

Cite this: *Catal. Sci. Technol.*, 2023,  
13, 5576

# Indirect design of OCM catalysts through machine learning of catalyst surface oxygen species†

Fumiya Nishino,<sup>a</sup> Hiroshi Yoshida,<sup>b</sup> Masato Machida,<sup>c</sup> Shun Nishimura,<sup>d</sup>  
Keisuke Takahashi<sup>e</sup>\* and Junya Ohyama<sup>c</sup>\*

Catalysts for oxidative coupling of methane (OCM) were designed through machine learning of a property of surface oxygen species on the basis of the knowledge that catalytic performance for the OCM is affected by catalyst surface oxygen species. To select the property of the surface oxygen species used as a guide of catalyst design *via* machine learning, the relationships between the total yield of ethylene and ethane (C<sub>2</sub> yield) and the O1s X-ray photoelectron spectral (XPS) features of the 51 catalysts prepared in our previous study were evaluated. Since a weak correlation was seen between the C<sub>2</sub> yield and the O1s XPS peak energy of CO<sub>3</sub><sup>2-</sup> species on the catalyst surface, the CO<sub>3</sub><sup>2-</sup> peak energy was chosen as the guiding parameter of catalyst design in this work. Machine learning was then performed on the dataset consisting of the CO<sub>3</sub><sup>2-</sup> peak energy (objective variable) and the physical quantities of elements in the catalysts (descriptor) to find the important physical quantities determining the CO<sub>3</sub><sup>2-</sup> peak energy. According to the important physical quantities, catalyst compositions were predicted. Based on the predicted compositions, 28 catalysts were synthesized to verify that their CO<sub>3</sub><sup>2-</sup> peak energies were in the range where high catalytic performance can be expected. Furthermore, the catalysts are tested for the OCM reaction. As a result, Ba–In–Rb/La<sub>2</sub>O<sub>3</sub> was found as a new highly active OCM catalyst having compatible activity to the conventional Mn–Na<sub>2</sub>WO<sub>4</sub>/SiO<sub>2</sub> catalyst. Therefore, it was demonstrated that the indirect catalyst through machine learning of the catalyst surface property is effective for development of catalysts.

Received 28th April 2023,  
Accepted 21st August 2023

DOI: 10.1039/d3cy00587a

rsc.li/catalysis

## Introduction

Oxidative coupling of methane is a direct conversion reaction of CH<sub>4</sub> to ethylene and ethane (C<sub>2</sub> compounds), which can offer a more energetically efficient and economical process than the conventional conversion of CH<sub>4</sub> to olefin *via* CO and CH<sub>3</sub>OH.<sup>1</sup> However, the OCM has a problem that the C<sub>2</sub> yield hardly exceeds 30%.<sup>1,2</sup> To overcome this problem, hundreds of catalysts have been developed.<sup>1–4</sup> Among the catalysts, Mn–

Na<sub>2</sub>WO<sub>4</sub>/SiO<sub>2</sub> shows relatively high performance for the OCM.<sup>5–7</sup> As seen in the case with Mn–Na<sub>2</sub>WO<sub>4</sub>/SiO<sub>2</sub>, combination of elements is effective in developing the OCM catalysts. Here, data science techniques have been applied to search for good combinations of catalyst elements from a vast combinatorial space.<sup>3,4,8–15</sup> In fact, in our previous studies, catalysts have been designed through machine learning and data mining of the direct relationship between the C<sub>2</sub> yield and the descriptors of elements (*e.g.*, element number, atomic weight, atomic radius, melting point, evaporation heat, electronegativity, *etc.*).<sup>13</sup> On the other hand, catalytic performance is strongly affected by the catalyst surface properties. In the case of the OCM catalysts, the properties of surface oxygen species are known to have an impact on the catalytic performance.<sup>16–19</sup> Accordingly, the OCM catalysts can also be designed by machine learning of the relationship between the surface oxygen properties and the descriptors of catalyst elements. In addition, it is expected that some of the catalyst element combinations predicted from the surface oxygen properties are the same as the ones predicted by the direct catalyst design from the relation between the catalytic performance and the descriptors of elements, but the others will be different because the properties of surface oxygen

<sup>a</sup> Department of Applied Chemistry and Biochemistry, Graduate School of Science and Technology, Kumamoto University, 2-39-1 Kurokami, Chuo-ku, Kumamoto, 860-8555, Japan

<sup>b</sup> College of Science and Engineering, Kanazawa University, Kakuma-cho, Kanazawa, 920-1192, Japan

<sup>c</sup> Faculty of Advanced Science and Technology, Kumamoto University, 2-39-1 Kurokami, Chuo-ku, Kumamoto, 860-8555, Japan. E-mail: ohyama@kumamoto-u.ac.jp

<sup>d</sup> Graduate School of Advanced Science and Technology, Japan Advanced Institute of Science and Technology (JAIST), 1-1 Asahidai, Nomi, 923-1292, Japan

<sup>e</sup> Department of Chemistry, Hokkaido University, N-15 W-8, Sapporo, 060-0815, Japan. E-mail: keisuke.takahashi@sci.hokudai.ac.jp

† Electronic supplementary information (ESI) available: O1s XPS data; supplementary Tables A and B listing predicted element combinations. See DOI: <https://doi.org/10.1039/d3cy00587a>

species do not completely describe the catalytic performance. Therefore, new catalyst combinations can be explored with a guide of the rational descriptors.

In this study, indirect design of OCM catalysts through machine learning of a property of catalyst surface oxygen species was performed. The property of catalyst surface oxygen species used as the guide for catalyst design was selected from the features of O1s X-ray photoelectron spectra (XPS). Based on the guide of the selected surface oxygen property, OCM catalysts were predicted using machine learning and were verified by XPS measurement and the OCM reaction tests.

## Methods

### Catalyst preparation

OCM catalysts predicted by machine learning,  $M_1$ - $M_2$ - $M_3$ /A ( $M_1$ ,  $M_2$ ,  $M_3$  = In, K, Rb, Cs, Ba, Yb, Bi, Sm, Ce, Li, W, Hf, Eu, Cu, Zn, Nd, A = SiO<sub>2</sub>, La<sub>2</sub>O<sub>3</sub>, Yb<sub>2</sub>O<sub>3</sub>, CeO<sub>2</sub>), were prepared by a wet impregnation method. In(NO<sub>3</sub>)<sub>3</sub>·3H<sub>2</sub>O, KNO<sub>3</sub>, RbNO<sub>3</sub>, CsNO<sub>3</sub>, Ba(NO<sub>3</sub>)<sub>2</sub>, Yb(NO<sub>3</sub>)<sub>3</sub>·5H<sub>2</sub>O, Bi(NO<sub>3</sub>)<sub>3</sub>·5H<sub>2</sub>O, Sm(NO<sub>3</sub>)<sub>3</sub>·6H<sub>2</sub>O, Ce(NO<sub>3</sub>)<sub>3</sub>·6H<sub>2</sub>O, LiNO<sub>3</sub>, (NH<sub>4</sub>)<sub>6</sub>H<sub>2</sub>W<sub>12</sub>O<sub>40</sub>·5H<sub>2</sub>O, HfOCl<sub>2</sub>·8H<sub>2</sub>O, Eu(NO<sub>3</sub>)<sub>3</sub>·6H<sub>2</sub>O, Cu(NO<sub>3</sub>)<sub>2</sub>·6H<sub>2</sub>O, Zn(NO<sub>3</sub>)<sub>2</sub>·6H<sub>2</sub>O, Nd(NO<sub>3</sub>)<sub>3</sub>·6H<sub>2</sub>O were used as the metal precursors. The loading amount of each of  $M_1$ ,  $M_2$ , and  $M_3$  was 1 wt%. After impregnation, the materials were dried at 110 °C overnight and calcined in air at 1000 °C for 3 h. The metals and supports used in the  $M_1$ - $M_2$ - $M_3$ /A preparation are shown in Table 1. Mn-Na<sub>2</sub>WO<sub>4</sub>/SiO<sub>2</sub> was prepared as a reference catalyst by impregnating SiO<sub>2</sub> in an aqueous

solution containing Mn(NO<sub>3</sub>)<sub>2</sub> and Na<sub>2</sub>WO<sub>4</sub>. After the impregnation, the material was dried overnight at 110 °C and calcined at 1000 °C for 3 h (10 °C min<sup>-1</sup>) to obtain Mn-Na<sub>2</sub>WO<sub>4</sub>/SiO<sub>2</sub> with 1.9 wt% Mn and 5.0 wt% Na<sub>2</sub>WO<sub>4</sub>.<sup>13</sup>

### Catalyst preparation

The OCM reaction was performed in a fixed-bed flow reactor. 50 mg of catalyst (10–20 mesh) was placed in a quartz glass tube reactor (inner diameter (ID) 4 mm) and fixed with quartz wool. Prior to the reaction, a catalyst bed was pretreated at 400 °C for 10 min under a O<sub>2</sub> flow of 8.3 mL min<sup>-1</sup>. After purging with a N<sub>2</sub> flow of 12 mL min<sup>-1</sup>, the reaction was started by flowing a reaction gas mixture of CH<sub>4</sub>/O<sub>2</sub>/N<sub>2</sub> = 24/7.5/3.0 mL min<sup>-1</sup>. Reaction temperature was increased to 500 °C, 600 °C, 700 °C, 800 °C, and 900 °C. The reaction gas was analyzed 15–30 min after the temperature reached to the set value. The outlet gas after cold traps at ca. 10 °C was analyzed using a gas chromatograph with a thermal conductivity detector (490 Micro GC Agilent Technologies). The carbon missing was determined by CH<sub>4</sub> conversion – (sum of C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, CO, and CO<sub>2</sub> yield) (%).

### XPS measurement

XPS measurements were performed using K-Alpha (Thermo Fisher Scientific) with monochromatic X-ray irradiation AlK $\alpha$  ( $h\nu$  = 1486.6 eV). The samples were analyzed as prepared without pretreatment. Binding energy was corrected with C1s peak at 285 eV. Peak deconvolution of O1s XPS spectra was performed after background removal.

### Machine learning and descriptors design

Random forest regression is implemented within scikit-learn in order to evaluate the importance of descriptors.<sup>20</sup> The number of tree is set to 100.

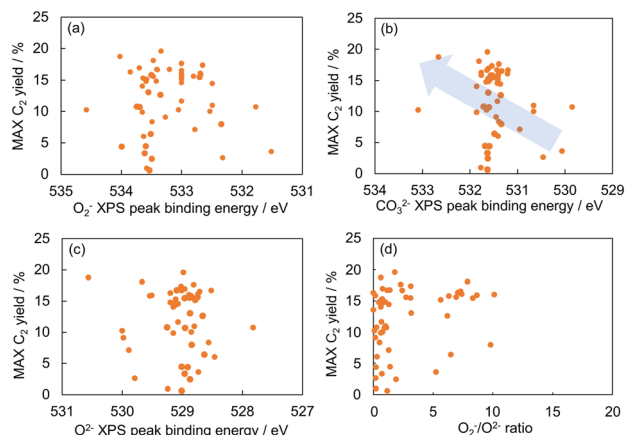
Catalysts descriptor is designed using physical quantities from periodic table. In particular, XenonPy is used to assign the physical quantity in order to define catalysts.<sup>21</sup> Here, catalyst descriptor is designed by weighted average physical quantities based on the following equation:  $\sum(P_i \times C_i)$ , where  $P_i$  is a physical quantity of element  $i$  and  $C_i$  is its composition (mol%).

## Results and discussion

The relationship between the C<sub>2</sub> yield and the O1s XPS spectra of the 51 OCM catalysts prepared in our previous study was examined to find property of surface oxygen species related to the catalytic activity for the OCM reaction (Table S1†).<sup>9</sup> Each of the O1s XPS spectra was deconvoluted into three peaks, which are assignable to superoxide (O<sub>2</sub><sup>-</sup>), carbonate (CO<sub>3</sub><sup>2-</sup>) and lattice oxygen (O<sup>2-</sup>) in order from high energy to low energy according to the literature.<sup>16,17</sup> The relationship between the C<sub>2</sub> yield and the peak binding energy of each oxygen species is presented in Fig. 1(a)–(c). Although all the three plots do not show strong correlations,

**Table 1** List of the 28 catalysts prepared in this study

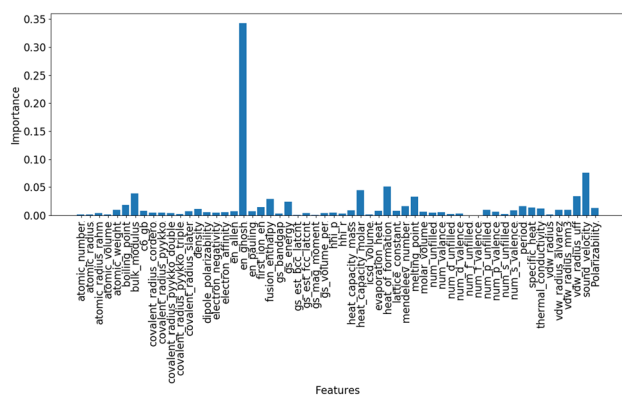
$M_1$	$M_2$	$M_3$	A
Ba	Ce	Cs	SiO <sub>2</sub>
Ba	Hf	Cs	CeO <sub>2</sub> or Yb <sub>2</sub> O <sub>3</sub>
Ba	Hf	Sm	CeO <sub>2</sub> or Yb <sub>2</sub> O <sub>3</sub>
Bi	Hf	Sm	Yb <sub>2</sub> O <sub>3</sub>
Eu	Hf	Sm	Yb <sub>2</sub> O <sub>3</sub>
Eu	Hf	Nd	Yb <sub>2</sub> O <sub>3</sub>
Hf	In	Sm	Yb <sub>2</sub> O <sub>3</sub>
Bi	In	Sm	Yb <sub>2</sub> O <sub>3</sub>
Hf	In	Nd	Yb <sub>2</sub> O <sub>3</sub>
Ba	Hf	—	Yb <sub>2</sub> O <sub>3</sub>
Cs	Hf	—	Yb <sub>2</sub> O <sub>3</sub>
Sm	Hf	—	Yb <sub>2</sub> O <sub>3</sub>
Ce	Hf	—	Yb <sub>2</sub> O <sub>3</sub>
Cu	Hf	—	Yb <sub>2</sub> O <sub>3</sub>
Zn	Hf	—	Yb <sub>2</sub> O <sub>3</sub>
Zn	W	—	Yb <sub>2</sub> O <sub>3</sub>
Ba	In	Rb	La <sub>2</sub> O <sub>3</sub>
In	Rb	Yb	La <sub>2</sub> O <sub>3</sub>
Bi	In	Rb	La <sub>2</sub> O <sub>3</sub>
Ba	In	Yb	La <sub>2</sub> O <sub>3</sub>
Ba	Rb	Yb	La <sub>2</sub> O <sub>3</sub>
Bi	In	Yb	La <sub>2</sub> O <sub>3</sub>
In	Rb	Sm	La <sub>2</sub> O <sub>3</sub>
Ce	In	Rb	La <sub>2</sub> O <sub>3</sub>
Cs	In	Rb	La <sub>2</sub> O <sub>3</sub>
Bi	Rb	Yb	La <sub>2</sub> O <sub>3</sub>



**Fig. 1** Plots of maximum  $C_2$  yield against the three deconvoluted  $O1s$  XPS peaks assignable to (a)  $O_2^-$ , (b)  $CO_3^{2-}$ , and (c)  $O^{2-}$  species. (d) Plot of maximum  $C_2$  yield against  $O_2^-/O^{2-}$  ratio. Maximum  $C_2$  yield: the highest  $C_2$  yield value obtained with each catalyst in the experiment protocol.

the plot in Fig. 1(b) exhibits better correlation than the those in Fig. 1(a) and (c). Thus, the  $CO_3^{2-}$  species can be a descriptor of the OCM catalyst. The correlation between the  $O_2^-/O^{2-}$  area ratio and the  $C_2$  yield was also investigated as presented in Fig. 1(d) because the  $O_2^-/O^{2-}$  area ratio has been suggested to be a descriptor of the  $C_2$  yield in the literature.<sup>17</sup> As a result, the  $CO_3^{2-}$  peak energy exhibited a better correlation with the  $C_2$  yield than the  $O_2^-/O^{2-}$  area ratio in the catalyst dataset of this study. Although the correlation in Fig. 1(b) is weak, the catalyst design based on the  $CO_3^{2-}$  peak energy is considered reasonable because the  $CO_3^{2-}$  peak energy can be related to the surface basicity, which contributes to the catalytic performance. Thus, the  $CO_3^{2-}$  peak energy was selected as the guiding parameter for catalyst design. Since there is a rough trend to increase the  $C_2$  yield with an increase of the  $CO_3^{2-}$  peak binding energy, a high  $CO_3^{2-}$  peak binding energy can guide the design of OCM catalysts.

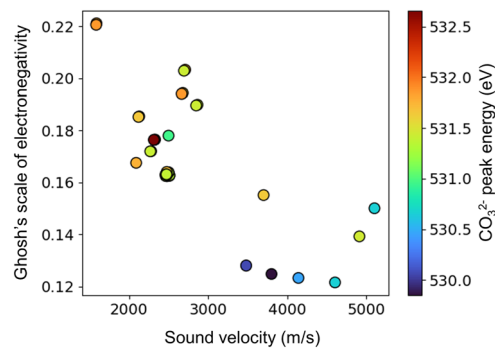
Random forest regression was performed on the dataset consisting of “ $\sum(P_i \times C_i)$ ” as the descriptor variables and



**Fig. 2** The importance of weighted average physical quantities against  $CO_3^{2-}$  peak energy.

$CO_3^{2-}$  peak energy as the objective variable to identify the important physical quantities of catalyst elements representing the  $CO_3^{2-}$  peak energy. The importance of descriptor variables is demonstrated in Fig. 2. Ghosh's scale of electronegativity ( $en\_ghosh$ ) and sound velocity have high importance. It should be noted that Ghosh's scale of electronegativity and sound velocity are not considered to directly relate to the  $CO_3^{2-}$  peak energy but indirectly represent some factors determining the  $CO_3^{2-}$  peak energy.

Here,  $en\_ghosh$  and sound velocity against  $CO_3^{2-}$  are visualized in Fig. 3. It shows that high  $en\_ghosh$  and low sound velocity tend to result high  $CO_3^{2-}$  energy peak. One can hypothesize that catalysts having high  $en\_ghosh$  and low sound velocity could result high  $C_2$  yield based on the fact high  $CO_3^{2-}$  peak results high  $C_2$  yield. Therefore, catalysts having high  $en\_ghosh$  and low sound velocity are explored from the calculated  $en\_ghosh$  and sound velocities of a variety of element combinations using  $\sum(P_i \times C_i)$ . The element combinations are created by selecting three elements from the 33 elements for  $M_1$ ,  $M_2$ , and  $M_3$ : Li, Na, Mg, Al, Si, Ni, K, Ca, Ti, V, Mn, Fe, Co, Cu, Zn, Rb, Sr, Y, Zr, Mo, Pd, In, Cs, Ba, La, Ce, Nd, Sm, Eu, Yb, Hf, W, Bi. These elements are selected from the elements found in the literature to find new combinations.<sup>2</sup> Here, 180 048 combinations of  $M_1$ ,  $M_2$ ,  $M_3$ , and support ( ${}_{33}C_3$  ( $M_1$ ,  $M_2$ ,  $M_3$  combinations)  $\times$  33 (support) = 180 048) where the mol% is set to 2, 4, 2, and 92, respectively are created and weighted average of  $en\_ghosh$  and sound velocity are calculated (ESI† Table A). Created catalysts combinations are then visualized as shown in Fig. 4. It must be noted that mol% of support is set to 92 which has a large impact on weighted average, therefore, data are aggregated by support. As it can be seen in Fig. 5, catalysts containing Si, Ce, Bi, Yb, and Sm in supports show high  $en\_ghosh$  and low sound velocity. Based on the data, catalysts listed in Table 2, which have high  $en\_ghosh$  and low sound velocity, are designed by randomly selecting a Si-based catalyst having  $\geq 0.177$   $en\_ghosh$  and  $\leq 2190$  of sound velocity, two Ce-based catalysts having  $\geq 0.167$   $en\_ghosh$  and  $\leq 2110$  of sound velocity, and eight Yb-based catalysts having  $\geq 0.218$  of  $en\_ghosh$  and  $\leq 1700$  of sound velocity. In addition, La-based catalysts are designed by selecting



**Fig. 3** Ghosh's scale of electronegativity ( $en\_ghosh$ ) and sound velocity against  $CO_3^{2-}$  peak energy.

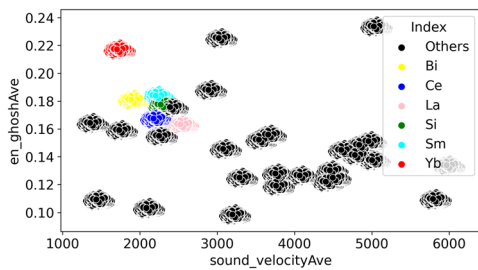


Fig. 4 Created 180 048 catalysts combination of Ghosh's scale of electronegativity ( $en\_ghosh$ ) and sound velocity.

combinations having  $\geq 0.161$  of  $en\_ghosh$  and  $\leq 2404$  of sound velocity since La-based catalysts are known to have relatively high activity for the OCM.<sup>13,22</sup> It should be noted that the prediction in this indirect design will not offer accurate or pinpoint prediction of metals–support combinations having high  $C_2$  yield because the prediction is based on the weak trend between the  $CO_3^{2-}$  peak energy and the  $C_2$  yield. However, the predicted catalyst group by this indirect method may contain good catalysts, which may be different from catalysts the direct prediction can find.

The catalysts were actually prepared based on the predicted elements listed in Table 2. The loading amount of each element was set to 1 wt%. Although the loading amount (1 wt%) is smaller than those calculated from the compositions in the prediction (2 or 4 mol%), the surface composition of the catalysts prepared by the impregnation method are considered to be compatible to or greater than the compositions used in the prediction. As a result, a total of 28 catalysts were prepared based on the prediction (Table 1). To verify the predictions, O1s XPS spectra of the 28 catalysts were measured. All the O1s XPS spectra are shown in Fig. S1.† Each of the spectra was deconvoluted into three peaks to evaluate the  $CO_3^{2-}$  peak energy (Fig. S2†). As a result, the  $CO_3^{2-}$  peak energies of all the 28 catalysts were larger than 531.2 eV (Table S2†), which means that the 28

Table 2 Designed catalysts having high  $en\_ghosh$  and low sound velocity within each supported catalyst

$M_1$	$C_1$	$M_2$	$C_2$	$M_3$	$C_3$	A	CA	$en\_ghosh$	Sound velocity ( $m\ s^{-1}$ )
Ba	0.02	Ce	0.04	Cs	0.02	Si	0.92	0.177	2184
Ba	0.02	Cs	0.04	Hf	0.02	Ce	0.92	0.168	2109
Ba	0.02	Cs	0.04	Sm	0.02	Ce	0.92	0.167	2094
Ba	0.02	Cs	0.04	Hf	0.02	Yb	0.92	0.218	1637
Ba	0.02	Hf	0.04	Sm	0.02	Yb	0.92	0.220	1651
Hf	0.02	In	0.04	Sm	0.02	Yb	0.92	0.219	1615
Hf	0.02	In	0.04	Nd	0.02	Yb	0.92	0.219	1618
Bi	0.02	In	0.04	Sm	0.02	Yb	0.92	0.218	1590
Bi	0.02	Hf	0.04	Sm	0.02	Yb	0.92	0.220	1654
Eu	0.02	Hf	0.04	Nd	0.02	Yb	0.92	0.220	1679
Eu	0.02	Hf	0.04	Sm	0.02	Yb	0.92	0.220	1676
Ba	0.02	Hf	0.04	Yb	0.02	Yb	0.92	0.220	1641
Cs	0.02	Hf	0.04	Yb	0.02	Yb	0.92	0.220	1652
Cu	0.02	Hf	0.04	Yb	0.02	Yb	0.92	0.220	1680
Hf	0.02	Yb	0.04	Zn	0.02	Yb	0.92	0.220	1656
Hf	0.02	Sm	0.04	Yb	0.02	Yb	0.92	0.220	1638
Ce	0.02	Hf	0.04	Yb	0.02	Yb	0.92	0.220	1651
W	0.02	Yb	0.04	Zn	0.02	Yb	0.92	0.220	1699
Ba	0.02	In	0.04	Rb	0.02	La	0.92	0.162	2392
In	0.02	Rb	0.04	Yb	0.02	La	0.92	0.162	2392
Bi	0.02	In	0.04	Rb	0.02	La	0.92	0.163	2396
Ba	0.02	In	0.04	Yb	0.02	La	0.92	0.164	2397
Ba	0.02	Rb	0.04	Yb	0.02	La	0.92	0.162	2400
Bi	0.02	In	0.04	Yb	0.02	La	0.92	0.165	2401
In	0.02	Rb	0.04	Sm	0.02	La	0.92	0.161	2402
Ce	0.02	In	0.04	Rb	0.02	La	0.92	0.162	2402
Cs	0.02	In	0.04	Rb	0.02	La	0.92	0.162	2403
Bi	0.02	Rb	0.04	Yb	0.02	La	0.92	0.162	2404

catalysts are in the high energy side of Fig. 1. Therefore, the predictions in Table 2 were verified.

Fig. 5 shows the results of the OCM reaction over the 28 catalysts together with only supports and Mn– $Na_2WO_4/SiO_2$  for comparison. The experimental error was evaluated by five blank (no catalyst) tests, which gave  $9.7 \pm 0.9\%$  of the  $C_2$  yield at 900 °C. The  $La_2O_3$ -based catalysts presented  $<5\%$  of the carbon missing at all reaction temperatures. The  $SiO_2$ -based catalysts exhibited  $<5\%$  at  $\leq 800$  °C and 10–15% at 900 °C. The  $Yb_2O_3$ - and  $CeO_2$ -based catalysts showed  $<5\%$  at  $\leq 800$  °C and 5–10% at 900 °C. The increase of the carbon missing at 900 °C and by using the  $SiO_2$ -based catalyst might be due to coke formation.

The  $La_2O_3$ - and  $Yb_2O_3$ -based catalysts showed higher activity at lower temperatures than those supported on  $SiO_2$  and  $CeO_2$ . More importantly, several predicted catalysts exhibited comparable  $C_2$  yields to that of Mn– $Na_2WO_4/SiO_2$ . Specifically, the most active catalyst was Ba–In–Rb/ $La_2O_3$ , which gave 19%  $C_2$  yield, 28%  $CH_4$  conversion, and 69%  $C_2$  selectivity at 700 °C, while the reference catalyst Mn– $Na_2WO_4/SiO_2$  gave 19%  $C_2$  yield, 34%  $CH_4$  conversion, and 57%  $C_2$  selectivity at 900 °C. Therefore, Ba–In–Rb/ $La_2O_3$  exhibited comparable or better catalytic performance to Mn– $Na_2WO_4/SiO_2$  at lower temperature. This demonstrates that the catalyst design by machine learning of catalyst surface properties is effective for development of catalysts.

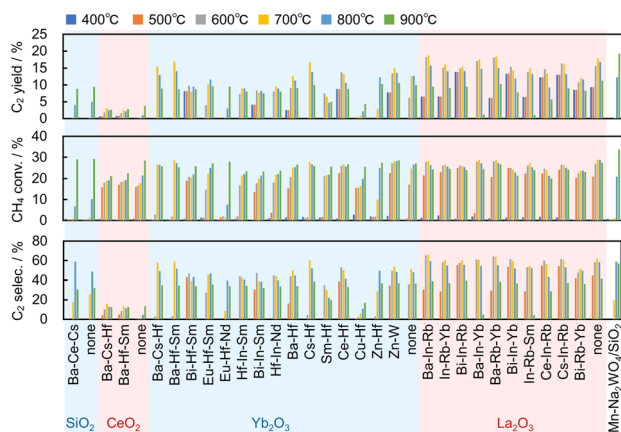


Fig. 5 Catalytic performance of the 28 predicted catalysts for the OCM reaction together with only supports and Mn– $Na_2WO_4/SiO_2$  for comparison.

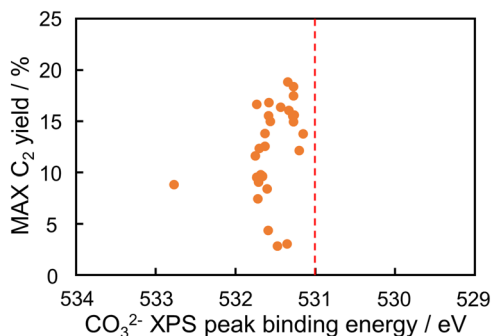


Fig. 6  $\text{CO}_3^{2-}$  XPS peak binding energy vs. maximum  $\text{C}_2$  yield of the predicted 28 catalysts.

The maximum  $\text{C}_2$  yields of the predicted catalysts at 400–900 °C were plotted against their  $\text{CO}_3^{2-}$  peak binding energies in Fig. 6. The figure does not show a correlation because of the limited kinds of catalysts in the verification experiment: the left lower data in the figure is derived from  $\text{SiO}_2$  supported catalyst, and the other data are  $\text{La}_2\text{O}_3$ ,  $\text{Yb}_2\text{O}_3$ , and  $\text{CeO}_2$  supported catalysts. This result suggests that the  $\text{CO}_3^{2-}$  peak binding energy is not the only descriptor of the OCM reaction. This is consistent with the rough trend in Fig. 1(b). However, the predicted catalysts designed from the not-strong descriptor contained Ba–In–Rb/ $\text{La}_2\text{O}_3$  having high catalytic performance. This result shows that the indirect catalyst design is effective in development of catalysts.

## Conclusions

The OCM catalysts were developed through machine learning of property of their surface oxygen species. In this study, the  $\text{CO}_3^{2-}$  peak energy was selected as the property of the surface oxygen species contributing to the OCM catalysts based on the rough trend of the  $\text{C}_2$  yield increasing with the  $\text{CO}_3^{2-}$  peak energy in the data of the previously reported catalysts. Machine learning was then performed to find as the important physical quantities of catalyst elements representing the  $\text{CO}_3^{2-}$  peak energy. Based on the relation between  $\text{CO}_3^{2-}$  peak energy and the important physical quantities, the catalyst compositions resulting high  $\text{CO}_3^{2-}$  peak energy were designed and the 28 catalysts were experimentally prepared. The O1s XPS spectral analysis verified that all the 28 catalysts have relatively high  $\text{CO}_3^{2-}$  peak energy. Furthermore, Ba–In–Rb/ $\text{La}_2\text{O}_3$ , one of the predicted catalysts, exhibited compatible  $\text{C}_2$  yield with a higher selectivity at lower reaction temperature compared to the conventional Mn– $\text{Na}_2\text{WO}_4/\text{SiO}_2$  catalyst. The results suggest that the indirect design of catalyst through machine learning of catalyst surface property is effective in developing catalysts. This means that, if one has an understanding or a hypothesis about how catalyst surface properties affect a target catalytic reaction and there is appropriate database of catalyst properties, one can obtain catalyst designs in a similar way to the indirect catalyst design method performed in this study.

## Author contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was supported by the Japan Science and Technology Agency (JST) CREST (grant no. JPMJCR17P2).

## References

- 1 E. V. Kondratenko, T. Peppel, D. Seeburg, V. A. Kondratenko, N. Kalevaru, A. Martin and S. Wohlrab, *Catal. Sci. Technol.*, 2017, **7**, 366–381.
- 2 U. Zavyalova, M. Holena, R. Schlögl and M. Baerns, *ChemCatChem*, 2011, **3**, 1935–1947.
- 3 T. N. Nguyen, T. T. P. Nhat, K. Takimoto, A. Thakur, S. Nishimura, J. Ohyama, I. Miyazato, L. Takahashi, J. Fujima, K. Takahashi and T. Taniike, *ACS Catal.*, 2020, **10**, 921–932.
- 4 S. Mine, M. Takao, T. Yamaguchi, T. Toyao, Z. Maeno, S. M. A. Hakim Siddiki, S. Takakusagi, K.-i. Shimizu and I. Takigawa, *ChemCatChem*, 2021, **13**, 3636–3655.
- 5 S. Arndt, T. Otremba, U. Simon, M. Yildiz, H. Schubert and R. Schomäcker, *Appl. Catal., A*, 2012, **425–426**, 53–61.
- 6 J. Ohyama, S. Nishimura and K. Takahashi, *ChemCatChem*, 2019, **11**, 4307–4313.
- 7 D. Kiani, S. Sourav, J. Baltrusaitis and I. E. Wachs, *ACS Catal.*, 2019, **9**, 5912–5928.
- 8 K. Takahashi, J. Ohyama, S. Nishimura, J. Fujima, L. Takahashi, T. Uno and T. Taniike, *Chem. Commun.*, 2023, **59**, 2222–2238.
- 9 K. Takahashi, I. Miyazato, S. Nishimura and J. Ohyama, *ChemCatChem*, 2018, **10**, 3223–3228.
- 10 L. Takahashi, T. N. Nguyen, S. Nakanowatari, A. Fujiwara, T. Taniike and K. Takahashi, *Chem. Sci.*, 2021, **12**, 12546–12555.
- 11 S. Nishimura, S. D. Le, I. Miyazato, J. Fujima, T. Taniike, J. Ohyama and K. Takahashi, *Catal. Sci. Technol.*, 2022, **12**, 2766–2774.
- 12 K. Sugiyama, T. N. Nguyen, S. Nakanowatari, I. Miyazato, T. Taniike and K. Takahashi, *ChemCatChem*, 2021, **13**, 952–957.
- 13 J. Ohyama, T. Kinoshita, E. Funada, H. Yoshida, M. Machida, S. Nishimura, T. Uno, J. Fujima, I. Miyazato, L. Takahashi and K. Takahashi, *Catal. Sci. Technol.*, 2021, **11**, 524–530.
- 14 K. Suzuki, T. Toyao, Z. Maeno, S. Takakusagi, K.-i. Shimizu and I. Takigawa, *ChemCatChem*, 2019, **11**, 4537–4547.
- 15 K. Takahashi, L. Takahashi, S. D. Le, T. Kinoshita, S. Nishimura and J. Ohyama, *J. Am. Chem. Soc.*, 2022, **144**, 15735–15744.

- 16 V. J. Ferreira, P. Tavares, J. L. Figueiredo and J. L. Faria, *Ind. Eng. Chem. Res.*, 2012, **51**, 10535–10541.
- 17 J. Xu, Y. Zhang, X. Xu, X. Fang, R. Xi, Y. Liu, R. Zheng and X. Wang, *ACS Catal.*, 2019, **9**, 4030–4045.
- 18 I. Kim, G. Lee, H. B. Na, J.-M. Ha and J. C. Jung, *Mol. Catal.*, 2017, **435**, 13–23.
- 19 P. Wang, G. Zhao, Y. Wang and Y. Lu, *Sci. Adv.*, 2017, **3**, e1603180.
- 20 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.
- 21 H. Yamada, C. Liu, S. Wu, Y. Koyama, S. Ju, J. Shiomi, J. Morikawa and R. Yoshida, *ACS Cent. Sci.*, 2019, **5**, 1717–1730.
- 22 Q. Zhou, Z.-Q. Wang, Z. Li, J. Wang, M. Xu, S. Zou, J. Yang, Y. Pan, X.-Q. Gong, L. Xiao and J. Fan, *ACS Catal.*, 2021, **11**, 14651–14659.