PAPER

Check for updates

Design, synthesis, anticancer activity, molecular docking and ADME studies of novel methylsulfonyl indole-benzimidazoles in comparison with ethylsulfonyl counterparts†

Fikriye Zengin Karadayi,‡[a] Murat Yaman,‡[b] Mehmet Murat Kisla, [ID] ‡[a]
Ozlen Konu★[bcd] and Zeynep Ates-Alagoz [ID] ★[a]

Cancer poses a world-wide healthcare problem, demanding selective and effective therapy protocols. To address that, a vast amount of therapeutic candidates are being investigated in the field of medicinal chemistry. Accordingly, indole-benzimidazole structures have recently gained considerable interest because of their anticancer properties and estrogen receptor (ER) modulatory actions. In this study, novel methylsulfonyl indole-benzimidazole derivatives have been synthesized upon substitution of respectively the first (R_1) and fifth (R_2) positions of benzimidazole and indole groups. Structure and activity relationships were then studied *via* ^{1}H NMR, ^{13}C NMR, mass spectral and *in silico* docking analyses, as well as cell viability measurements. We found that the compounds exhibited substantial affinity levels towards ER alpha (ERα). In addition, the correlation analysis of cytotoxicity profiles between ethyl- and methyl-sulfonyl indole-benzimidazoles revealed a collection of effective and consistent R_1 and R_2 substitutions. However, for some candidate derivatives, distinctive cytotoxicity levels and varying viability *versus* ERα affinity correlations were observable across the studies, suggesting that the sulfonyl side chain modifications themselves can also influence the ERα binding levels. These results demonstrated that our novel methylsulfonyl indole-benzimidazole derivatives, similar to their ethylsulfonyl counterparts, exhibit anticancer effects with potential estrogen receptor modulatory actions.

1. Introduction

ER positivity in breast cancer (BC) patients is higher than 70%.[1,2] ERα's role on BC therapy arises from its potential for inducing growth of the tumour cells in the presence of estrogen which can be targeted against, using antiestrogens or selective ER modulators (SERMs).[3] ERα expression influences disease progression by regulating multiple signaling pathways while maintaining the luminal phenotype, and results in increased susceptibility of breast cancer to hormone therapy.[4–6] Hence, modulation of ER signaling has been at the core of many

[a] Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Ankara University, Ankara, Turkey. E-mail: zeynep.ates@pharmacy.ankara.edu.tr

[b] Interdisciplinary Program in Neuroscience, Bilkent University, 06800, Ankara, Turkey. E-mail: konu@fen.bilkent.edu.tr

[c] Department of Molecular Biology and Genetics, Bilkent University, 06800, Ankara, Turkey

[d] UNAM-Institute of Materials Science and Nanotechnology, Bilkent University, 06800, Ankara, Turkey

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1nj01019k

‡ Equal contributors.

investigations bringing up novel therapeutic options for BC.[1,2] In this regard, development of SERMs gained considerable attention by utilizing ERα model cell lines in preclinic studies.[7,8] For example, an ERα-positive luminal cell line, MCF-7, has become a prominent *in vitro* model in conducting preliminary evaluations for novel SERMs against BC.[9]

SERMs are chemical substances that disrupt the action of estrogen by interacting with ER in cells.[10] More than 70 SERMs have been discovered to this day, and they can be classified as triphenylethylenes, benzothiophenes, tetrahydronaphtalenes, benzopyrans and indoles. In addition to the widely used first generation SERM known as tamoxifen,[11,12] there are other second and third generation SERMs developed to reduce side effects associated with tamoxifen.[11,13] For instance, the affinity of third generation bazedoxifene is more selective towards the ERα rather than ERβ, hence its inhibitory effect progresses through ERα downregulation and cell cycle arrest.[11] In addition, bazedoxifene with palbociclib combination is a promising therapeutic option in clinics, and tested against metastatic BC at stage IV.[14–16] Therefore, bazedoxifene has become a valuable therapeutic in clinic, and served as an appropriate reference for the studies assessing ER modulation, such as in molecular

docking studies.[17,18] Additionally, indole and benzimidazole cores are also main structural components of some SERMs. Besides bazedoxifene, melatonin and KB9520 carry indole moieties and pose SERM effects.[3,19] Methyl or naphthyl attached benzimidazole structures also demonstrate SERM-like actions on breast cancer cell lines.[3,19–21]

In line with the above-mentioned findings, we have previously identified several ethylsulfonyl indole-benzimidazole derivatives with antiestrogenic properties.[22] Interestingly, antiproliferative effects were distinct based on ER status where ERα-positive cell lines MCF-7 and HEPG2 resulted in overlapping cell viability profiles across various compound concentrations in comparison with ERα-negative cell line MDA-MB-231. Structure–activity relationship (SAR) studies have suggested that compounds possessing R_1: p-fluorobenzyl group have higher anticancer activity and moderate structural affinity towards ERα. Our results demonstrated that novel ethylsulfonyl indole-benzimidazole derivatives have ER modulatory activities similar to SERMs, and also had in common with dioxin-mediated aryl hydrocarbon receptor (AhR) signaling and amino acid deprivation-mediated integrated stress response pathways.[22] In a previous study, sulfonylmethyl moiety has been also investigated for its anti-BC activity identifying a cytotoxic agent against the MCF-7 cell line.[23] In addition, anticancer activities of methylsulfonyl benzothiazole derivatives have also been shown on a cervical cancer cell line, HeLA.[24] These studies suggest that sulfonylmethyl groups could also modify effectivity of indole-benzimidazoles, Thus, the methylsulfonyl bearing moieties can serve as candidate scaffolds in providing novel alternatives. Accordingly, our group designed, synthesized and tested a series of indole-benzimidazoles possessing methylsulfonyl group, on MCF-7. Obtained cytotoxicity estimates were then compared with the viability scores derived from those possessing ethylsulfonyl structures which were previously studied.[22] The meta-analysis enabled us to investigate the consistencies in the anticancer effects caused by specific side group modifications across methyl-/ethylsulfonyl indole-benzimidazoles. Furthermore, in silico docking analyses were proceeded with evaluating ERα binding affinities to see if the actions of the methylsulfonyls can also relate with ERα. Our findings suggest that both ethyl- and methyl-sulfonyl groups are active and may share common modulatory roles depending on the R_1 and R_2 chain diversities as well as the status of the sulfonyl modification. This is the first meta-analysis approach that allows for one-to-one anticancer activity comparisons between closely related counterparts of indole-benzimidazoles. Moreover, the study provides novel leads to identification of which R_1 and R_2 moieties can be more effective on viability of luminal breast cancer cells.

2. Results and discussion
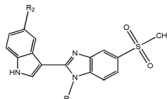
2.1. Biological activity

Anticancer activities of the derivatives were initially evaluated via MTT assays on MCF-7 treated at the 20 μM dose (Fig. 1); and relative cell viabilities as well as their significance were reported in Table 1 and Table S1 (ESI†). First, the assessments by one-way ANOVA/Dunnett showed that treatments with 28 out of 34 compounds (82%) have significantly reduced the MCF-7 viability (Table S1, ESI†). Besides, 19 of them resulted in lower cell growth than the overall mean viability score (66.47%), further
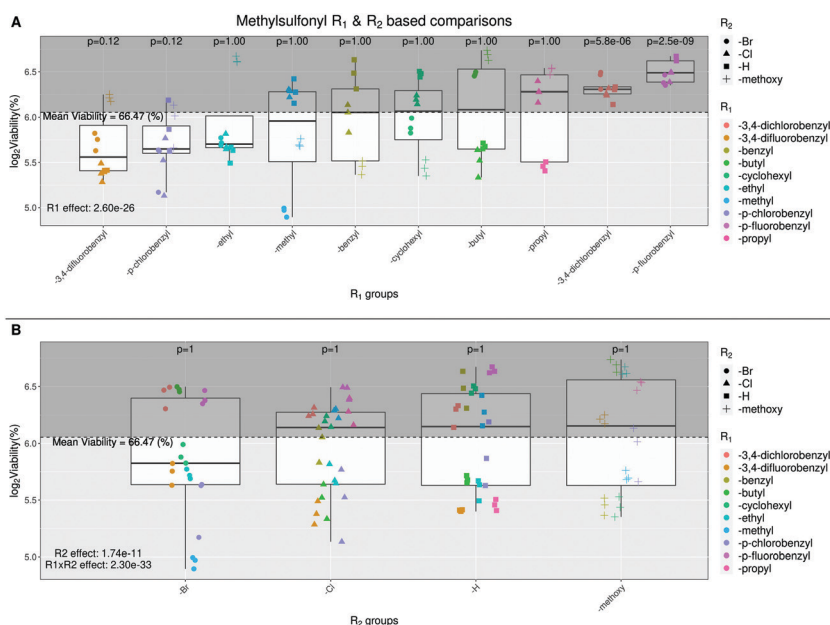
Fig. 1 (A) R_1 and (B) R_2 based cell viability comparisons for methylsulfonyl indole–benzimidazole derivatives. Log 2 transformed percent viability scores from each triplicate were given on the y-axis. R_1, R_2 and $R_1 \times R_2$ interaction effects were calculated via two-way ANOVA, as within group comparisons for either R_1 or R_2 moieties were performed via multiple t-tests with Holm corrections overall. R environment, ggplot2 and ggpubr packages were utilized for calculation and representation purposes.

Table 1 Percentages of relative cell viabilities and adjusted p-values (one-way ANOVA/Dunnett) obtained from 20 μM concentration of indole-benzimidazole derivatives containing methylsulfonyl. DMSO was used as the control group


Compound	R_1	R_2	%Viability	Adj. p-val.	Compound	R_1	R_2	%Viability	Adj. p-val.
23	$-CH_3$	$-H$	78.16	0.0013	**41**	cyclohexyl	$-Br$	59.72	<0.0001
24	$-CH_3$	$-OCH_3$	52.42	<0.0001	**42**	benzyl	$-H$	89.39	0.3467
25	$-CH_3$	$-Cl$	77.29	0.0008	**43**	benzyl	$-OCH_3$	43.64	<0.0001
26	$-CH_3$	$-Br$	31.00	<0.0001	**44**	benzyl	$-Cl$	64.56	0.0003
27	$-C_2H_5$	$-H$	48.54	<0.0001	**45**	4-fluorobenzyl	$-H$	99.91	>0.9999
28	$-C_2H_5$	$-OCH_3$	99.23	>0.9999	**46**	4-fluorobenzyl	$-Cl$	87.84	0.0148
29	$-C_2H_5$	$-Cl$	52.49	<0.0001	**47**	4-fluorobenzyl	$-Br$	84.36	0.0014
30	$-C_2H_5$	$-Br$	52.95	<0.0001	**48**	3,4-difluorobenzyl	$-H$	42.46	<0.0001
31	$-C_3H_7$	$-H$	43.93	<0.0001	**49**	3,4-difluorobenzyl	$-OCH_3$	74.21	<0.0001
32	$-C_3H_7$	$-OCH_3$	91.37	0.9621	**50**	3,4-difluorobenzyl	$-Cl$	41.82	<0.0001
33	$-C_3H_7$	$-Cl$	77.87	0.0702	**51**	3,4-difluorobenzyl	$-Br$	53.38	<0.0001
34	$-C_4H_9$	$-H$	51.32	<0.0001	**52**	4-chlorobenzyl	$-H$	60.26	<0.0001
35	$-C_4H_9$	$-OCH_3$	102.94	0.9991	**53**	4-chlorobenzyl	$-OCH_3$	61.84	<0.0001
36	$-C_4H_9$	$-Cl$	45.34	<0.0001	**54**	4-chlorobenzyl	$-Cl$	45.18	<0.0001
37	$-C_4H_9$	$-Br$	88.93	0.8268	**55**	4-chlorobenzyl	$-Br$	45.07	<0.0001
38	cyclohexyl	$-H$	88.99	0.8307	**56**	3,4-dichlorobenzyl	$-H$	76.72	<0.0001
39	cyclohexyl	$-OCH_3$	43.43	<0.0001	**57**	3,4-dichlorobenzyl	$-Cl$	77.28	<0.0001
40	cyclohexyl	$-Cl$	73.22	0.0139	**58**	3,4-dichlorobenzyl	$-Br$	85.90	0.0038

supporting that the substitutions caused significant cytotoxic effects. In addition, the results from multiple t-test/Holm and two-way ANOVA revealed the significant effects of different substitutions at the R_1 or R_2 side chains. Accordingly, R_1:-3,4-difluorobenzyl and -p-chlorobenzyl additions demonstrated a trend for exhibiting more sensitive profiles overall (Fig. 1A). Moreover, the number of –Cl and that of –F on the benzyl groups by the R_1 chain seemed to have an inverse relationship with respect to each other but acting oppositely for –Cl and –F. For instance R_1:-p-fluorobenzyl yielded the least anti-cancer activity across the screened derivatives (adj. p-value: 2.5×10^{-9}) while R_1:-p-difluorobenzyl was one of the most effective R_1 substitutions, approaching significance (adj. p-value: 0.12). Similarly, R_1:-3,4-dichlorobenzyl and R_1:-3,4-chlorobenzyl groups were located at the opposite ends of the viability spectrum (Fig. 1A). Moreover, the inclusion of –F by the third position of the R_1:-p-fluorobenzyl chain has led to an increase in cytotoxicity remarkably. In contrast, an addition of –Cl to the R_1:-p-chlorobenzyl has demonstrated a decrease in the anti-cancer activity. Although R_2-based multiple pairwise tests did not reveal noticeable differences, the results of ANOVA showed a significant influence of the R_2 groups on the cytotoxicity levels (Fig. 1B). Nonetheless, we also observed a significant inter-action between the R_1 and R_2 groups implicating that each R_2 group might influence cell viability dependent on the R_1 and *vice versa*. Therefore, this significant interaction term might actually stem from the differences generated by differential R_1:R_2 stereochemistry. We also noted that the median cytotoxicity scores of R_2:-Br bearing compounds were relatively lower than the others, suggesting improvements on cytotoxicity by the R_2:-Br side chains. In line with this, the compound **26** (R_1:-methyl & R_2:-Br) has shown the highest activity over all the methylsulfonyl derivatives screened in this study (Table 1). Judging by the cell viability values from this preliminary screening, we focused on this lead molecule, compound **26** for which we calculated the IC_{50} value (28.73 μM) (Fig. 3).
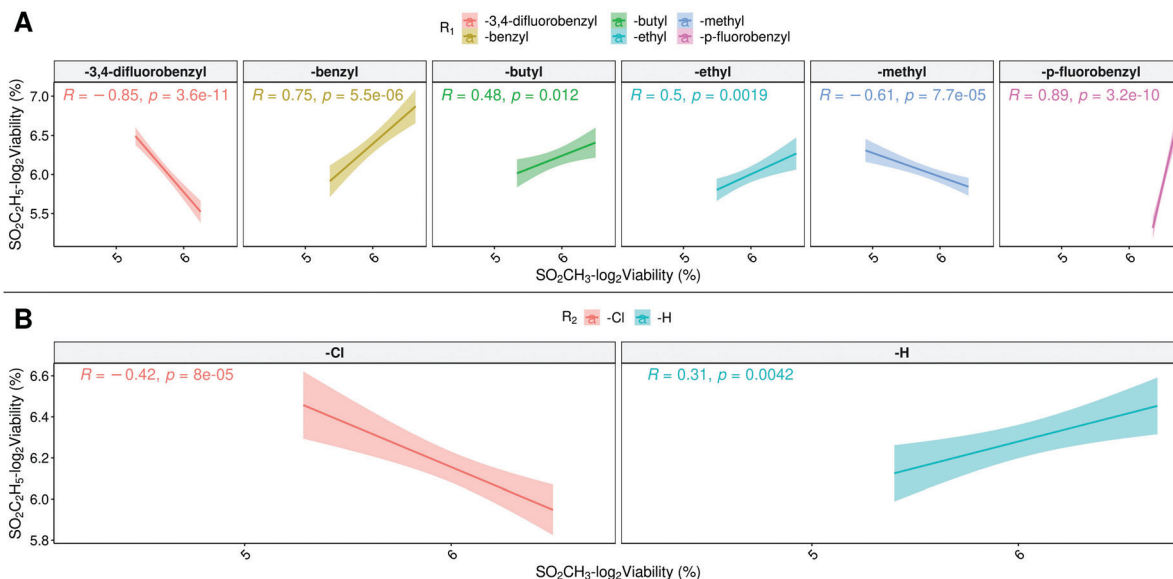
Fig. 2 Significant correlations observed for cellular viability between methylsulfonyl (20 μM) (x-axes) and ethylsulfonyl (16 μM) (y-axes) of indole-benzimidazoles based on 24 h treatment of MCF-7 cells. (A) R$_1$ and (B) R$_2$. Pearson's correlations were performed on the viability scores from the biological triplicates where the lines and shadow regions represent the linear correlations and confidence intervals, respectively. Correlation plots of non-significant groups (R$_1$:-3,4-dichlorobenzyl, -cyclohexyl and -propyl; R$_2$:-Br and -methoxy) are available in Fig. S1 (ESI†).
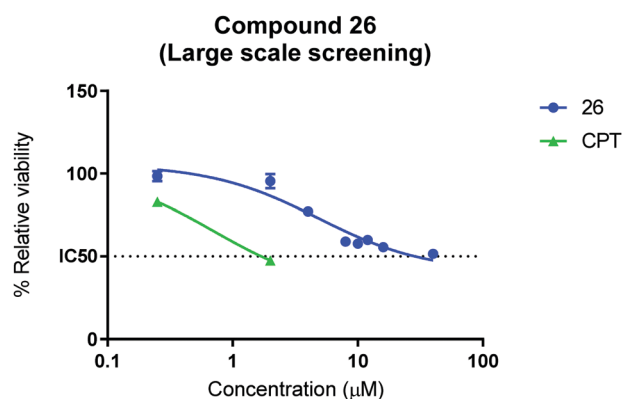
Fig. 3 Detailed MTT screening of eight concentration points for the lead molecule **26** on MCF-7 cell line for 24 h, across doses of 0.25 μM, 2 μM, 4 μM, 8 μM, 10 μM, 12 μM, 16 μM and 40 μM. CPT has served as the positive control.

Additionally, in our study we compared for the first time the viability results obtained from exposures to ethylsulfonyl derivatives (16 μM) with those of methysulfonyl derivatives (20 μM). This allowed us to test whether similar modifications exhibited similar effects when they were associated with ethyl- or methyl-sulphonyl carrying compounds. The cell viability scores between different sulfonyl moieties revealed both similar as well as inverse profiles depending on the R$_1$ and/or R$_2$ (Fig. 2 and Fig. S1, ESI†). For example R$_1$:-benzyl, -butyl, -ethyl, -p-fluorobenzyl in addition to R$_2$:-H demonstrated significantly positive correlations across the two sulfonyl moieties while R$_1$:-methyl and -3,4-difluorobenzyl additions and R$_2$:–Cl modifications were inversely correlated between the groups (Fig. 2A). Furthermore, some of the substitutions (R$_1$:-cyclohexyl and -propyl; R$_2$:-Br and -methoxy) showed no correlations, in line

with overall comparisons between the two different sulfonyl scaffolds (Fig. S1, ESI†). Similarly ethyl- and methylsulfonyl derivates were negatively or positively correlated respectively for –Cl and –H while other R$_2$ groups did not show any significant association (Fig. 2B and Fig. S1, ESI†). Therefore, based on which R$_1$ and/or R$_2$ group present, ethyl or methyl status of the sulfonyl groups were found to influence the viability levels differentially.

2.2. Molecular docking

According to the MTT assay results, compound **26** was elected as the lead candidate, and compared to the standard **CPT**. According to the complex of **CPT** with ERα ligand binding domain and reference study,[25] this area includes various hydrophobic residues (Fig. 4A). Therefore, incorporating hydrophobic groups such as indole and benzimidazole becomes essential for modulating the signalling mechanisms of this protein. According to our previous docking analyses with ERα, a hydrogen bond with some polar residues in this cavity could become important.[22] In hindsight, interaction diagram of **CPT** offers certain hydrophobic interactions with Leu387, Leu346, Ala350, Met343 and Leu343 residues. Amongst them, Met343 was previously addressed in publications. Through the examination of **26**'s interaction diagram (Fig. 4B), in addition to several steric interactions such as the ones with Leu354, Trp383 and Ala350, a hydrogen bond interaction by sulfonyl group is present. This binding profile increased the chance of **26** in being a prominent inhibitor of ERα.

The relevance of ERα binding and anticancer activity of the indole-benzimidazoles has been a crucial element in guiding our meta-analysis of the two types of sulfonyl groups. Accordingly for the methylsulfonyl derivatives, ERα docking scores

(A)

(B)

(C)
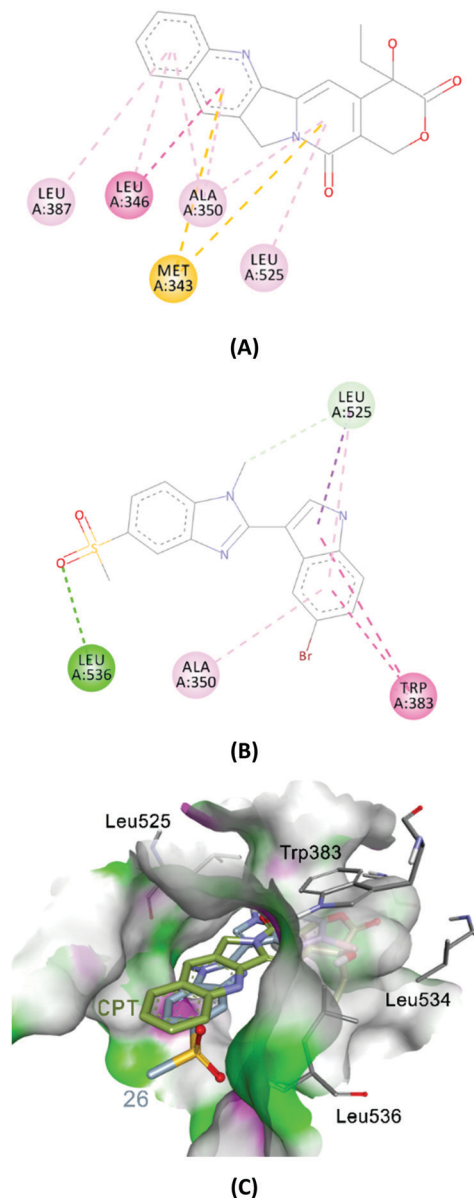
Fig. 4 The interaction patterns of reference **CPT** (A) and inhibitor **26** (B). In this figure, light pink corresponds to Pi-alkyl and alkyl–alkyl interactions whereas dark pink stands for Pi-Pi stacked interaction. Green is for the conventional H-bond interaction and light blue is for van der Waals interaction. Finally, the orange is for Pi-sulfur interaction. In panel (C), the aligned binding pose of these ligands in the cavity and interacting residues were given. Purple areas are for H-bond donor surfaces whilst green areas are for acceptor surfaces.

were further utilized in evaluating possible correlations they might exhibit with the percent viability scores (Fig. 5 and Fig. S2, ESI†). We have seen that docking scores of R_1:-3,4-difluorobenzyl and -3,4-dichlorobenzyl groups were positively correlated with the viability, as R_1:-ethyl moieties were also on a similar trend (Fig. 5A). Considering prominent effects by the R_1:-3,4-difluorobenzyl and -ethyl groups, the findings have implied a close relationship between ERα and MCF-7 toxicity. Findings on R_1:-3,4-dichlorobenzyl members supported the functional interaction for ERα and toxicity events, which was

also observable across the ethylsulfonyl counterparts (Fig. S2A, ESI†). Interestingly, R_1:-benzyl followed an inverse correlation regarding the viability *versus* docking score comparisons, as the remaining R_1 derivatives did not show a clear association in this sense. Thus, ERα modulatory activities of the methylsulfonyl compounds can differ depending on the side chain modifications. Similarly, we have seen that R_2: methoxy and -H substitutions can closely associate with ERα docking scores and MCF-7 viabilities while R_2:-Br and –Cl additions did not (Fig. 5B).

In parallel, we have employed a similar approach with ethylsulfonyl derivatives (Fig. S3, ESI†). By the R_1 side chain, -3,4-dichlorobenzyl groups again presented a positive relationship between docking scores and cytotoxicity levels. In addition, R_1: -butyl and –methyl members and R_2: -H groups indicated overlapping patterns with their methylsulfonyl counterparts when comparing their viability – ERα docking associations. Moreover, R_2:-Br and –Cl substitutions yielded significant effects on ERα binding and cell viability score comparisons, which were other distinctive features between ethyl/methylsulfonyl derivatives. Therefore, the findings have further implemented that ERα binding and cellular viability related events were likely to be restricted to some compound groups only.

2.3. Pharmacokinetic evaluation

As provided in the Table S2 (ESI†), corresponding inhibitor **CPT** has passed the pharmacokinetic filters such as Lipinski, Muegge and Leadlikeness, whereas **BZ** failed to do so. This outcome suggests that the former compound may be formulated into a drug dosage form and one shall encounter less bioavailability problems comparing to that of the latter. Compounds such as **46**, **47**, **51**, and **54–58** have given harshly unfavorable results since they have got stuck on these three filters. Consequently, when these molecules get formulated, certain pharmacokinetic problems such as low bioavailability, biotoxicity and higher incidence of side effects may occur. Regarding binding affinity values, all the compounds have had similar values comparing to that of **CPT**. All of the compounds including **CPT** have had the same amount of H-bond donor count, whereas H-bond acceptor values were drastically different. Interestingly, despite having less amount of donor count than **CPT**, compound **26** offered an additional H-bond. This finding may have possibly increased the activity of this potent compound.

BOILED-Egg representation is a diagram that demonstrates the bioavailability of the input that is used in SwissADME program. Using this diagram, one could potentially get an insight about central nervous system penetration and *P*-glycoprotein binding. In summary, all of our compounds tend to get absorbed through the gastrointestinal tract, meaning they get blood concentration buff. Unfortunately, some of them got stuck in BBB reabsorption hurdle. In this passage, the main concern will be the comparison of these compounds with **CPT** and **BZ**. Latter is a known ERα inhibitor, thus it should also be discussed in ADME evaluation. Herein we witness that **CPT** and **BZ** would get passively reabsorbed but would not pass
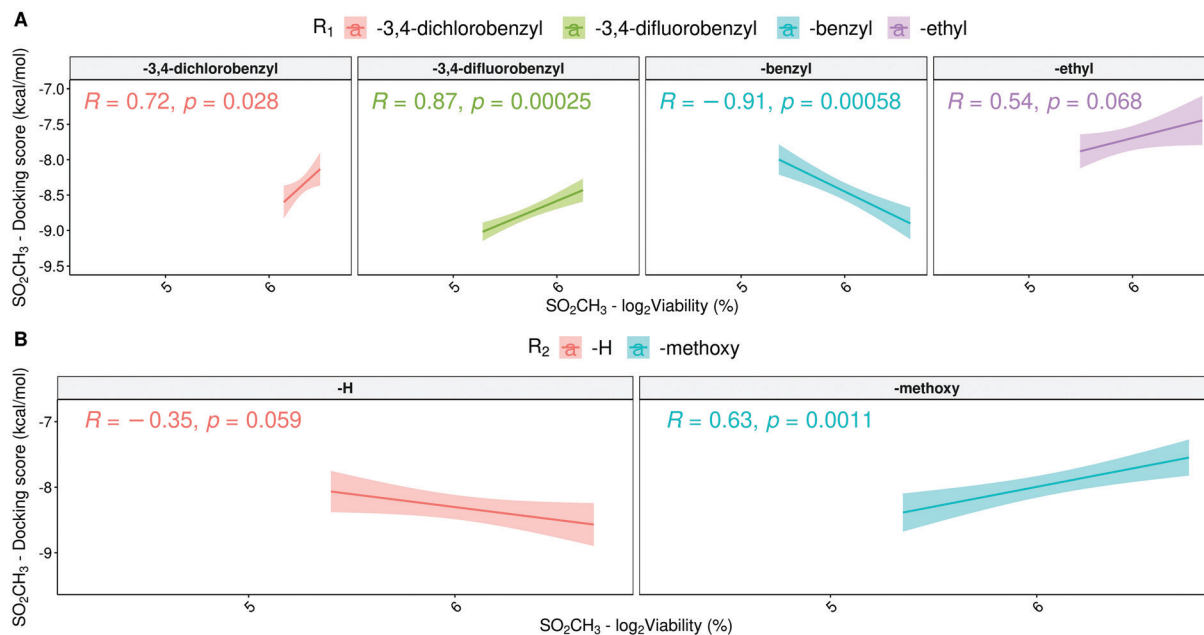
Fig. 5 Correlation maps between *in silico* docking scores (kcal mol^{-1}) (*y*-axes) and *in vitro* log 2 transformed viability percentiles (*x*-axes) obtained from methylsulfonyl derivative exposures on MCF-7 cell line in triplicates: (A) R$_1$ and (B) R$_2$. Each R$_1$ and R$_2$ groups are color coded along with the confidence intervals, as Pearson's correlation and significance scores are depicted inside the plots with the respective color codes for each group. Only the plots approaching significance are represented. Remaining plots are accessible in Fig. S2 (ESI†).

through BBB. This characteristic is ideal in terms of bioavailability and safety profile since reabsorption increases blood concentration of the drugs, and permeability through BBB causes certain CNS side effects which must be avoided. In this context, both **CPT** and **BZ** become ideal drugs. However, compounds **23, 25, 26, 27, 29, 31,** and **42** failed to ensure this characteristic since they can passively permeate, causing a risk for CNS side effects (Fig. 6).
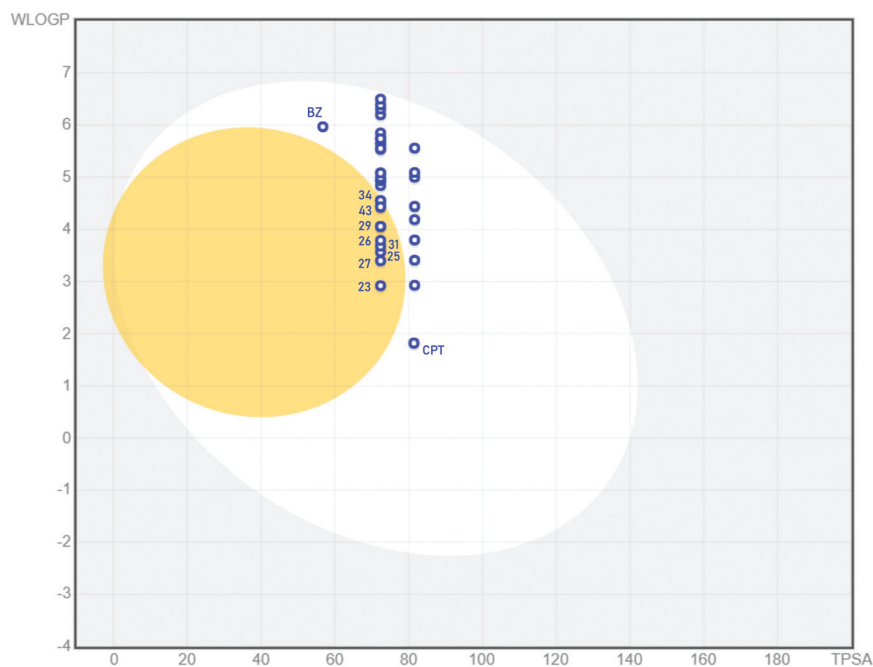

Fig. 6 BOILED-Egg representation of indole-benzimidazole derivatives and standards CPT (Camptothecin) and BZ (Bazedoxifene). In this diagram, *X*-axis is for Topological Polar Surface Area (TPSA) which indicates the polarity of compounds while the *Y*-axis is for the computational log *P* values. The white area is for molecules to be absorbed *via* the gastrointestinal tract whereas Egg yolk articulates compounds that are passively permeable through the blood–brain barrier.

3. Experimental part

3.1. Synthetic procedures

Synthesis of new compounds **23–58** were outlined in Scheme 1. Synthesis of 1-chloro-4-(methylsulfonyl)-2-nitrobenzene (**2**) was carried out with the nitration of 1-chloro-4-(methylsulfonyl)-benzene (**1**) on second position using potassium nitrate and sulfuric acid.[22,26] Aromatic nucleophilic substitution of the chlorine atom of **2** with appropriated amine to provide the corresponding 1-substituted-amino-4-(methylsulfonyl)-2-nitrobenzene **3–12** in good yields.[27] Reduction of the nitro group yielded **13–22** (Table 2).[28]

3.1.1. General procedure for synthesis of 23–58. A mixture of the appropriate *o*-phenylendiamine (1 mmol), indole-3-carboxaldehyde (1 mmol) and $Na_2S_2O_5$ (40%) (2 mL) in EtOH (4 mL) was refluxed until consumption of the starting material (determined by TLC, 4–12 h). The reaction mixture was poured into water, and the precipitate was filtered and washed with water. The residue was purified by column chromatography and crystallized with MeOH to give the final product (Scheme 1 and Table S3, ESI†).[27]

3.2. Biological activity assays

3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) (Molecular Probes) assay was utilized to obtain cell viability as percentages of the control DMSO (0.1%) samples. MCF-7 cells were seeded (96-well plate & 10 000 cells per well) and grown for 24 hours in replicates per group. Compounds listed in Table 1 (20 μM) were applied on cells for another 24 h. Campthotecin (CPT) was used as the positive control (at the concentrations of 0.25 μM and 2 μM). In addition to the large scale screening a candidate compound (**26**) was screened on a finer scale (0.25 μM, 2 μM, 4 μM, 8 μM, 10 μM, 12 μM, 16 μM and 40 μM). The cell growth media contained DMEM-low-glucose (phenol-free, GIBCO) supplemented with 10% FBS (Biowest), 1% penicillin–streptomycin (LONZA), 1% non-essential amino acid (LONZA), 2% L-glutamine (LONZA) and 1% Na-pyruvate (LONZA). Using BIO-TEK/μQuant Universal Microplate Spectrophotometer (BIO-TEK/KC junior tool

(v.1.418)), we obtained OD_{570} measurements and converted them to cell viability, by dividing blank subtracted average OD_{570} values for each concentration with those of DMSO control group OD_{570} averages. After multiplication with 100, the percentage viability estimations were analyzed *via* statistical approaches.

## 3.3. Material and methods

Melting points were determined and uncorrected using Buchi SMP-20 (BuchiLabortechnik, Flawil, Switzerland) and Electro-thermal 9100 capillary melting point apparatus (Electrother-mal, Essex, U.K.). The ^1H NMR spectra in DMSO-d_6 using Varian Mercury-400 FT-NMR spectrometer (Varian Inc., Palo Alto, CA, USA), and the Mass spectra based on ESI(+) method using Waters ZQ micromass LC-MS spectrometer (Waters Cor-poration, Milford, MA, USA) were measured. For elemental analysis and column chromatography (cc), respectively, LECO 932 CHNS (Leco-932, St. Joseph, MI, USA) instrument and Silica gel 60 (40–63 mm particle size) were used. 1-Chloro-4-(methyl-sulfonyl)benzene and indole-3-carboxaldehyde were obtained commercially. The 5-substituted-indole-3-carboxaldehyde deri-vatives were obtained by direct formylation of 5-substituted-indole with dimethylformamide, using phosphorous oxychloride as a catalyst.[29]

4. Computational methods

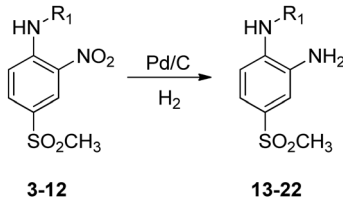## 4.1. Docking method and ADME property calculation

ERα (PDB ID:3ert, resolution: 2.5 Å) file was obtained from the RCSB Protein database website.[30] AutoDockTools 1.5.6. was used for deleting water molecules and defining the grid box.[31] After this process, polar hydrogens and Gasteiger charges were added and the grid was also prepared using the same software. Assigned grid's center was found at $X = 31.583$, $Y = -1.8$, $Z = 24.369$ and dimensions were $X = 40$, $Y = 40$, $Z = 48$. Spacing was defined as 0.375 Å. The 2D structures of the compounds were drawn on ChemDraw Ultra 12.0, minimized with MMFF94 and UFF forcefields (number of steps: 5000 with steepest descent algorithm and convergence value of 10×10^{-7}) to achieve the lowest energy conformer of the compounds. This conformer would also facilitate the occurrence of a ligand-protein complex with the lowest energy. After the minimization process, these files were converted to pdb files using Avogadro software.[32] Subsequently, Gasteiger charges and torsion were added to ligand files with AutoDockTools. Prepared ligands were docked with Autodock Vina,[33] interaction diagrams were created and interpreted using Discovery Studio Visualizer Ligand interaction module.[34]

The evaluation of physicochemical properties and the pre-diction of ADME parameters were determined with SwissADME online tools.[35] Besides, using BOILED-Egg representation,[36] brain penetration and gastrointestinal absorption of synthesized compounds were assessed. With the aid of the above-mentioned calculations, we aimed to gather medicinal chemistry information
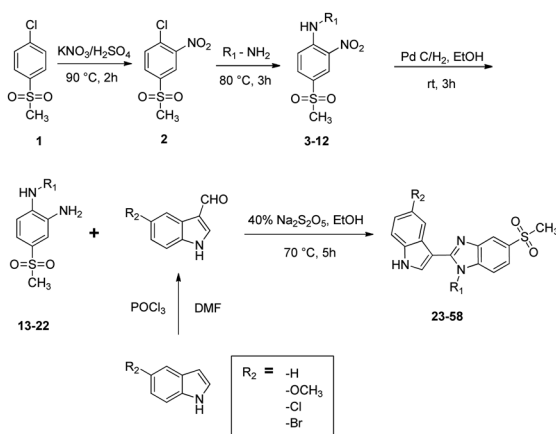
Scheme 1 Synthesis procedure of novel indole-benzimidazole deriva-tives **23–58**.

Table 2 Physicochemical data for compounds **3–22**

Comp.	R₁	^{1}H NMR	M.p. (°C)	Yield %	Comp.	Mass
3	–CH₃	3.00 (d, 3H, NH–C<u>H</u>₃), 3.17 (s, 3H, CH₃), 7.15 (d, J = 9.2 Hz, 1H, CH-6), 7.90 (dd, J = 9.2 Hz, J = 2.4 Hz, 1H, CH-5), 8.47 (d, J = 2.0 Hz, 1H, CH-3), 8.64 (brd d, J = 4.8 Hz, 1H, NH)	157	86	**13**	200
4	–C₂H₅	1.23 (t, 3H, CH₂–C<u>H</u>₃), 3.21 (s, 3H, CH₃), 3.45–3.52 (m, 2H, CH₂), 7.25 (d, J = 9.2 Hz, 1H, CH-6), 7.92 (dd, J = 9.2 Hz, J = 2.0 Hz, 1H, CH-5), 8.50 (d, J = 2.4 Hz, 1H, CH-3), 8.60 (brd t, 1H, NH)	164	83	**14**	215
5	–C₃H₇	0,92 (t, 3H, CH₂–C<u>H</u>₃), 1.59–1.64 (m,2H, C<u>H</u>₂–CH₃), 3.17 (s. 3H, CH₃), 3.36–3.41 (m, 2H, NH–C<u>H</u>₂), 7.24 (d, J = 9.2 Hz, 1H, CH-6), 7.88 (dd, J = 9.0 Hz, J = 2,4 Hz, 1H, CH-5), 8.47 (d, J = 2.4 Hz, 1H, CH-3), 8.6 (brd t, 1H, NH)	116	81	**15**	228
6	–C₄H₉	0.92 (t, 3H, CH₂–C<u>H</u>₃), 1.35- 1.41 (m, 2H, C<u>H</u>₂–CH₃), 1.57–1.64 (m, 2H, C<u>H</u>₂–CH₂), 3.20 (s, 3H, CH₃), 3.44 (q, 2H, NH–C<u>H</u>₂), 7.26 (d, J = 9.2 Hz, 1H, CH-6), 7.91 (dd, J = 9.4 Hz, J = 2.4 Hz, 1H, CH-5), 8.50 (d, J = 2.4 Hz, 1H, CH-3), 8.61 (brd t, 1H, NH)	122	87	**16**	243
7	-Cyclohexyl	1.24 (m, 1H, cyclohexyl CH), 1.41 (m, 4H, cyclohexyl), 1.56–70 (m, 3H, cyclohexyl), 1.93 (m, 2H, cyclohexyl CH₂), 3.18 (s, 3H, CH₃), 3.73 (m, 1H, cyclohexyl CH), 7.31 (d, J = 9.2 Hz, 1H, CH-6), 7.88 (dd, J = 8.6 Hz, J = 2.4 Hz, 1H, CH-5), 8.30 (brd d, 1H, NH), 8.45 (d, J = 2.4 Hz, 1H, CH-3)	124	85	**17**	269
8	-Benzyl	3.17 (s, 3H, CH₃), 4.72 (d, 2H, CH₂), 7.09 (d, J = 9.6 Hz, 1H, CH-6), 7.24–7.39 (m, 5H, CH-2',3',4',5',6'), 7.84 (dd, J = 9.2 Hz, J = 2.0 Hz, 1H, CH-5), 8.52 (d, J = 2.0 Hz, 1H, CH-3), 9.18 (t, 1H, NH)	132	85	**18**	277
9	-p-Fluorobenzyl	3.18 (s, 3H, CH₃), 4.71 (d, 2H, CH₂), 7.09 (d, J = 9.6 Hz, 1H, CH-6), 7.15–7.20 (m, 2H), 7.44 (m, 2H), 7.85 (dd, J = 9.6 Hz, J = 2.4 Hz, 1H, CH-5), 8.52 (d, J = 2.0 Hz, 1H, CH-3), 9.18 (t,1H, NH)	144	82	**19**	295
10	-3,4-Difluorobenzyl	3.16 (s, 3H, CH₃), 4.68 (d, 2H, CH₂), 7.05 (d, J = 8.8 Hz, 1H, CH-6), 7.21–7.24 (m, 1H), 7.35–7.49 (m, 2H), 7.83 (dd, J = 9.2 Hz, J = 2.4 Hz, 1H, CH-5), 8.49 (d, J = 2.4 Hz, 1H, CH-3), 9.16 (t,1H, NH)	128	80	**20**	313
11	-p-Chlorobenzyl	3.18 (s, 3H, CH₃), 4.71 (d, 2H, CH₂), 7.05 (d, J = 9.2 Hz, 1H, CH-6), 7.41 (s, 4H, CH-2',3',5',6'), 7.84 (dd, J = 8.8 Hz, J = 2.0 Hz, 1H, CH-5), 8.52 (d, J = 1.6 Hz, 1H, CH-3), 9.19 (t,1H, NH)	181	62	**21**	311
12	-3,4-Dichlorobenzyl	3.19 (s, 3H, CH₃), 4.73 (d, 2H, CH₂), 7.07 (d, J = 9.2 Hz, 1H, CH-6), 7.39 (dd, J = 8.0, Hz, J = 2.0 Hz, 1H), 7.60 (d, J = 8.0 Hz, 1H), 7.70 (d, J = 2.0 Hz, 1H), 7.86 (dd, J = 9.0 Hz, J = 2.4 Hz, 1H, CH-5), 8.53 (d, J = 2.4 Hz, 1H, CH-3), 9.21 (t,1H, NH).	129	77	**22**	345

about our compound set and to suggest further solutions for the improvement of these properties.

4.2. Statistical analyses

Percent viability value from each replicate was log 2 transformed. Initially, viability scores from 20 μM treatments in triplicates were compared to DMSO control group via one-way ANOVA followed by Dunnett tests (GraphPad Prism (v. 6.01)). Then, effects of the treatments with respect to their respective R_1 and R_2 groups were evaluated by using *ggplot2* and *ggpubr* packages in R statistical environment (3.6.1), while within group comparisons (R_1 or R_2) were employed via multiple *t*-test/Holm statistics.[37,38] The analyses were further supported with two-way ANOVA approaches in R. In addition, IC_{50} value of the compound **26** was estimated via GraphPad Prism tool with "log(inhibitor) vs. response (three parameters)" option. Moreover, *ggpubr* package was utilized for drawing the scatter plots for Pearson's correlations across viability values for each side-chain and ERα docking scores.[38] Meta-analysis was pursued with the data all of which were retrieved from our recent study.[22] Accordingly, (i) cell viability values of MCF-7 cells

exposed to 37 ethylsulfonyl indole-benzimidazole derivatives for 24 h (16 μM) as well as (ii) the ERα docking scores of the respective compounds from the Glide docking software were obtained.[39] Retrieved non-versatile docking scores from ethylsulfonyl data were capped to zero for making comparisons and representations possible.

5. Conclusions

Most of the novel derivatives yielded significant cytotoxic effects, which were likely to be above 20 μM IC_{50} values. However, the findings also revealed some candidate compounds and a subset of R_1 and R_2 modifications that can serve for lead molecule development. For instance, the compounds **26, 27, 48** and **50,** and R_1:-3,4-difluorobenzyl, -p-chlorobenzyl and ethyl groups as well as R_2:-Br additions can become useful for further improvements.

Results of the docking studies suggested that methylsulfonyl indole-benzimidazole derivatives have rendered similar affinity values and additional interactions in docking analysis, comparing to that of the biological activity standard **CPT**. Amongst

these derivatives, compound **26** which ensures certain interactions with the enzyme has shown some promise as a potential ERα inhibitor although with a potential risk of passing through BBB. Pharmacokinetic parameters were also calculated to evaluate whether our compounds lack certain requirements for their pharmaceutical development, or if they need any certain pharmacokinetic improvements. Major obstacles were observed to may involve high $\log P$ and molecular weight related issues. In practice, adjusting the molecular weight may not be an option, yet modifying the partition coefficient ($\log P$) can be pursued by several strategies.[40,41] As a result, both *in vitro* and *in silico* docking studies have provided crucial information in demonstrating potentials of novel antiproliferative agents and their possible interactions with ERα in support of the previous findings with indole-benzimidazoles bearing ethylsulfonyl moieties.[22]

In addition to unlikely pharmacokinetic properties, generalizing SERM characteristics for the compounds has become a challenging task, due to varying interrelationships between ERα binding and limited cellular toxicity levels. Since some of the side chain modifications also revealed significant effects on the cell viability percentages, the findings were in line with the knowledge that modulation status of ERα may influence the cancer cell viability. Yet, functional studies are required to address this issue with the lead compounds, across multiple cell lines with varying ER genotypes.

Initial SAR studies suggested that substituents with steric hindrance played a pivotal role for both sulfonylethyl and sulfonylmethyl derivatives, since our previous findings suggested that 3,4-difluorobenzyl group has led to increased inhibition. In a similar fashion, in this study we have concluded that lipophilic and bulky side chains such as 3,4-difluorobenzyl, *p*-chlorobenzyl and –Br greatly increase the inhibition rate, thus antiestrogenic activity. Likewise, a pioneering study corporating indole and benzimidazole rings by Singla *et al.* suggested that a bulky and/or lipophilic side chain is crucial in interactions with ERα ligand binding domain since this cavity primarily includes hydrophobic residues. They have found that other lipophilic substituents such as bromine and fluorine serve as bioisostere for –CF₃, which is another lipophilic group that leads to a decreased pi-electron cloud in the structure.[16,42]

Besides, not all the derivatives presented either positive or inverse relationship. First of all, *in silico* ERα binding levels can not completely reveal the activity state of the receptor, demanding functional analyses to be followed with. Nonetheless, the action mechanisms of all the derivatives may not be strictly ERα dependent *per se*. Therefore, indirect relationships between ERα and cellular viability should also be considered. In our previous study, we have observed this paradigm with the ethylsulfonyl-based indole-benzimidazoles. The candidates were able to modulate AhR and amino acid depletion pathways in addition to estrogen signaling where cell-cycle and stress pathways were strongly affected. Interestingly, R₁:-*p*-fluorobenzyl group was the most effective compound family there, unlike what we observed with the

methysulfonyl counterpart. Therefore, status of sulfonyl side on the indole-benzimidazole scaffold has been shown to play an important part in directing molecular and cellular effects in addition to the R₁ and R₂ positions. Yet, in line with the literature findings on indole-benzimidazoles, such comparative analyses on ERα binding and viability estimations can be a powerful approach in pinpointing candidate SERM derivatives for further studies.[43] Moreover, AutoDock and Maestro tools were utilized across the methylsulfonyl and ethylsulfonyl docking analyses. In this regard, *in silico* validations on the docking scores can be followed up by employing both platforms interchangeably and by deriving the docking correlations between each other.[44]

In conclusion, our findings demonstrate the prominent side chain modifications that are informative for ERα binding abilities of the derivatives and likely viability intervals upon exposure to these novel derivatives. This knowledge along with the literature on indole-benzimidazoles libraries can further pave the way for developing stricter SERMs with improved anti-breast cancer profiles. In this course, lead structures demand thorough *in vitro* and *in vivo* examinations for their modulatory activities on ERα and possible interactions with multiple targets.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

References

1 M. Xue, K. Zhang, K. Mu, J. Xu, H. Yang, Y. Liu, B. Wang, Z. Wang, Z. Li, Q. Kong, X. Li, H. Wang, J. Zhu and T. Zhuang, *Oncogenesis*, 2019, **8**, 30.

2 R. Siersbæk, S. Kumar and J. S. Carroll, *Genes Dev.*, 2018, **32**(17–18), 1141–1154.

3 I. Paterni, C. Granchi, J. A. Katzenellenbogen and F. Minutolo, *Steroids*, 2014, **90**, 13–29.

4 W. Tian, F. Teng, J. Gao, C. Gao, G. Liu, Y. Zhang, S. Yu, W. Zhang, Y. Wang and F. Xue, *Cancer Biol. Med.*, 2019, **16**(1), 55–70.

5 A. Arun, M. I. Ansari, P. Popli, S. Jaiswal, A. Dwivedi, K. Hajela and R. Konwar, *Cell Proliferation*, 2018, **51**(6), e12501.

6 X. W. Dou, Y. K. Liang, H. Y. Lin, X. L. Wei, Y. Q. Zhang, J. W. Bai, C. F. Chen, M. Chen, C. W. Du, Y. C. Li, J. Tian, K. Man and G. J. Zhang, *Theranostics*, 2017, **7**(16), 4041–4056.

7 G. Sflomos, V. Dormoy, T. Metsalu, R. Jeitziner, L. Battista, V. Scabia, W. Raffoul, J. F. Delaloye, A. Treboux, M. Fiche, J. Vilo, A. Ayyanan and C. Brisken, *Cancer Cell*, 2016, **29**(3), 407–422.

8 A. Bouzekri, A. Esch and O. Ornatsky, *FEBS Open Bio*, 2019, **9**(9), 1652–1669.

9 R. Callis, A. Rabow, M. Tonge, R. Bradbury, M. Challinor, K. Roberts, K. Jones and G. Walker, *J. Biomol. Screening*, 2015, **20**(6), 748–759.

10 J. Dandriyal, R. Singla, M. Kumar and V. Jaitak, *Eur. J. Med. Chem.*, 2016, **119**, 141e168.

11 J. S. Lewis-Wambi, H. Kim, R. Curpan, R. Grigg, M. A. Sarker and V. C. Jordan, *Mol. Pharmacol.*, 2011, **80**, 610e620.

12 S. J. Han, K. Begum, C. E. Foulds, R. A. Hamilton, S. Bailey, A. Malovannaya, D. Chan, J. Qin and B. W. O'Malley, *Mol. Pharmacol.*, 2016, **89**, 14e26.

13 F. Payton-Stewart, S. L. Tilghman, L. G. Williams and L. L. Winfield, *Biochem. Biophys. Res. Commun.*, 2014, **450**, 1358e1362.

14 R. Jeselsohn, in A Study of Palbociclib in Combination with Bazedoxifene in Hormone Receptor Positive Breast Cancer, ClinicalTrials.gov, Dana-Farber Cancer Institute, 2017 as downloaded on January 22, 2017 from, https://clinical trials.gov/ct2/show/record/NCT02448771.

15 R. Lindsay, J. C. Gallagher, R. Kagan, J. H. Pickar and G. Constantine, *Fertil. Steril.*, 2009, **92**, 1045e1052.

16 R. Singla, K. B. Gupta, S. Upadhyay, M. Dhiman and V. Jaitak, *Eur. J. Med. Chem.*, 2018, **146**, 206–219.

17 D. Bafna, F. Ban, P. S. Rennie, K. Singh and A. Cherkasov, *Int. J. Mol. Sci.*, 2020, **21**, 4193.

18 D. Sharma, S. Kumar and B. Narasimhan, *Chem. Cent. J.*, 2018, **12**, 107.

19 Y. M. Yoo and E. B. Jeung, *J. Pineal Res.*, 2009, **47**(4), 301–307.

20 S. Martinkovich, D. Shah, S. L. Planey and J. A. Arnott, *Clin. Interventions Aging*, 2014, **9**, 1437.

21 F. Payton-Stewart, S. L. Tilghman, L. G. Williams and L. L. Winfield, *Biochem. Biophys. Res. Commun.*, 2014, **450**(4), 1358–1362.

22 F. Zengin-Karadayi, M. Yaman, M. M. Kisla, A. G. Keskus, O. Konu and Z. Ates-Alagoz, *Bioorg. Chem.*, 2020, **100**, 103929.

23 R. Ghodsi, E. Azizi and A. Zarghi, *Iran. J. Pharm. Res.*, 2016, **15**(1), 169–177.

24 N. P. Lad, Y. Manohar, M. Mascarenhas, Y. B. Pandit, M. R. Kulkarni, R. Sharma, K. Salkar, A. Suthar and S. S. Pandit, *Bioorg. Med. Chem. Lett.*, 2017, **27**(5), 1319–1324.

25 A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard and G. L. Greene, *Cell*, 1998, **95**(7), 927–937.

26 Z. Ates-Alagoz, N. Coleman, M. Martin, A. Wan and A. Adejare, *Chem. Biol. Drug Des.*, 2012, **80**(6), 853–861.

27 H. Goker, M. Alp, Z. Ateş-Alagoz and S. Yildiz, *J. Heterocycl. Chem.*, 2009, **46**(5), 936–948.

28 H. Goker, C. Kuş, D. W. Boykin, S. Yildiz and N. Altanlar, *Bioorg. Med. Chem.*, 2002, **10**(8), 2589–2596.

29 Z. Ates-Alagoz, Z. Buyukbingol and E. Buyukbingol, *Pharmazie*, 2005, **60**, 643–647.

30 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.

31 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **16**, 2785–2791.

32 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminf.*, 2012, **4**, 17.

33 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.

34 Dassault Systèmes BIOVIA, Discovery Studio Visualizer v17.2.0.16349, Dassault Systèmes, San Diego, 2017.

35 A. Daina, O. Michielin and V. Zoete, *Sci. Rep.*, 2017, **7**, 42717.

36 A. Daina and V. Zoete, *ChemMedChem*, 2016, **11**(11), 1117–1121.

37 H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York, 2016, ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

38 Reference for ggpubr: https://cran.r-project.org/web/packages/ggpubr/index.html.

39 L. Schrödinger, New York, NY, 2019, Schrödinger Release 2019-3: Glide.

40 Y. Kawabata, K. Wada, M. Nakatani, S. Yamada and S. Onoue, *Int. J. Pharm.*, 2011, **420**(1), 1–10.

41 A. Singh, Z. A. Worku and G. V. Mooter, *Expert Opin. Drug Delivery*, 2011, **8**(10), 1361–1378.

42 W. O. Foye, *Foye's Principles of Medicinal Chemistry*, Lippincott Williams & Wilkins, 2008.

43 M. A. Iannone, C. A. Simmons, S. H. Kadwell, D. L. Svoboda, D. E. Vanderwall, S.-J. Deng, T. G. Consler, J. Shearin, J. G. Gray and K. H. Pearce, *Mol. Endocrinol.*, 2004, **18**(5), 1064–1081.

44 R. Gaddipati, G. K. Raikundalia and M. L. Mathai, *Int. J. Biosci., Biochem. Bioinf.*, 2014, **4**(2), 100–105.