

Article

Design of Protein based Biosensors for Selective Detection of Benzene Group of Pollutants

Shamayeeta Ray, Santosh Panjekar, and Ruchi Anand

ACS Sens., **Just Accepted Manuscript** • DOI: 10.1021/acssensors.8b00190 • Publication Date (Web): 07 Aug 2018

Downloaded from <http://pubs.acs.org> on August 7, 2018

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.



ACS Publications

is published by the American Chemical Society, 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

Design of Protein based Biosensors for Selective Detection of Benzene Group of Pollutants

Shamayeeta Ray¹, Santosh Panjikar^{2,3} and Ruchi Anand^{*1,4}

¹ Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, Maharashtra, India

² Department of Biochemistry and Molecular Biology, Monash University, Victoria 3800, Australia

³ Australian Synchrotron, Victoria 3168, Australia

⁴ Wadhvani Research Center for Bioengineering, IIT Bombay, Mumbai 400076, India

*To whom correspondence may be addressed at: Dr. Ruchi Anand, Structural Biochemistry Laboratory, Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, Maharashtra, India.

E-mail: ruchi@chem.iitb.ac.in

ABSTRACT

Benzene and its derivatives form a class of priority pollutants whose exposure poses grave risk to human health. Since benzene lacks active functional groups, devising specific sensors for its direct detection from a milieu of aromatics has remained a daunting task. Here, we report three engineered protein-based biosensors that exclusively and specifically detect benzene and its derivatives upto a detection limit of 0.3 ppm. Further, the biosensor design has been engineered to create templates that possess the ability to specifically discriminate between alkyl substituted benzene derivatives; such as toluene, m-xylene and mesitylene. Interference tests with simulated wastewater samples reveal that the engineered biosensors can selectively detect a specific benzene compound in water samples containing a milieu of high concentrations of commonly occurring pollutants. This work demonstrates the potential of structure guided protein engineering as a competent strategy towards design of selective biosensors for direct detection of benzene group of pollutants from real time environmental samples.

KEYWORDS: MopR, benzene, m-xylene, mesitylene, protein engineering, selective biosensor.

1
2
3 Benzene and its derivatives, BTEX (benzene, toluene, ethylbenzene, xylene) are listed amongst
4 the top ten priority pollutants as per the EPA pollution database¹. These xenobiotics contaminate
5 the sub-surface as well as groundwater due to leaching from underground gas storage tanks and
6 landfills as well as from hazardous waste sites. At these sites, benzene based pollutants are
7 present as a primary component of gasoline and petroleum products^{2,3}. Exposure to benzene
8 causes severe health hazards including central nervous system and immune system dysfunction,
9 cancer etc^{4,5}. In recent years, benzene spillage has caused havoc, such as in Lanzhou, China on
10 April, 2014, where ground water was contaminated, because of oil spillage, resulting in more
11 than 2.4 million people being severely affected. Moreover, benzene pollution poses to be a
12 serious threat as its hydrocarbon derivatives are highly resistant to degradation, and in absence of
13 competent bacteria which can degrade them, these xenobiotics can accumulate in highly toxic
14 concentrations in environmental wastewaters⁶.

15
16 From a sensing perspective, devising sensors for specific quantification of a specific
17 BTEX is challenging. This is mostly because these hydrocarbons lack active functional groups,
18 which makes direct detection methods difficult to implement. Currently, the most prevalent
19 benzene sensors for air pollution monitoring are metal oxide based gas sensors^{7,8}. A few
20 cavitands based sensors that exploit CH- π interaction to select between BTEX have also been
21 reported. However, in these systems, the sample has to be heated above 150°C to attain any
22 selectivity⁹. Fewer detection devices exist to monitor the levels of benzene compounds at
23 ambient temperatures, such as in ground water and wastewater samples. Gas chromatography –
24 flame ionization detector (GC-FID) coupled with mass spectrometry constitutes the current
25 established method of detection^{10,11,12}. However, this method has limited portability and requires
26 pre-treatment and concentration of samples prior to any reasonable quantitative analysis¹³.
27 Hence, there is a pressing need to develop simpler, easy to handle, specific and direct detection
28 methods for *in situ* BTEX monitoring in wastewater samples.

29
30 A sub-class of soil bacteria like *Pseudomonas sp.*^{14,15} have adapted to survive in toxic
31 aromatic pollutant environment by evolving proteins through natural selection that can
32 specifically detect these xenobiotics and subsequently degrade them¹⁶. These regulatory sensor
33 modules provide scaffolds for development of specific biosensor design^{17,18,19}. Nevertheless,
34 extraction of these proteins with high yield and solubility in a suitable *in vitro* format has
35 remained challenging²⁰. For instance, the XylR protein, which is the natural benzene sensor,
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

could only be extracted via urea denaturation and subsequent refolding techniques²¹. Hence, such proteins have been exploited till date in the frame of whole cell biosensors, allowing method of detection limits in the range of 0.1 μM within assays times of 2-3 hrs^{22,23,24,25,26,27}. Very recently, we successfully obtained soluble protein for the phenol regulator, MopR from *Acinetobacter calcoaceticus* NCIB8250 and also solved the crystal structure of its pollutant sensing domain (MopR^{AB})²⁸. The structure in complex with phenol and its derivatives serves as a representative of the XylR/DmpR subclass of pollutant sensing NtrC regulators and reveals the exact pollutant-binding pocket and the sensor residues (Figure 1A, S1). Members of this family possess a modular architecture consisting of an N-terminal signal sensing (A) domain, that upon pollutant binding, transmits the signal to the tandemly located AAA+ ATP dependent(C) domain^{29,30} (Figure 1B). It is only upon activation that ATP hydrolysis occurs, which provides the readout for quantification of pollutant monitoring (Figure 1B). Substrate profiling of native MopR protein reveals that it is only capable of detecting smaller phenolics and is non-responsive towards benzene and its derivatives (Figure 2D)^{28,31}. This is because the MopR pocket is specifically designed to anchor the phenolic OH group by hydrogen bonding with W134 and H106 (Figure 1A). Here, in this work, by utilizing the knowledge of the three-dimensional pocket architecture in conjunction with *in silico* and experimental approaches, we have tweaked the MopR sensor design and successfully created sensors that can exclusively sense BTEX and can even distinguish between alkyl substitutions on the benzene scaffold. Moreover, because these biosensors have been developed with purified proteins, this has reduced the response/assay time to 10-15 min making them more suitable candidates for pollutant monitoring in contaminated environmental wastewaters.

RESULTS AND DISCUSSION

The primary tool used for alteration of the substrate scope was structure guided protein engineering via targeted mutagenesis. The three main designed mutant constructs (listed in 'Experimental Section') were first screened *in silico* and only competent mutant protein-sensor pairs were further taken for experimentation. The benzene series of compounds lack the phenolic OH group. Therefore, in order to alter the substrate scope, the first mutation considered, required compensating for the interactions that anchor the primary phenolic ligand. The crystal structure reveals that both the imidazole group of H106 and the indole moiety of W134 partake in

hydrogen bonding with the hydroxyl group of phenol (Figure 1A, 2A). Hence, the suggested design should have these interactions intrinsically satisfied by replacement of an appropriate amino acid that is able to retain the aromatic character of histidine as well as compensate for the interactions that were designed to stabilize the hydroxyl moiety. Analysis reveals that tyrosine residue, which possesses an aromatic hydroxyl group satisfies both these requirement. Thus replacement of the histidine by a tyrosine residue, (H106Y mutant, MopR^{HY}) allows the tyrosine OH group to interact with the indole nitrogen atom of the W134 group, thereby leading to minimal change in pocket architecture (Figure 2B, 2C). Docking studies with an *in silico* MopR^{HY} mutation of the native pocket show that OH group of the tyrosine occupies the position of the phenolic OH and encounters a direct clash with the ligand, hence obliterating binding of all phenolic compounds (Figure 2B). It was observed that the overall pocket volume of the MopR^{HY} reduces to 143.6 Å³ in the mutant in comparison to 158.3 Å³ in the native structure and the pocket takes a shape such that the benzene can snugly be accommodated (Figure 2C,S3). Stacking interactions with the aromatic core of Y106 provide additional support, thereby yielding favorable binding energy scores (Table S1).

To validate the above hypothesis, supporting experiments were conducted and the mutant version of MopR^{HY} was cloned and purified (Experimental Section). The malachite green based *in vitro* colorimetric ATPase activity assay performed with a series of phenol and benzene derivatives (Figure S2) clearly showed that MopR^{HY} preferentially exhibits ATP hydrolysis in response to benzene and toluene and is completely non-responsive to phenol and most of its derivatives (Figure 2E). For the moderately sized hydrocarbons like m-xylene, the sensor response was substantially reduced by almost 70% and only about 10% of the activity was retained for the tri-substituted bulkier ligands like mesitylene (Figure 2E). These results clearly indicate that the MopR^{HY} sensor activity has been primed to respond exclusively to smaller benzene derivatives leading to creation of an exclusive sensor for benzene and toluene. One of the intriguing questions that arose from the sensor activity data was how a single mutation in the sensor scaffold results in the biosensor discriminating between a mono-substituted toluene and di-substituted m-xylene (Figure 2E). To explain these observations, docking of the mutants with the mono, di and tri substituted series was undertaken (Figure S3). Results reveal that the shape of the pocket changes even by a single substitution and the pocket architecture of the MopR^{HY} is altered enough to select out alkyl substitutions. Comparison of the docked structures of toluene

and m-xylene show that the presence of the additional methyl group in m-xylene results in a shift in the overall position of this compound, thereby creating a steric clash with residues like F132 that constitute the wall of the pocket (Figure 3A, S3). This yields a non-ideal fit and lowers the binding energy of di-substituted benzene derivatives resulting in selective sensing (Figure S3). Bulkier hydrocarbons like mesitylene fare worse and are unable to even gain entry into the modified pocket (Figure S3D).

Our next goal was to design sensor models, which can selectively sense bigger alkyl substituted compounds. This is because many of the bulkier compounds like di and tri-substituted benzenes constitute some of the commonly found environmental pollutants. To accommodate multiple alkyl group containing benzene derivatives, the primary strategy adopted was to increase the pocket volume of MopR^{HY} by retaining its hydrophobic character. Scanning through the pocket residues, it became apparent that the base of the pocket could be trimmed to create space. A series of *in silico* studies were performed and based on docking analysis, it was concluded that a minimal mutation of the tyrosine residue (Y176) to a phenylalanine is sufficient to create space for m-xylene and other similar sized compounds. Docking with a series of benzene derivatives revealed that the MopR^{HY-YF} pocket fits best the moderately sized hydrocarbons like m-xylene and ethylbenzene (Table S1, Figure 3B, S4A). The MopR^{HY-YF} pocket allows the m-xylene ligand to orient such that both its methyl groups are accommodated in the space created at the base of the pocket (Figure 3B). Both smaller (benzene) and bulkier (mesitylene) derivatives exhibited reduced binding energies (Table S1, Figure S4). Due to the additional mutation, the increase in the pocket volume (150.3 Å³) results in the benzene moiety no longer to be snugly fitting in the MopR^{HY-YF} pocket (Figure S4B). On the other hand, for the bulkier benzene derivative, mesitylene, although the ligand could enter the pocket, but steric hindrance by the pocket wall residue such as F132, resulted in a scenario that was energetically unfavorable (Figure S4D). *In silico* studies further indicated that among di-substituted aromatic derivatives, m-xylene is the most preferred isomer followed by o-xylene (Figure S5A-B). However, analysis reveals that the mutated pocket is shaped such that p-xylene is unable to fit in a orientation conducive for signal transduction (Figure S5C) (detailed in Supporting Information).

Experimental creation of the MopR^{HY-YF} double mutant, followed by purification confirms the *in silico* findings. *In vitro* ATPase assay with purified MopR^{HY-YF} confirmed that

the hydrolysis activity is maximum in response to m-xylene and ethylbenzene sized derivatives (Figure 2F). Both for the smaller benzene as well as bulkier hydrocarbons, the ATPase activity is reduced by 70-80% (Figure 2F). Similarly, among the xylene isomers, m-xylene shows maximum activity followed by o-xylene which exhibits 80% activity, whereas the sensor exhibits poor sensitivity for p-xylene (Figure S5D). Surprisingly, the sensor also has dramatically reduced binding for toluene but good sensitivity for ethylbenzene. Analysis of the docked structure with toluene shows that pocket mutation changes the shape such that its methyl group comes in close proximity with the OH of Y106 (Figure S4C). However, the ethyl group of ethylbenzene owing to its kink, tilts the ligand which adopts a conformation such that the ethyl group snugly fits in the space between Y106 and W134 (Figure S4A). These results reiterate the fact that subtle changes in pocket architecture can yield in high selectivity among the BTEX class of compounds. Hence, MopR^{HY_YF} sensor design is inductive in stabilization of moderate sized benzenes like m-xylene, providing a framework for selective sensor design for medium sized benzene compounds.

An appropriate sensor framework for the bulkier hydrocarbons like mesitylene still remained a question. Hence, search for an apt mutation in the MopR^{HY_YF} pocket, which leads to a volume increment as well as retains the stability of the hydrophobic core was undertaken using previously described tools. A substitution of the phenylalanine residue by alanine (F132A) leads to the creation of the MopR^{HY_YF_FA} mutant, which exhibits an appropriate increase in pocket volume (201.8 Å³) and is suitable to accommodate mesitylene (Figure S6A). Derivatives like benzene and m-xylene were no longer energetically favored, as the pocket volume was too large to lock these smaller ligands in unique orientations (Table S1, Figure 3C, S6). Creation of this triple mutation was experimentally undertaken and the *in vitro* ATPase assay with MopR^{HY_YF_FA} showed that maximum ATPase activity was observed in presence of the bulky compound mesitylene (Figure 2G). Further, the triple mutant exhibited highly reduced ATP hydrolysis in response to the smaller benzene derivatives and was completely non-responsive towards any of the phenol derivatives (Figure 2G). Based on these observations, it was inferred that MopR^{HY_YF_FA} serves as a model sensor design for selective sensing of bulkier toxic aromatic hydrocarbons. To further ascertain that all the designed sensor constructs were stable and changes in activity were not due to structural instability introduced by the mutations, confirmatory circular dichroism studies were performed on all the mutants (Figure S7). No

change or loss in secondary structure was observed for the experimentally tested mutants. Hence, it was concluded that all the observed alterations in activity for the MopR variants was a direct effect of the functional change and not attributed to any structural loss.

With the repertoire of these newly generated selective benzene based sensors, the next task was to test the viability and efficiency of these sensor designs in a realistic setting that is crowded by several pollutants of varied chemical profile. To test the sensor performance under harsh conditions, simulated waste water (SWW) was generated by adding diverse range of chemicals at high concentrations (1mM) like salts, metals, organics that are commonly found as pollutants in environmental wastewaters which are otherwise non-inducers of MopR sensor activity (composition detailed in Experimental Section). Interference analysis on the MopR^{HY} and MopR^{HY-YF} sensor model systems was performed by adding varying concentrations (0.5-20 μ M) of their best inducers (m-xylene in case of MopR^{HY-YF} and benzene in case of MopR^{HY}) to the SWW (black curves in Figure 4) and comparing the activity data with standard solutions(SS) of the inducers in the same concentration range (red curves in Figure 4). The concentration range chosen for testing interference was selected such that it spans limits that are well below the permissible toxicity limit of the contaminants as per EPA and OSHA. The aim was to gauge both selectivity and sensitivity of the biosensors when the random pollutants co-exist with the target inducers. The MopR^{HY} sensor pocket is efficiently tuned to only accept benzene and its closely related aromatics (toluene) and therefore presence of ions, salts, phenolics etc, only affected the signal by ~3-5% (Figure 4A). Overall, it was observed that for MopR^{HY}, there is negligible interference from the non-inducers (in SWW) across the sensing concentration range of benzene (0.5-20 μ M). This is likely because the MopR^{HY} sensor pocket is too small for non-specific interactions with other compounds thus, restricting undesired activity. Therefore, addition of random pollutants does not significantly seem to affect sensor performance in case of MopR^{HY}. On the other hand, in the case of MopR^{HY-YF} sensor design, for the non-inducing pollutants in SWW, 5-10% competitive negative interference towards m-xylene sensing was observed (Figure 4B). This could be attributed to increased pocket size that now allows more type of substrates to enter the pocket allowing them to co-exist with m-xylene.

Since it is possible that in certain type of pollution scenario, both m-xylene and benzene may co-exist, the interference of the MopR^{HY} and MopR^{HY-YF} sensor constructs in presence of both these structurally similar compounds were also gauged (Figure 4). Previous

analysis clearly indicates that the ATPase signal is always biased towards the designed inducer for a particular sensor system by 80% or more (Figure 2E-G). However, to quantitatively determine the effect, three different concentrations (10 μ M, 100 μ M, 1000 μ M) of a structurally similar interferent were added to the standard solutions(SS) of the inducers (detailed in 'Experimental Section'). In the case of MopR^{HY}, the interference from m-xylene and in MopR^{HY-YF}, effect from benzene were tested (green curves in Figure 4A). For MopR^{HY} benzene sensor, presence of m-xylene affected the signal minimally in the low to medium concentration ranges. Even at 10 times the m-xylene concentration the signal was only enhanced by about 7-10 % (Figure 4A). Thus, the MopR^{HY} sensor is very robust for sensing smaller aromatics. In contrast, for benzene as a competitor of m-xylene in the pocket of MopR^{HY-YF}, an additive effect was created which is about ~10% at low concentrations of benzene (Figure 4B). In a mixture that has ten times more benzene than m-xylene, the effect becomes more prominent (around ~15-20%) (Figure 4B). However, the advantage of creating both the benzene specific MopR^{HY} and m-xylene specific MopR^{HY-YF} sensors for monitoring this pollutant is that it allows for a direct comparison of their performance. The dual data enables the amount of each pollutant in the mixture to be evaluated. Especially, since the MopR^{HY} benzene sensor exhibits low level of interference, the data from both the sensors analyzed in conjunction allows estimation of individual amounts. The interference analysis highlights the importance of understanding the complex interactions and crosstalk between different pollutants and biosensors in a real-time scenario. Hence, for *in situ* monitoring, quantifying these effects aid in creation of efficient real-time sensors. Moreover, the MopR biosensor design is highly thermostable as well as requires low response time (minutes) to generate substantial signal (Figure S8) and possess both reasonable selectivity and sensitivity. All these factors coupled with exclusive shape-complementarity based bio-sensing design yields a great potential for this system towards quantitative estimation of these benzene derivatives in real time contaminated environmental water bodies that harbor a milieu of different pollutants.

CONCLUSION

In conclusion, here we demonstrate that structure guided protein engineering is a powerful tool for creation of selective biosensor models. Architectures that were originally only tuned to sense phenol and its derivatives were modified by logic based tweaking to exclusively accept select

BTEX compounds. Moreover, these direct detection protein-based sensors have the distinct advantage of possessing the ability to operate at ambient conditions. Furthermore, here, for the first time we have been able to create sensors that can differentiate between methyl substituted benzene derivatives such as toluene, xylene as well as mesitylene. These robust sensors are thermostable and do not exhibit substantial cross reactivity while retaining both selectivity and specificity in a milieu of environmental pollutants. The bio-sensing system serves as an important stepping-stone to fuel development of commercial direct detection based low cost BTEX biosensors.

EXPERIMENTAL SECTION

Docking Studies. Based on the detailed analysis of the pollutant binding pocket using the phenol bound crystal structure of the signal sensing domain of MopR (MopR^{AB})²⁸ (Figure S1), few site-specific mutations of the pocket residues were designed *in silico*. The first mutation designed was that of a tyrosine substitution of the key sensor residue H106 (MopR^{HY}). Further, to accommodate bulkier benzene derivatives with more than one substituent, two more mutations were designed sequentially in the MopR^{HY} construct. They include - (i) phenylalanine substitution of Y176 (MopR^{HY-YF}) and (ii) alanine substitution of F132 (MopR^{HY-YF-FA}). Docking experiments using AutoDock version 4.2³² of all these mutants were performed with phenol and a set of benzene derivatives (benzene, toluene, m-xylene, ethylbenzene and mesitylene) (Figure S2). This was performed to predict alterations in specificity of MopR. A monomeric subunit of the X-ray structure of the MopR^{AB}-effector complex²⁸ (PDB code: 5KBE) was used for docking so as to predict affinity of the MopR^{AB} mutants towards different pollutants. The template PDB used for the docking experiments was manually modified each time based on the mutants used. The final MopR^{AB} mutant PDBs used for docking had the previously present ligand coordinates deleted from them. All docking runs were conducted by using a genetic algorithm (GA) in AutoDock version 4.2³² against the target aromatic effectors. Each ligand for a particular mutant version of MopR^{AB} was scored according to a free energy cost function (ΔG^*) that accounts for van der waals, hydrogen bonding, electrostatic, solvation, and torsional free energy terms. The grid box for docking was selected in the binding pocket region, and rigid docking was performed with 250 runs for each ligand - MopR^{AB} mutant

combinations. On the basis of the estimated free energy of binding (ΔG^*) calculated in docking, the top-ranked ligand orientations were selected (listed in Table S1).

DNA manipulations, over expression and purification of the recombinant proteins.

To validate the docking results, the same set of *in silico* mutations of MopR (MopR^{HY}, MopR^{HY_YF} and MopR^{HY_YF_FA}) were experimentally incorporated into the native construct of MopR (MopR^{ABC}) from *Acinetobacter calcoaceticus* NCIB8250, that was previously cloned into modified pET vector³³. All the point mutations were performed employing standard site-directed mutagenesis protocol using the "site-directed mutagenesis kit" from Kapa biosystems. The different mutant protein constructs were subsequently transformed into *Escherichia coli* BL21(DE3)plysS cells and grown at 37°C till OD₆₀₀ reached 0.6-0.8 and then induced with 0.7mM IPTG (isopropyl- β -D- thiogalactopyranoside) and 5mM MgCl₂ at 16°C. All these mutants were expressed as C-terminal His tag fusion proteins and were purified using Ni-NTA resin employing standard His-tagged affinity purification protocol as described previously³³. The eluted fractions were further concentrated and desalted with buffer containing 25mM TRIS buffer, pH 7.5; 100mM NaCl using an Econo- Pac 10DG (Bio-Rad, CA, USA) column. All the desalted fractions were pooled together and concentrated upto 6-8mg/ml and were flash-frozen in liquid N₂, and stored at -80 °C until they were used.

Preparation of the chemical compounds. Saturated solutions (100mM) of the volatile benzene derivatives were prepared in DMSO by shaking for few hours in a 37°C incubator. Appropriate volumes of the saturated solutions were transferred to protein desalting buffer to achieve the desired liquid-phase concentrations (0.5 μ M-1000 μ M) for each of the benzene compounds based on their mass balance and gas-liquid equilibrium as predicted by Henry's Law coefficients at 37°C³⁴ to make 1ml working stocks in 1.5ml eppendorfs. The working stock concentrations of the phenol derivatives, being non-volatile, were diluted in protein buffer at appropriate concentrations from 100mM stock solutions in water. All the saturated and working stock concentrations were re-confirmed by measuring their UV absorbance at λ_{max} and calculating their concentrations based on the extinction coefficient for each compound at their previously reported λ_{max} values.

***In vitro* ATPase assay.** To perform the *in vitro* biosensing ATPase assay, 2 μ M of protein sample (native and mutated MopR^{ABC}) was first incubated with 10 μ M of each aromatic compound to be tested which, included the different phenol and benzene derivatives listed in

Figure S2. The concentrations of all the compounds used in the assay correspond to ~1-2 ppm which lie within the approximate environmental risk limits for these pollutants as per Occupational Safety and Health Administration (OSHA). The ATPase assay was performed based on protocol reported previously³³. All the absorbance values for the different MopR^{ABC} constructs directly represent their absolute ATPase activity. Each of the experiments was performed in triplicates and the error estimate has been represented as error bars within the figures.

Interference assay with simulated wastewater (SWW) sample. For performing the interference assay, the following types of solutions were first prepared- (a) standard solutions (SS) of inducers m-xylene (for the MopR^{HY-YF} sensor) and benzene (for the MopR^{HY} sensor) at a concentration range of (0.5-20 μ M) ; (b) simulated wastewater (SWW) sample (mimicking a real time contaminated environmental sample) having the same concentration range of the inducer pollutants along with a mixture of different commonly found environmental pollutants (at 1mM concentration of each interferent) including phenol, 3-chlorophenol, o-cresol (as representative phenol derivatives as part of volatile organic contaminants), NH₄⁺ (as a representative cationic contaminant), Cl⁻ (as a representative anionic contaminant) and Co²⁺ (as a representative metal contaminant) ; (c) standard solutions (SS) of the inducer pollutants mixed with fixed concentrations (10 μ M, 100 μ M and 1000 μ M) of structurally similar compounds; benzene (for the MopR^{HY-YF} sensor) and m-xylene (for the MopR^{HY} sensor). 2 μ M protein of the MopR^{HY-YF} mutant (which can preferentially sense m-xylene) and the MopR^{HY} mutant (which can preferentially sense benzene) were first induced with their corresponding solutions of all three types and then tested for interference from other pollutants using the *in vitro* ATPase assay protocol as described above. The resultant ATPase activities in all the solutions were compared for analysis purpose. All the absorbance values have been represented as percent ATPase activity. Each of the experiments was performed in triplicates and the error estimate has been represented as error bars within the figure.

Circular dichroism (CD) studies. The native and mutant MopR proteins were tested (at 0.2mg/ml each) for structural stability using CD studies. All the protein samples were prepared in phosphate buffer (25mM sodium phosphate pH-7.5, 80mMNaCl). Scans were performed at 20°C using 0.1 cm path length quartz cuvettes with 8 sec differential integration time at a scan rate of 50 nm/sec. For thermal stability analysis, the MopR^{ABC} protein at 0.5mg/ml

was tested using a CD based thermal denaturation experiment. Scans were performed at a temperature range of 20°C to 95°C using 0.1 cm path length quartz cuvettes with 16 sec differential integration time at a scan rate of 100 nm/sec with 3 min delay time per temperature change.

SUPPORTING INFORMATION

The Supporting Information is available free of charge on the ACS Publications Website.

Supporting Information includes Supporting results, eight figures and one table.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ruchii@chem.iitb.ac.in.

Author Contributions

S.R., R.A., S.P. designed research; S.R. performed research; R.A. contributed new reagents/analytic tools; S.R., S.P., R.A. analyzed data; and S.R., R.A. wrote the paper.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS

This work was funded by DST, Government of India (Grant Numbers EMR/2015/002121 and DST/TM/WTI/2K16/252) .

REFERENCES

1. Keith, L.; Telliard, W., ES&T Special Report: Priority pollutants: I-a perspective view. *Environ. Sci. Technol.* **1979**, *13* (4), 416-423.
2. Weisel, C. P., Benzene exposure: An overview of monitoring methods and their findings. *Chem. Biol. Interac.* **2010**, *184* (0), 58-66.
3. Verma, D.; Tombe, K., Benzene in Gasoline and Crude Oil: Occupational and Environmental Implications. *J. Occup. Environ. Hyg.* **2002**, *63*, 225-230.
4. Falzone, L.; Marconi, A.; Loreto, C.; Franco, S.; Spandidos, D. A.; Libra, M., Occupational exposure to carcinogens: Benzene, pesticides and fibers. *Mol. Med. Rep.* **2016**, *14* (5), 4467-4474.

5. Bahadar, H.; Mostafalou, S.; Abdollahi, M., Current understandings and perspectives on non-cancer health effects of benzene: A global concern. *Toxicol. Appl. Pharmacol.* **2014**, 276 (2), 83-94.
6. Fuchs, G.; Boll, M.; Heider, J., Microbial degradation of aromatic compounds — from one strategy to four. *Nat. Rev. Micro.* **2011**, 9 (11), 803-816.
7. Vaishnav, V. S.; Patel, S. G.; Panchal, J. N., Development of ITO thin film sensor for detection of benzene. *Sens. Actuators B-Chem.* **2015**, 206, 381-388.
8. Mabrook, M.; Hawkins, P., Benzene sensing using thin films of titanium dioxide operating at room temperature. *Sensors* **2002**, 2 (9), 374-382.
9. Trzcinski, J. W.; Pinalli, R.; Riboni, N.; Pedrini, A.; Bianchi, F.; Zampolli, S.; Elmi, I.; Massera, C.; Ugozzoli, F.; Dalcanale, E., In Search of the Ultimate Benzene Sensor: The EtQxBox Solution. *ACS Sens.* **2017**, 2 (4), 590-598.
10. Almeida, C. M. M.; Boas, L. V., Analysis of BTEX and other substituted benzenes in water using headspace SPME-GC-FID: method validation. *J. Environ. Monitor.* **2004**, 6 (1), 80-88.
11. Zwank, L.; Schmidt, T. C.; Haderlein, S. B.; Berg, M., Simultaneous Determination of Fuel Oxygenates and BTEX Using Direct Aqueous Injection Gas Chromatography Mass Spectrometry (DAI-GC/MS). *Environ. Sci. Technol.* **2002**, 36 (9), 2054-2059.
12. Grbić-Galić, D.; Vogel, T. M., Transformation of toluene and benzene by mixed methanogenic cultures. *Appl. Environ. Microbiol.* **1987**, 53 (2), 254-260.
13. Schweigkofler, M.; Niessner, R., Determination of Siloxanes and VOC in Landfill Gas and Sewage Gas by Canister Sampling and GC-MS/AES Analysis. *Environ. Sci. Technol.* **1999**, 33 (20), 3680-3685.
14. Timmis, K. N.; Pieper, D. H., Bacteria designed for bioremediation. *Trends Biotechnol.* **1999**, 17 (5), 201-204.
15. Diaz, E.; Prieto, M. A., Bacterial promoters triggering biodegradation of aromatic pollutants. *Curr. Opin. Biotechnol.* **2000**, 11 (5), 467-75.
16. Tropel, D.; van der Meer, J. R., Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol. Mol. Biol. Rev.* **2004**, 68 (3), 474-500.
17. O'Neill, E.; Ng, L. C.; Sze, C. C.; Shingler, V., Aromatic ligand binding and intramolecular signalling of the phenol-responsive σ^{54} -dependent regulator DmpR. *Mol. Microbiol.* **1998**, 28 (1), 131-141.

18. Rappas, M.; Bose, D.; Zhang, X., Bacterial enhancer-binding proteins: unlocking sigma54-dependent gene transcription. *Curr. Opin. Struct. Biol.* **2007**, *17* (1), 110-6.
19. Shingler, V.; Moore, T., Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolizing *Pseudomonas* sp. strain CF600. *J. Bacteriol.* **1994**, *176* (6), 1555-60.
20. Wikstrom, P.; O'Neill, E.; Ng, L. C.; Shingler, V., The regulatory N-terminal region of the aromatic-responsive transcriptional activator DmpR constrains nucleotide-triggered multimerisation. *J. Mol. Biol.* **2001**, *314* (5), 971-84.
21. Kim, M. N.; Park, H. H.; Lim, W. K.; Shin, H. J., Construction and comparison of *Escherichia coli* whole-cell biosensors capable of detecting aromatic compounds. *J. Microbiol. Meth.* **2005**, *60* (2), 235-245.
22. Kobatake, E.; Niimi, T.; Haruyama, T.; Ikariyama, Y.; Aizawa, M., Biosensing of benzene derivatives in the environment by luminescent *Escherichia coli*. *Biosens. Bioelectron.* **1995**, *10* (6), 601-605.
23. Ikeno, S.; Ogino, C.; Ito, T.; Shimizu, N., Detection of benzene derivatives by recombinant *E. coli* with Ps promoter and GFP as a reporter protein. *Biochem. Eng. J.* **2003**, *15* (3), 193-197.
24. Di Gennaro, P.; Bruzzese, N.; Anderlini, D.; Aiossa, M.; Papacchini, M.; Campanella, L.; Bestetti, G., Development of microbial engineered whole-cell systems for environmental benzene determination. *Ecotoxicol. Environ. Safety* **2011**, *74* (3), 542-549.
25. Stiner, L.; Halverson, L. J., Development and Characterization of a Green Fluorescent Protein-Based Bacterial Biosensor for Bioavailable Toluene and Related Compounds. *Appl. Environ. Microbiol.* **2002**, *68* (4), 1962-1971.
26. Tecon, R.; Beggah, S.; Czechowska, K.; Sentchilo, V.; Chronopoulou, P.-M.; McGenity, T. J.; van der Meer, J. R., Development of a Multistrain Bacterial Bioreporter Platform for the Monitoring of Hydrocarbon Contaminants in Marine Environments. *Environ. Sci. Technol.* **2010**, *44* (3), 1049-1055.
27. Shingleton, J. T.; Applegate, B. M.; Nagel, A. C.; Bienkowski, P. R.; Sayler, G. S., Induction of the tod Operon by Trichloroethylene in *Pseudomonas putida* TVA8. *Appl. Environ. Microbiol.* **1998**, *64* (12), 5049-5052.

28. Ray, S.; Gunzburg, M. J.; Wilce, M.; Panjikar, S.; Anand, R., Structural Basis of Selective Aromatic Pollutant Sensing by the Effector Binding Domain of MopR, an NtrC Family Transcriptional Regulator. *ACS Chem. Biol.* **2016**, *11*(8), 2357–2365.
29. Shingler, V., Signal sensing by σ 54-dependent regulators: derepression as a control mechanism. *Mol. Microbiol.* **1996**, *19* (3), 409-416.
30. Studholme, D. J.; Dixon, R., Domain architectures of sigma54-dependent transcriptional activators. *J. Bacteriol.* **2003**, *185* (6), 1757-67.
31. Schirmer, F.; Ehrt, S.; Hillen, W., Expression, inducer spectrum, domain structure, and function of MopR, the regulator of phenol degradation in *Acinetobacter calcoaceticus* NCIB8250. *J. Bacteriol.* **1997**, *179* (4), 1329-1336.
32. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785-91.
33. Ray, S.; Panjikar, S.; Anand, R., Structure Guided Design of Protein Biosensors for Phenolic Pollutants. *ACS Sens.* **2017**, *2* (3), 411-418.
34. Chang, H. L.; Alvarez-Cohen, L., Transformation capacities of chlorinated organics by mixed cultures enriched on methane, propane, toluene, or phenol. *Biotechnol. Bioeng.* **1995**, *45* (5), 440-9.

FIGURE LEGENDS

Figure 1. Design of *in vitro* biosensor. (A) Pollutant binding pocket of MopR where W134 and H106 anchor the phenolic OH group²⁸[Reprinted (in part) from 'Ray, S.; Gunzburg, M. J.; Wilce, M.; Panjikar, S.; Anand, R., Structural Basis of Selective Aromatic Pollutant Sensing by the Effector Binding Domain of MopR, an NtrC Family Transcriptional Regulator. *ACS Chem. Biol.* 2016, 11 (8), 2357–2365'. Copyright © 2016 American Chemical Society]. (B) Model design of *in vitro* biosensor based on MopR.

Figure 2. Selective *in vitro* biosensor designs for benzene and its derivatives. The panels (A-C) represent docked ligands for the following MopR mutants - (A) phenol with native MopR^{AB}, (B) phenol with MopR^{HY}, (C) benzene with MopR^{HY}. The H106 residue in panel (A) and Y106 in panels (B) and (C) are in surface representation. Panels (D-G) represent *in vitro* biosensing

ability of above sensor designs. The inducing concentrations of all the aromatic compounds used in the assay is 10 μ M.

Figure 3. Structure guided selective sensing of m-xylene. The panels (A-C) represent docked m-xylene (in yellow) in the MopR mutant pockets - (A) MopR^{HY}, (B) MopR^{HY-YF}, (C) MopR^{HY-YF-FA}. The MopR^{HY-YF} pocket exhibits the best geometry and energetics for m-xylene.

Figure 4. Interference tests on to gauge selectivity and sensitivity of biosensor designs. (A) Interference studies on MopR^{HY} sensor where sensing of benzene (0.5-20 μ M) in a simulated wastewater (SWW) sample (having 1mM each of various non-inducing pollutants, black curve) has been compared to standard solutions (SS) of benzene (red curve). The different green curves further represent interference tests in presence of 10 μ M(light green), 100 μ M(olive green) and 1000 μ M(bottle green) of a structurally homologous compound, m-xylene, mixed with benzene (0.5-20 μ M) in the SS. (B) Interference studies on MopR^{HY-YF} sensor where sensing of m-xylene (0.5-20 μ M) in a simulated wastewater (SWW) sample (having 1mM each of various non-inducing pollutants, black curve) has been compared to standard solutions (SS) of m-xylene (red curve). The different green curves further represent interference tests in presence of 10 μ M(light green), 100 μ M(olive green) and 1000 μ M(bottle green) of a structurally homologous compound, benzene, mixed with m-xylene (0.5-20 μ M) in the SS. Detailed composition of the solutions that were tested have been provided in 'Experimental Section'.

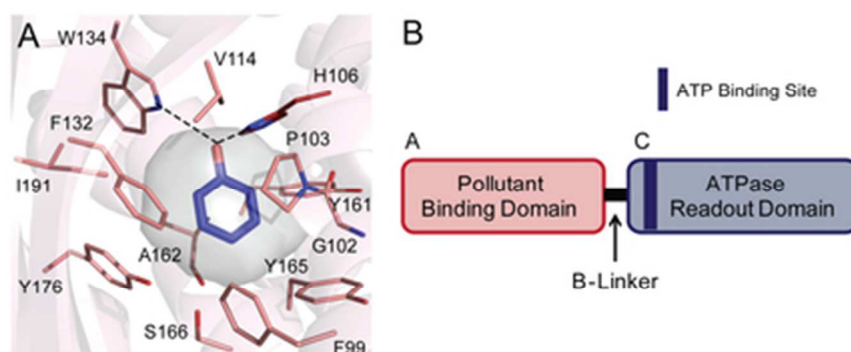


Figure 1. Design of in vitro biosensor. (A) Pollutant binding pocket of MopR where W134 and H106 anchor the phenolic OH group²⁸[Reprinted (in part) from 'Ray, S.; Gunzburg, M. J.; Wilce, M.; Panjikar, S.; Anand, R., Structural Basis of Selective Aromatic Pollutant Sensing by the Effector Binding Domain of MopR, an NtrC Family Transcriptional Regulator. ACS Chem. Biol. 2016, 11 (8), 2357–2365'. Copyright © 2016 American Chemical Society]. (B) Model design of in vitro biosensor based on MopR.

37x16mm (300 x 300 DPI)

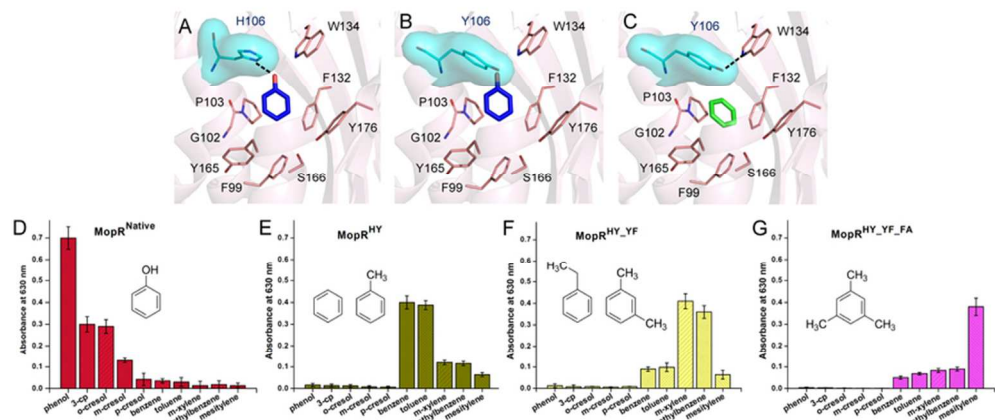


Figure 2. Selective in vitro biosensor designs for benzene and its derivatives. The panels (A-C) represent docked ligands for the following MopR mutants - (A) phenol with native MopRAB, (B) phenol with MopRHY, (C) benzene with MopRHY. The H106 residue in panel (A) and Y106 in panels (B) and (C) are in surface representation. Panels (D-G) represent in vitro biosensing ability of above sensor designs. The inducing concentrations of all the aromatic compounds used in the assay is 10 μ M.

76x33mm (300 x 300 DPI)

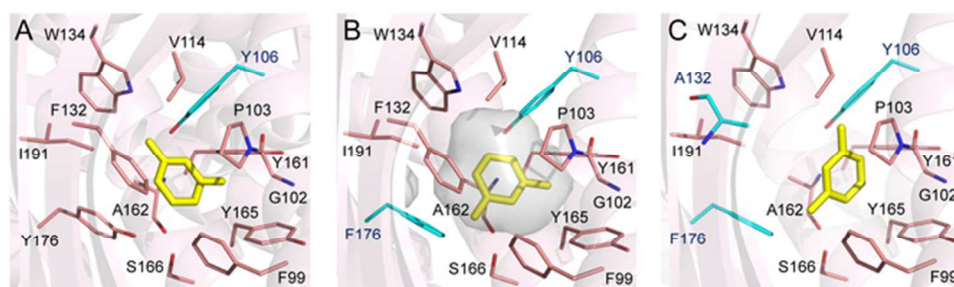


Figure 3. Structure guided selective sensing of m-xylene. The panels (A-C) represent docked m-xylene (in yellow) in the MopR mutant pockets - (A) MopRHY, (B) MopRHY_YF, (C) MopRHY_YF_FA. The MopRHY_YF pocket exhibits the best geometry and energetics for m-xylene.

58x19mm (300 x 300 DPI)

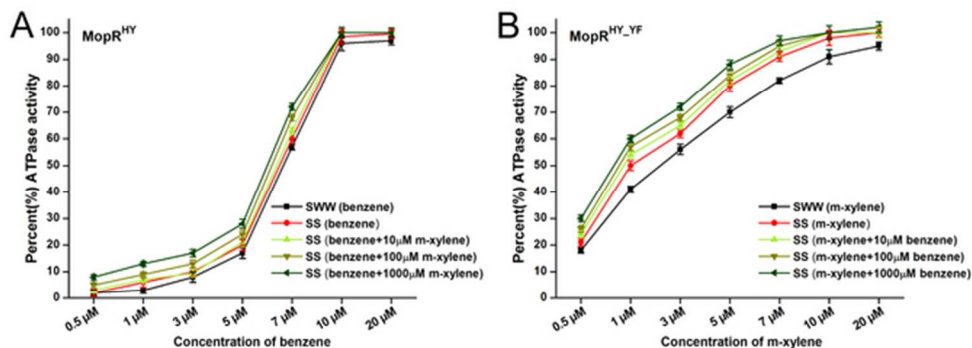


Figure 4. Interference tests on to gauge selectivity and sensitivity of biosensor designs. (A) Interference studies on MopRHY sensor where sensing of benzene (0.5-20 μ M) in a simulated wastewater (SWW) sample (having 1mM each of various non-inducing pollutants, black curve) has been compared to standard solutions (SS) of benzene (red curve). The different green curves further represent interference tests in presence of 10 μ M(light green), 100 μ M(olive green) and 1000 μ M(bottle green) of a structurally homologous compound, m-xylene, mixed with benzene (0.5-20 μ M) in the SS. (B) Interference studies on MopRHY_YF sensor where sensing of m-xylene (0.5-20 μ M) in a simulated wastewater (SWW) sample (having 1mM each of various non-inducing pollutants, black curve) has been compared to standard solutions (SS) of m-xylene (red curve). The different green curves further represent interference tests in presence of 10 μ M(light green), 100 μ M(olive green) and 1000 μ M(bottle green) of a structurally homologous compound, benzene, mixed with m-xylene (0.5-20 μ M) in the SS. Detailed composition of the solutions that were tested have been provided in 'Experimental Section'.

58x23mm (300 x 300 DPI)

