

# An end-to-end framework for reactivity in heterogeneous catalysis

Received: 8 May 2025

Accepted: 26 January 2026

Published online: 19 March 2026

Check for updates

Santiago Morandi <sup>1,2,12</sup>, Oliver Loveday <sup>1,2,12</sup>, Tim Renningholtz<sup>1</sup>, Sergio Pablo-García<sup>3,4,5</sup>, Rodrigo A. Vargas-Hernández<sup>6</sup>, Ranga Rohit Seemakurthi <sup>1</sup>, Pol Sanz Berman <sup>1,2</sup>, Rodrigo García-Muelas <sup>1,11</sup>, Alán Aspuru-Guzik<sup>3,4,5,7,8,9,10</sup> & Núria López <sup>1</sup> ✉

The rationalization of catalytic processes relies on the fundamental understanding of competing reaction mechanisms driving reactants to products. The list of elementary steps composing the reaction networks is proposed based on chemical intuition and evaluated via density functional theory. This approach is limited by the size of the network and disregards alternative paths. Here we present the Catalytic Automated Reaction Evaluator (CARE), a flexible end-to-end framework for heterogeneous catalysis composed of (1) a rule-based reaction network generator, (2) a thermodynamic and kinetic parameter evaluator powered by state-of-the-art machine learning models and (3) a fast microkinetic solver. CARE reproduces the experimental activity trends in methanol decomposition, identifies the selectivity to C<sub>3</sub> products in CO<sub>2</sub> electroreduction and generates the Fischer–Tropsch synthesis mechanism including 370,000 reactions reaching C<sub>6</sub> products. This comprehensive framework enables the exploration of thermal and electrocatalytic reactions previously not amenable to atomistic simulations.

Catalytic processes rely on the control of complex competing multistep chemical reaction networks (CRNs)<sup>1,2</sup>. The core of a CRN consists of a set of species **S**, linked by a set of reactions **R** defined by properties **P**, usually including, but not limited to, their reaction and activation energies<sup>3,4</sup>. For chemical processes in which the list of elementary steps can be manually generated, density functional theory (DFT) enables the evaluation of reaction profiles. However, this approach finds its limits as network complexity grows<sup>5</sup> and then it is necessary to leverage the CRN's abstract mathematical foundations<sup>6–11</sup>.

CRN algorithms fall into two categories: CRN potential energy surface and rule-based generators. The former require on-the-fly ab initio energy evaluation of the species involved along the reaction paths<sup>12–17</sup>. As such, these algorithms are constrained to chemical spaces

(CSs) of the order of 10<sup>2</sup> intermediates. Alternatively, rule-based CRN generators use templates to define the attainable CS via a list of possible transformations<sup>18,19</sup>, decoupling network generation from energy evaluation. Therefore, they can deal with larger networks compared with the CRN potential energy surface, including multiple paths that could contribute to the catalytic activity and rare events responsible for catalyst deactivation, thereby getting closer to achieving the completeness of the CS. However, the combinatorial explosion of the CS requires strategies to reduce the CRN complexity and/or substitute DFT with fast energy estimators.

The evaluation of the thermodynamic properties of the species in the network can be accelerated by using group additivity (GA) methods<sup>20–22</sup>, whereas linear scaling relationships<sup>23,24</sup> and

<sup>1</sup>Institute of Chemical Research of Catalonia, The Barcelona Institute of Science and Technology, Tarragona, Spain. <sup>2</sup>Department of Physical and Inorganic Chemistry, Universitat Rovira i Virgili, Campus Sescelades, Tarragona, Spain. <sup>3</sup>Department of Chemistry, University of Toronto, Lash Miller Chemical Laboratories, Toronto, Ontario, Canada. <sup>4</sup>Department of Computer Science, University of Toronto, Sandford Fleming Building, Toronto, Ontario, Canada. <sup>5</sup>Vector Institute for Artificial Intelligence, Toronto, Ontario, Canada. <sup>6</sup>Department of Chemistry & Chemical Biology, McMaster University, Hamilton, Ontario, Canada. <sup>7</sup>Department of Materials Science and Engineering, University of Toronto, Toronto, Ontario, Canada. <sup>8</sup>Lebovic Fellow, Canadian Institute for Advanced Research (CIFAR), Toronto, Ontario, Canada. <sup>9</sup>Director, Acceleration Consortium, Toronto, Ontario, Canada. <sup>10</sup>NVIDIA, Toronto, Ontario, Canada. <sup>11</sup>Present address: Iberian Centre for Research in Energy Storage, Cáceres, Spain. <sup>12</sup>These authors contributed equally: Santiago Morandi, Oliver Loveday. ✉e-mail: [nlopez@icicq.es](mailto:nlopez@icicq.es)

Brønsted–Evans–Polanyi (BEP) equations<sup>25–27</sup> can be used for characterizing the reaction properties. An example of this approach is the reaction mechanism generator<sup>28</sup>, which was applied to the CO<sub>2</sub> methanation and hydrocarbon (CH<sub>4</sub> and C<sub>2</sub>H<sub>6</sub>) oxidation on Pt(111) and Ni(111) surfaces<sup>29</sup>. On iterative species generation with graph-based approaches evaluated with the aforementioned estimators (GA–linear scaling relationship–BEP), a filtering step is applied after every network expansion to reduce the mechanism size. Although the reaction mechanism generator provides general reactivity trends, the accuracy of the energy predictions is low compared with DFT with a mean absolute error (MAE) of around 0.5 eV (ref. 30), limiting its applicability to mechanisms not involving selectivity issues.

The competition and correlation between the completeness of the CS and the associated computational burden arising from the characterization of CRNs are key to the development of the field. Expanding the boundaries of the investigated CSs requires more robust, accurate and efficient energy regressors, which can be obtained via ML acceleration strategies<sup>31–34</sup>. Ulissi et al.<sup>19</sup> built a surrogate model with GA fingerprints fed to a Gaussian process and combined it with BEP relationships to obtain the rate-limiting steps for the conversion of syngas to C<sub>2</sub> products on Rh catalysts. This model was used within an active learning strategy and tracking the uncertainty to predict the most relevant steps to be refined with DFT.

An alternative approach involves ML interatomic potentials (MLIPs) from the Open Catalyst Project<sup>32,35</sup>, trained on small CHNO fragments on metals and alloys, for producing initial guesses of the adsorption and transition-state (TS) geometries subsequently relaxed with DFT<sup>30,36</sup>, reaching an MAE of 0.1 eV. Other strategies target CHO species on metals including TSs via BEP and ML-based regression techniques, achieving MAEs of 0.2 eV (refs. 37,38). These models speed up DFT simulations, but are still too computationally demanding for high-throughput *in silico* exploration. MLIPs<sup>39–43</sup>, graph-based approaches<sup>33,44</sup> and diffusion models<sup>45</sup> reduce the prediction error, but have yet to be merged within CRN algorithms.

Here we present the Catalytic Automated Reaction Evaluator (CARE), an end-to-end framework tailored to investigate performance in heterogeneous catalysis that consists of (1) a catalyst-agnostic rule-based CRN generator for processes involving CHO species, (2) a module interfacing to state-of-the-art ML models for evaluating the thermodynamic and kinetic reaction parameters for thermal and electrochemical processes, mainly powered by GAME-Net-UQ, and (3) a microkinetic solver to predict the reactivity. CARE is benchmarked against experimental activity in methanol decomposition, used to study selectivity in electrochemistry and to generate large networks such as those describing the Fischer–Tropsch synthesis.

## Results

### CARE

CARE is an end-to-end framework for the generation and manipulation of CRNs in heterogeneous and electrocatalytic processes involving CHO species. The framework consists of three independent modules: a template-based network blueprint generator, an interface to state-of-the-art data-driven estimators for thermodynamic and kinetic reaction parameters, and a microkinetic solver (Fig. 1).

**Network blueprint generation.** The first step for generating the CRN blueprint in CARE is the definition of the CS. The size of the CS is determined by the backbone of the largest molecule, obtained by defining the maximum number of C ( $N_C$ ) and O ( $N_O$ ) atoms for the species, namely, the carbon and oxygen cut-offs. To set up the CS, a pool of alkanes with up to  $N_C$  carbon atoms is generated. Through a set of molecular templates, alkanes are then oxygenated and saturated molecules with up to  $N_O$  oxygen atoms are obtained (for example, alcohols, ethers and epoxides). In contrast to  $N_C$  and  $N_O$ , custom CSs can be defined by directly providing a list of SMILES.

Next, a bond-breaking template is applied to the molecules in the CS, simultaneously generating a set of reactions, open-shell fragments and unsaturated closed-shell molecules (alkenes, ketones and acids). This template is based on SMILES manipulations implemented with RDKit. The obtained species plus the CS define the extended CS (eCS).

In contrast to homogeneous catalysis, the number of possible reaction templates in heterogeneous catalysis is relatively small, but the number of events can be very large, as in polymerization/depolymerization processes. Therefore, the main reaction templates in CARE describe  $X-Y$  ( $X, Y = C, H, O$ ) bond-breaking/bond-forming reactions and the adsorption/desorption of closed-shell molecules. Additional available templates account for intramolecular rearrangements such as [1,2]-H shift and proton-coupled electron transfer (PCET), enabling modeling reaction networks under electrochemical conditions. The modularity of the network generator facilitates the inclusion of more reaction classes, for instance, reactivity in solution at the interfaces of electrodes.

The ensemble of the eCS and the set of reactions constitute the reaction network blueprint. It is catalyst agnostic and needs to be generated only once to model the same network on different catalysts. Additional details on the CRN blueprint generation are provided in Supplementary Note 1.

**Estimating thermodynamic and kinetic parameters.** To estimate the thermodynamic and kinetic properties of the reactions in the network, CARE provides access to many state-of-the-art ML energy estimators, but is mainly powered by GAME-Net-UQ (UQ, uncertainty quantification), a graph neural network targeting the DFT energy of adsorbed species and TSs. Originally, GAME-Net<sup>33</sup> was designed as a direct initial-structure-to-relaxed-energy model for closed-shell molecules on transition metals, starting from the simplest possible graph representation, without any spatial coordinate encoded in the nodes or edges.

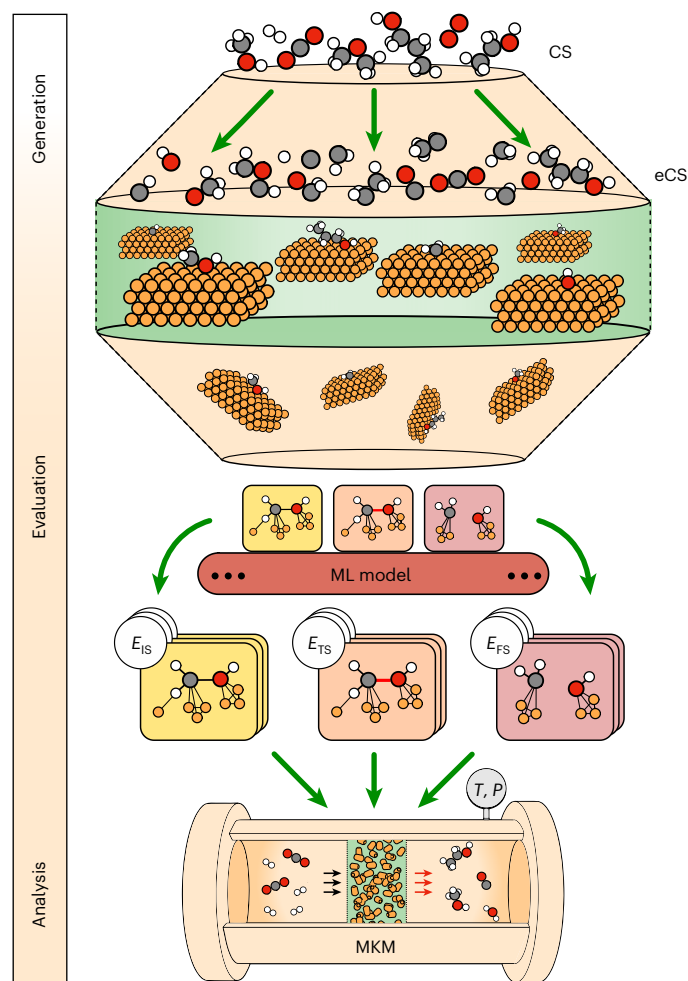
GAME-Net-UQ includes three crucial upgrades: (1) the extension to open-shell species, (2) the estimation of TS energies and (3) UQ. Including these features keeps the model lightweight (558,000 parameters), accurate and robust, enabling extensive direct energy evaluations at a minimal computational cost.

The graph dataset used to develop GAME-Net-UQ includes 12,303 structures (Supplementary Note 3). These represent the DFT-optimized adsorption of 184 closed-shell molecules from the GAME-Net training dataset, as well as intermediates and TSs included in the C<sub>2</sub>H<sub>6</sub>O<sub>2</sub> decomposition network (see the ‘DFT’ section). The final graph dataset comprises 80% intermediates and 20% TSs of the surface-bond-breaking events. During training, to distinguish TSs from intermediates graphs, a Boolean feature is added to the graph edges to label the bond involved in dissociation (Fig. 2a). Consequently, the graph neural network includes a topology-adaptive convolutional layer<sup>46</sup> utilizing the encoded edge attribute.

The representation of the surface has been expanded to enable the study of structure-sensitive processes. Surface effects are taken into account by including (1) species adsorbed on the second- and third-most-stable surface facets of the metals in the dataset (Supplementary Fig. 5), (2) the generalized coordination number<sup>24</sup> of surface atoms as the node feature and (3) the second-order surface neighbors to distinguish hollow adsorption sites.

As GAME-Net-UQ has been redefined as a mean variance estimator<sup>47</sup>, it returns predictions as normal distributions  $E \approx \mathcal{N}(\mu, \sigma^2)$ , with the standard deviation  $\sigma$  representing the uncertainty (Fig. 2a). Details about graph representation, UQ, model architecture and training are provided in Supplementary Notes 4–6.

GAME-Net-UQ yields an MAE of 0.21 eV on a test set of 2,460 data points randomly sampled from the initial dataset, 0.21 eV for intermediates and 0.22 eV for TSs. The outliers lie in the [−4; +4] eV range (Fig. 2b), which are the most relevant for practical catalytic reactions.



**Fig. 1 | CARE workflow from network generation to reactivity analysis.**

CARE starts from a CS defined by the maximum number of C and O atoms and a molecular template defining how to build the carbon backbone. The bond-breaking template returns the eCS and all potential reactions. Species in the eCS are automatically placed on the surface, and their energy and those of the linking TSs are evaluated with the ML model of choice. The reactivity of the network can be characterized using MKM under realistic reaction conditions.  $E_{IS}$ , initial state energy;  $T$ , temperature;  $P$ , pressure.

These outliers are small fragments adsorbed on Cd and Zn, typically not used as catalysts in their metallic form. The performed ablation study (Supplementary Note 7) shows that labeling the bond involved in the TS reduces the MAE by 0.23 eV, whereas encoding the generalized coordination number in the graph reduces the MAE by 0.04 eV. Predictions with small (big) errors are reflected in small (big) uncertainties, resulting in a model with a 1.3% miscalibration area (Supplementary Fig. 11). Small adsorbates on Cd, Zn and Fe show higher errors, but the uncertainty associated with these systems is positively correlated (Fig. 2c). The sharpness of the model, representing the magnitude of its uncertainty in the test set, is 0.23 eV, whereas the coefficient of variation, which quantifies the dispersion of the uncertainty estimates, is 0.34.

During inference, a streamlined version of DockOnSurf<sup>48</sup> is used for the generation of adsorption configurations before evaluation. Adsorbate atoms with coordination lower than their valence are defined as anchoring points, whereas surface sites are identified with ACAT<sup>49</sup> (Supplementary Note 2). As for the same intermediate multiple configurations screened within CARE, UQ can be leveraged to select the configurations with the lowest associated uncertainty, or to inform

active learning strategies for reevaluating key species at higher accuracy. This is particularly meaningful for intermediates participating in the highest number of elementary reactions<sup>11</sup>.

GAME-Net-UQ propagates the uncertainty to reaction energy  $\Delta E$  and activation barrier  $E_{act}$  (Supplementary Note 8). Surface reactions are evaluated in the bond-breaking direction  $A^* + * \rightarrow B^* + C^*$ . During inference, the TS graph is constructed from the reactant graph  $A^*$ . The broken bond involved in the reaction is tagged to denote the TS (Fig. 2a and Supplementary Note 9). The products  $B^*$  and  $C^*$  in the final state (FS) of the reaction are evaluated separately as the training data represent one adsorbate on the surface that lacks lateral interactions. This might result in GAME-Net-UQ predicting non-physical TS energies  $E_{TS} < E_{FS}$ . In this situation, CARE assumes exothermic steps to be barrierless ( $E_{act} = 0$  eV) and endothermic ones having  $E_{act} = \Delta E$ .

Although the initial development of CARE focused on GAME-Net-UQ, interfaces to MLIPs such as FAIRChem potentials<sup>32</sup>, MACE<sup>42</sup>, Orb<sup>43</sup>, PET-MAD<sup>50</sup> and SevenNet<sup>51</sup> are available. These allow both relaxation and TS search with the nudged elastic band (NEB) method, providing a seamless workflow to fully characterize the CRN energy landscape with any MLIP, surface reaction and catalyst. As of January 2026, users can access 91 MLIPs (Supplementary Table 9). Having the option to evaluate the same network with different evaluators within the same framework is a pivotal feature of CARE, providing flexibility to quantitatively compare different ML strategies in terms of accuracy with respect to DFT and the derived macroscale properties with respect to experimental activity.

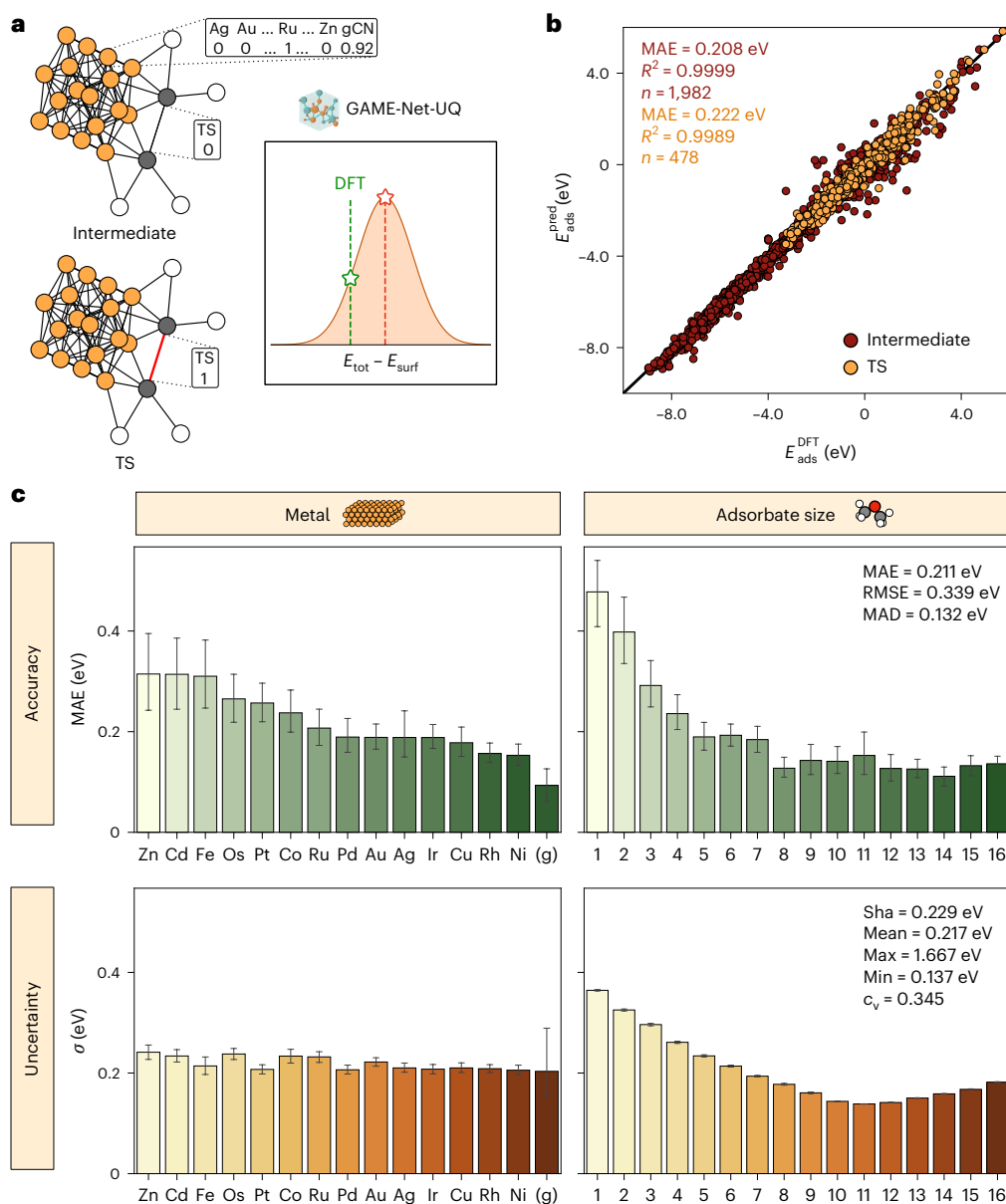
**CRN analysis via microkinetic modeling.** The prediction of the catalytic activity from CRNs in CARE with hundreds of reactions is not straightforward since these networks (1) have undefined directionality at the global level as this depends on the definition of the reactants, (2) are unconstrained to specific source and target species and (3) their catalytic activity depends on the applied operating conditions. Therefore, mean-field microkinetic modeling (MKM) functionalities<sup>52</sup> are present in CARE.

MKM relies on the numerical solution of the system of ordinary differential equations (ODEs) defining the mass balance for the species in the CRN. The default model in CARE is a zero-conversion differential plug-flow reactor, which enables the extraction of macroscopic steady-state reaction rates and selectivity directly comparable with experiments, surface coverages and rate-determining steps. The kinetic coefficients of the elementary reactions are evaluated via TS theory for all surface steps except adsorption, for which the Hertz–Knudsen equation<sup>53</sup> is used. The coefficients of the reverse reactions are obtained by imposing thermodynamic consistency (Supplementary Note 8). The ODE solvers in CARE automatically terminate the integration when the chemical steady state is reached. This occurs when all elemental balances are respected and the surface coverages sum to unity.

Elementary reactions possess characteristic times ranging multiple orders of magnitude, resulting in stiff ODEs requiring specialized solvers. Popular Python libraries such as SciPy<sup>54</sup> lead to a prohibitive computational cost for kinetic simulations on huge CRNs and/or stiff ODEs. Therefore, CARE additionally implements a solver based on the `DifferentialEquations.jl` Julia package<sup>55</sup>, outperforming SciPy solvers in terms of convergence and simulation time required. Details about code implementation are provided in the ‘MKM’ section.

### Application to industrially relevant problems

To highlight the capabilities of CARE, we have selected three industrially relevant applications of increasing complexity: methanol decomposition, propanol and propylene ( $C_3H_6$ ) production in electrochemical conditions, and the Fischer–Tropsch synthesis. First, we benchmark the activity against an experimental dataset<sup>56</sup>. Systematic activity measures are severely lacking but CatTestHub presents an excellent example, providing activity data for different catalysts and operating



**Fig. 2 | GAME-Net-UQ performance.** **a**, Graph representation of intermediates and TSs, and energy prediction as normal distribution.  $E_{\text{tot}}$ , energy of the total system;  $E_{\text{surf}}$ , energy of the surface slab. **b**, Parity plot of the predicted versus DFT adsorption energy for the test intermediates (brown) and TS (orange) data. **c**, MAE of the predictions and associated mean uncertainties  $\sigma$  grouped by the

metal and adsorbate size. Error bars represent the 95% confidence interval and '(g)' refers to the gas-phase graphs. RMSE, root mean squared error; MAD, median absolute deviation; Sha, sharpness;  $c_v$ , coefficient of variation (dimensionless). Statistics metrics refer to the whole test set ( $n = 2,460$ ), which has been randomly sampled from the original data distribution.

conditions. The second case study illustrates a reaction network in electrochemical environments. Selectivity is key to processes and typically depends on two competing steps with small energy differences, which is challenging in terms of accuracy, especially for data-driven energy estimators. The validity of CARE in predicting selectivity patterns in electrochemical applications is assessed by analyzing the steps toward  $C_3$  products on Cu(100) (ref. 57) in the  $C_3O_2$  ( $N_c = 3, N_o = 2$ ) CRN. Last, the capacity of CARE to tackle large networks has been assessed by analyzing the Fischer–Tropsch synthesis. This is the case of a reaction that cannot be addressed with DFT given the high number of potential steps involved. For this case study, we generated the  $C_6O_1$  CRNs on Co, Fe, Ni and Ru, resulting in networks of 40,000 intermediates and 370,000 elementary reactions and ran microkinetic simulations on their subnetworks, leading to linear products (37,000 reactions) to analyze the predicted product distribution.

**Methanol decomposition.** A benchmark of our framework against experimental activity is established by studying methanol decomposition to CO on different metal catalysts. Methanol is on the rise in sustainable energy applications as a hydrogen vector<sup>58,59</sup>, and developing efficient catalysts for its controlled decomposition is, therefore, crucial.

This example focuses on a small CS ( $N_c = 1, N_o = 2$ ) but covers a high-throughput investigation spanning multiple catalysts and operating conditions. The public experimental database CatTestHub<sup>56</sup> contains 96 kinetic experiments for this reaction, covering nine transition metal catalysts as well as multiple supports and operating conditions. As benchmark data, we selected a subset operating with the same nanoparticle size, inert carrier and operating conditions (Supplementary Table 11). We first built the  $C_1O_2$  CRN including 28 surface species, 10 gas-phase molecules and 62 elementary reactions (Fig. 3a). As metal

nanoparticles expose multiple surfaces, we rely on the Crystalium database<sup>60</sup> to obtain the nature and fraction of the exposed surfaces obtained through the Wulff construction. The corresponding surface slabs of the eight face-centered cubic metals (Au, Ag, Cu, Ir, Ni, Pt, Pd and Rh) and hexagonal close-packed Ru (Supplementary Table 13) have been generated, and the corresponding networks have been evaluated with GAME-Net-UQ, EquiformerV2-31M<sup>35</sup> and MACE-MP-0<sup>42</sup>. Last, kinetic simulations have been run at 473 K, 102.73 kPa and MeOH/ $N_2 = 10/90$  (Supplementary Note 10).

Figure 3a provides an overview of a  $C_1O_2$  CRN. CARE represents CRNs as bipartite graphs with nodes representing species and reactions, with edges  $e(s, r)$  embedding the consumption/formation rate of species  $s$  due to reaction  $r$ . Sequentially evaluating the 63 CRNs with GAME-Net-UQ took 47 min on a GPU. DFT simulation of all species in the networks without TS evaluation would require 286,168 CPU hours and 2,392 kWh of energy consumption (Supplementary Note 12).

Comparing GAME-Net-UQ predictions with the available DFT data<sup>61</sup> on a subset of the studied surfaces (Fig. 3b) shows an accuracy in terms of MAE of 0.47 eV for the reaction energy  $\Delta E$ , and 0.33 eV for the activation energy  $E_{act}$ . The higher prediction error for  $\Delta E$  arises from error propagation across the three involved species ( $A^* \rightarrow B^* + C^*$ ), whereas  $E_{act}$  depends only on the energies of  $A^*$  and TS. The outliers are associated with C–O-bond-breaking reactions typically occurring at steps<sup>62</sup>. Compared with EquiformerV2-31M and MACE-MP-0 (Supplementary Fig. 18), GAME-Net-UQ shows the highest error for  $\Delta E$  but appears to be the most robust when predicting  $E_{act}$ . Although EquiformerV2-31M shows a remarkable accuracy when predicting  $\Delta E$  (MAE = 0.40 eV), together with MACE-MP-0, it struggles with TSs, returning MAEs above 2 eV for  $E_{act}$ . Even though the CatTSunami framework<sup>36</sup> reports excellent performance on predicting  $E_{act}$  for similar reactions with EquiformerV2-31M (MAE = 0.06 eV), this metric has been obtained by using the DFT-relaxed initial and final states for the NEB ML-assisted simulation, plus a single-point DFT run to get the final TS energy. Here GAME-Net-UQ and MLIPs have been used in a fully ML-based scenario. The higher errors of the considered MLIPs appear because their training data are mainly near-equilibrium structures<sup>63,64</sup>. The automated TS search methodology in CARE (Supplementary Note 9) introduces further error, necessitating future refinements to reduce the divergence of NEB simulations.

DFT studies typically focus on modeling reactive processes on the most stable surface. However, the observed activity in metal nanoparticles is attributed to contributions from multiple exposed surfaces. CARE enables targeting multiple surfaces and predicting their associated activity in a systematic way. GAME-Net-UQ is able to capture a clear volcano trend on the CO formation rate, providing a ranking of the different metal surfaces (Fig. 3c). The predictions highlight Ru as the most active metal, in agreement with experimental evidence<sup>56</sup>. Rh(310) appears as an outlier, mainly due to the stepped nature of the surface, which is reflected in  $E_{ads,CO}$  1 eV higher than on its other facets (Supplementary Table 13), where CO should adsorb on the step. This less favorable interaction leads to a higher CO formation rate.

Defining the overall CO formation rate as a weighted sum of the rates of the individual surface facets<sup>60</sup> (Supplementary Fig. 16) leads to a general agreement with the experimental trend (Fig. 3d). The high deviation for Rh is due to the predicted rate on the (310) surface, which governs the metal activity and contributes 1% of the surface exposed by the nanoparticle. The offset between the predicted and experimental rates, which primarily originates from the additional ML error to that intrinsic to DFT, is a situation often encountered in mean-field kinetics based on atomistic data<sup>65,66</sup>. Other sources of discrepancy are the lack of lateral interactions and thermal contributions<sup>67</sup>.

Evaluating the CRNs with EquiformerV2-31M and MACE-MP-0 including TS search with NEB on GPU requires 100 times the resources of GAME-Net-UQ (Supplementary Fig. 17). The presence of a few reactions with very high barriers (Supplementary Table 10) and the general

overestimation of the CO adsorption energy leads to predicted formation rates many orders of magnitude below the experimental trends and experimental activity ranking not correctly reproduced (Supplementary Figs. 19 and 20).

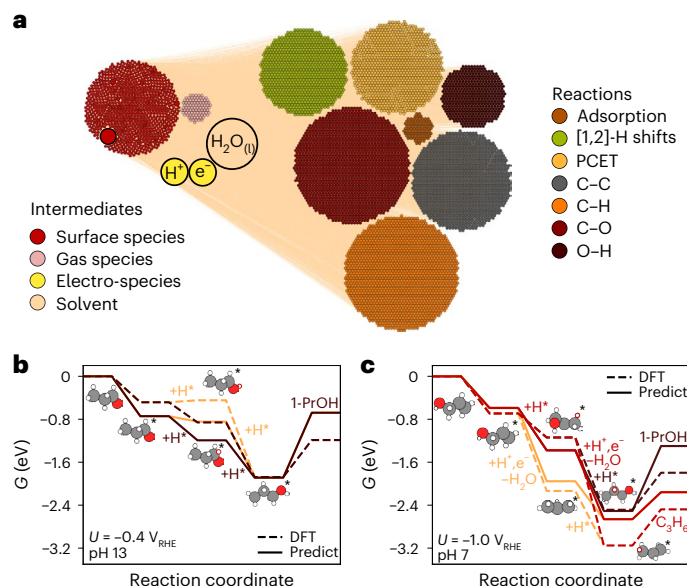
This case study highlights three important concepts. First, there is room for energy-targeted ML approaches not involving structural relaxation. Second, MLIPs for heterogeneous catalysis require data-sampling strategies enabling a better description of TSs<sup>63</sup>. Last, additional standardized experimental data are needed to define sound benchmarks for the computational heterogeneous catalysis community.

**Electrochemical reduction to  $C_3$  products.** This example focuses on the electrochemical reduction of  $CO_2$  on Cu toward  $C_3$  products. This process has been studied using CARE by evaluating a  $C_3O_2$  CRN on Cu(100) with GAME-Net-UQ. Electrochemical reactions require the consideration of potential and pH contributions, which CARE implements via the computational hydrogen electrode approach<sup>68</sup>. This defines a set of specific electrochemical reactions between the adsorbed species, charged particles (proton and electron) and the solvent. The CRN has been modeled in neutral (pH 7) and alkaline (pH 13) conditions, at applied potentials of  $-0.4$  and  $-1.0 V_{SHE}$ , and at ambient conditions (298 K, 1 bar). The network contains 893 surface species, 93 gas-phase molecules and 7,709 elementary reactions, classified in seven types: PCET; C–O-, C–C-, O–H-, and C–H-bond-breaking/bond-forming reactions; adsorption/desorption and [1,2]-H shifts (Fig. 4a). The CRN evaluation took 5 min on a 24-core CPU. Further characterization details of the  $C_3O_2$  network are provided in Supplementary Fig. 21.

The global analysis of the CRN (Supplementary Fig. 22) shows that PCET steps cluster into two distinct groups according to their uncertainty. The cluster with a higher uncertainty corresponds to  $-OH$  protonation, leading to water formation and elimination. Shifting toward a more alkaline pH (from pH 7 to pH 13) leads to an increase in the reaction energy for all PCET reactions. This behavior is expected since increasing the pH implies a lower concentration of protons, thermodynamically disfavoring the protonation of the intermediates. As expected, lowering the applied potential (from  $-1.0$  to  $-1.5 V_{SHE}$ ) leads to a favorable thermodynamic shift of the energies for all PCET reactions. A key problem in  $CO_2$  reduction on Cu is fine-tuning its selectivity toward long-chain products, such as 1-propanol (1-PrOH) versus  $C_3H_6$  production<sup>57</sup>. Experimentally, for  $C_2$  products, ethylene and ethanol are observed, but in the  $C_3$  fraction, only 1-PrOH is found and the route to  $C_3H_6$  is blocked. Following previous experimental and DFT investigations<sup>57</sup>, the selectivity of the network toward 1-PrOH and  $C_3H_6$  was analyzed with CARE starting from two key intermediates: allyl alcohol and propionaldehyde. In particular, the obtained trends are in agreement with those obtained from the DFT study, regardless of the absence of implicit solvation and Bader charge corrections in the energy evaluation workflow, which instead are included in the DFT treatment. This demonstrates the capacity of GAME-Net-UQ to capture essential reactivity trends without accounting for solvent and charge effects. At alkaline pH (Fig. 4b), thermodynamically favorable pathways toward the formation of 1-PrOH are identified. These findings align with DFT<sup>57</sup> and highlight the preferential formation of 1-PrOH over  $C_3H_6$ . Some discrepancies appear in the formation of  $CH_3CH_2CHOH^*$  species. Although DFT simulations suggest the oxygen hydrogenation from propionaldehyde to  $CH_3CH_2CHOH^*$  to be slightly endothermic ( $\Delta E = +0.04$  eV), GAME-Net-UQ yields a  $\Delta E$  value of  $-0.1$  eV.

Starting by allyl alcohol and under reductive potentials  $U = -1.0 V_{RHE}$  and pH 7 (Fig. 4c), GAME-Net-UQ still reproduces the DFT trends including solvation and charge corrections. Therefore, CARE can deal with electrochemical systems by integrating a specific template and evaluating the energies taking into account pH and applied potential contributions. In this way, selectivity trends comparable with DFT and experimental observations can be extracted.





**Fig. 4 | Electrochemical  $C_2O_2$  reaction network on Cu(100).** **a**, CRN visualization with intermediates (left) and elementary reactions (right) colored by phase and reaction type, respectively. **b, c**, Reaction paths for the electrocatalytic processes from propionaldehyde to 1-PrOH (**b**) and allyl alcohol to  $C_3H_6$  and 1-PrOH (**c**). The dashed lines represent the energy profiles obtained from ref. 57 with DFT, whereas solid lines show the profiles obtained using GAME-Net-UQ.

**Fischer–Tropsch synthesis.** CARE enables the generation of huge CRNs as those representing the Fischer–Tropsch process<sup>69</sup>, converting syngas into long-chain hydrocarbons. Such CRNs can neither be evaluated using DFT in reasonable times nor the MKM of networks with thousands of reactions has ever been performed before in heterogeneous catalysis<sup>11</sup>. Here we achieve these goals with CARE, using GAME-Net-UQ as the energy evaluator. To this end, a  $C_6O_1$  ( $N_C = 6$ ,  $N_O = 1$ ) CRN has been generated and evaluated for Co(0001), Fe(110), Ru(0001) and Ni(111). The CRNs include 38,913 surface species, 985 gas-phase molecules and 369,365 elementary reactions. GAME-Net-UQ, trained on fewer than 200 TS structures per metal, is used here to evaluate the TS energy of the surface-bond-forming steps, where 99.98% of them being completely new to the model (Supplementary Table 19). Considering that simulating an intermediate adsorbed on ferromagnetic surfaces with DFT on average required 307 CPU hours and about 2.5 kWh on supercomputing facilities (Supplementary Note 12), evaluating the species of this CRN with DFT would have taken  $10^6$  kWh and 1,360 CPU years, assuming sampling one configuration per adsorbate–surface pair. With CARE, the CRN blueprint generation plus the energy evaluation of its intermediates and TSs took 4 h on a 24-core CPU, resulting in a speed-up of  $10^6$ . Running microkinetic simulations on such large reaction networks fully energetically characterized (that is, both  $\Delta E$  and  $E_{act}$  evaluated for all elementary steps) has, to the best of our knowledge, never been reported in the literature of heterogeneous catalysis. We have accomplished to run kinetic simulations up to the chemical steady state on a subset of the original CRNs comprising the linear products (4,203 species and 36,843 reactions) at 523.15 K, 10 bar and  $H_2/CO = 2$ . These MKM simulations took between 6 and 50 min to converge, depending on the CRN stiffness. This achievement unlocks the possibility to quantitatively assess the power of ML models when deriving macroscale catalyst properties.

Our analysis at this point is semiquantitative to correlate the major trends experimentally observed. The energy trends obtained with GAME-Net-UQ, as well as the distribution of the derived formation rates of the products included in the CRN (Fig. 5), point to a general agreement with experimental evidence<sup>69</sup>.

Co and Fe are the industrially used catalysts. Our results (Fig. 5a) show that Co prefers the formation of C–C bonds over C–H and C–O, in agreement with its known selectivity for the formation of long-chain products. Fe, on the other hand, tends to produce more  $CO_2$  through the undesired water–gas shift reaction—a phenomenon that has been effectively captured by the kinetic results (Fig. 5b). Product distribution shows that Co and Fe are catalysts producing the highest amount of long-chain hydrocarbons.

Ni is a catalyst known for its high hydrogenation activity. This behavior is not well captured by GAME-Net-UQ, where although C–H bond-forming reactions appear as the most thermodynamically favorable, the kinetic results (Fig. 5b) show that methane is not among the main products. Ru shows a similar trend; however, its high cost hinders its industrial applicability. The Anderson–Schulz–Flory distributions (Fig. 5c) derived from the kinetic simulations are limited by the fact that the experimental linear trend is observed in the  $C_4$ – $C_{20}$  range. The absence of  $C_6$  products creates a boundary condition that leads to an accumulation of  $C_6$  species, distorting the expected slope associated with the chain-growth probability.

Kinetic simulations enable the ranking of products by their formation rates (Supplementary Table 16). Co is the only metal showing  $C_6$  species in the top ten, whereas Fe presents the highest  $CO_2$  formation rate. The kinetic simulations enable the investigation of the main steps of the chain-growth mechanism (Supplementary Tables 17 and 18). Although the initial steps of the chain growth ( $C_2 \rightarrow C_3/C_4$ ) are mainly governed by the addition of  $C_1$  intermediates, the formation of  $C_{5-6}$  species is primarily due to the combination of  $C_2 + C_3$ ,  $C_2 + C_4$  and  $C_3 + C_3$  fragments. Although these pathways are in contrast with the proposed mechanisms for the Fischer–Tropsch synthesis<sup>70</sup>, CARE represents a promising tool that enables researchers to unveil the key factors governing this critical process.

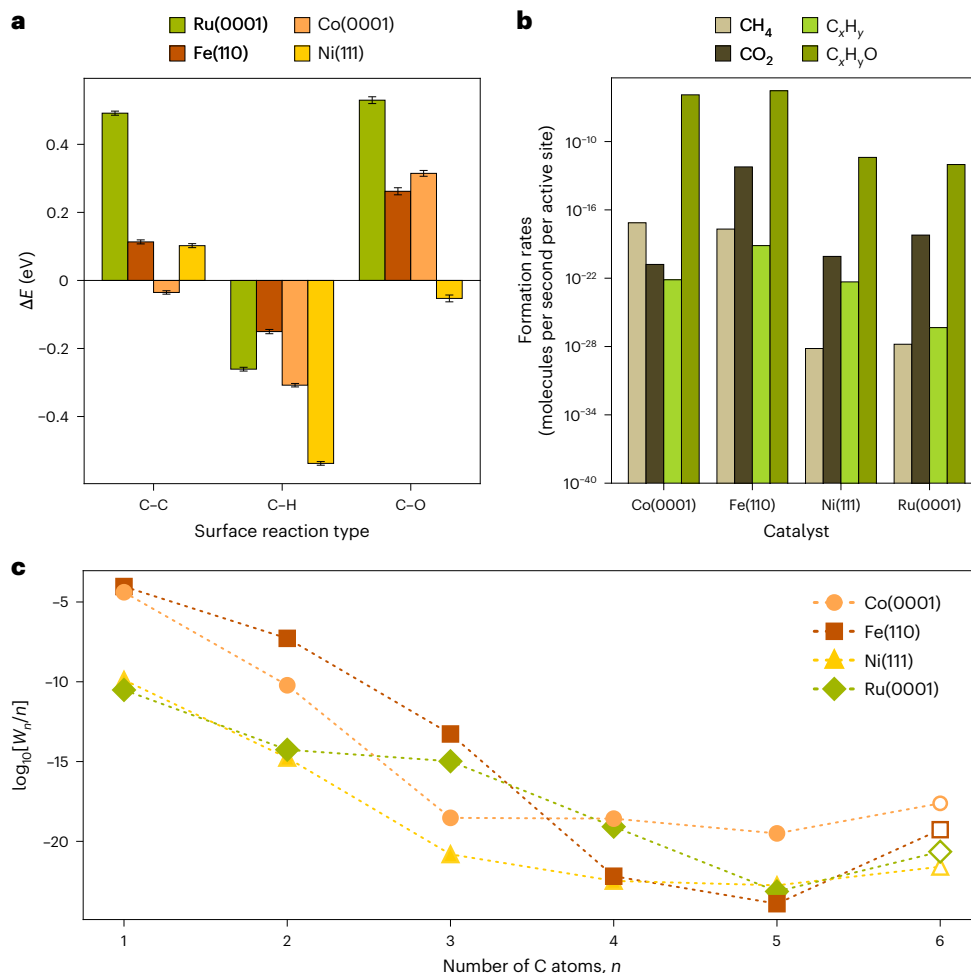
In terms of uncertainty, the reaction barriers obtained with GAME-Net-UQ show an associated uncertainty between 0.3 and 0.8 eV (Supplementary Fig. 23). The higher uncertainty in  $E_{act}$ , compared with the values reported for species and TSs, reflects the additive error of individual species and TS energies when deriving the reaction properties. The lowest uncertainty is obtained for C–C couplings with values of around 0.3 eV, whereas hydrogenation steps show an uncertainty distribution of around 0.6 eV for  $C_xH_y$  species, and a second peak at around 0.8 eV for small open-shell and oxygenated intermediates. When evaluating the TS energy  $E_{TS}$  of the surface reactions, GAME-Net-UQ predicts valid values ( $E_{IS} < E_{TS} < E_{FS}$ ) 83% of the times (Supplementary Table 14). The invalid evaluations mainly correspond to endothermic C–H steps, where the encountered situation is  $E_{TS} < E_{FS}$ , pointing to an overestimation of the FS energy due to absence of stabilizing lateral interactions in the treatment.

In summary, this case study highlights the capability of CARE to generate and evaluate huge networks cheaply as well as including the possible bond-breaking reactions that could eventually take place on the catalyst surface. Additionally, it demonstrates the power of the microkinetic module in CARE, unlocking the possibility to run simulations of networks containing thousands of reactions.

## Discussion

CARE represents a robust software for evaluating the activity of heterogeneous catalysts in thermal and electro-conditions, consisting of a reaction network generator powered by state-of-the-art ML energy estimators, and a reactor microkinetic model. Several potential challenges in the field must be considered.

For the definition of CRNs, we have approached the network completeness with a set of templates, ensuring the inclusion of all potential bond-breaking reactions that leads to all the available intermediate species—even those that might be initially discarded based on chemical intuition. CARE is modular in nature, and allows the user to implement additional templates for increasing the definition of the explored CS.



**Fig. 5 | Energy and kinetic trends obtained from the C<sub>6</sub>O<sub>1</sub> reaction network for the Fischer–Tropsch synthesis. a**, Mean reaction energy  $\Delta E$  obtained using GAME-Net-UQ for the surface reactions grouped by the bond-forming type and metal surface. Error bars correspond to the 95% confidence interval.  $n_{\text{C-C}} = 164,602$ ,  $n_{\text{C-H}} = 143,445$ ,  $n_{\text{C-O}} = 49,177$ . **b**, Steady-state formation rates

obtained from the kinetic simulations on the C<sub>6</sub>O<sub>1</sub> CRNs containing linear products on Co, Fe, Ni and Ru. **c**, Associated Anderson–Schulz–Flory distribution of the predicted products ( $W_n$ , mass formation rate of C<sub>*n*</sub> species). Operating conditions:  $T = 523.15$  K,  $P = 10$  bar,  $\text{H}_2/\text{CO} = 2$ .

This systematic approach reduces bias in CRNs, but leads to larger reaction lists that can be a bottleneck when analyzing and manipulating the constructed networks. However, there are three definitive advantages in the methodology. The first is that due to its modularity, CARE can be easily extended to tackle more exotic reactions, such as isomerizations, cycle formations, and those containing heteroatoms (N, S and P) by adding reaction templates and fine-tuning the ML evaluator of choice accordingly. Second, CARE can consider rare events easily, providing the possibility to study their implication in catalyst deactivation<sup>71</sup>, an aspect overlooked and difficult to treat with DFT simulations but crucial for inferring the long-term stability of the catalyst. The third aspect is the substantial reduction in computational cost by accelerating and automatizing the evaluation of reaction properties with faster ML, such as GAME-Net-UQ and MLIPs.

Our data-driven energy estimator is fast and relatively cheap due to its size and direct approach for predicting the DFT energy of relaxed structures. The UQ of the predictions provides additional information that can be propagated to reaction properties and MKM, or exploited for active learning purposes and Bayesian optimization. The low computational cost of GAME-Net-UQ and its balanced accuracy for intermediates and TSs remove the need for empirical rules like linear scaling relationships, since GAME-Net-UQ covers a wider diversity in terms of metals, surfaces and species. GAME-Net-UQ allows the energy estimation of thousands of intermediates/TSs in less than 1 CPU hour,

thereby being a viable ML strategy for fast property prediction. Additional challenges in this area are linked to the introduction of lateral interactions (high coverage or solvents), more detailed configurational analysis for adsorbates, and expansion of the domain of applicability to structurally complex materials such as alloys and oxides<sup>67</sup>. These aspects could be addressed by implementing lateral interactions via graph-based strategies<sup>72</sup>, active learning and  $\Delta$ -ML schemes<sup>73</sup>.

As for the challenges in the microkinetic part, the Julia solver implemented in CARE provides the computational scalability required to handle networks containing thousands of reactions. Future improvements would require strategies to include thermal and entropic contributions and a more detailed representation of electrochemical events. Clever network-pruning strategies are needed, particularly for huge networks, to facilitate the convergence of the associated kinetic simulations, and rate and selectivity control functionalities, essential to unveil important catalytic pathways. Finally, more complex reactor models able to capture relevant phenomena at the nanoparticle scale could be implemented.

Regarding CARE, the results highlight the power of direct energy estimators bypassing the need to perform the structural optimization of intermediates, to provide rankings for the catalytic performance. These models enable to focus on the overall catalytic reactivity, making substantial advances at 1% of the cost of MLIPs. The capability of CARE for reproducing experimental trends showcases its ability to rank potential materials and can be used to predict better formulations.

In summary, CARE paves the way for the automated exploration and characterization of heterogeneous catalytic processes with improved network completeness and reduced bias in an automated, robust and accurate manner. Our work opens a path toward the study of processes involving complex compounds such as plastics and biomass at a reasonable computational cost.

## Methods

### DFT

The DFT simulations for developing GAME-Net-UQ have been performed with the Vienna ab initio simulation package (v. 5.4.4)<sup>74</sup>. The Perdew–Burke–Ernzerhof<sup>75</sup> functional with reparameterized D2 (ref. 76) dispersion correction for metals was used<sup>77</sup>. Core electrons were represented by projector augmented-wave pseudopotentials<sup>78</sup> and valence electrons were represented by plane waves with a kinetic energy cutoff of 450 eV. Electronic convergence was set to 10<sup>-5</sup> eV and the atomic positions were converged until the residual forces fell below 0.03 eV Å<sup>-1</sup>.

The database provided in GAME-Net-UQ includes eight face-centered cubic, one body-centered cubic and five hexagonal close-packed metal structures on the three most stable *hkl* surface facets for each one (Supplementary Fig. 2). Metal surfaces were modeled by four-to-ten-layer slabs (Supplementary Table 1), where the half uppermost layers were fully relaxed and the bottom ones were fixed to the bulk distances. A surface coverage of 0.02 molecules Å<sup>-2</sup> was defined for all the adsorption structures, a reasonable value to neglect lateral interactions. The vacuum between the slabs was set between 13 and 16 Å, and the dipole correction was applied in the *z* direction<sup>79</sup>. The Brillouin zone was sampled by a  $\Gamma$ -centered 3 × 3 × 1 *k*-point mesh generated through the Monkhorst–Pack method<sup>80</sup>. TSs have been obtained using the improved dimer method<sup>81</sup> and validated using DFT frequency analysis displacing adsorbate atoms by 0.015 Å (Supplementary Fig. 6) following the criterion in ref. 82.

### MKM

Once the CRNs are constructed, three components are required for running microkinetic simulations.

(1) The stoichiometric matrix  $\mathbf{v} \in \mathbb{Z}^{n_s \times n_R}$  represents the reaction network with  $n_s$  species (including the surface site <sup>\*</sup>) and  $n_R$  elementary reactions.  $v_{i,j}$  is the stoichiometric coefficient of species *i* in reaction *j*. As elementary reactions are mono- or bimolecular, stoichiometric coefficients lie in the interval [-2;+2]. This matrix is intrinsically sparse, and since most of the reactions in the CRN are of the kind A + B → C + D, its sparsity  $s_v$  follows the relationship (Supplementary Fig. 12)

$$s_v = \left(1 - \frac{4}{n_s}\right) \times 100\% \quad (1)$$

Compressed sparse row/column formats and integer data type have to be preferred to optimize matrix multiplications for CRNs with more than 400 intermediates ( $s_v = 99\%$ ).

(2) The operating conditions such as temperature and pressure (and applied potential and electrolyte pH for electrochemical systems) define the thermodynamic and kinetic constants of the elementary reactions (Supplementary Note 8).

(3) The last building block is the reactor model. The default in CARE is a zero-conversion differential plug-flow reactor defined by the system of ODEs:

$$\begin{cases} \frac{d\theta_i}{dt} = \sum_{j=1}^{n_R} v_{i,j} r_j & \text{(adsorbed species)} \\ \frac{dP_i}{dt} = 0 & \text{(gas species)} \end{cases}, \quad (2)$$

where  $\theta_i$  is the fractional surface coverage of species *i*,  $v_{i,j}$  is the stoichiometric coefficient of species *i* in the reaction *j* and  $P_i$  is the partial

pressure of the gas-phase reactant *i*.  $r_j$  is the net rate of reaction *j*, defined as

$$r_j = \vec{r}_j - \overleftarrow{r}_j = \vec{k}_j \prod_{i=1}^{n_s} a_i^{|\min(v_{i,j},0)|} - \overleftarrow{k}_j \prod_{i=1}^{n_s} a_i^{|\max(v_{i,j},0)|}, \quad (3)$$

where  $\vec{r}_j$  and  $\overleftarrow{r}_j$  are the rates of reaction *j* in the forward and reverse directions, respectively;  $\vec{k}_j$  and  $\overleftarrow{k}_j$  are the corresponding kinetic coefficients; and  $a_i$  refers to the activity of species *i*, corresponding to  $\theta_i$  for adsorbed species and  $P_i$  for gas molecules. The initial conditions assume an empty surface exposed to the gas mixture defined by the user:

$$\begin{cases} \theta_* (t=0) = 1 & \text{(active site)} \\ \theta_k (t=0) = 0 & \text{(surface species)} \\ P_i (t=0) = P_i^0 & i \in \text{gas reactants} \\ P_j (t=0) = 0 & j \in \text{gas products} \end{cases} \quad (4)$$

To facilitate the convergence of highly stiff microkinetic simulations and to reduce the total ODE integration time, the analytical Jacobian of the ODE is defined and fallback strategies are additionally available. These include (1) removing surface reactions with energy barriers higher than a specified threshold, (2) considering all the reactions as barrierless (Supplementary Fig. 14) and (3) steering the model toward specified target products by removing species from the CRN that are unlikely to be formed based on user intuition. Steady state is reached when the following two conditions are satisfied:

$$\begin{cases} \sum_{i=1}^{n_s} \theta_i = 1 & \text{(adsorbed species)} \\ \sum_{i=1}^{\text{reactants}} R_i n_{i,k} = \sum_{i=1}^{\text{products}} R_i n_{i,k} & k = \text{C, H, O} \end{cases}, \quad (5)$$

where  $R_i$  is the consumption/formation rate of the gas reactant/product and  $n_{i,k}$  is the number of *k* atoms in the molecule. Here 64-bit floating-point arithmetic is implemented by default for the microkinetic simulations. For extremely stiff systems, the user has the option to run the simulation with higher precision when using the Julia solver (for example, BigFloat can be used for arbitrary-precision floating-point arithmetic). CARE uses the backward differentiation formula as the default ODE solver. Two essential ODE solver parameters are the absolute and relative tolerances ( $\text{atol}$  and  $\text{rtol}$ , respectively), which control the accuracy of the predicted surface coverages at each integration step. Initial  $\text{atol}$  and  $\text{rtol}$  values are conservatively set to 10<sup>-15</sup> and 10<sup>-12</sup>, respectively.

### Computational tools

CARE is written in Python (v. 3.12), and the CRN algorithm is mainly based on RDKit (v. 2023.09.6), ASE (v. 3.24.0)<sup>83</sup> and NetworkX (v. 3.2.1)<sup>84</sup>. Surfaces from the Materials Project have been obtained through `mp-api` (v. 0.45.3). GAME-Net-UQ has been developed with PyTorch Geometric (v. 2.5.3)<sup>85</sup> and PyTorch (v. 2.4.0)<sup>86</sup>, and it was trained on an NVIDIA RTX A2000 12-GB GPU with CUDA 12.3. MKM functionalities are implemented with SciPy (v. 1.12.0)<sup>54</sup> and DifferentialEquations.jl<sup>55</sup> (Julia (v. 1.10.2)). Visualizations have been created with Matplotlib (v. 3.8.0), Seaborn (v. 0.12.2) and Inkscape (v. 1.3.2). All runs with CARE have been performed on a 24-core 12th Gen Intel Core i9-12900K CPU and an NVIDIA RTX A2000 12-GB GPU with CUDA 12.2.

### Data availability

The DFT data used to develop GAME-Net-UQ are available via ioChem-BD at <https://doi.org/10.19061/iochem-bd-1-257> (FG-dataset from GAME-Net training dataset) and <https://doi.org/10.19061/iochem-bd-1-328> (new data). Results associated with GAME-Net-UQ and the

presented case studies are available via Zenodo at <https://doi.org/10.5281/zenodo.17977395> (ref. 87). Source data are provided with this paper.

## Code availability

CARE implementation has been publicly released under the MIT license and is available via GitHub at <https://github.com/LopezGroup-ICIQ/care>. The code for training and evaluating GAME-Net-UQ is also available via GitHub at [https://github.com/LopezGroup-ICIQ/gamenet\\_uq](https://github.com/LopezGroup-ICIQ/gamenet_uq). To improve the reproducibility of this work, CARE v0.1.3 has been frozen and uploaded to Zenodo at <https://doi.org/10.5281/zenodo.17977395> (ref. 87).

## References

1. Wen, M. et al. Chemical reaction networks and opportunities for machine learning. *Nat. Comput. Sci.* **3**, 12–24 (2023).
2. Unsleber, J. P. & Reiher, M. The exploration of chemical reaction networks. *Annu. Rev. Phys. Chem.* **71**, 121–142 (2020).
3. Broadbelt, L. J., Stark, S. M. & Klein, M. T. Computer generated pyrolysis modeling: on-the-fly generation of species, reactions, and rates. *Ind. Eng. Chem. Res.* **33**, 790–799 (1994).
4. Feinberg, M. *Foundations of Chemical Reaction Network Theory* (Springer, 2019).
5. Garay-Ruiz, D. & Bo, C. Revisiting catalytic cycles: a broader view through the energy span model. *ACS Catal.* **10**, 12627–12635 (2020).
6. Newman, M. E. J. The structure and function of complex networks. *SIAM Rev.* **45**, 167–256 (2003).
7. Tyson, J. J. & Novák, B. Functional motifs in biochemical reaction networks. *Annu. Rev. Phys. Chem.* **61**, 219–240 (2010).
8. Wong, A. S. Y. & Huck, W. T. S. Grip on complexity in chemical reaction networks. *Beilstein J. Org. Chem.* **13**, 1486–1497 (2017).
9. Hashemi, A., Bougueroua, S., Gaigeot, M.-P. & Pidko, E. A. ReNeGate: a reaction network graph-theoretical tool for automated mechanistic studies in computational homogeneous catalysis. *J. Chem. Theory Comput.* **18**, 7470–7482 (2022).
10. Steiner, M. & Reiher, M. Autonomous reaction network exploration in homogeneous and heterogeneous catalysis. *Top. Catal.* **65**, 6–39 (2022).
11. Stocker, S., Csányi, G., Reuter, K. & Margraf, J. T. Machine learning in chemical reaction space. *Nat. Commun.* **11**, 5505 (2020).
12. Unsleber, J. P., Grimmel, S. A. & Reiher, M. Chemoton 2.0: autonomous exploration of chemical reaction networks. *J. Chem. Theory Comput.* **18**, 5393–5409 (2022).
13. Zhao, Q. & Savoie, B. M. Simultaneously improving reaction coverage and computational cost in automated reaction prediction tasks. *Nat. Comput. Sci.* **1**, 479–490 (2021).
14. Zhao, Q., Xu, Y., Greeley, J. & Savoie, B. M. Deep reaction network exploration at a heterogeneous catalytic interface. *Nat. Commun.* **13**, 4860 (2022).
15. Jafari, M. & Zimmerman, P. M. Uncovering reaction sequences on surfaces through graphical methods. *Phys. Chem. Chem. Phys.* **20**, 7721–7729 (2018).
16. Maeda, S. et al. Implementation and performance of the artificial force induced reaction method in the GRRM17 program. *J. Comput. Chem.* **39**, 233–251 (2018).
17. Chen, D., Shang, C. & Liu, Z.-P. Machine-learning atomic simulation for heterogeneous catalysis. *NPJ Comput. Mater.* **9**, 2 (2023).
18. Rangarajan, S., Bhan, A. & Daoutidis, P. Language-oriented rule-based reaction network generation and analysis: description of ring. *Comput. Chem. Eng.* **45**, 114–123 (2012).
19. 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.* **8**, 14621 (2017).
20. Benson, S. W. & Buss, J. H. Additivity rules for the estimation of molecular properties. Thermodynamic properties. *J. Chem. Phys.* **29**, 546–572 (1958).
21. Sabbe, M. K. et al. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *J. Phys. Chem. A* **109**, 7466–7480 (2005).
22. Wittreich, G. R. & Vlachos, D. G. Python Group Additivity (pGrAdd) software for estimating species thermochemical properties. *Comput. Phys. Commun.* **273**, 108277 (2022).
23. Abild-Pedersen, F. et al. Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces. *Phys. Rev. Lett.* **99**, 016105 (2007).
24. Calle-Vallejo, F., Loffreda, D., Koper, M. T. M. & Sautet, P. Introducing structural sensitivity into adsorption–energy scaling relations by means of coordination numbers. *Nat. Chem.* **7**, 403–410 (2015).
25. Bronsted, J. N. Acid and basic catalysis. *Chem. Rev.* **5**, 231–338 (1928).
26. Evans, M. G. & Polanyi, M. Inertia and driving force of chemical reactions. *Trans. Faraday Soc.* **34**, 11–24 (1938).
27. Nørskov, J. K. et al. Universality in heterogeneous catalysis. *J. Catal.* **209**, 275–278 (2002).
28. Liu, M. et al. Reaction mechanism generator v3.0: advances in automatic mechanism generation. *J. Chem. Inf. Model.* **61**, 2686–2696 (2021).
29. Kreitz, B. et al. Automated generation of microkinetics for heterogeneously catalyzed reactions considering correlated uncertainties. *Angew. Chem. Int. Ed.* **62**, e202306514 (2023).
30. Kreitz, B. et al. Detailed microkinetics for the oxidation of exhaust gas emissions through automated mechanism generation. *ACS Catal.* **12**, 11137–11151 (2022).
31. Mou, T. et al. Bridging the complexity gap in computational heterogeneous catalysis with machine learning. *Nat. Catal.* **6**, 122–136 (2023).
32. Chanussot, L. et al. Open Catalyst 2020 (OC20) dataset and community challenges. *ACS Catal.* **11**, 6059–6072 (2021).
33. Pablo-García, S. et al. Fast evaluation of the adsorption energy of organic molecules on metals via graph neural networks. *Nat. Comput. Sci.* **3**, 433–442 (2023).
34. Suvarna, M. & Pérez-Ramírez, J. Embracing data science in catalysis research. *Nat. Catal.* **7**, 624–635 (2024).
35. Liao, Y.-L., Wood, B., Das, A. & Smidt, T. EquiformerV2: improved equivariant transformer for scaling to higher-degree representations. Preprint at <http://arxiv.org/abs/2306.12059> (2023).
36. Wander, B., Shuaibi, M., Kitchin, J. R., Ulissi, Z. W. & Zitnick, C. L. CatTSunami: accelerating transition state energy calculations with pretrained graph neural networks. *ACS Catal.* **15**, 5283–5294 (2025).
37. Göttl, F. & Mavrikakis, M. Generalized Brønsted-Evans-Polanyi relationships for reactions on metal surfaces from machine learning. *ChemCatChem* **14**, e202201108 (2022).
38. Hutton, D. J., Cordes, K. E., Michel, C. & Göttl, F. Machine learning-based prediction of activation energies for chemical reactions on metal surfaces. *J. Chem. Inf. Model.* **63**, 6006–6013 (2023).
39. 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.* **122**, 156001 (2019).
40. Stocker, S. et al. Estimating free energy barriers for heterogeneous catalytic reactions with machine learning potentials and umbrella integration. *J. Chem. Theory Comput.* **19**, 6796–6804 (2023).

41. 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.* **9**, 180 (2023).
42. Batatia, I. et al. A foundation model for atomistic materials chemistry. *J. Chem. Phys.* **163**, 184110 (2025).
43. Neumann, M. et al. Orb: a fast, scalable neural network potential. Preprint at <http://arxiv.org/abs/2410.22570> (2024).
44. Wang, B., Gu, T., Lu, Y. & Yang, B. Prediction of energies for reaction intermediates and transition states on catalyst surfaces using graph-based machine learning models. *Mol. Catal.* **498**, 111266 (2020).
45. Kim, S., Woo, J. & Kim, W. Y. Diffusion-based generative AI for exploring transition states from 2D molecular graphs. *Nat. Commun.* **15**, 341 (2024).
46. Du, J., Zhang, S., Wu, G., Moura, J. M. F. & Kar, S. Topology adaptive graph convolutional networks. Preprint at <http://arxiv.org/abs/1710.10370> (2018).
47. Hirschfeld, L., Swanson, K., Yang, K., Barzilay, R. & Coley, C. W. Uncertainty quantification using neural networks for molecular property prediction. *J. Chem. Inf. Model.* **60**, 3770–3780 (2020).
48. Martí, C. et al. DockOnSurf: a Python code for the high-throughput screening of flexible molecules adsorbed on surfaces. *J. Chem. Inf. Model.* **61**, 3386–3396 (2021).
49. Han, S., Lysgaard, S., Vegge, T. & Hansen, H. A. Rapid mapping of alloy surface phase diagrams via Bayesian evolutionary multitasking. *NPJ Comput. Mater.* **9**, 139 (2023).
50. Mazitov, A. et al. PET-MAD as a lightweight universal interatomic potential for advanced materials modeling. *Nat. Commun.* **16**, 10653 (2025).
51. Park, Y., Kim, J., Hwang, S. & Han, S. Scalable parallel algorithm for graph neural network interatomic potentials in molecular dynamics simulations. *J. Chem. Theory Comput.* **20**, 4857–4868 (2024).
52. Motagamwala, A. H. & Dumesic, J. A. Microkinetic modeling: a tool for rational catalyst design. *Chem. Rev.* **121**, 1049–1076 (2021).
53. Chorkendorff, I. & Niemantsverdriet, J. *Concepts of Modern Catalysis and Kinetics* (Wiley-VCH Verlag, 2003).
54. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
55. Rackauckas, C. & Nie, Q. DifferentialEquations.jl—a performant and feature-rich ecosystem for solving differential equations in Julia. *J. Open Res. Softw.* **5**, 15 (2017).
56. Burte, A. S. et al. CatTestHub: a benchmarking database of experimental heterogeneous catalysis for evaluating advanced materials. *J. Catal.* **442**, 115902 (2025).
57. Pablo-García, S. et al. Mechanistic routes toward C<sub>3</sub> products in copper-catalysed CO<sub>2</sub> electroreduction. *Catal. Sci. Technol.* **12**, 409–417 (2022).
58. Olah, G. A. Beyond oil and gas: the methanol economy. *Angew. Chem. Int. Ed.* **44**, 2636–2639 (2005).
59. Deka, T. J., Osman, A. I., Baruah, D. C. & Rooney, D. W. Methanol fuel production, utilization, and techno-economy: a review. *Environ. Chem. Lett.* **20**, 3525–3554 (2022).
60. Tran, R. et al. Surface energies of elemental crystals. *Sci. Data* **3**, 160080 (2016).
61. García-Muelas, R., Li, Q. & López, N. Density functional theory comparison of methanol decomposition and reverse reactions on metal surfaces. *ACS Catal.* **5**, 1027–1036 (2015).
62. Tuxen, A. et al. Size-dependent dissociation of carbon monoxide on cobalt nanoparticles. *J. Am. Chem. Soc.* **135**, 2273–2278 (2013).
63. Yang, C., Wu, C., Xie, W., Xie, D. & Hu, P. General reactive element-based machine learning potentials for heterogeneous catalysis. *Nat. Catal.* **8**, 891–904 (2025).
64. Deng, B. et al. Systematic softening in universal machine learning interatomic potentials. *NPJ Comput. Mater.* **11**, 9 (2025).
65. Honkala, K. et al. Ammonia synthesis from first-principles calculations. *Science* **307**, 555–558 (2005).
66. Sutton, J. E., Guo, W., Katsoulakis, M. A. & Vlachos, D. G. Effects of correlated parameters and uncertainty in electronic-structure-based chemical kinetic modelling. *Nat. Chem.* **8**, 331–337 (2016).
67. Chen, B. W. J. & Mavrikakis, M. Modeling the impact of structure and coverage on the reactivity of realistic heterogeneous catalysts. *Nat. Chem. Eng.* **2**, 181–197 (2025).
68. Nørskov, J. K. et al. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. *J. Phys. Chem. B* **108**, 17886–17892 (2004).
69. Mahmoudi, H. et al. A review of Fischer Tropsch synthesis process, mechanism, surface chemistry and catalyst formulation. *Biofuels Eng.* **2**, 11–31 (2017).
70. van Santen, R. A., Markvoort, A. J., Filot, I. A. W., Ghouri, M. M. & Hensen, E. J. M. Mechanism and microkinetics of the Fischer–Tropsch reaction. *Phys. Chem. Chem. Phys.* **15**, 17038–17063 (2013).
71. Martín, A. J., Mitchell, S., Mondelli, C., Jaydev, S. & Pérez-Ramírez, J. Unifying views on catalyst deactivation. *Nat. Catal.* **5**, 854–866 (2022).
72. Deshpande, S., Maxson, T. & Greeley, J. Graph theory approach to determine configurations of multidentate and high coverage adsorbates for heterogeneous catalysis. *NPJ Comput. Mater.* **6**, 79 (2020).
73. Ramakrishnan, R., Dral, P. O., Rupp, M. & von Lilienfeld, O. A. Big data meets quantum chemistry approximations: the  $\delta$ -machine learning approach. *J. Chem. Theory Comput.* **11**, 2087–2096 (2015).
74. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
75. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
76. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).
77. Almora-Barrios, N., Carchini, G., Błoński, P. & López, N. Costless derivation of dispersion coefficients for metal surfaces. *J. Chem. Theory Comput.* **10**, 5002–5009 (2014).
78. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
79. Neugebauer, J. & Scheffler, M. Adsorbate-substrate and adsorbate-adsorbate interactions of Na and K adlayers on Al(111). *Phys. Rev. B* **46**, 16067–16080 (1992).
80. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).
81. Heyden, A., Bell, A. T. & Keil, F. J. Efficient methods for finding transition states in chemical reactions: comparison of improved dimer method and partitioned rational function optimization method. *J. Chem. Phys.* **123**, 224101 (2005).
82. Brogaard, R. Y. et al. Methanol-to-hydrocarbons conversion: the alkene methylation pathway. *J. Catal.* **314**, 159–169 (2014).
83. Larsen, A. H. et al. The atomic simulation environment—a Python library for working with atoms. *J. Phys. Condens. Matter* **29**, 273002 (2017).
84. Hagberg, A. A., Schult, D. A. & Swart, P. J. Exploring network structure, dynamics, and function using networkx. In *Proc. 7th Python in Science Conference* 11–15 (SciPy, 2008).
85. Fey, M. & Lenssen, J. E. Fast graph representation learning with PyTorch geometric. Preprint at <http://arxiv.org/abs/1903.02428> (2019).

86. Paszke, A. et al. PyTorch: an imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems* (eds Wallach, H. et al.) 8024–8035 (Curran Associates, Inc., 2019).
87. Morandi, S. & Loveday, O. Data associated to CARE. *Zenodo* <https://doi.org/10.5281/zenodo.17977395> (2026).

## Acknowledgements

S.M., O.L., R.R.S., P.S.B. and N.L. thank the Spanish Ministry of Science and Innovation (PID2024-122516OBIOO and PRE2022-101291) for financial support. This publication was created as part of NCCR Catalysis (grant number 180544), a National Centre of Competence in Research funded by the Swiss National Science Foundation. O.L. acknowledges the Joan Oró Predoctoral Programme of the Generalitat de Catalunya, and the European Social Fund Plus (2023 FI-1 00769). T.R. acknowledges support from the Erasmus+ program of the European Union. S.P.-G. acknowledges support from the US Department of Energy, Office of Science, Subaward by University of Minnesota, with the project title 'Development of Machine Learning and Molecular Simulation Approaches to Accelerate the Discovery of Porous Materials for Energy-Relevant Applications' (DE-SC0023454). R.R.S. acknowledges funding from European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement number 754510. A.A.-G. acknowledges support from the Canada 150 Research Chairs Program as well as A. G. Frøseth. We thank BSC-RES for generously providing the computational resources.

## Author contributions

Conceptualization, S.M., O.L., S.P.-G., R.A.V.-H., N.L. Data generation and curation, S.M., T.R., R.G.-M. Formal analysis, S.M., O.L. Investigation, S.M., O.L. Methodology, S.M., O.L. Software, S.M., O.L., T.R., R.R.S., P.S.B. Visualization, S.M., O.L., P.S.B. Writing—original draft preparation, S.M., O.L., N.L. Writing—review and editing, S.M., O.L., T.R., S.P.-G., R.A.V.-H., R.R.S., P.S.B., R.G.-M., A.A.-G., N.L. Funding acquisition, A.A.-G., N.L. Project administration, N.L. Supervision, N.L.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44286-026-00361-8>.

**Correspondence and requests for materials** should be addressed to Núria López.

**Peer review information** *Nature Chemical Engineering* thanks Fanglin Che, Brett Savoie and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026