Volume 14
Number 17
17 March 2026
Pages 9731–10470

Journal of Materials Chemistry A

Materials for energy and sustainability

rsc.li/materials-a

ISSN 2050-7488

ROYAL SOCIETY OF CHEMISTRY

PAPER

Mikiya Fujii, Akihiko Kudo *et al.*

Machine-learning prediction of metal sulfide photocatalysts for sacrificial hydrogen evolution under visible light irradiation

Journal of Materials Chemistry A

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: J. Mater. Chem. A, 2026, 14, 9956

Machine-learning prediction of metal sulfide photocatalysts for sacrificial hydrogen evolution under visible light irradiation

Yuichi Yamaguchi, [ab] Fuga Kakami,[a] Ryuto Baba,[c] Wataru Takahara,[c] Yosuke Harashima, [cd] Tomoaki Takayama, [cd] Mikiya Fujii [*cde] and Akihiko Kudo [*ab]

The development of promising inorganic semiconductor photocatalysts for water splitting to produce green H_2 is required to achieve a sustainable society. Machine learning is expected to accelerate the exploration of novel inorganic semiconductor photocatalysts. We applied machine learning to explore novel metal sulfide photocatalysts for sacrificial H_2 evolution under visible light irradiation. A machine-learning model that exhibited good accuracy was successfully constructed using our original in-house dataset (not openly shared data) of metal sulfide photocatalysts developed by our group. Then, data on materials in the Inorganic Crystal Structure Database (ICSD) were input into the constructed machine-learning model, resulting in the identification of various metal sulfide candidates with high activities for H_2 evolution in the first screening. We selected Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) among the candidates for the second screening because many photocatalysts containing Cu(I) and/or Ag(I) ions and corner-shared MS_4 tetrahedra have been reported as visible-light-responsive photocatalysts for sacrificial H_2 evolution. Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) photocatalysts, prepared by a solid-state reaction, showed activities for sacrificial H_2 evolution under visible light irradiation. Thus, we developed novel visible-light-responsive metal sulfide photocatalysts for sacrificial H_2 evolution by employing machine learning on our original dataset.

Received 26th July 2025
Accepted 15th December 2025

DOI: 10.1039/d5ta06041a

rsc.li/materials-a

Introduction

Photocatalytic water splitting, as artificial photosynthesis, is a promising technology to produce low-cost green H_2 to solve resource, energy and environmental issues.[1–8] Since the report of the Honda–Fujishima effect of water splitting using a TiO_2 photoanode in the early 1970s,[9] various photocatalysts such as metal oxides,[1,2] metal (oxy)sulfides,[10–13] metal (oxy)nitrides[14–21] and polymeric materials[22,23] have been developed. In recent years, Domen and co-workers successfully developed a highly efficient Al-doped $SrTiO_3$ photocatalyst with a near-unity quantum yield, giving about 0.7% of a solar-to-hydrogen (STH) energy conversion efficiency.[24,25] In addition, they developed an efficient Z-schematic water splitting sheet composed of Rh and La co-doped $SrTiO_3$ as an H_2-evolving photocatalyst, Mo-doped $BiVO_4$ as an O_2-evolving photocatalyst and Au as a solid electron mediator, giving about 1.1% STH efficiency.[26] However, the efficiencies are insufficient for practical solar water splitting. Therefore, the development of novel efficient photocatalysts is still a challenging topic. Among various photocatalytic material groups, metal sulfides are attractive because many of them respond to visible light due to the valence band consisting of S 3p orbitals being located at a relatively negative energy level. We have reported various metal sulfide photocatalysts, including Cu-doped,[27] Ni-doped,[28] Pb and A (A = Cl, Br, and I)-codoped ZnS,[29] $AgInZn_7S_9$,[30] $NaInS_2$,[31] ZnS–$AgInS_2$,[32] ZnS–$CuInS_2$,[33] ZnS–$CuInS_2$–$AgInS_2$,[34,35] $A^I_2ZnA^{IV}S_4$ (A^I = Cu and Ag; A^{IV} = Sn and Ge),[36] $AGa_2In_3S_8$ (A = Cu and Ag),[37] $Zn_{1-2x}(CuGa)_xGa_2S_4$,[38] ZnS-$CuGaS_2$,[39] Ni-doped $AgGaS_2$,[40] $Cu_{1-x}Ag_xGaS_2$,[41] ZnS–CuCl,[11] $BaLaCuS_3$,[11] $ZnIn_2S_4$,[11] $ZnGa_xIn_{2-x}S_4$,[11] $MnGaInS_4$ [11] and Cu_3MS_4 (M = V, Nb and Ta),[11,42] for H_2 evolution in aqueous solutions containing S^{2-} and SO_3^{2-} ions under visible light irradiation. However, metal sulfide photocatalysts cannot be employed for water splitting as a single particle because they are not chemically stable due to self-oxidation in an aqueous medium under light irradiation. This drawback can be solved by utilizing them as photocathodes and H_2-evolving photocatalysts in photoelectrochemical and Z-schematic systems for artificial

[a]Department of Applied Chemistry, Faculty of Science, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan. E-mail: a-kudo@rs.tus.ac.jp
[b]Carbon Value Research Center, Research Institute for Science and Technology, Tokyo University of Science, 2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan
[c]Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma-shi, Nara 630-0192, Japan. E-mail: fujii.mikiya@ms.naist.jp
[d]Data Science Center, Nara Institute of Science and Technology, Ikoma-shi, Nara 630-0192, Japan
[e]Center for Material Research Platform, Nara Institute of Science and Technology, Ikoma-shi, Nara 630-0192, Japan

photosynthesis, respectively.[43,44] Various photoelectrochemical systems consisting of metal sulfide photocathodes have been constructed for water splitting under visible light irradiation.[39,41,42,45–50] Additionally, powdered Z-schematic water splitting under visible light irradiation has been achieved by using metal sulfides as a H_2-evolving photocatalyst, metal oxides as an O_2-evolving photocatalyst, and ionic or solid-state electron mediators such as Co complexes and a reduced graphene oxide.[39,46,51,52] Therefore, it is quite important to develop novel metal sulfide photocatalysts that are responsive to visible light.

So far, photocatalysts have mainly been developed by an intuitive empirical approach. Hence, it is expected that employing machine learning with an objective approach by data-driven science utilizing specialized physicochemical parameters will accelerate the exploration of promising novel photocatalysts. Material development utilizing a machine-learning method is generally referred to as materials informatics (MI) and has been applied to various materials systems, such as catalysts,[53,54] adsorbents,[55] magnetics[56] and batteries.[57,58] Recently, various functional materials, such as alloys, metal oxides, zeolites, metal complexes, metal–organic frameworks and polyoxometalates, have been developed for applications in which the properties of either the surface or the bulk play an important role.[54,59] In contrast, it remains challenging to apply MI to the field of photocatalysts, especially inorganic semiconductor photocatalysts for water splitting, because the photocatalytic reaction involves complicated factors such as photoabsorption, mobility of carriers, recombination, surface reactions, and so on. Although there are some reports of MI application to inorganic semiconductor photocatalysts, these works focus on optimization of photocatalytic reactions rather than developing new photocatalyst materials.[60] Application of MI to develop novel inorganic photocatalysts faces a significant barrier because it is difficult to employ open data from publications, as the photocatalytic activities have not been evaluated under a common set of experimental conditions. This means that the data cannot be compared with each other. Therefore, a large dataset needs to be prepared, including photocatalytic activities evaluated with the same experimental system, photocatalytic compounds and crystal structures, to construct a machine-learning model for training.

We have developed many photocatalysts of metal oxides and metal sulfides that are active for water splitting and sacrificial H_2 and O_2 evolutions.[2,61,62] We possess a large in-house dataset consisting of many photocatalyst compounds, crystal structures, preparation conditions, activities for water splitting, and sacrificial H_2 and O_2 evolutions evaluated under the same experimental conditions. Our dataset contains photocatalyst materials with not only high but also low activities, which is a great advantage for constructing an accurate predictive model. Recently, we have preliminarily demonstrated that a machine-learning model with acceptable accuracy was successfully constructed by employing our original dataset.[63,64] These reports are significant in terms of the effectiveness of our dataset.

In the present study, we constructed a machine-learning model based on our in-house dataset, which was then used to predict the activities for sacrificial H_2 evolution of metal sulfides listed in the Inorganic Crystal Structure Database (ICSD, a commercially licensed database). Then, we experimentally synthesized Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) from the candidates identified by machine learning and evaluated the activities for H_2 evolution under visible light irradiation.

Experimental

Machine-learning methods

We retrieved 8975 data entries of metal sulfide compounds from the ICSD[65,66] in order to confirm the material that had been experimentally identified. In addition, we used an in-house dataset of metal sulfide photocatalysts, including 357 experimental data points developed by our group, as the training data. A representative sample of the original dataset is included as a csv file named 'Example_of_original_dataset' in the SI. Each record in our dataset contains a material variable of chemical composition, 16 variables related to preparation conditions, dopants and cocatalysts, and an objective variable of activities for sacrificial H_2 evolution. The material variable of the chemical compositions in the in-house dataset of metal sulfide photocatalysts developed by our group was converted into numerical features by using XenonPy (version 0.6.5)[67] in order to process machine-learning algorithms. First, chemical compositions were converted into comp_dict, which is a dictionary of proportions by element, using Pymatgen.[68] Then, after normalization of the proportions, 232 compositional features on the basis of comp_dict were calculated, such as 58 features (atomic number, bond radius, van der Waals radius, electronegativity, and so on) of weighted average, weighted variance, and the minimum and maximum for each element.[69] The calculated features with high dimensions were projected to the figure with 2 dimensions by principal component analysis (PCA). The figure with axes called principal components shows the diversity of the data features.[70,71] Subsequently, LightGBM,[72] based on a gradient-boosting algorithm, was employed to construct the prediction model of sacrificial H_2 activities of metal sulfide photocatalysts because we focused on understanding the trend of the activities of metal sulfide photocatalysts in order to perform the rough material screening. Logarithmic transformations for the objective variables were used when constructing the machine-learning model. Double-cross-validation and evaluation metrics were carried out by Scikit-learn.[73] A 5-fold double cross-validation was used in this study.

Preparation of photocatalysts

Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn), which were predicted as new photocatalyst candidates by machine learning, were prepared by a solid-state reaction in evacuated quartz ampule tubes. Ag_2S (Kojundo Chemical, 99%), Cu_2S (Kojundo Chemical, 99%), CdS (Kojundo Chemical, 99.99%), SnS_2 (Kojundo Chemical, 99.9%) and GeS_2 (Kojundo Chemical, 99.99%) were used as starting materials. They were mixed in an agate mortar with 20% excess Ge for Ag_2CdGeS_4, 40% excess Ge

for Cu_2CdGeS_4, and 10% excess Cd and Sn for Cu_2CdSnS_4. The mixtures were heated under vacuum for 12 h at 923, 923 and 823 K for Ag_2CdGeS_4, Cu_2CdGeS_4 and Cu_2CdSnS_4, respectively.

Characterization of photocatalysts

X-ray diffraction (XRD; Rigaku; MiniFlex600) with Cu Kα radiation was used to identify the crystal structures of the obtained samples. Diffuse reflectance spectra were measured by using a UV-vis-NIR spectrometer (JASCO; UbeatV-570) with an integrating sphere and were converted to absorbance by the Kubelka–Munk method. Scanning electron microscopy (SEM; JEOL; JSM-6700F) was used to observe the morphologies of the particles. Photoelectron yield spectroscopy (PYS; Bunkoukeiki; BIP-KV100) was conducted under vacuum to determine the ionization potential of the samples.

Density functional theory calculations

Density functional theory (DFT) calculations were carried out using the CASTEP code.[74] ICSD #152753, #26150 and #619773 were employed for Ag_2CdGeS_4, Cu_2CdGeS_4 and Cu_2CdSnS_4, respectively. In geometric optimization, the Vanderbilt-type ultrasoft pseudopotentials,[75] the generalized gradient approximation (GGA)[76] and the Perdew–Burke–Ernzerhof (PBE)[77] as exchange–correlation functional were used under periodic boundary conditions. 600 eV of cutoff energy and a *k*-point mesh with a typical spacing of 0.07 $Å^{-1}$ given by the Monkhorst–Pack method[78] were used. The limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS) algorithm[79] was employed for geometric optimization of the supercells. Details of the conditions of the convergence criteria are summarized in the SI (Table S1). The Heyd–Scuseria–Ernzerhof (HSE03)[80] hybrid functional was used to analyze the density of states and the band structures of the optimized crystal structures. The valence electronic configurations for Cu, Ag, Cd, Ge, Sn and S were $3d^{10}4s^1$, $4d^{10}5s^1$, $4d^{10}5s^2$, $4s^24p^2$, $5s^25p^2$ and $3s^23p^4$, respectively. The number of electrons in the primitive unit cell were 156 and 124 for [Ag_2CdGeS_4] and [Cu_2CdMS_4 (M = Ge and Sn)], respectively.

Photocatalytic reactions

Photocatalytic H_2 evolution was carried out using a top-irradiation cell with a Pyrex window in a gas-closed circulation system. The sample was dispersed in an aqueous solution (150 mL) of 0.1 mol L^{-1} Na_2S ($Na_2S \cdot 9H_2O$; Kanto Chemical, 98%) and 0.5 mol L^{-1} K_2SO_3 (Kanto Chemical; 95%) as sacrificial reagents. The pH of the aqueous solution containing S^{2-} and SO_3^{2-} ions was 13. A Ru cocatalyst as an H_2-evolving site[36] was loaded on the prepared sulfide photocatalysts by *in situ* photodeposition in an aqueous solution of $RuCl_3 \cdot nH_2O$ (Tanaka Kikinzoku; 39% as Ru in $RuCl_3 \cdot nH_2O$) as a source of a cocatalyst. A 300 W Xe-arc lamp (PerkinElmer; Cermax PE300BF) was used as a light source. The wavelength of light irradiation was controlled by using a cut-off filter (HOYA; L42, λ > 420 nm). The amount of evolved H_2 was quantified using an online gas chromatograph (Shimadzu; GC-8A, MS-5A column, TCD, Ar carrier).

Apparent quantum yields (AQY) were measured using a 300 W Xe-arc lamp (PerkinElmer; Cermax PE300BF) and a 100 W Xe-arc lamp (Asahi Spectra; LAX 102) with band pass filters (Asahi Spectra). The photon flux of monochromatic light was measured using a silicon diode head (OPHIR; PD300-UV head) with a NOVA display. AQY was calculated by the following eqn (1).

$$[\text{AQY }\%] = 100 \times [\text{the number of reacted electrons}]/[\text{the number of incident photons}] = 100 \times [\text{the number of evolved } H_2 \text{ molecules} \times 2]/[\text{the number of incident photons}] \quad (1)$$

Results and discussion

Machine-learning-based prediction of metal sulfide photocatalyst candidates for sacrificial H_2 evolution under visible light irradiation

A flowchart for screening of metal sulfide photocatalysts for sacrificial H_2 evolution under visible light irradiation (λ > 420 nm) by machine learning is shown in Fig. 1. First, the prediction model of activities for H_2 evolution was constructed by machine learning on the basis of our original in-house dataset, as shown in Fig. 1(1). Fig. 2 shows a plot of the predicted *versus* experimental activities, which was used to evaluate the constructed machine-learning model using double cross-validation. The coefficient of determination (R^2) was more than 0.5, indicating that the model was constructed with acceptable accuracy for the initial (first/rough) materials screening, as shown in Fig. 1(1)–(3). The acceptable accuracy described here means that the constructed machine-learning model is suitable for understanding the trend of activities for H_2 evolution of metal sulfide photocatalysts, which leads to a subsequent (second/refined) screening based on chemical insights, as shown in Fig. 1(4). Next, we selected candidate materials for screening from the ICSD of metal sulfide compounds based on their similarity in the feature space to our in-house dataset of metal sulfide photocatalysts, as shown in Fig. 1(2). Fig. 3 shows the material space of metal sulfides in the ICSD and our in-house original dataset as red and blue dots, respectively. Then, we selected 765 metal sulfide compounds, located within the blue square in Fig. 3, as candidate materials for screening. This region was selected because it included many in-house experimental data that were used as training data, given that it was difficult to predict the activities of materials outside the range of the training data. Subsequently, the activities of the candidate sulfides for sacrificial H_2 evolution were predicted by using the constructed machine-learning model, as shown in Fig. 1(3) and 2. The selected metal sulfide candidate materials with their predicted activities are shown in Fig. 4. The 200 material sulfide candidates with their predicted activities were obtained, as shown in the 'list_of_candidates' csv file (SI). Thus, the metal sulfide photocatalyst candidates were identified in the first stage of screening by MI. As shown in Fig. 1(4), we selected three metal sulfides, Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn), for the second stage of screening based on empirical insights, given that a number of metal sulfides containing Cu(I) and/or Ag(I)

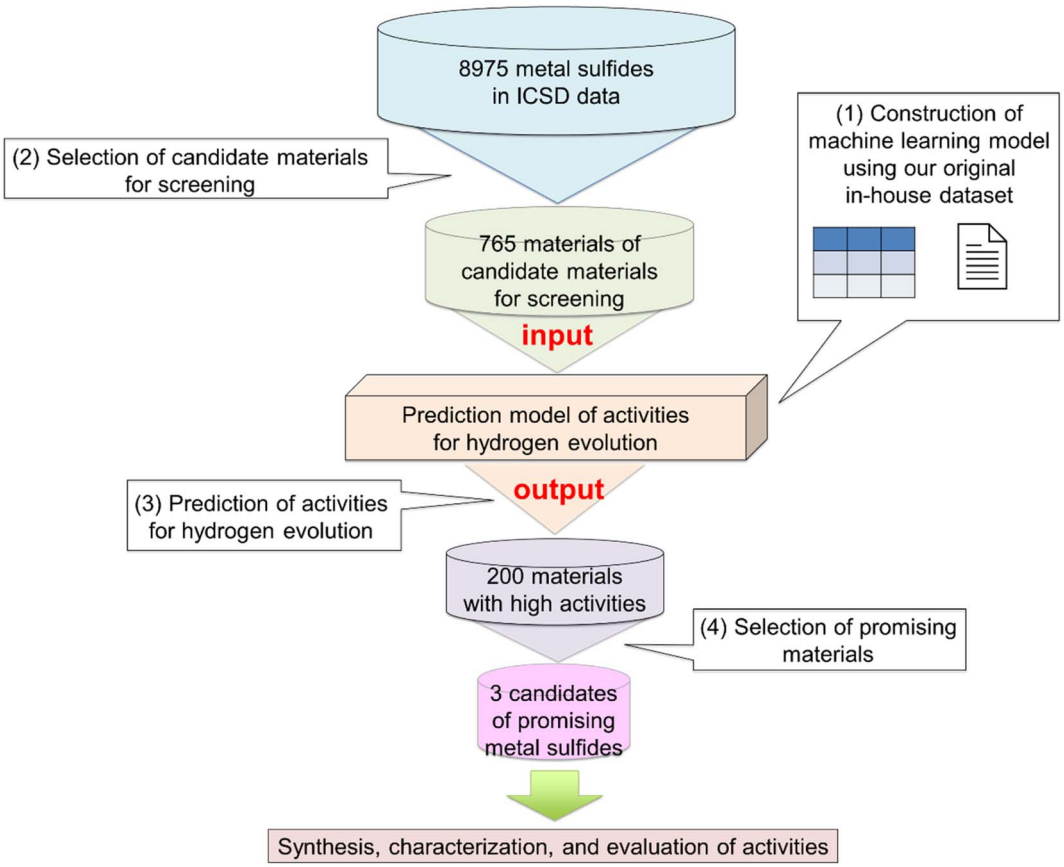

Fig. 1 Flowchart of exploring metal sulfides for H_2 evolution by machine-learning.

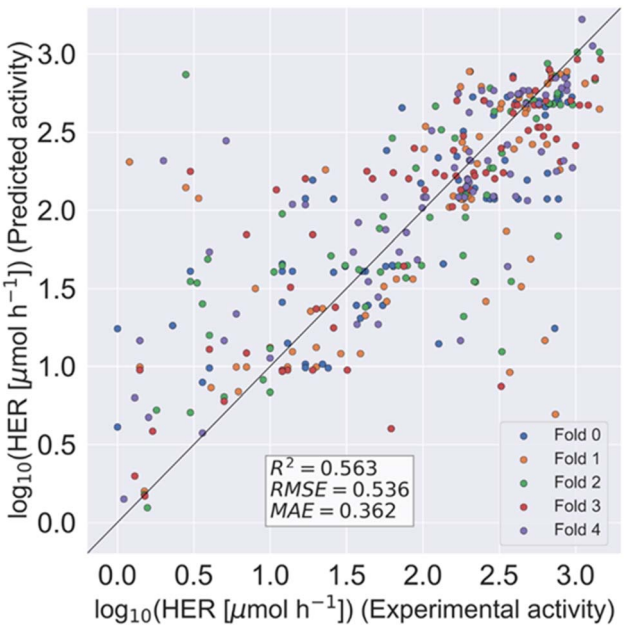

Fig. 2 Comparison between predicted activities obtained by the machine-learning model and experimental activities for sacrificial H_2 evolution under visible light irradiation ($\lambda > 420$ nm) on the basis of our original experimental dataset. In the 5-fold double cross-validation, the validation results from fold 0 to fold 4 are shown as blue, orange, green, red and purple circles, respectively.

ions and corner-shared MS_4 tetrahedra have been reported as promising visible-light-responsive photocatalysts for sacrificial H_2 evolution.[36,39,41]

Preparation and characterization of Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) photocatalysts

Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) were synthesized by a conventional solid-state reaction. XRD patterns revealed that orthorhombic Ag_2CdGeS_4, with the enargite structure, and tetragonal Cu_2CdSnS_4, with the stannite structure, were obtained in a single phase (Fig. S1). Orthorhombic Cu_2CdGeS_4, with the enargite structure in an almost single phase, was also obtained with a negligible amount of Cd_4GeS_6. We optimized the excess amount of GeS_2 (0–50 mol% ex.) with the aim of obtaining a single phase of Cu_2CdGeS_4. It was revealed that the lowest amount of Gd_4GeS_6 in the sample was achieved when a 40 mol% excess amount of GeS_2 was used, although the single phase of Cu_2CdGeS_4 was not obtained. Particulate morphologies were observed in the SEM images (Fig. S2). Sintered large particles with a size of several tens of micrometers were observed on Ag_2CdGeS_4, and aggregations of primary particles with a size of several micrometers were observed on Cu_2CdMS_4 (M = Ge and Sn). Diffuse reflectance spectra of the obtained samples are shown in Fig. 5(a). The band gaps of Ag_2CdGeS_4, Cu_2CdGeS_4 and Cu_2CdSnS_4 were estimated to be 2.21, 1.86 and

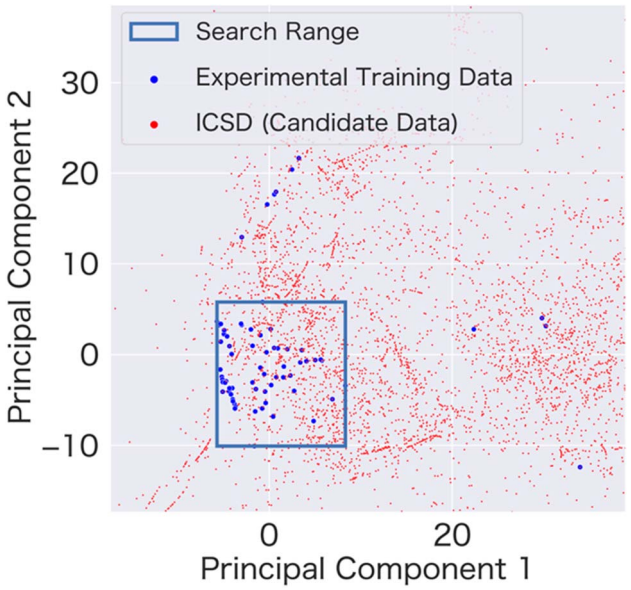

Fig. 3 Material space of metal sulfides in the ICSD and our original dataset by PCA. Principal components 1 and 2 (with no units), which form the axes, are new coordinate axes constructed by performing linear combinations of the original variables.

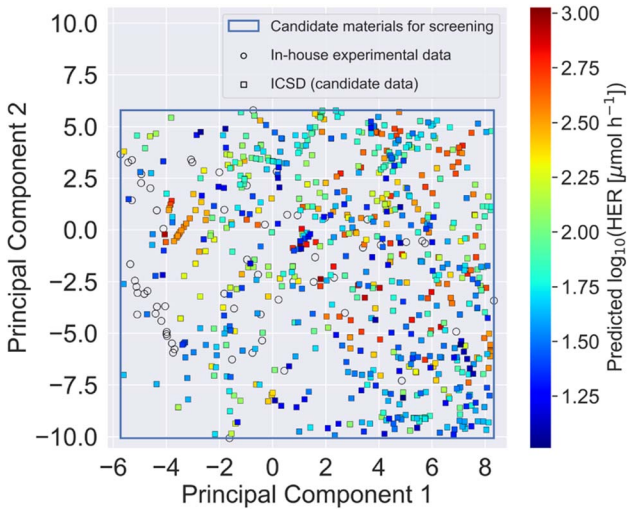

Fig. 4 Material space of metal sulfides in the ICSD and our original experimental dataset with predicted activities for sacrificial H_2 evolution under visible light irradiation (λ > 420 nm) by PCA. Principal components 1 and 2 (with no units), which form the axes, are new coordinate axes constructed by performing linear combinations of original variables. Open circles are in-house experimental data. Squares are the data from the ICSD, and the color indicates a predicted activity corresponding to the right vertical axis.

1.28 eV, respectively, from the absorption edges. It is notable that Cu_2CdSnS_4 absorbs over the whole range of visible light. PYS spectra were measured under vacuum to determine the ionization potentials, which can generally be assigned to the top of the valence band level of the samples. As shown in Fig. 5(b), the potentials were estimated to be −5.79 eV (*vs.* vacuum level) for Ag_2CdGeS_4 and −5.18 eV (*vs.* vacuum level) for Cu_2CdMS_4 (M = Ge and Sn), indicating that the valence band of Ag_2CdGeS_4 formed at a more positive potential than those of Cu_2CdMS_4 (M = Ge and Sn). This result is reasonable because the Ag 4d orbitals generally form the valence band at a more positive level than Cu 3d orbitals, as observed for other metal sulfide photocatalysts containing Cu(I) and Ag(I) ions.[41] DFT calculations were also performed to investigate the band structures of Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn), as shown in Fig. 6 and 7. All samples showed the direct transition property, and the calculated band gaps for Ag_2CdGeS_4, Cu_2CdGeS_4 and Cu_2CdSnS_4 were 1.89, 1.65 and 0.95 eV, respectively. The order of the band gaps agreed well with the experimental data estimated from absorption edges of diffuse reflectance spectra, as shown in Fig. 5(a). The partial density of states (PDOS) of Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) are shown in Fig. 7. The valence band maxima mainly consisted of Ag 4d + S 3p hybrid orbitals for Ag_2CdGeS_4 and Cu 3d + S 3p hybrid orbitals for Cu_2CdMS_4 (M = Ge and Sn), while the conduction band minima mainly consisted of Cd 5s5p + Ge 4s4p hybrid orbitals for Ag_2CdGeS_4 and Cu_2CdGeS_4. We previously reported DFT calculations for a Cu_2ZnGeS_4 photocatalyst with the stannite structure that was active for sacrificial H_2 evolution under visible light irradiation.[36] The results showed that Zn 4s4p orbitals were located at a higher energy than Ge 4s4p orbitals, indicating that the contribution of Zn 4s4p to the conduction band of Cu_2ZnGeS_4 was small. In contrast, the contribution of Cd 5s5p to the conduction band was relatively large for Ag_2CdGeS_4 and Cu_2CdGeS_4. The conduction band minima mainly consisted of Cd 5s5p + Sn 5s5p hybrid orbitals for Cu_2CdSnS_4, which were located at a lower energy than the Cd 5s5p + Ge 4s4p hybrid orbitals, leading to the narrow band gap of Cu_2CdSnS_4. Fig. 8 shows the band structures determined by diffuse reflectance spectra, the ionization potentials by PYS measurements, and the DFT calculations. The order of driving forces for the reduction of water was $Cu_2CdGeS_4 > Ag_2CdGeS_4 \gg Cu_2CdSnS_4$. The DRS, PYS and PDOS results suggest that Sn 5s5p orbitals contribute to the formation of the conduction band of Cu_2CdSnS_4 at a more positive level than for Ag_2CdSnS_4 and Cu_2CdGeS_4, leading to the decreased driving force for water reduction.

Photocatalytic H_2 evolution over Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) photocatalysts in the presence of sacrificial reagents under visible light irradiation

Photocatalytic H_2 evolution over Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) photocatalysts from an aqueous solution containing S^{2-} and SO_3^{2-} ions as sacrificial reagents, was investigated under visible light irradiation. H_2 continuously evolved for all samples for 3 h, without a noticeable deactivation, as shown in Fig. 9. As mentioned in the previous section, Cu_2CdGeS_4 contained a small amount of Cd_4GeS_6 as an impurity. However, the single Cd_4GeS_6 showed very low activity (16 μmol h^{-1}) compared to Cu_2CdGeS_4 (369 μmol h^{-1}). Therefore, the obtained Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) represent new visible-light-responsive metal sulfide photocatalysts for H_2 evolution. The Cu_2CdSnS_4 photocatalyst continuously evolved

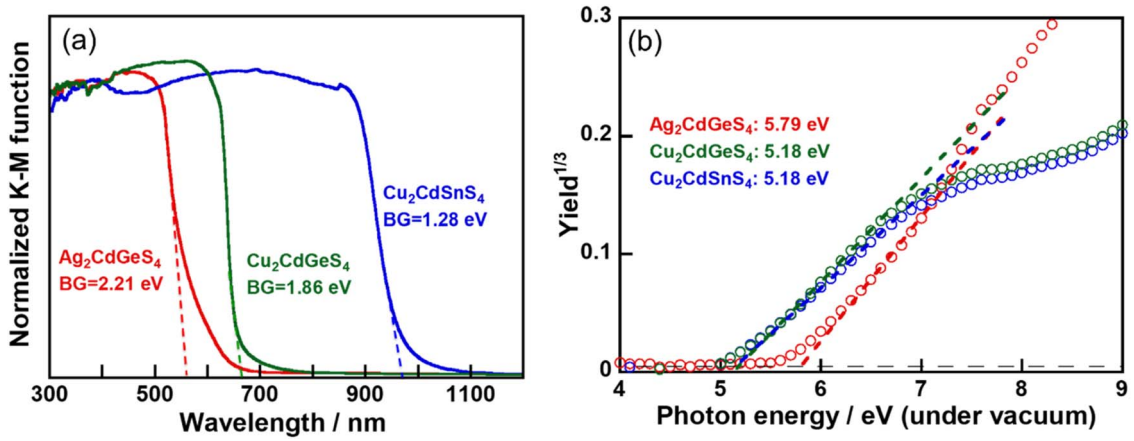

Fig. 5 (a) Diffuse reflectance spectra and (b) PYS spectra of Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) prepared by a solid-state reaction.

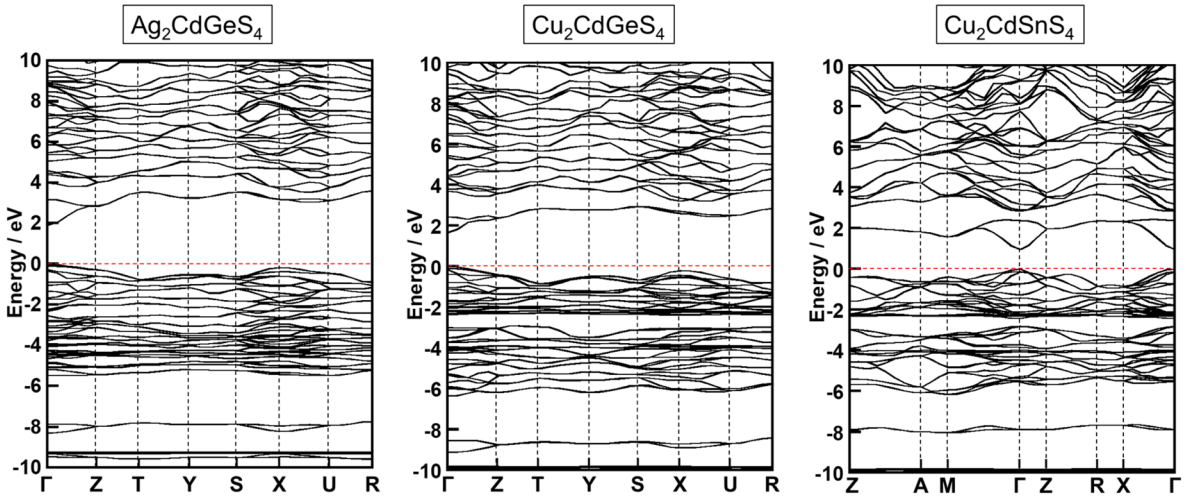

Fig. 6 Band structures of Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn).

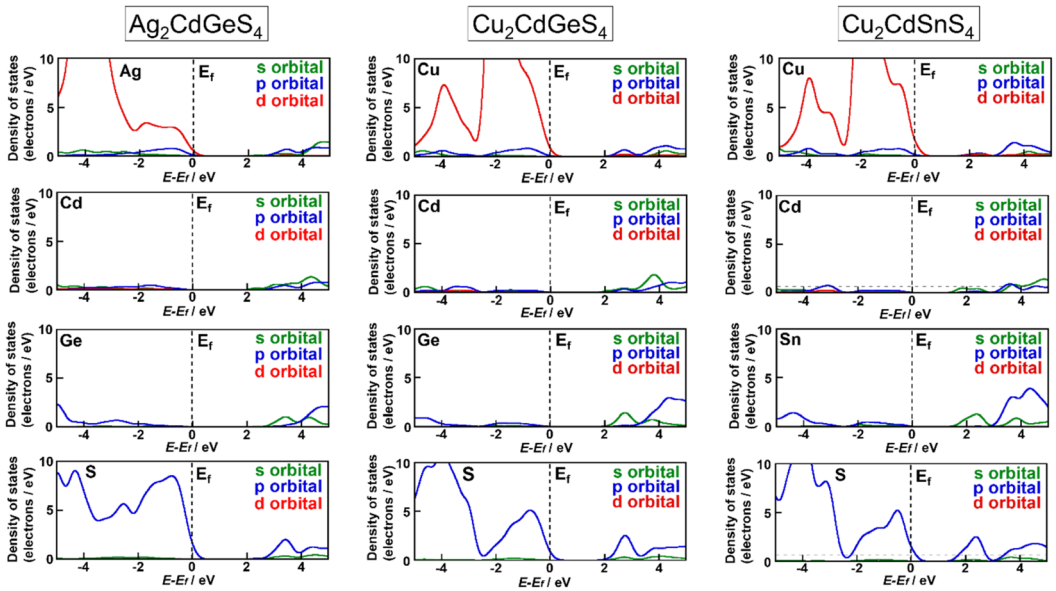

Fig. 7 Density of states of Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn).

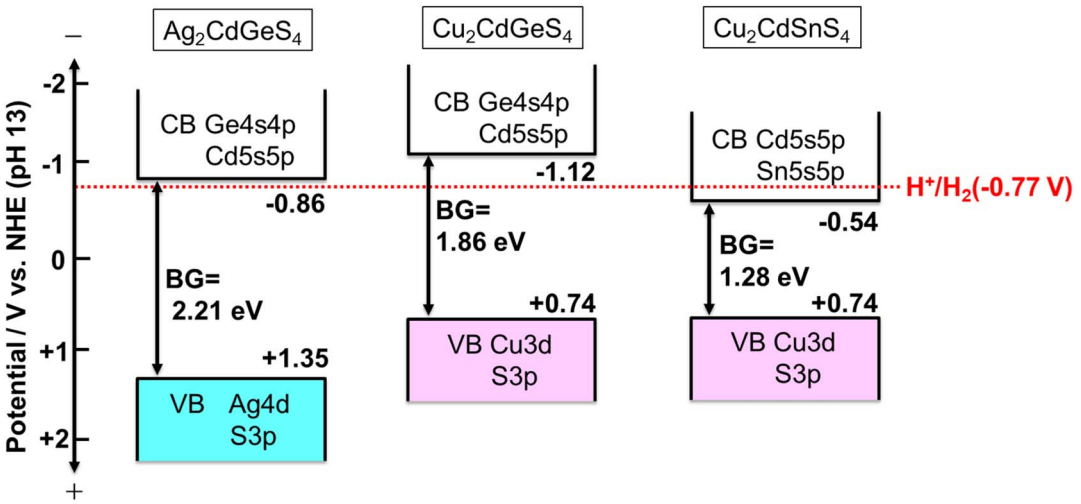

Fig. 8 Band structures of Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn).

H_2 for 22 h (Fig. S3). The turnover numbers of the molar quantity of reacted electrons for H_2 evolution to that of the photocatalyst host were 1.8 at 3 h, 3.1 at 3 h and 1.2 at 22 h for Ag_2CdGeS_4, Cu_2CdGeS_4 and Cu_2CdSnS_4, respectively. The turnover numbers were larger than unity, indicating that H_2 evolution proceeded photocatalytically. Cu_2CdGeS_4 showed the highest activity among the samples. Fig. 10 shows action spectra of sacrificial H_2 evolution over the Ru (0.5 wt%)-loaded Ag_2CdGeS_4, Cu_2CdGeS_4 and Cu_2CdSnS_4 photocatalysts. Ag_2CdGeS_4 and Cu_2CdGeS_4 responded to 570 and 780 nm, respectively. The onsets of the action spectra of Ag_2CdGeS_4 and Cu_2CdGeS_4 agreed well with those of the diffuse reflectance spectra, indicating that H_2 evolution proceeded by band gap excitation. Notably, Cu_2CdSnS_4 responded to near-infrared light at 900 nm. AQYs were 2.6% at 470 nm for Ag_2CdGeS_4, 3.4% at

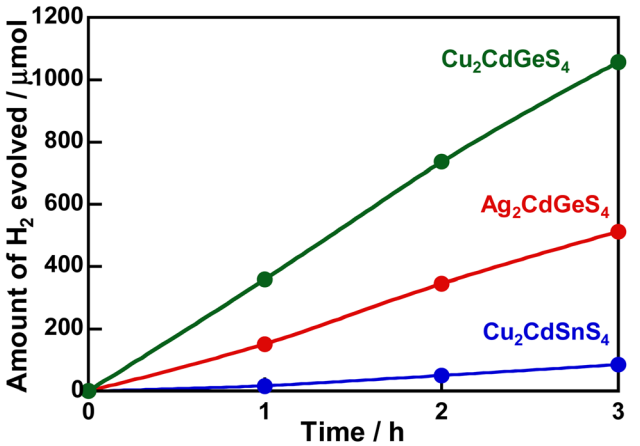

Fig. 9 H_2 evolution over Ru-loaded Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) photocatalysts prepared by a solid-state reaction from an aqueous solution containing sacrificial reagents under visible light irradiation. Photocatalyst: 0.3 g, cocatalyst: photodeposition *in situ*, reactant solution: 0.5 mol per L K_2SO_3 and 0.1 mol per L Na_2S aqueous solution (120 mL), cell: top-irradiation cell with a Pyrex window, light source: 300 W Xe lamp with a cut-off filter (HOYA: L42).

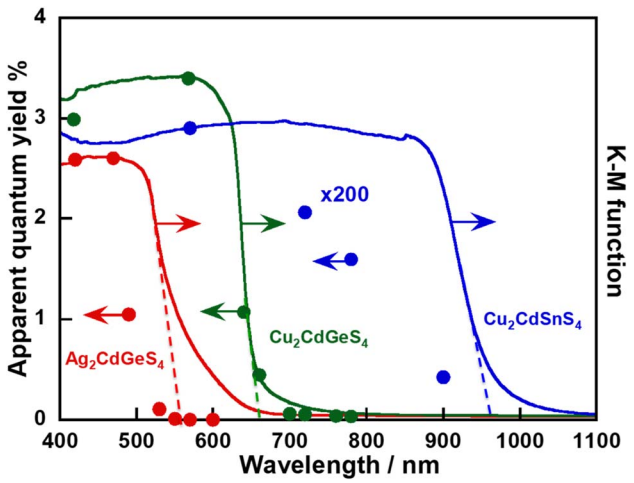

Fig. 10 Action spectra of sacrificial H_2 evolution over the Ru (0.5 wt%)-loaded Ag_2CdGeS_4, Cu_2CdMS_4 (M = Ge and Sn) photocatalysts. Photocatalyst: 0.3 g, cocatalyst: PD *in situ*, reactant solution: 0.5 mol per L K_2SO_3 and 0.1 mol per L Na_2S aqueous solution (120 mL), cell: top-irradiation cell with a Pyrex window, light source: 300 W Xe lamp with band-pass filters.

570 nm for Cu_2CdGeS_4 and 0.015% at 570 nm for Cu_2CdSnS_4. The order of AQYs was $Cu_2CdGeS_4 > Ag_2CdGeS_4 \gg Cu_2CdSnS_4$. Here, the order of the driving force of reduction of water was $Cu_2CdGeS_4 > Ag_2CdGeS_4 \gg Cu_2CdSnS_4$, judging from the position of the conduction band minima, as shown in Fig. 8. Therefore, Cu_2CdGeS_4, with the largest driving force for reduction of water, showed the highest activity among the samples.

Conclusions

We have successfully developed Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) as new visible-light-responsive metal sulfide photocatalysts for sacrificial H_2 evolution by utilizing machine

learning. The machine-learning model with acceptable accuracy was constructed by utilizing our original dataset of metal sulfide photocatalysts developed by our group. The features of material sulfides in the ICSD and our original dataset were plotted in a material space with two dimensions by using XenonPy and principal component analysis. We predicted the activities of 765 metal sulfides in the ICSD inside the material space (where many experimental data used as the training data existed) using the constructed machine-learning model, resulting in the identification of metal sulfide candidates for H_2 evolution. We experimentally synthesized Ag_2CdGeS_4 and Cu_2CdMS_4 (M = Ge and Sn) from the candidates because metal sulfides containing Cu(I) and/or Ag(I) ions and corner-shared MS_4 tetrahedra have been reported as promising photocatalysts for sacrificial H_2 evolution under visible light irradiation. The band gaps of Ag_2CdGeS_4, Cu_2CdGeS_4 and Cu_2CdSnS_4 were estimated to be 2.21, 1.86 and 1.28 eV, respectively, from the diffuse reflectance spectra. DFT calculations indicated that the conduction band minima of Ag_2CdGeS_4 and Cu_2CdGeS_4 mainly consisted of Cd 5s5p + Ge 4s4p hybrid orbitals, and Cu_2CdSnS_4 were mainly composed of Cd 5s5p + Sn 5s5p hybrid orbitals. Ag_2CdGeS_4, Cu_2CdGeS_4 and Cu_2CdSnS_4, prepared by a solid-state reaction, showed activities for sacrificial H_2 evolution under visible light irradiation up to 570, 780 and 900 nm, respectively. Among the prepared photocatalysts, Cu_2CdGeS_4 gave the highest AQY, possibly because it possessed a larger driving force for the reduction of water compared to Ag_2CdGeS_4 and Cu_2CdSnS_4. In this study, exploration of novel visible-light-responsive metal sulfides for sacrificial H_2 evolution by machine learning has been achieved. It is expected that this study will contribute to the accelerated exploration of novel metal semiconductor photocatalysts not only in the metal sulfide group but also in other material groups, such as metal oxides.

Author contributions

A. K. and M. F. designed and supervised the research. F. K. carried out the preparation, characterization and evaluation of the H_2-evolving activity of metal sulfide photocatalysts. R. B. and W. T. provided the candidates of metal sulfide photocatalysts for H_2 evolution by machine learning. Y. Y., F. K. and T. T. performed the DFT calculation. Y. Y. wrote the manuscript. Y. Y., F. K., R. B, W. T., Y. H., T. T., M. F. and A. K. discussed and reviewed the content of the research.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the supplementary information (SI). Supplementary information is available. See DOI: https://doi.org/10.1039/d5ta06041a.

Acknowledgements

This work was supported by JSPS KAKENHI, Grant Numbers 22K14770 and 23H00248, and the "Program for Promoting Research on the Supercomputer Fugaku" (realization of innovative light energy conversion materials utilizing the supercomputer Fugaku, Grant Number JPMXP1020210317). The computations were partly conducted using the facilities of the Supercomputer Center at the Institute for Solid State Physics at the University of Tokyo.

References

1 F. E. Osterloh, *Chem. Mater.*, 2008, **20**, 35–54.
2 A. Kudo and Y. Miseki, *Chem. Soc. Rev.*, 2009, **38**, 253–278.
3 R. Abe, *J. Photochem. Photobiol., C*, 2010, **11**, 179–209.
4 T. Hisatomi, J. Kubota and K. Domen, *Chem. Soc. Rev.*, 2014, **43**, 7520–7535.
5 T. Setoyama, T. Takewaki, K. Domen and T. Tatsumi, *Faraday Discuss.*, 2017, **198**, 509–527.
6 Q. Wang and K. Domen, *Chem. Rev.*, 2020, **120**, 919.
7 S. Nandy, T. Hisatomi, T. Takata, T. Setoyama and K. Domen, *J. Mater. Chem. A*, 2023, **11**, 20470–20479.
8 J. Xiao, T. Hisatomi and K. Domen, *Acc. Chem. Res.*, 2023, **56**, 878–888.
9 A. Fujishima and K. Honda, *Nature*, 1972, **238**, 37–38.
10 G. Ma, S. Chen, Y. Kuang, S. Akiyama, T. Hisatomi, M. Nakabayashi, N. Shibata, M. Katayama, T. Minegishi and K. Domen, *J. Phys. Chem. Lett.*, 2016, **7**, 3892–3896.
11 T. Takayama, I. Tsuji, N. Aono, M. Harada, T. Okuda, A. Iwase, H. Kato and A. Kudo, *Chem. Lett.*, 2017, **46**, 616–619.
12 S. Sun, T. Hisatomi, Q. Wang, S. Chen, G. Ma, J. Liu, S. Nandy, T. Minegishi, M. Katayama and K. Domen, *ACS Catal.*, 2018, **8**, 1690–1696.
13 Q. Wang, M. Nakabayashi, T. Hisatomi, S. Sun, S. Akiyama, Z. Wang, Z. Pan, X. Xiao, T. Watanabe, T. Yamada, N. Shibata, T. Takata and K. Domen, *Nat. Mater.*, 2019, **18**, 827–832.
14 K. Maeda, T. Takata, M. Hara, N. Saito, Y. Inoue, H. Kobayashi and K. Domen, *J. Am. Chem. Soc.*, 2005, **127**, 8286–8287.
15 K. Maeda, K. Teramura, D. Lu, T. Takata, N. Saito, Y. Inoue and K. Domen, *Nature*, 2006, **440**, 295.
16 Y. Lee, H. Terashima, Y. Shimodaira, K. Teramura, M. Hara, H. Kobayashi, K. Domen and M. Yashima, *J. Phys. Chem. C*, 2007, **111**, 1042–1048.
17 K. Maeda, D. Lu and K. Domen, *Chem.–Eur. J.*, 2013, **19**, 4986–4991.
18 C. Pan, T. Takata and K. Domen, *Chem.–Eur. J.*, 2016, **22**, 1854–1862.
19 Z. Wang, Y. Inoue, T. Hisatomi, R. Ishikawa, Q. Wang, T. Takata, S. Chen, N. Shibata, Y. Ikuhara and K. Domen, *Nat. Catal.*, 2018, **1**, 756–763.
20 H. Li, J. Xiao, J. J. M. Vequizo, T. Hisatomi, M. Nakabayashi, Z. Pan, N. Shibata, A. Yamakata, T. Takata and K. Domen, *ACS Catal.*, 2022, **12**, 10179–10185.

21 K. Chen, J. Xiao, J. J. M. Vequizo, T. Hisatomi, Y. Ma, M. Nakabayashi, T. Takata, A. Yamakata, N. Shibata and K. Domen, *J. Am. Chem. Soc.*, 2023, **145**, 3839–3843.
22 X. Wang, K. Maeda, A. Thomas, K. Takanabe, G. Xin, J. M. Carlsson, K. Domen and M. Antonietti, *Nat. Mater.*, 2009, **8**, 76–80.
23 Y. Bai, K. Hippalgaonkar and R. S. Sprick, *J. Mater. Chem. A*, 2021, **9**, 16222–16232.
24 T. Takata, J. Jiang, Y. Sakata, M. Nakabayashi, N. Shibata, V. Nandal, K. Seki, T. Hisatomi and K. Domen, *Nature*, 2020, **581**, 411–414.
25 H. Nishiyama, T. Yamada, M. Nakabayashi, Y. Maehara, M. Yamaguchi, Y. Kuromiya, H. Tokudome, S. Akiyama, T. Watanabe, R. Narushima, S. Okunaka, N. Shibata, T. Takata, T. Hisatomi and K. Domen, *Nature*, 2021, **598**, 304–307.
26 Q. Wang, T. Hisatomi, Q. Jia, H. Tokudome, M. Zhong, C. Wang, Z. Pan, T. Takata, M. Nakabayashi, N. Shibata, Y. Li, I. D. Sharp, A. Kudo, T. Yamada and K. Domen, *Nat. Mater.*, 2016, **15**, 611–615.
27 A. Kudo and M. Sekizawa, *Catal. Lett.*, 1999, **58**, 241–243.
28 A. Kudo and M. Sekizawa, *Chem. Commun.*, 2000, 1371–1372.
29 I. Tsuji and A. Kudo, *J. Photochem. Photobiol., A*, 2003, **156**, 249–252.
30 A. Kudo, I. Tsuji and H. Kato, *Chem. Commun.*, 2002, 1958–1959.
31 A. Kudo, A. Nagane, I. Tsuji and H. Kato, *Chem. Lett.*, 2002, **31**, 882–883.
32 I. Tsuji, H. Kato, H. Kobayashi and A. Kudo, *J. Am. Chem. Soc.*, 2004, **126**, 13406–13413.
33 I. Tsuji, H. Kato, H. Kobayashi and A. Kudo, *J. Phys. Chem. B*, 2005, **109**, 7323–7329.
34 I. Tsuji, H. Kato and A. Kudo, *Angew. Chem., Int. Ed.*, 2005, **117**, 3631–3634.
35 I. Tsuji, H. Kato and A. Kudo, *Chem. Mater.*, 2006, **18**, 1969–1975.
36 I. Tsuji, Y. Shimodaira, H. Kato, H. Kobayashi and A. Kudo, *Chem. Mater.*, 2010, **22**, 1402–1409.
37 H. Kaga, K. Saito and A. Kudo, *Chem. Commun.*, 2010, **46**, 3779–3781.
38 H. Kaga and A. Kudo, *J. Catal.*, 2014, **310**, 31–36.
39 T. Kato, Y. Hakari, S. Ikeda, Q. Jia, A. Iwase and A. Kudo, *J. Phys. Chem. Lett.*, 2015, **6**, 1042–1047.
40 K. Yamato, A. Iwase and A. Kudo, *ChemSusChem*, 2015, **8**, 2902–2906.
41 H. Kaga, Y. Tsutsui, A. Nagane, A. Iwase and A. Kudo, *J. Mater. Chem. A*, 2015, **3**, 21815–21823.
42 S. Ikeda, N. Aono, A. Iwase, H. Kobayashi and A. Kudo, *ChemSusChem*, 2019, **12**, 1977–1983.
43 Y. Wang, H. Suzuki, J. Xie, O. Tomita, D. J. Martin, M. Higashi, D. Kong, R. Abe and J. Tang, *Chem. Rev.*, 2018, **118**, 5201–5241.
44 S. Yoshino, T. Takayama, Y. Yamaguchi, A. Iwase and A. Kudo, *Acc. Chem. Res.*, 2022, **55**, 966–977.
45 S. Ikeda, T. Nakamura, S. M. Lee, T. Yagi, T. Harada, T. Minegishi and M. Matsumura, *ChemSusChem*, 2011, **4**, 262–268.
46 A. Iwase, S. Yoshino, T. Takayama, Y. H. Ng, R. Amal and A. Kudo, *J. Am. Chem. Soc.*, 2016, **138**, 10260.
47 T. Hayashi, R. Niishiro, H. Ishihara, M. Yamaguchi, Q. Jia, Y. Kuang, T. Higashi, A. Iwase, T. Minegishi, T. Yamada, K. Domen and A. Kudo, *Sustainable Energy Fuels*, 2018, **2**, 2016.
48 Y. Kageshima, S. Shiga, T. Ode, F. Takagi, H. Shiiba, M. T. Htay, Y. Hashimoto, K. Teshima, K. Domen and H. Nishikiori, *J. Am. Chem. Soc.*, 2021, **143**, 5698–5708.
49 H. Fukai, K. Nagatsuka, Y. Yamaguchi, A. Iwase and A. Kudo, *J. ECS Solid State Sci. Technol.*, 2022, **11**, 063002.
50 T. Takayama, A. Iwase and A. Kudo, *ACS Appl. Mater. Interfaces*, 2024, **16**, 36423–36432.
51 K. Iwashina, A. Iwase, Y. H. Ng, R. Amal and A. Kudo, *J. Am. Chem. Soc.*, 2015, **137**, 604–607.
52 S. Yoshino, A. Iwase, Y. H. Ng, R. Amal and A. Kudo, *ACS Appl. Energy Mater.*, 2020, **3**, 5684–5692.
53 A. F. Zahrt, J. J. Henle, B. T. Rose, Y. Wang, W. T. Darrow and S. E. Denmark, *Science*, 2019, **363**, eaau5631.
54 T. Toyao, Z. Maeno, S. Takakusagi, T. Kamachi, I. Takigawa and K.-i. Shimizu, *ACS Catal.*, 2020, **10**, 2260–2297.
55 P. G. Boyd, A. Chidambaram, E. García-Díez, C. P. Ireland, T. D. Daff, R. Bounds, A. Gładysiak, P. Schouwink, S. M. Moosavi, M. M. Maroto-Valer, J. A. Reimer, J. A. R. Navarro, T. K. Woo, S. Garcia, K. C. Stylianou and B. Smit, *Nature*, 2019, **576**, 253–256.
56 Y. Harashima, K. Tamai, S. Doi, M. Matsumoto, H. Akai, N. Kawashima, M. Ito, N. Sakuma, A. Kato, T. Shoji and T. Miyake, *Phys. Rev. Mater.*, 2021, **5**, 013806.
57 M.-F. Ng, J. Zhao, Q. Yan, G. J. Conduit and Z. W. Seh, *Nat. Mach. Intell.*, 2020, **2**, 161–170.
58 Y. Nomura, K. Yamamoto, M. Fujii, T. Hirayama, E. Igaki and K. Saitoh, *Nat. Commun.*, 2020, **11**, 2824.
59 B. Madika, A. Saha, C. Kang, B. Buyantogtokh, J. Agar, C. M. Wolverton, P. Voorhees, P. Littlewood, S. Kalinin and S. Hong, *ACS Nano*, 2025, **19**, 27116–27158.
60 H. Mai, T. C. Le, D. Chen, D. A. Winkler and R. A. Caruso, *Chem. Rev.*, 2022, **122**, 13478–13515.
61 A. Kudo, H. Kato and I. Tsuji, *Chem. Lett.*, 2004, **33**, 1534–1539.
62 Y. Yamaguchi and A. Kudo, *Front. Energy*, 2021, **15**, 568–576.
63 Y. Harashima, T. Miyake, R. Baba, T. Takayama, S. Takasuka, Y. Shigeta, Y. Yamaguchi, A. Kudo and M. Fujii, *arXiv*, 2024, preprint, arXiv:2408.08539, DOI: 10.48550/arXiv.2408.08539.
64 W. Takahara, R. Baba, Y. Harashima, T. Takayama, S. Takasuka, Y. Yamaguchi, A. Kudo and M. Fujii, *ACS Omega*, 2025, **10**, 14626–14639.
65 G. Bergerhoff, R. Hundt, R. Sievers and I. D. Brown, *J. Chem. Inf. Comput. Sci.*, 1983, **23**, 66–69.
66 A. Belsky, M. Hellenbrandt, V. L. Karen and P. Luksch, *Acta Crystallogr., Sect. B*, 2002, **58**, 364–369.
67 H. Yamada, C. Liu, S. Wu, Y. Koyama, S. Ju, J. Shiomi, J. Morikawa and R. Yoshida, *ACS Cent. Sci.*, 2019, **5**, 1717–1730.
68 S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson and G. Ceder, *Comput. Mater. Sci.*, 2013, **68**, 314–319.

69 C. Liu, E. Fujita, Y. Katsura, Y. Inada, A. Ishikawa, R. Tamura, K. Kimura and R. Yoshida, *Adv. Mater.*, 2021, **33**, 2102507.
70 K. Pearson, *London, Edinburgh Dublin Philos. Mag. J. Sci.*, 1901, **2**, 559–572.
71 W. Takahara and F. Masayuki, *Practical Handbook for Materials Informatics (In Japanese)*, Morikita Publishing Corp., 2025.
72 G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T.-Y. Liu, in: *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS 2017, Long Beach, CA, USA., 2017, pp. 3149–3157.
73 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 É. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.
74 M. D. Segall, J. D. L. Philip, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark and M. C. Payne, *J. Phys.:Condens. Matter*, 2002, **14**, 2717.
75 D. Vanderbilt, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1990, **41**, 7892–7895.
76 J. A. White and D. M. Bird, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **50**, 4954–4957.
77 J. P. Perdew, K. Burke and Y. Wang, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 16533–16539.
78 H. J. Monkhorst and J. D. Pack, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1976, **13**, 5188–5192.
79 D. Packwood, J. Kermode, L. Mones, N. Bernstein, J. Woolley, N. Gould, C. Ortner and G. Csányi, *J. Chem. Phys.*, 2016, **144**, 164109.
80 J. Heyd, G. E. Scuseria and M. Ernzerhof, *J. Chem. Phys.*, 2003, **118**, 8207–8215.