Biosensors and Bioelectronics 177 (2021) 112979

Contents lists available at ScienceDirect

Biosensors and Bioelectronics

journal homepage: http://www.elsevier.com/locate/bios

Optical bioelectronic nose of outstanding sensitivity and selectivity toward volatile organic compounds implemented with genetically engineered bacteriophage: Integrated study of multi-scale computational prediction and experimental validation

Jungyun Park [a,1], Jong-Min Lee [b,1], Hoje Chun [c], Yujin Lee [d], Sung Jun Hong [a,c], Hyunwook Jung [c], Ye-Ji Kim [d], Won-Geun Kim [d], Vasanthan Devaraj [b], Eun Jung Choi [b], Jin-Woo Oh [b,d,**], Byungchan Han [c,*]

[a] Integrated Science and Engineering Division, Yonsei University, Seoul, 03722, Republic of Korea
[b] Bio-IT Fusion Technology Research Institute, Pusan National University, Busan, 46241, Republic of Korea
[c] Department of Chemical and Biomolecular Engineering, Yonsei University, Seoul, 03722, Republic of Korea
[d] Department of Nano Fusion Technology, Pusan National University, Busan, 46241, Republic of Korea

ARTICLE INFO

Keywords:
Virtual screening
First-principles calculations
M13 bacteriophage
Electronic nose
Volatile organic compounds

ABSTRACT

Genetic engineering of a bacteriophage is a promising way to develop a highly functional biosensor. Almost countless configurational degree of freedom in the manipulation, considerable uncertainty and cost involved with the approach, however, have been huddles for the objective. In this paper, we demonstrate rapidly responding optical biosensor with high selectivity toward gaseous explosives with genetically engineered phages. The sensors are equipped with peptide sequences in phages optimally interacting with the volatile organic compounds (VOCs) in visible light regime. To overcome the conventional issues, we use extensive utilization of empirical calculations to construct a large database of 8000 tripeptides and screen the best for electronic nose sensing performance toward nine VOCs belonging to three chemical classes. First-principles density functional theory (DFT) calculations unveil underlying correlations between the chemical affinity and optical property change on an electronic band structure level. The computational outcomes are validated by *in vitro* experimental design and testing of multiarray sensors using genetically modified phage implemented with five selected tripeptide sequences and wild-type phages. The classification success rates estimated from hierarchical cluster analysis are shown to be very consistent with the calculations. Our optical biosensor demonstrates a 1 ppb level of sensing resolution for explosive VOCs, which is a substantial improvement over conventional biosensor. The systematic interplay of big data-based computational prediction and in situ experimental validation can provide smart design principles for unconventionally outstanding biosensors.

1. Introduction

Exact and rapid responses toward target materials are the prime concern for any type of sensing system (Mulchandani et al., 2001; Venugopal 2002; Turner and Magan 2004; Wang et al., 2018). Moreover, in the detection of volatile organic molecules, the sensing performance is of importance since many of them are easily flammable and not recognizable, often leading to critical disasters (Steinfeld and Wormhoudt 1998). These substances, including explosives, however, play key roles in a wide range of chemical industries, and their utilization has been ever increasing in modern society.

The design of outstanding sensors for VOCs of high sensitivity and selectivity with cost-effective materials faces serious, challenging issues. In particular, explosive VOCs easily leak into the open air (Dey et al., 2020) and show a substantially low evaporation rate under ambient conditions. Thus, despite the large number of explosives existing in a

* Corresponding author.
** Corresponding author. Bio-IT Fusion Technology Research Institute, Pusan National University, Busan, 46241, Republic of Korea.
E-mail addresses: ojw@pusan.ac.kr (J.-W. Oh), bchan@yonsei.ac.kr (B. Han).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.bios.2021.112979
Received 19 October 2020; Received in revised form 24 December 2020; Accepted 3 January 2021
Available online 13 January 2021

solid state, it is challenging to detect gas compounds. Detection level of conventional sensors is insufficient to identify VOCs below ppb level (Singh 2007). Accordingly, the design of sensors with outstanding selectivity and sensitivity toward explosives is imperative.

It is noteworthy that sensor array systems have been focused on highly functional alternatives (Rakow and Suslick 2000). It was reported that the integration of auxiliary sensors improves the detection capability of target molecules. The best example of the sensor arrays is an electronic nose, which is a multiarray sensor mimicking the mammalian olfactory system (Rock et al., 2008). The electronic noses consist of various cross-reactive elements, rather than analyte-specific receptors, to exhibit clear responses, which are analyzed using data processing algorithms to specify the substances (Li et al., 2019). If the electronic nose is lightweight, the whole sensor array can be a portable device usable for onsite accidents (Massie et al., 2006).

The major hurdle for electronic noses regarding VOC sensing is to improve the performance in terms of sensitivity and selectivity and the detection threshold up to the level of the human nose (ppb level or even lower) (Rock et al., 2008). It is not a smart way to simply add more sensing components due to the significant and complicated interference of informing signals (Oh et al., 2014).

It is known that the selectivity and precise identification of target molecules are keen to the transducer in the sensor. It is notable that an optical transducer working in the visible light regime is promising due to its clear real-time signal (Dickinson et al., 1996; Rock et al., 2008; Ko et al., 2020). Its sensing mechanism is ascribed to the fact that physicochemical adsorption of target molecules to an electronic nose causes the optical property change (Li et al., 2013; Oh et al., 2014). The variation in the optical property shift is proportional to the molecular interaction between the adsorbate and sensing component, and the color in the visible light regime can be readily observable without auxiliary tools.

There have been extensive efforts to maximize the binding affinity of a receptor with a specific target molecule to improve the performance of a whole sensor array. In particular, a peptide sensor has been of interest since it can easily convert molecular interactions into an identifiable signal via a bioconjugation mechanism (Liu et al., 2015). This implies that the design of exceptionally active sensors toward VOCs is possible based on a fundamental understanding of the molecular interaction mechanism between molecules and receptors. This is the first step toward accurate screening to determine the best adsorbent for target VOCs.

We think an optical biosensor implemented with peptide-based receptors can be promising for VOC detection. This is because the chemical affinity of peptides for volatile organic molecules can be tuned by sequence engineering of the amino acids (Smith 1985; Sidhu 2001; Bratkovic et al., 2005; Lunder et al., 2005; Moon et al., 2015). In fact, the peptide sensor attracted focused attention since its implementation in genetically engineered bacteriophages, research that was awarded a Nobel Prize in Chemistry in 2018 for phage display technology (Smith and Petrenko 1997; Oh et al., 2014). In particular, the M13 bacteriophage has been extensively utilized for integration of peptides since it has 2700 copies of pVIII separated by 3.3 nm for peptide substitution or addition processes (Chung et al., 2011).

The M13 phage implemented with peptides is known as a promising candidate for an optical biosensor in two following aspects (Sidhu 2001; Moon et al., 2015): (1) easy manipulation of the sequences using genetic engineering and (2) facile production of genetically modified phage bundles via spontaneous self-assembly process. The development of phage-based sensors and database has been largely attempted through experimental and computational approaches (Smith 1985; Kim et al., 2019) including our core technologies for the genetic manipulation on parts of the peptide sequences (three amino acids) of the pVIII coat protein in the wild-type M13 phage (Oh et al., 2014). In the studies, major issues were to precisely manipulate the phage complex on the level of genes, to significantly accelerate the response time and to improve sensing resolution in the aspect of the concentration of various target substances. The bottleneck of the objectives is that the proteinic configurational degree of freedom is practically infinite, and that the uncertainty involved in every experiments may be quite high due to hands-on experience. Also, typically the experiments require considerable cost and time (Kim et al., 2019).

Meanwhile, previous computational work for developing biosensors was typically based on empirical potential models (Mascini et al., 2004, 2017; Pizzoni et al., 2014), which are quick but sometimes erroneous and cannot give underlying mechanism of the sensor working principles. It is a promising methodology to propose that the systematic integration of (i) fast database analysis, (ii) precise quantum-mechanical first--principles computation, and (iii) experimental validation of genetically modified phages must be pursued to create a breakthrough solution to the traditional challenges.

We used the unconventional approach to design a highly active sensor, which can distinguish the types of detected molecules with a resolution of 1 ppb or lower. To accomplish this goal, we supposed that binding energy and optical gap shift as each VOC molecule interacts with the sequenced peptide are the key descriptors (Oh et al., 2014). If interaction of the tripeptide and target molecule is strong, the phage swells out, and thus the whole phage bundle expands. Since the tripeptides are located almost at the end of N-terminal of pVIII of M13 bacteriophage, we can expect coherent scattering of photons of the phage-based sensor. Thus, as VOCs bind, the optical gap shift of the sensor should appear. Underlying mechanism of the gap shift is verified with DFT calculations via calculating electronic configuration at HOMO/LUMO of the tripeptide when VOCs adsorb. In short, our hypothesis is that binding affinity will characterize the strength of mutual interaction of peptide and VOC, which is correlated with optical gap shift of our biosensor system.

Furthermore, we synthesized an optical biosensor and tested its sensing performance toward VOCs from computational prediction. From the database consisting 8000 different tripeptide sequences for M13 phages, we selected highly sensitive peptides with explosive molecules when assembled into a multiarray colorimetric M13 phage sensor. The performance was scored by hierarchical cluster analysis and affirmed that analogous molecules were well classified. The computational outcomes and testing results for the measurement of distinct optical responses (color) to specific explosives in the fabricated sensor were very consistent. Essentially, we designed a multiarray colorimetric sensor of bacterial virus (M13 phage) as an outstanding optical biosensor for explosive VOCs.

2. Methodology

2.1. Computations

As previously reported (Strasser 1999), peptide sequence engineering can significantly modify the chemical affinity of the whole amino acid (hydrophilicity, polarity, aromaticity, etc.), toward adsorbing molecules. We computationally established a large database of 8000 different M13 phages interacting with three amino acid sequences as substituents. We identified promising peptide-molecule pairs in each of 8000 phages for sensitive and selective optical sensor design using DFT and empirical computational methods for biochemistry (Pizzoni et al., 2014; Wolcott et al., 2015; Noh et al., 2018; Kim et al., 2020). Previously, Mascini (Mascini et al., 2004) designed efficient peptide-based piezoelectric sensors with the use of big data generated by *in silico* testing.

Conventionally, semicombinatorial virtual screening (Mascini et al., 2017) has been applied to screen the binding affinity of biosensing systems using OEDocking (Kelley et al., 2015). In this research, molecules in different chemical classes were largely examined. DFT calculations are used with the aim of understanding the electronic-structure level mechanism and acquiring key descriptors for the optical properties

in the designed biosensor. Our computational methods can predict the segregation of species in the same chemical classes, and more importantly, the reliability and accuracy of the database are high due to the application of first-principles calculations.

We utilized AutoDock VINA (Morris et al., 2009; Trott and Olson 2010) to calculate the binding scores of various tripeptides and validated the accuracy by comparing the available literature. The program was chosen because of the fast and accurate locating ability for bioactive conformation. In the docking step, the whole peptide was considered a receptor, and we selected five configurations with the lowest binding score. We generated a 27,000 $Å^3$ sized space to fully accommodate the largest receptor and the ligands. After the docking procedure, binding scores were analyzed to identify promising amino acid candidates for the components of a highly selective and sensitive multiarray colorimetric M13 phage sensor.

The benzene molecule was used as a prototype because it is the basis of all aromatic VOCs. Structural combinations of different peptides and benzene were calculated by AutoDock VINA. Then, their binding affinities were affirmed using DFT calculations, as implemented in the Vienna ab initio Simulation Package (VASP) (Kresse and Furthmuller 1996). Project Augmented Wave (PAW) was used to replace the interaction between the core electrons, and the Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation (GGA) (Perdew et al., 1996) was adopted to describe the exchange-correlation energy of the system. The cutoff energy of the plane wave basis set was set to 520 eV, while van der Waals (vdW) interactions between organic molecules were dealt with the DFT-D3 method (Grimme 2006). All structures were optimized until the total energy converged to 10^{-4} eV. Each tripeptide was surrounded with vacuum space of diameter 20 Å to exclude unnecessary interaction between periodic images, and gamma-point was used.

The binding scores calculated from AutoDock VINA were imported to obtain binding affinity, as defined in Eq. (1):

$$E_{bind} = E_{p\text{-}V} - (E_p + E_V) \quad (1)$$

where E_{bind}, $E_{p\text{-}V}$, E_p and E_V represent the binding energy between peptide and VOC, the total energy of peptide and VOC interacting in the same unit cell, and the total energy of peptide and VOC gas molecules, respectively.

The binding affinity ranging from −2.2 to −3.0 kcal/mol showed excellent consistency with the outcomes from VASP calculations, as shown in Fig. S1. The root mean square error was 0.899, implying that our database is highly reliable. Thus, we extensively utilized AutoDock Vina to establish a large database for the binding affinity of the peptides. In addition, the molecular orbitals and optical gaps of the peptides were calculated by time-dependent DFT (TD-DFT) calculations as implemented in Gaussian09 (Frisch et al., 2009), in which the peptide sequences were generated with Avogadro (Hanwell et al., 2012). All TD-DFT calculations and the optimization process were performed using Becke, 3-parameter, Lee-Yang-Parr (B3LYP) hybrid functional (Becke 1988) with a 6–31 + g (d,p) basis set. The peptide sequences in Fig. S1 marked in yellow with green centers were further validated by our *in vitro* experiments.

The structures for tripeptides and the target molecules were generated separately. Using the simplified molecular input line entry system (SMILES) (Weininger 1988), we assigned a text file to each of the 8000 tripeptide sequences. Then, hydrogen atoms were attached to each peptide sequence and converted into a 3D structure using the RDKit program. Then, 300 conformers were created and optimized. In this process, we applied the force field of MMFF94, which is known to be compatible with RDKit (Tosco et al., 2014). We selected the most energetically stable conformer as the optimal conformer for further experimental tests. Meanwhile, VOC structures were generated by Avogadro and further optimized with B3LYP functionals and the basis set of 6–31 + g (d,p).

2.2. Experiments

Selected types of genetically engineered M13 bacteriophages, along with wild-type phage, were prepared based on computational prediction. The major protein (pVIII) of the wild-type (WT) phage was modified into specific tripeptide sequences using site-directed mutagenesis PCR techniques (Oh et al., 2014). To verify phage amplification process stability, DNA sequences were confirmed at each step at the DNA Sequencing Facility (Bionics Co., Korea).

The synthesized phages were prepared in three films with different bundle sizes. To fabricate various phage-based self-assembled color band films, the extraction speeds were controlled between 10 and 100 μm/min using a programmable syringe pump (KD Scientific Co., LEGATO270), as reported previously (Oh et al., 2014). The gold-coated Si wafers (100 nm gold film over a titanium adhesion layer on 525 μm thick Si wafers, Platypus Co.) were pulled from 6 mg mL^{-1} phage suspensions in Tris-buffered saline (12.5 mM Tris and 37.5 mM NaCl, pH 7.5) with different controlled speeds to produce three distinct color bands.

Colorimetric film measurements were conducted using a home-built M13 bacteriophage-based color sensor system consisted of a 7000 mL chamber fitted with a charge-coupled device (CCD) camera (Celstron Inc.) and an inner cap lid space to position the sensor. The chamber was positioned on top of the heat block to control the temperature of the target substances. The camera settings and data collection were controlled in real time using MATLAB (Math Works, Inc.) program. Excess explosive powder (200 mg) was added to the chamber, which was heated for 10 min at the temperature corresponding to the vapor pressure of the chemical. We selected three pixels of each color band (different phage bundle sizes), and images were recorded every 10 s. Images were produced from the RGB signal differences between the initial and final images using Adobe Photoshop (Adobe System Inc.).

The optical color variation from the sensor was quantitatively evaluated as a function for each VOC molecule by analyzing the differences in the RGB values in the phage-based electronic nose. Hierarchical cluster analysis (HCA), which can categorize the variables into groups depending on full dimensional similarities, identified VOCs with their varying concentrations. Ward. D method defines the distance from a given cluster to others as the arithmetic mean.

A classification success rate (CSR) was introduced as a numerical index to compare the chemical classification of different characteristics. It is 100% if a compound with different concentrations belongs to the same branch; otherwise, the CSR is reduced.

The total score defined in Eq. (2) functioned as an indicator for the classification capability of the electronic nose, encompassing both major (aromatic or aliphatic group) and minor (molecular concentration) groups.

$$total\ score = average\ (CSRs\ of\ minor\ levels) \times average\ (CSRs\ of\ major\ levels) \quad (2)$$

3. Results and discussion

3.1. Design principle

It was reported (Oh et al., 2014) that the performance of the M13 phage-based biosensor is largely controlled by the interaction energy of the tripeptide and target molecule. We screened optimal sequences from the 8000 tripeptides database to enhance the sensitivity toward nine VOCs. These VOCs belong to the following three different classes: (i) inexplosive aromatic (benzene; toluene; 1,3-dimethylxylene; aniline; m-toluidine), (ii) explosive aromatic (dinitrotoluene (DNT); trinitrotoluene (TNT)), and (iii) explosive aliphatic compounds (1,3,5-trinitroperhydro-1,3,5-triazine (RDX); octahydro-1,3,5,7-tetranitro-1,3,5,

7-tetrazocine (HMX)). In addition to the explosives, the binding scores of five well-known aromatic VOCs, including benzene analogs, were calculated to validate the computational prediction.

Scheme 1 shows a process illustrating how to select the optimal tripeptide sequences for target VOCs. It is based on the binding affinity screened from all possible combinations of 8000 tripeptides with the nine VOCs. First, the binding score was calculated for each molecule, and the peptides were aligned in the order of increasing the binding scores. The more negative the binding score, the stronger the binding strength between the peptide and the VOC. Using the calculated binding scores, we set up a database of 8000 × 9 arrays to sort out promising peptides, possibly leading to outstanding performance for our optical biosensor when integrated into the M13 bacteriophage.

The binding scores of all 8000 tripeptides in the three chemical classes are illustrated in Fig. 1(a). The overall trend is that the explosive compounds show higher binding intensity toward tripeptides than nonexplosive aromatic counterparts. This is due to the abundance of electrons in the nitro groups and the stronger withdrawal of electrons from which they bind. It makes their molecular rings electron deficient compared to those without nitro groups. Consequently, more electrons are driven to transfer from peptides to the rings of VOC molecules. These characteristics can be further used for the selective detection of target molecules from others. For all ligands, binding scores show marginal differences after 500–1000 tripeptides and are more or less negative. While the binding scores differ by 0.9 kcal/mol within the top 5% peptide sequences, they do only by 0.6 kcal/mol for all sequences if the top and the bottom 5% are excluded. Since we were searching for the best peptide that is the most sensitive to the target molecule, we purposely sampled tripeptides possessing binding affinities of only the top 5% in making decisions regarding the promising peptide sequences for each VOC.

As shown in Fig. 1(b), almost 94% of peptides showing superior binding scores contain at least one aromatic amino acid. In addition, nearly all (98.5%) tripeptide sequences in the top 5% binding scores have aromatic amino acids when interacting with aromatic VOCs, regardless of explosiveness. It is more conspicuous in Fig. 1(c), which shows amino acid compositions of peptides in the top 5% binding score for each ligand. On average, 1.491 aromatic amino acids are required for the tripeptide to interact with aromatic VOCs. Likewise, aliphatic VOCs show the highest binding affinity to aromatic amino acids with an average composition of 1.094. However, they also do to polar and charged amino acids as well. These facts indicate that aromatic amino acids play a significant role in attracting VOCs, especially when the target VOC is aromatic. The electronic charge distribution for the molecular π-π interaction between the molecular rings of the aromatic amino acid and the explosives is depicted in Fig. S2a. It clearly shows that peptides with at least one or more aromatic amino acids strongly interact with aromatic VOCs.

The average binding scores of tripeptides with or without the specific amino acid are plotted in Fig. 1(d). There are 4096 peptide sequences without aromatic amino acids, while there are 3904 peptide sequences with aromatic amino acids. It is notable that the aromatic amino acid greatly enhances binding intensity, on average −0.397 kcal/mol, amounting to more than twice the difference of other kinds of amino acids. Again, this implies that the aromaticity in the peptide plays a substantial role in the high selectivity and sensitivity of our multiarray colorimetric phage sensor.

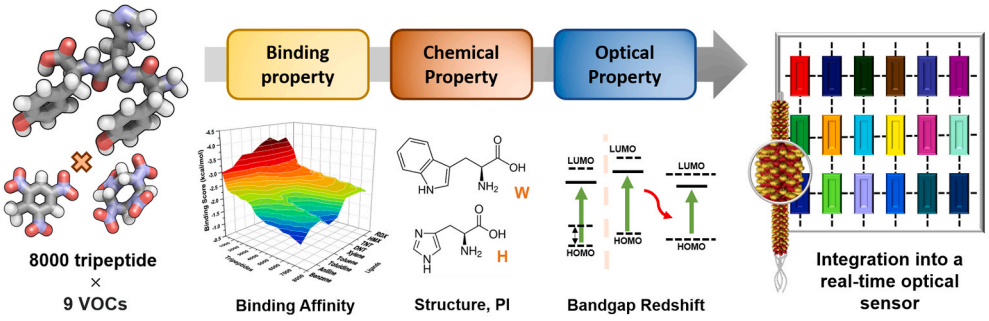
3.2. Peptide sequence selection

While all aromatic amino acids satisfy Hückel's rule (Frenking 2000), phenylalanine (F), tryptophan (W) and tyrosine (Y) contain a benzene-like molecular framework. Histidine (H) specifically has an imidazole ring, which makes the H heterocyclopentadienyl aromatic. Additionally, histidine is the only aromatic amino acid with a pI near neutral pH and R group that can release or gain a proton. In accordance with the different natures, we divided the aromatic amino acids into two groups: (1) F, W, Y and (2) H.

For the first group, we only investigated the tripeptides that have the top 5% binding scores toward explosives and checked the presence of each amino acid. Compared to F and Y, W showed a considerably high probability of being in peptides with preferable binding scores, as illustrated in Fig. S3. Thus, we decided to use W and H for our further *in vitro* experimental tests. Since only eight tripeptides are generated with pure W and H, we examined 27 species, including alanine (A). Alanine was used because its small size of the functional group is expected to have minimal effect on the overall interaction of tripeptide and VOCs.

To identify promising peptide sequences for our optical biosensor, we calculated how the bands of the peptides shift by VOC adsorption. The redshift of the band should be one of the key descriptors for the optical property considering the wavelength of visible light (390 nm–700 nm). As mentioned in section 3.1, the nitro group bound to the ring of the explosive molecule withdraws electrons. This means that electrons in the aromatic ring of H and W are transferred to the explosives. This mechanism renders the peptides unstable, increasing the HOMO energy level of the peptide-VOC complex, as shown in Fig. S2b. The modified energy level eventually appears as different colors in phage-based optical biosensors.

Our calculations indicated that peptides containing W demonstrated a larger band redshift than others, and the shift became larger as W was placed at the end edge of the peptide (N terminal). Table 1 represents the calculated band shifts for the nine functional peptides with W placed on the N terminal. WWW, WHW, and WAW are located at the top three polypeptides, showing the largest band redshift, with an average change of more than 15 nm. On the other hand, despite two Ws, the WWA and WWH show comparably smaller changes. From this, we inferred that peptides having W along with other amino acids are more effective in the band shift. From these results, WWW, WHW, and WAW were selected for our further experimental tests. WHA and AHW were also chosen to compare differences in sensor performance depending on the sequence arrangement. Wild-type phage was also prepared for controlled experiments.



Scheme 1. Overall schematization of the screening of 8000 tripeptides versus nine VOCs.

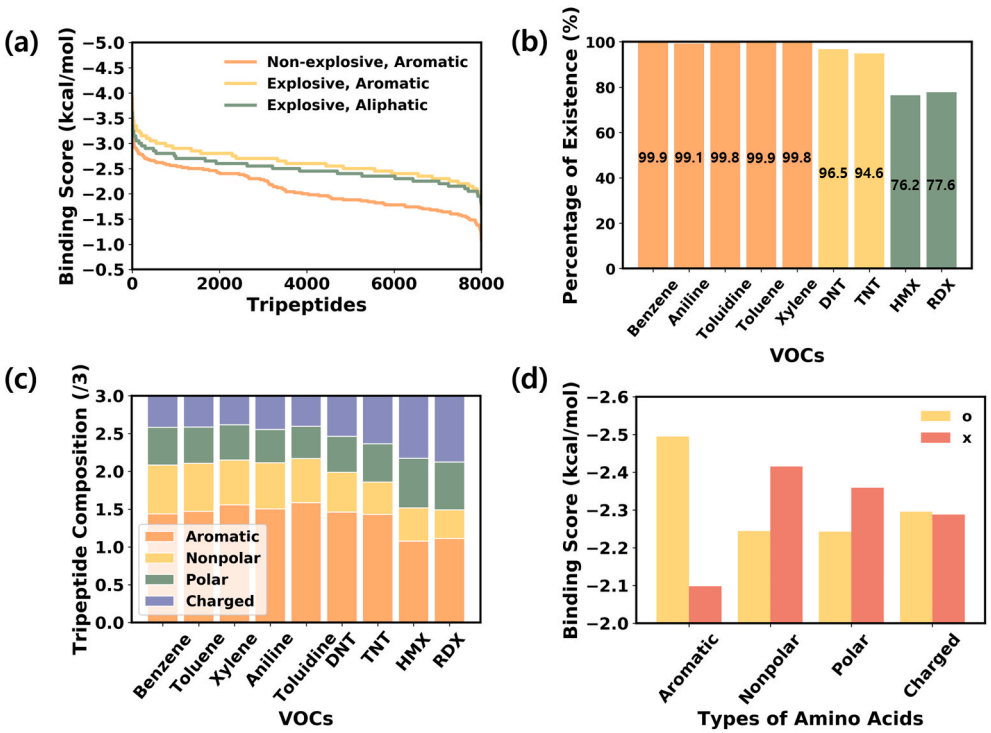


Fig. 1. Screening results. (a) Score distribution in the docking of 8000 tripeptide versus 3 chemical classes. (b) Existence of aromatic amino acids in peptides with the top 5% binding score for each VOC. (c) Amino acid composition of functional peptides with the top 5% binding score for each ligand. (d) Binding score of peptides assorted by the existence of certain amino acid types.

Table 1
Band redshift of functional peptides with W on the N terminus calculated by first-principles simulation.

	Band shift (nm)					
	Benzene	Toluene	Xylene	Aniline	Toluidine	Average
WHW	29.13	24.84	21.57	2.45	1.34	**15.866**
WWW	28.46	24.3	20.87	1.85	2.23	**15.542**
WAW	28.62	24.31	21.04	2.05	0.85	**15.374**
WHA	24.15	20.26	18.89	-2	-2.95	11.67
WAH	24.55	20.96	18.39	-2.29	-3.5	11.622
WHH	24.05	20.71	18.14	-2.1	-2.89	11.582
WWH	24.48	20.39	17.1	-1.59	-2.96	11.484
WAA	24.55	20.12	16.99	-1.75	-2.78	11.426
WWA	24.33	20.36	16.66	-2.56	-2.88	11.182

We analyzed the binding scores for all 8000 peptides in our large database, 27 peptides (consisting of A, H, and W), and the five selected peptide sequences (AHW, WAW, WHA, WHW, and WWW), as shown in Fig. S4. The average binding scores of the 8000 tripeptides and 27 tripeptides were −2.29 and −2.49 kcal/mol, respectively. In contrast, the five peptides showed a significantly stronger affinity toward VOCs with a binding score of −2.77 kcal/mol. Accordingly, we concluded that the selected peptides are important influential components for enhancing the performance of our phage-based optical biosensor.

3.3. Hierarchical cluster analysis and scoring

The structures of genetically engineered M13 phage and sequences of all peptides are shown in Fig. 2(a). The color patterns observed from the multiarray color films, consisting of WHW-, WAW-, and WWW-type functional phages, are displayed in Fig. 2(b). We evaluated the sensitivity of our electronic nose on three aspects: clear coloring pattern, low concentration limit, and the speed of responding time. It is noteworthy that the optical properties of explosive aromatic compounds are significantly different from those of explosive aliphatic compounds. Fig. 2(c) shows the fingerprint patterns of our phage-based sensor arrays when exposed to 100 ppb explosives. Our sensor system is superior to conventional biosensor system in that it clearly demonstrated unique color pattern for each of different VOC molecules, and even discerned substances belonging to the same chemical class, such as TNT and DNT. We carried out a limit of detection (LOD) analysis using the signal intensity of the CCD camera. As illustrated in Fig. S5, the WHW-type phage sensor shows a clear trend that the signal increases with concentrations of TNT

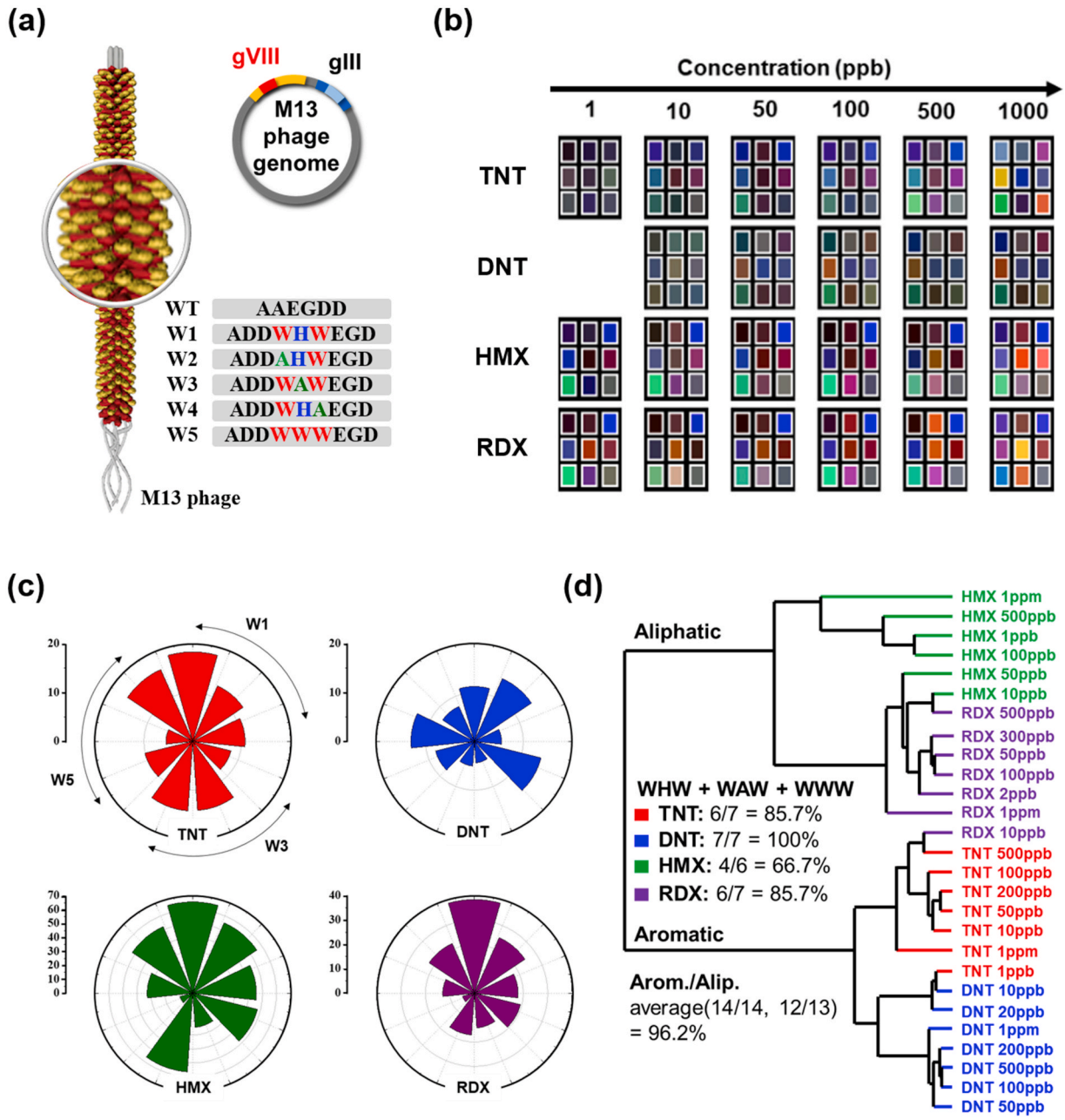


Fig. 2. Existence of specific aromatic amino acids in peptides with superior binding scores (top 5%) for target explosive molecules. High probability indicates that the specific amino acid is important in the binding of a target molecule.

and DNT being consistent with previous report (Oh et al., 2014). The LOD calculated from the trend lines was 79.6 ppb and 173.9 ppb for TNT and DNT, respectively. While the intensities of HMX and RDX also show linear correlation with their concentrations, the standard errors are rather larger than TNT and DNT.

Details of the fingerprint patterns of substances in the same chemical class were analyzed using the HCA method, as shown in Fig. S6. The CSR was used to score the phage sensor arrays, which are the combined systems of phages and peptides (sequences of WWW, WHW, WAW, AHW, and WHA), as well as a wild type. The experimental results were well explained by our computational prediction on the band redshift. The sensors composed of WWW, WHW, and WAW present the best sensing performance, as represented in Table 2. This means that band redshift is strictly correlated with the responsiveness of the colorimetric film. Furthermore, the HCA is shown in Fig. 2(d) is highly consistent in that the aromatic and aliphatic explosives are well classified by our optical biosensor. This score is almost equivalent to an electronic nose with all six genetic types, as illustrated in Fig. S13. This can significantly reduce the size of the electronic nose to almost a half of conventional biosensor. Our electronic nose equipped with highly sensitive and selective receptors is portable due to small sizes of sensing components and reduced number of receptors of high sensing capability.

In the analysis of the electronic nose composed of WHA and AHW peptides, the aromatic and aliphatic explosives were mixed. Thus, the total score of this electronic nose is lower than WHW-WHA-WWW scores. Electronic noses composed of WHW and other genetic types were further investigated. WHW showed the largest band shift among peptides in the sequence combination of A, H and W. Accordingly, the tripeptide WHW plays a decisive role in the performance of the electronic nose. Total scores of the electronic nose composed of WHW-WAW, WHW-WWW, and WHW-WHA integrated into our engineered phages were above 73. In short, the band shift predicted by first-

Table 2
Classification success rate (CSR) of phage-based electronic noses. All values were calculated based on the HCA classification method (Refer to Figs. S8, S9, S10, S11, and S12 for HCA dendrograms of WHW-WAW, WHW-WHA, WHW-WWW, WHA-AHW, and WHW-Wild, respectively.).

	Classification success rate (%)					
	TNT	DNT	HMX	RDX	Aromatic / Aliphatic	Total Score
WHW – WAW – WWW	85.7	100	66.7	85.7	96.2	81.3
WHW – WAW	85.7	100	66.7	85.7	92.3	78
WHW – WHA	71.4	71.4	66.7	85.7	100	73.8
WHW – WWW	71.4	71.4	66.7	85.7	100	73.8
WHA – AHW	71.4	100	66.7	85.7	82.1	66.5
WHW – Wild	57.1	71.4	33.3	85.7	92.6	57.3

principles calculations well describes experimentally measured electronic nose performance.

We further validated the accuracy of the electronic nose classification using controlled experimental tests. Introduction of WT sequences with no selectivity for explosives significantly reduced the total score when combined with other peptides. This result agrees well with our computational prediction and confirms that the binding affinity and the band shift are critical factors for the sensitivity of the electronic nose. The electronic nose composed of a single genetic type should be futile since the vector will respond with the same tendency. The response time for all the sensors was around 10 s, which is comparably rapid considering the real-time detecting nature of the phage-based colorimetric sensor (Oh et al., 2014). Also, the response time decreased with increasing VOC vapor pressure. It is another outstanding performance compared to conventional biosensor counterparts.

We also tested our sensing components under extreme environments as well. Previous results (Oh et al., 2014; Moon et al., 2017) have confirmed that the color changes of the phage films for target materials are the same even when exposed to severe EtOH (300 ppm), of which concentration is sufficient to saturate the color patterns of the sensor. It indicates that the addition of EtOH does not change the color of our phage film. The functional peptides on the phage surfaces, on the other hand, react selectively and effectively toward target molecules with excellent binding affinity. For chemical accidents, our sensor can be exposed to high thermal heat. We affirmed that our biosensor successfully works up to 200 °C and the structural integrity is sustained as high as 300 °C.

In our experimental tests, the phage-based multiple arrays detected even 1 ppb levels of TNT, RDX, and HMX with distinct colors for each of the specific target molecules. This is a remarkable improvement of conventional sensing resolution for 300 ppb TNT measured by the WHW peptide receptor alone (Kim et al., 2019). While optical property changes of the electronic nose occurred for the concentration below 1 ppb, our sensor clearly recognize color patterns down to 1 ppb, thus the resolution of our sensor is 1 ppb. Furthermore, our sensor is a real-time responding device, which is very useful for detecting specific VOC at the sites of accidents from the color signals with naked eyes. In addition, extra electronic wire or circuits are not necessary and can be easily implemented with cellphones (Oh et al., 2014). Our study provides a fundamental design principle for an outstanding optical biosensor in which the simple addition of sensing components of unspecified characteristics to the electronic nose is not enough to improve the sensing capability. Rather, understanding the molecular-level mechanism for the manifestation of optical properties is essential for making superior phage-based colorimetric sensors.

4. Conclusion

Interplay of big data driven computational screening of peptides followed by in-site experiments deployed unconventionally outstanding optical biosensor toward chemically inert VOCs. The electronic nose demonstrated in this study confirms that binding affinity and induced band shifts are the key descriptors for phage-based colorimetric sensors. The classification characteristics of the electronic nose were scored for quantitative analysis. Our colorimetric biosensors composed of phages (WWW-WAW-WHW) with superior chemical affinity and a large band shift demonstrated exceptionally high selectivity and detection resolution of 1 ppb. Even though our electronic nose can detect diverse molecules, the specification of the composition of VOC mixture is challenging since the simple statistical signal processing method used in this study may not be applicable. Regarding the issue, we are under development of a deep learning technique, which precisely mimics the mechanism of the human brain through large, multi-layer neural network. We expect that these insights into the phage-based electronic nose will play a significant role in the fields in need of electronic noses with precise sensing ability, including the detection of organic chemicals such as dangerous compounds, food, and disease diagnosis.

CRediT authorship contribution statement

Jungyun Park: Writing - original draft, Methodology, Investigation, Software, Validation, Formal analysis, Data curation. **Jong-Min Lee:** Writing - original draft, Formal analysis. **Hoje Chun:** Methodology, Validation, Software. **Yujin Lee:** Formal analysis. **Sung Jun Hong:** Validation, Software. **Hyunwook Jung:** Software. **Ye-Ji Kim:** Resources. **Won-Geun Kim:** Formal analysis. **Vasanthan Devaraj:** Formal analysis. **Eun Jung Choi:** Resources. **Jin-Woo Oh:** Conceptualization, Funding acquisition, Resources. **Byungchan Han:** Conceptualization, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the Creative Materials Discovery Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017M3D1A1039287) and the Global Frontier Program through the Global Frontier Hybrid Interface Materials (GFHIM) of National Research Foundation of Korea

(NRF) funded by the Ministry of Science and ICT (Project No. 2013M3A6B1078882).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bios.2021.112979.

References

Becke, A.D., 1988. Phys. Rev. 386, 3098–3100.
Bratkovic, T., et al., 2005. Biochem. Biophys. Res. Commun. 3323, 897–903.
Chung, W.J., et al., 2011. Nature 4787369, 364–368.
Dey, S., et al., 2020. Microsystems & Nanoengineering 635.
Dickinson, T.A., et al., 1996. Nature 3826593, 697–700.
Frenking, G., 2000. Theoretical Chemistry Accounts 187–189, 1033-4.
Frisch, M., et al., 2009. Gaussian Inc, Wallingford, CT.
Grimme, S., 2006. J. Comput. Chem. 2715, 1787–1799.
Hanwell, M.D., et al., 2012. J. Cheminf. 4.
Kelley, B.P., et al., 2015. J. Chem. Inf. Model. 558, 1771–1780.
Kim, B.H., et al., 2020. Science 3686486, 60-+.
Kim, W.G., et al., 2019. Org. Biomol. Chem. 1723, 5666–5670.
Ko, K., et al., 2020. Sci. Total Environ. 729, 138786.
Kresse, G., Furthmuller, J., 1996. Phys. Rev. B 5416, 11169–11186.
Li, M., et al., 2013. Biosens. Bioelectron. 43, 69–74.
Li, Z., et al., 2019. Chem. Rev. 1191, 231–292.
Liu, Q.T., et al., 2015. Talanta 136, 114–127.
Lunder, M., et al., 2005. Appl. Biochem. Biotechnol. 1272, 125–131.
Mascini, M., et al., 2004. Biosens. Bioelectron. 206, 1203–1210.
Mascini, M., et al., 2017. Biosens. Bioelectron. 93, 161–169.
Massie, C., et al., 2006. Sensor. Actuator. B Chem. 1132, 830–836.
Moon, J.S., et al., 2015. Mini-Reviews Org. Chem. 123, 271–281.
Moon, J.S., et al., 2017. Chem. Sci. 82, 1665-1665.
Morris, G.M., et al., 2009. J. Comput. Chem. 3016, 2785–2791.
Mulchandani, A., et al., 2001. Biosens. Bioelectron. 164–5, 225–230.
Noh, S.H., et al., 2018. J. Mater. Chem. 641.
Oh, J.W., et al., 2014. Nat. Commun. 5.
Perdew, J.P., et al., 1996. Phys. Rev. Lett. 7718, 3865–3868.
Pizzoni, D., et al., 2014. Biosens. Bioelectron. 52, 247–254.
Rakow, N.A., Suslick, K.S., 2000. Nature 4066797, 710–713.
Rock, F., et al., 2008. Chem. Rev. 1082, 705–725.
Sidhu, S.S., 2001. Biomol. Eng. 182, 57–63.
Singh, S., 2007. J. Hazard Mater. 1441–2, 15–28.
Smith, G.P., 1985. Science 2284705, 1315–1317.
Smith, G.P., Petrenko, V.A., 1997. Chem. Rev. 972, 391–410.
Steinfeld, J.I., Wormhoudt, J., 1998. Annu. Rev. Phys. Chem. 49, 203–232.
Strasser, B.J., 1999. Science 2865444, 1488–1490.
Tosco, P., et al., 2014. J. Cheminf. 6.
Trott, O., Olson, A.J., 2010. J. Comput. Chem. 312, 455–461.
Turner, A.P.F., Magan, N., 2004. Nat. Rev. Microbiol. 22, 161–166.
Venugopal, V., 2002. Biosens. Bioelectron. 173, 147–157.
Wang, J., et al., 2018. Sensor. Actuator. B Chem. 264, 279–284.
Weininger, D., 1988. J. Chem. Inf. Comput. Sci. 281, 31–36.
Wolcott, C.A., et al., 2015. J. Catal. 330, 197–207.