Contents lists available at ScienceDirect



Arabian Journal of Chemistry



journal homepage: www.ksu.edu.sa

# Design, synthesis and *in vitro* antitumor activity of $17\beta$ -estradiol-amino acid derivatives

Yu-qing Zhou<sup>a,b,1</sup>, Shi-chao Tian<sup>a,b,1</sup>, Li-xin Sheng<sup>a,b</sup>, Li-qiong Zhang<sup>a,b</sup>, Jing-jing Liu<sup>a,b</sup>, Wei-bin Mo<sup>b,c,\*</sup>, Quan-de Wang<sup>a,b,\*</sup>, Ke-guang Cheng<sup>a,b,d,\*</sup>

<sup>a</sup> School of Chemistry and Pharmaceutical Sciences, Guangxi Normal University, Guilin 541004, China

<sup>b</sup> State/Ministry of Education of China Key Laboratory for Chemistry and Molecular Engineering of Medicinal Resources, Guangxi Normal University, Guilin 541004,

China

<sup>c</sup> College of Physical and Health Education, Guangxi Normal University, Guilin 541006, China

<sup>d</sup> State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing 211198, China

ARTICLE INFO

Keywords: 17β-Estradiol Apoptosis Antitumor activity MDA-MB-231 cells

## ABSTRACT

A total of 28 derivatives were designed and synthesized, including 18 estradiol-amino acid (17 $\beta$ -OH) derivatives (**I-2** ~ **I-19**) and 10 estradiol-amino acid (16-C) derivatives (**II-2** ~ **II-11**). Amino acids used in synthesis include *L*-phenylalanine, *L*-serine, *L*-leucine, glycine, *L*-proline and *L*-glutamine. The results showed that **I-8** ~ **I-12** of 17 $\beta$ -OH conjugated amino acids in 17 $\beta$ -estradiol had good *in vitro* antitumor activity, indicating that the introduction of amino acids could improve the anti-proliferative effect. Among them, 3-benzyloxy-Estra-1,3,5 (10)-triene-17 $\beta$ -ol pyrrolidine-2-carboxylate hydrochloride (**I-12**) showed the best inhibitory activity against breast cancer (MDA-MB-231) cells (IC<sub>50</sub> = 4.89 µM). Further studies on apoptosis revealed that **I-12** could regulate the expression of apoptosis-related proteins in MDA-MB-231 cells. Network pharmacology and molecular docking analysis revealed that **I-12** may act on targets such as mechanistic target of rapamycin kinase (mTOR), estrogen receptor 1 (ESR1), pyruvate dehydrogenase kinase (PDK2) targets and their related pathways.

# 1. Introduction

With the incidence and mortality rates of cancer increasing year by year worldwide, it has become a major public health problem that seriously threatens the health of the population (Bray et al., 2012; Xie et al., 2015; Torre et al., 2016; Allemani et al., 2018; Bray et al., 2018). Breast cancer is the most common cancer in women (Veronesi et al., 2005; Torre et al., 2015; Chen et al., 2016; Siegel et al., 2018). Breast cancer is mainly treated by surgery and traditional chemotherapy. Since most anticancer drugs are characterized by drug resistance, high

toxicity, poor water solubility, low bioavailability and low selectivity, there is an urgent need to develop more novel and effective anticancer drugs.  $17\beta$ -Estradiol (Fig. 1) is an excellent candidate for steroidal estrogenic drugs, most likely as a treatment for breast cancer. It has good antitumor activity, such as in the treatment of metastatic breast cancer and advanced prostate cancer (Wright et al., 2014; Shull et al., 2018). It has also been in combination with other drugs for the treatment of related diseases (Stander et al., 2011).

It is a good research direction to design and synthesize a series of  $17\beta$ -estradiol derivatives for antitumor, aiming to achieve high

Peer review under responsibility of King Saud University.

\* Corresponding authors.

<sup>1</sup> These authors contribute equally to this work.

https://doi.org/10.1016/j.arabjc.2023.105539

Received 28 June 2023; Accepted 5 December 2023 Available online 11 December 2023



*Abbreviations*: mTOR, mechanistic target of rapamycin kinase; ESR1, estrogen receptor 1; PDK2, pyruvate dehydrogenase kinase; KIT, KIT proto-oncogene, receptor tyrosine kinase; RXRA, Retinoid X Receptor Alpha; MDA-MB-231, human breast cancer cells; MCF-7, human breast cancer cells; SKOV-3, human ovarian cancer cells; T24, human bladder cancer cells; BEL-7402, human liver cancer cells; HL-7702, human normal liver cells; MTT, 3-(4,5-Dimethylthiazol-2-yl) –2,5-diphenyltetrazolium bromide; Boc, *tert*-butoxycarbonyl; Cbz, carbobenzyloxy; DMF, *N*,*N*-dimethylformamide; DCC, dicyclohexylcarbodiimide; DMAP, 4-dimethylaminopyridine; THF, tetrahydrofuran; PCC, pyridinium chlorochromate; TFA, trifluoroacetic acid; PE, petroleum ether; DCM, dichloromethane; rt, room temperature; STITCH, search tool for interacting chemicals; SEA, similarity ensemble approach; DAVID, the database for annotation, visualization and integrated discovery; KEGG, kyoto encyclopedia of genes and genomes; DisGeNET, disease gene network; GEO, gene expression omnibus.

E-mail addresses: mogaol@163.com (W.-b. Mo), quande\_wang@163.com (Q.-d. Wang), kgcheng2008@163.com (K.-g. Cheng).

<sup>1878-5352/© 2023</sup> The Authors. Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

antitumor activity. Few 17 $\beta$ -estradiol derivatives are known to have good antitumor effects. The mitomycin C-estradiol benzoate coupling (Ishiki et al., 1997; Ishiki et al., 2004) was considered to be a potent hormone-drug coupled antitumor compound, but still had some toxicity. Irene B. Sorvik *et al.* examined the effects of novel 17 $\beta$ -estradiol analogs and 2-methoxyestradiol analogs on cells. These analogs had no or little estrogenic activity (Sorvik et al., 2018). Anke Löhndorf *et al.* (Lohndorf et al., 2021) synthesized a series of 2-methoxyestradiol derivatives that could antagonize store-operated Ca<sup>2+</sup> entry and subsequent signaling events in different T cells.

Amino acids are the basic units that constitute protein synthesis (Broer et al., 2017), participate in energy metabolism and material transport, and are closely related to a variety of vital activities in mammalian cells and organisms. Due to their advantages of good water solubility, low toxicity, biocompatibility, and certain biological activity, amino acids are widely used in the design and development of antitumor drugs (Vig et al., 2013). In 2010, Chrzanowski et al. designed a conjugate of melphalan and proline through imine bond, thus forming a good proteolytic substrate (Wu et al., 2010). In 2017, the N-phosphate alkyl bicyclic  $\beta$ -amino acid derivatives synthesized by Petar T. Todorov *et al.* showed good anti-proliferative effect on MDA-MB-231 cells (Pansare et al., 2017). It was also reported that thiazole derivatives containing amino acid groups showed good inhibitory activity against MCF-7 and BT-474 breast cancer cells (Harry et al., 2013). The coupling of amino acids with some natural products also showed superior antitumor activity than parent compounds, such as oleanolic acid-amino acid derivatives (van Meerloo et al., 2011) and celastrol-amino acid derivatives (Zhao et al., 2019).

In order to improve the water solubility and bioavailability of the compounds and to reduce the side effects, we synthesized new derivatives by linking  $17\beta$ -estradiol with amino acids. In this work, *L*-phenylalanine, *L*-serine, *L*-leucine, glycine, *L*-proline and *L*-glutamine were introduced into  $17\beta$ -estradiol using 3-benzyloxy-Estra-1,3,5 (10)-Triene- $17\beta$ -ol (Prokai et al., 2001) (I-1) and 3-benzyloxy- $16\alpha$ -bromo-Estra-1,3,5 (10)-Triene-17 $\beta$ -estradiol-amino acid derivatives. The anti-tumor activity of these derivatives was studied by MTT assay *in vitro* (Yin et al., 2022).

## 2. Experimental

#### 2.1. Materials and instruments

Unless otherwise indicated, all reagents and anhydrous solvents were purchased from commercial sources and used without purification. <sup>1</sup>H NMR and <sup>13</sup>C NMR spectra were recorded on a Bruker AV-500/400 spectrometer (Germany). Internal tetramethylsilane was used as chemical shift standard. HRMS was recorded on a Thermo Scientific Accela-Exactive High Resolution Accurate Mass Spectrometer (USA). Low and high resolution mass spectra (LRMS and HRMS) were given with electron impact mode. Melting points were measured on an RY-1 melting point apparatus (China). Column chromatography was performed using silica gel (300–400 mesh, Qingdao Haiyang Chemical Co., Ltd., China).

# 2.2. General procedure I for esterification

DCC (1.5 eq) and DMAP (0.3 eq) were added to a solution of I-1 (0.55 mmol, 1 eq) and amino acid protected by *N*-Boc/Cbz (1.5 eq) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (5 mL). After stirring at room temperature for 12 h, the reaction solution was filtrated, and the filter residue was washed by CH<sub>2</sub>Cl<sub>2</sub> (5 mL  $\times$  3). The filtrate was concentrated in vacuum, the residue was dispersed in EtOAc (25 mL), washed with saturated saline (10 mL  $\times$  3), dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered, concentrated, and purified by silica gel column chromatography.

# 2.3. General procedure II for de-boc reaction

Ethyl acetate solution of saturated hydrogen chloride (3 mL) was added to a solution of Boc-protected compound (0.66 mmol, 1 eq) in EtOAc (5 mL). The reaction solution was stirred at rt. for 2 h, and concentrated by rotary evaporation under reduced pressure. The residue was rotated twice with EtOAc, and then stirred in EtOAc (10 mL) for 1 h, filtered, and washed by EtOAc to get white solid.

# 2.4. General procedure III for debenzylation

A mixture of benzyl protected derivative (1 eq.), ammonium formate (10 eq.) and 10 % Pd/C (10 % weight) in MeOH/THF (4:1, V/V, 5 mL) was stirred at reflux for 12 h. The reaction mixture was filtered through Celite, and the insoluble substance was washed with MeOH (10 mL  $\times$  3). The filtrate was concentrated in vacuum to give an off-white solid, which was diluted in EtOAc (3 mL), and added ethyl acetate solution of saturated hydrogen chloride (2 mL). After stirring at room temperature for 8 h, the mixture was concentrated by rotary evaporation under reduced pressure. The residue was rotated three times with EtOAc, and then stirred in EtOAc (5 mL) at room temperature for 1 h, filtered, and washed by EtOAc to get white solid.

## 2.5. General procedure IV for esterification

To a solution of 3-benzyloxy- $16\alpha$ -bromo-Estra-1,3,5 (10)-Triene-17one (**II-1**, 0.100 g, 0.228 mmol, 1 eq) in DMF (5.0 mL), K<sub>2</sub>CO<sub>3</sub> (0.035 g, 0.25 mmol, 1.1 eq), KI (0.011 g, 0.068 mmol, 0.3 eq) and *N*-Boc protected amino acid (0.274 mmol, 1.2 eq) were added. After stirring at 60 °C for 12 h, the mixture was diluted with 2 N HCl (50 mL) and extracted with EtOAc (3 × 35 mL). The combined organic layers were washed successively with saturated aqueous NaHCO<sub>3</sub>, H<sub>2</sub>O, and saturated saline, dried over MgSO<sub>4</sub>, filtered, and concentrated. The residue was purified by silica gel column chromatography.

## 2.6. Synthesis of I-2 $\sim$ I-19 and II-2 $\sim$ II-11

#### 2.6.1. Compound I-2

According to the general procedure I for esterification, I-1 (0.200 g, 0.550 mmol) and *N*-Boc-*L*-phenylalanine (0.219 g, 0.825 mmol) were carried out to obtain I-2. Yield: 0.811 g, 80 %, white solid. Rf = 0.40 (PE: EtOAc = 6: 1), m.p. 120–122 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.77 (s, 3H, CH<sub>3</sub>), 1.27 – 2.30 (m, 13H), 1.42 (s, 9 H, 3 × CH<sub>3</sub>), 2.82 –



Fig. 1. Structures of 17β-estradiol, 3-benzyloxy-Estra-1,3,5 (10)-Triene-17β-ol (I-1) and 3-benzyloxy-16α-bromo-Estra-1,3,5 (10)-Triene-17-one (II-1).

2.87 (m, 2 H, H-6), 3.04 – 3.16 (m, 2 H, ArCH<sub>2</sub>), 4.60 (q, 1 H, J = 6.2 Hz, H-17), 4.67 (t, 1 H, J = 8.2 Hz, NCH), 4.98 (d, 1 H, J = 8.3 Hz, NH), 5.03 (s, 2 H, OCH<sub>2</sub>Ar), 6.72 (d, 1 H, J = 2.3 Hz, H-4), 6.80 (dd, 1 H, J = 2.5, 8.6 Hz, H-2), 7.16–7.44 (m, overlapping, 11*H*, H-1 and Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 12.2, 23.4, 26.3, 27.3, 27.7, 28.5, 29.9, 36.9, 38.6, 38.8, 43.1, 43.9, 49.8, 54.7, 70.1, 80.0, 84.1, 112.4, 115.0, 126.5, 127.1, 127.6, 128.0, 128.7, 129.5, 132.8, 136.3, 137.4, 138.1, 155.2, 156.9, 172.1. HRMS (ESI) m/z: calcd for C<sub>39</sub>H<sub>47</sub>NO<sub>5</sub>Na [M + Na]<sup>+</sup> 632.3352, found 632.3350.

# 2.6.2. Compound I-3

According to the general procedure I for esterification, I-1 (0.500 g, 1.38 mmol) and *N*-Boc-*O*-Bn-*L*-serine (0.611 g, 2.07 mmol) were carried out to obtain I-3. Yield: 0.771 g, 87 %, white solid. Rf = 0.34 (PE: EtOAc = 6: 1), m.p. 104–106 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.77 (s, 3H, CH<sub>3</sub>), 1.25 – 2.24 (m, 13H), 1.46 (s, 9 H, 3 × CH<sub>3</sub>), 2.84 – 2.88 (m, 2 H, H-6), 3.72 (dd, 1 H, *J* = 2.8, 9.2 Hz, OCH), 3.94 (dd, 1 H, *J* = 2.6, 9.2 Hz, OCH), 4.47 (m, 1 H, NCH), 4.53 (m, 2 H, OCH<sub>2</sub>Ar), 4.83 (t, 1 H, *J* = 8.6 Hz, H-17), 5.04 (s, 2 H, OCH<sub>2</sub>Ar), 5.45 (d, 1 H, *J* = 8.9 Hz, NH), 6.73 (d, 1 H, *J* = 2.4 Hz, H-4), 6.79 (dd, 1 H, *J* = 2.5, 8.5 Hz, H-2), 7.20 (d, 1 H, *J* = 8.6 Hz, H-1), 7.28 – 7.44 (m, 10*H*, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 12.1, 23.4, 26.3, 27.3, 27.4, 28.5, 29.9, 36.9, 38.7, 43.4, 43.9, 49.8, 54.3, 70.1, 70.8, 73.6, 80.0, 83.7, 112.4, 115.0, 126.5, 127.6, 127.9, 128.0, 128.5, 128.7, 132.9, 137.4, 137.6, 138.1, 155.7, 156.9, 170.8. HRMS (ESI) *m*/*z*: calcd for C<sub>40</sub>H<sub>49</sub>NO<sub>6</sub>Na [M + Na]<sup>+</sup> 662.3458, found 662.3444.

#### 2.6.3. Compound I-4

According to the general procedure I for esterification, I-1 (0.600 g, 1.66 mmol) and *N*-Boc-*L*-leucine monohydrate (0.531 g, 2.49 mmol) were carried out to obtain I-4. Yield: 0.838 g, 88 %, white solid. Rf = 0.55 (PE: EtOAc = 4: 1), m.p. 115–117 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.84 (s, 3H, CH<sub>3</sub>), 0.96 (s, 3H, CH<sub>3</sub>), 0.98 (s, 3H, CH<sub>3</sub>), 1.24 – 2.31 (m, 16H), 1.45 (s, 9H, 3 × CH<sub>3</sub>), 2.84 – 2.88 (m, 2H, H-6), 4.32 (q, 1H, *J* = 8.5 Hz, H-17), 4.72 (t, 1H, *J* = 8.5 Hz, NCH), 4.93 (d, 1H, *J* = 8.6 Hz, NH), 5.03 (s, 2H, OCH<sub>2</sub>Ar), 6.72 (d, 1H, *J* = 2.5 Hz, H-4), 6.78 (dd, 1H, *J* = 2.6, 8.5 Hz, H-2), 7.20 (d, 1H, *J* = 8.6 Hz, H-1), 7.30 – 7.44 (m, 5H, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 12.3, 22.3, 23.0, 23.4, 25.0, 26.3, 27.3, 27.6, 28.5, 29.9, 37.0, 38.7, 42.4, 43.2, 43.9, 49.9, 52.4, 70.1, 79.8, 83.7, 112.5, 115.0, 126.5, 127.6, 128.0, 128.7, 132.8, 137.4, 138.1, 155.5, 156.9, 173.7. HRMS (ESI) *m*/*z*: calcd for C<sub>36</sub>H<sub>49</sub>NO<sub>5</sub>Na [M + Na]<sup>+</sup> 598.3508, found 598.3501.

# 2.6.4. Compound I-5

According to the general procedure I for esterification, I-1 (0.600 g, 1.66 mmol) and *N*-Boc-glycine (0.436 g, 2.49 mmol) were carried out to obtain I-5. Yield: 0.766 g, 89 %, white solid. Rf = 0.25 (PE: EtOAc = 5: 1), m.p. 116–117 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.83 (s, 3H, CH<sub>3</sub>), 1.25 – 2.31 (m, 13H), 1.46 (s, 9 H, 3 × CH<sub>3</sub>), 2.84 – 2.88 (m, 2 H, H-6), 3.93 (d, 2 H, *J* = 5.4 Hz, NCH<sub>2</sub>CO), 4.76 (t, 1 H, *J* = 8.3 Hz, H-17), 5.01 (s, 1 H, NH), 5.03 (s, 2 H, OCH<sub>2</sub>Ar), 6.72 (d, 1 H, *J* = 2.5 Hz, H-4), 6.78 (dd, 1 H, *J* = 2.6, 8.6 Hz, H-2), 7.21 (d, 1 H, *J* = 8.6 Hz, H-1), 7.30 – 7.44 (m, 5 H, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 12.2, 23.4, 26.3, 27.3, 27.6, 28.5, 29.9, 37.0, 38.7, 42.7, 43.2, 43.9, 49.8, 70.1, 80.1, 83.9, 112.4, 115.0, 126.5, 127.6, 128.0, 128.7, 132.8, 137.4, 138.0, 155.8, 156.9, 170.5. HRMS (ESI) *m*/*z*: calcd for C<sub>32</sub>H<sub>41</sub>NO<sub>5</sub>Na [M + Na]<sup>+</sup> 542.2882, found 542.2876.

## 2.6.5. Compound I-6

According to the general procedure I for esterification, **I-1** (0.600 g, 1.66 mmol) and *N*-Boc-*L*-proline (0.536 g, 2.49 mmol) were carried out to obtain **I-6**. Yield: 0.851 g, 92 %, white solid. Rf = 0.45 (PE: EtOAc = 4: 1), m.p. 121–123 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.83 (d, 3H, J = 6.8, CH<sub>3</sub>), 1.26 – 2.29 (m, 17H), 1.43 (s, 9H, 3 × CH<sub>3</sub>), 2.84 – 2.87 (m, 2H, H-6), 3.37 – 3.60 (m, 2H, NCH<sub>2</sub>), 4.22 – 4.36 (m, 1H, COCH), 4.69 – 4.81 (m, 1H, H-17), 5.03 (s, 2H, OCH<sub>2</sub>), 6.72 (s, 1H, H-4), 6.79 (d, 1H, *J* 

= 8.5 Hz, H-2), 7.20 (d, 1H, J = 8.6 Hz, H-1), 7.30 – 7.44 (m, 5H, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 12.2, 22.3, 23.0, 23.4, 23.7, 24.4, 26.3, 27.3, 27.5, 27.7, 28.5, 28.6, 29.9, 30.4, 31.2, 37.0, 38.7, 43.4, 43.9, 46.6, 49.9, 59.5, 70.1, 80.0, 83.2, 112.5, 115.0, 126.5, 127.6, 128.0, 128.7, 133.0, 137.4, 138.0, 154.4, 156.9, 176.2. HRMS (ESI) *m/z*: calcd for C<sub>35</sub>H<sub>45</sub>NO<sub>5</sub>Na [M + Na]<sup>+</sup> 582.3195, found 582.3186.

#### 2.6.6. Compound I-7

According to the general procedure I for esterification, I-1 (0.300 g, 0.83 mmol) and *N*-Cbz-*L*-glutamine (0.349 g, 1.25 mmol) were carried out to obtain I-7. Yield: 0.363 g, 70 %, white solid, Rf = 0.35 (CH<sub>2</sub>Cl<sub>2</sub>: CH<sub>3</sub>OH = 10: 1), m.p. 118–120 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.82 (s, 3H, CH<sub>3</sub>), 0.87 – 0.90 (m, 1H, H-14), 1.26 – 2.37 (m, 16H), 2.81 – 2.88 (m, 2H, H-6), 4.37 – 4.42 (m, 1H, H-17), 4.75 (t, 1H, *J* = 8.4 Hz, NCH), 5.03 (s, 2H, OCH<sub>2</sub>Ar), 5.11 (s, 2H, OCH<sub>2</sub>Ar), 5.46 (s, 1H, NH), 5.68 (d, 1H, *J* = 8.0 Hz, NH), 5.93 (s, 1H, NH), 6.72 (d, 1H, *J* = 2.6 Hz, H-4), 6.78 (dd, 1H, *J* = 2.6, 8.5 Hz, H-2), 7.20 (d, 1H, *J* = 8.6 Hz, H-1), 7.30 – 7.44 (m, 10H, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 12.3, 23.4, 26.3, 27.3, 27.6, 29.0, 29.8, 29.9, 31.9, 37.0, 38.6, 43.3, 43.9, 49.8, 53.8, 67.3, 70.1, 84.3, 112.5, 115.0, 126.5, 127.6, 128.0, 128.3, 128.4, 128.7, 132.7, 136.3, 137.4, 138.0, 156.6, 156.9, 172.0, 174.4. HRMS (ESI) *m/z*: calcd for C<sub>38</sub>H<sub>44</sub>N<sub>2</sub>O<sub>6</sub>Na [M + Na]<sup>+</sup> 647.3097, found 647.3086.

#### 2.6.7. Compound I-8

According to the general procedure II for de-boc reaction, **I-2** (0.400 g, 0.656 mmol) were carried out to obtain **I-8**. Yield: 0.296 g, 89 %, white solid. Rf = 0.40 (PE: EtOAc = 1: 1), m.p. 96–97 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.78 (s, 3H, CH<sub>3</sub>), 1.26 – 2.30 (m, 13H), 2.82 – 2.86 (m, 2 H, H-6), 2.89 (dd, 1 H, *J* = 7.8, 13.5 Hz, ArCH<sub>a</sub>), 3.09 – 3.14 (dd, 1 H, *J* = 5.5, 13.5 Hz, ArCH<sub>b</sub>), 3.76 (dd, 1 H, *J* = 5.6, 7.7 Hz, NCH), 4.72 (t, 1 H, *J* = 8.9 Hz, H-17), 5.03 (s, 2 H, OCH<sub>2</sub>Ar), 6.72 (d, 1 H, *J* = 2.6 Hz, H-4), 6.78 (dd, 1 H, *J* = 2.6, 8.5 Hz, H-2), 7.19 – 7.44 (m, overlapping, 11*H*, Ar-H and H-1). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 12.2, 23.4, 26.3, 27.3, 27.7, 29.9, 36.9, 38.6, 41.4, 43.1, 43.9, 49.8, 56.0, 70.1, 83.4, 112.4, 114.9, 126.5, 126.9, 127.6, 128.0, 128.66, 128.71, 129.4, 132.8, 137.3, 137.4, 138.0, 156.9, 175.2. HRMS (ESI) *m*/*z*: calcd for C<sub>34</sub>H<sub>39</sub>NO<sub>3</sub> [M + H]<sup>+</sup>510.3008, found 510.3001.

#### 2.6.8. Compound I-9

According to the general procedure II for de-boc reaction, **I-3** (0.400 g, 0.625 mmol) were carried out to obtain **I-9**. Yield: 0.318 g, 88 %, white solid. Rf = 0.35 (PE: EtOAc = 1: 1), m.p. 158–160 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.74 (d, 3H, J = 6.7 Hz, CH<sub>3</sub>), 1.22 – 2.19 (m, 13H), 2.76 – 2.77 (m, 2 H, H-6), 3.87 – 3.95 (m, 2 H, OCH<sub>2</sub>), 4.35 (t, 1 H, J = 2.8 Hz, NCH), 4.44 – 4.49 (m, 1 H, OCH<sub>a</sub>Ar), 4.57 – 4.62 (m, 1 H, OCH<sub>b</sub>Ar), 4.78 (t, 1 H, J = 8.3 Hz, H-17), 5.04 (s, 2 H, OCH<sub>2</sub>Ar), 6.70 (d, 1 H, J = 2.6 Hz, H-4), 6.76 (dd, 1 H, J = 2.6, 8.5 Hz, H-2), 7.15 (d, 1 H, J = 8.7 Hz, H-1), 7.29 – 7.43 (m, 10H, Ar-H), 8.75 (s, 3 H, NH<sub>2</sub>HCl). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 11.7, 22.8, 25.7, 27.3, 26.7, 26.8, 29.2, 36.1, 38.1, 42.9, 43.1, 48.8, 52.5, 67.8, 69.0, 72.7, 83.6, 112.3, 114.5, 126.2, 127.5, 127.7, 127.8, 127.9, 128.2, 128.3, 128.4, 132.1, 137.2, 137.4, 156.2, 167.7. HRMS (ESI) m/z: calcd for C<sub>35</sub>H<sub>42</sub>ClNO<sub>4</sub> [M–Cl]<sup>+</sup> 540.3114, found 540.3107.

#### 2.6.9. Compound I-10

According to the general procedure II for de-boc reaction, **I-4** (0.400 g, 0.695 mmol) were carried out to obtain **I-10**. Yield: 0.340 g, 96 %, white solid. Rf = 0.45 (PE: EtOAc = 1: 1), m.p. 208–210 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.81 (s, 3H, CH<sub>3</sub>), 0.92 (dd, 6H, *J* = 5.1, 6.3 Hz, 2 × CH<sub>3</sub>), 1.25 – 2.28 (m, 16H), 2.75 – 2.77 (m, 2 H, H-6), 3.89 (t, 1 H, *J* = 7.2 Hz, NCH), 4.68 (t, 1 H, *J* = 8.2 Hz, H-17), 5.04 (s, 2 H, OCH<sub>2</sub>Ar), 6.70 (d, 1 H, *J* = 2.6 Hz, H-4), 6.75 (dd, 1 H, *J* = 2.6, 8.6 Hz, H-2), 7.16 (d, 1 H, *J* = 8.6 Hz, H-1), 7.28–7.43 (m, 5 H, Ar-H), 8.53 (brs, 3H, NH<sub>2</sub>HCl). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 12.0, 21.9, 22.4, 22.8, 24.0, 25.7, 26.7, 27.1, 29.2, 36.2, 38.1, 42.6, 43.2, 48.7, 50.7

69.0, 83.8, 112.3, 114.6, 126.2, 127.5, 127.7, 128.4, 132.1, 137.4, 156.2, 170.1. HRMS (ESI) m/z: calcd for  $C_{31}H_{42}ClNO_3~[M-Cl]^+$  476.3165, found 476.3159.

#### 2.6.10. Compound I-11

According to the general procedure II for de-boc reaction, **I-5** (0.400 g, 0.770 mmol) were carried out to obtain **I-11**. Yield: 0.324 g, 92 %, white solid. Rf = 0.21 (PE: EtOAc = 1: 1), m.p. 241–243 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.80 (s, 3H, CH<sub>3</sub>), 1.25 – 2.27 (m, 13H), 2.75 – 2.77 (m, 2 H, H-6), 3.79 and 3.80 (d, each 1 H, J = 17.2 Hz, NCH<sub>2</sub>), 4.73 (t, 1 H, J = 8.2 Hz, H-17), 5.04 (s, 2 H, OCH<sub>2</sub>Ar), 6.70 (d, 1 H, J = 2.6 Hz, H-4), 6.74 (dd, 1 H, J = 2.6, 8.5 Hz, H-2), 7.17 (d, 1 H, J = 8.6 Hz, H-1), 7.29 – 7.43 (m, 5 H, Ar-H), 8.49 (s, 3 H, NH<sub>2</sub>HCl). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 11.8, 22.8, 25.7, 26.7, 27.0, 29.2, 36.2, 38.1, 42.8, 43.2, 48.9, 69.0, 83.5, 112.3, 114.5, 126.2, 127.5, 127.7, 128.4, 132.1, 137.4, 156.2, 167.5. HRMS (ESI) *m/z*: calcd for C<sub>27</sub>H<sub>34</sub>ClNO<sub>3</sub> [M–CI]<sup>+</sup> 420.2539, found 420.2530.

## 2.6.11. Compound I-12

According to the general procedure II for de-boc reaction, **I-6** (0.400 g, 0.715 mmol) were carried out to obtain **I-12**. Yield: 0.346 g, 98 %, white solid. Rf = 0.40 (PE: EtOAc = 1: 1), m.p. 252–254 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.80 (s, 3H, CH<sub>3</sub>), 1.19–2.35 (m, 17H), 2.78 (d, 2H, *J* = 4.3 Hz, H-6), 3.16 – 3.26 (m, 2H, NCH<sub>2</sub>), 4.37 (t, 1H, *J* = 7.5 Hz, COCH), 4.77 (t, 1H, *J* = 8.1 Hz, H-17), 5.04 (s, 2H, OCH<sub>2</sub>Ar), 6.70 (d, 1H, *J* = 2.3 Hz, H-4), 6.75 (dd, 1H, *J* = 2.5, 8.5 Hz, H-2), 7.16 (d, 1H, *J* = 8.6 Hz, H-1), 7.29 – 7.42 (m, 5H, Ar-H), 9.01 (brs, 1H, HCl), 10.33 (brs, 1H, NH). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 11.9, 22.8, 23.0, 25.7, 26.7, 26.8, 28.1, 29.2, 36.3, 38.1, 42.9, 43.2, 45.1, 48.9, 58.6, 69.0, 83.9, 112.3, 114.5, 126.2, 127.5, 127.7, 128.4, 132.1, 137.4, 156.2, 168.8. HRMS (ESI) *m/z*: calcd for C<sub>30</sub>H<sub>38</sub>ClNO<sub>3</sub> [M–Cl]<sup>+</sup> 460.2852, found 460.2846.

# 2.6.12. Compound I-13

According to the general procedure III for debenzylation, **I-8** (0.116 g, 0.228 mmol) were carried out to obtain **I-13**. Yield: 0.079 g, 76 %, white solid. Rf = 0.24 (PE: EtOAc = 1: 1), m.p. 232–233 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.56 (s, 3H, CH<sub>3</sub>), 1.19 – 2.19 (m, 13H), 2.68 – 2.71 (m, 2 H, H-6), 3.03 – 3.09 (m, 1 H, ArCH<sub>a</sub>), 3.23 – 3.28 (m, 1 H, ArCH<sub>b</sub>), 4.26 (d, 1 H, *J* = 7.4 Hz, NCH), 4.55 (t, 1 H, *J* = 8.3 Hz, H-17), 6.44 (d, 1 H, *J* = 2.2 Hz, H-4), 6.51 (dd, 1 H, *J* = 2.4, 8.3 Hz, H-2), 7.03 (d, 1 H, *J* = 8.5 Hz, H-1), 7.25 – 7.36 (m, 5 H, Ar-H), 8.72 (s, 3 H, NH<sub>2</sub>HCl), 9.07 (s, 1 H, Ar-OH). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 12.0, 23.2, 26.2, 27.3, 27.6, 29.5, 36.3, 36.6, 38.6, 42.8, 43.6, 49.2, 53.7, 84.5, 113.3, 115.4, 126.5, 127.8, 129.1, 129.8, 130.0, 135.5, 137.5, 155.5, 169.5. HRMS (ESI) *m/z*: calcd for C<sub>27</sub>H<sub>34</sub>ClNO<sub>3</sub> [M–Cl]<sup>+</sup> 420.2539, found 420.2529.

## 2.6.13. Compound I-14

According to the general procedure III for debenzylation, **I-9** (0.200 g, 0.350 mmol) were carried out to obtain **I-14**. Yield: 0.055 g, 40 %, white solid. Rf = 0.11 (PE: EtOAc = 1: 1), m.p. 184–186 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.76 – 0.81 (m, 3H, CH<sub>3</sub>), 1.19 – 2.22 (m, 13H), 2.66 – 2.73 (m, 2 H, H-6), 3.83 – 3.87 (m, 2 H, OCH<sub>2</sub>), 4.09 (t, 1 H, J = 3.3 Hz, NCH), 4.72 (t, 1 H, J = 8.1, H-17), 5.64 – 5.66 (m, 1 H, OH), 6.45 (d, 1 H, J = 2.3 Hz, H-4), 6.51 (dd, 1 H, J = 2.4, 8.4 Hz, H-2), 7.03 (d, 1 H, J = 8.4 Hz, H-1), 8.59 (brs, 3 H, NH<sub>2</sub>HCl), 9.12 (s, 1 H, Ar-OH). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 11.8, 22.8, 25.9, 26.9, 27.1, 29.1, 36.2, 38.3, 42.8, 43.0, 43.2, 48.9, 54.4, 59.7, 83.9, 112.8, 115.0, 126.1, 130.0, 137.0, 155.1, 168.0. HRMS (ESI) *m/z*: calcd for C<sub>21</sub>H<sub>30</sub>ClNO<sub>4</sub> [M–Cl]<sup>+</sup> 360.2175, found 360.2166.

## 2.6.14. Compound I-15

According to the general procedure III for debenzylation, I-10 (0.100 g, 0.195 mmol) were carried out to obtain I-15. Yield: 0.056 g, 68 %, white solid. Rf = 0.25 (PE: EtOAc = 1: 1), m.p. 242–245 °C. <sup>1</sup>H

NMR (400 MHz,  $C_2D_6SO$ )  $\delta$  (ppm) 0.81 (s, 3H, CH<sub>3</sub>), 0.91 (dd, 6H, J = 4.9, 6.2 Hz, 2 × CH<sub>3</sub>), 1.22 – 2.25 (m, 16H), 2.65 – 2.73 (m, 2 H, H-6), 3.90 (t, 1 H, J = 7.1 Hz, NCH), 4.68 (t, 1 H, J = 8.2 Hz, H-17), 6.45 (d, 1 H, J = 2.4 Hz, H-4), 6.51 (dd, 1 H, J = 2.5, 8.4 Hz, H-2), 7.03 (d, 1 H, J = 8.5 Hz, H-1), 8.64 (s, 3H, NH<sub>2</sub>HCl), 9.13 (s, 1 H, Ar-OH). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 12.0, 21.9, 22.5, 22.9, 24.1, 25.8, 26.9, 27.1, 29.1, 36.3, 38.3, 42.6, 43.2, 48.8, 50.7, 69.0, 84.0, 112.9, 115.0, 126.1, 130.0, 137.1, 155.1, 170.0. HRMS (ESI) m/z: calcd for C<sub>24</sub>H<sub>36</sub>ClNO<sub>3</sub> [M–Cl]<sup>+</sup> 386.2695, found 386.2684.

# 2.6.15. Compound I-16

According to the general procedure III for debenzylation, I-11 (0.200 g, 0.195 mmol) were carried out to obtain I-16. Yield: 0.113 g, 99 %, white solid. Rf = 0.07 (PE: EtOAc = 1: 1), m.p. 241–243 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.80 (s, 3H, CH<sub>3</sub>), 1.23 – 2.24 (m, 13H), 2.67 – 2.73 (m, 2 H, H-6), 3.75 – 3.84 (m, 2 H, NCH<sub>2</sub>), 4.73 (t, 1 H, J = 8.2 Hz, H-17), 6.45 (d, 1 H, J = 2.3 Hz, H-4), 6.51 (dd, 1 H, J = 2.4, 8.3 Hz, H-2), 7.04 (d, 1 H, J = 8.5 Hz, H-1), 8.45 (brs, 3 H, NH<sub>2</sub>HCl), 9.10 (s, 1 H, Ar-OH). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 11.9, 22.8, 25.9, 26.9, 27.0, 29.1, 36.3, 38.3, 42.8, 43.2, 48.9, 83.6, 112.8, 115.0, 126.1, 130.0, 137.1, 155.1, 167.6. HRMS (ESI) *m*/*z*: calcd for C<sub>20</sub>H<sub>28</sub>ClNO<sub>3</sub> [M–Cl]<sup>+</sup> 330.2069, found 330.2061.

## 2.6.16. Compound I-17

According to the general procedure III for debenzylation, **I-12** (0.200 g, 0.357 mmol) were carried out to obtain **I-17**. Yield: 0.113 g, 99 %, white solid. Rf = 0.15 (PE: EtOAc = 1: 1), m.p. 254–256 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.79 (s, 3H, CH<sub>3</sub>), 1.22 – 2.34 (m, 17H), 2.65 – 2.73 (m, 2H, H-6), 3.16 – 3.28 (m, 2H, NCH<sub>2</sub>), 4.36 (t, 1H, *J* = 8.0 Hz, COCH), 4.75 (t, 1H, *J* = 8.0 Hz, H-17), 6.45 (d, 1H, *J* = 2.4 Hz, H-4), 6.52 (dd, 1H, *J* = 2.4, 8.4 Hz, H-2), 7.03 (d, 1H, *J* = 8.5 Hz, H-1), 9.10 (s, 1H, Ar-OH), 9.73 (brs, 2H, NHHCl). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 11.9, 22.8, 23.0, 25.8, 26.8, 28.1, 29.1, 36.4, 38.3, 42.9, 43.2, 45.1, 48.9, 58.6, 69.0, 83.9, 112.8, 115.0, 126.0, 130.0, 137.0, 155.1, 168.8. HRMS (ESI) *m/z*: calcd for C<sub>23</sub>H<sub>32</sub>ClNO<sub>3</sub> [M–Cl]<sup>+</sup> 370.2382, found 370.2373.

## 2.6.17. Compound I-18

According to the general procedure III for debenzylation, **I-6** (0.200 g, 0.357 mmol) were carried out to obtain **I-18**. Yield: 0.165 g, 98 %, white solid. Rf = 0.15 (PE: EtOAc = 4: 1), m.p. 176–178 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.80 (d, 3H, J = 7.3 Hz, CH<sub>3</sub>), 1.44 (s, 6H, 2 × CH<sub>3</sub>), 1.46 (s, 3H, CH<sub>3</sub>), 1.25 – 2.28 (m, 18H), 2.80 – 2.81 (m, 2H, H-6), 3.36 – 3.60 (m, 2H, NCH<sub>2</sub>), 4.30 (ddd, 1H, J = 2.9 Hz, 10.3 Hz, 43.0 Hz, COCHN), 4.73 (dt, 1H, J = 8.4 Hz, 33.7 Hz, H-17), 6.58 (d, 1H, J = 2.6 Hz, H-4), 6.62 – 6.66 (m, 1H, H-2), 7.12 (d, 1H, J = 8.5 Hz, H-1). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 12.3, 23.4, 23.7, 26.29, 27.3, 27.8, 28.5, 28.62, 29.7, 31.2, 37.1, 38.69, 43.2, 43.9, 46.7, 49.8, 59.6, 80.4, 83.4, 112.92, 115.5, 126.5, 132.1, 138.2, 154.0, 154.2, 173.4. HRMS (ESI) m/z: calcd for C<sub>28</sub>H<sub>39</sub>NO<sub>5</sub> [M + H]<sup>+</sup> 470.2906, found 470.2891.

#### 2.6.18. Compound I-19

According to the general procedure III for debenzylation, **I-7** (0.158 g, 0.250 mmol) were carried out to obtain **I-19**. Yield: 0.049 g, 49 %, white solid. Rf = 0.39 (CH<sub>2</sub>Cl<sub>2</sub>: CH<sub>3</sub>OH = 10: 1), m.p. 282–285 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.78 (s, 3H, CH<sub>3</sub>), 1.23 – 2.42 (m, 17H), 2.67 – 2.74 (m, 2H, H-6), 4.16 (dd, 1H, *J* = 4.0, 8.9 Hz, NCH), 4.68 (t, 1H, *J* = 8.6 Hz, H-17), 6.43 (d, 1H, *J* = 2.3 Hz, H-4), 6.50 (dd, 1H, *J* = 2.4, 8.4 Hz, H-2), 7.04 (d, 1H, *J* = 8.4 Hz, H-1), 8.00 (s, 1H, NH), 9.00 (s, 1H, Ar-OH). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 11.9, 22.8, 24.9, 25.8, 26.8, 27.0, 29.0, 29.1, 36.4, 38.3, 42.8, 43.2, 49.0, 55.0, 82.5, 112.8, 115.0, 126.1, 130.1, 137.1, 155.0, 172.9, 177.0. HRMS (ESI) *m/z*: calcd for C<sub>23</sub>H<sub>32</sub>N<sub>2</sub>O<sub>4</sub> [M + Na]<sup>+</sup> 423.2260, found 423.2282.

## 2.6.19. Compound II-2

According to the general procedure IV for esterification, II-1 (0.100

g, 0.228 mmol) and *N*-Boc-*O*-Bn-*L*-serine (0.081 g, 0.274 mmol) were carried out to obtain **II-2**. Yield: 0.090 g, 60 %, yellow oily substance. Rf = 0.34 (PE: EtOAc = 3: 1). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.99 (s, 3H, CH<sub>3</sub>), 1.45 (s, 9H, 3 × CH<sub>3</sub>), 1.50 – 2.67 (m, 11*H*), 2.89 – 2.91 (m, 2H, H-6), 3.74 (dd, 1 H, *J* = 3.1, 9.5 Hz, OCH<sub>a</sub>), 3.90 (dd, 1 H, *J* = 3.3, 9.5 Hz, OCH<sub>b</sub>), 4.46 – 4.50 (m, 1 H, NCH), 4.56 (q, 2H, *J* = 11.9 Hz, OCH<sub>2</sub>Ar), 5.04 (s, 2H, OCH<sub>2</sub>Ar), 5.10 (t, 1 H, *J* = 8.3 Hz, H-16), 5.43 (d, 1 H, *J* = 8.9 Hz, NH), 6.75 (d, 1 H, *J* = 2.6 Hz, H-4), 6.80 (dd, 1 H, *J* = 2.6, 8.5 Hz, H-2), 7.20 (d, 1 H, *J* = 8.6 Hz, H-1), 7.27 – 7.44 (m, 10*H*, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 14.7, 25.8, 26.8, 28.5, 29.2, 29.7, 31.9, 37.6, 44.2, 45.1, 47.4, 54.1, 70.1, 70.3, 73.6, 75.5, 80.2, 112.6, 115.0, 126.4, 127.6, 127.9, 128.0, 128.6, 128.7, 132.1, 137.3, 137.8, 155.6, 157.1, 170.3, 213.6. HRMS (ESI) *m*/*z*: calcd for C<sub>40</sub>H<sub>47</sub>NO<sub>7</sub> [M + Na]<sup>+</sup> 676.3250, found 676.3240.

# 2.6.20. Compound II-3

According to the general procedure IV for esterification, **II-1** (0.200 g, 0.455 mmol) and *N*-Boc-*L*-phenylalanine (0.145 g, 0.546 mmol) were carried out to obtain **II-3**. Yield: 0.115 g, 40 %, yellowish solid. Rf = 0.25 (PE: EtOAc = 6: 1), m.p. 72–74 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 1.00 (s, 3H, CH<sub>3</sub>), 1.26 – 2.62 (m, 11*H*), 1.42 (s, 9H, 3 × CH<sub>3</sub>), 2.89 – 2.92 (m, 2H, H-6), 3.08 – 3.20 (m, 2H, ArCH<sub>2</sub>), 4.60 – 4.65 (m, 1 H, NCH), 4.98 (d, 1 H, *J* = 8.4 Hz, NH), 5.04 (s, 2H, OCH<sub>2</sub>Ar), 5.12 (t, 1 H, *J* = 8.3 Hz, H-16), 6.74 (d, 1 H, *J* = 2.5 Hz, H-4), 6.80 (dd, 1 H, *J* = 2.6, 8.6 Hz, H-2), 7.21 (d, 1 H, *J* = 8.6 Hz, H-1), 7.23 – 7.44 (m, 10*H*, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 14.7, 25.8, 26.8, 28.4, 29.1, 29.6, 31.8, 37.6, 38.2, 44.2, 45.1, 47.5, 54.3, 70.1, 75.3, 80.1, 112.6, 115.0, 126.4, 127.2, 127.6, 128.0, 128.7, 129.7, 132.0, 135.8, 137.3, 137.8, 155.2, 157.1, 171.5, 213.9. HRMS (ESI) *m*/*z*: calcd for C<sub>39</sub>H<sub>45</sub>NO<sub>6</sub> [M + Na]<sup>+</sup> 646.3145, found 646.3132.

#### 2.6.21. Compound II-4

According to the general procedure IV for esterification, **II-1** (0.300 g, 0.683 mmol) and *N*-Boc-*L*-proline (0.177 g, 0.820 mmol) were carried out to obtain **II-4**. Yield: 0.170 g, 43 %, white solid. Rf = 0.13 (PE: EtOAc = 5:1), m.p. 113–115 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 1.00 (s, 3H, CH<sub>3</sub>), 1.44 – 1.47 (m, 9H, 3 × CH<sub>3</sub>), 1.53 – 2.65 (m, 15H), 2.90 (d, 2H, *J* = 4.1 Hz, H-6), 3.34 – 3.61 (m, 2H, NCH<sub>2</sub>), 4.26 – 4.36 (m, 1H, COCH), 5.04 (s, 2H, OCH<sub>2</sub>Ar), 5.13 (t, 1H, *J* = 8.4 Hz, H-16), 6.74 (d, 1H, *J* = 2.4 Hz, H-4), 6.80 (d, 1H, *J* = 8.6 Hz, H-2), 7.20 (d, 1H, *J* = 8.6 Hz, H-1), 7.30 – 7.44 (m, 5 H, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 14.8, 23.7, 24.3, 25.8, 26.9, 28.5, 28.6, 29.1, 29.6, 30.1, 31.2, 31.9, 37.6, 44.2, 45.1, 46.5, 46.8, 47.5, 49.9, 59.0, 70.1, 74.9, 80.2, 112.6, 115.1, 126.4, 127.6, 128.0, 128.7, 132.1, 137.3, 137.9, 154.6, 157.1, 172.7, 214.6. HRMS (ESI) *m/z*: calcd for C<sub>35</sub>H<sub>43</sub>NO<sub>6</sub> [M + Na]<sup>+</sup> 596.2988, found 596.2979.

## 2.6.22. Compound II-5

According to the general procedure IV for esterification, **II-1** (0.300 g, 0.683 mmol) and *N*-Boc-glycine (0.144 g, 0.820 mmol) were carried out to obtain **II-5**. Yield: 0.105 g, 29 %, yellow solid. Rf = 0.25 (PE: EtOAc = 4: 1), m.p. 140–143 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 1.00 (s, 3H, CH<sub>3</sub>), 1.46 (s, 9H, 3 × CH<sub>3</sub>), 1.53 – 2.65 (m, 11*H*), 2.88 – 2.91 (m, 2H, H-6), 3.85 – 4.07 (m, 2H, NCH<sub>2</sub>CO), 5.04 (s, 2H, OCH<sub>2</sub>Ar), 5.11 (t, 1 H, *J* = 8.4 Hz, H-16), 6.74 (s, 1 H, H-4), 6.80 (dd, 1 H, *J* = 2.7, 8.5 Hz, H-2), 7.20 (d, 1 H, *J* = 8.8 Hz, H-1), 7.30 – 7.44 (m, 5 H, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 14.1, 25.7, 26.8, 28.4, 29.1, 29.6, 31.8, 37.6, 42.5, 44.2, 45.0, 47.5, 70.1, 75.4, 80.2, 86.5, 112.6, 115.0, 126.4, 127.6, 128.0, 128.7, 132.0, 137.3, 137.7, 155.8, 157.1, 170.0, 217.1. HRMS (ESI) *m/z*: calcd for C<sub>32</sub>H<sub>32</sub>NO<sub>6</sub> [M + Na]<sup>+</sup> 556.2675, found 556.2666.

#### 2.6.23. Compound II-6

According to the general procedure IV for esterification, **II-1** (0.300 g, 0.683 mmol) and *N*-Boc-*L*-leucine monohydrate (0.190 g, 0.820 mmol) were carried out to obtain **II-6**. Yield: 0.148 g, 37 %, yellow oily

liquid. Rf = 0.28 (PE: EtOAc = 4: 1), m.p. 72–74 °C. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 0.95 (s, 3H, CH<sub>3</sub>), 0.96 (s, 3H, CH<sub>3</sub>), 1.01 (s, 3H, CH<sub>3</sub>), 1.26 – 2.63 (m, 14H), 1.46 (s, 9H, 3 × CH<sub>3</sub>), 2.89 – 2.92 (m, 2H, H-6), 4.30 – 4.36 (m, 1H, NCH), 4.94 (d, 1H, J = 8.6 Hz, NH), 5.04 (s, 2H, OCH<sub>2</sub>Ar), 5.12 (t, 1H, J = 8.6 Hz, H-16), 6.74 (d, 1H, J = 2.5 Hz, H-4), 6.80 (dd, 1H, J = 2.6, 8.6 Hz, H-2), 7.20 (d, 1H, J = 8.6 Hz, H-1), 7.30 – 7.44 (m, 5H, Ar-H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  (ppm) 14.7, 22.0, 22.9, 24.9, 25.7, 26.8, 28.4, 29.0, 29.6, 31.8, 37.6, 41.8, 44.2, 45.0, 47.4, 52.1, 70.1, 75.0, 80.0, 112.6, 115.0, 126.3, 127.6, 128.0, 128.7, 132.0, 137.3, 137.7, 155.5, 157.1, 173.2, 213.9. HRMS (ESI) *m/z*: calcd for C<sub>36</sub>H<sub>47</sub>NO<sub>6</sub> [M + Na]<sup>+</sup> 612.3301, found 612.3293.

#### 2.6.24. Compound II-7

According to the general procedure II for de-boc reaction, II-2 (0.080 g, 0.122 mmol) were carried out to obtain II-7. Yield: 0.071 g, 99 %, white solid. Rf = 0.17 (PE: EtOAc = 3: 1), m.p. 102–104 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.91 – 0.92 (m, 3H, CH<sub>3</sub>), 1.33 – 2.45 (m, 11*H*), 2.79 – 2.81 (m, 2H, H-6), 3.83 – 3.91 (m, 1 H, OCH<sub>2</sub>), 4.41 (t, 1 H, J = 3.2 Hz, NCH), 4.49 – 4.59 (m, 2H, OCH<sub>2</sub>Ar), 5.04 (s, 2H, OCH<sub>2</sub>Ar), 5.24 (t, 1 H, J = 8.6 Hz, H-16), 6.72 (d, 1 H, J = 2.2 Hz, H-4), 6.76 (dd, 1 H, J = 2.4, 8.5 Hz, H-2), 7.18 (d, 1 H, J = 8.6 Hz, H-1), 7.28 – 7.43 (m, 10*H*, Ar-H), 8.82 (s, 3H, NH<sub>2</sub>HCl). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 14.6, 25.3, 26.2, 28.3, 29.1, 31.2, 36.9, 43.3, 43.6, 46.8, 52.4, 67.5, 69.0, 72.7, 75.8, 112.4, 114.7, 126.2, 127.6, 127.7, 127.8, 128.3, 128.5, 131.8, 137.4, 137.5, 156.3, 167.3, 213.0. HRMS (ESI) *m/z*: calcd for C<sub>35</sub>H<sub>40</sub>ClNO<sub>5</sub> [M–CI]<sup>+</sup> 554.2906, found 554.2895.

#### 2.6.25. Compound II-8

According to the general procedure II for de-boc reaction, **II-3** (0.080 g, 0.128 mmol) were carried out to obtain **II-8**. Yield: 0.066 g, 92 %, white solid. Rf = 0.09 (PE: EtOAc = 6: 1), m.p. 296–298 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.83 – 0.88 (m, 3H, CH<sub>3</sub>), 1.31 – 2.42 (m, 11*H*), 2.78 – 2.80 (m, 2H, H-6), 3.25 (dd, 1 H, *J* = 4.8, 14.2 Hz, ArCH<sub>a</sub>), 3.14 (dd, 1 H, *J* = 7.7, 14.2 Hz, ArCH<sub>b</sub>), 4.31 – 4.34 (m, 1 H, NCH), 5.04 (s, 2H, OCH<sub>2</sub>Ar), 5.21 (t, 1 H, *J* = 8.6 Hz, H-16), 6.71 (d, 1 H, *J* = 2.5 Hz, H-4), 6.75 (dd, 1 H, *J* = 2.4, 8.5 Hz, H-2), 7.17 (d, 1 H, *J* = 8.7 Hz, H-1), 7.24 – 7.43 (m, 10*H*, Ar-H), 8.89 (brs, 3H, NH<sub>2</sub>HCl). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 14.5, 25.2, 26.2, 28.1, 29.1, 31.2, 35.6, 36.8, 43.2, 43.6, 46.7, 53.0, 69.0, 75.4, 112.4, 114.6, 126.2, 127.2, 127.5, 127.7, 127.9, 128.4, 128.5, 129.7, 131.8, 134.7, 137.4, 156.3, 168.4, 174.4, 213.0. HRMS (ESI) *m*/*z*: calcd for C<sub>34</sub>H<sub>38</sub>ClNO<sub>4</sub> [M–Cl]<sup>+</sup> 524.2801, found 524.2795.

## 2.6.26. Compound II-9

According to the general procedure II for de-boc reaction, **II-4** (0.098 g, 0.171 mmol) were carried out to obtain **II-9**. Yield: 0.061 g, 70 %, white solid. Rf = 0.09 (CH<sub>2</sub>Cl<sub>2</sub>: MeOH = 15: 1), m.p. 208–210 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.93 (s, 3H, CH<sub>3</sub>), 1.34 – 2.38 (m, 14H), 2.80 – 2.82 (m, 2H, H-6), 3.16 – 3.29 (m, 2H, NCH<sub>2</sub>), 4.43 – 4.47 (m, 1H, NCH), 5.05 (s, 2H, OCH<sub>2</sub>Ar), 5.25 – 5.29 (m, 1H, H-16), 6.72 (d, 1H, *J* = 2.5 Hz, H-4), 6.76 (dd, 1H, *J* = 2.6, 8.6 Hz, H-2), 7.17 (d, 1H, *J* = 8.6 Hz, H-1), 7.29 – 7.43 (m, 5H, Ar-H), 9.33 – 10.30 (m, 2H, NH<sub>2</sub>). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 14.6, 22.9, 25.2, 26.2, 27.9, 28.1, 29.1, 31.1, 36.9, 43.3, 43.6, 45.2, 46.7, 58.3, 69.0, 75.8, 112.4, 114.6, 126.2, 127.5, 127.7, 128.4, 131.7, 137.35, 137.38, 156.3, 168.3, 213.4. HRMS (ESI) *m*/*z*: calcd for C<sub>30</sub>H<sub>36</sub>ClNO<sub>4</sub> [M–Cl]<sup>+</sup> 474.2644, found 474.2626.

#### 2.6.27. Compound II-10

According to the general procedure II for de-boc reaction, **II-5** (0.054 g, 0.101 mmol) were carried out to obtain **II-10**. Yield: 0.032 g, 67 %, white solid. Rf = 0.13 (CH<sub>2</sub>Cl<sub>2</sub>: MeOH = 15: 1), m.p. 202–204 °C. <sup>1</sup>H NMR (600 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.92 (s, 3H, CH<sub>3</sub>), 1.33 – 2.49 (m, 11*H*), 2.80 – 2.83 (m, 2H, H-6), 3.89 (s, 2H, NCH<sub>2</sub>CO), 5.05 (s, 2H, OCH<sub>2</sub>Ar), 5.25 (t, 1 H, *J* = 8.6 Hz, H-16), 6.72 (d, 1 H, *J* = 2.3 Hz, H-4), 6.76 (dd, 1 H, *J* = 2.5, 8.6 Hz, H-2), 7.17 (d, 1 H, *J* = 8.6 Hz, H-1), 7.30 –

7.42 (m, 5 H, Ar-H), 8.55 (s, 3H, NH<sub>2</sub>HCl). <sup>13</sup>C NMR (150 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 14.5, 25.2, 26.2, 28.3, 29.1, 31.2, 36.8, 39.5, 43.2, 43.6, 46.7, 69.0, 75.5, 112.4, 114.6, 126.2, 127.5, 127.7, 128.4, 131.8, 137.36, 137.4, 156.3, 167.2, 213.3. HRMS (ESI) *m/z*: calcd for C<sub>27</sub>H<sub>32</sub>ClNO<sub>4</sub> [M–Cl]<sup>+</sup> 434.2331, found 434.2329.

#### 2.6.28. Compound II-11

According to the general procedure II for de-boc reaction, II-6 (0.069 g, 0.117 mmol) were carried out to obtain II-11. Yield: 0.040 g, 65 %, white solid. Rf = 0.11 (PE: EtOAc = 6: 1), m.p. 197–199 °C. <sup>1</sup>H NMR (400 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 0.88 – 0.98 (m, 9H, 3 × CH<sub>3</sub>), 1.34 – 2.37 (m, 14H), 2.80 – 2.82 (m, 2H, H-6), 3.97 – 4.01 (m, 1H, NCH), 5.05 (s, 2H, OCH<sub>2</sub>Ar), 5.26 – 5.30 (m, 1H, H-16), 6.72 (d, 1H, J = 2.2 Hz, H-4), 6.76 (dd, 1H, J = 2.4, 8.6 Hz, H-2), 7.17 (d, 1H, J = 8.6 Hz, H-1), 7.29 – 7.43 (m, 5H, Ar-H), 8.70 (s, 3H, NH<sub>2</sub>HCl). <sup>13</sup>C NMR (100 MHz, C<sub>2</sub>D<sub>6</sub>SO)  $\delta$  (ppm) 14.5, 21.8, 22.4, 23.8, 25.2, 26.2, 28.2, 29.1, 31.1, 36.9, 43.25, 43.5, 46.7, 47.6, 50.3, 69.0, 75.3, 112.4, 114.6, 126.17, 127.5, 127.7, 128.4, 131.8, 137.36, 137.38, 156.3, 169.4, 213.3. HRMS (ESI) m/z: calcd for C<sub>31</sub>H<sub>40</sub>ClNO<sub>4</sub> [M–Cl]<sup>+</sup> 490.2957, found 490.2941.

# 2.7. MTT assay

All cells were cultivated in Dulbecco's Modified Eagle Medium with 10 % Fetal Bovine Serum at 37 °C and 5 % CO<sub>2</sub>. The cells were seeded in 96-well plates (5 × 10<sup>3</sup> cells per well) and cultivated for 24 h. The compounds were treated with different concentrations (20, 10, 5, 2.5, 1.25  $\mu$ M) for 48 h. Each well was added 20  $\mu$ L MTT (5 %) solution and cultivated for 4 h, then removed the supernatant and added DMSO. Absorbance at 570 nm was recorded. The IC<sub>50</sub> value was calculated by SPSS software, and at least 3 parallel tests were performed.

## 2.8. Human apoptosis antibody assay

Apoptosis-related protein expression was detected and analyzed

using the Human Apoptosis Array Kit (ARY009, R&D Systems). Membranes containing immobilized apoptosis-related antibodies were sealed with bovine serum albumin for 1 h. Membranes were then incubated overnight at 2–8 °C with a cocktail of proteins extracted from cells and detection antibodies. The membranes were incubated with horseradish peroxidase labeled streptavidin for 30 min and images were visualized with chemiluminescent reagents. The average background signal was subtracted from each spot to obtain the average signal (pixel density) of paired replicate spots representing each apoptosis-related protein.

#### 2.9. Molecular docking

Structures of proteins mTOR (PDB: 4DRI), KIT (PDB: 6GQM), ESR1 (PDB: 1SJ0), PDK2 (PDB: 4MPC), and RXRA (PDB: 3OAP) were downloaded from the RCSB protein database (<u>https://www.rcsb.org/</u>). The structures of the protein molecules were opened with PyMOL, water molecules were removed, small molecule ligands and other molecules bound were removed, polar hydrogens were added, molecular docking experiments were performed using AutoDock software, the conformation with the best binding energy was selected, and the pictures were plotted using PyMOL.

## 3. Results and discussion

#### 3.1. Synthesis

#### 3.1.1. Synthesis of $17\beta$ -estradiol (17-OH)-amino acid derivatives

As shown in Scheme 1, I-1 (1 eq) was condensed with different amino acids (1.5 eq) protected by *N*-Boc/Cbz in the presence of DCC (1.5 eq) and DMAP (0.3 eq) to generate 17 $\beta$ -estradiol (17-OH)-amino acid conjugate compounds I-2 ~ I-7 (yield: 70 % to 92 %). Followed by deprotection in saturated HCl/EtOAc solution to afford compounds I-8 ~ I-12 (yield: 88 % to 98 %). Finally, debenzylation under HCOONH<sub>4</sub> (10 eq) and 10 % Pd/C (10 %, weight %) were carried out to get target



Scheme 1. Synthesis of  $17\rho$ -estradiol (17-OH)-amino acid derivatives. Reagents and conditions: (a) DCC/DMAP, CH<sub>2</sub>Cl<sub>2</sub>, rt. (b) Saturated HCl/EtOAc solution, EtOAc, rt. (c) Ammonium formate, 10 % Pd/C, MeOH/THF, reflux.

## compounds I-13 $\sim$ I-19 (yield: 40 % to 99 %).

#### 3.1.2. Synthesis of $17\beta$ -estradiol (16-C)-amino acid derivatives

As shown in Scheme 2, **II-1** was substituted by different *N*-Boc protected amino acids under KI and  $K_2CO_3$  to give compounds **II-2** ~ **II-6** (yield: 29 % to 60 %). Followed by deprotection in saturated HCl/EtOAc solution to afford compounds **II-7** ~ **II-11** (yield: 65 % to 99 %).

#### 3.2. Results of in vitro antitumor activity

The  $IC_{50}$  values of the tested compounds against tumor cells MDA-MB-231, MCF-7, SKOV-3, T24, BEL-7402 and normal cells HL-7702 were determined by MTT assay, using 5-fluorouracil (5-FU) as a positive control (Table 1).

The results showed that compounds I-8, I-9, I-10, I-11, and I-12 derived at the 17-OH position of  $17\beta$ -estradiol showed good inhibitory effects (4  $\sim$  20  $\mu$ M) against the tested tumor cells, but also had corresponding toxicity against HL-7702 normal cells. While compounds II-2 to **II-11** derived at position 16-C of  $17\beta$ -estradiol showed relatively no inhibitory activity. This difference could be viewed over the structure, indicating that  $17\beta$ -OH which oxidized to a carbonyl group maybe lead the inactivation. The hydrochloride form might have little effect on the antitumor activity (I-8 vs. I-9, I-10, I-11 and I-12). The 3-OH of compounds I-13  $\sim$  17 was bare and the 3-OH of compounds I-8  $\sim$  12 was protected by benzyl. The cytotoxicity of compounds I-13  $\sim$  17 was greatly reduced compared to compounds I-8  $\sim$  12. This indicated that the inhibitory activity of the compounds could be increased when the amino group was bare and the 3-OH was protected. Among them, compound I-12 (Fig. 2) (3-benzyloxy-Estra-1,3,5 (10)-triene- $17\beta$ -ol pyrrolidine-2-carboxylate hydrochloride) showed the best cytotoxicity against MDA-MB-231 cells with an IC<sub>50</sub> of 4.89  $\mu$ M. I-12 showed better cytotoxicity against MDA-MB-231 cells than estradiol derivatives 17*β*-{6-[(Pyridin-4-ylmethyl)amino]-hexyloxy}-estra-1,3,5(10)-triene-3-ol (IC\_{50} = 9.9  $\pm$  0.4  $\mu M$ ) and 16S-Bromo-3-benzyloxy-estra-1,3,5(10)triene-17-one (IC\_{50} = 5.3  $\pm$  0.2  $\mu M$ ) (Yin et al., 2023). Compound I-12 was selected for further mechanistic studies in antitumor activity on MDA-MB-231 cells.

## 3.3. Effect on apoptosis of MDA-MB-231 cell line

Compound **I-12** was selected for its expression on apoptosis-related proteins using the Proteome Analysis Tool (Human Apoptosis Array Kit). Fig. 3 showed the images at different exposure times (1–10 min). 6 min was the best exposure time, and the expression of apoptosis-related proteins was calculated based on its gray value (Fig. 4). The grayscale value is the color depth of the point in the black and white image, see Appendix A. for details of the grayscale value.

As shown in Figs. 3 and 4, compound I-12 could obviously cause differential expression of apoptosis related proteins on tested MDA-MB-231 cells, which indicating that it could induce apoptosis. Compared with the 0  $\mu$ M group, it was clearly exhibited that I-12 (5  $\mu$ M) could induce the upregulation of apoptosis-related proteins Bad, Bax, Cyto-chrome *c*, TRAIL R2/DR5, Fas/TNFRSF6/CD95, HIF-1 $\alpha$ , PON2, p21/CIP1/CDKN1A and Phospho-p53 (S46); while induce the down-regulation of cIAP, HO-2/HMOX2 and HSP27. This results suggested that I-12 may induce apoptosis in MDA-MB-231 cells through a mitochondria-mediated pathway.

# 3.4. Network pharmacological analysis and molecular docking

(see Appendix for details) A total of 106 targets for I-12 was found in Stitch, Swiss Target Prediction, SEA, PharmMapper, and TargetNet databases (PharmMapper: zscore > 0.5, TargetNet: prob > 0.5). See Appendix A for details of the 106 targets. These 106 targets were entered into DAVID for disease enrichment analysis, and breast cancer with a high score (enrichment score: 8.91) was selected as the study disease. Breast cancer genes were searched in GeneCard, DisGeNET and GEO databases, and 163 breast cancer disease genes were obtained by screening and summarizing (GeneCard: Relevance score > 70; Dis-GeNET: estrogen negative: Score gda  $\geq$  0.4, estrogen positive: Score\_gda  $\geq$  0.1; GEO: logFC > 0.2, high expression, logFC < -0.2, low expression). Following, these 163 disease genes were used to map in KEGG pathway mapping to select pathways with higher gene enrichment: Pathways in cancer, PI3K-Akt signaling pathway, Breast cancer pathway, EGFR tyrosine kinase inhibitor resistance, MAPK signaling pathway, Metabolic pathways, JAK-STAT signaling pathway and mTOR signaling pathway. Compound targets were then intersected with breast cancer disease genes using Venn diagrams to obtain four intersecting targets:



Scheme 2. Synthesis of 17*β*-estradiol (16-C)-amino acid derivatives. Reagents and conditions: (a) K<sub>2</sub>CO<sub>3</sub>, KI, DMF, 60 °C. (b) Saturated HCl/EtOAc solution, EtOAc, rt.

#### Table 1

IC<sub>50</sub> of tested  $17\beta$ -estradiol-amino acid derivatives on different cell lines.

Number	IC <sub>50</sub> (µM)					
	MDA-MB-231	SKOV-3	T24	BEL-7402	MCF-7	HL-7702
$17\beta$ -Estradiol	$13.97 \pm 1.08$	$45.63 \pm 5.01$	$12.14\pm2.07$	$111.51 \pm 23.29$	$13.89\pm0.16$	$12.94 \pm 1.30$
I-2	> 200	> 200	$46.54 \pm 5.59$	$92.51 \pm 17.69$	$99.81 \pm 18.38$	> 200
I-3	> 200	> 200	$79.33 \pm 40.31$	$183.48 \pm 46.35$	> 200	> 200
I-4	$\textbf{45.94} \pm \textbf{16.00}$	> 200	$102.74 \pm 48.89$	$104.08\pm27.83$	$38.78 \pm 19.25$	> 200
I-5	> 200	> 200	$40.32\pm3.95$	> 200	$55.78 \pm 8.75$	> 200
I-6	> 200	> 200	$69.18\pm17.44$	> 200	$177.38 \pm 69.07$	> 200
I-7	> 200	$\textbf{28.27} \pm \textbf{10.01}$	$58.68 \pm 8.58$	$30.42\pm5.05$	> 200	$35.50 \pm 1.96$
I-8	$14.55\pm0.68$	$19.64 \pm 2.48$	$13.24\pm0.81$	$24.49 \pm 1.62$	$38.43 \pm 21.12$	$\textbf{27.70} \pm \textbf{2.48}$
I-9	$11.38\pm0.74$	$14.74\pm0.83$	$11.26\pm0.86$	$14.48\pm0.52$	$10.62 \pm 1.42$	$12.73\pm0.60$
I-10	$9.21\pm0.37$	$12.73\pm0.71$	$11.14 \pm 1.19$	$13.49\pm0.24$	$14.54\pm0.84$	$11.36\pm0.74$
I-11	$19.74 \pm 1.50$	$19.10\pm0.51$	$27.87 \pm 4.33$	$93.91 \pm 38.57$	$40.37\pm22.85$	$17.35\pm1.37$
I-12	$4.89\pm0.25$	$11.49\pm0.87$	$6.74\pm0.85$	$11.49\pm0.83$	$19.10\pm3.49$	$\textbf{8.89} \pm \textbf{0.91}$
I-13	> 200	$96.36 \pm 15.64$	$95.48 \pm 35.55$	> 200	$178.74 \pm 42.87$	> 200
I-14	> 200	$72.41 \pm 32.22$	$63.73\pm32.63$	$52.25\pm32.39$	$56.28 \pm 3.03$	$50.88 \pm 10.80$
I-15	> 200	$80.16 \pm 19.53$	$106.60 \pm 43.64$	$149.60 \pm 23.57$	> 200	$98.57 \pm 17.35$
I-16	> 200	> 200	> 200	> 200	> 200	$14.38\pm5.10$
I-17	$\textbf{47.86} \pm \textbf{4.82}$	$76.57 \pm 33.30$	$\textbf{37.33} \pm \textbf{5.89}$	$\textbf{96.76} \pm \textbf{69.63}$	$93.27\pm76.32$	$\textbf{45.98} \pm \textbf{8.85}$
I-18	> 200	$36.35\pm0.56$	$46\pm17.70$	> 200	$47.60 \pm 10.52$	> 200
I-19	> 200	$52.87 \pm 4.00$	$18.30\pm2.98$	$60.51 \pm 12.65$	$93.02\pm40.21$	> 200
II-2	> 200	> 200	$167.25\pm7.99$	> 200	> 200	> 200
II-3	> 200	> 200	$129.52 \pm 21.42$	Nd	> 200	> 200
II-4	$56.96 \pm 13.81$	$82.65 \pm 63.09$	$46.20\pm4.84$	$62.90 \pm 14.99$	$80.28 \pm 12.58$	$49.10\pm9.67$
II-5	$112.17 \pm 48.97$	$61.08\pm30.63$	> 200	Nd	> 200	> 200
II-6	$142.74 \pm 42.84$	$86.96 \pm 48.60$	$33.66 \pm 10.96$	$55.20 \pm 26.95$	$109.73 \pm 12.27$	$50.91 \pm 15.54$
II-7	> 200	> 200	$178.25 \pm 79.02$	> 200	> 200	> 200
II-8	Nd	$\textbf{47.02} \pm \textbf{25.66}$	$197.90 \pm 63.67$	> 200	> 200	> 200
II-9	$83.11 \pm 16.57$	> 200	Nd	> 200	$57.43 \pm 5.63$	$24.38 \pm 6.03$
II-10	$80.35\pm 6.81$	Nd	> 200	> 200	$23.68 \pm 1.18$	$118.18\pm18.72$
II-11	> 200	$57.20 \pm 22.67$	$169.29 \pm 60.72$	> 200	> 200	$93.35 \pm 28.46$
5-FU	$\textbf{45.17} \pm \textbf{12.07}$	> 200	$124.05 \pm 28.55$	> 200	> 200	$\textbf{52.37} \pm \textbf{17.75}$

Note: Nd indicates no inhibitory activity.



Fig. 2. Structure of compound I-12.

mTOR, KIT, ESR1, PDK2.

The compound targets were entered into DAVID for enrichment. It was found that **I-12** may affect biological processes related to nucleoplasm, nucleus, zinc ion binding, transcription initiation from RNA polymerase II promoter, etc. The summarized targets were enriched and analyzed in Metascape. It indicated that the targets were mainly enriched in nuclear receptors, signaling by nuclear receptors, cellular response to nitrogen compounds, and steroid metabolic process. Combining the enrichment results from DAVID and Metascape, we found that compound **I-12** may have an effect on nuclear receptor signaling and its related biological processes, in which the target RXRA plays an important role. At last, the pathway enrichment was carried out on KEGG pathway mapping with selected targets, mainly looking for the relevant pathways of mTOR, KIT, PDK2, ESR1 and RXRA, which were

selected as the intersection targets, and selecting the pathways with relatively large number of target enrichment and interest. Targets were entered into String to see the interactions between the targets. Finally, Cytoscape software was used to constructed disease-pathway-gene network and compound-pathway-target network.

Jiameng Qu *et al.* investigated the mechanism of action of dandelion against triple-negative breast cancer using a network pharmacology approach and found that dandelion may act by inhibiting the cell cycle and affecting cellular metabolism (Qu *et al.*, 2022). Yonghuan Yan *et al.* revealed the potential bioactivities, active components and therapeutic targets of *T. aurantialba* using a network pharmacology approach, and then investigated the potential binding sites of the active compounds to the key targets by molecular docking (Yan *et al.*, 2022). Zhenyuan Yu *et al.* investigated the potential mechanism of action of tangeretin in oral squamous cell carcinoma by network pharmacology, and pathway enrichment analysis showed that PI3K-AKT was the most important pathway (Yu *et al.*, 2022).

The binding mode of compound I-12 to mTOR, KIT, ESR1, PDK2, RXRA was investigated by molecular docking. The mTOR (PDB: 4DRI), KIT (PDB: 6GQM), ESR1 (PDB: 1SJ0), PDK2 (PDB: 4MPC) and RXRA (PDB: 3OAP) were downloaded from the RCSB Protein Data Bank. Docking was carried out by AutoDock software and the images were drawn by PyMOL. The molecular docking results showed that the free energy of binding of compound I-12 to mTOR was -10.91 kcal/moL, to KIT was -7.96 kcal/mol, to ESR1 was -9.56 kcal/mol, to PDK2 was -8.28 kcal/mol and to RXRA was -8.07 kcal/mol. Compound I-12 had strong binding ability to mTOR, KIT, ESR1, PDK2, RXRA. The simulated docking results (Fig. 5) showed that compound I-12 was bound to protein residues through hydrogen bonding and hydrophobic interactions, respectively. Compound I-12 interacted with mTOR residues SER-2035 and ILE-87, KIT residues SER-639 and GLU-635, ESR1 residues SER-527 and ASP-351, PDK2 residue ASP-203, and RXRA residue GLU-394 through hydrogen bonding. This might explain the higher antitumor activity of compound I-12.



Fig. 3. Expression of apoptosis-related proteins at different exposure time (1, 2, 4, 6, 8 and 10 min) after treatment of MDA-MB-231 cells with I-12 for 48 h.



**Fig. 4.** Expression of apoptosis-related proteins (Bad, Bax, Bcl-2, Bcl-x, Pro-Caspase-3, Cleaved Caspase-3, Catalase, clAP-1, clAP-2, Claspin, Clusterin, Cytochrome c, TRAIL R1/DR4, TRAIL R2/DR5, FADD, Fas/TNFRSF6/CD95, HIF-1α, HO-1/HMOX1/HSP32, HO-2/HMOX2, HSP27, HSP60, HSP70, HTRA2/Omi, Livin, PON2, p21/CIP1/CDKN1A, p27/Kip1, Phospho-p53(S15), Phospho-p53(S46), Phospho-p53(S392), Phospho-Rad17(S635), SMAC/Diablo, Survivin, TNF RI/TNFRSF1A, XIAP) after 48 h of I-12 treatment.



Fig. 5. Binding poses of compound I-12 with mTOR (A), KIT (C), ESR1 (E), PDK2 (G) and RXRA (I). Docking sites of compound I-12 with mTOR (B), KIT (D), ESR1 (F), PDK2 (H), and RXRA (J), with blue solid lines representing hydrogen bonding and gray dashed lines representing hydrophobic interactions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



Fig. 6. Possible mechanism of compound I-12 against breast cancer.

In total, we found through network pharmacological analysis that compound **I-12** may act on nuclear receptors and affect biological processes related to nuclear receptor signals. The potential targets were mTOR, KIT, ESR1, PDK2, RXRA. The antitumor mechanism (Fig. 6) of **I-12** may be related to the pathways involved by the potential targets, such as PI3K/Akt/mTOR pathway, MAPK pathway, estrogen receptor signal pathway, etc.

## 4. Conclusion

A total of 28 derivatives were designed and synthesized, including 18 estradiol-amino acid (17 $\beta$ -OH) derivatives (I-2 ~ I-19) and 10 estradiol-amino acid (16-C) derivatives (II-2 ~ II-11). Amino acids used for synthesis include *L*-phenylalanine, *L*-serine, *L*-leucine, glycine, *L*-proline and *L*-glutamine. The results of antitumor activity *in vitro* showed that compounds I-8, I-9, I-10, I-11 and I-12 of 17 $\beta$ -OH conjugated amino acids of estradiol had good antitumor activity, indicating that the introduction of amino acids could improve the anti-proliferative effect to a certain extent. Among them, I-12 showed the best inhibitory activity against MDA-MB-231 breast cancer cells (IC<sub>50</sub> = 4.89  $\mu$ M). In contrast, the estradiol-amino acid conjugates derived at position 16 (16-C) of 17 $\beta$ -estradiol (II-2 ~ II-11 *vs.* I-2 ~ I-19) showed no inhibitory activity against tested tumor cell lines, which may be related to the fact that the hydroxyl group at position 17 was oxidized to carbonyl group.

Further research of **I-12** on cell apoptosis found that it could induce the apoptosis of MDA-MB-231 cells through the mitochondrial mediated apoptosis pathway, and induced the upregulation of apoptosis-related proteins Bad, Bax, Cytochrome c, TRAIL R2/DR5, Fas/TNFRSF6/ CD95, HIF-1 $\alpha$ , PON2, p21/CIP1/CDKN1A and Phospho-p53 (S46), as well as induced the downregulation of cIAP, HO-2/HMOX2 and HSP27. Furthermore, through network pharmacological analysis, it was found that **I-12** may act on mTOR, KIT, ESR1, PDK2, RXRA targets and their related pathways, such as PI3K/Akt/mTOR pathway, MAPK pathway, estrogen receptor signal pathway, etc. Compound **I-12** was found to bind well to mTOR, KIT, ESR1, PDK2 and RXRA by molecular docking. Of course, this prediction needs to be verified by further experiments. For future work, we will verify the conclusions by more experiments such as Western Blot.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

This work was supported by the talents program of Guangxi science and technology department (AD20297033), the open project of state key laboratory of natural medicines (SKLNMKF202312), the national natural science foundation of China (21562006), the BAGUI scholar program of Guangxi province of China (2016A13), start-up fund of state key laboratory for chemistry and molecular engineering of medicinal resources of Guangxi Normal University (CMEMR2020-A01).

#### Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.arabjc.2023.105539.

## References

- Allemani, C., Matsuda, T., Di Carlo, V., Harewood, R., Matz, M., Niksic, M., Bonaventure, A., Valkov, M., Johnson, C.J., Esteve, J., Ogunbiyi, O.J., Silva, G.A.E., Chen, W.Q., Eser, S., Engholm, G., Stiller, C.A., Monnereau, A., Woods, R.R., Visser, O., Lim, G.H., Aitken, J., Weir, H.K., Coleman, M.P., Grp, C.W., 2018. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 391, 1023–1075. https://doi. org/10.1016/S0140-6736(17)33326-3.
- Bray, F., Jemal, A., Grey, N., Ferlay, J., Forman, D., 2012. Global cancer transitions according to the Human Development Index (2008–2030): a population-based study. Lancet Oncol. 13, 790–801. https://doi.org/10.1016/S1470-2045(12)70211-5.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide

#### Y.-q. Zhou et al.

for 36 cancers in 185 countries. Ca-Cancer J. Clin. 68, 394–424. https://doi.org/ 10.3322/caac.21492.

Broer, S., Broer, A., 2017. Amino acid homeostasis and signalling in mammalian cells and organisms. Biochem. J. 474, 1935–1963. https://doi.org/10.1042/BCJ20160822.

- Chen, W.Q., Zheng, R.S., Baade, P.D., Zhang, S.W., Zeng, H.M., Bray, F., Jemal, A., Yu, X. Q., He, J., 2016. Cancer Statistics in China, 2015. Ca-Cancer J. Clin. 66, 115–132. https://doi.org/10.3322/caac.21338.
- Harry, A.G., Murphy, J., Butler, W.E., Tiedt, R., Mooney, A., Manton, J.C., Pryce, M.T., O'Donovan, N., Walsh, N., Crown, J., Rai, D.K., Kenny, P.T.M., 2013. The synthesis, structural characterization and in vitro anti-cancer activity of novel N-{6-(ferrocenyl) ethynyl-2-naphthoyl} amino acid and dipeptide ethyl esters. J. Organomet. Chem. 734, 86–92. https://doi.org/10.1016/j. jorganchem.2012.11.041.
- Ishiki, N., Onishi, H., Machida, Y., 1997. Biological properties of conjugates of mitomycin C with estradiol benzoate and estradiol: their stability characteristics in biological media and their binding abilities to estrogen receptor. Biol. Pharm. Bull. 20, 1096–1102. https://doi.org/10.1248/bpb.20.1096.
- Ishiki, N., Onishi, H., Machida, Y., 2004. Evaluation of antitumor and toxic side effects of mitomycin C-estradiol conjugates. Int. J. Pharm. X 279, 81–93. https://doi.org/ 10.1016/j.ijpharm.2004.04.017.
- Lohndorf, A., Hosang, L., Dohle, W., Odoardi, F., Waschkowski, S.A., Rosche, A., Bauche, A., Winzer, R., Tolosa, E., Windhorst, S., Marry, S., Flugel, A., Potter, B.V.L., Diercks, B.P., Guse, A.H., 2021. 2-Methoxyestradiol and its derivatives inhibit storeoperated Ca<sup>2+</sup> entry in T cells: Identification of a new and potent inhibitor. Biochim. Biophys. Acta, Mol. Cell Res. 1868, 118988 https://doi.org/10.1016/j. bhamer 2021 118988
- Pansare, D.N., Shelke, R.N., Shinde, D.B., 2017. A Facial Synthesis and Anticancer Activity of (Z)-2-((5-(4-nitrobenzylidene)-4-oxo-4,5-dihydrothiazol-2-yl)amino)substituted Acid. J. Heterocycl. Chem. 54, 3077–3086. https://doi.org/10.1002/ jhet.2919.
- Prokai, L., Oon, S.M., Prokai-Tatrai, K., Abboud, K.A., Simpkins, J.W., 2001. Synthesis and biological evaluation of 17 beta-alkoxyestra-1,3,5(10)-trienes as potential neuroprotectants against oxidative stress. J. Med. Chem. 44, 110–114. https://doi. org/10.1021/jm000280t.
- Qu, J.M., Ke, F., Liu, Z.R., Yang, X., Li, X.Z., Xu, H.R., Li, Q., Bi, K.S., 2022. Uncovering the mechanisms of dandelion against triple-negative breast cancer using a combined network pharmacology, molecular pharmacology and metabolomics approach. Phytomedicine 99, 153986. https://doi.org/10.1016/j.phymed.2022.153986.
- Shull, J.D., Dennison, K.L., Chack, A.C., Trentham-Dietz, A., 2018. Rat models of 17 betaestradiol-induced mammary cancer reveal novel insights into breast cancer etiology and prevention. Physiol. Genomics 50, 215–234. https://doi.org/10.1152/ physiolgenomics.00105.2017.
- Siegel, R.L., Miller, K.D., Jemal, A., 2018. Cancer Statistics, 2018. Ca-Cancer J. Clin. 68, 7–30. https://doi.org/10.3322/caac.21442.
- Sorvik, I.B., Solum, E.J., Labba, N.A., Hansen, T.V., Paulsen, R.E., 2018. Differential effects of some novel synthetic oestrogen analogs on oxidative PC12 cell death caused by serum deprivation. Free Radic. Res. 52, 273–287. https://doi.org/ 10.1080/10715762.2018.1430363.

- Stander, X.X., Stander, B.A., Joubert, A.M., 2011. In vitro effects of an in silico-modelled 17 beta-estradiol derivative in combination with dichloroacetic acid on MCF-7 and MCF-12A cells. Cell Prolif. 44, 567–581. https://doi.org/10.1111/j.1365-2184.2011.00789.x.
- Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J., Jemal, A., 2015. Global Cancer Statistics, 2012. Ca-Cancer J. Clin. 65, 87–108. https://doi.org/10.3322/ caac.21262.
- Torre, L.A., Siegel, R.L., Ward, E.M., Jemal, A., 2016. Global cancer incidence and mortality rates and trends-An update. Cancer Epidemiol. Biomark. Prev. 25, 16–27. https://doi.org/10.1158/1055-9965.EPI-15-0578.
- van Meerloo, J., Kaspers, G.J.L., Cloos, J., 2011. Cell sensitivity assays: the MTT assay. Methods Mol. Biol. (Clifton, N.J.) 731, 237–245. https://doi.org/10.1007/978-1-61779-080-5\_20.
- Veronesi, U., Boyle, P., Goldhirsch, A., Orecchia, R., Viale, G., 2005. Breast cancer. Lancet 365, 1727–1741. https://doi.org/10.1016/S0140-6736(05)66546-4.
- Vig, B.S., Huttunen, K.M., Laine, K., Rautio, J., 2013. Amino acids as promoieties in prodrug design and development. Adv. Drug Delivery Rev. 65, 1370–1385. https:// doi.org/10.1016/j.addr.2012.10.001.
- Wright, P.K., Jones, S.B., Ardern, N., Ward, R., Clarke, R.B., Sotgia, F., Lisanti, M.P., Landberg, G., Lamb, R., 2014. 17 beta-estradiol regulates giant vesicle formation via estrogen receptor-alpha in human breast cancer cells. Oncotarget 5, 3055–3065. https://doi.org/10.18632/oncotarget.1824.
- Wu, Z.Q., Shah, A., Patel, N., Yuan, X.D., 2010. Development of methotrexate proline prodrug to overcome resistance by MDA-MB-231 cells. Bioorg. Med. Chem. Lett. 20, 5108–5112. https://doi.org/10.1016/j.bmcl.2010.07.024.
- Xie, L., Semenciw, R., Mery, L., 2015. Cancer incidence in Canada: trends and projections (1983–2032). Health Promot Chron. 35 (Suppl 1), 2–186. https://doi.org/10.24095/ hpcdp.35.S1.02.
- Yan, Y.H., Wang, M.T., Gan, X.R., Wang, X., Fu, C.H., Li, Y.M., Chen, N., Lv, P., Zhang, Y., 2022. Evaluation of pharmacological activities and active components in *Tremella aurantialba* by instrumental and virtual analyses. Front. Nutr. 9 https://doi.org/ 10.3389/fnut.2022.1083581.
- Yin, Y.D., Sheng, L.X., Zhang, J.Z., Zhang, L.Q., Liu, J.J., Wen, X.A., Liu, Y.H., Si, Y., Cheng, K.G., 2022. Synthesis and *in vitro/in vivo* anticancer evaluation of pentacyclic triterpenoid derivatives linked with *L*-phenylalanine or *L*-proline. Bioorg. Chem. 126, 105865 https://doi.org/10.1016/j.bioorg.2022.105865.
- Yin, Y.D., Sun, L., Sheng, L.X., Zhang, L.Q., Liu, J.J., Wen, X.A., Mo, W.B., Wang, Q.D., Cheng, K.G., 2023. (4-Picolylamino)-17/β-Estradiol derivative and analogues induce apoptosis with death receptor trail R2/DR5 in MCF-7. Chem.-Biol. Interact. 369, 110286 https://doi.org/10.1016/j.cbi.2022.110286.
- Yu, Z.Y., Wu, Y.L., Ma, Y.J., Cheng, Y.H., Song, G.H., Zhang, F., 2022. Systematic analysis of the mechanism of aged citrus peel (Chenpi) in oral squamous cell carcinoma treatment via network pharmacology, molecular docking and experimental validation. J. Funct. Foods 91, 105012. https://doi.org/10.1016/j. jff.2022.105012.
- Zhao, J., Lv, C., Wu, Q.L., Zeng, H.W., Guo, X., Yang, J., Tian, S.S., Zhang, W.D., 2019. Computational systems pharmacology reveals an antiplatelet and neuroprotective mechanism of Deng-Zhan-Xi-Xin injection in the treatment of ischemic stroke. Pharmacol. Res. 147, 104365 https://doi.org/10.1016/j.phrs.2019.104365.