ROYAL SOCIETY
OF CHEMISTRY

PAPER

Check for updates

Quantum of selectivity testing: detection of isomers and close homologs using an AZO based e-nose without *a prior* training†

Boris V. Goikhman,[a] Fedor S. Fedorov, [ID] *[a] Nikolay P. Simonenko,[b] Tatiana L. Simonenko, [ID] [b] Nikita A. Fisenko, [ID] [bc] Tatiana S. Dubinina,[a] George Ovchinnikov,[a] Anna V. Lantsberg,[d] Alexey Lipatov, [ID] [e] Elizaveta P. Simonenko[b] and Albert G. Nasibulin*[af]

Tracing the chemical composition of the surrounding environment appeals to the design of highly sensitive and selective gas sensors. Primarily driven by IoT, miniaturized multisensor systems, like e-noses, are considered to address both selectivity and sensitivity issues. Although e-noses might enable discrimination between close homologs and isomers, they are required to be "trained", *i.e.* to project analyte-related signals into artificial space, prior to their in-field applications. In this study, using the programmed co-precipitation method, we synthesized aluminum-doped zinc oxide (AZO) and employed it as a sensing material in an e-nose to examine the sensing performance towards close C1–C5 alcohol homologs and isomers, *e.g.* 1-propanol and 2-propanol, 1-butanol and isobutanol in the frame of the multisensor paradigm. For the first time, we demonstrated selective recognition of the alcohol vapors without *prior* training of the e-nose. This was realized by matching projections of the known analytes' "fingerprints", used to build a chemical space, with the projections of analyte-related signals acquired using the e-nose in artificial space under machine learning algorithms. Moreover, the AZO based e-nose demonstrates a remarkable, up to 0.87, chemoresistive response to alcohol vapors, 0.9 ppm, in the mixture with air at 300 °C with a detection limit down to sub-ppb level. This opens a new avenue for the development of self-learning gas analytical systems, which might recognize new analytes whose profiles are not yet stored in their library.

Introduction

Information transmitted with gaseous species, smells, comprises a valuable source in many practical applications, such as air pollution monitoring,[1,2] methanol detection in beverages,[3] human health assessment,[4] detection of viruses[5] or explosives.[6] It was evolutionarily required for human survival, stimulating fast avoidance response to traces of dangerous smells.[7] Obvious demands for easy detection of these smells

have led to the appearance of a large family of sensors and gas analytical systems.

The analytical figure of merit of every sensor or gas analytical system is defined by several parameters, such as sensitivity, selectivity, stability, speed of operation, cross-sensitivity, limit of detection, *etc.*[8–10] A natural primary ambition is to achieve maximum performance for a sensor, *e.g.* detection of a single molecule is an ultimate characteristic of sensor sensitivity.[11] The other important characteristic, *i.e.* the ultimate selectivity, should address differentiation between two isomers or very close homologs, still complicated by undeveloped approaches to the description of the basis of the smell.[12,13]

Several analytical techniques meet the demands for selective detection of isomers or homologs, such as photoacoustic sensors, which also enable sensitivity down to sub-ppt level,[14] sensors based on Fourier-transform infrared spectroscopy (FTIR), which enable recognition of many analytes in the wavelength region of molecular fingerprints,[15] ion mobility mass spectrometry (IMS),[16] also mass spectrometry (MS) coupled with liquid or gas chromatography (GC).[17] While traditional MS methods handle the differentiation of particles using a mass-to-charge ratio, drift time IMS separation is directly based on particle collision cross-section,[18,19] enabling

[a]Skolkovo Institute of Science and Technology, 3 Nobel Str., Moscow, 121205, Russian Federation. E-mail: f.fedorov@skoltech.ru; a.nasibulin@skoltech.ru

[b]Kurnakov Institute of General and Inorganic Chemistry of the Russian Academy of Sciences, 31 Leninsky pr., Moscow 119991, Russian Federation

[c]D. Mendeleev University of Chemical Technology of Russia, 9 Miusskaya sq., Moscow 125047, Russian Federation

[d]Bauman Moscow State Technical University, 5/1 Baumanskaya 2-ya Str., Moscow, 105005, Russian Federation

[e]South Dakota School of Mines and Technology, 501 E. Saint Joseph St., Rapid City, SD 57701, USA

[f]Aalto University, Kemistintie 1, P.O. Box 16100, 00076 Aalto, Finland

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1ta10589b

the determination of isomers. In 2019 Ono *et al.* reported a vapochromic sensor for discrimination of small aromatic molecules.[20] Their device was proved to have good sensitivity to 1 ppm of toluene, *p*-xylene, 4-fluorotoluene, and anisole. The high selectivity of the device allowed one to differentiate between the different isomers of xylene. A combination of gas sensors with a chromatographic column has been utilized to approach selective recognition of close homologs. In 2006 Sanches *et al.* reported a volatile organic compound (VOC) detection device based on a single gas sensor combined with a GC micro-column with very high selectivity and with low response time.[21] Thirteen years later van den Broek *et al.* revived this approach for the separation of methanol over ethanol.[3] Recently, metal–organic frameworks have attracted considerable attention relying on highly selective adsorption capabilities.[22] In 2021 Kumar *et al.* demonstrated a highly selective ethanol sensor made by growing a surface-mounted metal–organic framework directly onto graphene field-effect transistors.[23]

Alternatively, selective gas analytical systems are designed employing the principles of biological systems. First proposed by Persaud and Dodd,[24] a classical electronic nose (e-nose) mimics the operation of the mammalian olfactory system, *i.e.* in e-nose, sensors collected in an array, provide a distinct response pattern towards a target analyte.[25,26] This vector response represents a "fingerprint" of the analyte, to be further processed by pattern recognition algorithms for its selective determination.[27–30] Such a gas analytical system has been proven suitable for the discrimination of close homologs.[31] Using MOFs, the e-nose approach enhanced with machine learning yields enantioselective detection and discrimination of chiral molecules.[32] Unlike many other techniques requiring bulky apparatus, an e-nose is often made following the lab-on-chip paradigm.[31,33] At this standpoint, an e-nose functionally occupies a niche between single-unit sensors,[2] possessing good sensitivity, fast response, reasonable stability and low power consumption, and GC-MS devices with good sensitivity and selectivity, able to decompose complex gas-mixtures, still having high power consumption, high price and low portability.

However, the e-nose is required to be "trained" for the task of recognition of a smell or analyte prior to its real application, *i.e.* demanding to project analyte-related signals into an artificial space. At the same time, such techniques with very sophisticated apparatus as NMR or FTIR offer the possibility of selective determination without *prior* training because they rely on the intrinsic characteristics of molecules; molecular fingerprints modeled by the DFT can be further matched with the acquired real data.

Nevertheless, portable made-on-chip e-noses find huge attention due to a boost in artificial intelligence (AI) and machine learning (ML) which facilitates related R&D activities. Moreover, electronic noses are selective enough to distinguish within close homologs projected in artificial space. At the same time, the modern AI methods are powerful enough to predict the compound's chemical structure based on its chemical space coordinates[34,35] by grouping chemical properties into high-dimensional virtual chemical space which is estimated to contain from 10^{12} to 10^{100} elements.[36,37] Thus, we hypothesize

that the implementation of ML methods enables the prediction of compound chemical structure based on e-nose-derived vector response towards analytes projected in artificial space coordinates. Specifically, matching chemical space built by AI processed data of chemical databases, *e.g.* database of molecular fingerprints, with e-nose built artificial space helps to design the ML model which predicts the structures of unknown compounds without *a prior* analyte mapping, *i.e.* e-nose training.

Metal oxides remain the most popular materials to be employed in e-noses due to their low price, easy-to-handle analytical signals, and rapid response time combined with good sensitivity.[38–41] Aluminum-doped ZnO, or simply AZO, has demonstrated a remarkable sensitivity down to a sub-ppm range of analyte concentrations,[42–45] stemming from its structural characteristics. Due to the small size of Al^{3+} ions compared to Zn^{2+}, Al-doping induces chemical defects and increases the conductivity of the sensor.[46,47] The sensing mechanism of AZO sensors belongs to the surface-controlled type where the gas sensitivity is determined by the number of surface adsorption sites.[42] Al-doping increases the number of oxygen adsorption vacancies, which yields improved sensitivity. In 2019 Yoo, Günter *et al.* reported about the sensing properties of AZO sensors towards acetone detection in breath. AZO sensors were evaluated to have a response time of 3 s, and sensitivity of 23 ppm^{-1} at 10 ppm of acetone, including remarkably good selectivity *versus* other gases persistent in breath.[48] Al-doped ZnO is usually obtained by sol–gel,[49] hydrothermal[50,51] and solvothermal[52] methods, atomic layer deposition,[53] reactive magnetron sputtering, pulsed laser deposition,[42,54,55] chemical vapor deposition,[56] and sintering.[57] One of the most convenient methods of synthesis of oxide nanoparticles of complex chemical composition is the co-precipitation method.[58] However, employing this method is often complicated by its reproducibility affecting the size and shape of the oxide nanoparticles. Recently, we introduced a programmed co-precipitation method for efficient control over the mentioned parameters.[59,60] The co-precipitation method enables automatization of the protocols of synthesis of nanomaterials of various chemical compositions and substantially increases its speed.

Here, we study the synthesis and sensing performance of aluminum-doped zinc oxide towards different alcohols, homologs, and isomers in the mixture with air. We demonstrate the remarkable sensitivity of this material towards alcohol molecules with a limit of detection down to a few ppt. The selectivity of this sensor system is approached by using the multisensor concept helping to differentiate within close homologs and isomers (1- and 2-propanol and 1-butanol and isobutanol). We also demonstrate the possibility of analyte detection using a simple e-nose system even without *prior* training by matching the chemical space built using molecular fingerprints with artificial space built with e-nose data.

Experimental

Synthesis of AZO

Synthesis of AZO nanoparticles was carried out using an aqueous solution of $Zn(NO_3)_2$ and $Al(NO_3)_3$ (supplied by

8414 | *J. Mater. Chem. A*, 2022, **10**, 8413–8423

This journal is © The Royal Society of Chemistry 2022

Khimmed Company, Russia). We used the solution with a total concentration of salts 0.3 M in a ratio corresponding to the target composition ZnO – 1.5% Al_2O_3. We added an aqueous 5% solution of NH_4OH with a controlled rate of 0.05 ml s^{-1} under constant steering using an automatic high precision potentiometric titrator ATP-02 (Aquilon JSC, Russia) until pH equal to 8 was reached. All chemicals were of analytical grade and used without further purification.

The synthesized AZO particles were separated by centrifugation, washed with distilled water, and then dried at 100 °C for 3 h. To facilitate crystallization we stabilized the obtained oxide powder at 350 °C for 1 h in air. We further prepared the dispersion of the obtained AZO nanoparticles using a mixture of ethanol with distilled water, $v(C_2H_5OH)/v(H_2O) = 1$, under ultrasonication. The dispersion contained 5% of the AZO nanoparticles. The chips with the sensitive AZO layer were made by drop-casting 1 μl of the dispersion followed by stepwise drying at a temperature range of 25–100 °C. To fully remove residual solvents and stabilize the prepared layer we annealed the chip at 350 °C for 1 h in air. A schematic representation of chip preparation is given in the ESI (Note 1, Fig. S1†).

Material characterization

X-ray diffraction analysis (XRD) of the powder was performed using a Bruker D8 Advance diffractometer (Germany), CuK_α = 1.5418 Å, Ni filter, E = 40 keV, I = 40 mA, signal accumulation time 0.3 s per point, increment 0.02. @X'Pert HighScore Plus software (PANalytical B.V. Almelo, The Netherlands) was used to perform phase identification and Rietveld refinement of the XRD pattern of the sample under study in order to calculate its lattice parameters.

To examine the morphology of the AZO layer on a chip, we utilized scanning electron microscopy (SEM) in a secondary electron detector regime. The high-resolution SEM images of the selected areas at the representative chip were taken using FEI Teneo VolumeScope device (FEI Comp., the USA) at 20 kV.

The chemical composition of materials under study was confirmed by time-of-flight secondary ion mass spectrometry (ToF-SIMS) analysis coupled with a scanning electron microscopy unit (TOF-Sims, Tofwerk AG, Switzerland; Tescan Solaris, Tescan Orsay Holding, Czech Republic). We evaluated a chip area of 30 × 30 μm² (or 35 × 35 μm²) between two Pt electrodes with the AZO material placed in-between. The material was bombarded by gallium ions employing a focused ion beam (FIB) source at an accelerating voltage of 20 kV and current 4 nA.

The ToF-SIMS data are acquired in positive ion mode due to the better sensitivity to Al and Zn.[61] Mass resolution, $m/\Delta m$, was ca. 700. Secondary ions were registered using the ToF-SIMS detector (TOF-Sims C-TOF). The intensity of detection is evaluated by ion counts related to the ToF-SIMS extraction coefficient (device software built-in function, @ToF SIMS Explorer, Version 1.12.2.0), i.e. counts per extraction (cts/ToF extraction). The extraction corresponds to one pulse of the ToF, or, equivalently, one pixel in the FIB scan. Scan resolution equals ca. 70 nm (i.e. 512 pixels over 35 μm). The pulse length was 30 ns.

The structure of the AZO material was also studied with the help of transmission electron microscopy (TEM). For the TEM studies, we prepared samples by placing droplets of dispersion of AZO in isopropanol onto a lacey carbon film supported by a copper grid. Bright-field (BF) and dark field (DF) images, and selected area electron diffraction (SAED) patterns, were acquired with FEI Tecnai G2 F20 S transmission electron microscope (FEI Comp., The Netherlands). The ring patterns were analyzed using @CrysTBox ringGUI software.[62]

Sensor chip

Here, we employed a chip, 10 × 10 mm², with 18 strip co-planar electrodes, two meander thermoresistors, and two meander heaters on its front side. The chip was made of thermally oxidized Si with a 300 nm SiO_2 layer. The electrodes, the thermoresistors, and the heaters were sputtered using magnetron sputtering of Pt and Ti targets to reach the thickness of ca. 150 nm. Ti was used as an adhesive layer with a thickness of about 5 nm. Electrodes' width equaled ca. 50 μm while the length reached 4 mm. A pair of electrodes comprised a sensor or a sensor segment, making 17 sensor segments on one chip. The electrodes were distanced by 50 μm what corresponds to the width of the sensing layer of a sensor. The chip operation temperature was maintained with the accuracy of ±5 °C controlled by the thermoresistors with feedback to resistive heating of the meander heater employing a home-made electric board. Before the measurement, we calibrated the chip temperature using an IR pyrometer Kelvin Compact 1200 D ("Euromix" CJSC, Russia) and stabilized it at 300 °C for 24 h. The chip was wired to a ceramic card with Au wire, 38 μm, leaving a 1 mm air gap between the chip and the card to mitigate the heat dissipation; the card was connected to a home-made electric board which enabled the temperature control and data acquisition, i.e. recording sensor segments' resistance transients. The sampling rate was ca. 0.39 Hz per segment. I–V curves were measured using the semiconductor analyzer Keysight B1500A (Keysight Technologies, the USA) in the range of voltages from −5 V to +5 V both in forward and backward directions. In our tests, the chips' performance was evaluated in the temperature range from RT up to 300 °C. Still, here we have shown the data for the range 200–300 °C, because of the prominent response observed in this range only. Here we have presented all the data for the chip with 13 operating segments, which we thoroughly characterized after pre-experimental testing of several chips with the AZO layer.

Gas-mixing setup and sensor performance evaluation

The card with a wired chip was placed in a gas-tight chamber with volume ca. 0.76 cm³ and tested in a dynamic flow mode, 200 sccm. We supplied dry air at RT from a pure air source to serve a background environment. The dry air was forwarded to three lines through mass flow controllers, MFCs (Bronkhorst®, The Netherlands), as shown in the ESI (Fig. S2†). One line contained dry air only and was utilized to set the background atmosphere. The other two lines were joined to ensure control of the concentration of the analytes. A diffusion vial

This journal is © The Royal Society of Chemistry 2022

J. Mater. Chem. A, 2022, **10**, 8413–8423 | **8415**

DYNACAL®, type A for methanol or type B for other analytes (VICI Metronics Inc., USA), was placed in a container incorporated in one of these lines. The vial is a vessel filled with an analyte to be released through a capillary tube neck.[63] The vial neck is long enough to consider the pressure difference inside and in the atmosphere as constant. Thus, the vial provides a continuous and stable diffusion rate of the investigated analytes, which is defined primarily by analyte vapor pressure, length, and diameter of the capillary. To control the vapor concentration, we pumped out the part of the flow passing through the line with the vial to be compensated by the flow in the other line. Finally, the line with dry air and the two joint lines were forked using a switch valve for flows to be forwarded either to a chamber with the chip or to exhaust. We checked the flow rates using flow-meters (M-500SCCM-D/5M, Alicat Scientific, USA), ensuring proper setup operation and flow stability. We set the exposure time to be 5 or 10 min for the analyte mixed with dry air and 15–20 min was used for exposure of the chip to pure dry air. We tested the following VOCs, close homologs, and isomers: methanol, ethanol, 1-propanol, 2-propanol, 1-butanol, isobutanol, isoamyl alcohol (3-methylbutan-1-ol), and acetone. Using the B(A) vial enabled us to adjust the concentration in the range from 0.01 ppm to 13.6 ppm at RT, as presented in the ESI (Fig. S3†). The maximum concentration of each analyte is defined by its vapor pressure, which slightly differs from one analyte to the other. To accurately describe the sensing performance, the following protocol was employed: 16 different analyte concentrations were measured at each operating temperature. Concentrations were defined with respect to maximum concentration, *i.e.* 1%, 5–50% with 5% step, 50–100% with 10% step. The concentration range has been calculated using our data collected in multiple gravimetric tests after 5 days of keeping the vials at a constant flow of dry air, 200 sccm. Based on these data we calculated the relative deviation of the concentration of an analyte.

We calculated chemiresistive response, S, facilitated by the appearance of VOC in accordance with the relationship $S = \Delta R / R_{air}$, where R_{air} is the resistance of the sensor segment in background air (MOhms) and ΔR represents the change in the resistance (MOhms) due to the appearance of VOC. Also, we estimated sensor response error accounting for fluctuations in the sensor segment resistance. The sensors' response and recovery time, t_{90}, has been calculated as the time to reach 90% of the response starting from when the sensor is exposed to an analyte in the mixture with air, or the time to achieve 10% difference with the baseline when dry air is pulsed, respectively.

Data processing

To selectively identify the analytes of interest and determine "unknown" alcohol vapors, *i.e.* without *prior* training, we utilized principal component analysis (PCA) and Decision Tree (DT) ML algorithm.

PCA is a linear dimension reduction technique based on the singular value decomposition of a feature matrix,[64] which is a matrix of vectors that contain independent variables to be processed. The method transforms an array of possibly correlated variables into an array of linear non-correlated features, called principal components. It discards less important variables and preserves more valuable ones finding principal component axes along which the dispersion of data is high.[65]

To build a chemical space as a basis for our model we used molecular "fingerprints", one of the most common representations of chemical structures. Those "fingerprints" are simplifications of the chemical information contained in any chemical entity through binary vectors. For this purpose (to codify the chemical structure), we utilized the PubChem Fingerprint database.[66] The PubChem fingerprint encodes molecular fragment information with 881 binary digits. It is based on substructure analysis, *i.e.* analysis of the fragments of a chemical structure.[67] A fingerprint is an ordered list of binary (1/0) bits. Each bit is a Boolean determination of, or test for, the presence of, for example, an element count, a type of ring system, atom pairing, atom environment (nearest neighbors), *etc.*, in a chemical structure. We applied PCA to the "fingerprints" of selected analytes taken from the database.

Also, we have used PCA to build a space of the tested VOCs, homologs, and isomers, using a vector signal of 13 sensors acquired at 300 °C where each combination of sensor signals (resistance or response) gives a unique "fingerprint" for the particular analyte. Training data represented the recorded resistance transients of sensors in an array (or their response, S) when an analyte in the mixture with air is pulsed, while test data are obtained in a similar way, though, utilizing the data of another pulse.

We hypothesized that the chemical structure is correlated with the sensors' response. To construct a bridge between chosen PubChem "fingerprints" of VOCs projected to space made by PC components and PCA results of e-nose signal processing, *i.e.* an artificial space built from sensors' array responses, we employed the DT supervised non-parametric learning method. We utilized DecisionTreeRegressor from the Scikit-learn[68] Python package with default parameters. k-Fold cross-validation is applied to estimate the performance of the model employing the data which was not used for the training, namely, we trained a regression model to match and relate chemical "fingerprints" based on a molecular structure (PubChem) to sensors array "fingerprints" (vector response) for chosen analytes.

While the space of PubChem "fingerprints" is too large, the number of points is rather small ($n = 8$) to train our model. To overcome this issue, we reduced the dimension of the PubChem fingerprint vectors from 881 to dimension of 2 using PCA and sensor array data, *i.e.* vector of responses of 13 sensors, to dimension of 1, further referred to as "reduced sensor data". We found the DT model suitable for such a small dataset, *i.e.* to prove the proposed concept (for parameters see the ESI, Note 3†).

Results

Material characterization

According to XRD results, the synthesized material is crystallized in the hexagonal wurtzite structure, space group $P6_3mc$
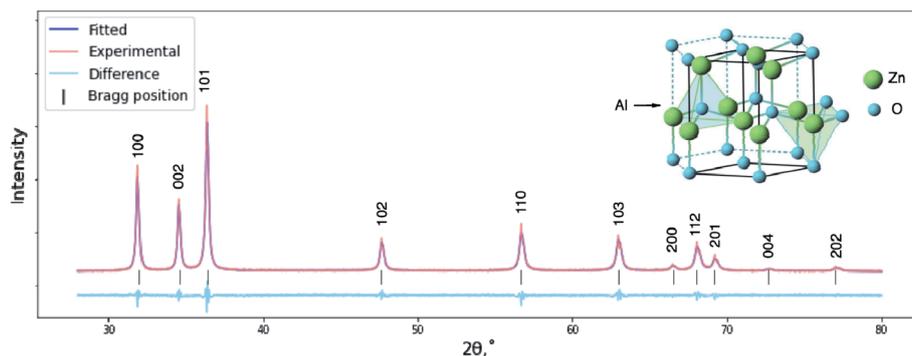
8416 | *J. Mater. Chem. A*, 2022, **10**, 8413–8423

This journal is © The Royal Society of Chemistry 2022

Fig. 1 XRD pattern and corresponding refinement by Rietveld method data for the synthesized ZnO − 1.5% Al_2O_3 powder.

without any crystalline impurities (JCPDF #75-0576), Fig. 1. Implementing Rietveld analysis to the obtained XRD pattern, we have refined the parameters of the crystal lattice ($a = b = 3.251$ Å, $c = 5.207$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$) of synthesized Al-doped ZnO. By estimating the ratio of c/a (1.602), which characterizes the doping of the crystal with a hexagonal lattice,[50] we have evaluated the content of Al atoms to be *ca.* 3% in a distorted ZnO lattice that confirms the target composition of ZnO − 1.5% Al_2O_3.

The material obtained by the method of controlled co-precipitation, placed on a chip, is represented by nearly round-shaped nanoflakes with a mean width of 0.5 μm (Fig. 2a). We have also identified agglomerated particles whose close

inspection evidences the highly porous structure (see the ESI, Fig. S4a†). ToF-SIMS mapping has been conducted to evaluate the spatial distribution of Zn and Al over the AZO layer placed on the chip. The ToF-SIMS maps for Al and Zn, 35×35 μm^2, are presented in Fig. 2b and c. While ToF-SIMS is considered to be a very surface sensitive (1–2 monolayers of sampling depth) technique,[69] the maps indicate even distribution of Zn and Al elements that supports the homogeneous composition of AZO flakes and agglomerates obtained by the programmed co-precipitation method. The positive mass spectrum for Zn and Al is presented in the ESI (Fig. S5†). The results of TEM studies corroborate with SEM data suggesting the material to be crystallized in the form of round plates with a width of up to 400–
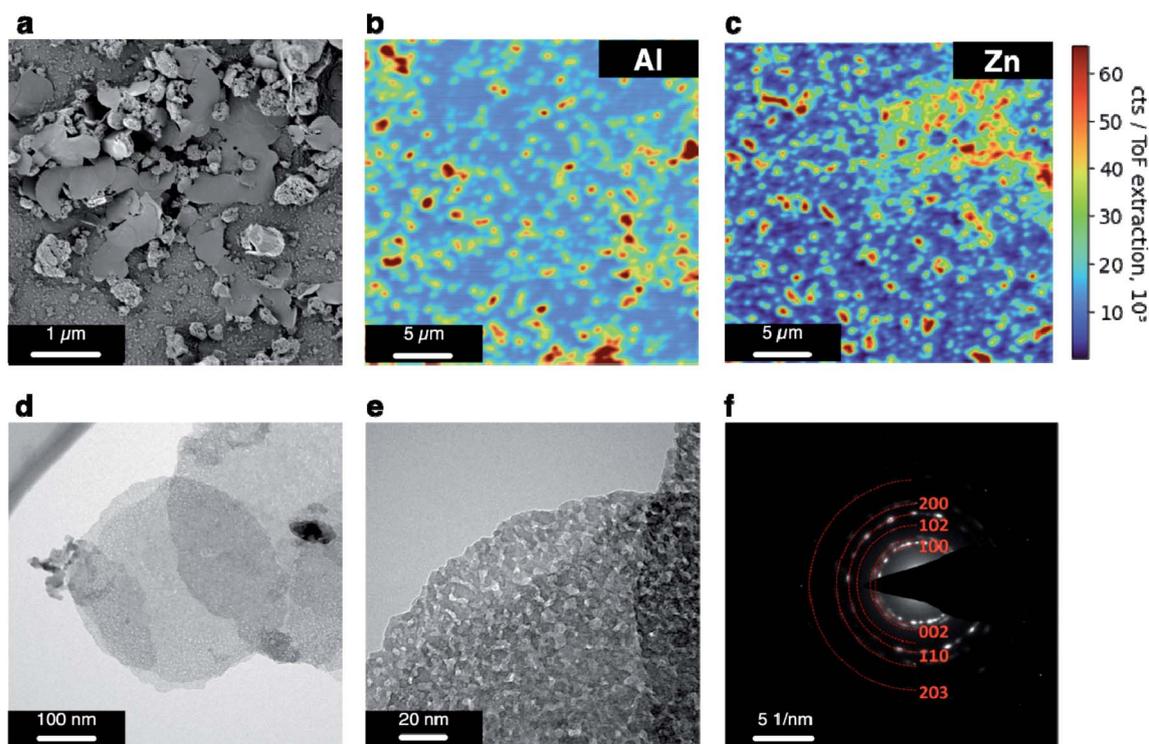

Fig. 2 Characterization of aluminum doped zinc oxide. (a) SEM image of flakes' clusters, *i.e.* agglomerates of aluminum doped zinc oxide particles; ToF-SIMS positive ion atomic maps of (b) Al and (c) Zn on the chip surface; (d) and (e) BF TEM images of AZO flake; (f) SAED patterns of the sample annealed directly at the grid in a muffle oven at 300 °C for 3 h in air.

600 nm (Fig. 2d). The flakes are composed of interconnected crystallites forming a rather porous structure (Fig. 2e and S4b in the ESI†). The crystallite size is about 20 nm proved by DF TEM results given in the ESI, Fig. S4c.† The acquired ring pattern indicates the ZnO crystal phase, space group $P6_3mc$ ($a = b = 3.251$ Å, $c = 5.207$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$), as depicted in Fig. 2f. The phase is identified by (100), (002), (102), (110), (200), and (203) planes characterized by 2.80, 2.58, 2.00, 1.65, 1.40 and 1.09 Å d-spaces, respectively (JCPDF #75-0576).

Sensing performance

We have examined the electrical characteristics and sensing performance of the chip with the AZO layer towards eight different analytes, methanol, ethanol, 1-propanol, 2-propanol, 1-butanol, isobutanol, isoamyl alcohol, and acetone, at three operating temperatures (200 °C, 250 °C, 300 °C). Measured I–V curves for the sensors in pure dry air at 300 °C are presented in Fig. 3a. The results indicate linear I–V correlation in a wide
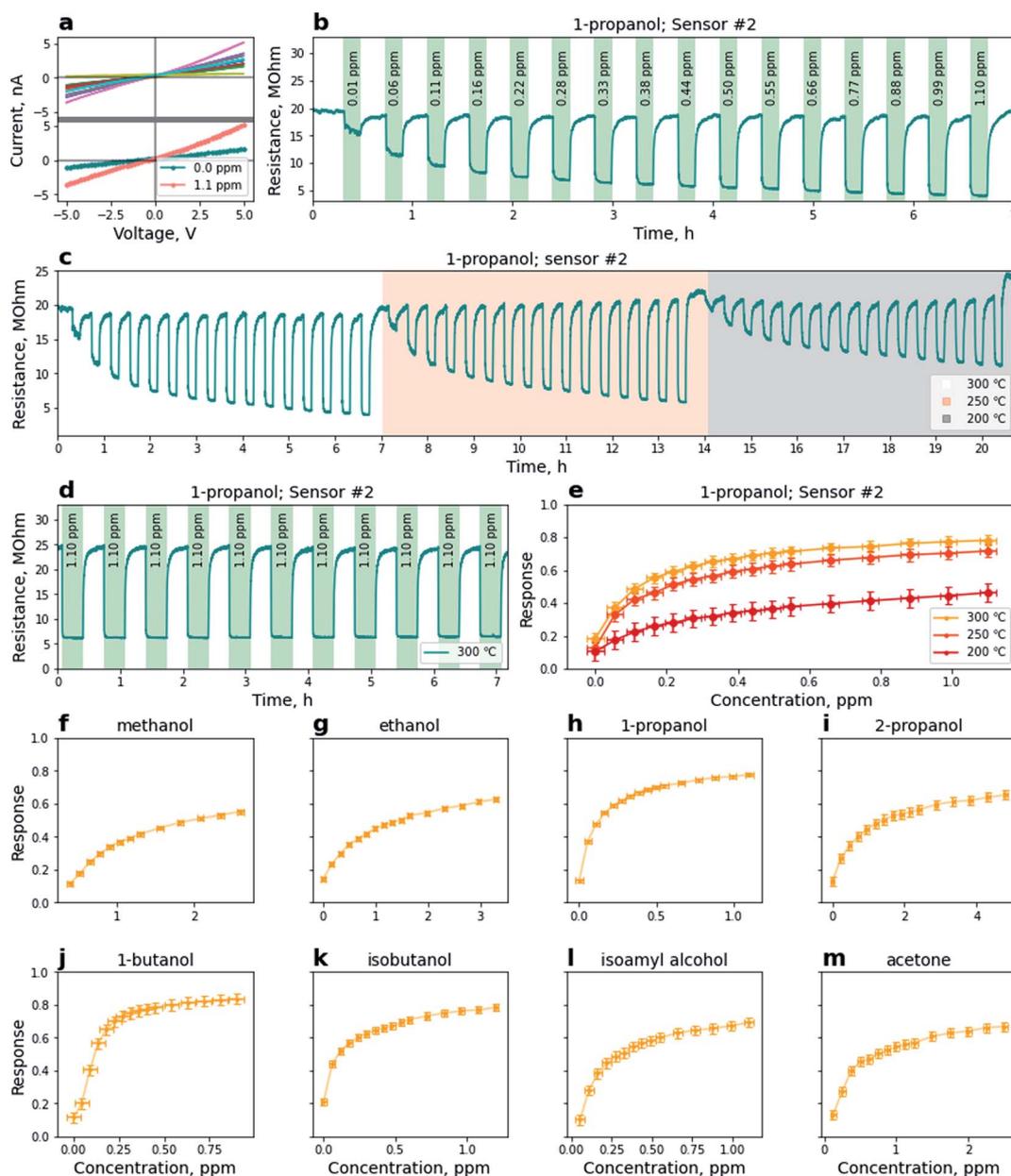

Fig. 3 Performance of AZO based sensors: (a) I–V curves of the sensors in pure dry air (top) for several exemplary sensor segments, I–V curves of the representative sensor in pure dry air and dry air with the addition of 1-propanol (bottom); (b) resistance transients of a representative sensor at different concentrations of 1-propanol [0.01; 1.10] ppm mixed with air at 300 °C; (c) resistance transients of the sensor exposed to 1-propanol [0.01; 1.10] ppm mixed with air at 300 °C, 250 °C, and 200 °C, (d) results of stability tests of sensor segment after 3 months of shelf-life to respond towards 1.1 ppm of 1-propanol in the mixture with air at 300 °C; (e) concentration dependence of the response of the representative sensor to 1-propanol [0.01; 1.10] ppm mixed with air at 300 °C, 250 °C, and 200 °C; (f)–(m) concentration dependence of the chemoresistive response of the representative sensor at 300 °C to methanol, ethanol, 1-propanol, 2-propanol, 1-butanol, isobutanol, isoamyl alcohol, and acetone, respectively.

8418 | *J. Mater. Chem. A*, 2022, **10**, 8413–8423

This journal is © The Royal Society of Chemistry 2022

Table 1 Analytical characteristics of the multisensor chip towards the studied analytes at 300 °C

	Acetone	2-Propanol	Ethanol	Methanol	Isobutanol	1-Propanol	Isoamyl alcohol	1-Butanol
Median LoD, ppb	1.91	1.43	1.96	36.12	0.05	0.15	2.49	1.21
Mean t_{90} (response), s	69 ± 2	33 ± 2	36 ± 2	54 ± 2	20 ± 2	28 ± 2	43 ± 2	15 ± 2
Mean t_{90} (recovery), s	155 ± 12	162 ± 12	103 ± 12	113 ± 7	186 ± 12	173 ± 12	258 ± 12	333 ± 12
Median S (0.94 ± 0.05 ppm)	0.595	0.508	0.531	0.434	0.761	0.789	0.681	0.860
Median n ($S = kC^n$)	0.197	0.246	0.293	0.357	0.169	0.197	0.253	0.174

voltage range $[-5; 5]$ V, which suggests that the sensors can be characterized as resistors with an ohmic contact with Pt electrodes. Fig. 3a (bottom) depicts *I–V* curves recorded for one representative sensor in pure dry air and in the presence of 1.1 ppm of 1-propanol in air at 300 °C. The appearance of 1-propanol vapors facilitates a drop of sensor resistance typical for *n*-type oxide based sensors, while the ohmic behavior remains stable.

A change in analyte concentration in air facilitates a corresponding change in resistance (Fig. 3b) demonstrated for the representative sensor at the chip heated to 300 °C. The results suggest a nonlinear dependence of the sensor's electrical properties on concentration with an obvious trend to saturation at high concentrations of the analyte in air, typical for semiconductor oxides due to a limited number of active surface sites (see the ESI, Fig. S6† for all on-chip sensors' resistance transients). Fig. 3c shows the concentration dependence of the sensor recorded at 200, 250, and 300 °C. While the best performance is achieved at 300 °C, the change to lower temperatures is accompanied by an increase of resistance, possessing smaller changes of the sensor resistance under adsorption of the analyte. We also notice longer response and recovery times in the results recorded at 200 °C. Mean response and recovery times, t_{90}, for the presented sensor upon the appearance of 1-propanol, are correspondingly 28 and 173 s, 43 and 173 s, 98 and 167 s at 300 °C, 250 °C, and 200 °C. We have further employed this sensor to demonstrate the stability of the sensing performance of the AZO material. Stability tests have been handled in 3 months after acquiring isotherms presented in Fig. 3b and c. We have checked the properties of the sensor by the repetitive pulsing of 1.1 ppm of 1-propanol mixed with dry air followed by dry air pulses (Fig. 3d). We notice that the absolute resistance of the sensor has changed slightly over the time of storage. However, despite the differences in absolute resistance value, the sensitivity (change in resistance normalized by resistance in the air) remains the same and the sensor demonstrates a highly stable reproducible signal. In Fig. 3e one finds three representative isotherms (200, 250, and 300 °C) for 1-propanol mixed with air in a range of concentrations, 0.01–1.10 ppm, calculated using the data presented in Fig. 3c. The temperature decay yields a decrease of sensitivity, emphasizing the best operating temperature to be 300 °C.

The chemiresistive response has been carefully investigated for all eight analytes. In Fig. 3f–m we present isotherms for eight analytes in the mixture with air used in this study at 300 °C. Interestingly, for the analytes with greater molecular mass, like butanol, isobutanol, and isoamyl alcohol, saturation is achieved at lower concentrations. At the same time, the sensor demonstrates a weaker response to the analytes with lower molecular mass like methanol and ethanol. We calculated the limit of detection, LoD, as the triple amplitude of signal noise.[70] The median LoD values belong mainly to the region of concentrations in ppb or even ppt domain, *e.g.* 50 ppt for isobutanol and 150 ppt for 1-propanol at 300 °C. Using the Freundlich isotherm, $S = kC^n$, where C is the concentration, we fitted the obtained isotherms to further estimate n values which vary from 0.17 to 0.36 for the analytes at 300 °C which confirms non-linear response characteristics typical of semiconductor oxide sensors.[31] The mean response and recovery times, along with median sensor response calculated for 0.94 ± 0.05 ppm of analyte mixed with air, the median value of power in $S = kC^n$, and median LoD at 300 °C are given in Table 1.

Discussion

Machine learning for analyte prediction

Oxide based gas sensors usually possess fundamentally high cross-sensitivity to various VOCs. While the sensor response might be different for analytes at the same concentration in air, the isotherms might intersect at other concentrations (see Fig. 3f–m) leaving no chance for selective identification of the analyte with one sensor. Even though the interaction of the analyte with the surface of the oxide is rather specific, it is influenced by many factors, like microstructure,[71] number of junctions, and particle size (*e.g.* in the case of nanowires).[72–76] Direct mass dependence of homologs shown earlier[77] is not followed in the case of isomers as seen from Fig. 4a–h because their branched structure might influence adsorption, and further the charge exchange. Here we approach the selectivity by combining the sensors into an array and processing their vector signal by pattern recognition algorithms. The response of sensors of the multisensor chip is presented in Fig. 4a–h for the close concentrations of the studied VOCs in the mixture with air. Along with a response, we have calculated the so-called sensitivity coefficient by normalizing the presented responses by the concentration of an analyte mixed with air. Although the concentration dependence is nonlinear, using concentrations close to unity supports the assumption made.

Both chemoresistive responses and sensitivity coefficients for an array of sensors represent a unique combination, *i.e.* fingerprint, that favors using the pattern recognition algorithms to handle the data and projecting it to artificial space to meet selective discrimination of the analytes. Usually, selective classification of analytes is approached by using dimension

This journal is © The Royal Society of Chemistry 2022

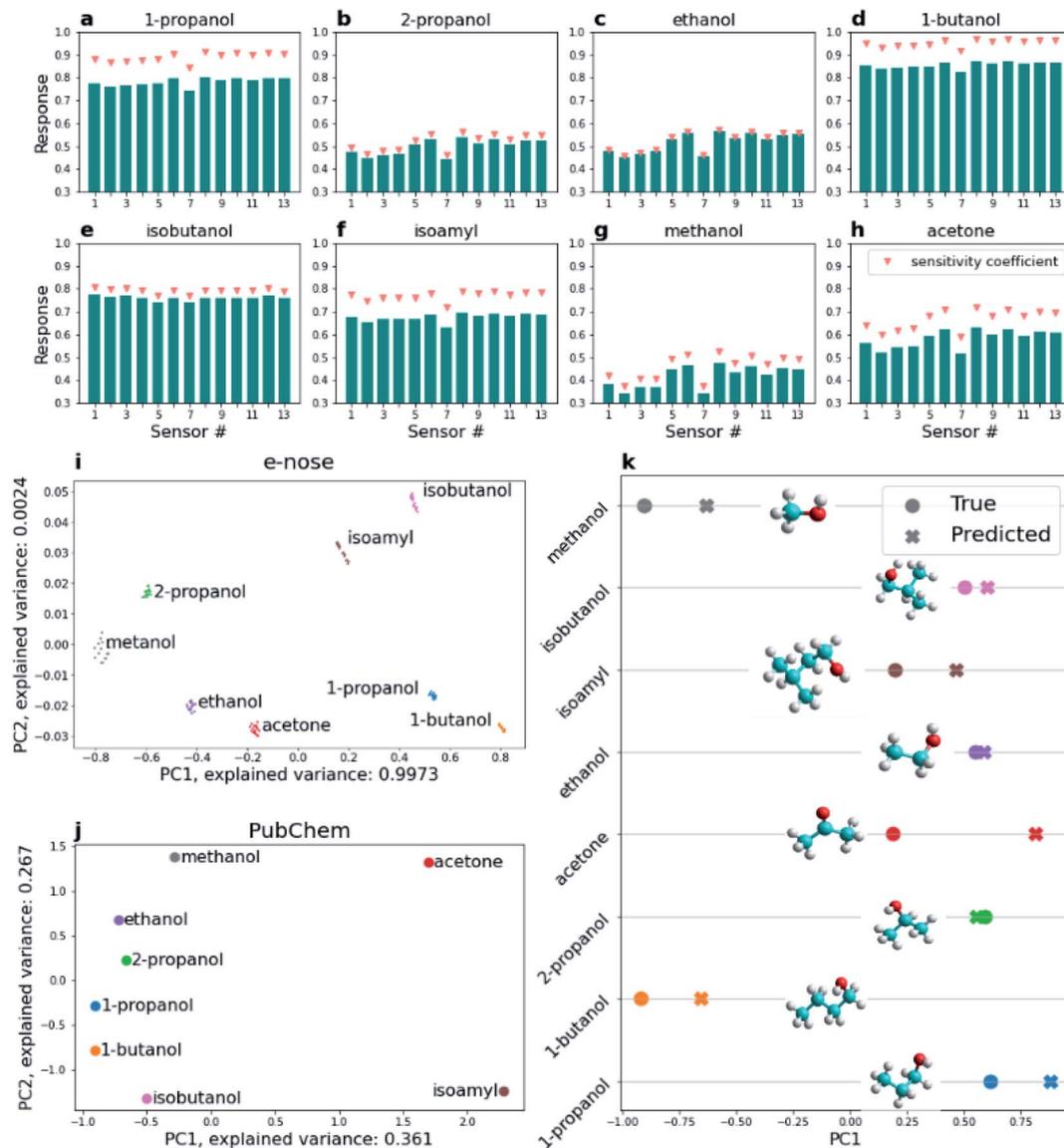*J. Mater. Chem. A*, 2022, **10**, 8413–8423 | **8419**

Fig. 4 Selective recognition and analysis of "fingerprints" of the VOCs. (a)–(h) Chemoresistive responses of the sensors of the multisensor chip towards methanol (0.91 ppm), ethanol (0.99 ppm), 1-propanol (0.88 ppm), 2-propanol (0.96 ppm), butanol (0.90 ppm), isobutanol (0.96 ppm), isoamyl alcohol (0.88 ppm) and acetone (0.88 ppm) at 300 °C, with calculated sensitivity coefficients; (i)–(k) PCA representation of data acquired from the multisensor system, and from PubChem "fingerprints" of chosen analytes with the corresponding prediction of the "unknown" analyte in the mixture with air: (i) PCA of vector signal recorded using the multisensor system for the tested analytes at concentrations of 0.94 ± 0.05 ppm in the mixture with air; (j) PCA projection of PubChem "fingerprints"; (k) Decision Tree prediction of "reduced sensor data" from PubChem "fingerprints" by training using PCA data of multisensor vector signals shown in (i) to 1D PCA accordingly, circles denote true values, crosses – predicted ones; compounds are color-coded.

reduction methods or ML algorithms employing train and test datasets, or cross-validation routine. Applying the PCA to project vector responses of 8 studied analytes using acquired responses of an array of 13 sensors at 300 °C, reduced to the analyte concentration of 0.94 ± 0.05 ppm in the mixture with air, enables their spatial separation in the space of two principal components (Fig. 4i). The PCA is performed for the data obtained for concentrations of vapors close to unity, which helps to mitigate concentration dependence targeting the primary task of selective discrimination of an analyte. The concentration of the particular analyte can be evaluated after its selective

discrimination using one calibrated sensor. PCA featuring the responses normalized to concentration dependence (C^n) is given in the ESI, Fig. S7.† Notably, a recent study approaches this issue by using a multi-task convolutional neural network[78] which enables us to obtain the concentration value, type of analyte, and device status. There is a good separation of projected vector responses of the studied analytes, confirming their selective identification. Some data tend to position rather closely in 2D PC space related to the structural similarity of the analyte molecules (Fig. 4i). The studied n-alcohols, i.e. methanol, ethanol, 1-propanol, and butanol, are projected mostly

8420 | J. Mater. Chem. A, 2022, 10, 8413–8423

This journal is © The Royal Society of Chemistry 2022

along the first PC component, meaning that most information that helps to differentiate these analytes is described by this component, but the similarity is supported by a small contribution of the second one. The results of PCA data reduction for molecular substructure features of analytes' "fingerprints" taken from the PubChem database are presented in Fig. 4j. The later data show that almost all alcohol molecule-related clusters are positioned along the second PC component, *i.e.* favoring its spatial separation, with a clear trend for *n*-chain alcohol molecules. Notably, the 2-propanol related cluster is projected close to ethanol and 1-propanol, hinting at the similarities of these alcohols; accordingly, isobutanol is projected close to *n*-butanol.

Furthermore, to evaluate the possibility of analyte selective detection without *prior* training of the e-nose, we used the ML method, DT, to relate reduced sensor data and results of PCA of molecular "fingerprints". Training a full model to predict sensor response from PubChem "fingerprints" (or, to predict molecular "fingerprint" from sensor data, which would be even better from a practical standpoint) is rather complicated due to the small number of data points, but not the principal impossibility. We further demonstrate the principle feasibility of this approach by reducing data dimensionality.

We fitted (trained) DecisionTreeRegressor for each analyte on 7 other molecular PubChem "fingerprints" projected into 2D space *via* PCA and reduced sensor data projected into 1D space *via* PCA (Fig. 4k) and tried to predict that the reduced sensor data for the 8th analyte which was not used for fitting, *i.e.* corresponded to an unknown analyte. This operation was performed for all studied analytes. The results indicate how successfully analytes are predicted, showing a good match between sensor data for the analyte (point), and predicted value (cross). 2-Propanol, isobutanol, and 1-propanol molecules seem to have the best prediction accuracy. Training the DecisionTreeRegressor with a leave-one-out cross-validation strategy enables us to obtain the R^2 value for the prediction to be 0.69, where R^2 denotes the coefficient of determination. The current results suggest that the best prediction is for 1-propanol for 1D PCA results indicating closer Euclidean distances to the predicted point. Given the information presented in the figures, such a value suggests a good opportunity to predict the analyte of interest using the proposed approach. We have also evaluated Mol2vec "fingerprints" for the determination of alcohols and other VOCs using other chips prepared by a similar method, see the ESI (Note 6, Fig. S8†).[79,80]

Detection of close homologs or even isomers is complicated by the type of interaction of the molecule with the surface, *i.e.* shift of Fermi level, or catalytic process with an injection of charge carriers as in the case of Co oxide.[81,82] A particular response is defined by several parameters of the material, its receptor or transduction functions, which stem from surface states, crystal size, type of activation, *etc.*[83,84] It makes the response of intrinsically cross-sensitive oxides an outcome of many different material-related input parameters. Although there is a direct dependence on the molecular mass of *n*-homologs of alcohols as shown in the work of Fedorov *et al.*,[77] in the case of isomers it might be complicated by the way the

molecule is adsorbed on the surface, *i.e.* deviating from this direct trend. Thus, we expect no strong linear dependence of response value on molecular mass when branched isomers are tested compared to *n*-alcohols. A particular configuration or orientation of the isomer molecule near the oxide surface might be very similar to the orientation of some *n*-alcohols on the surface, *i.e.* the way the isomer molecule approaches the surface might originate some overlapping of "fingerprints" of sensor responses in artificial space that explains the obtained prediction accuracy. Surely, one might improve it by obtaining a larger database of sensor array vector signals, *i.e.* testing a greater number of various analytes. Still, here we have shown, for the first time, that using a miniature e-nose based on simple resistive cross-sensitive oxide sensors might allow approaching selective recognition of the analytes' close homologs and isomers even without *prior* training, but just with the support of ML employed for connecting the chemical "fingerprint" databases with sensor vector data projected in artificial space.

Conclusions

In this study, we have first tested programmed co-precipitation for the synthesis of aluminum-doped zinc oxide. The method enables precipitation of AZO with a high accuracy towards target chemical composition, ZnO – 1.5% Al$_2$O$_3$. The material is crystallized in the form of thin porous flakes made of nanocrystals with a slightly distorted wurtzite structure.

We show a remarkable, up to 0.87, chemoresistive response of AZO sensors at the e-nose to alcohol vapors, 0.9 ppm, in the air at 300 °C with a detection limit down to sub-ppb and even ppt level. The sensors indicate good stability of response and low response time. We have confirmed a direct dependence of chemiresistive response on analyte molecular weight for *n*-alcohols, disrupted in the case of isomers.

By using PCA analysis and the Decision Tree ML method, we demonstrate not only the selective recognition of close alcohol homologs and alcohol isomers but also the feasibility of their detection without *prior* e-nose "training" just by matching available molecular "fingerprints" with e-nose analyte-related vector responses.

Author contributions

Boris V. Goikhman: investigation, conceptualization, writing – original draft, data curation; Fedor S. Fedorov: conceptualization, methodology, writing – original draft, supervision, project administration; Nikolay P. Simonenko: investigation, formal analysis, conceptualization, methodology, writing – original draft; Tatiana L. Simonenko: investigation, formal analysis, writing – review & editing; Nikita A. Fisenko: investigation, formal analysis; Tatiana S. Dubinina: investigation, formal analysis; George Ovchinnikov: methodology, investigation; Anna V. Lantsberg: methodology; Alexey Lipatov: investigation, formal analysis; Elizaveta P. Simonenko: conceptualization, methodology, writing – original draft, writing – review & editing, supervision; Albert G. Nasibulin: supervision, funding acquisition, writing – review & editing.

This journal is © The Royal Society of Chemistry 2022

J. Mater. Chem. A, 2022, **10**, 8413–8423 | **8421**

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 A. Lewis and P. Edwards, *Nature*, 2016, **535**, 29–31.

2 R. A. Potyrailo, *Chem. Rev.*, 2016, **116**, 11877–11923.

3 J. van den Broek, S. Abegg, S. E. Pratsinis and A. T. Güntner, *Nat. Commun.*, 2019, **10**, 4220.

4 S. Das and M. Pal, *J. Electrochem. Soc.*, 2020, **167**, 037562.

5 K. Kalantar-Zadeh, S. A. Ward, K. Kalantar-Zadeh and E. M. El-Omar, *ACS Nano*, 2020, **14**, 5179–5182.

6 D. Marder, N. Tzanani, H. Prihed and S. Gura, *Anal. Methods*, 2018, **10**, 2712–2721.

7 B. Iravani, M. Schaefer, D. A. Wilson, A. Arshamian and J. N. Lundström, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**(42), 2101209118.

8 X. Liu, T. Ma, N. Pinna and J. Zhang, *Adv. Funct. Mater.*, 2017, **27**, 1702168.

9 L. A. Currie, *Pure Appl. Chem.*, 1995, **67**, 1699–1723.

10 J. F. Fennell Jr, S. F. Liu, J. M. Azzarelli, J. G. Weis, S. Rochat, K. A. Mirica, J. B. Ravnsbæk and T. M. Swager, *Angew. Chem., Int. Ed. Engl.*, 2016, **55**, 1266–1281.

11 F. Schedin, A. K. Geim, S. V. Morozov, E. W. Hill, P. Blake, M. I. Katsnelson and K. S. Novoselov, *Nat. Mater.*, 2007, **6**, 652–655.

12 A. Keller, R. C. Gerkin, Y. Guan, A. Dhurandhar, G. Turu, B. Szalai, J. D. Mainland, Y. Ihara, C. W. Yu, R. Wolfinger, C. Vens, L. Schietgat, K. De Grave, R. Norel, DREAM Olfaction Prediction Consortium, G. Stolovitzky, G. A. Cecchi, L. B. Vosshall and P. Meyer, *Science*, 2017, **355**, 820–826.

13 R. Haddad, R. Khan, Y. K. Takahashi, K. Mori, D. Harel and N. Sobel, *Nat. Methods*, 2008, **5**, 425–429.

14 T. Tomberg, M. Vainio, T. Hieta and L. Halonen, *Sci. Rep.*, 2018, **8**, 1848.

15 T. M. G. Salerno, P. Donato, G. Frison, L. Zamengo and L. Mondello, *Front. Chem.*, 2020, **8**, 624.

16 A. B. Kanu, P. Dwivedi, M. Tam, L. Matz and H. H. Hill Jr, *J. Mass Spectrom.*, 2008, **43**, 1–22.

17 M. N. Eckberg, L. E. Arroyo-Mora, D. R. Stoll and A. P. DeCaprio, *J. Anal. Toxicol.*, 2019, **43**, 170–178.

18 K. B. McAfee, D. Sipler and D. Edelson, *Phys. Rev.*, 1967, **160**, 130–135.

19 L. G. McKnight, K. B. McAfee and D. P. Sipler, *Phys. Rev.*, 1967, **164**, 62–70.

20 T. Ono, Y. Tsukiyama, S. Hatanaka, Y. Sakatsume, T. Ogoshi and Y. Hisaeda, *J. Mater. Chem. C*, 2019, **7**, 9726–9734.

21 J.-B. Sanchez, F. Berger, M. Fromm and M.-H. Nadal, *Sens. Actuators, B*, 2006, **119**, 227–233.

22 M. A. Andrés, M. T. Vijjapu, S. G. Surya, O. Shekhah, K. N. Salama, C. Serre, M. Eddaoudi, O. Roubeau and I. Gascón, *ACS Appl. Mater. Interfaces*, 2020, **12**, 4155–4162.

23 S. Kumar, Y. Pramudya, K. Müller, A. Chandresh, S. Dehm, S. Heidrich, A. Fediai, D. Parmar, D. Perera, M. Rommel, L. Heinke, W. Wenzel, C. Wöll and R. Krupke, *Adv. Mater.*, 2021, **33**, 2103316.

24 K. Persaud and G. Dodd, *Nature*, 1982, **299**, 352–355.

25 M. A. Ryan and N. S. Lewis, *Enantiomer*, 2001, **6**, 159–170.

26 M. A. Ryan, A. V. Shevade, H. Zhou and M. L. Homer, *MRS Bull.*, 2004, **29**, 714–719.

27 K. J. Albert, N. S. Lewis, C. L. Schauer, G. A. Sotzing, S. E. Stitzel, T. P. Vaid and D. R. Walt, *Chem. Rev.*, 2000, **100**, 2595–2626.

28 I. Lundström, R. Erlandsson, U. Frykman, E. Hedborg, A. Spetz, H. Sundgren, S. Welin and F. Winquist, *Nature*, 1991, **352**, 47–50.

29 J. W. Gardner and P. N. Bartlett, *Sens. Actuators, B*, 1994, **18**, 210–211.

30 A. Hierlemann and R. Gutierrez-Osuna, *Chem. Rev.*, 2008, **108**, 563–613.

31 F. S. Fedorov, N. P. Simonenko, V. Trouillet, I. A. Volkov, I. A. Plugin, D. P. Rupasov, A. S. Mokrushin, I. A. Nagornov, T. L. Simonenko, I. S. Vlasov, E. P. Simonenko, V. G. Sevastyanov, N. T. Kuznetsov, A. S. Varezhnikov, M. Sommer, I. Kiselev, A. G. Nasibulin and V. V. Sysoev, *ACS Appl. Mater. Interfaces*, 2020, **12**, 56135–56150.

32 S. Okur, P. Qin, A. Chandresh, C. Li, Z. Zhang, U. Lemmer and L. Heinke, *Angew. Chem., Int. Ed. Engl.*, 2021, **60**, 3566–3571.

33 J. Goschnick, *Microelectron. Eng.*, 2001, **57–58**, 693–704.

34 P. Schwaller, D. Probst, A. C. Vaucher, V. H. Nair, D. Kreutter, T. Laino and J.-L. Reymond, *Nature Machine Intelligence*, 2021, **3**, 144–152.

35 T. I. Oprea and J. Gottfries, *J. Comb. Chem.*, 2001, **3**, 157–166.

36 H. O. Villar and R. T. Koehler, *Mol. Diversit*, 2000, **5**, 13–24.

37 W. P. Walters, M. T. Stahl and M. A. Murcko, *Drug Discovery Today*, 1998, **3**, 160–178.

38 G. Korotcenkov, *Nanomaterials*, 2020, **10**(7), 1392.

39 T. Seiyama, A. Kato, K. Fujiishi and M. Nagatani, *Anal. Chem.*, 1962, **34**, 1502–1503.

40 N. Taguchi, *US Pat.*, 3695848, 1972.

41 T. Wang, H. Ma, W. Jiang, H. Zhang, M. Zeng, J. Yang, X. Wang, K. Liu, R. Huang and Z. Yang, *Phys. Chem. Chem. Phys.*, 2021, **23**, 23933–23944.

42 J. F. Chang, H. H. Kuo, I. C. Leu and M. H. Hon, *Sens. Actuators, B*, 2002, **84**, 258–264.

8422 | *J. Mater. Chem. A*, 2022, **10**, 8413–8423

This journal is © The Royal Society of Chemistry 2022

43 S. C. Navale, V. Ravi, I. S. Mulla, S. W. Gosavi and S. K. Kulkarni, *Sens. Actuators, B*, 2007, **126**, 382–386.

44 M. Hjiri, L. El Mir, S. G. Leonardi, A. Pistone, L. Mavilia and G. Neri, *Sens. Actuators, B*, 2014, **196**, 413–420.

45 D. Kohl, *Sens. Actuators, B*, 1990, **1**, 158–165.

46 P. P. Sahay and R. K. Nath, *Sens. Actuators, B*, 2008, **134**, 654–659.

47 M. Zhao, X. Wang, J. Cheng, L. Zhang, J. Jia and X. Li, *Curr. Appl. Phys.*, 2013, **13**, 403–407.

48 R. Yoo, A. T. Güntner, Y. Park, H. J. Rim, H.-S. Lee and W. Lee, *Sens. Actuators, B*, 2019, **283**, 107–115.

49 H. Layoul, F. Meriche, Y. Bouznit and A. Boukerika, *Appl. Phys. A: Mater. Sci. Process.*, 2021, **127**, 625.

50 S. Jantrasee, P. Moontragoon and S. Pinitsoontorn, *J. Semicond.*, 2016, **37**, 092002.

51 R. Nuryadi, L. Aprilia, M. Hosoda, M. A. Barique, A. Udhiarto, D. Hartanto, M. B. Setiawan, Y. Neo and H. Mimura, *Sensors*, 2020, **20**(7), 2013.

52 A. Šutka, M. Timusk, N. Döbelin, R. Pärna, M. Visnapuu, U. Joost, T. Käämbre, V. Kisand, K. Saal and M. Knite, *RSC Adv.*, 2015, **5**, 63846–63852.

53 A. Mošková, M. Moško, M. Precner, M. Mikolášek, A. Rosová, M. Mičušík, V. Štrbík, J. Šoltýs, F. Gucmann, E. Dobročka and K. Fröhlich, *J. Appl. Phys.*, 2021, **130**, 035106.

54 M. Chen, X. Wang, Y. H. Yu, Z. L. Pei, X. D. Bai, C. Sun, R. F. Huang and L. S. Wen, *Appl. Surf. Sci.*, 2000, **158**, 134–140.

55 H. Agura, A. Suzuki, T. Matsushita, T. Aoki and M. Okuda, *Thin Solid Films*, 2003, **445**, 263–267.

56 J. Hu and R. G. Gordon, *J. Appl. Phys.*, 1992, **71**, 880–890.

57 T. Tsubota, M. Ohtaki, K. Eguchi and H. Arai, *J. Mater. Chem.*, 1997, **7**, 85–90.

58 B. Zhou, L. Chen, C. Li, N. Qi, Z. Chen, X. Su and X. Tang, *ACS Appl. Mater. Interfaces*, 2020, **12**, 51669–51678.

59 T. L. Simonenko, N. P. Simonenko, P. Y. Gorobtsov, I. S. Vlasov, V. R. Solovey, A. V. Shelaev, E. P. Simonenko, O. V. Glumov, N. A. Melnikova, M. G. Kozodaev, A. M. Markeev, A. A. Lizunova, I. A. Volkov, V. G. Sevastyanov and N. T. Kuznetsov, *J. Colloid Interface Sci.*, 2021, **588**, 209–220.

60 A. S. Mokrushin, T. L. Simonenko, N. P. Simonenko, P. Y. Gorobtsov, N. C. Kadyrov, E. P. Simonenko, V. G. Sevastyanov and N. T. Kuznetsov, *J. Alloys Compd.*, 2021, **868**, 159090.

61 R. G. Wilson, *Int. J. Mass Spectrom. Ion Processes*, 1995, **143**, 43–49.

62 M. Klinger, *J. Appl. Crystallogr.*, 2017, **50**, 1226–1234.

63 Y.-H. Lin, K. G. Sexton, I. Jaspers, Y.-R. Li, J. D. Surratt and W. Vizuete, *Environ. Sci.: Processes Impacts*, 2014, **16**, 2703–2710.

64 J. E. Jackson and J. Edward Jackson, *Wiley Series in Probability and Statistics*, 1991.

65 A. Tharwat, *International Journal of Applied Pattern Recognition*, 2016, **3**, 197–240.

66 S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang and E. E. Bolton, *Nucleic Acids Res.*, 2021, **49**, D1388–D1395.

67 Data Specification, https://pubchemdocs.ncbi.nlm.nih.gov/data-specification, accessed 2 December 2021.

68 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, *Journal of Machine Learning Research*, 2011, **12**, 2825–2830.

69 R. Kohli and K. L. Mittal, *Developments in Surface Contamination and Cleaning, Volume 4: Detection, Characterization, and Analysis of Contaminants*, William Andrew, 2011.

70 J. Burgués, J. M. Jiménez-Soto and S. Marco, *Anal. Chim. Acta*, 2018, **1013**, 13–25.

71 C. Xu, J. Tamaki, N. Miura and N. Yamazoe, *Sens. Actuators, B*, 1991, **3**, 147–155.

72 A. Rothschild and Y. Komem, *J. Appl. Phys.*, 2004, **95**, 6374–6380.

73 F. Fedorov, M. Vasilkov, A. Lashkov, A. Varezhnikov, D. Fuchs, C. Kübel, M. Bruns, M. Sommer and V. Sysoev, *Sci. Rep.*, 2017, **7**, 9732.

74 V. V. Sysoev, B. K. Button, K. Wepsiec, S. Dmitriev and A. Kolmakov, *Nano Lett.*, 2006, **6**, 1584–1588.

75 V. V. Sysoev, T. Schneider, J. Goschnick, I. Kiselev, W. Habicht, H. Hahn, E. Strelcov and A. Kolmakov, *Sens. Actuators, B*, 2009, **139**, 699–703.

76 E. Comini, C. Baratto, G. Faglia, M. Ferroni, A. Vomiero and G. Sberveglieri, *Prog. Mater. Sci.*, 2009, **54**, 1–67.

77 F. S. Fedorov, M. A. Solomatin, M. Uhlemann, S. Oswald, D. A. Kolosov, A. Morozov, A. S. Varezhnikov, M. A. Ivanov, A. K. Grebenko, M. Sommer, O. E. Glukhova, A. G. Nasibulin and V. V. Sysoev, *J. Mater. Chem. A*, 2020, **8**, 7214–7228.

78 T. Wang, H. Zhang, Y. Wu, W. Jiang, X. Chen, M. Zeng, J. Yang, Y. Su, N. Hu and Z. Yang, *Sens. Actuators, B*, 2022, **351**, 130915.

79 B. Ramsundar, P. Eastman, P. Walters, V. Pande, K. Leswing and Z. Wu, *Deep Learning for the Life Sciences*, O'Reilly Media, 2019.

80 S. Jaeger, S. Fulle and S. Turk, *J. Chem. Inf. Model.*, 2018, **58**, 27–35.

81 J. R. Stetter, *J. Colloid Interface Sci.*, 1978, **65**, 432–443.

82 T. Wolkenstein, *Electronic Processes on Semiconductor Surfaces during Chemisorption*, Springer US, 1991.

83 G. Korotcenkov and V. Sysoev, *Chemical sensors: comprehensive sensor technologies*, 2011, vol. 4, pp. 39–186.

84 N. Barsan and U. Weimar, *J. Electroceram.*, 2001, **7**, 143–167.

This journal is © The Royal Society of Chemistry 2022

J. Mater. Chem. A, 2022, **10**, 8413–8423 | **8423**