Environment International 147 (2021) 106339

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

# Identification of toxicity pathway of diesel particulate matter using AOP of PPARγ inactivation leading to pulmonary fibrosis

Jaeseong Jeong, Su-yong Bae, Jinhee Choi *

*School of Environmental Engineering, University of Seoul, 163 Seoulsiripdae-ro, Dongdaemun-gu, Seoul 02504, Republic of Korea*

ARTICLE INFO

Handling Editor: Martí Nadal

*Keywords:*
Diesel particulate matter
Adverse Outcome Pathway
ToxCast
Deep learning
Mixture toxicity
Molecular docking

ABSTRACT

Diesel particulate matter (DPM), a major subset of urban fine particulate matter (PM2.5), raises huge concerns for human health and has therefore been classified as a group 1 carcinogen by the International Agency for Research on Cancer (IARC). However, as DPM is a complex mixture of various chemicals, understanding of DPM's toxicity mechanism remains limited. As the major exposure route of DPM is through inhalation, we herein investigated its toxicity mechanism based on the Adverse Outcome Pathway (AOP) of pulmonary fibrosis, which we previously submitted to AOPWiki as AOP ID 206 (AOP206). We first screened whether individual chemicals in DPM have the potential to exert their toxicity through AOP206 by using the ToxCast database and deep learning models approach, then confirmed this by examining whether DPM as a mixture alters the expression of the molecular initiating event (MIE) and key events (KEs) of AOP206. For identifying the activeness of the component chemicals of DPM, we used 24 ToxCast assays potentially related to AOP206 and deep learning models based on these assays, which were identified and developed in our previous study. Of the 100 individual chemicals in DPM, 34 were active in PPARγ (MIE)-related assay, of which 17 were active in one or more KEs. To further identify whether individual chemicals in DPM are related to the MIE of AOP206, we performed molecular docking simulation on PPARγ for the chemicals showing activeness. Benzo[e]pyrene, benzo[a]pyrene and other related chemicals were the most likely to bind to PPARγ. In *in vitro* experiments, PPARγ activity increased with exposure of the DPM mixture, and the protein expression of PPARγ (MIE), and fibronectin (AO) also tended to be increased. Overall, we have demonstrated that AOP206 can be applied to identify the toxicity pathway of DPM. Further, we suggest that applying the AOP approach using ToxCast and deep learning models is useful for identifying potential toxicity pathways of chemical mixtures, such as DPM, by determining the activity of individual chemicals.

## 1. Introduction

Fine particulate matter (PM2.5) is defined as "fine" sized particulate matter with an aerodynamic diameter of less than 2.5 μm, and its small size causes a number of adverse effects on human health, which is of increasing interest in recent years (Achilleos et al., 2017; Schlesinger, 2007). Various epidemiology and toxicology studies have been conducted using PM2.5 sampled in the real world as well as using standard materials (Park et al., 2018; Piao et al., 2018; Wang et al., 2019). Those studies suggested PM2.5 could be related to development of various diseases, such as cardiovascular and respiratory diseases (Song et al., 2019; Wang et al., 2019). Oxidative stress and inflammation were most frequently reported as underlying mechanism of toxicity of PM2.5 (He et al., 2017). Most PM2.5 derives from combustion, such as gasoline and diesel fuels by motor vehicles, burning of natural gas to generate electricity, and wood burning (Robinson et al., 2010). Diesel engines emit a complex mixture of air pollutants, including both gaseous and solid materials. The solid material in diesel exhaust is known as diesel particulate matter (DPM). More than 90% of DPM is less than 1 μm in diameter (Zheng et al., 2017). The International Agency for Research on Cancer (IARC) classified diesel engine exhaust as Group 1 (carcinogenic to humans), based on sufficient evidence of carcinogenicity (IARC, 2014). In addition, DPM is known to cause lung diseases such as pulmonary fibrosis, asthma and chronic obstructive pulmonary disease through an inflammatory response (Reynolds et al., 2011; Ristovski et al., 2012).

However, understanding of DPM's toxicity mechanism remains limited due to complex nature of DPM as a mixture of various chemicals.

* Corresponding author.
*E-mail address:* jinhchoi@uos.ac.kr (J. Choi).

https://doi.org/10.1016/j.envint.2020.106339
Received 24 July 2020; Received in revised form 9 December 2020; Accepted 14 December 2020
Available online 12 January 2021

DPM is a mixture of hundreds of chemicals including various metals and organic compounds such as polycyclic aromatic hydrocarbon (PAH), sulfur, nitrogen, etc. due to its very small size, large surface area and hence easy absorption of chemicals. The chemical composition of DPM is very variable depending on the combustion time and location, and is also influenced by various factors such as the operating condition of the engine and the type of fuel (Ristovski et al., 2012). Since the toxic effects and pathways of DPM can vary depending on the composition of the chemicals, studying the toxicity of DPM itself has been limited in identifying the mechanism of toxicity.

Adverse Outcome Pathway (AOP) is a framework for capturing existing knowledge and can be applied to chemical toxicity screening (Groh et al., 2015). AOP is a toxicological construct that connects mechanistic information to apical endpoints in a formalized way for regulatory purposes (Wittwehr et al., 2017). AOP has several potential applications in hazard identification and risk assessment of chemicals. Among them, prioritization of chemicals for full toxicity testing and screening level hazard identification is currently the most popular application (LaLone et al., 2017). Due to various technical difficulties and excessive time consumption for testing inhalation toxicity testing, AOPs can support a more efficient approach to screening level evaluations of inhaled toxicants, such as DPM.

Pulmonary fibrosis is one of the major health outcomes of inhalation exposure of hazardous substances. In our previous study, we developed AOP on PPARγ inactivation leading to lung fibrosis for inhalation toxicity screening (Jeong et al., 2019). Based on a literature search and compilation of relevant information, PPARγ inactivation was proposed as a molecular initiating event (MIE); activation of TGF-β, inflammation, epithelial mesenchymal transition (EMT), and collagen deposition were proposed as key events (KEs); and pulmonary fibrosis is the adverse outcome (AO). We have submitted this AOP to AOPWiki (https://aopwiki.org/aops/206), which is included in Organization for Economic Cooperation and Development (OECD) program.

In the same study, to validate the AOP, we have identified ToxCast assays relevant to this AOP, and deep learning classification models were subsequently developed that can classify chemical activity based on ToxCast database. Once trained to recognize chemical structures associated with active hit calls in selected ToxCast assays, the models can be used to predict what other structures could be expected to be active. Consequently, the models can help select novel chemicals that may be useful for the experimental validation of AOP206 (Jeong et al., 2019).

As the major exposure route of DPM is through inhalation, in this study, we investigated DPM's toxicity mechanism using the AOP of pulmonary fibrosis, which we previously developed. We therefore conducted a tiered approach to address whether DPM exerts toxicity through the AOP206, first by using the ToxCast database and deep learning models approach for screening the activity of individual chemical components of DPM, followed by experiments on the MIE and KEs of AOP using DPM as a mixture. For identifying the activeness of the component chemicals of DPM, we used 24 ToxCast assays potentially related to the AOP206 and deep learning models based on these assays, which were identified and developed in our previous study (Jeong et al., 2019). To further identify whether individual chemicals in DPM are related to the MIE of the AOP206, we performed molecular docking simulation on PPARγ using the chemicals that showed activeness from ToxCast assay and deep learning model. We then conducted *in vitro* experiments on the MIE and KEs of the AOP206 using a DPM mixture by examining the binding activity of PPARγ and expression of KE proteins.

## 2. Materials and methods

### 2.1. Preparation of PM2.5

DPM is chemical mixture of more than 100 PAHs. The experiments were conducted using DPM NIST 1650b, purchased from Sigma-Aldrich, Inc. (St. Louis, MO, USA). Based on information from the National Institute of Standards & Technology (NIST), individual chemicals in DPM are listed in Table S1.

DPM stock solution (10 mg/mL) was prepared in dimethyl sulfoxide (DMSO) and sonicated for 60 min to avoid agglomeration of the suspended DPM particles. Experiments were performed within 1 h of stock preparation to avoid variability in DPM composition in solution.

### 2.2. ToxCast assays relevant to the AOP of pulmonary fibrosis

As described previously (Jeong et al., 2019), ToxCast assays relevant to the AOP of pulmonary fibrosis were identified. In total, 24 assays were selected: 4 of PPARγ-, 4 of TGF-β-, 3 of NF-κB-, 10 of Inflammation- and 3 of EMT-related targets (Table S2). We used ToxCast & Tox21 INVITRODB_V2 summary files (https://www.epa.gov/chemical-research/exploring-toxcast-data-downloadable-data) to select relevant assays.

### 2.3. Deep learning models

#### 2.3.1. Data preparation

Using the ToxCast *in vitro* assays data, we built 24 artificial neural network models as described previously (Jeong et al., 2019). The canonical simplified molecular-input line-entry system (SMILES) strings describing the structure of the chemicals were collected from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/). The SMILES strings and ToxCast hit call data were then prepared as text files. Because each assay had a different number of experimental data, the number of input data to build the model varied.

#### 2.3.2. Data imbalance

The ToxCast hit call dataset is highly imbalanced, so that in many assays the ratio of inactive data far exceeds that of active data. A bias present between the classes can hinder the performance of the deep learning classification model (Wang et al., 2019), which classifies active and inactive chemicals based on chemical structure information. Therefore, we used the Synthetic Minority Oversampling Technique (SMOTE) to generate synthetic minority data by interpolating between existing minority data and their nearest neighbors to oversample the minor class (Kass and Raftery, 1995).

#### 2.3.3. Multilayer perceptron modeling

For each ToxCast assay, we built 24 multi-layer perceptron (MLP) models. The MLP, a deep neural network, is a widely used deep learning algorithm due to its simplicity (Hamadache et al., 2017). The structure of MLP consists of an input layer, a dense layer with a RELU activator, and a dense single neuron layer as the output with a sigmoidal activator. MLP was implemented by using Python 3.6, with Keras, the Python open source neural network library, running on top of TensorFlow library. The SMILES code was transformed into Morgan Fingerprints with a radius of two bonds using an RDKit (http://www.rdkit.org). Sci-kit Learn toolkit was used to perform the stratified 5-fold cross-validation. In 5-fold cross-validation, the dataset is randomly partitioned into 5 blocks of equal size, and the learning algorithm runs 5 times with each of the blocks being used as the test set. The average of the 5 results gives the test accuracy of the models (Diamantidis et al., 2000).

### 2.4. Docking simulations

#### 2.4.1. Preparation of ligands

The 3D structures of all ligands were collected in .mol2 format from the ZINC database (Irwin et al., 2012). These files could not be directly used for docking simulation; thus they were converted into .pdbqt format using AutoDockTools v1.5.6 (Morris et al., 2009; Sanner, 1999).

#### 2.4.2. Preparation of PPARγ

The coordinates of the X-ray crystal structures of LBDs of the

receptors (PDB ID: 6C5T) were retrieved from the Protein Data Bank (PDB) (Berman et al., 2002). The structures were edited to remove ligands and heteroatoms (HETATM) using Discovery Studio Visualizer v4.5.

#### 2.4.3. Docking simulations

We used AutoDock Vina v1.1 (Trott and Olson, 2010) programs to investigate the binding of ligands to receptors. Required input files for AutoDock Vina were prepared using AutoDockTools v1.5.6 (The Scripps Research Institute, La Jolla, CA, USA). Preparation of files involved changing atom type, removing water molecules, and adding polar hydrogen atoms and Gasteiger charges. The grid box size was kept as 62, 52, and 58 for X, Y, and Z, respectively, and the grid points spacing was 1 Å. The structure files were saved in PDBQT format. Molecular docking analysis was performed using AutoDock Vina v1.1 (The Scripps Research Institute). The exhaustiveness was set to 128 and the maximum number of simultaneous threads was set to 2. The results with best conformation and energetic were selected for analysis. Discovery Studio Visualizer v4.5 (BIOVIA, San Diego, CA, USA) was used for visualization and analysis of the protein–ligand complexes.

### 2.5. PPARγ binding assay

To identify the binding activity of DPM to the *PPARγ* receptor, TransFACTORIAL assay (Romanov et al., 2008) was performed by Attagene (Morrisville, NC, USA) for assessing the agonist/antagonist properties of the compounds. The assay uses HepG2 cells (human liver carcinoma cells) to assess the transfected nuclear receptor activity. The assay was conducted in triplicate at two concentrations (50 ppm and 100 ppm) of DPM after 24 h exposure. The detailed methods are provided in the Supplementary materials.

### 2.6. Cell culture and treatment

BEAS-2B cells (human bronchial epithelial cells) were purchased from the American Type Culture Collection (ATCC; Manassas, USA), cultured in DMEM/F12 (GIBCO, Invitrogen, USA) and supplemented with 10% (v/v) fetal bovine serum and 1% (v/v) antibiotics at 37 °C a 5% $CO_2$ atmosphere. BEAS-2B cells were exposed to 0, 0.1, 1 and 10 ppm of DPM for 24 h. From the stock solutions, experimental concentrations of DPM were obtained by dilution in the cell culture medium.

### 2.7. Western blot

After harvesting, the cell extract was prepared in protein extraction buffer and the protein concentration was measured by Bradford method. Equal amounts (20 mg) of proteins were separated on 12% Mini-PROTEAN@ TGX Gels, Stain-Free Gels (Bio-Rad Laboratories, Inc., CA, USA) and transferred with Trans-Blot Turbo Transfer Pack (Bio-Rad Laboratories, Inc.). The membranes were blocked with 3% bovine serum albumin (BSA) in Tween-20 tris-buffered saline (TTBS) at room temperature. The primary antibody was used at 1:1000 dilutions and the secondary antibody was used at 1:10000 dilutions. Clarity Max Western ECL Substrate (Bio-Rad Laboratories, Inc.) and ChemiDoc XRS$^+$ (Bio-Rad Laboratories, Inc.) were used to detect bands. Blots were developed using an enhanced chemiluminescence western blotting detection kit (Amersham, Little Chalfont, England). The tested proteins were PPARγ (catalog: ab59256, 57 kDa), TGF-β (catalog: ab179695, 44 kDa), TNF-α (catalog: ab9635, 26 kDa), fibronectin (catalog: ab32419, 263 kDa), and β-actin (catalog: sc-47778, 43 kDa). The β-actin antibody was purchased from Santa Cruz Biotechnology, Inc. (CA, USA), and other antibodies were purchased from abcam (England). Three biological replicates were used for each analysis.

### 2.8. Statistics

The statistical significance of differences among treatment groups was determined using the one-way ANOVA test and followed by a post-hoc test (Tukey, $p < 0.05$). All statistical analyses were carried out using SPSS 13.0 (SPSS Inc., Chicago, Il, USA) and graphs were plotted using SigmaPlot 12.0 (Systat Software, Inc., San Jose, CA, USA).

## 3. Results and discussion

Previously, we developed an AOP where antagonism of PPARγ leads to pulmonary fibrosis (Jeong et al., 2019). The MIE is an antagonism of PPARγ, which increases the profibrotic effect of TGF-β/Smad3 signaling (KE1). Then, the TGF-β signaling pathway and oxidative stress pathway lead to increased inflammatory cytokine production (KE2), which in turn drives EMT (KE3) to result in deposition of an interwoven network of collagens (KE4). Increasing amounts of collagen lead to increased tissue stiffness and to tissue damage and scarring or fibrosis, the AO (Fig. S1).

As the major exposure route of DPM is through inhalation, we investigated the general mechanism of toxicity related to the hazardous potential of DPM using AOP206 by examining the activity of individual chemicals in DPM to the MIE and KEs of AOP206 using an *in silico-in vitro* tiered approach. We first conducted *in silico* screening of individual chemical components of DPM using the ToxCast database-deep learning approach and molecular docking, followed by *in vitro* experiments on the MIE and KEs of AOP using DPM as a mixture. The workflow of this study is presented in Fig. 1.

### 3.1. Preparation of ToxCast assays and deep learning models relevant to AOP206

First, we used 24 previously identified assays related to AOP206 from the ToxCast database: 4 PPARγ, 4 TGF-β, 3 NF-κB, 10 Inflammation and 3 EMT (Jeong et al., 2019). The target family of inflammation is cytokines, including IL1α, IL6, and TNF-α. The target family of EMT is protease, including MMP2 and MMP9. A detailed description of each assay is summarized in Table S2.

Next, we used 24 MLP models, also developed in our previous study (Jeong et al., 2019), to classify the activities of each assay related to AOP206. As described previously, the model accuracy obtained using stratified 5-fold cross-validation ranged from 87.16 to 99.76 and the average accuracy was 95.73 (Table S3). In addition to the model accuracy, other metrics that determine whether the model is well developed include the true positive rate (sensitivity), which ranged from 0.91 to 1.00, and the average true negative rate (specificity), which ranged from 0.83 to 1.00.

### 3.2. Identification of individual chemicals of DPM interfering with the AOP related assays

We identified whether individual chemicals in DPM were tested on selected ToxCast assays. For the tested chemicals, active chemicals were identified for each assay, whereas for chemicals not tested, their activity was predicted using the MLP deep learning models. Taking into account the experimental results from the ToxCast database or the predicted results from the deep learning models, the number of active chemicals on the 24 relevant assays was identified (Table 1). In assays related to PPARγ activity, a MIE, a total of 34 chemicals out of 100 were active. This fairly high ratio suggests that the DPM mixture is also likely to be active against PPARγ. Since the proportion of active chemicals was higher in the ToxCast experimental data than in the deep learning predictions, the potential of PPARγ activity was more reliable than the assays showed higher in deep learning prediction than in assay itself (i.e. TGF-β and inflammation). In the case of TGF-β, KE1, 17 chemicals were shown to be active, equivalent to a ratio of one sixth, which is considered

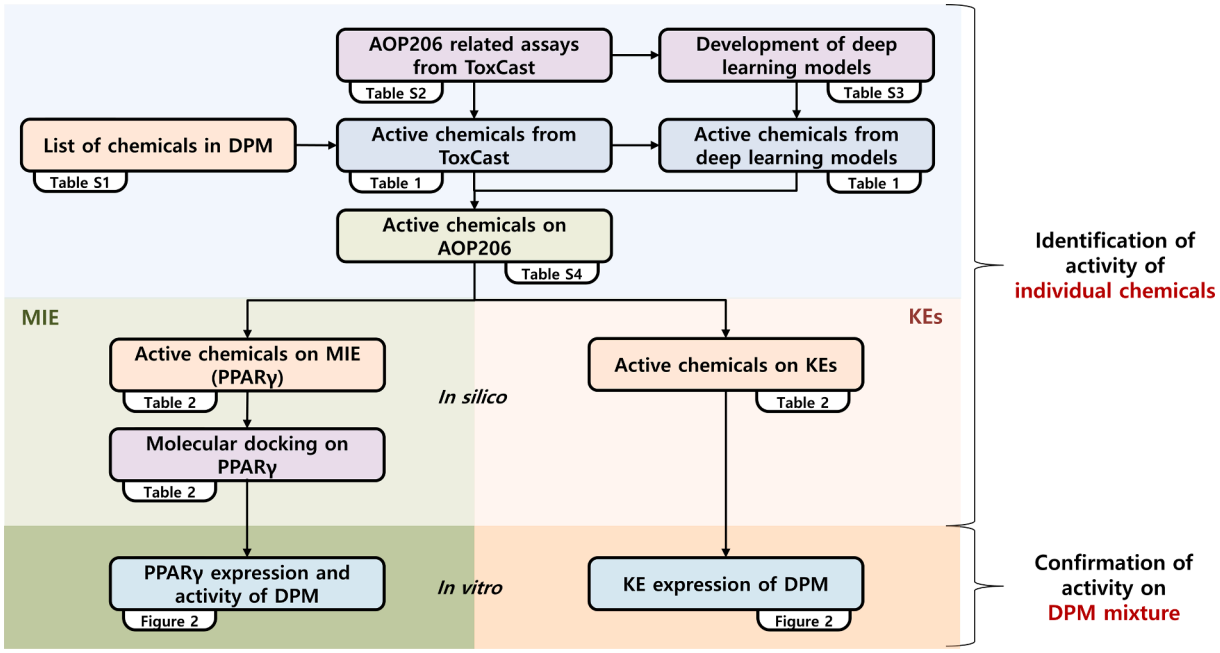


**Fig. 1.** Schematic overview of the workflow of this study.

**Table 1**
Identification of individual chemicals interfering with the AOP of pulmonary fibrosis (AOP206)-related assays.

| AOP206 | | Assay | ToxCast | | Deep learning | | Total |
|---|---|---|---|---|---|---|---|
| | | | Number of chemicals | Number of Active chemicals (%) | Number of chemicals | Number of Active chemicals (%) | Number of Active chemicals (%) |
| MIE | PPARγ | ATG_PPARg_TRANS_up | 19 | 5 (26) | 81 | 13 (16) | 34 (34) |
| | | TOX21_PPARg_BLA_Agonist_ratio | 28 | 5 (18) | 72 | 14 (19) | |
| | | TOX21_PPARg_BLA_antagonist_ratio | 28 | 1 (4) | 72 | 5 (7) | |
| | | NVS_NR_hPPARg | 3 | 0 (0) | 97 | 3 (3) | |
| KE1 | TGF-β | ATG_TGFb_CIS_dn | 19 | 1 (5) | 81 | 0 (0) | 17 (17) |
| | | ATG_TGFb_CIS_up | 19 | 1 (5) | 81 | 1 (1) | |
| | | BSK_BE3C_TGFb1_down | 18 | 0 (0) | 82 | 1 (1) | |
| | | BSK_KF3CT_TGFb1_down | 18 | 1 (6) | 82 | 12 (15) | |
| KE2 | NF-κB | ATG_NF_kB_CIS_dn | 19 | 0 (0) | 81 | 0 (0) | 10 (10) |
| | | ATG_NF_kB_CIS_up | 19 | 0 (0) | 81 | 1 (1) | |
| | | TOX21_NFkB_BLA_agonist_ratio | 28 | 3 (11) | 72 | 6 (8) | |
| | Inflammation | BSK_BE3C_IL1a_down | 18 | 0 (0) | 82 | 11 (13) | 26 (26) |
| | | BSK_BE3C_IL1a_up | 17 | 0 (0) | 83 | 0 (0) | |
| | | BSK_KF3CT_IL1a_down | 18 | 2 (11) | 82 | 16 (20) | |
| | | BSK_KF3CT_IL1a_up | 17 | 0 (0) | 83 | 0 (0) | |
| | | BSK_LPS_IL1a_down | 18 | 0 (0) | 82 | 0 (0) | |
| | | BSK_LPS_IL1a_up | 18 | 0 (0) | 82 | 0 (0) | |
| | | BSK_CASM3C_IL6_down | 18 | 1 (6) | 82 | 1 (1) | |
| | | BSK_CASM3C_IL6_up | 18 | 0 (0) | 82 | 0 (0) | |
| | | BSK_LPS_TNFa_down | 18 | 0 (0) | 82 | 0 (0) | |
| | | BSK_LPS_TNFa_up | 18 | 1 (6) | 82 | 3 (4) | |
| KE3 | EMT | NVS_ENZ_hMMP2 | 0 | 0 (0) | 100 | 0 (0) | 18 (18) |
| | | NVS_ENZ_hMMP9 | 1 | 0 (0) | 99 | 4 (4) | |
| | | BSK_KF3CT_MMP9_down | 18 | 3 (17) | 82 | 11 (13) | |

to be a ratio that can sufficiently affect the toxicity of the mixture. Most chemicals were predicted to be active chemicals through deep learning models. In the case of NF-κB, corresponding to KE2, 10 chemicals were analyzed to be active. Compared with ToxCast, the deep learning models identified more active chemicals. In inflammation, KE2, 26 chemicals were active. Similarly, significant numbers of chemical were predicted to be active in the deep learning models. The number of active chemicals per inflammatory marker was 20 for IL1α, 2 for IL6, and 4 for TNF-α. In EMT, KE3, 18 chemicals were analyzed as active. Due to the lack of ToxCast experimental data, most were predicted to be active chemicals through deep learning models.

### 3.3. Confirmation of the MIE and KEs activity

#### 3.3.1. Validation of the MIE of individual chemicals in DPM using in silico molecular docking

To validate whether DPM could affect the activity of PPARγ, we performed molecular docking simulation on PPARγ. Molecular docking simulation was carried out with the 34 ligands that were active in at least one assay of PPARγ, including ToxCast data and deep learning prediction. For each ligand, among many docking positions, only those with the highest docking score were chosen. Since the affinity data represent the free energy of the coupling in the AutoDock Vina v1.1 docking software, a large absolute value of affinity energy means that

the response of the corresponding ligand receptor has a great affinity.

Of the 34 individual chemicals in DPM activated on PPARγ, 17 were active in one or more KEs, nine were tested in ToxCast, and the remaining eight were predicted using the deep learning models (Table 2; the full list is in Table S4). Molecular docking of ToxCast active chemicals revealed that benzo[e]pyrene and benzo[a]pyrene were most likely to bind to PPARγ, with binding affinities of −10.5 and −10.4 kcal/mol, respectively. For active chemicals derived from deep learning models, dibenzo[a,e]pyrene was most likely to bind to PPARγ with a binding affinity of −12.4 kcal/mol, followed by naphtho[2,3-e]pyrene with binding affinities of −12.0 kcal/mol. 1-Nitronaphthalene from ToxCast had the lowest binding affinity of −7.2 kcal/mol. In general, a binding affinity of lower than −6.0 kcal/mol (Shityakov and Förster, 2014) indicates a capacity to bind to the receptor, so individual chemicals in DPM are more likely to bind to PPARγ.

In the AOP concept, MIE is the only chemical-related event among the components of AOP, as long as the MIE is derived, so it can be connected to the AO along the subsequent KEs (Allen et al., 2016; Villeneuve et al., 2014). It is therefore the most important event to identify which individual chemicals contribute significantly to the toxicity of the mixture. The binding energy calculated by molecular docking can not only confirm the binding to PPARγ, but also the ranking of the binding and toxicity potential of individual chemicals. In addition, ToxCast and deep learning models can be used to provide further evidence of the potential for these individual chemicals to follow the AOP206 pathway through information on the activity of subsequent KEs.

### 3.3.2. *Experiments on the MIE and KEs of DPM mixture using in vitro assays*

Next, *in vitro* experiments were performed using the DPM mixture (Fig. 2). The PPARγ (MIE) protein expression following DPM exposure tended to increase with increasing concentration. In addition, the expression levels of fibronectin, a fibrosis marker (AO), also increased with DPM exposure. However, the expression levels of TGF-β (KE1) and the inflammation marker TNF-α (KE2) showing no statistically significant difference with DPM exposure (Fig. 2A). For further experiments on the activity of the MIE, we evaluated the PPARγ activity of DPM at two concentrations and at 24 h using the *trans*-FACTORIAL assay. As shown in Fig. 2B, DPM stimulated PPARγ in a concentration-dependent manner.

In total, we analyzed the activity of individual chemicals included in DPM on AOP-related assays and identified the potential mechanism of toxicity of DPM through the AOP206. Our study suggests the AOP-based approach seems to be useful for expanding prior knowledge of the potential toxicity pathways of the mixture by identifying the activity of individual chemicals along AOP. In particular, by identifying potential chemicals relevant to the specific AOP, candidate chemicals that contribute primarily to the toxic effects of the mixture can be identified.

However, there are several limitations of this approach. First limitation lies on inconclusive result on MIE. Functional inactivation of PPARγ, an MIE of AOP206, is an event in which an external chemical binds to PPARγ and acts as an antagonist, thereby deactivating the function of PPARγ. In addition to the agonistic and antagonistic effects, the effect of chemicals on PPARγ can also affect the level of PPARγ protein through the upstream effect, and can also exert agonistic and antagonistic effects through physical interaction by binding to PPARγ. It is not easy to distinguish these various interactions or effects using one assay, and therefore, in this study, all ToxCast assays targeting PPARγ were used regardless of agonist or antagonist for screening potential chemicals affecting PPARγ. In the results of *in vitro* experiments, DPM increased the activity and expression of PPARγ. Even if the function of the receptor is inhibited by the antagonist, the expression of the receptor itself may increase. Also, it can be assumed that the role of the antagonist is also partially functional in view of the increased activity of the subsequent KEs. Therefore, further experiments need to be conducted to identify whether the increased expression and activity of PPARγ by DPM mixture that were observed in this study are related to the functional inactivation of the PPARγ receptor.

Another limitation of this study is lack of quantitative aspect of DPM chemicals. This approach is based on the ratio of active chemicals, and therefore could not address quantitative information on the concentrations or mass fraction in DPM. There are 32 of 100 chemicals were tested in 4 PPARγ assays, and only 9 of them showed activity with concentration at 50% of maximum activity or the AC50. Unfortunately, the number of AC50 data is insufficient, making it difficult to create deep learning models that guarantee high accuracy to predict quantitative values. Still, the AC50 of nine chemicals from ToxCast can be compared with the effective concentration from *in vitro* experiments with DPM. The mass fraction of nine chemicals active in the PPARγ assays ranged from 0.0368 to 18.4 mg/kg (Table S1). Using this, the concentration of the chemicals contained in the DPM of 10 ppm, the exposure concentration of this study, is estimated from 0.0018 nM to 0.00074 μM, and this range is much lower compared to the AC50 of the ToxCast assays ranged from 1.06 to 78.72 μM (Table S5). However, since the *in vitro*

**Table 2**
Calculated binding affinity on PPARγ and activity on PPARγ and key events of the AOP206 of individual chemicals in DPM.

| CAS No. | Chemical name | PPARγ binding affinity (kcal/mol) | Number of active assays | | | | |
|---|---|---|---|---|---|---|---|
| | | | PPARγ | TGF-β | NF-κB | Inflammation | EMT |
| *ToxCast* | | | | | | | |
| 192-97-2 | Benzo[e]pyrene | −10.5 | 1 (TOX21_Ago) | 0 | 1 | 6 | 0 |
| 50-32-8 | Benzo[a]pyrene | −10.4 | 1 (TOX21_Ago) | 0 | 0 | 2 | 0 |
| 5522-43-0 | 1-Nitropyrene | −9.9 | 1 (TOX21_antago) | 0 | 0 | 2 | 0 |
| 207-08-9 | Benzo[k]fluoranthene | −9.8 | 1 (TOX21_Ago) | 0 | 0 | 4 | 0 |
| 205-99-2 | Benzo[b]fluoranthene | −9.6 | 2 (ATG, TOX21_Ago) | 0 | 0 | 4 | 0 |
| 56-55-3 | Benz[a]anthracene | −9.6 | 1 (ATG) | 0 | 0 | 2 | 0 |
| 602-60-8 | 9-Nitroanthracene | −8.3 | 2 (ATG, TOX21_Ago) | 0 | 0 | 2 | 0 |
| 602-87-9 | 5-Nitroacenaphthene | −8.2 | 1 (ATG) | 0 | 1 | 7 | 0 |
| 86-57-7 | 1-Nitronaphthalene | −7.2 | 1 (ATG) | 1 | 1 | 4 | 0 |
| *Deep learning* | | | | | | | |
| 192-65-4 | Dibenzo[a,e]pyrene | −12.4 | 1 (TOX21_Ago) | 0 | 1 | 0 | 0 |
| 193-09-9 | Naphtho[2,3-e]pyrene | −12.0 | 1 (TOX21_Ago) | 0 | 1 | 0 | 0 |
| 63041-90-7 | 6-Nitrobenzo[a]pyrene | −10.7 | 1 (TOX21_antago) | 0 | 1 | 0 | 0 |
| 84030-79-5 | Dibenzo[a,k]fluoranthene | −10.4 | 1 (TOX21_Ago) | 0 | 1 | 0 | 0 |
| 57835-92-4 | 4-Nitropyrene | −9.9 | 1 (TOX21_Ago) | 0 | 1 | 0 | 0 |
| 75321-20-9 | 1,3-Dinitropyrene | −9.8 | 2 (ATG, TOX21_antago) | 0 | 1 | 0 | 0 |
| 82064-15-1 | 4-Nitrophenanthrene | −8.7 | 1 (TOX21_antago) | 0 | 1 | 0 | 0 |
| 892-21-7 | 3-Nitrofluoranthene | −8.5 | 2 (ATG, TOX21_antago) | 0 | 2 | 0 | 0 |

ATG, ATG_PPARg_TRANS_up; NVS, NVS_NR_hPPARg; TOX21_Ago, TOX21_PPARg_BLA_Agonist_ratio; TOX21_antago, TOX21_PPARg_BLA_antagonist_ratio.

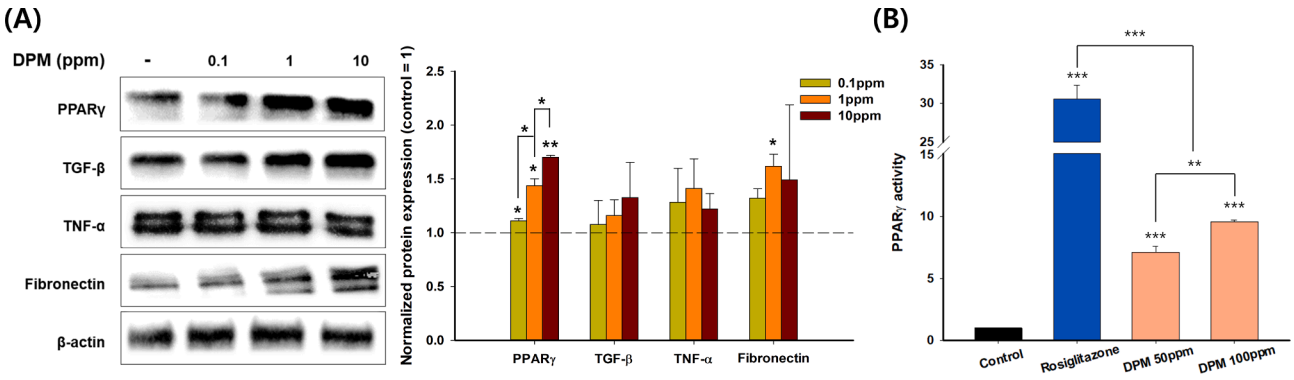


**Fig. 2.** Expressions of MIE and KEs proteins (A) and activity of PPARγ (B) on DPM exposure. (A) Beas2B cells were exposed to 0, 0.1, 1 and 10 ppm of DPM for 24 h. (B) HepG2 cells that transiently transfected with the optimized nuclear receptor library treated with DPM for 24 h. The profile of the PPARγ activities was determined as fold of induction values versus vehicle-treated (DMSO) control. Rosiglitazone, a PPARγ agonist, was used as a positive control. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$ in one-way ANOVA test and followed by a post-hoc test (Tukey, $p < 0.05$).

experimental condition of this study is different from the ToxCast assays, direct comparison is not desirable. As can be seen from these calculations, this approach only confirms the activity of individual chemicals and therefore does not reflect the combined effects that may occur when chemicals are mixed. In addition, since it does not reflect the toxic effect of the entire DPM mixture, including the physical effect as particles, there is a limitation in fully predicting the toxic pathway of DPM through the activity of individual chemicals. Due to these limitations, the results from the *in vitro* experiments showed that TNF-α and TGF-β expressions did not match the predicted results of individual chemical. Therefore, identifying the quantitative contribution of each individual chemical and the potential combined effects will require further research to develop a model capable of predicting the quantitative effects for each individual chemical and their mixture.

Nevertheless, our results from *in vitro* experiments using the DPM mixture are consistent with analyses of the assay activity of individual chemicals, suggesting that the proposed approach using the ToxCast database and deep learning models could potentially identify the toxicity pathway of the DPM. However, to determine which individual chemical activity affects the toxicity of the DPM mixture, it is necessary to conduct *in vitro* experiments on the individual chemicals prioritized in ToxCast and deep learning prediction.

## 4. Conclusion

In this study, we have demonstrated that the AOP of PPARγ inactivation leading to pulmonary fibrosis can be applied to identify the toxicity pathway of DPM. Our proposed approach using ToxCast and deep learning models may be used to identify potential chemicals from the mixtures, such as DPM, by determining the activity of individual chemicals. Our study also suggests that *in silico* assays combined with *in vitro* experiments is a useful and efficient tool for identifying the toxicity pathway of chemical mixtures.

## Funding

This work was supported by the Korean Ministry of Environment under the 'Environmental Health R&D Program' (2017001370001).

## CRediT authorship contribution statement

**Jaeseong Jeong:** Writing - original draft, Methodology, Software, Formal analysis, Writing - review & editing. **Su-yong Bae:** Writing - original draft, Software, Data curation, Investigation. **Jinhee Choi:** Supervision, Conceptualization, 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.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2020.106339.

## References

Achilleos, S., Kioumourtzoglou, M.A., Wu, C.Da, Schwartz, J.D., Koutrakis, P., Papatheodorou, S.I., 2017. Acute effects of fine particulate matter constituents on mortality: A systematic review and meta-regression analysis. Environ. Int. 109, 89–100. https://doi.org/10.1016/j.envint.2017.09.010.

Allen, T.E.H., Goodman, J.M., Gutsell, S., Russell, P.J., 2016. A history of the molecular initiating event. Chem. Res. Toxicol. 29, 2060–2070. https://doi.org/10.1021/acs.chemrestox.6b00341.

Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Bourne, P.E., Burkhardt, K., Feng, Z., Gilliland, G.L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J.D., Zardecki, C., 2002. The Protein Data Bank. Acta Crystallogr. Sect. D Biol. Crystallogr. 58, 899–907. https://doi.org/10.1107/S0907444902003451.

Diamantidis, N.A., Karlis, D., Giakoumakis, E.A., 2000. Unsupervised stratification of cross-validation for accuracy estimation. Artif. Intell. 116, 1–16. https://doi.org/10.1016/S0004-3702(99)00094-6.

Groh, K.J., Carvalho, R.N., Chipman, J.K., Denslow, N.D., Halder, M., Murphy, C.A., Roelofs, D., Rolaki, A., Schirmer, K., Watanabe, K.H., 2015. Development and application of the adverse outcome pathway framework for understanding and predicting chronic toxicity: I. Challenges and research needs in ecotoxicology. Chemosphere 120, 764–777. https://doi.org/10.1016/j.chemosphere.2014.09.068.

Hamadache, M., Benkortbi, O., Hanini, S., Amrane, A., 2017. Application of multilayer perceptron for prediction of the rat acute toxicity of insecticides. Energy Procedia 139, 37–42. https://doi.org/10.1016/j.egypro.2017.11.169.

He, M., Ichinose, T., Yoshida, S., Ito, T., 2017. PM2. 5-induced lung inflammation in mice: Differences of inflammatory response in macrophages and type II alveolar cells. J. Appl. Toxicol. 1203–1218. https://doi.org/10.1002/jat.3482.

IARC, I.A. for R. on C., 2014. Diesel and gasoline engine exhausts and some nitroarenes.

Irwin, J.J., Sterling, T., Mysinger, M.M., Bolstad, E.S., Coleman, R.G., 2012. ZINC: A free tool to discover chemistry for biology. J. Chem. Inf. Model. 52, 1757–1768. https://doi.org/10.1021/ci3001277.

Jeong, J., Garcia-Reyero, N., Burgoon, L.D., Perkins, E.J., Park, T., Kim, C., Roh, J.-Y., Choi, J., 2019. Development of Adverse Outcome Pathway for PPARγ antagonism leading to pulmonary fibrosis and chemical selection for its validation: ToxCastTM database and a deep learning artificial neural network model based approach. Chem. Res. Toxicol. 32, 1212–1222. https://doi.org/10.1021/acs.chemrestox.9b00040.

Kass, R.E., Raftery, A.E., 1995. Bayes Factors. J. Am. Stat. Assoc. 90, 773–795. https://doi.org/10.2307/2291091.

LaLone, C.A., Ankley, G.T., Belanger, S.E., Embry, M.R., Hodges, G., Knapen, D., Munn, S., Perkins, E.J., Rudd, M.A., Villeneuve, D.L., Whelan, M., Willett, C., Zhang, X., Hecker, M., 2017. Advancing the adverse outcome pathway framework—An international horizon scanning approach. Environ. Toxicol. Chem. 36, 1411–1421. https://doi.org/10.1002/etc.3805.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: Automated docking with

selective receptor flexibility. J. Comput. Chem. 30, 2785–2791. https://doi.org/10.1002/jcc.21256.
Park, M., Joo, H.S., Lee, K., Jang, M., Kim, S.D., Kim, I., Borlaza, L.J.S., Lim, H., Shin, H., Chung, K.H., Choi, Y., Park, S.G., Bae, M., Lee, J., Song, H., Park, K., 2018. Differential toxicities of fine particulate matters from various sources. Sci. Rep. 8, 1–11. https://doi.org/10.1038/s41598-018-35398-0.
Piao, J.M., Ahn, J., Ah, K., Yea, K., Ryu, S., Jae, Y., Kristina, H., 2018. Particulate matter 2. 5 damages skin cells by inducing oxidative stress, subcellular organelle dysfunction, and apoptosis. Arch. Toxicol. 92, 2077–2091. https://doi.org/10.1007/s00204-018-2197-9.
Reynolds, P.R., Wasley, K.M., Allison, C.H., 2011. Diesel particulate matter induces receptor for advanced glycation end-products (RAGE) expression in pulmonary epithelial cells, and RAGE signaling influences NF-κB-mediated inflammation. Environ. Health Perspect. 119, 332–336. https://doi.org/10.1289/ehp.1002520.
Ristovski, Z.D., Miljevic, B., Surawski, N.C., Morawska, L., Fong, K.M., Goh, F., Yang, I. A., 2012. Respiratory health effects of diesel particulate matter. Respirology 17, 201–212. https://doi.org/10.1111/j.1440-1843.2011.02109.x.
Robinson, A.L., Grieshop, A.P., Donahue, N.M., Hunt, S.W., Robinson, A.L., Grieshop, A. P., Donahue, N.M., Hunt, S.W., Robinson, A.L., Grieshop, A.P., Donahue, N.M., Hunt, S.W., 2010. Updating the Conceptual Model for Fine Particle Mass Emissions from Combustion Systems Allen L . Robinson Updating the Conceptual Model for Fine Particle Mass Emissions from Combustion Systems. J. Air Waste Manage. Assoc. 60, 1204–1222. https://doi.org/10.3155/1047-3289.60.10.1204.
Romanov, S., Medvedev, A., Gambarian, Maria, Poltoratskaya, N., Moeser, M., Medvedeva, L., Gambarian, Mikhail, Diatchenko, L., Makarov, S., 2008. Homogeneous reporter system enables quantitative functional assessment of multiple transcription factors. Nat. Methods 5, 253–260. https://doi.org/10.1038/nmeth.1186.
Sanner, M.F., 1999. Python: a programming language for software integration and development. J. Mol. Graph. Model. 17, 57–61.
Schlesinger, R.B., 2007. The health impact of common inorganic components of fine particulate matter (PM2.5) in ambient air: A critical review. Inhal. Toxicol. 19, 811–832. https://doi.org/10.1080/08958370701402382.
Shityakov, S., Förster, C., 2014. In silico predictive model to determine vector-mediated transport properties for the blood-brain barrier choline transporter. Adv. Appl. Bioinforma. Chem. 7, 23–36. https://doi.org/10.2147/AABC.S63749.
Song, Y., Li, R., Zhang, Y., Wei, J., Chen, W., Kong, C., Chung, A., Cai, Z., 2019. Mass spectrometry-based metabolomics reveals the mechanism of ambient fi ne particulate matter and its components on energy metabolic reprogramming in BEAS-2B cells. Sci. Total Environ. 651, 3139–3150. https://doi.org/10.1016/j.scitotenv.2018.10.171.
Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461. https://doi.org/10.1002/jcc.21334.
Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C. A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L., Whelan, M., 2014. Adverse outcome pathway (AOP) development I: Strategies and principles. Toxicol. Sci. 142, 312–320. https://doi.org/10.1093/toxsci/kfu199.
Wang, G., Zhang, X., Liu, X., Zheng, J., Chen, R., Kan, H., 2019. Ambient fine particulate matter induce toxicity in lung epithelial- endothelial co-culture models. Toxicol. Lett. 301, 133–145. https://doi.org/10.1016/j.toxlet.2018.11.010.
Wang, H., Liu, R., Schyman, P., Wallqvist, A., 2019. Deep neural network models for predicting chemically induced liver toxicity endpoints from transcriptomic responses. Front. Pharmacol. 10, 1–12. https://doi.org/10.3389/fphar.2019.00042.
Wang, Y., Zou, L., Wu, T., Xiong, L., Zhang, T., Kong, L., 2019. Identification of mRNA-miRNA crosstalk in human endothelial cells after exposure of PM2. 5 through integrative transcriptome analysis. Ecotoxicol. Environ. Saf. 169, 863–873. https://doi.org/10.1016/j.ecoenv.2018.11.114.
Wittwehr, C., Aladjov, H., Ankley, G., Byrne, H.J., de Knecht, J., Heinzle, E., Klambauer, G., Landesmann, B., Luijten, M., MacKay, C., Maxwell, G., Meek, M.E.B., Paini, A., Perkins, E., Sobanski, T., Villeneuve, D., Waters, K.M., Whelan, M., 2017. How adverse outcome pathways can aid the development and use of computational prediction models for regulatory toxicology. Toxicol. Sci. 155, 326–336. https://doi.org/10.1093/toxsci/kfw207.
Zheng, Y., Lan, H., Thiruvengadam, M., Tien, J.C., Li, Y., 2017. Effect of single dead end entry inclination on DPM plume dispersion. Int. J. Min. Sci. Technol. 27, 401–406. https://doi.org/10.1016/j.ijmst.2017.03.003.