Investigations on Anticancer and Antimalarial Activities of Indole-Sulfonamide Derivatives and In Silico Studies

Ratchanok Pingaew,* Prasit Mandi, Veda Prachayasittikul,* Anusit Thongnum, Supaluk Prachayasittikul, Somsak Ruchirawat, and Virapong Prachayasittikul

ABSTRACT: A library of 44 indole-sulfonamide derivatives (1–44) were investigated for their cytotoxic activities against four cancer cell lines (i.e., HuCCA-1, HepG2, AS49, and MOLT-3) and antimalarial effect. Most of the studied indoles exhibit anticancer activity against the MOLT-3 cell line, whereas only hydroxyl-bearing bisindoles displayed anticancer activities against the other tested cancer cells as well as antimalarial effect. The most promising anticancer compounds were noted to be CF3, Cl, and NO2 derivatives of hydroxyl-bearing bisindoles (30, 31, and 36), while the most promising antimalarial compound was an OCH3 derivative of non-hydroxyl-containing bisindole 11. Five quantitative structure–activity relationship (QSAR) models were successfully constructed, providing acceptable predictive performance (training set: \( R = 0.6186–0.9488, \text{RMSE} = 0.0938–0.2432 \); validation set: \( R = 0.4242–0.9252, \text{RMSE} = 0.1100–0.2785 \)). QSAR modeling revealed that mass, charge, polarizability, van der Waals volume, and electronegativity are key properties governing activities of the compounds. QSAR models were further applied to guide the rational design of an additional set of 22 compounds (P1–P22) in which their activities were predicted. The prediction revealed a set of promising virtually constructed compounds (P1, P3, P9, P10, and P16) for further synthesis and development as anticancer and antimalarial agents. Molecular docking was also performed to reveal possible modes of bindings and interactions between the studied compounds and target proteins. Taken together, insightful structure–activity relationship information obtained herein would be beneficial for future screening, design, and structural optimization of the related compounds.

INTRODUCTION

An indole scaffold is a significant privileged structural motif found in a wide range of natural products and clinically used drugs such as nonsteroidal anti-inflammatory agents (i.e., indomethacin), anticancer agents (i.e., vincristine and vinblastine), anti-HIV drugs (i.e., delavirdine and atezivir), and antiviral drugs (i.e., arbidol).1−3 Both natural-derived and synthetic indole derivatives have been reported for diverse pharmacological activities. For example, indole-3-carbinol (I3C), a constituent of cruciferous vegetables (such as broccoli, Brussels sprouts, cauliflower, and cabbage), and its metabolite 3,3′-diindolylmethane (DIM) have been used as chemoprevention for delaying the progression of many cancers.4−7 In addition, indole derivatives have gained attention for their antimalarial effects. Flindersol B and C isolated from Flindersia amboinensis showed selective growth inhibition against Dd2 Plasmodium falciparum malaria strain (IC50 values range = 0.15–1.42 \( \mu M \)).8 Similarly, NITD609, a compound in the spiroindolone class, was proposed as an antimalarial agent (displaying IC50 = 10 nM toward P. falciparum) and is now undergoing clinical trials.9,10

Sulfonamide derivatives are pharmacologically active agents that elicit a wide range of biological activities (i.e., anticancer, antimalarial, antimicrobial, and antiviral activities).11−15 The sulfonamide group could interact with various biological targets16 and was noted to maximize therapeutic efficacy and minimize side effects.17

Compounds containing indole and sulfonamide moieties (Figure 1) have been reported to exhibit anticancer activity in which they acted as inhibitors against many biological targets such as tubulins I−III,16−18 carboxic anhydrase IX IV,19 MET tyrosine kinase V,20 estrogen receptor-α VI,21 phosphatidylinositol-5-phosphate 4-kinase (PISP4K) VII,22 etc. Moreover, indole-sulfonamide derivatives VIII23 and IX24 displayed antiplasmodial activity. Recently, two sets of bisindole-
trisindole-containing sulfonamide derivatives (10−44, Figure 2), except for compounds 27, 31, 33, and 40, have been reported as aromatase inhibitors by our research group.25 In addition, some indole derivatives (1, 2, 4, 7−12, 27−29, and 31−44) were noted for their cytotoxic activities against four cancer cell lines, i.e., HuCCA-1, HepG2, A549, and MOLT-3.26,27

Computational (in silico) approaches have been currently utilized as fundamental tools to facilitate the drug discovery process. Quantitative structure−activity relationship (QSAR) modeling is one of the most commonly used approaches to elucidate the structure−activity relationship (SAR) of the compounds and their biological activities. QSAR modeling is crucial for further design, screening, and structural optimization to obtain derivatives with improved potency and minimized side effects.28,29 QSAR modeling has been employed for the discovery of various drugs including anticancers, antimalarials, and others.30−33 Additionally, molecular docking is a widely used method to find possible modes of protein−ligand binding and interactions, which would be beneficial for screening and design of various types of bioactive compounds.34−37

To find novel lead candidates as promising anticancer and antimalarial agents, our in-house indole-sulfonamide derivatives (Figure 3) were evaluated for their in vitro cytotoxic activities against four cancer cell lines and antimalarial activity. Furthermore, QSAR modeling was performed to construct predictive models, which were further applied for prediction of an additional set of structurally modified compounds. Additionally, molecular docking was performed to elucidate possible binding modes and interactions of the compounds against the protein targets.

■ RESULTS AND DISCUSSION

Chemistry. Three sets of indole-sulfonamide derivatives comprising monoindoles (1−9, series A), bisindoles (10−39, Figure 3, series B), and trisindoles (40−44, Figure 3, series J) were synthesized by our previous works.25−27 The monoindoles (1−9) were prepared by treatment of tryptamine with various N-benzenesulfonyl chlorides. Alkylation of the monoindoles (1−9) with the corresponding benzaldehyde...
derivatives or isatin catalyzed by 20 mol % H₄O₄SiW₁₂·aq in acetonitrile provided bisindoles (10–39) and trisindoles (40–44), respectively.

**Biological Activities. Cytotoxic Activity.** Herein, all of them (1–44) were investigated for their in vitro cytotoxic effects against four human cancer cell lines including HuCCA-1 (cholangiocarcinoma), HepG2 (hepatocellular carcinoma), A549 (lung carcinoma), and MOLT-3 (lymphoblastic leukemia), as summarized in Table 1.

Figure 3. In-house indole-sulfonamide derivatives for biological investigations in this study.
Table 1. Anticancer and Antimalarial Activities (IC_{50}, μM) of Compounds 1–44

| Compound | X    | R    | Anticancer activity* | Antimalarial activity* |
|----------|------|------|----------------------|------------------------|
|          |      |      | HuCCA-1 | HepG2 | A549 | MOLT-3 |          |          |
| 1        | -    | 4-CH₃| 128.18±1.53 | 83.75±4.73 | 132.00±2.12 | 50.41±1.69 | ND      |
| 2*       | -    | 4-OCH₃| 134.69±6.36 | 87.73±2.08 | 119.55±0.71 | 50.15±1.41 | ND      |
| 3        | -    | 4-Br  | NC       | 82.11±2.87 | NC       | 74.78±0.84 | IA      |
| 4*       | -    | 4-Cl  | NC       | 69.68±2.89 | 71.68±0.64 | 46.23±1.65 | IA      |
| 5        | -    | 4-F   | NC       | 114.05±1.41 | NC       | 91.56±1.14 | IA      |
| 6        | -    | 4-CF₃| NC       | 113.85±4.23 | NC       | 59.02±0.97 | IA      |
| 7*       | -    | 4-NO₂| 136.09±1.14 | 95.55±2.82 | 124.50±1.41 | 68.91±2.63 | ND      |
| 8*       | -    | 3-NO₂| NC       | 99.40±4.04 | 128.85±4.95 | 54.38±2.34 | IA      |
| 9*       | -    | 2-NO₂| NC       | NC       | NC       | 59.21±3.13 | IA      |
| 10*      | H    | 4-CH₃| NC       | NC       | NC       | NC       | 4.28    |
| 11*      | H    | 4-OCH₃| NC       | NC       | NC       | NC       | 2.79    |
| 12*      | H    | 4-NO₂| NC       | NC       | NC       | NC       | 8.17    |
| 13       | 4-NO₂| 4-Br  | NC       | NC       | NC       | NC       | IA      |
| 14       | 4-NO₂| 4-F   | NC       | NC       | NC       | 56.39±2.77 | 4.40   |
| 15       | 4-NO₂| 4-CF₃| NC       | NC       | NC       | 13.60±3.88 | IA      |
| 16       | 4-CN | 4-Cl  | NC       | NC       | NC       | NC       | IA      |
| 17       | 4-CN | 4-F   | NC       | NC       | NC       | 18.22±0.23 | 4.08   |
| 18       | 4-CN | 4-CF₃| NC       | NC       | NC       | 23.83±10.08 | 7.59 |
| 19       | 4-Br | 4-Cl  | NC       | NC       | NC       | 44.37±2.50 | IA      |
| 20       | 4-Br | 4-CF₃| NC       | NC       | NC       | 14.58±3.60 | IA      |
| 21       | 4-F  | 4-Br  | NC       | NC       | NC       | NC       | 5.39    |
| 22       | 4-F  | 4-Cl  | NC       | NC       | NC       | 24.84±4.80 | 3.71   |
| 23       | 4-F  | 4-CF₃| NC       | NC       | NC       | 10.65±2.65 | 6.80   |
| 24       | 4-CF₃| 4-Br  | NC       | NC       | NC       | 19.83±3.60 | IA      |
| 25       | 4-CF₃| 4-Cl  | NC       | NC       | NC       | 10.73±1.07 | 6.41   |
| 26       | 4-CF₃| 4-CF₃| NC       | NC       | NC       | 14.75±1.99 | IA      |
| 27*      | 4-OH | 4-CH₃| 14.33±2.12 | 9.55±1.73 | 15.69±0.71 | 6.30±0.91 | ND      |
| 28*      | 4-OH | 4-OCH₃| 33.99±1.41 | 13.30±1.44 | IA       | 7.22±1.21 | 4.17   |
| 29*      | 4-OH | 4-Cl  | 9.69±0.16 | 9.91±1.04 | 12.39±0.58 | 7.13±0.34 | IA      |
| 30       | 4-OH | 4-CF₃| 8.56±1.64 | 7.37±0.41 | 13.30±2.45 | 4.90±0.35 | 3.46   |
| 31*      | 4-OH | 4-NO₂| NC       | 22.02±2.80 | NC       | 2.04±0.10 | ND      |
| 32*      | 4-OH | 3-NO₂| 26.00±1.15 | NC       | 8.49±1.19 | 3.59    |
| 33*      | 4-OH | 2-NO₂| 20.93±3.76 | 9.94±1.93 | NC       | 5.06±0.56 | ND      |
| 34*      | 2-OH | 4-CH₃| 15.69±0.71 | 11.15±1.44 | 64.13±1.41 | 6.94±0.90 | 3.92   |
| 35*      | 2-OH | 4-OCH₃| 30.72±2.08 | 18.73±2.08 | 47.06±2.83 | 7.35±0.54 | 3.65   |
| 36*      | 2-OH | 4-Cl  | 7.75±0.37 | 8.82±1.04 | 8.74±0.79 | 6.06±0.81 | IA      |
| 37*      | 2-OH | 4-NO₂| NC       | 22.23±2.08 | NC       | 4.97±0.13 | 3.16   |
| 38*      | 2-OH | 3-NO₂| NC       | 26.00±1.04 | NC       | 5.47±0.68 | 3.64   |
| 39*      | 2-OH | 2-NO₂| 61.65±1.41 | 22.02±3.79 | NC       | 9.95±0.82 | 4.10   |
Monoidoles (1–9, series A) displayed weak to moderate cytotoxic effects against all four tested cancer cell lines (IC50 value range = 46.23–136.09 μM). Among all monoidoles, compound 4 with a 4-chloro group was noted to be the most potent one against HepG2, A549, and MOLT-3 cell lines (IC50 = 69.68, 71.68, and 46.23 μM, respectively). Inversely, most of the monoidoles were inactive against the HuCCA-1 cell line, except for three compounds (1, 2, and 7), exhibiting a weak cytotoxic effect.

Bisindoles (10–12, series B) with an unsubstituted central benzene ring (X = H) were inactive against all tested cell lines. Bisindole series C–G seem to be inactive anticancer agents toward all tested cell lines, except for MOLT-3. The cytotoxic effect against the MOLT-3 cell line was observed when an electron-withdrawing group is substituted on the benzene ring. This effect was observed for most of the compounds in series C–G (X = 4-NO2: 13–15, X = 4-CN: 16–18, X = 4-Br: 19 and 20, X = F: 21–23, X = CF3: 24–26) with the IC50 range of 10.65–56.39 μM. Distinctively, the replacement of the electron-withdrawing groups of compounds 13–26 (series C–G) with an electron-donating hydroxyl group afforded the derivatives with improved cytotoxic activities, as observed for compounds in series H (27–33, X = 4-OH) and series I (34–39, X = 2-OH). It was noted that compounds 27, 29, 30, and 34–36 displayed a broad range of cytotoxic activities against all cancer cell lines.

Most of the trisindoles (40–43, series J) exhibited selective cytotoxic activity toward the MOLT-3 cell line (IC50 = 5.49–53.84 μM), except for the compound 42, which was also active against HuCCA-1 and HepG2 cell lines (IC50 = 15.95 and 12.52 μM, respectively).

For the HuCCA-1 cell line, bisindole series H and I (27, 28, 33–35, and 39) with R = 4-CH3, 4-OCH3, and 2-NO2 displayed moderate cytotoxic activity (IC50 = 14.33–61.65 μM). Notably, the enhancement of the cytotoxic effect was observed when the halogen substituents (R = 4-Cl and 4-CF3) were introduced to provide more potent derivatives 29, 30, and 36 (IC50 = 7.75–9.69 μM).

For the HepG2 cell line, all bisindoles with a hydroxyl group (27–39, series H and I) exhibited cytotoxic activity with the IC50 range of 7.37–26.00 μM. Interestingly, all of them showed more potent anticancer effect against HepG2 cells when compared to the reference drug etoposide. Among these compounds, the derivative 30 with a 4-trifluoromethyl substituent was shown to be the most potent compound (IC50 = 7.37 μM) with 4.6-fold more potency than the etoposide.

For the A549 cell line, bisindole series H and I (27, 29, 30, and 34–36) with R = 4-CH3, 4-OCH3, 4-Cl, and 4-CF3 displayed cytotoxic activity with IC50 in the range of 8.74–64.13 μM, whereas all nitro derivatives (31–33 and 37–39) were inactive. At this point, it seemed reasonable that lipophilicity could play an important role for the observed activity.

For the MOLT-3 cell line, all of compounds in series H and I (27–39) exhibited cytotoxic activities with comparable IC50 values lesser than 10 μM. Apparently, the 4-NO2 derivative 31 was shown to be the most potent compound with an IC50 value of 2.04 μM.

**Antimalarial Activity.** The antimalarial activity of the indole-sulfonamide derivatives was evaluated against *P. falciparum* (K1, multidrug-resistant strain) (Table 1). Results showed that all monoidoles (series A, I–9) were inactive, but antimalarial activity was shown in most of bisindoles (series B–I, 10–39) and trisindoles (series J, 40–44), in which both types of these tested indoles exhibited comparable IC50 values in the range of 2.79–8.17 μM. The most potent antimalarial agent was noted to be a 4-OCH3 derivative (11) of series B with IC50 = 2.79 μM.

**QSAR Study.** Multiple linear regression (MLR) is an algorithm widely used in computer-aided drug discovery and development due to its interpretable nature. The MLR model displays the linear relationship in which the biological activity pIC50 value (Y variable) is a function of multiple predictors (descriptors or independent X variables). This characteristic of the model provides the informative hint for guiding the rational design of new derivatives as demonstrated in many of our previous works. According to the investigated activities against four cancer cell lines and antimalarial effect, chemical structures of active compounds and experimentally obtained biological activity (IC50 values) were used for the preparation of five separated datasets.
most potent compounds (Table 1) were ranked as Cl derivative of series I (36) > CF3 derivative of series H (30) > Cl derivative of series H (29). It was noted that the substitution of Cl or CF3 groups on both distal benzene rings containing a sulfonamide moiety may influence the values of key descriptors to afford more potent activities. In contrast, other bisindoles with other substituted (X) moieties, rather than the hydroxyl, on the central benzene ring mostly were inactive. This also suggested that the substitution of the hydroxyl group at the central benzene ring was crucial for cytotoxic activity against the HuCCA-1 cell line. It was also observed that only three monoindoles 1, 2, and 7 (series A) showed weak activity, whereas the rest of the series were inactive. These three active monoindoles possessed high values of all descriptors (less negative and more positive values). Similarly, chloro-trisindole 42 is an only active trisindole compound to a HuCCA-1 cell line due to its lower GATS2c and GATS4m values when compared to others in the series. HepG2.

\[
pI_{C_{50}} = 2.7278 \times (AATS7p) - 2.0051 \times (GATS8m) - 3.285
\]

The constructed QSAR model indicated that polarizability (AATS7p) and mass (GATS8m) descriptors influence cytotoxic activity against the HepG2 cell line, in which the polarizability played a slightly more influence as shown by its higher regression coefficient value. A high positive value of polarizability descriptor (AATS7p) along with a low positive value of mass descriptor (GATS8m) was essential for the compound to afford potent activity.

Only two classes of bisindoles containing an X = OH-substituted group (series H and I) exhibited anticancer activity against the HepG2 cell line. The top three most potent compounds were ranked as 30 (R = 4-CF3) > 36 (R = 4-Cl) > 27 (R = 4-CH3), which may be due to their high polarizability along with low mass descriptor values. For series H, it was found that types of substituted mieties on the central benzene ring that were noted for potent activities were ranked to be 4-CF3 > 4-CH3 > 4-Cl > 2-NO2 > 4-OCH3 > 4-NO2 > 3-NO2: 36 > 27 > 29 > 33 > 28 > 31 > 32. This showed that types of substituted mieties on the central benzene ring could influence activities of the compounds by affecting mass descriptor values. This was noted for the compounds, which were ranked as the four least potent compounds (33, 28, 31, 32).

Bioactivity values were normalized by converting to pIC50 values and all chemical structures were subjected to descriptor calculation. Correlation-based feature selection was performed to obtain final sets of informative descriptors for model construction. Finally, five QSAR models were successfully constructed using MLR (eqs 1–5). Definitions and values of descriptors used for model construction are presented in Table 2 and Table S1, respectively.

All constructed models displayed acceptable predictive performance (training set: R = 0.6186–0.9488, RMSE = 0.0938–0.2432; validation set: R = 0.4242–0.9252, RMSE = 0.1100–0.2785) as summarized in Table 3. Plots between experimental and predicted pIC50 values from five constructed QSAR models are provided in Figure 4, and values are provided in Table S2.

### Table 2. Definitions of Informative Descriptors for QSAR Modeling

| Descriptor | Definition | Type |
|------------|------------|------|
| ATSC1c     | centered Broto–Moreau autocorrelation - lag 1/weighted by charges | 2D autocorrelation descriptor |
| GATS2c     | Geary autocorrelation - lag 2/weighted by charges | 2D autocorrelation descriptor |
| GATS4m     | Geary autocorrelation - lag 4/weighted by mass | 2D autocorrelation descriptor |
| AATS7P     | average Broto–Moreau autocorrelation - lag 7/weighted by polarizabilities | 2D autocorrelation descriptor |
| GATS8m     | Geary autocorrelation - lag 8/weighted by mass | 2D autocorrelation descriptor |
| AATS1Cm    | average centered Broto–Moreau autocorrelation - lag 1/weighted by mass | 2D autocorrelation descriptor |
| ATSC5p     | centered Broto–Moreau autocorrelation - lag 5/weighted by polarizabilities | 2D autocorrelation descriptor |
| ATSC0m     | centered Broto–Moreau autocorrelation - lag 0/weighted by mass | 2D autocorrelation descriptor |
| AATS8v     | average centered Broto–Moreau autocorrelation - lag 8/weighted by van der Waals volumes | 2D autocorrelation descriptor |
| AATS1v     | average centered Broto–Moreau autocorrelation - lag 1/weighted by van der Waals volumes | 2D autocorrelation descriptor |
| MATS1v     | Moran autocorrelation - lag 1/weighted by van der Waals volumes | 2D autocorrelation descriptor |
| GATS8c     | Geary autocorrelation - lag 8/weighted by charges | autocorrelation descriptor |
| AATS8e     | Average centered Broto–Moreau autocorrelation - lag 8/weighted by Sanderson electronegativities | autocorrelation descriptor |
| SpMin3, Bhe| Smallest absolute eigenvalue of Burden modified matrix - n 3/weighted by relative Sanderson electronegativities | burden-modified eigenvalues descriptor |

### Table 3. Predictive Performance of Constructed QSAR Models

| Model     | Training pIC50 | LOOCV pIC50 | Training RMSE | LOOCV RMSE |
|-----------|----------------|-------------|---------------|------------|
| HuCCA-1   | 0.9488         | 0.9094      | 0.1389        | 0.1916     |
| HepG2     | 0.9395         | 0.9252      | 0.1432        | 0.1589     |
| AS49      | 0.9431         | 0.8784      | 0.1478        | 0.2170     |
| MOLT-3    | 0.8337         | 0.7811      | 0.2432        | 0.2785     |
| antimalarial | 0.6186       | 0.4242      | 0.0938        | 0.1100     |

\[
pIC_{50} = -0.9731 \times (ATSC1c) - 1.3868 \times (GATS2c) - 1.6772 \times (GATS4m) + 1.26 \tag{1}
\]

Key predictors presented in the model indicated that charges (i.e., ATSC1c and GATS2c) and mass (GATS4m) are key properties that influence cytotoxic activity against the HuCCA-1 cell line, in which the mass descriptor GATS4m played the most influencing role, as indicated by its highest regression coefficient values. Additionally, more negative values (or less positive values) of all descriptors were required to afford a high pIC50 value.

In overview, Cl and CF3 derivatives (R group) of the bisindole-containing hydroxyl-substituted (X group) central benzene ring exhibited the most promising activities. The three most potent compounds (Table 1) were ranked as Cl derivative of series I (36) > CF3 derivative of series H (30) > Cl derivative of series H (29).
and 32) in which their GATS8m values were observed to be increased greater than 0.9 (Table S1).

Similar with the HuCCA-1 cell line, most of the trisindoles (series J) were inactive, except for the Cl derivative 42. This suggested that the substitution of the 4-Cl (R) group on the benzene ring of sulfonamides could increase polarizability descriptor (AATS7p = 1.505) values as well as decrease mass descriptor (GATS8m = 0.883) values. This phenomenon was also noticed when comparing the values of descriptors in the Cl derivative with other compounds in the same hydroxyl-containing bisindole series H and I (29: AATS7p = 1.519, GATS8m = 0.849; 36: AATS7p = 1.513, GATS8m = 0.867).

All investigated monoindoles were active cytotoxic agents against the HepG2 cell line, except that of compound 9. Halogen derivatives of monoindoles 4 and 3 elicited the most potent activities among the rest of series, which may be due to their high polarizability along with low descriptor values (4: AATS7p = 1.390, GATS8m = 1.179; 3: AATS7p = 1.424, GATS8m = 1.202; Table S1). Additionally, changing the position of substituted NO2 on the central benzene ring could affect the activity of the compounds, as demonstrated by the inactive activity of 2-NO2 derivative 9, while 4-NO2 (7) and 3-NO2 (8) derivatives elicited weak activities.

**A549.**

\[ \text{pIC}_{50} = -0.1329 \text{ (AATSC1m)} + 0.0305 \text{ (ATSC5p)} + 0.0002 \text{ (ATSC0m)} - 1.0154 \quad (3) \]

The QSAR model revealed that mass (AATSC1m and ATSC0m) and polarizability (ATSC5p) influenced the cytotoxic effect against the A549 cell line. The mass descriptor AATSC1m was noted to be the most influential predictor as seen by its highest regression coefficient value. The lower positive/higher negative value of AATSC1m is required for potent activity.

The top three most potent compounds were ranked as 36 > 29 (R = 4-Cl) > 30 (R = 4-CF3). All of these compounds possessed high ATSC5p and ATSC0m values but low AATSC1m values (36: ATSC5p = 5.979, ATSC0m = 5333.171, AATSC1m = 8.102; 29: ATSC5p = 6.204, ATSC0m = 5333.171, AATSC1m = 8.102; 30: ATSC5p = 6.695, ATSC0m = 4550.110, AATSC1m = 7.350). It should be noted that both chloro derivatives 36 (X = 2-OH) and 29 (X = 4-OH) had the same values of mass descriptors (ATSC0m and AATSC1m) but different polarizability descriptor values (ATSC5p). This suggested that the position of the substituted OH group on the central benzene ring influenced the activity of the compounds by affecting the polarizability. For both OH-containing bisindole series H and I, chloro derivatives with high polarizability were ranked as the most potent of the series (29 for series H and 36 for series I), but CH3 derivatives ranked the least (27 for series H and 34 for series I). The impaired activity of the CH3 derivatives may be affected via an increase in AATSC1m values when compared to the most potent ones (29: AATSC1m = 8.102, pIC50 = -1.093 vs 27: AATSC1m = 9.526, pIC50 = -1.196; 36: AATSC1m = 8.102, pIC50 = -0.942 vs 34: AATSC1m = 9.526, pIC50 = -1.807).

Like other cell lines, the most potent monoindole (series A) with a chloro moiety (4) possessed the low positive value of AATSC1m (9.624) along with the high ATSC5p (3.436) and ATSC0m (2482.528) values when compared to the others in the same series (Table S1).

**MOLT-3.**

\[ \text{pIC}_{50} = -0.0869 \text{ (AATSC8v)} - 0.0696 \text{ (AATSC1v)} + 28.3224 \text{ (MATS1v)} - 2.8263 \quad (4) \]
All predictors of the QSAR model are van der Waals volume descriptors, which indicated that this property is the main characteristic governing the cytotoxic effect against the MOLT-3 cell line. The QSAR equation revealed that the descriptor MATS1v played the main role (regression coefficient value = 28.3224), whereas another two descriptors (AATSC8v and AATSC1v) played lesser influence on pIC50 values.

In overview, the cytotoxic effect of the studied compounds (1−44) showed different patterns against MOLT-3 cell lines, in which they were shown to be more sensitive to MOLT-3 than the others, especially the series (i.e., series A, C, D, E, F, and G) that were inactive agents against other three cancer cell lines (i.e., HuCCA-1, HepG2, and A549) but displayed activities against the MOLT-3. The most potent compounds were ranked as 31 (R = 4-NO2 of series H) > 30 (R = 4-CF3 of series H) > 37 (R = 4-NO2 of series I).

Monoindole series A (1−9) displayed the less potent activities than the other series due to their lower MATS1v values (all of them possessed MATS1v values not greater than 0.061). For series C, E, and F, the 4-CF3 derivatives displayed more potent activities than that of halogen derivatives (15 > 14, 20 > 19, and 23 > 22), which may be due to their lower AATSC1v values (AATSC1v: 15 = 3.044, 14 = 3.296, 20 = 3.090, 19 = 3.221, 23 = 3.179, 22 = 3.319; Table S1).

Additionally, the effect of changing the substituted position of NO2 (R group) on the benzene sulfonamide rings of bisindoles was observed for the compounds in series H and I. For both series, substitution on the 4-position of the benzene ring (R = 4-NO2) provided the compounds with the best activities when compared to those on 3- and 2-positions (31 > 33 > 32 and 37 > 38 > 39). Considering these NO2 bisindole analogs, all of them had the same values of AATSC1v (3.669) and MATS1v (0.076) but different values of AATSC8v. For example, compound 39 with 2-NO2 substitution exhibited the least potent activity when compared to 38 with 3-NO2 and 37 with 4-NO2 substitutions due to their AATSC8v values as shown (39: AATSC8v = −0.678, pIC50 = −0.998; 38: AATSC8v = −2.043, pIC50 = −0.738; 37: AATSC8v = −2.004, pIC50 = −0.696; Table S1).

Antimalarial.

\[
pIC_{50} = 0.3918 \ (GATS8c) - 9.4117 \ (AATSC8e) - 1.6871 \ (SpMin3\_Bhe) + 2.1086 \quad (5)
\]

QSAR analysis indicated that charge (GATS8c) and electronegativity (AATSC8e and AATSC8e) are important descriptors. The most potent compounds were ranked as 11 (X = H, R = 4-OCH3 derivative series B) > 37 (4-NO2 derivative series I) > 30 (CF3 derivative series H). It was also noticed that the most potent/or the solely active compounds of each series contained halogen (R = 4-F, 4-Cl) substituents on their molecules (i.e., compounds 14, 17, 22, 25, and 42). These halogen analogs displayed the lesser positive values of the electronegativity descriptor SpMin3_Bhe when compared to the others within the same series, therefore affording the higher pIC50 values (Table S1). However, this is an exception for series H and I (Figure 1), in which the most potent compounds were noted to be derivatives of 4-CF3 (30) with the low value of the electronegativity descriptor AATSC8e (−0.006), leading to the more potent activity.
Q SAR-Driven Rational Design and Prediction of Newly Designed Compounds (P1–P22). The constructed QSAR models were further applied for rational design and structural modification to obtain a set of 22 newly designed derivatives (structurally modified compounds P1–P22, Figure 5). Compounds in series H and I were used as parent compounds for structural modification to provide modified compounds of series M1–M4. Compounds of M1 series were obtained by substitution of OCH3 at the S-position of both indole rings. Similarly, compounds of M3 series were designed by substitution of one OCH3 and two OCH3 at the S-position of the indole rings as well as an exclusion of both aromatic sulfonyl groups. For M4 series, the compounds were obtained by several modification strategies, i.e., including and removing the alkyl amine groups, changing the position of the substituted OH group on the central benzene ring, and substituting one CH3/OCH3 group on both indole rings.

Similar with the original compounds, all modified compounds were drawn, geometrically optimized, and calculated to obtain key descriptor values, which were required as predictors; values of the key descriptors are provided in Table S3. Subsequently, the constructed models were utilized to predict activities of these modified compounds (P1–P22); the predicted activities (predicted pIC50 values) are shown in Table 4. Additionally, the novelty of these virtually constructed compounds was ensured by structural similarity searching using a web-based tool against various available libraries. Results showed that all modified compounds are novel and have not existed or been reported elsewhere; the similarity score is summarized in Table S4.

Most potent modified compounds against HuCCA-1 and HepG2 cell lines were predicted to be compounds from series M1 and M2, while those against AS49 and MOLT-3 cell lines were from M3 series. Additionally, the most potent antimalarial agents were predicted to be compounds of M4 series.

For the HuCCA-1 cell line, two most potent compounds were predicted to be CF3-containing compound P3 followed by P7 (predicted pIC50 = −0.964, and −0.979, respectively). According to the QSAR model (eq 1), compound P3 had the highest negative values of the charge descriptor ATSC1c (−0.992) when compared with other modified compounds. Similarly, P7 displayed a comparable value of ATSC1c (−0.991).

For the HepG2 cell line, two most potent compounds were predicted to be Cl-containing derivatives P1 and P5 (predicted pIC50 = −0.740 and −0.788, respectively). According to the QSAR model (eq 2), the potent predicted activity of both compounds could be due to their highest polarizability descriptor GATS8m value.

Compounds P10 and P9 were predicted as the most potent cytotoxic agent against AS49 and MOLT-3 (P10: AS49 predicted pIC50 = −0.682; P9: MOLT-3 predicted pIC50 = 0.310). Considering the predicted activities against the MOLT-3 cell line, compound P9 bearing a single OCH3 moiety displayed a high positive pIC50 value (0.310). However, only the substitution of one or more OCH3 moieties to another indole ring caused a marked change activity of the compound P10 as shown by a negative pIC50 value (−0.071). For the antimalarial effect, compound P16 from M4 series was noted as the most potent compound (predicted pIC50 = −0.440) due to its highest charge GATS8c descriptor (1.211). The summary of the most potent compounds for each activity (both tested and newly designed ones) is shown in Figure 6.

**Molecular Docking.** Molecular docking is widely recognized as an essential tool for structure-based drug discovery to reveal possible binding modes and interactions of the ligands against their protein targets. Herein, molecular docking was performed against two putative target proteins for anticancer and antimalarial activities. The epithelial growth factor receptor (EGFR) is a receptor tyrosine kinase that plays essential functions in cell growth and cell proliferation. Overexpression, overactivation, and mutation of the EGFR have been reported in many types of cancers and are associated with tumor initiation, angiogenesis, metastasis, and drug resistance. The EGFR was noted to be a putative target for indole-based anticancer agents. Cyclin-dependent protein kinases (CDKs) play an essential role in the control of cell growth and differentiation of the parasite; thus, an inhibition of CDKs is well known as one strategy for discovery of antimalarial agents. The P. falciparum cyclin-dependent protein kinase PpPK5 was reported as a putative target for indole-based antimalarial agents. Herein, the compounds were docked against EGFR (PDB ID: 5YU9) and P. falciparum PpPK5 (PDB ID: 1VO0).

Redocking of cocrystallized ligands was performed against the target protein. Both of the redocked cocystalized ligands provided root-mean-square deviation (RMSD) less than 2 Å, which ensures that the docking condition of each model is reliable for the investigation of the studied compounds (RMSD values: redocked EGFR = 0.777 Å and PpPK5 = 0.352 Å). Redocked poses of the cocrystallized ligands and the target proteins are provided in Figure S1. Docking poses and 2D protein–ligand interactions of the most promising compounds for anticancer (compounds 36 and 31) and antimalarial activities (compounds P1 and P22) are provided in Figures S4 and S5, respectively.
Antimalarial (compound 11) agents are summarized in Figures 7 and 8. Docking poses and 2D interaction diagrams of original compounds and virtually designed compounds displaying the most potent activities against each cancer cell line (as summarized in Figure 6) are provided in Figures S2 and S3. The binding energy of the docked compounds is summarized in Tables S5 and S6.

Molecular docking revealed that two of the promising anticancer compounds 36 and 31 could occupy in the active site of the EGFR in a similar manner with that of the cocrystallized ligand ibrutinib. The 2D diagram of ibrutinib–EGFR interactions showed that the N-atom of the pyrimidine ring and its substituted amino group play a role in forming two hydrogen bond interactions with Met793 and Gln791 residues, respectively, whereas other hydrophobic parts of the drug molecule form hydrophobic interactions with several residues (Figure 8A). For compound 36, displaying the most potent activities against HuCCA-1 and A549 cell lines, it was observed that the indole ring and hydroxyl group substituted on the ortho position of the central benzene ring are involved in the formation of hydrogen interactions with Met793 and Cys797 residues, respectively. Additionally, two indole rings, central benzene as well as terminal chloro-substituted benzene rings, played a part in hydrophobic interactions with Val726, Leu844, Leu718, and Ala743 (similar residues of those of ibrutinib) together with additional residues such as Pro794 and Leu792.

Figure 6. Summary of the most potent compounds (original tested and newly designed).

Figure 7. Possible binding modes of the most potent compounds and target proteins. (A) Binding modes of the cocrystallized ligand ibrutinib (orange), compound 36 (green), and 31 (magenta) against the anticancer target EGFR (PDB ID: 5YU9). (B) Binding modes of the cocrystallized ligand indirubin-5-sulfonate (orange) and compound 11 (green) against the antimalarial target PPK5 (PDB ID: 1V0O).
For the most potent compound against the MOLT-3 cell line (compound 31), four hydrogen bond interactions were observed between NO2, SO2, OH, and NH groups of the molecule with His805, Lys745, Gln791, and Phe795 residues of the target protein as well as with several hydrophobic interactions with other residues (Figure 8C). For the modified compounds P1 and P3, it was revealed that the OH group substituted on the para position of the central benzene ring is essential for interaction with the Gly724 residue (Figure S2E,F), whereas the NH2 groups located at both terminal branches of the substituted alkyl chain on indole rings of compounds P9 and P10 are required for hydrogen bond interactions with Asn842, Gln791, and Arg841 (Figure S2G,H).

For the antimalarial action, it was observed from the cocrystallized ligand indirubin-5-sulfonate that both indole rings form hydrogen bond interactions with Glu80 and Leu82 residues, whereas one of them could additionally form a \( \pi-\pi \) interaction with Phe79. The SO3\(^-\) (sulfonate) group in the molecule is also essential for hydrogen bonding with Lys32; however, none of hydrophobic interactions were observed (Figure 8D). In contrast, the binding mode of the most potent indole-based antimalarial compound 11 seems to be dominated by several hydrophobic interactions rather than the hydrogen interactions. It was demonstrated that hydrophobic parts of the compound (i.e., indole ring, central benzene ring, and substituted terminal benzene ring) play a part in the formation of hydrophobic interactions with several amino acid residues. A hydrogen bond formation between the SO2 group and the Asn130 residue was also noted (Figure 8E).

For the antimalarial action, it was observed from the cocrystallized ligand indirubin-5-sulfonate that both indole rings form hydrogen bond interactions with Glu80 and Leu82 residues, whereas one of them could additionally form a \( \pi-\pi \) interaction with Phe79. The SO3\(^-\) (sulfonate) group in the molecule is also essential for hydrogen bonding with Lys32; however, none of hydrophobic interactions were observed (Figure 8D). In contrast, the binding mode of the most potent indole-based antimalarial compound 11 seems to be dominated by several hydrophobic interactions rather than the hydrogen interactions. It was demonstrated that hydrophobic parts of the compound (i.e., indole ring, central benzene ring, and substituted terminal benzene ring) play a part in the formation of hydrophobic interactions with several amino acid residues. A hydrogen bond formation between the SO2 group and the Asn130 residue was also noted (Figure 8E).

For the modified compound series, it was found that besides several hydrophobic interactions and hydrogen bonding with the Asn130 residue, the most potent modified compound P16 could form an additional hydrogen bond interaction with the Gln129 residue (Figure S3).

### CONCLUSIONS

A set of 44 in-house indole-sulfonamide derivatives (series A–J, 1–44) were investigated for their cytotoxic activities against four cancer cell lines and antimalarial effect. Experimental results revealed that most of the tested compounds (i.e., monoindoles, bisindoles, and trisindoles) exhibited anticancer activity against the MOLT-3 cell line, except for that of series B. Unlikely, most of the compounds were inactive against HuCCA-1, HepG2, and A549 cell lines in which only hydroxyl-containing bisindoles (series H and I) were noted to be promising active classes, while some of monoindoles (series A) and trisindoles (series J) showed weak to moderate activity. Antimalarial activities were observed for most of the studied compounds except for monoindoles. Five QSAR models were successfully constructed, providing acceptable predictive performance. Key predictors in the constructed models revealed a set of key properties, which influence the activity of the compounds including charge, mass, polarizability, van der Waals volume, and electronegativity. The constructed models were further applied to guide the rational design and activity prediction of an additional set of 22 structurally modified compounds (series M1–M4, compounds P11–P22). Finally, a set of promising indole-based compounds (Figure 6) were highlighted for their potential to be further developed as anticancer and antimalarial agents. QSAR modeling indicated some key chemical properties required for potent activity. Molecular docking demonstrated that an insertion of hydrophobic moieties (i.e., aromatic rings and
halogen atoms) could play an essential role for the investigated indoles to accommodate into the active site of the target proteins through hydrophobic interactions. Hydrogen bond interactions between key functional groups (i.e., indole ring, SO₂, NO₂, and NH₃) and target protein residues for anticancer (i.e., Met793, Cys797, His805, Lys745, Gln791, and Phe795) and antimalarial (i.e., Asn130 and Gln129) actions were also revealed. In summary, the obtained structure–activity relationship knowledge would be beneficial for further studies, design, and optimization of the indole-based compounds as well as other related classes.

## Experimental Section

**Cytotoxic Assay.** The cells suspended in the corresponding culture medium were inoculated in 96-well microtiter plates (Corning Inc., NY, USA) at a density of 10,000–20,000 cells per well and were incubated for 24 h at 37 °C in a humidified atmosphere with 95% air and 5% CO₂. An equal volume of additional medium containing either the serial dilutions of the test compounds, positive control (etoposide acid), 32 mM NaHCO₃, and 10% heat-activated human serum were further incubated for an additional 48 h. The number of surviving cells in each well was determined using MTT assay, and XTT assay (for adherent cells: HuCCA-1, HepG2, and A549 cells) and XTT assay (for suspended cells: MOLT-3 cells). The IC₅₀ value is defined as the drug (or compound) concentration that inhibits cell growth by 50% (relative to negative control).

**Antimalarial Assay.** *P. falciparum* (K1, multidrug-resistant strain) was cultivated under in vitro conditions, according to Trager and Jensen, in RPMI 1640 medium containing 20 mM HEPES (N-2-hydroxyethylpiperazine-N’-2-ethanesulfonic acid), 32 mM NaHCO₃, and 10% heat-activated human serum with 3% erythrocytes in a humidified 37 °C incubator with 3% CO₂. The culture was passaged with a fresh mixture of erythrocytes and medium every day to maintain cell growth. Quantitative assessment of antimalarial activity in vitro was determined by microculture radiotype techniques based upon the methods described by Desjardins et al. Briefly, a mixture of 200 μL of 1.5% erythrocytes with 1% parasitemia at the early ring stage was pre-exposed to 25 μL of the medium containing a test sample dissolved in 1% DMSO (0.1% final concentration) for 24 h. Subsequently, 25 μL of [³H] hypoxanthine (Amersham, USA) in culture medium (0.5 μCi) was added to each well and the plates were incubated for an additional 24 h. Levels of incorporated radioactive-labeled hypoxanthine, indicating parasite growth, were determined using a Top Count microplate scintillation counter (Packard, USA). The percentage of parasite growth was calculated using the signal count per minute of treated (CPMT) and untreated (CPMU) as shown by eq 6.

\[
\text{%parasite growth} = \frac{\text{CPMT}}{\text{CPMU}} \times 100 \tag{6}
\]

**QSAR Modeling.** Datasets. According to investigated activities (i.e., cytotoxic activities against four cancer cell lines and antimalarial effect), five separated datasets were prepared for QSAR modeling. Experimental results (IC₅₀ values) and chemical structures of active tested compounds were used for dataset preparation. Only active compounds were included for modeling, while all the inactive ones were excluded from the datasets. To normalize data points, experimental IC₅₀ values were converted to pIC₅₀ values by taking negative logarithm to the base of 10 (−log IC₅₀). Chemical structures of the tested compounds (1–4) were further subjected to descriptor calculation.

**Molecular Structure Optimization and Descriptor Calculation.** Chemical structures of the compounds (1–4) were drawn using GaussView software. To obtain the low-energy conformers for subsequent descriptor calculations, the structures underwent geometrical optimization by initial optimization using Gaussian 09 at the semi-empirical Austin Model 1 (AM1) level followed by density functional theory (DFT) calculation using Becke’s three-parameter hybrid method with the Lee–Yang–Parr correlation functional (B3LYP) together with the 6-31g(d) level. An in-house developed script was used to extract a set of 13 quantum chemical descriptors (i.e., mean absolute atomic charge (Qma), total energy (E_total), total dipole moment (μ), highest occupied molecular orbital energy (E_HOMO), lowest unoccupied molecular orbital energy (E_LUMO), energy difference of HOMO and LUMO (HOMO-LUMO)), electron affinity (EA), ionization potential (IP), Mulliken electronegativity (χ), hardness (η), softness (S), electrophilic index (αₜ), and electrophilicity (α). The optimized structures were subsequently used as input data for molecular descriptor calculation using PADEL software (version 2.1) to obtain an additional set of 144 descriptors (55 classes) and 3055 fingerprints. Finally, a set of 4499 calculated descriptors and fingerprints were subjected for feature selection to obtain a set of important descriptors for QSAR model construction.

**Feature Selection.** Initially, descriptors with constant values were excluded. A set of 480 descriptors were included for the preparation of datasets including ALOGP descriptor, auto-correlation descriptor, PaDEL bond count descriptor, burden-modified eigenvalues descriptor, constitutional descriptor, and ring count descriptor. Correlation-based feature selection was performed in which Pearson’s pair correlation coefficient value (r) was calculated for each pair of descriptor and bioactivity (pIC₅₀ values). Cutoff values of 0.5–0.8 were used to select a set of highly correlated descriptors whose |r| ≥ 0.5–0.8 (a cutoff value of 0.8 for HuCCA-1, HepG2, and AS49, a cutoff value of 0.7 for MOLT-3, and a cutoff value of 0.5 for antimalarial). Then, these sets of selected descriptors were subjected to a further feature selection process using attribute selection (CfsSubsetEval, Best First) and/or stepwise multiple linear regression (MLR) as implemented by Waikato Environment for Knowledge Analysis (WEKA) version 3.4.5. Additionally, intercorrelation between all pairs of descriptors was calculated to remove highly correlated and redundant descriptors. Finally, a set of informative descriptors were obtained to prepare final datasets for model construction.

**Multivariate Analysis.** The final datasets were used as input files for multivariate analysis using the multiple linear regression (MLR) algorithm as implemented by Waikato Environment for Knowledge Analysis (WEKA) version 3.4.5. Descriptor values were assigned as independent variables (X) to predict biological activity pIC₅₀ values (dependent variable Y) as shown in eq 7.

\[
Y = B_0 + \sum B_n X_n \tag{7}
\]

where Y is the pIC₅₀ value of the compound, B₀ is the intercept, and Bₙ are the regression coefficients of descriptors Xₙ.
Data Sampling. The dataset was randomly divided into two sets (i.e., training set and testing leave-one-out cross-validation (LOO-CV) set). One sample was excluded from the whole dataset to be used as a testing set, while the remaining samples \((N - 1)\) were used as a training set. The same process was continued until all of the samples in the dataset were selected as a testing set to predict activity \((Y\) variable).

Evaluating the Performance of QSAR Models. Predictive performance and predictive errors of the constructed models were assessed by two statistical parameters, i.e., correlation coefficient \((R)\) and the root-mean-square error \((\text{RMSE})\), respectively.

Application of QSAR Models for Rational Design and Prediction of Structurally Modified Compounds \((P1−P22)\). The constructed QSAR models were used for guiding the rational design of an additional set of 22 structurally modified compounds \((P1−P22, \text{Figure 5})\). Structural modification was performed following the key predictors (descriptors appeared in the model). All newly designed compounds were drawn and subjected to descriptor calculation in the same way as the original compounds to obtain key descriptor values \((X\) variables), which were further used for the prediction of biological activities \((pIC_{50}: Y\) variables).

Similarity Search to Define the Novelty of the Newly Designed Compounds. To ensure that the newly design compounds \((P1−P22)\) are novel, a web-based tool, namely, SwissSimilarity \((\text{http://www.swisssimilarity.ch/})\), was used to screen whether our new compounds have already been reported elsewhere. The search is performed based on the structural similarity of the interest compounds against the compounds in databases \(\text{(i.e., drugs, bioactive compounds, and commercially available compound databases)}\). Results were returned in the form of similarity score in which a higher score indicates high similarity \((\text{full score } = 1)\), which indicates that the newly designed compound is identical with the existing compound in the databases.

Molecular Docking. Molecular docking was performed to elucidate interactions of highly potent compounds toward their anticipated target proteins. The epidermal growth factor receptor \((\text{EGFR})\) and \(P. falciparum\) protein kinase \(5\) \((\text{PPPKS})\) were selected as anticancer and antimalarial targets, respectively. Initially, the crystal structure of EGFR 696-1022 \((\text{T790M in complex with ibrutinib and the structure of } P. falciparum\text{ PPPKS in complex with indirubin-5-sulfonate ligand were retrieved from the Protein Data Bank (PDB IDs: SY9 and 1V00, respectively)}\). Geometry-optimized chemical structures of the original \((1−44)\) and modified compounds \((P1−P22)\) were used as investigated ligands. Prior to docking, the protein structure was prepared by adding hydrogen atoms, optimizing hydrogen bonds, removing atomic clashes, and assigning optimal protonation states using the “Protein Prep” function incorporated in Flare software version 5.0 \((\text{Cresset software)}\). All ligands were imported into Flare software using the “Do full preparation on proteins and ligands” option, which will simply assign protonated or deprotonated states \(\text{(assuming a pH of 7.0)}\) as well as given charge to charged groups for imported ligands. An energy grid box with a radius of 6 Å was created around the cocrystallized ligand within the active site of the protein. Molecular docking was performed using the “Normal” docking option in Flare software. The docking calculation used “Lead Finder”, a built-in genetic algorithm-based docking algorithm, in which the pool size and population size were set as \(1.0\). The cocrystallized ligands ibrutinib and indirubin-5-sulfonate were redocked into the EGFR and PPPKS, respectively, as to validate the docking protocol. The redocking was evaluated by the calculation of root-mean-standard deviation \((\text{RMSD})\) between the original and redocked positions of the cocrystallized ligand using the online tool DockRMSD version 1.1. If the calculated RMSD values \(< 2\) Å, the protocol could be further used for investigation of the studied compounds. Docking poses of the investigated compounds were visualized and analyzed using the open source PyMoL software version 2.5, and the 2D ligand–protein interaction diagram was generated using an online available tool, PoseView.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.1c04552.

Values of molecular descriptors used for the construction of predictive four QSAR models \(\text{(original compounds } 1−44)\), experimental and predicted activity \((pIC_{50})\) values of compounds \(1−44\) against four cancer cell lines and antimalarial activities, values of molecular descriptors of structurally modified compounds \((P1−P22)\), similarity scores of structurally modified compounds \((P1−P22)\), summary binding energy of the original compounds and the structurally modified compounds \((\text{kcal/mol)}\), redocking poses of cocrystallized ligands against their target, diagrams showing two-dimensional protein–ligand interactions of the cocrystallized ligand, the most potent tested compounds, and the structurally modified compounds predicted as potent inhibitors against the anticancer target EGFR, and diagrams showing two-dimensional protein–ligand interactions of the cocrystallized ligand, the most potent tested compound, and the structurally modified compound predicted as a potent inhibitor against the antimalarial target PPPKS \((\text{PDF)}\).

AUTHOR INFORMATION

Corresponding Authors
Ratchanok PINGAEW — Department of Chemistry, Faculty of Science, Srinakharinwirot University, Bangkok 10110, Thailand; \(\text{orcid.org/0000-0003-4977-5854}; \) Phone: +66-2-649-5000 ext 18253; Email: ratchanok@g.swu.ac.th; Fax: 662-260-0128

Veda Prachayasittikul — Center of Data Mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand; \(\text{orcid.org/0000-0001-6338-3721}; \) Phone: +66-2-441-4376; Email: veda.pra@mahidol.ac.th

Authors
Prasit Mandal — Department of Community Medical Technology, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand
Anusit Thongnum — Department of Physics, Faculty of Science, Srinakharinwirot University, Bangkok 10110, Thailand; \(\text{orcid.org/0000-0003-3272-5797}\)
Supaluk Prachayasittikul — Center of Data Mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand

https://doi.org/10.1021/acsomega.1c04552
ACS Omega 2021, 6, 31864−31866
Somsak Ruchirawat — Laboratory of Medicinal Chemistry, Chulabhorn Research Institute, and Program in Chemical Sciences, Chulabhorn Graduate Institute, Bangkok 10210, Thailand; Center of Excellence on Environmental Health and Toxicology (EHT), Commission on Higher Education, Ministry of Education, Bangkok 10400, Thailand

Virapong Prachayasittikul — Department of Clinical Microbiology and Applied Technology, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.1c04552

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This project was financially supported from National Research Council of Thailand (NRCT: grant no. 474/2563). We are also indebted to the Chulabhorn Research Institute for bioactivity testing.

REFERENCES

(1) de Sa Alves, F. R.; Barreiro, E. J.; Manssour Fraga, C. A. From nature to drug discovery: the indole scaffold as a ‘privileged structure’. Mini-Rev. Med. Chem. 2009, 9, 782–793.

(2) Suzen, S. Recent studies and biological aspects of substantial indole derivatives with anti-cancer activity. Curr. Org. Chem. 2017, 21, 2068–2076.

(3) Kaushik, N. K.; Kaushik, N.; Attrii, P.; Kumar, N.; Kim, C. H.; Verma, A. K.; Choi, E. H. Biomedical importance of indoles. Molecules 2013, 18, 6620–6662.

(4) Singh, A. A.; Patil, M. P.; Kang, M.-J.; Niyonzigieyi, I.; Kim, G.-D. Biomedical application of indole-3-carbinol: A mini-review. Phytochem. Lett. 2021, 41, 49–54.

(5) Jiang, Y.; Fang, Y.; Ye, Y.; Xu, X.; Wang, B.; Gu, J.; Aschner, M.; Chen, J.; Lu, R. Anti-cancer effects of 3,3′-dindolylmethane on human hepatocellular carcinoma cells is enhanced by calcium ionophore: the role of cytosolic Ca2+and p38 MAPK. Front. Pharmacol. 2019, 10, 1167.

(6) Banerjee, S.; Kong, D.; Wang, Z.; Bao, B.; Hillman, G. G.; Sarkar, F. H. Attenuation of multi-targeted proliferation-linked signaling by N-acylsulfonamides: Synthetic routes and biological potential in medicinal chemistry. Chem. Biol. Drug Des. 2017, 90, 1094–1105.

(7) Man, R.-J.; Tang, D.-J.; Lu, X.-Y.; Duan, Y.-T.; Tso, X.-X.; Yang, M.-R.; Wang, L.-L.; Wang, B.-Z.; Xu, C.; Zhu, H.-L. Synthesis and biological evaluation of novel indole derivatives containing sulfonamide scaffold as potential tubulin inhibitor. MedChemComm 2016, 7, 1759–1767.

(8) Aceves-Luquero, C.; Galiana-Rosseló, C.; Ramis, G.; Villalonga-Planells, R.; García-Espera, E.; Fernández de Mattos, S.; Pelaez, R.; Llinares, J. M.; González-Rosende, M. E.; Villalonga, P. N-(2-methyl-indol-1H-5-yl)-1-naphthalenesulfonamide: A novel reversible antimitotic agent inhibiting cancer cell motility. Biochem. Pharmacol. 2016, 115, 28–42.

(9) Chang, J. Y.; Hsieh, H. P.; Chang, C. Y.; Hus, K. S.; Chiang, Y. F.; Chen, C. M.; Kuo, C. C.; Liou, J. P. 7-Aryl-aminooindoline-1-sulfonamides as a novel class of potent antitubulin agents. J. Med. Chem. 2006, 49, 6656–6659.

(10) Peerzada, M. N.; Khan, P.; Ahmad, K.; Hassan, M. I.; Azam, A. Synthesis, characterization and biological evaluation of tertiary sulfonamide derivatives of pyridyl-indole based heteroaryl chalcone as potential carbonic anhydrase IX inhibitors and anticancer agents. Eur. J. Med. Chem. 2018, 155, 13–23.

(11) Sattler, M.; Pride, Y. B.; Ma, P.; Gramlich, J. L.; Chu, S. C.; Quinnan, L. A.; Shirazian, S.; Liang, C.; Podor, K.; Christensen, J. G.; Salgia, R. A novel small molecule met inhibitor induces apoptosis in cells transformed by the oncogenic TPR-MET tyrosine kinase. Cancer Res. 2003, 63, 5462–5469.

(12) Hendy, M. S.; Ali, A. A.; Ahmed, L.; Hossam, R.; Mostafa, A.; Elmaraz, M. M.; Naguib, B. H.; Attia, Y. M.; Ahmed, M. S. Structure-based drug design, synthesis, In vitro, and In vivo biological evaluation of indole-based biomimetic analogs targeting estrogen receptor-α inhibition. Eur. J. Med. Chem. 2019, 166, 281–290.

(13) Manz, T. D.; Sivakumaren, S. C.; Ferguson, F. M.; Zhang, T.; Yangar, A.; Sea, H. S.; Piccaro, S. B.; Card, J. D.; Shim, H.; Miduturu, C. V.; Simeonov, A.; Shen, M.; Marto, J. A.; Dhe-Paganon, S.; Hall, M. D.; Cantley, L. C.; Gray, N. S. Discovery and structure-activity relationship study of (Z)-5-methylenthiazolidin-4-one derivatives as potent and selective pan-phosphatidylinositol 5-phosphate 4-kinase inhibitors. J. Med. Chem. 2020, 63, 4880–4895.

(14) Devender, N.; Gunjan, S.; Tripathi, R.; Tripathi, R. P. Synthesis and antiplasmodial activity of novel indoleamide derivatives bearing sulfonamide and triazole pharmacophores. Eur. J. Med. Chem. 2017, 131, 171–184.

(15) Yadav, R. R.; Khan, S. I.; Singh, S.; Khan, I. A.; Vidwakarma, R. A.; Bhazate, S. B. Synthesis, antimalarial and antitubercular activities of meridianin derivatives. Eur. J. Med. Chem. 2015, 98, 160–169.

(16) Pingaew, R.; Mandi, P.; Prachayasittikul, V.; Prachayasittikul, S.; Ruchirawat, S.; Prachayasittikul, V. Synthesis, molecular docking, and QSAR study of sulfonamide-based indoles as aromatase inhibitors. Eur. J. Med. Chem. 2018, 143, 1604–1615.

(17) Pingaew, R.; Prachayasittikul, S.; Ruchirawat, S. Prachayasittikul, V. Synthesis and cytotoxicity of novel 2,2′-bis- and 2,2′,2″-tris-indolylmethanes-based benzagcarbonile analogs. Arch. Pharmacal Res. 2012, 35, 949–954.

(18) Pingaew, R.; Prachayasittikul, S.; Ruchirawat, S.; Prachayasittikul, V. Synthesis and structure-activity relationship of mono-indole-, bis-indole-, and tris-indole-based sulfonamides as potential anticancer agents. Mol. Diversity 2013, 17, 595–604.

(19) Cui, W.; Aoudate, A.; Wang, S.; Yu, Q.; Li, Y.; Yuan, S. Discovering anti-cancer drugs via computational methods. Front. Pharmacol. 2020, 11, 733.
