

Data-driven massive reaction networks reveal mechanistic pathways underlying catalytic CO₂ hydrogenation

Corresponding Author: Professor Ananth Govind Rajan

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Anand M. Verma et. al combined DFT and machine learning to prediction activation barrier of CO₂ hydrogenation on copper, and a Cu/SiO₂ sample was synthesized to validate products prediction from MKM model. It was seen that the involved reaction network was exhaustive and the database for activation barriers for CO₂ hydrogenation was massive. I thought this work was meaningful for the methodology to explore catalysis mechanism, nevertheless, I have several issues to confirm for this manuscript on this version:

- (1) In the Introduction section, authors mentioned: "when Cu combined with metals such as Fe, C-C coupling reactions would happen and generate >C₂ products." Why did they mention Fe species here? Was there any specific reason?
- (2) I thought one contribution of this study was that authors developed an optimization framework to predict the elementary reactions for CO₂ hydrogenation reaction. Was this prediction process influenced by employed specific catalyst actually? I was wondering the efficiency of this approach to unveil the reaction rationale and then to guide the catalyst optimization compared with current method (experimental results + theoretical calculation for mechanism exploration). Since methanol presented higher value than CO as CO₂ hydrogenation products, the dominating CO selectivity over Cu/SiO₂ system in this study really raised my interest to ask "why didn't authors employ well known Cu/ZnAl or other well estimated oxide system for methanol synthesis via thermocatalysis to predict the elementary steps?"
- (3) In this study, authors proposed many detailed reaction pathways and intermediates for CO₂ hydrogenation, such as hydrogenation of CO₂ to formate via surface hopping of the hydrogen atom of an adsorbed hydrogen molecule, and more detailed routes for CH₃OH and CO were proposed in section of "Revealing the precise mechanistic pathways of the major products". For energy barriers, it was not easy to examined from experimental results, however, I though the experiments to record those aforementioned intermediates and their variation trend were necessary to be conducted to validate authors' prediction results.
- (4) For the experimental results, authors mentioned "for the automated MKM model based on the extensive set of reactions, the production rates and product distribution closely followed experimental trends (Figure 5)". The products distribution was seen in Fig. 5i, but where were the products rates, which was also imperative to validate the number in Fig. 4c. In addition, HCOOH presented comparable rates with methanol in Fig. 4c, but no HCOOH product was detected in Fig. 5i. Was this induced by strong adsorption of HCOOH on sample or the detection limitation of HCOOH by GC?

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

In this work, DFT calculations are combined with machine learning for kinetic modeling of the CO₂ hydrogenation process. The data-driven approach is extensive and valuable; however, the experimental validation remains limited, and the novel insights derived from the modeling are not sufficiently emphasized. My specific comments are as follows:

1. The use of an in-house DFT dataset (comprising 304 reaction barriers) for training the ML model effectively mitigates issues related to data heterogeneity, thereby enhancing model accuracy and reliability. However, this approach may also constrain the model's generalizability to broader catalytic reaction systems. The authors are suggested to discuss the potential impact of using a relatively homogeneous dataset on the model's generalization capability.
2. The AFE-AE model does not include a feature importance analysis. Given that activation free energy is predicted from

activation energy, a high-performing model might achieve comparable accuracy using only activation energy and temperature as input features. Additionally, the relatively low R^2 score of the AE-RE model suggests limited generalization capability, which warrants further discussion.

3. The introduction section primarily focuses on describing the present work, rather than providing a concise overview of relevant prior studies. It would be beneficial to include a brief summary of previously reported works on similar topics to establish the research context. Additionally, content specific to this study should be shifted to later sections rather than being included in the introduction.

4. The model assumes that the catalyst surface remains static, which may not accurately reflect real-world catalytic conditions. To enhance the model's realism, it would be valuable to incorporate surface reconstruction or dynamic changes in catalyst structure.

5. Although CO_2 hydrogenation to methanol has been extensively studied, the dataset used for ML model training is relatively small (302 reactions). This limitation may hinder the model's generalization capability. Expanding the training dataset could significantly improve predictive accuracy.

6. While automated reaction enumeration eliminates human bias, it also generates a substantial number of impractical or redundant reactions. The authors are suggested to refine the reaction network by drawing upon experimentally validated reaction pathways for CO_2 hydrogenation to methanol. This approach would enhance computational efficiency and improve the model's practical relevance.

7. The experimental validation is conducted using Cu/SiO_2 catalysts, but the range of catalysts examined remains limited. Expanding the selection to include other supported metal catalysts (e.g., Cu/ZnO , Cu/ZrO_2) or metal oxide catalysts (e.g., ZnO-ZrO_2 , $\text{In}_2\text{O}_3\text{-ZrO}_2$) would strengthen the reliability and broader applicability of the findings.

8. The scope of experimental validation is rather limited. Certain conclusions, such as the increase/decrease in methanol/ CO selectivity with rising temperature, are well-established and do not necessarily require validation through modeling. The study would benefit from highlighting novel predictions derived from the model. Moreover, a more quantitative comparison between experimental and model results would be desirable. Additionally, the experimental conditions explored—particularly the temperature range—are somewhat narrow. Expanding the experimental parameters (e.g., pressure, space velocity) would enhance the robustness of model validation.

9. The modeling effort is extensive, but readers would expect to gain clear guidance or new insights from the results to facilitate rational catalytic system design for targeted product formation. The discussion should emphasize actionable insights derived from the model.

10. The title states that the study “reveals new pathways underlying catalytic CO_2 hydrogenation.” However, the discussion on newly identified reaction pathways appears limited. The authors should clarify and substantiate this claim.

11. Additional comments: The manuscript contains subjective descriptions and potentially disputable statements that should be reconsidered. For example:

- In the introduction: “it would take ≈ 47 years to compute all activation barriers with 240 CPU cores.”
- On page 7: “Overall, the proposed reaction network is the most comprehensive one for C1 to C4 pathways so far...”

(Remarks on code availability)

The code provide a README file with enough instructions for installing and running the application.

Reviewer #3

(Remarks to the Author)

This manuscript introduces a multi-faceted approach to elucidate complex catalytic mechanisms. The authors combine several advanced techniques in an end-to-end workflow: extensive density functional theory (DFT) calculations, an automated reaction enumeration algorithm (inspired by human chemical intuition), machine-learning (ML) models to predict activation energy barriers, and comprehensive microkinetic modeling. Previous studies of CO_2 hydrogenation mechanisms were forced to examine only a few dozen reactions due to computational limits, or to rely on simplified scaling relations for kinetics. By contrast, this work's data-driven pipeline enables the automated construction and evaluation of a complex reaction network (on the order of 10^3 – 10^4 reactions), which marks a clear departure from conventional, small-scale microkinetic models. Based the complex reaction network, they found that direct H_2 (gas-phase molecule) can participate in surface hydrogenation steps (an Eley–Rideal-type mechanism) is novel for CO_2 hydrogenation on Cu, as such steps are rarely considered in traditional mechanisms. The following issues need to be addressed for this work.

(1) The “elementary reaction identifier” was validated on only 388 manually labeled reactions, yet is then applied to extract 9,237 elementary steps from an initial pool of 104,723 reactions. Without further benchmarking on a larger or more chemically diverse test set, there is a risk that many non-elementary (concerted or multi-step) reactions slipped through, or genuine elementary steps were discarded.

(2) The microkinetic modeling treats all elementary steps on $\text{Cu}(111)$ as happening on isolated, identical sites under mean field conditions, ignoring lateral adsorbate–adsorbate interactions and coverage dependent energetics. At the high coverages implicit in the 50 bar, such interactions can significantly alter barrier heights and site availability, potentially invalidating quantitative predictions.

(3) Although a handful of molecular H_2 –assisted hydrogenation steps were “validated post-facto” via select DFT calculations, the manuscript does not provide experimental evidence (e.g., isotopic labeling, in situ spectroscopy) for these unconventional Eley–Rideal-type or inter-species hydrogen transfer mechanisms. Their true relevance under reaction conditions thus remains speculative.

(4) The reported barrier of 0.18 eV for the direct H_2 reaction with CO_2 on Cu is extraordinarily low. The authors should provide full NEB pathways, transition state structures, and energetics to justify this value.

(5) It is unclear why the 152-reaction, DFT-based mechanism predicts greater formic-acid formation (and much lower CO_2

conversion) compared to the 9,389-step, ML-augmented network. A discussion of rate-determining steps and their barriers in both models is essential for readers to understand these contrasting outcomes.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Anand et al. revised their manuscript submitted to Nature Communications on data-driven reaction pathway identification. After reviewing their responses to the three reviewers, I believe that there was a significant inconsistency between the theoretical calculations and experimental validation in this manuscript. Therefore, I considered the paper not suitable for publication in its current form. My major concerns were as follows:

(1) Microkinetic results in Section S6: The authors claimed that the microkinetic results of the 152-reaction network using DFT- and ML-based barriers matched well in Fig. S6. What was the authors' criterion to determine this match? In my opinion, especially in the results with CO correction in Fig. S6b, the trends for HCOOH and CH₃OH changed significantly. Therefore, the credibility of the proposed model results needed to be improved.

(2) Experimental Validation with Cu/SiO₂ Catalyst: I didn't not understand the authors' intentions to employ Cu/SiO₂ catalyst to validate the theoretical predictions. According to the prediction results in Fig. 4c, CO, CH₄, CHOOH, and CH₃OH should be the main products, especially the product rate from automatically enumerated reactions. However, in Fig. 5i, only CH₃OH and CO were observed. Should this discrepancy be attributed to the choice of catalyst, or was it due to a mismatch between the computational and experimental results? The authors stated that "the objective of this study is not to quantitatively match experimental results via fitting parameters but to match them qualitatively and highlight the importance of accounting for all possible elementary reactions rather than just a few tens of reactions." However, there should still be some reliability in the qualitative trends of the main products. If the model catalyst system used here cannot effectively predict this, how can the proposed strategy be generalized to other catalyst systems or fields of catalysis research? As the model catalyst chosen by the authors, the standard or significance of Cu/SiO₂ as a model should be clarified. Currently, the computational predictions and experimental validation were not convincing and might even have an adverse effect. Additionally, the manuscript also considered C₂+ products over the Cu(111) surface in the calculations and predictions, but no C₂+ products were present in the chosen catalyst system. How can the reliability of the proposed strategy be demonstrated? This suggested a major flaw in the choice of the model catalyst.

(3) Hydrogenation via Molecular H₂: One of the new findings in this study was that CO₂ hydrogenation might occur via molecular H₂. If this was the conclusion, how did this reconcile with many recent experimental researches indicating that noble metal-modified methanol synthesis catalysts significantly improve the rate of methanol generation via CO₂ hydrogenation (comparison between InOx and Pt/InOx system), which was closely related to the ability of noble metals to dissociate hydrogen into hydrogen atoms? Additionally, the authors speculated: "Catalysts that can enhance H₂ adsorption via alkali metals or alkali earth metals as promoters by increasing H₂ binding energy may act as better CH₃OH catalysts, and this could be probed in the future." I believed that if alkali earth metals enhance H₂ adsorption and improve methanol synthesis, the authors should confirm this in their current work through experiments, rather than leaving it for future studies.

(Remarks on code availability)

I have reviewed the relevant code and believed that it could be reproduced.

Reviewer #2

(Remarks to the Author)

My concerns have been well addressed in the revised version. I have no further comment.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

The revision addressed my previous concerns and I have no issues with the current version. It can be published as it is.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have carefully addressed the reviewer's comments. After two rounds of revisions, they have also provided explanations regarding the inconsistencies between the experimental and computational results. Given that the other two reviewers have recommended publication, I also agree to the paper's publication. However, I still maintain some of my concerns and hope that the authors' future work can provide a more thorough explanation of the potential inconsistencies.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

I have no further comments on the revised version.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to Reviewers' Comments

Manuscript ID: NCOMMS-25-14446-T

Title: Data-driven massive reaction networks reveal new pathways
underlying catalytic CO₂ hydrogenation

Response to Reviewer 1

Overall Comment: Anand M. Verma et. al combined DFT and machine learning to prediction activation barrier of CO₂ hydrogenation on copper, and a Cu/SiO₂ sample was synthesized to validate products prediction from MKM model. It was seen that the involved reaction network was exhaustive and the database for activation barriers for CO₂ hydrogenation was massive. I thought this work was meaningful for the methodology to explore catalysis mechanism, nevertheless, I have several issues to confirm for this manuscript on this version.

Response: We are grateful to the reviewer for their thorough review of our manuscript. We are glad the reviewer finds our methodology meaningful for exploring catalysis mechanisms. Below, we respond to each of the comments and hope that the revised manuscript is suitable for publication in Nature Communications.

Comment 1: In the Introduction section, authors mentioned: “when Cu combined with metals such as Fe, C-C coupling reactions would happen and generate >C2 products.” Why did they mention Fe species here? Was there any specific reason?

Response: We appreciate this comment from the reviewer. The inclusion of Fe in pure Cu can increase the selectivity towards higher carbon products, as shown in previous work (App. Cat. B 2016, 202, 605–610 and J. Phys. Chem. C 2017, 121, 13164–13174). We acknowledge that exclusively highlighting Fe may appear selective, and therefore, we have rephrased the sentence, which can be seen on the page 4 in the revised version of the manuscript.

Comment 2: I thought one contribution of this study was that authors developed an optimization framework to predict the elementary reactions for CO₂ hydrogenation reaction. Was this prediction process influenced by employed specific catalyst actually? I was wondering the efficiency of this approach to unveil the reaction rationale and then to guide the catalyst optimization compared with current method (experimental results + theoretical calculation for mechanism exploration). Since methanol presented higher value than

CO as CO₂ hydrogenation products, the dominating CO selectivity over Cu/SiO₂ system in this study really raised my interest to ask “why didn’t authors employ well known Cu/ZnAl or other well estimated oxide system for methanol synthesis via thermocatalysis to predict the elementary steps?”

Response: We thank the reviewer for their comment. We developed a framework that is capable of enumerating and identifying all possible elementary reactions, followed by microkinetic modeling to predict the reaction outcome based on DFT-guided and ML-predicted barriers. For our case study, we utilized CO₂ hydrogenation on Cu(111), a well-established reaction for producing multiple carbon-based products, to predict its reaction outcome and validate the prediction with experiments. We note here that though we examined CO₂ hydrogenation on Cu(111), a prototypical catalyst, the proposed workflow is equally applicable to any catalytic reaction on any type of catalyst. Thus, in future work, this approach can be applied to industrial catalysts, such as Cu/Zn-Al. We have clarified this point on page 18 of the revised manuscript.

Comment 3: In this study, authors proposed many detailed reaction pathways and intermediates for CO₂ hydrogenation, such as hydrogenation of CO₂ to formate via surface hopping of the hydrogen atom of an adsorbed hydrogen molecule, and more detailed routes for CH₃OH and CO were proposed in section of “Revealing the precise mechanistic pathways of the major products” . For energy barriers, it was not easy to examined from experimental results, however, I though the experiments to record those aforementioned intermediates and their variation trend were necessary to be conducted to validate authors’ prediction results.

Response: We thank the reviewer for this comment. The identification of intermediate compounds through experiments is indeed highly challenging, as many of these intermediates are short-lived species and others require advanced in situ characterization facilities such as operando IR spectroscopy. Nevertheless, to address the comment raised by the reviewer, we conducted high-pressure in-situ DRIFTS experiments on Cu/SiO₂ catalyst at 20 bar and varying temperature (483-523 K). The pressure of 20 bar is the closest possible experimental condition that we can achieve. We observed several surface species, including carbonate, formate, and surface-attached formic acid (please refer to Figure S13 and Table S14 of the revised Supporting Information file). This new set of experimental information, obtained from DRIFTS, supports the DFT-based predictions of the presence of various intermediates, such as formate and formic acid, involved in the production of CO and methanol. In addition, the presence of formate on the catalyst surface also supports the DFT-based prediction of CO₂ hydrogenation to formate instead of carboxylic species due to lower activation energy. Please note that the presence of formic acid is detected in our microkinetic modeling, but it was absent from the experimental list of CO₂ hydrogenated gas-phase products. Using DFT, we affirmed the adsorption free energy of formic acid to be negative (-0.21 eV) on Cu(111), which indicates that the formic acid should be bound to Cu(111) catalyst as per the DRIFTS experiment (Figure S13 and Table S14 of the revised Supporting Information file).

Comment 4: For the experimental results, authors mentioned “for the automated MKM model based on the extensive set of reactions, the production rates and product distribution closely followed experimental trends (Figure 5)”. The products distribution was seen in Fig. 5i, but where were the products rates, which was also imperative to validate the number in Fig. 4c. In addition, HCOOH presented comparable rates with methanol in Fig. 4c, but no HCOOH product was detected in Fig. 5i. Was this induced by strong adsorption of HCOOH on sample or the detection limitation of HCOOH by GC?

Response: We appreciate the reviewer’s comment. In response, as per the MKM-based product distribution (see Figure 4c), major products, such as methanol and CO, are in qualitative agreement with the experimentally obtained product distribution (see Figure 5i). Though this workflow successfully predicts the experimental trends, as demonstrated in the original manuscript, the difference in the production rates between computations and experiment is in the order of 10^2 - 10^3 (see Section S7.4 and SI Figure S11), which corresponds to an error of 0.2-0.3 eV in activation barriers. Predicting activation barriers with errors lower than these numbers is highly challenging without the use of fitting parameters, as employed in previous work. Furthermore, experiments involve various complexities, including mass transport limitations, the availability of multiple active sites/facets, adsorbate-adsorbate interactions, and the influence of catalyst support, among others. Regarding the point on formic acid, it can be very well detected by our GC analyser; however, its presence in the gas-phase product mixture was not detected. To understand this finding, we also conducted experiments with stoichiometric conditions of $H_2:CO_2=1:1$; however, we still couldn’t detect any formic acid in the gas-phase. We believe that the absence of formic acid might be due to its strong adsorption on the catalyst surface. On Cu(111), the adsorption free energy of HCOOH was computed to be -0.21 eV, which indicates unfavorable desorption of formic acid. In other words, formic acid might still be on the catalyst surface. To test this hypothesis, we conducted a high-pressure in situ DRIFTS experiment, which indeed demonstrates the presence of formate and surface-bound formic acid species (please refer to Figure S13 and Table S14 of the revised Supporting Information file). Thus, it reaffirms our hypothesis that formic acid may act as an intermediate rather than desorbing to the gas phase. This also confirms our reaction mechanism predictions that report formic acid as one of the key intermediates in the formation of methanol from CO_2 hydrogenation.

Response to Reviewer 2

Overall Comment: In this work, DFT calculations are combined with machine learning for kinetic modeling of the CO_2 hydrogenation process. The data-driven approach is extensive and valuable; however, the experimental validation remains limited, and the novel insights derived from the modeling are not sufficiently emphasized. My specific comments are as follows

Response: We wholeheartedly thank the reviewer for their extensive review of our work and for providing an opportunity for us to improve our manuscript. We included new

calculations and experimental data to address the concerns raised by the reviewer in the following points.

Comment 1: The use of an in-house DFT dataset (comprising 304 reaction barriers) for training the ML model effectively mitigates issues related to data heterogeneity, thereby enhancing model accuracy and reliability. However, this approach may also constrain the model’s generalizability to broader catalytic reaction systems. The authors are suggested to discuss the potential impact of using a relatively homogeneous dataset on the model’s generalization capability.

Response: The reviewer raises an important concern here and we thank them to provide a chance to address this issue. We agree that a specific dataset constrains the model’s generalizability to broader catalytic reaction systems. However, generalizability is challenging even for the present state-of-the-art ML potentials such as MACE-OMAT and MACE-MPA. We took this opportunity to highlight the trade-off between generalizability and accuracy by computing and comparing the activation barriers of manually curated DFT-based reactions with those computed using MACE-OMAT and MACE-MPA models, as discussed in SI section 2.1 and Table S5. According to the results presented in Table S5, with errors exceeding 1 eV for MACE-MPA, compared to only 0.13 eV for our model, we can conclude that these models demonstrate significantly low accuracy, indicating a considerable trade-off between generalizability and predictive performance. Further details about the MACE-OMAT and MACE-MPA calculations are available in SI Section 2.1. A brief description of the comparison is provided on page 9 of the revised manuscript.

Comment 2: The AFE-AE model does not include a feature importance analysis. Given that activation free energy is predicted from activation energy, a high-performing model might achieve comparable accuracy using only activation energy and temperature as input features. Additionally, the relatively low R^2 score of the AE-RE model suggests limited generalization capability, which warrants further discussion.

Response: We appreciate this comment from the reviewer. In response, we have added the feature importance analysis of the AFE-AE model in Figure S5 for both conditions (one with all features and another with only E_a and T as features) as well. Next, we tested the AFE-AE model with only E_a and T as input features and observed that adding only E_a and T as input features achieves comparable accuracy. It shows only a slight decrease in the performance between both models, i.e., a difference of 0.013 eV in the train MAE and 0.006 eV in the test MAE with comparable R^2 . Therefore, we retained the additional features, along with E_a and T , as these features depict the physics of the reactions, while also capturing subtle trends and increasing model robustness at no additional computational cost. Furthermore, as per the feature importance analysis (see Figure S5 in SI), other features such as molecular weight, C_n -H bond, and others show their considerable dependence on the target variable. Please refer to Section 2.2 of the revised SI for further information. Related texts are added on page 11 of the revised manuscript. Regarding the low R^2 score of the AE-RE model, we would like to clarify

that we primarily focused on mean absolute error (MAE) as it provides a more direct and interpretable measure of predictive accuracy in energy units. Though R^2 is another metric to predict the fitting between actual and predicted values, it is often quite sensitive to the variance of target values and, therefore, can be misleading in datasets with narrow ranges. For the AE-RE model, though the R^2 for the test set is reported to be 0.821, the corresponding MAE (0.13 eV) after hyperparameter tuning is well within the permissible range/accuracy. Furthermore, we have already demonstrated above that even machine learning potentials face significant difficulties with generalizability.

Comment 3: The introduction section primarily focuses on describing the present work, rather than providing a concise overview of relevant prior studies. It would be beneficial to include a brief summary of previously reported works on similar topics to establish the research context. Additionally, content specific to this study should be shifted to later sections rather than being included in the introduction.

Response: In response to the valuable comment raised by the reviewer, we restructured our Introduction section and added several previous developments that can be seen on page 3 of the revised main text.

Comment 4: The model assumes that the catalyst surface remains static, which may not accurately reflect real-world catalytic conditions. To enhance the model’s realism, it would be valuable to incorporate surface reconstruction or dynamic changes in catalyst structure.

Response: This is another interesting concern raised by the reviewer, and we also feel that this issue must be discussed in the manuscript. It is true that in a realistic environment, there is a possibility of surface reconstruction of the catalyst. To verify this, we conducted an ab initio molecular dynamics (AIMD) simulation of a hydrogen-covered Cu(111) catalyst at a temperature of 1000 K, as our MKM predicts the surface to be largely covered with hydrogen, reaching a coverage of 0.3125 ML. We observe no significant surface reconstruction even after a 10-ps timeframe (see SI Figure S17), which affirms that the catalyst remains static in terms of its configuration even at the experimental conditions. It is briefly mentioned on page 19 of the revised manuscript, while more details can be found in SI Section 11.

Comment 5: Although CO_2 hydrogenation to methanol has been extensively studied, the dataset used for ML model training is relatively small (302 reactions). This limitation may hinder the model’s generalization capability. Expanding the training dataset could significantly improve predictive accuracy.

Response: We appreciate this comment from the reviewer. The reviewer is correct that CO_2 hydrogenation to methanol has been extensively studied, but unlike previous studies, we have incorporated almost all possible mechanisms of CO_2 hydrogenation, comprehensively covering all major reaction classes including (de)hydrogenation, (de)oxygenation, (de)hydration, (de)hydroxylation, C-C coupling and isomerization reactions, leading to

C1 to C4-based hydrocarbons/alcohols, including methanol. Based on this extensive investigation, we manually curated 302 reactions (forward and reverse) using DFT, which is possibly the largest dataset generated yet by any research group. Therefore, the dataset is sufficient to train the ML model for activation barrier predictions without bringing the infamous issue of data heterogeneity. The robustness of our AE-RE model is also evident from the metrics (MAE and R^2) employed in this study, which yield a test MAE of 0.13 eV, compared to as high as 1.2 eV for state-of-the-art machine learning interatomic potentials (MLIPs). Increasing the dataset size could be explored in future work, and we have indicated this as a direction for future work on page 18 to aid in the fine-tuning of MLIPs.

Comment 6: While automated reaction enumeration eliminates human bias, it also generates a substantial number of impractical or redundant reactions. The authors are suggested to refine the reaction network by drawing upon experimentally validated reaction pathways for CO₂ hydrogenation to methanol. This approach would enhance computational efficiency and improve the model’s practical relevance.

Response: We appreciate the reviewer’s concern regarding the validity of the reactions generated by the automated framework. We would like to clarify that our automated reaction enumerator (ARE) code initially reported 104,723 reactions, which were subsequently refined by our elementary reaction identifier (ERI) framework, resulting in a reduction to 9,237 reactions. The set of genuine reactions reported by ERI was validated by randomly choosing 388 reactions and verifying them. However, to further ensure the accuracy and robustness of our ERI framework, we expanded our validation set to 523 reactions, comprising 276 elementary and 247 non-elementary reactions. Note that these random reactions included all classes of reactions, e.g., hydrogenation/dehydrogenation (145 reactions), oxygenation/deoxygenation (69 reactions), C–C coupling/decoupling (21 reactions), and isomerization (12 reactions). We would like to report here that our ERI framework successfully classified each random reaction with 100 % accuracy. We believe that such a diverse and comprehensive dataset supports the claim of robustness of the ERI framework, which accurately identifies elementary reactions. The related change can be seen on page 11 of the revised manuscript. The updated dataset has been uploaded to our GitHub repository. Moreover, the 105 intermediates chosen in this study were included in our DFT dataset and were thermodynamically stable.

Comment 7: The experimental validation is conducted using Cu/SiO₂ catalysts, but the range of catalysts examined remains limited. Expanding the selection to include other supported metal catalysts (e.g., Cu/ZnO, Cu/ZrO₂) or metal oxide catalysts (e.g., ZnO-ZrO₂, In₂O₃-ZrO₂) would strengthen the reliability and broader applicability of the findings.

Response: We appreciate the reviewer’s suggestion regarding the inclusion of a broader range of catalysts for the experimental validation. We would like to clarify that the experimental validation in this manuscript is primarily intended to demonstrate the applicability

of the proposed computational workflow that systematically and efficiently accounts for all possible elementary reactions that may occur simultaneously or in tandem on a particular catalyst surface. Exploring all possible elementary reactions associated with a vast reaction network would otherwise be infeasible using DFT due to its significant computational expense and complexity in locating the transition state. Hence, the focus of this manuscript is primarily on developing and validating a robust computational framework, rather than conducting an extensive experimental survey that includes several types of catalysts. In addition, please note that the present experiment on Cu/SiO₂ is considered only to mimic the computational model of (111) facet of Cu catalyst, which is abundant in copper nanoparticles. Therefore, the experimental validation using Cu/SiO₂ serves as a suitable proof of concept. We agree that extending experiments to other supported metal catalysts (e.g., Cu/ZnO, Cu/ZrO₂) and mixed metal oxides (e.g., ZnO–ZrO₂, In₂O₃–ZrO₂) could be considered as an important direction for future work, as indicated on page 18 of the revised manuscript.

Comment 8: The scope of experimental validation is rather limited. Certain conclusions, such as the increase/decrease in methanol/CO selectivity with rising temperature, are well-established and do not necessarily require validation through modeling. The study would benefit from highlighting novel predictions derived from the model. Moreover, a more quantitative comparison between experimental and model results would be desirable. Additionally, the experimental conditions explored—particularly the temperature range—are somewhat narrow. Expanding the experimental parameters (e.g., pressure, space velocity) would enhance the robustness of model validation.

Response: We thank the reviewer for their comment. In response, we have expanded the experimental conditions to focus on different pressures, temperatures, and space velocities (see SI section 8.4). Our framework successfully matches the qualitative trends shown by the experimental results, validating its performance (SI section 7.4). However, a quantitative comparison requires consideration of various complexities, as explained in Reviewer 1’s comment 4, and can be addressed in future work. In this regard, state-of-the-art MLIP frameworks incur errors on the order of 1 eV in activation barriers, whereas an error of just 0.1 eV in the barrier corresponds to a change in rates by a factor of 10. Thus, our model, with a mean absolute error of 0.13 eV for barriers, is currently one of the only tractable approaches to resolving the complexity of a massive reaction network.

Comment 9: The modeling effort is extensive, but readers would expect to gain clear guidance or new insights from the results to facilitate rational catalytic system design for targeted product formation. The discussion should emphasize actionable insights derived from the model.

Response: As per the valuable suggestions by the reviewer, we have discussed the insights based on the workflow and its outcomes on pages 14, 17, and 18 of the revised manuscript. Furthermore, we also discussed possible actionable insights for the readers, derived from the current model.

Comment 10: The title states that the study “reveals new pathways underlying catalytic CO₂ hydrogenation.” However, the discussion on newly identified reaction pathways appears limited. The authors should clarify and substantiate this claim.

Response: We acknowledge the reviewer’s point and agree that there should be additional discussions on newly found pathways. Our workflow reveals several novel reactions that have a significant impact on overall performance, which have not been previously reported elsewhere. For instance, our ML-based MKM predicted that some hydrogenation reactions should occur in a concerted fashion rather than a sequential manner (see Figure 4d). In the methanol production pathway based on an automated workflow (Figure 4d), CO₂*, HCOOH*, and CH₃O* hydrogenation reactions are facilitated by adsorbed molecular hydrogen, whereas the remaining reactions hydrogenate via atomic H. We further discuss this on page 14 of the revised manuscript. Additionally, we also highlight the importance of cross-species interaction/hydrogenation reactions such as those listed in Table S4. Our ERI framework based on manually curated reactions reports many such reactions with lower barriers, e.g., CH₃CHCH₂* + H₂COO* = CH₃CH₂CH₃* + HCOO* (ΔG_{act} = 0.0802), CH₂* + OCCHO* = CH₂CHOCO* + * (ΔG_{act} = 0.051), and others.

Comment #11: The manuscript contains subjective descriptions and potentially disputable statements that should be reconsidered. For example:

- In the introduction: “it would take $\approx$ 47 years to compute all activation barriers with 240 CPU cores.”
- On page 7: “Overall, the proposed reaction network is the most comprehensive one for C1 to C4 pathways so far...”

Response: We thank the reviewer for pointing this out. We have now carefully revised the manuscript to avoid potentially overstated or subjective claims. For instance:

- The statement regarding “it would take $\approx$ 47 years to compute all activation barriers with 240 CPU cores” has been softened to emphasize the computational challenges without exaggeration. The revised sentence, which can be seen on page 2 of the revised manuscript, reads: “Considering a network with 10,000 reactions, it would require decades of runtime on a single workstation, highlighting the magnitude of the associated computational cost.”
- We have revised it to “Overall, the proposed reaction network forms one of the most comprehensive datasets...”, which can be seen on page 7 of the revised manuscript. Nevertheless, we stand by our statement, which reads “Overall, the proposed reaction network is the most comprehensive one for C1 to C4 pathways so far...”. It is majorly because we have searched the literature substantially, and we were unable to find a single source that considered these many reactions obtained via DFT for CO₂ hydrogenation.

Response to Reviewer 3

Overall Comment: This manuscript introduces a multifaceted approach to elucidating complex catalytic mechanisms. The authors combine several advanced techniques in an end-to-end workflow: extensive density functional theory (DFT) calculations, an automated reaction enumeration algorithm (inspired by human chemical intuition), machine-learning (ML) models to predict activation energy barriers, and comprehensive microkinetic modeling. Previous studies of CO₂ hydrogenation mechanisms were limited by computational constraints, which prevented them from examining a large number of reactions or forced reliance on simplified scaling relations for kinetics. By contrast, this work’s data-driven pipeline enables the automated construction and evaluation of a complex reaction network (on the order of 10^3 - 10^4 reactions), which marks a clear departure from conventional, small-scale microkinetic models. Based the complex reaction network, they found that direct H₂ (gas-phase molecule) can participate in surface hydrogenation steps (an Eley–Rideal-type mechanism) is novel for CO₂ hydrogenation on Cu, as such steps are rarely considered in traditional mechanisms. The following issues need to be addressed for this work.

Response: We are thankful to the reviewer for their efforts in comprehensively reviewing our work and appreciate for the same. We do understand the concerns raised by the reviewer, and below we respond to each of their comments pointwise.

Comment 1: The “elementary reaction identifier” was validated on only 388 manually labeled reactions, yet is then applied to extract 9,237 elementary steps from an initial pool of 104,723 reactions. Without further benchmarking on a larger or more chemically diverse test set, there is a risk that many nonelementary (concerted or multistep) reactions slipped through, or genuine elementary steps were discarded.

Response: The reviewer raised an important concern, and, in fact, a similar concern was also raised by Reviewer 2. We initially worked with 388 reactions randomly selected from the predicted list of reactions by our ERI (Elementary Reaction Identifier) framework. The model was shown to be working perfectly, as our ERI framework had accurately predicted each reaction. However, we understand the concern here, which is quite important in making the model robust and accurate; therefore, we have increased the testing set from 388 to 523 reactions, comprising each class of reaction. We are pleased to report that our ERI framework has again demonstrated 100 % accuracy as each reaction was correctly classified as being elementary or non-elementary, which suggests that the ERI model is robust enough to apply to a diverse set of reactions with perfect accuracy. Further information can be found in the response to comment 6 raised by reviewer 2.

Comment 2: The microkinetic modeling treats all elementary steps on Cu(111) as happening on isolated, identical sites under mean field conditions, ignoring lateral adsorbate–adsorbate interactions and coverage dependent energetics. At the high coverages implicit in the 50 bar, such interactions can significantly alter barrier heights and site availability, potentially invalidating quantitative predictions

Response: We thank the reviewer for this comment. According to our MKM results, the catalyst surface remains largely occupied with hydrogen atoms (see Figure S9). Therefore, to assess the influence of lateral adsorbate–adsorbate interactions, we explicitly evaluated hydrogen adsorption energetics on Cu(111) at various surface coverages (1/16–5/16 ML) using DFT. We found that while adsorption energies varied at lower coverages, the differences became negligible at higher H coverages, such as 5/16 ML (see Figures S18 and S19), which closely corresponds to the MKM-predicted H coverage. Therefore, the mean-field approximation used in our MKM framework remains valid and does not compromise model accuracy. A detailed discussion is provided in Section S12.

Comment 3: Although a handful of molecular H₂-assisted hydrogenation steps were “validated postfacto” via select DFT calculations, the manuscript does not provide experimental evidence (e.g., isotopic labeling, in situ spectroscopy) for these unconventional Eley–Rideal-type or interspecies hydrogen transfer mechanisms. Their true relevance under reaction conditions thus remains speculative.

Response: We appreciate the reviewer’s comment. We agree with the reviewer’s suggestions; however, we would like to state that our model is not only robust but also validated on multiple fronts. The ML-based predictions are trained and tested with promising accuracy and are significantly more accurate than state-of-the-art machine-learned interatomic potentials. MKM-based predictions follow similar trends to our experimental outcomes. Nevertheless, as per the reviewer’s suggestions, we also performed advanced characterization, namely in situ DRIFTS experiments at 20 bar and various temperatures (483–523 K). We indeed observed surface-bound carbonate and formate species, as also reported in our DFT-based calculations. These newer experiments support the choice of intermediates used in the study and validate the prediction of the reaction pathway for methanol formation, which involves HCOOH, also observed as an intermediate in DRIFTS. More details can be found in Figure S13 and Table S14 of the revised SI. Although we agree with the reviewer’s suggestions regarding the validation of Eley–Rideal type or interspecies hydrogen transfer mechanisms using isotopic labeling, we regret to inform you that, at this stage, we are unable to conduct such experiments due to resource limitations. However, we remain confident about our model and welcome experimentalists to further investigate.

Comment 4: The reported barrier of 0.18 eV for the direct H₂ reaction with CO₂ on Cu is extraordinarily low. The authors should provide full NEB pathways, transition state structures, and energetics to justify this value.

Response: We thank the reviewer for this comment. In response, we performed nudged elastic band (NEB) calculations for the reactions listed in Table 1 of the revised manuscript. For the reaction CO₂* to HCOO*, the low barrier arises due to the combination of molecular H₂ as the reactant with atomic H as the transition state. Accordingly, we have revised Table 1 in the main text to remove this particular activation barrier. However, for the other cases, we now provide full NEB pathways in Figure S4 of the revised SI

along with transition state structures, which support molecular hydrogenation of the intermediate species, particularly those involving O-site hydrogenation. This could be due to the higher electronegativity of oxygen compared to carbon, which would allow one to break the H-H bond more easily. We have discussed this point on page 14 of the revised manuscript.

Comment 5: It is unclear why the 152 reaction, DFT-based mechanism predicts greater formic acid formation (and much lower CO₂ conversion) compared to the 9389-step, ML-augmented network. A discussion of the rating-determining steps and their barriers in both models is essential for readers to understand these contrasting outcomes.

Response: We thank the reviewer for this important comment. The manual reaction network for the methanol is based on the atomic hydrogenation (see Figure S10) for which the rate-determining step is identified as $\text{CH}_3\text{O}^* \rightarrow \text{CH}_3\text{OH}^*$ ($\Delta G_{Act} = 1.06$ eV). However, further hydrogenation of HCOOH is hindered due to a higher kinetic barrier, which explains why the manual network yields more formic acid than methanol. On the other hand, in the automated network, which is based on the combination of atomic and molecular hydrogenation (see Figures 4d and S10), the rate-determining step is identified as $\text{HCOO}^* \rightarrow \text{HCOOH}^*$ ($\Delta G_{Act} = 0.85$ eV), which is 0.2 eV lower than the RDS in manual reaction. Therefore, the automated network proceeds more easily to produce methanol compared to the manual network. This is now discussed on page 15 of the revised manuscript and in Section S7.3.

Response to Reviewers' Comments II

Manuscript ID: NCOMMS-25-14446A

Title: Data-driven massive reaction networks reveal new pathways
underlying catalytic CO₂ hydrogenation

Response to Reviewer 1

Overall Comment: Anand et al. revised their manuscript submitted to Nature Communications on data-driven reaction pathway identification. After reviewing their responses to the three reviewers, I believe that there was a significant inconsistency between the theoretical calculations and experimental validation in this manuscript. Therefore, I considered the paper not suitable for publication in its current form. My major concerns were as follows:

Response: We are thankful to the reviewer for evaluating our revised manuscript and for providing further feedback on our paper. We acknowledge the importance of consistency between theoretical and experimental calculations. In response to this concern, we have carefully re-examined our manuscript and revised it accordingly. We have also added additional simulation and experimental analyses, which strengthen the paper considerably. We hope that these revisions resolve the perceived inconsistency and demonstrate that the computational and experimental results are complementary rather than contradictory. Below, we respond to each comment point by point and hope that the revised manuscript is suitable for publication in Nature Communications.

Comment 1: Microkinetic results in Section S6: The authors claimed that the microkinetic results of the 152-reaction network using DFT- and ML-based barriers matched well in Fig. S6. What was the authors' criterion to determine this match? In my opinion, especially in the results with CO correction in Fig. S6b, the trends for HCOOH and CH₃OH changed significantly. Therefore, the credibility of the proposed model results needed to be improved.

Response: We thank the reviewer for this comment. We re-emphasize that both DFT- and ML-based models show qualitatively consistent behavior and capture the same dominant trends in overall activity, major product identity, and temperature dependence.

Specifically, our criterion for consistency is based on (i) preservation of the relative ranking of the major products over the investigated temperature range (CO>HCOOH>CH₃OH>CH₂O),

and (ii) comparable production rates within the same order of magnitude for key species (Table 1). We acknowledge that, particularly in Fig. S6b with gas-phase CO energy correction, the temperature-dependent trends for CH₃OH differ more noticeably between the DFT and ML-based models. In this regard, we note that in the full 9k+ microkinetic reaction network, which also includes the manually enumerated 152 reactions, the exact DFT-calculated activation barriers are used for 152 reactions rather than ML-predicted values. As a result, any residual discrepancies arising from ML-based barriers in the reduced comparison model do not propagate into the final large-scale microkinetic simulations, thereby ensuring the robustness and credibility of the reported model predictions. To address this point, we have revised Section S6 to explicitly specify our criteria for a good match, acknowledge these deviations, and avoid overstatement of quantitative agreement.

Table 1: Prediction error metrics for production rates using ML-predicted values compared with DFT values.

Species	MAE (mol m ⁻³ s ⁻¹)	RMSE (mol m ⁻³ s ⁻¹)
HCOOH	8.99×10^{-7}	9.28×10^{-7}
CO	2.39×10^{-4}	7.27×10^{-4}
CH ₃ OH	1.08×10^{-6}	1.29×10^{-6}

Comment 2: Experimental Validation with Cu/SiO₂ Catalyst: I didn’t not understand the authors’ intentions to employ Cu/SiO₂ catalyst to validate the theoretical predictions. According to the prediction results in Fig. 4c, CO, CH₄, CHOOH, and CH₃OH should be the main products, especially the product rate from automatically enumerated reactions. However, in Fig. 5i, only CH₃OH and CO were observed. Should this discrepancy be attributed to the choice of catalyst, or was it due to a mismatch between the computational and experimental results? The authors stated that “the objective of this study is not to quantitatively match experimental results via fitting parameters but to match them qualitatively and highlight the importance of accounting for all possible elementary reactions rather than just a few tens of reactions.” However, there should still be some reliability in the qualitative trends of the main products. If the model catalyst system used here cannot effectively predict this, how can the proposed strategy be generalized to other catalyst systems or fields of catalysis research? As the model catalyst chosen by the authors, the standard or significance of Cu/SiO₂ as a model should be clarified. Currently, the computational predictions and experimental validation were not convincing and might even have an adverse effect. Additionally, the manuscript also considered C₂+ products over the Cu(111) surface in the calculations and predictions, but no C₂+ products were present in the chosen catalyst system. How can the reliability of the proposed strategy be demonstrated? This suggested a major flaw in the choice of the model catalyst.

Response: We thank the reviewer for the detailed and thoughtful critique regarding the choice of Cu/SiO₂ as the experimental validation platform and the apparent discrepancies between the computational predictions and experimental observations. In response, we would like to provide a point-by-point rebuttal.

On the purpose of Cu/SiO₂ as a model catalyst: The primary objective of employing Cu/SiO₂ was not to achieve a one-to-one quantitative validation of all predicted products, but rather to test whether the dominant qualitative trends emerging from a large, automatically enumerated reaction network are reflected in a realistic experimental system. Cu-based catalysts are widely regarded as prototypical systems for CO₂ hydrogenation, and Cu/SiO₂ in particular is a well-established model catalyst. For this reason, we found that Cu/SiO₂ is suitable for probing intrinsic Cu-driven reaction pathways rather than for maximizing product diversity. Future work can examine other catalyst systems. In fact, in the revised manuscript, we also include experimental results using another model system, Cu/Al₂O₃, which also indicates the role of the metal-support interaction and provides avenues for future work. To address this point, we have revised Section S8.1 in the Supplementary Information.

On the absence of CH₄ and its predicted rate: Although CH₄ appears as a possible product in the microkinetic predictions, its predicted turnover frequency is extremely low. For example, at 523 K, the predicted TOF for CH₄ formation is on the order of 10⁻⁸ s⁻¹, which is well below the experimental detection limit. Therefore, the absence of detectable methane in Fig. 5i is fully consistent with the theoretical predictions and does not represent a mismatch between computation and experiment. To address this point, we have revised Section S7.4 in the Supplementary Information.

On formic acid (HCOOH): As discussed in detail in our previous response to Reviewer 1’s Comment 4 (in the first round of revisions), the absence of gas-phase HCOOH does not contradict the model predictions. Microkinetic modeling identifies HCOOH as a key surface intermediate rather than a desorbing product. DFT calculations show that HCOOH binds relatively strongly to Cu(111) ($\Delta G_{ads} = -0.21$ eV), making desorption unfavorable. This hypothesis is supported by our high-pressure in situ DRIFTS measurements, which clearly detect surface-bound formate and formic acid species (Figures S13 and Table S15), thereby experimentally validating the predicted reaction mechanism. To address this point, we have revised Section S7.4 in the Supplementary Information.

On C₂⁺ products: Please note that it is essential to include C₂+ species in the DFT dataset in order to have a sufficient number of data points (152) to train our ML model. In the absence of C₂+ species, we would only have a low number of reactions (50) in the network, which would prevent the training of a fairly generalizable ML model. Now, although the computational framework automatically enumerates and includes C₂⁺ reaction pathways, it is remarkable that the microkinetic simulations still predict C₁ products as the dominant species across the studied conditions. The negligible formation rates of C₂⁺ products even after expanding the reaction network demonstrate that the model does not overpredict higher-carbon products as an artifact of reaction enumeration. This prediction is in direct qualitative agreement with the experimental results obtained on Cu/SiO₂, where no C₂⁺ products were observed. Such consistency confirms that the proposed strategy captures intrinsic kinetic selectivity rather than producing nonphysical or spurious reaction outcomes. To address this point, we have revised Section S7.4 in the

Supplementary Information.

Comment 3: Hydrogenation via Molecular H₂: One of the new findings in this study was that CO₂ hydrogenation might occur via molecular H₂. If this was the conclusion, how did this reconcile with many recent experimental researches indicating that noble metal-modified methanol synthesis catalysts significantly improve the rate of methanol generation via CO₂ hydrogenation (comparison between InO_x and Pt/InO_x system), which was closely related to the ability of noble metals to dissociate hydrogen into hydrogen atoms? Additionally, the authors speculated: "Catalysts that can enhance H₂ adsorption via alkali metals or alkali earth metals as promoters by increasing H₂ binding energy may act as better CH₃OH catalysts, and this could be probed in the future." I believed that if alkali earth metals enhance H₂ adsorption and improve methanol synthesis, the authors should confirm this in their current work through experiments, rather than leaving it for future studies.

Response: We thank the reviewer for raising this important point regarding hydrogen activation pathways and their relationship to prior experimental studies on catalysts with promoter atoms. In response, we have now conducted both modeling and experiments to investigate the effect of alkali promoter atoms on Cu catalyst activity. The case of alkaline earth metals can be considered in future work, and thus, we have removed any claims regarding those from the paper.

Our new DFT calculations show that hydrogen adsorption is more favorable in the presence of K promoter atoms, and is less favorable in the presence of Na promoter atoms. We have also carried out experiments on Na/K-promoted Cu/Al₂O₃ catalysts, indicating the K promoter atoms increase the selectivity of methanol formation (see Sections S13.3 *Experimental insights into the effect of alkali metals as promoters on copper* and S13.4 *Adsorption and dissociation free energies for molecular H₂ on Na/K/pure-Cu(111) using DFT* in Supplementary Information). This suggests a link between more favorable H₂ adsorption and a greater tendency toward methanol formation. In addition, our DFT simulations show that dissociation of H₂ has the following order of favorability: K-promoted Cu > Na-promoted Cu > pure Cu, which could also be contributing to the reaction mechanism. Importantly, these findings do not contradict experimental observations that noble metal-modified catalysts (e.g., Pt/InO_x) improve methanol synthesis by promoting H₂ dissociation. Rather, they indicate that two mechanistic regimes (hydrogenation by atomic vs. molecular hydrogen) may coexist. Overall, we have not only provided experimental validation of the alkali-promoted enhancement of methanol formation but also a theoretical basis for these findings. Accordingly, we have revised Section S13 in the Supplementary Information.

Response to Reviewer 2

Overall Comment: My concerns have been well addressed in the revised version. I have no further comment.

Response: We are thankful to the reviewer for evaluating our revised manuscript and for recommending its publication in Nature Communications.

Response to Reviewer 3

Overall Comment: The revision addressed my previous concerns and I have no issues with the current version. It can be published as it is.

Response: We are thankful to the reviewer for evaluating our revised manuscript and for recommending its publication in Nature Communications as it is.