



Identification of potential descriptors of water-soluble fullerene derivatives responsible for antitumor effects on lung cancer cells via QSAR analysis



Hung-Jin Huang^a, Margarita Chetyrkina^b, Chui-Wei Wong^a, Olga A. Kraevaya^{b,c}, Alexander V. Zhilenkov^c, Ilya I. Voronov^c, Pei-Hwa Wang^d, Pavel A. Troshin^{c,*}, Shan-hui Hsu^{a,e,f,*}

^a Institute of Polymer Science and Engineering, National Taiwan University, Taipei, Taiwan, ROC

^b Skolkovo Institute of Science and Technology, Moscow, Russian Federation

^c Institute for Problems of Chemical Physics of Russian Academy of Sciences, Chernogolovka, Russian Federation

^d Department of Animal Science and Technology, National Taiwan University, Taipei, Taiwan, ROC

^e Institute of Cellular and System Medicine, National Health Research Institutes, Miaoli, Taiwan, ROC

^f Research and Development Center for Medical Devices, National Taiwan University, Taipei, Taiwan, ROC

ARTICLE INFO

Article history:

Received 20 October 2020

Received in revised form 4 January 2021

Accepted 8 January 2021

Available online 16 January 2021

Keywords:

NSCLC

Genetic algorithm

QSAR

Antitumor

Water-soluble fullerene derivatives

ABSTRACT

Water-soluble fullerene derivatives are actively investigated as potential drugs for cancer treatment due to their favorable membranotropic properties. Herein, cytotoxic effects of twenty fullerene derivatives with different solubilizing addends were evaluated in three different types of non-small-cell lung carcinoma (NSCLC). The potential structural descriptors of the solubilizing addends related to the inhibitory activities on each type of lung cancer cell were investigated by the quantitative structure–activity relationship (QSAR) approach. The determination coefficient r^2 for the recommended QSAR model were 0.9325, 0.8404, and 0.9011 for A549, H460, and H1299 cell lines, respectively. The results revealed that the chemical features of the fullerene-based compounds including aromatic bonds, sulfur-containing aromatic rings, and oxygen atoms are favored properties and promote the inhibitory effects on H460 and H1299 cells. Particularly, thiophene moiety is the key functional group, which was positively correlated with strong inhibitory effects on the three types of lung cancer cells. The useful information obtained from our regression models may lead to the design of more efficient inhibitors of the three types of NSCLC.

© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Human lung cancer is the most common leading cause of cancer-associated deaths for both men and women [1,2]. Approximately 18.4% of all deaths caused by cancer were associated with lung cancer, and the number of deaths from lung cancer is more than that of colon, prostate, and breast cancers [3]. Most cases of lung cancer can be divided into two categories, about 13% of lung cancers being small-cell lung carcinoma (SCLC) and 84% belonging to non-small cell lung cancer (NSCLC) [4]. The survival rate of patients with lung cancer is still pretty low despite the advances in drug therapy and

diagnosis [5,6]. The common treatment strategies of lung cancer include chemotherapy [7], radiation therapy [8], systemic therapy [9], targeted therapy [10], and immunotherapy [11]. For the current cancer treatments, there is no effective therapeutic strategy for most patients with metastatic lung cancer. Metastatic lung cancer occurs when the originated lung tumors spread and travel to other parts of tissues via the blood and lymphatic system. The typically invaded areas of the human body include adrenal glands [12], liver [13], brain [14], bone [15], and lymph nodes. [16]. To date, lung transplantation is one of the accepted treatments to prolong survival of the end-stage patients with lung disease [17]. According to a recent report, pulmonary drug delivery offers an opportunity for the drug to be delivered directly to the lung [18]. The nanoparticle drug delivery systems [19] have potential effects that can be applied to improve the antitumor treatment [20–23]. For instance, the nanoparticle-based platform can serve as a highly efficient way of delivering drugs for lung cancer therapy [24–26].

* Corresponding authors at: Institute of Polymer Science and Engineering, National Taiwan University, No. 1, Sec. 4 Roosevelt Road, Taipei 10617, Taiwan, ROC (Shan-hui Hsu), Institute of Problems of Chemical Physics of Russian Academy of Sciences, Chernogolovka 142432, Russian Federation (Pavel A. Troshin).

E-mail addresses: troshin2003@inbox.ru (P.A. Troshin), shhsu@ntu.edu.tw (S.-h. Hsu).

Epithelial-mesenchymal transition (EMT) is a process that promotes the immotile epithelial cells to lose cell–cell adhesion and converts epithelial cells into mesenchymal cells during cancer progression and metastasis [27]. The mesenchymal cells have invasive properties that can migrate from the originated epithelial layer to other parts of tissues through blood vessels [28]. There are many different types of lung cancer cell lines, such as A549, H460, and H1299. According to previously published results, H1299 cell line can be characterized as mesenchymal cells [29]. The H1299 cell line represents a heterogeneous and aggressive form that contributes to the metastatic property and drug-resistivity of the cells due to the expression of CD44 [30]. Regarding the invasive potential of NSCLC cancer cell lines, literature has reported that H1299 is more aggressive than A549 in the result of cell invasion assay [31]. The A549 cells have more epithelial characteristics than H1299 cells, while the H1299 cells have more mesenchymal characteristics and less effective epithelial marker expression [32]. Meanwhile, H460 exhibits EMT characteristics and contains both epithelial and mesenchymal markers to promote cell elongation and other morphological changes [33]. For the balance between epithelial and mesenchymal properties, H460 showed weak mesenchymal features, unlike H1299 that displayed increased mesenchymal phenotype [34]. The above information reveals that the H1299 cell line is the most aggressive and most resistant to drug therapy.

Fullerenes and their derivatives are promising nanomaterials with a wide range of biomedical applications. For example, fullerene C₆₀ and C₇₀ derivatives can act as anti-HIV drugs [35], cancer cell inhibitors [36], DNA photocleavage [37], superconducting materials [38], radionuclide chelating agents [39], drug delivery carriers [40–42], and strong antioxidants [43–46]. The membrane transportation of fullerenes and their derivatives has been actively investigated during the past decade, and distinct penetration through lipid bilayers of cell membranes was revealed [47–50]. Fullerenes with hydrophobic surface can easily penetrate into the lipid membrane because their translocation kinetics are similar to those of dioleoylphosphatidylcholine (DOPC) and dipalmitoylphosphatidylcholine (DPPC) in the lipid bilayer. Because fullerenes are extremely insoluble in water, functionalization of fullerenes with solubilizing functional groups helps overcome the limitation for drug development [51]. Meanwhile, the mechanism of nanoparticle distribution is highly important for the development of less toxic drug delivery systems based on fullerenes.

The quantitative structure–activity relationship (QSAR) modeling is an efficient tool for rapid identification of correlations between chemical structures of compounds and their biological activities based on statistical analysis and generation of classification models. The QSAR approach can be used to identify potential chemical or physical properties among the various types of synthetic compounds for leading drug discovery and optimization as well. In this study, a series of twenty different water-soluble fullerene derivatives were synthesized and their cytotoxicity was evaluated for A549, H460, and H1299 lung cancer cells. Obtained experimental data was used to generate QSAR models and reveal the influence of various solubilizing addends attached to the fullerene cage on cytotoxicity of the fullerene derivatives for various cancer cell lines. The information obtained using predictive computer models might guide the design of water-soluble fullerene derivatives toxic to various types of lung cancer cells.

2. Materials and methods

2.1. Synthesis of the Water-Soluble fullerene derivatives

Compounds **1–6** have been synthesized using the reaction of chlorofullerene C₆₀Cl₆ with *tert*-butyl esters of amino acids fol-

lowed by hydrolysis of *tert*-butyl protective groups using method reported earlier by our group [52]. Compounds **1–6** have been characterized in the form of *tert*-butyl esters. Spectral data for the new compounds are given below.

Compound **1**. ¹H NMR (500 MHz, CDCl₃, δ, ppm): 0.90–1.14 (m, 30H, (CH₃)₂CH), 1.21–1.53 (m, 45H, (CH₃)₃C), 1.58–1.74 (m, 10H, CH₂), 1.85–2.46 (m, 5H, (CH₃)₂CH), 3.21–4.79 (m, 10H, CH, NH).

¹³C NMR (126 MHz, CDCl₃, δ, ppm): 22.02 (CH₃)₂CH, 22.33 (CH₃)₂CH, 22.53 (CH₃)₂CH, 22.78 (CH₃)₂CH, 22.86 (CH₃)₂CH, 22.90 (CH₃)₂CH, 22.96 (CH₃)₂CH, 23.06 (CH₃)₂CH, 23.12 (CH₃)₂CH, 23.37 (CH₃)₂CH, 24.57 (CH₃)₂CH, 24.84 (CH₃)₂CH, 24.92 (CH₃)₂CH, 25.04 (CH₃)₂CH, 25.19 (CH₃)₂CH, 27.89 (CH₃)₃C, 27.94 (CH₃)₃C, 28.00 (CH₃)₃C, 28.05 (CH₃)₃C, 43.65 (CH₂), 43.83 (CH₂), 44.08 (CH₂), 44.17 (CH₂), 44.38 (CH₂), 57.32 (NHCH), 57.36 (NHCH), 57.62 (NHCH), 57.89 (NHCH), 58.39 (NHCH), 63.86 (C_{sp3}, cage), 64.07 (C_{sp3}, cage), 65.14 (C_{sp3}, cage), 67.29 (C_{sp3}, cage), 68.21 (C_{sp3}, cage), 80.69 (CH₃)₃C, 80.87 (CH₃)₃C, 80.91 (CH₃)₃C, 81.05 (CH₃)₃C, 81.25 (CH₃)₃C, 141.26, 141.73, 142.03, 142.27, 142.45, 142.79, 142.91, 143.37, 143.61, 143.65, 143.68, 143.73, 143.95, 144.02, 144.09, 144.13, 144.16, 144.29, 144.58, 144.92, 145.40, 145.46, 145.70, 146.28, 146.51, 146.83, 146.94, 147.04, 147.06, 147.14, 147.19, 147.26, 147.41, 147.43, 147.99, 148.29, 148.37, 148.42, 148.52, 148.56, 148.60, 148.74, 148.76, 148.79, 149.08, 149.22, 149.50, 151.12, 151.68, 152.83, 154.39, 154.67, 154.99, 155.31, 176.22 (COO), 177.01 (COO), 177.07 (COO), 177.16 (COO), 177.38 (COO).

Compound **2**. ¹H NMR (500 MHz, CDCl₃, δ, ppm): 1.06–1.29 (m, 30H, (CH₃)₂CH), 1.34–1.53 (m, 45H, (CH₃)₃C), 2.11–2.46 (m, 5H, (CH₃)₂CH), 3.37 (d, *J* = 11.7 Hz, 1H, NH), 3.46 (d, *J* = 12.0 Hz, 1H, NH), 3.55 (d, *J* = 11.6 Hz, 1H, NH), 3.64 (d, *J* = 12.3 Hz, 1H, NH), 3.71 (d, *J* = 11.6 Hz, 1H, NH), 4.44–4.56 (m, 3H, C_{sp3}, cage-H, CH), 4.73 (dd, *J* = 11.6, 4.9 Hz, 1H, CH), 4.80 (dd, *J* = 12.0, 5.0 Hz, 1H, CH), 4.87 (dd, *J* = 11.9, 5.4 Hz, 1H, CH).

¹³C NMR (126 MHz, CDCl₃, δ, ppm): 17.81 ((CH₃)₂CH), 17.94 ((CH₃)₂CH), 18.07 ((CH₃)₂CH), 18.36 ((CH₃)₂CH), 18.56 ((CH₃)₂CH), 19.60 ((CH₃)₂CH), 19.82 ((CH₃)₂CH), 19.87 ((CH₃)₂CH), 20.02 ((CH₃)₂CH), 27.99 ((CH₃)₃C), 28.02 ((CH₃)₃C), 28.10 ((CH₃)₃C), 28.24 ((CH₃)₃C), 32.16 ((CH₃)₂CH), 32.31 ((CH₃)₂CH), 32.42 ((CH₃)₂CH), 32.57 ((CH₃)₂CH), 60.63 (C_{sp3}, cage-H), 63.19 (NHCH), 63.62 (NHCH), 64.33 (NHCH), 64.42 (NHCH), 64.46 (NHCH), 64.47 (C_{sp3}, cage), 64.90 (C_{sp3}, cage), 66.20 (C_{sp3}, cage), 66.57 (C_{sp3}, cage), 67.49 (C_{sp3}, cage), 80.94 ((CH₃)₃C), 81.02 ((CH₃)₃C), 81.07 ((CH₃)₃C), 81.19 ((CH₃)₃C), 81.23 ((CH₃)₃C), 142.82, 142.85, 142.90, 143.19, 143.32, 143.45, 143.47, 143.68, 143.74, 143.75, 144.00, 144.06, 144.28, 144.30, 144.33, 144.67, 144.81, 144.88, 145.07, 145.44, 145.50, 145.52, 145.58, 146.14, 146.93, 147.00, 147.09, 147.14, 147.17, 147.21, 147.27, 147.36, 147.76, 148.00, 148.08, 148.10, 148.17, 148.26, 148.38, 148.43, 148.46, 148.49, 148.56, 148.70, 148.77, 148.82, 150.95, 151.52, 151.83, 152.12, 153.17, 154.66, 154.68, 155.20, 175.58 (COO), 175.59 (COO), 175.79 (COO), 175.88 (COO), 176.15 (COO).

Compound **3**. ¹H NMR (500 MHz, CDCl₃, δ, ppm): 0.96–1.21 (m, 30H, CH₃CH₂, CH₃CH), 1.26–1.57 (m, 45H, (CH₃)₃C), 1.57–2.18 (m, 15H, CH₃CH₂, CH₃CH), 3.35 (d, *J* = 11.8 Hz, 1H, NH), 3.45 (d, *J* = 12.0 Hz, 1H, NH), 3.54 (d, *J* = 12.2 Hz, 1H, NH), 3.64 (d, *J* = 12.3 Hz, 1H, NH), 3.70 (d, *J* = 11.7 Hz, 1H, NH), 4.41–4.47 (m, 2H, CH), 4.49 (s, 1H, C_{sp3}, cage-H), 4.67 (dd, *J* = 11.7, 5.4 Hz, 1H,

CH), 4.76 (dd, $J = 11.8, 5.4$ Hz, 1H, CH), 4.80 (dd, $J = 12.2, 6.2$ Hz, 1H, CH).

^{13}C NMR (126 MHz, CDCl_3 , δ , ppm): 11.76 (CH₃CH₂), 11.95 (CH₃CH₂), 12.08 (CH₃CH₂), 12.21 (CH₃CH₂), 12.28 (CH₃CH₂), 15.01 (CH₃CH), 15.65 (CH₃CH), 15.79 (CH₃CH), 16.12 (CH₃CH), 16.23 (CH₃CH), 24.99 (CH₂), 25.32 (CH₂), 25.41 (CH₂), 25.49 (CH₂), 25.70 (CH₂), 27.98 ((CH₃)₃C), 28.02 ((CH₃)₃C), 39.18 (CH₃CH), 39.44 (CH₃CH), 39.47 (CH₃CH), 39.58 (CH₃CH), 39.87 (CH₃CH), 60.57 (C_{sp3, cage}-H), 63.12 (NHCH), 63.59 (NHCH), 64.22 (NHCH), 64.35 (NHCH), 64.37 (NHCH), 64.39 (C_{sp3, cage}), 64.90 (C_{sp3, cage}), 66.20 (C_{sp3, cage}), 66.58 (C_{sp3, cage}), 67.47 (C_{sp3, cage}), 80.91 ((CH₃)₃C), 81.03 ((CH₃)₃C), 81.07 ((CH₃)₃C), 81.18 ((CH₃)₃C), 81.23 ((CH₃)₃C), 141.31, 141.79, 142.83, 142.93, 143.17, 143.31, 143.32, 143.47, 143.49, 143.69, 143.70, 143.76, 144.03, 144.07, 144.12, 144.30, 144.32, 144.43, 144.59, 144.79, 144.89, 144.95, 145.03, 145.18, 145.39, 145.54, 145.62, 145.71, 146.17, 146.89, 146.94, 147.01, 147.11, 147.16, 147.28, 147.82, 148.02, 148.10, 148.17, 148.26, 148.37, 148.49, 148.56, 148.71, 148.78, 149.14, 151.01, 151.46, 151.76, 152.20, 153.22, 154.68, 155.28, 175.64 (COO), 175.70 (COO), 175.96 (COO), 175.98 (COO), 176.28 (COO).

Compound 4. ^1H NMR (500 MHz, CDCl_3 , δ , ppm): 0.86–1.20 (m, 15H, CH₃CH₂), 1.25–1.50 (m, 45H, (CH₃)₃C), 1.70–2.12 (m, 10H, CH₃CH₂), 3.07–3.87 (m, 5H, NH), 4.07–5.00 (m, 5H, CH).

^{13}C NMR (126 MHz, CDCl_3 , δ , ppm): 9.97 (CH₃CH₂), 10.26 (CH₃CH₂), 10.44 (CH₃CH₂), 10.56 (CH₃CH₂), 10.59 (CH₃CH₂), 27.66 (CH₃CH₂), 27.83 ((CH₃)₃C), 27.96 ((CH₃)₃C), 27.99 ((CH₃)₃C), 28.03 ((CH₃)₃C), 28.08 ((CH₃)₃C), 59.45 (NHCH), 59.65 (NHCH), 60.52 (NHCH), 61.46 (NHCH), 61.89 (NHCH), 64.18 (C_{sp3, cage}), 64.42 (C_{sp3, cage}), 65.65 (C_{sp3, cage}), 67.33 (C_{sp3, cage}), 68.53 (C_{sp3, cage}), 77.13 (C-Cl), 80.61 ((CH₃)₃C), 80.87 ((CH₃)₃C), 80.90 ((CH₃)₃C), 81.00 ((CH₃)₃C), 81.22 ((CH₃)₃C), 140.78, 141.87, 142.38, 142.53, 142.62, 142.67, 142.78, 143.29, 143.61, 143.68, 143.72, 144.00, 144.04, 144.07, 144.09, 144.20, 144.30, 144.34, 144.57, 145.20, 145.43, 145.49, 146.45, 147.15, 147.20, 147.32, 147.34, 147.41, 147.97, 148.23, 148.32, 148.42, 148.50, 148.56, 148.71, 148.77, 148.79, 149.67, 150.56, 151.03, 151.69, 154.43, 154.72, 155.11, 175.77 (COO), 176.02 (COO), 176.21 (COO), 176.29 (COO), 176.45 (COO).

Compound 5. ^1H NMR (500 MHz, CDCl_3 , δ , ppm): 0.80–1.16 (m, 15H, CH₃CH₂), 1.23–1.53 (m, 45H, (CH₃)₃C), 1.54–1.98 (m, 20H, CH₂CH₂), 3.15–3.73 (m, 5H, NH), 4.30–5.09 (m, 5H, CH).

^{13}C NMR (126 MHz, CDCl_3 , δ , ppm): 14.15 (CH₃CH₂), 14.16 (CH₃CH₂), 14.22 (CH₃CH₂), 14.33 (CH₃CH₂), 14.49 (CH₃CH₂), 18.82 (CH₃CH₂), 19.04 (CH₃CH₂), 19.29 (CH₃CH₂), 19.37 (CH₃CH₂), 27.97 ((CH₃)₃C), 28.04 ((CH₃)₃C), 28.05 ((CH₃)₃C), 36.65 (CH₂CH₂), 36.81 (CH₂CH₂), 37.11 (CH₂CH₂), 37.40 (CH₂CH₂), 58.01 (NHCH), 58.72 (NHCH), 59.32 (NHCH), 60.33 (NHCH), 60.62 (NHCH), 64.08 (C_{sp3, cage}), 64.28 (C_{sp3, cage}), 65.44 (C_{sp3, cage}), 67.40 (C_{sp3, cage}), 68.45 (C_{sp3, cage}), 80.59 ((CH₃)₃C), 80.81 ((CH₃)₃C), 80.82 ((CH₃)₃C), 80.98 ((CH₃)₃C), 81.24 ((CH₃)₃C), 140.93, 141.93, 142.19, 142.42, 142.59, 142.67, 142.83, 143.30, 143.58, 143.60, 143.69, 144.00, 144.04, 144.16, 144.18, 144.27, 144.31, 144.40, 144.56, 145.12, 145.34, 145.52, 146.56, 147.17, 147.20, 147.32, 147.36, 147.41, 147.97, 148.23, 148.29, 148.32, 148.41, 148.51, 148.58, 148.72, 148.76, 148.79, 149.54, 150.72, 151.10, 151.61, 154.47, 154.74, 155.16, 176.00 (COO), 176.23 (COO), 176.65 (COO), 176.72 (COO).

Compound 6. ^1H NMR (500 MHz, CDCl_3 , δ , ppm): 0.05–0.32 (m, 10H, CH₂ cycle), 0.40–0.67 (m, 10H, CH₂ cycle), 0.80–1.17 (m, 5H, CH

cycle), 1.23–1.60 (m, 45H, (CH₃)₃C), 1.63–1.95 (m, 10H, CH₂), 3.23–4.07 (m, 5H, NHCH), 4.48–5.15 (m, 5H, NHCH).

^{13}C NMR (126 MHz, CDCl_3 , δ , ppm): 4.49 (CH₂ cycle), 4.53 (CH₂ cycle), 4.75 (CH₂ cycle), 4.79 (CH₂ cycle), 4.88 (CH₂ cycle), 5.07 (CH₂ cycle), 5.32 (CH₂ cycle), 5.37 (CH₂ cycle), 7.19 (CH cycle), 7.54 (CH cycle), 7.60 (CH cycle), 7.69 (CH cycle), 7.83 (CH cycle), 27.96 ((CH₃)₃C), 28.01 ((CH₃)₃C), 39.60 (CH₂), 39.75 (CH₂), 39.98 (CH₂), 40.02 (CH₂), 40.36 (CH₂), 58.76 (NHCH), 59.10 (NHCH), 59.78 (NHCH), 60.79 (NHCH), 61.21 (NHCH), 64.14 (C_{sp3, cage}), 64.36 (C_{sp3, cage}), 65.53 (C_{sp3, cage}), 67.39 (C_{sp3, cage}), 68.49 (C_{sp3, cage}), 77.11 (C-Cl), 80.65 ((CH₃)₃C), 80.93 ((CH₃)₃C), 80.97 ((CH₃)₃C), 81.10 ((CH₃)₃C), 81.18 ((CH₃)₃C), 140.99, 141.83, 142.20, 142.51, 142.58, 142.71, 142.83, 143.32, 143.61, 143.65, 143.70, 143.75, 143.93, 144.02, 144.06, 144.08, 144.12, 144.17, 144.29, 144.43, 144.57, 145.16, 145.41, 145.66, 146.43, 147.17, 147.20, 147.33, 147.36, 147.41, 147.98, 148.24, 148.32, 148.43, 148.52, 148.58, 148.74, 148.80, 149.65, 150.64, 151.17, 151.66, 154.44, 154.71, 155.29, 175.76 (COO), 175.89 (COO), 176.15 (COO), 176.20 (COO), 176.37 (COO).

Compounds 7–10 have been synthesized using method of direct arylation of chlorofullerene C₆₀Cl₆ with aromatic acids reported earlier [53].

Compound 7. ^1H NMR (500 MHz, (CD₃)₂SO, δ , ppm): 2.27 – 2.45 (m, 20H, CH₂), 2.62 – 2.90 (m, 10H, CH₂), 3.13 – 3.30 (m, 10H, CH₂), 6.92 – 7.83 (m, 20H, CH), 7.89 – 8.02 (m, 5H, NH), 12.17 (br. s, 5H, COOH), ^{13}C NMR (126 MHz, (CD₃)₂SO, δ , ppm): 30.90 (CH₂), 31.14 (CH₂), 34.35 (CH₂), 35.18 (CH₂), 35.24 (CH₂), 35.25 (CH₂), 36.87 (CH₂), 37.11 (CH₂), 37.20 (CH₂), 57.85 (C_{sp3, cage}), 60.58 (C_{sp3, cage}), 63.15 (C_{sp3, cage}), 76.53 (C_{sp3, cage}-Cl), 128.31, 128.48, 128.64, 128.72, 129.35, 129.43, 134.53, 136.00, 141.29, 141.60, 141.79, 141.87, 142.02, 142.78, 143.57, 143.64, 143.76, 143.90, 144.06, 144.10, 144.17, 144.22, 144.40, 144.43, 145.50, 147.18, 147.22, 147.33, 147.78, 148.07, 148.17, 148.41, 148.53, 148.57, 148.61, 148.72, 150.69, 151.40, 154.10, 156.85, 171.74 (CO), 171.82 (CO), 171.85 (CO), 173.31 (CO), 173.37 (CO).

Spectral data for compounds 8–10 have been published earlier [53]. Compounds 11–20 have been obtained and characterized as reported previously by our group [54].

2.2. Dynamic light-scattering (DLS) measurements

DLS measurements were performed on a Photocor Complex photon correlation spectrometer at 23 °C using the near-infrared laser (790 nm). Compounds 1–20 (4 mg) were dissolved in deionized water (4 mL), and filtered through a PES syringe filter (0.45 μm). The measurements were performed in glass vials (d = 10 mm).

2.3. Cell culture

A549, H460 and H1299 NSCLC cell lines were obtained from the American Type Culture Collection (ATCC). The human dermal fibroblasts were obtained from human foreskin (IRB#100–05–251) [55]. All types of NSCLC cells were grown in the Roswell Park Memorial Institute medium (RPMI 1640, Gibco, USA), while the human fibroblasts were cultured in the high-glucose Dulbecco's modified Eagle medium (DMEM, Gibco, USA). The culture medium contains 10% fetal bovine serum (FBS, Gibco, USA) and 10 mg/L Antibiotic-Antimycotic (Gibco, USA). Cells were maintained in a humidified 5% CO₂ chamber at 37 °C. The cultured cells were resuspended using 0.25% trypsin/EDTA solution (Gibco) every 3 days for subculture process. The medium replacement was performed two times every week.

2.4. Cell viability analysis

A total number of 1500 cells for A549, H460 and H1299 NSCLC cell lines were respectively seeded in each well of 96-well plates and then maintained in the humidified incubator for 24 h before treatment with fullerene derivatives. After 1 d of incubation, all of the cultured medium was replaced by the cell culture medium containing dissolved fullerene derivatives and then cultured for another 72 h. 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) powder (CAS# 298–93–1, Alfa Aesar) in deionized water (0.5 mg/mL) was added to each well of the culture plates and then reacted for 60 min under the condition of 37 °C. Subsequently, the purple formazan products were dissolved by dimethyl sulfoxide, and the colored solution was measured at a test wavelength of 570 nm via a multimode microplate reader (SpectraMax® iD3).

2.5. QSAR modeling

The 2D and 3D geometry structure of twenty fullerene derivatives were drawn and constructed by ChemDraw (PerkinElmer, USA) and BIOVIA Discovery Studio software (BIOVIA, USA), respectively. The quotative molecular descriptors of the constructed geometry structures were calculated by the Calculate Molecular Property program of the Discovery Studio software. For regression models generation, the cell survival rate of each fullerene compound was defined as the dependent property while the quotative descriptors were regarded as independent property. The appropriated molecular descriptors related to the cell survival rate were selected using the Genetic Function Approximation (GFA) algorithm. [56] The reasonable statistical parameters including coefficient of determination r^2 and leave-one-out cross-validation q^2 were used to determine the recommended equations from results of QSAR modeling.

2.6. Statistical analyses

The measurements of all experimental data were presented as the mean \pm standard deviation (SD). Each experimental groups were performed independently at least three times for reproducibility. The value of cross-validation (q^2) to estimate the accuracy of each QSAR model was calculated by the following equation:

$$q^2 = 1 - \frac{PRESS}{TSS} \quad (1)$$

where *PRESS* means the predicted residual error sum of squares, and the *TSS* indicates the total sum of squares of the differences.

3. Results

3.1. Synthesis and aggregational behavior of water-soluble fullerene derivatives

Although the synthesis of compounds **8–20** was reported previously, water-soluble fullerene derivatives **1–7** were synthesized and characterized for the first time in this work (Table 1). As discussed above, amphiphilic water-soluble fullerene derivatives tend to form various supramolecular structures including micelles, multilayer vesicles, and clusters of micelles and vesicles in aqueous media [57,58]. To investigate the aggregational behavior of compounds **1–20**, the DLS measurements were performed for their aqueous solutions. The obtained size distribution profiles are presented in Fig. 1. The hydrodynamic radii of the observed particles ranged from several nanometers to hundreds of micrometers, which matched the data reported before for other water-soluble

fullerene derivatives [57]. Most of the investigated compounds formed several types of supramolecular associates. However, compounds **7–10** with the solubilizing addends represented by the aromatic amides of amino acids formed mostly one type of aggregates with very narrow size distribution profiles.

3.2. Cell viability of synthesized water-soluble fullerene derivatives

Chemical structures of twenty water-soluble fullerene derivatives **1–20** [54] and the survival rates for A549, H460, and H1299 NSCLC lung cancer cells treated with the aforementioned compounds are listed in Table 1. The standard deviation (SD) of each survival rate was measured from three independent experimental groups via MTT assay. According to cytotoxicity analysis, several fullerene derivatives had significant inhibitory effects on three different lung cancer cells in a concentration of 400 μM (Fig. 2). In particular, the compounds **8, 9, 11, 12, 14, 15, 18, 19**, and **20** were more effective than the other compounds, and sufficiently reduced the survival rate below $\sim 50\%$. Compounds **1, 3, 5, 7, 13, 16**, and **17** reduced the cell growth beyond 15% for at least two different kinds of cancer cell lines. The viabilities of H460 and H1299 cells were decreased after exposure to compound **1**. The viabilities of A549 and H460 cells were inhibited by compound **3**. The viabilities of A549 and H1299 cells were reduced by compound **5**, but only slightly decreased by compounds **6** and **7**. However, compounds **2, 6, 16**, and **17** induced no significant cell death but increased the growth rate of H460 and H1299 cells. Compounds **2, 6, 16**, and **17** inhibited A549 cells, while no effective toxicity on H460 and H1299 cells was observed. The half-lethal inhibition concentrations (IC_{50}) on three kinds of lung cancer cells could only be estimated for compounds **8, 9, 11, 12, 14, 15, 18, 19**, and **20** among the twenty fullerene derivatives.

In order to determine the suitable concentration of fullerene derivatives, different concentrations of the compounds **7, 8**, and **9** were chosen to evaluate the cytotoxic effects on human fibroblasts for the initial screening. As shown in Fig. S1, the measured IC_{50} values of compounds **8** and **9** were 138.3 μM and 182.4 μM , respectively. However, the IC_{50} value of compound **7** was not observed due to low toxicity. The data showed that compounds **7, 8**, and **9** in the concentration of 50 μM did not have toxic effects on normal human cells. To investigate biological effects of the above three fullerene derivatives on cancer cells, compounds **7, 8**, and **9** of the concentration 50 μM were used to confirm the cell viability of the cancer cells. The results revealed that compounds **8** and **9** induced more dead cells than the compound **7** (Fig. S2). Because compounds **8** and **9** in 200 μM decreased the survival rate of normal cells to below 50% (Fig. S1), compounds **7, 8** and **9** of 200 μM were analyzed for effects on cell cycle and apoptosis of cancer cells. Results of cell cycle showed the arrested G0/G1 for cancer cells after the treatment with compounds **7, 8** or **9** (Fig. S3). As for cell apoptosis assay, compounds **8** and **9** significantly reduced the number of live cells, while the compound **7** did not induce strong apoptosis in cancer cells (Fig. S4).

3.3. QSAR analysis

The values of survival rate for A549, H460, and H1299 NSCLC cells were regarded as the activity to generate three different regression models for the twenty fullerene derivatives. In order to build a more reasonable QSAR model, we included the experimental data of ten fullerene derivatives from our previous report and then combined with data from the current study for model constructions. Because the ten fullerene derivatives synthesized in our previous literature only had bio-activity data for A549 cells [54], the survival rate of the H460 and H1299 cells were evaluated in the current study after treatment by the ten fullerene com-

Table 1

Molecular structures of fullerene derivatives and the viability of A549, H460, and H1299 after exposure to various fullerene derivatives at 400 μM for 72 h. The cell viability was estimated by MTT assay.

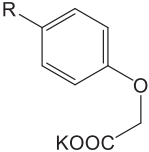
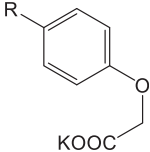
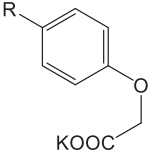
Comp.	R1	R2	R3	Survival rate (%)		
				A549	H460	H1299
1	Cl			92.10 \pm 5.30	69.62 \pm 3.72	76.37 \pm 8.31
2	H			84.70 \pm 3.79	97.63 \pm 3.35	101.00 \pm 4.01
3	H			71.62 \pm 9.68	84.54 \pm 6.34	87.58 \pm 5.48
4	Cl			88.54 \pm 3.10	85.44 \pm 2.56	104.88 \pm 2.82
5	Cl			62.84 \pm 4.58	92.84 \pm 1.36	83.51 \pm 2.34
6	Cl			79.01 \pm 5.83	105.59 \pm 6.47	102.41 \pm 3.87
7	Cl			85.12 \pm 8.50	107.75 \pm 9.31	75.90 \pm 2.56
8	Cl			50.74 \pm 6.52	21.90 \pm 5.58	23.27 \pm 3.12
9	Cl			40.57 \pm 5.87	40.68 \pm 5.28	19.16 \pm 3.36
10	Cl			102.20 \pm 7.50	105.50 \pm 4.26	104.37 \pm 4.79

Table 1 (continued)

Comp.	R1	R2	R3	Survival rate (%)		
				A549	H460	H1299
11	Cl			47.32 ± 1.78	26.76 ± 2.45	53.32 ± 3.70
12	Et			3.99 ± 0.22	26.17 ± 3.89	46.88 ± 4.69
13	Cl			4.22 ± 0.44	34.73 ± 2.34	71.34 ± 3.03
14	H			4.53 ± 0.11	19.76 ± 2.89	56.74 ± 4.12
15	H			56.38 ± 6.32	25.00 ± 2.68	58.06 ± 1.96
16	-			56.21 ± 5.01	110.91 ± 6.42	100.95 ± 6.36
17	Cl			41.69 ± 4.17	110.60 ± 4.80	84.36 ± 4.53
18	Cl			9.54 ± 0.97	26.13 ± 1.58	51.34 ± 5.91
19	H			12.58 ± 1.25	16.57 ± 2.20	56.19 ± 2.02

(continued on next page)

Table 1 (continued)

Comp.	R1	R2	R3	Survival rate (%)		
				A549	H460	H1299
20				12.05 ± 0.47	14.91 ± 2.82	40.57 ± 4.79

pounds in the concentration of 400 μM (Fig. 2). A total number of 204 molecular descriptors were calculated by the module of *Calculate Molecular Properties* of BIOVIA Discovery Studio software. Top ten linear mathematical models with seven suggested molecular properties that correlated with the survival rate of A549, H460, and H1299 are listed in Table S1–S3. The recommended equations among the generated QSAR models were suggested based on two criteria including the least squares fitting r^2 and cross-validated correlation coefficient q^2 to define three best reasonable models for each kind of NSCLC cancer cell (Table 2). The three suggested QSAR models of A549, H460, and H1299 had r^2 of 0.9325, 0.8404, and 0.9011, respectively. The comparison of the experimental and the predicted values for three recommended QSAR models are displayed in Fig. 3. For the accuracy of generated models, the value of q^2 should be beyond 0.6 to make sure that the individual equations for the three different cancer cells have high predictive power (Table 2).

3.4. Quantitative descriptors analysis

The recommended molecular descriptors of the twenty fullerene derivatives relative to survival rate of A549 cells were **O_Count**, **ES_Count_sCl**, **ES_Count_ssO**, **ES_Count_sssCH**, **ES_Count_ssssC**, **Num_Rings6**, and **Num_StereoAtoms**. The quantitative values of each molecular descriptor selected by GFA algorithm are listed in Table 3. **O_Count** indicates the number of oxygen atoms on the solubilizing addends of fullerene derivatives. The four descriptors with electrotopological states (E-state) including **ES_Count_sCl**, **ES_Count_ssO**, **ES_Count_sssCH**, and **ES_Count_ssssC** are the sums of E-state values for halogens (Cl) and organic elements (C, O, and CH) with single bonds. **Num_Rings6** descriptor means the number of rings with six atoms. Finally, **Num_StereoAtoms** defines the number of atoms belong stereo state. Comparing cell viability with the molecular descriptors, **ES_Count_sssCH** and **Num_StereoAtoms** were related to the high survival rate of A549 cells.

For H460 cells, the seven descriptors correlating with the cell viability included **ES_Count_aaS**, **ES_Count_sCH3**, **ES_Count_sssCH**, **Num_AromaticBonds**, **Num_DoubleBonds**, **Num_RingAssemblies**, and **Num_Rings6** (Table 4). The **ES_Count_aaS** means the calculated E-state sums for sulfur-containing aromatic rings. **ES_Count_sCH3** and **ES_Count_sssCH** are calculated for methyl groups with single bonds. **Num_AromaticBonds** and **Num_DoubleBonds** indicate the number of bonds in aromatic ring

systems and double bonds between a pair of atoms, respectively. **Num_RingAssemblies** denotes numbers of benzene rings joined directly by single bonds [59]. For instance, the ring assemble number of naphthalene is one, while biphenyl is two. Regarding the highly toxic fullerene compounds, higher values of **ES_Count_aaS** and **Num_AromaticBonds** were linked to the survival rate of H460 cells.

For H1299 cells, **N_Count**, **O_Count**, **ES_Count_aaS**, **ES_Count_sCl**, **ES_Count_ssssC**, **HBA_Count**, and **Num_AliphaticSingleBonds** were suggested descriptors from the best reasonable QSAR model that related to the cell survival rate (Table 5). **N_Count** and **O_Count** are the number of nitrogen and oxygen atoms, respectively. **ES_Count_aaS** defines the sums of E-state values of aromatic rings coupled a sulfur atom, while **ES_Count_sCl** and **ES_Count_ssssC** mean single bonds containing a chlorine and a carbon atom, respectively. **HBA_Count** is the number of hydrogen bond accepting atoms in soluble addends of a fullerene derivative. For the last descriptor, **Num_AliphaticSingleBonds** indicates the number of single bond in aliphatic chains of a fullerene compound. The data showed that the four fullerene derivatives with a sulfur-containing aromatic ring (compounds **8**, **9**, **11**, and **12**) demonstrated efficient inhibition effect and reduced the survival rate of H1299 down to below ~ 50%. Among the sulfur-containing compounds, compounds **8** and **9** have more number of oxygen atoms and more inhibitory activity than the previously reported compounds (compounds **11** and **12**). Interestingly, the quantitative data of **Num_AliphaticSingleBonds** and **ES_Count_aaS** descriptors revealed that more numbers of aliphatic chains and no sulfur atom in the structure of fullerene derivatives decreased the inhibitory effects for H1299. However, the fullerene derivatives with a sulfur-containing aromatic ring had distinct results, i.e. increasing the value of aliphatic chains in these compounds could promote the toxic effects for H1299 cells.

The favored and disfavored descriptors of fullerene compounds against A549, H460, and H1299 lung cancer cells are summarized in Table 6. The oxygen and the stereo atoms are disfavored properties for A549 cells, and the fullerene compounds with these descriptors had low inhibitory effects. The aromatic bond and sulfur-containing aromatic ring are favored chemical functionalities and can facilitate the cytotoxicity of fullerene compounds against H460 cells. For H1299 cells, the oxygen atoms and sulfur-containing aromatic ring are favored properties for the fullerene compounds to increase the inhibitory effect on cells, while the number of aliphatic chain is the disfavored property.

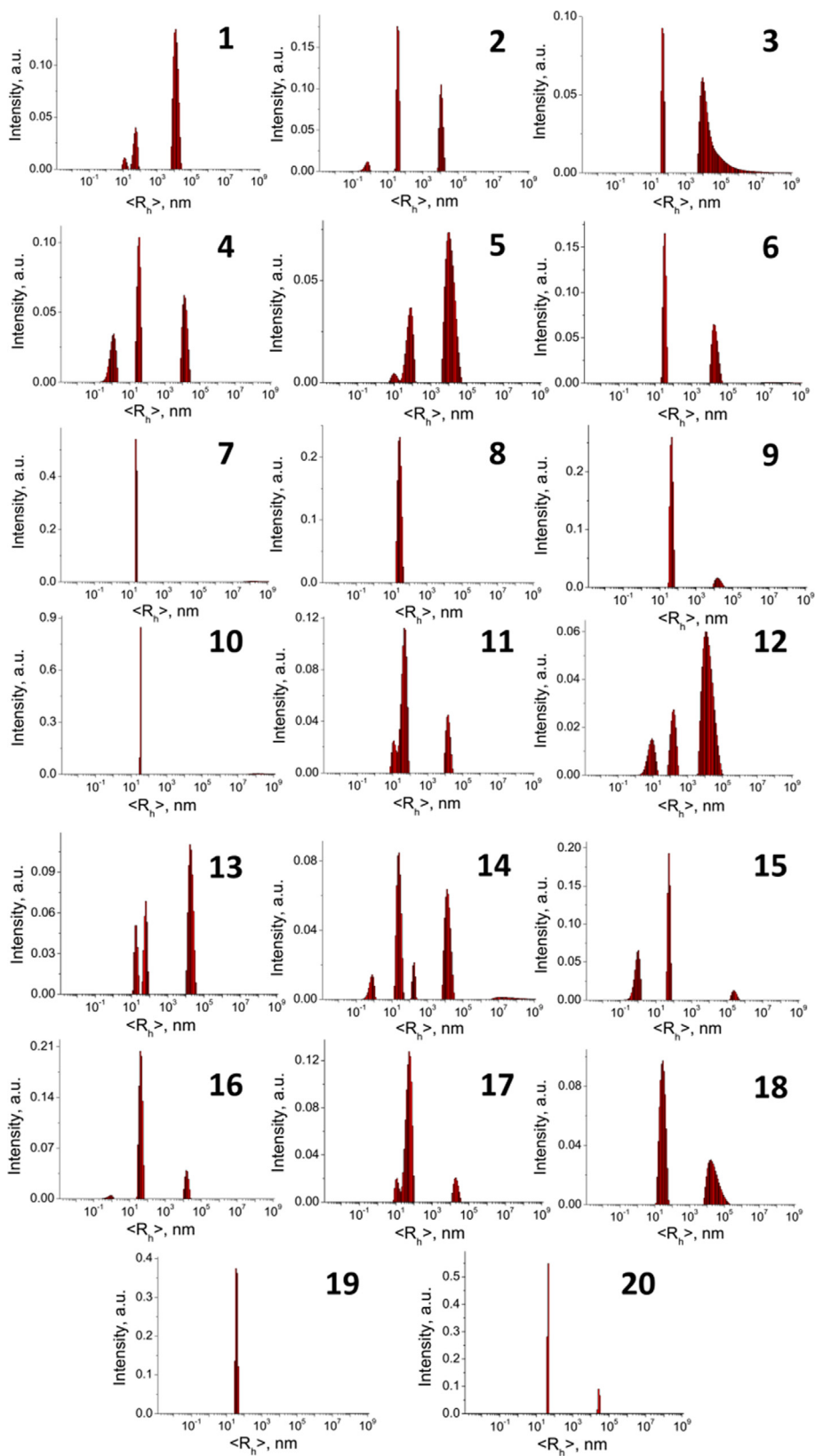


Fig. 1. DLS profiles showing the particle size distribution in aqueous solutions of the fullerene derivatives 1–20.

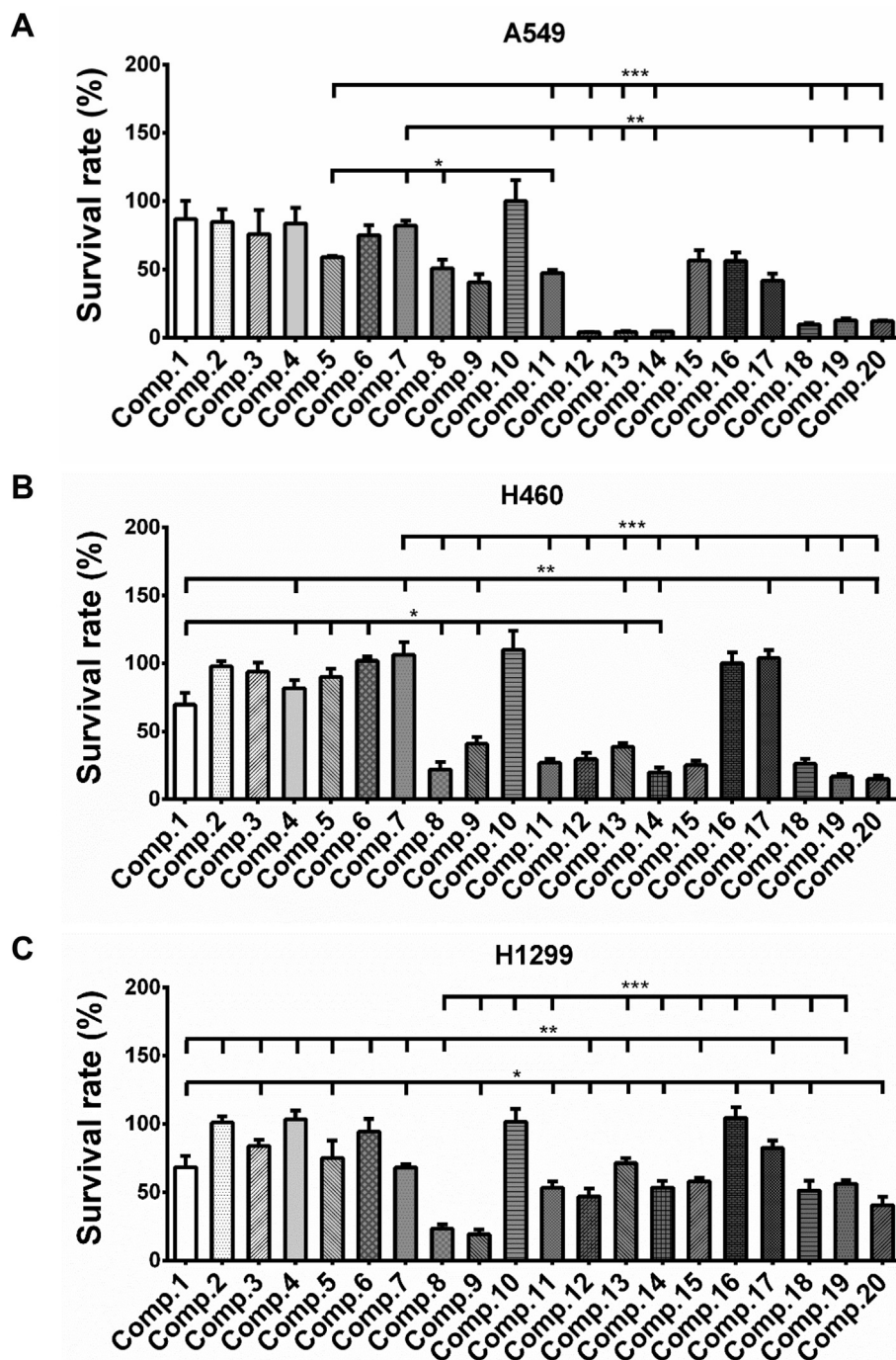


Fig. 2. The survival rate of A549, H460, and H1299 lung cancer cells upon exposure to ten water-soluble fullerene derivatives at 400 μM , measured by the MTT assay. Significant difference was represented by asterisks: * $p < 0.05$, ** $p < 0.01$; ***, $p < 0.001$.

Table 2

The recommended equation models of the survival rate (%) for A549, H460, and H1299.

QSAR equations	
Model (A549)	$= -202.67 + 4.7734 \times \text{O_Count} + 14.913 \times \text{ES_Count_sCl} - 5.8748 \times \text{ES_Count_ssO} + 6.7969 \times \text{ES_Count_sssCH} + 45.856 \times \text{ES_Count_ssssC} + 7.5748 \times \text{Num_Rings6} - 2.1975 \times \text{Num_StereoAtoms}$ ($r^2 = 0.9325$, $q^2 = 0.8319$)
Model (H460)	$= 1000.3 - 62.582 \times \text{ES_Count_aaS} + 12.458 \times \text{ES_Count_sCH3} - 13.399 \times \text{ES_Count_sssCH} - 0.6373 \times \text{Num_AromaticBonds} + 9.1893 \times \text{Num_DoubleBonds} + 29.379 \times \text{Num_RingAssemblies} - 61.785 \times \text{Num_Rings6}$ ($r^2 = 0.8404$, $q^2 = 0.6428$)
Model (H1299)	$= 34.969 - 7.6348 \times \text{N_Count} - 7.6348 \times \text{N_Count} + 3.1813 \times \text{O_Count} - 11.625 \times \text{ES_Count_aaS} + 11.727 \times \text{ES_Count_sCl} + 33.851 \times \text{ES_Count_ssssC} - 5.1685 \times \text{HBA_Count} + 0.77298 \times \text{Num_AliphaticSingleBonds}$ ($r^2 = 0.9011$, $q^2 = 0.7006$)

4. Discussion

Fullerenes and their derivatives are one of the most promising types of nanomaterials and have attracted a lot of attention for their perspective biomedical applications. Functionalized fullerene molecules can be utilized as platforms for drug delivery [60], promising antiviral [61], and antitumor drugs [62]. For example, the soluble fullerene derivatives can be used as anti-HIV drugs inhibiting viral replication [63]. Moreover, the functionalized fullerene derivatives have been widely investigated for antitumor effects on brain or lung cancer cells in recent years [52,58]. The QSAR approaches can offer a useful strategy to provide key information from the relationship between physico-chemical properties of investigated compounds and their biological activities for drug design. However, the anticancer activity of the water-soluble fullerene derivatives is still underinvestigated. In this study we utilized twenty water-soluble fullerene derivatives and three lung cancer cell lines in order to generate reasonable QSAR models for design of novel antitumor agents.

In our previous study, only ten fullerene derivatives (compounds **11–20**) were investigated for their inhibitory effects on pulmonary epithelial cell line A549. More importantly, the characteristic of EMT has been displayed in the progression of lung cancer cells [64]. The other two kinds of lung cancers with mesenchymal features, H460 and H1299, were used to estimate the inhibitory activity of ten reported fullerene derivatives in the current study. However, IC_{50} values of certain fullerene derivatives such as compounds **1–6**, **9**, **10**, **13**, **16**, and **17** were not estimated on three lung cancer cell lines due to the low cytotoxicity. Using an alternative method to obtain the inhibitory effects, we utilized MTT assay to obtain the cell survival rate of each type of lung cancer cells after exposure to the tested fullerene derivatives at the concentration of 400 μ M. The obtained survival rate of individual lung cancer cells was considered as bioactivity for building the QSAR models instead of IC_{50} values.

We further analyzed the descriptors suggested from the three different recommended equation models responsible for the individual survival rate of A549, H460, and H1299. Regarding the linear regression model of A549 cells from our previous work [54], the suggested molecular descriptors were **S_Count**, **ES_Count_dO**, **ES_Count_sCl**, **AliphaticSingleBonds**, **AromaticRings**, **H_Acceptors**, and **RingBonds**. Comparing two equation models from previous and current studies on A549, we noted that the key features including oxygen atom and chlorine atom are given by both equations. More effective fullerene compounds for A549 cells have less oxygen and chlorine atoms than the other fullerene derivatives. In addition, we also discovered more critical features such as **ES_Count_sssCH**, **ES_Count_ssssC**, and **Num_StereoAtoms** in the current equation model. Fullerene derivatives with a lower number of carbon atoms with single bond and less atoms with stereo state could reduce the cell viability of A549 to below ~ 50%. Based on our equation model, it was not suitable to increase the number of atoms with single bond such as oxygen, chlorine, and carbon for designing fullerene derivatives as A549 inhibitors.

As the result of QSAR model generation for H460, the sixth equation was the most accurate among the ten models suggested by GFA algorithm (Table S2). The sixth QSAR model of H460 containing key features such as **ES_Count_aaS** and **Num_AromaticBonds**. These two molecular descriptors illustrate that the sulfur atom and the aromatic ring are both critical functional groups among the twenty fullerene derivatives. Moreover, the sulfur atom and the aromatic ring were also regarded as key features in the QASR equation of A549 previously reported. Hence, the 6th equation of H460 mentioned above may be reasonable to predict the relationship between inhibitory effects and chemical proper-

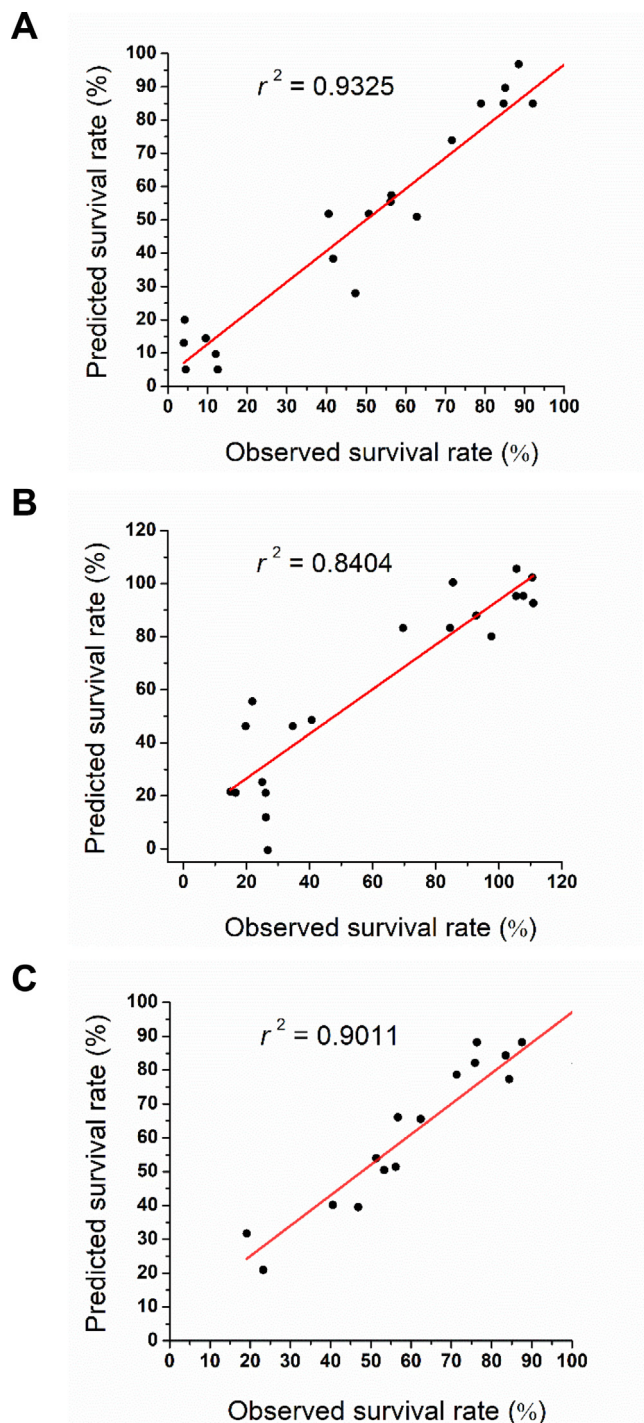


Fig. 3. Correlation between the observed and predicted survival rate of A549, H460, and H1299 from the recommended QSAR model.

ties. We further discussed the ten fullerene derivatives that had inhibitory effects to reduce the survival rate of H460 below ~ 50%, including compounds **8**, **9**, **11–15**, and **18–20**. Fullerene derivatives **8**, **9**, **11**, and **12** contain five numbers of **ES_Count_aaS** descriptor, indicating that sulfur-containing aromatic ring is a potential functional group to increase the anticancer activities of these four fullerene derivatives. In addition, compounds **13–15** and **18–20** have more numbers of **Num_AromaticBonds** descriptor than the other fullerene compounds. Interestingly, compounds **7** and **10** had similar features with compounds **11** and **12**. For instance, all of the

Table 3

The quantitative values of the chemical properties from the recommended equation of QSAR model compared to cell viabilities of A549 estimated by MTT assay.

Comp.	Survival rate (%)		Structural parameters						
	Observation	Prediction	O_Count	ES_Count_sCl	ES_Count_ssO	ES_Count_sssCH	ES_Count_ssssC	Num_Rings6	Num_StereoAtoms
10	102.20	89.69	15	1	5	0	1	25	0
1	92.10	84.93	10	1	5	10	1	20	5
4	88.54	96.80	10	1	5	5	2	20	5
7	85.12	89.69	15	1	5	0	1	25	0
2	84.70	84.93	10	0	5	10	1	20	5
6	79.01	84.93	10	1	5	10	1	20	5
3	71.62	73.94	10	0	5	10	1	20	10
5	62.84	50.95	10	1	5	5	1	20	5
15	56.38	57.40	20	0	10	5	0	25	0
16	56.21	55.50	16	0	0	0	0	24	0
8	50.74	51.82	15	1	5	0	1	20	0
11	47.32	27.95	10	1	5	0	1	20	0
17	41.69	38.33	20	1	10	0	0	25	0
9	40.57	51.82	15	1	5	0	1	20	0
19	12.58	5.06	10	0	5	0	0	25	0
20	12.05	9.69	18	0	12	0	0	26	0
18	9.54	14.46	15	1	10	0	0	25	0
14	4.53	5.06	10	0	5	0	0	25	0
13	4.22	19.97	10	1	5	0	0	25	0
12	3.99	13.04	10	0	5	0	1	20	0

Table 4

The quantitative values of the chemical properties from the recommended equation of QSAR model compared to cell viabilities of H460 estimated by MTT assay.

Comp.	Survival rate (%)		Structural parameters						
	Observation	Prediction	ES_Count_aaS	ES_Count_sCH3	ES_Count_sssCH	Num_Aromatic Bonds	Num_Double Bonds	Num_Ring Assemblies	Num_Rings6
16	110.91	92.62	0	0	0	49	50	5	24
17	110.60	102.33	0	0	0	55	55	6	25
7	107.75	95.32	0	0	0	66	55	6	25
6	105.59	105.59	0	0	10	36	35	6	20
10	105.50	95.32	0	0	0	66	55	6	25
2	97.63	80.09	0	10	10	41	35	1	20
5	92.84	87.98	0	5	5	36	35	1	20
4	85.44	100.44	0	6	5	36	35	1	20
3	84.54	83.28	0	10	10	36	35	1	20
1	69.62	83.28	0	10	10	36	35	1	20
9	40.68	48.58	5	0	0	61	50	6	20
13	34.73	46.19	0	0	0	71	50	6	25
11	26.76	-0.56	5	0	0	66	45	6	20
12	26.17	11.90	5	1	0	66	45	6	20
18	26.13	21.07	0	0	0	96	49	6	25
15	25.00	25.14	0	0	5	71	55	6	25
8	21.90	55.59	5	0	0	50	50	6	20
14	19.76	46.19	0	0	0	71	50	6	25
19	16.57	21.07	0	0	0	96	49	6	25
20	14.91	21.60	0	0	0	102	53	7	26

above four compounds had sixty-six aromatic bonds in attached solubilizing addends (**Num_AromaticBonds** = 66). In the results of cell viability of H460, the compounds **11** and **12** displayed more cytotoxic effects than compounds **7** and **10**. Because compounds **7** and **10** had no sulfur-containing aromatic rings on the soluble addends, suggesting that this key feature increased the inhibitory activity of fullerene derivatives to reduce the survival rate of H460 cells. According to our data analysis, the aromatic rings containing sulfur and their aromatic bonds are favored chemical properties that can enhance the bioactivity of fullerene compounds on H460 cells.

The recommended equation model of H1299 had two critical molecular descriptors that has been displayed in a previously reported equation model of A549 [54], including **ES_Count_sCl**, and **Num_AliphaticSingleBonds** descriptors. Our previous study demonstrated that lower numbers of chlorine atoms and aliphatic single bonds on the soluble addends of fullerene compounds promoted the antitumor activities on A549 cells. In the current study,

we found that fullerene compounds containing **ES_Count_aaS** descriptor had significant cytotoxic effects on both of H460 and H1299 cells (Table 5), indicating that the sulfur-containing aromatic ring of fullerene derivatives could help to reduce the cell viability of these two cell lines. Notably, the favored properties for the cytotoxicity of the fullerene compounds without **ES_Count_aaS** descriptor are contrary to those of the fullerene compounds with **ES_Count_aaS** descriptor. Among the fullerene derivatives with no sulfur-containing aromatic rings, more numbers of nitrogen atoms, chlorine atoms, and aliphatic single bonds reduced the cytotoxic effects on H1299. For example, compound **15** has a similar quantitative value of suggested descriptors to compound **20** but different cytotoxicity. The compound **20** has lower numbers of aliphatic single bonds than that of compound **15**. We further investigate the fullerene derivatives with similar numbers of aliphatic single bonds such as compounds **18** and **19**. The compound **18** have slightly more inhibitory effects than compound **19** due to the difference in **O_Count** and **HBA_Count** descriptors. The com-

Table 5

The quantitative values of the chemical properties from the recommended equation of QSAR model compared to cell viabilities of H1299 estimated by MTT assay.

Comp.	Survival rate (%)		Structural parameters						
	Observation	Prediction	N_Count	O_Count	ES_Count_aaS	ES_Count_sCl	ES_Count_ssssC	HBA_Count	Num_AliphaticSingleBonds
4	104.88	115.10	5	10	0	1	2	10	76
10	104.37	93.73	5	15	0	1	1	15	105
6	102.41	92.07	5	10	0	1	1	10	90
2	101.00	82.02	5	10	0	0	1	10	77
16	100.95	100.95	0	16	0	0	0	8	73
3	87.58	88.20	5	10	0	0	1	10	85
17	84.36	77.29	0	20	0	1	0	20	91
5	83.51	84.34	5	10	0	1	1	10	80
1	76.37	88.20	5	10	0	1	1	10	85
7	75.90	82.13	5	15	0	1	1	15	90
13	71.34	78.61	0	10	0	1	0	10	67
15	62.40	65.56	0	20	0	0	0	20	91
14	56.74	66.11	0	10	0	0	0	10	66
19	56.19	51.43	0	10	0	0	0	10	47
11	53.32	50.47	0	10	5	1	1	10	62
18	51.34	53.99	0	15	0	1	0	15	48
12	46.88	39.52	0	10	5	0	1	10	63
20	40.57	40.17	0	18	0	0	0	18	53
8	23.27	20.91	5	15	5	1	1	15	86
9	19.16	31.74	5	15	5	1	1	15	100

Table 6

The favored and disfavored chemical properties for three kinds of lung cancer cells.

Lung cancer	Chemical properties		
	Favored		Disfavored
A549	–	–	O_Count
H460	Aromatic Bonds	aaS	–
H1299	O_Count	aaS	Aliphatic Single Bond
			Num_StereoAtoms
			–
			–

Note: Favored properties: a chemical property that enhances the toxicity of compounds. Disfavored properties: a chemical property that decreases the toxicity of compounds.

compound **18** has more numbers of oxygen atoms and H-bond acceptors than compound **19**, suggesting that **O_Count** and **HBA_Count** descriptors are favored properties for the cytotoxic effects of fullerene derivatives with no sulfur-containing aromatic ring. In contrast, the favored chemical descriptors for the cytotoxicity of the fullerene compounds bearing sulfur-containing aromatic ring are **N_Count**, **O_Count**, **HBA_Count**, and **Num_AliphaticSingleBonds**. For instance, compounds **8** and **9** have more numbers of these favored properties than compounds **11** and **12**. Hence, increasing the numbers of sulfur-containing aromatic rings, nitrogen atoms, oxygen atoms, H-bond acceptors, and aliphatic single bonds on the addend of fullerene derivatives can promote the cytotoxic effect on H1299 cells.

We further analyzed if compounds **8** and **9** had inhibitory effects on cancer cells under non-toxic dose for normal human cells. Our data showed that 50 μ M of fullerene compounds had no toxic effects on human cells (Fig. S1). Interestingly, compounds **8** and **9** in the concentration of 50 μ M increased the number of dead cells by live/dead cell viability staining assay (Fig. S2). In addition, both of compounds **8**, and **9** could induce cell cycle arrest at the G0/G1 stage and cell apoptosis. In comparison with the *in vitro* results and functional groups on fullerene cage, compounds **8** and **9** with the thiophene moiety have strongly inhibitory effect. On the other hand, the functional group of compound **7** contains benzene ring, which have less toxic effect than compounds **8** and **9**. These data suggested that the thiophene moiety may have potent effects in reducing the tumor growth.

Thiophene is a monocyclic heteroarene that consists of sulfur atom in a planar five-membered heterocyclic ring [65]. The thiophene moiety is commonly present in various thiophene-based

drugs including nonsteroidal anti-inflammatory (NSAID), antihistaminic, antiasthma, and antitumor drugs [66,67]. Therefore, thiophene derivatives are an important class of heterocyclic compounds in the field of medical applications. Among our tested fullerene derivatives, compounds **8**, **9**, **11**, and **12** are thiophene-containing fullerene derivatives. The thiophene moiety in the solubilizing addends of these four fullerene derivatives was represented by the **ES_Count_aaS** descriptor in the quantitative data of QSAR models. Regarding the fullerene compounds with thiophene moiety, compound **12** had more inhibitory effects on A549 than compounds **8**, **9**, and **11**. The compound **12** has an ethyl group at R1 position of the fullerene derivative structure, while compounds **8**, **9**, and **11** have a chlorine atom at R1 position. Our data indicated that the ethyl group improved the cytotoxicity of thiophene-based fullerene derivatives on A549 cells. However, compound **12** also had lower numbers of nitrogen atoms, oxygen atoms, H-bond acceptors, and aliphatic single bonds than compounds **8** and **9**, which might be responsible for its lower cytotoxic effects on H1299. Comparing the inhibitory effects between compounds **8** and **9** on H460 cells, we noticed that compound **8** with five aliphatic single CH₂ groups had lower anticancer effects than compound **9**. Thus, thiophene-containing fullerene derivative with shorter aliphatic spacers enables an enhancement of the anticancer effect, and this finding was also supported by our previous reported model [54]. The compound **12**, bearing an ethyl group at R1 position, had strong inhibitory effect on A549 cells. Moreover, the compound **8** with a chlorine atom at R1 position had strong cytotoxic effects on both H460 and H1299, but did not significantly reduce the survival rate of A549 cells. Hence, we assumed that replacing the ethyl group at R1 position of the compound **8** might

strongly enhance the inhibitory effects on A549, which may be used as an inhibitor for all three kinds of NSCLC cells.

5. Conclusion

The QSAR models were generated based on twenty synthesized water-soluble fullerene derivatives and their antitumor activity identified key structural features on the three different kinds of lung cancer cells. Based on the equations of QSAR models, fullerene derivatives with thiophene moiety such as compounds **8**, **9**, **11**, and **12** had stronger cytotoxic effects on A549, H460, and H1299 than the other fullerene compounds. In addition, increase in the number of nitrogen atoms, oxygen atoms, H-bond acceptors, and aliphatic chains in the structures of solubilizing addends of thiophene-based fullerene compound could promote the cytotoxic effect on H1299 cells. Meanwhile, increase in the number of nitrogen atoms, chlorine atoms, and aliphatic chains on the fullerene derivatives with no thiophene moiety reduced the cytotoxic effect on H1299 cells. The current study demonstrated that thiophene-containing fullerene derivatives had strong anticancer activities on three different types of lung cancer cells based on QSAR modeling, which may help to provide important information on a rational design of the water-soluble fullerene derivatives for lung cancer treatment in the future.

CRedit authorship contribution statement

Hung-Jin Huang: Conceptualization, Writing - original draft, Investigation. **Margarita Chetyrkina:** Methodology, Validation, Visualization. **Chui-Wei Wong:** Methodology, Validation, Visualization. **Olga A. Kraevaya:** Methodology, Investigation. **Alexander V. Zhilenkov:** Methodology, Investigation. **Ilya I. Voronov:** Methodology, Investigation. **Pei-Hwa Wang:** Resources. **Pavel A. Troshin:** Conceptualization. **Shan-hui Hsu:** Conceptualization.

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.

Acknowledgements

This research was supported by the bilateral Taiwanese-Russian research project, Ministry for Science and Education of the Russian Federation (RFBR No. 16-53-52030; MOST 105-2923-E-002-003-MY3), and Ministry of Science and Technology (MOST 109-2622-E-002-012-CC1). This work was also supported by National Taiwan University Core Consortium (NTU-CC-109L891001). Synthesis of water-soluble fullerene derivatives was supported by the Russian Science Foundation (project 19-13-00411). We are thankful to the Joint Center for Instruments and Researches, College of Bioresearches and Agriculture, National Taiwan University for providing technical assistance in confocal microscopy.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2021.01.012>.

References

- [1] Li X, Zhang Q, Yang Z. Knockdown of hsa_circ_0058124 inhibits the proliferation of human lung cancer cells by up-regulation of miR-1297. *Artif Cells Nanomed Biotechnol* 2020;48(1):584–93.
- [2] Xu WT, Li TZ, Li SM, Wang C, Wang H, et al. Cytisine exerts anti-tumour effects on lung cancer cells by modulating reactive oxygen species-mediated signalling pathways. *Artif Cells Nanomed Biotechnol* 2020;48(1):84–95.
- [3] de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, et al. Reduced lung-cancer mortality with volume ct screening in a randomized trial. *N Engl J Med* 2020;382(6):503–13.
- [4] Goebel C, Loudon CL, McKenna Jr R, Onugha O, Wachtel A, et al. Diagnosis of non-small cell lung cancer for early stage asymptomatic patients. *Cancer Genomics Proteomics* 2019;16(4):229–44.
- [5] Wu H, Zhou C. Long non-coding RNA UCA1 promotes lung cancer cell proliferation and migration via microRNA-193a/HMGBl axis. *Biochem Biophys Res Commun* 2018;496(2):738–45.
- [6] Song Y, Kong L, Sun B, Gao L, Chu P, et al. Induction of autophagy by an oleanolic acid derivative, SZC017, promotes ROS-dependent apoptosis through Akt and JAK2/STAT3 signaling pathway in human lung cancer cells. *Cell Biol Int* 2017;41(12):1367–78.
- [7] Zhang C, Leigh NB, Wu Y-L, Zhong W-Z. Emerging therapies for non-small cell lung cancer. *J Hematol Oncol* 2019;12(1):45.
- [8] Cheng M, Jolly S, Quarshie WO, Kapadia N, Vigneau FD, et al. Modern radiation further improves survival in non-small cell lung cancer: an analysis of 288,670 patients. *J Cancer* 2019;10(1):168–77.
- [9] Arbour KC, Riely GJ. Systemic therapy for locally advanced and metastatic non-small cell lung cancer: a review. *JAMA* 2019;322(8):764–74.
- [10] Dong J, Li B, Lin D, Zhou Q, Huang D. Advances in targeted therapy and immunotherapy for non-small cell lung cancer based on accurate molecular typing. *Front Pharmacol* 2019;10:230.
- [11] Doroshov DB, Sanmamed MF, Hastings K, Politi K, Rimm DL, et al. Immunotherapy in non-small cell lung cancer: Facts and hopes. *Clin Cancer Res* 2019;25(15):4592–602.
- [12] Da Dalt G, Friziero A, Grego A, Serafini S, Fassina A, et al. Adrenal metastasis from endometrial cancer: A case report. *World J Clin Cases* 2019;7(14):1844–9.
- [13] Funazo T, Nomizo T, Kim YH. Liver metastasis is associated with poor progression-free survival in patients with non-small cell lung cancer treated with nivolumab. *J Thorac Oncol* 2017;12(9):e140–1.
- [14] Schulz M, Salamero-Boix A, Niesel K, Alekseeva T, Sevenich L. Microenvironmental regulation of tumor progression and therapeutic response in brain metastasis. *Front Immunol* 2019;10:1713.
- [15] Hiraga T. Bone metastasis: Interaction between cancer cells and bone microenvironment. *J Oral Biosci* 2019;61(2):95–8.
- [16] Liu Z, Liang H, Lin J, Cai X, Pan Z, et al. The incidence of lymph node metastasis in patients with different oncogenic driver mutations among T1 non-small-cell lung cancer. *Lung Cancer* 2019;134:218–24.
- [17] Dako F, Hota P, Kahn M, Kumaran M, Agosto O. Post-lung transplantation abdominopelvic complications: the role of multimodal imaging. *Abdom Radiol* 2020;45(4):1202–13.
- [18] Zhang P, Tao H, Yu L, Zhou L, Zhu C. Developing protein arginine methyltransferase 1 (PRMT1) inhibitor TC-E-5003 as an antitumor drug using INEL drug delivery systems. *Drug Deliv* 2020;27(1):491–501.
- [19] Skamrova GB, Laponogov I, Buchelnikov AS, Shckorbatov YG, Prylutska SV, et al. Interceptor effect of C60 fullerene on the in vitro action of aromatic drug molecules. *Eur Biophys J* 2014;43(6):265–76.
- [20] Mukherjee A, Paul M, Mukherjee S. Recent progress in the theranostics application of nanomedicine in lung cancer. *Cancers (Basel)* 2019;11(5):597.
- [21] Wang X, Chen H, Zeng X, Guo W, Jin Y, et al. Efficient lung cancer-targeted drug delivery via a nanoparticle/MSC system. *Acta Pharmaceutica Sinica B* 2019;9(1):167–76.
- [22] Prylutskyi YI, Evstigneev MP, Cherepanov VV, Kyzyma OA, Bulavin LA, et al. Structural organization of C60 fullerene, doxorubicin, and their complex in physiological solution as promising antitumor agents. *J Nanopart Res* 2015;17(1):45.
- [23] Prylutska S, Politenkova S, Afanasieva K, Korolovych V, Bogutskaya K, et al. A nanocomplex of C(60) fullerene with cisplatin: design, characterization and toxicity. *Beilstein J Nanotechnol* 2017;8:1494–501.
- [24] Sivarajakumar R, Mallukaraj D, Kadavakollu M, Neelakandan N, Chandran S, et al. Nanoparticles for the treatment of lung cancers. *J Young Pharm* 2018;10(3):276–81.
- [25] Prylutska S, Grynyuk I, Skaterna T, Horak I, Grebinyk A, et al. Toxicity of C60 fullerene-cisplatin nanocomplex against Lewis lung carcinoma cells. *Arch Toxicol* 2019;93(5):1213–26.
- [26] Prylutska SV, Lynchak OV, Kostjukov VV, Evstigneev MP, Remeniak OV, et al. Antitumor effects and hematotoxicity of C(60)-Cis-Pt nanocomplex in mice with Lewis lung carcinoma. *Exp Oncol* 2019;41(2):106–11.
- [27] Roche J. The epithelial-to-mesenchymal transition in cancer. *Cancers (Basel)* 2018;10(2).
- [28] Liu D, Ertay A, Hill C, Zhou Y, Li J, et al. ASP1 deficiency promotes epithelial-mesenchymal transition, invasion and metastasis in colorectal cancer. *Cell Death Dis* 2020;11(4):224.
- [29] Schliekelman MJ, Taguchi A, Zhu J, Dai X, Rodriguez J, et al. Molecular portraits of epithelial, mesenchymal, and hybrid states in lung adenocarcinoma and their relevance to survival. *Cancer Res* 2015;75(9):1789–800.
- [30] Leung EL, Fiscus RR, Tung JW, Tin VP, Cheng LC, et al. Non-small cell lung cancer cells expressing CD44 are enriched for stem cell-like properties. *PLoS ONE* 2010;5(11):e14062.
- [31] Li CH, Liu CW, Tsai CH, Peng YJ, Yang YH, et al. Cytoplasmic aryl hydrocarbon receptor regulates glycogen synthase kinase 3 beta, accelerates vimentin

- degradation, and suppresses epithelial-mesenchymal transition in non-small cell lung cancer cells. *Arch Toxicol* 2017;91(5):2165–78.
- [32] Cruz-Bermúdez A, Laza-Briviesca R, Vicente-Blanco RJ, García-Grande A, Coronado MJ, et al. Cancer-associated fibroblasts modify lung cancer metabolism involving ROS and TGF- β signaling. *Free Radic Biol Med* 2019;130:163–73.
- [33] Treesuwan S, Sritularak B, Chanvorachote P, Pongrakhananon V. Cypripedin diminishes an epithelial-to-mesenchymal transition in non-small cell lung cancer cells through suppression of Akt/GSK-3 β signalling. *Sci Rep* 2018;8(1):8009.
- [34] Andriani F, Bertolini G, Facchinetti F, Baldoli E, Moro M, et al. Conversion to stem-cell state in response to microenvironmental cues is regulated by balance between epithelial and mesenchymal features in lung cancer cells. *Mol Oncol* 2016;10(2):253–71.
- [35] Bakry R, Vallant RM, Najam-ul-Haq M, Rainer M, Szabo Z, et al. Medicinal applications of fullerenes. *Int J Nanomedicine* 2007;2(4):639–49.
- [36] Prylutska S, Grynuk I, Matyshevska O, Prylutsky Y, Evstigneev M, et al. C60 fullerene as synergistic agent in tumor-inhibitory Doxorubicin treatment. *Drugs R D* 2014;14(4):333–40.
- [37] Ikeda A, Doi Y, Hashizume M, Kikuchi J-i, Konishi T. An extremely effective DNA photocleavage utilizing functionalized liposomes with a fullerene-enriched lipid bilayer. *J Am Chem Soc* 2007;129(14):4140–4141.
- [38] Joseph J, Sivasankarapillai VS, Nikazar S, Shanawaz MS, Rahdar A, et al. Borophene and Boron Fullerene Materials in Hydrogen Storage: Opportunities and Challenges. *ChemSusChem* 2020;13(15):3754–65.
- [39] Zheng T, Wan X, Zhang Q, Jin B, Peng R-F. Catechol amide derivatized polyhydroxylated fullerene as potential chelating agents of radionuclides: Synthesis, reactive oxygen species scavenging, and cytotoxic studies. *J Inorg Biochem* 2020;203:110921.
- [40] Xiaoyan Z, Hailin C, Bing Y, Qun C. Recent Advances of Water-Soluble Fullerene Derivatives in Biomedical Applications. *Mini Rev Org Chem* 2019;16(1):92–9.
- [41] Grebinyk A, Prylutska S, Grebinyk S, Prylutsky Y, Ritter U, et al. Complexation with C60 fullerene increases doxorubicin efficiency against leukemic cells in vitro. *Nanoscale Res Lett* 2019;14(1):61.
- [42] Grebinyk A, Prylutska S, Buchelnikov A, Tverdokhle N, Grebinyk S, et al. C60 fullerene as an effective nanoplatform of alkaloid berberine delivery into leukemic cells. *Pharmaceutics* 2019;11(11):586.
- [43] Sergeeva V, Kraevaya O, Ershova E, Kameneva L, Malinovskaya E, et al. Antioxidant properties of fullerene derivatives depend on their chemical structure: a study of two fullerene derivatives on HELFs. *Oxid Med Cell Longev* 2019;2019:4398695.
- [44] Eswaran SV. Water soluble nanocarbon materials: A panacea for all?. *Curr Sci* 2018;114:1846–50.
- [45] Vereshchaka IV, Bulgakova NV, Maznychenko AV, Gonchar OO, Prylutsky YI, et al., C(60) fullerenes diminish muscle fatigue in rats comparable to N-acetylcysteine or β -Alanine. *Front Physiol* 2018;9: 517–517..
- [46] Gonchar OO, Maznychenko AV, Bulgakova NV, Vereshchaka IV, Tomiak T, et al. C60 fullerene prevents restraint stress-induced oxidative disorders in rat tissues: possible involvement of the Nrf2/ARE-antioxidant pathway. *Oxid Med Cell Longev* 2018;2018.
- [47] Wong-Ekkabut J, Baoukina S, Triampo W, Tang IM, Tieleman DP, et al. Computer simulation study of fullerene translocation through lipid membranes. *Nat Nanotechnol* 2008;3(6):363–8.
- [48] Prylutsky Y, Bychko A, Sokolova V, Prylutska S, Evstigneev M, et al. Interaction of C60 fullerene complexed to doxorubicin with model bilipid membranes and its uptake by HeLa cells. *Mater Sci Engineer C* 2016;59:398–403.
- [49] Prylutska SV, Grebinyk AG, Lynchak OV, Byelinska IV, Cherepanov VV, et al. In vitro and in vivo toxicity of pristine C60 fullerene aqueous colloidal solution. *Fullerenes, Nanotubes, Carbon Nanostruct* 2019;27(9):715–28.
- [50] Schuetze C, Ritter U, Scharff P, Fernekorn U, Prylutska S, et al. Interaction of N-fluorescein-5-isothiocyanate pyrrolidine-C60 with a bimolecular lipid model membrane. *Mater Sci Eng, C* 2011;31(5):1148–50.
- [51] Kraevaya OA, Peregodov AS, Fedorova NE, Klimova RR, Godovikov IA, et al. Thiophene-based water-soluble fullerene derivatives as highly potent antiherpetic pharmaceuticals. *Org Biomol Chem* 2020;18(42):8702–8.
- [52] Hsieh F-Y, Zhilenkov AV, Voronov II, Khakina EA, Mischenko DV, et al. Water-soluble fullerene derivatives as brain medicine: surface chemistry determines if they are neuroprotective and antitumor. *ACS Appl Mater Interfaces* 2017;9(13):11482–92.
- [53] Kraevaya OA, Peregodov AS, Godovikov IA, Shchurik EV, Martynenko VM, et al. Direct arylation of C60Cl6 and C70Cl8 with carboxylic acids: a synthetic avenue to water-soluble fullerene derivatives with promising antiviral activity. *ChemComm* 2020;56(8):1179–82.
- [54] Huang H-J, Kraevaya OA, Voronov II, Troshin PA, Hsu S-h, Fullerene derivatives as lung cancer cell inhibitors: investigation of potential descriptors using QSAR approaches. *Int J Nanomedicine* 2020;15:2485.
- [55] Luo P-W, Han H-W, Yang C-S, Shrestha LK, Ariga K, et al. Optogenetic modulation and reprogramming of bacteriorhodopsin-transfected human fibroblasts on self-assembled fullerene C60 nanosheets. *Adv Biosyst* 2019;3(2):1800254.
- [56] Rogers D, Hopfinger AJ. Application of genetic function approximation to quantitative structure-activity relationships and quantitative structure-property relationships. *J Chem Inf Comput Sci* 1994;34(4):854–66.
- [57] Zhou S, Burger C, Chu B, Sawamura M, Nagahama N, et al. Spherical bilayer vesicles of fullerene-based surfactants in water: a laser light scattering study. *Science* 2001;291(5510):1944–7.
- [58] Wong C-W, Zhilenkov AV, Kraevaya OA, Mischenko DV, Troshin PA, et al. Toward understanding the antitumor effects of water-soluble fullerene derivatives on lung cancer cells: apoptosis or autophagy pathways?. *J Med Chem* 2019;62(15):7111–25.
- [59] *Hydrocarbon Ring Assemblies, Nomenclature of organic compounds, American chemical society* 1974, pp. 43–48..
- [60] Kazemzadeh H, Mozafari M. Fullerene-based delivery systems. *Drug Discov Today* 2019;24(3):898–905.
- [61] Zhu X, Xiao S, Zhou D, Sollogoub M, Zhang Y. Design, synthesis and biological evaluation of water-soluble per-O-methylated cyclodextrin-C60 conjugates as anti-influenza virus agents. *Eur J Med Chem* 2018;146:194–205.
- [62] Dzhemilev UM, Khuzin AA, Akhmetov AR, D'yakovon VA, Dzhemileva LU, et al. Synthesis of C60 fullerene-quadracyclane hybrid compound and Its preliminary in vitro antitumor activity in combination with cisplatin. *ACS Omega* 2019;4(14):15929–34.
- [63] Klimova R, Andreev S, Momotyuk E, Demidova N, Fedorova N, et al. Aqueous fullerene C60 solution suppresses herpes simplex virus and cytomegalovirus infections. *Fuller Nanotub Carbon Nanostructures* 2020;28(6):487–99.
- [64] Karacosta LG, Anchang B, Ignatiadis N, Kimmey SC, Benson JA, et al. Mapping lung cancer epithelial-mesenchymal transition states and trajectories with single-cell resolution. *Nat Commun* 2019;10(1):5587.
- [65] Kellogg RM. 3.13 - Thiophenes and their Benzo Derivatives: (i) Structure. In: Katritzky AR, Rees CW, editors. *Comprehensive Heterocyclic Chemistry*. Oxford: Pergamon; 1984. p. 713–40.
- [66] Zhao M, Cui Y, Zhao L, Zhu T, Lee RJ, et al. Thiophene Derivatives as New Anticancer Agents and Their Therapeutic Delivery Using Folate Receptor-Targeting Nanocarriers. *ACS Omega* 2019;4(5):8874–80.
- [67] Gramac D, Peterlin Mašič L, Sollner Dolenc M. Bioactivation Potential of Thiophene-Containing Drugs. *Chem Res Toxicol* 2014;27(8):1344–58.