nature communications

Article

https://doi.org/10.1038/s41467-022-32514-7

Deep reaction network exploration at a heterogeneous catalytic interface

Received: 13 June 2022

Accepted: 3 August 2022

Published online: 18 August 2022

Check for updates

Qiyuan Zhao [1,2], Yinan Xu [1,2], Jeffrey Greeley [1] ✉ & Brett M. Savoie [1] ✉

Characterizing the reaction energies and barriers of reaction networks is central to catalyst development. However, heterogeneous catalytic surfaces pose several unique challenges to automatic reaction network characterization, including large sizes and open-ended reactant sets, that make ad hoc network construction the current state-of-the-art. Here, we show how automated network exploration algorithms can be adapted to the constraints of heterogeneous systems using ethylene oligomerization on silica-supported single-site Ga^{3+} as a model system. Using only graph-based rules for exploring the network and elementary constraints based on activation energy and size for identifying network terminations, a comprehensive reaction network is generated and validated against standard methods. The algorithm (re)discovers the Ga-alkyl-centered Cossee-Arlman mechanism that is hypothesized to drive major product formation while also predicting several new pathways for producing alkanes and coke precursors. These results demonstrate that automated reaction exploration algorithms are rapidly maturing towards general purpose capability for exploratory catalytic applications.

Establishing the kinetic details of complex reaction networks is central to understanding heterogeneous catalytic surfaces[1–3]. The development of such networks for new systems is often painstaking, even when good hypotheses exist for the governing reactions and cycles[4,5]. Yet the domain knowledge necessary for hypothesis-driven exploration is often outpaced by the high throughput of synthetic and experimental efforts currently driving exploratory catalyst development. For these reasons, establishing reaction networks is typically a retrospective activity performed on already promising catalysts, rather than a predictive component of catalyst exploration workflows. To address this gap, it is urgent to develop computational methods to accelerate and automate the exploration, characterization, and refinement of complex reaction networks at surfaces[6].

In the context of heterogeneous catalysis, computational methods are relatively mature for characterizing the transition states (TSs) of targeted reactions[7–13], performing microkinetic modeling on established reaction networks, and using descriptor-based methods for optimizing catalysts[14–17]. However, a central challenge in characterizing new catalytic interfaces lies in establishing the kinetically relevant reaction network, which is often based on intuition and can be time-consuming and error prone to characterize ad hoc[4–6,18]. Even seemingly simple heterogeneous reactions, such as methane activation on metal oxide surfaces, can be decomposed into numerous elementary steps[19,20]. Catalytic cycles can also involve many intermediates, or even open-ended reactant lists, such that brute force enumeration and characterization are infeasible. Canonical examples of this include hydrocarbon forming reactions such as the oxidative coupling of methane and olefin oligomerization, each potentially involving the formation/dissociation of long carbon backbones as intermediates and an open set of olefins as adsorbed reactants[21,22].

For these reasons, the recent advent of automated reaction prediction approaches is potentially promising for elucidating reaction networks involving heterogeneous interfaces[23–25]. These methods can be categorized on the basis of whether the potential energy surface (PES) is explored in detail to locate TSs or whether the reaction networks are enumerated using a closed set of reaction templates. The latter class includes packages such as Network Generation (NetGen)[26] and Reaction Mechanism Generator (RMG)[27]. This strategy is less

[1]Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN 47906, USA. [2]These authors contributed equally: Qiyuan Zhao, Yinan Xu. ✉e-mail: jgreeley@purdue.edu; bsavoie@purdue.edu

relevant to characterizing exploratory catalysts where established reaction templates are typically absent. Methods that directly explore the PES circumvent this limitation, at least in principle. This class includes several approaches that are under active development, including the artificial force induced reaction (AFIR) method[28,29], stochastic surface walking reaction sampling (SSW)[30], the ZStucture method[31,32], and Yet Another Reaction Program (YARP)[33], our recently developed methodology. All of these approaches are intrinsically more expensive than template-based methods because they sample the PES (e.g., using quantum chemistry calculations), which has been a major obstacle to applying them to heterogeneous systems in an exploratory context. For example, both SSW and AFIR have been applied to successfully (re)discover the relatively simple heterogeneous water-gas shift reaction occurring at a copper surface[29,34,35]. However, this analysis required millions of density functional theory (DFT) gradient calls, despite the small reactive system sizes. Due to the combinatorial scaling of possible reactions with the number of reacting atoms, a highly efficient reaction exploration scheme is indispensable to mitigate these costs. Heterogeneous applications also have several other technical obstacles to applying automated approaches, including the larger system sizes that are typical of surface models, the occurrence of spectator atoms that do not participate in reactions but nevertheless play important non-covalent or structural roles in the reaction pathways, and the use of periodic versus molecular models of the reacting systems. The optimal manner of overcoming these obstacles is an outstanding research question.

Here, we show how these problems can be addressed by combining a graph-based reaction exploration scheme, YARP, with a cluster model of a reactive interface. Ethylene oligomerization on silica-supported single site Ga^{3+} catalysts is used as a benchmark system for this approach based on the fact that, while considerable reaction data exists for this system, it still exhibits several unaccounted for product pathways. In particular, it has been previously reported that single site Ga^{3+} performs oligomerization chemistry via the classic Ga-alkyl-centered Cossee-Arlman mechanism with reasonably high selectivity to short linear alpha-olefins at 250 °C and 1 atm. However, side products, such as light alkanes and coke, have also been detected, especially at higher temperatures, and the mechanistic details of these side pathways remain unexplored[36,37]. Here, YARP not only (re)discovers the established Ga-alkyl-centered Cossee-Arlman catalytic cycle producing 1-butene, but also discovers relatively low-barrier TSs for side reactions leading to the formation of isomers of 1-butene, odd-number oligomers, alkanes, and coke precursors. These pathways exhibit diverse mechanisms including carbon-backbone lengthening, oligomer liberation, and hydrogen transfer to form alkanes. The kinetic significance of the TSs is rationalized by detailed analysis of the energy surfaces of three representative catalytic cycles.

Results and discussion

YARP workflow

The YARP workflow for automatically exploring a reaction network on silica-supported single site Ga catalysts consists of four components: Ga-silica cluster model construction, graph-based product enumeration, transition state localization and characterization, and reaction cycle investigation (Fig. 1). In brief, a $Si_8O_{12}(OH)_8$ cluster is adapted from Ugliengo et al.[38]. A Ga-ethyl site is then created by substitution of a Si-OH moiety with a Ga atom and adding an ethyl group to the Ga site, which serves as the starting reactant (Fig. 1a). Once the cluster model is built, graph-based product enumeration is recursively applied to a subset of reactive atoms, involving the gallium, carbon, and hydrogen atoms attached to carbon in the cluster model (shown as pink, gray, and white spheres in Fig. 1b) to generate potential products. Product enumeration is performed using generic elementary reaction steps (ERS) consisting of bonding changes in the reactants that can be applied combinatorially to define closed reaction spaces. After the ERS-based product enumeration, transition states of the enumerated reactions are characterized using the same procedure as our earlier

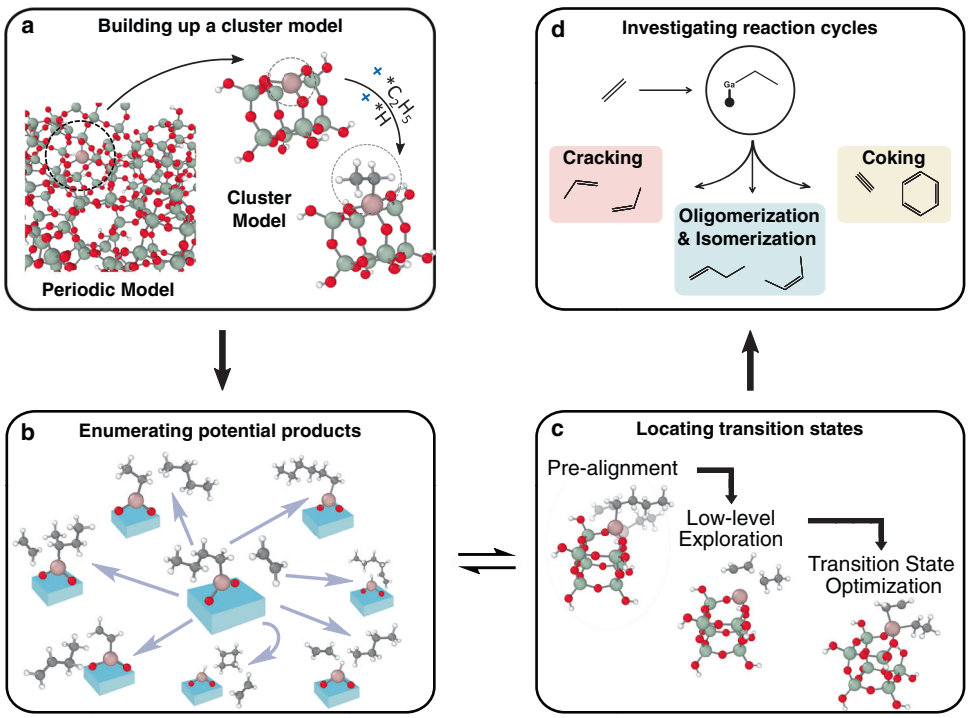

Fig. 1 | Overview of automated reaction network characterization applied to ethylene oligomerization on single site Ga^{3+} catalysts supported on silica. a A cluster model of a Ga^{3+} single site is built from a conventional periodic model. **b** Possible products are recursively enumerated from reactants/intermediates produced from elementary reaction steps on the cluster model. **c** Transition state localization and characterization are applied to each enumerated reaction. **d** Once the network exploration recursion terminates, detailed reaction mechanisms and relevant reaction cycles are summarized.

https://doi.org/10.1038/s41467-022-32514-7

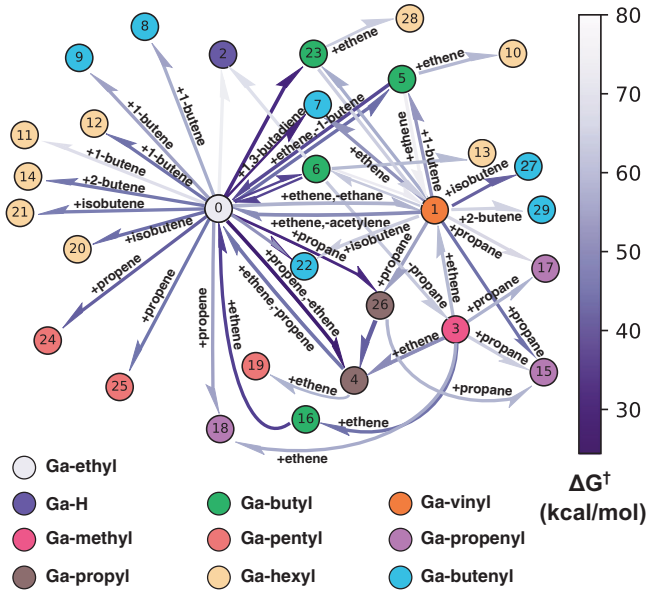

Fig. 2 | Gallium catalyzed olefin oligomerization reaction network obtained from YARP exploration. The edge colors reflect the activation free energy ($\Delta G^{\dagger}$) of each pathway as a measure of kinetic accessibility. Intermediate types are classified based on the alkyl and alkenyl attached to Ga and are denoted by different node colors.

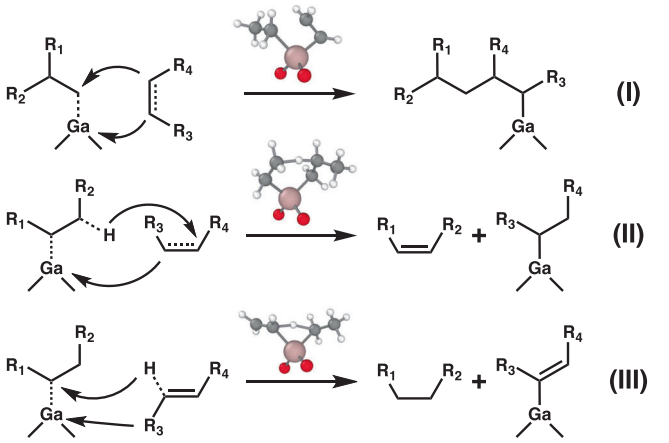

Fig. 3 | Three elementary reaction types identified during reaction network exploration. (I) olefin insertion; (II) β-hydride transfer; (III) α-hydride transfer. R_i refers to hydrogen, methyl, and ethyl groups. The presented TS structures represent the simplest examples of I-III reactions.

study[33], consisting of reaction geometry initialization and growing string method (GSM)[39] search performed with the GFN2-xTB semi-empirical model chemistry[40], followed by Berny optimization and intrinsic reaction coordinate (IRC) calculations performed using DFT (Fig. 1c). All of the kinetically-relevant products explored in this procedure serve as inputs for further reaction exploration until the discovery of new reactions is exhausted. In the end, important reaction pathways and catalytic cycles are analyzed in detail based on the reaction mechanisms discovered during the exploration (Fig. 1d). Detailed descriptions of each component are provided in the Methods section.

The complete reaction network involving Ga^{3+}

The overall reaction network that was generated by YARP for ethylene oligomerization on silica-supported single site Ga^{3+} is shown in Fig. 2. Network exploration was initialized with the Ga-ethyl species (node 0 in Fig. 2), which has been proposed as a key intermediate in the Cossee-Arlman ethylene oligomerization cycle[37]. After a single step of reaction enumeration and TS characterization, Ga-n-butyl, Ga-vinyl + ethane, and Ga-hydride + 1-butene, were identified as intended products of reactions between Ga-ethyl and ethylene. The free energies of activation ($\Delta G^{\dagger}$) of forming Ga-n-butyl, Ga-vinyl, and Ga-hydride are 44.1, 59.8, and 93.5 kcal/mol, respectively. Based on its high activation energy of formation, YARP excluded Ga-hydride from further exploration, whereas Ga-n-butyl and Ga-vinyl were included as active nodes for further reaction exploration. The high activation energy of β-hydrogen elimination to form Ga-hydride has also been observed in previous studies using conventional periodic DFT analysis[36,37]. The second step of exploration identifies Ga-n-butenyl (from Ga-vinyl, $\Delta G^{\dagger} = 53.2$ kcal/mol), acetylene (formed with Ga-ethyl from Ga-vinyl, $\Delta G^{\dagger} = 51.4$ kcal/mol), Ga-hexyl (from Ga-butyl, $\Delta G^{\dagger} = 61.3$ kcal/mol), 1-butene (formed with Ga-ethyl from Ga-butyl, $\Delta G^{\dagger} = 36.0$ kcal/mol) and butane (formed with Ga-vinyl from Ga-butyl, $\Delta G^{\dagger} = 76.4$ kcal/mol) as intended products. Notably, the lowest barrier step yielding 1-butene constitutes a rediscovery by the algorithm of the classic Cossee-Arlman mechanism that has previously been studied as the likely pathway for major product formation in this system. Based on the activation energies of the reactions at this iteration, Ga-n-butenyl was included as a new active node for further exploration (node 7), Ga-n-hexyl was classified as a terminal node (node 13) due to its size (see Methods for termination criteria), and 1-butene was added to the free-olefin list as a candidate for further reactions with the active nodes, Ga-ethyl (node 0) and Ga-vinyl (node 1). YARP recursively explored the reaction space via the same approach that was employed in the first and second iteration until all reactions within the prescribed constraints had been explored. All reactions explored with $\Delta G^{\dagger} < 80$ kcal/mol are presented in Fig. 2, and detailed geometries of each node can be found in the SI.

Three key reaction types occurring on Ga^{3+}

Three distinct types of reactions were discovered during the network exploration that are distinguished by their reactions with the adsorbed carbon species. All instances of each class exhibit $\Delta G^{\dagger} < 70$ kcal/mol. The first type is responsible for lengthening (or breaking, in the case of the reverse reaction) the carbon backbone (Type I in Fig. 3). The TS of the Type I reaction involves a "C=C" moiety bonding to the catalyst to form a four-coordinated Ga intermediate that precedes bond formation with an adsorbed alkyl species. The second type of reaction is β-hydride transfer that liberates an oligomer and closes an oligomerization cycle (Type II in Fig. 3). In the TS of this reaction type, the β-hydrogen of the adsorbed alkyl species transfers to an incoming olefin, which binds to the Ga center and becomes a new adsorbate. An oligomerization cycle can also be completed by a β-hydride elimination step to form Ga-hydride, but YARP predicts a much higher activation energy for this pathway (see Supporting Information for details). The facile β-hydride transfer step on Ga is a fundamentally different reaction channel from those occurring on traditional transition metal catalytic sites, where the hydrogen being transferred is not interacting with the metal center[41]. The third type of reaction produces an alkane, leaving a hydrogen-deficient adsorbed species, like Ga-vinyl (Type III in Fig. 3). The TS of the Type III reaction resembles that of Type II, except that the hydrogen transfers to the α-carbon. Alkane formation has been reported in multiple olefin oligomerization experiments[36,41–43], which may be explained by moderate barrier Type III pathways. Further, we hypothesize that the products of type III reactions may undergo additional type I and type II steps. The combination of type I-III reactions may eventually liberate alkynes and aromatics that are commonly considered to be coke precursors[44,45]

https://doi.org/10.1038/s41467-022-32514-7

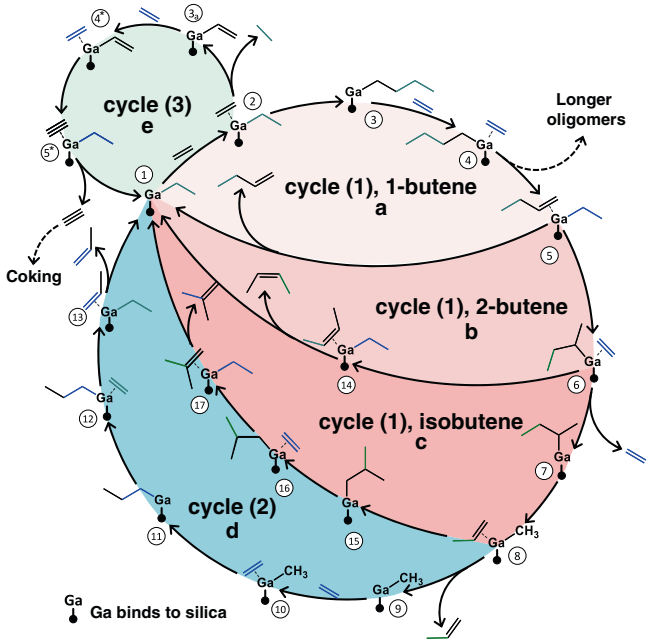

Fig. 4 | Proposed pathways of oligomerization (light red), isomerization (medium and dark red), cracking (blue), and coking (green) reactions. **a** Oligomerization pathway to 1-butene. Isomerization pathways forming (**b**) 2-butene and **c** isobutene. **d** Cracking pathway for propylene formation. **e** An example of a pathway forming alkanes and hydrogen-deficient products starting from Ga-ethyl. Similar alkane formation cycles can also occur for species ②, ④, ⑥, and ⑩.

(region e in Fig. 4). For example, a low barrier (~42 kcal/mol) Ga-catalyzed conversion of ethyne to benzene is compared with a non-catalyzed conversion in Fig. S3.

In addition to participating in the expected Cossee-Arlman oligomerization cycle, the elementary reaction types described above can contribute to several catalytic cycles for olefin isomerization and chain cracking (Fig. 4). In particular, following the formation of 1-butene and the recovery of the Ga-ethyl intermediate (species ⑤ in Fig. 4), where a Cossee-Arlman oligomerization cycle is nearly complete, the 1-butene molecule can be re-adsorbed with a simple rotation to react with Ga-ethyl through another β-hydride transfer (type II), producing Ga-2-butyl (species ⑥). This newly reported intermediate can undergo a facile type II reaction, forming cis- or trans-2-butene (only cis-2-butene formation is considered here, species ⑭). Butenes have been previously detected in experiments at 523 K and 1 atm, but distinguishing between cis and trans isomers is challenging due to their similar retention times in gas chromatography[36]. In an alternative pathway, Ga-2-butyl can undergo additional type I and II reactions to form Ga-methyl with physisorbed propylene (species ⑧). In addition, there can be another re-adsorption step of propylene on Ga-methyl, resulting in a Ga-isobutyl species (species ⑮), which eventually leads to isobutene (species ⑰). Throughout the isomerization and cracking pathways, the type III step can occur on each Ga-alkyl species, including Ga-ethyl, Ga-propyl, and Ga-butyl species. For example, a plausible pathway involving the type III reaction is outlined in the green circle of Fig. 4, where the resulting Ga-vinyl intermediate undergoes additional β-hydride transfer, leading to the formation of acetylene (a coke precursor).

Kinetic significance of type I-III transition states

Potential energy diagrams were used to compare the kinetic relevance of reactions cycles discovered for oligomerization, isomerization, cracking, and coking pathways (Fig. 5):

$$2C_2H_4 \longrightarrow C_4H_8 \quad (1)$$

$$3C_2H_4 \longrightarrow 2C_3H_6 \quad (2)$$

$$\frac{n+2}{2}C_2H_4 \longrightarrow C_2H_2 + C_nH_{2n+2}, \text{ where n} = 1, 2, \text{and } 3. \quad (3)$$

The competition between these cycles determines the selectivity of producing gaseous products and coke precursors. Cycle (1) involves ethylene dimerization products (Fig. 4a–c), including 1-butene and associated isomers. One catalytic cycle closes through an ethylene insertion (denoted as type I) and a β-hydride transfer (denoted as type II). Following additional type I-II steps occurring on Ga-ethyl with an adsorbed 1-butene (species ⑤), cis-2-butene and isobutene can also form (Fig. 4b, c). Cycle (2) involves the formation of cracking products, such as propylene, which are not favorable at a relatively low temperature (250 °C, 1 atm). One propylene molecule can be obtained through C-C bond breaking of a Ga-2-butyl species (reverse type I). The production of a second propylene molecule occurs via the same Cossee-Arlman oligomerization cycle initiated by the Ga-methyl intermediate (species ⑨, Fig. 4d). In cycle (3), the type III elementary step generates an alkane, which may occur for all Ga-alkyl intermediates, and an alkyne, such as acetylene, is formed that balances the stoichiometry. A relatively facile acetylene formation pathway occurs through a type II step occurring on the Ga-vinyl species from the type III reaction (species ⑪, Fig. 4e). Many other relatively low barrier pathways (≤70 kcal/mol) are discovered by YARP, including the formation of various C_nH_{2n} species, and C_nH_{2n-2} isomers. As discussed later, although these pathways are not as favorable as the primary Cossee-Arlman cycle, they may be responsible for coking and deactivation pathways and result in a broad diversity of possible products.

To validate the accuracy of the cluster model results, they were benchmarked against conventional periodic DFT with the NEB-Lanczos algorithm for localizing TSs of reaction cycles (1) and (2) (Fig. 5a). For this comparison, the geometries and energies for the cluster model were recalculated at the B3LYP-D3/6-311G(d,p) level of theory to minimize the DFT errors as a confounding factor when comparing the cluster and slab results (see the Supporting Information for additional details on selecting a suitable DFT level). Overall, periodic DFT and the cluster model predict similar binding energies, reaction energies, and reaction barriers, while some systematic deviations can be observed, including modestly higher activation energies predicted by the cluster model. For example, the cluster and periodic models predict activation energies of 1.8 and 1.6 eV, respectively, for the ethylene insertion step (species ② to species ③). The difference may be attributed to long-range order and reconstruction effects in the silica support, which may systematically lower activation energies, but are absent in the cluster model. Another systematic difference is that the binding energies obtained from the periodic model are consistently lower (less negative) than those obtained from the cluster model (Fig. 5a). In the periodic model, the Ga site is surrounded by siloxane frameworks with various ring sizes, which can contribute to a weakened binding due to steric effects, especially for the species with longer carbon chains or with physisorbed olefins. This effect is particularly significant for species ⑤, where 1-butene, the largest physisorbed reactant in our analysis, is involved. The comparison is also affected by the distinct functionals that were used due to their differing availability in the reference molecular and periodic quantum chemistry packages. Nevertheless, the two approaches predict similar relative barriers for all of the TSs under each elementary step type. The mean difference between activation energies for type I versus type II reactions are both

https://doi.org/10.1038/s41467-022-32514-7

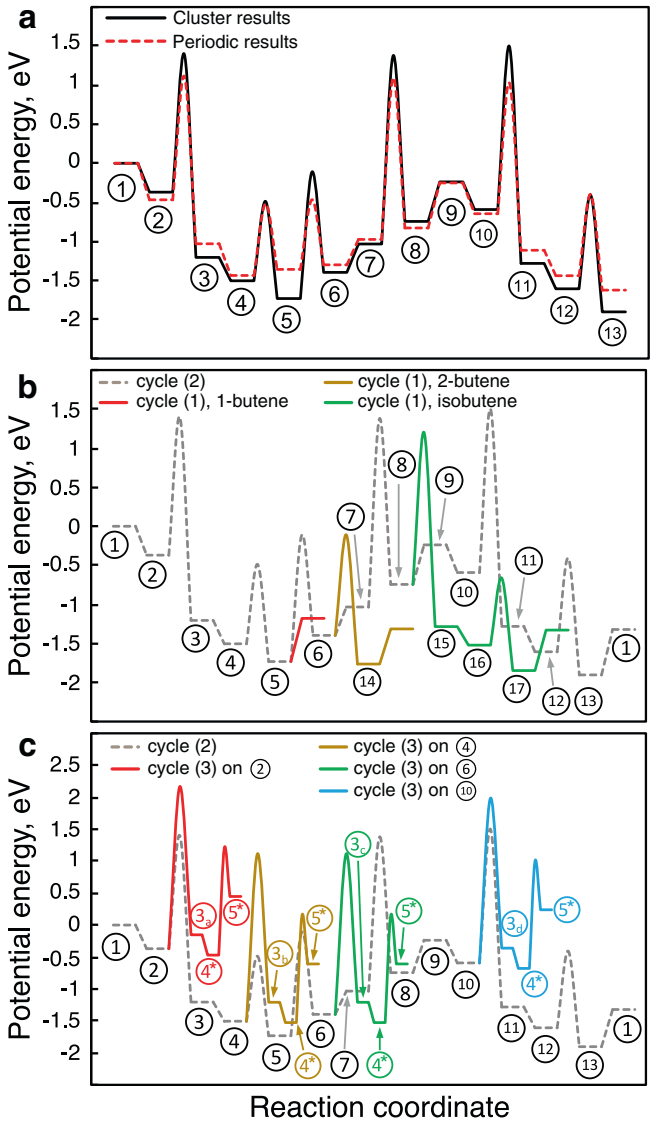

Fig. 5 | Energy diagrams of three kinetically relevant reaction cycles discovered within the reaction network. a Comparison of the energy landscape for cycle (2) using the cluster model and periodic slab. **b** Comparison of competing olefin formation pathways (colored) and cracking pathways (gray). **c** Comparison of competing acetylene formation pathways (colored) and cracking pathways (gray). The species are numbered based on the pathway diagram in Fig. 4, and the energies are reported with respect to the single point energy of Ga-ethyl plus a gaseous ethylene molecule.

0.8 eV, calculated by the cluster model and the periodic model. Further, all type I transition states are similarly accessible (i.e. barriers within 0.1 eV), and both models predict type II reactions to have consistently lower barriers. The overall agreement between the cluster and periodic models with respect to relative barrier heights validates the usefulness of the cluster models for performing reaction network exploration on silica surfaces.

Figure 5b, c outline the energy landscape comparison between the overall reaction cycles (1)-(3) using the cluster results. In cycle (1), where the carbon chain length doubles and 1-butene is formed (species ① - ⑤), the ethylene insertion involves a higher activation energy (1.76 eV) than the olefin liberation step (1.02 eV). This is consistent with a previous study showing that ethylene insertion is rate-determining in this system[37]. In the energy landscape of cycle (2), three type I elementary steps have relatively high activation energies: the ethylene insertion shared with cycle (1), the cracking of Ga-2-butyl (species ⑦, 2.40 eV), and the step forming Ga-1-propyl (species ⑩) from Ga-methyl and ethylene (2.08 eV). The cracking of Ga-2-butyl to form Ga-methyl and propylene involves the highest activation energy since it is a reversed type I step. Both periodic and YARP-cluster results predict that type I reactions are exothermic. Therefore, cycle (2) will not dominate the reaction network. Indeed, previous experimental results of Ga single sites show a strong selectivity to olefin oligomerization at 250 °C and 1 atm, forming short linear oligomers[36]. However, the activation free energy of cracking reduces as temperature increases (due to the entropically favored reverse type I step) in comparison to the formation of longer Ga-alkyl carbon chains, thus narrowing the energy difference between Ga-2-butyl (species ⑦) and the cracking TS. Based on the reaction entropies, the selectivity for propylene production over butene should increase with increased temperature or reduced pressure since cycle (2) produces a higher number of gas molecules than cycle (1). Finally, the high barrier of the reverse type I step provides a basis for the competition between type I and III reactions starting from the Ga-2-butyl species. In particular, the formation of 1-butane (species ⑥ - ⑰) can be competitive with cracking reactions (species ⑦ - ⑧). Subsequently, acetylene formation can occur via facile type II reactions (species ⑰ - ⑬, 1.68 eV). Therefore, our pathway analysis suggests that type III reactions are kinetically less favorable, but nevertheless represent side-reaction channels that become accessible as they compete with the reverse of type I step. With the formation of alkynes, other side reactions, such as aromatization and coking, may occur as subsequent thermodynamic products, especially at higher temperatures[46,47]. For example, the selectivities for producing alkynes, aromatics, and coke may increase with the pressure of ethylene since it shifts the reaction free energy of acetylene condensation to benzene. These insights into the reaction network and TSs downstream of Ga-ethyl formation also generalize to other metal single sites with similar electronic properties, such as Zn and Al[48,49]. In the Supporting information section 2, we have included the analogous energy diagrams for the reactions involving an Al single site. These results show energy barriers and a competition between cycles (1)-(3) that are similar to those of Ga.

Given the generic reaction rules and size constraints that were used to generate the ethylene oligomerization reaction network in this work, there are many opportunities for applying this approach to other heterogeneous systems. Among the salient details of the implementation to consider for future applications are the use of a cluster model as a surrogate for a periodic slab and the major speedup provided by semi-empirical quantum chemistry. Neither detail is intrinsic to applying YARP, and indeed, the cluster assumption was validated here and adopted out of convenience. There are, however, no obstacles to applying YARP using a periodic code, outside of cost. The applicability of this approach to other heterogeneous surfaces is therefore anticipated and is currently under investigation.

These results demonstrate how automatic exploration can be applied to heterogeneous catalytic networks using ethylene oligomerization catalyzed by a silica-supported Ga single site as a benchmark. The method (re)discovered the classic Ga-ethyl-centered Cossee-Arlman oligomerization cycle and several side-product pathways using generic graphical reaction rules and an ultra-low cost TS localization framework. The reaction network elaborated here represents the largest (in terms of both the number of reactive atoms, 20, and the branched alkyl intermediates up to C_6) that has been reported for a heterogeneous catalyst using an automated exploration algorithm based on quantum chemically characterized transition states. We do not expect this record to last, since there are relatively few impediments to applying this approach to other systems. Among the foreseeable challenges are that we have not considered adsorption and desorption steps that are often rate-limiting within catalytic cycles. When these steps are non-covalent and molecular in nature, the presented framework can already support them in scenarios where they

apply. A more fundamental limitation is that the catalytic surface was treated as static, save for the reactive atom and its nearest neighbors. This assumption will break down for cycles where the surface significantly restructures, such as for nanocatalysts, or where electrodeposition/desorption occurs. This could be accommodated by expanding the space of reactive atoms, but this will increase costs and presents opportunities for further innovation.

As network exploration becomes a predictive tool for catalyst screening, analyzing the reaction data and extracting general insights will become a bottleneck. As one example, characterizing the full reaction network reported here took around one-week of computational time on minimal resources (one node with 128 cores), but the costly manual process of classifying reaction mechanisms and closed cycles belonging to the same family of transformations took the majority of the time in this study. Although knowledge extraction of this kind will always be manual to a degree, there are several opportunities for streamlining this through automated mechanism classification based on orbital analyses and microkinetic modeling, both of which will be leveraged in the future.

Methods

Ga-silica cluster model construction

Ethylene oligomerization on single-site Ga^{3+}/SiO_2 was modeled based on a $Si_8O_{12}(OH)_8$ cluster that was adapted from Ugliengo et al.[38], wherein a Ga^{3+} single-site was created by substitution of a Si-OH moiety with a Ga atom. A Ga-ethyl site was, in turn, created by adding an ethyl group to the Ga site and a proton to the adjacent oxygen atom to maintain charge balance (Fig. 1a). The focus of this effort was to establish the reaction network downstream of Ga-ethyl formation since this is where several mechanistic gaps exist in the catalytic cycle, especially with respect to side-product formation. In particular, previous work has established facile initiation of the Ga-ethyl species from a bare single site and gaseous ethylene, as well as a low-barrier Ga-ethyl-centered Cossee-Arlman pathway to ethylene oligomerization. We have previously examined the competing Ga-hydride-centered Cossee-Arlman cycle and the empty Ga site-centered proton transfer cycle, and facile formation of the Ga-ethyl was observed in all cases[37]. Nevertheless, the choice to use Ga-ethyl as a starting point for the exploration is non-essential, since the network exploration used here automatically discovered all possible C_1-C_4 Ga sites and gas-phase olefins during the reaction network exploration.

The cluster model can be viewed as a finite portion of the solid silica surface, with the dangling oxygen atoms passivated by hydrogen atoms[38,50]. The localized nature of oxides and the Ga^{3+} center make the cluster model a credible approximation for assessing surface reactivity, as demonstrated by our previous analysis on ethylene oligomerization[37]. Cluster models provide a potentially useful bridge between the active methods development occurring for molecular reaction network exploration and reaction networks occurring at reactive surfaces. Here, comparisons of the energies and barrier heights calculated on periodic surfaces and the cluster model were used to further validate this assumption (typical differences in activation energies were around 0.3 eV, Table S2).

Reaction network characterization

The YARP methodology was used to enumerate the reactions and characterize the TSs associated with the Ga-ethyl species modeled in the presence of excess ethylene. For a more detailed description of the YARP methodology, we direct readers to our previous publication[33]. In the following sections, we focus on the modifications that were implemented to the reaction enumeration and reaction pathway construction steps to adapt YARP to explore ethylene oligomerization on single-site Ga^{3+}/SiO_2. The salient features of YARP are that all exploration steps are automated and based on quantum chemically characterized transition states rather than particular user-defined reaction mechanisms. For these reasons, minimal human intervention is required for the network exploration after the initial setup of the reacting system and the termination criteria.

Product enumeration. The YARP methodology consists of recursively applying graph-based elementary reaction steps (ERS) of the form break *m* bonds and form *n* bonds (bmfn). These rules are sufficiently generic to recapitulate many reactions without relying on explicit reaction templates, and they define reaction spaces that can be comprehensively explored (as an example, all b2f2 pathways of a given set of reactants constitute a well-defined set). For neutral closed-shell systems, the simplest reaction that yields non-trivial closed-shell products is b2f2 (describing, for example, an E2 reaction); however, single-step b3f3 reactions, such as Diels-Alder and Claisen rearrangements, are also common. Here, we applied a compromise scheme, including all b2f2 reactions and the subset of b3f3 reactions that involved at least one double-bond breaking. These ERSs were applied to the gallium, carbon, and hydrogen atoms attached to carbon in the cluster model (shown as pink, gray, and white spheres in Fig. 1b) to enumerate all products for each reactant in the network. Reactions that did not involve Ga (e.g., non-catalytic reactions between ethane and other alkyl products), and reactions that yielded species with more than five carbons were discarded from consideration.

Transition state localization. After product enumeration, YARP attempts to localize TSs for each reaction. This procedure consists of initializing a reaction geometry, estimating the TS geometry at the semi-empirical GFN2-xTB[40] level using the growing string method (GSM)[51], TS optimization at the DFT level using Berny optimization, and intrinsic reaction coordinate (IRC) calculations to classify the resulting TSs (Fig. 1c). For the geometry initialization, the joint-optimization algorithm reported in the original YARP publication was retained, with the exception that the positions of all silica atoms except the two oxygen atoms attached to gallium were fixed to preserve the initial DFT-level cluster structure. The joint-optimization algorithm is designed to produce minimally displaced reactant-product conformers that are well-conditioned for a double-ended TS search[33]. These optimized structures were then used as the fixed endpoints for GSM calculations. After convergence, the highest energy node along the reaction pathway was selected as the initial guess for an unconstrained DFT level Berny transition state optimization. The final TS, after successful convergence of previous steps with a structure exhibiting a single imaginary frequency, was characterized by an IRC calculation to ensure its correspondence to the attempted reaction. When the two end nodes obtained by the IRC calculation matched the input reactant and product, the attempted reaction was classified as "intended" and included in the reaction network.

Reaction network construction. To construct the reaction network, interleaved product enumeration and TS localization was performed until the discovery of new reactions was exhausted. At each stage of this iteration, the Ga-products of the previous iteration served as potential reactants for the next iteration subject to conditions that were designed to manage the size of the reaction network while being relatively permissive in terms of exploring new reactivities. Specifically, Ga-adsorbed species were only included as potential reactants at the next iteration if they were connected to the rest of the network by an intended reaction with an activation energy less than 3 eV (~70 kcal/mol, above which the reaction will require extremely high temperatures to be kinetically competitive, and under which conditions thermal reactions will, in any case, be dominant[52,53]). It is possible for a Ga-adsorbed species to fail the activation energy constraint at an early iteration, but then to be included later if an alternative pathway is discovered. Additionally, the size of the reactant species attached to the gallium site was limited to butyl and smaller moieties to avoid the

trivial growth of the network due to lengthening of the carbon backbone. All of the Ga intermediates obtained without violating these constraints were included as species capable of participating in reactions in the next iteration. At each iteration, the set of explored reactants consisted of all combinations of the active Ga-adsorbed species and any free olefins that were produced as products during previous iterations of exploration. Thus, a newly generated Ga-adsorbed species would participate in up to $n+1$ separate reactant sets, where n is the collection of free olefins discovered up until that point of exploration, and the plus one corresponds to unimolecular reactions involving the Ga-species. Reactant combinations involving more than six carbons were discarded to avoid uninteresting growth of the network. For each set of reactants, the ERS-generated reactions were characterized, and the recursion ended after no new reactions were discovered.

Periodic DFT calculations

The reaction energies and reaction barriers for a subset of pathways were recalculated on an amorphous silica slab model with a large unit cell (21.6 Å × 21.6 Å × 34.5 Å) and compared with cluster model results for validation. These calculations were performed on the amorphous structure reported by Comas-Vives, generated from an annealing process using classical molecular dynamics and multiple dehydration processes that result in a high level of dehydroxylation (1.1 silanol moieties nm^{-1}) and siloxane rings with different sizes[54]. The Ga-ethyl moiety was created using the same approach as was employed in the cluster models. Previous studies indicate that the less-constrained, three-coordinated Ga sites are responsible for the oligomerization chemistry, whereas the constrained four-coordinated sites are relatively inactive due to strong steric hindrance effects[37]. Therefore, the periodic slab calculations focused on the less constrained Ga site, as represented in our cluster model.

Computational details

YARP used Gaussian 16 as the reference quantum chemistry engine for the DFT calculations associated with the Berny optimizations and IRC calculations[55]. Calculations were performed at the B3LYP/6-31G level of theory during network exploration. The geometries of the TSs, reaction energies, and reaction barriers were recalculated at the B3LYP-D3/6-311G(d,p) level of theory for validation and comparison with periodic calculations. This level of theory showed a balance of accuracy and cost describing the Ga^{3+}/SiO_2 system (Figs. S1 and S2). The GSM calculations were performed by the pyGSM package using eleven images, fixed reactant and product geometries, and other default hyperparameters[39]. All GFN2-xTB calculations were performed with the xTB program (version 6.2.3)[40].

Periodic DFT calculations were performed using Vienna Ab-initio Simulation Package (VASP, 5.4.1), where planewave basis sets describe the Kohn-Sham orbitals, and the Kohn-Sham equations were solved self-consistently[56–60]. The BEEF-VdW exchange-correlation functional with projector augmented wave (PAW) pseudopotentials was employed[60–62]. A Monkhorst-Pack k-sampling was used, and a k point grid of 2 × 2 × 1 was applied. A cutoff energy of 400 eV and a force-convergence criterion of 20 meV $Å^{-1}$ for energy local minima were used. The climbing image nudged elastic band (CI-NEB) method was used as a first step to locate transition states[63,64]. Seven images were used in each NEB calculation, as generated by the Image Dependent Pair Potential (IDPP) tool[65]. Following each NEB calculation, Lanczos diagonalization was used to identify the transition state with a greater accuracy[66]. The force convergence criterion of 20 meV $Å^{-1}$ was used for TS optimization. All TSs were confirmed to exhibit only one imaginary frequency.

Data availability

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files. Source data for Figs. 2 and 5 are available in Source Data. The raw data generated in this study have been deposited in the YARP database and are available at https://doi.org/10.6084/m9.figshare.14766624[67], including raw output files and molecular (reactants, products and transition states) geometries.

Code availability

The version of YARP used in this study is available through GitHub under the GNU GPL-3.0 License [https://github.com/zhaoqy1996/YARP]. The specific version of the package used to generate the results in this study can be found at https://doi.org/10.5281/zenodo.6828628 [68].

References

1. Ertl, G. Elementary Steps in Heterogeneous Catalysis. *Angew. Chem., Int. Ed.* **29**, 1219–1227 (1990).
2. Chorkendorff, I. & Niemantsverdriet, J. W. *Concepts of modern catalysis and kinetics*. (John Wiley & Sons, 2003)
3. Medford, A. J. et al. From the Sabatier principle to a predictive theory of transition-metal heterogeneous catalysis. *J. Catal.* **328**, 36–42 (2015).
4. Mazeau, E. J., Satpute, P., Blöndal, K., Goldsmith, C. F. & West, R. H. Automated Mechanism Generation Using Linear Scaling Relationships and Sensitivity Analyses Applied to Catalytic Partial Oxidation of Methane. *ACS Catal.* **11**, 7114–7125 (2021).
5. Xu, J., Cao, X.-M. & Hu, P. Perspective on computational reaction prediction using machine learning methods in heterogeneous catalysis. *Phys. Chem. Chem. Phys.* **23**, 11155–11179 (2021).
6. Steiner, M. & Reiher, M. Autonomous Reaction Network Exploration in Homogeneous and Heterogeneous Catalysis. *Top. Catal.* **65**, 1–34 (2022).
7. Abild-Pedersen, F. et al. Scaling Properties of Adsorption Energies for Hydrogen-Containing Molecules on Transition-Metal Surfaces. *Phys. Rev. Lett.* **99**, 016105 (2007).
8. Nørskov, J. et al. Universality in Heterogeneous Catalysis. *J. Catal.* **209**, 275–278 (2002).
9. Greeley, J. Theoretical Heterogeneous Catalysis: Scaling Relationships and Computational Catalyst Design. Annu. Rev. Chem. *Biomol. Eng.* **7**, 605–635 (2016).
10. Bligaard, T. et al. The Brønsted-Evans-Polanyi relation and the volcano curve in heterogeneous catalysis. *J. Catal.* **224**, 206–217 (2004).
11. Kropp, T. & Mavrikakis, M. Brønsted-Evans-Polanyi relation for CO oxidation on metal oxides following the Mars-van Krevelen mechanism. *J. Catal.* **377**, 577–581 (2019).
12. Vojvodic, A. et al. On the behavior of Brønsted-Evans-Polanyi relations for transition metal oxides. *J. Chem. Phys.* **134**, 244509 (2011).
13. Loffreda, D., Delbecq, F., Vigné, F. & Sautet, P. Fast Prediction of Selectivity in Heterogeneous Catalysis from Extended Brønsted-Evans-Polanyi Relations: A Theoretical Insight. *Angew. Chem.* **121**, 9140–9142 (2009).
14. Greeley, J. & Mavrikakis, M. Alloy catalysts designed from first principles. *Nat. Mater.* **3**, 810–815 (2004).
15. Greeley, J., Jaramillo, T. F., Bonde, J., Chorkendorff, I. & Nørskov, J. K. Computational high-throughput screening of electrocatalytic materials for hydrogen evolution. *Nat. Mater.* **5**, 909–913 (2006).
16. Greeley, J. et al. Alloys of platinum and early transition metals as oxygen reduction electrocatalysts. *Nat. Chem.* **1**, 552–556 (2009).
17. Studt, F. et al. Discovery of a Ni-Ga catalyst for carbon dioxide reduction to methanol. *Nat. Chem.* **6**, 320–324 (2014).
18. Unsleber, J. P. & Reiher, M. The exploration of chemical reaction networks. *Annu. Rev. Phys. Chem.* **71**, 121–142 (2020).
19. Nilsson, A. & Pettersson, L. G. *Chemical Bonding at Surfaces and Interfaces*. (Elsevier, 2008). pp 57–142

20. Li, J. et al. Mechanistic Variants in Gas-Phase Metal-Oxide Mediated Activation of Methane at Ambient Conditions. *J. Am. Chem. Soc.* **138**, 11368–11377 (2016).
21. Mleczko, L. & Baerns, M. Catalytic oxidative coupling of methane-reaction engineering aspects and process schemes. *Fuel Process. Technol.* **42**, 217–248 (1995).
22. Vernuccio, S., Bickel, E. E., Gounder, R. & Broadbelt, L. J. Microkinetic Model of Propylene Oligomerization on Brønsted Acidic Zeolites at Low Conversion. *ACS Catal.* **9**, 8996–9008 (2019).
23. Goldsmith, C. F. & West, R. H. Automatic generation of microkinetic mechanisms for heterogeneous catalysis. *J. Phys. Chem. C.* **121**, 9970–9981 (2017).
24. Ismail, I., Stuttaford-Fowler, H. B., OchanAshok, C., Robertson, C. & Habershon, S. Automatic proposal of multistep reaction mechanisms using a graph-driven search. *J. Phys. Chem. A* **123**, 3407–3417 (2019).
25. Blondal, K. et al. Computer-Generated Kinetics for Coupled Heterogeneous/Homogeneous Systems: A Case Study in Catalytic Combustion of Methane on Platinum. *Ind. Eng. Chem. Res.* **58**, 17682–17691 (2019).
26. 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).
27. Gao, C. W., Allen, J. W., Green, W. H. & West, R. H. Reaction Mechanism Generator: Automatic construction of chemical kinetic mechanisms. *Comput. Phys. Commun.* **203**, 212–225 (2016).
28. Maeda, S., Taketsugu, T. & Morokuma, K. Exploring transition state structures for intramolecular pathways by the artificial force induced reaction method. *J. Comput. Chem.* **35**, 166–173 (2014).
29. Jafari, M. & Zimmerman, P. M. Uncovering reaction sequences on surfaces through graphical methods. *Phys. Chem. Chem. Phys.* **20**, 7721–7729 (2018).
30. Shang, C. & Liu, Z. P. Stochastic surface walking method for structure prediction and pathway searching. *J. Chem. Theory Comput.* **9**, 1838–1845 (2013).
31. Zimmerman, P. M. Automated discovery of chemically reasonable elementary reaction steps. *J. Comput. Chem.* **34**, 1385–1392 (2013).
32. Sugiyama, K., Sumiya, Y., Takagi, M., Saita, K. & Maeda, S. Understanding CO oxidation on the Pt (111) surface based on a reaction route network. *Phys. Chem. Chem. Phys.* **21**, 14366–14375 (2019).
33. Zhao, Q. & Savoie, B. M. Simultaneously improving reaction coverage and computational cost in automated reaction prediction tasks. *Nat. Comput. Sci.* **1**, 479–490 (2021).
34. Zhang, X.-J., Shang, C. & Liu, Z.-P. Stochastic surface walking reaction sampling for resolving heterogeneous catalytic reaction network: A revisit to the mechanism of water-gas shift reaction on Cu. *J. Chem. Phys.* **147**, 152706 (2017).
35. Iwasa, T. et al. Combined Automated Reaction Pathway Searches and Sparse Modeling Analysis for Catalytic Properties of Lowest Energy Twins of Cu13. *J. Phys. Chem. A* **123**, 210–217 (2018).
36. LiBretto, N. J. et al. Olefin oligomerization by main group Ga^{3+} and Zn^{2+} single site catalysts on SiO_2. *Nat. Commun* **12**, 2322 (2021).
37. Xu, Y., LiBretto, N., Zhang, G., Miller, J. & Greeley, J. First Principles Analysis of Ethylene Oligomerization on Single-site Ga^{3+} Catalysts Supported on Amorphous Silica. *ACS Catal.* **12**, 5416–5424 (2022).
38. Roggero, I., Civalleri, B. & Ugliengo, P. Modeling physisorption with the ONIOM method: the case of NH3 at the isolated hydroxyl group of the silica surface. *Chem. Phys. Lett.* **341**, 625–632 (2001).
39. Aldaz, C., Kammeraad, J. A. & Zimmerman, P. M. Discovery of conical intersection mediated photochemistry with growing string methods. *Phys. Chem. Chem. Phys.* **20**, 27394–27405 (2018).
40. Bannwarth, C., Ehlert, S. & Grimme, S. GFN2-xTB?An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).
41. Sydora, O. L. Selective Ethylene Oligomerization. *ACS Catal.* **38**, 997–1010 (2019).
42. Moussa, S., Concepción, P., Arribas, M. A. & Martínez, A. Nature of Active Nickel Sites and Initiation Mechanism for Ethylene Oligomerization on Heterogeneous Ni-beta Catalysts. *ACS Catal.* **8**, 3903–3912 (2018).
43. Dagle, V. L. et al. Production and fuel properties of iso-olefins with controlled molecular structure and obtained from butene oligomerization. *Fuel* **277**, 118147 (2020).
44. Cai, H., Krzywicki, A. & Oballa, M. C. Coke formation in steam crackers for ethylene production. *Chem. Eng. Process.* **41**, 199–214 (2002).
45. Kopinke, F. D., Zimmermann, G., Reyniers, G. C. & Froment, G. F. Relative rates of coke formation from hydrocarbons in steam cracking of naphtha. 2. paraffins, naphthenes, mono-, di-, and cycloolefins, and acetylenes. *Ind. Eng. Chem. Res.* **32**, 56–61 (1993).
46. Graff, M. J. & Albright, L. F. Coke deposition from acetylene, butadiene and benzene decompositions at 500-900° C on solid surfaces. *Carbon* **20**, 319–330 (1982).
47. Trotuş, I.-T., Zimmermann, T. & Schüth, F. Catalytic Reactions of Acetylene: A Feedstock for the Chemical Industry Revisited. *Chem. Rev.* **114**, 1761–1782 (2014).
48. Werghi, B. et al. Well-defined silica supported aluminum hydride: another step towards the utopian single site dream? *Chem. Sci.* **6**, 5456–5465 (2015).
49. Werghi, B. et al. Exploiting the interactions between the ruthenium Hoveyda-Grubbs catalyst and Al-modified mesoporous silica: the case of SBA15 vs. KCC-1. *Chem. Sci.* **9**, 3531–3537 (2018).
50. Tielens, F., Gierada, M., Handzlik, J. & Calatayud, M. Characterization of amorphous silica based catalysts using DFT computational methods. *Catal. Today* **354**, 3–18 (2020).
51. Zimmerman, P. M. Growing string method with interpolation and optimization in internal coordinates: Method and examples. *J. Chem. Phys.* **138**, 184102 (2013).
52. Beadle, P. C., Golden, D. M., King, K. D. & Benson, S. W. Pyrolysis of cyclobutane. *J. Am. Chem. Soc.* **94**, 2943–2947 (1972).
53. Vreeland, R. W. & Swinehart, D. Mass Spectometric Investigation of the Thermal Decomposition of Cyclobutane at Low Pressures. *J. Am. Chem. Soc.* **85**, 3349–3353 (1963).
54. Comas-Vives, A. Amorphous SiO_2 surface models: energetics of the dehydroxylation process, strain, ab initio atomistic thermodynamics and IR spectroscopic signatures. *Phys. Chem. Chem. Phys.* **18**, 7475–7482 (2016).
55. Frisch, M. J. et al. *Gaussian 16 Revision C.01*. (Gaussian Inc, Wallingford CT, 2016).
56. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
57. 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).
58. Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
59. Kresse, G. & Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251–14269 (1994).
60. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
61. Wellendorff, J. et al. Density functionals for surface science: Exchange-correlation model development with Bayesian error estimation. *Phys. Rev. B* **85**, 235149 (2012).

62. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
63. Henkelman, G., Uberuaga, B. P. & Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **113**, 9901–9904 (2000).
64. Henkelman, G. & Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.* **113**, 9978–9985 (2000).
65. Smidstrup, S., Pedersen, A., Stokbro, K. & Jónsson, H. Improved initial guess for minimum energy path calculations. *J. Chem. Phys.* **140**, 214106 (2014).
66. Olsen, R. A., Kroes, G. J., Henkelman, G., Arnaldsson, A. & Jónsson, H. Comparison of methods for finding saddle points without knowledge of the final states. *J. Chem. Phys.* **121**, 9776–9792 (2004).
67. Zhao, Q. YARP reaction dataset. https://doi.org/10.6084/m9.figshare.14766624 (2022).
68. Zhao, Q. YARP: Yet Another Reaction Program. https://doi.org/10.5281/zenodo.6828628 (2022).

Acknowledgements

The work performed by Q.Z. and B.M.S was made possible by the Office of Naval Research (ONR) through support provided by the Energetic Materials Program (MURI grant number: N00014-21-1-2476, Program Manager: Dr. Chad Stoltz). Y.X. and J.G. acknowledge support from the National Science Foundation through the Center for Innovative and Sustained Transformation of Alkane Resources (CISTAR) under Cooperative Agreement No. EEC-1647722. Use of the Center for Nanoscale Materials, an Office of Science user facility, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02- 06CH11357. Use of the National Energy Research Scientific Computing Center is also gratefully acknowledged. Q.Z., Y.X., J.G., and B.M.S. would like to thank Prof. Jeffrey T. Miller from Purdue University and Dr. Nicole J. LiBretto from National Renewable Energy Laboratory (NREL) for the insightful discussions.

Author contributions

The project was conceived by Q.Z., Y.X., J.G., and B.M.S. Q.Z. and Y.X. designed and performed the simulations. Y.X. established the structures of catalytic sites on cluster and periodic silica models. Q.Z. performed the automated reaction network exploration and the related cluster DFT calculations. Y.X. performed the periodic DFT calculations. All authors analyzed the results and prepared the manuscript.

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/s41467-022-32514-7.

Correspondence and requests for materials should be addressed to Jeffrey Greeley or Brett M. Savoie.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer review reports are available.

Reprints and permission information is available at 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 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. The images or other third party material in this article 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 http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022