Supporting Information

Exploring heteroaryl-pyrazole carboxylic acids as human carbonic anhydrase XII inhibitors

Roberta Cadoni,[a] Nicolino Pala,[a] Carrie Lomelino,[b] Brian P. Mahon,[b] Robert McKenna,[b] Roberto Dallocchio,[c] Alessandro Dessì,[c] Mauro Carcelli,[d] Dominga Rogolino,[d] Vanna Sanna,[a] Mauro Rassu,[e] Ciro Iaccarino,[e] Daniela Vullo,[f] Claudiu T. Supuran,[f,*] and Mario Sechi[a,*]

[a]*Department of Chemistry and Pharmacy, University of Sassari, Via Vienna 2, 07100 Sassari, Italy*
[b]*Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, 1600 SW Archer Road, PO Box 100245, Gainesville, Florida 32610, United States*
[c]*Istituto CNR di Chimica Biomolecolare, Traversa La Crucca 3, I-07100, Sassari, Italy*
[d]*Department of Chemical, Life Science and Environmental Sustinability, University of Parma, Parco Area delle Scienze 17/A, 43124 Parma, Italy*
[e]*Department of Biomedical Sciences, University of Sassari, Via Muroni 25, 07100 Sassari, Italy*
[f]*Polo Scientifico, Neurofarba Department and Laboratorio di Chimica Bioinorganica, Università degli Studi di Firenze, Rm. 188, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy*

Contents

Experimental section

Chemistry

General methods. Unless otherwise noted, all solvents, including anhydrous solvents and chemicals, were purchased from Aldrich Co. and/or Alfa Aesar, and used without further purification. All reactions involving air- or moisture-sensitive compounds were performed under a nitrogen atmosphere using oven-dried glassware and syringes to transfer solutions. Silica gel thin-layer chromatography (TLC) sheets from Merck silica gel F-254 plates (silica gel precoated aluminum sheets with fluorescent indicator at 254 nm) were used for TLC. Flash chromatography purifications were performed on Merck Silica gel 60 (230-400 mesh ASTM) as a stationary phase. Melting points (m.p.) were determined using an Electrothermal melting point or a Köfler apparatus and are uncorrected. Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded in $CDCl_3$ or DMSO-d_6 on 400 MHz Bruker Avance III nanobay. Chemical shifts (δ) are reported in parts per million (ppm) downfield from tetramethylsilane (TMS), used as an internal standard. Splitting patterns are designated as follows: broad singlet (bs), singlet (s), doublet (d), triplet (t), quartet (q), septet (sep), and multiplet (m). Mass spectra were obtained on a triple-quadrupole QqQ Varian 310-MS mass spectrometer, or a MALDI micro MX (Waters, Micromass) equipped with a reflectron analyser. Elemental analyses were performed on a Perkin-Elmer Elemental Analyzer 2400-CHN at Laboratorio di Microanalisi, Dipartimento di Chimica, Università di Sassari (Italy), and were within ±0.4% of the theoretical values.

General procedure for the preparation of *N*-alkyl-indoles-1*H*-pyrazole-5-carboxylic acids 1b-d.[1,2] A solution of the appropriate ester **2b-d** (0.5 mmol) and NaOH powder (2.0 mmol) in ethanol (16 mL) was stirred under reflux for about 1.5 h. Next, the reaction mixture was poured into water and ice and acidified with 1N HCl to afford a solid precipitate that was filtered under reduced pressure, washed with water, and then recrystallized from water/ethanol to give the desired compound as white/ yellow crystals. **1a** has been previously synthesized and characterized.[1]

3-(1-butyl-1H-indol-3-yl)-1H-pyrazole-5-carboxylic acid (1b). White solid; m.p. 250-252 °C; yield: 86%. ^{1}H NMR 400 MHz (DMSO-d_6): δ 8.01 (bs, 1H), 7.85 (s, 1H), 7.53 (d, 1H), 7.21 (t, 1H), 7.13 (t, 1H), 6.97 (s, 1H), 4.22 (t, 2H), 1.79 (m, 2H), 1.28 (m, 2H), 0.91 (t, 3H); ^{13}C NMR 101 MHz (DMSO-d6): δ 162.3, 144.6, 136.1, 126.8, 124.9, 123.6, 121.7, 120.2, 119.8, 110.1, 106.1, 104.1, 45.3, 31.8, 19.5, 13.5. MALDI: m/z 284 $[M+H]^+$. Anal. Calcd for $C_{16}H_{17}N_3O_2$: C 67.83 H 6.05, N 14.83 Found: C 67.55, H 6.05, N 14.62.

3-(1-isopropyl-1*H*-indol-3-yl)-1*H*-pyrazole-5-carboxylic acid (1c). Yellow solid; m.p. 287 °C dec; yield: 64%. ^{1}H NMR 400 MHz (DMSO-d_6): δ 7.99 (bs, 2H), 7.57 (d, 1H), 7.21 (t, 1H), 7.13 (t, 1H), 7.00 (s, 1H), 4.81 (m, 1H), 1.50 (d, 6H); ^{13}C NMR 101 MHz (DMSO-d_6): δ 162.4, 135.6, 133.7, 125.0, 123.1, 121.6, 120,2, 119.9, 110.2, 109.3, 106.2, 104.0, 46.6, 22.4. LC/MS: m/z 270 $[M+H]^+$. Anal. Calcd for $C_{15}H_{15}N_3O_2$: C 66.90, H 5.61, N 15.60. Found: C 66.73, H 5.81, N 15.48.

3-(1-ethyl-1*H*-indol-3-yl)-1-methyl-1*H*-pyrazole-5-carboxylic acid (1d). White solid; m.p. 224-225 °C; yield: 64%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 8.11 (d, 1H), 7.54 (s, 1H), 7.35 (d, 1H), 7.30-7.20 (m, 3H), 4.27 (s, 3H), 4.22 (q, 2H), 1.52 (t, 3H); ^{13}C NMR 101 MHz (DMSO-d_6): δ 160.8, 145.6, 135.3, 133.7, 126.1, 125.2, 121.5, 120.9, 119.5, 109.8, 107.7, 107.1, 40.4, 38.9, 15.3. MALDI: m/z 270 $[M+H]^+$. Anal. Calcd for $C_{15}H_{15}N_3O_2$: C 66.90 H 5.61, N 15.60. Found: C 66.71, H 5.36, N 15.27.

General procedure for the preparation of *N*-alkyl-indoles-1*H*-pyrazole-5-carboxylates 2a-c [17,23]. Hydrazine monohydrate 98% (28.1 mmol) was added dropwise to a mixture of isopropyl alcohol (7.9 mL) and glacial acetic acid (3.4 mL) cooled at 0 °C. Then the appropriate *N*-alkyl-indoles ß-diketoester **5a-c** (3.2 mmol) was added portionwise. After stirring for 2 h, the mixture was poured into water and ice to afford the product as a brown precipitate that was filtered under reduced pressure. The solid was purified by silica gel flash chromatography using a 8:2 petroleum ether and ethyl acetate mixture as eluent, to give (pale) yellow solid.

Methyl 3-(1-ethyl-1*H*-indol-3-yl)-1*H*-pyrazole-5-carboxylate (2a) [17]. Yellow solid; m.p. 176-178 °C; yield: 93%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 7.95 (d, 1H), 7.47 (s, 1H), 7.28 (d, 1H), 7.32-7.21 (m, 2H), 7.07 (s, 1H), 4.23 (q, 2H), 3.97 (s, 3H), 1.52 (t, 3H); ^{13}C NMR 101 MHz ($CDCl_3$): δ 175.9, 146.2, 136.5, 127.8, 125.2, 122.4, 120.6, 120.0, 117.3, 109.8, 105.7, 105.3, 52.1, 41.2, 15.4. MALDI: m/z 269 $[M]^+$. Elemental analysis: $C_{15}H_{15}N_3O_2$. Calculate C 66.90, H 5.61, N 15.60. Found: C 66.86, H 5.57, N 15.02.

Methyl 3-(1-butyl-1*H*-indol-3-yl)-1*H*-pyrazole-5-carboxylate (2b). Pale yellow solid; m.p. 173-175 °C; yield: 46%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 10.45 (bs, 1H), 7.94 (d, 1H), 7. 44 (s, 1H), 7.40 (d, 1H), 7.32-7.22 (m, 2H), 7.07 (s, 1H), 4.17 (t, 2H), 3.97 (s, 3H), 1.87 (m, 2H), 1.37 (m, 2H), 0.96 (t, 3H); ^{13}C NMR 101 MHz ($CDCl_3$): δ 161.9, 149.1, 143.5, 136.5, 129.9, 126.1, 125.7, 122.3, 120.5, 120.0, 109.9, 105.2, 52.1, 46.3, 32.2, 20.2, 13.7. LC/MS: m/z 298 $[M+H]^+$. Anal. Calcd for $C_{17}H_{19}N_3O_2$: C 68.67, H 6.44, N 14.13. Found: C 68.30, H 6.54, N 13.99.

Methyl 3-(1-isopropyl-1*H*-indol-3-yl)-1*H*-pyrazole-5-carboxylate (2c). Pale yellow solid; m.p. 174-175 °C; yield: 66%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 7.94 (d, 1H), 7.58 (s, 1H), 7.43 (d, 1H), 7.29 (t, 1H), 7.23 (t, 1H), 7.09 (s, 1H), 4.73 (m, 1H), 3.96 (s, 3H), 1.58 (d, 6H); ^{13}C NMR 101 MHz ($CDCl_3$): δ 161.8, 143.7, 140.4, 136.1, 125.7, 122.3, 122.1, 120.6, 119.9, 109.9, 105.9, 105.3, 52.1, 47.4, 22.8. LC/MS: m/z 284 $[M+H]^+$. Anal. Calcd for $C_{16}H_{17}N_3O_2$: C 67.83, H 6.05, N 14.83. Found: C 67.60, H 6.21, N 14.81.

Preparation of methyl 3-(1-ethyl-1*H*-indol-3-yl)-1-methyl-1*H*-pyrazole-5-carboxylate (2d). To a solution of methyl 3-(1-ethyl-1*H*-indol-3-yl)-1*H*-pyrazole-5-carboxylate (**2a**, 1.8 mmol) in anhydrous DMF (2.6 mL) under nitrogen atmosphere, NaH 60% oil dispersion (2.6 mmol) was added portionwise at 0 °C. The mixture thus obtained was stirred for 10 min at 0 °C, then CH_3I (2.6 mmol) was added dropwise, and the stirring was continued for another 1 h at room temperature. The mixture was poured into water and extracted with ethyl acetate. The combined organic layers were then washed with water, dried over Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by flash chromatography (8:2 petroleum ether / ethyl acetate) to give the desired product. Pale brown solid; m.p. 110 °C; yield: 39%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 8.09 (d, 1H), 7.52 (s, 1H), 7.37 (d, 1H), 7.28-7.19 (m, 2H), 7.11 (s, 1H), 4.24 (s, 3H), 4.21 (q, 2H), 3.92 (s, 3H), 1.50 (t, 3H); ^{13}C NMR 101 MHz ($CDCl_3$): δ 160.5, 146.0, 136.2, 132.7, 126.0, 124.9, 121.9, 120.7, 120.0, 109.4, 108.4, 107.9, 51.9, 41.1, 39.4, 15.4. LC/MS: m/z 284 $[M+H]^+$. Anal. Calcd for $C_{16}H_{17}N_3O_2$: C 67.83 H 6.05, N 14.83. Found: C 67.85, H 5.97, N 14.79.

General procedure for the preparation of *N*-alkyl-indoles β-diketoesters 5a-c. A solution of the appropriate 3-alkylacetylindole **4a-c** (3.8 mmol) and diethyl oxalate (5.0 mmol) in methanol (7.45 mL) was added to a solution of sodium methoxide (12.0 mmol), freshly prepared dissolving sodium (12.0 mmol) in methanol (5.2 mL). The mixture was refluxed under nitrogen atmosphere. After 4 h, 6.0 mmol of MeONa and 2.5 mmol of diethyl oxalate were added, and the mixture was refluxed for further 1 h. The reaction was quenched with water and acidified with 1N HCl. The precipitate formed was then recovered by filtration under reduced pressure as yellow solid, which was purified by silica gel flash chromatography using petroleum ether / ethyl acetate (7:3) as eluent.

(*Z*)-methyl 4-(1-ethyl-1*H*-indol-3-yl)-2-hydroxy-4-oxobut-2-enoate (5a). Yellow solid; m.p. 160 °C; yield: 85%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 8.40-8.36 (m, 1H), 7.94 (s, 1H), 7.42-7.40 (m, 1H),

7.36-7.34 (m, 2H), 6.87 (s, 1H), 4.28 (q, 2H), 3.94 (s, 3H), 1.58 (t, 3H). MALDI: m/z 274 $[M+H]^+$. Anal. Calcd for $C_{15}H_{15}NO_4$: C 65.92, H 5.53, N 5.13. Found: C 66.15, H 5.48, N 5.24.

(*Z*)-methyl 4-(1-butyl-1*H*-indol-3-yl)-2-hydroxy-4-oxobut-2-enoate (5b). Yellow solid; m.p. 104-105 °C; yield: 54%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 8.39-8.37 (m, 1H), 7.90 (s, 1H), 7.41-7.39 (m, 1H), 7.36-7.33 (m, 2H), 6.87 (s, 1H), 4.20 (t, 2H), 3.94 (s, 3H), 1.90 (m, 2H), 1.38 (m, 2H), 0.98 (t, 3H). LC/MS: m/z 302 $[M+H]^+$. Anal. Calcd for $C_{17}H_{19}NO_4$: C 67.76, H 6.36, N 4.65. Found: C 67.50, H 6.54, N 4.41.

(*Z*)-methyl 4-(1-isopropyl-1*H*-indol-3-yl)-2-hydroxy-4-oxobut-2-enoate (5c). Yellow solid; m.p. 90-92 °C; yield: 93%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 8.40-8.38 (m, 1H), 8.00 (s, 1H), 7.45-7.43 (m, 1H), 7.36-7.33 (m, 2H), 6.89 (s, 1H), 4.73 (m, 1H), 3.95 (s, 3H), 1.62 (d, 6H). LC/MS: m/z 288 $[M+1]^+$. Anal. Calcd for $C_{16}H_{17}NO_4$: C 66.89, H 5.96, N 4.88. Found: C 66.89, H 6.07, N 4.74.

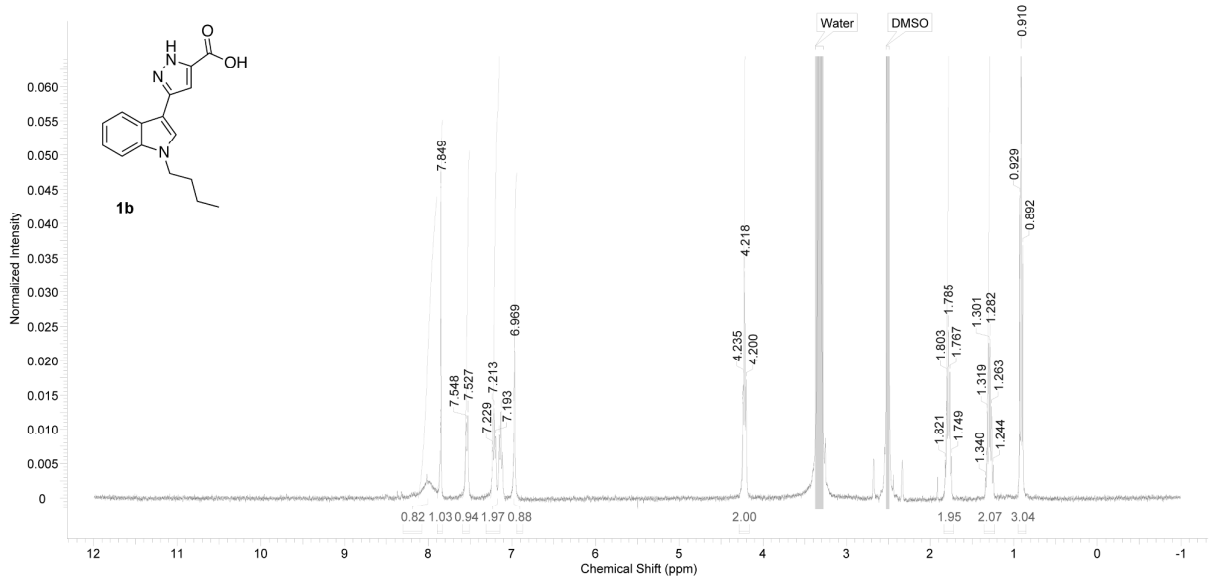
General procedure for the preparation of *N*-alkyl-3-acetyl indoles 4a-c. Anhydrous DMSO (25 mL) was added to KOH crushed pellets (50 mmol), and the mixture was stirred at room temperature for 5 min. 3-Acetylindole (**3**, 13 mmol) was then added, and the mixture was stirred at room temperature for 45-60 min. An appropriate alkyl halide (bromoethane, 2-iodopropane, 1-iodobutane, 26 mmol) was added and the stirring was continued for further 45-60 min. After reaction completion, water was added and the white precipitate that formed was filtered under reduced pressure and washed with water. In case of compound 1-(1-butyl-1*H*-indol-3-yl)ethanone (**3c**), product was extracted from the aqueous solution with diethyl ether, the organic layer was washed with water and dried over Na_2SO_4, and the solvent was evaporated under reduced pressure. The solid or oil obtained was purified by silica gel flash chromatography (8:2 petroleum ether / ethyl acetate).

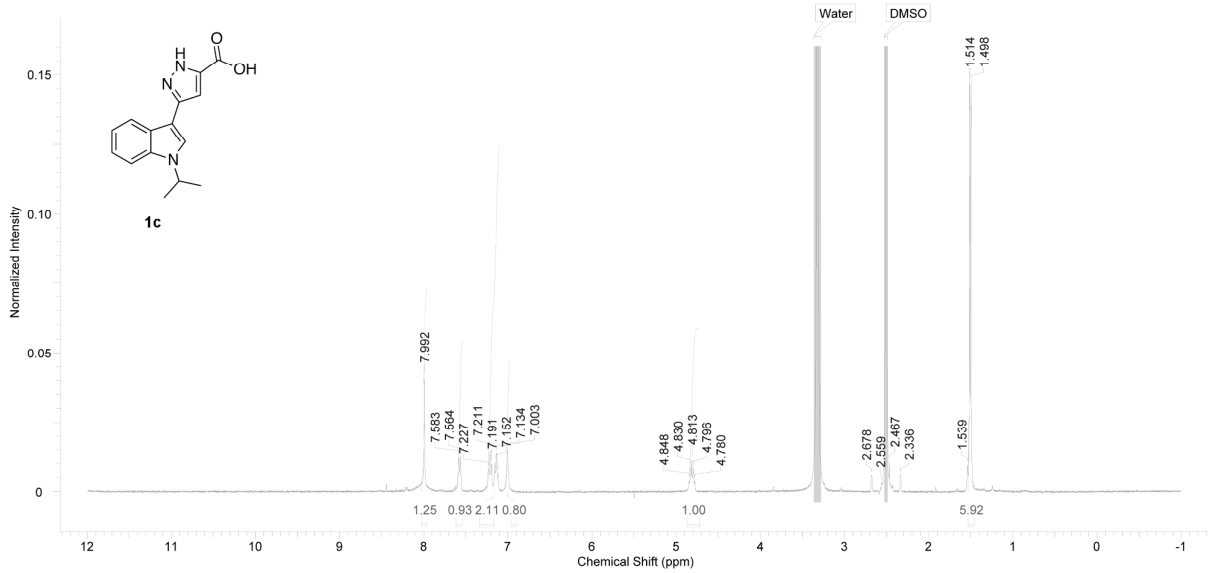
1-(1-ethyl-1*H*-indol-3-yl)ethanone (4a).[2] White crystals; m.p. 87-89 °C; yield: 89%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 8.39-8.36 (m, 1H), 7.77 (s, 1H), 7.38-7.36 (m, 1H), 7.32-7.27 (m, 2H), 4.21 (q,

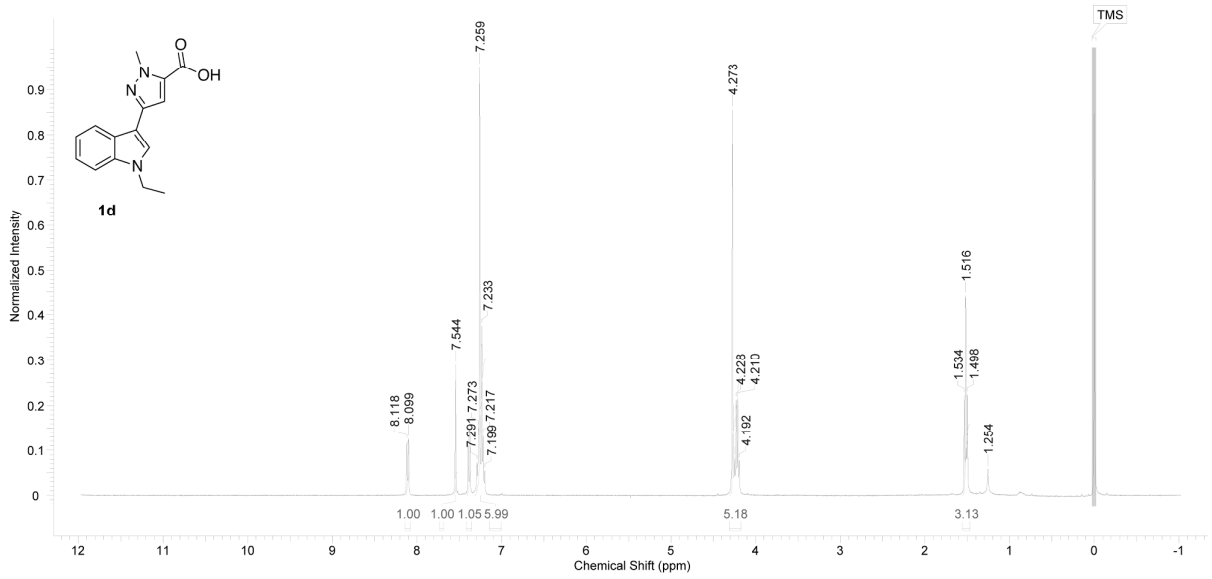
2H), 2.53 (s, 3H), 1.52 (t, 3H). LC/MS: m/z 188.1 $[M+H]^+$, 210.1 $[M+Na]^+$. Anal. Calcd for $C_{12}H_{13}NO$: C 76.98, H 7.00, N 7.48. Found: C 77.12, H 7.09, N 7.35.

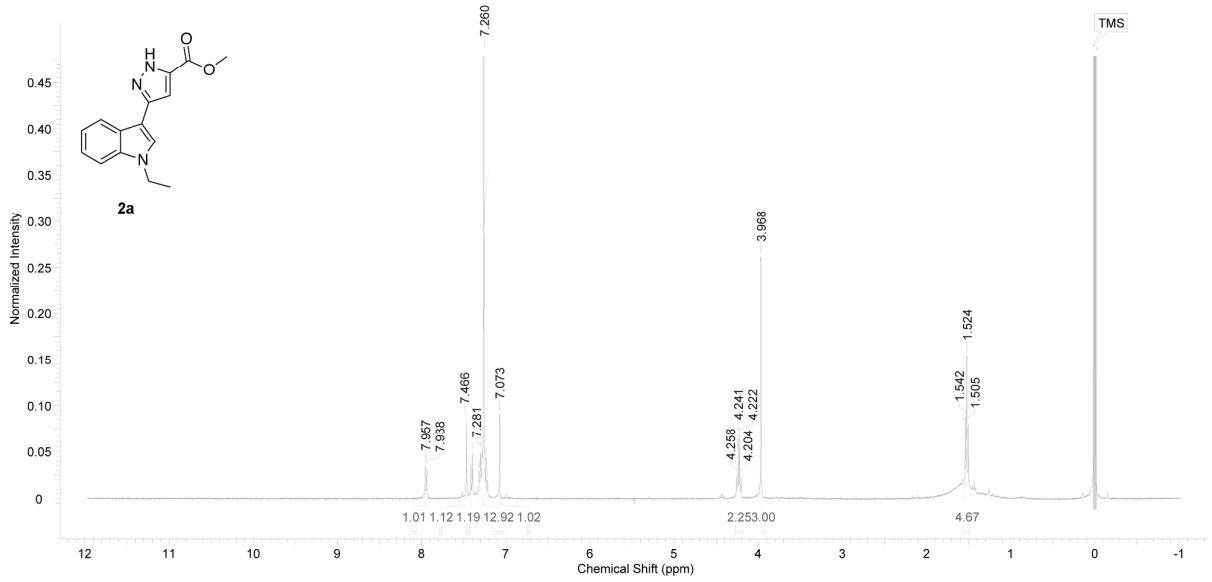
1-(1-butyl-1*H*-indol-3-yl)ethanone (4b). Pale yellow oil; yield: 91%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 8.39-8.36 (m, 1H), 7.73 (s, 1H), 7.37-7.35 (m, 1H), 7.30-7.28 (m, 2H), 4.15 (t, 2H), 2.53 (s, 3H), 1.87 (m, 2H), 1.37 (m, 2H), 0.96 (t, 3H). LC/MS: m/z 215.1 $[M+H]^+$. Anal. Calcd for $C_{14}H_{17}NO$: C 78.10, H 7.96, N 6.51. Found: C 77.94, H 8.06, N 6.32.

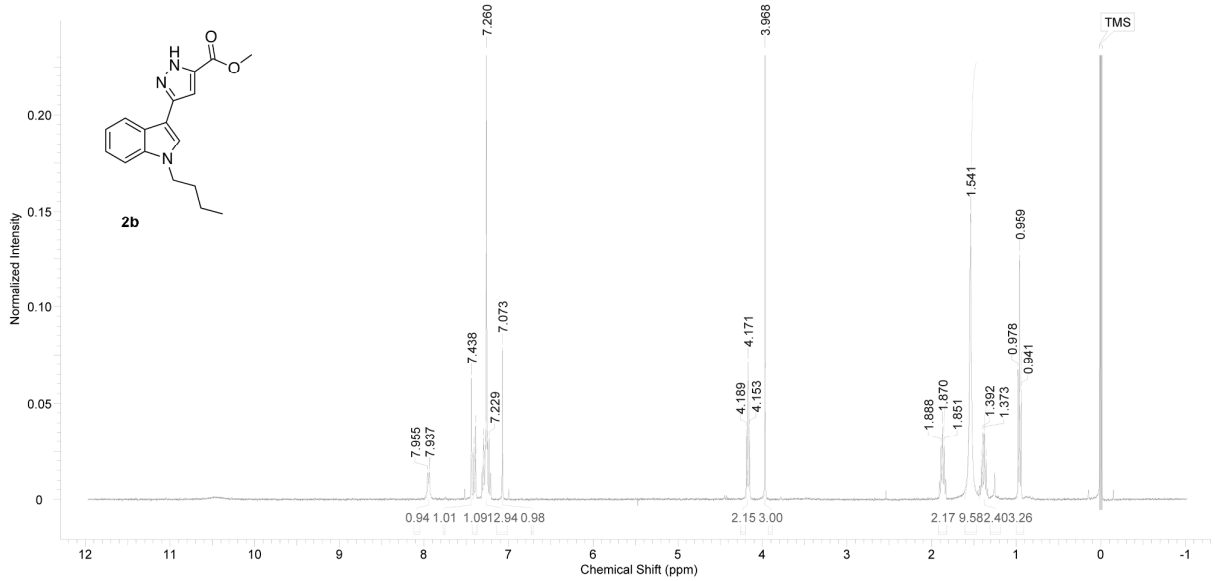
1-(1-isopropyl-1*H*-indol-3-yl)ethanone (4c). Beige solid; m.p. 82-83 °C; yield: 50%. ^{1}H NMR 400 MHz ($CDCl_3$): δ 8.38-8.36 (m, 1H), 7.86 (s, 1H), 7.41-7.39 (m, 1H), 7.31-7.29 (m, 2H), 4.71 (m, 1H), 2.55 (s, 3H), 1.60 (d, 6H). LC/MS: m/z 202.1 $[M+H]^+$ Anal. Calcd for $C_{13}H_{15}NO$: C 77.58, H 7.51, N 6.96. Found: C 77.30, H 7.68, N 6.73.

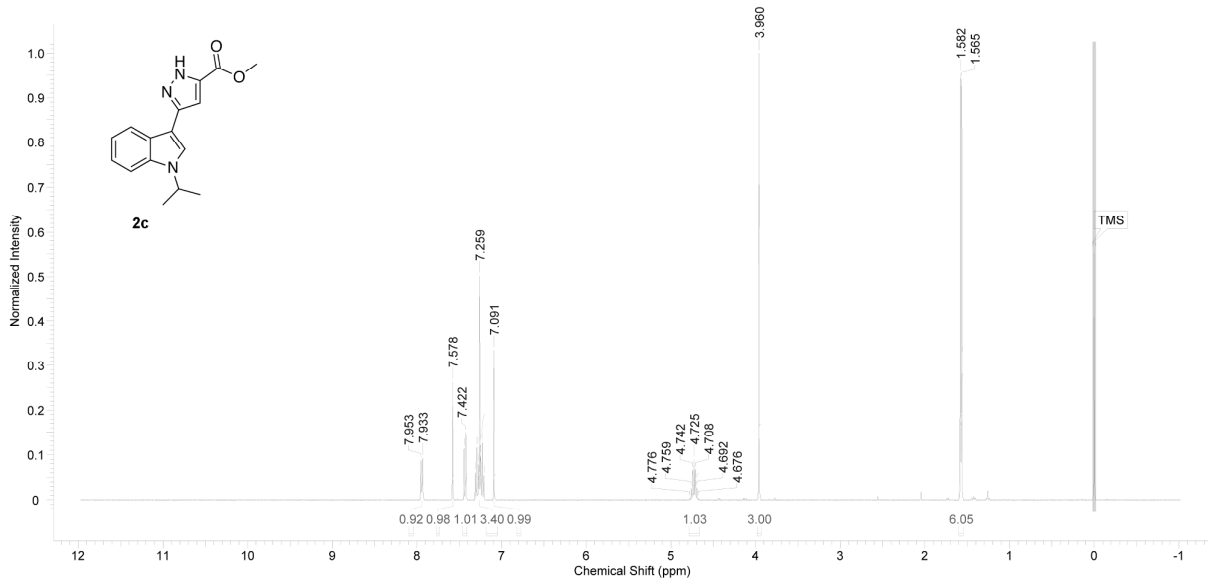

1b
Water
DMSO
7.849
7.548
7.527
7.229
7.213
7.193
6.969
4.235
4.218
4.200
1.803
1.785
1.767
1.749
1.321
1.340
1.319
1.301
1.282
1.263
1.244
0.929
0.910
0.892
0.82 1.03 0.94 1.97 0.88
2.00
1.95 2.07 3.04
Normalized Intensity
Chemical Shift (ppm)

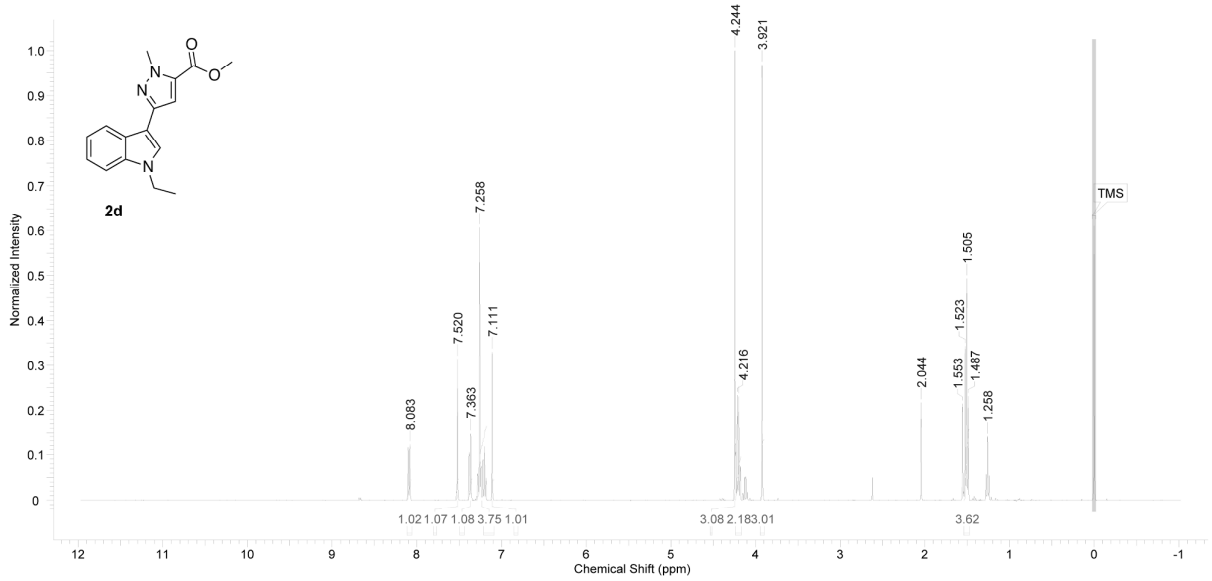

Water
DMSO
1.514
1.498
7.992
7.583
7.564
7.227
7.211
7.191
7.152
7.134
7.003
4.848
4.830
4.813
4.796
4.780
2.678
2.559
2.467
2.336
1.539
1.25
0.93
2.11
0.80
1.00
5.92
Normalized Intensity
Chemical Shift (ppm)
1c

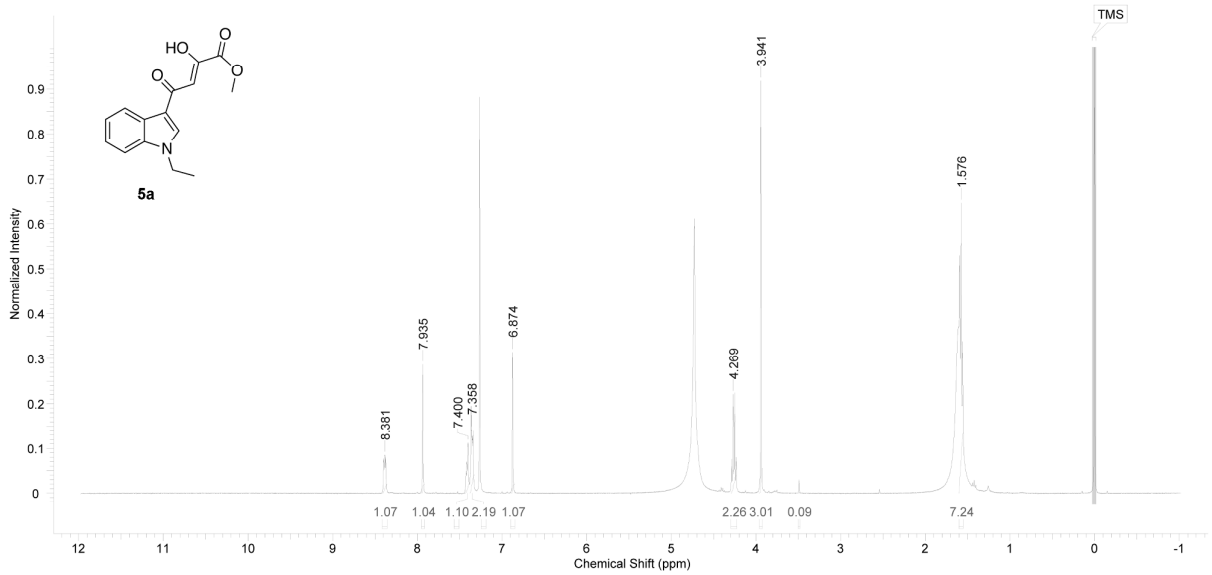

1d
Normalized Intensity
Chemical Shift (ppm)
TMS
8.118
8.099
7.544
7.291
7.273
7.259
7.233
7.217
7.199
4.273
4.223
4.210
4.192
1.516
1.534
1.498
1.254
1.00
1.00
1.05
5.99
5.18
3.13

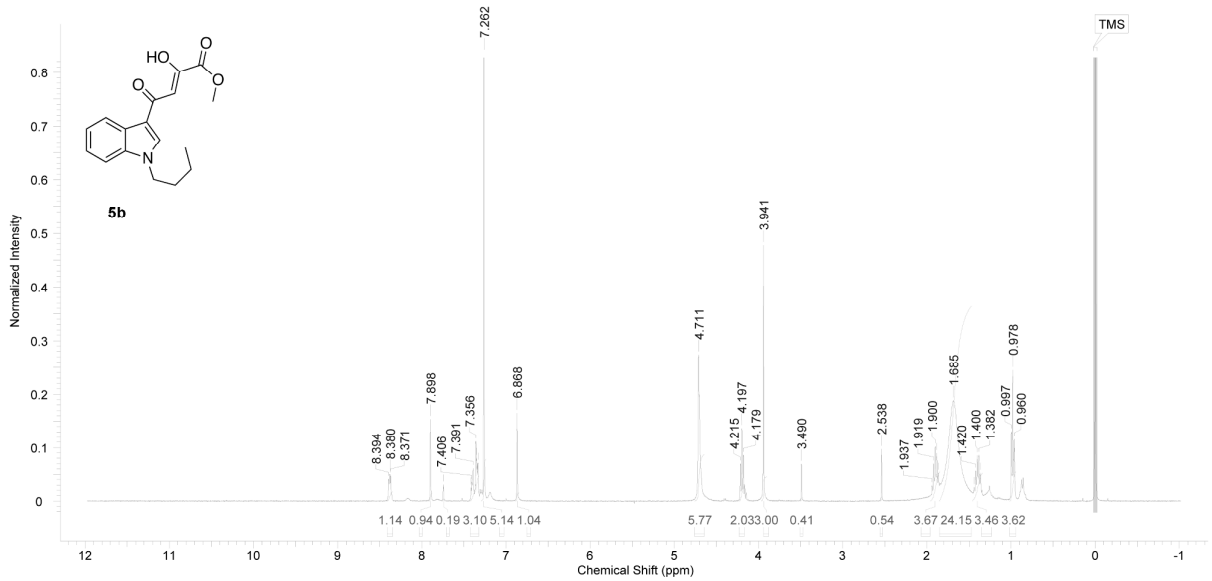

2a
Normalized Intensity
Chemical Shift (ppm)
7.260
TMS
3.968
1.524
7.957
7.938
7.466
7.281
7.073
4.258
4.241
4.222
4.204
1.542
1.505
1.01 1.12 1.19 12.92 1.02
2.253.00
4.67

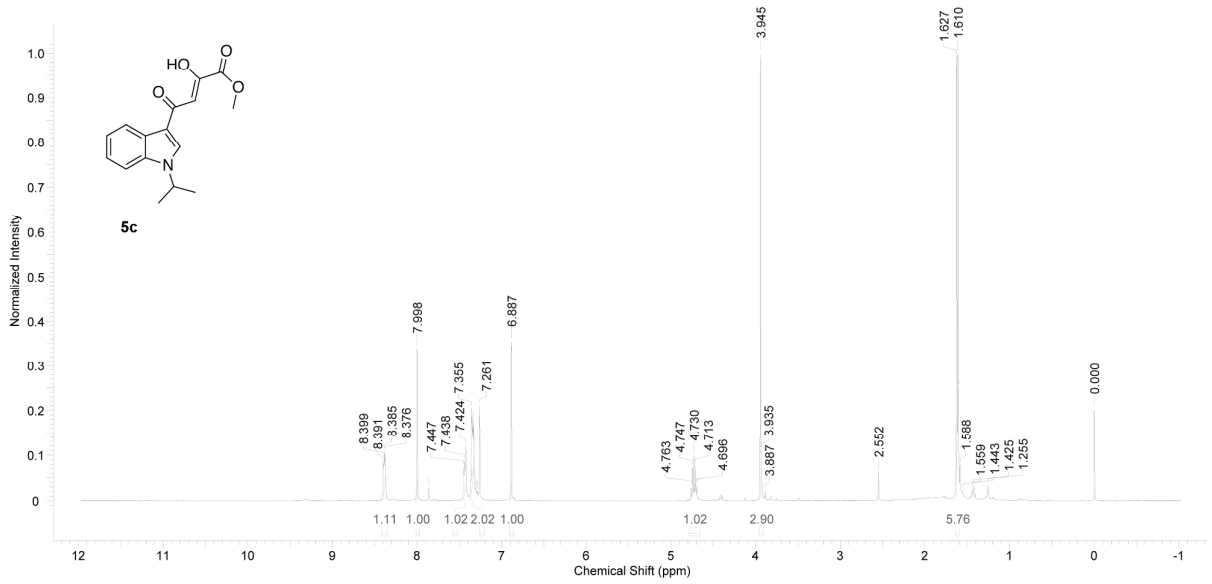

2b
Normalized Intensity
Chemical Shift (ppm)
TMS
7.260
3.968
7.955
7.937
7.438
7.229
7.073
4.189
4.171
4.153
1.888
1.870
1.851
1.541
1.392
1.373
0.978
0.959
0.941
0.94 1.01 1.09 12.94 0.98
2.15 3.00
2.17 9.58 2.40 3.26

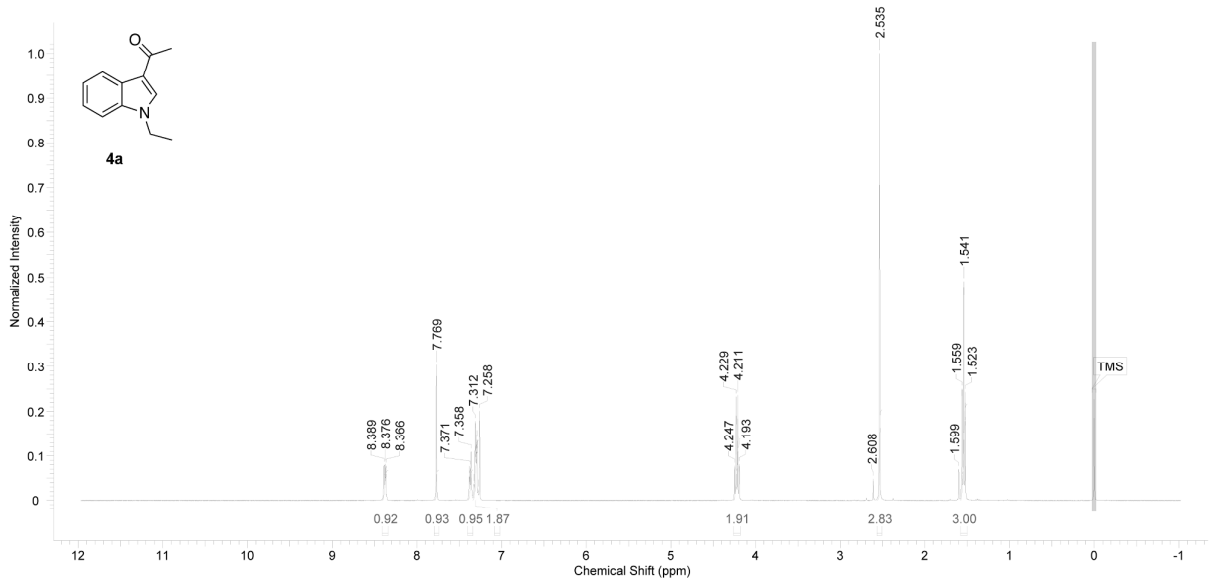

2c
Normalized Intensity
Chemical Shift (ppm)
7.953
7.933
7.578
7.422
7.259
7.091
4.776
4.759
4.742
4.725
4.708
4.692
4.676
3.960
1.582
1.565
TMS
0.92 0.98 1.01 3.40 0.99
1.03
3.00
6.05

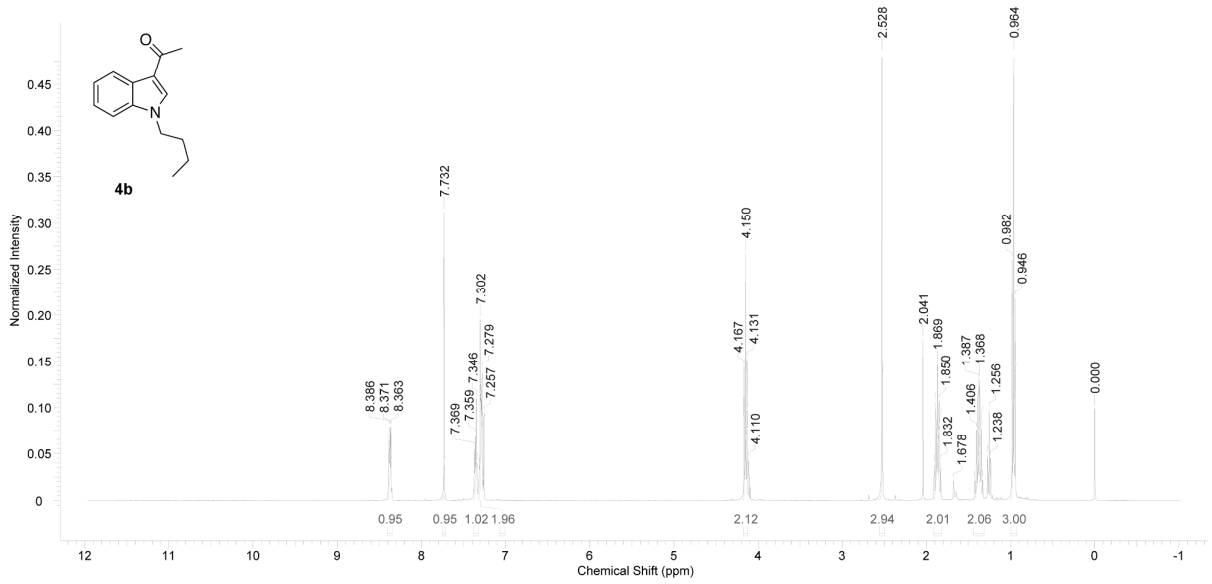

2d
Normalized Intensity
Chemical Shift (ppm)
8.083
7.520
7.363
7.258
7.111
4.244
3.921
4.216
2.044
1.553
1.523
1.505
1.487
1.258
TMS
1.02 1.07 1.08 3.75 1.01
3.08 2.18 3.01
3.62

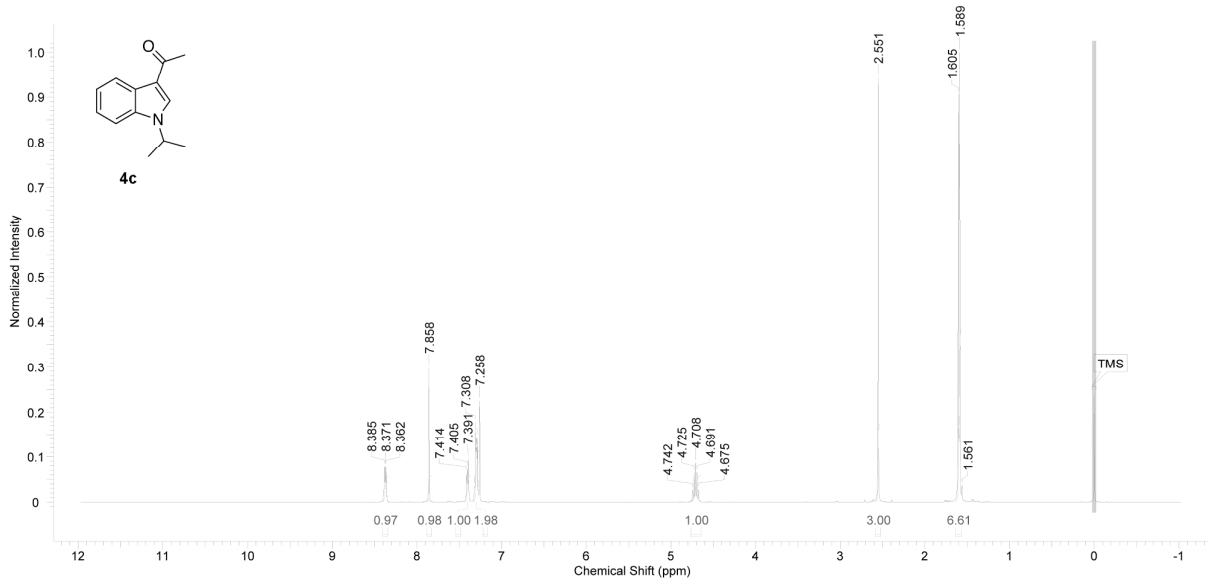

HO
O
O
O
N
5a
TMS
3.941
1.576
8.381
7.935
7.400
7.358
6.874
4.269
1.07
1.04
1.10
2.19
1.07
2.26
3.01
0.09
7.24
Normalized Intensity
Chemical Shift (ppm)

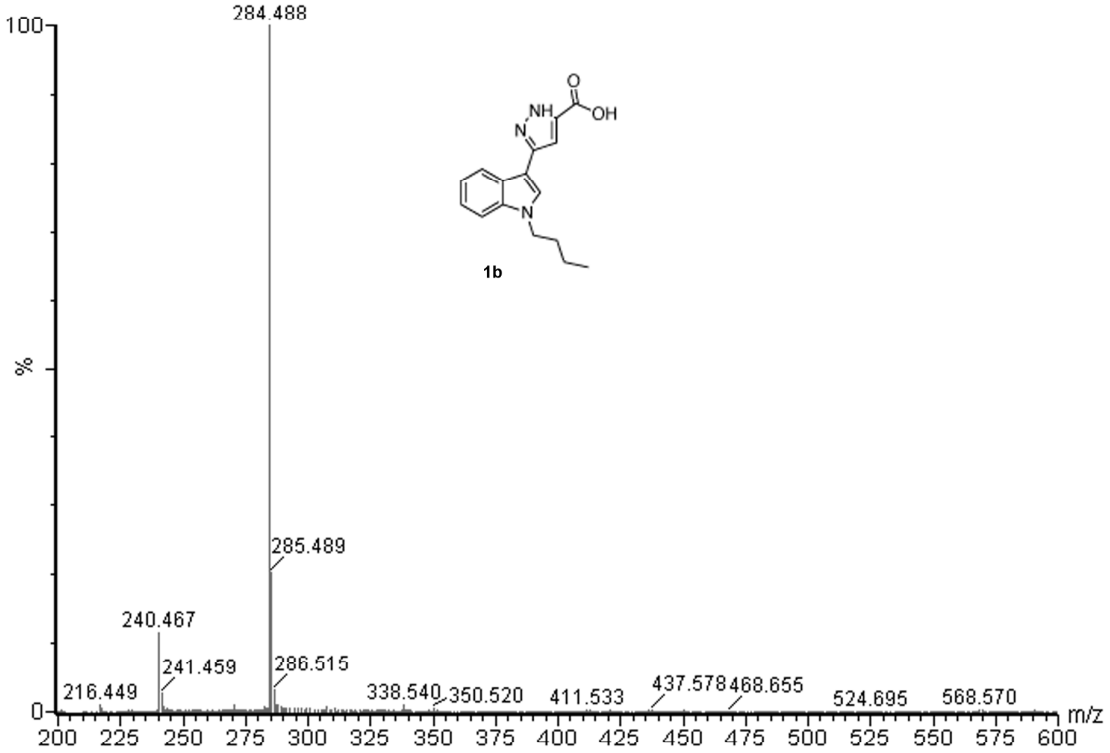

HO
O
O
O
N
5b
Normalized Intensity
Chemical Shift (ppm)
TMS
7.262
3.941
4.711
7.898
6.868
7.356
4.197
4.215
4.179
0.978
1.665
3.490
2.538
1.900
1.919
1.937
1.420
1.400
1.382
0.997
0.960
8.394
8.380
8.371
7.406
7.391
1.14 0.94 0.19 3.10 5.14 1.04
5.77 2.033.00 0.41
0.54 3.67 24.15 3.46 3.62

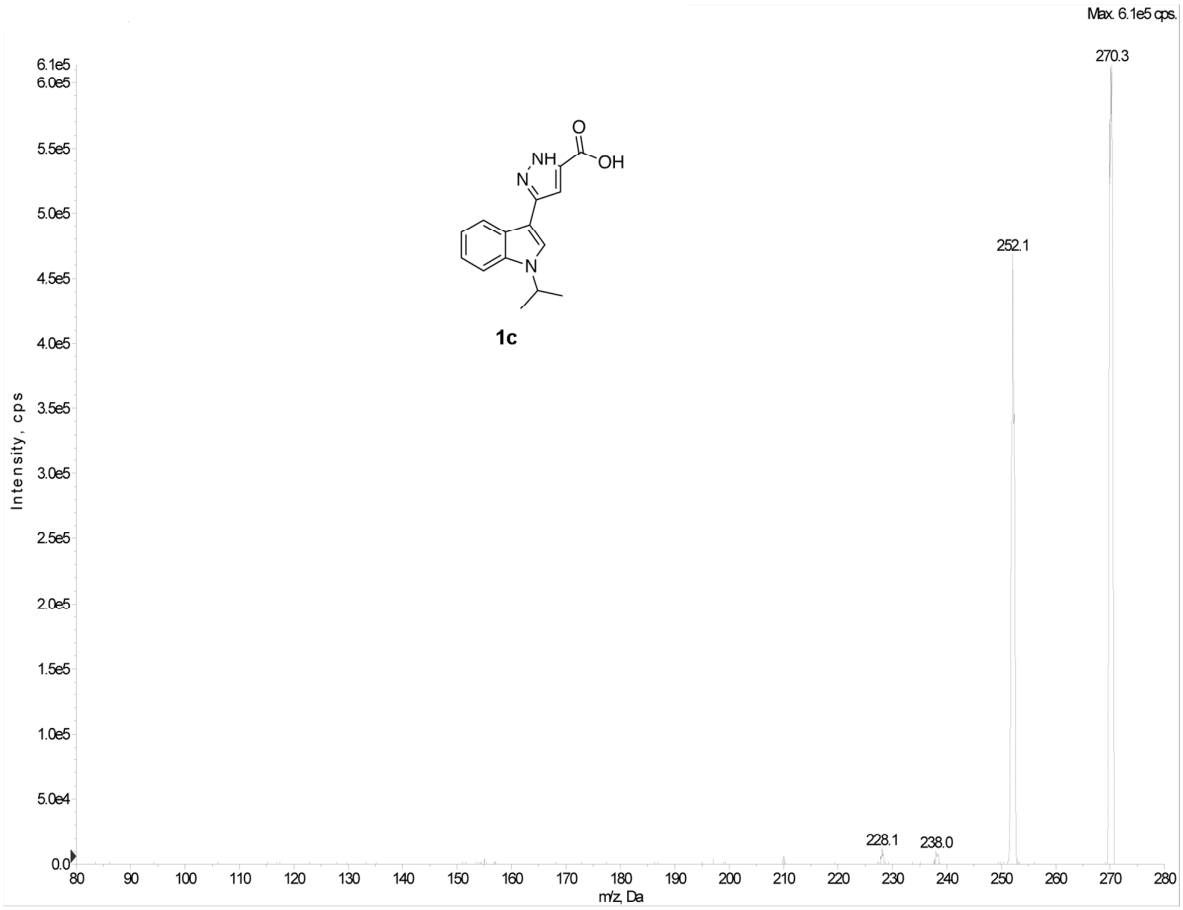

5c
Normalized Intensity
Chemical Shift (ppm)
8.399
8.391
8.385
8.376
7.998
7.447
7.438
7.424
7.355
7.261
6.887
4.763
4.747
4.730
4.713
4.696
3.945
3.935
3.887
2.552
1.627
1.610
1.588
1.559
1.443
1.425
1.255
0.000
1.11
1.00
1.02
2.02
1.00
1.02
2.90
5.76

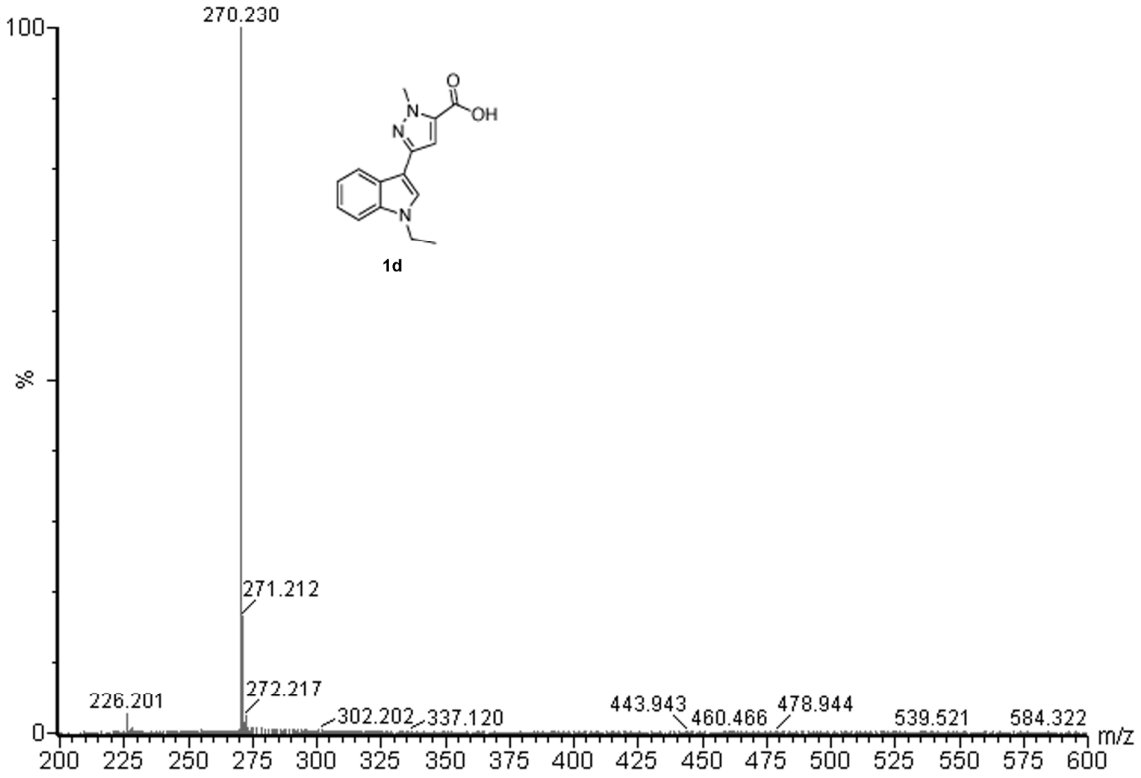

4a
Normalized Intensity
Chemical Shift (ppm)
8.389
8.376
8.366
7.769
7.371
7.358
7.312
7.258
4.229
4.211
4.247
4.193
2.608
2.535
1.599
1.559
1.541
1.523
TMS
0.92
0.93
0.95
1.87
1.91
2.83
3.00

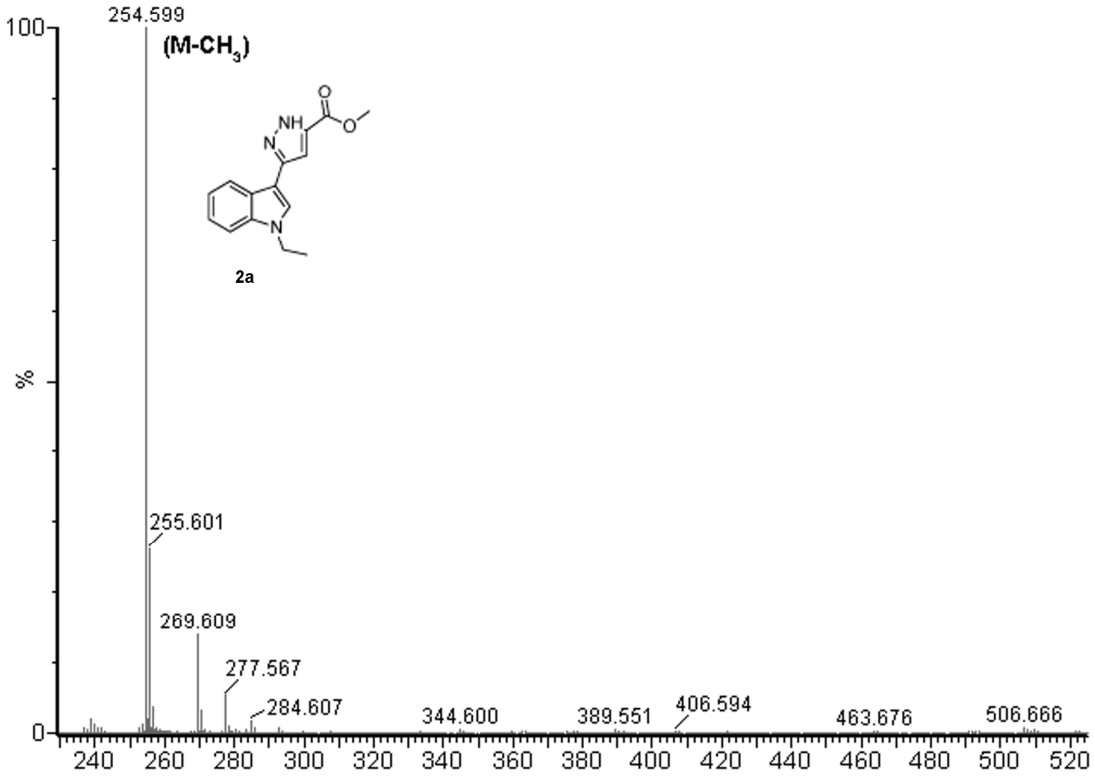

4b
Normalized Intensity
Chemical Shift (ppm)
8.386
8.371
8.363
7.732
7.369
7.359
7.346
7.302
7.279
7.257
4.167
4.150
4.131
4.110
2.528
2.041
1.869
1.850
1.678
1.406
1.387
1.368
1.332
1.256
1.238
0.982
0.964
0.946
0.000
0.95
0.95
1.02
1.96
2.12
2.94
2.01
2.06
3.00

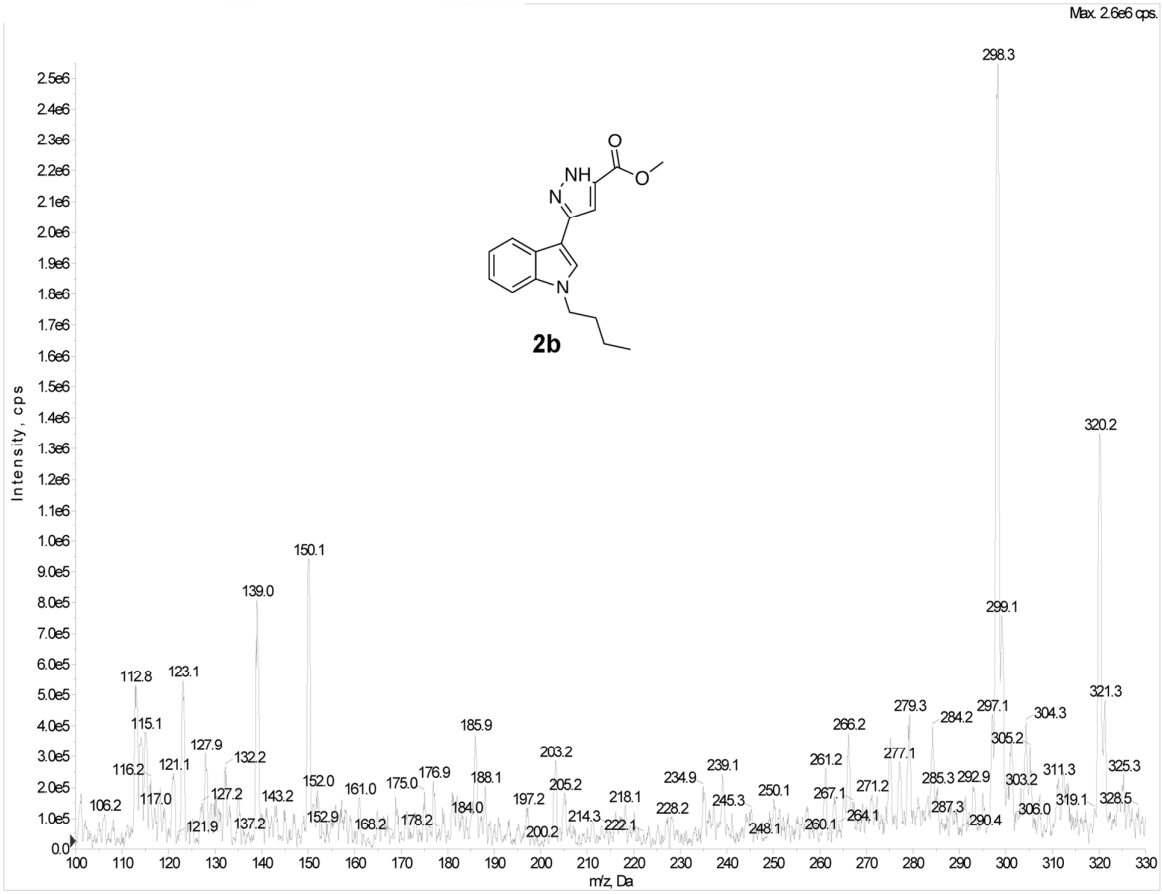

4c
Normalized Intensity
Chemical Shift (ppm)
8.385
8.371
8.362
7.858
7.414
7.405
7.391
7.308
7.258
4.742
4.725
4.708
4.691
4.675
2.551
1.605
1.589
1.561
TMS
0.97
0.98
1.00
1.98
1.00
3.00
6.61

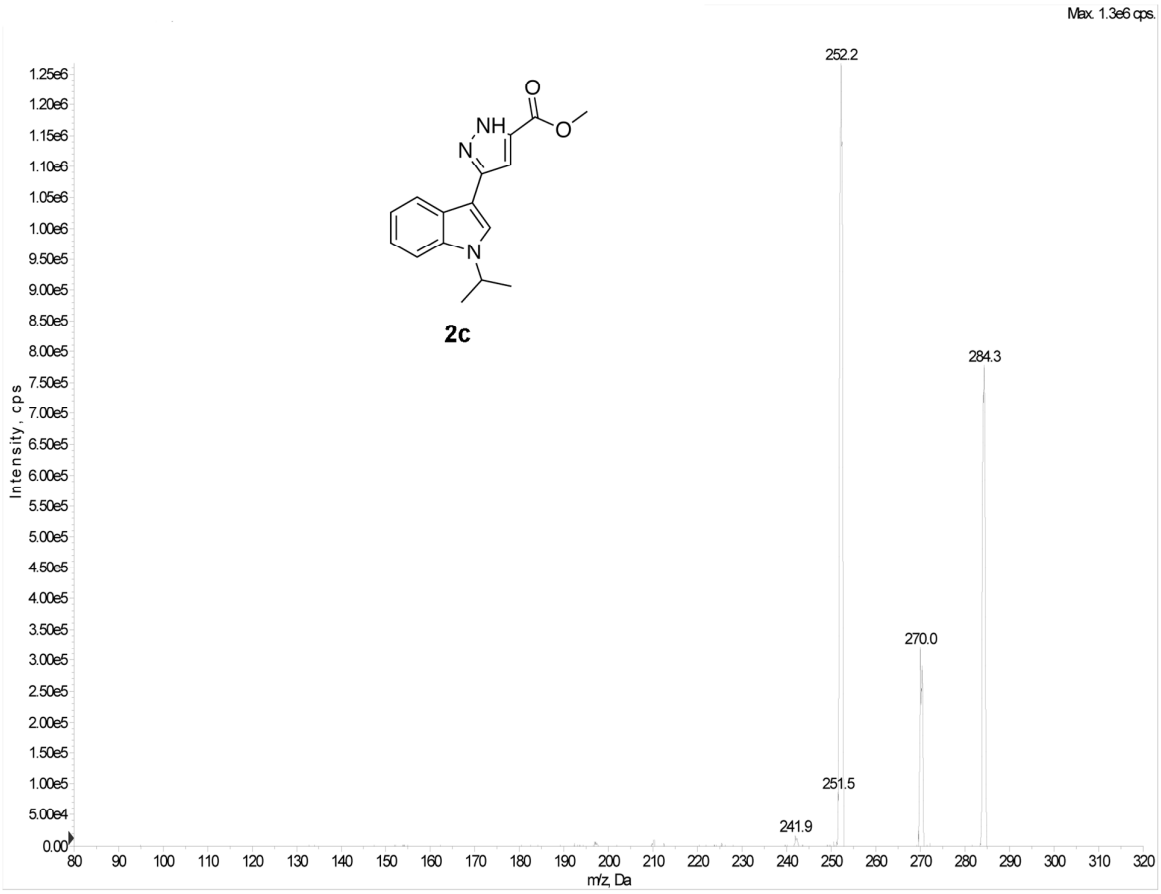

284.488
100
%
0
285.489
240.467
241.459
286.515
216.449
338.540
350.520
411.533
437.578
468.655
524.695
568.570
m/z
200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600
O
OH
NH
N
N
1b


Max. 6.1e5 cps.
270.3
252.1
228.1
238.0
Intensity, cps
m/z, Da
1c

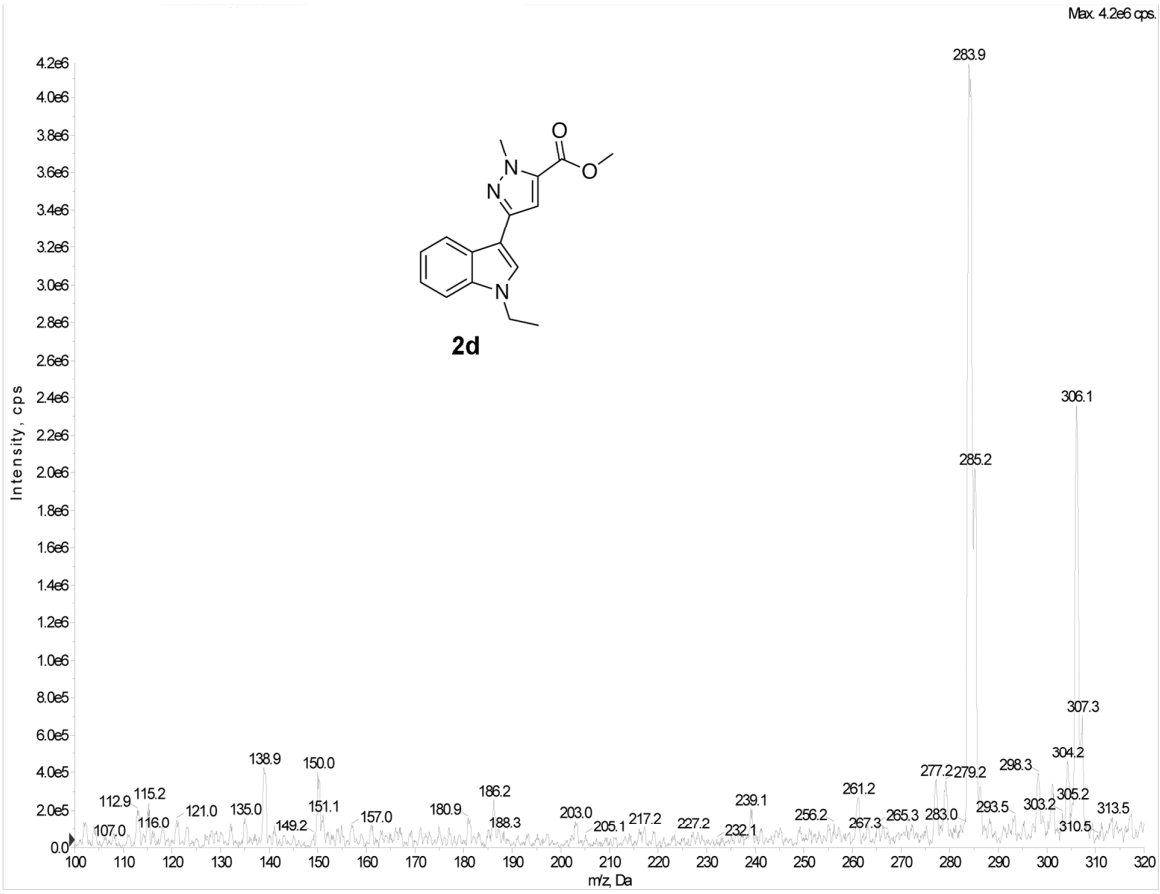

100
%
0
270.230
271.212
272.217
226.201
302.202
337.120
443.943
460.466
478.944
539.521
584.322
m/z
200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600
O
OH
N
N
N
1d

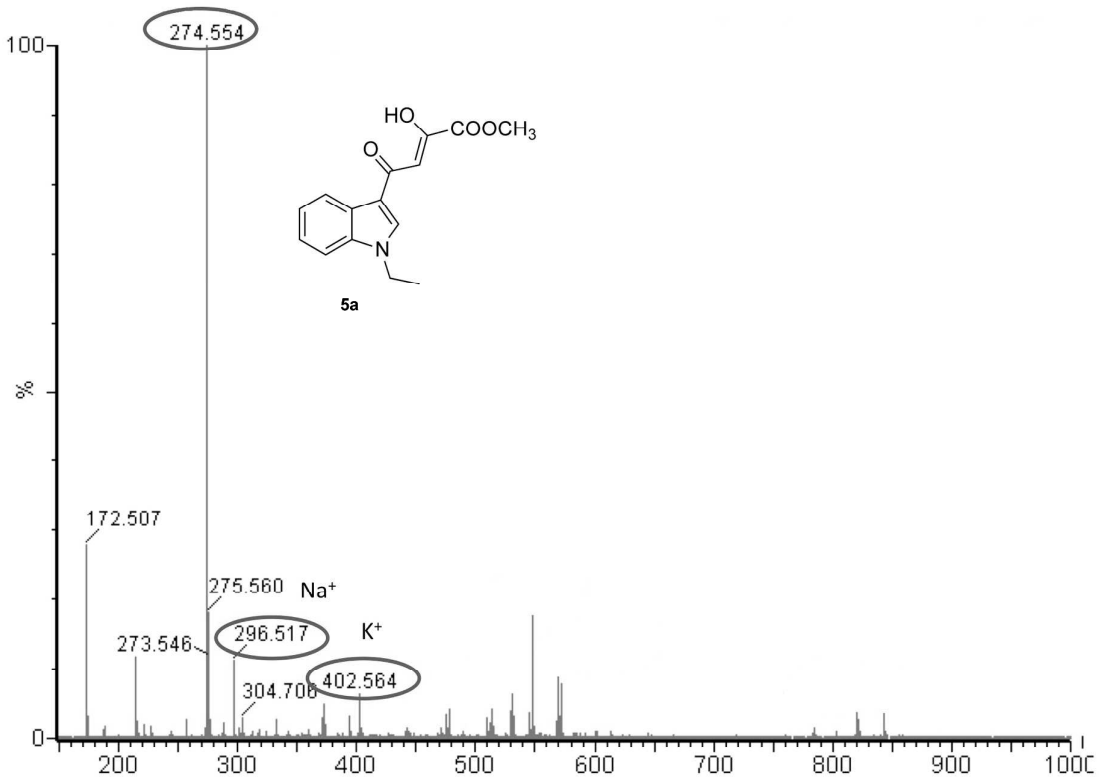

254.599
(M-CH3)
NH
N
O
O
N
2a
100
%
0
255.601
269.609
277.567
284.607
344.600
389.551
406.594
463.676
506.666
240
260
280
300
320
340
360
380
400
420
440
460
480
500
520

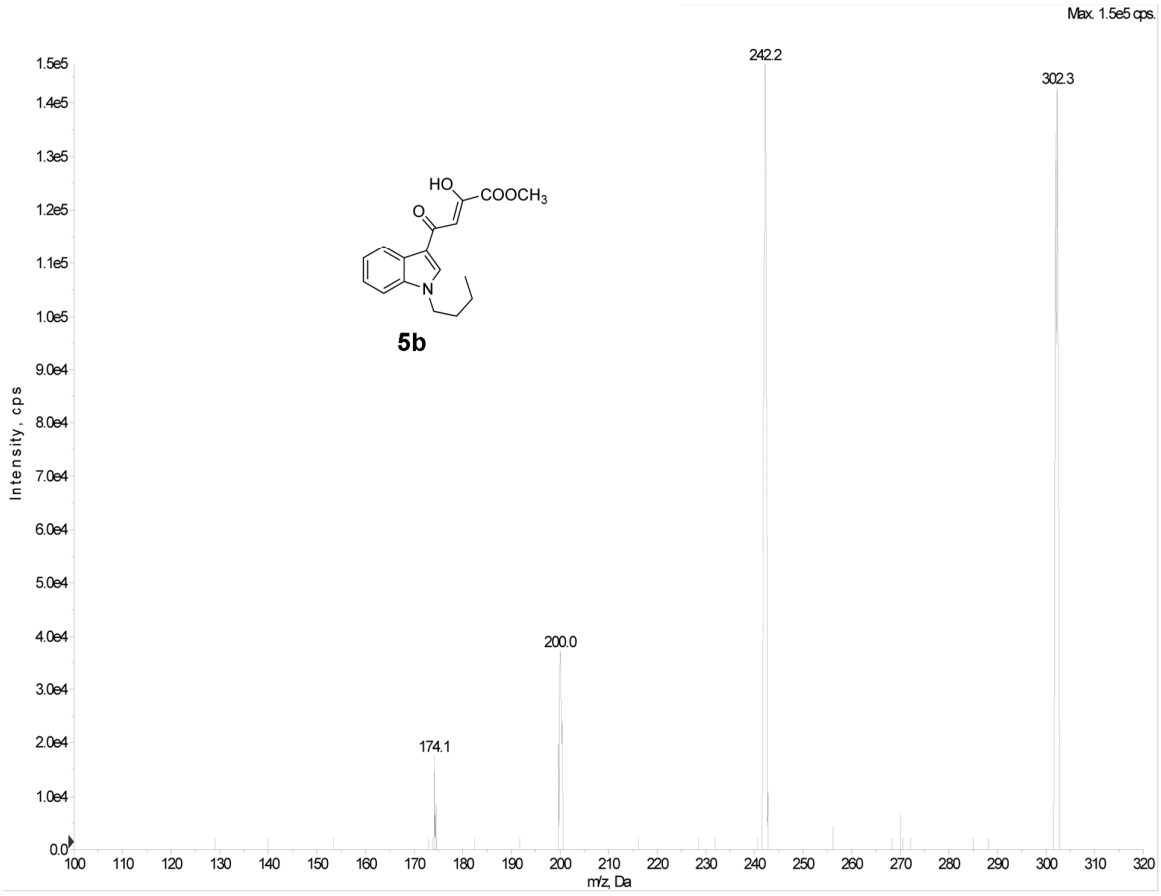

Max. 2.6e6 cps.
Intensity, cps
m/z, Da
2b
298.3
320.2
150.1
139.0
299.1
112.8
123.1
321.3
279.3
297.1
284.2
304.3
185.9
266.2
203.2
305.2
277.1
239.1
325.3
115.1
127.9
132.2
116.2
121.1
188.1
261.2
234.9
311.3
152.0
175.0
176.9
205.2
285.3
292.9
303.2
106.2
117.0
127.2
143.2
161.0
197.2
218.1
245.3
250.1
267.1
271.2
319.1
328.5
121.9
137.2
152.9
168.2
178.2
184.0
200.2
214.3
222.1
228.2
248.1
260.1
264.1
287.3
290.4
306.0

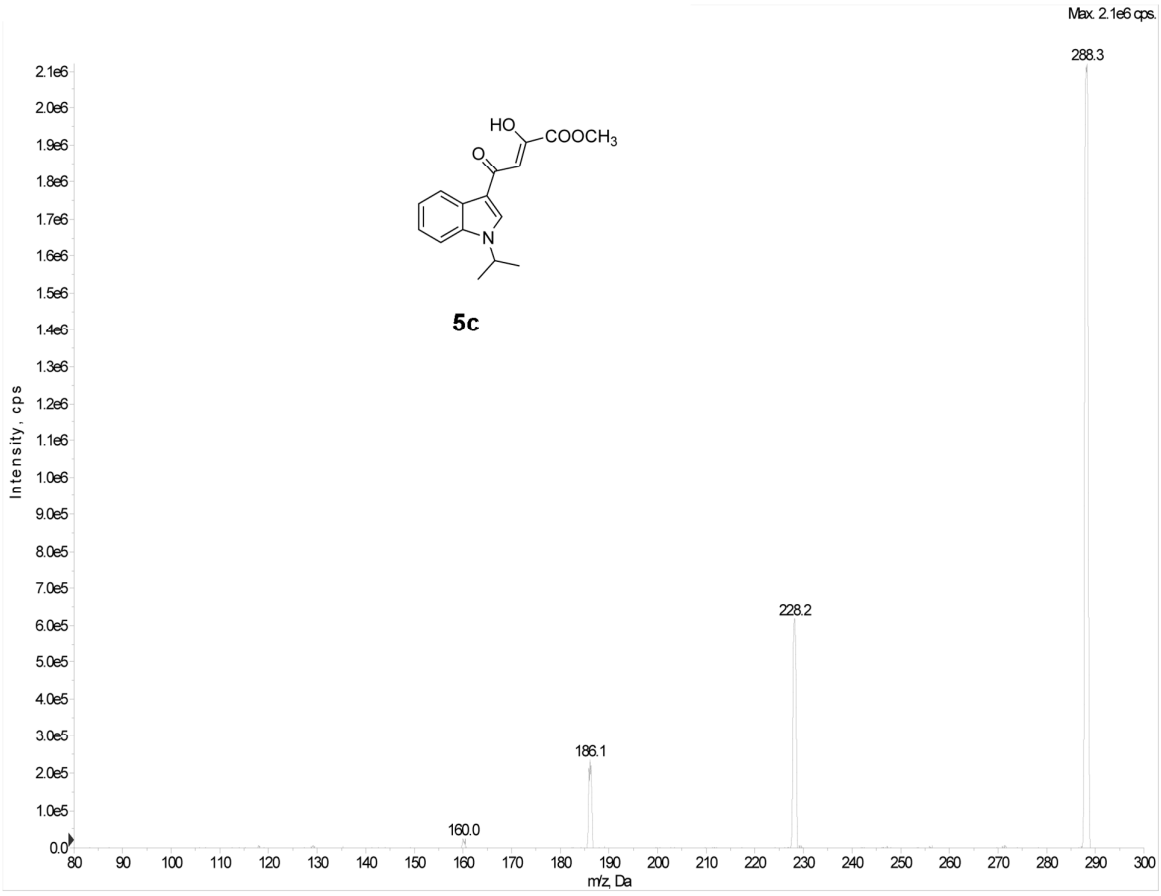

Max. 1.3e6 cps.
252.2
284.3
270.0
251.5
241.9
Intensity, cps
m/z, Da
NH
N
O
O
N

2c


Max. 4.2e6 cps.
Intensity, cps
m/z, Da
2d
283.9
306.1
285.2
307.3
304.2
277.2
279.2
298.3
261.2
305.2
239.1
138.9
150.0
186.2
151.1
112.9
115.2
116.0
121.0
135.0
149.2
157.0
180.9
188.3
203.0
205.1
217.2
227.2
232.1
256.2
267.3
265.3
283.0
293.5
303.2
310.5
313.5
107.0


274.554
HO
COOCH3
O
N
5a
100
%
0
172.507
275.560
Na+
296.517
K+
273.546
304.708
402.564
200
300
400
500
600
700
800
900
1000


Max. 1.5e5 cps.
242.2
302.3
200.0
174.1
HO
COOCH3
O
N
5b
Intensity, cps
m/z, Da


Max. 2.1e6 cps.
288.3
228.2
186.1
160.0
HO
COOCH3
O
N
5c
Intensity, cps
m/z, Da

Biology

CA inhibition assays. An Applied Photophysics stopped-flow instrument has been used for assaying the CA catalysed CO_2 hydration activity.[3] Phenol red (at a concentration of 0.2 mM) has been used as indicator, working at the absorbance maximum of 557 nm, with 20 mM Hepes (pH 7.4) and 20 mM $NaBF_4$ (for maintaining constant the ionic strength), following the initial rates of the CA-catalyzed CO_2 hydration reaction for a period of 10-100 s. The CO_2 concentrations ranged from 1.7 to 17 mM for the determination of the kinetic parameters and inhibition constants. For each inhibitor, at least six traces of the initial 5-10% of the reaction have been used for determining the initial velocity.[3] The uncatalyzed rates were determined in the same manner and subtracted from the total observed rates. Stock solutions of inhibitor (10 mM) were prepared in distilled-deionized water and dilutions up to 0.01 nM were done thereafter with distilled-deionized water. Inhibitor and enzyme solutions were preincubated together for 15 min at RT prior to assay, in order to allow for the formation of the E-I complex. The inhibition constants were obtained by non-linear least-squares methods using PRISM 3, whereas the kinetic parameters for the uninhibited enzymes from Lineweaver-Burk plots, as reported earlier,[4] and represent the mean from at least three different determinations. All CAs were recombinant proteins obtained as reported earlier by these groups.[5,6]

Cell culture. Hormone-independent prostate cells (PC-3), human embryonic kidney cells (HEK 293), and human neuroblastoma cells (SH-SY5Y, American Type Culture Collection number CRL-2266) were cultured in Dulbecco's Modified Eagle Medium F-12 (DMEM/F12) medium (Life Technologies, Carlsbad, CA, USA) in the presence of 10% fetal calf serum (Life Technologies) inactivated at 56 °C for 30 minutes. The cells are grown in an incubator at 37 °C in a humidified atmosphere containing 5% CO_2.

Assessment of cell viability. The viability of the different cancer cells was calculated after 24 hours, for PC-3 and HEK 293 that grow very rapidly, or 48 hours, for SH-SY5Y that have a slow doubling time, of compound treatmentat the indicate concentrations,[7] with and without overnight treatment of the cells by $CoCl_2$.[8] Cell viability was assessed by a colorimetric assay using the MTS assay (CellTiter 96 Aqueous One Solution Assay; Promega Corporation, Madison, WI, USA), according to the manufacturer's instructions. Absorbance at 490 nm was measured in a multilabel counter (Victor X5; Perkin Elmer, Waltham, MA, USA).

X-Ray Crystallography and Data Collection

Human CA II (*h*CA II) was expressed in BL21DE3 competent cells and purified as discussed previously.[9,10] In brief, purified *h*CA II was crystallized in a precipitation solution of 1.6 M Na-Citrate, 50 mM Tris, pH 7.8 using the hanging drop vapor diffusion method. Well-formed *h*CA II crystals were observed in 5 days.[9,10] Crystals of *h*CA II were soaked in stock solution of **1d** that was made by diluting a 20 mM, 100% DMSO solution of **1d** into 100 µL of precipitant solution. Crystals were soaked with a final concentration of ~2 mM of **1d** in the crystallization drop for 24 hours prior to data collection. X-ray diffraction data was collected at the Cornell High Energy Synchrotron Source (CHESS) on beamline F1 using a wavelength of 0.9177 Å. The data sets were collected using an ADSC Quantum 270 CCD detector at a crystal-to-detector distance of 120 mm with a 0.5° oscillation angle and an exposure time of 1 sec per image. A total of 360 images were collected. The data were indexed, integrated, and scaled using HKL200.[11] Data was scaled to the monoclinic C2 space group (unit cell parameters a = 42.4, b = 41.5, c = 72.3 Å, β = 104.3°), to 1.20 Å resolution with a completeness of 91.1 %, and an R_{sym} of 7.0 % (Summarized in Table S1). Phases were determined via molecular replacement using PDB: 3KS3,[12] as a search model and calculated using the program *Phenix*.[13] Refinements and ligand restraint files were also generated in the *Phenix*. Modification to Models and ligand PDB files were generated in Coot,[14,15] along with the determination of bond lengths of observable interactions. All figures were generated using *PyMol*.[16]

Table S1. X-ray crystallography statistics for data processing and refinement of **1d** in complex with *h*CA II.

CA II/1d PDB ID: 5CJL	
Space Group Cell Dimensions (Å;°)	$P2_1$ a = 42.4, b = 41.5, c = 72.3; β= 104.3°
Resolution (Å)	19.9 – 1.20 (1.23 – 1.20)
Total Reflections	67579
R_{sym}[a] (%)	7.0 (26.0)
I/Iσ	29.5 (5.73)
Redundancy	2.6 (2.7)
Completeness (%)	91.1 (90.7)
R_{cryst}[b] (%)	13.9 (18.9)
R_{free}[c] (%)	15.5 (21.4)
# of Protein Atoms	4205 (All Chains)
# of Water Molecules	339 (All Chains)
# of Ligand Atoms (includes hydrogens)	35
Ramachandran stats (%): Favored, allowed, generously allowed	97.3, 2.7, 0.0
Avg. B factors (Å^2): Main-chain, Side-chain, Ligand	10.3, 16.7, 20.6
rmsd for bond lengths, angles (Å,°)	0.009, 1.440

[a] $R_{sym} = (\sum|I - \langle I\rangle|/\sum \langle I\rangle) \times 100$

[b] $R_{cryst} = (\sum|F_o - F_c|/\sum |F_o|) \times 100$

[c] *R_{free} is calculated in the same way as R_{cryst} except it is for data omitted from refinement (5% of reflections for all data sets).*

[e] *Values in parenthesis correspond to the highest resolution shell.*

Hydrogens were added to structure since there was observable density at this resolution.

Molecular Modelling

All computational modeling studies were carried out on multiprocessor machines equipped with OS Ubuntu 14.04 or Windows Seven.

Preparation of the ligands. Ligand structures were built using standard bond lengths and angles from the fragment database with MacroModel 6.0.[17] Minimization of structures was performed with the MacroModel/BachMin 6.0 program using the AMBER force field. An extensive conformational search was further carried out using the Monte Carlo/energy minimization (Ei-E min <5 Kcal/mol, energy difference between the generated conformation and the current minimum).[18] The atomic charges were assigned using the Gasteiger-Marsili method,[19] in particular acid derivatives were considered as mono deprotonated species, whereas ester derivatives were kept in their neutral form. Representative minimum energy conformations of each compound were optimized using density functional theory (DFT) quantum chemistry program Gaussian 09W with the method B3LYP/6-311G basis set.[20] Visual analysis was performed with GaussView version 5.0.[21]

Structure refinement and computational docking procedures. Both *h*CAII and *h*CAXII protein models were prepared from the respectively 1.2 Å and 1.55 Å high resolution crystal structure. For CAII model our crystal complex between *h*CAII and compound **1d** (pdb code: 5CJL) was used, whereas for CAXII the free protein with pdb accession code 1JCZ was selected.[22] All ligands and water, except the water linked to the tetrahedral zinc, were stripped, and hydrogen atoms were added using the ADT module of MGLTools 1.5.7rc1,[23] then Gasteiger charges[19] were assigned. Ligands were docked using the Lamarckian genetic algorithm (LGA) defined through a grid box (coordinates: X, -7.0; Y, 2.0; Z, 15.0, and grid points: X, 50; Y, 50; Z, 50) with default grid spacing of 0.375 Å. The atom type set was modified including oxygen acceptor (OA) and hydrogen donor (HD) parameters. W maps were calculated according to Forli et al.[24] Docking runs were performed considering the active site as a rigid molecule and the ligands as flexible (i.e., all non-ring torsions were allowed), and using the Lamarckian genetic algorithm (LGA). Docking parameters for each compounds were set to 100 runs, population size of 150 individuals, maximum number of

generations and energy evaluations of 27,000 and 25,000,000, respectively. Docking results were analyzed using MGLTools 1.5.7rc1,[21] and AutoDock 4.2.6 docking programs.[25,26] From the estimated free energy of ligand binding (ΔG), the inhibition constant (K_i) for each ligand was calculated. K_i was calculated by the equation: K_i =exp [(ΔG*1000)/(R*T)]. Where ΔG is docking energy, R (gas constant) is 1.98719 cal $K^{-1}mol^{-1}$ and T (Temperature) is 298.15 K.

Comparison between the Forli and Olson's method and the standard docking procedure. To determine the best docking approach, in the early stage of our modelling study we re-docked compound **1d** into the *h*CAII protein model derived from the X-ray study using both the method of Forli and Olson, and the standard method implemented on AutoDock. Comparison between the best pose derived from each docking method and the experimental coordinates showed that Forli and Olson's procedure gave better results rather than the standard method in terms of rmsd (0.76 and 0.88, respectively) and visual inspection.

Figure S1. (A) Comparison of experimental structure of **1d** and docking obtained with adapted Forli and Olson's method within *h*CA II depicted as cartoon. (B) Best pose from docking **1d** into the *h*CA XII catalytic pocket. *h*CA II and XII are coloured as light brown and pale green, respectively; side chain of significant residues are represented as yellow sticks and labelled. Zinc cofactor, zinc-bound water and pseudoatoms are depicted as spheres and coloured in gray, red, and pale blue or cyan, respectively.

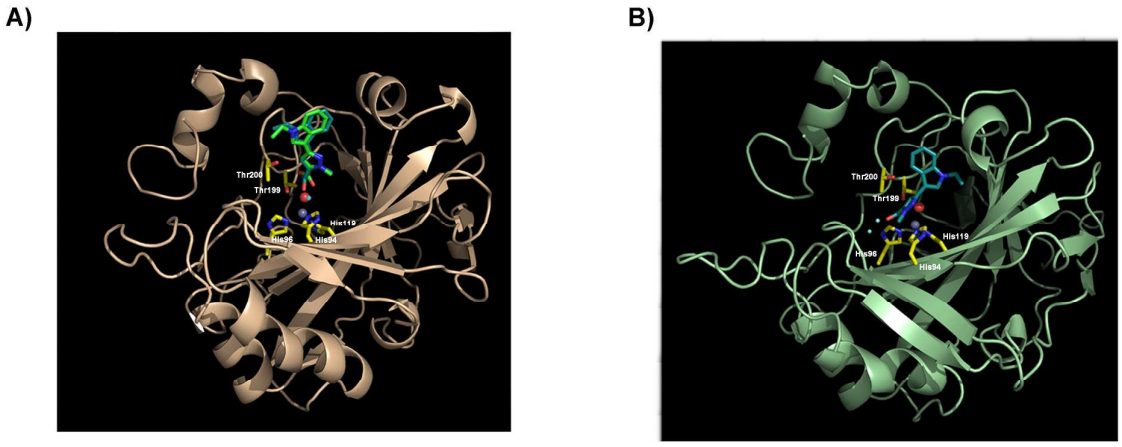


Figure S2. Overall view of protein surface with the top-ranked binding modes of titled compounds superimposed. *h*CA II (A) and XII (B) isoforms are coloured as light brown and pale green, respectively; side chain of significant residues are represented as yellow sticks and labelled. Ligand are in sticks (**1a**, white; **1b**, yellow; **1c**, magenta; **1d**, deep teal; **2a**, cyan; **2b**, orange; **2c**, pink; **2d**, blue violet). Zinc cofactor, zinc-bound water and pseudoatoms are depicted as spheres and coloured in gray, red, and pale blue or cyan, respectively.

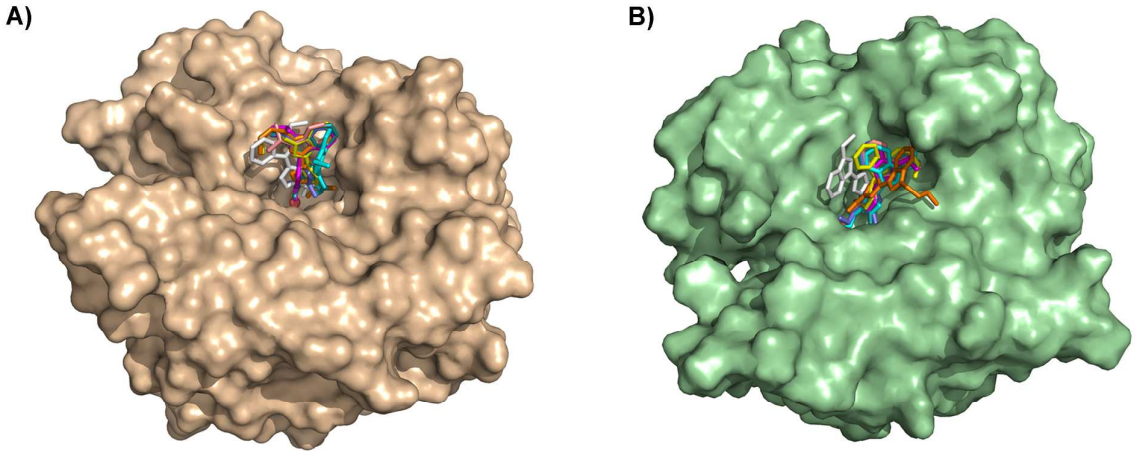


Figure S3. Molecular docking poses of compounds **1a-d** and **2a-d** in the *h*CA II binding site (pdb code: 5CJL) depicted as sticks. Catalytic and interacting residues are also represented as sticks, and are coloured in light apricot, and mantis, respectively. Zinc ion (gray) as well as the respective water bound molecule (red) are represented as spheres. Monoatomic pseudoatoms accounting for explicit water molecules are depicted as cyan spheres.

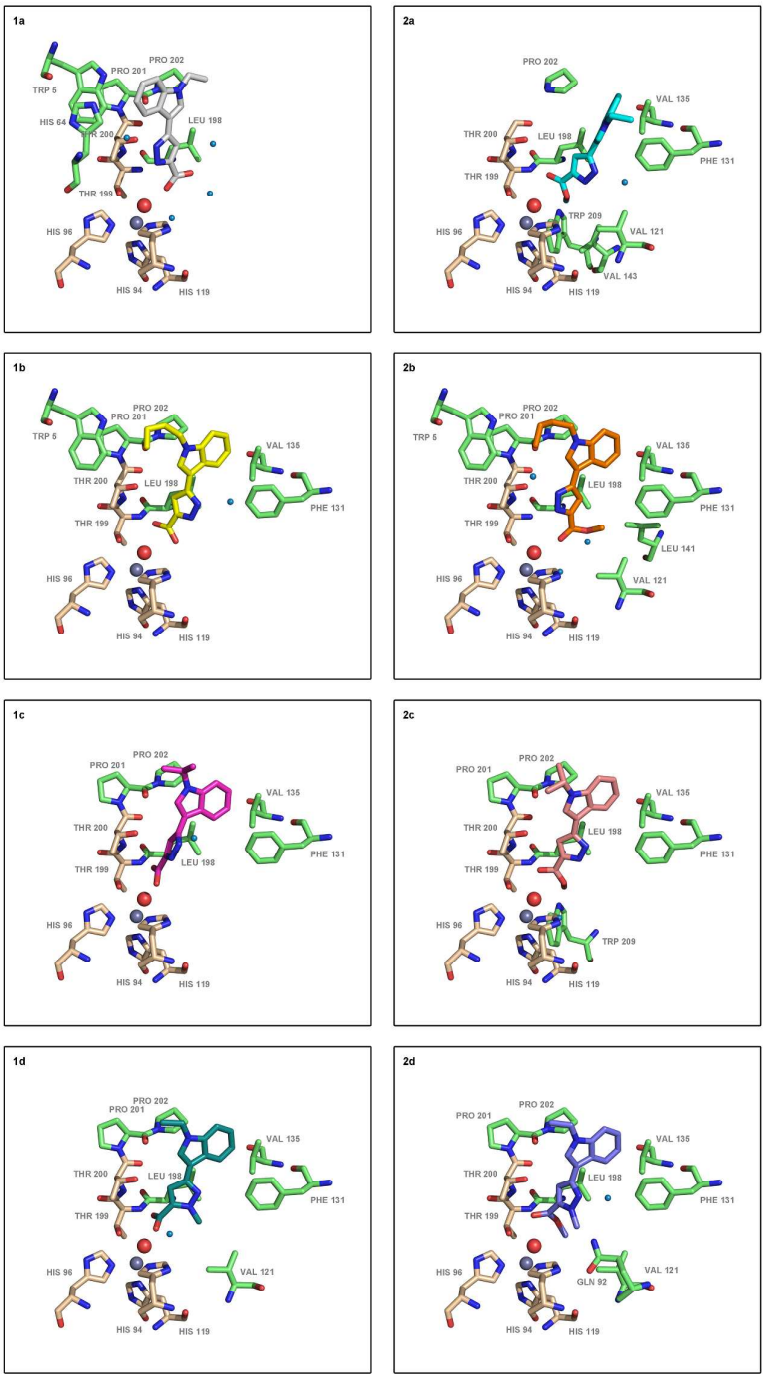


Figure S4. Molecular docking poses of compounds **1a-d** and **2a-d** in the *h*CA XII binding site (pdb code: 1JCZ) depicted as sticks. Catalytic and interacting residues are also represented as sticks, and are coloured in light apricot, and mantis, respectively. Zinc ion (gray) as well as the respective water bound molecule (red) are represented as spheres. Monoatomic pseudoatoms accounting for explicit water molecules are depicted as cyan spheres.

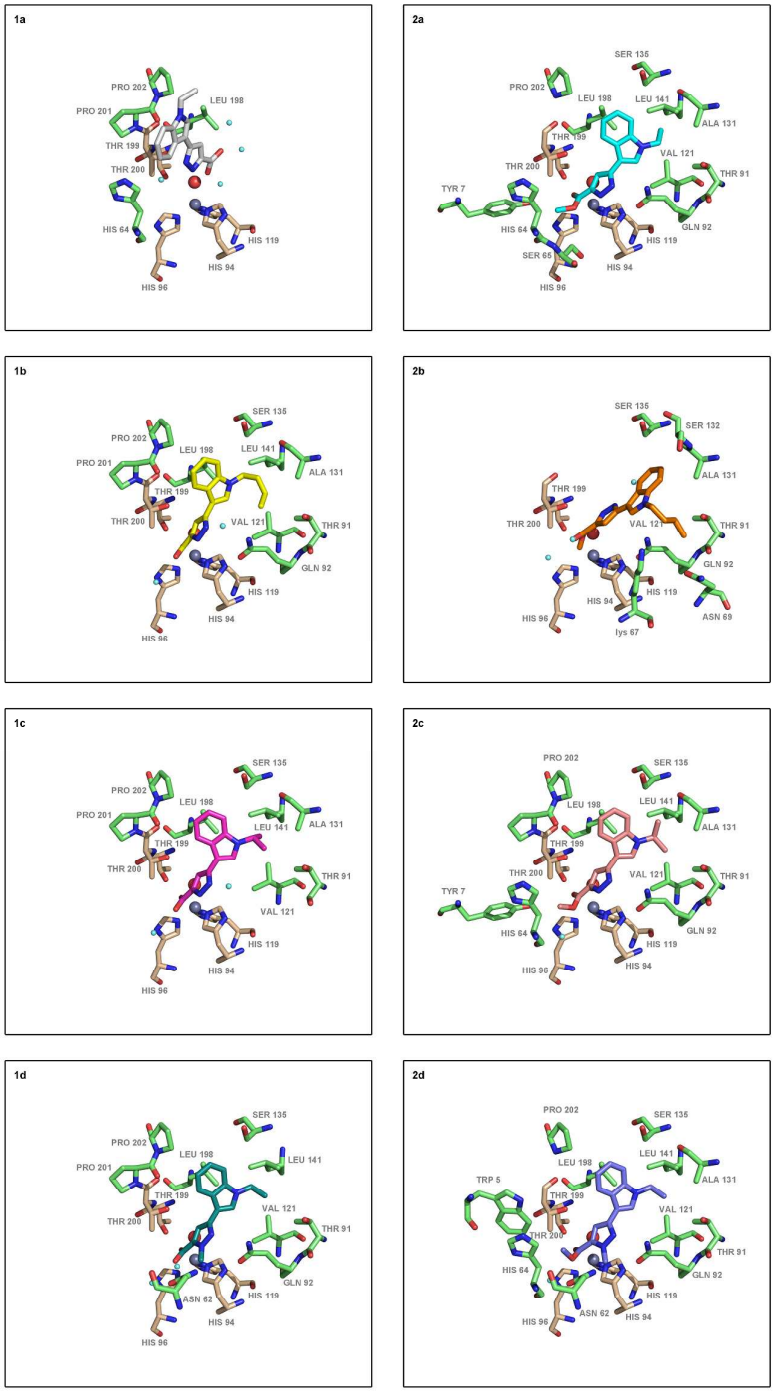


Table S2. Interactions of **1a-d** and **2a-d** with amino acid residues on *h*CA II and hCA XII binding pockets.

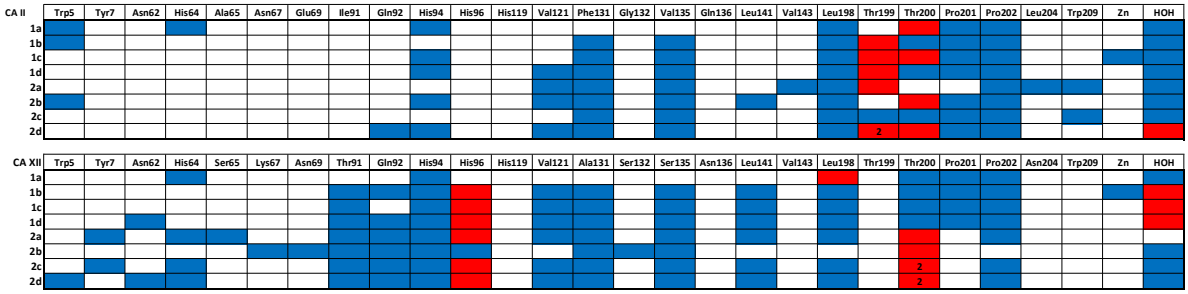


Figure S5. Superimposition of *h*CA II and *h*CA XII X-ray structures used for docking calculation. Focusing on the catalytic region, the primary sequence appears highly conserved, with the exception of nine residues (ie, A65S, N67K, E69D, I91T, P131A, G132S, V135S, Q136N, and L204N.

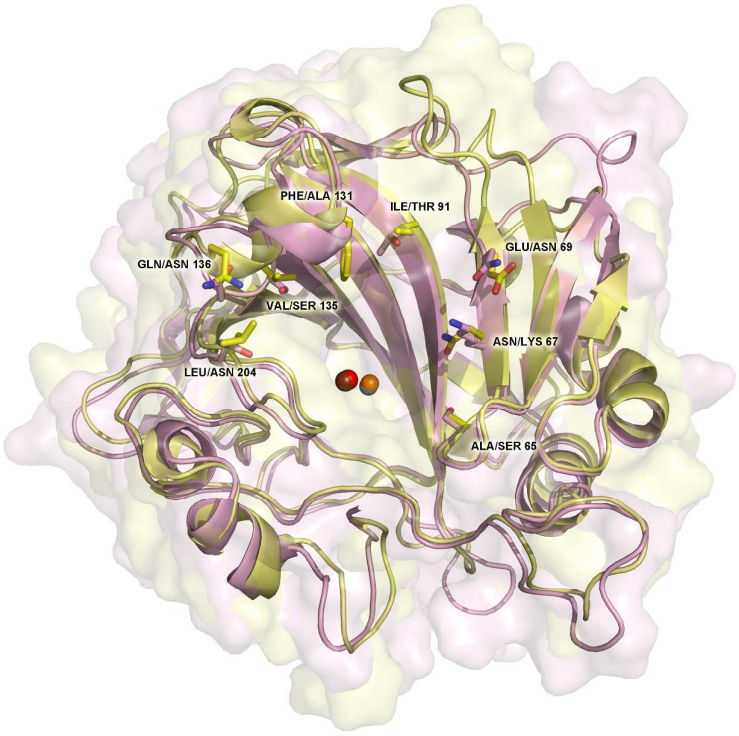
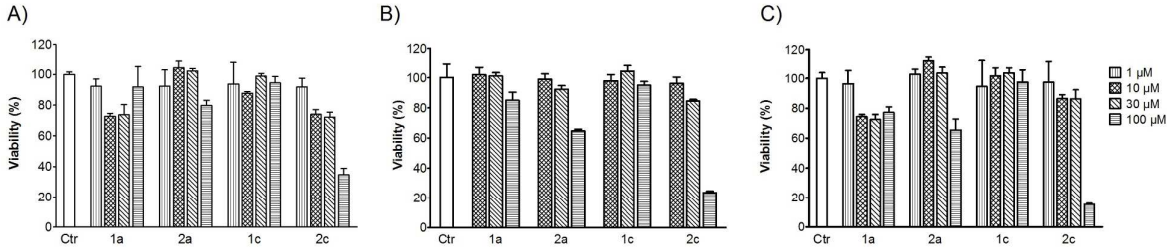


Figure S6. Antiproliferative activity of compounds **1a**, **2a**, **1c**, and **2c** at various concentrations assessed by percentage of cell viability on (A) human renal HEK 293 cells treated for 24 hours; (B) humane prostate PC-3 cells treated for 24 hours; and (C) human neuronal SH-SY5Y cells treated for 48 hours.



References

1. Sechi, M.; Innocenti, A.; Pala, N.; Rogolino, D.; Carcelli, M.; Scozzafava, A.; Supuran, C. T. Inhibition of α-class cytosolic human carbonic anhydrases I, II, IX and XII, and β-class fungal enzymes by carboxylic acids and their derivatives: New isoform-I selective nanomolar inhibitors. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 5801-5806.
2. Zhang, D.; Wang, G.; Tan, C.; Xu, W.; Pei, Y.; Huo, L. Synthesis and biological evaluation of 3-(1H-indol-3-yl)pyrazole-5-carboxylic acid derivatives. *Arch. Pharm. Res.* **2011**, *34*, 343-355.
3. Khalifah, R. G. The carbon dioxide hydration activity of carbonic anhydrase. I. Stop-flow kinetic studies on the native human isoenzymes B and C. *Biol. Chem.* **1971**, *246*, 2561-2573.
4. Sharma, A.; Tiwari, M.; Supuran, C. T. Novel coumarins and benzocoumarins acting as isoform-selective inhibitors against the tumor-associated carbonic anhydrase IX. *J. Enzyme Inhib. Med. Chem.* **2014**, *29*, 292-296.
5. Alterio, V.; Hilvo, M.; Di Fiore, A.; Supuran, C. T., Pan, P.; Parkkila, S.; Scaloni, A.; Pastorek, J.; Pastorekova, S.; Pedone, C.; Scozzafava, A.; Monti, S. M.; De Simone, G. Crystal structure of the catalytic domain of the tumor-associated human carbonic anhydrase IX. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16233-16238.
6. La Regina, G.; Coluccia, A.; Famiglini, V.; Pelliccia, S.; Monti, L.; Vullo, D.; Nuti, E.; Alterio, V.; De Simone, G.; Monti, S. .; Pan, P.; Parkkila, S.; Supuran, C. T.; Rossello, A.; Silvestri, R. Discovery of 1,1'-biphenyl-4-sulfonamides as a new class of potent and selective carbonic anhydrase XIV inhibitors. *J. Med. Chem.* **2015**, *58*, 8564-8572.
7. Sanna, V.; Pala, N.; Dessì, G.; Manconi, P.; Mariani, A.; Dedola, S.; Rassu, M.; Crosio, C.; Iaccarino, C.; Sechi, M. Single-step green synthesis and characterization of gold-conjugated polyphenol nanoparticles with antioxidant and biological activities. *Int. J. Nanomed.* **2014**, *9*, 4935-4951.
8. Al Okail, M. S. Cobalt chloride, a chemical inducer of hypoxia-inducible factor-1α in U251 human glioblastoma cell line. *J. Saudi Chem. Soc.* **2010**, *14*, 197-201.
9. Pinard, M. A.; Boone, C. D.; Rife, B. D.; Supuran, C. T.; McKenna, R. Structural study of interaction between brinzolamide and dorzolamide inhibition of human carbonic anhydrases. *Bioorg. Med. Chem.* **2013**, *21*, 7210-7215.
10. Moeker, J.; Mahon, B. P.; Bornaghi. L. F.; Vullo, D.; Supuran, C. T.; McKenna, R.; Poulsen. S. A. Structural insights into carbonic anhydrase IX isoform specificity of carbohydrate-based sulfamates. *J. Med. Chem.* **2014** ,*57*, 8635-8645.
11. Otwinowski, Z., Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307-326.
12. Avvaru, B. S.; Kim, C. U.; Sippel, K. H.; Gruner, S. M.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R. A short, strong hydrogen bond in the active site of human carbonic anhydrase II. *Biochemistry* **2010**, *49*, 249-251.
13. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Echols, N.; Headd, J. J.; Hung, L. W.; Jain, S.; Kapral, G. J.; Grosse Kunstleve, R. W.; McCoy, A. J., Moriarty, N. W.; Oeffner, R. D.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. The Phenix software for automated determination of macromolecular structures. *Methods* **2011**, *55*, 94-106.
14. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr*. **2004**, *60*, 2126-2132.
15. Debreczeni, J. É.; Emsley, P. Handling ligands with Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 425-430.
16. Schrodinger LLC. The PyMOL Molecular graphics system, version 1.2r3pre, Schrödinger, LLC.
17. Mohamadi, F.; Richards, N. G.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufiel, C.; Chang, G.; Hendrickson, T.; Still, W. C. Macromodel—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.,* **1990**, *11*, 440-467.
18. Chang, G.; Guida, W. C., Still, W. C. An internal-coordinate Monte Carlo method for searching

conformational space. *J. Am. Chem. Soc*. **1989**, *111*, 4379–4386.
19. Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity – a rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219–3228.
20. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B., Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R., Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X., Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A. Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith ,T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.;. Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, Gaussian 09, Revision B.01, 2010.
21. Dennington, R.; Keith T.; Millam, J. GaussView, Version 5.0 Semichem Inc., Shawnee Mission KS, 2009.
22. Whittington, D. A.; Waheed, A.; Ulmasov, B.; Shah, G. N.; Grubb, J. H.; Sly, W. S.; Christianson, D. W. Crystal structure of the dimeric extracellular domain of human carbonic anhydrase XII, a bitopic membrane protein overexpressed in certain cancer tumor cells. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9545-9550.
23. Sanner, M. F. Python: a programming language for software integration and development. *J. Mol. Graph. Model* **1999**, *17*, 57-61.
24. Forli, S.; Olson, A. J. A force field with discrete displaceable waters and desolvation entropy for hydrated ligand docking. *J. Med. Chem.* **2012**, *55*, 623-638.
25. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew R. K.; Olson, A. J. Automated docking using a lamarckian genetic algorithm and and empirical binding free energy function. *J. Comput. Chem.,* **1998**, *19*, 1639-1662.
26. Huey, R.; Morris, G. M.; Olson A. J.; Goodsell, D. S. A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem.,* **2007**, *28*, 1145-1152.