Liquid Organic Hydrogen Carriers: High-throughput Screening of Homogeneous Catalysts

Gaurav Vishwakarma[1, *] and Johannes Hachmann[1, 2, 3, †]

[1] *Department of Chemical and Biological Engineering, University at Buffalo, The State University of New York, Buffalo, NY 14260, United States*
[2] *Computational and Data-Enabled Science and Engineering Graduate Program, University at Buffalo, The State University of New York, Buffalo, NY 14260, United States*
[3] *New York State Center of Excellence in Materials Informatics, Buffalo, NY 14203, United States*

In the times of an ever-increasing rate of global warming and rapidly depleting fossil fuels, renewable sources of energy are attracting vast attention and numerous efforts have been directed towards achieving a hydrogen-based economy over the past few years. However, the biggest technical challenge so far has been the development of materials and the required infrastructure for efficient storage and transportation of hydrogen. To this end, liquid organic hydrogen carriers (LOHCs) have been extensively studied as they provide a safer alternative to storing high-purity hydrogen using the existing fuel infrastructure. However, commercial applications of LOHCs are only feasible when expensive, noble metal catalysts are substituted with inexpensive but equally efficient alternatives. In this work, we employ our group's cyberinfrastructure for the data-driven discovery and design of novel catalysts for LOHCs. We screen a library of homogeneous Ir-based pincer catalysts for the dehydrogenation of perhydro-N-ethyl carbazole. We develop a computational protocol to evaluate these catalysts based on thermodynamic parameters calculated using Density Functional Theory. Next, we use this data to train machine learning models for predicting the Gibbs free energies of the reactions and analyze hidden structure-property relationships in these systems.

I. HYDROGEN CARRIERS

Our present-day energy system is dominated by the systematic exploitation of large reservoirs of carbon-based fossil fuels. Although fossil fuel reserves are expected to last several centuries at the current rate of consumption, the rate of global climate change is the primary driving force behind the search for large-scale renewable energy sources. However, a major shortcoming in the utility of renewable sources of energy is the disparity in the geographic locations (and/or timings) associated with the production and demand for energy. This necessitates the development of energy storage solutions in different media, which can then be transported to the region of demand in order to ensure an uninterrupted supply of energy. Even batteries that promise large storage capacities fall short in the context of storage for extensive time periods, as well as their low gravimetric energy density. On the other hand, hydrogen which only yields water upon combustion, has a high gravimetric energy density of 28.68 $kcal/g$ (approx. 3× gasoline), and is thus a prime candidate for large-scale applications in the automotive industry.

Several aspects of a hydrogen-powered economy, such as production, storage, and distribution of hydrogen are an active field of research [1]. However, rapid storage of hydrogen in sufficient quantities under mild operating conditions is perhaps the biggest challenge in its commercialization [2, 3]. Of the several strategies for hydrogen storage, the focus of this work is on chemically-bound hydrogen storage in liquid organic hydrogen carriers (LOHCs) [4–9].

LOHCs provide a flexible media for the storage and transportation of high-purity hydrogen, without losses even for long distances, using the existing fuel infrastructure. They have a higher gravimetric as well as a higher volumetric energy density as compared to other fuels used for commercial applications. Hydrogen is stored in these liquid-phase organic compounds for transportation, and is extracted on-demand. LOHCs retain their core molecular structure after both catalytic reactions, hydrogenation and dehydrogenation, thus eliminating the need for a new energy carrier after every cycle. While the hydrogenation reaction is exothermic in most cases, the dehydrogenation reaction is endothermic, which allows for coupling of these reactions within a reactor in order to maximize efficiency.

The stability of the hydrogenated LOHC is ensured by favorable hydrogenation enthalpy, which in turn ensures acceptable dehydrogenation temperatures. The desired enthalpy range has been stated to be in the range of 40-55 $kJ/molH_2$ [7, 10]. For most applications, the thermodynamic reversibility of the hydrogenation/dehydrogenation reactions is the primary factor in the choice of LOHCs. The storage density, stability, cost, performance and operational safety have also been weighed in this context. The U.S. Department of Energy aims to achieve 0.055 kg H_2/kg system or 0.04 kg H_2/L system for onboard hydrogen storage for light-duty fuel cell vehicles by 2025. As energy storage media, the present research focus is on the safety and robustness of the overall process, along with the selectivity and kinetics of the catalysts.

* gvishwak@buffalo.edu
† hachmann@buffalo.edu

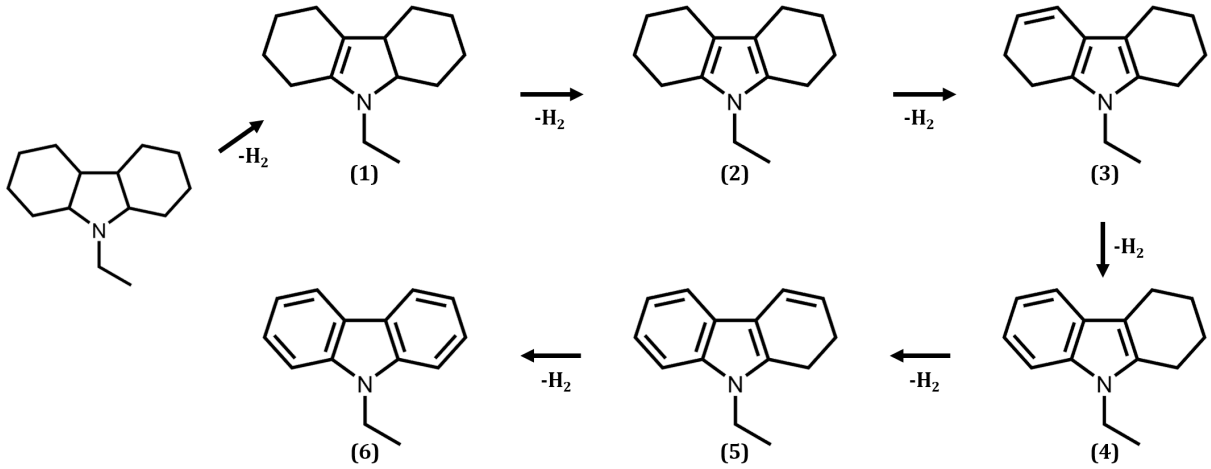

FIG. 1: Schematic for the dehydrogenation mechanism of perhydro-N-ethyl carbazole.

Over the past few years, various classes of cycloalkanes have been studied and reported as potential LOHC candidates. Abundantly utilized in industries (especially oil refineries), chemicals such as benzene, toluene and decalin are prime examples of this category. With a high hydrogen carrying capacity and a high boiling point, these compounds are quite suitable for large-scale applications. However, they are associated with high dehydrogenation enthalpies which requires high operational temperatures of up to $300°C$. It has been observed that the dehydrogenation reaction is more favorable thermodynamically with the introduction of heteroatoms such as nitrogen, oxygen, phosphorous or boron in the cycloalkanes. N-substituted five-membered rings, polycyclic hydrocarbons, lower steric hindrance around heteroatoms are other factors that also favor the reaction thermodynamics [4, 11–16].

Among heterocycles, N-ethyl carbazole (NEC) is perhaps the most prominent example that has been studied in the context of reversible dehydrogenation [17–24]. NEC has a hydrogen carrying capacity of 5.8 $wt.\%$, and a dehydrogenation enthalpy in the range 50-53 $kJ/mol\ H_2$. Upon complete dehydrogenation, perhydro-NEC (or H12-NEC) releases 6 moles of hydrogen in three stages, with H8-NEC and H4-NEC as the most stable intermediates [25–27]. The dehydrogenation reaction starts from the five-membered ring and subsequently proceeds in the two six-membered rings as shown in Fig. 1 [28, 29].

For practical applications, the design of catalysts for both hydrogenation and dehydrogenation reactions is aimed at high selectivity, but under sufficiently mild reaction conditions since LOHCs are susceptible to thermal degradation [30]. A large number of heterogeneous catalysts (with or without supports), that are associated with high selectivity and yield, have been reported in literature for several classes of cycloalkanes as well as heterocycles [31–45]. However, these noble-metal catalysts are expensive and deactivation of highly active catalysts due to coking is a major concern. Thus, stable homogeneous catalysts with lower manufacturing costs and comparable catalytic activities are promising alternatives [46–60]. One of the primary advantages of homogeneous catalysts is their tunability at the molecular level, which allows their properties to be tailored towards specific requirements. For the dehydrogenation of H12-NEC, Wang *et al.* reported three homogeneous Ir based tridentate (PCP pincer) ligand complexes at $200°C$ [61].

Since a systematic exploration of catalyst scope is very time- and resource- intensive, this work aims to build upon the results of Wang *et al.* to propose novel homogeneous pincer complexes (that are thermally stable even at higher temperatures) for the complete dehydrogenation of H12-NEC. Using computational studies, the process of characterization and discovery of novel catalysts for LOHCs can be accelerated at a significantly faster pace and a substantially lower cost. These calculations can then help in guiding the experimentalists towards only the most promising set of candidates.

With respect to organic hydrogen carriers, only a handful of iridium, ruthenium and iron based pincer complexes have been discovered for dehydrogenation. Therefore, we address this non-systematic exploration of the catalyst scope for hydrogen carriers, using the software ecosystem developed in our research group [62]. We build a library of candidate molecules and cast them into a high-throughput screening (HTPS) protocol using our group's software ecosystem for the data-driven discovery of materials. In Sec. II, we introduce the physical foundations of our proposed protocol and discuss the details of the computational approach employed. We present the results of the quantum chemical calculations and the performance of predictive models trained on this data in Sec. III. Finally, we summarize our findings in Sec. IV.

II. METHODS AND COMPUTATIONAL DETAILS

In this work, we follow the accepted mechanism for the catalytic dehydrogenation of cycloalkanes by the Ir-PCP type pincer complexes as reported in Refs. [48, 63]. The basis for our HTPS protocol is the calculation of thermodynamic barriers ($\Delta G_r = \Delta G_{f,products} - \Delta G_{f,reactants}$), for each of the six dehydrogenation steps of H12-NEC (shown in Fig. 1) using the aforementioned reaction mechanism. We use our group's library generator code, *ChemLG* [64], to generate initial geometries for 1) catalyst molecules, 2) NEC structures corresponding to the six dehydrogenation reactions, and 3) six catalyst-NEC adducts for every catalyst.

A. Virtual Screening Library

Using SMILES strings as the input to *ChemLG*, we create a library of 3066 candidate molecules for catalysing the dehydrogenation reaction of H12-NEC. We run *ChemLG* in a constrained combinatorial mode to achieve sequential linking of molecular fragments (or building blocks) at the four sites on the reference pincer catalyst shown in Fig. 2(a). We use a set of 18 building blocks compiled from homogeneous catalysts reported in literature for C-H bond activation reactions. The building blocks, as well as the specific sites on the reference catalyst at which they can be attached, are shown in Fig. 2(b).

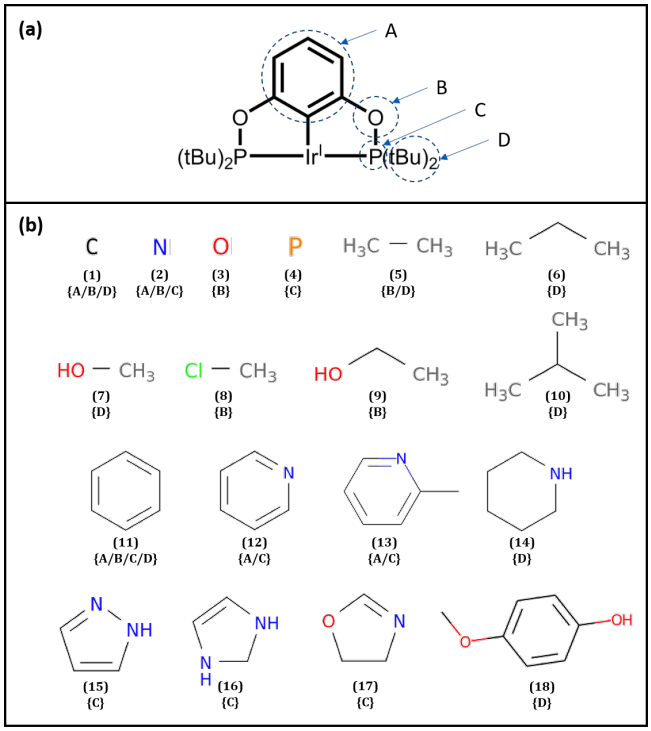

FIG. 2: (a) shows the four sites on the reference pincer catalyst that are considered for substitution in the library generation process. (b) shows the 18 building blocks that are provided as initial fragments to our library generator code, *ChemLG*, along with the sites on the reference catalyst where each building block can be added.

ChemLG utilizes the OpenBabel 3.0 software [65] to generate 3D structures for these catalysts using the UFF force field [66] and further optimizes the initial structures for 2000 steps. Additionally, we generate 3D structures for the adducts of every catalyst in this library with the six NEC structures (a total of 7 structures corresponding to each catalyst) using the same protocol for initial geometry optimization.

B. Quantum Chemistry Protocol

We model the six dehydrogenation steps of H12-NEC following this mechanism and perform *first-principles* quantum chemistry calculations within the Kohn-Sham density functional theory (DFT) framework.

For each of the pre-optimized 3D catalyst structures obtained from *ChemLG*, we further optimize the geometries using the B3LYP hybrid functional [67] and the LANL2DZ basis set [68–70] (including Grimme's D3 correction [71] to account for dispersion interactions) available from the ORCA 4.0.2 quantum chemistry package [72]. This is done in two consecutive optimization steps defined by the 'LOOSEOPT' and 'TIGHTOPT' convergence threshold parameters in ORCA along with the 'SlowConv' convergence strategy. We then obtain thermochemical properties *via* single point numerical frequency ('NUMFREQ') calculations for the final optimized geometries.

C. Hypergeometric Distribution Analysis

In addition to identifying the best catalyst molecules arising from the HTPS process, we also look for structural patterns in these molecules that result in both favorable as well as unfavorable reaction thermodynamics. In this case, a hypergeometric distribution analysis of the data is useful for uncovering relationships between intuitive individual descriptors and the target property, i.e., ΔG_r. For this, we compute the Z-scores (Z_i) of each building block (irrespective of the sites at which they are attached) given by:

$$Z_i = \frac{k_i - m\frac{K_i}{M}}{\sigma_i}$$

$$\sigma_i = \left[\frac{mK_i}{M} \times \left(\frac{M-K_i}{M}\right) \times \left(\frac{M-m}{M-1}\right)\right]^{\frac{1}{2}}$$

where M is the total number of compounds in the generated library, m is the subset of compounds being considered (e.g., the best catalysts discovered), K_i is the frequency of building block i in M molecules and k_i is its frequency in the subset of m molecules.

It follows that a large positive Z-score for a building block indicates its statistical over-expression in that subset of molecules relative to the overall screening library. Thus, Z-scores can be used to identify (and quantify the extent of correlation of) the most important building blocks for a catalyst with the least activation barriers for a reaction. Similarly, a large negative Z-score implies a statistical under-expression (or inverse correlation), while Z-scores close to zero imply no statistical correlation of that descriptor to the target property.

D. Machine Learning Protocol

In order to accelerate the screening of molecular catalysts, we take the help of machine learning models to make guiding predictions in the search for the most promising catalysts. We utilize several 2-D and 3-D descriptors to featurize the relevant chemical space for revealing hidden structure-property relationships within this data. To generate 2-D descriptors, we use the SMILES strings of the catalysts as the input, which is useful in capturing all information related to atom types and their connections to neighboring atoms. From the RDKit [73] code, we use the following fingerprints: Morgan fingerprints with circular radius 2 [74, 75], hashed atom pair (HAP) fingerprints [76], hashed topological torsion (HTT) fingerprints [77], and MACCS keys. Additionally, we compute the synthetic feasibility of the catalysts in our library using two scoring functions: SA-score [78] from RDKit and SC-score [79].

We generate 3-D descriptors for each of the three sets of optimized geometries obtained from *ChemLG* and DFT calculations mentioned in Sec. II B. These descriptors include:

- OpenBabel descriptors: FP2, FP3 and FP4 (including six extended-connectivity fingerprints – ECFP 1-6) path-based fingerprints that are similar to Daylight fingerprints. We also use Spectrophores [80, 81] that are created from four property fields surrounding the molecules – atomic partial charges, shape deviations, lipophilicities, and electrophilicities. Since Spectrophores encode the 3-D structures of the catalysts independently of their position and orientation, structures with similar 3-D shape and properties yield similar Spectrophores.

- Coulomb matrices [82]: We use all five Coulomb matrix representations available through *ChemML* that include eigenspectrum (Coul_Mat_1), random coulomb (Coul_Mat_2), sorted coulomb (Coul_Mat_3), unsorted coulomb (Coul_Mat_4), and unsorted triangular coulomb matrix (Coul_Mat_5).

- Bag of bonds [83]: These descriptors which are inspired from the bag-of-words descriptor often used in natural language processing applications are also generated *via ChemML*.

- Fractional buried volumes [84]: In a transition metal complex, the fraction of buried volume can be used to measure the steric hindrance on the metal center that is induced by a ligand.

For each of these descriptors, we train two Keras deep neural network (DNN) architectures using *ChemML* [85] – 1) for predicting the ΔG of individual reaction steps, and 2) for predicting the ΔG of all reaction steps simultaneously. For the first architecture, the input descriptors correspond to the catalyst-substrate adduct structures, whereas for the second architecture, they only take into account the catalyst structures.

For training the model, the input data is split into a 90:10 train:test ratio, and all descriptors other than fingerprints, i.e.non-binary descriptors, are scaled to unit variance. The model hyperparameters – regularization, activation function and the number and size of hidden layers, are optimized *via ChemML*'s genetic algorithm module, which optimizes the mean absolute error (MAE) of the validation set obtained from a five-fold cross-validation step in the training process [86–88].

III. RESULTS AND DISCUSSION

In Ref. [61], the dehydrogenation of H12-NEC has been reported without the use of any solvent. Thus, we initially setup our DFT calculations in a similar fashion. To account for the solvent effects of liquid phase H12-NEC, we use the CPCM model for solvation [89]. We use a dielectric constant of 1.3 and a refractive index of 1.6394 for H12-NEC. For a small subset of 200 catalyst molecules, we also run the same calculations using gas-phase H12-NEC, i.e.without the use of a solvent field, and compare the values of ΔG in both cases. We note that there is only a marginal difference between the two values, with the minimum and maximum values at -0.3 and +0.81 $kcal/mol$ respectively. A scatter plot of these calculations for the 200 catalysts is shown in Fig. 1 of the SI. Thus, for the remainder of the molecules in our library, we utilize the latter approach to calculate the properties of interest as these calculations have a lower computational overhead.

For the 3066 catalysts generated through our library generation process, the SMILES codes and the corresponding building blocks for the top 20 catalysts in terms of the lowest ΔG_r for all of the six dehydrogenation reactions is tabulated in Table 1 of the SI. The distribution of Gibbs free energies (blue) and enthaplies (red) in $kcal/mol$, along with the corresponding kernel density

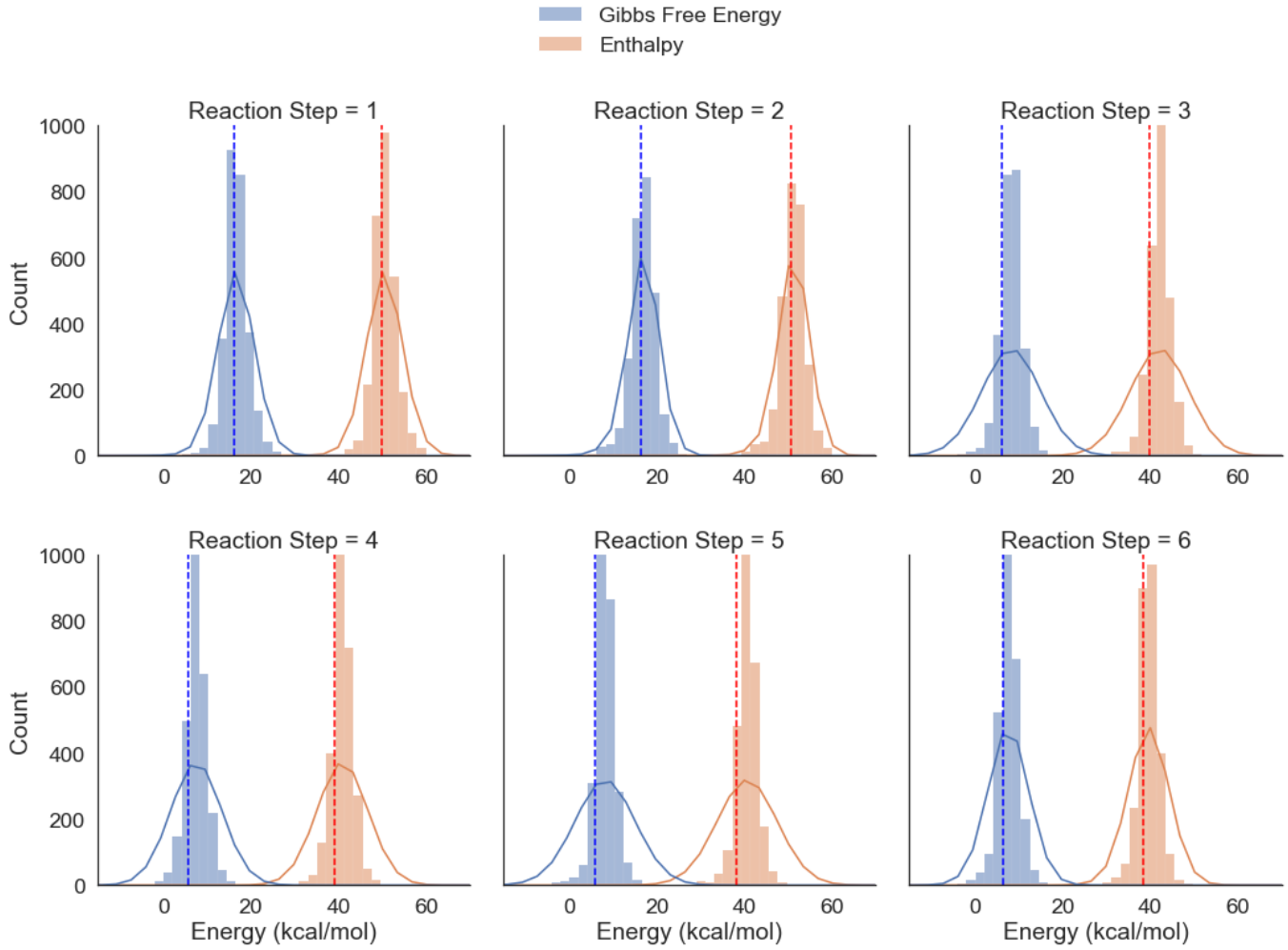

FIG. 3: A distribution of the Gibbs free energies (ΔG in kcal/mol) and enthalpies (ΔH in kcal/mol) of the 3066 catalyst molecules in our virtual library, for the six dehydrogenation reactions of perhydro-N-ethyl carbazole. The dotted blue and red lines indicate the mean ΔG and ΔH values respectively for each reaction step, whereas the line curves show the respective kernel density estimates.

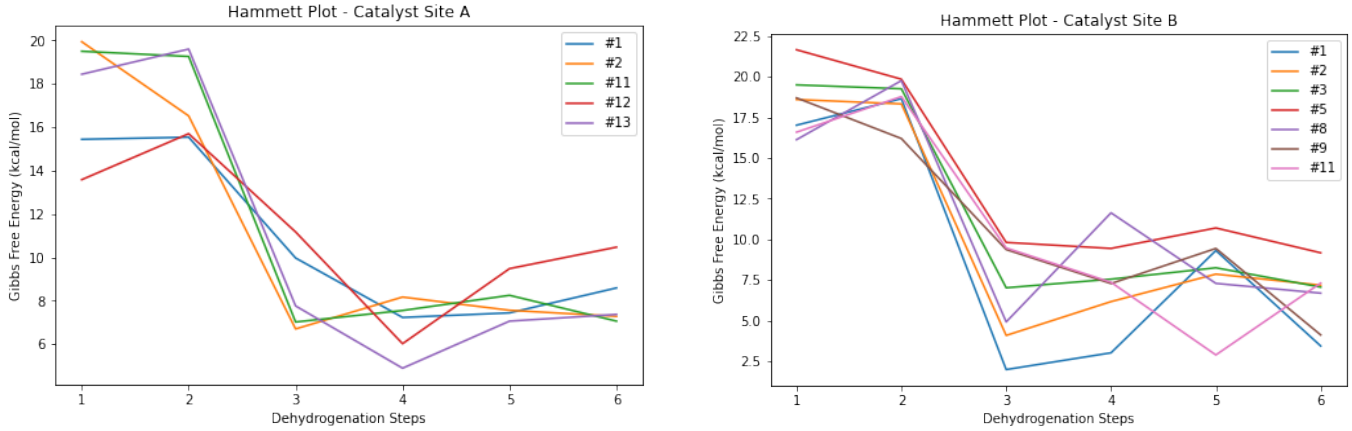

FIG. 4: Hammett plots for single site substitutions w.r.t. sites A and B on the reference catalyst. Each curve in the plot corresponds to the building block substituted at the respective site.

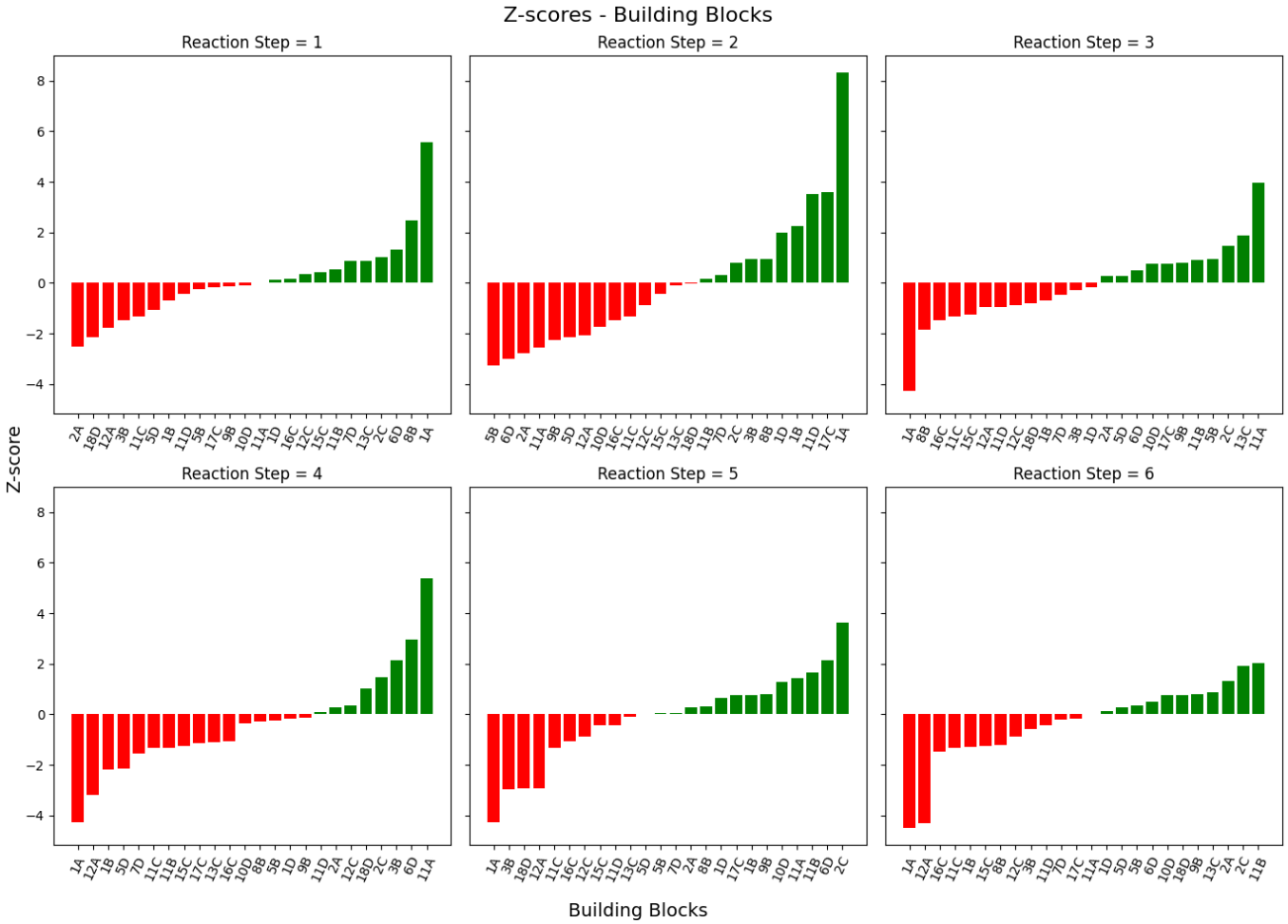

FIG. 5: Z-scores of the 18 building blocks. The x-axis labels are comprised of the building block number and the corresponding site where they are attached.

estimates (line curves) and mean values (dotted vertical lines) for both thermochemical properties, for the six dehydrogenation reactions of H12-NEC is shown in Fig. 3. For both ΔG and ΔH, we see a decreasing trend in the corresponding values as the reaction progresses towards complete dehydrogenation (overall shift of nearly 10-15 $kcal/mol$).

To analyze this data further, we first look at single building block substitutions at each of the four sites shown in Fig. 2(a), i.e., building blocks used at all the remaining sites on the catalyst are held constant. For sites A and B, Fig. 4 shows the change in free energy barriers (ΔG in kcal/mol) due to the different building blocks. Although different building blocks are dominant at different steps of the reaction, we observe a substantial fluctuation in the ΔG even from single substitutions. In particular, substitutions on site B show larger fluctuations in ΔG across all reaction steps as compared to those on site A. This implies a stronger electronic effect of site B on the metal center compared to site A. We make similar observations for sites C and D, where a large number of cases arise as we also account for non-symmetric substitutions on the pincer arm.

In addition to this, we also report the ΔG for substitutions made on the para-phenyl position of the reference catalyst in Table 2 of the SI. We select a set of 8 substituents for which the electronic effects span from electron withdrawing to electron donating effect. Compared to the reference catalyst, we do not see a significant shift in the ΔG values with the introduction of these groups into the catalyst. However, we note that electron donating groups are slightly more favorable for the dehydrogenation reactions compared to electron withdrawing groups.

For a site-agnostic analysis of the building blocks, we calculate the z-scores for the 18 building blocks as shown in Fig. 5. For this figure, we look at the top 5% of the catalysts (w.r.t. ΔG) for each of the six reactions, and evaluate the contribution (presence/absence in the catalyst molecule) of each building block in this subset *vs* their contribution in the remainder of the library. For catalyst sites A and B, we see a high positive as well as high negative correlation, whereas for sites C and D, the correlation goes from moderately positive to weakly neg-

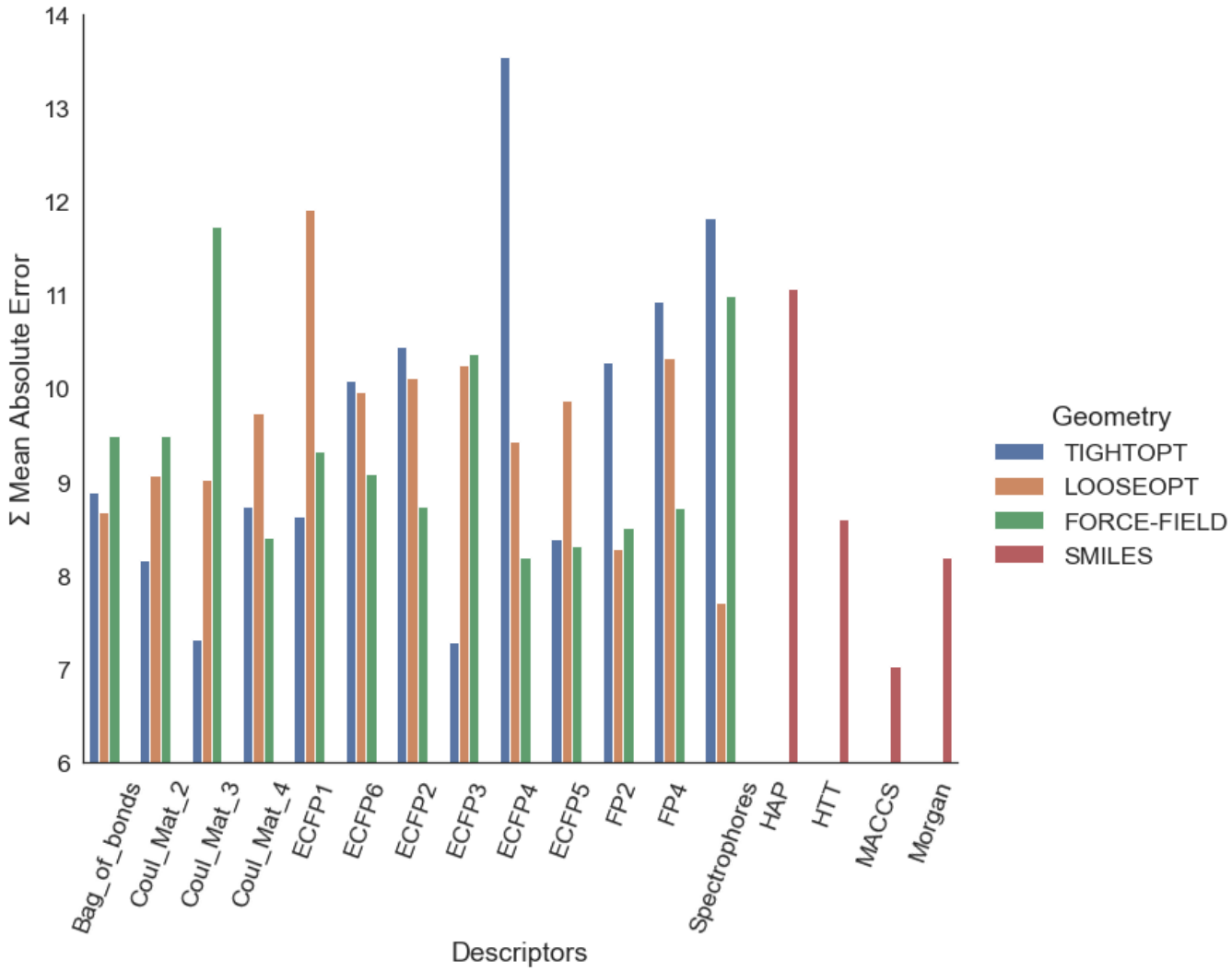

FIG. 6: A comparison of the performance of machine learning models trained using various 2-D and 3-D descriptors for predicting the ΔG_r of all six dehydrogenation steps simultaneously. The y-axis corresponds to the sum of the mean absolute errors (test set) of the six predictions for each catalyst molecule. The blue, orange, green, and red lines for each descriptor refer to the four types of optimized geometries used for generating the 3-D descriptors.

ative. This implies a stronger effect of substitutions on sites A and B as compared to sites C and D. Furthermore, we see a flipping of the building blocks from the red zone to the green zone, and *vice versa*, after the first two reaction steps. With this result, we infer two things: 1) we cannot select a single set of the strongest correlated building blocks to construct a catalyst that will yield superior performance across all reaction steps, and 2) the selection of the best building blocks for designing a catalyst should instead be based on the rate determining step(s) in the reaction mechanism.

We also plot the z-scores for the cyclometallated ring sizes of the catalysts in Fig. 2 of the SI. These refer to the number of atoms in each of the two cyclic rings that are formed by the metal center and the ligand atoms. For instance, in Fig. 2(a), the two rings are comprised of $Ir-C-C-O-P$ atoms, and have an identical size of 5-5. Across all pincer complexes reported in literature, symmetric ring sizes for pincer complexes largely outnumber non-symmetric ones. However, the plot indicates that non-symmetric ring sizes are also strongly correlated to ΔG. Similar to the z-scores for building blocks, we see a flipping of the cyclometallated ring sizes from the red zone to the green zone, and *vice versa*, as we move further through the reaction steps.

The generation of thermochemical data for the 3066 catalysts in our library following the computational protocol described in Sec. II B is a computationally demanding procedure. For each catalyst molecule, it took almost an average of 3 days (approx. few minutes for force-field optimized geometries, 8-10 hours for LOOSEOPT optimization, and 50-70 hours for TIGHTOPT optimization) for all the calculations to complete on a 40-core node, resulting in a total of a little over 2 years for gen-

erating the entire data. Thus, in order to accelerate the screening/hyper-screening of additional libraries of pincer catalysts, we train supervised ML models for predicting the thermochemical data.

Of the two model architectures described in Sec. II D, models trained using the second architecture uniformly outperformed those trained using the first architecture for all descriptors. For the second model architecture, Fig. 6 shows the performance of ML models generated from various 2-D (using SMILES strings) and 3-D descriptors (using each of the three optimized geometries – 'TIGHTOPT', 'LOOSEOPT', and 'FORCE-FIELD'). The y-axis of the plot corresponds to the sum of the MAEs of the six predictions for each molecule of the test set.

The best models corresponding to the 'TIGHTOPT', 'LOOSEOPT', 'FORCE-FIELD', and 'SMILES' descriptors arise from the sorted coulomb matrix, Spectrophores, ECFP4 and the MACCS keys respectively. Within 3-D descriptors, the most accurate predictions arise from the models trained on 'TIGHTOPT' geometries. When considering all three geometries simultaneously, the best models are those corresponding to the random coulomb matrix. Furthermore, for the bag of bonds and random coulomb matrix descriptors, we see very close predictions between the 3 geometries.

For most descriptors, we don't see a significant difference in the model predictions w.r.t. the three geometries, thus, models trained on less accurate geometries can potentially be used to make predictions on a large dataset. Since the creation of descriptors using 'FORCE-FIELD' and 'LOOSEOPT' geometries takes substantially lower time than that for 'TIGHTOPT' geometries, the primary bottleneck, i.e. the computational overhead in the HTPS protocol can be alleviated.

IV. CONCLUSIONS

In this study, we employ a data-driven *in silico* approach to investigate liquid organic hydrogen carriers that provide an excellent media for the storage and transportation of pure hydrogen. We conduct a high-throughput screening study to identify lead molecular catalysts for extracting hydrogen pairs from N-ethyl carbazole. Our methodology serves as a rapid and efficient scheme to assess the thermochemical properties of catalysts, as well as to reveal structure-property relationships that are crucial to establish rational design principles.

We identify the most promising catalysts, as well as the top building blocks (both site-specific and site-agnostic) and structural patterns that contribute favorably to the reaction thermodynamics. To accelerate the high-throughput screening, we also train machine learning models that accurately model the Gibbs free energies of these reactions. Even with models trained on partially optimized geometries, we are able to achieve similar model performance, thus significantly reducing the time required for *first-principles* computational studies. We aim to utilize these valuable insights into these systems to pursue promising candidates *via* future experiments.

SUPPLEMENTARY MATERIAL

Electronic supplementary material accompanies this paper. It provides details of all computational and experimental values displayed in the figures throughout this paper or that were used in the statistical analysis.

COMPETING FINANCIAL INTERESTS

The authors declare to have no competing financial interests.

ACKNOWLEDGMENTS

This work was supported by start-up funds provided through the University at Buffalo (UB), the National Science Foundation (NSF) CAREER program (grant No. OAC-1751161), and the New York State Center of Excellence in Materials Informatics (grants No. CMI-1122381, CMI-1140384, CMI-1148092), the National Science Foundation CBET Energy for Sustainability (grant No. CBET-1511528), and the Army Research Office SBIR program (W15QKN-17-C-0078).

Computing time on the high-performance computing clusters '*Alpha*', '*Beta*', and '*Vortex*' was provided by the UB Center for Computational Research (CCR).

[1] Li Zhou, "Progress and problems in hydrogen storage methods," Renewable and Sustainable Energy Reviews **9**, 395–408 (2005).

[2] Sunita Satyapal, John Petrovic, Carole Read, George Thomas, and Grace Ordaz, "The us department of energy's national hydrogen storage project: Progress towards meeting hydrogen-powered vehicle requirements," Catalysis today **120**, 246–256 (2007).

[3] Ping Chen and Min Zhu, "Recent progress in hydrogen storage," Materials today **11**, 36–43 (2008).

[4] Karsten Müller, Johannes Völkl, and Wolfgang Arlt, "Thermodynamic evaluation of potential organic hydrogen carriers," Energy Technology **1**, 20–24 (2013).

[5] Patrick Preuster, Christian Papp, and Peter Wasserscheid, "Liquid organic hydrogen carriers (lohcs): toward a hydrogen-free hydrogen economy," Accounts of chemical research **50**, 74–85 (2017).

[6] M Markiewicz, YQ Zhang, A Bösmann, N Brückner, J Thöming, P Wasserscheid, and Stefan Stolte, "Environmental and health impact assessment of liquid organic hydrogen carrier (lohc) systems–challenges and preliminary results," Energy & Environmental Science **8**, 1035–1045 (2015).

[7] Alan C Cooper, Karen M Campbell, and Guido P Pez, "An integrated hydrogen storage and delivery approach using organic liquid-phase carriers," in *Proc. 16th World Hydrogen Energy Conference, Lyon, France* (2006).

[8] Abdennour Bourane, Mohamed Elanany, Thang V Pham, and Sai P Katikaneni, "An overview of organic liquid phase hydrogen carriers," International journal of hydrogen energy **41**, 23075–23091 (2016).

[9] Karsten Müller, Rabya Aslam, Armin Fischer, Katharina Stark, Peter Wasserscheid, and Wolfgang Arlt, "Experimental assessment of the degree of hydrogen loading for the dibenzyl toluene based lohc system," international journal of hydrogen energy **41**, 22097–22103 (2016).

[10] J Von Wild, T Friedrich, A Cooper, B Toseland, G Muraro, Ward TeGrotenhuis, Yong Wang, Paul Humble, and Ayman Karim, "Liquid organic hydrogen carriers (lohc): An auspicious alternative to conventional hydrogen storage technologies," in *18th World Hydrogen Energy Conference, Essen, Germany* (2010).

[11] Robert H Crabtree, "Hydrogen storage in liquid organic heterocycles," Energy & Environmental Science **1**, 134–138 (2008).

[12] Farnaz Sotoodeh, Benjamin JM Huber, and Kevin J Smith, "The effect of the n atom on the dehydrogenation of heterocycles used for hydrogen storage," Applied Catalysis A: General **419**, 67–72 (2012).

[13] Darrell Dean, Boyd Davis, and Philip G Jessop, "The effect of temperature, catalyst and sterics on the rate of n-heterocycle dehydrogenation for hydrogen storage," New Journal of Chemistry **35**, 417–422 (2011).

[14] Robert H. Crabtree, "Nitrogen-Containing Liquid Organic Hydrogen Carriers: Progress and Prospects," ACS Sustainable Chemistry and Engineering **5**, 4491–4498 (2017).

[15] T He, Q J Pei, and P Chen, "Liquid organic hydrogen carriers," Journal of Energy Chemistry **24**, 587–594 (2015).

[16] Guido P Pez, Aaron R Scott, Alan C Cooper, and Hansong Cheng, "Hydrogen storage by reversible hydrogenation of pi-conjugated substrates," (2006), uS Patent 7,101,530.

[17] A. Mehranfar, M. Izadyar, and A. A. Esmaeili, "Hydrogen storage by n-ethylcarbazol as a new liquid organic hydrogen carrier: A dft study on the mechanism," International Journal of Hydrogen Energy **40**, 5797–5806 (2015).

[18] Farnaz Sotoodeh and Kevin J. Smith, "Structure sensitivity of dodecahydro-n-ethylcarbazole dehydrogenation over pd catalysts," Journal of Catalysis **279**, 36–47 (2011).

[19] A. Mehranfar and M. Izadyar, "N-ethylcarbazole-doped fullerene as a potential candidate for hydrogen storage, a kinetics approach," RSC Advances **5**, 49159–49167 (2015).

[20] L. M. Kustov, A. L. Tarasov, and O. A. Kirichenko, "Microwave-activated dehydrogenation of perhydro-n-ethylcarbazol over bimetallic pd-m/tio2 catalysts as the second stage of hydrogen storage in liquid substrates," International Journal of Hydrogen Energy **42**, 26723–26729 (2017).

[21] Zhao Jiang, Xiang Gong, Bin Wang, Zhiqiang Wu, and Tao Fang, "A experimental study on the dehydrogenation performance of dodecahydro-n-ethylcarbazole on m/tio2 catalysts," International Journal of Hydrogen Energy **44**, 2951–2959 (2019).

[22] Bin Wang, Tie yan Chang, Zhao Jiang, Jin jia Wei, Yong hai Zhang, Sen Yang, and Tao Fang, "Catalytic dehydrogenation study of dodecahydro-n-ethylcarbazole by noble metal supported on reduced graphene oxide," International Journal of Hydrogen Energy **43**, 7317–7325 (2018).

[23] Ming Yang, Yuan Dong, Shunxin Fei, Hanzhong Ke, and Hansong Cheng, "A comparative study of catalytic dehydrogenation of perhydro-n-ethylcarbazole over noble metal catalysts," International Journal of Hydrogen Energy **39**, 18976–18983 (2014).

[24] Hongen Yu, Xue Yang, Yong Wu, Yanru Guo, Shuan Li, Wei Lin, Xingguo Li, and Jie Zheng, "Bimetallic runi/tio2 catalysts for hydrogenation of n-ethylcarbazole: Role of tio2 crystal structure," Journal of Energy Chemistry **40**, 188–195 (2020).

[25] Farnaz Sotoodeh, Liang Zhao, and Kevin J Smith, "Kinetics of h2 recovery from dodecahydro-n-ethylcarbazole over a supported pd catalyst," Applied Catalysis A: General **362**, 155–162 (2009).

[26] Ming Yang, Yuan Dong, Shunxin Fei, Hanzhong Ke, and Hansong Cheng, "A comparative study of catalytic dehydrogenation of perhydro-n-ethylcarbazole over noble metal catalysts," International journal of hydrogen energy **39**, 18976–18983 (2014).

[27] Päivi T Aakko-Saksa, Chris Cook, Jari Kiviaho, and Timo Repo, "Liquid organic hydrogen carriers for transportation and storing of renewable energy–review and discussion," Journal of Power Sources **396**, 803–823 (2018).

[28] Max Amende, Stefan Schernich, Marek Sobota, Ioannis Nikiforidis, Wolfgang Hieringer, Daniel Assenbaum, Christoph Gleichweit, Hans-Jörg Drescher, Christian Papp, Hans-Peter Steinrück, *et al.*, "Dehydrogenation mechanism of liquid organic hydrogen carriers: dodecahydro-n-ethylcarbazole on pd (111)," Chemistry–A European Journal **19**, 10854–10865 (2013).

[29] Ming Yang, Chaoqun Han, Gang Ni, Jinping Wu, and Hansong Cheng, "Temperature controlled three-stage catalytic dehydrogenation and cycle performance of perhydro-9-ethylcarbazole," International journal of hydrogen energy **37**, 12839–12845 (2012).

[30] Audrey Moores, Macarena Poyatos, Yi Luo, and Robert H Crabtree, "Catalysed low temperature h 2 release from nitrogen heterocycles," New Journal of Chemistry **30**, 1675–1678 (2006).

[31] Vladimir Sobolev, Igor Asanov, and Konstantin Koltunov, "The role of support in formic acid decomposition on gold catalysts," Energies **12** (2019), 10.3390/en12214198.

[32] Manuel Iglesias and Luis A. Oro, "Mechanistic considerations on homogeneously catalyzed formic acid dehydrogenation," European Journal of Inorganic Chemistry **2018**, 2125–2138 (2018).

[33] Fedor S. Golub, Sergey Beloshapkin, Artem V. Gusel'Nikov, Vasily A. Bolotov, Valentin N. Parmon, and Dmitri A. Bulushev, "Boosting hydrogen production

from formic acid over pd catalysts by deposition of n-containing precursors on the carbon support," Energies **12** (2019), 10.3390/en12203885.

[34] G Do, P Preuster, R Aslam, A Bösmann, K Müller, W Arlt, and P Wasserscheid, "Hydrogenation of the liquid organic hydrogen carrier compound dibenzyltoluene–reaction pathway determination by 1 h nmr spectroscopy," Reaction chemistry & engineering **1**, 313–320 (2016).

[35] Katarzyna Morawa Eblagon, Daniel Rentsch, Oliver Friedrichs, Arndt Remhof, Andreas Zuettel, AJ Ramirez-Cuesta, and Shik Chi Tsang, "Hydrogenation of 9-ethylcarbazole as a prototype of a liquid hydrogen carrier," International journal of hydrogen energy **35**, 11609–11621 (2010).

[36] A Bulgarin, H Jorschick, P Preuster, A Bösmann, and P Wasserscheid, "Purity of hydrogen released from the liquid organic hydrogen carrier compound perhydro dibenzyltoluene by catalytic dehydrogenation," International Journal of Hydrogen Energy **45**, 712–720 (2020).

[37] Cecil N.M. Ouma, Phillimon M. Modisha, and Dmitri Bessarabov, "Catalytic dehydrogenation onset of liquid organic hydrogen carrier, perhydrodibenzyltoluene: The effect of pd and pt subsurface configurations," Computational Materials Science **172** (2020), 10.1016/j.commatsci.2019.109332.

[38] Max Amende, Christoph Gleichweit, Kristin Werner, Stefan Schernich, Wei Zhao, Michael P A Lorenz, Oliver Höfert, Christian Papp, Marcus Koch, and Peter Wasserscheid, "Model catalytic studies of liquid organic hydrogen carriers: Dehydrogenation and decomposition mechanisms of dodecahydro-N-ethylcarbazole on Pt (111)," ACS catalysis **4**, 657–665 (2014).

[39] M Sobota, I Nikiforidis, M Amende, B S Zanon, T Staudt, O Hofert, Y Lykhach, C Papp, W Hieringer, M Laurin, D Assenbaum, P Wasserscheid, H P Steinruck, A Gorling, and J Libuda, "Dehydrogenation of Dodecahydro-N-ethylcarbazole on Pd/Al2O3 Model Catalysts," Chemistry-a European Journal **17**, 11542–11552 (2011).

[40] M Amende, S Schernich, M Sobota, I Nikiforidis, W Hieringer, D Assenbaum, C Gleichweit, H J Drescher, C Papp, H P Steinruck, A Gorling, P Wasserscheid, M Laurin, and J Libuda, "Dehydrogenation Mechanism of Liquid Organic Hydrogen Carriers: Dodecahydro-N-ethylcarbazole on Pd(111)," Chemistry-a European Journal **19**, 10854–10865 (2013).

[41] F Sotoodeh and K J Smith, "An overview of the kinetics and catalysis of hydrogen storage on organic liquids," Canadian Journal of Chemical Engineering **91**, 1477–1490 (2013).

[42] K M Eblagon and S C E Tsang, "Structure-reactivity relationship in catalytic hydrogenation of heterocyclic compounds over ruthenium black; Part B: Effect of carbon substitution by heteroatom," Applied Catalysis B-Environmental **163**, 599–610 (2015).

[43] Audrey Moores, Macarena Poyatos, Yi Luo, and Robert H Crabtree, "Catalysed low temperature H2 release from nitrogen heterocycles," New Journal of Chemistry **30**, 1675–1678 (2006).

[44] Paul Crawford, Robbie Burch, Christopher Hardacre, K T Hindle, Peijun Hu, B Kalirai, and D W Rooney, "Understanding the dehydrogenation mechanism of tetrahydrocarbazole over palladium using a combined experimental and density functional theory approach," The Journal of Physical Chemistry C **111**, 6434–6439 (2007).

[45] Farnaz Sotoodeh, Benjamin J M Huber, and Kevin J Smith, "Dehydrogenation kinetics and catalysis of organic heteroaromatics for hydrogen storage," International Journal of Hydrogen Energy , 2715–2722 (2012).

[46] Chao Guan, Yupeng Pan, Tonghuan Zhang, Manjaly J. Ajitha, and Kuo Wei Huang, "An update on formic acid dehydrogenation by homogeneous catalysis," Chemistry - An Asian Journal **15**, 937–946 (2020).

[47] Ryohei Yamaguchi, Chikako Ikeda, Yoshinori Takahashi, and Ken-ichi Fujita, "Homogeneous catalytic system for reversible dehydrogenation- hydrogenation reactions of nitrogen heterocycles with reversible interconversion of catalytic species," Journal of the American Chemical Society **131**, 8410–8412 (2009).

[48] Zachary X Giustra, Jacob SA Ishibashi, and Shih-Yuan Liu, "Homogeneous metal catalysis for conversion between aromatic and saturated compounds," Coordination Chemistry Reviews **314**, 134–181 (2016).

[49] Eric D Cline, Samantha E Adamson, and Stefan Bernhard, "Homogeneous catalytic system for photoinduced hydrogen production utilizing iridium and rhodium complexes," Inorganic Chemistry **47**, 10378–10388 (2008).

[50] Monica Trincado, Jonas Bösken, and Hansjörg Grützmacher, "Homogeneously catalyzed acceptorless dehydrogenation of alcohols: A progress report," Coordination Chemistry Reviews **443**, 213967 (2021).

[51] Gábor Laurenczy and Paul J Dyson, "Homogeneous catalytic dehydrogenation of formic acid: progress towards a hydrogen-based economy," Journal of the Brazilian Chemical Society **25**, 2157–2163 (2014).

[52] Matthias W Haenel, Stephan Oevers, Klaus Angermund, William C Kaska, Hua-Jun Fan, and Michael B Hall, "Thermally stable homogeneous catalysts for alkane dehydrogenation," Angewandte Chemie International Edition **40**, 3596–3600 (2001).

[53] Matthew E Sloan, Timothy J Clark, and Ian Manners, "Homogeneous catalytic dehydrogenation/dehydrocoupling of amine-borane adducts by the rh (i) wilkinson's complex analogue rhcl (phcy2) 3 (cy= cyclohexyl)," Inorganic chemistry **48**, 2429–2435 (2009).

[54] Manuel Iglesias and Francisco J Fernández-Alvarez, "Advances in nonprecious metal homogeneously catalyzed formic acid dehydrogenation," Catalysts **11**, 1288 (2021).

[55] Peter J Bonitatibus Jr, Sumit Chakraborty, Mark D Doherty, Oltea Siclovan, William D Jones, and Grigorii L Soloveichik, "Reversible catalytic dehydrogenation of alcohols for energy storage," Proceedings of the National Academy of Sciences **112**, 1687–1692 (2015).

[56] Zhuofeng Ke, Yinwu Li, Cheng Hou, and Yan Liu, "Homogeneously catalyzed hydrogenation and dehydrogenation reactions–from a mechanistic point of view," Physical Sciences Reviews **3** (2018).

[57] Martin Nielsen, "Hydrogen production by homogeneous catalysis: alcohol acceptorless dehydrogenation," in *Hydrogen production and remediation of carbon and pollutants* (Springer, 2015) pp. 1–60.

[58] Xiang-Biao Zhang and Zhao Xi, "A theoretical study of the mechanism for the homogeneous catalytic reversible dehydrogenation—hydrogenation of nitrogen heterocycles," Physical Chemistry Chemical Physics **13**, 3997–

4004 (2011).
[59] Yujie Wang, Mingyang Wang, Yibiao Li, and Qiang Liu, "Homogeneous manganese-catalyzed hydrogenation and dehydrogenation reactions," Chem **7**, 1180–1223 (2021).
[60] Arup Mukherjee and David Milstein, "Homogeneous catalysis by cobalt and manganese pincer complexes," ACS Catalysis **8**, 11435–11469 (2018).
[61] Zhaohui Wang, Ian Tonks, Jack Belli, and Craig M Jensen, "Dehydrogenation of n-ethyl perhydrocarbazole catalyzed by pcp pincer iridium complexes: evaluation of a homogenous hydrogen storage system," Journal of Organometallic Chemistry **694**, 2854–2857 (2009).
[62] Johannes Hachmann, Mohammad Atif Faiz Afzal, Mojtaba Haghighatlari, and Yudhajit Pal, "Building and deploying a cyberinfrastructure for the data-driven design of chemical systems and the exploration of chemical space," Molecular Simulation **44**, 921–929 (2018).
[63] CraigáM Jensen, "Iridium pcp pincer complexes: highly active and robust catalysts for novel homogeneous aliphatic dehydrogenations," Chemical communications , 2443–2449 (1999).
[64] Mohammad Atif Faiz Afzal, Gaurav Vishwakarma, Janhavi Abhay Dudwadkar, Mojtaba Haghighatlari, and Johannes Hachmann, "Chemlg–a program suite for the generation of compound libraries and the survey of chemical space," (2019).
[65] Noel M O'Boyle, Michael Banck, Craig A James, Chris Morley, Tim Vandermeersch, and Geoffrey R Hutchison, "Open babel: An open chemical toolbox," Journal of cheminformatics **3**, 1–14 (2011).
[66] Anthony K Rappé, Carla J Casewit, KS Colwell, William A Goddard III, and W Mason Skiff, "Uff, a full periodic table force field for molecular mechanics and molecular dynamics simulations," Journal of the American chemical society **114**, 10024–10035 (1992).
[67] Axel D Becke, "Density-functional thermochemistry. i. the effect of the exchange-only gradient correction," The Journal of chemical physics **96**, 2155–2160 (1992).
[68] P Jeffrey Hay and Willard R Wadt, "Ab initio effective core potentials for molecular calculations. potentials for k to au including the outermost core orbitals," The Journal of chemical physics **82**, 299–310 (1985).
[69] Willard R Wadt and P Jeffrey Hay, "Ab initio effective core potentials for molecular calculations. potentials for main group elements na to bi," The Journal of chemical physics **82**, 284–298 (1985).
[70] P Jeffrey Hay and Willard R Wadt, "Ab initio effective core potentials for molecular calculations. potentials for the transition metal atoms sc to hg," The Journal of chemical physics **82**, 270–283 (1985).
[71] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu," The Journal of chemical physics **132**, 154104 (2010).
[72] Frank Neese, "Software update: the orca program system, version 4.0," Wiley Interdisciplinary Reviews: Computational Molecular Science **8**, e1327 (2018).
[73] Gregory Landrum, "Rdkit: A software suite for cheminformatics, computational chemistry, and predictive modeling," Release **1**, 4 (2013).
[74] H L Morgan, "The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service." Journal of Chemical Documentation **5**, 107–113 (1965).
[75] David Rogers and Mathew Hahn, "Extended-Connectivity Fingerprints," Journal of Chemical Information and Modeling **50**, 742–754 (2010).
[76] Raymond E Carhart, Dennis H Smith, and R Venkataraghavan, "Atom pairs as molecular features in structure-activity studies: definition and applications," Journal of Chemical Information and Computer Sciences **25**, 64–73 (1985).
[77] Ramaswamy Nilakantan, Norman Bauman, J Scott Dixon, and R Venkataraghavan, "Topological torsion: a new molecular descriptor for SAR applications. Comparison with other descriptors," Journal of Chemical Information and Computer Sciences **27**, 82–85 (1987).
[78] Peter Ertl and Ansgar Schuffenhauer, "Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions," Journal of cheminformatics **1**, 1–11 (2009).
[79] Connor W Coley, Luke Rogers, William H Green, and Klavs F Jensen, "Scscore: synthetic complexity learned from a reaction corpus," Journal of chemical information and modeling **58**, 252–261 (2018).
[80] Patrick Bultinck, Wilfried Langenaeker, Philippe Lahorte, Frank De Proft, Paul Geerlings, Christian Van Alsenoy, and JP Tollenaere, "The electronegativity equalization method ii: applicability of different atomic charge schemes," The Journal of Physical Chemistry A **106**, 7895–7901 (2002).
[81] Patrick Bultinck, Wilfried Langenaeker, Ramon Carbó-Dorca, and Jan P Tollenaere, "Fast calculation of quantum chemical molecular descriptors from the electronegativity equalization method," Journal of chemical information and computer sciences **43**, 422–428 (2003).
[82] Matthias Rupp, Alexandre Tkatchenko, Klaus-Robert Müller, and O Anatole Von Lilienfeld, "Fast and accurate modeling of molecular atomization energies with machine learning," Physical review letters **108**, 058301 (2012).
[83] Katja Hansen, Franziska Biegler, Raghunathan Ramakrishnan, Wiktor Pronobis, O Anatole Von Lilienfeld, Klaus-Robert Müller, and Alexandre Tkatchenko, "Machine learning predictions of molecular properties: Accurate many-body potentials and nonlocality in chemical space," The journal of physical chemistry letters **6**, 2326–2331 (2015).
[84] Laura Falivene, Raffaele Credendino, Albert Poater, Andrea Petta, Luigi Serra, Romina Oliva, Vittorio Scarano, and Luigi Cavallo, "Sambvca 2. a web tool for analyzing catalytic pockets with topographic steric maps," Organometallics **35**, 2286–2293 (2016).
[85] Mojtaba Haghighatlari, Gaurav Vishwakarma, Doaa Altarawy, Ramachandran Subramanian, Bhargava U Kota, Aditya Sonpal, Srirangaraj Setlur, and Johannes Hachmann, "ChemML: A machine learning and informatics program package for the analysis, mining, and modeling of chemical and materials data," Wiley Interdisciplinary Reviews: Computational Molecular Science **10**, e1458 (2020).
[86] Gaurav Vishwakarma, *Machine Learning Model Selection for Predicting Properties of High-Refractive-Index Polymers*, Master's thesis, University at Buffalo (2018).
[87] Gaurav Vishwakarma, Mojtaba Haghighatlari, and Johannes Hachmann, "Towards autonomous machine learning in chemistry via evolutionary algorithms," ChemRxiv

, 9782387 (2019).

[88] Gaurav Vishwakarma, Aditya Sonpal, and Johannes Hachmann, "Metrics for benchmarking and uncertainty quantification: Quality, applicability, and best practices for machine learning in chemistry," Trends in Chemistry (2021).

[89] Vincenzo Barone and Maurizio Cossi, "Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model," The Journal of Physical Chemistry A **102**, 1995–2001 (1998).