## PAPER

# Breakdown of linear correlations in methane activation on doped nickel nanoparticles: physical origins and nonlinear relations revealed by symbolic regression

Yang Wei, [iD] †[a] Yang Wang,†[a] Yibo Chang,[a] Simone Lang, [iD] [b]
Shiru Lin [iD] *[b] and Junwei Lucas Bao [iD] *[a]

Dry reforming of methane (DRM) both mitigates greenhouse gases (CH$_4$, CO$_2$) and produces syngas (H$_2$, CO) for fuels and chemicals. Designing efficient DRM catalysts requires quantitative links between reaction barriers and adsorption/reaction energies. Traditional Brønsted–Evans–Polanyi (BEP) scaling, although widely assumed, may fail to predict quantitative kinetics, motivating the discovery of nonlinear relations that better couple thermodynamics and kinetics. Here, we investigate CH$_4$ and CO$_2$ decomposition on 12 doped Ni nanoparticles (M–Ni) using density functional theory. We show that BEP breakdown in DRM arises not only from binding-site preferences but also from the intrinsic electronic properties of the dopant, as analyzed by diabatic state calculations; kinetic modeling further amplifies these deviations. To uncover governing connections, we employ symbolic regression to derive physically interpretable expressions linking dopant descriptors to kinetic and thermodynamic parameters. Compared with linear scaling, these relations capture nonlinear effects, reveal hidden correlations, and provide predictive formulas that accelerate high-throughput screening. Validation on 12 previously unexplored doped nanoparticles confirms transferability: the best candidate, Mo–Ni, shows overall reaction energy and rate-determining barrier predictions within 0.05 eV of DFT. This integrated mechanistic and data-driven framework offers design rules for highly active, coke-resistant DRM nanocatalysts beyond BEP limitations.

## 1. Introduction

Dry reforming of methane (DRM) is a promising process that converts two major greenhouse gases, methane (CH$_4$) and carbon dioxide (CO$_2$), into syngas, a valuable mixture of hydrogen (H$_2$) and carbon monoxide (CO).[1] The resulting syngas has a wide range of applications, including fuel production, power generation, and the production of various chemicals such as methanol, ammonia, and synthetic fuels. Owing to this ability, DRM has the potential to address pressing energy and environmental challenges while contributing to a more sustainable and low-carbon future.

The reaction involves the cleavage of strong C–H and C≡O bonds, requiring the use of catalysts to lower the activation energy. By choosing suitable catalyst materials and optimizing their properties through doping, it is possible to enhance the catalytic activity and selectivity for CH$_4$ and CO$_2$ decomposition. Ni metallic nanoparticle is a commonly used catalyst in DRM due to its high activity and cost-effectiveness.[2,3] The doping of other metals onto Ni (M–Ni) has been shown to improve catalytic performance and alleviate coke formation issues associated with Ni catalysts.[4–7] Our previous studies have demonstrated the effectiveness of doping other metals onto Ni in enhancing catalytic performance and mitigating coke formation.[8,9]

Understanding the factors that affect the reaction barriers is crucial in the context of CH$_4$ and CO$_2$. This knowledge enables rapid assessment and facilitates accelerated catalyst screening with reduced computational demands. For an elementary reaction, the Brønsted–Evans–Polanyi (BEP) correlation (eqn (1)) directly establishes a linear correlation between the change in the reaction barrier ($E_{act}$) to the change in reaction energy, $\Delta E_{rxn}$, for different surfaces:[10–13]

$$E_{act} = \alpha \Delta E_{rxn} + \beta \qquad (1)$$

This relationship is an empirical linear free-energy relation consistent with transition-state theory (TST) under

[a] *Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02467, USA. E-mail: lucas.bao@bc.edu*

[b] *Department of Chemistry and Biochemistry, Texas Woman's University, Denton, Texas, 76204, USA. E-mail: slin6@twu.edu*

† Equal contribution.

**7358** | *Catal. Sci. Technol.*, 2025, **15**, 7358–7370

This journal is © The Royal Society of Chemistry 2025

Hammond-like and small-entropy/prefactor-variation assumptions across a homologous reaction family, and it is a valuable tool for approximating activation energies without explicitly locating transition states (TSs). This greatly simplifies the evaluation process and has been widely applied in surface elementary reaction steps.[10,13–15] Importantly, the BEP relationship also provides the kinetic foundation of volcano plots by linearly linking activation energy to reaction energy, which in turn often scales with adsorption energy. Volcano plots emerge from these linear relations, capturing the trade-off between overly weak and overly strong binding.[11] Consequently, BEP scaling relationships not only enable simplified catalyst screening but also define intrinsic limitations in catalytic performance.

However, the applicability of BEP relationships and volcano plots becomes limited in the case of doped nanoparticle catalysts. For example, Nwaokorie and Montemore illustrated the breakdown of linear scaling relations between adsorption energies and reaction barriers in steam reforming of methane that attempt to circumvent the traditional trade-offs imposed by these relations.[16] On alloy surfaces, such deviations often arise from variations in coordination environments and site-specific bonding, which disrupt the uniform electronic behavior assumed in classical BEP scaling. For instance, intermediates like $CH_3$ and C may adsorb on different atomic sites, each with distinct reactivity. In certain alloys, $CH_3$ preferentially adsorbs on a specific metal dopant, while C interacts with multiple metal atoms, leading to stronger $CH_3$ adsorption without over-stabilization of C.

A recent DFT study by Zhou *et al.* supports this observation through an investigation of C–H bond activation in methane over 12 single-atom alloy (SAA) catalysts with Cu and Ag as hosts.[17] The authors reported substantial deviation from the linear BEP correlation and revealed that the breakdown is mainly due to the differentiated adsorption behaviors of the H atom and the $CH_3$ group products across different SAAs.

These findings highlight the potential of engineering nanoparticle catalysts to break the linear BEP relationship, thereby offering a promising strategy to surpass the intrinsic activity limits defined by volcano plots. Nevertheless, the fundamental mechanistic origins of BEP breakdown remain insufficiently understood, warranting further systematic investigation. Given the complexity of reaction mechanisms and the interplay of surface phenomena, accurate prediction in such systems demands an advanced computational framework beyond conventional linear scaling.

Machine learning (ML) is increasingly utilized in materials science as a powerful tool for identifying quantitative relationships between reaction barriers and properties, accelerating materials design. However, conventional ML models are often criticized for their "black box" nature and overfitting, as they tend to generate highly complex mathematical expressions, making it challenging to derive straightforward physical interpretations from the model.

As an interpretable alternative, symbolic regression (SR) simultaneously identifies the optimal mathematical expression that describes underlying physical relationships, thus offering direct insights for materials design.[18] For instance, Weng and co-workers used SR to derive the expression that is used to guide the design of oxide perovskite catalysts for the oxygen evolution reaction (OER).[19] From a feature set containing electronic parameters as well as structural parameters of 23 selected oxide perovskites, SR identified that the octahedral factor over the tolerance factor provided the best balance between complexity and accuracy among 43 200 000 candidates. The descriptor exhibited a linear correlation with the potential $V_{RHE}$, revealing that the activity of the OER is related to the structural parameters of catalysts. Based on these insights, five new oxide catalysts were synthesized, and four of them exhibited higher experimental activities than previously reported perovskite oxide catalysts.

In addition to SR, other models such as the least absolute shrinkage and selection operator (LASSO) regression[20] have also been employed to derive interpretable correlations. For example, the existing linear scaling relationships between the adsorption energy of metal adatoms on various supports and the metal adatom's oxide formation enthalpy suggest that the interactions of adatoms with surface oxygen atoms largely determine the strength of their interfacial bond with the metal oxides. O'Connor and co-workers utilized LASSO regression to identify key property descriptors that predict the interaction strengths between 13 single metal atoms and 7 oxide supports.[21] Their analysis revealed that the ratio of the oxide formation enthalpy of the metal adatom to the oxygen vacancy formation energy consistently appeared in the predicted models with low RMSE, indicating its strong correlation. This descriptor clearly suggests that metal atoms exhibit stronger binding to surfaces when they have higher oxide formation enthalpies and weaker binding to surfaces with increased oxygen vacancy formation energies.

A comparative study of multiple ML models was presented by Liu and co-workers, who aimed to predict the adsorption energies of CO on nine different transition metal planes.[22] Four models, including adaptive boosting (Ada), decision tree (DT), LASSO regression, and random forest (RF), were trained with 20 selected features. Among them, RF exhibited the best performance with an $R^2$ value of 0.885 and identified the valency of adsorbate and boiling point of the catalyst element as the most important features that affect the adsorption. Feature engineering, which involved identifying highly correlated feature pairs and removing low-importance features, only slightly improved the $R^2$ of the RF model by 0.01. In comparison, the utilization of SR enabled the construction of a new composite feature derived from three less important variables ($\geq 10$ among all features), which are the surface energy of the substrates ($E_s$), the coordination number of the adsorption sites (CN), and the interplanar crystal spacing ($D$). The resulting expression $(CN + D)/E_s$ replaced the original variables in the RF model and improved the $R^2$ to 0.921, while the mean absolute error (MAE) decreased by 0.048 eV.

These examples highlight the potential of ML models for building accurate and interpretable expressions. In this study, we employed symbolic regression in Python (PySR)[23] to derive mathematical expressions that reveal the underlying complex relationships between the target properties and the input features of M–Ni catalysts.

The features of 12 well-studied dopants, including Ni (the prinstine system), Cr, Fe, Mn, Co, Pt, Rh, Ru, Ti, V, Y, and Zn, obtained from quantum-level calculations, were used to train PySR models for regression, as shown in the dataset box in Fig. 1. PySR identifies one or two of the most important features and effectively uses simple formulas to correlate them with (i) the overall reaction energy of $CH_4$ decomposition (excluding the coke formation step, *i.e.*, the 4th step initiated with C(ads) formation), which is denoted as $\Delta E_{1-3}(CH_4)$: $CH_4(ads) \rightarrow CH(ads) + 3H(ads)$, (ii) the rate-determining barriers for the $CH_4$ decomposition ($V^{\ddagger}_{rds}(CH_4)$), (iii) the $CO_2$ direct decomposition energy ($\Delta E_{CO_2}$), and (iv) its barriers $\left(V^{\ddagger}_{CO_2}\right)$. The inner loop of PySR follows a classic evolutionary algorithm, where each population evolves through tournament selections, mutations, and crossovers for generating new formulas. The process begins with the generation of an initial population with selected features as variables, followed by the random selection from the population. In each tournament, the parent formulas are evaluated based on the fitness function, and the formula with the lowest loss undergoes a random crossover, which exchanges variables with other individuals, and mutation, which alters operators, to create offspring. The parents and the offspring then compete, and the formula that best balances the models' complexity and loss at a certain complexity is saved. If the termination criteria are met, the saved formula is printed along with its MAE and complexity. Otherwise, a copy of the winner formula replaces the weakest member in the original population, ensuring iterative improvement through selection and variation. Subsequently, the ranking metric, normalized discounted cumulative gain (nDCG),[24] kinetic Monte Carlo (kMC) simulation, and the interpretation of the formula were used to assess the most

meaningful formula generated by PySR. The selected formula is then chosen to screen 12 candidate dopants from untrained transition metals, including Sc, Mo, Zr, Nb, Tc, Pd, Cd, Hf, Ta, W, Os, and Ir, followed by the validation through DFT calculations.

We model and compare the reaction kinetics of DRM using DFT-calculated energy, PySR-predicted nonlinear correlation, and BEP-based linear correlation. By comparing the results from these approaches, we showcase the critical importance of accounting for nonlinear effects in scaling relations for accurate prediction of reaction kinetics and efficient catalyst design. The PySR framework not only enabled us to uncover the intricate correlations between these variables but also offers a predictive tool in catalysis research, reactor design, and the development of sustainable chemical processes.

## 2. Computational methods

### 2.1 DFT calculations

We employed the Vienna *Ab initio* Simulation Package (VASP, version 5.4)[25] along with the projector-augmented wave (PAW) method[26] and the generalized gradient approximation (GGA) in the form proposed by Perdew, Burke, and Ernzerhof (PBE) to gain physical insights into the computed reaction energetics.[27] To account for van der Waals attraction,[28,29] we incorporated Grimme's D3 correction with the Becke–Johnson (BJ) damping function.[30]

For single-atom doped M–Ni supercell optimization consisting of 125 atoms in five layers, we used a $2 \times 2 \times 1$ Monkhorst–Pack $k$-point mesh.[31,32] All computations were spin-polarized with an energy cutoff of 400 eV for the plane wave basis set. Local minimum structure was achieved using a force convergence tolerance of 0.023 eV Å$^{-1}$, while transition-state searches employed a tolerance of 0.05 eV Å$^{-1}$. To prevent unphysical interactions between adjacent layers, the M–Ni surfaces were oriented in the $xy$-plane with the $z$-direction perpendicular to the layer plane, adopting a vacuum spacing of over 15 Å and applying a $z$-direction dipole correction.[33]
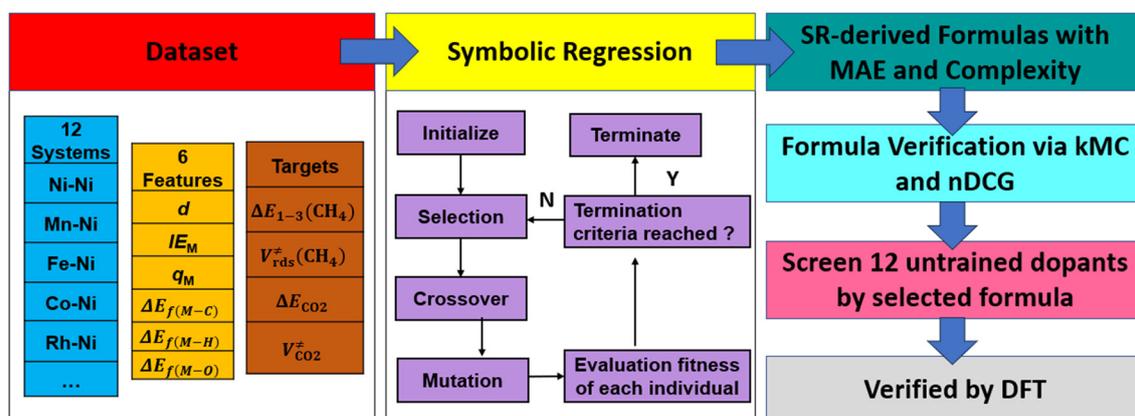


**Fig. 1** Workflow of the symbolic regression.

The adsorption energy of $CH_4$ ($\Delta E_{CH_4(ads)}$) was calculated using eqn (2),

$$\Delta E_{CH_4(ads)} = E_{CH_4(ads \; on \; M-Ni)} - E_{M-Ni} - E_{CH_4(GS,G)} \quad (2)$$

where $E_{CH_4(ads \; on \; M-Ni)}$ is the electronic structure energy of the adsorbed $CH_4$ on M–Ni, $E_{M-Ni}$ is the energy of a clean M–Ni slab, and $E_{CH_4(GS,G)}$ is the ground-state (GS) energy of an isolated $CH_4$ in the gas (G) phase.

The $CH_4$ dissociation energy ($\Delta E_{CH_4}$) is computed by

$$\Delta E_{CH_4} = E_{CH_4(ads \; on \; M-Ni)} - E_{CH_3(ads \; on \; M-Ni)} - E_{H(ads \; on \; M-Ni)} \quad (3)$$

To determine transition states and reaction barriers ($V^{\ddagger}$), we employed the climbing-image nudged elastic band (CI-NEB) method.[34] To ensure the validity of the transition states, we further computed vibrational normal modes with a convergence tolerance of $10^{-6}$ eV, confirming the presence of a single imaginary mode. Detailed reaction profiles for all elementary steps are provided in Tables S1 and S2.

## 2.2 CDFT-CI calculations

To investigate the breakdown of the BEP linear correlation, we employed constrained DFT with configuration interactions (CDFT-CI) to generate the diabatic potential energy curves and the corresponding electronic couplings of the rate-determining step across various dopant systems.[35–37] All geometries along the rate-determining dissociation pathway on the M–Ni slab were obtained from CI-NEB. For each image, we extracted the atomic positions from the optimized periodic structures, and a $5 \times 5$ surface cluster was constructed by selecting only the top atomic layer.

Although the subsurface layers may contribute to reactivity,[38,39] we chose to use cluster-based models for performing the CDFT-CI analysis, aiming to elucidate qualitative diabatic energy trends across dopants. To evaluate the validity of our cluster-based approach, single-point Kohn–Sham density functional theory (KS-DFT) calculations were performed using Q-Chem,[40] employing the PBE-D3BJ functional with def2-TZVP basis set,[41] to determine the most favorable spin state by using the barrier and reaction energies from periodic calculations as the reference. As shown in Tables S1 and S6, the trends predicted by KS-DFT using the top-layer cluster model are in strong agreement with those obtained from periodic VASP calculations. This consistency suggests that the top-layer cluster is sufficient to capture the key local electronic reorganizations responsible for BEP deviations. In the CDFT-CI calculations, the total spin multiplicity was fixed between the reactant state and the product states. Although the constraint eliminates the spin-crossing effects observed in the periodic calculations, the fixed-spin diabatic model provides a controlled framework to isolate the effect of geometric and electronic structure variations on the BEP correlation. DFT-D3(BJ) dispersion correction and def2-ECP effective core potential (ECP) were

used to account for the effects on heavy atoms. Due to the small HOMO–LUMO gaps for the metal slab, the relaxed constraint algorithm (RCA)[42] and level-shifting method[43,44] were applied to facilitate SCF convergence. Population analyses were carried out using the Becke scheme to obtain atomic charges and spin densities, while natural bond orbital (NBO)[45] analysis was conducted to further investigate changes in the valence orbital character along the reaction pathway.

The Becke charges and spins from the initial and final states along the reaction path were compared to compute atomic charge and spin differences. Based on these differences, the system was partitioned and reordered into three fragments, which are the donor region of the slab, the acceptor region of the slab, and the molecular moiety ($CH_4$ or $CH_3$). We tested alternative fragmentation schemes, including slab/dopant/$CH_4$ and spin- or charge-constrained multi-atom partitions. These either failed to converge because the constraints were unstable or produced unphysical diabatic profiles without a crossing. Increasing the number of fragments also substantially raised computational cost and hindered convergence. We therefore chose the donor/acceptor/$CH_4$ partitioning for its physical relevance and numerical stability.

The diabatic states were constructed by constraining the total charge and spin of the three fragments defined above using the Becke partitioning scheme, and fragment populations were taken directly from the KS-DFT results. The CI was performed between the two constrained diabatic states to compute state energies and the state couplings along the reaction path.

## 2.3 Symbolic regression

The PySR script is provided in the SI. Five populations were generated, each containing 27 individuals. To prioritize physical interpretability over mathematical complexity, only the basic arithmetic operators (+, −, ×, ÷) were used, which also reduces the risk of overfitting. In PySR, the complexity is a numerical measure of how complicated a symbolic expression is. Each operation is assigned a predefined weight, and the total complexity is the sum of the weights of all operations and variables in the expression. By default, most operators and functions count as 1, although this can be customized. Lower complexity corresponds to simpler, more interpretable formulas, while higher complexity indicates more elaborate expressions.

In each tournament, 15 formulas were considered, and the MAE was used as the loss function in this study:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| E_{DFT} - E_{PySR} \right| \quad (4)$$

where $E_{DFT}$ refers to the results (reaction energies or barriers) from DFT, and $E_{PySR}$ refers to the energy predicted by PySR. The number of nodes in an expression tree was set to 30 to control the complexity of the equation. As shown in Fig. S1,

This journal is © The Royal Society of Chemistry 2025

*Catal. Sci. Technol.*, 2025, **15**, 7358–7370 | 7361

the loss function for $V_{rds}^{\ddagger}$ of $CH_4$ has reached a plateau once the complexity is higher than 10, indicating a sufficient setting. After the variation, candidate models were scored based on the balance between loss and complexity, where expression has a much better loss at a slightly higher complexity, resulting in a higher score. The termination of PySR is solely determined by the total iteration number, which was set to 50 000 in this study.

The features selected for training the PySR in this study were based on the Pearson correlation coefficient between each pair of features from our previous work.[9] These features along with their corresponding physical interpretations are summarized in Table 1.

The X–M formation energy $\Delta E_{f(M–X)}$ (X = C, O, H) of the dopant metal, which indicates the thermodynamic feasibility for M to bind with atom X, can be calculated by

$$\Delta E_{f(M–X)} = (E_{X–C} - n_X E_X - n_M E_M)/(n_X + n_M) \quad (5)$$

where $E_{X–C}$ is the electronic structure energy of the most stable M–X bulk structure, $E_X$ and $E_M$ are the averaged electronic structure energy of a single atom in their corresponding surface, and $n_X$ and $n_M$ are the number of atoms in the computational unit cell.

The values of all features across various dopants are provided in the SI (see Table S8) and were computed at the PBE/6-311+G** level of theory[27,49,50] using the Gaussian 16 program.[51]

**2.4 Expression validation**

The ideal discounted cumulative gain (IDCG)[24] and nDCG were used to evaluate the predicted ranking from PySR using

$$(I)DCG = \sum_{i=1}^{p} \frac{rel_{i'}}{\log_2(i+1)} \quad (6)$$

$$nDCG = \frac{DCG}{IDCG} \quad (7)$$

where $p$ is the number of terms for ranking; $i'$ is the position number of M–Ni in the DFT ranking results, and $i$ in IDCG is also the DFT position number, but it is the predicted position number from PySR in DCG; $rel_{i'}$ is the relevance value of M–Ni, which is defined as $p - i' + 1$. If a promising catalyst appears lower in a predicted ranking, the nDCG value decreases.

**Table 1** Features used in this study and their meaning

| Features | Explanations |
|---|---|
| $d$ | The number of d electrons in the dopant (M) |
| $q_M$ | The atomic Bader charges (in $e$) of the dopant in the clean surface of M–Ni (without adsorbates)[46,47] |
| $IE_M$ | The first ionization energy of the dopant (in eV) |
| $\Delta E_{f(M–C)}$ | Carbide formation energy of the dopant (in eV) |
| $\Delta E_{f(M–H)}$ | Hydride formation energy of the dopant (in eV) |
| $\Delta E_{f(M–O)}$ | Oxide formation energy of the dopant[21,48] (in eV) |

To solve the rate equations of the kinetic model, we employed the Gillespie kinetic Monte Carlo (kMC) method,[52] revealing substantial discrepancies between linear and nonlinear correlation models. The rate constants associated with $CH_4$ and $CO_2$ dissociations are obtained by

$$k^{TST} = \frac{k_B T}{h} e^{-\frac{\Delta G^{\ddagger}}{k_B T}} \quad (8)$$

where $k^{TST}$ is the rate constant predicted by transition state theory (TST), $k_B$ is Boltzmann's constant, $T$ is set to 1000 K, which is a typical condition of DRM,[53] $h$ is Planck's constant, and $\Delta G^{\ddagger}$ is the thermal free energy barrier. We approximate the free energy of activation as the barrier height herein to reveal the trend of the catalytic kinetics. This assumption is reasonable because our analysis targets dopant-dependent trends rather than absolute values.

The kMC simulations were performed using GillesPy2,[54] with a total simulation time of 10 ns, at which point the concentration of all species had reached steady state.

# 3. Results and discussion

## 3.1 Correlation in $CH_4$ decomposition

We initiated our investigation by examining the correlation between the adsorption energies of intermediates ($\Delta E_{ads}$) in 12 single-doped Ni alloys and the rate-determining step (RDS) barriers ($V_{rds}^{\ddagger}$) for $CH_4$ decomposition, as shown in Table 2 and Fig. 2. Given that the dissociation of H from $CH_4(ads)$ is the RDS for most of the dopants, except for Pt and Zn, where the dominant step is the dissociation of H from $CH_3(ads)$, as shown in Table S1, we initially investigated $\Delta E_{CH_4(ads)}$ and $\Delta E_{CH_3(ads)}$. As depicted in Fig. 2a and b, neither $\Delta E_{CH_4(ads)}$ nor $\Delta E_{CH_3(ads)}$ could differentiate the decomposition barriers with an $R^2$ value of 0.0024 and 0.1745, respectively.

Upon plotting $V_{rds}^{\ddagger}$ against $\Delta E_{CH_4(ads)}$ and $\Delta E_{CH_3(ads)}$ we observed a relatively compact clustering of points within the range of −0.41 eV to −0.18 eV for $\Delta E_{CH_4(ads)}$ and −2.54 eV to −2.00 eV for $\Delta E_{CH_3(ads)}$. We turned our attention to correlating the adsorption energies of $\Delta E_{C(ads)}$ with $V_{rds}^{\ddagger}$, which has been reported to obey the scaling relation on pure metal

**Table 2** Adsorption energies and the forward reaction barriers of the rate-determining step for $CH_4$ decomposition on various doped Ni surfaces (unit: eV)

| M–Ni | $\Delta E_{CH_4(ads)}$ | $\Delta E_{CH_3(ads)}$ | $\Delta E_{C(ads)}$ | $\Delta E_{CH_4(ads)–CH_3(ads)}$ | $V_{rds}^{\ddagger}$ |
|---|---|---|---|---|---|
| Ni | −0.20 | −2.27 | −7.08 | 2.07 | 0.78 |
| Cr | −0.35 | −2.52 | −7.19 | 2.17 | 0.73 |
| Fe | −0.22 | −2.27 | −7.06 | 2.05 | 0.87 |
| Mn | −0.25 | −2.30 | −6.88 | 2.05 | 0.84 |
| Co | −0.21 | −2.29 | −7.11 | 2.08 | 0.83 |
| Pt | −0.18 | −2.00 | −7.06 | 1.81 | 0.90 |
| Rh | −0.24 | −2.23 | −7.18 | 1.98 | 0.66 |
| Ru | −0.26 | −2.34 | −7.40 | 2.07 | 0.57 |
| Ti | −0.41 | −2.48 | −7.13 | 2.08 | 0.86 |
| V | −0.32 | −2.54 | −6.94 | 2.22 | 0.86 |
| Y | −0.35 | −2.09 | −7.31 | 1.73 | 1.43 |
| Zn | −0.18 | −2.24 | −7.07 | 2.06 | 1.30 |
| $R^2$ | 0.0024 | 0.1745 | 0.0007 | 0.2689 | |

**7362** | *Catal. Sci. Technol.*, 2025, **15**, 7358–7370

This journal is © The Royal Society of Chemistry 2025
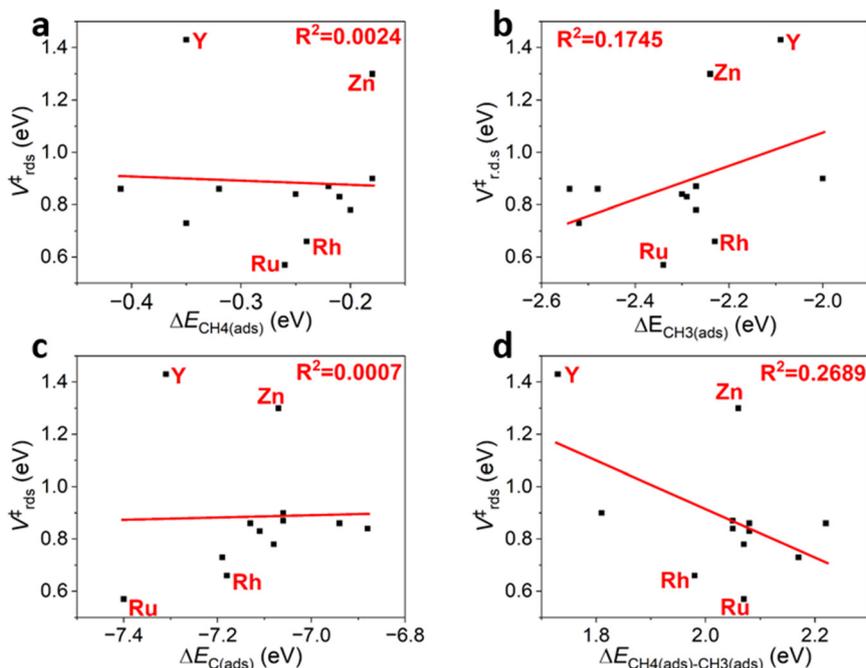
**Fig. 2** Scatter plot of CH$_4$ direct decomposition barriers ($V^{\ddagger}_{rds}$) vs. adsorption energy of (a) CH$_4$ ($\Delta E_{CH_4(ads)}$), (b) CH$_3$ ($\Delta E_{CH_3(ads)}$), (c) C ($\Delta E_{C(ads)}$) and (d) the adsorption energy difference between CH$_4$ and CH$_3$ ($\Delta E_{CH_4(ads)-CH_3(ads)}$) on M–Ni. The four dopants showing the largest deviation from the best-fitting line are highlighted in red.

surfaces.[16] Although the resulting distribution was relatively scattered from −7.40 eV to −6.88 eV, we found an even weaker linear correlation with an $R^2$ value of 0.0007 eV (Fig. 2c). Given a poor correlation between individual intermediate adsorption energies and $V^{\ddagger}_{rds}$, we explored the energy difference between $\Delta E_{CH_4(ads)}$ and $\Delta E_{CH_3(ads)}$ ($\Delta E_{CH_4(ads)-CH_3(ads)}$). Yet this approach failed to distinguish between the different M–Ni alloys in terms of CH$_4$ decomposition barriers, with an $R^2$ value of 0.2689 (Fig. 2d). We further tested the square and cubic adsorption energy of $\Delta E_{CH_3(ads)}$ and $\Delta E_{CH_4(ads)-CH_3(ads)}$ due to their relatively large $R^2$. While the adsorption energy for M–Ni alloys began to separate more as the degree of these expressions increased,

linear fitting remained inadequate, with the highest $R^2$ value reaching only 0.2559 (Table S3).

We sought to identify a BEP correlation between the reaction energy and the $V^{\ddagger}_{rds}$, as reported in Table S4, to assess the capability of a dopant to promote reaction activation. Our screening results reveal that dopants that show large deviation from BEP scaling in the rate-determining step (Fig. 3a) also exhibit significant deviation in the overall dissociation energy (Fig. 3b). Notably, in both absorption energy and reaction energy analysis, Ru, Rh, Y, and Zn show the largest deviation from the expected BEP-derived linear trend. Specifically, Ru and Rh are below the best-fitting line, indicating an enhanced catalytic activity with
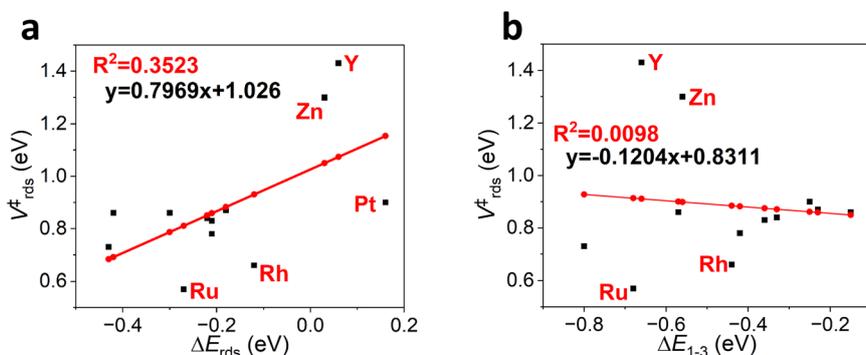


**Fig. 3** Scatter plot and the best fitting line of (a) the rate-determining step barriers of the reaction ($V^{\ddagger}_{rds}$) vs. the reaction energy of the rate-determining step ($\Delta E_{rds}$). (b) The rate-determining step barriers of the reaction ($V^{\ddagger}_{rds}$) vs. the overall reaction energy of CH$_4$ decomposition (steps 1–3) ($\Delta E_{1-3}$(CH$_4$)). Black squares: DFT results; red dots: linear regression (LR) result based on BEP correlation; red color elements: the elements exhibit the largest deviation from the DFT results.

lower actual activation barriers than predicted by the BEP linear relations. These observations suggest that violations of the BEP relationship in the rate-determining step often coincide with deviations in the overall $CH_4$ dissociation energy, indicating a possible coupling between kinetic and thermodynamic scaling violations.

As all the elementary-step dissociation barriers of $CH_4$ were higher than those of $CO_2$ (Table S2), this demonstrates that the dissociation of $CH_4$ should be the RDS on M–Ni. The detailed information for $CO_2$ is listed in the SI. Similarly to $CH_4$ decomposition, no good correlation was observed for $CO_2$ decomposition. Among the tested descriptors, CO adsorption energy showed the most notable degree of correlation, with an $R^2$ value of only 0.58, indicating the complexity of this relationship, as shown in Table S5.

### 3.2 Breakdown of the BEP principle

To gain a deeper insight into the breakdown of BEP linear correlation, we first investigated the geometric differences induced by various dopants, inspired by the study of Nwaokorie and Montemore,[16] which states that the differences in coordination between adsorbed intermediates in SAAs could break the scaling relation. In particular, we investigated the adsorption configurations of Mn–Ni, Fe–Ni, Co–Ni, and Rh–Ni, as these dopants are adjacent in the periodic table and share the same RDS. Moreover, Mn, Fe, and Co are close to those predicted by the BEP relationship, while the energy of Rh significantly deviates, suggesting a breakdown of BEP scaling (Fig. 3).

As shown in Fig. 4a–c, Mn, Fe, and Co show a similar structure for $CH_x$ upon $CH_4$ dissociation, where at the

reactant state, $CH_4$ is located above the dopant (d1), and the $CH_3$ fragment is tri-coordinated with two Ni atoms and one dopant atom at the product state. However, at the product state of Fe and Co, the H atom is bound to three Ni atoms with a distance of about 3.0 Å away from the dopant (d2), while for Mn, the H atom is tri-coordinated with two Ni atoms and one Mn atom at a distance of 2.91 Å. Despite the very close $V_{rds}^{\ddagger}$ among Mn (0.84 eV), Fe (0.87 eV), and Co (0.83 eV), notable geometric distinctions are observed. Interestingly, Rh, which exhibits a significantly lower $V_{rds}^{\ddagger}$ of 0.66 eV, displays a similar binding geometry and absorption site to Fe and Co, as shown in Fig. 4d. These observations suggest that the adsorption site geometry alone does not fully account for the violation of the BEP rules.

To further examine the factors underlying the bonding pattern variations, we performed CDFT-CI calculations on four systems. By imposing constraints on the wavefunction (spin and charge distribution), we constructed diabatic states to gain insight into the electronic structure. Along the same diabatic state, the wavefunction character remains unchanged over the entire reaction coordinate. Two diabatic states were generated: one by constraining the wavefunction at the reactant state while varying the geometry toward the product, and the other by constraining the wavefunction at the product state while varying the geometry from reactant to product. Adiabatic states, in contrast, fully relax the wavefunctions to their respective eigen-solutions and can thus be viewed as constructed from diabatic states. Generally speaking, the crossing point of the two diabatic states marks a significant change in wavefunction character at a specific geometry (an image) on its adiabatic state.
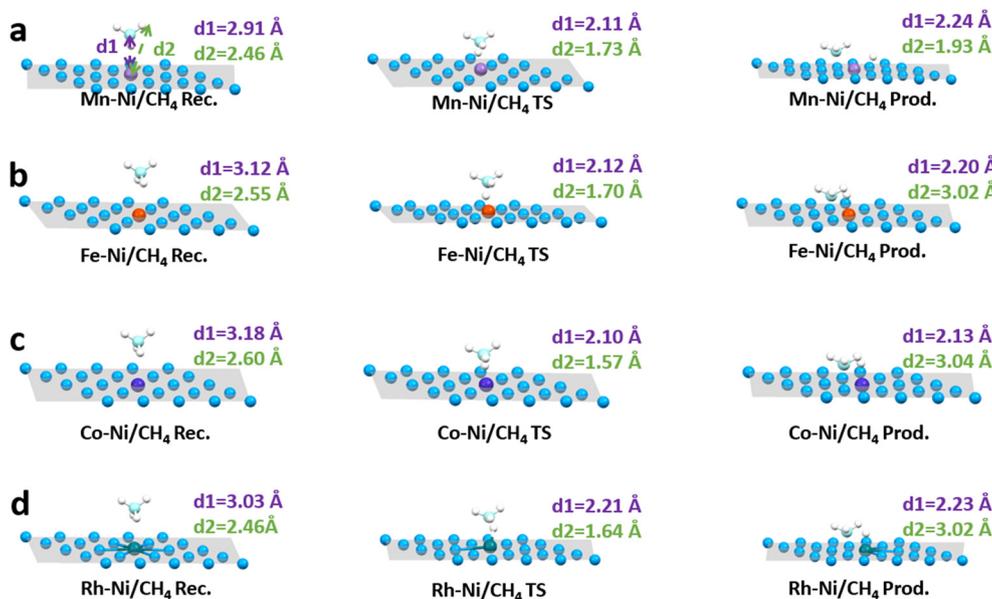


**Fig. 4** The adsorption geometries of $CH_4$, $CH_3\cdots H$ and $CH_3$ on (a) Mn–Ni, (b) Fe–Ni, (c) Co–Ni, and (d) Rh–Ni, d1 represents the distance (in angstrom) between the C atom and the dopant; d2 represents the distance between the dissociated H atom and the dopant. Element colors: blue – Ni, cyan – C, white – H, red – Fe, purple – Mn, violet – Co, green – Rh, grey – Ni plane.
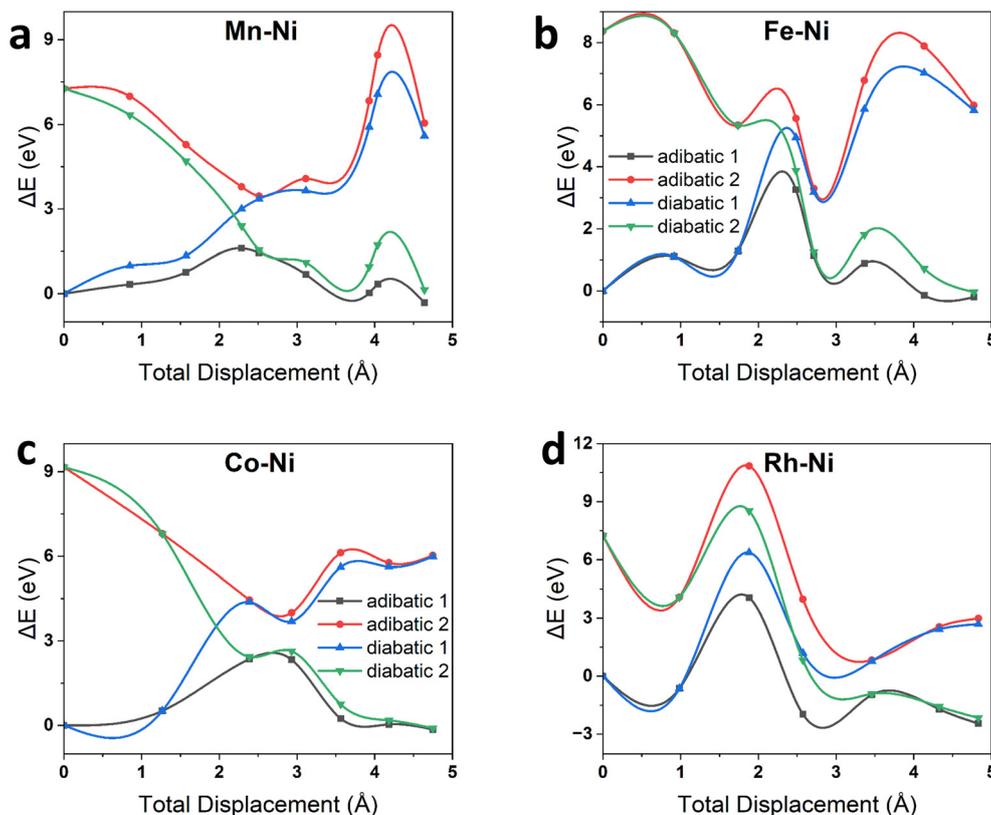
**7364** | *Catal. Sci. Technol.*, 2025, **15**, 7358–7370

This journal is © The Royal Society of Chemistry 2025

**Fig. 5** The CDFT-CI results of (a) Mn–Ni, (b) Fe–Ni, (c) Co–Ni and (d) Rh–Ni. State color scheme: black – ground adiabatic state, red – excited adiabatic state, blue – diabatic state 1 (constraining on reactant electronic structure), green – diabatic state 2 (constraining on product electronic structure).

The resulting diabatic and adiabatic potential energy surfaces reveal distinct crossing behaviors depending on the dopant. For Mn, Fe, and Co (Fig. 5a–c), the diabatic crossing point appears slightly before or coincident with the ground-state adiabatic transition state (TS) at image 4, image 4, and image 3, respectively (the starting image is labeled as image 1). In contrast, for the off-line metal Rh (Fig. 5d), the adiabatic TS occurs at image 3, while the diabatic crossing point is shifted beyond the TS, appearing near image 4. A similar behavior was observed in Ni–Ni and Ru–Ni (Fig. S2), where Ni follows the BEP relationship while Ru deviates from it.

NBO analysis provides a complementary perspective on electronic structure changes along the reaction coordinate, revealing detailed variations in valence orbital occupations for selected atoms in the different doped-Ni systems (Table S7). For Mn–Ni, the NBO electron configuration of the dopant and the dissociated H atom exhibits abrupt occupancy changes at the TS (Table 3), with the dopant's 4s and 3d orbitals' occupations changing by ~0.1 and ~0.2, respectively, and the H 1s orbital increasing by ~0.2. In contrast, the corresponding changes before the TS (image 3) and after the TS (image 5) are much smaller. This observation, consistent with the localized nature of 3d

**Table 3** Electron configurations of the NBO valence orbitals of selected atoms in various doped-Ni systems

|  |  | Image 1 | Image 2 | Image 3 | Image 4 | Image 5 | Image 6 | Image 7 | Image 8 | Image 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| Mn | Dopant | $4s^{0.88}$ | $4s^{0.86}$ | $4s^{0.83}$ | $4s^{0.70}$ | $4s^{0.70}$ | $4s^{0.66}$ | $4s^{0.64}$ | $4s^{0.62}$ | $4s^{0.62}$ |
|  |  | $3d^{5.82}$ | $3d^{5.86}$ | $3d^{5.91}$ | $3d^{6.09}$ | $3d^{5.99}$ | $3d^{5.98}$ | $3d^{5.92}$ | $3d^{6.00}$ | $3d^{5.91}$ |
|  | H | $1s^{0.78}$ | $1s^{0.78}$ | $1s^{0.80}$ | $1s^{1.02}$ | $1s^{1.20}$ | $1s^{1.20}$ | $1s^{1.20}$ | $1s^{1.17}$ | $1s^{1.20}$ |
| Fe | Dopant | $4s^{0.90}$ | $4s^{0.89}$ | $4s^{0.85}$ | $4s^{0.71}$ | $4s^{0.70}$ | $4s^{0.66}$ | $4s^{0.65}$ | $4s^{0.74}$ |  |
|  |  | $3d^{6.74}$ | $3d^{6.76}$ | $3d^{6.81}$ | $3d^{7.02}$ | $3d^{6.97}$ | $3d^{6.91}$ | $3d^{6.87}$ | $3d^{6.77}$ |  |
|  | H | $1s^{0.78}$ | $1s^{0.78}$ | $1s^{0.81}$ | $1s^{1.05}$ | $1s^{1.21}$ | $1s^{1.21}$ | $1s^{1.23}$ | $1s^{1.19}$ |  |
| Co | Dopant | $4s^{0.91}$ | $4s^{0.87}$ | $4s^{0.67}$ | $4s^{0.66}$ | $4s^{0.63}$ | $4s^{0.68}$ | $4s^{0.72}$ |  |  |
|  |  | $3d^{7.92}$ | $3d^{7.96}$ | $3d^{8.23}$ | $3d^{8.19}$ | $3d^{8.09}$ | $3d^{7.99}$ | $3d^{7.98}$ |  |  |
|  | H | $1s^{0.78}$ | $1s^{0.80}$ | $1s^{1.01}$ | $1s^{1.21}$ | $1s^{1.22}$ | $1s^{1.18}$ | $1s^{1.19}$ |  |  |
| Rh | Dopant | $5s^{0.83}$ | $5s^{0.78}$ | $5s^{0.68}$ | $5s^{0.69}$ | $5s^{0.64}$ | $5s^{0.60}$ | $5s^{0.67}$ |  |  |
|  |  | $4d^{8.45}$ | $4d^{8.53}$ | $4d^{8.66}$ | $4d^{8.58}$ | $4d^{8.58}$ | $4d^{8.57}$ | $4d^{8.47}$ |  |  |
|  | H | $1s^{0.78}$ | $1s^{0.79}$ | $1s^{0.90}$ | $1s^{1.14}$ | $1s^{1.15}$ | $1s^{1.18}$ | $1s^{1.20}$ |  |  |

orbitals and the overlap of diabatic and adiabatic transition points, indicates that in these dopants the electronic configuration is highly sensitive to geometric distortions along the reaction pathway. A similar behavior is found for Fe and Co, where electron configuration changes at the TS are robust.

In contrast, Rh shows a more gradual redistribution, with the dopant 4d orbital occupations evolving smoothly from 8.45, 8.53, 8.66, to 8.58 from image 1 to 4 (Table 3), reflecting greater delocalization. Notably, the dissociated H atom on Rh undergoes its most significant orbital occupancy change only after passing the adiabatic TS at image 4, as the 1s orbital increases from 0.90 to 1.14. This change is larger than the corresponding variations at the TS (0.113) and at image 5 (0.009) (Table S7). The delayed occurrence of this major electronic reorganization explains why the diabatic crossing is shifted beyond the adiabatic TS. These results underscore a key distinction: while 3d dopants exhibit geometry-driven electronic changes, 4d dopants display electronically driven reorganization. Together, these findings indicate that BEP deviations arise not only from structural differences but also from fundamental differences in the electronic nature and reactivity of the dopants.

### 3.3 PySR-discovered analytic nonlinear correlations

The mathematical formulas that best balance the accuracy and simplicity among all candidates are listed in Table 4. $\Delta E_{1-3}$ is inversely proportional to the sum of $\Delta E_{f(M-H)}$ and $d$ (Fig. 6a), with a MAE less than 0.1 eV and moderate complexity. The descriptor reveals that a dopant with a sum of $\Delta E_{f(M-H)}$ and $d$ less than about −4 always exhibits a lower $\Delta E_{1-3}$ than that with a larger summation, indicating that lower hydride formation energy and fewer d electrons are generally desirable for the reaction thermodynamically. However, excessive M–H interaction prevents the spillover of H atoms, blocking M–C bond formation and preventing further C–H bond cleavage, resulting in an increase in $\Delta E_{1-3}$. On the other hand, a larger deviation between the d-orbital characteristics of the dopant metal and Ni typically results in a more positive $q_M$ (Table S8), reflecting stronger perturbation of the surface electronic structure. This property is captured by the following formula:

$$\Delta E_{1-3} = -0.436 \text{ eV} + 0.00687 \text{ eV}/(0.596 - q_M/e) \quad (9)$$

which shows that a dopant with its $q_M$ higher than ~0.60 $e$, which makes the 2nd term negative and further lowers the reaction energy, outperforms those below this threshold.

Notably, dopants with higher $q_M$ tend to exhibit lower $\Delta E_{f(M-H)}$, which again prevents the spillover of H atoms. It is worth noting that eqn (9), with a complexity value of 7, significantly elevated the MAE by 45% compared to the optimal analytic formula (Table S9), indicating that $q_M$ alone cannot predict the reaction energy, even though it is still able to provide valuable insight. Ultimately, Cr and Ru, with a moderate number of d electrons, best balance $\Delta E_{f(M-H)}$ and $q_M$, and thus exhibited the lowest overall reaction energy.

The $V_{rds}^{\ddagger}$ of $CH_4$ decomposition shows a linear correlation with the summation of $IE_M$ and $q_M$, with a MAE value of 0.046 eV (Table 4). These two features also appear in other formulas with complexities 5 and 9 (Table S9), indicating that they are the most sufficient features in predicting the reaction barrier. The descriptor clearly illustrates that a lower $IE_M$, facilitating electron transfers from the dopant to the antibonding orbital of the adsorbed species, and a more electron-rich dopant, which helps to stabilize the product states, promote C–H decomposition. Ru, Rh, and Cr were predicted as the top three candidates, which aligns with the DFT conclusion, as shown in Fig. 6b.

The complete ranking based on PySR equations is listed in Table S10. For $\Delta E_{1-3}$ and $V_{rds}^{\ddagger}$, the predicted ranking generally aligns with the DFT result with nDCG values of 0.969 and 0.980. Therefore, the PySR provides a reliable prediction for screening promising DRM catalysis activity.

For completeness, we also examined the optimal equation for $CO_2$ decomposition (Table S9). The $CO_2$ decomposition energy is expressed by the equation:

$$\Delta E_{CO_2(ads) \rightarrow CO(ads) + O(ads)} = -0.781 \text{ eV} + \frac{0.456 \text{ eV}}{6.04 - IE_M/eV} \quad (10)$$

with a MAE value of 0.083 eV. A lower $IE_M$ facilitates electron transfer from the dopant to the antibonding orbital of the adsorbed species, thereby promoting C–O decomposition. Consistently, Cr and V, which have the lowest $IE_M$, yield the lowest reaction energy (see Table S8 and Fig. S3a).

The barrier of $CO_2$ decomposition is expressed by the equation:

$$V_{CO_2}^{\ddagger} = 0.562 \text{ eV} - \left[ \frac{0.0217 \text{ eV}}{(0.158 - q_M/e) \times (1.515 - q_M/e)} \right] \quad (11)$$

with a MAE value of 0.053 eV. While the expression with a complexity value of 13 yields a slightly higher nDGG score of 0.966 and an acceptable MAE of 0.063 eV, it involves three features, making its physical interpretation less straightforward. In contrast, the current expression achieves a comparable nDGG of 0.959 using only a single descriptor, suggesting that the barrier primarily depends on $q_M$.

**Table 4** The formulas derived from PySR for $CH_4$ activation on doped Ni

| Target | Complexity | Formulas | MAE (eV) | nDCG |
|---|---|---|---|---|
| $\Delta E_{1-3}$ | 9 | $-0.481 \text{ eV} + \left[ \frac{0.258 \text{ eV}}{\frac{\Delta E_{f(M-H)}/eV + d + 4.11}} \right]$ | 0.084 | 0.969 |
| $V_{rds}^{\ddagger}$ | 7 | $0.261 \text{ eV} \times \left( \frac{IE_M}{eV} + \frac{q_M}{e} \right) - 1.22 \text{ eV}$ | 0.046 | 0.980 |

**7366** | *Catal. Sci. Technol.*, 2025, **15**, 7358–7370

This journal is © The Royal Society of Chemistry 2025
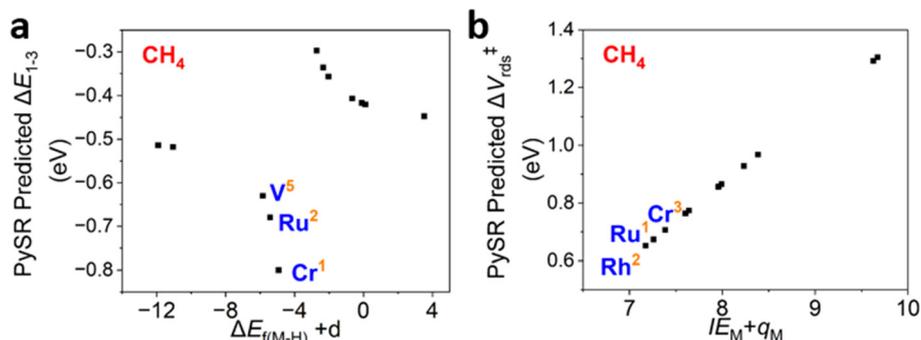
**Fig. 6** The scatter plot of (a) PySR-predicted $\Delta E_{1-3}$ of $CH_4$ vs. the sum of $\Delta E_{f(M-H)}$ and $d$ and (b) PySR-predicted $V_{rds}^{\ddagger}$ of $CH_4$ vs. the sum of $IE_M$ and $q_M$; dopants with the lowest predicted reaction energies or barriers are labeled in blue and the orange superscript shows the corresponding DFT ranking (taken as the ground truth).

### 3.4 Kinetic modeling

The BEP-predicted $V_{CH_4}^{\ddagger}$ of Y and $V_{CH_3}^{\ddagger}$ of Zn show the largest deviation from the DFT-calculated values, with a difference of 0.30 eV and 0.62 eV, respectively (see Table S11). Moreover, BEP prediction identified the $CH_4$ decomposition step (BEP-$V_{CH_4}^{\ddagger}$) of Zn as the rate-determining step with a barrier of 0.84 eV, which is inconsistent with the DFT results. To test this, we calculated the reaction rate using (1) DFT-calculated (ground truth) adiabatic energy barriers of all steps, (2) DFT-calculated adiabatic energy barriers for non-RDSs, and for the RDS, the nonlinear correlation-derived barrier from PySR, and (3) BEP-predicted adiabatic energy barriers of all steps. For all approaches, the reaction energies of each elementary step were obtained from DFT calculations. Indeed, the

underestimated BEP-predicted barriers for the RDS on Y–Ni and Zn–Ni significantly accelerate the decomposition process, resulting in an unrealistically high concentration of the H product, as illustrated by the dashed pink line in the upper panel of Fig. 7a and b. In contrast, the DFT and PySR results show nearly overlapping curves, indicating that PySR provides highly accurate predictions.

These differences span multiple orders of magnitude, which qualitatively affect our understanding of reaction kinetics and pathways. The choice between linear and nonlinear correlations not only impacts the predicted turnover rates but also drastically shifts the relative distribution of intermediates. Importantly, nonlinear correlation models align more closely with the actual DFT results, highlighting the critical role they play in accurately
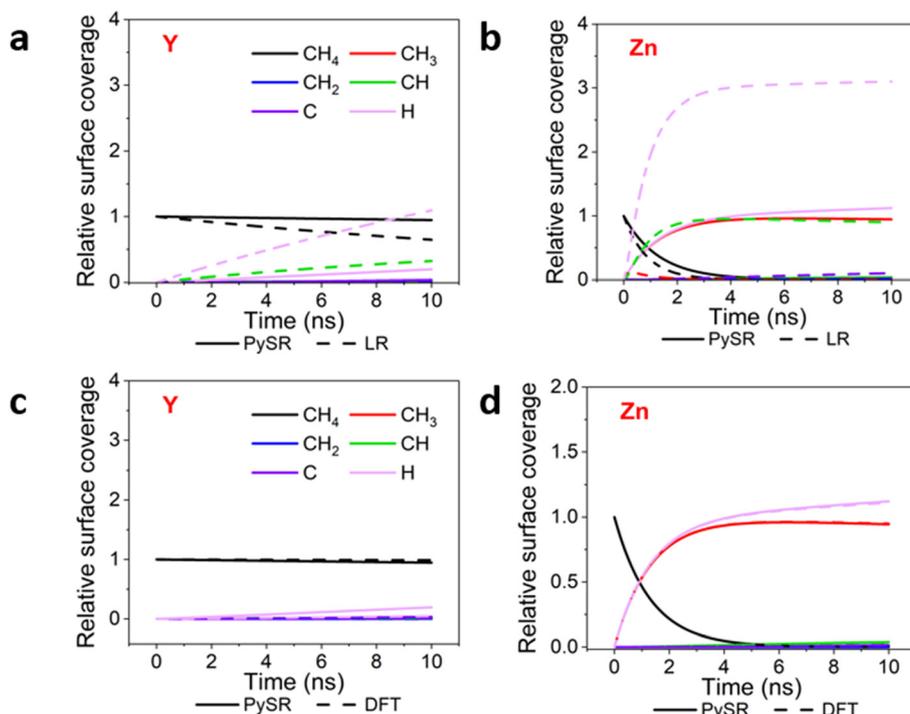


**Fig. 7** Reaction kinetics of DRM, in terms of the time evolution of the intermediate surface coverage (in a relative scale), by using PySR-predicted barriers, linear-regression (BEP) barriers based on DFT reaction energies, and the ground-truth DFT reaction barriers for (a and c) Y–Ni and (b and d) Zn–Ni.

This journal is © The Royal Society of Chemistry 2025

Catal. Sci. Technol., 2025, **15**, 7358–7370 | **7367**

capturing the complexities of catalytic behavior and ensuring reliable kinetic predictions. This emphasizes the limitations of traditional linear approaches and underscores the need for more sophisticated modeling in catalytic studies.

### 3.5 Validation of PySR selected dopants

We further applied our trained models to predict $\Delta E_{1-3}$ and $V_{rds}^{\ddagger}$ for 12 unexplored doped metals, including Sc, Mo, Zr, Nb, Tc, Pd, Cd, Hf, Ta, W, Os and Ir, as shown in Table S12. According to the predicted $\Delta E_{1-3}$, Ta and Mo exhibit the lowest reaction energy at −0.56 eV and −0.53 eV, respectively. For $V_{rds}^{\ddagger}$, Tc and Mo exhibit the lowest value at 0.72 eV and 0.77 eV, respectively. Among all predicted dopants, only Mo ranks highly in both $\Delta E_{1-3}$ and $V_{rds}^{\ddagger}$, making it a strong candidate for further DFT computation. The model prediction for Mo indicates that $\Delta E_{1-3}$ and $V_{rds}^{\ddagger}$ are −0.53 eV and 0.77 eV, differing only slightly from the DFT results of −0.56 eV and 0.73 eV, respectively.

We also perform DFT to calculate the $\Delta E_{1-3}$ of Ta, W, and Ir and the $V_{rds}^{\ddagger}$ of Nb to assess the predicted results due to their high predicted ranking in $\Delta E_{1-3}$ or $V_{rds}^{\ddagger}$. Although Tc has the highest ranking in $V_{rds}^{\ddagger}$, its radiation characteristics preclude its industrial application. As a result, the $\Delta E_{1-3}$ of Ta and Ir were found to be −0.4760 eV and −0.4700 eV, which differ by 0.081 eV and 0.047 eV from the predicted value. The $V_{rds}^{\ddagger}$ of Nb is 0.7667 eV, which differs by 0.049 eV from the predicted value.

For unexplored systems, the nonlinear model provides rapid screening to prioritize candidates, but DFT validation remains essential for final selection. The SR model generally agrees with DFT, although discrepancies, such as for W, warrant caution. The relatively large error for W likely reflects limitations of the current training dataset, which focuses on 4th- and 5th-period elements. As a 6th-period transition metal, W may exhibit features not fully captured by the model. Future work will expand the training dataset to include heavier and more diverse elements to improve generalizability.

Overall, the close agreement between PySR predictions and DFT results highlights the potential of symbolic regression in catalyst design, where precise predictions of reaction barriers are essential for optimizing catalytic activity.

## 4. Conclusion

We investigated the scaling relationship between DFT-computed reaction barriers and adsorption energies of intermediates in $CH_4$ and $CO_2$ direct decomposition to determine if BEP correlations are applicable to $CH_4$ and $CO_2$ direct decomposition on M–Ni. The diabatic energy analysis indicates that the breakdown of BEP linear correlation arises not just from structural differences but also from fundamental differences in the electronic nature and reactivity of the dopants.

Additionally, we utilized PySR to establish correlations between DFT-computed adsorption energies of intermediates on M–Ni and easily obtained physical properties of M–Ni (in separate models) with reaction barriers in the decomposition of $CH_4$ and $CO_2$. Furthermore, we employed PySR to make predictions for the remaining M–Ni systems and selected the most promising systems for further DFT calculations.

Establishing a quantitative and predictive scaling relationship between reaction barriers and reaction/adsorption energies in $CH_4$ and $CO_2$ direct decomposition has proven to be a challenging task. Therefore, using PySR to fit nonlinear relationships between various variables is essential. This approach not only enhances predictability but also offers a simpler and more straightforward method compared to black-box machine learning models with a high complexity. Moreover, we leveraged PySR formulas to predict the barriers for $CH_4$ and $CO_2$ direct decomposition in additional transition metal-doped Ni systems that are out of the training set. Concurrently, we identified the Mo–Ni alloys, with high ranking in both barrier and reaction energy, as the best candidate. DFT calculations were also performed to validate the nonlinear correlation results for other highly ranked doped M–Ni systems with an average error of less than 0.1 eV.

Our findings have significant implications for the design of more efficient DRM catalysts, the reduction of computational costs, and broader applications. These results contribute to advancing sustainable energy solutions, reducing greenhouse gas emissions, and benefiting various industries. By providing a more accurate and efficient method *via* automated nonlinear correlations for predicting reaction barriers, our work has the potential to accelerate the development of new catalysts and materials for energy-related applications that require a framework that goes beyond BEP linear relationships.

## Author contributions

Conceptualization: J. L. B. Data curation: Y. W., Y. W., S. L., S. L. Formal analysis: Y. W., S. L., J. L. B. Funding acquisition: J. L. B. Investigation: Y. W., Y. W., Y. C., S. L., S. L., J. L. B. Project administration: J. L. B. Resources: S. L., J. L. B. Software: S. L., J. L. B. Supervision: J. L. B. Validation: Y. W., Y. W., Y. C., S. L., S. L., J. L. B. Visualization: Y. W., Y. W., Y. C., S. L., S. L., J. L. B. Writing – original draft: Y. W., S. L. Writing – review & editing: J. L. B. All authors reviewed and approved the final manuscript.

## Conflicts of interest

The authors declare no conflicts of interest.

## Data availability

The data supporting this article have been included as part of the SI.

Supplementary information is available. See DOI: https://doi.org/10.1039/d5cy01066g.

# Acknowledgements

# References

1 B. Agún and A. Abánades, Comprehensive review on dry reforming of methane: Challenges and potential for greenhouse gas mitigation, *Int. J. Hydrogen Energy*, 2025, **103**, 395–414, DOI: 10.1016/j.ijhydene.2025.01.160.

2 L. Wu, X. Xie, H. Ren and X. Gao, A short review on nickel-based catalysts in dry reforming of methane: Influences of oxygen defects on anti-coking property, *Mater. Today: Proc.*, 2021, **42**, 153–160.

3 Z. Wang, Z. Mei, L. Wang, Q. Wu, C. Xia, S. Li, T. Wang and C. Liu, Insight into the activity of Ni-based thermal catalysts for dry reforming of methane, *J. Mater. Chem. A*, 2024, **12**, 24802–24838.

4 C. Fan, Y.-A. Zhu, M.-L. Yang, Z.-J. Sui, X.-G. Zhou and D. Chen, Density functional theory-assisted microkinetic analysis of methane dry reforming on Ni catalyst, *Ind. Eng. Chem. Res.*, 2015, **54**(22), 5901–5913.

5 W.-B. Zhang, C. Chen and S.-Y. Zhang, Equilibrium crystal shape of Ni from first principles, *J. Phys. Chem. C*, 2013, **117**(41), 21274–21280.

6 J. C. Swart, P. van Helden and E. van Steen, Surface energy estimation of catalytically relevant fcc transition metals using DFT calculations on nanorods, *J. Phys. Chem. C*, 2007, **111**(13), 4998–5005.

7 S. J. Carey, W. Zhao, A. Frehner, C. T. Campbell and B. Jackson, Energetics of adsorbed methyl and methyl iodide on Ni (111) by calorimetry: comparison to Pt (111) and implications for catalysis, *ACS Catal.*, 2017, **7**(2), 1286–1294.

8 S. Lin, C. Teng and J. L. Bao, CO2 adsorptions on d-block-metal-doped nickel nanoparticles: Unexpected adsorption configurations predicted by machine intelligence, *J. Phys. Chem. C*, 2021, **125**(36), 19839–19846.

9 S. Lin, J.-B. Tristan, Y. Wang and J. L. Bao, Dry reforming of methane on doped Ni nanoparticles: Feature-assisted optimizations and ranking of doping metals for direct activations of CH4 and CO2, *Nano Res.*, 2022, **15**(10), 9670–9682.

10 T. Bligaard, J. K. Nørskov, S. Dahl, J. Matthiesen, C. H. Christensen and J. Sehested, The Brønsted–Evans–Polanyi relation and the volcano curve in heterogeneous catalysis, *J. Catal.*, 2004, **224**(1), 206–217.

11 S. M. Stratton, S. Zhang and M. M. Montemore, Addressing complexity in catalyst design: From volcanos and scaling to more sophisticated design strategies, *Surf. Sci. Rep.*, 2023, **78**(3), 100597, DOI: 10.1016/j.surfrep.2023.100597.

12 J. Cheng, P. Hu, P. Ellis, S. French, G. Kelly and C. M. Lok, Brønsted–Evans–Polanyi relation of multistep reactions and volcano curve in heterogeneous catalysis, *J. Phys. Chem. C*, 2008, **112**(5), 1308–1311.

13 R. A. Van Santen, M. Neurock and S. G. Shetty, Reactivity theory of transition-metal surfaces: a Brønsted–Evans–Polanyi linear activation energy–free-energy analysis, *Chem. Rev.*, 2009, **110**(4), 2005–2048.

14 P. Ferrin, D. Simonetti, S. Kandoi, E. Kunkes, J. A. Dumesic, J. K. Nørskov and M. Mavrikakis, Modeling ethanol decomposition on transition metals: a combined application of scaling and Brønsted–Evans–Polanyi relations, *J. Am. Chem. Soc.*, 2009, **131**(16), 5809–5815.

15 J. L. Fajín, M. N. D. Cordeiro, F. Illas and J. R. Gomes, Generalized Brønsted–Evans–Polanyi relationships and descriptors for O–H bond cleavage of organic molecules on transition metal surfaces, *J. Catal.*, 2014, **313**, 24–33.

16 C. F. Nwaokorie and M. M. Montemore, Alloy catalyst design beyond the volcano plot by breaking scaling relations, *J. Phys. Chem. C*, 2022, **126**(8), 3993–3999.

17 W. Zhou, Y. Chen, L. Tan, Q. Tang, C. Lu and L. Dong, Understanding the C–H activation of methane over single-atom alloy catalysts by density functional theory calculations, *AIChE J.*, 2023, **69**(8), e18118.

18 L. Billard and E. Diday, Symbolic regression analysis, in *Classification, clustering, and data analysis: recent advances and applications*, Springer, 2002, pp. 281–288.

19 B. Weng, Z. Song, R. Zhu, Q. Yan, Q. Sun, C. G. Grice, Y. Yan and W.-J. Yin, Simple descriptor derived from symbolic regression accelerating the discovery of new perovskite catalysts, *Nat. Commun.*, 2020, **11**(1), 3513.

20 T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, R. Tibshirani and J. Friedman, Linear methods for regression, *The elements of statistical learning: Data mining, inference, and prediction*, 2009, pp. 43–99.

21 N. J. O'Connor, A. Jonayat, M. J. Janik and T. P. Senftle, Interaction trends between single metal atoms and oxide supports identified with density functional theory and statistical learning, *Nat. Catal.*, 2018, **1**(7), 531–539.

22 S. Liu, Z.-L. Wang, L. Zhang, G.-X. Chen, H.-F. Yang, X.-N. Liang and J. Qiu, Investigations on symbol regression for improving the prediction accuracy of gas-metal adsorption energies in machine learning, *Surf. Interfaces*, 2024, **55**, 105469.

23 M. Cranmer, Interpretable machine learning for science with PySR and SymbolicRegression.jl, *arXiv*, 2023, preprint, arXiv:2305.01582, DOI: 10.48550/arXiv.2305.01582.

24 K. Järvelin and J. Kekäläinen, Cumulated gain-based evaluation of IR techniques, *ACM Trans. Inf. Syst.*, 2002, **20**(4), 422–446.

25 G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**(16), 11169.

26 P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **50**(24), 17953.

27 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.*, 1996, **77**(18), 3865–3868.

This journal is © The Royal Society of Chemistry 2025

*Catal. Sci. Technol.*, 2025, **15**, 7358–7370 | **7369**

28   S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A consistent and accurate ab initio parameterization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, 2010, **132**, 154104.

29   S. Grimme, Semiempirical GGA-type density functional constructed with a long-range dispersion correction, *J. Comput. Chem.*, 2006, **27**, 1787–1799, DOI: 10.1016/j.ssc.2006.04.027.

30   E. R. Johnson and A. D. Becke, A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections, *J. Chem. Phys.*, 2006, **124**(17), 174104.

31   H. J. Monkhorst and J. D. Pack, Special points for Brillouin-zone integrations, *Phys. Rev. B: Solid State*, 1976, **13**(12), 5188.

32   G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **59**(3), 1758.

33   G. Makov and M. C. Payne, Periodic boundary conditions in ab initio calculations, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1995, **51**(7), 4014.

34   G. Henkelman, B. P. Uberuaga and H. Jónsson, A climbing image nudged elastic band method for finding saddle points and minimum energy paths, *J. Chem. Phys.*, 2000, **113**(22), 9901–9904.

35   Q. Wu and T. Van Voorhis, Constrained Density Functional Theory and its Application in Long-Range Electron Transfer, *J. Chem. Theory Comput.*, 2006, **2**(3), 765–774.

36   Q. Wu and T. Van Voorhis, Direct Calculation of Electron Transfer Parameters Through Constrained Density Functional Theory, *J. Phys. Chem. A*, 2006, **110**(29), 9212–9218.

37   Q. Wu, C.-L. Cheng and T. Van Voorhis, Configuration Interaction Based on Constrained Density Functional Theory: A Multireference Method, *J. Chem. Phys.*, 2007, **127**(16), 164119.

38   M. Vitale-Sullivan and K. A. Stoerzinger, Interplay of surface and subsurface contributions in electrocatalysis, *Curr. Opin. Electrochem.*, 2023, **39**, 101252.

39   R. Ohmann, G. Levita, L. Vitali, A. De Vita and K. Kern, Influence of subsurface layers on the adsorption of large organic molecules on close-packed metal surfaces, *ACS Nano*, 2011, **5**(2), 1360–1365.

40   E. Epifanovsky, A. T. Gilbert, X. Feng, J. Lee, Y. Mao, N. Mardirossian, P. Pokhilko, A. F. White, M. P. Coons and A. L. Dempwolff, Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package, *J. Chem. Phys.*, 2021, **155**(8), 084801.

41   F. Weigend and R. Ahlrichs, Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy, *Phys. Chem. Chem. Phys.*, 2005, **7**(18), 3297–3305, DOI: 10.1039/B508541A.

42   E. Cancès and C. Le Bris, Can we outperform the DIIS approach for electronic structure calculations?, *Int. J. Quantum Chem.*, 2000, **79**(2), 82–90.

43   V. R. Saunders and I. H. Hillier, A "Level–Shifting" method for converging closed shell Hartree–Fock wave functions, *Int. J. Quantum Chem.*, 1973, **7**(4), 699–705.

44   M. Guest and V. Saunders, On methods for converging open-shell Hartree-Fock wave-functions, *Mol. Phys.*, 1974, **28**(3), 819–828.

45   F. Weinhold and E. D. Glendening, *NBO 5.0 program manual: natural bond orbital analysis programs*, Theoretical Chemistry Institute and Department of Chemistry, University of Wisconsin, Madison, WI, 2001, p. 53706.

46   R. F. W. Bader, Atoms in molecules, *Acc. Chem. Res.*, 1985, **18**(1), 9–15, DOI: 10.1021/ar00109a003.

47   R. F. Bader, A quantum theory of molecular structure and its applications, *Chem. Rev.*, 1991, **91**(5), 893–928.

48   C. T. Campbell and J. R. Sellers, Anchored metal nanoparticles: Effects of support and size on their energy, sintering resistance and reactivity, *Faraday Discuss.*, 2013, **162**, 9–30.

49   R. Krishnan, J. Binkley, R. Seeger, J. Pople, R. Krishnan, J. Binkley, R. Seeger and J. Pople, Self-consistent molecular orbital methods, XX. A basis set for correlated wave functions, *J. Chem. Phys.*, 1980, **72**, 650–654.

50   M. J. Frisch, J. A. Pople and J. S. Binkley, Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets, *J. Chem. Phys.*, 1984, **80**(7), 3265–3269.

51   *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

52   D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *J. Comput. Phys.*, 1976, **22**(4), 403–434.

53   E. D. López and A. Comas-Vives, Kinetic Monte Carlo simulations of the dry reforming of methane catalyzed by the Ru (0001) surface based on density functional theory calculations, *Catal. Sci. Technol.*, 2022, **12**(13), 4350–4364.

54   S. Matthew, F. Carter, J. Cooper, M. Dippel, E. Green, S. Hodges, M. Kidwell, D. Nickerson, B. Rumsey and J. Reeve, Gillespy2: A biochemical modeling framework for simulation driven biological discovery, *Lett. Biomath.*, 2023, **10**(1), 87.

**7370** | *Catal. Sci. Technol.*, 2025, **15**, 7358–7370

This journal is © The Royal Society of Chemistry 2025