Utrecht Journal of Chemical Research

Analysis of DFT calculated reaction energies for CO bonding and dissociation on metal alloy surfaces.

Menno R.A. Heijnen*[ab], August N.A. Ammerlaan[ab], Koen C. Pons[ab], Ilker Tezsevi[b], and Süleyman Er+[b].

a) Utrecht University, Padualaan 8, 3584 CH Utrecht, the Netherlands, Department of Chemistry.
b) Autonomous Energy Materials Discovery [AMD] Research Group - DIFFER - Dutch Institute for Fundamental Energy Research & CCER – Center for Computational Energy Research, De Zaale 20, 5612 AJ Eindhoven, the Netherlands.

*Corresponding author, e-mail: m.r.a.heijnen@students.uu.nl, student number: 5610397
+Corresponding supervisor, e-mail: s.er@differ.nl

Keywords: Carbon monoxide, Heterogenous Catalysis, Data Curation, Density Functional Theory, and Surface reactions; Number of words indication: 1479

ABSTRACT: The Catalysis-Hub.org web-platform database that has been generated by the Center for Interface Science and Catalysis contain a vast amount of information about CO surface reactions on catalysts. Currently the database is not well organized to be useful for machine learning. We report a logical analysis of the database focused on CO bonding and dissociation in a tree chart. With a data analysis we discovered a correlation between increasing group numbers of atoms in the chemical compositions and an increasing reaction energy. The processed data can be used after more qualitative research by machine learning surveys.

Introduction

We live in an era where we are at the start of a major change from fossil fuels to renewable resources. One of the reasons is the exhaustible nature of fossil fuels [1]. For this we must look for alternatives to energy sources. One of these alternatives is the production of hydrocarbons from carbon monoxide (CO) using catalysts and is known as the Fischer-Tropsch synthesis [2].

$$n\mathrm{CO} + 2n\mathrm{H_2} \rightarrow (-\mathrm{CH_2})_n + n\mathrm{H_2O} \qquad (1)$$

The Fischer-Tropsch synthesis mainly produces *n*-alkanes and 1-alkenes according to equation 1 with a catalyst [3]. Figure 1 explains the reaction visually [4].
An alternative catalyst can be cheaper or more effective, so it is valuable to gain more insight into this. Fortunately, a lot of research has been done into the bonding and dissociation of CO on a catalyst and is brought together in the Catalysis-Hub.org web-platform database [5]. This is a web-platform database designed by SUNCAT's Center for Interface Science and Catalysis. However, the interconnected information in database need to be organized in a way to use it for further studies, such as machine learning. This research aims to extract the vast amount of information in the catalysis-hub.org database and sorting so that it can be used for machine learning. A side aim of this research is to make a data analysis to see if there are any trends visible.

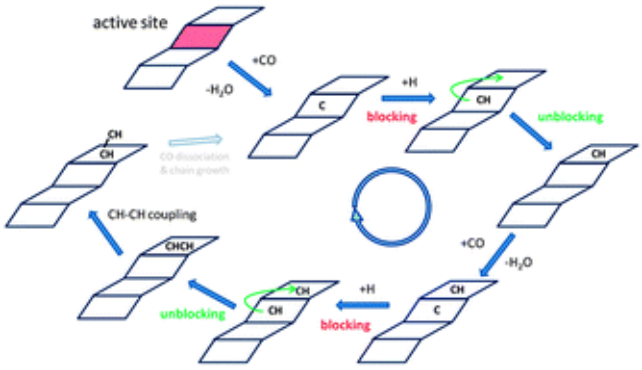

Figure 1 The Fischer-Tropsch synthesis of hydrocarbons on a corrugated Ru surface according to Shetty et al. [4]. Blocking and unblocking refers to the accessibility of the active site.

Experimental

Data source

As a data source we downloaded information [6-13] from Catalysis-Hub.org on the 20nd of May 2019. The Catalysis-Hub.org database structure is built of several interconnected tables. The most important being the 'Reaction', 'Publication', and 'Systems' table. Each giving information about the reaction, publication, and system respectively. For this article only information from these tables are used. At http://docs.catalysis-hub.org/ more details can be found.
Calculations in these researches were made with the Density Functional Theory (DFT) and is widely used to calculate the reaction energies of catalysis reactions [14]. DFT uses in principle exact electron density distribution *n(r)* and is

preferable to the intrinsically approximate electron wave function theories $\Psi(r_1,\ldots,r_N)$ with higher accuracy. [15]
The database allows to search on the following chemical properties: reactants, products, surface of catalyst, and/or 'facet'. However, the equations, the DFT code and DFT functional which were used to make the calculations, and the chemical composition of the catalyst are needed to compare the vast amount of data. The main used equations were CO binding and dissociation as described by equations 2 and 3. The * is the binding energy.

$$CO(g) + * \rightarrow CO* \quad (2)$$

$$CO(g) + 2* \rightarrow C* + O* \quad (3)$$

The main used DFT codes were DACAPO and Quantum Espresso and both uses pseudopotentials and a plane wave basis set [16]. The main used DFT functionals were RPBE and BEEF-vdW, which were the functionals with the lowest error for CO reaction energies (9-11 kJ/mol) [17]. The chemical composition is the arrangement, type, and ratio of atoms in the catalysts. We have generalized the chemical composition into classes in which the metal was simply replaced by M (Metal). I.e. Zr3Au falls in class M3M1 and Au36 in M36

Raw datasheet
On Catalysis-Hub.org we searched in the database of reaction energies and barriers (Surface Reactions) for 'CO' reactants. Then the GRAPHQL QUERY was opened to see the JSON code of the database. A query was entered containing all the possible variables from the 'systems' and 'reactions' tables. From the JSON format 2091 data instances were converted online (http://convertcsv.com/json-to-csv.htm) into an excel sheet.

Processing of raw datasheet
The data processing was done in 5 steps:

1. The data containing equations with only 'CO(g)' as a reactant was preserved, and the other equations were removed.
   Remaining data instances: 1994
2. The data containing more than 50 data instances in the combination of DFT code and DFT functional were preserved and the other equations were removed.
   Remaining data instances: 1965
3. The data containing facets with less than 40 data instances were discarded (except for the '211' facets).
   Remaining data instances: 1928
4. The chemical compositions were divided in designated groups. This was done by replacing the atoms with an M (from Metal) and then putting the same names in a group. (e.g. Ag3Pd9 became M3M9 and therefore ended up in group M3M9)
5. The data was sorted on chemical composition and how the CO*, C* or O* are bound to the catalyst (i.e. binding sites).

Curated data is available on request from the corresponding author.

Tree chart
The tree chart was made according to a logical sequence, starting with a square (bucket) of the total number of data instances. The choice was made to then separate the data in the following order: (1) Equation, (2) DFT code, (3) DFT functional, (4) facet of the catalyst surface, (5) chemical composition, and the (6) binding sites. Numbers indicate the layer in the structure graph where (1) is the highest layer below the total bucket

Data analysis
From the CO bonding and dissociation reaction, the buckets with the highest data instances were chosen to be analysed. The CO dissociation reaction bucket was formed by combining the OH binding site bucket of M3M9 and M9M3. The M9M3 data was inversed, so this also became M3M9 data. The combined M3M9 bucket was called A3B9 to distinguish between the first and the second M.
For the CO bonding reaction bucket, the { } binding site of the M3M1 bucket was chosen. Although there is no given binding site, this bucket has the most data instances and can therefore be used for a plot. The M3M1 was called A3B1 to distinguish between the first and the second M.
The 3D plots were made with MATLAB.

Results
Tree Chart
Each square in the tree chart (figure 4 in appendix A) represents a collection of data instances that meet the bucket in which they fall and above.

Data analysis
Data comes from the buckets as explained in the 'Data analysis' in Experimental. In the 3D plots (figure 2 and 3) it is clearly visible that with increasing both group numbers, the reaction energy increases.
Discussion and implications

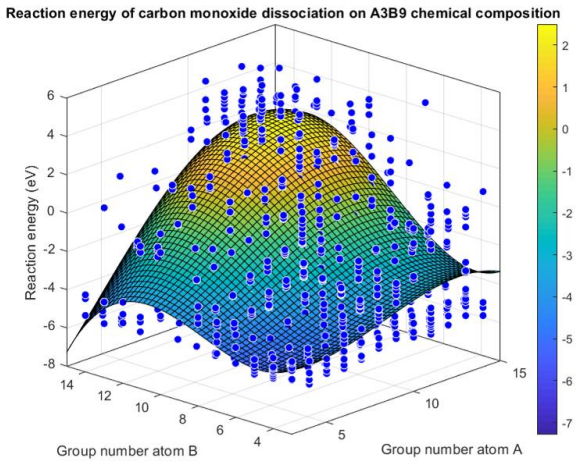

Figure 2 A 3D plot with the dissociation reaction energy (eV) plotted against the periodic group of atom A versus B of the chemical composition with a hollow binding site. A third-degree polynomial was used in MATLAB to generate a fit through all the data points.

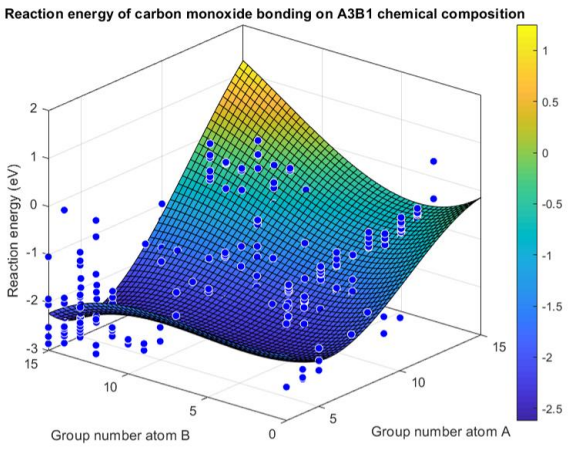

Figure 3 A 3D plot with the bonding reaction energy (eV) plotted against the periodic group of atom A versus B of the chemical composition with no given binding site. A third-degree polynomial was used in MATLAB to generate a fit through all the data points.

The tree chart is useful to navigate through the curated data of the Catalysis-Hub.org database. However, the processed data required a lot of work to get it i n the right buckets. When a new research is uploaded to the database, the tree chart with the data will have to be updated manually at this time. A possibility to solve this is by prescribing a code to automate the process.
The data analysis was useful to see if useful data could follow from our research. It can be said with some certainty that with increasing periodic groups of the chemical composition atoms, the reaction energy will also i ncrease with CO bonding and dissociation. However, it is clearly visible that the plots missed a lot of data points in figure 2 and 3 and therefore less trustworthy. Furthermore, the CO binding data points (Figure 3) did not have a given binding site, so it is pos sible that different binding sites were compared with each other.
More qualitative research will have to follow to be able to say something about the data with more certainty.

Conclusions

This research has extracted and sorted a vast amount of information in the Catalysis-Hub.org database and has made it more usable for machine learning. This research has also shown periodic trends in chemical composition and reaction energy, but more qualitative research is needed to get more certainty in the trends.

ACKNOWLEDGMENT

Utrecht University is kindly acknowledged for giving the opportunity to follow an internship in the Autonomous Energy Materials Discovery [AMD] research group. We thank the AMD group and, in particular, dr. Süleyman Er (supervisor) and dr. Ilker Tezsevin (daily supervisor) for assistance through the project. We also thank the Dutch Institute for Fundamental Energy Research (DIFFER) for the workspace they have given us.

REFERENCES

1. Zou, C.; Zhao, Q.; Zhang, G.; Xiong, B. Energy Revolution: From a Fossil Energy Era to a New Energy Era. Natural Gas Industry B **2016**, 3 (1), 1–11. doi:10.1016/j.ngib.2016.02.001
2. Sarup, B.; Wojciechowski, B. W. Studies of the Fischer-Tropsch Synthesis on a Cobalt Catalyst II. Kinetics of Carbon Monoxide Conversion to Methane and to Higher Hydrocarbons. *The Canadian Journal of Chemical Engineering.* **1989**, 67 (1), 62–74. doi:10.1002/cjce.5450670110
3. Yates, I. C.; Satterfield, C. N. Intrinsic Kinetics of the Fischer-Tropsch Synthesis on a Cobalt Catalyst. *Energy & Fuels.* **1991**, 5 (1), 168–173. doi:10.1021/ef00025a029
4. Shetty, S. G.; Ciobica, I. M.; Hensen, E. J. M.; Santen, R. A. V. Site Regeneration in the Fischer–Tropsch Synthesis Reaction: a Synchronized CO Dissociation and C–C Coupling Pathway. *Chemical Communications.* **2011**, 47 (35), 9822. doi:10.1039/c1cc11499a
5. Winther, Kirsten T.; Hoffmann, Max J.; Boes, Jacob R.; Mamun, Osman; Bajdich, Michal; Bligaard, Thomas. Catalysis-Hub.org, an open electronic structure database for surface reactions. *Scientific Data.* **2019,** *6. doi:10.1038/s41597-019-0081-y*
6. Fields, M.; Tsai, C.; Chen, L. D.; Abild-Pedersen, F.; Nørskov, J. K.; Chan, K. Scaling Relations for Adsorption Energies on Doped Molybdenum Phosphide Surfaces. *ACS Catalysis.* **2017**, 7 (4), 2528–2534. doi:10.1021/acscatal.6b03403
7. Hoffmann, M. J.; Medford, A. J.; Bligaard, T. Framework for Scalable Adsorbate–Adsorbate Interaction Models. *The Journal of Physical Chemistry C.* **2016**, 120 (24), 13087–13094.
8. Jiang, T.; Mowbray, D. J.; Dobrin, S.; Falsig, H.; Hvolbæk, B.; Bligaard, T.; Nørskov, J. K. Trends in CO Oxidation Rates for Metal Nanoparticles and Close-Packed, Stepped, and Kinked Surfaces. *The Journal of Physical Chemistry C.* **2009**, 113 (24), 10548–10553. doi:10.1021/jp811185g
9. Mamun, O.; Winther, K. T.; Boes, J. R.; Bligaard, T. High-Throughput Calculations of Catalytic Properties of Bimetallic Alloy Surfaces. *Scientific Data.* **2019**, 6 (1). doi:10.1038/s41597-019-0080-z
10. Nieddu, M.; Boatto, G.; Pirisi, M. A.; Dessì, G. Determination of Four Thiophenethylamine Designer Drugs (2C-T-4, 2C-T-8, 2C-T-13, 2C-T-17) in Human Urine by Capillary Electrophoresis/Mass Spectrometry. *Rapid Communications in Mass Spectrometry.* **2010**, 24 (16), 2357–2362. doi:10.1002/rcm.4656.
11. Sandberg, R. B.; Hansen, M. H.; Nørskov, J. K.; Abild-Pedersen, F.; Bajdich, M. Strongly Modified Scaling of CO Hydrogenation in Metal Supported TiO Nanostripes. *ACS Catalysis.* **2018**, 8 (11), 10555–10563. doi:10.1021/acscatal.8b03327
12. Schumann, J.; Medford, A. J.; Yoo, J. S.; Zhao, Z.-J.; Bothra, P.; Cao, A.; Studt, F.; Abild-Pedersen, F.; Nørskov, J. K. Selectivity of Synthesis Gas Conversion to C2 Oxygenates on Fcc(111) Transition-Metal Surfaces. *ACS Catalysis.* **2018**, 8 (4), 3447–3453. doi:10.1021/acscatal.8b00201
13. Kohn, W.; Becke, A.; Parr, R. Density Functional Theory of Electronic Structure. *The Journal of Physical Chemistry.* **1996,** *100* (31), 12974-12980. doi:10.1021/jp960669l
14. Ulissi, Zachary W.; Medford, Andrew J.; Bligaard, Thomas; Nørskov, Jens K. To address surface reaction network complexity using scaling relations machine learning and DFT calculations. *Nature Communications,* **2017,** 8. doi:10.1038/ncomms14621
15. Molmod.ugent.be. (**2019**). Comparing Solid State DFT Codes, Basis Sets and Potentials | Center for Molecular Modeling. [online] Available at: https://molmod.ugent.be/deltacodesdft [Accessed 19 Jun. 2019].
16. Wandelt, K. (**2018**). Encyclopedia of interfacial chemistry. Amsterdam [i pozostałe]: Elsevier, p.686. ISBN:9780128097397

Appendix A – Tree Chart

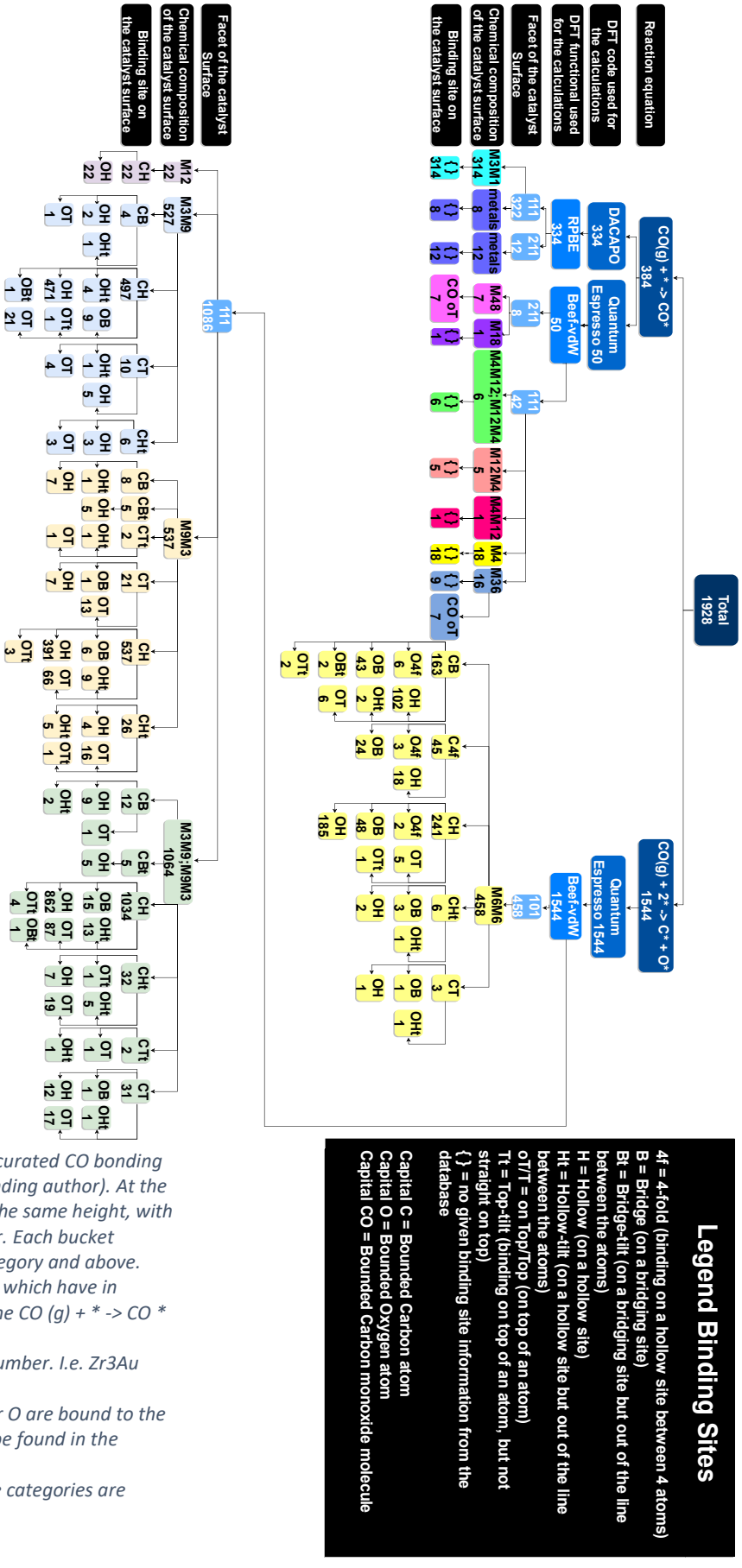

*Figure 4 The tree chart contains a logical distribution of all curated CO bonding and dissociation data (can be requested from the corresponding author). At the far left are categories that include all squares (buckets) at the same height, with the binding site category containing the entire bottom layer. Each bucket contains the total data instances of the relevant bucket category and above. (E.g. the most left RPBE bucket contains 334 data instances which have in common the RPBE DFT functional, DACAPO DFT code and the CO (g) + * -> CO * reaction equation.)*
The chemical composition consists of an M (Metal) and a number. I.e. Zr3Au becomes M3M1 and Au36 becomes M36.
The binding sites are the places where and how the CO, C or O are bound to the catalyst. Explanation of the binding site abbreviations can be found in the attached legend.
NB the tree chart is cut in the middle to fit on this page. The categories are repeated on the left.