CCS Chemistry

RESEARCH ARTICLE

Received: Sept. 3, 2024 | Accepted: Nov. 20, 2024 | Published: Dec. 16, 2024

# Automating Discovery of Electrochemical Urea Synthesis Reaction Paths via Active Learning and Graph Theory

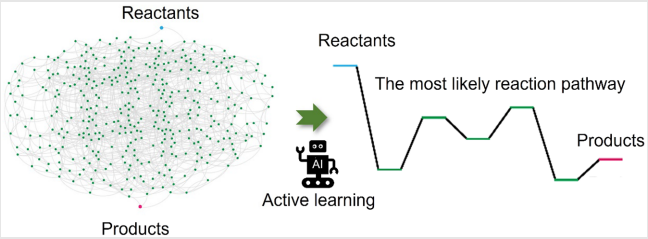
Shisheng Zheng[1,2], Haowen Ding[1], Xinzhe Yang[1], Shunning Li[1]* & Feng Pan[1]*

[1]*School of Advanced Materials, Peking University Shenzhen Graduate School, Shenzhen 518000,* [2]*College of Energy, Xiamen University, Xiamen 361000*

**Corresponding authors: lisn@pku.edu.cn; panfeng@pkusz.edu.cn*

**Cite this:** *CCS Chem.* **2025**, 7, 2822–2834
**DOI:** 10.31635/ccschem.024.202404955

**The exploration of the most probable reaction path is a central task in electrocatalysis, often limited by the computational burdens when evaluating via density functional theory (DFT) with growing reaction network complexity. Various machine learning algorithms have been developed to minimize computational costs and have made progress, but most of them have been demonstrated on metals and alloys with static, intact surfaces. The dynamic effect of catalyst surface reconstruction, playing an essential role in the electrocatalysis process, has been commonly overlooked in current surrogate models. Here, we introduce a workflow combining graph theory and active learning loops to address this long-standing challenge. Our streamlined workflow covers most, if not all, of the bond rearrangement types in the electrochemical reactions, and the computational costs for DFT are reduced by a pipeline successively performing the stability and formation energy prediction of intermediates. We demonstrate its efficiency in the context of urea electrosynthesis on a nitrogen-doped graphene with substantial dynamic structural reconstructions of the catalytic center. This framework can be extended to other complex electrochemical reactions and facilitates the rapid estimation of overpotential with minimal reliance on precise quantum chemical calculations, paving the way for automated computational analysis of catalytic mechanisms under realistic conditions.**




*Keywords:* **active learning, graph theory, urea synthesis, density functional theory, nitrogen-doped graphene, reaction network**

## Introduction

Electrochemical reactions that involve multiple electron transfers are generally characteristic of a vast reaction network that encompasses a considerable number of elementary steps.[1–3] Early studies primarily relied on heuristic rules to explore the most probable reaction path from the network, which could lead to a subjective, rather than objective, understanding of the underlying reaction mechanisms.[4] To enable more stringent determination of the reaction path, computational analysis via quantum chemical methods has experienced a booming development in the past decade as a means to deliver valuable details on the mechanisms for complex electrochemical

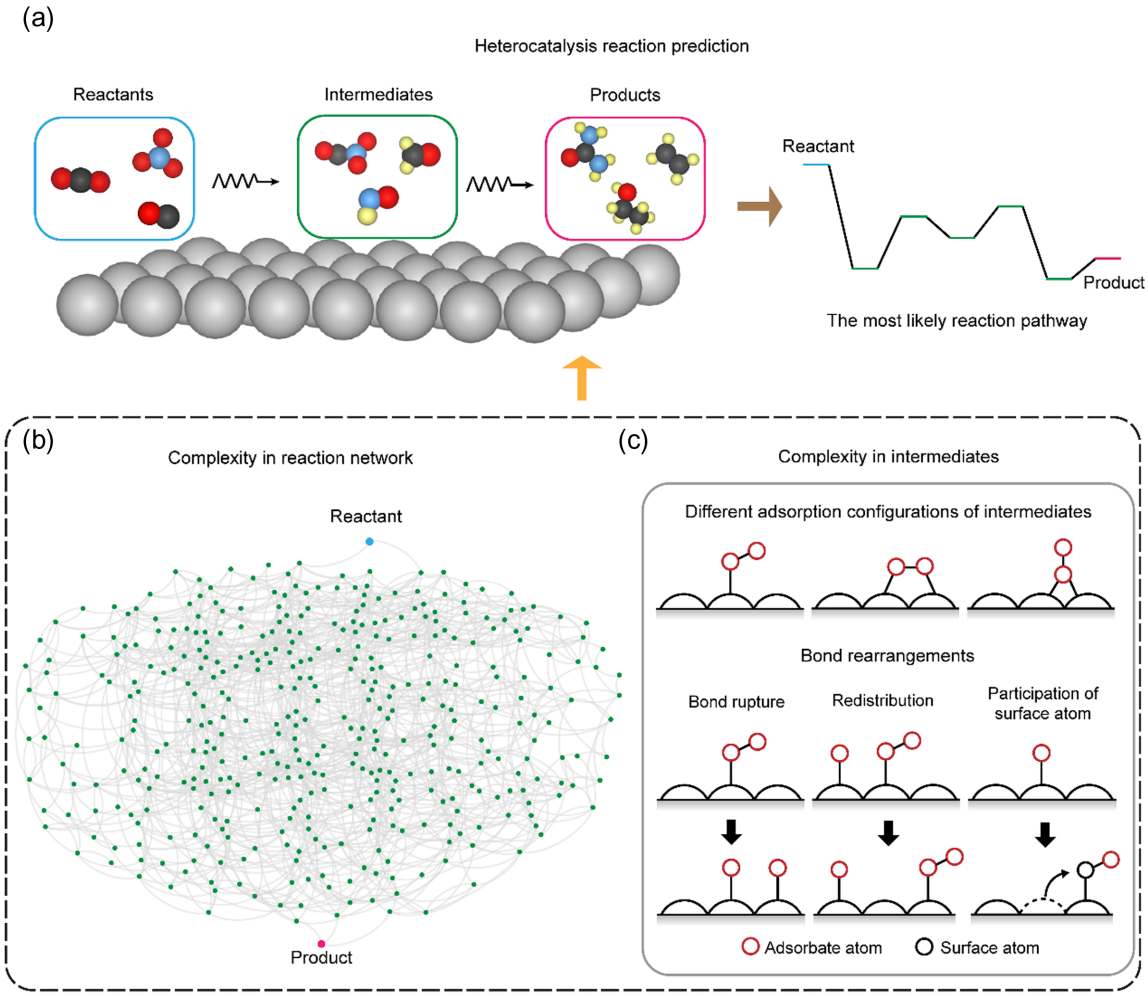



**Figure 1** | *Complex reaction network in electrocatalysis. (a) Free energy profile for the most likely reaction pathway indicated by DFT calculations. (b) A complex reaction network encompassing multiple reaction pathways involving hundreds of intermediates. (c) Different adsorption configurations and possible bond rearrangements on the catalyst surface.*

reactions (Figure 1a).[5–15] However, for reactions that employ catalysts with nonuniform structures or compositions, the diversity of reaction intermediates and possible catalytic active sites will result in a combinatorial explosion that makes conventional density functional theory (DFT) calculations of the whole reaction network intractably expensive.[16] This problem has been encountered in recent theoretical studies of $CO_2$/CO electroreduction to multicarbon products, and electrosynthesis of amides and their derivatives.[2,17–30] Even worse, the wide range of adsorption motifs for the relevant reaction intermediates (including mono-, bi-, and higher-dentate adsorption modes) on the catalyst surface will further exacerbate the challenge of computational analysis[31–35] (Figure 1b,c), leaving the comprehensive exploration of the configurational space for reaction intermediates a nearly hopeless task.

Lately, a series of theoretical studies have resorted to surrogate models using machine learning (ML) techniques to overcome the above difficulty.[4,16,36–39] One of the strategies is to replace DFT with ML force fields in structural optimization calculations, which can speed up calculations that take into account diverse chemical environments for the complicated adsorbate-catalyst interactions.[40–42] However, the construction of ML force fields often needs immense DFT training data, and the capacity for interatomic potentials to extrapolate beyond their training data remains questionable, particularly for complex systems.[16] When moving to a higher level of abstraction, another popular strategy is to directly predict the adsorption energies of reaction intermediates on different catalyst facets using ML models trained on physics-inspired descriptors.[43–45] For example, the pioneering work by Nørskov et al.[6] has demonstrated the efficiency of ML models in providing the estimates of transition-state free energies and determining the most likely reaction pathway for syngas conversion on metal surfaces. Despite the success of these previous

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

studies on prototypical catalytic reactions, mostly utilizing metal or alloy catalysts with their surfaces remaining intact in the reaction,[6,46,47] it is noted that the dynamic effect of catalyst surface reconstruction has been nearly completely overlooked. The surface dynamics are particularly important when investigating electrochemical reactions involving large-size molecules,[48–50] high-coverage adsorbates,[51–53] and reactive surfaces (such as metal oxides and low-dimensional materials).[21,28,29,54,55] Atoms at the catalyst surfaces can participate in the reactions, widely known as the Mars-van Krevelen (MvK) mechanism,[42,56–60] in addition to various possible changes in the adsorbate configurations, such as bond rupture and site redistribution (Figure 1c). These conditions make it unfeasible to accurately predict the free energy of the adsorption configurations using features extracted from the unrelaxed structures, and therefore false identification of reaction pathways would naturally be expected from the previous ML models.

Herein, we introduce a workflow combining graph theory and active learning loops to address this long-standing challenge and demonstrate its efficiency in the context of urea electrosynthesis on a nitrogen-doped carbon material. The complexity of this reaction manifests in the dynamic structural reconstruction of the catalytic center, wherein the intrinsic surface hydrogen atoms will participate in the electrochemical reduction reaction, as justified by our previous experimental study.[18] Our streamlined workflow covers most, if not all, of the bond rearrangement types in the electrochemical reactions, and the computational cost for identifying the most probable reaction path can be considerably reduced as compared to conventional high-throughput DFT calculations. This method can be generalized to other complex electrochemical reactions and will enable the rapid estimate of overpotential using a small fraction of accurate quantum chemical calculations, which could pave the way for the realization of automatic computational analysis of catalytic mechanisms under realistic conditions.

# Computational Methods

## DFT calculation methods

All the calculations were performed by means of spin-polarized DFT using Vienna ab initio simulation pack (VASP) with projector augmented wave method.[61] A plane-wave cutoff energy of 400 eV was adopted and the Perdew–Burke–Ernzerhof exchange-correlation functional was employed. To complement the deficiencies of DFT in dealing with dispersion interactions, DFT-D3[62] semiempirical van der Waals corrections were included. Supercells consisting of 5 × 5 unit cells of graphene were constructed for the nitrogen-doped graphene catalysts. A vacuum space of at least 20 Å was used to separate the slabs in the *z* direction. The Brillouin zone was sampled by Γ-centered Monkhorst–Pack *k*-point mesh of 2 × 2 × 1 for free energy calculations. To take into account the contribution of solvent effect, we introduce the VASPsol[63] module during the energy calculation. The Gibbs free energy in every elementary step was calculated by the computational hydrogen electrode model.[64] More details can be found in Supporting Information.

## The species generation

The RDKit,[65] which is an open-source tool for cheminformatics studies based on molecular graph, is used for the species generation in the classical mechanism. The 2*NO and $CO_2$ are considered as the reactants to generate the possible spices involved in the classical mechanism. The species involved in the reaction are generated by applying defined elementary step types to the reactants which are converted to molecular graph. Simultaneously, the new species generated in this process would be verified whether they meet the boundary conditions since we focus on the double-ended mechanism search. These boundary conditions can ensure reasonable bond principles and will guarantee that the species is a precursor of urea. The valid species are recorded and new species would be further generated iteratively until the urea is found. For the MvK mechanism, the nitrogen source originates from the substrate, thus involving only two carbon-containing species: *COOH and *CO. More details and workflow can be found in Supporting Information.

## The adsorption structure enumeration

The species should be associated with the catalysis surface to construct the geometric adsorption structures, which is a necessary step for the investigation of reaction mechanism by DFT. There are various possible adsorption types of intermediates on the surface, among which the formation energy may show a large gap. We developed an automated algorithm to enumerate the adsorption configurations based on graph theory. Firstly, we convert the species recorded in the simplified molecular input line entry system (SMILES) representation to molecular graphs and enumerate possible interaction modes with the surface in the graph with user-defined rules. Subsequently, we identify the active sites of catalysis surface that adsorb the species, and finally, the positions of atoms of species that interact with the active sites are determined. The positions of the left part of the species are determined to generate the whole adsorption structures based on the connectivity of molecular graphs and geometric rules. This algorithm can automatically enumerate the possible adsorption structures and can easily integrate chemical principles to balance the enumeration scales and computational costs. After the construction of adsorption structures of one species, graph isomorphism algorithm is used to eliminate equivalent

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

structures due to the discrimination of the same element atoms. The eight-electron rule is also applied to remove the chemical unfavorable structures. For the classical mechanism, the total adsorption configuration is 1256 due to the species having many possible atoms that can interact with surface. For the MvK mechanism, it requires passivation of the carbon atoms near the nitrogen atoms to facilitate the break of the carbon-nitrogen bonds. The two nitrogen atoms and nearby five carbon atoms are considered to be active sites and each active site will be hydrogenated once. In this regard, we construct a set of 478 possible intermediates, and each intermediate has only one adsorption configuration. The details are presented in Supporting Information Figures S3–S7.

### The construction of reaction network

With all species and adsorption structures generated, the reaction network can be easily constructed by the graph isomorphism algorithm. The elementary reaction step can be considered as an operation to the graph of the species. Depending on different reaction types, the graph of new species would be a subgraph of the graph of the old species, or vice versa. In this approach, all elementary reaction steps can be determined, and then the construction of reaction network can be easily realized. The details are presented in Supporting Information Figure S8.

### The machine learning methodology

The ML methodology includes two parts: the gradient boosting classification to predict the intermediate stability and the ridge regression to predict the formation energy of chemical stable intermediate. They both use the subgraph fingerprints as presented in the Supporting Information Figure S9. Generally, the adsorbate, $N_4$ moiety, and surrounding carbon atoms are used to model the microenvironment of the nitrogen-doped graphene surface. Subsequently, the microenvironment is decomposed into a number of fragments according to the first-neighbor atoms of each central atom. Thus, each fragment represents the local environment of each atom. The output feature vector contains the number of occurrences for each fragment and is used to correlate the stability or formation energy of intermediates. Both the gradient boosting classification and the ridge regression model are constructed by the *sklearn* module in python with default parameters. More details are presented in Supporting Information.

## Results and Discussion

### Workflow overview

The overall architecture of the proposed framework is shown in Figure 2. The reaction network connecting the

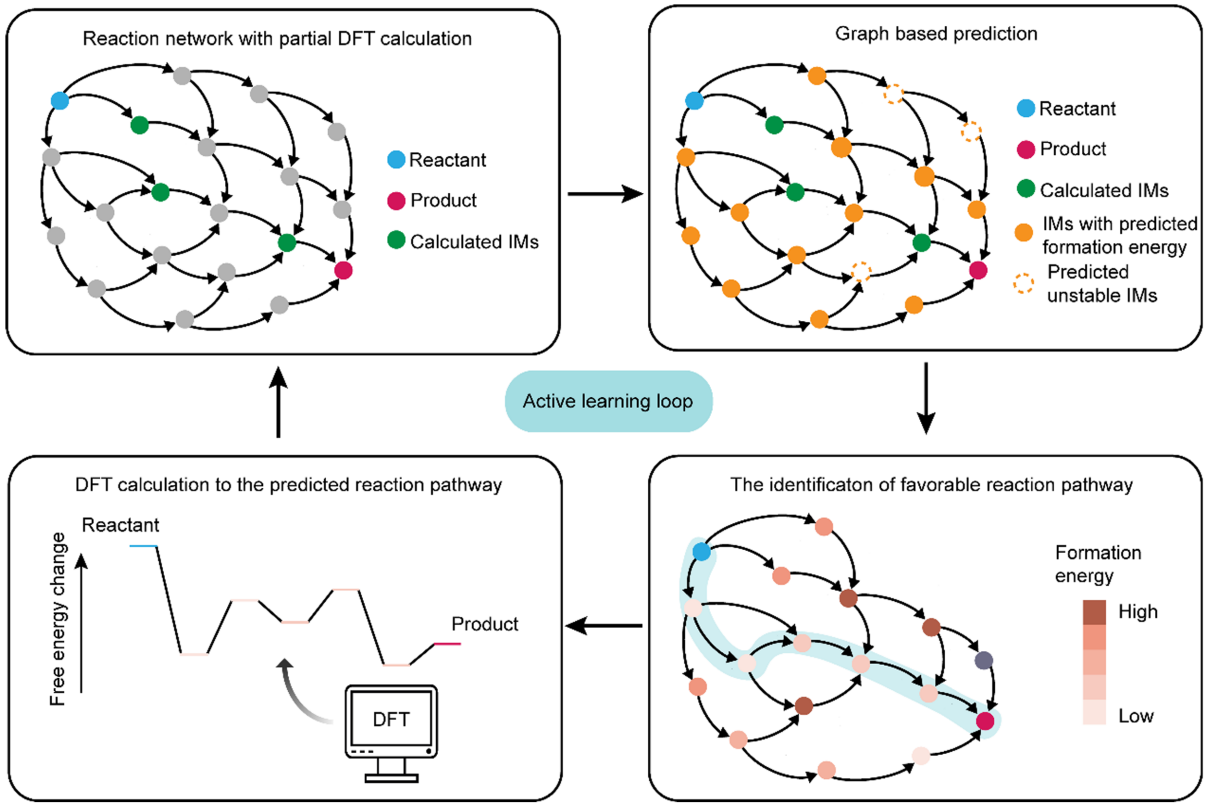



**Figure 2** | *The graph-based ML framework to rapidly predict the most likely reaction pathway from a complex reaction network.*

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

reactant and product is firstly constructed by enumerating all the adsorption configurations of each intermediate. Subsequently, structural optimizations by DFT are performed for a subset of the intermediates, and their stabilities are evaluated by comparing the initial and optimized structures using their structure graphs. An ML classification model is then employed to predict the stability of other intermediates that have not been included for DFT calculations. Those that are classified as unstable will be discarded in the reaction network, while the stable ones will be fed into a ridge regression model to estimate the formation energies. The optimal reaction pathway at this stage can be determined according to the predicted largest free energy change in each pathway, which corresponds to the limiting potential in this path. The one with the lowest limiting potential (when the limiting potential is identical, the lowest sum of uphill energy changes in the energy profile will be adopted) can be regarded as optimal. The above procedure will be recursively conducted by performing DFT calculations for the predicted optimal reaction pathway and implementing the corresponding data to the training set of the ML model. After several iteration, the most favorable reaction pathway will be obtained, which avoids the need of high-throughput DFT calculations for all the intermediates.

## Generation of reaction network

The reaction network for urea electrosynthesis using $NO_3^-$ and $CO_2$ as the reactants involves several hundreds of reaction intermediates, which, when in combination with a catalyst that undergoes reconstruction during the reaction, will substantially hinder calculations at the DFT level for identifying the most likely reaction pathway. Nitrogen-doped carbon (Supporting Information Figure S1) is one such example,[18] showing high catalytic performance on C–N coupling, which is a critical step for urea synthesis. It has been recently justified that the number of N–H bonds at the active sites is switchable during the reaction, which may steer the reaction between $NO_3^-$ reduction and $CO_2$ reduction towards the formation of urea. The dynamic evolution of the catalytic centers cannot be left out of consideration in this system.

To fully elucidate the reaction pathway for this complex catalytic reaction with minimal DFT calculations, we employed the graph-based ML framework as described above. The first step is to enumerate the intermediate species connecting the reactants and the products, and two different kinds of reaction mechanisms are evaluated here (Figure 3a). In the classical mechanism, the N atoms at the catalyst surface remain intact during reaction, while in the MvK mechanism, the surface N atoms could participate in forming urea and be replenished by the $NO_3^-$ feedstock in the electrolyte. The MvK mechanism on nitrogen-doped carbon catalysts has been experimentally confirmed in the ammonia electrosynthesis reaction.[66] For both mechanisms, all the possible reaction intermediates are recursively generated using a molecule-graph–based automated enumeration algorithm with RDKit (Supporting Information Figure S2). The graph-based notation of molecules is conducive to the computer recognition and operation due to the simple transformation of graphs into SMILES strings or matrix representations. Each elementary step corresponds to an edit of the molecule graph, including the addition and subtraction of the nodes, which represent the hydrogenation steps, the C–N coupling steps, and dihydroxylation steps. The graph edit operations will proceed until the urea product is obtained, and the structural information of bond changes between reactants and urea product could be captured by this series of edit operations. We note that some thresholds have been set to identify the cases where the principles of molecular bonding have been violated (details in computational methods).

For each reaction intermediate, we have enumerated all the possible configurations that conform to the geometric rules of organic molecules (details in computational methods), such as the rule of a maximum of four atoms connected to the N or C atom. The obtained reaction network is displayed in Figure 3b, which highlights the different intermediates corresponding to the elementary steps of the two reaction mechanisms. Notably, in the MvK mechanism, the $CO_2$ reactant will be transformed into *CO adsorbate after two-electron reduction steps, and *CO will then interact with the N atoms at the catalytic center to form the N–C–N backbone of urea. Rearrangement of surface H atoms is indispensable during this process, which can be directly reproduced in our graph-based algorithm. Overall, the reaction network consists of 901 intermediates and 1734 adsorption configurations (Supporting Information Table S1), implying a significant computational burden if conventional high-throughput DFT calculations are performed for this network.

## Bond rearrangement of adsorption configurations

The stability of an adsorption configuration is evaluated by whether significant bond rearrangement has taken place during its structural optimization in DFT calculations. If there is a change in the connectivity of adsorbate or catalyst surface, this configuration is defined as unstable. The graph isomorphism algorithm is employed to accomplish this task (Supporting Information Figure S10), which is powerful for structure identification as demonstrated in the previous studies.[31,67–70] The stable and unstable groups of intermediates can be assigned from the DFT results, and they are used as the training set for a gradient boosting classification model to determine the stability of other intermediates, donated as the

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

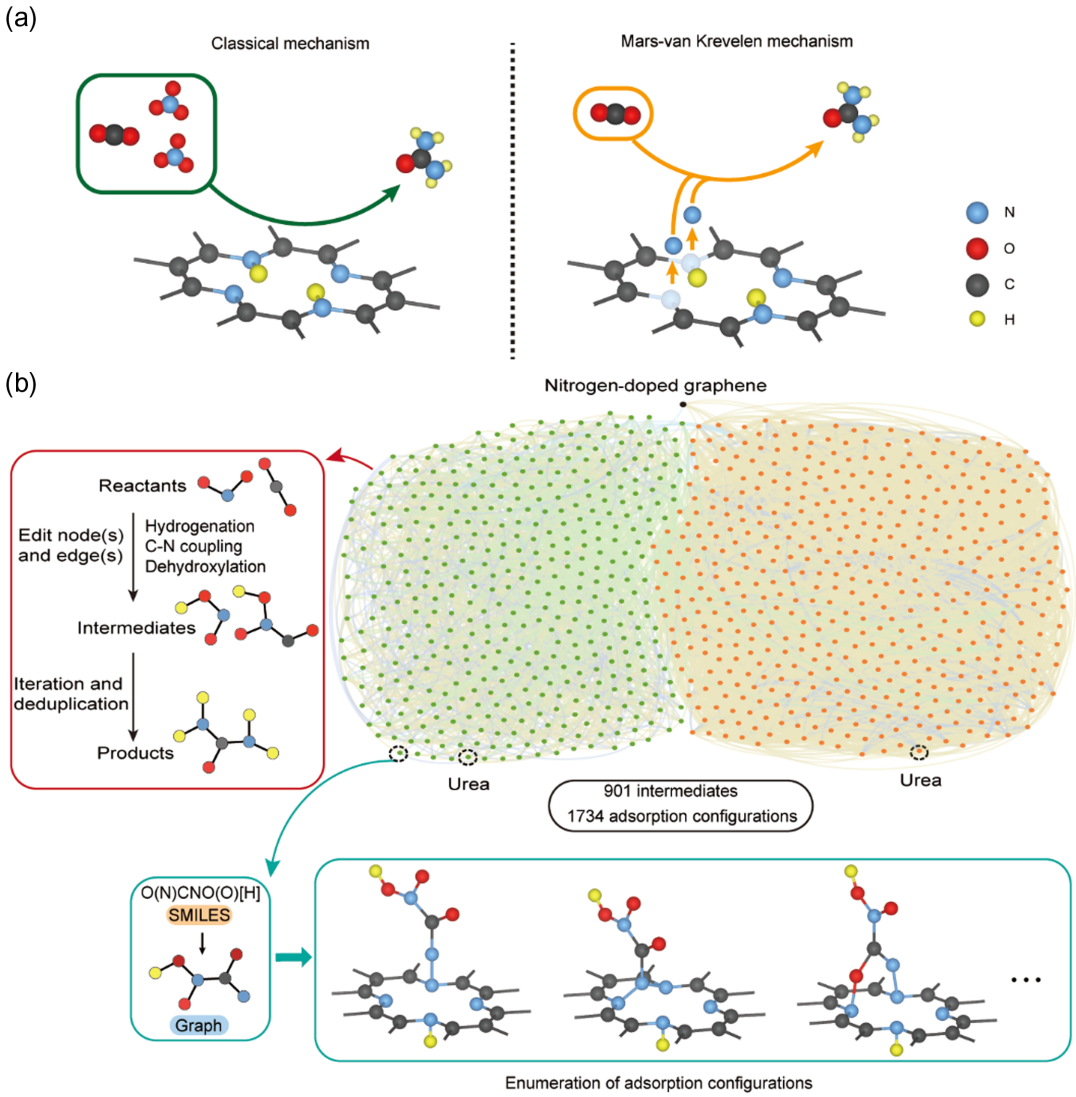



**Figure 3** | *Construction of reaction network for urea electrosynthesis on a nitrogen-doped carbon catalyst. (a) Two reaction mechanisms for urea formation on the catalyst surface. (b) The scheme of overall reaction network, the typical elementary steps, and the example of adsorption configuration enumeration. Green dots represent reaction intermediates in classical mechanism, while orange dots represent those in MvK mechanism.*

graph-based stability prediction (GSP) algorithm, that have not been calculated via DFT. Subgraph fingerprints are employed for this model, with the local structure of the adsorbate and surrounding microenvironments decomposed into a number of fragments according to the nearest-neighbor atoms of each central atom (details in Supporting Information).

For all the reaction intermediates, there are three types of bond rearrangements in the two reaction mechanisms, as presented in Figure 4. The first one is bond rupture of the adsorbate, which occurs in the classical mechanism. Although the constructed initial adsorption configurations satisfy the valency rules, it is still probable that delocalized electron will be energetically more favorable, leading to the appearance of atoms that do not conform to the eight-electron rule. The other two types of bond rearrangements take place in the MvK mechanism. Adsorption of additional H atoms to the catalytic center is required for the passivation of dangling bonds due to the participation of the surface N atoms in forming urea product. This will lead to frequent occurrence of redistribution of these H adsorbates, which could facilitate the breakage of C–N bond at the catalyst surface and the formation of C–N bond for the reaction intermediate. The reaction intermediate can be regarded as stable if one of the adsorption configurations does not undergo any rearrangement upon structural optimization.

A graph-based gradient boosting classification model is employed for stability prediction, involving three types of bond rearrangements in the two reaction mechanisms. From the gradient boosting classification model, we can determine which reaction intermediates are highly

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

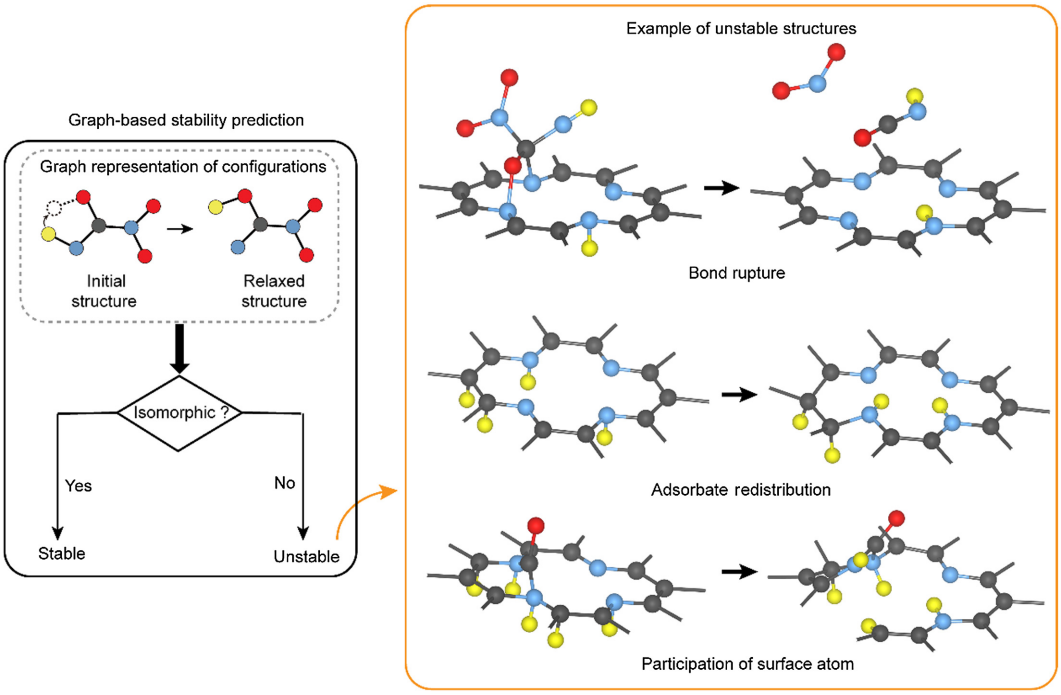



**Figure 4** | *The graph-based stability prediction of intermediates (GSP algorithm).*

unstable in the reaction network, and they will be discarded in the following procedure. Then, a ridge regression model is employed to estimate the formation energies of the stable intermediates, also using the DFT results as the training data set and relying on the same fingerprints as those in the gradient boosting classification model. The most likely reaction pathway is determined from the ridge regression model, and the second active loop will be initiated by supplementing DFT calculations for the intermediates in this path.

## Model performance after refinement

There may be multiple reaction pathways that exhibit the same limiting potential, that is, the same value for the maximum uphill free energy change ($\Delta G_{max}$). This necessitates another criterion for the evaluation of optimal reaction pathway. Given that a flattened energy landscape is highly preferable from the kinetic perspective, an attribute that corresponds to the flatness of the free energy profile in the reaction pathway will be an option.[71,72] Here, we use the sum of free energy changes for those uphill elementary reaction steps ($\Sigma\ \Delta G_i$) as the second determinant for the optimal reaction pathway (Figure 5a). Different from thermocatalysis, in which the reaction steps can proceed in both forward and backward directions, the investigation in electrocatalysis will only focus on the forward direction. Therefore, only when the free energy change is positive should we take this elementary step into consideration, and a smaller $\Sigma\ \Delta G_i$ will apparently imply a flatter energy landscape that is more kinetically feasible for the reaction to take place.

To initiate the active learning loop, we began by performing DFT calculations on all species involved in the pathway for the reduction of $NO_3^-$ to $NH_3$, as well as the reduction of $CO_2$ to CO. Additionally, we randomly selected 10% of the species from the reaction network for further calculations. The two parts of calculations served as the initial dataset. In each active learning loop, DFT calculations for the most likely reaction pathway predicted in the last loop will be conducted, and the corresponding data will be fed into the model for the present loop. Yet, this is not sufficient for a significant improvement in model accuracy. Hence, we select an additional group of reaction intermediates for DFT calculations, which is according to their connectivity to other intermediates in the network. By treating the reaction network as a graph, the nodes represent intermediates and edges represent elementary steps. A higher degree of the node indicates that it is connected with a larger number of intermediates, and therefore there is a larger opportunity that the optimal path will come across this node. By performing a degree analysis to the graph of reaction network (Supporting Information Figure S11), we select the corresponding intermediates for DFT calculations in the new loop, which, alongside those from the optimal path in the last loop, can add up to a total of around 10% of all the reaction intermediates in the network.

We take an iteration process shown in Figure 5b as an example, which involves intermediates belonging to the MvK mechanism. After four iterations (around 40% of all

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

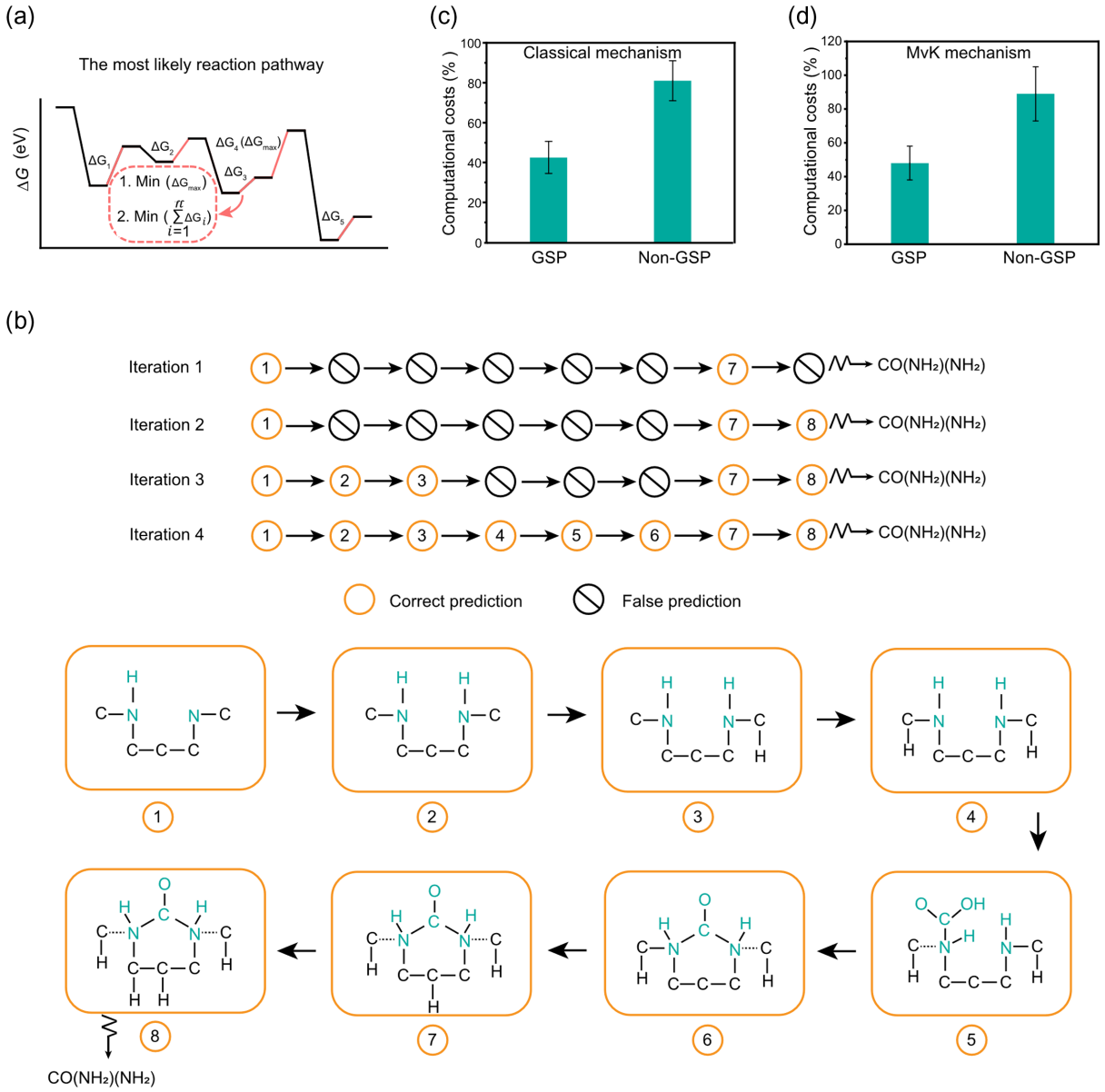



**Figure 5** | *Performance of the graph-based model. (a) Definition of the most likely reaction pathway. (b) An example of the iteration process for searching the most likely reaction pathway of the MvK mechanism as benchmarked by the DFT calculations. After four iterations, the most likely reaction pathways can be fully derived. The average DFT calculation costs (percentage of reaction intermediates calculated by DFT) for the prediction of (c) classical mechanism and (d) MvK Mechanism with and without GSP algorithm. The large error bars primarily stem from the randomness of certain prediction results of 100 rounds of testing, and each 10% increase in data also contributes to the data dispersion. The 10% increment can be further reduced in practical applications.*

the intermediates are calculated via DFT), the most likely reaction pathway predicted by the surrogate model is found identical to that obtained from DFT calculations of the entire reaction network (Supporting Information Figure S12). We have performed similar tests 100 times, starting with different sets of intermediates employed in the initial loop. The results show that in most circumstances, only around 40% of the entire network should be computed at DFT level before the most likely reaction pathway is identified (Figure 5c,d and Supporting Information Figure S13). In contrast, about 90% of the intermediates need to be evaluated via DFT, if we do not apply the GSP procedure. According to the energy profiles of reaction pathways for classical and MvK mechanisms (Supporting Information Figures S14 and S15), we find that the classical mechanism is more preferable ($\Delta G_{max}$ = 0.46 eV) for urea electrosynthesis on the nitrogen-doped carbon catalyst.

Our workflow can also be adopted for the identification of multiple thermodynamic probable reaction pathways

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

in a complex reaction network. Given that a low limiting potential or a relatively flat energy landscape does not always guarantee superior kinetics as compared to other routes,[73] there is a need to pinpoint multiple probable paths that are thermodynamically more favorable than the rest. This entails extracting a small subnetwork form the entire reaction network that correspond to a few optimal pathways.[74] Here, we predict the ranking of the top-five favorable reaction pathways according to $\Delta G_{max}$, with $\Sigma\ \Delta G_i$ as small as possible for each $\Delta G_{max}$ value. The results (Supporting Information Figure S16) show that the computational costs are increased by less than 5% in the classical mechanism and less than 10% in the MvK one, suggesting that the intermediates involved in top-five favorable reaction pathways have almost been examined in the procedure of finding the most likely reaction pathway. This also indicates that our workflow is robust and effective for the prediction of subnetworks rather than only the most likely reaction pathway.

## Discussion

Complex electrochemical reactions involving multiple electron transfer steps have become a focus in recent years, which spurs the research needed to reduce the amount of time and labor spent on high-throughput calculations at the DFT level. Previous pioneering studies have relied on scaling-relation–based ML models,[6] which, after several years of modifications, have seen an unprecedented success in different electrocatalytic reactions. However, an aspect that is often overlooked in previous ML models is the dynamics of catalyst surface structure.[4] This can be ascribed to the difficulty to configure DFT calculations for the reconstructed surfaces, given that the combination of various adsorption structures and diverse surface structures could lead to an enormous number of possible configurations. Here, we manage to establish a graph-based ML framework to tackle this challenge, using urea electrosynthesis on a nitrogen-doped carbon catalyst as an example. At the core of our methodology lies the application of graph edit operations and graph isomorphism matching that can enable marked reduction in the computational burden for identifying unstable structures. The unstable reaction intermediates mainly stem from the subtle interaction between the intermediate and the catalytic center, which may undergo significant bond rearrangements such as those displayed in the MvK mechanism, exhibiting strong dynamic characteristics. The graph-based algorithm can capture the intrinsic difference and relationship between two adsorption configurations, which offers great efficiency in our active learning loop for exploration of the most likely reaction pathway.

Our workflow is highly automated and can be transferred to other electrocatalytic reactions. While ML models are employed in this workflow, the output pathway and the corresponding free energy changes are quantitatively robust, because our framework guarantees that all the elementary steps in the final output pathway have been calculated via DFT. Combined with the complexity of the system discussed above, these explain why the DFT computational cost can only be reduced to around 40%. Unless we consider it unnecessary that the predicted pathway must be within DFT accuracy, it is hardly achievable to further reduce the DFT cost in the present workflow. Although we believe that the computational cost could be further reduced by simply using the output of ML models as the final result, it is worth noting that ML models trained for specific catalytic systems are generally customized with regard to specific feature extraction methods. This will imply that a significant drop in accuracy will probably be seen when it is extended to other catalytic systems, because of the inherent heavy reliance on the performance of the ML models. However, our on-the-fly active learning workflow can mitigate such reliance, since the iteration will proceed on the basis that the optimal pathway has been evaluated at the DFT level.

This scheme can be further improved in the direction of kinetic aspects for the catalytic reactions. Activation energy calculations can be integrated into the iteration process or be performed on the derived subnetwork to obtain pathways that are feasible in both kinetics and thermodynamics, and this can be typically achieved by the nudged elastic band methods.[75] However, it will considerably increase the computational burden. Although some recent studies have attempted to speed up the calculations via the universal Brønsted–Evans–Polyani relationship[6] or the ML force fields,[76] a fully automated and time-efficient scheme still remains to be developed, especially for the complex catalytic systems as exemplified in this work. For DFT calculations, different levels of theory and different parameterizations may yield different results, and therefore the ensemble-based approached[77] can be used to evaluate the uncertainty in the calculation process.

## Conclusion

Collectively, we develop a graph-based ML framework that integrates the function of reaction network generation, adsorption configuration enumeration, intermediate stability classification, and formation energy prediction to effectively identify the most likely reaction pathway in a complex electrocatalytic system with much surface dynamics. The stability prediction is essential for the reduction of DFT computational costs, as demonstrated by the example of urea electrosynthesis on the nitrogen-doped carbon catalyst. The calculation results show that the framework is effective with the DFT cost decreased

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

by more than one half. Our proposed workflow enables researchers to rapidly discover favorable reaction pathways of complex catalytic process and thus can be leveraged to high-throughput DFT-based screening and prediction of various kinds of electrocatalysts.

## Supporting Information

Supporting Information is available and includes additional DFT computational details, the details of the framework, the overall free energy profiles of two mechanism, and the Github repo of code to generate results.

## Conflict of Interest

The authors declare no competing interests.

## Funding Information

This work was supported by the Soft Science Research Project of Guangdong Province (grant no. 2017B030301013), the National Natural Science Foundation of China (grant no. 22109003), the Basic and Applied Basic Research Foundation of Guangdong Province (grant no. 2021B1515130002), the Shenzhen Key Laboratory of New Energy Resources Genome Preparation and Testing (grant no. ZDSYS201707281026184), and the Major Science and Technology Infrastructure Project of Material Genome Big-science Facilities Platform supported by Municipal Development and Reform Commission of Shenzhen.

## References

1. Nitopi, S.; Bertheussen, E.; Scott, S. B.; Liu, X.; Engstfeld, A. K.; Horch, S.; Seger, B.; Stephens, I. E. L.; Chan, K.; Hahn, C.; Nørskov, J. K.; Jaramillo, T. F.; Chorkendorff, I. Progress and Perspectives of Electrochemical $CO_2$ Reduction on Copper in Aqueous Electrolyte. *Chem. Rev.* **2019**, *119*, 7610–7672.
2. Peng, X.; Zeng, L.; Wang, D.; Liu, Z.; Li, Y.; Li, Z.; Yang, B.; Lei, L.; Dai, L.; Hou, Y. Electrochemical C–N Coupling of $CO_2$ and Nitrogenous Small Molecules for the Electrosynthesis of Organonitrogen Compounds. *Chem. Soc. Rev.* **2023**, *52*, 2193–2237.
3. Xiong, Y.; Wang, Y.; Zhou, J.; Liu, F.; Hao, F.; Fan, Z. Electrochemical Nitrate Reduction: Ammonia Synthesis and the Beyond. *Adv. Mater.* **2024**, *36*, e2304021.
4. Margraf, J. T.; Jung, H.; Scheurer, C.; Reuter, K. Exploring Catalytic Reaction Networks with Machine Learning. *Nat. Catal.* **2023**, *6*, 112–121.
5. Miao, L.; Jia, W.; Cao, X.; Jiao, L. Computational Chemistry for Water-Splitting Electrocatalysis. *Chem. Soc. Rev.* **2024**, *53*, 2771–2807.
6. Ulissi, Z. W.; Medford, A. J.; Bligaard, T.; Nørskov, J. K. To Address Surface Reaction Network Complexity Using Scaling Relations Machine Learning and DFT Calculations. *Nat. Commun.* **2017**, *8*, 14621.
7. Zheng, S.; Liang, X.; Pan, J.; Hu, K.; Li, S.; Pan, F. Multi-Center Cooperativity Enables Facile C–C Coupling in Electrochemical $CO_2$ Reduction on a $Ni_2P$ Catalyst. *ACS Catal.* **2023**, *13*, 2847–2856.
8. Hossain, M. D.; Huang, Y.; Yu, T. H.; Goddard Iii, W. A.; Luo, Z. Reaction Mechanism and Kinetics for $CO_2$ Reduction on Nickel Single Atom Catalysts from Quantum Mechanics. *Nat. Commun.* **2020**, *11*, 2256.
9. Osella, S.; Goddard Iii, W. A. $CO_2$ Reduction to Methane and Ethylene on a Single-Atom Catalyst: A Grand Canonical Quantum Mechanics Study. *J. Am. Chem. Soc.* **2023**, *145*, 21319–21329.
10. Singstock, N. R.; Musgrave, C. B. How the Bioinspired $Fe_2Mo_6S_8$ Chevrel Breaks Electrocatalytic Nitrogen Reduction Scaling Relations. *J. Am. Chem. Soc.* **2022**, *144*, 12800–12806.
11. Hu, X.; Chen, S.; Chen, L.; Tian, Y.; Yao, S.; Lu, Z.; Zhang, X.; Zhou, Z. What is the Real Origin of the Activity of Fe-N-C Electrocatalysts in the $O_2$ Reduction Reaction? Critical Roles of Coordinating Pyrrolic N and Axially Adsorbing Species. *J. Am. Chem. Soc.* **2022**, *144*, 18144–18152.
12. Vijay, S.; Ju, W.; Brückner, S.; Tsang, S.-C.; Strasser, P.; Chan, K. Unified Mechanistic Understanding of $CO_2$ Reduction to CO on Transition Metal and Single Atom Catalysts. *Nat. Catal.* **2021**, *4*, 1024–1031.
13. Skulason, E.; Bligaard, T.; Gudmundsdottir, S.; Studt, F.; Rossmeisl, J.; Abild-Pedersen, F.; Vegge, T.; Jonsson, H.; Nørskov, J. K. A Theoretical Evaluation of Possible Transition Metal Electro-Catalysts for $N_2$ Reduction. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1235–1245.
14. Rohr, B. A.; Singh, A. R.; Gauthier, J. A.; Statt, M. J.; Nørskov, J. K. Micro-Kinetic Model of Electrochemical Carbon Dioxide Reduction over Platinum in Non-Aqueous Solvents. *Phys. Chem. Chem. Phys.* **2020**, *22*, 9040–9045.
15. Yoo, J. S.; Christensen, R.; Vegge, T.; Nørskov, J. K.; Studt, F. Theoretical Insight into the Trends that Guide the Electrochemical Reduction of Carbon Dioxide to Formic Acid. *ChemSusChem* **2016**, *9*, 358–363.
16. Wen, M.; Spotte-Smith, E. W. C.; Blau, S. M.; McDermott, M. J.; Krishnapriyan, A. S.; Persson, K. A. Chemical Reaction Networks and Opportunities for Machine Learning. *Nat. Comput. Sci.* **2023**, *3*, 12–24.
17. Yuan, M.; Chen, J.; Bai, Y.; Liu, Z.; Zhang, J.; Zhao, T.; Wang, Q.; Li, S.; He, H.; Zhang, G. Unveiling Electrochemical Urea Synthesis by Co-Activation of $CO_2$ and $N_2$ with Mott-Schottky Heterostructure Catalysts. *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 10910–10918.
18. Li, Y.; Zheng, S.; Liu, H.; Xiong, Q.; Yi, H.; Yang, H.; Mei, Z.; Zhao, Q.; Yin, Z. W.; Huang, M.; Lin, Y.; Lai, W.; Dou, S. X.; Pan, F.; Li, S. Sequential Co-Reduction of Nitrate and Carbon Dioxide Enables Selective Urea Electrosynthesis. *Nat. Commun.* **2024**, *15*, 176.
19. Meng, N.; Huang, Y.; Liu, Y.; Yu, Y.; Zhang, B. Electrosynthesis of Urea from Nitrite and $CO_2$ over Oxygen Vacancy-Rich ZnO Porous Nanosheets. *Cell Rep. Phys. Sci.* **2021**, *2*, 100378.

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

20. Chen, X.; Lv, S.; Gu, H.; Cui, H.; Liu, G.; Liu, Y.; Li, Z.; Xu, Z.; Kang, J.; Teobaldi, G.; Liu, L. M.; Guo, L. Amorphous Bismuth-Tin Oxide Nanosheets with Optimized C–N Coupling for Efficient Urea Synthesis. *J. Am. Chem. Soc.* **2024**, *146*, 13527–13535.
21. Lv, C.; Zhong, L.; Liu, H.; Fang, Z.; Yan, C.; Chen, M.; Kong, Y.; Lee, C.; Liu, D.; Li, S.; Liu, J.; Song, L.; Chen, G.; Yan, Q.; Yu, G. Selective Electrocatalytic Synthesis of Urea with Nitrate and Carbon Dioxide. *Nat. Sustain.* **2021**, *4*, 868–876.
22. Wang, C.; Wang, X.; Ren, H.; Zhang, Y.; Zhou, X.; Wang, J.; Guan, Q.; Liu, Y.; Li, W. Combining Fe Nanoparticles and Pyrrole-Type Fe-$N_4$ Sites on Less-Oxygenated Carbon Supports for Electrochemical $CO_2$ Reduction. *Nat. Commun.* **2023**, *14*, 5108.
23. Feng, J.; Zhang, L.; Liu, S.; Xu, L.; Ma, X.; Tan, X.; Wu, L.; Qian, Q.; Wu, T.; Zhang, J.; Sun, X.; Han, B. Modulating Adsorbed Hydrogen Drives Electrochemical $CO_2$-to-$C_2$ Products. *Nat. Commun.* **2023**, *14*, 4615.
24. Dai, Y.; Li, H.; Wang, C.; Xue, W.; Zhang, M.; Zhao, D.; Xue, J.; Li, J.; Luo, L.; Liu, C.; Li, X.; Cui, P.; Jiang, Q.; Zheng, T.; Gu, S.; Zhang, Y.; Xiao, J.; Xia, C.; Zeng, J. Manipulating Local Coordination of Copper Single Atom Catalyst Enables Efficient $CO_2$-to-$CH_4$ Conversion. *Nat. Commun.* **2023**, *14*, 3382.
25. Zhou, Y.; Martín, A. J.; Dattila, F.; Xi, S.; López, N.; Pérez-Ramírez, J.; Yeo, B. S. Long-Chain Hydrocarbons by $CO_2$ Electroreduction Using Polarized Nickel Catalysts. *Nat. Catal.* **2022**, *5*, 545–554.
26. Wang, X.; Ou, P.; Ozden, A.; Hung, S.-F.; Tam, J.; Gabardo, C. M.; Howe, J. Y.; Sisler, J.; Bertens, K.; García de Arquer, F. P.; Miao, R. K.; O'Brien, C. P.; Wang, Z.; Abed, J.; Rasouli, A. S.; Sun, M.; Ip, A. H.; Sinton, D.; Sargent, E. H. Efficient Electrosynthesis of n-Propanol from Carbon Monoxide Using a Ag–Ru–Cu Catalyst. *Nat. Energy* **2022**, *7*, 170–176.
27. Xu, H.; Rebollar, D.; He, H.; Chong, L.; Liu, Y.; Liu, C.; Sun, C.-J.; Li, T.; Muntean, J. V.; Winans, R. E.; Liu, D.-J.; Xu, T. Highly Selective Electrocatalytic $CO_2$ Reduction to Ethanol by Metallic Clusters Dynamically Formed from Atomically Dispersed Copper. *Nat. Energy* **2020**, *5*, 623–632.
28. Voiry, D.; Shin, H. S.; Loh, K. P.; Chhowalla, M. Low-Dimensional Catalysts for Hydrogen Evolution and $CO_2$ Reduction. *Nat. Rev. Chem.* **2018**, *2*, 0105.
29. Zhou, Y.; Yeo, B. S. Formation of C–C Bonds During Electrocatalytic $CO_2$ Reduction on Non-Copper Electrodes. *J. Mater. Chem. A* **2020**, *8*, 23162–23186.
30. Tao, Z.; Rooney, C. L.; Liang, Y.; Wang, H. Accessing Organonitrogen Compounds via C–N Coupling in Electrocatalytic $CO_2$ Reduction. *J. Am. Chem. Soc.* **2021**, *143*, 19630–19642.
31. Deshpande, S.; Maxson, T.; Greeley, J. Graph Theory Approach to Determine Configurations of Multidentate and High Coverage Adsorbates for Heterogeneous Catalysis. *npj Comput. Mater.* **2020**, *6*, 79.
32. Xu, W.; Reuter, K.; Andersen, M. Predicting Binding Motifs of Complex Adsorbates Using Machine Learning with a Physics-Inspired Graph Representation. *Nat. Comput. Sci.* **2022**, *2*, 443–450.
33. Boes, J. R.; Mamun, O.; Winther, K.; Bligaard, T. Graph Theory Approach to High-Throughput Surface Adsorption Structure Generation. *J. Phys. Chem. A* **2019**, *123*, 2281–2285.
34. Pablo-García, S.; Morandi, S.; Vargas-Hernández, R. A.; Jorner, K.; Ivković, Ž.; López, N.; Aspuru-Guzik, A. Fast Evaluation of the Adsorption Energy of Organic Molecules on Metals via Graph Neural Networks. *Nat. Comput. Sci.* **2023**, *3*, 433–442.
35. Gu, G. H.; Lee, M.; Jung, Y.; Vlachos, D. G. Automated Exploitation of the Big Configuration Space of Large Adsorbates on Transition Metals Reveals Chemistry Feasibility. *Nat. Commun.* **2022**, *13*, 2087.
36. Li, H.; Li, X.; Wang, P.; Zhang, Z.; Davey, K.; Shi, J. Q.; Qiao, S. Z. Machine Learning Big Data Set Analysis Reveals C–C Electro-Coupling Mechanism. *J. Am. Chem. Soc.* **2024**, *146*, 22850–22858.
37. Jiao, Y.; Li, H.; Jiao, Y.; Qiao, S.-Z. Activity and Selectivity Roadmap for C–N Electro-Coupling on MXenes. *J. Am. Chem. Soc.* **2023**, *145*, 15572–15580.
38. Schaaf, L. L.; Fako, E.; De, S.; Schäfer, A.; Csányi, G. Accurate Energy Barriers for Catalytic Reaction Pathways: An Automatic Training Protocol for Machine Learning Force Fields. *npj Comput. Mater.* **2023**, *9*, 180.
39. Cohen, M.; Goculdas, T.; Vlachos, D. G. Active Learning of Chemical Reaction Networksviaprobabilistic Graphical Models and Boolean Reaction Circuits. *React. Chem. Eng.* **2023**, *8*, 824–837.
40. Shi, Y. F.; Kang, P. L.; Shang, C.; Liu, Z. P. Methanol Synthesis from $CO_2$/CO Mixture on Cu-Zn Catalysts from Microkinetics-Guided Machine Learning Pathway Search. *J. Am. Chem. Soc.* **2022**, *144*, 13401–13414.
41. Chen, D.; Shang, C.; Liu, Z.-P. Machine-Learning Atomic Simulation for Heterogeneous Catalysis. *npj Comput. Mater.* **2023**, *9*, 2.
42. Yang, M.; Raucci, U.; Parrinello, M. Reactant-Induced Dynamics of Lithium Imide Surfaces During the Ammonia Decomposition Process. *Nat. Catal.* **2023**, *6*, 829–836.
43. Andersen, M.; Reuter, K. Adsorption Enthalpies for Catalysis Modeling Through Machine-Learned Descriptors. *Acc. Chem. Res.* **2021**, *54*, 2741–2749.
44. Tran, K.; Ulissi, Z. W. Active Learning Across Intermetallics to Guide Discovery of Electrocatalysts for $CO_2$ Reduction and $H_2$ Evolution. *Nat. Catal.* **2018**, *1*, 696–703.
45. Garcia-Muelas, R.; Lopez, N. Statistical Learning Goes Beyond the d-Band Model Providing the Thermochemistry of Adsorbates on Transition Metals. *Nat. Commun.* **2019**, *10*, 4687.
46. Kang, P. L.; Shi, Y. F.; Shang, C.; Liu, Z. P. Artificial Intelligence Pathway Search to Resolve Catalytic Glycerol Hydrogenolysis Selectivity. *Chem. Sci.* **2022**, *13*, 8148–8160.
47. Walker, E. A.; Mohammadi, M. M.; Swihart, M. T. Graph Theory Model of Dry Reforming of Methane Using Rh(111). *J. Phys. Chem. Lett.* **2020**, *11*, 4917–4922.
48. Calvinho, K. U. D.; Laursen, A. B.; Yap, K. M. K.; Goetjen, T. A.; Hwang, S.; Murali, N.; Mejia-Sosa, B.; Lubarski, A.; Teeluck, K. M.; Hall, E. S.; Garfunkel, E.; Greenblatt, M.; Dismukes, G. C. Selective $CO_2$ Reduction to $C_3$ and $C_4$

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

Oxyhydrocarbons on Nickel Phosphides at Overpotentials as Low as 10 mV. *Energy Environ. Sci.* **2018**, *11*, 2550–2559.
49. Zheng, T.; Zhang, M.; Wu, L.; Guo, S.; Liu, X.; Zhao, J.; Xue, W.; Li, J.; Liu, C.; Li, X.; Jiang, Q.; Bao, J.; Zeng, J.; Yu, T.; Xia, C. Upcycling $CO_2$ into Energy-Rich Long-Chain Compounds via Electrochemical and Metabolic Engineering. *Nat. Catal.* **2022**, *5*, 388–396.
50. Choi, M.; Bong, S.; Kim, J. W.; Lee, J. Formation of 1-Butanol from $CO_2$ Without *CO Dimerization on a Phosphorus-Rich Copper Cathode. *ACS Energy Lett.* **2021**, *6*, 2090–2095.
51. Jin, J.; Wicks, J.; Min, Q.; Li, J.; Hu, Y.; Ma, J.; Wang, Y.; Jiang, Z.; Xu, Y.; Lu, R.; Si, G.; Papangelakis, P.; Shakouri, M.; Xiao, Q.; Ou, P.; Wang, X.; Chen, Z.; Zhang, W.; Yu, K.; Song, J.; Jiang, X.; Qiu, P.; Lou, Y.; Wu, D.; Mao, Y.; Ozden, A.; Wang, C.; Xia, B. Y.; Hu, X.; Dravid, V. P.; Yiu, Y. M.; Sham, T. K.; Wang, Z.; Sinton, D.; Mai, L.; Sargent, E. H.; Pang, Y. Constrained $C_2$ Adsorbate Orientation Enables CO-to-Acetate Electroreduction. *Nature* **2023**, *617*, 724–729.
52. Wei, P.; Gao, D.; Liu, T.; Li, H.; Sang, J.; Wang, C.; Cai, R.; Wang, G.; Bao, X. Coverage-Driven Selectivity Switch from Ethylene to Acetate in High-Rate $CO_2$/CO Electrolysis. *Nat. Nanotechnol.* **2023**, *18*, 299–306.
53. Li, J.; Wang, Z.; McCallum, C.; Xu, Y.; Li, F.; Wang, Y.; Gabardo, C. M.; Dinh, C.-T.; Zhuang, T.-T.; Wang, L.; Howe, J. Y.; Ren, Y.; Sargent, E. H.; Sinton, D. Constraining CO Coverage on Copper Promotes High-Efficiency Ethylene Electroproduction. *Nat. Catal.* **2019**, *2*, 1124–1131.
54. Dorakhan, R.; Grigioni, I.; Lee, B.-H.; Ou, P.; Abed, J.; O'Brien, C.; Sedighian Rasouli, A.; Plodinec, M.; Miao, R. K.; Shirzadi, E.; Wicks, J.; Park, S.; Lee, G.; Zhang, J.; Sinton, D.; Sargent, E. H. A Silver–Copper Oxide Catalyst for Acetate Electrosynthesis from Carbon Monoxide. *Nat. Synth.* **2023**, *2*, 448–457.
55. Wang, K.; Mao, R.; Liu, R.; Zhang, J.; Zhao, H.; Ran, W.; Zhao, X. Intentional Corrosion-Induced Reconstruction of Defective NiFe Layered Double Hydroxide Boosts Electrocatalytic Nitrate Reduction to Ammonia. *Nat. Water* **2023**, *1*, 1068–1078.
56. Ferreira de Araujo, J.; Dionigi, F.; Merzdorf, T.; Oh, H. S.; Strasser, P. Evidence of Mars-Van-Krevelen Mechanism in the Electrochemical Oxygen Evolution on Ni-Based Catalysts. *Angew. Chem. Int. Ed. England* **2021**, *60*, 14981–14988.
57. Qiao, Z.; Johnson, D.; Djire, A. Challenges and Opportunities for Nitrogen Reduction to Ammonia on Transitional Metal Nitrides via Mars-van Krevelen Mechanism. *Cell Rep. Phys. Sci.* **2021**, *2*, 100438.
58. Efremenko, I.; Neumann, R. Computational Insight into the Initial Steps of the Mars-van Krevelen Mechanism: Electron Transfer and Surface Defects in the Reduction of Polyoxometalates. *J. Am. Chem. Soc.* **2012**, *134*, 20669–20680.
59. Wang, C.; Gu, X.-K.; Yan, H.; Lin, Y.; Li, J.; Liu, D.; Li, W.-X.; Lu, J. Water-Mediated Mars–Van Krevelen Mechanism for CO Oxidation on Ceria-Supported Single-Atom $Pt_1$ Catalyst. *ACS Catal.* **2016**, *7*, 887–891.
60. Zhai, G.; Cai, L.; Ma, J.; Chen, Y.; Liu, Z.; Si, S.; Duan, D.; Sang, S.; Li, J.; Wang, X.; Liu, Y. A.; Qian, B.; Liu, C.; Pan, Y.; Zhang, N.; Liu, D.; Long, R.; Xiong, Y. Highly Efficient, Selective, and Stable Photocatalytic Methane Coupling to Ethane Enabled by Lattice Oxygen Looping. *Sci. Adv.* **2024**, *10*, eado4390.
61. Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59*, 1758–1775.
62. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
63. Mathew, K.; Kolluru, V. S. C.; Mula, S.; Steinmann, S. N.; Hennig, R. G. Implicit Self-Consistent Electrolyte Model in Plane-Wave Density-Functional Theory. *J. Chem. Phys.* **2019**, *151*, 234101.
64. Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L.; Kitchin, J. R.; Bligaard, T.; Jónsson, H. Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode. *J. Phy. Chem. B* **2004**, *108*, 17886–17892.
65. Lovric, M.; Molero, J. M.; Kern, R. PySpark and RDKit: Moving Towards Big Data in Cheminformatics. *Mol. Inform.* **2019**, *38*, e1800082.
66. Mukherjee, S.; Cullen, D. A.; Karakalos, S.; Liu, K.; Zhang, H.; Zhao, S.; Xu, H.; More, K. L.; Wang, G.; Wu, G. Metal-Organic Framework-Derived Nitrogen-Doped Highly Disordered Carbon for Electrochemical Ammonia Synthesis Using $N_2$ and $H_2O$ in Alkaline Electrolytes. *Nano Energy* **2018**, *48*, 217–226.
67. Li, S.; Chen, Z.; Wang, Z.; Weng, M.; Li, J.; Zhang, M.; Lu, J.; Xu, K.; Pan, F. Graph-Based Discovery and Analysis of Atomic-Scale One-Dimensional Materials. *Natl. Sci. Rev.* **2022**, *9*, nwac028.
68. Weng, M.; Wang, Z.; Qian, G.; Ye, Y.; Chen, Z.; Chen, X.; Zheng, S.; Pan, F. Identify Crystal Structures by a New Paradigm Based on Graph Theory for Building Materials Big Data. *Sci. China Chem.* **2019**, *62*, 982–986.
69. Zhang, W.; Weng, M.; Zhang, M.; Chen, Z.; Wang, B.; Li, S.; Pan, F. Rapid Mining of Fast Ion Conductors via Subgraph Isomorphism Matching. *J. Am. Chem. Soc.* **2024**, *146*, 18535–18543.
70. Zheng, S.; Ding, H.; Li, S.; Chen, D.; Pan, F. Application of Topology-Based Structure Features for Machine Learning in Materials Science. *Chinese J. Struct. Chem.* **2023**, *42*, 100120.
71. Robertson, C.; Habershon, S. Fast Screening of Homogeneous Catalysis Mechanisms Using Graph-Driven Searches and Approximate Quantum Chemistry. *Catal. Sci. Technol.* **2019**, *9*, 6357–6369.
72. Raugei, S.; DuBois, D. L.; Rousseau, R.; Chen, S.; Ho, M. H.; Bullock, R. M.; Dupuis, M. Toward Molecular Catalysts by Computer. *Acc. Chem. Res.* **2015**, *48*, 248–55.
73. Govind Rajan, A.; Carter, E. A. Discovering Competing Electrocatalytic Mechanisms and Their Overpotentials: Automated Enumeration of Oxygen Evolution Pathways. *J. Phys. Chem. C* **2020**, *124*, 24883–24898.
74. Kim, Y.; Kim, J. W.; Kim, Z.; Kim, W. Y. Efficient Prediction of Reaction Paths Through Molecular Graph and Reaction Network Analysis. *Chem. Sci.* **2018**, *9*, 825–835.
75. Henkelman, G.; Uberuaga, B. P.; Jo´nsson, H. A Climbing Image Nudged Elastic Band Method for Finding Saddle

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955

Points and Minimum Energy Paths. *J. Chem. Phys.* **2000**, *113*, 9901.
76. Garrido Torres, J. A.; Jennings, P. C.; Hansen, M. H.; Boes, J. R.; Bligaard, T. Low-Scaling Algorithm for Nudged Elastic Band Calculations Using a Surrogate Machine Learning Model. *Phys. Rev. Lett.* **2019**, *122*, 156001.
77. Wellendorff, J.; Lundgaard, K. T.; Møgelhøj, A.; Petzold, V.; Landis, D. D.; Nørskov, J. K.; Bligaard, T.; Jacobsen, K. W. Density Functionals for Surface Science: Exchange-Correlation Model Development with Bayesian Error Estimation. *Phys. Rev. B* **2012**, *85*, 235149.

**DOI:** 10.31635/ccschem.024.202404955
**Citation:** *CCS Chem.* **2025**, 7, 2822–2834
**Link to VoR:** https://doi.org/10.31635/ccschem.024.202404955