



Hairpin DNA-AuNPs as molecular binding elements for the detection of volatile organic compounds

Marcello Mascini^{a,b,*}, Sara Gaggiotti^a, Flavio Della Pelle^a, Joseph Wang^b, José M. Pingarrón^c, Dario Compagnone^{a,**}

^a Faculty of Bioscience and Technology for Food, Agriculture and Environment, University of Teramo, 64100 Teramo, Italy

^b Department of Nanoengineering, University of California, San Diego, La Jolla, CA 92093, United States

^c Department of Analytical Chemistry, Faculty of Chemistry, University Complutense of Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain

ARTICLE INFO

Keywords:

Hairpin DNA
In-silico screening
AuNPs
QCMs
Multivariate analysis
VOCs
Gas-sensors
E-nose

ABSTRACT

Hairpin DNA (hpDNA) loops were used for the first time as molecular binding elements in gas analysis. The hpDNA loops sequences of unpaired bases were studied in-silico to evaluate the binding versus four chemical classes (alcohols, aldehydes, esters and ketones) of volatile organic compounds (VOCs). The virtual binding score trend was correlated to the oligonucleotide size and increased of about 25% from tetramer to hexamer. Two tetramer and pentamer and three hexamer loops were selected to test the recognition ability of the DNA motif. The selection was carried out trying to maximize differences among chemical classes in order to evaluate the ability of the sensors to work as an array. All oligonucleotides showed similar trends with best binding scores for alcohols followed by esters, aldehydes and ketones. The seven ssDNA loops (CCAG, TTCT, CCCGA, TAAGT, ATAATC, CATGTC and CTGCAA) were then extended with the same double helix stem of four base pair DNA (GAAG to 5' end and TTTC to 3' end) and covalently bound to gold nanoparticles (AuNPs) using a thiol spacer attached to 5' end of the hpDNA. HpDNA-AuNPs were deposited onto 20 MHz quartz crystal microbalances (QCMs) to form the gas piezoelectric sensors. An estimation of relative binding affinities was obtained using different amounts of eight VOCs (ethanol, 3-methylbutan-1-ol, 1-pentanol, octanal, nonanal, ethyl acetate, ethyl octanoate, and butane-2,3-dione) representative of the four chemical classes. In agreement with the predicted simulation, hexamer DNA loops improved by two orders of magnitude the binding affinity highlighting the key role of the hpDNA loop size. Using the sensors as an array a clear discrimination of VOCs on the basis of molecular weight and functional groups was achieved, analyzing the experimental with principal components analysis (PCA) demonstrating that HpDNA is a promising molecular binding element for analysis of VOCs.

1. Introduction

In the past decade, DNA was extensively used in sensors design, fabrication, characterization and application providing new impulses to analytical research (Bettazzi et al., 2017; Rasheed and Sandhyarani, 2017). Through the selection of the DNA sequence, a wide variety of analytical applications have been proposed, the majority of which were applied to liquid samples. Currently, very few gas sensors propose DNA as functional material (Wasilewski et al., 2017). The first attempts to use DNA in gas sensing was reported few years ago and was obtained by decoration of carbon nanotubes (Khamis et al., 2012; Kybert et al., 2013; Staii et al., 2005; Su et al., 2013). In a recent work DNA extracted from fish sperm was introduced between a gate dielectric and an

organic semiconducting layer to build up an organic field-effect transistor sensor for NO₂ detection (Shi et al., 2016). Another interesting work evaluated the ability to detect odors delivered in the vapor phase of double-stranded (dsDNA) and single-stranded DNA (ssDNA); ssDNA exhibited sequence-specific responses for a variety of volatile compounds (White et al., 2008).

To date no gas sensor works has explored the use of hpDNA for sensing of VOCs. In fact, hpDNA has been used for sensor applications only in liquid media, mainly coupled to electrochemical transducers (Martín-Fernández et al., 2015; Wang et al., 2014). The particular shape of hpDNA is very interesting since it appears ideal to maximize the orientation of the binding element (via the stem) using the loop as sensing element.

* Corresponding author at: Faculty of Bioscience and Technology for Food, Agriculture and Environment, University of Teramo, 64100 Teramo, Italy.

** Corresponding author.

E-mail addresses: mmascini@unite.it (M. Mascini), dcompagnone@unite.it (D. Compagnone).

<https://doi.org/10.1016/j.bios.2018.07.028>

Received 5 May 2018; Received in revised form 2 July 2018; Accepted 13 July 2018

Available online 17 July 2018

0956-5663/ © 2018 Elsevier B.V. All rights reserved.

In this work, hpDNA conjugated with AuNPs was used as a molecular binding element in piezoelectric gas detection (Fig. 1S). Piezoelectric transducers can monitor the frequency change of functionalized QCMs when gaseous molecules are adsorbed, providing the relationship between mass and resonant frequency shift (Skládal, 2016). In gas piezoelectric sensors, the use of AuNPs as platform for VOCs binding was found to increase the sensitivity by two orders of magnitude versus monolayer modified QCMs (Compagnone et al., 2013).

The new hpDNA-AuNP piezoelectric gas detection strategy, described in this paper, is based on in-silico calculation of the hpDNA loop binding. HpDNA loop shape is ideal to build sensing molecules since provides a large combinatorial complexity of structures; the latter can be tailored with the help of computer-aided methods to respond to different volatile chemicals. In-silico rationally designed molecular traps have been demonstrated to have a strong impact on the development of analytical strategies since they minimize experimental issues such as reagent stability and nonspecific recognition also for separation procedures (Baggiani et al., 2013; Mascini et al., 2013; Narcisi et al., 2011; Uzun and Turner, 2016). A computational approach was also recently used to reduce the large number of attempts necessary to select the right combination of tools for different gas sensing application (Gustafson and Wilmer, 2017; Mascini et al., 2017; Pizzoni et al., 2014).

AuNPs were used as immobilization platform for the hpDNA sequences. The relative binding affinities of the hpDNA loops vs. different VOCs belonging to relevant chemical classes were evaluated. The results showed a significant increase of the binding affinity versus VOCs with the increase of the hpDNA loop size highlighting the key role of the molecular geometry.

The hpDNA sensors were then used in a sensor array format using the principle of combinatorial selectivity. The combination of such sensors into arrays was thought to overcome the limited selectivity of a single oligonucleotide that is common, to the majority of gas sensing devices. When complemented by a multivariate data analysis technique, this sensor array allows for the classification and the identification of compounds with a performance well beyond that of a single selective sensor.

As a practical use, the pattern recognition of these new sensors was estimated by using the unsupervised multivariate algorithm PCA, a convenient tool often used in sensors post-processing analysis (Akamatsu et al., 2017; Compagnone et al., 2015; Di Natale et al., 2014; Imamura et al., 2017). Data obtained demonstrated that the hpDNA sensors, used as array, were able to discriminate the eight molecules on the basis of molecular weight and functional groups.

2. Materials and methods

2.1. In-silico screening

The in-silico screening procedure was aimed to test the virtual binding affinities of all possible combinations of tetramer, pentamer and hexamer single strand DNA (ssDNA) of the hairpin loop versus four chemical classes. The classes were represented by the following compounds: 14 Alcohols ((1S,2R,5R)-2-isopropyl-5-methyl-cyclohexanol; (2S)-propane-1,2-diol; (2Z)-3,7-dimethylocta-2,6-dien-1-ol; (3R)-3,7-dimethylocta-1,6-dien-3-ol; (3R,6Z)-3,7,11-trimethyldodeca-1,6,10-trien-3-ol; (3S)-3,7-dimethyloct-6-en-1-ol; (3S)-oct-1-en-3-ol; (4S,4aR,8aR)-4,8a-dimethyldecal-4a-ol; 2-Propanol; 3-methylbutan-1-ol; Ethanol; Hex-3-en-1-ol; hexan-1-ol; Terpinen-4-ol; 13 aldehydes ((2E,6Z)-nona-2,6-dienal; (2S)-2-methylbutanal; (2Z)-3,7-dimethylocta-2,6-dienal; (3S)-3,7-dimethyloct-6-enal; (E)-non-2-enal; (Z)-hex-3-enal; 2-methylpropanal; 3-methylbutanal; acetaldehyde; decanal; hexanal; nonanal; octanal; 18 esters (5-methylhexanoate; butanoate; ethyl (2S)-2-methylbutanoate; ethyl 3-methylbutanoate; ethyl acetate; ethyl butanoate; ethyl hexanoate; ethyl octanoate; ethyl propanoate; hexyl acetate; Isopentyl acetate; methyl acetate; methyl butanoate; methyl formate; methyl propanoate; octyl acetate; pentyl butanoate; pentyl

pentanoate); 5 ketones (5 molecules: (1S,4S)-1,7,7-trimethylnorbornan-2-one; acetone; butane-2,3-dione; cyclopentadecanone; pentane-2,3-dione). The 50 molecules were selected in order to possess different functional groups and size.

The entire DNA library of tetramer, pentamer and hexamer ssDNA was generated using Hyperchem 8.0.5 software on a Microsoft Windows 10 laptop. Calculations of the in-silico screening process, including molecular docking run and data preparation were performed using a desktop computer with 19 processors Intel Xeon X5690 at 3.47 GHz each, with 94.5 GiB RAM, running Kernel Linux 2.6.32-642.1.1.el6.x86_64, GNOME 2.28.2. Tools from OpenEye Scientific Software package under academic license, were used at different stages of the in-silico procedure. VOCs were obtained via LEXICHEM 2.1.0 package, by converting ligands standard IUPAC names into their corresponding structures (LEXICHEM version 2.1.0). SZYBKI 1.5.7 with default parameterization was used to optimize molecular geometries (SZYBKI version 1.5.7). Conformational space for both ssDNA and VOCs was taken into account with OMEGA 2.4.6 (Hawkins and Nicholls, 2012; Hawkins et al., 2010; OMEGA version 2.4.6). Multi-conformer rigid body docking was carried out using OEDocking 3.0.0, having Chemgauss4 as scoring function (Kelley et al., 2015; OEDocking version 3.0.0). Structures visualization and generation of molecular surfaces were performed using VIDA 4.1.1 (VIDA version 4.1.1).

The entire DNA molecular surface was included in the active site box defining the area where VOCs were expected to bind. For each ssDNA receptor, a dedicated box (10–20 nm³) was generated. The time elapsed for processing each DNA conformer was about 2 min per processor, from the initial 3D structures generation to final docking results. Ten conformers per ssDNA and a maximum of 200 conformers for each of the 50 VOCs were considered. The binding score average for each DNA was calculated over all the conformers. The entire process was automated using a bash script and using a freeware BASIC-like scripting language (AutoIT V3) for post processing data analysis.

2.2. Experimental procedures

All the reagents and the eight VOCs were purchased from Sigma-Aldrich (Italy). The eight VOCs (ethanol, 3-methylbutan-1-ol, 1-pentanol, octanal, nonanal, ethyl acetate, ethyl octanoate, and butane-2,3-dione) were of analytical grade. HpDNA with unpaired tetramer loop were purchased from Thermo Fischer Scientific (Italy); the others were from Integrated DNA technologies (USA). Standard desalted purified oligonucleotides were bought with a thiol spacer having six carbons (C6).

The piezoelectric measurements were carried out using an Enose-UTV from Sensor group, University of Rome Tor Vergata (Italy). 20 MHz QCM sensors, were from KVG GmbH (Germany).

Colloidal AuNPs were synthesized using the trisodium citrate reduction method (Frens, 1973). In brief, 50 mL of 0.3 mM tetrachloroauric acid solution (in water) was stirred vigorously and heated. At boiling point, 1.5 mL of 40 mM trisodium citrate solution was added. The mixture was left boiling for 20 min (the color turned from clear liquid to wine red). The colloidal suspension was then cooled down to 4 °C. Ultraviolet-visible spectrophotometry was used to confirm the AuNPs formation and verify the AuNPs dispersion. The concentration of AuNPs was 3.5×10^{-9} M considering an average diameter of 15 nm as reported by Sanghavi et al. (2016).

Immobilization of the oligonucleotides on the AuNPs surface was carried out covalently using a C6 thiol modifier group attached to 5' phosphate end of the hpDNA. Each hpDNA was dissolved in deionized water and added to 1 mL of the AuNPs colloidal solution at a final concentration of 0.678 μM. The hpDNA-AuNPs colloidal suspensions were incubated at +5 °C for 12 h. HpDNA-AuNPs were then centrifuged at 13,000 rpm for 30 min at 4 °C. The colorless supernatant was discarded and the solid pellet was resuspended in 1 mL of deionized water. All steps were monitored via UV-Vis spectrophotometry.

The QCM sensors modification was achieved by drop casting 5 μ L of the hpDNA-AuNP - suspension on each side of the crystal and let dry for few minutes. Before the first use, the QCM sensors were completely dried under N_2 at a flow rate of 2 L/h and stored at room temperature in the dark when not in use. The shape of a single crystal unit resonating at about 20 MHz was a circular plate with a diameter of 8 mm (Zampetti et al., 2008). The QCM background noise in all cases was ± 1 Hz.

The piezoelectric measurements were carried out using N_2 as carrier gas at a flow rate of 2 L/h. Measurements of the VOCs were carried out using different amounts of compounds introduced in a gas-tight lab bottle (100 mL) connected to the measuring chamber containing sensor array via three-way stop-cocks. The liquid organic compound was completely evaporated placing the bottle at 45 $^{\circ}$ C for 10 min. The temperature was then brought back to 25 $^{\circ}$ C and the measurement started opening the stop-cocks and, then, flowing the analyte to the sensor chamber. The frequency shift (ΔF), taken as analytical signal was recorded. Steady state was reached between 100 and 200 s after opening the stop-cocks. After each measurement, a complete recovery of the signal was achieved under N_2 flow in about 400 s. Piezoelectric responses dataset was analyzed by the unsupervised multivariate technique principal component analysis (PCA) using MatLab R2011 (USA). Dataset were autoscaled (zero mean and unitary variance) before analysis. PCA was applied to inspect the multivariate data structure by decomposing a data matrix of eight rows (the VOCs) and seven columns (the hpDNA-AuNP sensors).

3. Results and discussion

3.1. In-silico screening: ssDNA vs chemical classes

The binding properties of the ssDNA libraries were calculated against 50 VOCs molecules belonging to four different chemical classes (alcohols, aldehydes, esters and ketones). Only the four natural bases adenine (A), cytosine (C), guanine (G) and thymine (T) were used to build the ssDNA libraries. The minimum oligomer size to have a significant library of loops was using four DNA bases. The tetrameric structure was then the starting library tested. Considering that more bases can contribute synergistically in binding the VOCs, the size of the oligomer library was increased adding in every library an additional base. All the possible combinations of the four DNA bases were tested;

the libraries consisted, then, in 256 elements for tetramers, 1024 elements for pentamers and 4096 elements for hexamers. Increasing the oligomer size by using a pure combinatorial approach generates too much structures to be calculated, therefore hexamer was the largest structure tested in this work.

The molecular docking functions used screened compounds that potentially interacted with the binding site predominantly through non-covalent interactions, particularly hydrogen bonds. Therefore, only the hpDNA loops having unpaired bases were virtually screened. Performing an in-silico screening of the entire hairpin DNA increased enormously the machine time consumption; this was avoided considering that the shape of double strain DNA has no preferential sites for the VOCs binding.

Fig. 1 reports the binding score trend of tetramer, pentamer and hexamer hpDNA loop libraries for the four chemical classes tested. The binding score was reported as the average calculated over 10 conformers for each ssDNA sequence. The score values were calculated using Chemgauss4 scoring function, thus lower values represented higher ssDNA–ligand affinity. The Chemgauss4, the new scoring function from OpenEye software is a modification of the Chemgauss3 that has improved hydrogen bonding and metal chelator functions. This scoring function was particularly suitable for the focus of this work based on the comparison between unpaired DNA bases in binding the VOCs, particularly via electrostatic interactions (hydrogen bond, van der Waals forces).

The molecular docking functions used screened compounds that potentially interacted with the binding site predominantly through non-covalent interactions, particularly hydrogen bonds. Therefore, only the hpDNA loops having unpaired bases were virtually screened. Fig. 1 reports the binding score trend of tetramer, pentamer and hexamer hpDNA loop libraries for the four chemical classes tested. The binding score was reported as the average calculated over 10 conformers for each ssDNA sequence. The score values were calculated using chemgauss4 scoring function, thus lower values represented higher ssDNA–ligand affinity. The oligonucleotides virtual binding score trend was correlated to the oligonucleotide size for all chemical classes, with values increasing of about 25% from tetramer to hexamer. All oligonucleotides had common trend with best binding scores for alcohols followed by esters, aldehydes and ketones. In all libraries, alcohols were 2 times higher than ketones. The minimum-maximum dynamic range

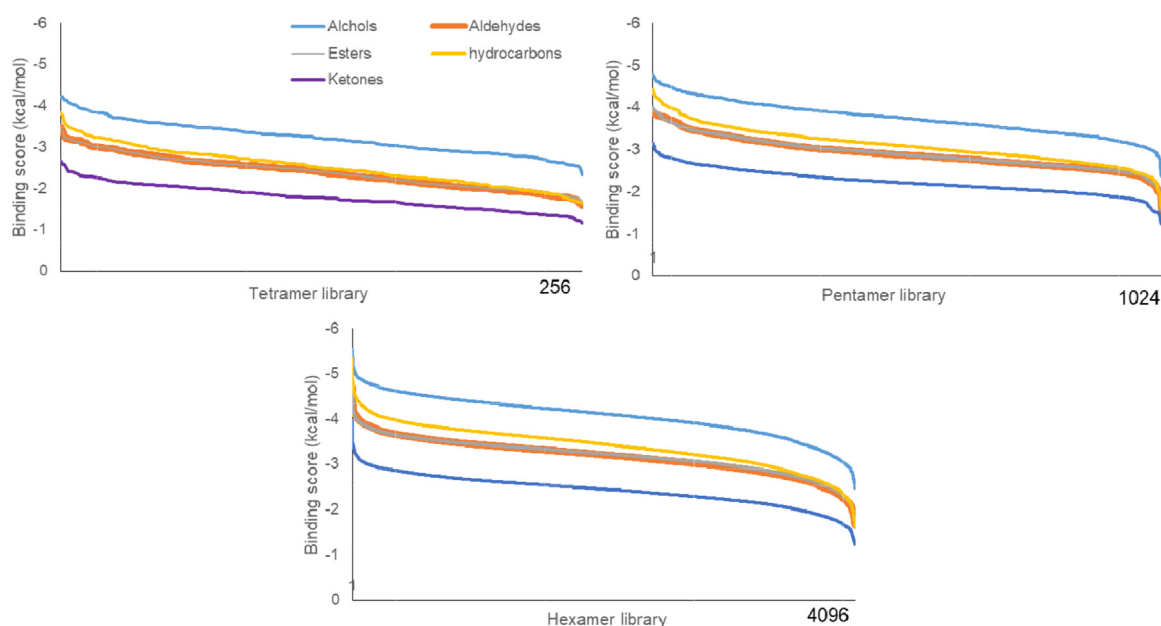


Fig. 1. Binding score trend of tetramer, pentamer and hexamer hpDNA loop libraries for the four chemical classes tested. The data were sorted in ascending order of score, thus not necessarily a correspondence must exist between the positions of the ssDNA in each curve.

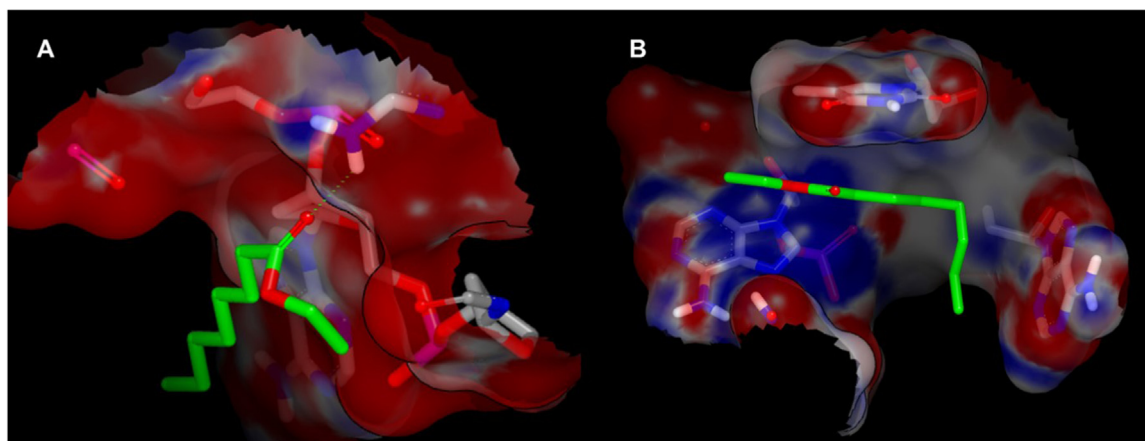


Fig. 2. Electrostatic molecular surfaces of the ssDNA CTGCAA, with a planar interaction surface (binding score -2.26 Kcal/mol) (A), and ATAATC with a saddle shaped binding pocket (binding score -6.28 Kcal/mol) (B) in complex with ethyl octanoate (highlighted in green).

for each chemical class was quite narrow for tetramers becoming relevant only for the hexamer DNA library (-3.07 kcal/mol). In all cases, average and median were very close to each other demonstrating a good symmetry in normal distribution.

Structural analysis was carried out to study the occurrence of the four DNA bases in each oligonucleotide position. The 5% top ranked structures of the tetramer, pentamer and hexamer unpaired DNA were tested versus the four chemical classes. The structural data exhibited a very high level of similarity in DNA bases distribution. Top ranked tetramer and pentamer DNA had higher amount of adenine and thymine, however, in the hexamer DNA the occurrence of both purines was higher than pyrimidines.

Due to the small combinations generated by only four different DNA bases, the binding difference within the DNA library was likely due to the steric/conformational effects. Increasing the DNA in size enhanced the internal flexibility of specific DNA regions and the target accessibility to the binding box conformational space.

Fig. 2 reports the specific positions of the DNA bases contributing cooperatively to target binding. Top binding scores were obtained when DNA docked VOCs with a saddle shaped binding pocket, allowing oligonucleotide to bury the entire ligand in its surface. On the other hand, inefficient binding was found when DNA docked VOCs with a planar interaction. This confirms that the degree of freedom to move around the DNA backbone of the single bases was the major effect to explain the binding score data; this is particularly true for hexamer DNA where the probability of synergic cooperation is higher.

The results of the virtual screening were used to select some oligonucleotides with different affinities for the VOCs in order to evaluate their potential applicability in gas analysis by using QCM sensors.

Since the final aim is the use of the sensors as array (as electronic nose), the loops for the experimental data were selected taking into account not the absolute “best” binding scores but the minimum cross reactivity. This was done looking at the largest differences among the chemical classes.

The selection was finalized to maximize the recognition properties of DNA motif between chemical classes. Thus, two tetramers, two pentamers and three hexamer DNA were finally chosen. **Table 1** reports the binding score of the DNA versus the VOCs selected in experimental part. The binding score average obtained by the simulations of the ssDNA versus the chemical classes (14 alcohols, 13 aldehydes, 18 esters and 5 ketones) was also reported in order to emphasise the differences between chemical class, average and single compounds of the same class.

The selected oligonucleotides have the same trend of the entire DNA library with better interaction for alcohols followed by esters, aldehydes, and ketones showing always the lowest interactions. According

to the binding score data, all the DNA sequences exhibited similar trend for alcohols except for ethanol; binding scores varied significantly for the interaction with esters and aldehydes and, in the case of one hexamer, also for ketones. Three of the seven oligonucleotides selected for the experimental part, CCAG, TAAGT and CTGCAA, were oligonucleotides supposed to be poor candidates in binding VOCs. These oligonucleotides were selected to test the matching between in-silico and experimental data.

The oligonucleotide TTCT showed a good interaction particularly for aldehydes and ethyl octanoate. The other tetramer CCAG exhibited a clear difference in binding alcohols and the other VOCs selected in experimental part. The pentamer TAAGT was selected because of the very low interaction with all the molecules compared to its counterpart CCCGA that had almost two-fold more interaction energy for each of the VOC.

A clear difference in affinity scores was observed using the hexameric DNA. As reported also considering the entire DNA library, increasing the number of bases, there was a considerable increase of docking scores. The hexamer ATAATC showed better binding score than the other oligonucleotide receptors for most of the ligands and, in particular, for ethyl octanoate and both aldehydes (nonanal and octanal). This hexamer and CATGTC exhibited the same pattern in docking the alcohols, aldehydes and esters, showing a significant difference among small compounds, as ethanol and ethyl acetate, and the other molecules. All oligonucleotides exhibited affinity properties inversely correlated to the molecular weight except the hexamer CTGCAA that had good affinity only for ethanol and half interaction energy for all the other VOCs when compared to the other two hexamer DNA.

It should be noted that the same stem DNA sequence was used for the realization of the hpDNAs in order to evaluate the contribution of the loop. Thus, some oligonucleotides, particularly in hexamer DNA, were discarded due to stem-loop intramolecular base pairing.

3.2. AuNPs-DNA functionalization and QCM sensors modification

The selected sequences were extended with the same double helix stem of four base pair DNA attaching to the 5' end the sequence GAAG and to the 3' end the sequence CTTC. Each secondary structure was analyzed using the Mfold Web Server (www.unafold.rna.albany.edu) to check the stem-loop intramolecular base pairing. All selected DNA had unpaired loop in standard conditions.

The AuNP functionalization with hpDNA was followed by UV-Vis spectroscopy. The amount of hpDNA for the AuNPs functionalization was selected testing different concentrations of hpDNA (0.136, 0.271, 0.678 and 1.355 μM).

The UV-VIS spectra after AuNPs functionalization with hpDNA are

Table 1

Binding score average (Kcal/mol) of the tetramer, pentamer and hexamer DNA versus the VOCs tested in experimental part. In *italic-bold*, the binding score obtained by the simulations of the ssDNA versus the chemical classes (14 alcohols, 13 aldehydes, 18 esters and 5 ketones). The average and standard deviation was calculated over 10 conformers.

	CCAG	TTCT	CCCGA	TAAGT	ATAATC	CATGTC	CTGCAA
Ethanol	−2.72 ± 0.16	−2.41 ± 0.17	−2.77 ± 0.28	−2.28 ± 0.23	−3.21 ± 0.29	−3.09 ± 0.28	−3.24 ± 0.32
3-methylbutan-1-ol	−3.23 ± 0.29	−2.77 ± 0.17	−3.26 ± 0.20	−2.74 ± 0.19	−4.63 ± 0.23	−4.12 ± 0.29	−2.72 ± 0.19
1-pentanol	−3.30 ± 0.23	−3.22 ± 0.32	−3.53 ± 0.32	−2.48 ± 0.20	−5.14 ± 0.36	−4.21 ± 0.42	−2.82 ± 0.14
<i>Alcohols</i>	−2.94 ± 0.15	−3.12 ± 0.16	−3.38 ± 0.30	−2.39 ± 0.12	−5.55 ± 0.28	−4.75 ± 0.38	−3.01 ± 0.21
Octanal	−2.39 ± 0.19	−3.07 ± 0.15	−2.62 ± 0.16	−1.54 ± 0.08	−5.13 ± 0.26	−3.71 ± 0.22	−2.40 ± 0.22
Nonanal	−2.32 ± 0.16	−3.24 ± 0.26	−2.56 ± 0.26	−1.53 ± 0.09	−5.54 ± 0.44	−4.05 ± 0.36	−2.56 ± 0.26
<i>Aldehydes</i>	−2.25 ± 0.18	−2.78 ± 0.28	−2.42 ± 0.22	−1.55 ± 0.14	−4.67 ± 0.23	−3.39 ± 0.17	−2.27 ± 0.14
Ethyl acetate	−2.17 ± 0.15	−2.19 ± 0.15	−2.49 ± 0.22	−1.51 ± 0.14	−3.64 ± 0.18	−2.68 ± 0.16	−1.90 ± 0.17
Ethyl octanoate	−1.95 ± 0.14	−3.14 ± 0.16	−2.49 ± 0.17	−1.46 ± 0.07	−6.28 ± 0.57	−4.20 ± 0.38	−2.26 ± 0.11
<i>Esters</i>	−2.19 ± 0.20	−2.73 ± 0.22	−2.53 ± 0.23	−1.58 ± 0.16	−4.60 ± 0.23	−3.37 ± 0.24	−2.26 ± 0.14
Butane-2,3-dione	−2.02 ± 0.14	−2.51 ± 0.15	−1.88 ± 0.11	−1.44 ± 0.09	−3.13 ± 0.28	−2.17 ± 0.17	−1.95 ± 0.14
<i>Ketones</i>	−1.84 ± 0.15	−2.17 ± 0.11	−1.94 ± 0.16	−1.24 ± 0.12	−3.55 ± 0.18	−2.43 ± 0.15	−1.81 ± 0.16

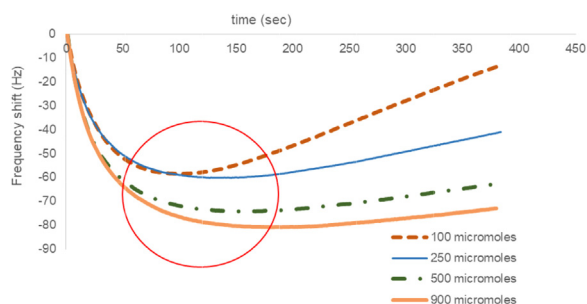


Fig. 3. Frequency shifts recorded by the sensor modified with CTGCAA after introducing in the 100-mL glass bottle different micromolar amounts of 1-pentanol. In all cases, the relative binding affinity between hpDNA and VOCs were calculated taking the frequency shift before desorption of the compound adsorbed on the QCM surface modified with hpDNA-AuNP.

reported in [supplementary material \(Fig. 2S\)](#). Similar absorption spectra were obtained in the 350–800 nm range for bare AuNPs and all the different amounts of hpDNA-AuNPs tested demonstrating that the functionalization did not cause AuNPs aggregation. Similar results were obtained using all DNA loops.

The UV–Vis spectra of the hpDNA-AuNPs resuspended in the same volume of water after centrifugation showed that centrifugation was essential to remove chemicals excess after nanoparticles functionalization. The spectra ([Fig. S2B](#)) showed that the AuNPs were stabilized by the negatively charged DNA that acted as electrostatic repulsing capping agent among the AuNPs as reported in the literature ([Baldock and Hutchison, 2016; Xu et al., 2016](#)). Moreover, the nanoparticles functionalization with hpDNA was confirmed by the presence of a sharp peak at 260 nm indicating the presence of DNA. Unmodified AuNPs were not easily resuspended in water, showing an irreversible aggregation due to the centrifugation step. The spectra of the supernatant ([Fig. 2S C](#)) gave an indirect indication of the maximum amount of hpDNA necessary to saturate the binding sites of AuNPs. In fact, the peak at 260 nm, gave indication of unbound hp-DNA. The latter increased significantly for hpDNA concentrations higher than 0.678 μM , indicating saturation of the binding sites. Therefore this concentration was chosen for the functionalization of all the AuNPs.

After functionalization, the 20 MHz QCM sensors surfaces were modified by drop-casting of 2.5 μL of hpDNA-AuNP suspension on each side of the crystal and let drying at room temperature. This procedure was repeated to assess the maximum loadable amount. Every 2.5 μL addition of hpDNA-AuNP suspension on each side of the sensor led to a variation of approximately 2.5 kHz for all the sensors realized demonstrating the reproducibility of the deposition procedure. After four times (20 μL total volume) QMC crystals frequency crashed and no variation was detectable. Thus, a total amount of 20 μL of hpDNA-AuNP

suspension was selected for further work, leading to a variation of 10 kHz in all cases.

3.3. QCM sensors response to VOCs

QCMs frequency shifts (Δf) were used to calculate the relative experimental binding constants of the eight VOCs and to assess the correlation between the virtual screening and real binding data. For this reason, pure VOCs were tested by using N_2 as carrier gas directly in the measuring chamber.

The relative binding affinities of the complex hpDNA VOC were calculated by adding to the gas-tight lab bottle different amount of liquid pure VOCs. After complete evaporation of the analyte (achieved incubating at 45 $^{\circ}\text{C}$ for 10 min), this was sent to the sensors measuring chamber using N_2 as carrier. The VOC binding to the sensor surface was estimated by recording the frequency shift. A quantitative evaluation of the mass captured by the QCM sensors was achieved through the oscillation constant ($K_q = -4.8 \text{ Hz/ng}$). Using the estimated nanograms it was possible to calculate the moles bound by the sensor. [Fig. 3](#) shows, as an example, the frequency shifts measured with the sensor modified with CTGCAA as loop, for different amounts of 1-pentanol. The piezoelectric sensorgram was similar for all hpDNA-AuNP and VOCs, showing a rapid decrease of the signal after the stop-cocks opening, followed by a slower raise up to the steady state.

The steady state was reached between 100 and 200 s after the start of the measurement. The adsorption kinetics was similar for all the VOCs tested. The frequency shift (Δf), taken as analytical signal, was recorded for all cases before desorption.

The bound compound was determined assuming 1:1 complexation stoichiometry. Using the Scatchard model, the ratio between bound and free compound versus the bound was plotted and the relative binding affinity was calculated by linear regression fitting. The results are reported in [Table 2](#).

Despite their different structure both tetramer DNA loops had very similar binding affinity for all VOCs. The tetramer TTCT exhibited slight better affinity for aldehydes leading to a significant correlation with simulated results. On the contrary, there was no correlation for the other DNA tetramer loop looking at 1-pentanol and ethyl octanoate, respectively. These two molecules and 3-methylbutan-1-ol were bound by the pentamer CCCGA with an affinity of one order of magnitude higher than both the DNA tetramers. The other DNA pentamer loop, TAAGT, had the lowest binding affinity for all molecules. The correlation coefficient of this pentamer DNA was only 0.37 because of the lack of correlation for alcohols. Affinity was high in-silico and low for the experimental data.

The DNA loop size played an important role in the observed experimental behavior improving the binding affinities, as revealed by the DNA hexamer loop binding data. Both hpDNA having as loop ATAATC

Table 2

HpDNA-AuNP sensors relative binding affinities vs the VOCs, estimated using piezoelectric response. The correlation coefficient between experimental and simulated binding is reported in the last row. The standard deviation was calculated using three measurements taken in three different days.

	CCAG Experimental K binding $\times 10^3$ (Moles ⁻¹)	TTCT	CCCGA	TAAGT	ATAATC	CATGTC	CTGCAA
Ethanol	4.3 $\pm$ 0.6	2.3 $\pm$ 0.1	7.4 $\pm$ 0.5	0.9 $\pm$ 0.1	6.5 $\pm$ 0.5	8.4 $\pm$ 0.8	18.4 $\pm$ 2.8
3-methylbutan-1-ol	6.7 $\pm$ 0.5	6.5 $\pm$ 0.6	23.6 $\pm$ 3.3	7.0 $\pm$ 0.6	91.2 $\pm$ 11.9	70.1 $\pm$ 3.5	21.7 $\pm$ 2.2
1-pentanol	3.3 $\pm$ 0.4	3.1 $\pm$ 0.3	60.3 $\pm$ 7.8	2.0 $\pm$ 0.1	530.6 $\pm$ 58.4	352.8 $\pm$ 17.6	10.3 $\pm$ 1.3
Octanal	2.6 $\pm$ 0.3	6.3 $\pm$ 0.6	4.7 $\pm$ 0.3	2.5 $\pm$ 0.1	582.2 $\pm$ 81.5	348.0 $\pm$ 41.8	3.8 $\pm$ 0.3
Nonanal	2.6 $\pm$ 0.2	7.9 $\pm$ 0.6	6.5 $\pm$ 0.8	1.9 $\pm$ 0.1	283.1 $\pm$ 14.2	97.9 $\pm$ 8.8	4.8 $\pm$ 0.4
Ethyl acetate	0.9 $\pm$ 0.1	2.4 $\pm$ 0.3	9.8 $\pm$ 1.5	3.1 $\pm$ 0.4	19.6 $\pm$ 1.2	17.8 $\pm$ 2.3	0.9 $\pm$ 0.0
Ethyl octanoate	8.8 $\pm$ 0.8	12.8 $\pm$ 0.6	43.1 $\pm$ 4.3	2.5 $\pm$ 0.2	519.2 $\pm$ 41.5	345.8 $\pm$ 17.3	6.9 $\pm$ 0.5
Butane-2,3-dione	1.4 $\pm$ 0.2	3.4 $\pm$ 0.3	2.9 $\pm$ 0.3	3.4 $\pm$ 0.4	9.2 $\pm$ 1.1	4.1 $\pm$ 0.3	2.5 $\pm$ 0.4
Correlation with Simulated results	0.16	0.63	0.65	0.37	0.83	0.67	0.80

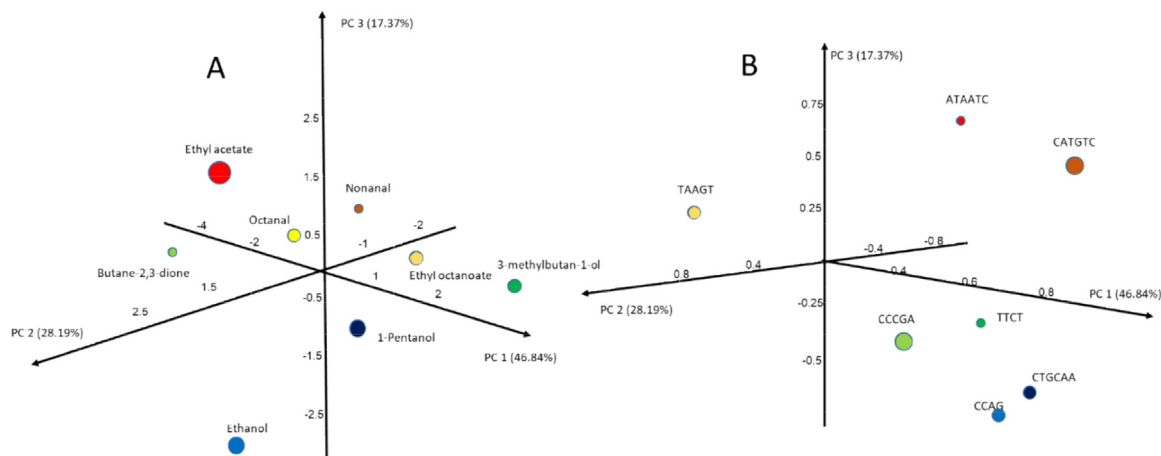


Fig. 4. PCA of the piezoelectric responses of hpDNA-AuNP sensors obtained using 900 μ moles of each VOC. The plot of Scores (A) and the plot of loading (B) of the first three principal components showed 92.40% of the cumulative variance. Data were autoscaled before PCA.

and CATGTC showed a significant interaction with ligands, which was approximately two fold higher than the smaller DNA loop. This was in good agreement with the prediction by virtual screening. Strong interaction with larger molecular weight molecules such as 1-pentanol, octanal, nonanal and ethyl octanoate was observed. On the other hand, the binding behavior of the other DNA hexamer, CTGCAA, was in agreement with the virtual screening results only for alcohols. The different responses of these DNA hexamers to the VOCs emphasized the importance of the chemical nature of the DNA loop. Such heterogeneous data set demonstrated that the binding affinities did not depend on the presence of the stem that was the same for all hpDNA. It should be noted that the seven ssDNA were extended with the same double helix stem of four base pair DNA, that played as spacer. The hairpin structure was chosen to have unpaired DNA bases in the loop therefore with more probability to bind the target via electrostatic interactions (hydrogen bond, van der Waals forces). By increasing the oligomer size, it was supposed that more bases contributed synergistically in binding the VOCs. From the results obtained was clear that larger structures gave better results in terms of chemical classes detection and discrimination. However, if compared with peptides from a previous work (Mascini et al., 2017) detection specificity of oligonucleotides appeared limited. As a consequence, oligonucleotides proposed in this work can be only used for the classification of different VOCs patterns rather than to track individual VOC.

The inter-relationships between the sequence-specific responses of hpDNA to VOCs were highlighted considering all sensors measurements in multivariate analysis format. The data set was represented by the hpDNA-AuNPs-QCMs frequency shifts obtained using 900 μ moles of each VOC.

The data were autoscaled and then analyzed making use of unsupervised PCA. Fig. 4 shows the scores and loading plots of the first

three principal components. The first component represented 46.84% of the variance, the second 28.19% and the third 17.37% displaying together a cumulative variance of 92.40%.

The score points, representing the new coordinates of the VOCs were interpreted assuming that close distance in plot plane is a measure of the similitude between samples. PC 1 separated well both aldehydes and the ketone butane-2,3-dione from alcohols and esters. PC 2 highlighted the differences within alcohol and ester classes grouping the low molecular weight molecules ethanol and ethyl acetate. PC 3 contributed to the separation between ethanol and ethyl acetate, the two small molecules of the group. This separation was influenced by the synergic contribution of all sensors.

The loadings, representing the contribution of each DNA sensor to the principal components, contributed mostly to the scores spread on the PC 2. The PC 2 axis highlighted the differences among sensors. Both pentamer DNA contributed significantly to the separation of the small alcohol and ester to the other family members. On the other hand, the hexamer ATAATC and the tetramer TTCT played an important role in clustering on the PC 2 the molecules with higher molecular weight confirming the predictions obtained by virtual data. The other two hexamers had very similar pattern recognition performance contributing only in spreading the VOCs on PC 1. All sensors contributed to the spread on the PC 3.

It is important to note that hpDNA-AuNPs-QCM sensors can discriminate molecular classes and separate molecules on the basis of the molecular weight. The PCA algorithm highlighted that the DNA sensors, used as array, can be effectively applied to those cases where the difference between VOC patterns plays a crucial role in classification purposes.

4. Conclusions

This work contributes to the growth of the DNA applications in biotechnological and analytical field. For the first time, the interaction between hpDNA loops and VOCs were rationally calculated by virtual assessment and then experimentally tested. A good matching between in-silico selection and experimental results was found especially with hexamer hpDNA.

Multivariate data elaboration showed that beyond interesting differences between chemical classes, molecules could also be clearly discriminated based on the molecular weight.

The key parameter for increasing the affinities of sensors versus VOCs was found to be the size of the DNA loop within the hairpin structure.

This work represents the starting point for the selection hpDNA used as molecular binding elements in gas sensors. In near future, taking advantage of the fast progress in computing, larger ssDNA loops with more complex shapes can be screened in short times, tailoring the efficiency and effectiveness of the gas analysis.

Acknowledgement

This work was supported in part from the European Union's Horizon 2020 Research And Innovation Programme under the Marie Skłodowska-Curie grant agreement No 656687. The authors would like also to acknowledge NBCR award number (P41 GM103426).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.bios.2018.07.028.

References

- Akamatsu, T.I.T., Tsuruta, A., Shin, W., 2017. Selective detection of target volatile organic compounds in contaminated humid air using a sensor array with principal component analysis. *Sensors* 17 (7), 1662.
- Baggiani, C., Biagioli, F., Anfossi, L., Giovannoli, C., Passini, C., Giraudi, G., 2013. Effect of the mimic structure on the molecular recognition properties of molecularly imprinted polymers for ochratoxin A prepared by a fragmental approach. *React. Funct. Polym.* 73 (6), 833–837.
- Baldock, B.L., Hutchison, J.E., 2016. UV–visible spectroscopy-based quantification of unlabeled DNA bound to gold nanoparticles. *Anal. Chem.* 88 (24), 12072–12080.
- Bettazzi, F., Marrazza, G., Minunni, M., Palchetti, I., Scarano, S., 2017. Biosensors and related bioanalytical tools. *Compr. Anal. Chem.*
- Compagnone, D., Faieta, M., Pizzoni, D., Di Natale, C., Paolesse, R., Van Caelenberg, T., Beheydt, B., Pittia, P., 2015. Quartz crystal microbalance gas sensor arrays for the quality control of chocolate. *Sens. Actuators B: Chem.* 207, 1114–1120.
- Compagnone, D., Fusella, G.C., Del Carlo, M., Pittia, P., Martinelli, E., Tortora, L., Paolesse, R., Di Natale, C., 2013. Gold nanoparticles-peptide based gas sensor arrays for the detection of food aromas. *Biosens. Bioelectron.* 42, 618–625.
- Di Natale, C., Paolesse, R., Martinelli, E., Capuano, R., 2014. Solid-state gas sensors for breath analysis: a review. *Anal. Chim. Acta* 824, 1–17.
- Frens, G., 1973. Controlled nucleation for the regulation of the particle size in mono-disperse gold suspensions. *Nature* 241 (105), 20–22.
- Gustafson, J.A., Wilmer, C.E., 2017. Computational design of metal–organic framework arrays for gas sensing: influence of array size and composition on sensor performance. *J. Phys. Chem. C* 121 (11), 6033–6038.
- Hawkins, P.C., Nicholls, A., 2012. Conformer generation with OMEGA: learning from the data set and the analysis of failures. *J. Chem. Inf. Model.* 52 (11), 2919–2936.
- Hawkins, P.C., Skillman, A.G., Warren, G.L., Ellingson, B.A., Stahl, M.T., 2010. Conformer generation with OMEGA: algorithm and validation using high quality structures from the protein databank and Cambridge structural database. *J. Chem. Inf. Model.* 50 (4), 572–584.
- Imamura, R., Murata, N., Shimanouchi, T., Yamashita, K., Fukuzawa, M., Noda, M., 2017. A label-free fluorescent array sensor utilizing liposome encapsulating calcein for discriminating target proteins by principal component analysis. *Sensors* 17 (7), 1630.
- Kelley, B.P., Brown, S.P., Warren, G.L., Muchmore, S.W., 2015. POSIT: flexible shape-guided docking for pose prediction. *J. Chem. Inf. Model.* 55 (8), 1771–1780.
- Khamis, S., Jones, R., Johnson, A.C., Preti, G., Kwak, J., Gelperin, A., 2012. DNA-decorated carbon nanotube-based FETs as ultrasensitive chemical sensors: discrimination of homologues, structural isomers, and optical isomers. *Aip Adv.* 2 (2), 022110.
- Kybert, N.J., Lerner, M.B., Yodh, J.S., Preti, G., Johnson, A.C., 2013. Differentiation of complex vapor mixtures using versatile DNA–carbon nanotube chemical sensor arrays. *ACS Nano* 7 (3), 2800–2807.
- LEXICHEM, version 2.1.0. OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>.
- Martín-Fernández, B., de-los-Santos-Álvarez, N., Lobo-Castañón, M.J., López-Ruiz, B., 2015. Hairpin-based DNA electrochemical sensor for selective detection of a repetitive and structured target codifying a gliadin fragment. *Anal. Bioanal. Chem.* 407 (12), 3481–3488.
- Mascini, M., Montesano, C., Sergi, M., Perez, G., De Cicco, M., Curini, R., Compagnone, D., 2013. Peptides trapping cocaine: docking simulation and experimental screening by solid phase extraction followed by liquid chromatography mass spectrometry in plasma samples. *Anal. Chim. Acta* 772, 40–46.
- Mascini, M., Pizzoni, D., Perez, G., Chiarappa, E., Di Natale, C., Pittia, P., Compagnone, D., 2017. Tailoring gas sensor arrays via the design of short peptides sequences as binding elements. *Biosens. Bioelectron.* 93, 161–169.
- Narcisi, V., Mascini, M., Perez, G., Del Carlo, M., Tiscar, P.G., Yamanaka, H., Compagnone, D., 2011. Electrochemical genosensors for the detection of Bonamia parasite. Selection of single strand-DNA (ssDNA) probes by simulation of the secondary structure folding. *Talanta* 85, 1927–1932.
- OEDocking, version 3.0.0. OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>.
- OMEGA, version 2.4.6. OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>.
- Pizzoni, D., Mascini, M., Lanzone, V., Del Carlo, M., Di Natale, C., Compagnone, D., 2014. Selection of peptide ligands for piezoelectric peptide based gas sensors arrays using a virtual screening approach. *Biosens. Bioelectron.* 52, 247–254.
- Rasheed, P.A., Sandhyarani, N., 2017. Carbon nanostructures as immobilization platform for DNA: a review on current progress in electrochemical DNA sensors. *Biosens. Bioelectron.* 97, 226.
- Sanghavi, B.J., Moore, J.A., Chávez, J.L., Hagen, J.A., Kelley-Loughnane, N., Chou, C.F., Swami, N.S., 2016. Aptamer-functionalized nanoparticles for surface immobilization-free electrochemical detection of cortisol in a microfluidic device. *Biosensors Bioelectron.* 78, 244–252.
- Shi, W., Yu, X., Zheng, Y., Yu, J., 2016. DNA based chemical sensor for the detection of nitrogen dioxide enabled by organic field-effect transistor. *Sens. Actuators B: Chem.* 230, 127–133.
- Stail, C., Johnson, A.T., Chen, M., Gelperin, A., 2005. DNA-decorated carbon nanotubes for chemical sensing. *Nano Lett.* 5 (9), 1774–1778.
- Su, H.C., Zhang, M., Bosze, W., Lim, J.H., Myung, N.V., 2013. Metal nanoparticles and DNA co-functionalized single-walled carbon nanotube gas sensors. *Nanotechnology* 24 (50), 505502.
- SZYBK, version 1.5.7. OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>.
- Uzun, L., Turner, A.P., 2016. Molecularly-imprinted polymer sensors: realising their potential. *Biosens. Bioelectron.* 76, 131–144.
- VIDA, version 4.1.1. OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>.
- Wang, L., Hua, E., Liang, M., Ma, C., Liu, Z., Sheng, S., Liu, M., Xie, G., Feng, W., 2014. Graphene sheets, polyaniline and AuNPs based DNA sensor for electrochemical determination of BCR/ABL fusion gene with functional hairpin probe. *Biosens. Bioelectron.* 51, 201–207.
- Wasilewski, T., Gebicki, J., Kamysz, W., 2017. Advances in olfaction-inspired biomaterials applied to bioelectronic noses. *Sens. Actuators B: Chem.*
- White, J., Truesdell, K., Williams, L.B., Atkisson, M.S., Kauer, J.S., 2008. Solid-state, dye-labeled DNA detects volatile compounds in the vapor phase. *PLoS Biol.* 6 (1), e9.
- Xu, Q., Lou, X., Wang, L., Ding, X., Yu, H., Xiao, Y., 2016. Rapid, surfactant-free, and quantitative functionalization of gold nanoparticles with thiolated DNA under physiological pH and its application in molecular beacon-based biosensor. *ACS Appl. Mater. Interfaces* 8 (40), 27298–27304.
- Zampetti, E., Pantalei, S., Macagnano, A., Proietti, E., Di Natale, C., D'amico, A., 2008. Use of a multiplexed oscillator in a miniaturized electronic nose based on a multi-channel quartz crystal microbalance. *Sens. Actuators B: Chem.* 131 (1), 159–166.