

UCLA

UCLA Previously Published Works

Title

Developing a Descriptor-Based Approach for CO and NO Adsorption Strength to Transition Metal Sites in Zeolites

Permalink

<https://escholarship.org/uc/item/2x39v587>

Journal

Chemistry of Materials, 29(15)

ISSN

0897-4756

Authors

Götl, Florian
Müller, Philipp
Uchupalanun, Pajeau
[et al.](#)

Publication Date

2017-08-08

DOI

10.1021/acs.chemmater.7b01860

Peer reviewed

Developing a Descriptor-Based Approach for CO and NO Adsorption Strength to Transition Metal Sites in Zeolites

Florian Göttl,^{*,†,‡} Philipp Müller,[†] Pajeon Uchupalanun,[†] Philippe Sautet,[¶] and Ive Hermans^{*,†,‡}

[†]*University of Wisconsin - Madison, Department of Chemistry, University Avenue 1101, 53706 Madison, Wisconsin, USA*

[‡]*University of Wisconsin - Madison, Department of Chemical and Biological Engineering, 1415 Engineering Drive, 53706 Madison, Wisconsin, USA*

[¶]*University of California Los Angeles, Department of Chemical and Biomolecular Engineering, Los Angeles, CA 90005, United States*

E-mail: fgoeltl@wisc.edu; hermans@chem.wisc.edu

June 27, 2017

Abstract

The discovery of new materials tailored for a given application typically requires the screening of a large number of compounds and this process can be significantly accelerated by computational analysis. In such an approach the performance of a compound is correlated to a materials property, a so called descriptor. Here we develop a descriptor-based approach for the adsorption of CO and NO to Cu, Ni, Co and Fe sites in zeolites. We start out by discussing a possible design strategy for zeolite catalysts, define the studied test set of sites in the zeolites SSZ-13 and Mordenite, and define a

set of appropriate descriptors. In a subsequent step we use these descriptors in single-, two- and multi-parameter regression analysis and finally use a machine-learning genetic algorithm to reduce the number of variables. We find that one or two descriptors are not sufficient to accurately capture the interactions between molecules and metal centers in zeolites and indeed a multi-parameter approach is necessary. Even though many of the descriptors are directly correlated, we identify the position of the s-orbital and the number of valence electrons of the active site as well as the HOMO-LUMO gap of the adsorbate as most important descriptors. Furthermore the reconstruction of the active sites upon adsorption plays a crucial role and when it is explicitly included in the analysis, correlations improve significantly. In the future we expect that the fundamental methodology developed here will be adapted and transferred to selected problems in adsorption and catalysis and will assist the rational design of materials for the given application.

Introduction

One of the key challenges in many technological fields is the discovery of new materials with optimum properties. The most intuitive approach in this context is to use high-throughput screening, i.e. test a series of different compounds and identify the most efficient one. At the same time, the synthesis and testing of more complex materials can be exceedingly time consuming, which limits the size of the search space.¹ The progress in electronic structure calculations over the last decades has allowed to circumvent these problems. The comparatively efficient treatment of entire classes of materials allows the determination of fundamental material parameters, so called descriptors, which can be correlated to the performance in a given application.²⁻⁸ Today databases for these descriptors for different types of materials exist,⁹ and it is possible to screen them to identify promising compounds and predict their efficiency.

A class of materials of particular interest are zeolites, which are among the most impor-

tant industrial catalysts today and also find applications in other, technologically relevant fields.¹¹ Their main feature is a crystalline, microporous structure, resulting in a large surface area per volume in combination with diffusional limitations with respect to molecular size. While purely siliceous materials are chemically inert and mainly suitable for applications in separation,¹² they can be functionalized in various ways to give them properties tailored towards the desired applications in catalysis or as storage medium.^{13,14} Most commonly this functionalization is associated with the introduction of Al in the framework, which leads to the creation of Brønsted acid sites, but also single metal atoms or clusters, acting as Lewis acids, can be introduced.

Despite their importance in industry, comparatively few descriptor-based, predictive approaches have been presented in the field of zeolites so far.¹⁵⁻¹⁹ This is mainly due to the complexity of these materials and their comparatively large unit cells, which are associated with a significant computational cost. So far, the focus was put on predicting the synthesis process and only over the past few years, research has focused on developing descriptor-based approaches and scaling relationships for catalytic applications so far,¹⁵⁻²⁰ methods that have been applied successfully to predict the activity of compounds in heterogeneous surface catalysis.⁷ However, previous work in the field has focused on a very narrow class of possible metal sites and an accurate, general approach has not been achieved.

In this work we develop an accurate, descriptor-based approach for the adsorption of CO and NO to extra-framework metal cations in zeolites. Zeolites containing these types of active sites have been successfully used in a variety of relevant catalytic reactions and the prediction of the efficiency of so far unexplored materials would allow for an efficient design of new catalysts or storage materials. CO and NO are molecular probes for key functional properties of zeolites,²¹⁻²³ and hence predicting their adsorption strength sheds light on their chemical activity in a variety of reactions and molecular storage problems. We start our discussion with a possible design scheme, continue with a detailed discussion of the chosen descriptors, and finally arrive at molecular adsorption and how descriptor-based approaches

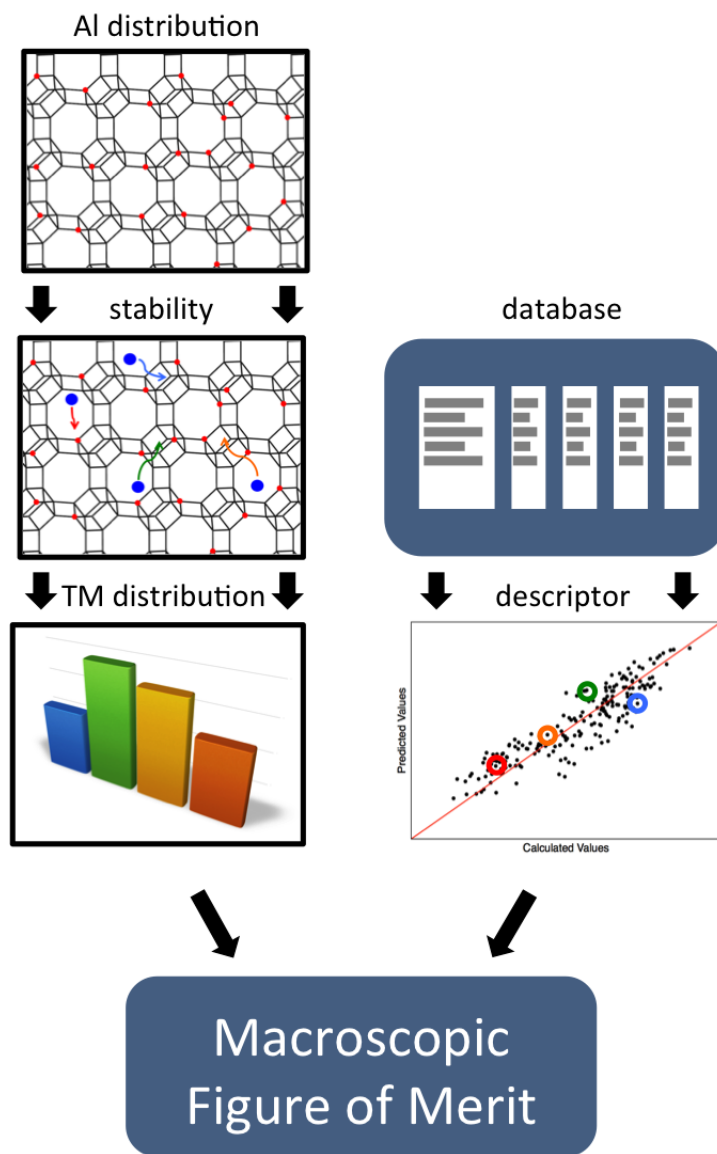
are able to predict the adsorption strength. We show that one and two parameter regression analysis is not sufficient to accurately predict the adsorption strength. We therefore closely follow the strategy of Ras et al.,³ who used a multi-parameter regression and a genetic algorithm to accurately predict the adsorption strength of a set of small molecules to metal surfaces, and finally arrive at an accurate prediction of the adsorption strength.

A design-based approach for transition metal exchanged zeolite catalysts

While the targeted design of heterogeneous surface catalysts and homogeneous catalysts based upon theoretical and screening approaches has been shown to be highly successful,^{8,24,25} similar studies for transition metal exchanged zeolites are rare.^{15–19} This originates from the presence of a distribution of active sites,^{21,22,28} which leads to a larger complexity in predicting macroscopically measurable observables.^{21,26}

In the following we will describe a possible catalyst design route which is visualized in Scheme 1. The underlying property of a zeolite is the distribution of Al atoms, which is determined during the synthesis of the material and changing it afterwards very often involves the partial destruction and reconstruction of the framework.²⁷ This substitution of Si by Al creates a local charge balance, which is compensated for by the presence of counter cations. Furthermore, thermodynamics indicates that upon ion exchange the energetically most stable positions will be occupied by transition metal cations.²⁸ Kinetic restrictions in the preparation process aside, combining the Al-distribution and ion exchange level with this information will allow to understand the distribution of metal sites in the material. However, it is important to keep in mind that this is an idealized image, since not all the information is known today.

Similar to other problems within materials science or catalysis, it is then possible to create a database containing descriptors for the different metal centers. Developing a descriptor-



Scheme 1: A design scheme for transition metal exchanged zeolite catalysts. If the distribution of Al in the material is known, thermodynamics will dictate the distribution of metal centers in the material. Combining this knowledge with a descriptor based approach for each single site, will then allow to arrive at a macroscopic figure of merit.

based approach allows to predict the contribution of a selected descriptor to a certain figure of merit. Combining this with the distribution of active sites discussed above leads to the prediction of a macroscopic observable (such as adsorption strength of a molecule or catalytic conversion). Following this logic, it is possible to identify four contributions determining the catalytic activity of a zeolite catalyst: (i) the framework type, (ii) the Al-distribution, (iii) the ion exchange level, and (iv) the nature of the exchanged cation.

Gathering information about the distribution of Al atoms in zeolites is notoriously difficult. This originates in a challenging characterization process, particularly for low to medium Si/Al ratios,^{21,22,26} and a lack of predictive tools, which take the complex synthesis process into account. We will therefore focus on creating a database for a zeolite test-set and develop a descriptor-based approach to predict the adsorption strength of CO and NO. In the future, these considerations might be used to arrive at an adsorption capacity of this material or to develop Brønsted-Evans-Polanyi relationships for catalytic reactions.²⁹

The test set and descriptors

Similar to quartz, the basic chemical composition of zeolites is SiO_2 . The building blocks are SiO_4 tetrahedra (T-sites), which are linked via O atoms and form microporous channels and pores. The framework is activated when, instead of Si, an Al-atom is present in the center of the T-site. To saturate all O bonds a local negative charge surrounding the Al atom is created, which is compensated for by the presence of a positive ion. The most simple positive ion is a proton, but it can be exchanged by other cations and in this study we will focus on divalent Cu, Ni, Co and Fe cations. To host one metal atom in one unit cell and create the formal charge state of +2 it is necessary to substitute 2 Al atoms in each unit cell.

The exact chemical activity of the exchanged cations will depend on the exchange site and the local distribution of Al atoms close to it.^{21,22} A variety of different zeolite structures exists, which again leads to a large variety of different ion exchange sites. However, only

comparatively few secondary building units (groups of several T-sites) exist and the most important industrial zeolites can be grouped into two large ensembles: (i) pentasil zeolites, which are formed from five membered rings and show a clear one- or two-dimensional channel structure and (ii) structures based on double six-membered ring (D6R) prisms, which form three-dimensional cavities.

For pentasil zeolites three basic types of active sites have been suggested by Wichterlowa *et al.*, the α , β and γ sites.³⁰ Even though the exact geometry of particularly the β site varies significantly with the zeolite structure, they seem to lead to rather similar spectroscopic features for a variety of characterization methods. However, this model is not complete, since sites in larger rings, which will be occupied at high exchange levels, are neglected. For structures based on the D6R prism, active sites are either in six rings (similar to the β site in pentasils), or larger rings.

To represent these two general types of zeolite structures we model the zeolite SSZ-13, a zeolite in the chabazite structure, as D6R prism based zeolite and the pentasil zeolite Mordenite (MOR), which are displayed in Fig. 1. As shown in Fig. 1, the SSZ-13 structure has three different exchange sites and the MOR structure eight.³¹ However, work by Grybos and Hafner on divalent Pd shows that this element is only stable in five of these positions, named Mortier sites A through E.³² We assume similar trends for the divalent metals studied in this work and focus on the same exchange sites. To arrive at the divalent nature of the active sites we substitute two Al atoms in the unit cell, which are distributed following the Löwenstein rule,³³ i.e. that neighboring T-sites must not be occupied by Al atoms. For each active site we investigate different Al distributions, which lead to cases where two, one or none Al atoms are close to the studied cation. In the case of SSZ-13, this leads to 11 different configurations, while 15 different configurations for MOR are studied. The different, local Al positions for the two zeolite frameworks are given in the structural files in the Supporting Information. We furthermore focus on Fe, Co, Ni, and Cu, metals which are relevant for various catalytic applications in zeolites and have been studied extensively.

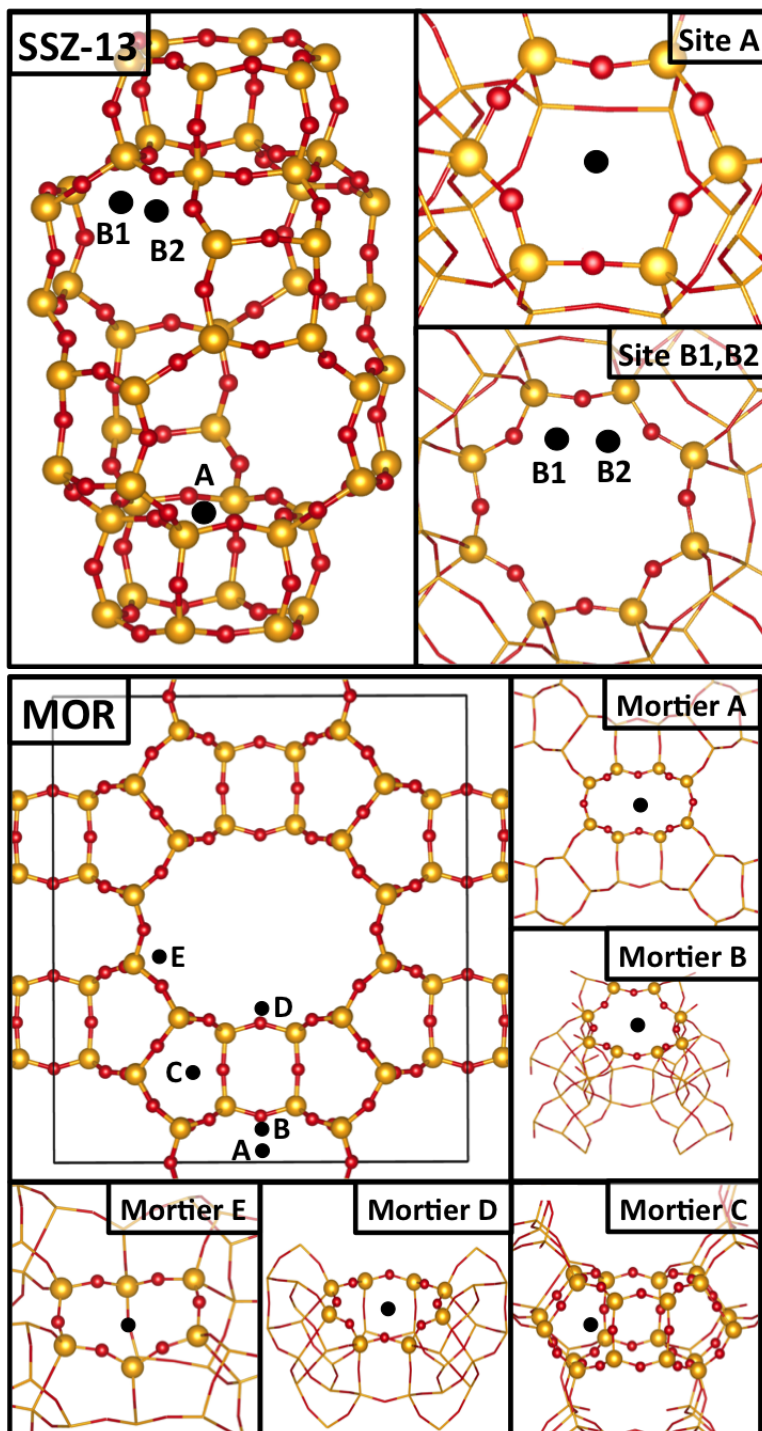


Figure 1: The crystal structures and studied exchange sites of SSZ-13 and MOR. We study different Al distributions and all the studied configurations are given in the structural files in the Supporting Information. In the atomistic figures, red atoms correspond to O and yellow atoms to Si. For the active sites, only atoms in the neighboring ring-structure are displayed and the rest of the framework is displayed as bonds.

Descriptors for the metal centers

The first task in arriving at a predictive approach that links the properties of the metal active site to adsorption properties is the definition of descriptors. In the literature, descriptors based on structural parameters, the stability of the active site, electronic structure related parameters, the volume and the charge of the active site have been suggested.^{2,3,25,35,36,38–40}

It is a priori not clear which descriptors are necessary to accurately describe the molecular adsorption of CO and NO to transition metal exchanged zeolites. We therefore focus on a series of descriptors. To capture structural effects we study three parameters, i.e. the coordination number of the metal (n), the number of bonds to activated O bonds (n_{act}) and the average bond length $\bar{\rho}$. We also use the stability of the active site (E_{stab}), as well as the position of the lowest unoccupied molecular orbital (E_{LUMO}), the center of mass of all unoccupied d orbitals (ϵ_{LUMO}), the position of the spin up and down s-orbitals (E_{s-up} and E_{s-down}) and the center of the d-orbitals (ϵ_d). We furthermore add the Bader charge of the active center (Q) and the Bader volume (V). We furthermore add two tabulated values with the number of valence electrons of the metal centers (n_{val}) and the atomic mass of the adsorbates (M_w) when necessary in our fit. The full definition and a brief discussion of these parameters is given in the supporting information section S1.

Stability of the active sites

We start our discussion by focusing on the stability of the active sites. Due to bond-order conservation arguments, it is expected that the stability of the active site is inversely proportional to its ability to bind adsorbates and we therefore introduce the stability with respect to metal-chloride and the bare zeolite framework (E_{stab}) as a descriptor. However, it is also important in understanding which sites will be present in a realistic system after ion exchange.

A graphical representation of this parameter is shown in Fig. 2 and the values are given in the supplementary information. As a general trend, we observe that the relative stabilities

of the different exchange sites are very similar for all metals. We find that the stability within each group of active sites increases with the amount of Al atoms close to them. In almost all cases Co and Fe have comparable energy differences for all those situations. At the same time, Cu and Ni are more stabilized if no Al atoms are close to the active sites, an effect that is stronger for Cu than for Ni. However, except for site A in SSZ-13, all the stabilization energies for all those sites with no Al close are positive, which confirms the chemical intuition that none of those sites will be occupied during the ion exchange process and every cation will be located close to an Al atom.

Correlation between the descriptors

Physical chemistry indicates that the introduced descriptors have certain interdependencies. To explore them in more detail we investigate the correlation map, which is displayed in Fig. 3. In our opinion the most interesting information is contained in the correlation of E_{stab} with the other descriptors, since it gives direct information about the bonding between metal cation and the framework. On one hand, the stability is strongly correlated to E_{LUMO} , which indicates a strong covalent bond between TM-metal and framework, since this empty d-state is the anti-bonding orbital of the O-TM interaction. On the other hand, the correlation between s-states and stability indicates that TM-sites are more stable in more tightly confined environments, where the Pauli repulsion pushes the s-state up. Clearly the electrostatic interaction between TM atom and framework should also play a major role in the stabilization of the sites. However, the correlation map shows only a very weak correlation between Q and E_{stab} . This indicates that the bond-strength is mainly influenced by the covalent and not by the ionic components. As expected, the stability is also directly correlated to n_{act} .

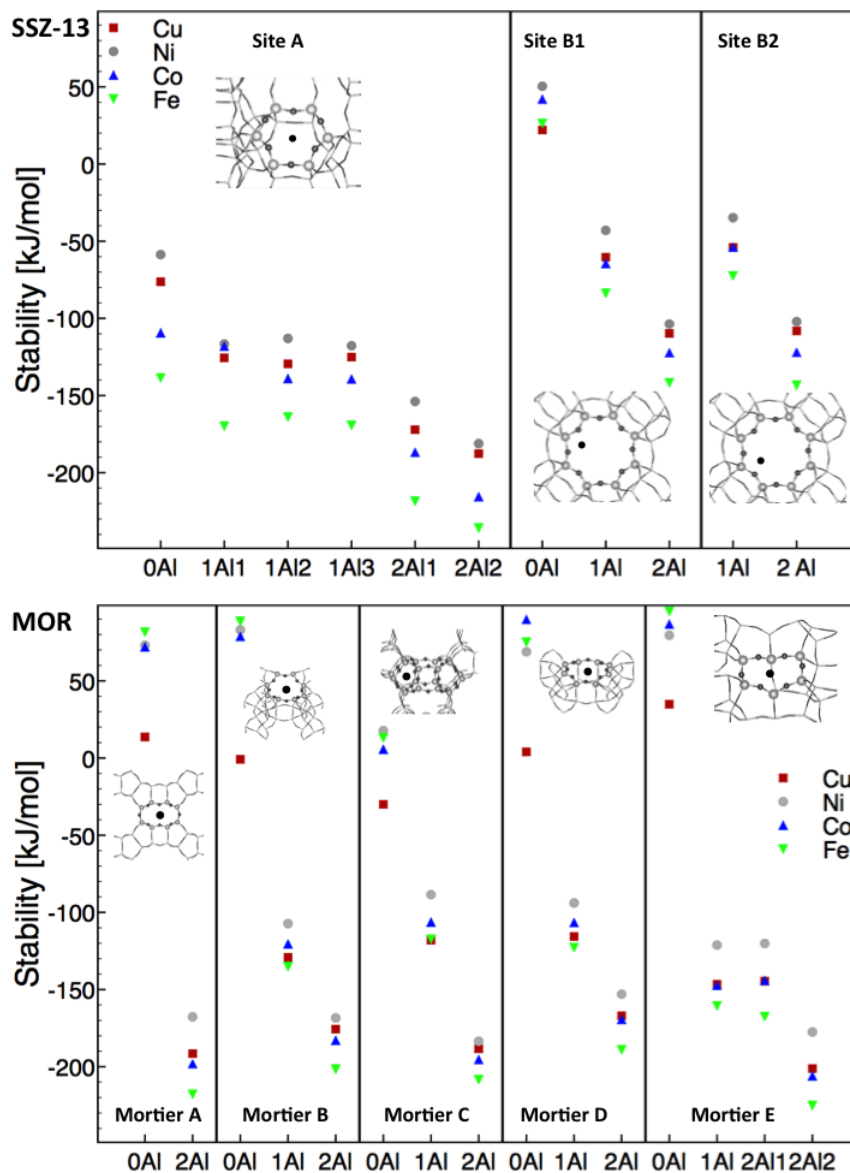


Figure 2: Stability of transition metal atoms in the different exchange sites for SSZ-13 and MOR. Red squares correspond to Cu, grey diamonds to Ni, blue triangles to Co and green triangles to Fe. The different Al positions for the two zeolite frameworks are given in the structural files provided in the Supporting Information.

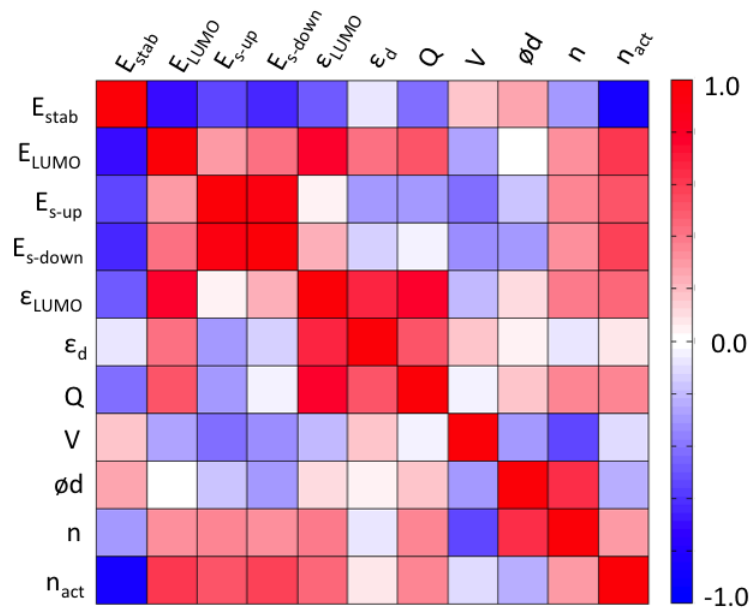


Figure 3: Correlation map of the studied descriptors. Red and blue color correspond to strong, direct correlation, white to no correlation. The descriptors are the stability of the active site (E_{stab} , the position of the LUMO orbital (E_{LUMO} , the position of the s-orbitals (E_{s-up} for the spin-up component, E_{s-down} for the spin-down component), the average energy of the unoccupied d-orbitals (ϵ_{LUMO}), the charge at the TM cation (Q), the Bader volume of the site (V), the average bond length to the framework (\AA), the coordination number (n) and the number of bonds to activated O atoms (n_{act}).

Molecular adsorption

In this work we focus on the molecular adsorption of CO and NO to the transition metal centers discussed above and to understand the differences between these molecules, we first want to discuss the fundamental bonding principles involved.

CO has a total of 10 electrons, which leads to a full occupation of the bonding p- σ and p- π states. At the same time the anti-bonding p- σ^* and p- π^* states are unoccupied. The bonding of CO to TM atoms is usually described within the Blyholder model.⁴¹ A bonding interaction between the molecular σ orbital with the dz^2 state of the TM cation can be formed, i.e. charge is donated from the molecule to the metal, and at the same time charge is donated from the TM d-states to the anti-bonding π^* state, which typically leads to an increase in the bond-length with adsorption strength.

NO, on the other hand, has one more valence electron compared to CO and this leads to significant differences in the bonding. The extra electron occupies the anti-bonding π^* -orbital and is donated to the TM cation upon bonding. Formally speaking there is a bond between a NO^+ and TM^I -cation. On top of this donation also the bonding mechanisms already discussed for CO are present.

Clearly these differences in the bonding lead to a significantly different behavior when these molecules adsorb to active centers, which is reflected in the structures and the energetics. To keep the discussion focussed, we give the structural files in the Supporting Information and will, in the following, focus on the energetics.

When studying the adsorption energies, displayed in Fig. 4, several significant differences are observed between CO and NO. First of all, NO adsorbs more strongly to the active centers. For most of the metals this difference is on the order of magnitude of 100 kJ/mol. Another important point to notice is that the adsorption strength typically increases when moving from sites with two adjacent Al atoms to sites with no Al atoms close. However, it is important to point out that differences in adsorption strength are far larger when moving from no Al in the vicinity to one Al atom, than moving from one neighboring Al atoms to

two Al atoms. It is also worth noting that this effect is significantly smaller for site A in SSZ-13 compared to all the other sites.

While these trends generally hold for all metals when studying CO adsorption, this is not true anymore for NO adsorption to Cu. Here adsorption strengths are significantly decreased and for some of the sites in MOR trends in adsorption strength with adjacent Al atoms are reversed. Such differences are not entirely surprising and can be traced back to the bonding mechanism. In this bond, NO donates an electron to a divalent Cu, which leads to a filled d-shell of the TM. Clearly this ion interacts significantly weaker with the framework and shows a significant reconstruction. Both effects weaken the observed adsorption strength.

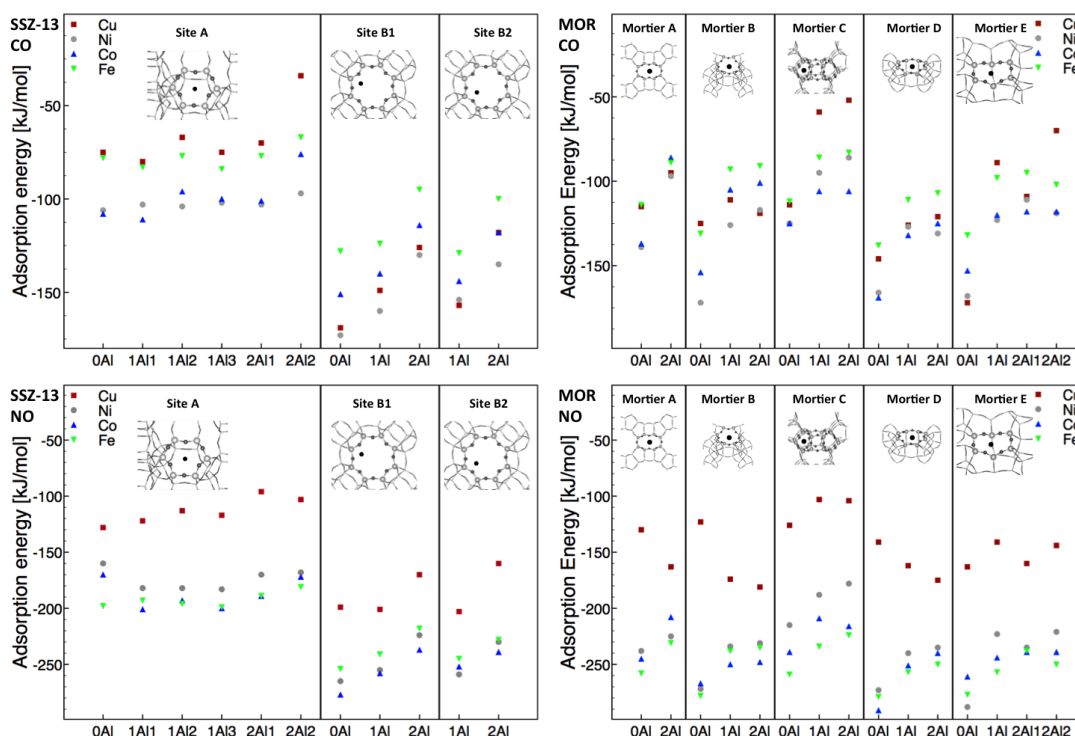


Figure 4: Adsorption energies for CO and NO in the ion exchanged zeolites SSZ-13 and MOR. The legend in the top left corner indicates the zeolite and molecule, adsorption energies for Cu are shown as dark-red squares, Ni in grey circles, Co in blue triangles and Fe in green upside down triangles.

Regression analysis

One of the key goals of this work is to establish relationships between the adsorption strength and the descriptors defined above. The most simple way to achieve a predictive approach is a regression analysis. Such an analysis is very often performed for a single parameter, but can also be extended to multiple parameters. Here we start our analysis focusing on one parameter, followed by a two parameter regression and finally, a multi-parameter regression is performed for the dataset consisting of adsorption of two different molecules to four metals, eleven different adsorption sites in SSZ-13 and 15 adsorption sites in MOR, leading to a total of 208 data points combined with a total of 13 descriptors.

One parameter regression

We start our analysis using a single parameter fit using linear and square terms for CO using all descriptors and dividing our dataset into subsets for different metals and sites, containing between 8 and 24 datapoints. The R^2 values for CO, as measure of the quality of the fit, are visualized in A and B of Fig. 5.

Except for one subset we find correlation coefficients (R^2) below 0.80. When comparing the R^2 values for different metals and sites for CO adsorption, the best performing descriptors vary when different subsets are chosen. For most of the sites the stability of the active site seems to be a good criterion and it seems to work especially well for sites where the reconstruction of the active site upon adsorption is small, such as Mortier site A and the eight membered ring sites in SSZ-13 and MOR (Fig. 5 A). For Mortier site A we also find other descriptors to show weak correlation. When moving to the different metals, the stability of the active site still shows a weak correlation with the adsorption strength (Fig. 5 B). The best correlations are found when Ni, Co and Fe are analyzed on their own, actually the E_{s-up} descriptor leads to the best fit, with an R^2 above 0.75. Especially for Mortier site A we also find good correlation for other descriptors, but no general trend for the other sites can be

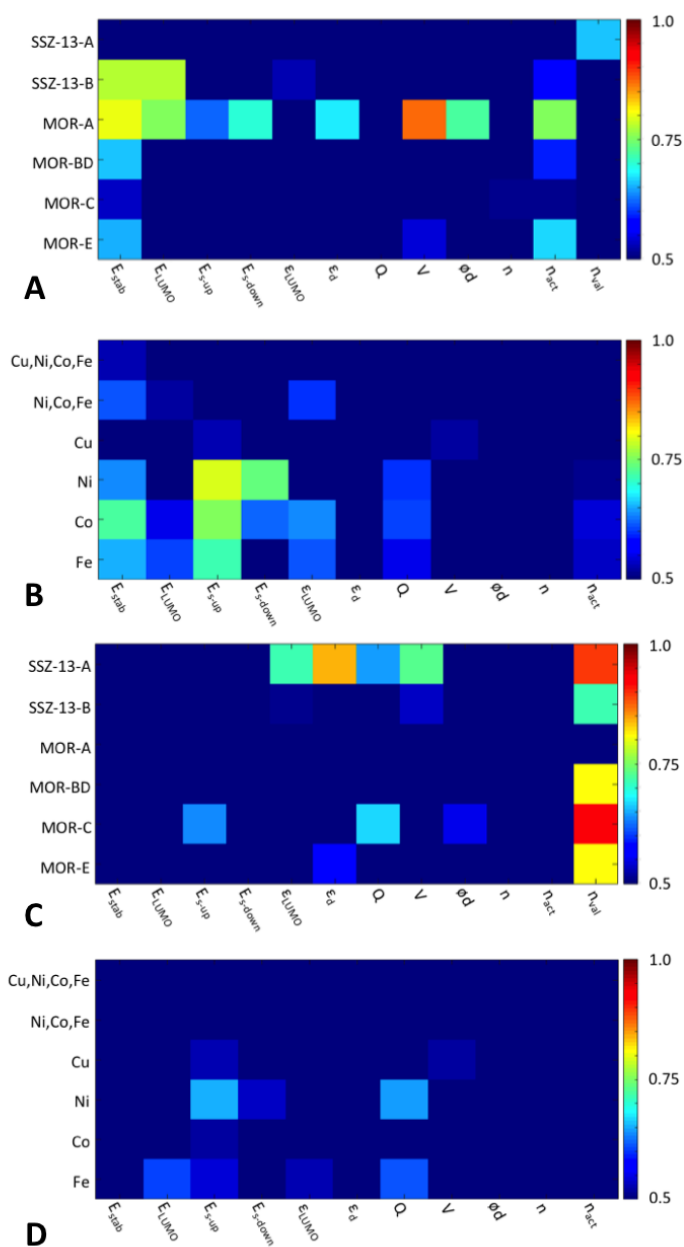


Figure 5: R^2 values for regression analysis using linear and square terms for single parameters for CO (A,B) and NO (C,D) adsorption to TM-sites in SSZ-13 and MOR. The dataset is split up into the different sites (A, C) and the different metals (B, D). Dark-blue corresponds to R^2 values lower than 0.5, dark-red to 1.0.

found here.

The picture for NO adsorption is significantly different from CO (Fig. 5 C and D). While generally no good correlations can be found when splitting the dataset with respect to metals, for some of the different sites excellent correlations can be found with the number of valence electrons when studying the different sites. These results are encouraging, since they indicate that the choice of metal seems to be one of the most important factors for adsorption strength. However, this quantity is not able to distinguish between different local environments of the metal sites.

Two parameter regression

As a next step we performed a square two-parameter regression, which includes two linear terms, two square terms and one cross term (explained in detail in the computational methods section) and the results for the total sets for CO and NO adsorption are displayed in Fig. 6 A and B. To test for the reliability of our predictions we split our dataset into two thirds as test set and one third as validation set, which in the end leads to a R^2 value, describing the quality of the fit of the test set, and also a Q^2 value, describing the quality of the fit when extrapolating to the validation set. For CO (Fig 6 A) correlation for the full dataset is still comparatively bad, with all R^2 below 0.7.

When now also considering a two-parameter model for NO (Fig. 6 B) we find trends similar to CO, only with worse correlations. In general all the R^2 values are below 0.6 for the full set. However, the Q^2 values show that the fit is still not reliable.

Multi-parameter approach

The low quality of the regression, represented by low R^2 values, is rather discouraging and as a next step we systematically increase the number of variables used in the fit. The first step is probing for the number of variables necessary to arrive at an accurate approach. We therefore studied the quality of the regression analysis for stepwise increasing the number of

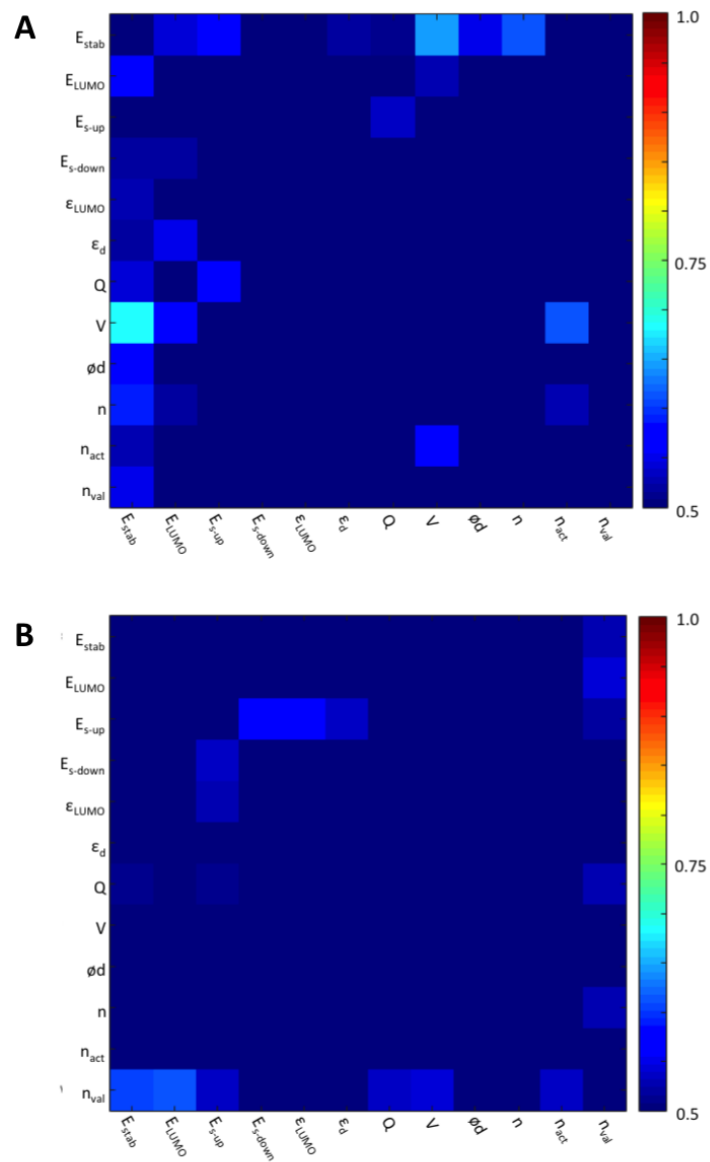


Figure 6: Here R^2 values for linear and square two-parameter regression for the adsorption of CO (A) and NO (B) are displayed. Only the full dataset is studied and we did not find good correlation with any two parameters.

variables from 2 to 10. In this approach we used linear, square and cross terms and explore all the different combinations of descriptors, and only the highest R^2/Q^2 values are displayed in Fig. 7. Generally speaking, we find the R^2 and Q^2 values for the full dataset with CO and NO to be higher than those for the separate molecules. Furthermore we see a continuous increase in R^2 values for all three options, i.e. CO, NO and both molecules, until it plateaus around a value $R^2=0.88$ for CO and both molecules and a value of $R^2=0.84$ for NO. At the same time the Q^2 values show a larger variation with values significantly below $Q^2=0.8$ for the separate molecules with slightly higher values for both molecules together.

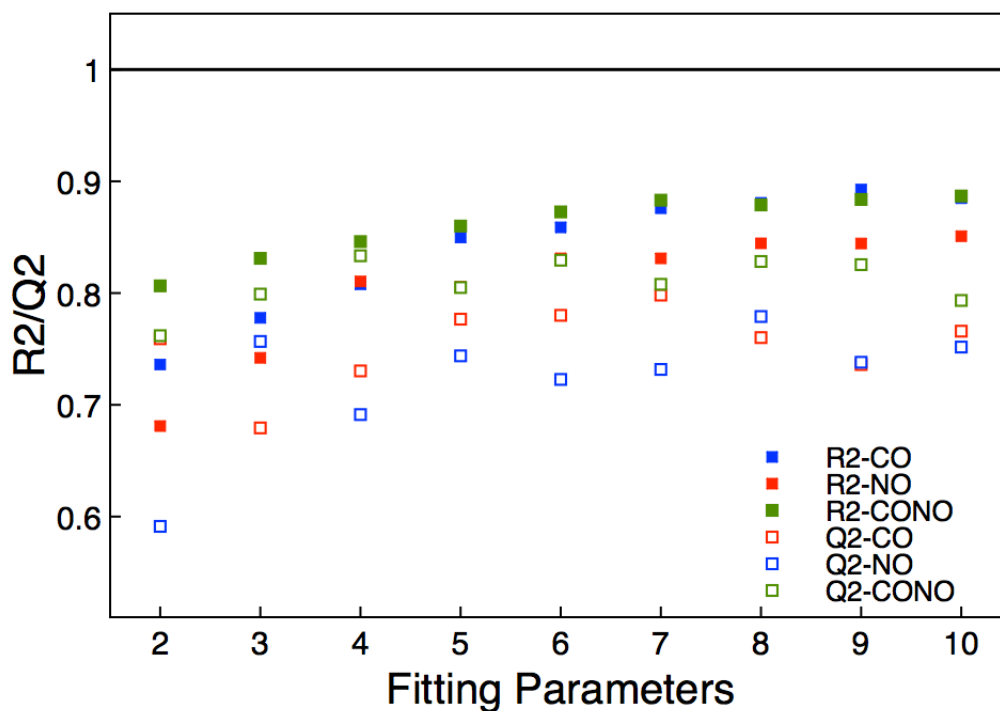


Figure 7: R^2 (filled squares) and Q^2 (empty squares) values for regression analysis for 2 to 10 parameters for adsorption of CO (blue), NO (red) and CO and NO (green) molecules.

We furthermore performed a more thorough analysis of the regression approach using all descriptors. The R^2 values and the Q^2 values are graphically displayed as orange squares in Fig. 8. When we use only linear terms, the fit for CO and NO adsorption (see Fig. 8 C and in Tab. S1 in the Supporting Information) to all metals is, with a $R^2(Q^2)$ of 0.84(0.68) still comparatively bad. Adding the square terms improves the quality of the fit

significantly and leads to an $R^2(Q^2)$ of 0.83(0.85). Linear and cross terms already give an excellent $R^2(Q^2)$ of 0.90(0.85), and finally including all terms improves the $R^2(Q^2)$ values to 0.92(0.89). Interestingly, the quality of the fit gets worse when only CO (see Fig. 8 A) or NO (see Fig. 8 B) are considered. This can likely be explained by the additional descriptor introduced to distinguish between the different molecules (see methods section for details). The respective $R^2(Q^2)$ values are 0.70(0.68), 0.83(0.78), 0.89(0.76) and 0.83(0.76) for linear/linear+square/linear+cross/linear+square+cross for CO, and 0.84(0.68), 0.80(0.57), 0.89(0.63) and 0.84(0.77). We believe that these comparatively large differences between R^2 and Q^2 values are most likely due to the small size of our data-set (see below).

As already discussed above, predicting the adsorption of NO to Cu^{II}-sites is complicated, since Cu changes from an open shell to a closed shell configuration. Therefore, we also only investigated a subset of the data describing the adsorption to the metals Fe, Co and Ni (see Fig. 8). We see a dramatic increase in the accuracy for the linear terms ($R^2(Q^2)$ of 0.92(0.93)), which remains similar, when square ($R^2(Q^2)$ of 0.93(0.89) terms are added and improves upon inclusion of cross ($R^2(Q^2)$ of 0.95(0.94) and both terms ($R^2(Q^2)$ of 0.97(0.92) are added.

Genetic algorithm

The values, especially for the reduced data-set, already show excellent correlation, but, especially when all the different terms are included, over hundred variables are used in the fit. It was pointed out above that the different descriptors are correlated and not all of them might be necessary to accurately capture this problem. Therefore, we have applied a genetic algorithm to reduce the number of descriptors in the regression analysis. The first benchmark for this method was whether it can reproduce the accuracy shown in the full regression analysis for CO, NO and both molecules (results shown as triangles in Fig. 8 A, B and C and in Tab S1 in the Supporting information). Generally speaking, the quality of the

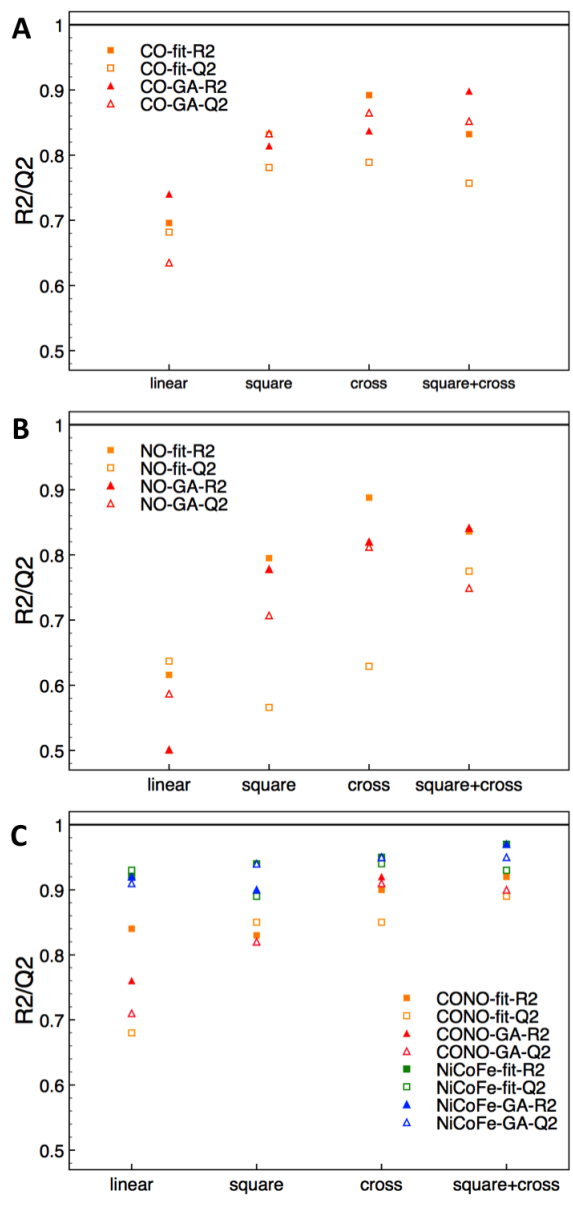


Figure 8: R^2 (filled symbols) and Q^2 (empty symbols) for regression analysis (squares) and genetic algorithm (triangles) using linear, square, cross or all terms for CO (A), NO (B) and CO and NO (C) adsorption. Correlations improve significantly, if Cu is omitted (green and blue in (C)).

fit increases when using only linear, linear+square, linear+cross and all terms. For almost all the studied fits we find excellent agreement when using a genetic algorithm compared to the full dataset. Especially when only the adsorption of one molecule is studied the fit gets very often better with Q^2 values closer to the R^2 values. The only case where differences to the full regression analysis are large, is the analysis including only linear terms for CO, where we find little correlation. However, already the full fit had a comparatively bad correlation with an R^2 of about 0.75. This deficiency is already compensated for when more variables are included, where Q^2 are higher and far closer to R^2 values compared to the full fit.

In general the quality of the fit using the genetic algorithm is very similar to the full regression analysis. However, the number of variables used in it is significantly reduced and the numeric values are given in Tab. S1 in the Supporting Information. The number of variables gets reduced to at least half and goes to a value in the mid-20s for the fits using the full dataset. However, the numbers of variables do not always follow a clear trend, i.e. not always less variables are needed when the dataset gets smaller. Such a behavior is not unexpected, since a genetic algorithm will not always find the ideal solution, but a solution that is sufficiently close to the minimum. Due to the randomness in the choice of variables also different parameters will become important.

Testing the reliability of the genetic algorithm

Especially when only Fe, Co and Ni are used and cross and square terms are included the discovered correlations are impressive. However, the fits get significantly worse when we split the dataset into too small subsets. This might be due to a significant overfitting, i.e. the inclusion of too many descriptors for the number of datapoints. Such a behavior is not surprising and it would be important to extend our screening space. However, the accuracy for the full fit comes at a high cost, i.e. the inclusion of a large number of variables. It is encouraging to see that they can successfully be reduced using a genetic algorithm. While the solution is not unique, it is able to provide a reasonably good approximation, but it is

impossible to tell how well the obtained fit agrees with a hypothetical ideal fit.

A certain ambiguity is not unexpected, since, as shown above in Fig. 3, the different variables are not linearly independent. However, we would expect that the most important ones are included in every fit. We therefore analyzed the choice of variables obtained from the genetic algorithm when different numbers of descriptors are used for the full dataset (Fig. 9 A). Of all the variables only two, namely the position of the s-up orbital as well as the volume of the active site are used in every fit. But it is important to mention that, depending on the included terms, always M_w is either included as linear and square term. In the next step of the analysis we compared the performance for the adsorption of CO, NO and both molecules (Fig. 9 B) and again, the s-orbital was used in all the fits. Now we also found the square term for the number of valence electrons of the metal and two cross terms in every fit.

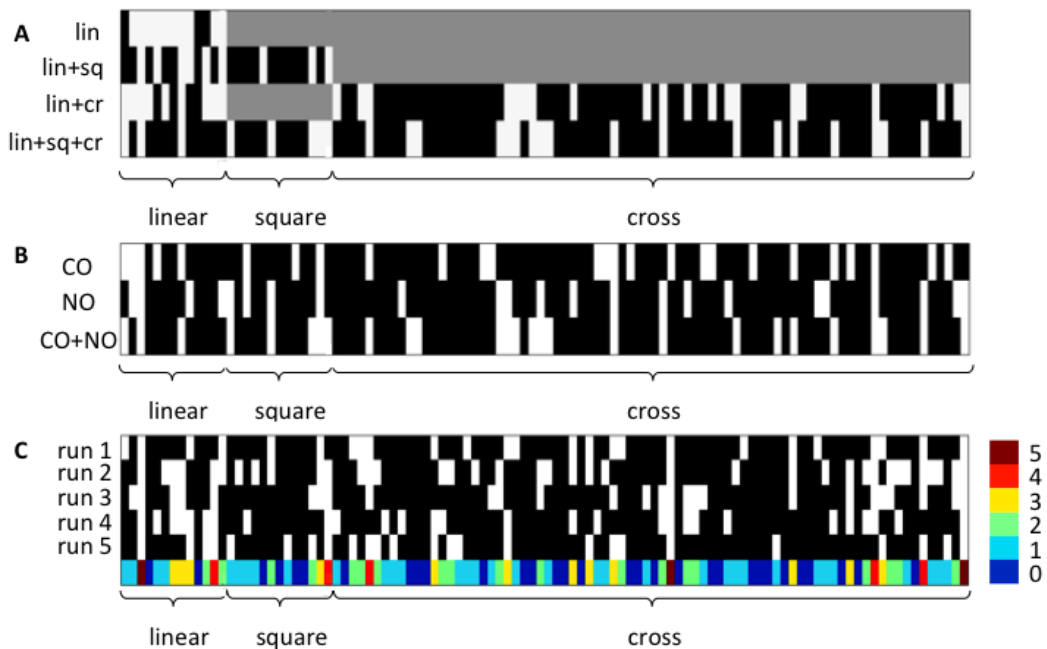


Figure 9: Analysis of descriptors used in the genetic algorithm when (A) different descriptor sets are used, (B) different molecules are used and (C) for five different runs. White lines indicate the use of the descriptor, black indicates it is not used and the colors in (C) show how often the descriptor was used in the five runs. The x-axis corresponds to the different variables used in the fit.

Another key issue when using a genetic algorithm is to test its reliability. To address this issue we performed five runs applying the genetic algorithm to the full dataset. The first thing to notice is that we obtain a certain spread in the quality of the fit. R^2 values vary significantly with values between 0.90 and 0.96, a trend that is also observed for the Q^2 values, which range between 0.84 and 0.94. Additionally the number of variables used to obtain this fit also varies between 25 and 38. In Fig. 9 C all the used variables are shown and one can see that only three variables (the position of the s-orbital, and two cross terms) appear in all five runs. However, again n_{val} and M_w are used either as linear or square terms in every fit, indicating that these three variables are of particular importance in this analysis.

The PLS analysis also identifies the latent variables, i.e. the linearly independent variables that would be necessary for the fit. Identifying them from the full dataset is hardly possible, however, it is possible to study the sum of the projections of the used parameters on them and we have displayed such a projection in section S3 of the Supporting Information for the adsorption of CO, NO, CO and NO for the full set as well as for the reduced number of metals. As can be seen, in all the fits E_{s-up} and the square of n_{val} (except when Cu is omitted) is used and as soon as both molecules are studied, the mass of the adsorbates, i.e. the factor that can differentiate between the different adsorbates, is decisive.

The impact of the reconstruction of the active site upon adsorption

As mentioned above, the coordination of the active site changes significantly upon adsorption of the molecule. This influences the adsorption in two ways, namely the adsorption energy is composed of the interaction of the molecule with the adsorption site as well as the energy necessary for reconstruction and, on top of that, the related descriptors change for the different local geometry of the active site. The approach applied above therefore tries to describe the adsorption energy by implicitly taking this reconstruction into account.

The discussion above shows that, following this strategy, a multi-parameter approach is necessary to capture the adsorption strength at a reasonable accuracy. However, it is not clear whether the large number of descriptors is needed to capture the reconstruction of the site reasonably accurately or whether to complexity of the interactions between the active site and the adsorbate require the use of many descriptors. To test for this, we also evaluated all the descriptors for the reconstructed site, i.e. a calculation without structural optimization for the adsorption geometry, where we removed the molecule. The full discussion of the impact of the reconstruction on all variables would go far beyond the scope of the main text. We therefore only discuss the main points in the main text and all descriptors and a detailed discussion are given in the supporting information section S4.

Especially for single and dual descriptors the quality of the regression analysis improves significantly. When only one metal and one descriptor are considered for CO, significant correlations with E_{s-up} , E_{s-down} and E_{stab} can be obtained. When two parameters are used, even R^2 -values larger than 0.9 can be found for combinations including the s-orbitals. For NO adsorption, single parameter approaches are still not able to accurately capture adsorption strength. This changes when two-parameter regressions are considered. Here especially combination of s-orbitals and n_{val} work well for subsets of different metals. However, when multiparameter regression or a genetial algorithm are considered, only a slight improvement can be achieved. This indicates that the complexity of the reconstruction can be captured reasonably accurately by a multi-parameter approach.

Conclusion, and Outlook

In this work we studied the adsorption of CO and NO to extra framework, transition metal sites in the zeolites SSZ-13 and MOR. After discussing a possible design-based approach for zeolite catalysts and the definition of the used test set and descriptors, we explored different possibilities for regression analysis and finally applied a genetic algorithm and arrived at an

accurate approach to predict the adsorption strength based on the introduced descriptors. By analyzing the correlation between the different descriptors, we identified a significant covalent component in the bonding of the transition metal site to the zeolite framework. In our analysis of single, two and multiple parameter regression analysis we found that the problem is highly complex and reasonable agreement can only be achieved, when multiple parameters are included in the analysis. The main challenge here is that the active site is reconstructed upon adsorption of a molecule, an effect that has to be implicitly included in the regression analysis. Correlations improve significantly when this effect is either explicitly included by studying the reconstructed sites or Cu sites, which show the largest reconstruction upon adsorption, are omitted. We finally used a genetic algorithm to reduce the number of parameters required for the fit. Indeed, the quality of the fit stays constant and analysis of the chosen variables reveals that in particular the metal s-orbital, the number of valence electrons and the HOMO-LUMO gap of the adsorbates are parameters included in every fit.

The approach presented in this work is an important step towards a design-based approach for this family of materials. However, there are several important obstacles that still have to be addressed in the future. Even though we studied more than 200 data points, it will be necessary to extend this dataset by adding data for more and different transition metals and molecules to allow for a larger generality in the future. Furthermore the calculations here are based on typical semi-local DFT calculations and it is well known that this methodology overestimates the strength of the bond between metal and adsorbate. It might be necessary to correct for this shortcoming by either calculating adsorption energies using hybrid functionals or even benchmark them against experimental measurements. In combination with an approach that allows for the prediction of the Al atoms in dependence of the synthesis process, this will allow for a reliable prediction of the activity of transition metal zeolites in the future.

Methods

DFT calculations

The aim of this paper is to develop a descriptor based approach. However, these descriptors vary with computational setup. Therefore, it is of prime importance to describe it in detail. All calculations were performed using the Vienna Ab-Initio Simulation Package (VASP),^{43,44} a plane wave code using PAW pseudo potentials.^{45,46} We used the PBE density functional⁴² in combination with latest released PBE pseudo potentials optimized for GW calculations. The energy cut-off was chosen as 420 eV, which corresponds to the minimum energy cut-off of the pseudo-potentials. Since relatively large unit cells are encountered, bonding is highly covalent and a certain degree of disorder is expected in the distribution of active sites, the calculations were restricted to the Γ point. To avoid problems with electronic convergence, a gaussian smearing of 0.02 eV was applied. To arrive at the right spin state the difference in the occupation of the different bands was kept constant, which (after careful testing) lead to a total spin of 1/2 for Cu, 1 for Ni, 3/2 for Co and 2 for Fe.

To assure comparability between the different metals, stabilities were calculated as

$$E_{stab}^{A+B} = E^{A+B} - E^A - E^{BCl_2} + E^{Cl_2}, \quad (1)$$

where A stands for the zeolite system and B for the transition metal cation. As can be seen above the Cl compounds of the divalent cations were used. As zeolite reference we used the zeolites without compensating cation. This makes the approach more generally applicable and it is independent of the countercation prior to ion exchange. Charges and atomic volumes were calculated using Bader charge analysis.⁴⁷ In a successive step, CO and NO were adsorbed to the TM sites and the adsorption strength was calculated with respect to the gas phase molecules.

For the zeolite structures we chose the most basic unit cells, which corresponds to a cell

containing 12 T-sites for SSZ-13. The basic geometry has been described elsewhere⁴⁸ and we fixed the volume of the cell to 830 Å³, a minimum volume that has been reported for SSZ-13 with one or two Al atoms in one unit cell. The unit cell of Mordenite contains 24 T-sites, and the geometry has been reported by Grybos et al. previously.³² The volume of the unit cell was chosen as 1430 Å³.

Regression analysis

Throughout the manuscript, the 11 basic descriptors that are described in detail above (E_{stab} , E_{LUMO} , E_{s-up} , E_{s-down} , ϵ_{LUMO} , ϵ_d , Q , V , ϕ_d , n , and n_{act}) were used. An additional descriptor, distinguishing between the different metals (number of valence electrons n_{val}), was used whenever we did regressions on data sets with several metals. Whenever we studied the adsorption of both CO and NO, an additional descriptor to distinguish between the two molecules was used (molecular mass of adsorbate (M_w)). For the reconstructed data sets, an additional descriptor (E_{recon}) was introduced describing the reconstruction energy. Hence, the maximum number of basic descriptors was 14. The data set consists of 11 SSZ-13 sites per metal, and 15 MOR sites per metal. Hence, the maximum data set size for CO and NO is 208 data points.

All of the calculations were performed using the statistics toolbox in Matlab.⁴⁹ Before any modeling was done, the data was autoscaled in order to yield mean-zero unit variance matrices. This ensures that any bias due to different units/scales in the variable matrices is excluded. We then applied a multi-variable regression to the data, where the linear problem can be written as follows:

$$y = \sum_{i=1}^m a_i x_i + e \tag{2}$$

In this expression y denotes the dependent variable, x_i are the descriptors, the a_i are the respective sensitivities and e is a constant offset. We increased the variable scope (linear

terms) by adding cross terms

$$\sum_{i=1}^m \sum_{j=1}^m b_{ij} x_i y_j \quad (3)$$

and square terms

$$\sum_{i=1}^m c_i x_i^2. \quad (4)$$

Here the b_{ij} and c_i represent the respective sensitivities. Four different scenarios were tested, namely (i) linear terms only, (ii) linear terms and cross terms, (iii) Linear terms and square terms, and (iv) linear terms, square terms and cross terms.

In a second step, we applied a genetic algorithm (GA), as implemented in the PLS toolbox (Eigenvector Research Inc.) in Matlab,⁵⁰ to all of these scenarios to identify the important variables. This evolutionary algorithm evaluates a model with a selection of variables in each generation. If a variable has a positive impact on the model, it is passed on to the next generation, while detrimental or useless variables get excluded. We used the genetic algorithm with partial least square regression (PLS), which is a multivariate extension of the multiple linear regression model without imposing restrictions.

To test for the reliability of our approaches, we split the data randomly into a test (two thirds of total data points) and a validation set (one third of total data points) and calculated corresponding R^2 and Q^2 values. Since our data set is still comparatively small, we performed the fitting procedure multiple times and averaged over the obtained R^2 and Q^2 values. We found reasonable convergence of these numbers after five iterations, which is what we report.

Acknowledgement

F. Göttl, P. Müller and I. Hermans acknowledge financial support from the University of Wisconsin Madison and the Wisconsin Alumni Research Foundation (WARF). P. Müller and I. Hermans acknowledge support from the Donors of the American Chemical Society

Petroleum Research Fund under Grant No. PRF 56136-ND5. The authors acknowledge computational time at Phoenix Supercomputer, which is in part supported by National Science Foundation Grant CHE-0840494 and the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Project No. m2070-Zeo-genome. This research was in part performed using the compute resources and assistance of the UW-Madison Center For High Throughput Computing (CHTC) in the Department of Computer Sciences. The CHTC is supported by UW-Madison, the Advanced Computing Initiative, the Wisconsin Alumni Research Foundation, the Wisconsin Institutes for Discovery, and the National Science Foundation, and is an active member of the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science. F. Göttl acknowledges computational time at the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. The authors furthermore acknowledge computational time at the Pôle Scientifique de Modélisation Numérique (PMSN).

Supporting Information

Supporting Information contains structural files for all geometries, numerical values for descriptors, numerical values for multiparameter regression analysis and genetic algorithm, a definition of the descriptors, the sum of the projection of descriptors on latent variables and a discussion of fitting for the reconstructed sites.

References

- (1) Ras, E.-J.; Rothenberg, G. Heterogeneous Catalyst Discovery using 21st Century Tools: A Tutorial *RSC Adv.* **2014**, *4*, 5963-5974

- (2) Calle-Vallejo, F.; Loffreda, D.; Koper, M.T.M.; Sautet, P. Introducing Structural Sensitivity Into Adsorption-Energy Scaling Relations by Means of Coordination Numbers, *Nat. Chem.* **2015**, *7*, 403-410
- (3) Ras, E.-J.; Louwse, M.J.; Mittelmeijer-Hazeleger, M.C.; Rothenberg, G.; Predicting Adsorption on Metals: Simple yet Effective Descriptors for Surface Catalysis, *Phys. Chem. Chem. Phys.* **2013**, *15*, 4436-4443
- (4) Nørskov, J.K.; Bligaard, T.; Rossmeisl, J.; Christensen, C.H. Towards the Computational Design of Solid Catalysts, *Nat. Chem.* **2009**, *1*, 37-46
- (5) Seh, Z.W.; Kibsgaard, J.; Dickens, C.F.; Chorkendorff, I.; Nørskov, J.K.; Jaramillo, T.F. Combining Theory and Experiment in Electrocatalysis: Insights Into Materials Design, *Science* **2017**, *355*, 146
- (6) Butler, K.T.; Frost, J.M.; Skelton, J.M.; Svane, K.L.; Walsh, A. Computational Materials Design of Crystalline Solids, *Chem. Soc. Rev.* **2016**, *45*, 6138
- (7) Greeley, J. Theoretical Heterogeneous Catalysis: Scaling Relationships and Computational Catalyst Design, *Anno. Rev. Chem. Biomol. Eng.* **2016**, *7*, 605-635
- (8) Calle-Vallejo, F.; Tymoczko, J.; Colic, V.; Vu, Q.H.; Pohl, M.D.; Morgenstern, K.; Loffreda, D.; Sautet, P.; Schuhmann, W.; Bandarenka, A.S.; Finding Optimal Surface Sites on Heterogeneous Catalysts by Counting Nearest Neighbors, *Science* **2015**, *350*, 185-189
- (9) Jain, A.; Ong, S.P.; Hautier, G.; Chen, W.; Richards, W.D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K.A. The Materials Project: A materials genome approach to accelerating materials innovation, *APL Materials* **2013**, *1*, 011002
- (10) Hummelshøj, J.S.; Abild-Pedersen, F.; Studt, F.; Bligaard, T.; Prof. Nørskov, J.K., Cat-

- App: A Web Application for Surface Chemistry and Heterogeneous Catalysis, *Angew. Chem. Int. Ed.* **2012**, *51*, 272-274
- (11) Davis, M.E. Ordered Porous Materials for Emerging Applications, *Nature* **2002**, *417*, 813-821
- (12) Caro, J.; Noack, M.; Kolsch, P.; Schafer, R. Zeolite Membranes - State of Their Development and Perspective, *Microp. Mesop. Mat.* **2000**, *38*, 3-24
- (13) Corma, A. State of the Art and Future Challenges of Zeolites as Catalysts, *J. Catal* **2002**, *216*, 298-312
- (14) Choi, S.; Drese, J.H.; Jones, C.W. Adsorbent Materials for Carbon Dioxide Capture from Large Anthropogenic Point Sources, *ChemSusChem* **2009**, *7*, 796-854
- (15) Brogaard, R.Y.; Wang, C.-M.; Studt, F. MethanolAlkene Reactions in Zeotype Acid Catalysts: Insights from a Descriptor-Based Approach and Microkinetic Modeling *ACS Catal.* **2014**, *4*, 4504-4509
- (16) Wang, C.-H.; Brogaard, R.Y.; Weckhuysen, B.M.; Nørskov, J.K.; Studt, F. Reactivity Descriptor in Solid Acid Catalysis: Predicting Turnover Frequencies for Propene Methylation in Zeotypes *J. Phys. Chem. Lett* **2014**, *5*, 1516-1521
- (17) Wang, C.-H.; Brogaard, R.Y.; Xie, Z.-K.; Studt, F. Transition-State Scaling Relations in Zeolite Catalysis: Influence of Framework Topology and Acid-Site Reactivity *Catal. Sci. Technol.* **2015**, *5*, 2814-2820
- (18) Sahrostami, S.; Falsig, H.; BEato, P.; Moses, P.G.; Nørskov, J.K.; Studt, F. Exploring Scaling Relations for Chemisorption Energies on Transition-Metal-Exchanged Zeolites ZSM-22 and ZSM-5 *ChemCatChem* **2016**, *8*, 767-772
- (19) Bukowski, B.C.; Greeley, J Scaling Relationships for Molecular Adsorption and Dissociation in Lewis Acid Zeolites, *J. Phys. Chem. C* **2016**, *120*, 6714-6722

- (20) Latimer, A.A; Kulkarni, A.R.; Aljama, H.; Montoya, J.H.; Yoo, J.S.; Tsai, C.; Abild-Pedersen, F.; Studt, F.; Nørskov, J.K. Understanding Trends in CH Bond Activation in Heterogeneous Catalysis *Nat. Mat.* **2017**, *16*, 225-229
- (21) Göltl, F.; Sautet, P.; Hermans, I.; Can Dynamics Be Responsible for the Complex Multippeak Infrared Spectra of NO Adsorbed to Copper(II) Sites in Zeolites? *Angew. Chem. Int. Ed.* **2015**, *54*, 7799-7804
- (22) Göltl, F.; Bulo, R.E.; Hafner, J.; Sautet, P.; What Makes Copper-Exchanged SSZ-13 Zeolite Efficient at Cleaning Car Exhaust Gases? *J. Phys. Chem. Lett.* **2013**, *4*, 2244-2249
- (23) Göeltl, F.; Hafner, J.; Structure and Properties of Metal-Exchanged Zeolites Studied Using Gradient-Corrected and Hybrid Functionals. III. Energetics and Vibrational Spectroscopy of Adsorbates, *J. Chem. Phys.* **2012**, *136*, 064503
- (24) Greeley, J.; Jaramillo, T.F.; Bonde, J.; Chorkendorff, I.; Nørskov J.K. Computational High-Throughput Screening of Electrocatalytic Materials for Hydrogen Evolution, *Nat. Mat.* **2006**, *5*, 909-913
- (25) Andrikopoulos, P.C.; Michel, C.; Chouzier, S.; Sautet, P. In Silico Screening of Iron-Oxo Catalysts fo CH Bond Cleavage, *ACS Catal.* **2015**, *5*, 2490-2499
- (26) Göltl, F.; Sautet, P.; Hermans, I.; The impact of finite temperature on the coordination of Cu cations in the zeolite SSZ-13, *Catal. Today* **2016**, *267*, 41-46
- (27) Silaghi, M.-C.; Chizallet, C.; Petracovschi, E.; Kerber, T.; Sauer, J.; Pascal Raybaud, P.; Regioselectivity of Al-O Bond Hydrolysis during Zeolites Dealumination Unified by Brønsted-Evans-Polanyi Relationship, *ACS Catalysis* **2015**, *5*, 11-15
- (28) Göltl, F.; Love, A.M.; Hermans, I. Developing a Thermodynamic Model for the Inter-

- actions between Water and Cu in the Zeolite SSZ-13, *J. Phys. Chem. C* **2017**, *121*, 6160-6169
- (29) van Santen, R.A.; Neurock, M.; Shetty, S.G. Reactivity Theory of Transition-Metal Surfaces: A Brønsted-Evans-Polanyi Linear Activation Energy-Free-Energy Analysis, *Chem. Rev.* **2010**, *110*, 2005-2048
- (30) Wichterlova, B.; Dedecek, J.; Sobalik, Z. Single Metal Ions in Host Zeolite Matrices. Structure-Activity-Selectivity-Relationships. In *Catalysis by Unique Metal Structures in Solid Matrices; From Science to Applications*; G. Centi, A.T. Bell, B. Wichterlova, Eds.; Kluwer Academic Press: Dordrecht, The Netherlands, **2001**, 31
- (31) Baerlocher, C.; McCusker, L.B.; Olson, D.H. *Atlas of Zeolite Framework Types (Sixth Edition)*, Elsevier Science, 2007
- (32) Grybos, R.; Hafner, J.; Benco, L.; Toulhoat, H. Structure of Active Sites in Pd-Exchanged Mordenite: A Density Functional Investigation, *J. Phys. Chem. C* **2007**, *111*, 6454-6464
- (33) Löwenstein, W. The Distribution of Aluminum in the Tetrahedra of Silicates and Aluminates, *Am. Miner.* **1954**, *39*, 92
- (34) Göltl, F.; Hafner, J. Structure and Properties of Metal-Exchanged Zeolites Studied Using Gradient-Corrected and Hybrid Functionals. II. Electronic Structure and Photoluminescence Spectra *J. Chem. Phys.* **136**, 064502 (2012)
- (35) Burello, E.; Rothenberg, G. In Silico Design in Homogeneous Catalysis Using Descriptor Modeling *Int. J. Mol. Sci.* **2006**, *7*, 375-404
- (36) Hageman, J.A.; Westerhuis, J.A.; Frhauf, H.-W.; Rothenberg G. Design and Assembly of Virtual Homogeneous Catalyst Libraries Towards in silico Catalyst Optimisation, *Adv. Synth. Catal.* **2006**, *348*, 361-369

- (37) Paolucci, C.; Parekh, A.A.; Khurana, I.; Di Iorio, J.R.; Li, H.; Albarracin Caballero, J.D.; Shih, A.J.; Anggara, T.; Delgass, W.N.; Miller, J.T.; Ribeiro, F.H.; Gounder, R.; Schneider W.F.; Catalysis in a Cage: Condition-Dependent Speciation and Dynamics of Exchanged Cu Cations in SSZ-13 Zeolites, *J. Am. Chem. Soc.* **2016**, *138*, 6028-6048
- (38) Hammer, B.; Nørskov J.K. Theoretical Surface Science and Catalysis - Calculations and Concepts, *Adv. Catal.* **2000**, *45*, 71-129
- (39) Occhipinti, G.; Bjørsvik, H.-R.; Jensen, V.R. Quantitative Structure-Activity Relationships of Ruthenium Catalysts for Olefin Metathesis, *J. Am. Chem. Soc.* **2006**, *128*, 6951-6964
- (40) Notario-Estevez, A.; Kozlov, S.M.; Viñes, F.; Illas, F. Electronic-Structure-Based Material Descriptors: (In)Dependence on Self-Interaction and HartreeFock Exchange, *Chem. Commun.* **51**, 5602-5605
- (41) Blyholder, G.; Molecular Orbital View of Chemisorbed Carbon Monoxide *J. Phys. Chem.* **1964**, *68*, 2772
- (42) Perdew J.P., Burke K., Ernzerhof M., Generalized Gradient Approximation Made Simple *Phys. Rev. Lett.* **1996**, *77*, 3865
- (43) Kresse, G.; Hafner, J.; Ab Initio Molecular Dynamics for Open-Shell Transition Metals; *Phys. Rev. B* **1993** *48*, 13115
- (44) Kresse, G.; Furthmüller, J; Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set, *Comput. Mat. Sci.* **1996**, *6*, 15-50
- (45) Blöchl P.E.; Projector Augmented-Wave Method; *Phys. Rev. B* **1994**, *50*, 17953
- (46) Kresse G.; Joubert, D.; From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method; *Phys. Rev. B* **1999**, *59*, 1758

- (47) Yu, M.; Trinkle, D.R. Accurate and Efficient Algorithm for Bader Charge Integration, *J. Chem. Phys.* **2011**, *134*, 064111
- (48) Göttl, F.; Hafner, J. Structure and Properties of Metal Exchanged Zeolites Studied Using Gradient-Corrected and Hybrid Functionals. I. Structure and Energetics, *J. Chem. Phys.* **2012**, *136*, 064501
- (49) MATLAB 2014a, Statistics and Machine Learning Toolbox, The MathWorks, Natick, 2014
- (50) PLS Toolbox v4.2 for Matlab, Eigenvector Research Inc.

