RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

RSC Advances

Synthesis and biological evaluation of novel 2,3-pyrazole ring-substituted-4,4-dimethyl lithocholic acid derivatives as selective protein tyrosine phosphatase 1B (PTP1B) inhibitors with cellular efficacy

Shi-Wei Mao [a,ℜ], Lin Shuai [b,ℜ], Hai-Bing He [c,ℜ], Na Pan [a], Li-Xin Gao [b], Li-Fang Yu [a], Jia Li [b], Jing-Ya Li [b,*] Fan Yang [a,*]

[a]*MTDD, School of Chemistry and Molecular Engineering, East China Normal University, Shanghai 200062,China*
[b]*National Center for Drug Screening, Shanghai Institute of Materia Medica, Chinese Academy of Science, Shanghai 201203, China*
[c]*Chemistry and Chemical Engineering, Nantong University, Jiangsu 226019, China*

Abstract

In our continued efforts to develop lithocholic acid (LCA) analogues as selective PTP1B inhibitors, 14 novel 2,3-pyrazole ring-substituted-4,4-dimethyl derivatives were synthesized and evaluated against PTP1B, as well as the homologous protein tyrosine phosphatases (PTPs). All compounds were shown to be more potent and selective PTP1B inhibitors than LCA (IC_{50} = 12.74 μM) with IC_{50} values ranging between 0.42 to 4.49 μM. Moreover, treatment of CHO/hIR cells with 4,4-dimethyl-2'-(*p*-fluoro phenyl)-2'*H*-chola-2,5-dieno[3,2-*c*]pyrazol-24-oic acid (**30**) or 4,4-dimethyl-2'-(*o*-chloro phenyl)-2'*H*-chola-2,5-dieno[3,2-*c*]pyrazol-24-oic acid (**34**) increased the phosphorylation levels of IR and Akt in a dose dependent manner. The promising findings in this study suggest that further investigation of these compounds for the treatment of metabolic disorders is warranted.

Keywords:

2,3-Pyrazole ring-substituted lithocholic acid, PTP1B inhibitors, Selectivity, Cellular efficacy, Diabetes and obesity

* Corresponding author. Tel.: +86 21 62232764; fax: +86 21 62232100 (F. Yang); +86 21 50801313ext132; fax: +86 21 50801552 (J.-Y. Li);

E-mail address: fyang@chem.ecnu.edu.cn (F. Yang); jyli@mail.shcnc.ac.cn (J.-Y. Li)

ℜ These authors contributed equally to this work.

Introduction

Type 2 diabetes and obesity are metabolic disorders that are characterized by insulin and leptin resistance[1]. Protein tyrosine phosphatase 1B (PTP1B) is an enzyme that belongs to the protein tyrosine phosphatase (PTPs) family, which is involved in the regulation of several important physiological pathways. PTP1B dephosphorylates the insulin receptor (IR), IR substrate 1 (IRS1), and Janus kinase 2, resulting in a negative regulation of insulin and leptin signal transduction[2,3]. Moreover, PTP1B activity and expression are increased in insulin resistant humans and rodents[4,5], whereas PTP1B deficient mice were found to resistant to weight gain and remain the insulin sensitivity as compared to the wild type mice[6,7]. Therefore, PTP1B has been proposed as novel target for the treatment of type 2 diabetes and obesity. It is worth noting that PTPs share a high degree of structural conservation in the active site[8]. T-cell protein tyrosine phosphatase (TCPTP) has a sequence identity of about 74% in the catalytic domains with PTP1B[9]. Although it is unknown whether the combined inhibition of PTP1B and TCPTP will lead to severe side effects, pancreatic-TCPTP knockout mice exhibited impaired glucose tolerance and attenuated glucose-stimulated insulin secretion when challenged with prolonged high fat feeding[10]. Another drawback limited the application of most inhibitors for further development is poor intracellular PTP1B inhibitory potency. Because of the highly cationic character of the active site of PTP1B, most potent phosphotyrosine (pTyr) mimetics are highly negative charged providing these compounds with poor membrane permeability[11].

Natural products have been widely regarded as privileged structures in drug discovery efforts, and scaffolds derived from natural sources have led to the nearly half of the approved drugs in the past 2 decades. Many natural products and their derivatives were found possessing PTP1B inhibitory activities with moderate to good selectivities[12,13]. Trodusquemine (MSI-1436, **1**, **Figure 1**) is a noncompetitive and selective PTP1B inhibitors with moderate potency (IC_{50} = 1 μM) which induces weight loss without significant associated toxicities in animal models[14,15]. Recently, several pentacyclic acid triterpenoids (ursolic, oleanolic, moronic and morolic acids, **2**-**5**, **Figure 1**) were reported to present significant antidiabetic activities in non-insulin dependent diabetic rat models[16]. In our previous study[17], natural steroid compound lithocholic acid (LCA, **6**, **Figure 1**) was identified as a PTP1B inhibitor with an IC_{50} value of 12.74 μM. Embedding of fused heterocycles (substituted pyrazole, oxazole, substituted pyrimidine, aminothiazole and pyrazine) on ring A resulted in an improvement in potency and selectivity over TCPTP as exemplified by compound **7** (IC_{50} = 1.62 μM; TCPTP/PTP1B = 14.1-fold, **Figure 1**). Herein, we reported our continued efforts in the structure-activity relationship (SAR) studies and biological evaluations of 2,3-pyrazole ring-substituted-4,4-dimethyl lithocholic acid derivatives.

Results and discussion

Chemistry

As shown in **Scheme 1**, compound **10** with two methyls at C-4 position was obtained from lithocholic acid in 5 steps [17]. A typical formylation at C-2 of **10** was carried out by using sodium hydride and ethyl formate to yield **11**, which was then reacted with various substituted hydrazine hydrochlorides in ethanol to afford substituted pyrazoles over two steps in 36-62% overall yields. The substitution position of phenyl at the pyrazole ring was determined via X-ray crystallography. Single crystal of compound **33** was obtained by slow evaporation method using mixed solvent of chloroform and methanol (4/1 v/v). The crystal structure[18] (**Figure 2**) shows that the phenyl group was at the N-1 position, and this is well agreement with previous report [19]. When it comes to the methyl substituted pyrazoles, both of the isomers (**12a** and **12b**, structures determined by NOE, see supplemental information) were obtained. **12a**, **12b** and **13**-**24** were hydrolyzed in the presence of lithium hydroxide to g[ive] the desired carboxylic acids **25a**, **25b** and **26**-**37**. All the fina[l] compounds were characterized by ^{13}CNMR, ^{1}HNMR and HRMS.

Structure-activity relationship

Compounds **25a**, **25b** and **26**-**37** were evaluated in the enzym[e] inhibition assay (p-NPP assay) against PTP1B. Oleanolic acid (OA) was used as reference compound (**Table 1**). The inhibitory activities on the homologous enzyme TCPTP were investigated. We also tested the inhibitory activities on other homologous PTPs (LAR, SHP1 and SHP2).

It was intriguing that all of the 14 compounds were found to b[e] more potent PTP1B inhibitors than that of **LCA** (**6**, IC_{50} = 12.74 μM) suggesting that substitutions on 2,3-pyrazole-substituted-4,4-dimethyl lithocholic acid are well tolerated. 10 Derivatives (**27**, **28**, **30**-**37**) with substituted phenyl groups presented improved PTP[1B] inhibitory activities than compound **7** (IC_{50} = 1.62 μM). **30** (IC[50] = 0.42 μM) emerged as the most potent PTP1B inhibitor with almost 4-fold improved inhibitory potency compared to compound **7**. Replacement of phenyl by methyl group (**25a**) reduced the inhibitor[y] activity with IC_{50} value of 2.82 μM, which suggested that the phenyl group is favorable for the activity. Among the *para*-positio[n] substituted compounds (**26**, **27**, **28**, **29** and **30**), the para-methyl substituted compound **27** (IC_{50} = 0.81 μM) and the *para*-chlo[ro] substituted compound **28** (IC_{50} = 1.12 μM) showed improve[d] PTP1B inhibitory activities compared to compound **7** (IC_{50} = 1.6[2] μM). The *para*-fluoro substituted compound **30** was most activ[e] towards PTP1B with an IC_{50} value of 0.42 μM. However, the *para*-methoxy substituent **26** and the *para*-bromo substituent **29** reduce[d] the activity to 2.66 μM and 2.35 μM, respectively. Compared to th[e] *para*-methyl substituted compound **27**, *ortho*-methyl compound **(32)** and *meta*-methyl compound **(31)** exhibited similar PTP1B inhibitory activities, with IC_{50} values of 1.17 μM and 0.93 μM, respectively. The 3,4-dimethyl-substituted compound **36**, in the combination [of] *para*-methyl substitution and *meta*-methyl substitution, showed an increased inhibitory activity (IC_{50}=0.73 μM). All the chloro-substituted compounds (**28**, **33**, **34** and **37**) exhibited simila[r] inhibitory activities at around 1 μM.

Selectivity over other homologous PTPs

The methyl substituted compounds (**25a** and **25b**) did not show noticeable inhibitory activities against TCPTP. When it comes to th[e] phenyl substituted compounds, the para-fluoro substituted compound **30**, which emerged as the most potent PTP1B inhibitor, remained about 11-fold selectivity over TCPTP. All the methyl substituted compounds (**27**, **31**, **32**, **35** and **36**) had good selectivity over TCPTP, **27**, **31** and **32** did not exhibit noticeable inhibition towards TCPTP under 40 μM. The inhibition (%) of the chloro substituted compounds **28** and **33** against TCPTP at 40 μM wa[s] below 5%, while the dichloro-substituted compound (**37**) showed decreased selectivity over TCPTP. It seemed that the electron density of the phenyl ring had influence on the selectivity. In detail, fluoro and dichloro substituted compounds (**30** and **37**) with lower electron density on the phenyl showed more potent inhibition to TCPTP, while remaining similar inhibition to PTP1B wh[en] compared with methyl, bromo or monochloro substituted compounds with one exception (**34**). Encouragingly, all the tested six compounds (**27**, **28**, **30**, **31**, **32** and **33**) exhibited no obvious inhibition on LAR, SHP1 and SHP2 at 40 μM (**Table 2**).

Cellular efficacy

PTP1B is a negative regulator in the insulin signaling pathway by dephosphorylating specific phosphotyrosine residues of IR. Inhibition of PTP1B activity could enhance the phosphorylation of IR and downstream protein Akt, and thus activate the downstream signaling pathway. Four compounds (**30**, **31**, **34** and **36**) with good potency and/or selectivity were selected to evaluate the cellular efficacy in CHO/hIR cells. As shown in **Figure 3**, two compounds (**30** and **34)** were found to remarkably increase the phosphorylation levels of IR and its downstream protein Akt in a concentration-dependent manner (3, 10 μM) , suggesting that **30** and **34** may effectively stimulate the insulin signaling by inhibiting the activity of PTP1B.

Enzyme kinetic study and molecular docking analysis

In our previous study, the binding modality of compound **7** with PTP1B was investigated [17], since the substituent on the phenyl group had a crucial influence on potency and selectivity, it is necessary to find out whether the binding modality was changed after substituent introduction.

Firstly, the most active compound **30** was selected for PTP1B enzyme kinetic study. As shown in **Figure 4**, the Vmax value retained constant while Km value increased with the mounting compound concentration, indicating that **30** was a competitive PTP1B inhibitor (Ki = 0.26 μM). Afterwards, a molecular docking analysis of **30** was carried out using LibDock available with Discovery Studio 2.1.[20] The preferred coordination mode of **30** with PTP1B was presented in **Figure 5**, **Figure 5**(b) shows the binding interactions of **30** with PTP1B. The carboxylic acid group of **30** is bound into the active site. Similar to that of **7**, the carbonyl of the –COOH may interact with Arg221 via a salt bridge and the hydroxyl of the –COOH shows H-bond interaction with Glu115. However, the para-fluoro substituted phenyl group of **30** binds in the second phosphotyrosine (pTyr) binding site of PTP1B by H-bond interaction with guanidine group of Arg254, which was different from the ion-π interaction between **7** and PTP1B. The key H-bond interaction with Arg254 might be crucial for tightening the complex of PTP1B and compound **30**. It indicated that the substituted phenyl group of these compound series had a diverse binding modality with PTP1B enzyme.

Conclusion

14 novel pyrazole-fused 4,4-dimethyl LCA derivatives were synthesized, the potency against PTP1B and selectivity over homologous PTPs were explored. Intriguingly, all of the 14 compounds exhibited improved activities (IC_{50} = 0.42-4.49 μM) toward PTP1B compared to that of **LCA** (**6**, IC_{50} = 12.74 μM). In particular, 4-fluoro phenyl substituted compound **30** (IC_{50}=0.42 μM) was the most potent PTP1B inhibitor, which also possessed good selectivity towards TCPTP, SHP1, SHP2 and LAR. Moreover, both **30** and **34** (3, 10, 30 μM) significantly increased the insulin-induced phosphorylation of IR and Akt in cell-based assays. In summary, we identified a series of 2,3-pyrazole-4,4-dimethyl lithocholic acid derivatives as novel PTP1B inhibitors with good potency, selectivity as well as cell permeability. Further investigation of these compounds may lead to the development of novel antidiabetic agents derived from natural steroids.

Experimental section

General methods

Starting materials, reagents and chemicals were purchased from commercial suppliers and used without further purification unless otherwise stated. The progress of reactions was monitored by silica gel thin layer chromatography (TLC) plates, visualized under UV or charred using concentrated H_2SO_4 in the solution of EtOH followed by heating. Flash column chromatography was performed using Qingdao Haiyang silica gel (200-300) with distilled solvents. At 26.5 ℃, ^{1}H NMR (400MHz) spectra were recorded on Bruker Avance 400 spectrometers in $CDCl_3$ or DMSO-d6 [using TMS as internal standard] and ^{13}C NMR (100 MHz) spectra on Bruker Avance 400 spectrometers in $CDCl_3/CD_3OD$ = 2/1 v/v [the residual peak of $CDCl_3$ (^{1}H NMR δ 7.26, ^{13}C NMR δ 77.16); the residual peak of DMSO-d6 (^{1}H NMR δ 2.50); the residual peak of CD_3OD (^{13}C NMR δ 49.00)]. The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, m = multiplet, br = broad, coupling constant (in Hz) and integration. High resolution mass spectra were performed using a Bruker ESI-TOF high-resolution mass spectrometer. Melting points were uncorrected and were recorded on a Buchi B-54 melting point apparatus.

Synthesis of intermediate compound 10

A solution of lithocholic acid (**6**) (3.76 g, 10.0 mmol) in methanol (150 ml) at 0 ℃ was treated with $SOCl_2$ (2.92 ml, 40.0 mmol) over 30 min, then the mixture was stirred for 4h at room temperature. The solvent was concentrated in vacuo, then the residue was dissolved in H_2O (30 mL) and EtOAc (30 mL). The EtOAc phase was separated, and the aqueous phase was extracted with EtOAc (2×20 mL). The combined organic extract was washed with saturated $NaHCO_3$, brine, dried over anhydrous Na_2SO_4, concentrated to give white solid and used for the next step without further purification.

To a solution of the white solid obtained above in CH_2Cl_2 (60 ml) was added PCC (4.32 g, 20.0 mmol), after stirring for 12h at room temperature, the mixture was filtered and the filtrate was washed with saturated aqueous $NaHSO_3$, brine, dried over anhydrous Na_2SO_4 and concentrated. The residue was purified by silica gel chromatography (petroleum ether/ EtOAc 7/1 v/v) to give compound **8** (3.56 g, 93%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.67 (s, 3H), 2.70 (t, *J* = 14.2 Hz, 1H), 2.40 - 2.29 (m, 2H), 2.26 - 2.20 (m, 1H), 2.20 - 2.13 (m, 1H), 2.06 - 1.99 (m, 3H), 1.92 - 1.19 (m, 16H), 1.16 - 1.05 (m, 4H), 1.02 (s, 3H), 0.92 (d, *J* = 6.5 Hz, 3H), 0.68 (s, 3H); HRMS (ESI): calcd for $C_{25}H_{40}NaO_3$ [M+Na]$^+$ 411.2870, found [M+Na]$^+$ 411.2898.

Compound **8** (3.56 g, 9.2 mmol) was dissolved in CH_2Cl_2 (150 ml) and a solution of Br_2 (0.52 ml, 10.1 mmol) in HOAc (20 ml) was added at 0 ℃ over 30 min, the stirring was continued for 1h. The mixture was washed with water, saturated aqueous $NaHCO_3$, brine, dried over anhydrous Na_2SO_4, concentrated to give a yellow oil mixture which was used for the next step without further purification.

To a solution of the yellow oil mixture in DMF (60 ml) was added Li_2CO_3 (2.72 g, 36.8 mmol) and $LiBr{\cdot}H_2O$ (1.93 g, 18.4 mmol) under N_2, the reaction mixture was stirred for 6h at 90 ℃. Then the mixture was cooled to room temperature, and added H_2O (50 mL) and EtOAc (50 mL). The EtOAc phase was separated, and the aqueous phase was extracted with EtOAc (3×20 mL). The organic phase were combined and washed with brine, dried over anhydrous Na_2SO_4 and concentrated. The residue was purified by silica gel chromatography (petroleum ether/ EtOAc 10/1 v/v) to give compound **9** (1.83 g, 48%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.72 (s, 1H), 3.67 (s, 3H), 2.46 - 2.30 (m, 4H), 2.30 - 2.18 (m, 2H), 2.06 - 1.98 (m, 2H), 1.93 - 1.75 (m, 3H), 1.73 - 1.37 (m, 9H), 1.18 (s, 3H), 1.17 - 0.97 (m, 4H), 0.92 (d, *J* = 6.5 Hz, 3H), 0.71 (s, 3H); HRMS (ESI): calcd for $C_{25}H_{38}NaO_3$ [M+Na]$^+$ 409.2713, found [M+Na]$^+$ 409.2755.

A solution of compound **9** (1.83 g, 4.7 mmol) in dry *t*-BuOH (30 ml) was treated with *t*-BuOK (2.13 g, 19.0 mmol) under N_2, the reaction mixture was stirred at room temperature for 1h, and then CH_3I was slowly introduced. After another 24h at room temperature the reaction mixture was poured into ice water (50 ml), and the pH value was adjusted to 5 by 1M HCl aqueous solution. The mixture was extracted with EtOAc (3×20 mL). The organic phase was washed with brine, dried over anhydrous Na_2SO_4 and concentrated. The residue was purified by silica gel chromatography (petroleum

ether/ EtOAc 20/1 v/v) to give compound **10** (0.98 g, 50%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.56 (dd, *J* = 2.0, 4.8 Hz, 1H), 3.67 (s, 3H), 2.63 - 2.18 (m, 4H), 2.15 - 1.75 (m, 5H), 1.72 - 1.29 (m, 9H), 1.23 (s, 6H), 1.20 - 1.00 (m, 5H), 0.93 (d, *J* = 6.4 Hz, 3H), 0.85 (s, 3H), 0.69 (s, 3H); HRMS (ESI): calcd for $C_{27}H_{42}NaO_3$ $[M+Na]^+$ 437.3015, found $[M+Na]^+$ 437.3005.

General procedure for the synthesis of 2,3-pyrazole ring-substituted-4,4-dimethyl lithocholic acid ethyl esters 12a, **12b**, **and 13-24**

To a solution of compound **10** (300 mg, 0.7 mmol) in dry toluene (20 ml) was added ethyl formate (3.0 ml, 36.0 mmol) and NaH (579 mg, 60%, 14.1 mmol) under N_2, the reaction mixture was stirred for 6h at room temperature. The mixture was poured into ice water and then the pH value was adjusted to 7 by 1M HCl aqueous solution. The aqueous phase was extracted with EtOAc (3×20 mL). The combined organic phase was washed with brine, dried over anhydrous Na_2SO_4 and concentrated to give a light yellow solid (compound **11**) which was used for the next step without further purification.

Compound **11** (320 mg, 0.7 mmol) was dissolved in ethanol (15 ml) and substituted hydrazine hydrochloride (3.5 mmol) was added under N_2, the reaction mixture was heated to reflux for 2h. After cooling to room temperature, the solvent was removed in vacuo, then H_2O (10 ml) was added and the mixture was extracted with EtOAc (2×10 mL). The organic phase was separated and washed with brine, dried over anhydrous Na_2SO_4 and concentrated, The crude product was purified by a flash chromatography (petroleum ether/ EtOAc 10/1 v/v) to afford the corresponding pure 2,3-pyrazole ring-substituted-4,4-dimethly lithocholic acid ethyl esters **12a**, **12b** and **13-24**.

Characterization data

Compound **12a**: isolated in 32% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.11 (s, 1H), 5.71 (dd, *J* = 5.4, 2.2 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 3.89 (s, 3H), 2.60 (d, *J* = 14.6 Hz, 1H), 2.32 - 2.24 (m, 1H), 2.19 - 1.95 (m, 4H), 1.85 - 1.70 (m, 2H), 1.66 - 1.42 [m, 8H, containing 1.48 (s, 3H)], 1.33 (s, 3H), 1.31 - 1.21 (m, 3H), 1.19 (t, *J* = 7.1 Hz, 3H), 1.14 (d, *J* = 4.6 Hz, 1H), 1.10 - 0.94 (m, 4H), 0.88 (d, *J* = 6.5 Hz, 3H), 0.84 (s, 3H), 0.63 (s, 3H).

Compound **12b**: isolated in 29% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.91 (s, 1H), 5.67 (dd, *J* = 5.2, 2.2 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 3.77 (s, 3H), 2.61 (d, *J* = 14.4 Hz, 1H), 2.32 - 2.24 (m, 1H), 2.18 - 1.95 (m, 4H), 1.87 - 1.69 (m, 3H), 1.65 - 1.47 (m, 4H), 1.43 (s, 3H), 1.38 (m, 1H), 1.29 (s, 3H), 1.26 - 1.24 (m, 2H), 1.18 (t, *J* = 7.1 Hz, 3H), 1.15 - 0.93 (m, 5H), 0.87 (d, *J* = 6.4 Hz, 3H), 0.78 (s, 3H), 0.62 (s, 3H).

Compound **13**: isolated in 43% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.28 (s, 1H), 7.22 (m, 2H), 6.85 (d, *J* = 8.7 Hz, 2H), 5.61 (d, *J* = 3.3 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 3.78 (s, 3H), 2.68 (d, *J* = 14.6 Hz, 1H), 2.32 - 1.97 (m, 5H), 1.87 - 1.72 (m, 2H), 1.64 - 1.35 (m, 6H), 1.30 - 1.17 [m, 9H, containing 1.22 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H)], 1.11 - 0.94 [m, 7H, containing 1.12 (s, 3H)], 0.88 (d, *J* = 8.5 Hz, 3H), 0.87 (s, 3H), 0.63 (s, 3H).

Compound **14**: isolated in 43% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 (s, 1H), 7.20 - 7.13 (m, 4H), 5.60 (dd, *J* = 5.3, 2.1 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 2.68 (d, *J* = 14.6 Hz, 1H), 2.34 (s, 3H), 2.32 - 2.24 (m, 1H), 2.19 - 1.97 (m, 4H), 1.87 - 1.34 (m, 10H), 1.31 - 1.24 (m, 2H), 1.22 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H), 1.15 (m, 1H), 1.08 (s, 3H), 1.04 - 0.94 (m, 2H), 0.88 (d, *J* = 6.5 Hz, 3H), 0.87 (s, 3H), 0.63 (s, 3H).

Compound **15**: isolated in 47% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 - 7.33 (m, 2H), 7.30 (s, 1H), 7.28 - 7.24 (m, 2H), 5.62 (dd, *J* = 5.4, 2.3 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 2.68 (d, *J* = 14.6 Hz, 1H), 2.32 - 2.24 (m, 1H), 2.19 - 1.97 (m, 4H), 1.87 - 1.70 (m, 2H), 1.64 - 1.34 (m, 7H), 1.31 - 1.23 (m, 2H), 1.21 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H), 1.16 (m, 1H) , 1.09 (s, 3H), 1.07 - 0.94 (m, 3H), 0.88 (d, *J* = 6.5 Hz, 3H), 0.86 (s, 3H), 0.63 (s, 3H).

Compound **16**: isolated in 47% yield as a brown solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.49 (d, *J* = 8.4 Hz, 2H), 7.30 (s, 1H), 7.20 (d, *J* = 8.4 Hz, 2H), 5.61 (d, *J* = 3.2 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 2.67 (d, *J* = 14.6 Hz, 1H), 2.32 - 1.97 (m, 5H), 1.87 - 1.70 (m, 2H), 1.64 - 1.34 (m, 6H), 1.31 - 1.24 (m, 2H), 1.21 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H), 1.16 (m, 1H), 1.13 - 0.94 [m, 7H, containing 1.09 (s, 3H)], 0.89 (d, *J* = 8.4 Hz, 3H), 0.86 (s, 3H), 0.63 (s, 3H).

Compound **17**: isolated in 46% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.34 - 7.27 (m, 3H), 7.05 (t, *J* = 8.5 Hz, 2H), 5.62 (d, *J* = 3.2 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 2.68 (d, *J* = 14.6 Hz, 1H), 2.28 (m, 1H), 2.19 - 1.98 (m, 4H), 1.87 - 1.70 (m, 2H), 1.65 - 1.34 (m, 7H), 1.31 - 1.24 (m, 2H), 1.21 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H), 1.16 - 1.11 (m, 2H), 1.08 (s, 3H), 1.05 - 0.96 (m, 2H), 0.88 (d, *J* = 6.7 Hz, 3H), 0.87 (s, 3H), 0.63 (s, 3H).

Compound **18**: isolated in 43% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29 (s, 1H), 7.25 - 7.12 (m, 2H), 7.11 (m, 2H), 5.60 (dd, *J* = 5.2, 2.0 Hz, 1H), 4.04 (q, *J* = 7.1 Hz, 2H), 2.68 (d, *J* = 14.6 Hz, 1H), 2.31 (s, 3H), 2.29 - 2.23 (m, 1H), 2.18 - 1.96 (m, 4H), 1.86 - 1.34 (m, 9H), 1.31 - 1.25 (m, 2H), 1.22 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H), 1.17 (m, 1H), 1.08 (s, 3H), 1.06 - 0.95 (m, 3H), 0.88 (d, *J* = 8.4 Hz, 3H), 0.86 (s, 3H), 0.63 (s, 3H).

Compound **19**: isolated in 45% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33 (m, 1H), 7.31 - 7.25 (m, 2H), 7.21 - 7.1 (m, 2H), 5.60 (dd, *J* = 5.2, 2.1 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 2.69 (d, *J* = 14.6 Hz, 1H), 2.32 - 2.24 (m,1H), 2.19 - 1.97 (m, 4H), 1.89 (s, 3H), 1.85 - 1.70 (m, 2H), 1.64 - 1.35 (m, 6H), 1.31 - 1.16 [m, 9H, containing 1.30 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H)], 1.14 - 1.02 [m, 4H, containing 1.05 (s, 3H)], 1.00 - 0.86 [m, 9H, containing 0.88 (d, *J* = 6.5 Hz, 3H), 0.86 (s, 3H)], 0.63 (s, 3H).

Compound **20**: isolated in 46% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 - 7.34 (m, 2H), 7.31 - 7.27 (m, 2H), 7.2 - 7.20 (m, 1H), 5.61 (dd, *J* = 5.2, 1.9 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 2.68 (d, *J* = 14.6 Hz, 1H), 2.31 - 2.24 (m, 1H), 2.18 - 1.97 (m, 4H), 1.87 - 1.70 (m, 2H), 1.64 - 1.34 (m, 6H), 1.31 - 1.25 (m, 2H), 1.23 (, 3H), 1.18 (t, *J* = 7.1 Hz, 3H), 1.16 (m, 1H), 1.11 (s, 3H), 1.08 - 0.9 (m, 4H), 0.88 (d, *J* = 6.5Hz, 3H), 0.86 (s, 3H), 0.63 (s, 3H).

Compound **21**: isolated in 47% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.47 - 7.40 (m, 2H), 7.38 (s, 1H), 7.35 - 7.2 (m, 2H), 5.61 (dd, *J* = 5.4, 2.2 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 2.69 (d, *J* = 14.6 Hz, 1H), 2.32 - 1.98 (m, 5H), 1.87 - 1.70 (m, 2H), 1.64 - 1.35 (m, 6H), 1.31 - 1.27 (m, 1H), 1.24 (s, 3H), 1.20 - 1.11 [m, 6H, containing 1.18 (t, *J* = 7.1 Hz, 3H)], 1.10 - 1.02 (m, 3H), 1.00 (s, 3H), 0.89 (d, *J* = 6.2 Hz, 3H), 0.88 (s, 3H), 0.63 (s, 3H).

Compound **22**: isolated in 49% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.31 (d, *J* = 4.9 Hz, 1H), 7.20 - 7.11 (m, 1H), 7.02 - 6.94 (m, 2H), 5.60 (d, *J* = 3.5 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 2.68 (d, *J* = 14.5 Hz, 1H), 2.29 (s, 3H), 2.26 (m, 1H), 2.19 - 1.97 (m, 4H), 1.84 (s, 3H), 1.82 - 1.70 (m, 2H), 1.64 - 1.35 (m, 6H), 1.30 - 1.16 [m, 9H, containing 1.27 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H)], 1.14 - 0.99(m, 4H), 0.93 (s, 3H), 0.88 (d, *J* = 6.4 Hz, 3H), 0.85 (s, 3H), 0.63 (s, 3H).

Compound **23**: isolated in 50% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.27 (s, 1H), 7.10 - 7.03 (m, 3H), 5.60 (d, *J* = 3.2 Hz, 1H), 4.05 (q, *J* = 7.1 Hz, 2H), 2.67 (d, *J* = 14.5 Hz, 1H), 2.32 - 2.10 [m, 9H, containing 2.24 and 2.21, both (s, 3H)], 2.07 - 1.93 (m, 2H), 1.87 - 1.71 (m, 2H), 1.64 - 1.35 (m, 6H), 1.31 - 1.16 [m, 9H containing 1.23 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H)], 1.13 - 0.95 [m, 7H,

containing 1.09 (s, 3H)], 0.88 (d, *J* = 8.8 Hz, 3H), 0.87 (s, 3H), 0.63 (s, 3H).

Compound **24**: isolated in 47% yield as a white solid; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 (m, 2H), 7.30 (s, 1H), 7.20 (d, *J* = 8.3 Hz, 1H), 5.63 (d, *J* = 3.1 Hz, 1H), 4.05 (q, *J* = 7.0 Hz, 2H), 2.67 (d, *J* = 14.6 Hz, 1H), 2.32 - 1.97 (m, 5H), 1.87 - 1.72 (m, 2H), 1.64 - 1.35 (m, 6H), 1.28 - 1.16 [m, 9H, containing 1.22 (s, 3H), 1.18 (t, *J* = 7.1 Hz, 3H)], 1.12 (s, 3H), 1.09 - 0.96 (m, 4H), 0.88 (d, *J* = 6.4 Hz, 3H), 0.86 (s, 3H), 0.63 (s, 3H).

General procedure for the synthesis of 2,3-pyrazole ring-substituted-4,4-dimethyl lithocholic acid derivatives 25a, **25b and 26-37**

To a solution of the 2,3-pyrazole ring-substituted-4,4-dimethly lithocholic acid ethyl esters (0.5 mmol) in MeOH/H_2O (20 ml/2 ml v/v) was added $LiOH{\cdot}H_2O$ (420 mg, 10.0 mmol), the reaction mixture was stirred at room temperature for 24h. Aqueous HCl (1M) was introduced to adjust the pH = 5. After solvent removal, the residue was extracted by EtOAc (3×10 mL). The combined organic phase was dried over anhydrous Na_2SO_4 and concentrated. The crude product was flash-chromatographed on a silica gel column (CH_2Cl_2/ MeOH 30/1 v/v) to give the corresponding pure 2,3-pyrazole ring-substituted-4,4-dimethly lithocholic acid derivatives **25a**, **25b** and **26-37**.

Characterization data

Compound **25a**: isolated in 76% yield as a white solid; mp: 141.5 ℃- 142.4℃;^{1}H NMR (400 MHz, DMSO-d6) δ 11.93 (br, 1H), 7.09 (s, 1H), 5.82 (d, *J* = 2.8 Hz, 1H), 3.87 (s, 3H), 2.63 (d, *J* = 14.5 Hz, 1H), 2.24 (m, 1H), 2.15 - 1.97 (m, 4H), 1.83 (m, 1H), 1.73 - 1.63 (m, 2H), 1.62 - 1.54 (m, 2H), 1.51 (s, 3H), 1.49 - 1.38 (m, 2H), 1.36 (s, 3H), 1.30 - 1.16 (m, 4H), 1.13 - 1.00 (m, 4H), 0.90 (d, *J* = 6.2 Hz, 3H), 0.85 (s, 3H), 0.66 (s, 3H)；^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.49, 151.58, 144.65, 136.20, 121.44, 114.91, 57.19, 56.07, 49.93, 42.56, 40.03, 38.89, 38.65, 36.32, 35.69, 33.58, 32.52, 31.72, 31.61, 31.40, 31.34, 29.54, 28.39, 24.46, 21.39, 21.25, 18.46, 12.03; HRMS (ESI): calcd for $C_{28}H_{43}N_2O_2$ $[M+H]^+$ 439.3319, found $[M+H]^+$ 439.3313.

Compound **25b**: isolated in 80% yield as a white solid; mp: 142.6℃ - 143.7℃;^{1}H NMR (400 MHz, DMSO-d6) δ 7.27 (s, 1H), 5.73 (d, *J* = 3.1 Hz, 1H), 3.74 (s, 3H), 2.66 (d, *J* = 14.4 Hz, 1H), 2.24 (m, 1H), 2.15 - 1.97 (m, 4H), 1.83 (m, 1H), 1.70 - 1.44 (m, 6H), 1.43 - 1.35 [m, 4H, containing 1.39 (s, 3H)], 1.29 - 1.16 [m, 6H, containing 1.26 (s, 3H)], 1.14 - 1.00 (m, 4H), 0.90 (d, *J* = 6.5 Hz, 3H), 0.77 (s, 3H), 0.66 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.50, 156.30, 149.65, 129.67, 121.70, 114.83, 57.27, 56.18, 49.71, 42.64, 40.12, 38.88, 38.81, 36.17, 35.77, 33.61, 32.77, 32.61, 32.36, 31.96, 31.47, 31.43, 28.45, 24.51, 21.53, 21.44, 18.50, 12.10; HRMS (ESI): calcd for $C_{28}H_{43}N_2O_2$ $[M+H]^+$ 439.3319, found $[M+H]^+$ 439.3330.

Compound **26**: isolated in 74% yield as a white solid; mp: 233.1℃ - 234.2℃;^{1}H NMR (400 MHz, DMSO-d6) δ 7.31 (s, 1H), 7.29 (d, *J* = 8.7 Hz, 2H), 7.01 (d, *J* = 8.7 Hz, 2H), 5.71 (d, *J* = 3.1 Hz, 1H), 3.82 (s, 3H), 2.72 (d, *J* = 14.7 Hz, 1H), 2.23 (m, 1H), 2.14 - 1.98 (m, 4H), 1.81 (m, 1H), 1.71 - 1.36 (m, 8H), 1.31 - 1.17 [m, 6H, containing 1.24 (s, 3H)], 1.14 - 1.08 [m, 4H, containing 1.09 (s, 3H)], 1.06 - 1.00 (m, 2H), 0.91 (d, *J* = 6.7 Hz, 3H), 0.89 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.80, 160.32, 151.30, 146.69, 137.46, 134.74, 130.27 (2C), 121.25, 114.85, 113.98 (2C), 57.20, 56.10, 55.73, 49.90, 42.61, 40.06, 38.75, 37.23, 35.73, 33.59, 32.47, 32.39, 31.74, 31.56, 31.48, 31.40, 28.41, 24.47, 21.45, 21.32, 18.50, 12.09; HRMS (ESI): calcd for $C_{34}H_{47}N_2O_3$ $[M+H]^+$ 531.3581, found $[M+H]^+$ 531.3561.

Compound **27**: isolated in 64% yield as a white solid; mp: 235.8℃ - 236.9℃;^{1}H NMR (400 MHz, DMSO-d6) δ 11.94 (br, 1H), 7.32 (s, 1H), 7.29 (d, *J* = 8.3 Hz, 2H), 7.25 (d, *J* = 8.3 Hz, 2H), 5.70 (d, *J* = 3.1 Hz, 1H), 2.73 (d, *J* = 14.6 Hz, 1H), 2.39 (s, 3H), 2.24 (m, 1H), 2.16 - 2.00 (m, 4H), 1.80 (m, 1H), 1.72 - 1.35 (m, 8H), 1.28 - 1.18 [m, 5H, containing 1.24 (s, 3H)], 1.15 - 1.01 [m, 7H, containing 1.0 (s, 3H)], 0.91 (d, *J* = 6.5 Hz, 3H), 0.88 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.49, 151.33, 146.80, 139.85, 139.44, 137.50, 129.57 (2C), 128.94 (2C), 121.35, 115.01, 57.27, 56.17, 49.99, 42.67, 40.13, 38.81, 37.31, 35.77, 33.62, 32.52, 32.47, 31.81, 31.58, 31.46, 31.41, 28.47, 24.52, 21.49, 21.39, 21.35, 18.54, 12.14; HRMS (ESI): calcd for $C_{34}H_{47}N_2O_2$ $[M+H]^+$ 515.3632, found $[M+H]^+$ 515.3652.

Compound **28**: isolated in 63% yield as a white solid; mp: 228.5 ℃ - 229.7℃;^{1}H NMR (400 MHz, DMSO-d6) δ 11.93 (br, 1H), 7.57 (d, *J* = 8.6 Hz, 2H), 7.43 (d, *J* = 8.6 Hz, 2H), 7.37 (s, 1H), 5.72 (d, *J* = 3.1 Hz, 1H), 2.73 (d, *J* = 14.7 Hz, 1H), 2.24 (m, 1H), 2.16 - 1.97 (m, 4H), 1.82 (m, 1H), 1.73 - 1.34 (m, 8H), 1.29 - 1.18 [m, 5H, containing 1.24 (s, 3H)], 1.14 - 1.01 [m, 7H, containing 1.10 (s, 3H)], 0.91 (d, *J* = 6.5 Hz, 3H), 0.88 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.73, 151.10, 147.02, 140.71, 138.19, 135.74, 130.61 (2C), 129.30 (2C), 121.60, 115.45, 57.28, 56.19, 49.99, 42.68, 40.13, 38.81, 37.29, 35.79, 33.57, 32.52 (2C), 31.82, 31.71, 31.46, 31.44, 28.47, 24.53, 21.49, 21.40, 18.53, 12.13; HRMS (ESI): calcd for $C_{33}H_{44}ClN_2O_2$ $[M+H]^+$ 535.3086, found $[M+H]^+$ 535.3095.

Compound **29**: isolated in 71% yield as a brown solid; mp: 236 ℃ - 237.4℃;^{1}H NMR (400 MHz, DMSO-d6) δ 11.94 (br, 1H), 7.70 (d, *J* = 8.2 Hz, 2H), 7.37 (s, 1H), 7.36 (d, *J* = 8.2 Hz, 2H), 5.72 (d, *J* = 3.0 Hz, 1H), 2.73 (d, *J* = 14.4 Hz, 1H), 2.24 (m, 1H), 2.17 - 1.98 (m, 4H), 1.83 (m, 1H), 1.72 - 1.35 (m, 8H), 1.29 - 1.19 [m, 5H, containing 1.24 (s, 3H)], 1.16 - 0.99 [m, 7H, containing 1.10 (s, 3H)], 0.91 (d, *J* = 6.5 Hz, 3H), 0.88 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.66, 151.18, 147.27, 141.28, 138.27, 132.51 (2C), 131.09 (2C), 123.98, 121.81, 115.68, 57.45, 56.31, 50.16, 42.83, 40.29, 38.95, 37.44, 35.96, 33.69, 32.66, 32.62, 31.98, 31.81, 31.60, 28.62, 24.66, 21.58, 21.55, 18.62, 12.22; HRMS (ESI): calcd for $C_{33}H_{44}BrN_2O_2$ $[M+H]^+$ 579.2581, found $[M+H]^+$ 579.2553.

Compound **30**: isolated in 75% yield as a white solid; mp: 231.8℃ - 232.6℃;^{1}H NMR (400 MHz,DMSO-d6) δ 11.93 (br, 1H), 7.48 - 7.43 (m, 2H), 7.37 - 7.30 (m, 3H), 5.72 (dd, *J* = 5.0, 1.9 Hz, 1H), 2.73 (d, *J* = 14.7 Hz, 1H), 2.24 (m, 1H), 2.16 - 1.99 (m, 4H), 1.82 (m, 1H), 1.73 - 1.33 (m, 8H), 1.28 - 1.18 [m, 5H, containing 1.24 (s, 3H)], 1.15 - 1.00 [m, 7H, containing 1.09 (s, 3H)], 0.91 (d, *J* = 6.5 Hz, 3H), 0.89 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.78, 163.55 (d, $J_{C,F}$ = 249.5 Hz), 151.38, 147.20, 138.42, 138.17, 131.46, 131.37, 121.82, 116.28, 116.06, 115.56, 57.55, 56.47, 50.29, 42.91, 40.39, 39.04, 37.51, 36.05, 33.79, 32.75, 32.58, 32.08, 31.78, 31.70 (2C), 28.69, 24.73, 21.62 (2C), 18.65, 12.25; HRMS (ESI): calcd for $C_{33}H_{44}FN_2O_2$ $[M+H]^+$ 519.3381, found $[M+H]^+$ 519.3360.

Compound **31**: isolated in 80% yield as a white solid; mp: 236.7℃ - 237.5℃;^{1}H NMR (400 MHz, DMSO-d6) δ 7.41 - 7.29 (m, 3H), 7.22 - 7.14 (m, 2H), 5.71 (d, *J* = 3.0 Hz, 1H), 2.73 (d, *J* = 14.7 Hz, 1H), 2.36 (s, 3H), 2.22 (m, 1H), 2.16 - 1.99 (m, 4H), 1.83 (m, 1H), 1.71 - 1.56 (m, 4H), 1.51 - 1.36 (m, 3H), 1.29 - 1.19 [m, 6H, containing 1.25 (s, 3H)], 1.14 - 1.01 [m, 7H, containing 1.09 (s, 3H)], 0.91 (d, *J* = 6.6 Hz, 3H), 0.89 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.86, 151.32, 146.60, 142.00, 139.08, 137.54, 130.33, 129.72, 128.71, 126.16, 121.31, 114.97, 57.24, 56.14, 49.95, 42.63, 40.09, 38.78, 37.29, 35.77, 33.57, 32.49, 32.33, 31.78, 31.60 (2C), 31.44, 28.44, 24.49, 21.46, 21.35, 21.26, 18.54, 12.10; HRMS (ESI): calcd for $C_{34}H_{47}N_2O_2$ $[M+H]^+$ 515.3632, found $[M+H]^+$ 515.3647; $C_{34}H_{46}NaN_2O_2$ $[M+Na]^+$ 537.3457, found $[M+Na]^+$ 537.3444.

Compound **32**: isolated in 73% yield as a white solid; mp: 236.2℃ - 239.0℃;^{1}H NMR (400 MHz, DMSO-d6) δ 11.97 (br, 1H), 7.51 - 6.99 (m, 5H), 5.71 (d, *J* = 3.2 Hz, 1H), 2.74 (d, *J* = 14.5 Hz, 1H), 2.24 (m, 1H), 2.15 - 1.99 (m, 4H), 1.86 (s, 3H), 1.80 (m, 1H), 1.72 - 1.54 (m, 4H), 1.51 -1.36 (m, 3H), 1.31 (s, 3H), 1.27 - 1.15 (m, 4H), 1.13 - 1.01 [m, 5H, containing 1.05 (s, 3H)], 0.97 (m, 1H), 0.91 (s, 3H), 0.89 (d, *J* = 8.5 Hz, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.26, 150.91, 146.02, 140.85, 137.66, 130.85, 129.66, 128.67, 126.14, 121.15, 120.82, 114.70, 57.01, 55.90, 49.89, 42.42, 39.86, 38.61, 36.88, 35.50, 33.30, 32.23, 32.13, 31.55, 31.19, 31.13, 29.29, 28.22, 24.29, 21.20, 21.14, 18.36, 17.14, 11.95; HRMS (ESI): calcd for $C_{34}H_{47}N_2O_2$ $[M+H]^+$ 515.3632, found $[M+H]^+$ 515.3655.

Compound **33**: isolated in 72% yield as a white solid; mp: 232.9℃ - 233.8℃;^{1}H NMR (400 MHz, DMSO-d6) δ 7.62 (d, *J* = 8.1 Hz, 1H), 7.57 - 7.50 (m, 2H), 7.41 (d, *J* = 8.5 Hz, 1H), 7.38 (s, 1H), 5.72 (d, *J* = 3.2 Hz, 1H), 2.73 (d, *J* = 14.7 Hz, 1H), 2.21 (m, 1H), 2.14 - 1.98 (m, 4H), 1.81 (m, 1H), 1.71 - 1.36 (m, 8H), 1.29 - 1.19 [m, 5H, containing 1.25 (s, 3H)], 1.15 - 1.00 [m, 7H, containing 1.11 (s, 3H)], 0.91 (d, *J* = 6.7 Hz, 3H), 0.89 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.80, 151.09, 146.99, 143.29, 138.32, 134.61, 130.14, 130.01, 129.55, 127.61, 121.63, 115.52, 57.30, 56.21, 50.00, 42.69, 40.14, 38.82, 37.33, 35.83, 33.56, 32.51, 31.83, 31.80, 31.78, 31.66, 31.52, 28.49, 24.54, 21.51, 21.41, 18.55, 12.14; HRMS (ESI): calcd for $C_{33}H_{44}ClN_2O_2$ $[M+H]^+$ 535.3086, found $[M+H]^+$ 535.3087; $C_{33}H_{43}NaClN_2O_2$ $[M+Na]^+$ 557.2911, found $[M+Na]^+$ 557.2897.

Compound **34**: isolated in 68 % yield as a white solid; mp: 232.8℃ - 234.4℃;^{1}H NMR (400 MHz, DMSO-d6) δ 11.95 (br, 1H), 7.68 - 7.62 (m, 2H), 7.61 - 7.47 (m, 2H), 7.42 (s, 1H), 5.72 (d, *J* = 2.9 Hz, 1H), 2.74 (d, *J* = 14.7 Hz, 1H), 2.22 (m, 1H), 2.16 - 1.99 (m, 4H), 1.82 (m, 1H), 1.71 - 1.36 (m, 7H), 1.28 (s, 3H), 1.25 - 1.17 (m, 3H), 1.16 - 1.04 (m, 4H), 0.99 (s, 3H), 0.92 (d, *J* = 6.5 Hz, 3H), 0.89 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.54, 151.08, 147.25, 139.84, 138.73, 135.07, 132.01, 131.39, 130.63, 127.61, 121.51, 115.41, 57.32, 56.23, 49.99, 42.73, 40.18, 39.04, 37.32, 37.16, 35.83, 33.58, 32.52, 31.88, 31.50, 31.47, 29.38, 28.51, 24.57, 21.46, 21.20, 18.57, 12.16; HRMS (ESI): calcd for $C_{33}H_{44}ClN_2O_2$ $[M+H]^+$ 535.3086, found $[M+H]^+$ 535.3112; $C_{33}H_{43}NaClN_2O_2$ $[M+Na]^+$ 557.2911, found $[M+Na]^+$ 557.2917 .

Compound **35**: isolated in 69% yield as a white solid; mp: 228.4℃ - 229.4℃;^{1}H NMR (400 MHz, DMSO-d6) δ 11.94 (br, 1H), 7.36 (s, 1H), 7.23 (m, 1H), 7.18 - 7.06 (m, 2H), 5.70 (d, *J* = 3.0 Hz, 1H), 2.73 (d, *J* = 14.6 Hz, 1H), 2.35 (s, 3H), 2.23 (m, 1H), 2.15 - 1.98 (m, 4H), 1.82 (m, 1H), 1.80 (s, 3H), 1.71 - 1.55 (m, 4H), 1.51 - 1.37 (m, 3H), 1.30 (s, 3H), 1.25 - 1.18 (m, 3H), 1.14 - 1.04 (m, 4H), 0.92 (s, 3H), 0.91 (d, *J* = 6.6 Hz, 3H), 0.87 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.44, 151.09, 146.43, 139.96, 138.39, 137.46, 131.69, 129.72, 128.57, 126.96, 121.38, 114.93, 57.24, 56.14, 49.89, 42.62, 40.09, 38.80, 38.70, 37.27, 37.11, 35.73, 33.80, 33.50, 32.58, 32.43, 32.22, 31.79, 31.37, 29.39, 28.41, 24.47, 21.35, 18.54, 12.11; HRMS (ESI): calcd for $C_{35}H_{49}N_2O_2$ $[M+H]^+$ 529.3832, found $[M+H]^+$ 529.3813; $C_{35}H_{48}NaN_2O_2$ $[M+Na]^+$ 551.3613, found $[M+Na]^+$ 551.3615

Compound **36**: isolated in 70% yield as a white solid; mp: 229.4℃ - 230.5℃;^{1}H NMR (400 MHz, DMSO-d6) δ 7.31 (s, 1H), 7.24 (d, *J* = 7.9 Hz, 1H), 7.14 (s, 1H), 7.08 (dd, *J* = 7.7, 1.8 Hz, 1H), 5.70 (d, *J* = 3.1 Hz, 1H), 2.72 (d, *J* = 14.6 Hz, 1H), 2.29 (s, 3H), 2.26 (s, 3H), 2.22 (m, 1H), 2.15 - 1.99 (m, 4H), 1.82 (m, 1H), 1.71 - 1.35 (m, 8H), 1.28 - 1.20 [m, 5H, containing 1.25 (s, 3H)], 1.14 - 1.01 [m, 7H, containing 1.09 (s, 3H)], 0.91 (d, *J* = 6.5 Hz, 3H), 0.88 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.92, 151.73, 146.93, 140.06, 138.71, 137.79, 137.65, 130.35, 130.24, 126.72, 121.63, 115.21, 57.65, 56.57, 50.39, 42.98, 40.49, 39.11, 37.62, 36.13, 33.92, 32.83, 32.56, 32.17, 31.81 (2C), 31.75, 28.75, 24.80, 21.69, 21.65, 19.84, 19.70, 18.69, 12.29; HRMS (ESI): calcd for $C_{35}H_{49}N_2O_2$ $[M+H]^+$ 529.3832, found $[M+H]^+$ 529.3849.

Compound **37**: isolated in 71% yield as a white solid; mp: 237.8℃ - 238.8℃; ^{1}H NMR (400 MHz, DMSO-d6) δ 11.91 (br, 1H), 7.82 - 7.74 (m, 2H), 7.46 (dd, *J* = 8.5, 2.3 Hz, 1H), 7.40 (s, 1H), 5.73 (d, *J* = 3.1 Hz, 1H), 2.74 (d, *J* = 14.7 Hz, 1H), 2.23 (m, 1H), 2.16 - 1.99 (m, 4H), 1.82 (m, 1H), 1.71 - 1.55 (m, 4H), 1.52 - 1.37 (m, 3H), 1.34 - 1.18 [m, 6H, containing 1.25 (s, 3H)], 1.15 - 1.00 [m, 7H, containing 1.13 (s, 3H)], 0.91 (d, *J* = 6.5 Hz, 3H), 0.89 (s, 3H), 0.67 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$ /CD_3OD) δ 177.50, 150.88, 147.05, 141.48, 138.62, 134.11, 132.92, 131.16, 130.75, 128.64, 121.69, 115.68, 57.21, 56.12, 49.91, 42.63, 40.06, 38.74, 37.25, 35.74, 33.48, 32.60, 32.46, 31.79, 31.74, 31.43, 31.38, 28.43, 24.48, 21.47, 21.34, 18.50, 12.10; HRMS (ESI): calcd for $C_{33}H_{43}Cl_2N_2O_2$ $[M+H]^+$ 569.2696, found $[M+H]^+$ 569.2698; $C_{33}H_{42}NaCl_2N_2O_2$ $[M+Na]^+$ 591.2521, found $[M+Na]^+$ 591.2520 .

Biological assays

Enzyme-based assay of PTP1B and related homologous PTPs

PTP1B hydrolyzes pNPP to pNP which can be detected at 405 nm. Briefly, the tested compounds were solubilized in DMSO and serially diluted into series of concentrations for the inhibitory test. The assays which were performed in 96-well plates were carried out in a final volume of 100 μL containing 50 mmol/L MOPS, pH 6.5, 2 mmol/L pNPP, 30 nmol/L GST-PTP1B, and 2% DMSO, and the catalysis of pNPP was continuously monitored on a SpectraMax 340 microplate reader at 405 nm for 2 min at 30 ℃. The IC_{50} value was calculated from the nonlinear curve fitting of the percent inhibition [inhibition (%)] vs the inhibitor concentration [I] using the following equation: % inhibition = 100/{1 + (IC_{50}/[I]) *k*}, where *k* is the Hill coefficient. The inhibition assay of TCPTP was conducted by the same procedure.

To study the inhibition on the other homologous PTPs, SHP1, SHP2 and LAR assays were performed according to procedures described previously[21].

Effect of PTP1B inhibitors on the phosphorylation level of IR and Akt in CHO-hIR cells

CHO-hIR cells were cultured in Ham's F-12 medium (Invitrogen, Carlsbad, CA) supplemented with 10% (v/v) FBS, 100 units/mL penicillin and 100 mg/mL streptomycin at 37 ℃ in 5% CO_2. The cells were starved in serum-free medium for 2 h and then treated with PTP1B inhibitors for 3 h, followed by stimulation with 10 nM insulin for 10 min. Then cells were washed three times with ice cold PBS and lysed with lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% NP-40, 1 mM Na_3VO_4, 1 mM PMSF, 1 mM DTT, 1 mM EDTA, 1 mM EGTA) containing complete protease inhibitors (Roche). The immunoblots were visualized by chemiluminescence using the enhanced chemiluminescence Western Blotting System.

Characterization of the inhibitor on enzyme kinetics [22]

In the time-independent inhibition experiment, PTP1B was preincubated with compounds (2% DMSO) on the ice for different times, and then add 10 μL mixture of enzyme and compounds to 90 μL assay system. To characterize the inhibitor of PTP1B, the assay was carried out in a 100 μL system containing 50 mmol/L MOPS, pH 6.5, 30 nmol/L PTP1B, *p*NPP in 2-fold dilution from 80 mmol/L, and different concentrations of the inhibitor. In the presence of the competitive inhibitor, the Michaelis-Menten equation is described as $1/v = (K_m/[V_{max}[S]])(1+[I]/K_i)+1/V_{max}$, where K_m is the Michaelis constant, *v* is the initial rate, *V*max is the maximum rate, and [S] is the substrate concentration. The K_i value was obtained by the linear replot of apparent K_m/V_{max} (slope) from the primary reciprocal plot versus the inhibitor concentration [I] according to the equation $K_m/V_{max} = 1 + [I]/K_i$.

Acknowledgments

This research was financially supported by Shanghai Science and Technology Council (No. 13142200900, 14DZ0511800, 14142201200), National Science and Technology Major Projects for Major New Drugs Innovation and Development (2012ZX09304011 and 2013ZX09507002), National science Fund for Distinguished Young Scholars (81125023). We also thank the Laboratory of Organic Functional Molecules, Sino-French Institute of ECNU for support.

References

1. A. R. Saltiel, C. R. Kahn, *Nature* 2001, 414, 799-806.
2. D. Bandyopadhyay, A. Kusari, K. A. Kenner, F. Liu, J. Chernoff, T. A. Gustafson, J. Kusari, *J. Biol. Chem.* 1997, 272, 1639-1645.
3. B. J. Goldstein, A. Bittner-Kowalczyk, M. F. White, M. Harbeck, *J. Biol. Chem.* 2000, 275, 4283-4289.
4. F. Ahmad, J. L. Azevedo, J. R. Cortright, G. L. Dohm, B. J. Goldstein, *J. Clin. Invest.* 1997, 100, 449-458.
5. F. Ahmad, B. J. Goldstein, *Metabolism* 1995, 44, 1175-1184.
6. M. Elchebly, P. Payette, E. Michaliszyn, W. Cromlish, S. Collins, A. L. Loy, et al., *Science* 1999, 283, 1544-1548.
7. L. D. Klaman, O. Boss, O. D. Peroni, J. K. Kim, J. L. Martino, J. M. Zabolotny, et al. *Mol. Cell. Biol.* 2000, 20, 5479-5489.
8. J. Montalibet, B. P. Kennedy, *Drug Discov. Today: Therap. Strat.* 2005, 2, 129-135.
9. L. F. Iversen, K. B. Moller, A. K. Pedersen, G. H. Peters, A. S. Petersen, H. S. Andersen, et al., *J. Biol. Chem.* 2002, 277, 19982-19990.
10. Y. Xi, S. Liu, A. Bettaieb, K. Matsuo, I. Matsuo, E. Hosein, et al., *Diabetologia* 2015, 58, 122-131.
11. A. J. Barr, *Future Med. Chem.* 2010, 2, 1563-1576.
12. C. S. Jiang, L. F. Liang, Y. W. Guo, *Acta Pharmaco. Sinica* 2012, 33, 1217-1245.
13. L. J. Wang, B. Jiang, N. Wu, S. Y. Wang and D. Y. Shi, *RSC Adv.*, 2015, 5, 48822-48834
14. K. A. Lantz, S. G. E. Hart, S. L. Planey, M. F. Roitman, I. A. Ruiz-White, H. R. Wolfe, M. P. Mclane, *Obesity* 2010, 18, 1516-1523.
15. M. F. Roitman, S. Wescott, J. J. Core, M. P. Mclane, H. R. Wolfe, *Pharmacology, Biochem. Behav.* 2010, 97, 138-143.
16. J. J. Ram ńez-Espinosa, M. Y. Rios, S. L ópez-Mart ńez, F. L ópez-Vallejo, J. L. Medina-Franco, P. Paoli, et al., *Eur. J. Med. Chem.* 2011, 46, 2243-2251.
17. H. B. He, L. X. Gao, Q. F. Deng, W. P. Ma, C. L. Tang, W. W. Qiu, et al., *Bioorg. Med. Chem. Lett.* 2012, 22, 7237-7242.
18. CCDC 1401731 contains the supplementary crystallographic data of compound **33** for this paper. These data can be obtained free of charge via http://www.ccdc.cam.ac.uk/cgi-bin/catreq.cgi or by emailing data-request@ccdc.cam.ac.uk or by contacting the Cambridge Crystallographic Data Centre, 12, Union Road, Cambridge CB21EZ, U.K. (fax, +44 1223 336 33).
19. C. J. Bungard, G. D. Hartman, J. J. Manikowski, J. J. Perkins, C. Bai, et al., *Bioorg. Med. Chem.* 2011, 19, 7374-7386.
20. The X-ray crystal structure of PTP1B was taken from PDB (1NNY) with 2.40 Å resolution used for the docking studies.
21. L. Shi, H. P. Yu, Y. Y. Zhou, J. Q. Du, Q. Shen, J. Y. Li, J. Li, *Acta Pharmaco. Sinica* 2008, 29, 278-284.
22. W. Zhang, D. Hong, Y. Zhou, Y. Zhang, Q. Shen, J. Y. Li, L. H. Hu, J. Li, *Biochim. Biophys. Acta.* 2006, 1760, 1505-1512.

Table captions:

Table 1. Inhibitory activities against PTP1B and TCPTP

Table 2. Inhibition (%) of compounds **27**, **28**, **30**, **31**, **32** and **33** against other homologous PTPs at 40 μM

Figure captions:

Figure 1. Structures of selected pentacyclic triterpenoids and steroids as PTP1B inhibitors

Figure 2. X-ray structure of compound **33**

Figure 3. Effect of **30** and **34** on IR phosphorylation and Akt phosphorylation in CHO/hIR cells.

Figure 4. Characterization of **30** to PTP1B

Figure 5. Molecular docking analysis of **30** and PTP1B

Scheme 1 General synthesis of 2,3-pyrazole ring-substituted-4,4-dimethyl derivatives

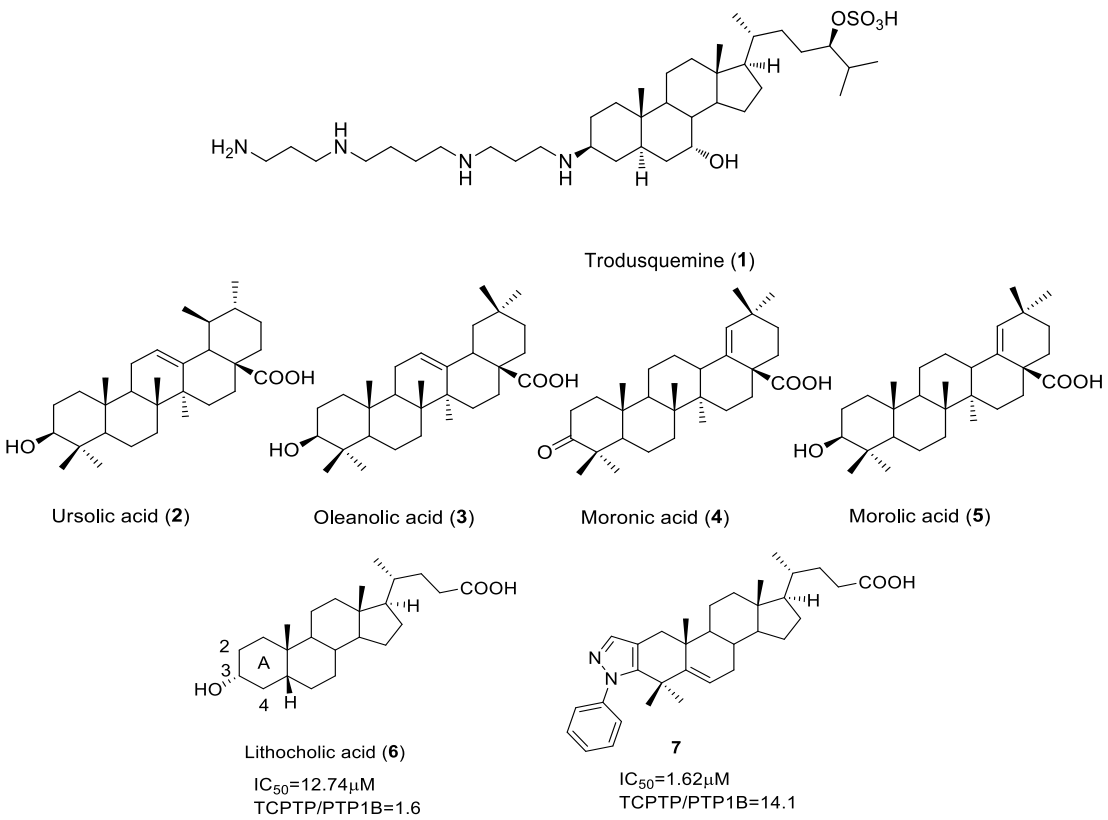
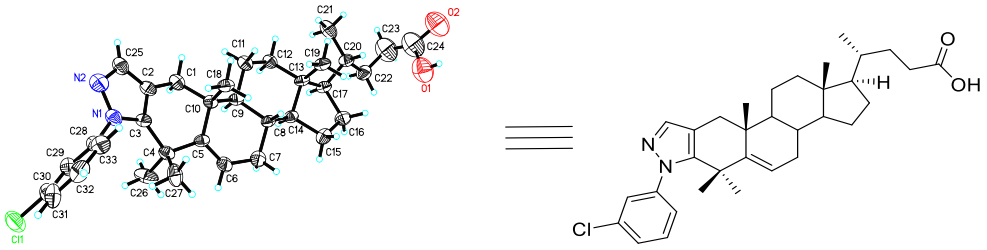


Figure 1. Structures of selected pentacyclic triterpenoids and steroids as PTP1B inhibitors



33

Figure 2. X-ray structure of compound **33**

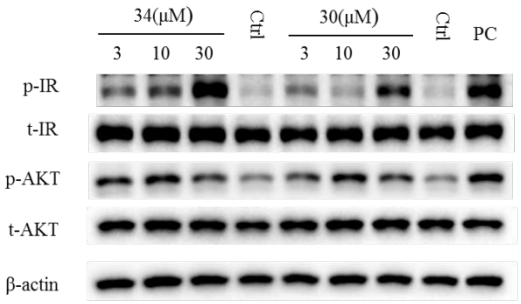


Figure 3. Effect of **30** and **34** on IR phosphorylation and Akt phosphorylation in CHO/hIR cells. CHO/hIR cells were incubated with 200 μM sodium orthovanadate, 0.2% DMSO, or compound **30** (or **34**) for 3h, and then treated with 10 nM insulin for 10 min. V: 200 μM sodium orthovanadate used as a positive control (PC).

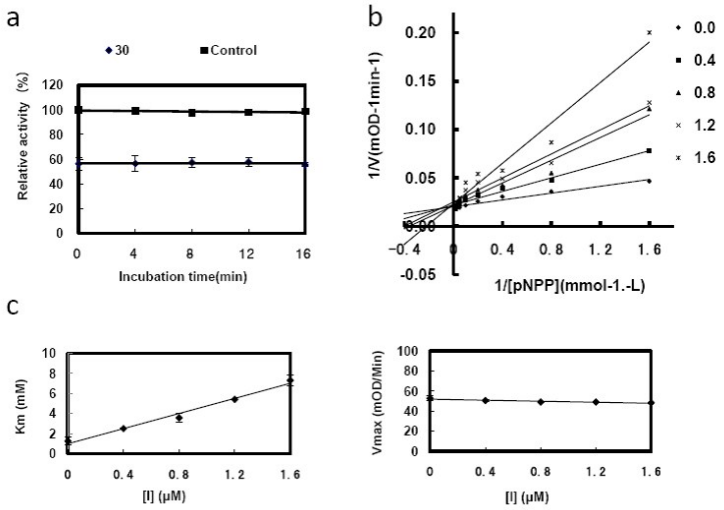


Figure 4. Characterization of **30** to PTP1B. (a)Time-independent inhibition of PTP1B by **30**. (b) Typical competitive inhibition of **30** shown by Lineweaver-Burk plot. (c) At various fixed concentrations of **30** the initial velocity was determined with various concentrations of *p*NPP.

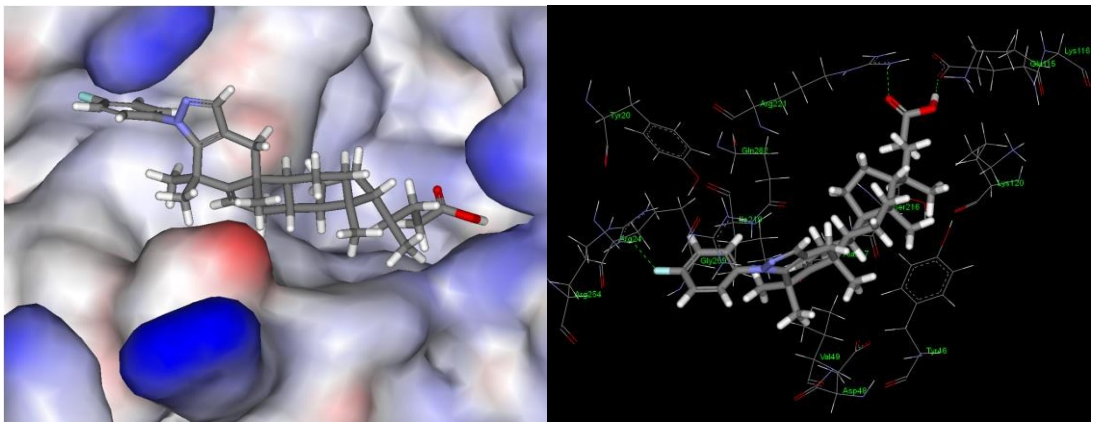


Figure 5. (a) Binding pose of **30** in the protein surface of PTP1B; (b) key residues of PTP1B binding site surrounding **30**.

Table 1
Inhibitory activities against PTP1B and TCPTP

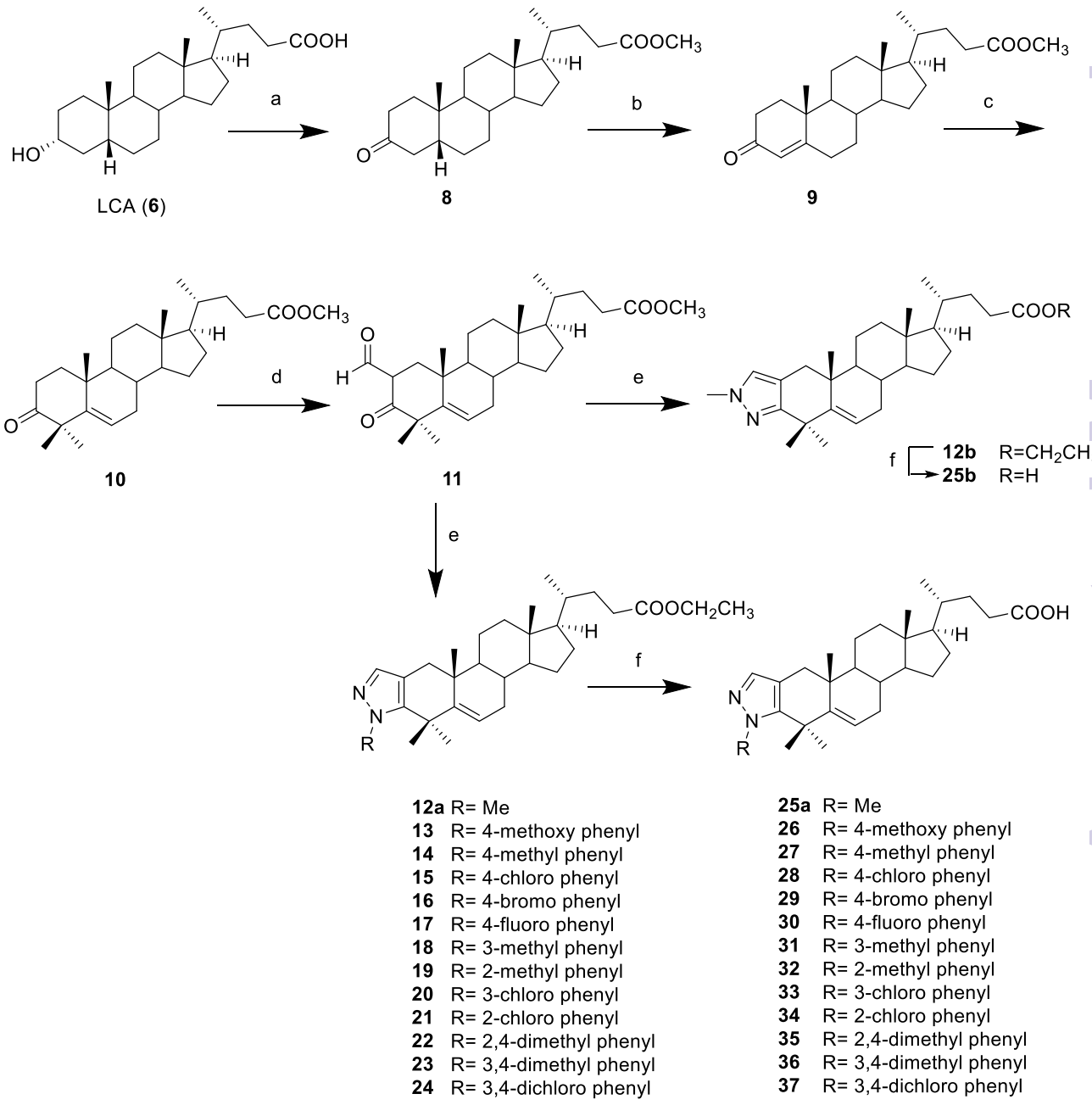
Compounds	IC_{50} (μM)		TCPTP/PTP1B[a]
	PTP1B	TCPTP	
25a	2.82±0.57	>40[b]	>14.2
25b	4.49±0.68	NA[c]	>8.9
26	2.66±0.51	>40	>15
27	0.81±0.06	>40	>49.4
28	1.12±0.11	NA	>35.7
29	2.35±0.27	>40	>17
30	0.42±0.07	4.53±0.19	10.8
31	0.93±0.19	NA	>43
32	1.17±0.27	>40	>34.2
33	1.24±0.22	NA	>32.3
34	0.86±0.03	3.99±0.18	4.6
35	1.34±0.25	28.99±1.61	21.6
36	0.73±0.09	23.41±0.53	32.1
37	1.03±0.05	5.70±0.36	5.5
6	12.74±1.41	20.50±2.10	1.6
7	1.62±0.08	22.78±4.36	14.1
OA[d]	2.78±0.19	6.00±0.15	2.1

[a] TCPTP/PTP1B, the ratio of IC_{50} of TCPTP and PTP1B
[b] IC_{50}＞40μM, the inhibition (%) on TCPTP was between 5% and 50% at 40μM
[c] The inhibition (%) on TCPTP was below 5% at 40μM
[d] Oleanolic acid as positive control

Table 2
Inhibition (%) of compounds **27**, **28**, **30**, **31**, **32** and **33** against other homologous PTPs at 40μM

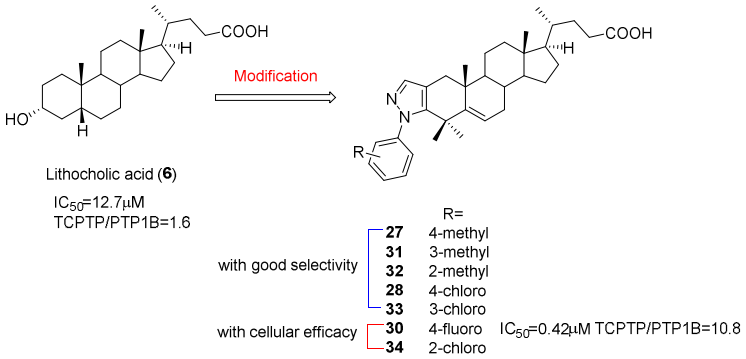
Compounds	Inhibition (%)[a]		
	LAR	SHP1	SHP2
27	20.62	23.79	12.31
28	6.25	5.29	2.15
30	24.56	19.82	22.25
31	24.75	6.26	4.12
32	1.02	3.90	5.59
33	4.09	8.30	3.90

[a] Inhibition at 40 μM



Scheme 1. Reagents and conditions: (a) (i) MeOH, $SOCl_2$, 99%, (ii) PCC, DCM, 93%; (b) (i) HOAc, DCM, Br_2,(ii) Li_2CO_3, $LiBr{\cdot}H_2O$, DMF, 48%; (c) *t*-BuOK, *t*-BuOH, MeI, 50%; (d) NaH, HCOOEt, 98%; (e) R-$NHNH_2{\cdot}$HCl, EtOH, 36%-62%; (f) MeOH/H_2O, $LiOH{\cdot}H_2O$, 62%-81%.

Graphical Abstract



Abstract

In our continued efforts to develop lithocholic acid (LCA) analogues as selective PTP1B inhibitors, 14 novel 2,3-pyrazole ring-substituted-4,4-dimethyl derivatives were synthesized and evaluated against PTP1B, as well as the homologous protein tyrosine phosphatases (PTPs). All compounds were shown to be more potent and selective PTP1B inhibitors than LCA (IC_{50} = 12.74 μM) with IC_{50} values ranging between 0.42 to 4.49 μM. Moreover, treatment of CHO/hIR cells with 4,4-dimethyl-2'-(*p*-fluoro phenyl)-2′*H*-chola-2,5-dieno[3,2-*c*]pyrazol-24-oic acid (**30**) or 4,4-dimethyl-2'-(*o*-chloro phenyl)-2′*H*-chola-2,5-dieno[3,2-*c*]pyrazol-24-oic acid (**34**) increased the phosphorylation levels of IR and Akt in a dose dependent manner. The promising findings in this study suggest that further investigation of these compounds for the treatment of metabolic disorders is warranted.