CO Oxidation Catalyzed by Perovskites: The Role of Crystallographic Distortions Highlighted by Systematic Experiments and AI

Giulia Bellini,[a] Frank Girgsdies,[a] Gregor Koch,[a] Spencer J. Carey,[a] Olaf Timpe, [a]

Gudrun Auffermann,[b] Matthias Scheffler,[c] Robert Schlögl,[a] Lucas Foppa,*[c] and Annette Trunschke*[a]

---

[a]   Dr. Giulia Bellini, Dr. Frank Girgsdies, Dr. Gregor Koch, Dr. Spencer J. Carey, Dr. Olaf Timpe, Dr. Gudrun Auffermann, Prof. Dr. Robert Schlögl, Dr. Annette Trunschke*
Inorganic Chemistry Department
Fritz-Haber-Institut der Max-Planck-Gesellschaft
Faradayweg 4-6, 14195 Berlin, Germany
E-mail: trunschke@fhi-berlin.mpg.de

[b]   Dr. Gudrun Auffermann
Max-Planck-Institut für Chemische Physik Fester Stoffe
Nöthnitzer Straße 40, 01187 Dresden, Germany

[c]   Prof. Dr. Matthias Scheffler, Dr. Lucas Foppa*
The NOMAD Laboratory at the Fritz-Haber-Institut of the Max-Planck-Gesellschaft and IRIS-Adlershof of the Humboldt-Universität zu Berlin
Faradayweg 4-6, 14195 Berlin, Germany
E-mail: foppa@fhi-berlin.mpg.de

Abstract: The identification of key parameters that correlate with catalytic performance through a combination of experiments and model calculations can accelerate the development of improved catalysts and reveal the relevant underlying processes. However, the analysis of correlations in heterogeneous catalysis is often hindered by inconsistent data. Besides, nontrivial, yet unknown relationships may be important, and the intricacy of the various processes may be significant. Here, we address these challenges for perovskite-catalyzed CO oxidation by linking systematic experiments and artificial intelligence (AI). For this purpose, 14 $AMn_{(1-x)}Cu_xO_3$ phase-pure perovskites with A = Pr, La were synthesized, characterized, and tested according to rigorous experiments. To the so-generated consistent dataset, we applied the symbolic-regression SISSO approach and identified a descriptor for CO consumption rates as an analytical expression containing the bulk and surface A content, the copper surface content, and the deviation (D) of the normalized lattice constants from the cubic root of the normalized cell volume. Crucially, D reflects a crystallographic distortion that depends on the element at the A site. Thus, in addition to the relative abundance of redox-active species on the surface, the species at the A-sites also influence the performance by modulating the properties of the material and the surface.

Introduction

Perovskites[1] are a promising materials class[2] for energy-related application and key to future catalysis technologies.[2b] Mixed oxides perovskites have been applied, for instance, to environmental catalysis[3] and to electrocatalysis.[2b] Due to the high degree of flexibility of the perovskite structure with respect to the composition, an immense number of compounds can be obtained by combining different elements and by varying the stoichiometry. Double perovskite oxides with the formula $AB_{(1-x)}B'_xO_3$ can accommodate different A, B, and B' elements at different proportions. In this perovskite structure, one large A cation (e.g., lanthanide) occupies a site that is coordinated with 12 oxygen anions (O^{2-}) and the smaller B and B' cations (e.g., transition metals) are coordinated with 6 oxygen anions forming BO_6 (or $B'O_6$) octahedra. The properties and catalytic performance of these materials can be adjusted by the composition. However, elucidating the underlying mechanisms and efficiently exploring the practically infinite space of possible perovskites are extremely challenging tasks. This is because heterogeneous catalysis is governed by a high intricacy of many underlying processes such as surface reactions, the material's dynamic restructuring, and the interplay of bulk and surface properties. This prevents the explicit, atomistic modelling of the full catalytic progression.[4] Indeed, the surface motifs that are present under reaction conditions and the reaction mechanisms that operate on those surfaces are typically unknown and continuously adapt to the progress of the reaction and the associated changes in chemical potential.

The CO oxidation to CO_2 by molecular oxygen is an example of an important reaction for environmental catalysis and for the purification of hydrogen for energy generation. This reaction is catalyzed by perovskite oxides.[3, 5] Despite being one of the most investigated heterogeneous reactions,[6] the mechanisms governing CO oxidation on perovskites are not well understood. The performance of perovskites in oxidation catalysis has been related to different underlying processes such as the availability and redox potential of surface B sites and the materials' capacity

to provide lattice oxygen. In particular, series of double perovskites containing lanthanides at the A site and different proportions of manganese and a second metal at the B and B' sites, respectively, (e.g., $LaMn_{(1-x)}Cu_xO_3$) have been studied in detail to probe the influence of the composition on the catalytic performance.[7] Some studies report that the CO consumption rates are proportional[7k, 7m] or present a volcano-type relationship[7b, 7c] with respect to the content of copper in $LaMn_{(1-x)}Cu_xO_3$. Accordingly, it has been suggested that the function of manganese is the activation of molecular oxygen and that of copper the adsorption of carbon monoxide.[7c] The influence of the A element has been investigated for cobalt-based materials with the general formula $ACoO_3$, where A is a lanthanide.[8] The activity was related to spin and valence states of cobalt as well as to the reducibility and to the strength of oxygen binding on the surface. The A sites influence the bulk and surface properties of perovskites, for instance by inducing crystallographic strain,[9] i.e., when the structure deviates from the ideal cubic perovskite structure. This also results in changes in the surface structure and electronic properties that might affect catalysis. In particular, small A cations induce octahedral tilting, creating more covalent bonding structures compared to those obtained with larger A cations.[10] Indeed, we recently demonstrated this phenomenon for the case of perovskites $AMn_{(1-x)}Cu_xO_3$ with A = La, Pr.[11] However, the effect of crystallographic distortions, i.e., distortions in the bulk, not only at the surface, has not yet been investigated in CO oxidation catalyzed by double perovskite oxides. Furthermore, the correlations and key parameters proposed in the literature are partly contradictory. In particular, the conclusions depend on the composition of the respective catalyst series investigated and do not apply to CO oxidation over perovskites in general.

In order to unveil the underlying mechanisms and to design new materials, the identification of key parameters correlated with the catalytic performance is crucial. Artificial intelligence (AI) has been increasingly applied to identify nonlinear correlations and complex patterns in data in materials science and catalysis.[12] However, the inconsistency of experimental data in heterogeneous-catalysis research hinders the application of AI to reported data (see ref. [13] and references therein). Indeed, the measured materials properties, e.g., surface compositions, depend on the experimental workflow of sample synthesis, the pretreatment, and the measurement conditions. These workflows often involve different synthetic procedures and multi-step treatments such as calcination or annealing, which are performed before the catalytic reaction is carried out. Besides, the measured catalytic performance is strongly influenced by the reaction conditions, e.g., start-up procedure (induction period under the reaction feed also referred to as catalyst formation or activation), temperature, contact time, or feed composition. This is because the surface and solid-state chemistry of the material is coupled with the chemistry of the catalytic reaction. Finally, the presence of small amounts of by-phases formed as minor impurities during synthesis can also lead to inconsistencies when searching for structure-performance relationships. Such minor amounts of additional catalyst components can significantly influence the performance. They are often only detectable by electron microscopic or other methods of spatial resolved or trace analysis

and might be overlooked by (bulk) techniques for materials' characterization such as XRD and therefore not taken into account. In addition to data inconsistency, the extremely small number of materials that can be accessed by detailed, consistent experiments prevents the application of widely used AI methods such as neural networks. These AI methods typically require orders of magnitude more data points (e.g., >10^4) compared to the number of materials that can be studied by consistent experiments (e.g., <10^2).

To address these issues, we recently proposed a combination of rigorous experimental protocols, which produce consistent data,[14] with the data-efficient symbolic-regression AI approach SISSO (sure-independence screening and sparsifying operator).[15] This strategy enables the identification of the key physicochemical parameters correlated with the performance, out of many candidate parameters, also referred to as primary features, characterizing the materials and possibly relevant underlying processes. In analogy to genes in biology, these key parameters were called "materials genes", since they describe the materials behavior similar to how many genes in biology and medicine describe characteristics such as hair color, i.e., they capture complex patterns without providing the full, explicit description of all underlying processes and their interplay. Furthermore, the analytical expressions identified by SISSO, hereafter referred to as descriptors, and AI in general provide a statistical description, i.e., they provide mean values with a certain variance or uncertainty. Importantly, the SISSO approach can be applied even to the small number of materials that can be accessed by consistent experiments.

In this paper, we utilize the materials-genes concept to model and design perovskite catalysts, using the CO oxidation as an example. In particular, we consider double perovskite oxides of the type $AMn_{(1-x)}Cu_xO_3$ (A = La, Pr). By simultaneously analyzing different copper contents and A-site elements with different sizes, we capture the influence of both copper substitutions and A element on the performance of the perovskites. We identify a descriptor for the measured reaction rates in the form of an analytical expression, which enables the prediction of the activity of materials and compositions that were not included in the training set.

Results and Discussion

Synthesis, characterization and catalytic testing of the perovskite materials

We synthesized 14 $AMn_{(1-x)}Cu_xO_3$ (A = La, Pr, $0 \leq x \leq 0.4$) phase-pure materials by solution combustion synthesis,[16] and measured 30 physicochemical parameters per material by the characterization techniques inductively coupled plasma optical emission spectroscopy (ICP OES), oxygen analysis, x-ray diffraction (XRD), x-ray photoelectron spectroscopy (XPS) and N_2 adsorption. The full list of the 30 parameters used is provided in Table S1. Raw data of CO oxidation experiments are shown in Figure S1. Further details on the preparation procedure, characterization techniques, and catalyst testing are described in the Supporting Information. All numerical physicochemical

catalyst characterization results and the CO oxidation rates are given electronically in the Supporting Information in the file dataset.xls.

X-ray diffraction reveals that all samples present a perovskite-like phase-pure structure (Table 1). The orthorhombic crystal structure is observed for the majority of the synthesized materials. Only the lanthanum-based perovskites with low copper content exhibit the rhombohedral crystal symmetry. In order to quantify the distortions of the perovskites with respect to the ideal cubic structure, we determined the observed tolerance factor t_{obs} based on the Rietveld refinement analysis and on the expression

$$t_{obs} = \frac{<AO>_\mu}{<BO>_\mu * \sqrt{2}} \quad . \qquad \text{(Eq. 1)}$$

Here, $<AO>_\mu$ and $<BO>_\mu$ are the mean values of the interatomic A-O and B-O distances, respectively. The observed tolerance factor for a cubic perovskite is equal to the unity. For the synthesized materials, the values of t_{obs} are between 0.984 and 0.993, where the values decrease with increasing copper amount (see dataset.xls in the Supporting Information). Besides, t_{obs} values are lower for the praseodymium-based materials compared to the lanthanum-based ones. This indicates that the smaller size of praseodymium induces higher octahedral tilting and higher crystallographic strain. Three additional parameters calculated by the Rietveld refinement analysis express the crystallographic distortion: the deviation (D) of the normalized lattice constants from the cubic root of the normalized cell volume, the deviation of single A-O distance values from the mean ($<AO>_\sigma$), and deviation of single B-O distance values from the mean. ($<BO>_\sigma$). D quantifies the geometric distortion of the unit cell from the ideal cubic state (Table 1). It is defined so that distortions in both orthorhombic and rhombohedral crystal symmetries can be expressed by the same quantity:

$$D = \sqrt{\frac{(a'-a^*)^2+(b'-a^*)^2+(c'-a^*)^2}{3}} \quad . \qquad \text{(Eq. 2)}$$

In Eq. 2, normalized lattice parameters are used that depend on the crystal symmetry. For the orthorhombic crystal structure in setting *Pbnm*, we use:

$$a' = \frac{a}{\sqrt{2}} \text{ (Eq. 3)}, \ b' = \frac{b}{\sqrt{2}} \text{ (Eq. 4)}, \ c' = \frac{c}{2} \text{ (Eq. 5)}.$$

For the rhombohedral structure in a hexagonal coordinate system (*R*-3*c*H), we use:

$$a' = \frac{a}{\sqrt{2}} \text{ (Eq. 6) and } c' = \frac{c}{\sqrt{12}} \text{ (Eq. 7)}.$$

The parameter *a** is the pseudo-cubic lattice parameter defined by the formula

$$a^* = \sqrt[3]{V'} \quad \text{(Eq. 8)}.$$

The parameter *a** represents the geometric mean of the normalized lattice parameters but may also be regarded as the cell parameter *a* of a virtual cubic cell with the same normalized

volume V' which is the actual unit cell volume divided by the number of ABO_3 formula units:

$$V' = \frac{V}{Z_{ABO_3}} \quad \text{(Eq. 9)}.$$

The parameters $<AO>_\sigma$ and $<BO>_\sigma$ reflect cuboctahedral, and octahedral/Jahn-Teller distortions, respectively. In addition to the mentioned quantities, we obtained other parameters from the XRD analysis such as the normalized unit cell volume. In total, 13 parameters characterizing the perovskite structures and their distortions were obtained from XRD (Table S1 and dataset.xls). A more detailed discussion of complex crystallographic characteristics of the materials studied here is provided elsewhere.[11]

The ICP OES and oxygen analysis on the one hand, and XPS on the other hand were used to measure the composition of the bulk x^{bulk} and surface x^{surf} of the perovskites, respectively. In order to quantify the relative amount of redox-active species (*B* and *B'*) with respect to the atoms of element *A* in the structure, we determined the ratio between the sum of atomic fractions of *B* and *B'* (x_B and $x_{B'}$, respectively) and the atomic fraction of *A* (x_A) as

$$x_{B/A} = \frac{x_B + x_{B'}}{x_A}. \text{ (Eq. 10)}$$

The ratio for bulk atomic fractions (denoted $x_{B/A}^{\text{bulk}}$) is higher for lanthanum-based samples compared to the praseodymium-based ones (Table 1). In contrast, the ratio for surface atomic fractions (denoted $x_{B/A}^{\text{surf}}$) is higher for the praseodymium series compared to the lanthanum series. This can be attributed to the enhanced crystallographic distortion of praseodymium-based materials, which hinders the homogeneous embedding of the two *B* sites into the octahedral of the bulk in the praseodymium-based perovskites. However, except for the concentration gradients between surface and bulk (Table 1), the materials are homogeneous. Only in the praseodymium-based material with the highest copper content, $PrMn_{0.6}Cu_{0.4}O_3$, small amounts of exsolved CuO nanoparticles were detected in a minority of cases by scanning-transmission-electron-microscopy energy-dispersive x-ray analysis. [11] Thus, the x^{surf} values correspond to the average of different surface terminations in the near-surface region that is detected by laboratory XPS and they also include the contribution of CuO, in the case of the material $PrMn_{0.6}Cu_{0.4}O_3$. In addition to x_B^{surf}, $x_{B'}^{\text{surf}}$, and x_A^{surf}, the atomic fraction of oxygen on the surface x_O^{surf}, and the onset of the valance band relative to the Fermi level were also measured by XPS (see dataset.xls in the Supporting Information). These parameters reflect the surface and electronic properties of the materials. We also analyzed the oxygen 1*s* spectra measured by laboratory XPS and determined the relative amount of different oxygen species such as regular lattice oxygen, oxygen located near to defects, surface oxygen, oxygen in adsorbed carbonates, hydroxyl groups, water and adsorbed oxygen-containing species (Table S2, Figures S2 and S3, and dataset.xls). It should be mentioned that the above-mentioned segregation of CuO in the Pr-based material with the highest Cu content seems to affect the O 1s spectrum. In total, chemical analyses (ICP OES and oxygen analysis) and XPS provided 4 and 13 parameters, respectively (Table S1). Finally, the N_2 physisorption provided the specific surface area of the

perovskites. The values of surface area vary between 2.5 and 12 $m^2 \cdot g^{-1}$ (see dataset.xls file in the Supporting Information).

The perovskites were tested in CO oxidation at 130 °C under steady state in a plug flow reactor (Figure S1 and measured rates in the file dataset.xls in the Supporting Information). We compared the reactivity of the tested materials by analyzing the CO consumption rates $r_{CO}^{measured}$, in $mmol \cdot h^{-1} \cdot g^{-1}$. In particular, we plot the reaction rates as function of x in AMn$_{(1-x)}$Cu$_x$O$_3$ (A = La, Pr) in Figure 1A. The measured rate values range from 74.2 mmol·g^{-1}·h^{-1} to 349 mmol·g^{-1}·h^{-1}. Thus, the considered perovskites present significantly different reactivity. The less and most active materials, among all tested ones, are PrMn$_{0.75}$Cu$_{0.25}$O$_3$ and LaMn$_{0.60}$Cu$_{0.40}$O$_3$, respectively. The lanthanum-based materials achieve higher rates compared to the praseodymium-based ones. In particular, the lanthanum-based perovskites with high x display the highest rates for CO consumption. For the lanthanum-based materials, the rates are in general directly proportional to x. This behavior is in line with previous reports for perovskites containing manganese and copper as B and B' elements.[7k, 7m] However, this behavior is not observed for the praseodymium series. Indeed, the rates are inversely proportional to x for the praseodymium-based materials. This shows that the nominal stoichiometry is not enough to capture the reactivity of all the considered materials.

We note that the materials LaMn$_{0.65}$Cu$_{0.35}$O$_3$, PrMn$_{0.75}$Cu$_{0.25}$O$_3$, and PrMn$_{0.60}$Cu$_{0.40}$O$_3$ deviate from the linear trend, which indicates that x might not be an appropriate descriptor. We have also investigated the evolution of the reaction rates with respect to the relative copper surface atomic fraction ($x_{Cu,rel}^{surf}$), as the relative copper content has been suggested to correlate with the activity of LaMn$_{(1-x)}$Cu$_x$O$_3$ perovskites towards CO oxidation.[7b, 7c, 7k, 7m] The value $x_{Cu,rel}^{surf}$ is defined by

$$x_{Cu,rel}^{surf} = \frac{x_{Cu}^{surf}}{x_A^{surf} + x_{Cu}^{surf} + x_{Mn}^{surf}}. \text{ (Eq. 11)}$$

In Eq. 11, x_A^{surf}, x_{Cu}^{surf}, and x_{Mn}^{surf} are the surface atomic fraction of A, copper, and manganese with respect to all metals present on the surface. The reaction rates are directly and inversely proportional to $x_{Cu,rel}^{surf}$ for the lanthanum- and praseodymium-based materials, respectively (Figure 1B). Thus, the trends observed in Figure 1A also emerge when the influence of the relative copper surface content on the rates is analyzed. Therefore, nor the degree of exchange of manganese by copper (x) neither the relative copper surface atomic fraction is able to describe the reactivity of the double perovskites across different A elements and proportions of B and B' elements simultaneously.

Table 1. Summary of the properties of perovskite materials with general formula AMn$_{(1-x)}$Cu$_x$O$_3$ (A = La, Pr, Sm, 0 ≤ x ≤ 0.4) measured by XRD, ICP OES, XPS, and N$_2$ physisorption.

A	x	Space group	D	$x_{B/A}^{bulk}$	$x_{B/A}^{surf}$	x_A^{bulk} (at%)[c]	x_A^{surf} (at%)[c]	x_{Cu}^{surf} (at%)[c]	Material ID[d]
La	0	R-3cH [a]	0.0203	1.02	0.76	0.20	0.23	0.00	S30649
La	0.10	R-3cH	0.0252	0.97	0.77	0.19	0.23	0.03	S30867
La	0.20	R-3cH/$Pbnm$	0.0242	1.00	0.84	0.20	0.20	0.04	S31180
La	0.25	$Pbnm$[b]	0.0126	1.06	0.83	0.19	0.21	0.06	S30635
La	0.30	$Pbnm$	0.0128	1.07	0.82	0.20	0.22	0.07	S30659
La	0.35	$Pbnm$	0.0114	1.02	0.71	0.19	0.25	0.07	S30624
La	0.40	$Pbnm$	0.0116	1.01	1.03	0.20	0.21	0.12	S31285
Pr	0	$Pbnm$[b]	0.018	0.95	1.19	0.20	0.18	0.00	S31345
Pr	0.10	$Pbnm$	0.0266	0.95	1.13	0.20	0.20	0.03	S31070
Pr	0.20	$Pbnm$	0.0324	0.96	1.11	0.20	0.21	0.06	S31021
Pr	0.25	$Pbnm$	0.0311	0.98	0.92	0.21	0.23	0.07	S30637
Pr	0.30	$Pbnm$	0.0309	0.95	0.91	0.21	0.23	0.07	S30934
Pr	0.35	$Pbnm$	0.0273	0.98	0.78	0.21	0.25	0.08	S31176
Pr	0.40	$Pbnm$	0.0268	1.01	0.64	0.18	0.25	0.08	S29908
Sm	0	$Pbnm$	0.167	1.00	0.94	0.18	0.18	0.00	S30103

[a]R-3cH : rhombohedral symmetry. [b] $Pbnm$: orthorhombic symmetry. [c] Percentage of atoms. [d] Samples ID according to the internal database of Fritz Haber Institute for unique identification of the sample batches.
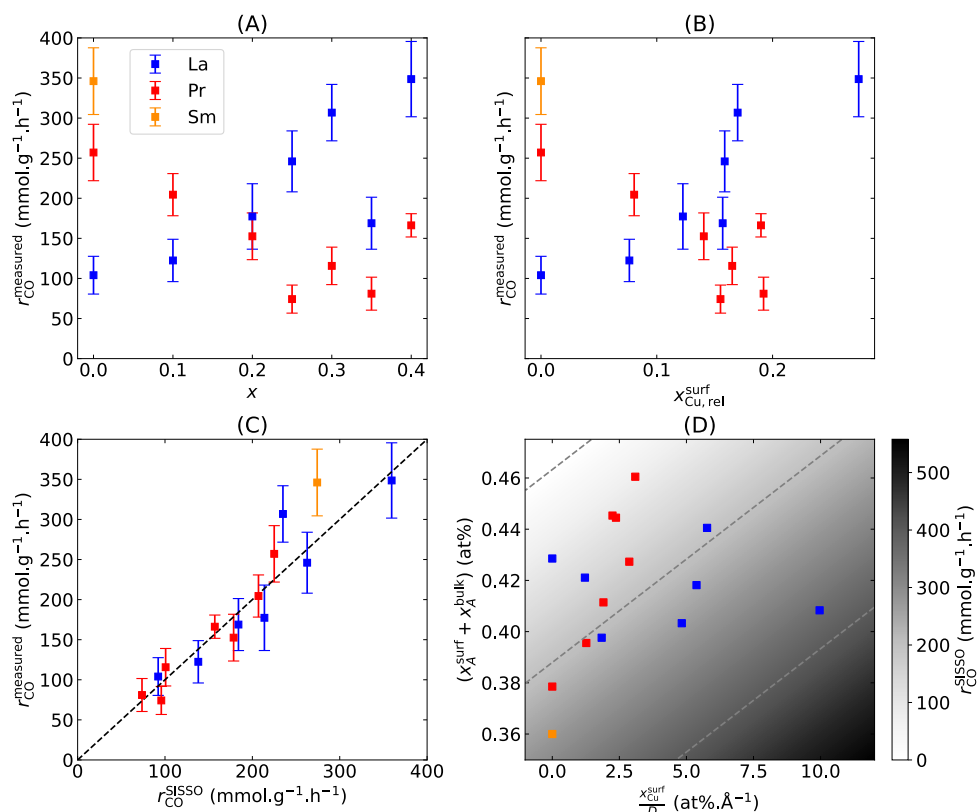
Figure 1. (A) CO consumption vs. the degree of exchange of manganese by copper (x); (B) CO consumption rate vs. the copper relative surface fraction ($x_{Cu,rel}^{surf}$); (C) Correlation between the model identified by the SISSO (r_{CO}^{SISSO}) and the CO consumption rates; (D) Map of potential catalysts created using the SISSO analysis; The SISSO model correctly captures the reactivity of the SmMnO$_3$ material, which was not used for training; Red, blue, and orange markers show results for lanthanum, praseodymium, and samarium-based materials, respectively; CO oxidation was carried out at 130 °C, a W/F = 0.036 g·s·mL^{-1}, where W is the mass of the catalyst in g and F is the total flow rate in mL·h^{-1}, with a feed CO/O$_2$/inert: 1/20/79; The dashed lines in D indicate the isovalue curves associated to 0, 200 and 400 mmol·g^{-1}·h^{-1}.

AI analysis

In order to identify a descriptor that captures the reactivity towards CO oxidation for both lanthanum- and praseodymium-based perovskites, we apply the SISSO approach.[17] This method identifies models describing a target of interest T (e.g., reaction rates) as analytical expressions depending on key parameters, out of many offered ones. The SISSO approach starts with the collection of its input parameters, termed primary features. Then, a large number (e.g., up to millions) of candidate analytic expressions are generated by combining the primary features according to (nonlinear) mathematical operators such as addition, multiplication, division, and exponential. Finally, compressed sensing is applied in order to identify the few expressions, which combined by means of weighing coefficients, best correlate with the target for the dataset. The models identified by SISSO have the form

$$T^{(SISSO)} = c_0 + \sum_{i=1}^{N} c_i d_i \quad , \text{(Eq. 12)}$$

where c_i are fitted coefficients and d_i are the selected analytical expressions, which correspond to components of a descriptor vector. Because the functional forms of SISSO models are flexible and can easily fit the rather small number of data points generated

in this work, the appropriate model complexity needs to be determined in order to avoid models that fit the provided data but are not generalizable. We evaluate the optimal complexity with respect to the predictability through a leave-one-material-out-cross-validation (CV) approach (method described in the Supporting Information, Figure S4).[15a] The 30 parameters obtained from XRD, ICP OES, oxygen analysis, XPS, and N$_2$ physisorption were used as primary features in the SISSO analysis (Table S1, dataset.xls file in the Supporting Information). The best model identified by SISSO for the CO oxidation steady-state rates, corresponding to the model obtained using the entire dataset at the optimal complexity, correctly describes the reactivity trend across both lanthanum and praseodymium series of perovskites (Figure 1C). This model has the following expression:

$$r_{CO}^{(SISSO)} = 1230 + 21.43\left(\frac{x_{Cu}^{surf}}{D}\right) - 2655\left(x_A^{surf} + x_A^{bulk}\right). \text{ (Eq. 13)}$$

The key parameters that appear in the model expression of Eq. 13 are the surface copper content (x_{Cu}^{surf}) determined by XPS, the deviation (D) of the normalized lattice parameters from the cubic root of the normalized unit cell volume determined by XRD, and the concentration of A in the bulk and on the surface of the

materials (x_A^{bulk} and x_A^{surf}, respectively), determined by ICP OES and XPS, respectively (Table 1). The relevance of $x_{\text{Cu}}^{\text{surf}}$ can be related to the function of copper as an adsorption center for CO, as suggested by previous works.[7c] The parameter D, in turn, points at a crucial role of bulk crystallographic distortions in the reactivity. Indeed, octahedral tilting induced by different compositions[10] can alter the electronic structure of the perovskites and thus the bonding patterns, [9] in particular the balance between ionic and covalent bonding contributions. These different bulk bonding patterns consequently translate into different surface reactivity. Noteworthy, SISSO selects D as a key parameter rather than t_{obs} or the mean values and standard deviations of interatomic distances (e.g,. metal-oxygen or oxygen-oxygen) and B-O-B angles, which were also offered as candidate parameters. Finally, x_A^{bulk} and x_A^{surf} reflect the influence of bulk and surface composition in the catalytic performance, in particular the relative A content. It should be noted that the surface compositions measured by XPS are not independent, since $x_A^{\text{surf}} + x_{\text{Cu}}^{\text{surf}} + x_{\text{Mn}}^{\text{surf}} + x_O^{\text{surf}} = 1$. Thus, the lanthanum or praseodymium surface atomic contents are indirectly related to the atomic contents of copper, manganese, and oxygen. The surface oxygen contents or the relative amount of oxygen species measured by XPS, which were also offered as parameters in our analysis, were not selected by SISSO as important parameters. This may imply that oxygen activation is no limiting factor under the given reaction conditions in an oxygen-rich feed. The total surface area measured by N_2 physisorption was not selected by SISSO either. Overall, the identified materials genes suggest that a combination of the abundance of certain adsorption centers with their local electronic structure is the key for CO oxidation on the perovskites. These insights agree with the previous reports in the sense that complex relationships are involved.[7c, 7g, 7k, 7l] However, Eq. 13 highlights that parameters characterizing crystallographic distortions such as D are required to describe CO oxidation rates across different A elements, in particular to simultaneously capture the lanthanum- and praesodymium-based materials, which experimentally show different trends. Interestingly, the rate decreases with increasing crystallographic strain, which again underlines that oxygen activation is less important than the activation of CO on metal cations at the surface. The crucial impact of the mentioned distortions in the CO oxidation activity has been overlooked by previous works, but it was identified here via systematic experiments and AI. Based on the model identified by SISSO, we construct a map of potential catalyst (Figure 1D). This map shows how the reaction rates are predicted (grey color scale) based on the two descriptor components, i.e., the two terms containing analytical expressions in Eq. 13.

In order to assess the generalizability of the descriptor identified by SISSO, we synthesized, characterized and tested the $SmMnO_3$ material, which contains a different A element (samarium) with respect to those employed to train the model. This catalyst (orange markers in Figure 1) was subjected to the same characterization methods and catalytic experiment applied to the (La,Pr)Mn$_{(1-x)}$Cu$_x$O$_3$ series for the rate determination. The rate of $SmMnO_3$ (346 mmol·g^{-1}·h^{-1}) is correctly predicted to be among the highest ones (orange markers in Figures 1C and 1D). This result suggests that the identified descriptor is valid beyond the compositional space used for its generation. Indeed, the

model of Eq. 13 is expected to be valid for new materials containing different elements in their compositions as long as the reactivity of the examined compounds is governed by the same processes governing the catalytic performance in the perovskites used to train the model. Even though the candidate descriptive parameters included in this work only reflect the thermodynamic properties of the catalysts, the obtained SISSO models describe the reaction rates obtained at fixed reaction conditions with good accuracy. The inclusion of parameters that correlate with the catalyst dynamics and with the reaction environment, i.e., parameters determined in operando experiments, might be crucial for the description of reactions and materials over multiple reaction conditions.[15a, 15b]

Conclusion

In summary, we synthesized, characterized, and tested in CO oxidation 14 phase-pure AMn$_{(1-x)}$Cu$_x$O$_3$ perovskites (A = La, Pr) according to a systematic experimental procedure. The activity of AMn$_{(1-x)}$Cu$_x$O$_3$ perovskites towards the catalytic CO oxidation cannot be easily correlated with the copper content, but it is captured by the SISSO AI approach. By applying a SISSO analysis, we identified a descriptor for CO consumption rates that contains the copper surface content, the deviation of the normalized lattice parameters from the cubic root of the normalized unit cell volume, and the concentration of the A in the bulk and on the surface of the materials as key parameters. The descriptor is of more general validity than previously proposed ones and it allows predictions also for materials that do not contain copper. Crucially, this descriptor highlights the so far overlooked role of crystallographic strain, which depends on the A element. This example clearly shows that the relationships between the material properties of perovskites and catalysis are of a complex nature and that an optimal catalyst cannot be found simply by varying the composition or a crystallographic parameter. The approach described in this paper accelerates the development of improved perovskite materials for catalysis by providing insights into the key processes underlying catalysis and thus proposing approaches to the inorganic chemist or materials scientist that are not readily apparent.

Supporting Information

The Supporting Information contains the list of physicochemical parameters used in the AI analysis, details on the materials synthesis, characterization, and CO oxidation experiments as well as details on the cross-validation strategy. Additional references are cited within the Supporting Information.[16, 18] The values of all physicochemical parameters used in the AI analysis for the 14 materials considered in our study are provided in the file dataset.xls.

Acknowledgements

Keywords: perovskite • CO oxidation • heterogeneous catalysis • artificial intelligence • symbolic regression

[1] a) V. M. Goldschmidt, *Naturwissenschaften* **1926**, *14*, 477-485; b) F. Fujishiro, C. Sasaoka, T. Ina, T. Sakuragi, M. Oishi, *J. Phys. Chem. C.* **2021**, *125*, 13283-13290; c) S. A. Oskoui, A. Niaei, H.-H. Tseng, D. Salari, B. Izadkhah, S. A. Hosseini, *ACS Comb. Sci.* **2013**, *15*, 609-621.

[2] a) J. Zhu, H. Li, L. Zhong, P. Xiao, X. Xu, X. Yang, Z. Zhao, J. Li, *ACS Catal.* **2014**, *4*, 2917-2940; b) J. Hwang, R. R. Rao, L. Giordano, Y. Katayama, Y. Yu, Y. Shao-Horn, *Science* **2017**, *358*, 751-756; c) M. A. Peña, J. L. G. Fierro, *Chem. Rev.* **2001**, *101*, 1981-2018.

[3] S. Royer, D. Duprez, *ChemCatChem* **2011**, *3*, 24-65.

[4] A. Bruix, J. T. Margraf, M. Andersen, K. Reuter, *Nat. Catal.* **2019**, *2*, 659-670.

[5] H. Tanaka, M. Misono, *Curr. Opin. Solid St.* **2001**, *5*, 381-387.

[6] H.-J. Freund, G. Meijer, M. Scheffler, R. Schlögl, M. Wolf, *Angew. Chem. Int. Ed.* **2011**, *50*, 10064-10094.

[7] a) B. C. Tofield, W. R. Scott, *J. Solid State Chem.* **1974**, *10*, 183-194; b) N. Mizuno, Y. Fujiwara, M. Misono, *Chem. Comm.* **1989**, 316-318; c) H. Yasuda, Y. Fujiwara, N. Mizuno, M. Misono, *Faraday Trans.* **1994**, *90*, 1183-1189; d) K. S. Chan, J. Ma, S. Jaenicke, G. K. Chuah, J. Y. Lee, *Appl. Cat. A.* **1994**, *107*, 201-227; e) J. Töpfer, J. B. Goodenough, *J. Solid State Chem.* **1997**, *130*, 117-128; f) K. Tabata, Y. Hirano, E. Suzuki, *Appl. Cat. A.* **1998**, *170*, 245-254; g) N. Taihei, I. Tatsumi, M. Makoto, *Bull. Chem. Soc. Jpn.* **1988**, *61*, 621-626; h) Y. Zhang-Steenwinkel, J. Beckers, A. Bliek, *Appl. Cat. A.* **2002**, *235*, 79-92; i) S. Miyoshi, J.-O. Hong, K. Yashiro, A. Kaimai, Y. Nigara, K. Kawamura, T. Kawada, J. Mizusaki, *Solid State Ion.* **2003**, *161*, 209-217; j) T. Kuznetsova, V. Sadykov, L. Batuev, L. Kurina, *React. Kinet. Catal. L.* **2005**, *86*, 257-265; k) M. Abdolrahmani, M. Parvari, M. Habibpoor, *Chin. J. Catal.* **2010**, *31*, 394-403; l) A. Tarjomannejad, A. Niaei, A. Farzi, D. Salari, P. R. Zonouz, *Catal. Lett.* **2016**, *146*, 1544-1551; m) P. R. Zonouz, A. Niaei, A. Tarjomannejad, *Int. J. Environ. Sci.* **2016**, *13*, 1665-1674.

[8] a) O. Parkash, P. Ganguly, G. R. Rao, C. N. R. Rao, D. S. Rajoria, V. G. Bhide, *Mater. Res. Bull.* **1974**, *9*, 1173-1176; b) M. Futai, C. Yonghua, Louhui, *React. Kinet. Catal. L.* **1986**, *31*, 47-54; c) S. George, B. Viswanathan, *J. Colloid Interface Sci.* **1983**, *95*, 322-326.

[9] a) E. Niwa, T. Sato, Y. Watanabe, Y. Toyota, Y. Hatakeyama, K. Judai, K. Shozugawa, M. Matsuo, T. Hashimoto, *J. Ceram. Soc. Jpn.* **2015**, *123*, 501-506; b) B. Levasseur, S. Kaliaguine, *J. Solid State Chem.* **2008**, *181*, 2953-2963.

[10] S. A. Hallweger, C. Kaußler, G. Kieslich, *Phys. Chem. Chem. Phys.* **2022**, *24*, 9196-9202.

[11] G. Koch, G. Bellini, F. Girgsdies, S. J. Carey, O. Timpe, T. Götsch, J. Kröhnert, G. Auffermann, M. Hashagen, R. Schlögl, A. Trunschke, *Chem. Mater.*, under review.

[12] a) Q. Tao, P. Xu, M. Li, W. Lu, *npj Comput. Mater.* **2021**, *7*, 23; b) B. Weng, Z. Song, R. Zhu, Q. Yan, Q. Sun, C. G. Grice, Y. Yan, W.-J. Yin, *Nat. Comm.* **2020**, *11*, 3513; c) T. Toyao, Z. Maeno, S. Takakusagi, T. Kamachi, I. Takigawa, K.-i. Shimizu, *ACS Catal.* **2020**, *10*, 2260-2297; d) S. Ma, Z.-P. Liu, *ACS Catal.* **2020**, *10*, 13213-13226; e) K. Takahashi, L. Takahashi, I. Miyazato, J. Fujima, Y. Tanaka, T. Uno, H. Satoh, K. Ohno, M. Nishida, K. Hirai, J. Ohyama, T. N. Nguyen, S. Nishimura, T. Taniike,

ChemCatChem **2019**, *11*, 1146-1152; f) P. Schlexer Lamoureux, K. T. Winther, J. A. Garrido Torres, V. Streibel, M. Zhao, M. Bajdich, F. Abild-Pedersen, T. Bligaard, *ChemCatChem* **2019**, *11*, 3581-3601; g) R. Palkovits, S. Palkovits, *ACS Catal.* **2019**, *9*, 8383-8387; h) J. R. Kitchin, *Nat. Catal.* **2018**, *1*, 230-232; iJ K. Nørskov, T. Bligaard, *Angew. Chem. Int. Ed.* **2013**, *52*, 776-777; j) L. Foppa, C. Sutton, L. M. Ghiringhelli, S. De, P. Löser, S. A. Schunk, A. Schäfer, M. Scheffler, *ACS Catal.* **2022**, *12*, 2223-2232; k) L. Foppa, L. M. Ghiringhelli, *Top. Catal.* **2022**, *65*, 196-206.

[13] Y. Liu, T. Zhao, W. Ju, S. Shi, *J. Materiomics* **2017**, *3*, 159-177.

[14] A. Trunschke, G. Bellini, M. Boniface, S. J. Carey, J. Dong, E. Erdem, L. Foppa, W. Frandsen, M. Geske, L. M. Ghiringhelli, F. Girgsdies, R. Hanna, M. Hashagen, M. Hävecker, G. Huff, A. Knop-Gericke, G. Koch, P. Kraus, J. Kröhnert, P. Kube, S. Lohr, T. Lunkenbein, L. Masliuk, R. Naumann d'Alnoncourt, T. Omojola, C. Pratsch, S. Richter, C. Rohner, F. Rosowski, F. Rüther, M. Scheffler, R. Schlögl, A. Tarasov, D. Teschner, O. Timpe, P. Trunschke, Y. Wang, S. Wrabetz, *Top. Catal.* **2020**, *63*, 1683-1699.

[15] a) L. Foppa, L. M. Ghiringhelli, F. Girgsdies, M. Hashagen, P. Kube, M. Hävecker, S. J. Carey, A. Tarasov, P. Kraus, F. Rosowski, R. Schlögl, A. Trunschke, M. Scheffler, *MRS Bull.* **2021**, *46*, 1016-1026; b) L. Foppa, F. Rüther, M. Geske, G. Koch, F. Girgsdies, P. Kube, S. J. Carey, M. Hävecker, O. Timpe, A. V. Tarasov, M. Scheffler, F. Rosowski, R. Schlögl, A. Trunschke, *J. Am. Chem. Soc.* **2023**, *145*, 3427-3442; c) R. Miyazaki, K. S. Belthle, H. Tüysüz, L. Foppa, M. Scheffler, *J. Am. Chem. Soc.* **2024**, *146*, 5433-5444.

[16] A. Varma, A. S. Mukasyan, A. S. Rogachev, K. V. Manukyan, *Chem. Rev.* **2016**, *116*, 14493-14586.

[17] a) R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, L. M. Ghiringhelli, *Phys. Rev. Mat.* **2018**, *2*, 083802; b) L. Foppa, T. A. R. Purcell, S. V. Levchenko, M. Scheffler, L. M. Ghringhelli, *Phys. Rev. Lett.* **2022**, *129*, 055301; c) T. A. R. Purcell, M. Scheffler, C. Carbogno, L. M. Ghiringhelli, *J. Open Source Soft.* **2022**, *7 (71)*, 3960.

[18] a) J. J. Yeh, I. Lindau, *Atomic Data and Nuclear Data Tables* **1985**, *32*, 1-155; b) S. Tanuma, C. J. Powell, D. R. Penn, *Surf. Interf. Anal.* **1993**, *20*, 77-89; c) M. C. Biesinger, B. P. Payne, A. P. Grosvenor, L. W. M. Lau, A. R. Gerson, R. S. C. Smart, *Appl. Surf. Sci.* **2011**, *257*, 2717-2730; d) M. F. Sunding, K. Hadidi, S. Diplas, O. M. Løvvik, T. E. Norby, A. E. Gunnæs, *J. Elec. Spec. Rel. Phenom.* **2011**, *184*, 399-409; e) M. C. Biesinger, *Surf. Interf. Anal.* **2017**, *49*, 1325-1334; f) H. Ogasawara, A. Kotani, R. Potze, G. A. Sawatzky, B. T. Thole, *Physical Review B* **1991**, *44*, 5465; g) G. Koch, M. Hävecker, D. Teschner, S. J. Carey, Y. Wang, P. Kube, W. Hetaba, T. Lunkenbein, G. Auffermann, O. Timpe, F. Rosowski, R. Schlögl, A. Trunschke, *ACS Catal.* **2020**, *10*, 7007-7020; h) J. Walton, P. Wincott, N. Fairley, A. Carrick, *Peak Fitting with CasaXPS*, Accolyte Science, Knutsford, UK, **2010**; i) A. Moshantaf, P. Oppermann, H. Junkes, *Vol. 2023*, **2022**.