Supporting Information

Rapid Targeted Screening and Identification of Active Ingredients in Herbal Extracts through Ligand-Detected NMR and Database Matching

Tao Huang[a], Xin Chai[a], Shuangli Li[a], Biao Liu[a,c], Jianhua Zhan[a], Xiaohua Wang[a], Xiong Xiao[a,b], Qinjun Zhu[a], Caixiang Liu[a,b], Danyun Zeng[a,b], Bin Jiang[a,b,c,d], Xin Zhou[a,b,c,d], Lichun He[a,b], Zhou Gong[a,b], Maili Liu*[a,b,c,d], Xu Zhang*[a,b,c,d]

[a] Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement of Science and Technology, Chinese Academy of Sciences, Wuhan, 430071, China.

[b] University of Chinese Academy of Sciences, Beijing, 100049, China.

[c] Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, 430071, China.

[d] Optics Valley Laboratory, Wuhan, 430074, China.

* Corresponding Author

E-mail: zhangxu@wipm.ac.cn (X. Zhang); ml.liu@wipm.ac.cn (M. Liu)

Table of Contents

Experimental procedures

Isothermal Titration Calorimetry

Isothermal Titration Calorimetry (ITC) experiments were conducted using a MicroCal PEAQ-ITC instrument (Malvern Panalytical Ltd.). Two different injection protocols were used: one with 13 injtections (a single injection of 0.6 µL followed by 12 injections of 3 µL) and another with 19 injections (a single injection of 0.4 µL followed by 18 injections of 2 µL). All experiments were performed at 298 K. The ligand at a concentration of 2.0 mM was titrated into the cell containing the protein at a concentration of 200 µM. ITC measurements were conducted in a buffer consisting of 50 mM sodium phosphate and 100 mM NaCl at pH 7.4. To minimize the heat of dilution, the DMSO concentrations in both the cell and the syringe were carefully matched. The experiments were performed with a spinning speed of 750 rpm, a reference power of 10 µcal/s, and a time spacing of 150 s between injections. Data from the ITC experiments were analyzed using the PEAQ-ITC Analysis Software (Malvern Panalytical Ltd.). Each set of measurements was fitted to a single-site binding model, and the first injection was omitted from the data analyses. The dissociation constant K_D was calculated as an average of independent measurements with the reported values, including standard deviations.

R_2 Determination

The T_2 relaxation experiments were performed using the pulse program peCPMG. The spin-echo delay was set to 400 µs, and the recycle delay was adjusted to $5*T_1$ of the methyl protons. A pseudo-2D experiment was implemented with various spin-echo periods (different n values). The experimental data were processed using TopSpin 4.4.0 software. R_2 values were fitted using no-linear regression with a monoexponential equation.

An example (Sinomenine) of the data structure for an entry in our customized database:

Name: Sinomenine

CAS: 115-53-7

Solvent: CDCl3

^{1}H-^{13}C HSQC peak table:

Peak	^{1}H chemical shift	^{13}C chemical shift	CHn[a]
1	6.56	119.23	CH
2	6.68	108.94	CH
3	2.54	49.23	CH2
4	4.32	49.23	CH2
5	5.37	115.15	CH
6	3.18	56.70	CH
7	3.03	24.23	CH2
8	2.69	24.23	CH2
9	2.98	45.97	CH
10	1.92	36.05	CH2
11	1.87	36.05	CH2

12	2.06	47.14	CH2
13	2.53	47.14	CH2
14	3.76	56.05	CH3
15	3.45	54.77	CH3
16	2.42	42.81	CH3

Note: a) n stands the number of protons attached to the carbon

SMILES: OC1=C2[C@@](CC3=O)(CCN4C)[C@@](C=C3OC)([H])[C@@H]4CC2=CC=C1OC

Reference: Bao, G.-H.; Qin, G.-W.; Wang, R.; Tang, X.-C. Morphinane alkaloids with cell protective effects from *Sinomenium acutum*. *Journal of Natural Products* **2005**, *68*, 1128-1130.

Supplementary Tables

Table S1. The candidate scores of the active component in the methanol extract of *Sinomenii Caulis* evaluated by Database Matching

NO.	Name	Solvent[a]	Score[b]
1	sinomenine	CDCl3	0.75
2	(+)-1S, 2R-laudanidine-Nα-oxide	CD3OD	0.56
3	racemosidine A	CDCl3	0.38
4	(+)-1S, 2R-laudanidine-Nβ-oxide	CD3OD	0.38
5	nelumboferine	CDCl3	0.37
6	3(R)-1,7-di(3,4-dihydroxyphenyl)-3-O-beta-D-[6-(Z-3,4-dimethoxycinnamoylglucopyranosyl)]heptane	CD3OD	0.34
7	cyanidin 3-O-[6-O-(E)-caffeoyl-2-O-(6-(E)-feruloyl-2-O-beta-D-glucopyranosyl-(1-2)-beta-D-glucopyranoside)-5-O-beta-D-glucopyranoside	DMSO-d6/TFA-d (9:1)	0.34
8	3-O-[6-O-(E)-feruloyl-2-O-{6-O-(E)-p-coumaroyl-beta-D-glucopyranosyl}-beta-D-glucopyranosyl]-5-O-(6-O-malonyl-beta-D-glucopyranosyl)-pelargonidin	DMSO-d6: CF3COOD(9:1)	0.31
9	3beta-O-(E)-feruloyl-D:C-friedooleana-7,9(11)-dien-29-ol	CDCl3	0.29
10	racemosidine C	CDCl3	0.29
11	wattisine A	CDCl3	0.29
12	acutissimatriterpene C	CDCl3	0.28
13	racemosidine B	CDCl3	0.28
14	acutissimatriterpene A	CDCl3	0.26
15	malvidin 3-O-(6-O-(4-O-(4-O-(6-O-feruloyl-beta-D-glucopyranosyl)-E-p-coumaroyl)-alpha-rhamnosyl)-beta-D-glucopyranoside)-5-beta-D-glucopyranoside	CD3OD	0.24
16	stephalonine E	CDCl3	0.24
17	7-O-methylaloeresin A	acetone-d6	0.24
18	cimicifugic acid N	CD3OD at 35C	0.23
19	wattisine B	CDCl3-CD3OD	0.22
20	3alpha-E-feruloyltaraxerol	CDCl3	0.22
21	3'-hydroxy-N,Ndimethylcoclaurinium trifluoroacetate	DMSO-d6	0.21
22	sebestenoid C	CD3OD	0.19
23	cepharanthine-2'alpha-N-oxide	CD3OD	0.17
24	11-O-demethylmarchantin I	acetone-d6	0.17
25	stebisimine	CDCl3	0.17
26	3'-nor-4'-oxocepharanthine	CD3OD+CDCl3 (1:1)	0.16
27	racemosinine A	CDCl3+CD3OD	0.15
28	6-cinnamoylhernandine	DMSO-d6	0.14
29	3beta-E-feruloyltaraxerol	CDCl3	0.14
30	marchantin I	CDCl3	0.12
31	Neuroprotectin B	DMSO-d6	0.08
32	leuconoline	CDCl3	0.08
33	11-O-methylisocorniculatolide A	CDCl3	0.08
34	marchantin H	CDCl3	0.07
35	Neuroprotectin A	DMSO-d6	0.04
36	vatalbinoside A	acetone-d6	0.03

37	marchantin C	CDCl3	0.00
38	fesumtuorin H	CDCl3	0.00

Note: (a) NMR solvent in the literature; (b) the chemical shift tolerances were set to [0.175, 1.75] ppm.

Table S2. The candidate scores of the active component in the methanol extract of *Celastrus orbiculatus* evaluated by Database Matching

NO	Name	Solvent	Score
1	(-)-epicatechin	DMSO-d6	1.00
2	vitisinol	acetone-d6	0.77
3	epiafzelechin	acetone-d6 at 303K	0.63
4	epicatechin-(4beta-8)-4'-O-methylgallocatechin	CD3OD	0.60
5	6-(2-pyrrolidinone-5-yl)-(-)-epicatechin	CD3OD	0.58
6	8-(2-pyrrolidinone-5-yl)-(-)-epicatechin	CD3OD	0.58
7	epigallocatechin	acetone-d6 at 303K	0.57
8	Procyanidin B2	acetone-d6	0.50
9	ent-guibourtinidol-(4beta-6)-catechin	DMSO-d6 + D2O	0.50
10	Tupichinol B	acetone-d6	0.50
11	Anachelin-2	DMSO-d6	0.48
12	grincamycin D	CDCl3	0.46
13	8-hydroxygenkwanol A	CD3OD	0.45
14	flueggenine B	CD3OD	0.44
15	fibrosterol sulfate A	CD3OD	0.44
16	4',3''-di-O-methylapocynin-D	CD3OD	0.43
17	gelseiridone	CDCl3	0.40
18	(2S)-4',5,7-trihydroxyflavan-(4beta-8)-epiafzelechin	CD3OD	0.40
19	epigallocatechin-(4beta-8)-4'-O-methylgallocatechin	CD3OD	0.38
20	corbulain Ia	CD3OD	0.38
21	withalongolide N	C5D5N	0.38
22	(4alpha-8)-bis-4'-O-methylgallocatechin	CD3OD	0.36
23	5alpha-hydroperoxyivalin	CD3OD	0.33
24	bacilosarcin A	CDCl3	0.32
25	(6S,7S,8R)-2-(3,4-dihydroxyphenyl)-6-(4-hydroxyphenyl)-8-(2,4-dihydroxyphenyl)-2,3-trans-6,7-cis-7,8-trans-3,4,9,10-tetrahydro-2H,6H-pyrano[2,3-f]chromene-3,7,9-triol	DMSO-d6 + D2O	0.31
26	pentalinonside	CDCl3	0.31
27	tupichigenin A	C5D5N	0.31
28	caprazamycin B	DMSO-d6:C5D5N:D2O (5:5:1)	0.30
29	Bugbanoside D	C5D5N containg a few drops of D2O	0.30
30	catiguanin A	CD3OD	0.29
31	lucilianoside D	C5D5N	0.28
32	teaseedsaponin C	C5D5N	0.28
33	Deoxytrillenoside B	C5D5N	0.28
34	angiopterlactone B	CDCl3	0.27
35	3'-O-acetyl-4'-O-sulphodeglucoruscin	C5D5N	0.27

36	(3R,4S,6R)-p-menth-1-ene-3,6-diol 6-O-beta-D-glucopyranoside	C5D5N	0.27
37	Trillenogenin	C5D5N	0.27
38	11alpha-hydroxyacetylfawcettine	CD3OD	0.26
39	Trillenoside C	C5D5N	0.26
40	albopilosin J	C5D5N	0.25
41	trikamsteroside C	C5D5N	0.24
42	inuloxin D	CDCl3	0.24
43	Hydrocotyloside I	C5D5N	0.23
44	Xindongnin O	C5D5N	0.22
45	kahiricoside II	C5D5N	0.22
46	trikamsteroside D	C5D5N	0.22
47	ajugacetalsterone C	C5D5N	0.21
48	angudracanoside C	C5D5N	0.21
49	mersiloscine	CDCl3	0.21
50	teaseedsaponin L	C5D5N	0.20
51	Picrodendrin Y	C5D5N	0.14
52	amurensisin	CD3OD	0.11
53	tupisteroide B	C5D5N	0.10
54	taxane enolate	CDCl3	0.05

Note: (a) NMR solvent in the literature; (b) the chemical shift tolerances were set to [0.175, 1.75] ppm.

Table S3. The candidate scores of the major active component in the total alkaloid extract of *Stephania tetrandra* evaluated by Database Matching

NO.	Name	Solvent[a]	Score[b]
1	tetrandrine	CDCl3	1.00
2	fangchinoline	CDCl3	1.00
3	viniphenol A	acetone-d6 at 310 K	0.06

Note: (a) NMR solvent in the literature; (b) the chemical shift tolerances were set to [0.175, 1.75] ppm.

Table S4. The candidate scores of the minor active component in the total alkaloid extract of *Stephania tetrandra* evaluated by Database Matching

NO.	Name	Solvent[a]	Score[b]
1	tetrandrine	CDCl3	1.00
2	fangchinoline	CDCl3	1.00
3	(-)-cyclogalgravin	CDCl3	0.50
4	racemosinine A	CDCl3 + CD3OD	0.50
5	racemosidine B	CDCl3	0.48
6	wattisine B	CDCl3-CD3OD	0.48
7	kaerophyllin	C6D6	0.33
8	stebisimine	CDCl3	0.33
9	viniphenol A	acetone-d6 at 310 K	0.09

Note: (a) NMR solvent in the literature; (b) the chemical shift tolerances were set to [0.175, 1.75] ppm.

Table S5. K_D values of EB-3D and sinomenine binding to ChoKα1 measured by ITC and STD

Inhibitor	ITC[a]	ITC[a]	STD
	K_D (μM) in condition 1 (0% DMSO)	K_D (μM) in condition 2 (6% DMSO)	K_D (μM) in condition 2 (6% DMSO)
EB-3D	1.21±0.02	15.5±4.3	42+9
Sinomenine	0.34±0.08	34±14	30±9

Note: (a) Other inhibitors listed in table 1 of the main manuscript were not measured by ITC due to their poor water solubility.

Supplementary Figures

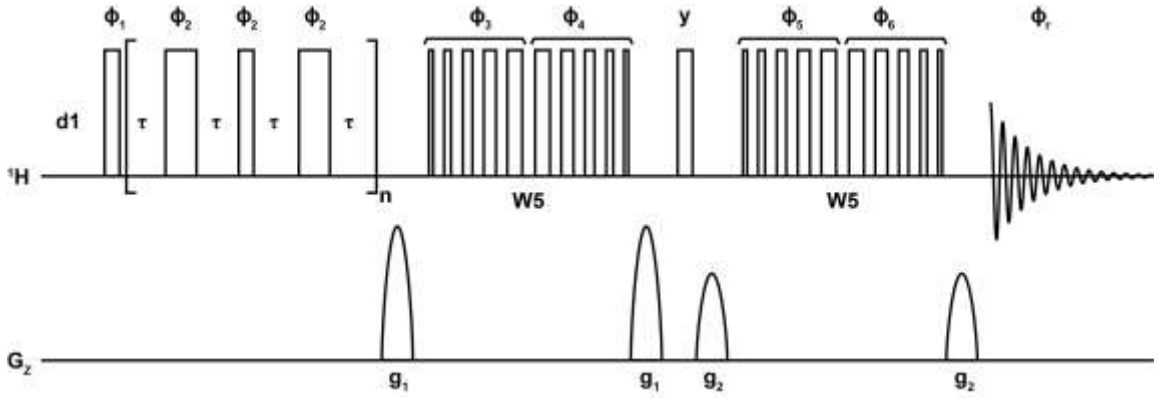

Figure S1. T_2-weighted peCPMG pulse sequence with solvent suppression and homonuclear scalar coupling artifacts elimination. Thin and thick bars represent π/2 and π pulses, respectively, semi-ellipse are gradient pulses on Z-axis, and clustered bars corresponding to W5 binomial π pulses. The phase cycles are ϕ_1= x, -x; ϕ_2= 16(y, -y), 16(-y, y); ϕ_3= 2(x), 2(y), 2(-x), 2(-y); ϕ_4= 2(-x), 2(-y), 2(x), 2(y); ϕ_5= 8(x), 8(y), 8(-x), 8(-y); ϕ_6= 8(-x), 8(-y), 8(x), 8(y); ϕ_r= 2(x, -x, -x, x), 2(-x, x, x, -x). Gradients g1 and g2 used for solvent dephasing are in different amplitudes.

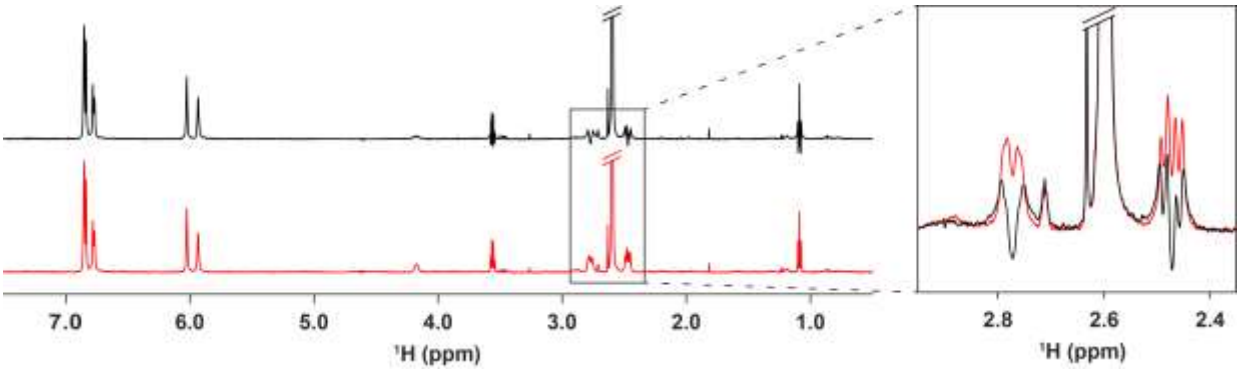

Figure S2. CPMG spectra of Chokα1-catechin-ethanol using the sequences with conventional echo (black) or perfect echo (red), as illustrated in figure S1, both in CPMG pulse train and in watergate-W5 module. The spin-echo delay τ and the total echo time were set to 1.6 ms and 25.6 ms (n = 8), respectively.

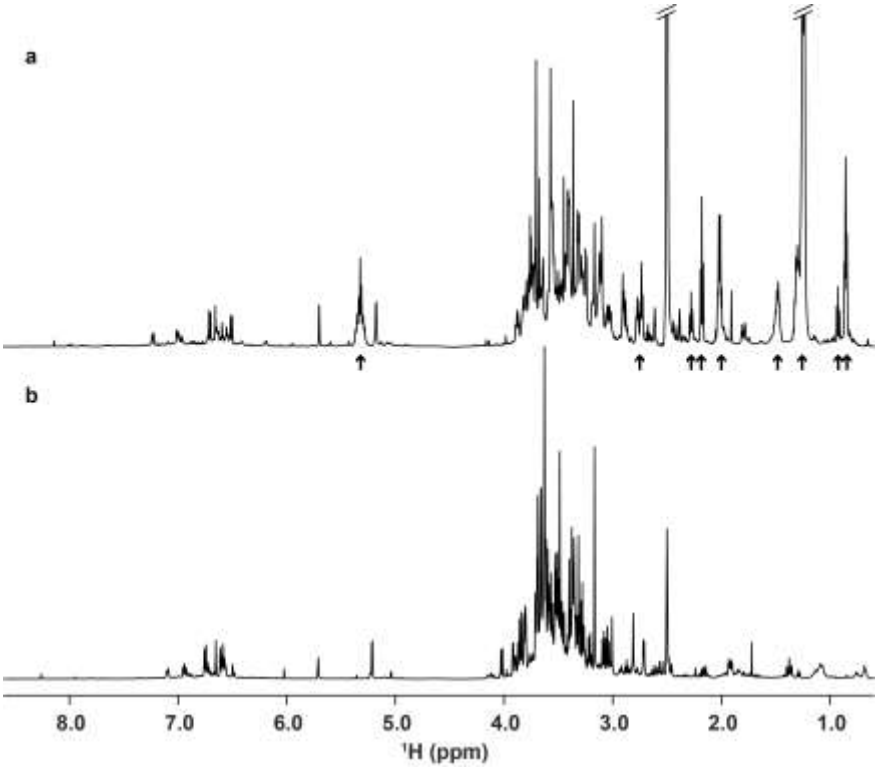

Figure S3. Comparison ^{1}H peCPMG spectra of *Sinomenii Caulis* extract dissolved in different solvents: (a) DMSO-d6; (b) phosphate buffer. The signals indicated by the arrows are originate lipids. For purposes of comparison, the pulse sequences and parameters used in in the two spectra are exactly the same.

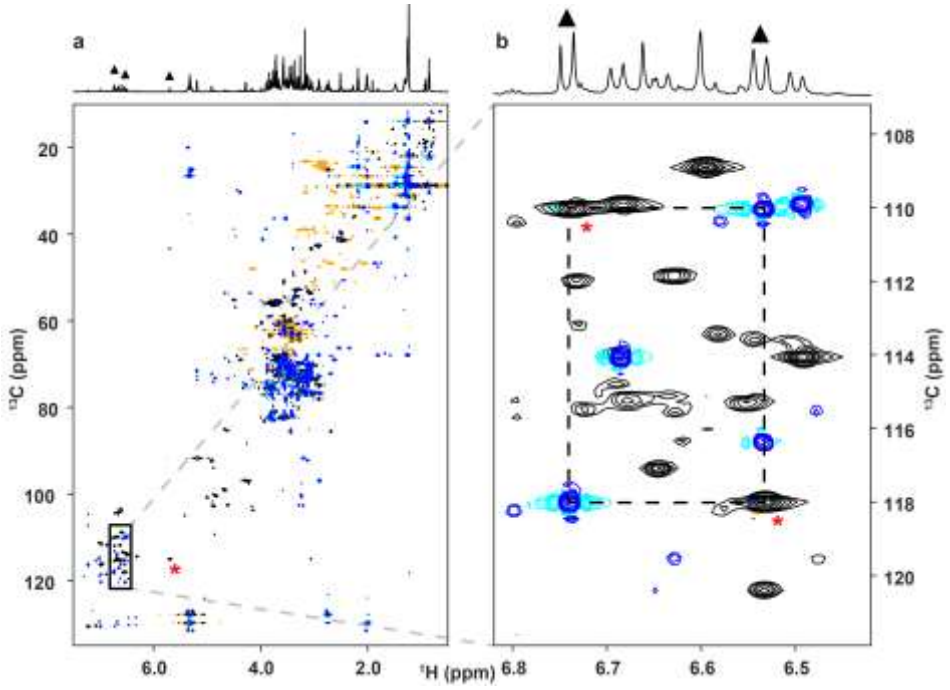

Figure S4. Overlay of multiplicity-edited 2D ^{1}H-^{13}C psHSQC and H2BC spectra of the methanol extract of *Sinomenii Caulis*, with the above corresponding 1D ^{1}H peCPMG spectrum. An expanded view (b) of the delineated section in (a) is presented for detailed examination. The edited psHSQC is depicted with black-orange contours, and the H2BC spectrum with blue-cyan contours. Black triangles (▲) mark the characteristic peaks of the active ingredient identified in the 1D ^{1}H peCPMG screening, while peaks labeled with a red asterisk (*) are distinctive C-H signals of the active ingredient within the edited psHSQC spectrum.

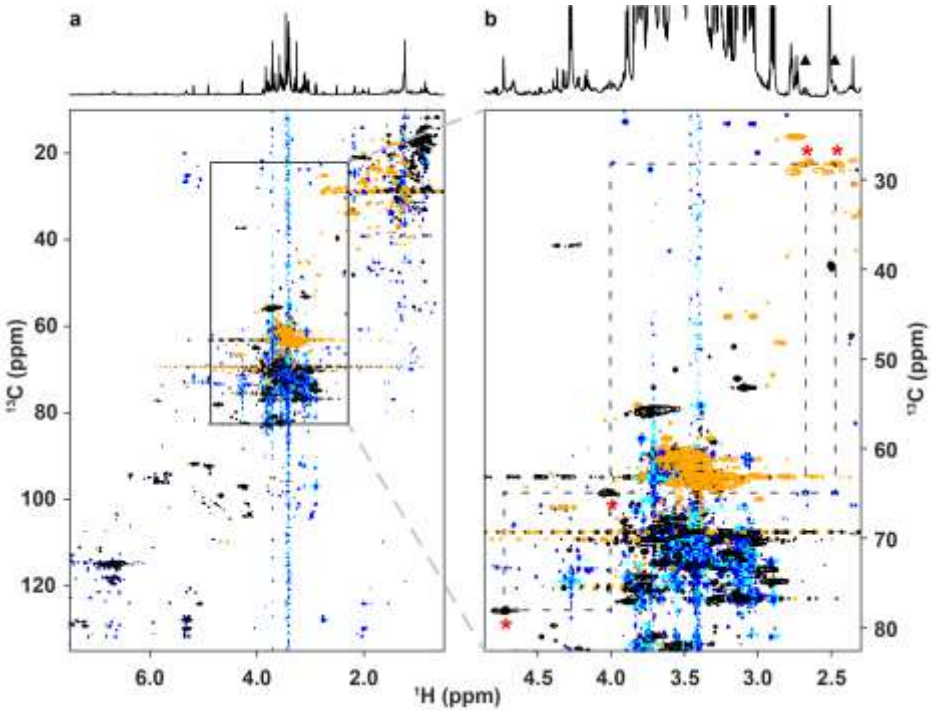

Figure S5. Overlay of multiplicity-edited psHSQC and H2BC 2D ^{1}H-^{13}C spectra of the methanol extract of *Celastrus orbiculatus*, displayed as a contour plot underneath the corresponding perfect-CPMG ^{1}H spectrum. An expanded view (b) of the region indicated by the black rectangle in (a) is provided on the right for clarity. In this overlay plot, the multiplicity-edited psHSQC spectrum is represented in black-orange contours, while the H2BC spectrum is shown in blue-cyan. The black triangles (▲) indicated the characteristic peaks of the active ingredient that was found by 1D ^{1}H pe-CPMG screening. Peaks labeled with a red asterisk (*) are characteristic C-H signals of the active ingredient in the edited psHSQC spectrum.

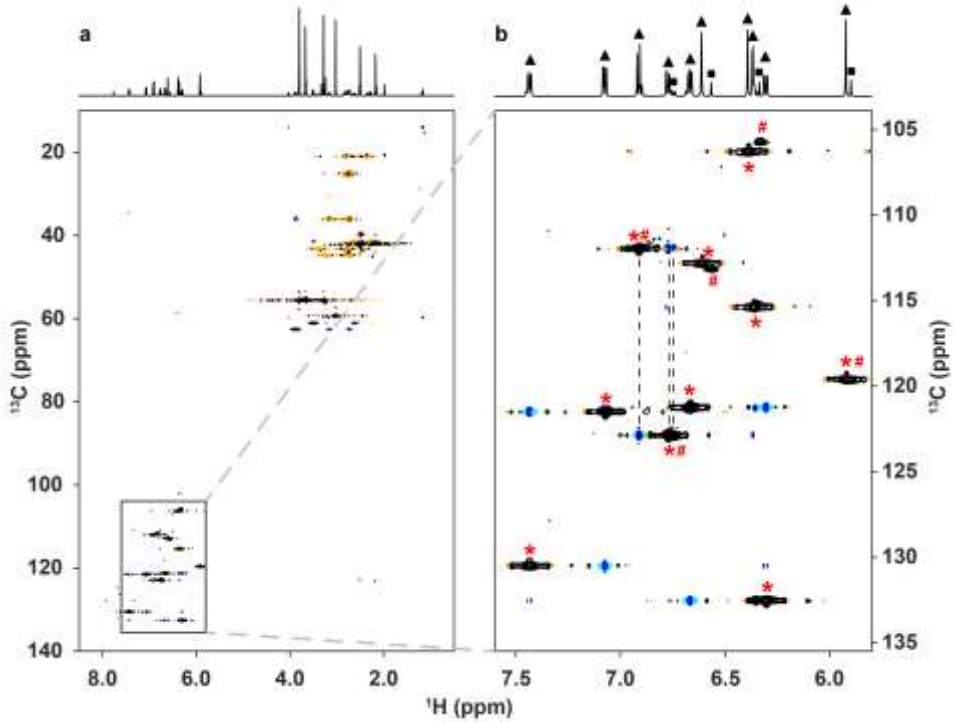

Figure S6. Overlay of multiplicity-edited psHSQC and H2BC 2D ^{1}H-^{13}C spectra of the total alkaloid extract of *Stephania tetrandra*, displayed as a contour plot underneath the corresponding perfect-CPMG ^{1}H spectrum. An expanded view (b) of the region indicated by the black rectangle in (a) is provided on the right for clarity. In this overlay plot, the multiplicity-edited psHSQC spectrum is represented in black-red contours, while the H2BC spectrum is shown in blue-cyan. The black triangles (▲) and squares (■) indicated the characteristic peaks of the major and minor active ingredients, respectively, that were found by 1D ^{1}H pe-CPMG screening. Peaks labeled with a red asterisk (*) or a red pound (#) sign are corresponding characteristic C-H signals of the active ingredients in the edited psHSQC spectrum.

Figure S7. Chemical structures of the inhibitors.

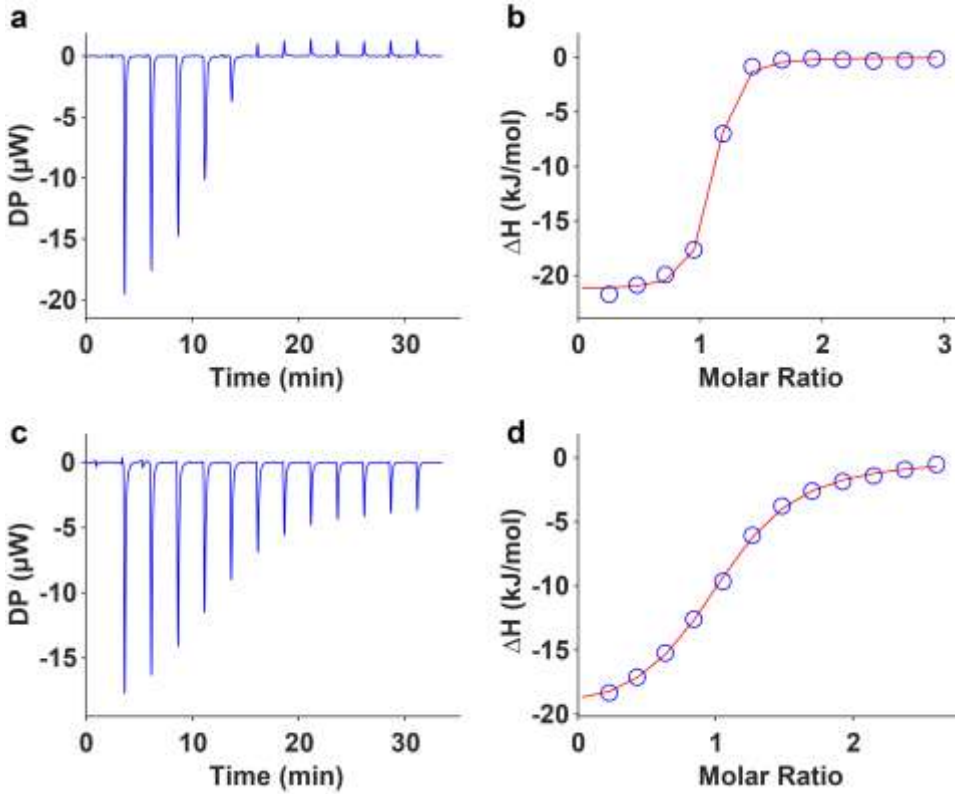

Figure S8. Representative ITC thermograms of EB-3D binding to ChoKα1 at two DMSO concentrations: 0% (a, b) and 6% (c, d) v/v.

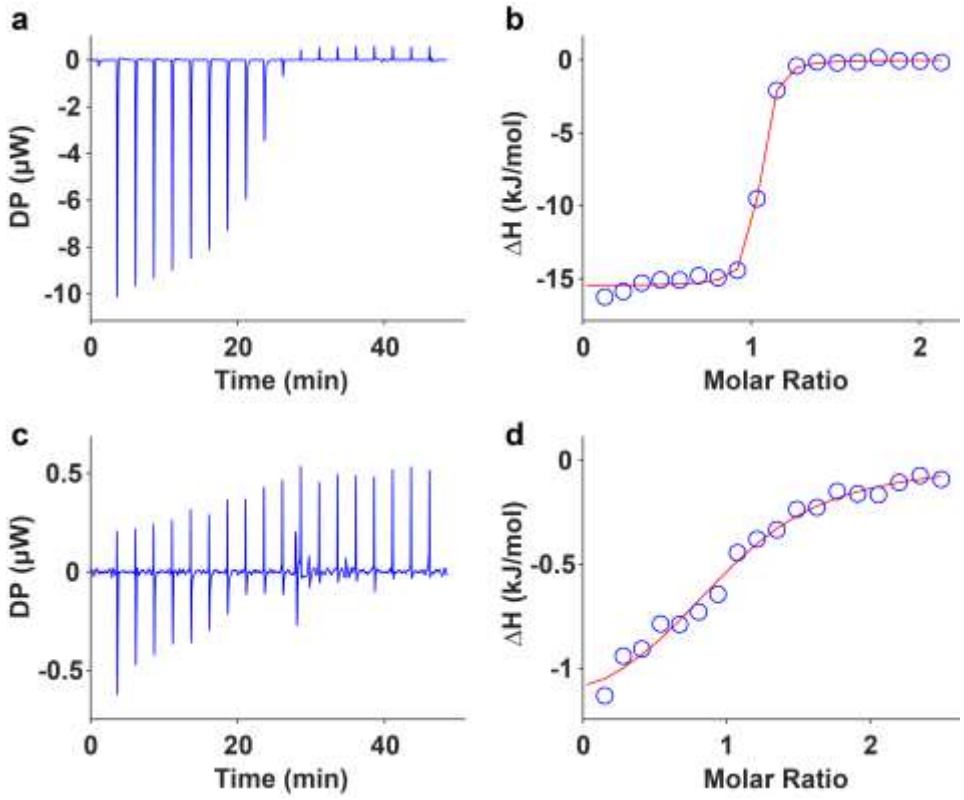

Figure S9. Representative ITC thermograms of sinomenine binding to ChoKα1 at two DMSO concentrations: 0% (a, b) and 6% (c, d) v/v.

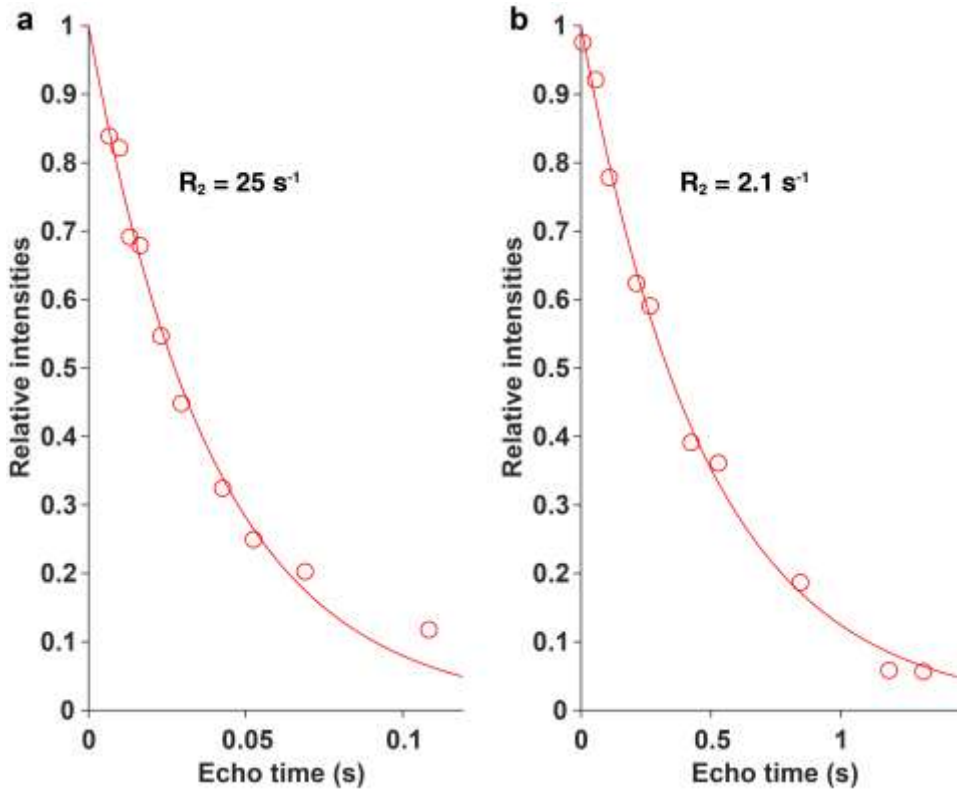

Figure S10. peCPMG Measurement of methyl ^{1}H R_2 values of 200 µM RSM-932A in the presence (a) and absence (b) of 10 µM ChoKα1. Values of fitted relaxation rates are shown in the subplots.

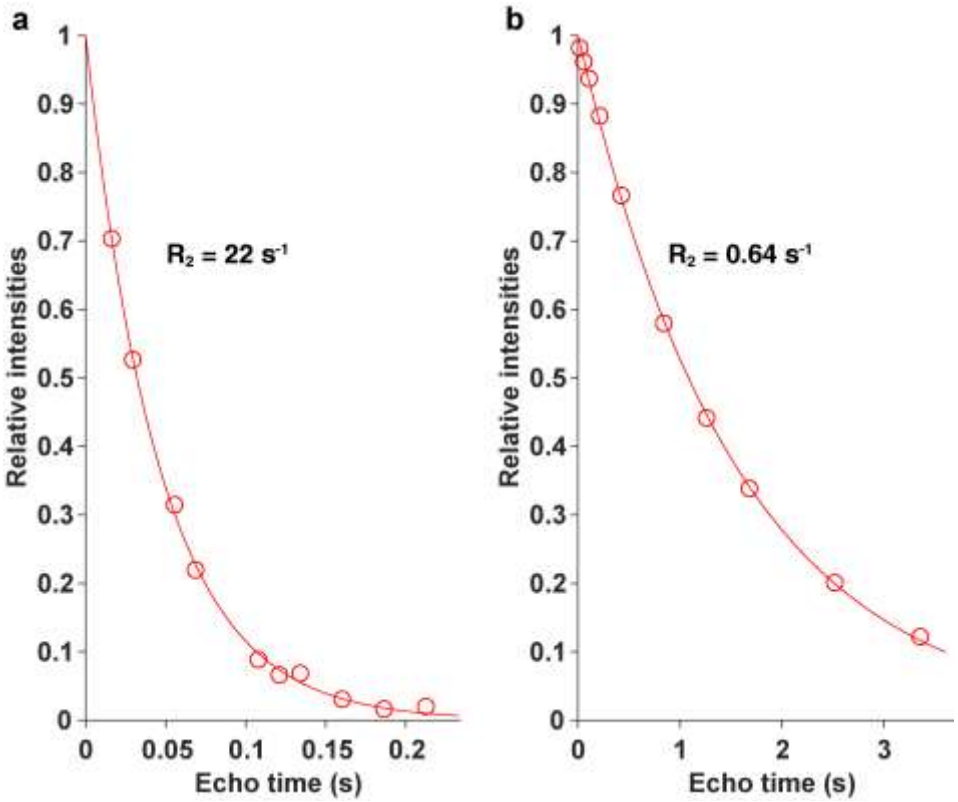

Figure S11. peCPMG Measurement of methyl ^{1}H R_2 values of 200 μM EB-3D in the presence (a) and absence (b) of 10 μM ChoKα1. Values of fitted relaxation rates are shown in the subplots.

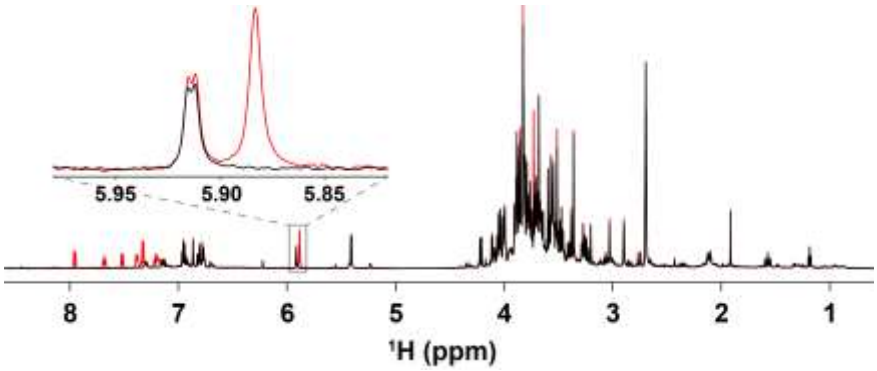

Figure S12. peCPMG spectra of the methanol extract of *Sinomenii Caulis* spiked with 100 μM RSM-932A in the presence (black) and absence (red) of 10 μM ChoKα1. An inset is provided to highlight the zoomed region.