Available online at www.sciencedirect.com
ScienceDirect
Chinese Journal of Natural Medicines 2023, **21**(1): 65-80
doi: 10.1016/S1875-5364(23)60386-3
Chinese Journal of Natural Medicines

•Original article•

Flavonoids from the roots and rhizomes of *Sophora tonkinensis* and their *in vitro* anti-SARS-CoV-2 activity

LI Zhuo[1, 2Δ], XIE Hang[1, 4Δ], TANG Chunping[2], FENG Lu[2], KE Changqiang[2], XU Yechun[3, 4, 5], SU Haixia[4*], YAO Sheng[2*], YE Yang[1, 2, 3*]

[1] *School of Chinese Materia Medica, Nanjing University of Chinese Medicine, Nanjing 210023, China;*
[2] *State Key Laboratory of Drug Research, and Natural Products Chemistry Department, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China;*
[3] *School of Pharmacy, University of Chinese Academy of Sciences, Beijing 100049, China;*
[4] *Drug Discovery and Design Center, CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China;*
[5] *School of Pharmaceutical Science and Technology, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China*

Available online 20 Jan., 2023

[ABSTRACT] Acute respiratory infection caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) had caused a global pandemic since 2019, and posed a serious threat to global health security. Traditional Chinese medicine (TCM) has played an indispensable role in the battle against the epidemic. Many components originated from TCMs were found to inhibit the production of SARS-CoV-2 3C-like protease (3CLpro) and papain-like protease (PLpro), which are two promising therapeutic targets to inhibit SARS-CoV-2. This study describes a systematic investigation of the roots and rhizomes of *Sophora tonkinensis*, which results in the characterization of 12 new flavonoids, including seven prenylated flavanones (**1**−**7**), one prenylated flavonol (**8**), two prenylated chalcones (**9**−**10**), one isoflavanone (**11**), and one isoflavan dimer (**12**), together with 43 known compounds (**13**−**55**). Their structures including the absolute configurations were elucidated by comprehensive analysis of MS, 1D and 2D NMR data, and time-dependent density functional theory electronic circular dichroism (TDDFT ECD) calculations. Compounds **12** and **51** exhibited inhibitory effects against SARS-CoV-2 3CLpro with IC_{50} values of 34.89 and 19.88 μmol·L^{-1}, repectively while compounds **9**, **43** and **47** exhibited inhibitory effects against PLpro with IC_{50} values of 32.67, 79.38, and 16.74 μmol·L^{-1}, respectively.

[KEY WORDS] *Sophora tonkinensis*; Flavonoid; Anti-SARS-CoV-2 activity; SARS-CoV-2 3CLpro; SARS-CoV-2 PLpro
[CLC Number] R284, R965 **[Document code]** A **[Article ID]** 2095-6975(2023)01-0065-16

Introduction

Coronavirus disease 2019 (COVID-19) [1] is a pandemic disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Two cysteine proteases, a chymotrypsin-like cysteine protease called 3C-like protease (3CLpro) and a papain-like protease (PLpro), responsible for proteolysis of new virions, are essential for the design of therapeutic targets [2, 3]. Specifically, 3CLpro is considered to act as the main protease (Mpro) due to its key role in mediating viral replication and transcription [4]. Thus, it is necessary to develop more structurally diverse small molecules such as 3CLpro and PLpro inhibitors for the treatment of COVID-19.

Traditional Chinese medicines have provided effective strategies to promote the prevention and treatment of COVID-19. Many natural compounds originated from medicinal herbs showed inhibitory activity against SARS-CoV-2 (Fig. 1) [5-7], and much attention has been drawn on the inhibitory activity against SARS-CoV-2 3CLpro and PLpro. For example, baicalin and baicalein, two natural flavonoids isolated from *Scutellaria baicalensis* [8], and four polyphenols from the fer-

[Received on] 26-Jul.-2022
[Research funding] This work was supported from the National Natural Science Foundation of China (Nos. 82173696 and 21920102003), the Science and Technology Commission of Shanghai Municipality (Nos. 20431900200 and 20430780300), the Strategic Priority Research Program of Chinese Academy of Sciences (Nos. SIMM010110 and SIMM040302) and the Sustainable Development of Precious Traditional Chinese Medicine Resources (No. 2060302-2001-01).
[*]**Corresponding author]** E-mails: suhaixia1@simm.ac.cn (Su Haixia); yaosheng@simm.ac.cn (YAO Sheng); yye@simm.ac.cn (YE Yang)
[Δ]These authors contributed equally to this work.
These authors have no conflict of interest to declare.

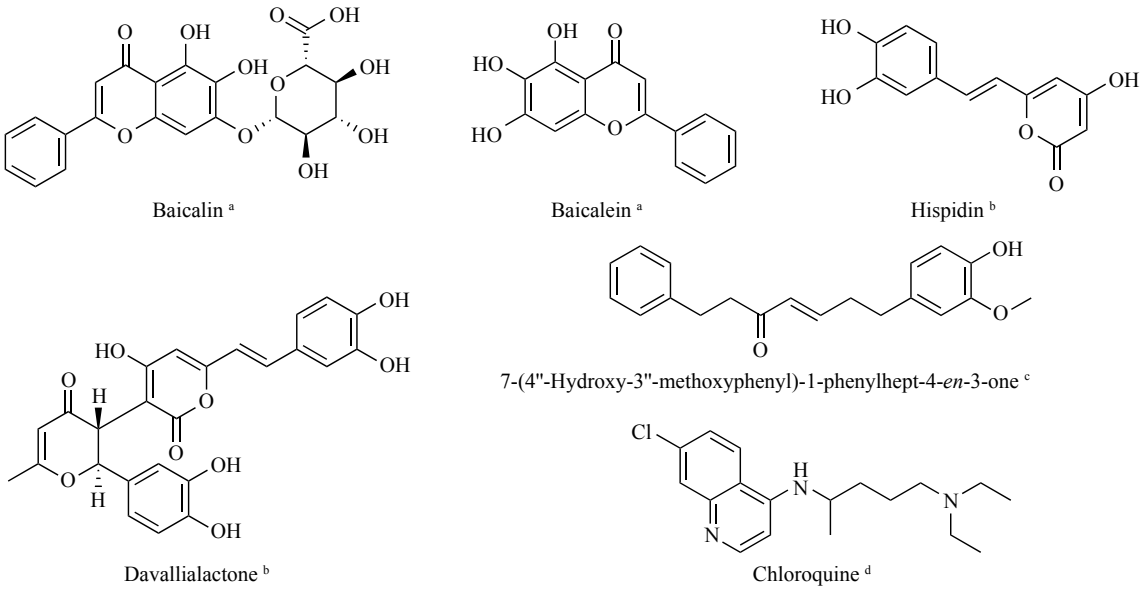

Fig. 1 Representative natural products showing inhibitory activity against SARS-CoV-2 3CLpro and PLpro. Compounds were isolated from a: *Scutellaria baicalensis*, b: *Sanghuangporus sanghuang*, c: *Alpinia officinarum*, and d: *Cinchona officinalis*

mentation products of *Sanghuangporus sanghuang* [9] exhibited promising SARS-CoV-2 3CLpro inhibitory activity. Eight natural compounds found in rhizomes of *Alpinia officinarum* and ginger were identified as potential inhibitors of SARS-CoV-2 PLpro [10]. Additionally, chloroquine [11] and glycyrrhizin [12], two natural compounds obtained from *Cinchona officinalis* and *Glycyrrhiza glabra*, respectively, were both efficient to block the interaction of RBD of SARS-CoV to ACE-2.

In our previous research, the oral liquid and the lyophilized powder of Shuang-Huang-Lian injection, and their bioactive components were found to dose-dependently inhibit SARS-CoV-2 3CLpro and the replication of SARS-CoV-2 in Vero E6 cells. Another investigation showed that the pyrogallol group of a natural flavonoid named myricetin emerged as a warhead that was covalently linked to cysteine under oxidative conditions, which provides a vital clue to understand the diverse bioactivities of natural products and phenolic natural products are identified as covalent ligands [13]. In addition, a native MS-based affinity-selection method named bio-affinity-MS (BA-MS) was established in our lab, and has been successfully utilized to discover three flavonoids, baicalein, baicalein and glycyrrhizin from herbal extracts as potential noncovalent inhibitors against SARS-CoV-2 3CLpro [14].

In the high-throughput screening of natural products in our in-house compound library, we found that the extract of a plant named *Sophora tonkinensis* showed inhibitory effects against 3CLpro. *Sophora tonkinensis* Gagnep. (Leguminosae), known as “Shan-Dou-Gen” in Chinese, is widely distributed in Southern China and Vietnam. Its rhizomes and roots have long been used as a traditional Chinese medicine for the treatment of throat-swelling and acute pharyngolaryngeal infections [15]. Previous investigations revealed that the plant is rich in prenylated flavonoids [16-18] and quinolizidine alkaloids [19-21]. To search for new flavonoids with anti-SARS-CoV-2 activity, the EtOH extract of this plant was investigated, and the active fractions with inhibitory effects on 3CLpro and PLpro were selected for isolation, resulting in the characterization of 12 undescribed flavonoids (Fig. 2), including seven prenylated flavanones (**1**−**7**), one prenylated flavonol (**8**), two prenylated chalcones (**9**−**10**), one isoflavanone (**11**), and one isoflavan dimer (**12**), together with 43 known analogs (**13**−**55**) (Fig. S1, Supporting Information). Their planar structures were elucidated by extensive analysis of spectroscopic data and DFT NMR calculations. Their absolute configurations were further determined by TDDFT ECD calculations. It is noteworthy that **2** and **11** are two pairs of enantiomers, and their racemic nature were resolved by chiral HPLC, which permitted online stereochemical characterization when used in combination with circular dichroism (CD) spectroscopy (LC-CD coupling). All the compounds were evaluated for anti-SARS-CoV-2 activities through suppressing the production of 3CLpro and PLpro.

Experimental

Instruments, reagents and chemicals

IR spectra were measured on a Thermo Nicolet FTIR IS5 spectrometer (Thermo Fisher, USA). UV spectra were measured on an Agilent Cary 300 Series UV-Vis spectrophotometer (Agilent, USA). Optical rotation values were recorded by a Rudolph Research Analytical Autopol VI 90079 polarimeter (Hackettstown, NJ). ECD spectra were recorded on a Jasco-815 CD spectropolarimeter (JASCO, Japan). HR-ESI-MS spectra were obtained by a Waters Synapt G2-Si Q-TOF Mass Spectrometer. NMR spectra were collected on a Bruker AVANCE III 500 or 600 MHz instrument (Bruker Biospin AG, Switzerland). Chemical shifts were reported in ppm (δ) with coupling constants (*J*) in Hertz. Analytical HPLC was

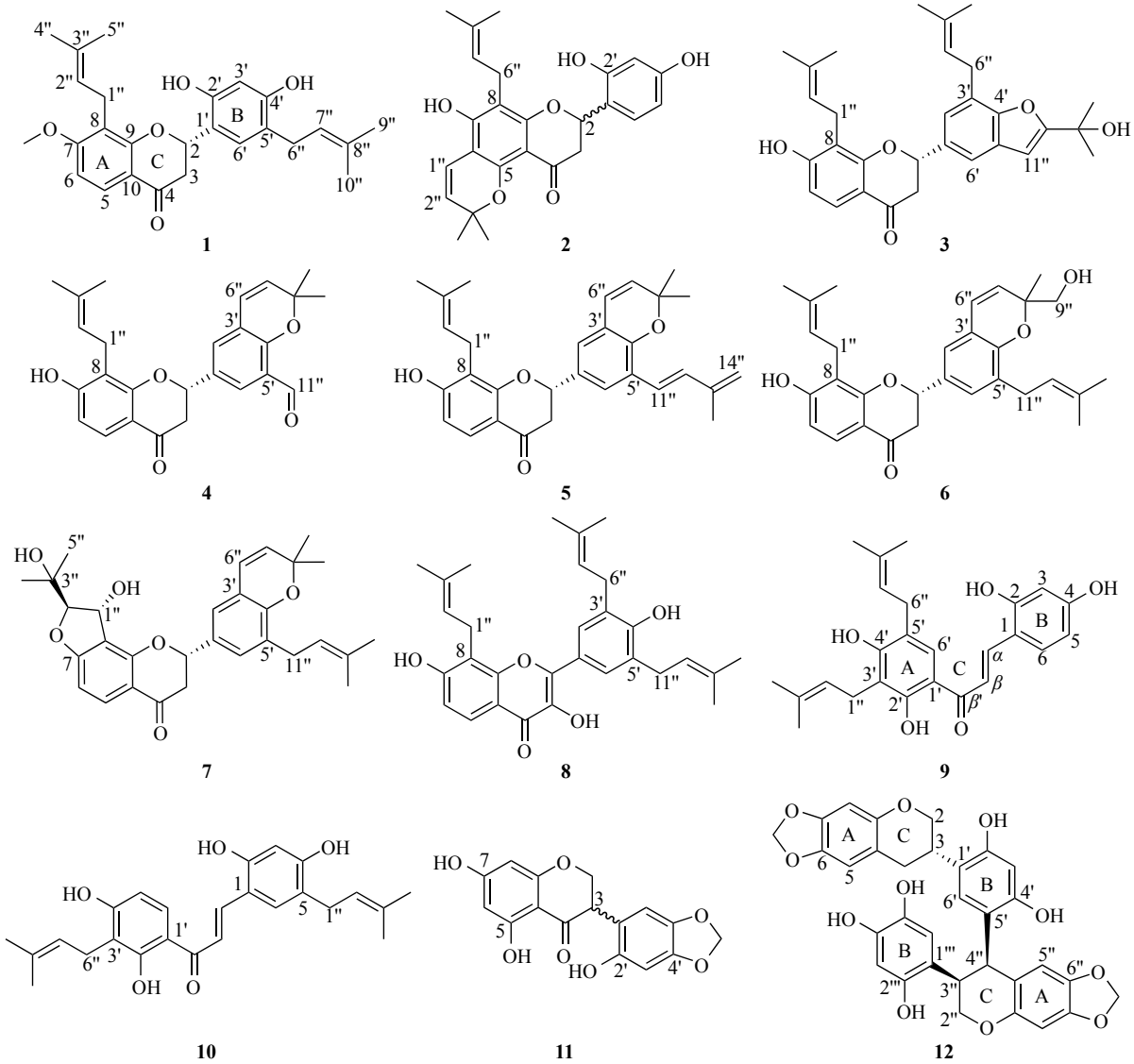

Fig. 2 Structures of compounds 1−12

performed on a Waters 2695 instrument with a Waters 2998 PDA detector, a Waters 2424 ELSD, and Waters 3100 MS detectors. Preparative HPLC was run on a Waters system with an Alltech 2424 ELSD and 2489 PDA using a Waters Sunfire RP C_{18} column (5 μm, 30 mm × 150 mm, acetonitrile−water and at a flow rate of 30 mL·min^{-1}). MCI gel CHP20P (75−150 μm, Mitsu bishi Chemical Industries, Japan), Sephadex LH-20 (Pharmacia Biotech AB, Uppsala, Sweden), and Silica gel (100−200, 200−300 and 300−400 mesh, Qingdao Marine Chemical Co., China) were used for column chromatography (CC). TLC analysis was carried out on precoated GF_{254} silica gel plates (Yantai Chemical Industrials, Yantai, China), and the spots were observed at 254 nm and visualized by 5% sulfuric acid in alcohol containing 10 mg·mL^{-1} of vanillin followed by heating. All solvents were of analytical grade (Sinopharm Chemical Reagents Co., Ltd., Shanghai, China) for column chromatography, and of HPLC grade (Merck KGaA, Darmstadt, Germany and Ourchem, Shanghai, China) for HPLC and preparative HPLC.

Plant material

The roots and rhizomes of *Sophora tonkinensis* Gapnep. were purchased from Shanghai Tongrentang Pharmaceutical Co., Ltd. (Shanghai, China), and authenticated by ZHANG Xiaoming, Shanghai Kangqiao Medicine Yinpian Co., Ltd. (Shanghai, China). A voucher specimen (No. 20210401) was deposited at the herbarium of Shanghai Institute of Materia Medica, Chinese Academy of Sciences (Shanghai, China).

Extraction, isolation and spectroscopic data

The dried roots and rhizomes of *S. tonkinensis* (10 kg) were extracted with 95% EtOH at room temperature three times (7 days for each). The percolates were combined and evaporated under reduced pressure to give a crude extract (680 g). The extract was then suspended in water, and partitioned successively with petroleum ether (PE) and ethyl acetate (EA) three times, yielding fractions PE (Fr. A, 120.0 g), EA (Fr. B, 244.0 g), and water-soluble fraction (Fr. C). Frac-

tion B was chromatographed over an MCI column eluted with aqueous MeOH (30%, 50%, 70% and 100%) to give Frs. B1−B4. Fr. B4 (126 g) was then subjected to MCI column chromatography (CC), eluting with EtOH in water (70%, 75%, 80%, 85%, 90%, and 95%) to afford Frs. B4A to B4J. Fr. B4C (14.0 g) was separated by a Sephadex LH-20 gel column ($CHCl_3$: MeOH, 1 : 1, *V/V*) to afford Frs. B4C1−B4C7. All fractions were evaluated for inhibitory effects against SARS-CoV-2 3CLpro and PLpro. Among them, Fr. B4C6 exhibited inhibitory effects against SARS-CoV-2 3CLpro with an IC_{50} value of 6.22 ± 0.36 μg·mL^{-1}, while Fr. B4G and Fr. B4H exhibited inhibitory effects against PLpro with IC_{50} values of 17.52 ± 0.01, 36.10 ± 0.45 μg·mL^{-1}, respectively. Fr. B4C6 (1.6 g) was then subjected to silica gel CC using a solvent mixture of PE : EA (5 : 1 to 1 : 1, *V/V*), and further purified by preparative HPLC (30%−50% CH_3CN in H_2O, 25 min), yielding seven compounds **11** (3.8 mg), **12** (6.5 mg), **41** (3.5 mg), **42** (1.4 mg), **44** (9.2 mg), **54** (98.2 mg), and **55** (20.2 mg). Fr. B4H (24.1 g) was fractionated using a Sephadex LH-20 gel (MeOH) column to provide Frs. B4H1−B4H6. Fr. B4H4 (345.0 mg) was applied to a Sephadex LH-20 gel column ($CHCl_3$: MeOH, 1 : 1, *V/V*) to give subfractions B4H4A−B4H4F. Fr. B4H4A (10.0 mg) was purified by preparative HPLC (70%−90% CH_3CN in H_2O, 25 min), yielding compound **46** (2.1 mg). Frs. B4H4C (40.9 mg) and B4H4D (18.3 mg) were further purified using preparative HPLC (55%−75% CH_3CN in H_2O, 25 min), respectively, to yield compounds **16** (1.9 mg), **18** (2.1 mg), **27** (2.0 mg), and **29** (8.5 mg). Fr. B4G (14.3 g) was combined with Fr. B4H6 (4.1 g), and then purified by Sephadex LH-20 column eluting with MeOH to give subfractions B4G1−B4G6. Fr. B4G2 (3.8 g) was separated using a Sephadex LH-20 gel (MeOH) column to yield Frs. B4G2A−B4G2F. Fr. B4G2E (2.4 g) was repeatedly separated by silica gel CC using CH_2Cl_2 : MeOH (200 : 1 to 50 : 1, *V/V*), and finally purified by preparative HPLC (60%−80% CH_3CN in H_2O, 25 min), yielding compounds **3** (44.0 mg), **7** (3.5 mg), **17** (20.5 mg), **20** (4.1 mg), **21** (4.7 mg), **28** (9.7 mg), **33** (178.8 mg), **34** (123.5 mg), **37** (14.2 mg), and **48** (11.1 mg). Fr. B4G3 (2.1 g) was repeatedly separated by silica gel CC using a solvent mixture of CH_2Cl_2 : acetone (150 : 1 to 50 : 1, *V/V*), and finally by preparative HPLC (65%−85% CH_3CN in H_2O, 25 min), affording compounds **1** (22.6 mg), **4** (5.6 mg), **5** (2.4 mg), **6** (17.5 mg), **14** (24.0 mg), **19** (5.1 mg), **22** (23.4 mg), **35** (3.6 mg) and **38** (8.9 mg). Fr. B4G4 (4.4 g) was fractionated using a Sephadex LH-20 gel column ($CHCl_3$: MeOH, 1 : 1, *V/V*) to obtain Frs. B4G4A−B4G4C. Fr. B4G4C (785.7 mg) was repeatedly separated by silica gel CC using CH_2Cl_2 : MeOH (200 : 1 to 50 : 1, *V/V*), and finally purified by preparative HPLC (60%−80% CH_3CN in H_2O, 25 min) to give compounds **15** (3.3 mg), **30** (31.5 mg), **31** (21.9 mg), and **36** (8.7 mg). Fr. B4G4D (638.2 mg) and B4G4E (1.1 g) were subjected to similar procedures to yield compounds **8** (11.0 mg), **23** (22.9 mg), **24** (11.3 mg), **25** (2.2 mg), **26** (16.9 mg), **32** (75.1 mg) and **2** (2.4 mg), **13** (7.1 mg), **43** (1.3 mg), **52** (55.5 mg), **53** (15.1 mg).

Sophotonin A (**1**). Yellow powder; $[\alpha]_D^{20}$ −77 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε): 203 (5.78), 283 (5.36) nm; ECD (MeOH) ($\Delta\varepsilon$) 305 (−5.66), 335 (+2.61) nm; IR (KBr) ν_{max} 3337, 2967, 2913, 1655, 1594, 1437, 1374, 1272, 1215, 1174, 1118, 1085 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 1 and 2; HR-ESI-MS *m/z* 423.2174 $[M + H]^+$ (Calcd. for $C_{26}H_{31}O_5$, 423.2171).

Sophotonin B (**2**). Brown solid; $[\alpha]_D^{20}$ −2 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε): 203 (5.53), 274 (5.32) nm; ECD (**2a**; hexane/isopropanol) ($\Delta\varepsilon$) 303 (+0.22), 333 (−0.04) nm; ECD (**2b**; hexane/isopropanol) ($\Delta\varepsilon$) 303 (−0.35), 333 (+0.06) nm; IR (KBr) ν_{max} 3371, 2975, 1624, 1444, 1410, 1383, 1305, 1167, 1117 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 1 and 2; HR-ESI-MS *m/z* 423.1795 $[M + H]^+$ (Calcd. for $C_{25}H_{27}O_6$, 423.1808).

Sophotonin C (**3**). White powder; $[\alpha]_D^{20}$ −76 (*c* 0.2, MeOH); UV (MeOH) λ_{max} (log ε): 216 (5.48), 241 (5.13), 287 (4.91) nm; ECD (MeOH) ($\Delta\varepsilon$) 297 (−10.27), 330 (+6.40) nm; IR (KBr) ν_{max} 3271, 2980, 2915, 1661, 1588, 1444, 1286, 1169, 1135, 1106, 1046 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 1 and 2; HR-ESI-MS *m/z* 475.2480 $[M + H]^+$ (Calcd. for $C_{30}H_{35}O_5$, 475.2484).

Sophotonin D (**4**). Yellow gum; $[\alpha]_D^{20}$ −90 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε): 218 (5.58), 236 (5.46), 277 (5.19) nm; ECD (MeOH) ($\Delta\varepsilon$) 288 (−8.29), 320 (+4.99) nm; IR (KBr) ν_{max} 3282, 2974, 2925, 1683, 1589, 1443, 1375, 1337, 1286, 1208, 1150, 1109, 1070, 1045 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 1 and 2; HR-ESI-MS *m/z* 419.1859 $[M + H]^+$ (Calcd. for $C_{26}H_{27}O_5$, 419.1858).

Sophotonin E (**5**). White powder; $[\alpha]_D^{20}$ −80 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε): 202 (5.56), 281 (5.41) nm; ECD (MeOH) ($\Delta\varepsilon$) 297 (−7.81), 329 (+4.10) nm; IR (KBr) ν_{max} 3269, 2973, 2925, 1660, 1588, 1442, 1378, 1336, 1282, 1151, 1107, 1045 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 1 and 2; HR-ESI-MS *m/z* 457.2372 $[M + H]^+$ (Calcd. for $C_{30}H_{33}O_4$, 457.2379).

Sophotonin F (**6**). White powder; $[\alpha]_D^{20}$ −69 (*c* 0.2, MeOH); UV (MeOH) λ_{max} (log ε): 228 (5.45), 281 (4.92) nm; ECD (MeOH) ($\Delta\varepsilon$) 297 (−9.81), 330 (+4.33) nm; IR (KBr) ν_{max} 3272, 2969, 2923, 1660, 1588, 1441, 1374, 1337, 1284, 1209, 1150, 1106, 1046 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 1 and 2; HR-ESI-MS *m/z* 475.2842 $[M + H]^+$ (Calcd. for $C_{30}H_{35}O_5$, 475.2484).

Sophotonin G (**7**). White powder; $[\alpha]_D^{20}$ −67 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε): 223 (5.54), 279 (5.03) nm; ECD (MeOH) ($\Delta\varepsilon$) 214 (+22.67), 230 (−10.34), 240 (+2.87), 300 (−6.48), 330 (+1.64) nm; IR (KBr) ν_{max} 3408, 2973, 2926, 1668, 1607, 1457, 1376, 1308, 1258, 1211, 1150, 1096, 1070 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 3 and 4; HR-ESI-MS *m/z* 491.2437 $[M + H]^+$ (Calcd. for $C_{30}H_{35}O_6$, 491.2434).

Sophotonin H (**8**). Yellow powder; $[\alpha]_D^{20}$ +1 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε): 206 (5.81), 251 (5.32), 363 (5.40) nm; IR (KBr) ν_{max} 3330, 2971, 2914, 1602, 1560, 1443, 1282, 1181, 1123, 1049 cm^{-1}; ^{1}H and ^{13}C NMR data

Table 1 ^{1}H NMR (500 MHz) data of compounds 1–6 (δ in ppm, *J* in Hz)

No.	1[b]	2[b]	3[a]	4[a]	5[a]	6[a]
2	5.64, dd (12.4, 3.2)	5.62, dd (13.1, 3.0)	5.50, dd (13.1, 3.0)	5.36, dd (13.2, 3.0)	5.35, dd (13.3, 2.9)	5.33, dd (13.2, 3.0)
3	2.94, dd (17.0, 12.4), *α*	3.08, dd (17.2, 13.1), *α*	3.05, dd (16.9, 13.1), *α*	2.98, dd (16.9, 13.2), *α*	3.02, dd (16.9, 13.3), *α*	2.99, dd (16.9, 13.2), *α*
	2.79, dd (17.0, 3.2), *β*	2.76, dd, 17.2, 3.0), *β*	2.86, dd (16.9, 3.0), *β*	2.79, dd (16.9, 3.0), *β*	2.81, dd (16.9, 2.9), *β*	2.80, dd (16.9, 3.0), *β*
5	7.75, d (8.8)		7.76, d (8.6)	7.75, d (8.6)	7.76, d (8.6)	7.73, d (8.6)
6	6.73, d (8.8)		6.54, d (8.6)	6.55, d (8.6)	6.54, d (8.6)	6.54, d (8.6)
2'			7.16, d (1.8)	7.29, d (2.4)	7.42, d (2.2)	6.95, d (2.2)
3'	6.35, s	6.33, d (2.3)				
5'		6.35, dd (8.3, 2.3)				
6'	7.12, s	7.26, d (8.3)	7.48, d (1.8)	7.73, d (2.4)	6.96, d (2.2)	7.09, d (2.2)
1"	3.34, d (7.3)	6.54, d (10.0)	3.42, d (7.4)	3.40, d (7.3)	3.43, d (7.3)	3.41, d (7.2)
2"	5.17, t (7.3)	5.53, d (10.0)	5.26, t (7.4)	5.24, t (7.3)	5.26, t (7.3)	5.26, t (7.2)
4"	1.62, s	1.44, s	1.74, s	1.73, d (1.3)	1.75, d (1.4)	1.76, s
5"	1.67, s	1.41, s	1.74, s	1.73, d (1.3)	1.74, d (1.4)	1.73, s
6"	3.20, d (7.3)	3.21, d (7.3)	3.63, d (7.4)	6.38, d (10.0)	6.33, d (9.8)	6.47, d (9.9)
7"	5.27, t (7.3)	5.17, t (7.3)	5.41, t (7.4)	5.77, d (10.0)	5.68, d (9.8)	5.61, d (9.9)
9"	1.70, s	1.66, d (1.3)	1.76, d (1.4)	1.53, s	1.48, s	3.70, d (11.8) 3.53, d (11.8)
10"	1.64, s	1.78, d (1.3)	1.79, d (1.4)	1.53, s	1.48, s	1.42, s
11"			6.60, s	10.46, s	6.83, d (16.4)	3.41, d (15.4) 3.18, dd (15.4, 7.2)
12"					6.93, d (16.4)	5.26, t (7.2)
14"			1.69, s		5.09, each d (16.5)	1.74, d (1.4)
15"			1.69, s		2.00, s	1.72, d (1.4)
7-OCH_3	3.90, s					

[a] Measured in $CDCl_3$; [b] Measured in CD_3OD

see Tables 3 and 4; HR-ESI-MS *m/z* 475.2480 $[M + H]^+$ (Calcd. for $C_{30}H_{35}O_5$, 475.2484).

Sophotonin I (**9**). Yellow gum; $[\alpha]_D^{20}$ −4 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log*ε*): 203 (5.53), 391 (5.30) nm; IR (KBr) ν_{max} 3378, 2912, 1617, 1542, 1471, 1371, 1310, 1257, 1148 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 3 and 4; HR-ESI-MS *m/z* 409.2010 $[M + H]^+$ (Calcd. for $C_{25}H_{29}O_5$, 409.2015).

Sophotonin J (**10**). Yellow gum; $[\alpha]_D^{20}$ −10 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log *ε*): 203 (5.60), 315 (4.94) nm; IR (KBr) ν_{max} 3381, 2921, 1615, 1549, 1441, 1373, 1244, 1091, 1040 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 3 and 4; HR-ESI-MS *m/z* 409.2013 $[M + H]^+$ (Calcd. for $C_{25}H_{29}O_5$, 409.2015).

Sophotonin K (**11**). White powder; $[\alpha]_D^{20}$ +1 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log *ε*): 201 (5.61), 292 (5.34) nm; ECD (**11a**; hexane/isopropanol) (Δ*ε*) 210 (−0.05), 280 (+0.08), 310 (−0.10) nm; ECD (**11b**; hexane/isopropanol) (Δ*ε*) 210 (+0.05), 280 (−0.05), 310 (+0.07) nm; IR (KBr) ν_{max} 3346, 1637, 1503, 1483, 1447, 1385, 1262, 1163, 1037 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 3 and 4; HR-ESI-MS *m/z* 317.0664 $[M + H]^+$ (Calcd. for $C_{16}H_{13}O_7$, 317.0661).

Sophotonin L (**12**). Brown solid; $[\alpha]_D^{20}$ −110 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log *ε*): 202 (5.90), 298 (5.15) nm; ECD (MeOH) (Δ*ε*) 216 (−14.14), 298 (−1.92) nm; IR (KBr) ν_{max} 3447, 1618, 1501, 1438, 1383, 1169, 1037 cm^{-1}; ^{1}H and ^{13}C NMR data see Tables 3 and 4; HR-ESI-MS *m/z* 585.1412 $[M + H]^+$ (Calcd. for $C_{32}H_{25}O_{11}$, 585.1397).

LC-CD analysis

HPLC was performed on a Jasco HPLC system on a Chiralcel AD-3 column (3 μm, 250 mm × 4.6 mm). The enantiomers of **2** [retention time (t_R) 6.6 min for (2*R*)-**2a**, 7.6 min for (2*S*)-**2b**] were separated with a solvent mixture of hexane : isopropanol (80 : 20, *V/V*, 0.8 mL·min^{-1}) as mobile phase, and the ECD spectra were recorded at 270 nm with a Jasco J-1500 CD spectrometer. The enantiomers of **11** [t_R 11.5 min for (3*S*)-**11a**, 17.0 min for (3*R*)-**11b**] were run with a solvent mixture of hexane : isopropanol (75 : 25, *V/V*, 0.8 mL·min^{-1}) as mobile phase, and the ECD spectra were recorded at 290 nm. The chromatogram was zeroed right after the start of recording, and hence relative absorbance was meas-

Table 2 ^{13}C NMR (126 MHz) data of compounds 1–6 (δ in ppm)

No.	1[b]	2[b]	3[a]	4[a]	5[a]	6[a]
2	76.5, CH	76.2, CH	80.2, CH	79.1, CH	79.8, CH	79.5, CH
3	44.2, CH_2	43.2, CH_2	44.8, CH_2	44.0, CH_2	44.3, CH_2	44.1, CH_2
4	195.2, C	199.1, C	191.7, C	191.1, C	191.6, C	192.0, C
5	127.2, CH	157.2, C	126.7, CH	126.7, CH	126.6, CH	126.6, CH
6	105.9, CH	103.0, C	110.7, CH	110.9, CH	110.7, CH	110.6, CH
7	164.9, C	160.8, C	161.5, C	161.6, C	161.4, C	161.7, C
8	118.9, C	110.3, C	114.7, C	114.8, C	114.7, C	115.0, C
9	162.4, C	162.1, C	160.9, C	160.7, C	160.9, C	161.1, C
10	116.3, C	103.4, C	115.2, C	115.1, C	115.2, C	115.0, C
1'	118.0, C	117.9, C	134.2, C	131.6, C	131.2, C	131.5, C
2'	154.2, C	156.9, C	122.1, CH	129.8, CH	124.1, CH	122.6, CH
3'	103.3, CH	103.5, CH	125.6, C	122.9, C	121.8, C	120.9, C
4'	56.7, C	159.8, C	153.3, C	156.5, C	150.6, C	150.3, C
5'	120.4, C	107.8, CH	128.4, C	124.1, C	125.8, C	129.4, C
6'	128.4, CH	128.7, CH	116.4, CH	125.0, CH	123.4, CH	128.0, CH
1"	23.0, CH_2	117.0, CH	22.4, CH_2	22.4, CH_2	22.5, CH_2	22.4, CH_2
2"	123.2, CH	127.2, CH	121.2, CH	121.2, CH	121.2, CH	121.4, CH
3"	132.5, C	79.1, C	135.3, C	135.4, C	135.5, C	134.7, C
4"	26.0, CH_3	28.7, CH_3	25.9, CH_3	25.9, CH_3	26.0, CH_3	26.0, CH_3
5"	18.0, CH_3	28.4, CH_3	18.0, CH_3	18.0, CH_3	18.1, CH_3	18.0, CH_3
6"	28.6, CH_2	21.7, CH_2	28.4, CH_2	121.4, CH	122.4, CH	125.2, CH
7"	124.5, CH	123.8, CH	121.4, CH	132.1, CH	131.4, CH	127.3, CH
8"	132.3, C	131.7, C	133.7, C	78.3, C	77.0, C	79.7, C
9"	26.0, CH_3	25.9, CH_3	26.0, CH_3	28.5, CH_3	28.3, CH_3	68.3, CH_2
10"	17.8, CH_3	18.0, CH_3	18.1, CH_3	28.5, CH_3	28.3, CH_3	22.9, CH_3
11"			100.8, CH	189.1, CH	122.8, CH	29.1, CH_2
12"			163.9, C		132.7, CH	122.8, CH
13"			69.6, C		142.7, C	133.2, C
14"			28.9, CH_3		117.3, CH_2	25.8, CH_3
15"			28.9, CH_3		18.7, CH_3	18.0, CH_3
7-OCH_3	56.4, CH_3					

[a] Measured in $CDCl_3$; [b] Measured in CD_3OD

ured. The on-line CD spectra were recorded simultaneously at the maximal UV absorption.

Computational section

The DFT NMR and TDDFT ECD calculations for compounds **7** and **12** were performed with the Gaussian 16 program (Revision A.03, Gaussian Inc., Wallingford CT, 2016) [22]. Conformational searching was conducted by the Conflex 8.0 software (CONFLEX Corporation, Tokyo, Japan, 2017) using the MMFF force field within an energy window of 5.0 kcal·mol^{-1} [23]. The conformers with the Boltzmann population above 0.1% were re-optimized at the level of M06-2X/6-31G(d) in vacuo or with the SMD solvent model for methanol. TDDFT ECD calculations were run at the M06-2X/def2-TZVP level with the SMD solvent model for methanol. Calculated ECD spectra were generated using the SpecDis Version 1.71 with Gaussian broadening after UV correction [24, 25]. DFT NMR calculations were run at the mPW1PW91/6-311G(d, p) level with the PCM solvent model for chloroform or the SMD solvent model for methanol. The possible configuration was specified by DP4+ probability [26, 27].

Table 3 ^{1}H NMR data of compounds 7–12 (δ in ppm, *J* in Hz)

No.	7[a, e]	8[c, e]	9[b, e]	10[a, e]	11[b, d]	12[b, f]
α			7.98, d (15.4)	8.07, d (15.3)		
β			7.73, d (15.4)	7.61, d (15.3)		
2	5.37, dd (13.3, 2.9)				4.52, dd (10.9, 10.9) 4.39, dd (10.9, 5.5)	4.14, ddd (10.1, 3.3, 1.9) 3.87, dd (10.1, 10.1)
3	3.09, dd (16.9, 13.3), *α* 2.78, dd (16.9, 2.9), *β*		6.37, d (2.4)	6.34, s	4.24, dd (10.9, 5.5)	3.42, dddd (11.0, 10.1, 5.4, 3.3)
4						2.78, dd (15.9, 11.0) 2.66, ddd (15.9, 5.4, 1.9)
5	7.90, d (8.6)	7.87, d (8.6)	6.35, dd (8.6, 2.4)			6.55, s
6	6.57, d (8.6)	7.04, d (8.6)	7.41, d (8.6)	7.28, s	5.90, d (2.2)	
8					5.88, d (2.2)	6.37, s
2'	6.92, d (2.3)	8.01, s				
3'					6.41, s	6.24, s
5'				6.40, d (8.9)		
6'	7.04, d (2.3)	8.01, s	7.57, s	7.72, d (8.9)	6.55, s	6.54, s
1"	5.52, d (4.7)	3.71, d (7.2)	3.37, d (7.1)	3.30, d (7.2)		
2"	4.43, d (4.7)	5.40, t (7.2)	5.20, t (7.1)	5.29, t (7.2)		4.17, dd (10.7, 3.3) 4.02, dd (10.7, 7.1)
3"						3.62–3.56, m
4"	1.37, s	1.83, s	1.67, d (1.4)	1.78, s		4.45, d (7.3)
5"	1.27, s	1.66, s	1.79, d (1.4)	1.78, s		6.74, s
6"	6.30, d (9.8)	3.46, d (7.3)	3.28, d (7.4)	3.47, d (7.2)		
7"	5.64, d (9.8)	5.40, t (7.3)	5.35, t (7.4)	5.29, t (7.2)		
8"						6.32, s
9"	1.44, s	1.76, s	1.82, d (1.3)	1.76, s		
10"	1.44, s	1.75, s	1.75, d (1.3)	1.84, s		
11"	3.29, d (7.4)	3.46, d (7.3)				
12"	5.28, t (7.4)	5.40, t (7.3)				
14"	1.73, s	1.76, s				
15"	1.74, s	1.75, s				
3"'						6.28, s
6"'						6.22, s
2'-OH				13.96, s		
4'-OCH_2O					5.84, each d (1.2)	
6-OCH_2O						5.79, each d (1.2)
6"-OCH_2O						5.76, each d (1.3)

[a] Measured in $CDCl_3$; [b] Measured in CD_3OD; [c] Measured in acetone-d_6; [d] Recorded at 400 MHz; [e] Recorded at 500 MHz; [f] Recorded at 600 MHz

The inhibition assay of proteases

SARS-CoV-2 3CLpro and SARS-CoV-2 PLpro expression and purification were performed as previously described [8]. The fractions or compounds against SARS-CoV-2 3CLpro were measured by the fluorescence resonance energy transfer (FRET) protease assay as previously reported [8]. In briefly, the fluorogenic substrate Dacyl-KTSAVLQSGFRKME-Edans was obtained from GenScript (Nanjing, China). SARS-CoV-2 3CLpro (50 nmol·L^{-1} final concentration) was mixed with different fractions or compounds in 80 μL assay buffer (50 mmol·L^{-1} Tris-HCl, pH 7.3, 1 mmol·L^{-1} EDTA) and incubated for 10 min. The reaction

Table 4 ^{13}C NMR (126 MHz) data of compounds 7–12 (δ in ppm)

No.	7[a]	8[c]	9[b]	10[a]	11[b]	12[b]
α			142.2, CH	140.5, CH		
β			118.5, CH	118.5, CH		
β'			194.6, C	193.2, C		
1			115.7, C	115.4, C		
2	80.5, CH	145.6, C	161.0, C	156.1, C	71.4, CH_2	71.0, CH_2
3	44.2, CH_2	137.9, C	103.7, CH	104.1, CH	48.1, CH	33.4, CH
4	190.5, C	173.4, C	162.7, C	158.1, C	199.0, C	31.4, CH_2
5	131.3, CH	124.4, CH	109.0, CH	120.3, C	165.9, C	107.8, CH
6	105.2, CH	114.8, CH	133.2, CH	131.6, CH	97.1, CH	142.2, C
7	166.9, C	160.2, C			168.3, C	147.7, C
8	115.9, C	116.3, C			96.0, CH	98.7, CH
9	159.9, C	155.7, C			165.2, C	150.7, C
10	115.7, C	115.4, C			103.8, C	120.9, C
1'	130.1, C	124.4, C	114.6, C	114.4, C	114.3, C	125.2, C
2'	122.5, CH	127.8, CH	163.3, C	163.9, C	151.6, C	154.6, C
3'	121.3, C	128.9, C	116.7, C	114.2, C	98.9, CH	103.5, CH
4'	151.3, C	154.9, C	160.9, C	161.6, C	149.0, C	155.1, C
5'	129.8, C	128.9, C	120.7, C	107.9, CH	142.1, C	114.7, C
6'	127.7, CH	127.8, CH	129.1, CH	129.6, CH	110.3, CH	131.6, CH
1"	71.2, CH	22.9, CH_2	22.8, CH_2	29.2, CH_2		
2"	97.9, CH	122.8, CH	123.5, CH	121.8, CH		69.0, CH_2
3"	71.4, C	132.7, C	132.4, C	135.5, C		39.1, CH
4"	26.1, CH_3	18.2, CH_3	26.0, CH_3	26.0, CH_3		
5"	24.4, CH_3	25.9, CH_3	18.0, CH_3	18.1, CH_3		108.7, CH
6"	122.3, CH	29.4, CH_2	29.0, CH_2	21.9, CH_2		142.0, C
7"	131.3, CH	123.0, CH	123.2, CH	121.4, CH		147.6, C
8"	76.7, C	133.8, C	134.5, C	135.9, C		98.4, CH
9"	28.3, CH_3	25.9, CH_3	26.0, CH_3	26.0, CH_3		150.8, C
10"	28.3, CH_3	17.9, CH_3	17.9, CH_3	18.1, CH_3		120.8, C
11"	28.4, CH_2	29.4, CH_2				
12"	122.3, CH	123.0, CH				
13"	132.7, C	133.8, C				
14"	18.0, CH_3	25.9, CH_3				
15"	25.9, CH_3	17.9, CH_3				
1'"						117.1, C
2'"						149.2, C
3'"						104.0, CH
4'"						145.7, C
5'"						140.2, C
6'"						117.5, CH
6-OCH_2O						102.0, CH_2
4'-OCH_2O					102.3, CH_2	
6"-OCH_2O						101.9, CH_2

[a] Measured in $CDCl_3$; [b] Measured in CD_3OD; [c] Measured in acetone-d_6

was triggered by adding the fluorogenic substrate (40 μL) at a final concentration of 5 μmol·L^{-1}. Then, the fluorescence absorbance was measured at 340 nm (excitation)/490 nm (emission) every 35 s for 3.5 min using a Bio-Tek SynergyH1plate reader. The inhibitory effect of the fractions or compounds on SARS-CoV-2 PLpro was measured according to previous reports [28] with a fluorogenic peptide (RLRGG-AMC) synthesized by GenScript (Nanjing, China). The reactions were performed in a total volume of 120 μL. First, 50 nmol·L^{-1} PLpro was incubated with the indicated concentrations of the tested fractions or compounds with the presence of 50 mmol·L^{-1} HEPES, pH 7.5, 0.1 mg·mL^{-1} BSA, 5 mmol·L^{-1} DTT for 20 min. The reactions were initiated by the addition of 10 μmol·L^{-1} fluorogenic peptide. Then, the fluorescence absorbance was immediately measured at 360 nm (excitation)/460 nm (emission) every 1 min for 5 min using a Bio-Tek SynergyH1plate reader. Compounds PF-07321332 [29] and GRL0617 [30] were used as positive controls for SARS-CoV-2 3CLpro and PLpro, respectively.

The initial velocities of reactions with fractions or compounds at various concentrations compared with the addition of DMSO were calculated.

Results and Discussion

*Structural elucidation of new compounds **1**−**12***

Sophotonin A (**1**) was obtained as yellow powder. Its molecular formula was assigned as $C_{26}H_{30}O_5$ with 12 indices of hydrogen deficiency on the basis of its HR-ESI-MS analysis at *m/z* 423.2174 $[M + H]^+$ (Calcd. for $C_{26}H_{31}O_5$, 423.2171). The UV spectrum exhibited maximum absorptions at 230 nm and 283 nm, which implied a flavanone skeleton [31]. The IR spectrum exhibited absorption bands for hydroxy (3337 cm^{-1}), conjugated carbonyl (1655 cm^{-1}), and aromatic (1500, 1437 cm^{-1}) functionalities.

The ^{1}H NMR spectrum (Table 1) displayed characteristic signals for a flavanone moiety at δ_H 5.64 (1H, dd, *J* = 12.4, 3.2 Hz), 2.94 (1H, dd, *J* = 17.0, 12.4 Hz), and 2.79 (1H, dd, *J* = 17.0, 3.2 Hz), two chemically inequivalent isoprenyl groups at δ_H 3.34 (2H, d, *J* = 7.3 Hz), 5.17 (1H, t, *J* = 7.3 Hz), 1.62 (3H, s), 1.67 (3H, s) and 3.20 (2H, d, *J* = 7.3 Hz), 5.27 (1H, t, *J* = 7.3 Hz), 1.70 (3H, s), 1.64 (3H, s), a methoxy group at δ_H 3.90 (3H, s), a set of *ortho*-coupled aromatic signals at δ_H 7.75 (1H, d, *J* = 8.8 Hz), 6.73 (1H, d, *J* = 8.8 Hz), and a pair of *para*-coupled ones at δ_H 6.35 (1H, s) and 7.12 (1H, s). The ^{13}C NMR (Table 2) and DEPT spectra indicated that 26 carbon signals were ascribed to one carbonyl, 10 sp^2 quaternary carbons, six sp^2 methines, one sp^3 methine, three sp^3 methylenes, one methoxy, and four methyls. The ^{1}H and ^{13}C NMR spectroscopic data indicated that **1** is a prenylated flavanone structurally related to shandougenine D [32], also isolated from *S. tonkinensis*, except that one additional methoxy group was observed for **1**.

The structure of **1** was determined by comprehensive analysis of ^{1}H–^{1}H COSY, HSQC, and HMBC correlations (Fig. 3). Ring A was established from the HMBC correlations from H-5 to C-4/C-7/C-9, and from H-6 to C-8/C-10. The methoxy group was located at C-7 in view of the HMBC correlation from the methoxy (δ_H 3.90) to C-7. One isoprenyl group was designated at C-8 by the HMBC correlations from H_2-1" to C-7/C-8/C-9. Ring B was confirmed from the HMBC correlations from H-3' to C-2'/C-4', and from H-6' to C-2. The other isoprenyl group was assigned to C-5' by the HMBC correlations from H_2-6" to C-5'/C-6'.

The ECD spectrum of **1** gave a positive Cotton effect at 335 nm and a negative one at 305 nm, indicating the absolute configuration of C-2 being *S*. Typical 2*S*-flavanones are well known to show a positive CE at the n→π* transition (320−330 nm) and a negative CE at the π→π* transition (270−290 nm), respectively [33]. Therefore, the full structure of compound **1** was proposed, and named sophotonin A.

Compound **2** was isolated as a brown solid. Its molecular formula was determined to be $C_{25}H_{26}O_6$ with 13 indices of hydrogen deficiency. Its UV and IR data were very similar to those of **1**. The ^{1}H NMR spectrum (Table 1) showed signals for a flavanone skeleton at δ_H 5.62 (1H, dd, *J* = 13.1, 3.0 Hz), 3.08 (1H, dd, *J* = 17.2, 13.1 Hz), and 2.76 (1H, dd, *J* = 17.2, 3.0 Hz), an isoprenyl group at δ_H 3.21 (2H, d, *J* = 7.3 Hz), 5.16 (1H, t, *J* = 7.3 Hz), 1.78 (3H, d, *J* = 1.3 Hz), and 1.66 (3H, d, *J* = 1.3 Hz), a set of 1,2,4-trisubstituted phenyl signals at δ_H 7.26 (1H, d, *J* = 8.3 Hz), 6.35 (1H, dd, *J* = 8.3, 2.3 Hz), and 6.33 (1H, d, *J* = 2.3 Hz), and a *gem*-dimethylpyran moiety at δ_H 6.54 (1H, d, *J* = 10.0 Hz), 5.53 (1H, d, *J* = 10.0 Hz), 1.44 (3H, s), and 1.41 (3H, s). The ^{13}C NMR (Table 2) and DEPT spectra illustrated 25 resonances including one carbonyl, 10 sp^2 quaternary, one sp^3 quaternary, six sp^2 methine, one sp^3 methine, two sp^3 methylene, and four methyl carbons. The ^{1}H and ^{13}C NMR spectra of **2** were similar to those of kushenol E [34], isolated from *Sophora fla*vescens, except for a *gem*-dimethylpyran moiety rather than an isoprenyl group observed for **2**.

The structure of **2** was further confirmed by comprehensive analysis of ^{1}H–^{1}H COSY, HSQC, and HMBC correlations (Fig. 3). In the ^{1}H–^{1}H COSY spectrum, fragments of H-2/H_2-3, H-5'/H-6', H-1"/H-2" and H_2-6"/H-7" were established. Ring A was constructed from the HMBC correlations from H-1" to C-5/C-7, and from H-2" to C-3". The isoprenyl group was assigned to C-8 by the HMBC correlation from H_2-6" to C-7/C-8/C-9. Ring B was established from the HMBC correlations from H-3' to C-4', from H-5' to C-4', and from H-6' to C-2/C-2'/C-4'.

The optical rotation value of compound **2** was close to zero, suggesting that it may exist as a racemic mixture. Chiral HPLC separation of **2** afforded two anticipated peaks of **2a** and **2b** in an approximate ratio of 1 : 1, which were evidenced as a pair of enantiomers by the opposite Cotton effects at 303 nm and 333 nm (Fig. 4). By comparing their experimental ECD spectra, the absolute configurations of **2a** and **2b** were assigned to be (2*R*)- and (2*S*)- configurated, respectively [33]. Thus, **2** was determined to be a racemic mixture, and named sophotonin B.

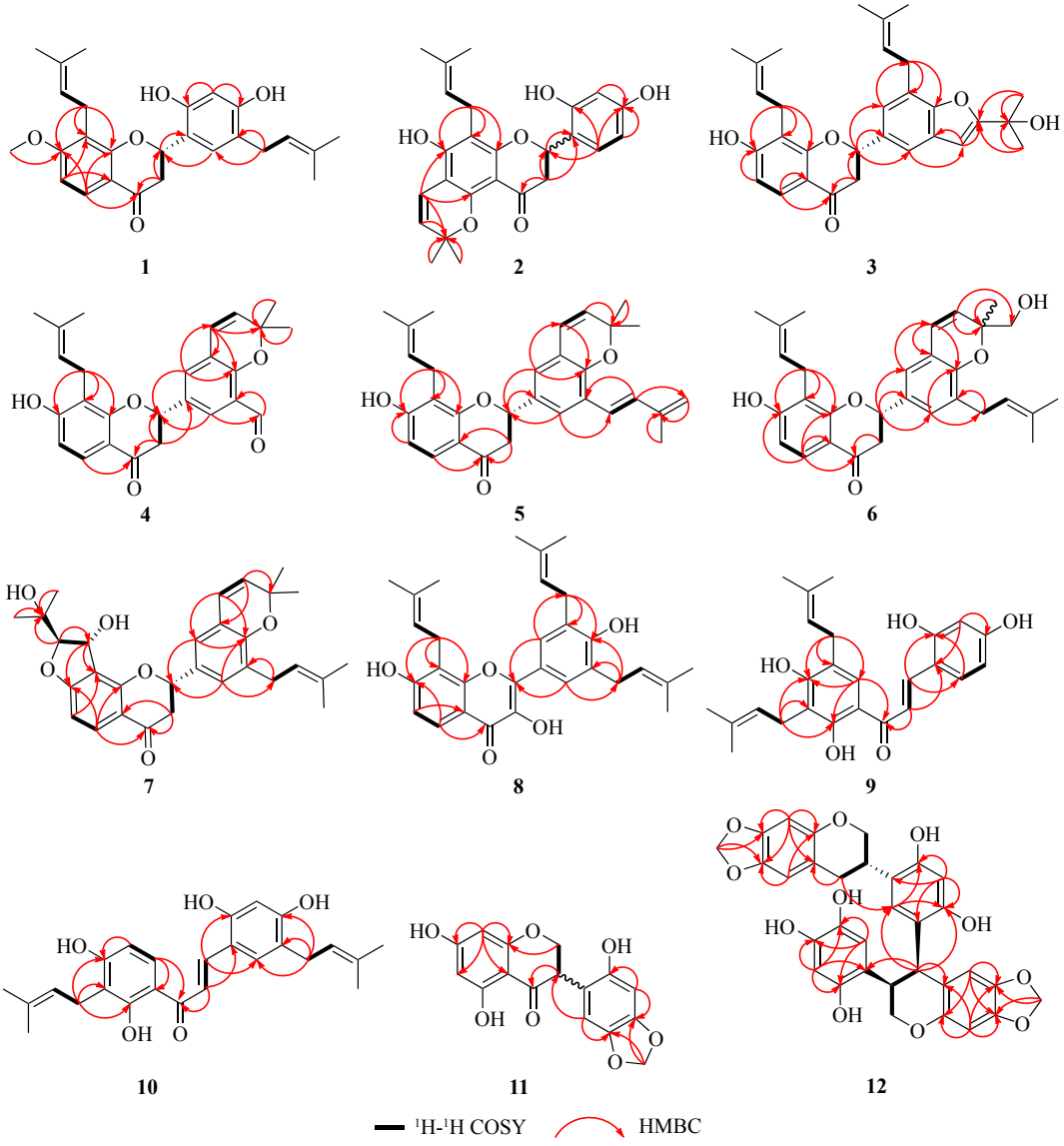

Fig. 3 Key ^{1}H–^{1}H COSY and HMBC correlations (H→C) for compounds 1–12

Compound **3**, obtained as white powder, had a molecular formula of $C_{30}H_{34}O_5$ with 14 indices of hydrogen deficiency. Its UV and IR spectra were similar to those of **2**. The ^{1}H and ^{13}C NMR data (Tables 1 and 2) indicated that **3** is a prenylated flavanone, structurally related to the known compound 2-[{2'-(1-hydroxy-1-methylethyl)-7'-(3-methyl-2-butenyl)-2',3'-dihydrobenzofuran}-5'-yl]-7-hydroxy-8-(3-methyl-2-butenyl) chroman-4-one [35], also isolated in this study. The major differences lay in the signals for a 5-(1-hydroxy-1,1-dimethylmeth-1-yl)-2,3-dihydrofuran moiety [δ_H 6.60 (1H, s), 1.69 (6H, s) and δ_C 163.9, 100.8, 69.6, 28.9, 28.9] in **3**, which was connected with C-4' inferred from the HMBC correlations from H-6' to C-11", and from H-11" to C-5'/C-12". Compound **3** was optically active, and the absolute configuration at C-2 was determined as *S* by the ECD spectrum, showing a positive Cotton effect at 330 nm and a negative one at 297 nm [33]. Therefore, the structure of compound **3** was proposed, and named sophotonin C.

Compound **4** was obtained as a yellow gum. Its molecular formula was determined as $C_{26}H_{26}O_5$ with 14 indices of hydrogen deficiency. Its UV and IR spectra were indicative of a flavanone skeleton. The ^{1}H NMR data of **4** (Table 1) revealed characteristic signals for a flavanone moiety, a isoprenyl group, two *ortho*-coupled aromatic protons at δ_H 7.75 (1H, d, J = 8.6 Hz), 6.55 (1H, d, J = 8.6 Hz), two *meta*-coupled aromatic protons at δ_H 7.73 (1H, d, J = 2.4 Hz), 7.29 (1H, d, J = 2.4 Hz), a *gem*-dimethylpyran ring at δ_H 6.38 (1H, d, J = 10.0 Hz), 5.77 (d, J = 10.0 Hz, 1H), and 1.53 (6H, s), and a formyl group at δ_H 10.46 (1H, s). According to the ^{13}C NMR (Table 2) and DEPT spectra, 26 carbon resonances were ascribed to one carbonyl, one formyl, nine sp^2 quaternary, one sp^3 quaternary carbon, seven sp^2 methine, one sp^3 methine, two sp^3 methylene, and four methyl carbons.

The planar structure of **4** was finally established by com-

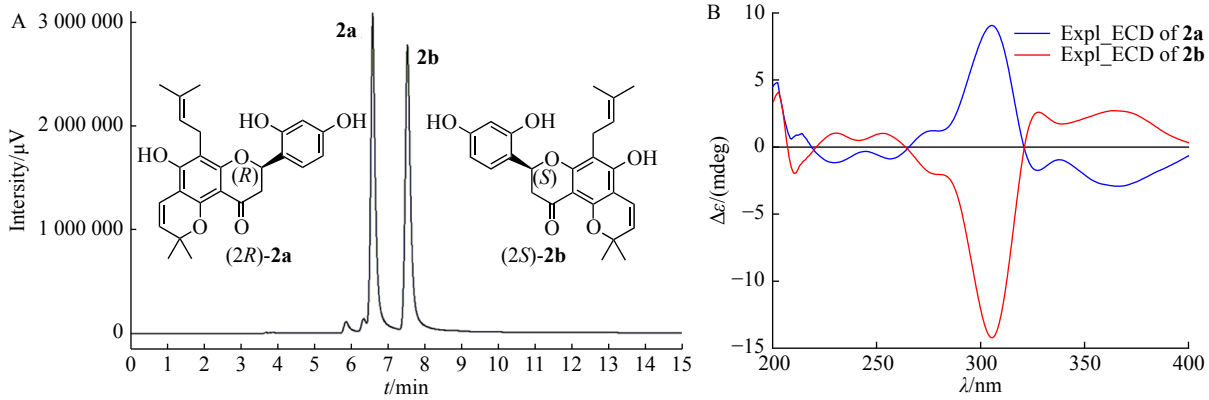

Fig. 4 A. Chiral HPLC separation of 2a and 2b in a solvent mixture of hexane : isopropanol (80 : 20, *V/V*); B. Experimental ECD spectra of 2a (blue line) and 2b (red line) recorded by on-line LC-CD coupling in a solvent mixture of hexane/isopropanol

prehensive analysis of ^{1}H–^{1}H COSY, HSQC, and HMBC correlations (Fig. 3). The isoprenyl group was attached to C-8 of ring A by the HMBC correlations from H_2-1" to C-7/C-8/C-9. The *gem*-dimethylpyran ring was specified to be fused to C-3' and C-4' of ring B deduced by the HMBC correlation from H-6" to C-3'/C-4', and from H-9" and H-10" to C-8". The position of formyl group was placed at C-5' by the HMBC correlations from H-11" to C-5'. In the same way, the absolute configuration at C-2 was determined as *S* from its ECD spectrum [33]. Therefore, the full structure of **4** was proposed, and named sophotonin D.

The molecular formula of compound **5** was established as $C_{30}H_{32}O_4$ with 15 indices of hydrogen deficiency. Its UV and IR spectra were very similar to those of **4**. Its ^{1}H and ^{13}C NMR data (Tables 1 and 2) showed high similarities to those of the known prenylated flavanone sophoranochromene [35], which was previously isolated from the same plant. The major difference lay in a *trans*-2-methyl-1,3-pentadiene moiety [δ_H 6.83 (1H, d, J = 16.4 Hz), 6.93 (1H, d, J = 16.4 Hz), 5.09 (2H, each d, J = 16.5 Hz), 2.00 (3H, s) ; δ_C 142.7, 132.7, 122.8, 117.3, 18.7] in **5**, which was supported by the ^{1}H–^{1}H COSY correlations between H-11"/H-12", and the HMBC correlations from H-6' to C-11", H-12" to C-5' and C-14", and both H_2-14" and H_3-15" to C-13" as well (Fig. 3). The 2*S*-configurtaion of **5** was determined based on the ECD spectrum [33]. Therefore, the full structure of **5** was established, and named sophotonin E.

Compound **6** had a molecular formula of $C_{30}H_{34}O_5$ determined by its HR-ESI-MS data, corresponding to 14 indices of hydrogen deficiency. Its UV and IR spectra were similar to those of **4** and **5**. The ^{1}H and ^{13}C NMR data (Tables 1 and 2) indicated that **6** shared a similar skeleton with sophoranochromene [35], except that one hydroxymethyl group [δ_H 3.70 (1H, d, J = 11.8 Hz), 3.53 (1H, d, J = 11.8 Hz); δ_C 68.3 ppm] rather than a methyl was observed in the structure of **6**. The hydroxymethyl group was further fixed at C-8" by the HMBC correlations from H-7" to C-9" (Fig. 3). The absolute configuration at C-2 was determined as *S* from its ECD spectrum, showing a positive Cotton effect at 330 nm, and a negative one at 297 nm [33]. Therefore, the structure of **6** was defined, and named sophotonin F.

Compound **7** was obtained as white powder. Its molecular formula was established as $C_{30}H_{34}O_6$ with 14 indices of hydrogen deficiency. Its UV and IR spectra were similar to those of **6**. Compared with **6**, the ^{1}H and ^{13}C NMR data (Tables 3 and 4) of **7** showed the presence of a 4-hydroxy-5-(2-hydroxyisopropyl) dihydrofuran moiety [δ_H 5.52 (1H, d, J = 4.7 Hz), 4.43 (1H, d, J = 4.7 Hz), 1.37 (3H, s), 1.27 (3H, s) and δ_C 97.9, 71.4, 71.2, 26.1, 24.4 ppm] rather than a normal perylene group that is located at C-8 in **7**. HMBC correlations from H-1" to C-7 and C-8 suggested that the dihydrofuran moiety was fused at C-7 and C-8 (Fig. 3).

The absolute configuration of C-2 was determined to be *S* from its Cotton effects in the ECD spectrum [33]. However, the stereochemistry of C-1" and C-2" remained unclear. The NOESY correlation of H-1"/H_3-5" suggested that H-1" and H-2" were in a *trans* form (Fig. 5). Thus, there are two possible configurations at C-1" and C-2" for **7**, with a certain 2*S* configuration. To figure out the relative configuration, DFT NMR calculations for two structures, *rel*-(2*S*,1"*S*,2"*S*)-**7** and *rel*-(2*S*,1"*R*,2"*R*)-**7**, were carried out. To simplify the calculations, truncated structures with the isoprenyl group replaced by an ethyl were used. Conformational searching was conducted using the Conflex in a 5.0 kcal·mol^{-1} energy window [23]. All conformers with the Boltzmann percentage above 0.1% were taken out for re-optimization at the level of M06-2X/6-31G(d) in vacuo. After removal of the duplicated conformers, the ^{1}H and ^{13}C NMR chemical shifts of the left optimized conformers were calculated at the level of mPW1PW91/6-311G(d, p) with the PCM solvent mode for chloroform. The NMR calculation method at the level of mPW1PW91/6-311G(d, p)/SMD (chloroform)//M062X/6-31G(d) was established in our previous study [36], so a specific set of statistical parameters [μ, σ, ν] was directly used for the custom DP4+ analysis in this study [26, 27]. The result gave 100% (H data, all data) and 94.28% possibilities (C data) for the (2*S*,1"*R*,2"*R*)-**7**, suggesting that both H-1" and H-2" R^*-configurated. To further determine its absolute configuration, the TDDFT ECD

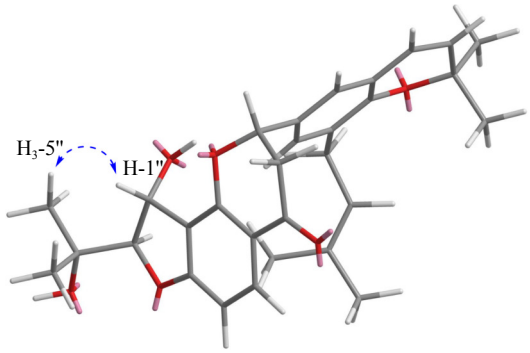

Fig. 5 Key NOESY correlations (H↔H) for compound 7

calculation was performed on the (2*S*,1"*R*,2"*R*)-**7** at the level of M06-2X/def2TZVP with the SMD solvent model for methanol. The ECD spectrum was generated by SpecDis Version 1.71 with 0.25 sigma/gamma (eV) broadening after UV correction (+15 nm) [24, 25]. The result showed that the calculated curve matched well with the experimental one (Fig. 6), both having positive Cotton effects around 215 and 240 nm, and negative ones around 230 and 300 nm. Accordingly, the full structure of **7** was constructed, and named sophotonin G.

Compound **8** was obtained as yellow powder. Its molecular formula was established as $C_{30}H_{34}O_5$ with 14 indices of hydrogen deficiency. The IR spectrum exhibited vibration bands for hydroxy (3330 cm^{-1}), conjugated carbonyl (1602 cm^{-1}), and aromatic (1560, 1500, 1443 cm^{-1}) functionalities. The maximal UV absorptions at 251 nm and 363 nm, in combination with the characteristic carbons at δ_C 145.6 (C-2), 137.9 (C-3) and 173.4 (C-4), was indicative of a flavonol skeleton [31]. The ^1H and ^{13}C NMR data (Tables 3 and 4) of **8** were compared with the known compound sophoranone [35], isolated previously from *S. tonkinensis*, which revealed that they had the same substituent pattern. Therefore, the structure of compound **8** was established, and named sophotonin H.

Compound **9**, obtained as a yellow gum, possessed a molecular formula of $C_{25}H_{28}O_5$ with 12 indices of hydrogen deficiency. The IR spectrum exhibited vibration bands for hydroxy (3378 cm^{-1}), conjugated carbonyl (1617 cm^{-1}), and aromatic (1542, 1471 cm^{-1}) functionalities. The UV spectrum exhibited the maximal absorptions at 260 nm and 391 nm, suggesting the presence of a chalcone ring [31]. The ^1H and ^{13}C NMR data (Tables 3 and 4) indicated that **9** is a prenylated chalcone, whose structure was closely related to the previously reported compound 6-[3-(2',4'-dihydroxyphenyl)acryloyl]-7-hydroxy-2,2-dimethyl-8-(3-methyl-2-butenyl)-*2H*-benzopyran [37], previously isolated from the same plant, except that an isoprenyl group [δ_H 3.28 (2H, d, J = 7.4 Hz), 5.35 (1H, t, J = 7.4 Hz), 1.82 (3H, d, J = 1.3 Hz), 1.75 (3H, d, J = 1.3 Hz); δ_C 134.5, 123.2, 29.0, 26.0, 17.9 ppm] rather than the *gem*-dimethylpyran moiety was present in **9**. Such elucidation was further supported by the ^1H-^1H COSY correlations between H_2-6"/H-7", and the HMBC correlations from H_2-6" to C-4" and C-6' (Fig. 3). Therefore, the structure of compound **9** was fully established, and named sophotonin I.

Compound **10** had a molecular formula of $C_{25}H_{28}O_5$ with 12 indices of hydrogen deficiency, which was the same with that of **9**. Its UV and IR spectra were similar to those of **9**. Their ^1H and ^{13}C NMR data (Tables 1 and 2) revealed that they had the same skeleton and substituent groups, but different substituent patterns. The HMBC correlations from H_2-1" to C-4 and C-6, and from H_2-6" to C-2' and C-4' suggested that two isoprenyl groups were linked to C-5 and C-3', respectively. Therefore, the structure of compound **10** was proposed, and named sophotonin J.

Compound **11** was obtained as white powder. Its molecular formula was established as $C_{16}H_{12}O_7$ with 11 indices of hydrogen deficiency. The UV spectrum exhibited the maximal absorptions at 201 nm, 235 nm, and 292 nm. The IR absorption bands indicated the presence of hydroxy (3346 cm^{-1}), conjugated carbonyl (1637 cm^{-1}), aromatic (1503, 1447 cm^{-1}), and methylenedioxy (-OCH_2O-) (1037 and 940 cm^{-1}) groups.

The ^1H NMR (Table 3) spectrum showed typical signals of an isoflavanone skeleton at δ_H 4.52 (1H, dd, J = 10.9, 10.9 Hz), 4.39 (1H, dd, J = 10.9, 5.5 Hz), and 4.24 (1H, dd, J = 10.9, 5.5 Hz), a set of *meta*-coupled aromatic signals at δ_H 5.90 (1H, d, J = 2.2 Hz), and 5.88 (1H, d, J = 2.2 Hz), a pair of *para*-coupled signals at δ_H 6.55 (1H, s), and 6.41 (1H, s), and a doublet signal for a methylenedioxy group at δ_H 5.84 (2H, d, J = 1.2 Hz). The ^{13}C NMR (Table 4) and DEPT data

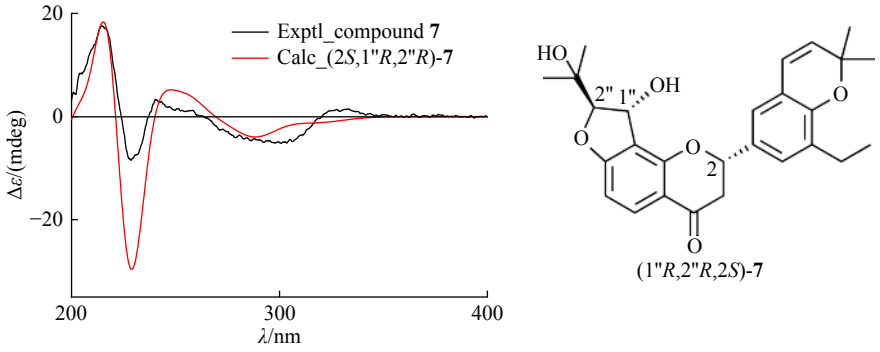

Fig. 6 Experimental ECD spectrum of compound 7 in MeOH compared with the Boltzmann-weighted M06-2X/def2TZVP SMD/MeOH ECD spectrum of the truncated structure (2*S*,1"*R*,2"*R*)-7 computed for the M06-2X/6-31G(d) optimized conformers

illustrated that 16 carbon signals were ascribed to one carbonyl, eight sp^2 quaternary, four sp^2 methine, one sp^3 methine, and two sp^3 methylene carbons. The structure of **11** was evidenced by comprehensive analysis of its ^1H–^1H COSY, HSQC, and HMBC correlations (Fig. 3). The methylenedioxy group was assigned to be fused to C-4' and C-5' by the HMBC correlations from the oxygenated methine at δ_H 5.84 to C-4'/C-5'. Therefore, the full structure of compound **11** was proposed, and named sophotonin K.

The optical rotation value of **11** was close to zero, suggesting that it may exist as a racemic mixture like compound **2**. Using the same method as described for **2**, the anticipated enantiomers of **11a** and **11b** in an approximate ratio of 1 : 1 were measured by chiral HPLC, and their enantiomeric nature was evidenced by the opposite Cotton effects at 210 nm, 280 nm and 310 nm (Fig. 7). Thus, the absolute configurations of **11a** and **11b** were assigned to be (3*S*)- and (3*R*)-configurated, respectively, by comparing their experimental ECD spectra with previous data [33].

Compound **12** was obtained as a brown solid. Its molecular formula was established as $C_{32}H_{26}O_{11}$ with 20 indices of hydrogen deficiency, based on its HR-ESI-MS analysis at *m/z* 585.1412 $[M - H]^-$ (Calcd. for $C_{32}H_{25}O_{11}$, 585.1397). The UV spectrum exhibited the maximal absorptions at 202 nm, 235 nm, and 298 nm. The IR spectrum indicated the presence of hydroxy (3347 cm^{-1}), aromatic (1618, 1501, 1438 cm^{-1}), and methylenedioxy (-OCH_2O-) (1037 cm^{-1}) groups.

The ^1H NMR data of **12** (Table 3) displayed characteristic signals for two isoflavan moiety [δ_H 4.14 (1H, ddd, J = 10.1, 3.3, 1.9 Hz), 3.87 (1H, dd, J = 10.1, 10.1 Hz), 3.42 (1H, dddd, J = 11.0, 10.1, 5.4, 3.3 Hz), 2.78 (1H, dd, J = 15.9, 11.0 Hz), 2.66 (1H, ddd, J = 15.9, 5.4, 1.9 Hz), and 4.17 (1H, dd, J = 10.7, 3.3 Hz), 4.02 (1H, dd, J = 10.7, 7.1 Hz), 3.62−3.56 (m, 1H), 4.45 (1H, d, J = 7.3 Hz)], four pairs of *para*-coupled aromatic signals [δ_H 6.55 (1H, s), 6.37 (1H, s); 6.54 (1H, s), 6.24 (1H, s); 6.74 (1H, s), 6.32 (1H, s); 6.27 (1H, s), 6.22 (1H, s)], and two methylenedioxy groups [δ_H 5.79 (2H, each d, J = 1.2 Hz), 5.76 (2H, each d, J = 1.3 Hz)]. The ^{13}C NMR (Table 4) and DEPT spectra resolved only 31 carbon signals, ascribed to 16 sp^2 quaternary carbons, eight sp^2 methines, two sp^3 methines and five sp^3 methylenes. One sp^3 methine, absent in the ^{13}C NMR spectrum, was discovered from the HMBC spectrum. All these data suggested the presence of an isoflavan dimer.

The planar structure of **12** was further evidenced by comprehensive analysis of ^1H–^1H COSY, HSQC, and HMBC correlations (Fig. 3). The two methylenedioxy groups were assigned to be fused to C-6 and C-7, and C-6" and C-7", respectively, from the HMBC correlation from one oxygenated methine at δ_H 5.79 to C-6/C-7, and from the other one at δ_H 5.76 to C-6"/C-7". The unambiguous HMBC correlations from H-4" to C-4' and C-6' revealed a unique C-5'/C-4" linkage between two isoflavan moieties.

Compound **12** was optically active, and eight possible configurations were generated from three chiral carbons C-3, C-3", and C-4". To figure out the relative configuration, DFT NMR calculations were performed on four structures, namely, *rel*-(3*R*,3"*R*,4"*R*)-**12**, *rel*-(3*R*,3"*R*,4"*S*)-**12**, *rel*-(3*R*,3"*S*,4"*R*)-**12**, and *rel*-(3*R*,3"*S*,4"*S*)-**12**. The similar method as described for compound **7** was carried out at the level of M06-2X/6-31G(d)/SMD (methanol)//mPW1PW91/6-311G (d, p)/SMD (methanol). A set of new statistical parameters [μ, σ, ν] was also calculated at the same level, which was further used for the custom DP4+ analysis [26, 27]. The result gave 99.98% (H data) and 99.96% (all data) possibilities for *rel*-(3*R*,3"*S*,4"*R*)-**12**.

The arbitrarily chosen enantiomer (3*R*,3"*S*,4"*R*)-**12** was further taken for TDDFT ECD calculation at the level M06-2X/def2TZVP with the SMD solvent model for methanol. The ECD spectrum was generated by SpecDis Version 1.71 with 0.45 sigma/gamma (eV) broadening without UV correction [24, 25]. The mirror image of the calculated curve showed a good agreement with the experimental one (Fig. 8), suggesting that compound **12** possessed a (3*S*,3"*R*,4"*S*)-configuration. Therefore, the full structure of compound **12** was proposed, and named sophotonin L.

In addition to new compounds, a total of 43 known compounds were characterized from *S. tonkinensis* in this study as

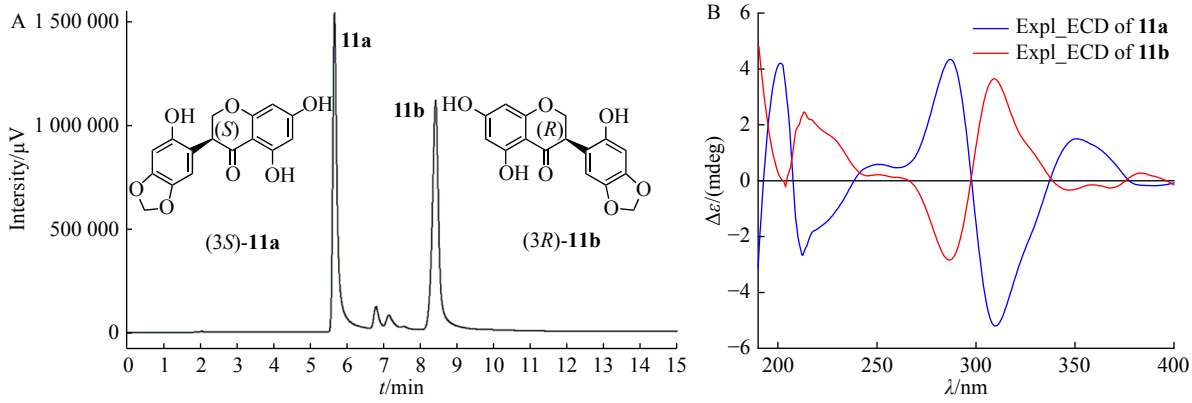

Fig. 7 A. Chiral HPLC separation of 11a and 11b in a solvent mixture of hexane : isopropanol (75 : 25, *V/V*); B. Experimental ECD spectra of 11a (blue line) and 11b (red line) recorded by on-line LC-CD coupling in a solvent mixture of hexane/isopropanol

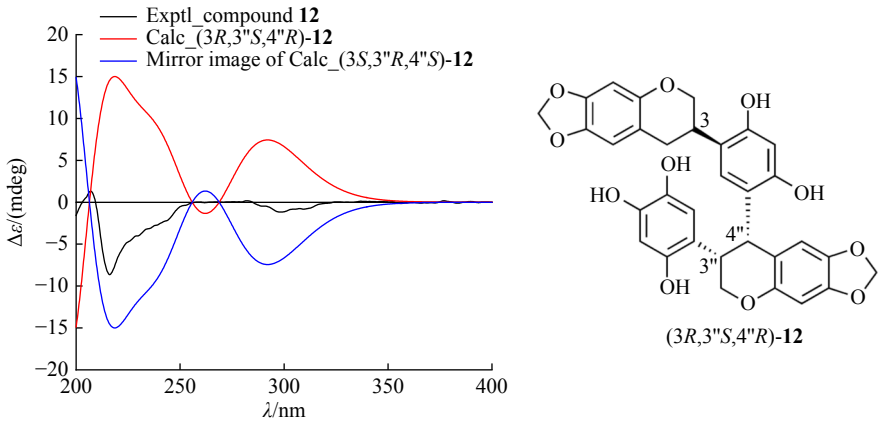

Fig. 8 Experimental ECD spectrum of compound 12 in MeOH compared with the Boltzmann-weighted M06-2X/def2TZVP SMD/MeOH ECD spectrum of the truncated structure (3*R*,3"*S*,4"*R*)-12 and its mirror image computed for the M06-2X/6-31G(d) SMD/MeOH optimized conformers

2-(2,4-dimethoxyphenyl)-5,7-dihydroxy-2,3-dihydrochromen-4-one (**13**) [38], glabrol (**14**) [39], euchrenone α_8 (**15**) [40], (2*S*)-7,4'-dihydroxy-5'-aldehyde-8,3'-(3"-methyl-but-2"-enyl)-flavanone (**16**) [41], tonkinochromane F (**17**) [35], (−)-sophoranone (**18**) [35], (2*S*)-7,2',4'-trihydroxy-8,3',5'-tri(3-methyl-but-2-enyl)-flavanone (**19**) [41], tonkinochromane G (**20**) [35], tonkinochromane I (**21**) [42], 6,8-di(γ,γ-dimethylallyl)-4',7-dihydroxyflavanone (**22**) [17], 6,8-diprenylnaringenin (**23**) [43], kushenol E (**24**) [34], lespedezaflavanone H (**25**) [44], paratocarpin L (**26**) [45], euchrenone α_5 (**27**) [45], tonkinochromane H (**28**) [35], 5-dehydroxyupinifolin (**29**) [32, 46], lupinifolin (**30**) [47], 2-(2',4'-dihydroxyphenyl)-8,8-dimethyl-10-(3-methyl-2-butenyl)-8*H*-pyrano[2,3-d]chroman-4-one (**31**) [48], flemichin D (**32**) [49], 2-[{3'-hydroxy-2',2'-dimethyl-8'-(3-methyl-2-butenyl)}chroman-6'-yl]-7-hydroxy-8-(3-methyl-2-butenyl)chroman-4-one (**33**) [50], 2-[(2'-(1-hydroxy-1-methylethyl)-7'-(3-methyl-2-butenyl)-2',3'-dihydrobenzofuran)-5'-yl]-7-hydroxy-8-(3-methyl-2-butenyl) chroman-4-one (**34**) [50], lonchocarpol (**35**) [51], exiguaflavanone B (**36**) [52], (2*R*,3*R*)-7,4'-dihydroxy-8,3',5'-triprenyldihydroflavanol (**37**) [53], lupinifolinol (**38**) [54], tonkinensisol (**39**) [55], 6,8-diprenylkaempferol (**40**) [49], genistein (**41**) [56], 3',5,7-trihydroxy-4'-methoxyisoflavone (**42**) [57], 7-*O*-methylluteone (**43**) [57], isoliquiritigenin (**44**) [58], kanzonol C (**45**) [59], sophoradin (**46**) [60], 6-[3-(2',4'-dihydroxyphenyl)acryloyl]-7-hydroxy-2,2-dimethyl-8-(3-methyl-2-butenyl)-2*H*-benzopyran (**47**) [37], sophoratonin H (**48**) [18], maackiain (**49**) [34], (−)-4-methoxymaackiain (**50**) [61], sophotokin (**51**) [41], maackiapterocarpan B (**52**) [62], maackiapterocarpan A (**53**) [62], 7,2'-dihydroxy-4',5'-methylenedioxyisoflavan (**54**) [63], and bolusanthin IV (**55**) [64].

Anti-SARS-CoV-2 activity of the isolated compounds

3CLpro and PLpro are essential for SARS-CoV-2 transcription and replication. To investigate the inhibitory activity of the isolated compounds against SARS-CoV-2 3CLpro and PLpro, all compounds were tested by the FRET-based protease assay. The results are shown in Table 5. Compounds **12** and **51** exhibited inhibitory effects against SARS-CoV-2 3CLpro with IC_{50} values of 34.89, and 19.88 μmol·L^{-1}, respectively, and compounds **9**, **43** and **47** exhibited inhibitory effects against PLpro with IC_{50} values of 32.67, 79.38, and 16.74 μmol·L^{-1}, respectively. Compounds PF-07321332 [29] and GRL0617 [30] were used as positive controls for SARS-CoV-2 3CLpro and PLpro, respectively. The determined IC_{50} value of PF-07321332 against SARS-CoV-2 3CLpro was 27.36 nmol·L^{-1}, and the measured IC_{50} value of GRL0617 against SARS-CoV-2 PLpro was 1.77 μmol·L^{-1}.

Table 5 Inhibitory activity of compounds 9, 12, 43, 47, and 51 against SARS-CoV-2 3CLpro and PLpro

Compound	IC_{50}/(μmol·L^{-1})	
	SARS-CoV-2 3CLpro	SARS-CoV-2 PLpro
9	-	32.67 ± 4.34
12	34.89 ± 4.48	-
43	-	79.38 ± 1.56
47	-	16.74 ± 1.40
51	19.88 ± 5.71	-
PF-07321332	27.36 nmol·L^{-1}	-
GRL0617	-	1.77 μmol·L^{-1}

Among the active compounds, sophotokin (**51**) is one of the major constituents in the rhizomes and roots of *S. tonkinensis*, which, to some extent, explains for the inhibitory effect against 3CLpro of the plant extract. At the same time, we did not observe the common structural features of the active compounds that could distinguish them from the inactive ones. More work on the structure-activity relationship is needed for future investigation.

Conclusion

In summary, a total of 12 undescribed flavonoids are separated from *S. tonkinensis*, along with 43 known compounds. The full structures including the absolute configurations for all new compounds are proposed by extensive analysis of spectroscopic data, DFT NMR calculations coupled

with DP4+ statistical analysis, and TDDFT ECD calculations. It is noteworthy that **2** and **11** are racemic mixtures, and their racemic nature is resolved by chiral HPLC sparation, which permits online stereochemical characterization when used in combination with circular dichroism (CD) spectroscopy (LC-CD coupling). Some compounds exhibit inhibitory effects against SARS-CoV-2 3CLpro and PLpro with IC_{50} values ranging from 16.74 to 79.38 μmol·L^{-1}, but there is not enough data supporting a structure-activity relationship. Our finding enriches the chemical diversity of *S. tonkinensis*, and provides evidence to support its usage as an anti- SARS-CoV-2 agent.

References

[1] Hasoksuz M, Kilic S, Sarac F. Coronaviruses and SARS-COV-2 [J]. *Turk J Med Sci*, 2020, **50**(SI-1): 549-556.

[2] Zumla A, Chan J F, Azhar E I, *et al*. Coronaviruses: drug discovery and therapeutic options [J]. *Nat Rev Drug Discov*, 2016, **15**(5): 327-347.

[3] Pillaiyar T, Manickam M, Namasivayam V, *et al*. An overview of severe acute respiratory syndrome-coronavirus (SARS-CoV) 3CL protease inhibitors: peptidomimetics and small molecule chemotherapy [J]. *J Med Chem*, 2016, **59**(14): 6595-6628.

[4] Dai WH, Zhang B, Jiang XM, *et al*. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease [J]. *Science*, 2020, **368**(6497): 1331-1335.

[5] Shagufta AI. An update on pharmacological relevance and chemical synthesis of natural products and derivatives with anti SARS-CoV-2 activity [J]. *ChemistrySelect*, 2021, **6**(42): 11502-11527.

[6] Liu W, Huang J, Zhang F, *et al*. Comprehensive profiling and characterization of the absorbed components and metabolites in mice serum and tissues following oral administration of Qing-Fei-Pai-Du decoction by UHPLC-Q-Exactive-Orbitrap HRMS [J]. *Chin J Nat Med*, 2021, **19**(4): 305-320.

[7] Wei WL, Wu SF, Li HJ, *et al*. Chemical profiling of Huashi Baidu prescription, an effective anti-COVID-19 TCM formula, by UPLC-Q-TOF/MS [J]. *Chin J Nat Med*, 2021, **19**(6): 473-480.

[8] Su HX, Yao S, Zhao WF, *et al*. Anti-SARS-CoV-2 activities *in vitro* of Shuanghuanglian preparations and bioactive ingredients [J]. *Acta Pharmacol Sin*, 2020, **41**(9): 1167-1177.

[9] Zhang JJ, Chen BS, Dai HQ, *et al*. Sesquiterpenes and polyphenols with glucose-uptake stimulatory and antioxidant activities from the medicinal mushroom *Sanghuangporus sanghuang* [J]. *Chin J Nat Med*, 2021, **19**(9): 693-699.

[10] Goswami D, Kumar M, Ghosh S, *et al*. Natural product compounds in *Alpinia officinarum* and ginger are potent SARS-CoV-2 papain-like protease inhibitors [Z]. ChemRxiv, 2020, doi.org/10.26434/chemrxiv.12071997.v1

[11] Gao J, Tian Z, Yang X. Breakthrough: chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies [J]. *Biosci Trends*, 2020, **14**(1): 72-73.

[12] Chrzanowski J, Chrzanowska A, Graboń W. Glycyrrhizin: an old weapon against a novel coronavirus [J]. *Phytother Res*, 2021, **35**(2): 629-636.

[13] Su HX, Yao S, Zhao WF, *et al*. Identification of pyrogallol as a warhead in design of covalent inhibitors for the SARS-CoV-2 3CL protease [J]. *Nat Commun*, 2021, **12**(1): 3623.

[14] Zhu DF, Su HX, Ke CQ, *et al*. Efficient discovery of potential inhibitors for SARS-CoV-2 3C-like protease from herbal extracts using a native MS-based affinity-selection method [J]. *J Pharm Biomed Anal*, 2022, **209**: 114538.

[15] The Chinese Pharmacopoeia Commission. *Pharmacopoeia of the People's Republic of China*, Vol. 1 [M]. Beijing: China Medical Science and Technology Press, 2020.

[16] Zhang SN, Li XZ, Tan LY, *et al*. A review of pharmacological and toxicological effects of *Sophora tonkinensis* with bioinformatics prediction [J]. *Am J Chinese Med*, 2021, **49**(02): 359-389.

[17] Zhang W, Wei X, Zhang LY, *et al*. Chemical constituents of *Sophora tonkinensis* [J]. *Chem Nat Compd*, 2020, **56**(6): 1140-1142.

[18] Ahn J, Kim YM, Chae HS, *et al*. Prenylated flavonoids from the roots and rhizomes of *Sophora tonkinensis* and their effects on the expression of inflammatory mediators and proprotein convertase subtilisin/kexin type 9 [J]. *J Nat Prod*, 2019, **82**(2): 309-317.

[19] Wu C, He LJ, Yi X, *et al*. Three new alkaloids from the roots of *Sophora tonkinensis* [J]. *J Nat Med*, 2019, **73**(3): 667-671.

[20] He LJ, Liu JS, Luo D, *et al*. Quinolizidine alkaloids from *Sophora tonkinensis* and their anti-inflammatory activities [J]. *Fitoterapia*, 2019, **139**: 104391.

[21] Pan QM, Zhang GJ, Huang RZ, *et al*. Cytisine-type alkaloids and flavonoids from the rhizomes of *Sophora tonkinensis* [J]. *J Asian Nat Prod Res*, 2016, **18**(5): 429-435.

[22] Duchemin I, Guido CA, Jacquemin D, *et al*. The Bethe-Salpeter formalism with polarisable continuum embedding: reconciling linear-response and state-specific features [J]. *Chem Sci*, 2018, **9**(19): 4430-4443.

[23] Goto H, Osawa E. Corner flapping: a simple and fast algorithm for exhaustive generation of ring conformations [J]. *J Am Chem Soc*, 1989, **111**(24): 8950-8951.

[24] Bruhn T, Schaumloffel A, Hemberger Y, *et al*. SpecDis: quantifying the comparison of calculated and experimental electronic circular dichroism spectra [J]. *Chirality*, 2013, **25**(4): 243-249.

[25] Pescitelli G, Bruhn T. Good computational practice in the assignment of absolute configurations by TDDFT calculations of ECD spectra [J]. *Chirality*, 2016, **28**(6): 466-474.

[26] Grimblat N, Zanardi MM, Sarotti AM. Beyond DP4: an improved probability for the stereochemical assignment of isomeric compounds using quantum chemical calculations of nmr shifts [J]. *J Org Chem*, 2015, **80**(24): 12526-12534.

[27] Zanardi MM, Sarotti AM. Sensitivity analysis of DP4+ with the probability distribution terms: development of a universal and customizable method [J]. *J Org Chem*, 2021, **86**(12): 8544-8548.

[28] Ghosh AK, Takayama J, Aubin Y, *et al*. Structure-based design, synthesis, and biological evaluation of a series of novel and reversible inhibitors for the severe acute respiratory syndrome−coronavirus papain-like protease [J]. *J Med Chem*, 2009, **52**(16): 5228-5240.

[29] Owen DR, Allerton CMN, Anderson AS, *et al*. An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19 [J]. *Science*, 2021, **374**(6575): 1586-1593.

[30] Fu Z, Huang B, Tang J, *et al*. The complex structure of GRL0617 and SARS-CoV-2 PLpro reveals a hot spot for antiviral drug discovery [J]. *Nat Commun*, 2021, **12**(1): 488.

[31] Xu RS. *The Chemistry of Natural Products* [M]. Beijing: Science Press, 1993.

[32] Luo GY, Yang Y, Zhou M, *et al*. Novel 2-arylbenzofuran dimers and polyisoprenylated flavanones from *Sophora tonkinensis* [J]. *Fitoterapia*, 2014, **99**: 21-27.

[33] Slade D, Ferreira D, Marais JP. Circular dichroism, a powerful tool for the assessment of absolute configuration of flavonoids [J]. *Phytochemistry*, 2005, **66**(18): 2177-2215.

[34] Li P, Chai WC, Wang ZY, *et al*. Bioactivity-guided isolation of compounds from *Sophora flavescens* with antibacterial activity against *Acinetobacter baumannii* [J]. *Nat Prod Res*, 2022, **36**(17): 4340-4348.

[35] Li XN, Lu ZQ, Chen GT, *et al*. NMR spectral assignments of isoprenylated flavanones from *Sophora tonkinensis* [J]. *Magn Reson Chem*, 2008, **46**(9): 898-902.

[36] Chen ZA, Ke CQ, Zhou S, *et al*. Ten undescribed cadinane-type sesquiterpenoids from *Eupatorium chinense* [J]. *Fitoterapia*, 2022, **156**: 105091.

[37] Kyogoku K, Hatayama K, Yokomori S, *et al*. Studies on the constituents of *Sophora* species. VI. Constituents of the root of *Sophora subprostrata* CHUN et T. CHEN. (4) [J]. *Chem Pharm Bull*, 1973, **21**(6): 1192-1197.

[38] Guan LP, Peng DX, Zhang L, *et al*. Design, synthesis, and cholinesterase inhibition assay of liquiritigenin derivatives as anti-Alzheimer's activity [J]. *Bioorganic Med Chem Lett*, 2021, **52**: 128306.

[39] Yoo H, Chae HS, Kim YM, *et al*. Flavonoids and arylbenzofurans from the rhizomes and roots of *Sophora tonkinensis* with IL-6 production inhibitory activity [J]. *Bioorganic Med Chem Lett*,

2014, **24**(24): 5644-5647.
[40] Mizuno M, Tanaka T, Matsuura N, *et al*. Two flavanones from *Euchresta horsfieldii* [J]. *Phytochemistry*, 1990, **29**(8): 2738-2740.
[41] Xia W, Luo P, Hua P, *et al*. Discovery of a new pterocarpan-type antineuroinflammatory compound from *Sophora tonkinensis* through suppression of the TLR4/NFκB/MAPK signaling pathway with PU.1 as a potential target [J]. *ACS Chem Neurosci*, 2019, **10**(1): 295-303.
[42] Li XN, Yan HX, Pang X, *et al*. Chemical constituents of flavonoids from rhizome of *Sophora tonkinensis* [J]. *China J Chin Mater Med*, 2009, **34**(3): 282-285.
[43] Meragelman KM, McKee TC, Boyd MR. Anti-HIV prenylated flavonoids from *Monotes africanus* [J]. *J Nat Prod*, 2001, **64**(4): 546-548.
[44] Sin MG, Li JR, Ding NP, *et al*. Chemical constituents of *Lespedeza davidii* [J]. *J China Pharm Univ*, 1991, **22**(3): 148-149.
[45] Lin Y, Kuang Y, Li K, *et al*. Nrf2 activators from *Glycyrrhiza inflata* and their hepatoprotective activities against CCl4-induced liver injury in mice [J]. *Bioorg Med Chem*, 2017, **25**(20): 5522-5530.
[46] Sun XG, Pang X, Liang HZ, *et al*. New prenylated flavonoid glycosides derived from *Epimedium wushanense* by β-glucosidase hydrolysis and their testosterone production-promoting effects [J]. *Chin J Nat Med*, 2022, **20**(9): 712-720.
[47] Simões MA, Pinto DC, Neves BM, *et al*. Flavonoid profile of the *Genista tridentata* L., a species used traditionally to treat inflammatory processes [J]. *Molecules*, 2020, **25**(4): 812.
[48] Yang XZ, Deng SH, Huang M, *et al*. Chemical constituents from *Sophora tonkinensis* and their glucose transporter 4 translocation activities [J]. *Bioorganic Med Chem Lett*, 2017, **27**(6): 1463-1466.
[49] Kim JY, Wang Y, Song YH, *et al*. Antioxidant activities of phenolic metabolites from *Flemingia philippinensis* Merr. et Rolfe and their application to DNA damage protection [J]. *Molecules*, 2018, **23**(4): 816.
[50] Inoue M, Tanabe H, Nakashima KI, *et al*. Rexinoids isolated from *Sophora tonkinensis* with a gene expression profile distinct from the synthetic rexinoid bexarotene [J]. *J Nat Prod*, 2014, **77**(7): 1670-1677.
[51] Tahara S, Katagiri Y, Ingham J L, *et al*. Prenylated flavonoids in the roots of yellow lupin [J]. *Phytochemistry*, 1994, **36**(5): 1261-1271.
[52] Ruangrungsi N, Iinuma M, Tanaka T, *et al*. Three flavanones with a lavandulyl group in the roots of *Sophora exigua* [J]. *Phytochemistry*, 1992, **31**(3): 999-1001.
[53] Baruah P, Barua NC, Sharma RP, *et al*. Flavonoids from *Millettia pulchra* [J]. *Phytochemistry*, 1984, **23**(2): 443-447.
[54] Chepkirui C, Ochieng PJ, Sarkar B, *et al*. Antiplasmodial and antileishmanial flavonoids from *Mundulea sericea* [J]. *Fitoterapia*, 2021, **149**: 104796.
[55] Sutthivaiyakit S, Thongnak O, Lhinhatrakool T, *et al*. Cytotoxic and antimycobacterial prenylated flavonoids from the roots of *Eriosema chinense* [J]. *J Nat Prod*, 2009, **72**(6): 1092-1096.
[56] Cheung S, Fang W, Li X, *et al*. Chemical constituents of *Dalbergia odorifera* [J]. *Chem Nat Compd*, 2021, **57**(6): 1122-1124.
[57] Ji S, Li Z, Song W, *et al*. Bioactive constituents of *Glycyrrhiza uralensis* (Licorice): discovery of the effective components of a traditional herbal medicine [J]. *J Nat Prod*, 2016, **79**(2): 281-292.
[58] Dat LD, Tu NTM, Duc NV, *et al*. Anti-inflammatory secondary metabolites from the stems of *Millettia dielsiana* Harms ex Diels [J]. *Carbohydr Res*, 2019, **484**: 107778.
[59] Li Y, Sun B, Zhai J, *et al*. Synthesis and antibacterial activity of four natural chalcones and their derivatives [J]. *Tetrahedron Lett*, 2019, **60**(43): 151165.
[60] Komatsu M, Tomimori T, Hatayama K, *et al*. Studies on the constituents of *Sophora* species. I. Constituents of *Sophora subprostrata* CHUN et T. CHEN. (1). Isolation and structure of new flavonoids, sophoradin and sophoranone [J]. *Chem Pharm Bull*, 1970, **18**(3): 602-607.
[61] Máximo P, Lourenço A, Feio SS, *et al*. Flavonoids from *Ulex* species [J]. *Z Naturforsch*, 2000, **55c**: 506-510.
[62] Li X, Wang D, Xia MY, *et al*. Cytotoxic prenylated flavonoids from the stem bark of *Maackia amurensis* [J]. *Chem Pharm Bull*, 2009, **57**(3): 302-306.
[63] Posri P, Suthiwong J, Thongsri Y, *et al*. Antifungal activity of compounds from the stems of *Dalbergia stipulacea* against *Pythium insidiosum* [J]. *Nat Prod Res*, 2021, **35**(17): 2823-2830.
[64] Erasto P, Bojase-Moleta G, Majinda RRT. Antimicrobial and antioxidant flavonoids from the root wood of *Bolusanthus speciosus* [J]. *Phytochemistry*, 2004, **65**(7): 875-880.

Cite this article as: LI Zhuo, XIE Hang, TANG Chunping, FENG Lu, KE Changqiang, XU Yechun, SU Haixia, YAO Sheng, YE Yang. Flavonoids from the roots and rhizomes of *Sophora tonkinensis* and their *in vitro* anti-SARS-CoV-2 activity [J]. *Chin J Nat Med*, 2023, **21**(1): 65-80.

SU Haixia received her B.S. degree (2016) from East China University of Science and Technology and Ph.D. degree from Shanghai Institute of Materia Medica (SIMM), Chinese Academy of Sciences (CAS). She is interested in the determination of crystal structures of drug targeted proteins, and structure-based drug design.

YAO Sheng was graduated with his Ph.D. degree of Medicinal Chemistry from Shanghai Institute of Materia Medica, Chinese Academy of Sciences in 2008. He currently serves as professor of Natural Product Research Center in Shanghai Institute of Materia Medica. His research mainly focuses on the discovery of natural products with novel structures and potential bioactivities from Chinese herbal medicines or plants. To date, he has published over 70 SCI papers with over 900 citations.

YE Yang, the full professor of Shanghai Institute of Materia Medica, CAS, was awarded the National Science Fund for Distinguished Young Scholars in 2009, and the special government allowance in 2016. He was selected as a leading talent of Shanghai in 2020. Professor YE leads a research group focusing on isolation and structural elucidation of bioactive secondary metabolites from medicinal plant resources by establishing a multi-technology platform for quick analysis and separation of natural products. He conducted systematic investigations of commonly used medicinal, and provided scientific evidence to explain their traditional therapeutic effects.