, *J. Am. Chem. Soc.* **2017**, *139*, 17281-17284; c) W.

Ye, S. Chen, Y. Lin, L. Yang, S. Chen, X. Zheng, Z. Qi, C. Wang, R. Long, M. Chen, J. Zhu, P. Gao, L. Song, J. Jiang, Y. Xiong, *Chem* **2019**, *5*, 2865-2878.

[12] a) S. Tian, Q. Fu, W. Chen, Q. Feng, Z. Chen, J. Zhang, W. C. Cheong, R. Yu, L. Gu, J. Dong, J. Luo, C. Chen, Q. Peng, C. Draxl, D. Wang, Y. Li, *Nat. Commun.* **2018**, *9*, 2353; b) Y. Yang, Y. Qian, H. Li, Z. Zhang, Y. Mu, D. Do, B. Zhou, J. Dong, W. Yan, Y. Qin, L. Fang, R. Feng, J. Zhou, P. Zhang, J. Dong, G. Yu, Y. Liu, X. Zhang, X. Fan, *Sci. Adv.* **2020**, *6*, eaba6586; c) L. Bai, C. S. Hsu, D. T. L. Alexander, H. M. Chen, X. Hu, *J. Am. Chem. Soc.* **2019**, *141*, 14190-14199.

[13] S. Wei, A. Li, J. C. Liu, Z. Li, W. Chen, Y. Gong, Q. Zhang, W. C. Cheong, Y. Wang, L. Zheng, H. Xiao, C. Chen, D. Wang, Q. Peng, L. Gu, X. Han, J. Li, Y. Li, *Nat. Nanotechnol.* **2018**, *13*, 856-861.

[14] a) M. Sun, T. Wu, Y. Xue, A. W. Dougherty, B. Huang, Y. Li, C.-H. Yan, *Nano Energy* **2019**, *62*, 754-763; b) M. Sun, A. W. Dougherty, B. Huang, Y. Li, C. H. Yan, *Adv. Energy Mater.* **2020**, *10*.

[15] a) H. Yu, Y. Xue, B. Huang, L. Hui, C. Zhang, Y. Fang, Y. Liu, Y. Zhao, Y. Li, H. Liu, Y. Li, *iScience* **2019**, *11*, 31-41; b) L. Hui, Y. Xue, H. Yu, Y. Liu, Y. Fang, C. Xing, B. Huang, Y. Li, *J. Am. Chem. Soc.* **2019**, *141*, 10677-10683.

[16] Y. Zhou, E. Song, W. Chen, C. U. Segre, J. Zhou, Y. C. Lin, C. Zhu, R. Ma, P. Liu, S. Chu, T. Thomas, M. Yang, Q. Liu, K. Suenaga, Z. Liu, J. Liu, J. Wang, *Adv. Mater.* **2020**, e2003484.

[17] a) J. K. Norskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard, H. Jonsson, *J. Phys. Chem. B* **2004**, *108*, 17886-17892; b) J. Greeley, I. E. Stephens, A. S. Bondarenko, T. P. Johansson, H. A. Hansen, T. F. Jaramillo, J. Rossmeisl, I. Chorkendorff, J. K. Norskov, *Nat. Chem.* **2009**, *1*, 552-556; c) T. Wu, M. Sun, B. Huang, *Small* **2020**, *16*, e2002434.

[18] a) J. Suntivich, K. J. May, H. A. Gasteiger, J. B. Goodenough, Y. Shao-Horn, *Science* **2011**, *334*, 1383-1385; b) J. Suntivich, H. A. Gasteiger, N. Yabuuchi, H. Nakanishi, J. B. Goodenough, Y. Shao-Horn, *Nat. Chem.* **2011**, *3*, 546-550.

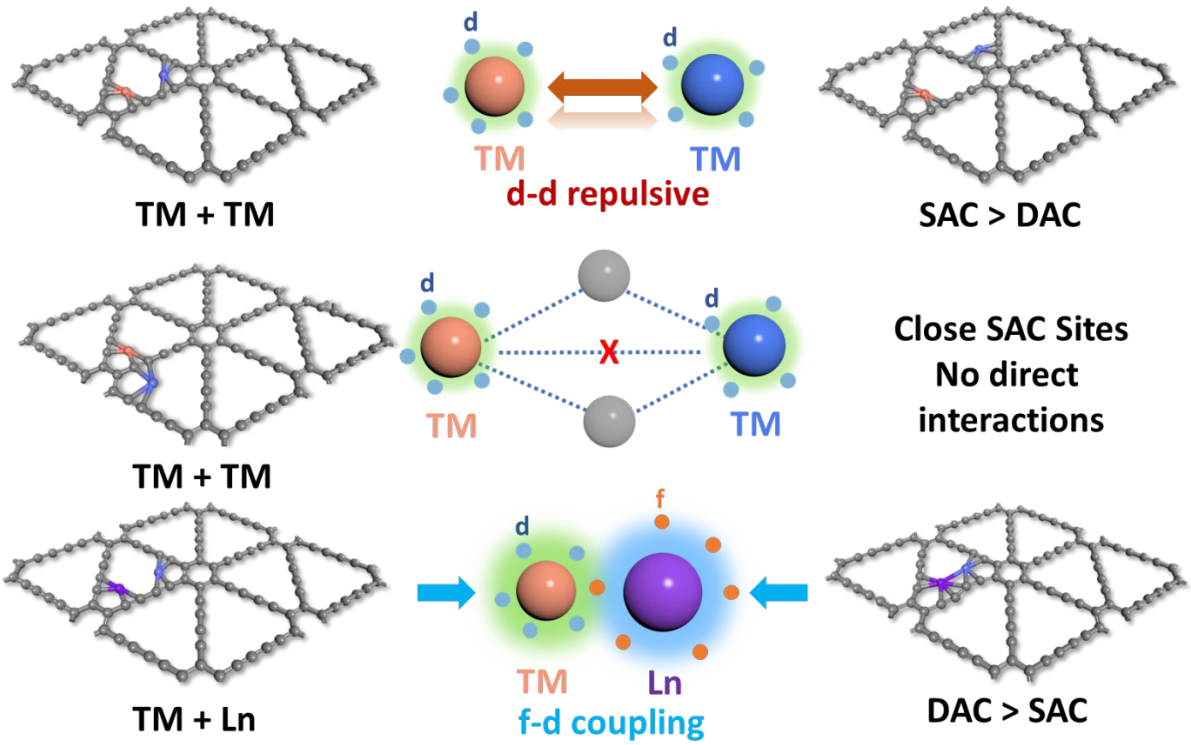

Figure 1. The schematic diagram of developments of the DAC. The different combinations of metals lead to varied interactions and synthesis products.

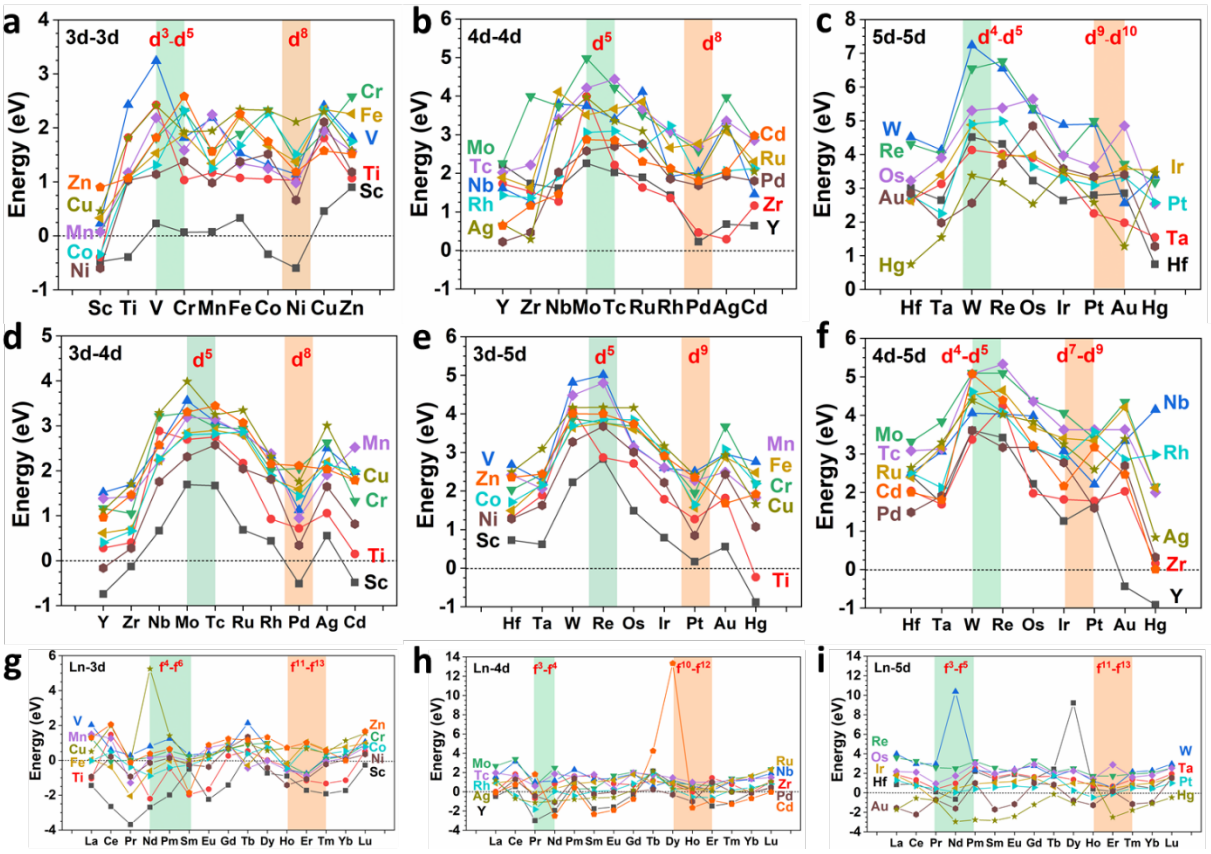

Figure 2. The formation energies of GDY based DACs between **(a)** 3d-3d, **(b)** 4d-4d and **(c)** 5d-5d TM combinations. The formation energies of GDY based DACs by different TM combinations of **(d)** 3d-4d, **(e)** 3d-5d, and **(f)** 4d-5d. The formation energies of GDY based DACs by TM and Ln metal combinations of **(g)** Ln-3d, **(h)** Ln-4d, and **(i)** Ln-5d.

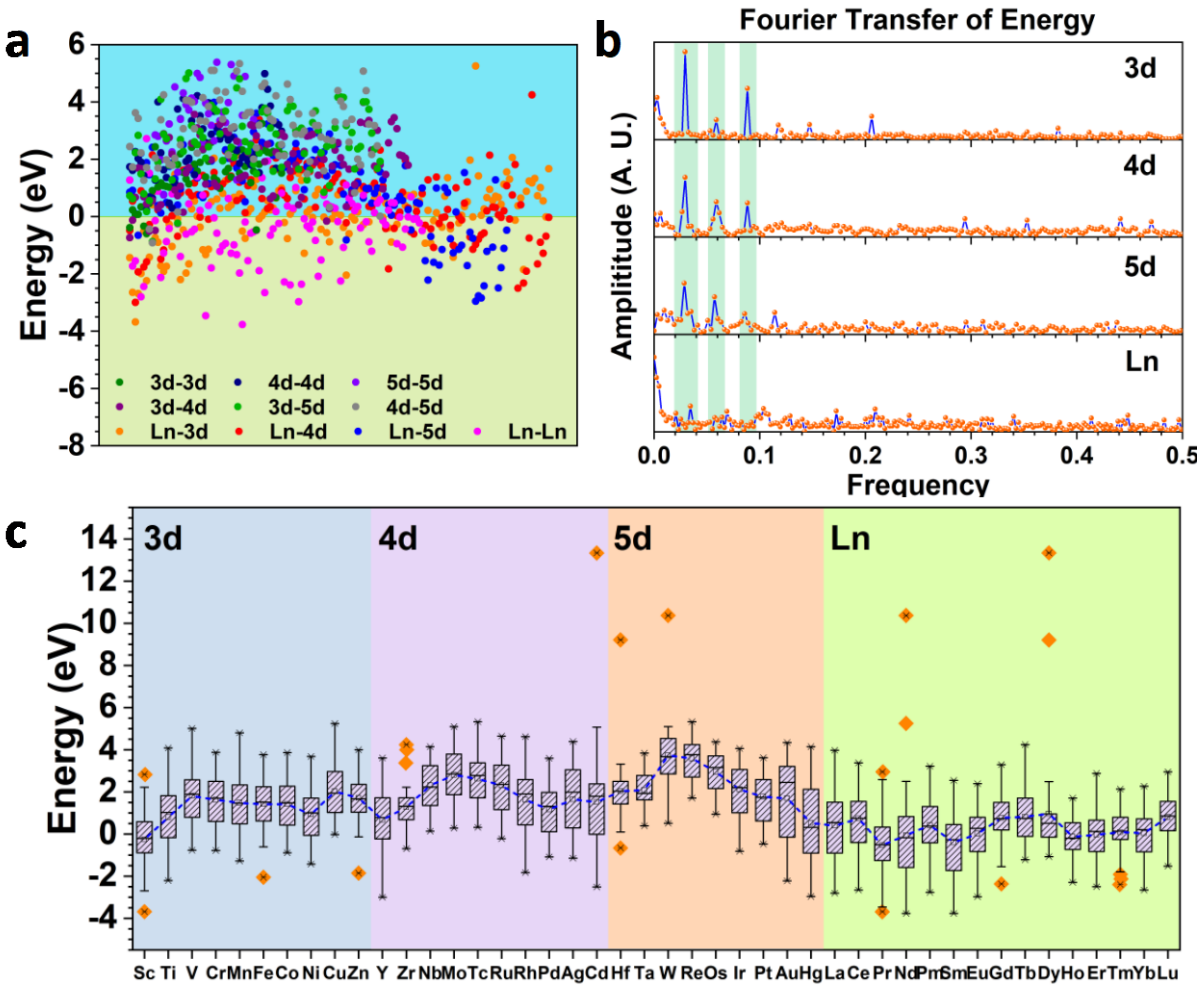

Figure 3. (a) The formation energy demonstration of all the GDY-based DACs combinations, including 990 models. Apparently, the combination of TM and Ln significantly lowers the formation energies. **(b)** The Fourier transfer of formation energies of different groups. Noted, the evident peak at 0.03, 0.05, and 0.09 gradually disappear from 3d-based DACs to Ln-based DACs. **(c)** The box plots of metal combinations of each TM and Ln, which mainly indicates the distribution range of the formation energies and their variation range.

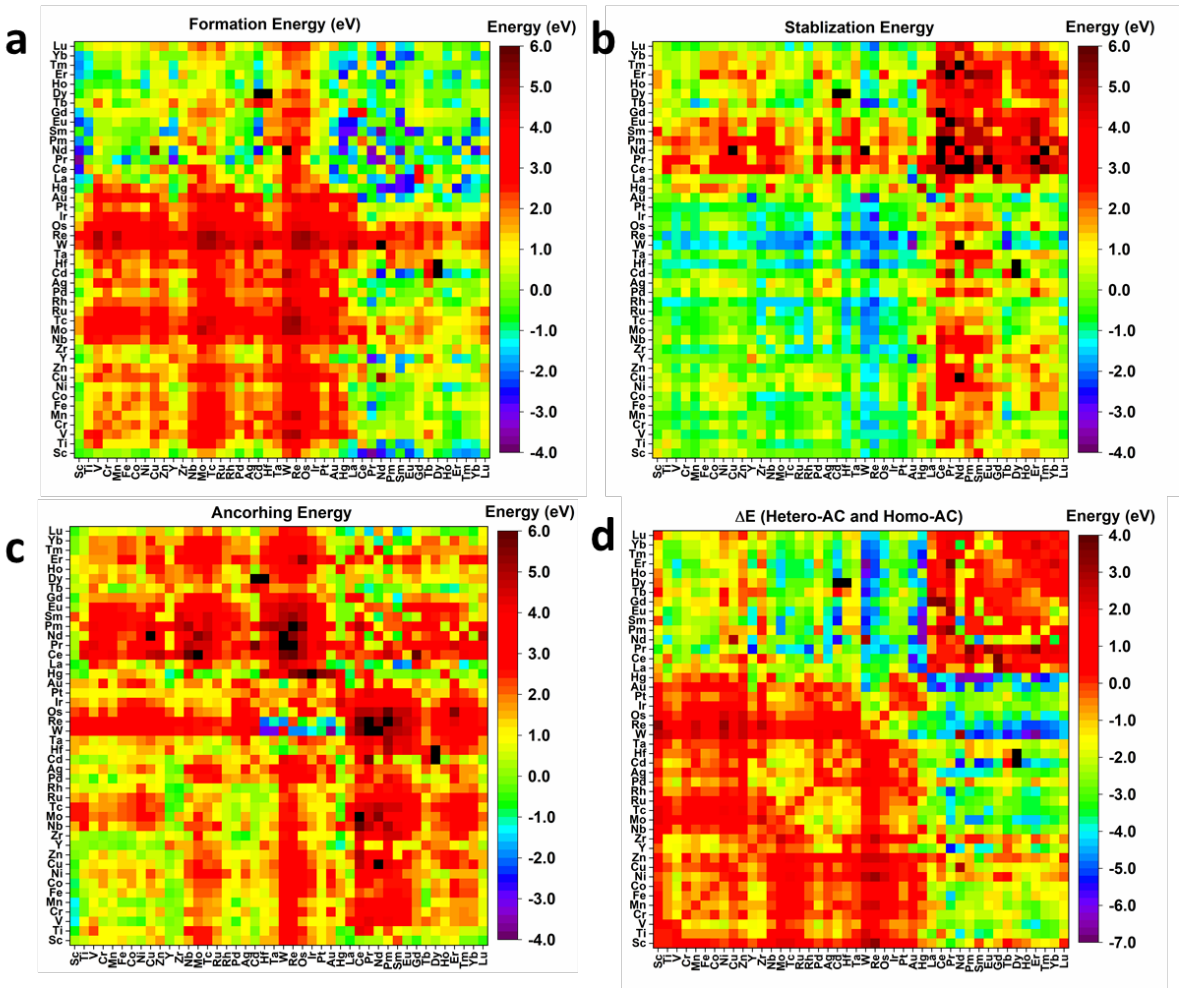

Figure 4. **(a)** The formation energy mappings of all the GDY-based DAC combinations, including 990 models. **(b)** The stabilization energy mappings of all the GDY-based DAC combinations. The stabilization energies are calculated by the energy comparison between the DAC and two individual SACs. **(c)** The anchoring energy mappings of all the GDY-based DACs combinations, which is equal to the energy requirement of anchoring the second metal. **(d)** The mappings of the energy difference between the hetero-coupling DACs and homo-coupling DACs, indicating the TM-TM combinations are highly unstable for experimental synthesis.

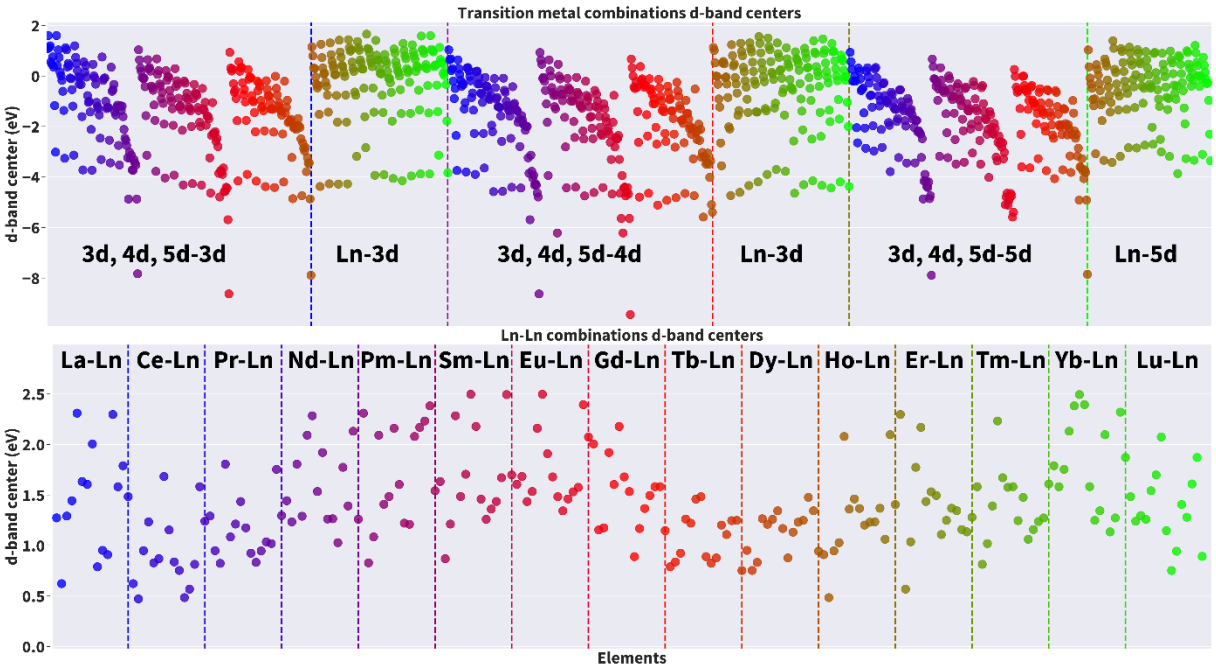

Figure 5. The d-band center mapping of all GDY based DACs. The upper figure demonstrates the d-band center variation trend including 3d-(3d, 4d, 5d), 3d-Ln, 4d-(3d, 4d, 5d), 4d-Ln, 5d-(3d, 4d, 5d) and 5d-Ln DACs. The bottom figure demonstrates the d-band center variation of Ln-Ln DACs. Based on the different combinations between Ln metals, the figure has been separated into 14 different groups.

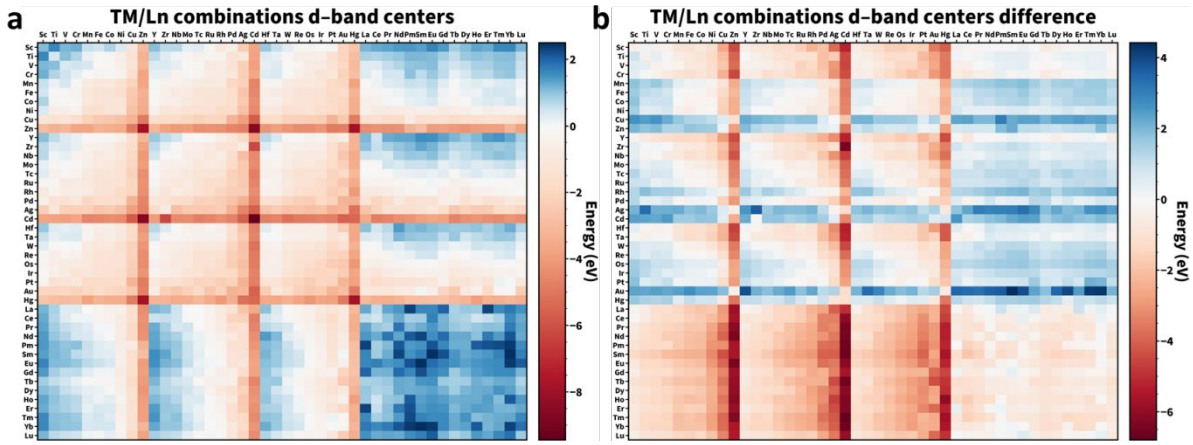

Figure 6. (a) The d-band mapping for all the combinations of GDY-based DACs. **(b)** The mapping of d-band difference for all the combinations of GDY-based DACs. Each data is obtained by the d-band center comparison between the DAC and the SAC of the left column metals.

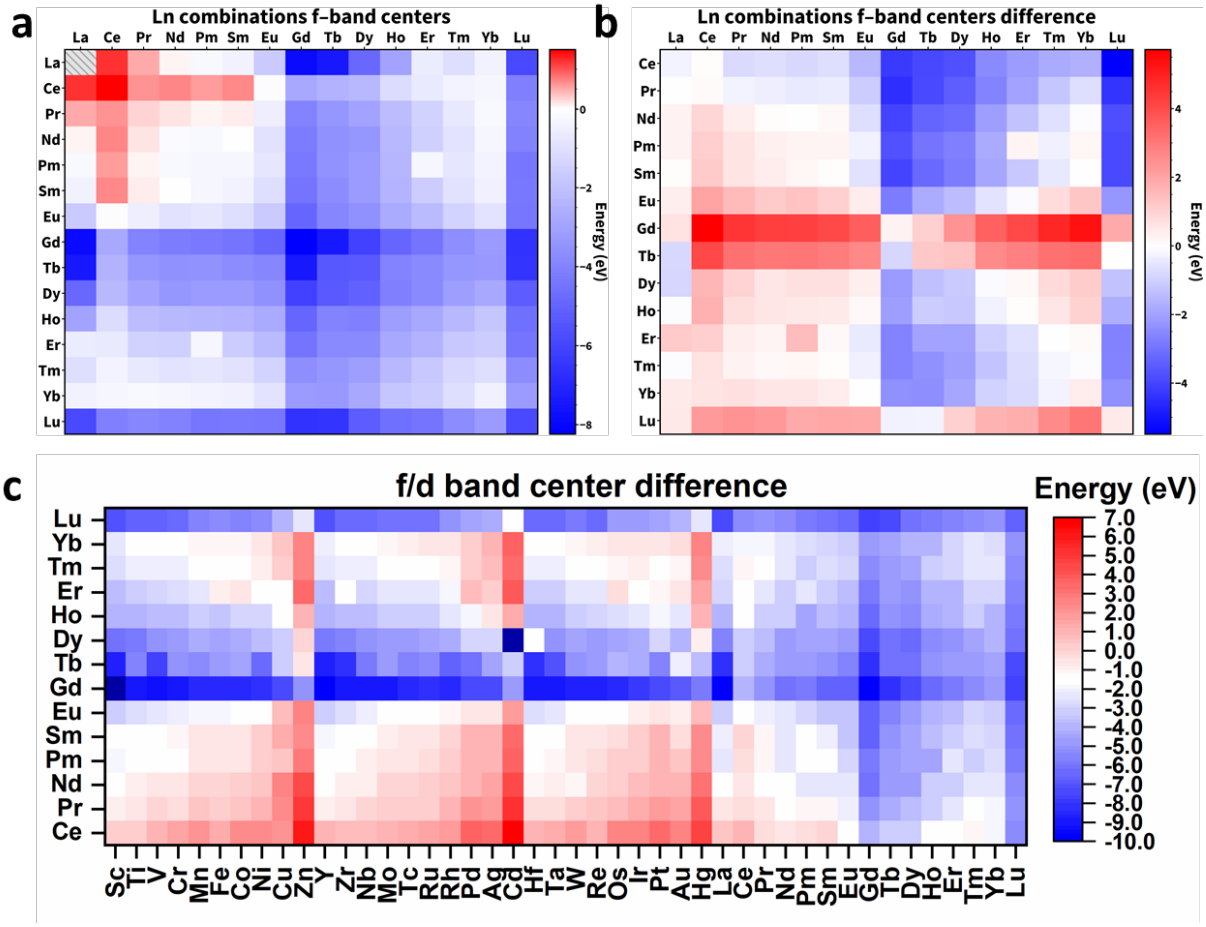

Figure 7. (a) The *f*-band mapping for all the Ln based DACs. **(b)** The mapping of *f*-band difference for all the Ln based DACs. Each data is obtained by the *f*-band center comparison between the DAC and the SAC of the left column metals. **(c)** The mapping of *f/d* band center difference. The data is calculated by the position difference of *f*-band center and *d*-band center in the DAC.

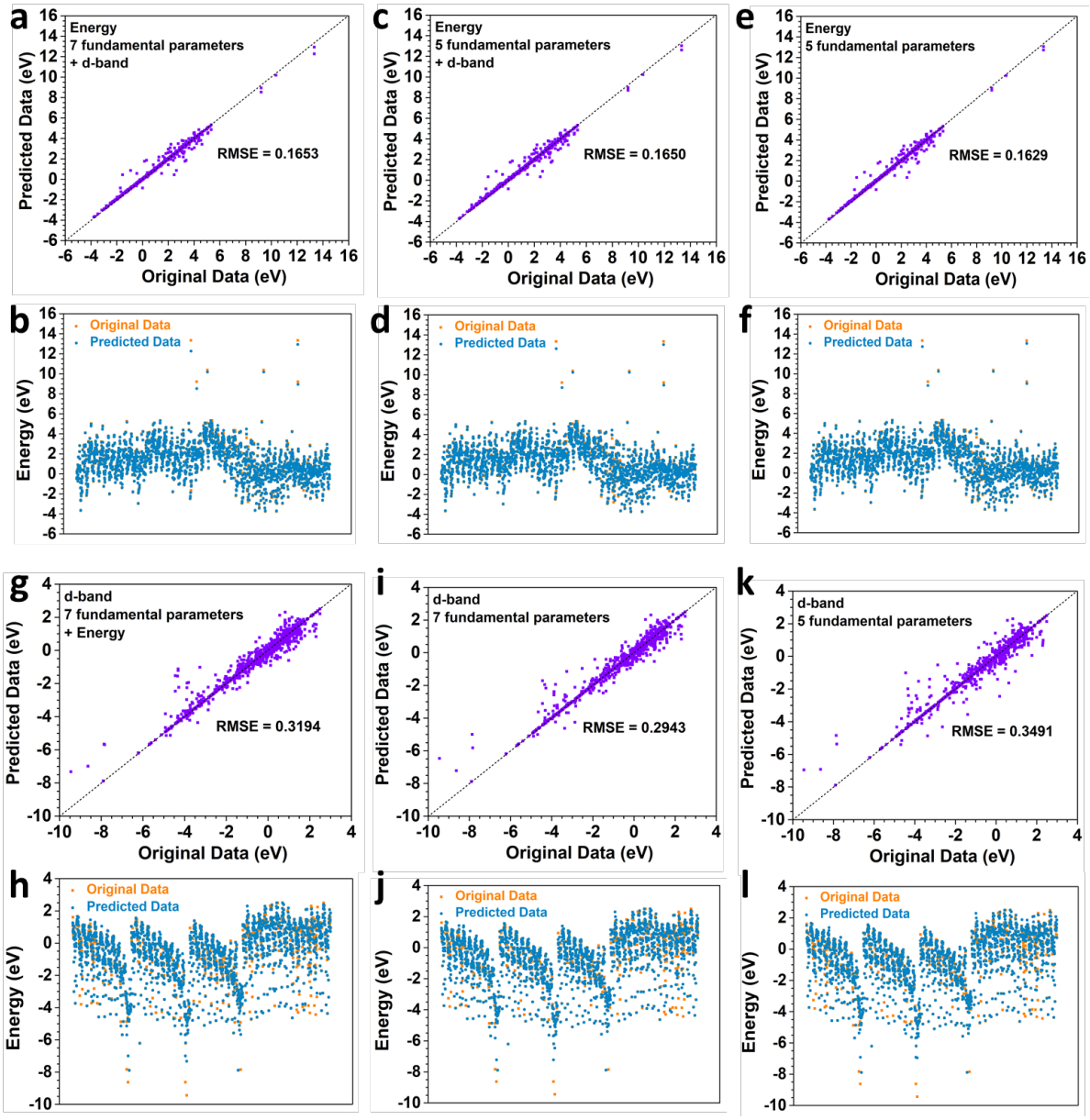

Figure 8. The comparison and scattering plot of GPR model predicted data and original data of the formation energy of **Figure 3a**. **(a,b)** Prediction based on 7 intrinsic parameters and d-band center. **(c,d)** Prediction based on 5 intrinsic parameters and d-band center. **(e,f)** Prediction based on 5 intrinsic parameters. The energy is mostly based on the intrinsic parameters of the material. Too many parameters and d-band cannot improve prediction accuracy. The comparison and scattering plot of GPR model predicted data and original data of d-band in **Figure 5**. **(g,h)**

Prediction based on 7 intrinsic parameters and energy. **(i,j)** Prediction based on 7 intrinsic parameters. **(e,f)** Prediction based on 5 intrinsic parameters.

The explorations of dual atomic catalysts are emerging in the electrocatalysts. This work has mapped out all the combinations between transition and lanthanide metals regarding both stability and electroactivity. The introduction of machine learning technique has further confirmed the complicated interactions between anchoring metals and graphdiyne support. This work provides significant references for the future synthesis of dual atomic catalysts.

Keyword: Electrcatalysis

M. Sun, T. Wu, A. W. Dougherty, M. Lam, B. Huang, Y. Li, C.-H. Yan*

Self-validated Machine Learning Study of Graphdiyne-based Dual Atomic Catalyst